3-2011

Isolation by Distance Explains Genetic Structure of Buggy Creek Virus, a Bird-Associated Arbovirus

Abinash Padhi
University of Tulsa, apadhi@umd.edu

Amy T. Moore
University of Tulsa

Mary Bomberger Brown
University of Nebraska–Lincoln, mbrown9@unl.edu

Jerome E. Foster
University of Tulsa, jerome.foster@sta.uwi.edu

Martin Pfeffer
University of Leipzig, pfeffer@vetmed.uni-leipzig.de

See next page for additional authors

Published in *Evolutionary Ecology* 25:2 (March 2011), pp. 403–416; doi: 10.1007/s10682-010-9419-9

Submitted May 7, 2010; accepted August 16, 2010; published online September 10, 2010.

Isolation by Distance Explains Genetic Structure of Buggy Creek Virus, a Bird-Associated Arbovirus

Abinash Padhi,[1] Amy T. Moore,[1] Mary Bomberger Brown,[1]

Jerome E. Foster,[1] Martin Pfeffer,[2] and Charles R. Brown[1]

1. Department of Biological Sciences, University of Tulsa, Tulsa, OK 74104, USA
2. Institute for Animal Hygiene and Veterinary Public Health, University of Leipzig, 04103 Leipzig, Germany

Corresponding author – Charles R. Brown, email charles-brown@utulsa.edu

Note: A. Padhi is at the Center for Infectious Disease Dynamics, Department of Biology, 208 Mueller Laboratory, The Pennsylvania State University, University Park, PA 16802, USA; M. B. Brown is with the Tern and Plover Conservation Partnership, University of Nebraska, 3310 Holdrege St, Lincoln, NE 68583, USA; and J. E. Foster is with the Department of Preclinical Sciences, Faculty of Medical Sciences, University of the West Indies, St. Augustine, Trinidad, West Indies.

Abstract

Many of the arthropod-borne viruses (arboviruses) show extensive genetic variability and are widely distributed over large geographic areas. Understanding how virus genetic structure varies in space may yield insight into how these pathogens are adapted to and dispersed by different hosts or vectors, the relative importance of mutation, drift, or selection in generating genetic variability, and where and when epidemics or epizootics are most likely to occur. However, because most arboviruses tend to be sampled opportunistically and often cannot be isolated in large numbers at a given locale, surprisingly little is known about their spatial genetic structure on the local scale at which host/vector/virus interactions typically occur. Here, we examine fine-scale spatial structure of two sympatric lineages of Buggy Creek virus (BCRV, Togaviridae), an alphavirus transmitted by the ectoparasitic swallow bug (*Oeciacus vicarius*) to colonially nesting cliff swallows (*Petrochelidon pyrrhonota*) and invasive house sparrows (*Passer domesticus*) in North America. Data from 377 BCRV isolates

at cliff swallow colony sites in western Nebraska showed that both virus lineages were geographically structured. Most haplotypes were detected at a single colony or were shared among nearby colonies, and pair-wise genetic distance increased significantly with geographic distance between colony sites. Genetic structure of both lineages is consistent with isolation by distance. Sites with the most genetically distinct BCRV isolates were occupied by large numbers of house sparrows, suggesting that concentrations of invasive sparrows may represent foci for evolutionary change in BCRV. Our results show that bird-associated arboviruses can show genetic substructure over short geographic distances.

Keywords: arbovirus, Buggy Creek virus, cliff swallow, house sparrow, swallow bug, virus evolution, virus population genetics

Introduction

The arthropod-borne RNA viruses (arboviruses) include many that directly affect either human or animal health, and most of them tend to be distributed over wide geographic areas (Calisher 1994; Gubler 2002). Knowing the genetic structure of virus populations allows us to better understand the extent and mechanisms of immigration between subpopulations, colonization history, co-evolution with specific vectors or hosts, and selection in response to environmental variation as well as predicting where virus may respond to environmental change in ways that promote disease outbreaks (Moya et al. 2004; Bertolotti et al. 2008; Holmes 2009). For example, genetic similarity of some arboviruses from widely separated locations has suggested that mobile hosts such as birds could be responsible for disseminating these viruses on a continental scale and thus maintaining genetic homogeneity over wide areas (Weaver et al. 1994; Cilnis et al. 1996; Norder et al. 1996; Kramer et al. 2008; van den Hurk et al. 2009). In other cases, strong spatial structure over small areas has indicated that most arbovirus transmission is highly localized (Bertolotti et al. 2008) and may lead to the establishment of locally restricted genotypes (Kramer et al. 1997; Chandler et al. 2001; Reisen et al. 2002).

Most of what we know about the phylogeography of arboviruses comes from studies of virus isolates taken from across entire continents or from different continents. This in part reflects the difficulty in getting samples: endemic arboviruses usually exist in cryptic, enzootic cycles in (mostly) mosquito vectors and wild bird or mammal hosts, and periodically some of them seem to disappear (Forrester et al. 2008; Reisen et al. 2008). Field surveys tend to require massive effort to yield even a relatively few isolates of the better studied arboviruses (Stamm and Newman 1963; Lord and Calisher 1970; Monath 1980; Reeves 1990; Crans et al. 1994; Howard et al. 2004; Dusek et al. 2009), and consequently most of the spatially based genetic analyses of arboviruses (with few exceptions; Bertolotti et al. 2008) typically have had to rely on opportunistically collected virus isolates from widely separated locations and sometimes over a considerable span of years (Weaver et al. 1994, 1997; Cilnis et al. 1996; Sammels et al. 1999; Kramer and Fallah 1999; Davis et al. 2005). We know almost nothing about how arbovirus genetic structure varies at the fine (highly localized) scale where most interactions between virus, hosts, and vectors typically occur (Bertolotti et al. 2008) and where the effects of selection, drift, mutation, or migration may lead to disease emergence (Holmes 2009).

Here, we use a large genetic dataset on an arbovirus that can be sampled systematically over small geographic regions to (1) explore how genetic structure varies in space, and from that (2) infer mechanisms generating genetic variability and (3) predict spatial foci for evolutionary change in this virus. Buggy Creek virus (BCRV; Togaviridae, *Alphavirus*) is a single-stranded, positive sense RNA virus that is part of the western equine encephalomyelitis virus antigenic complex (Hopla et al. 1993) and in North America occurs in two distinct lineages that differ by >6% at the nucleotide level (Pfeffer et al. 2006; Padhi et al. 2008). The primary vector for BCRV and related viruses (Calisher et al. 1980; Scott et al. 1984; Brault et al. 2009) is the swallow bug (Hemiptera: Cimicidae: *Oeciacus vicarius*), an ectoparasite of the colonially nesting cliff swallow (*Petrochelidon pyrrhonota*). Vertebrate amplifying hosts for BCRV include cliff swallows and introduced house sparrows (*Passer domesticus*) that occupy nests in swallow colonies (Scott et al. 1984; Hopla et al. 1993; O'Brien et al. 2010a). Because the wingless swallow bugs are sedentary throughout the year and confined to cliff swallow colony sites, the spatial foci for BCRV occurrence are discrete and predictable (Brown et al. 2001, 2008; Moore et al. 2007). BCRV can be isolated relatively frequently from swallow bug vectors at all times of the year (Brown et al. 2009a,b, 2010b).

In this study, we used 377 isolates of BCRV collected in 1998–2006 from swallow bug vectors in summer at cliff swallow colony sites in southwestern Nebraska (Brown et al. 2008, 2009b) to investigate whether genetic similarity of virus varies with distance between sampling sites, whether genetic divergence of virus at certain sites is related to presence of one or both potential host species, and whether the two ecologically different lineages (designated A and B; [Brown et al. 2009b]) differ in spatial structure, genetic diversity, and potential rates of gene flow. No spatial genetic analyses have been done previously for BCRV, and we use the results to gain insight into patterns of fine-scale virus spatial structure and factors that generate spatial variation. Similar studies exist for only a few arboviruses and for no other alphaviruses.

As in earlier analyses (Brown et al. 2008, 2009b), we examine variation in the 1,269 bp of the E2 glycoprotein-coding region, which in alphaviruses codes for glycoproteins that are responsible for cell receptor binding (Navaratnarajah and Kuhn 2007), and is the region of the genome most sensitive to selection brought about by the immune systems of different hosts (Strauss and Strauss 1994; Powers et al. 2001; Pfeffer et al. 2006). The E2 gene has been used in studies of other alphaviruses to infer phylogeographic patterns over wide geographic areas (Oberste et al. 1998; Kramer and Fallah 1999; Kondig et al. 2007).

Materials and methods

Study organisms

Swallow bugs are nest-based parasites primarily of cliff swallows that overwinter in the swallows' nests or in the cracks and crevices of the nesting substrate near the nests. They are hematophagous, feeding on the birds mostly at night, and they travel on the adult birds only briefly (Brown and Brown 1996, 2004, 2005). Swallow bugs are long-lived and begin to reproduce as soon as they feed in the spring, and they can survive at a site for up to 3 years in the absence of hosts (Smith and Eads 1978; Loye and Carroll 1991; Rannala 1995).

The bugs also parasitize house sparrows that nest in some cliff swallow colonies (Hopla et al. 1993; O'Brien et al. 2010a). BCRV overwinters in bugs at temperate latitudes as far north as southern North Dakota (Brown et al. 2009a, 2010b).

Cliff swallows are highly colonial passerines that breed throughout most of western North America (Brown and 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. Although swallows are important as hosts for swallow bugs and move BCRV-infected bugs from site to site (Brown et al. 2008), cliff swallows do not amplify the virus well when infected, and may not be important in BCRV transmission cycles (O'Brien et al. 2010a).

House sparrows were introduced into North America from Europe in the late 1800s and are found in all parts of the United States (Robbins 1973; Lowther and 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 nonmigratory and resident around the swallow colonies throughout the year. Nestling house sparrows are relatively frequently infected by BCRV and amplify the virus to high titers (O'Brien et al. 2010a, b).

Study site

Our study site was in southwestern Nebraska, centered 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 Garden, Lincoln, and Morrill counties (fig. 1a). Cliff swallows have been studied there since 1982 (Brown and 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. Presence of house sparrows was noted at each cliff swallow colony site each year, and for some sites the maximum number of active sparrow nests was recorded by inspecting nest contents with a flashlight and dental mirror.

Field collections, virus isolation, and sequencing

Collections of swallow bugs for virus isolation were done each year from 1998 to 2006 in May, June, and July. 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.

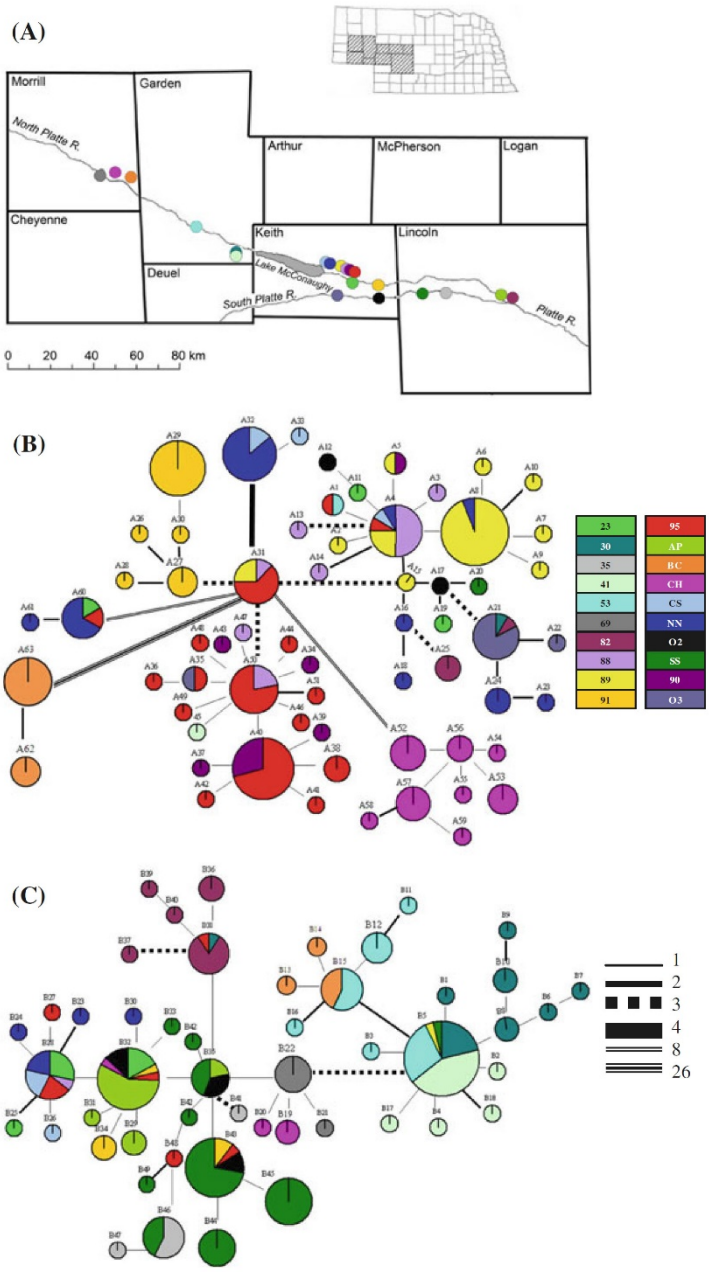

Figure 1. Locations of colony sites in southwestern Nebraska sampled for Buggy Creek virus in this study (a). Minimum spanning network showing relationships among BCRV isolates of lineage A (b) and B (c) found at each colony site. Each circle represents a unique E2 gene sequence (haplotype), and the size of the circle is proportional to the number of isolates of that haplotype. Colony sites are indicated by unique colors. Lines connecting the haplotypes show by how many mutations they differ. Spatial relationships between haplotypes (b, c) are inferred based on colony color codes and geographic positions of colonies (a).

Virus was isolated from bug pools following the methods given previously (Brown et al. 2008). RNA from each isolate (n = 377) was used in an alphavirus reverse transcription-polymerase chain reaction (RT–PCR) to amplify the entire 1,269 bp of the E2 gene (Brown et al. 2008). Sequences were aligned against the corresponding region in a 1981 BCRV reference sequence (strain 81V1822, GenBank no. AF339474) and fragments combined for a given isolate using SEQMAN 6.1 (DNAStar, Lasergene) to obtain a contiguous nucleotide sequence for each sample. Additional details on field sampling, virus isolation, and sequencing are given in Moore et al. (2007), Brown et al. (2008), and Padhi et al. (2008). The GenBank accession numbers of the sequences used in this paper are EU483667–EU484043.

Analyses

Using the previously published sequences as reference (Pfeffer et al. 2006; Padhi et al. 2008), the lineage of each BCRV isolate was determined from a maximum likelihood phylogeny constructed using PAUP* 4.0b10 (Swofford 2002). The number of unique sequences of a lineage was determined with COLLAPSE version 1.2 (Posada 2004). Each unique sequence that differed from all others by at least one nucleotide change was designated as a haplotype. The appropriate model of nucleotide substitution for the combined data set and for each lineage data set was selected using the hierarchical likelihood ratio test as implemented in MODELTEST version 3.7 (Posada and Crandall 1998). Only colony sites (n = 20) for which we had at least five isolates identified to haplotype were used in the analyses here. ARLEQUIN version 3.1 (Schneider et al. 2000) was used to reconstruct the minimum spanning network; to perform Mantel tests to determine the significance of the relationship between Slatkin's linearized F_{st} ($F_{st}/(1-F_{st})$ estimates and the natural log of the geographic distance among each pair of colony sites (with significance testing based on 10,000 simulations); and to estimate the number of effective migrants (N_m) for colony pairs. Population pair-wise F_{st} estimates were derived in ARLEQUIN using Kimura 2-Parameter (K2P) distances with a gamma correction of 0.0618 for lineage A and 0.172 for lineage B. We also used ARLEQUIN to estimate the overall nucleotide diversity ($\pi \pm$ SE) and the mean number of pair-wise nucleotide differences (d ± SE) among the isolates within each lineage. To determine whether any codons of the E2 gene in either lineage were under positive selection, we performed selection analyses using the Single Likelihood Ancestor Counting (SLAC) and Fix Effect Likelihoods (FEL) methods using DATAMONKEY (Kosakovsky Pond and Frost 2005).

Results

The Tamura-Nei model with equal frequency (TrNef) and a gamma distribution shape parameter (G = 0.2026) was the best-fit nucleotide substitution model for the combined BCRV data set. The rate matrix for these data was A–C, 1.0000; A–G, 5.0826; A–T, 1.0000; C–G, 1.0000; C–T, 7.9673; and G–T, 1.0000. The K80 (the Kimura-two-parameter) model with gamma distribution shape parameters of 0.0618 for lineage A and 0.172 for lineage B was the best-fit nucleotide substitution model for each BCRV lineage. The transition to transversion ratio (Ti/Tv) was 2.2974 and 3.4396 for lineage A and B, respectively. The nucleotide diversity ($\pi \pm 1$ SE) of the E2 gene and the mean number of pair-wise nucleotide

differences (d ± 1 SE) among the BCRV isolates belonging to lineage A ($\pi = 0.0074 \pm 0.0038$; $d = 9.42 \pm 4.34$) were higher than the comparable values estimated for lineage B ($\pi = 0.0033 \pm 0.0018$; $d = 4.16 \pm 2.08$), although overlapping confidence intervals indicated that these differences were not statistically significant. The overall ratio of the rates of nonsynonymous (d_n) to synonymous (d_s) substitutions ($d_n/d_s = \omega$) for lineages A and B were 0.222 (95% CI, 0.161–0.297) and 0.225 (0.141–0.335), respectively. Codon-specific selection analyses detected no evidence of positive selection in either lineage.

From the 377 BCRV isolates (191 from lineage A, 186 from lineage B), we obtained 63 unique E2 gene sequences (haplotypes) of lineage A and 50 of lineage B. Each lineage was found throughout the study area (fig. 1). Most haplotypes were specific to a single colony site: 52 lineage A and 42 lineage B haplotypes were found at only one site (fig. 1). Seven lineage A haplotypes and 2 lineage B haplotypes were isolated at two colonies each, with the remaining 10 found at 3 or more colony sites. The most widely distributed haplotype (from lineage B) was found at 6 different colony sites (B32; fig. 1c), with these sites spanning an east–west distance of 180 km (colony AP to colony CH), which was essentially the entire study area (fig. 1a).

Both lineages showed geographically structured populations (fig. 1). Within lineage A, colonies BC, CH, and 91 had no haplotypes that were shared with other colonies. Two of these sites (BC, CH) were only 8.9 km apart and about 95 km from the cluster of sites where the next closest haplotype was found (fig. 1). The third site, 91, that shared no haplotypes with other colonies was 17.5 km from the colonies containing the next closest haplotype. Virus from sites BC and CH differed by 26 and 8 mutational steps, respectively, from the next most similar haplotype, while that at site 91 differed by 3 mutational steps (fig. 1b). In contrast, other lineage A haplotypes were shared between several sites: haplotype A21 was found mostly at site 03 in the center of the study area, but also occurred at site 82, 76 km to the east, and at site 30, 48 km to the west (fig. 1). Haplotype A4 was found at 5 sites all within 11.5 km of each other (fig. 1). Sites 95 and NN had less geographic structure, each containing multiple, divergent virus isolates from across the network (fig. 1).

Isolates from lineage B showed less geographic substructure, with only site 69 (at the edge of the study area) not sharing any haplotypes with other sites (fig. 1c): virus at this site differed by only one nucleotide from the next most similar. Lineage B isolates exhibited fewer mutational-step differences among haplotypes/colony populations, on average, than did lineage A. Most shared lineage B haplotypes, however, still occurred mostly among nearby sites (fig. 1c), with a notable exception being haplotype B38 that was primarily found at site 82 but also appeared at site 30, 122 km to the west (fig. 1). Overall, colony sites that were dominated by lineage A isolates had relatively few or no isolates from lineage B (fig. 1b,c).

Mantel tests showed a strong association between the genetic distance, as measured by pair-wise linearized F_{st}, and the geographic distance between each pair of colony sites in the study area (fig. 2). This relationship held for both lineages, indicating that even though lineage B haplotypes were less differentiated from each other than were those of lineage A, each still exhibited strong spatial structure.

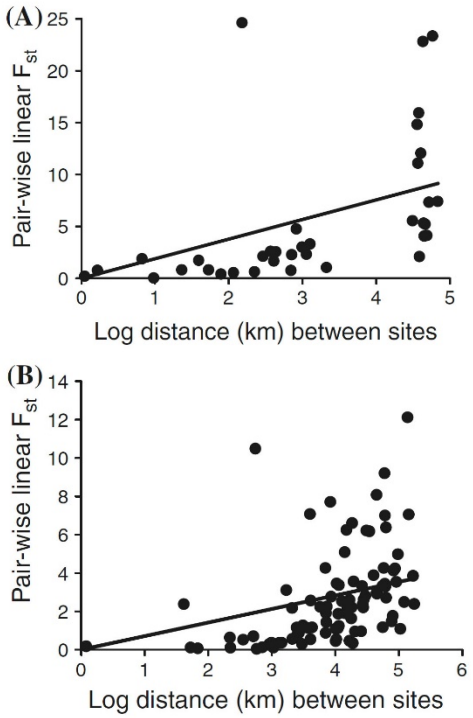

Figure 2. Pair-wise linearized F_{st} values ($F_{st}/(1-F_{st})$) for Buggy Creek virus between different colony sites in the Nebraska study area in relation to natural log of the geographic distance (in km) between each pair, for lineage A isolates (a) and lineage B isolates (b). Genetic distance increased significantly with the geographic distance between sites for lineage A ($r = 0.524$, regression slope 0.274, $P = 0.001$) and lineage B ($r = 0.436$, regression slope 0.190, $P = 0.0005$; Mantel test for each).

BCRV of both lineages tended to show significant ($P < 0.05$) genetic differentiation by colony site; however, lineage A isolates showed a relatively higher degree of genetic differentiation by site and a lower level of gene flow between colony sites (as measured by the number of effective migrants) than did virus of lineage B (fig. 3). For lineage A, only one pair of colonies (sites 90–95) showed no significant genetic differentiation among isolates ($P > 0.05$), while for lineage B, six pairs of sites (95–23, NN–23, 91–95, 02–91, 02–95, NN–95) showed no significant differentiation. These results are consistent with those of the haplotype network (fig. 1). Because earlier work showed little temporal structure in BCRV isolates (Padhi et al. 2008), we have not considered year of isolation in these analyses.

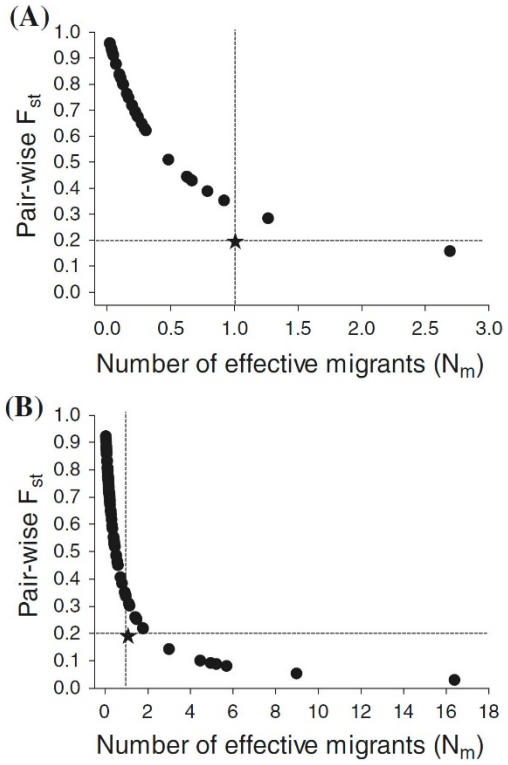

Figure 3. Pair-wise F_{st} values (fixation indices) for Buggy Creek virus between different colony sites in the Nebraska study area in relation to the estimated number of effective migrants (N_m) between each pair of sites, for lineage A isolates (a) and lineage B isolates (b). Note difference in scale of *x*-axes. In each, star and intersection of dotted lines indicate the point at which $N_m \approx 1$ and $F_{st} \approx 0.2$, a value conventionally interpreted as the approximate demarcation point between "high" and "low" gene flow (Avise 2000).

Presence of house sparrows was related to the extent of BCRV divergence at a cliff swallow colony site. For 16 sites we knew the maximum number of active house sparrow nests per season. At 10 sites where ≥ 50% of isolates belonged to lineage A, the mean (± SE) number of sparrow nests was 4.9 (± 2.4), compared to 0.0 (± 0.0) at 6 sites where < 50% of isolates belonged to lineage A; this difference was significant (Wilcoxon test, $Z = -2.19$, $P = 0.029$). The three colony sites with the most genetically diverse virus of lineage A (sites BC, CH, and 91; fig. 1) contained 10, 20, and 16 sparrow nests, respectively, compared to 0 or 1 sparrow nests at all other sites. Sites BC and 91 contained only house sparrows in most years of the study, with cliff swallows occupying these colony sites infrequently. Two of these sparrow sites (BC, CH) exhibited a pair-wise (non-linearized) F_{st} of 0.961, the highest of any of the colony pairs, despite being only 8.9 km apart.

Discussion

Our results illustrate strong spatial structuring of BCRV. Many BCRV haplotypes occurred at only one cliff swallow colony site, and the shared haplotypes were most often found at sites in relatively close geographic proximity. This spatial structure applied to both BCRV lineages, although lineage A showed greater haplotype divergence across the study area, a relatively higher degree of genetic differentiation by site, and a lower level of gene flow between colony sites.

Gene flow and isolation by distance

BCRV gene flow between neighboring colony sites is likely caused by the movement of transient cliff swallows (especially nonbreeders) that circulate among colonies and introduce BCRV to sites (Brown et al. 2007), probably via introduction of infected swallow bugs that travel on the birds' feet (Brown and Brown 2004). Although no formal analyses of bird movement and distance have been done, some evidence indicates that swallows moving between colony sites within a summer are more likely to visit colonies that are nearby (Brown and Brown 1996). Sites closer together are thus probably more likely, on average, to have virus haplotypes exchanged among them by bird movement, and such area-restricted virus movement presumably accounts in part for the relationship between genetic and geographical distance reported here. Thus, spatial structure in both BCRV lineages is consistent with classical isolation by distance models (Wright 1943; Slatkin 1993).

Viruses of the two lineages, however, do apparently differ in the extent to which they are dispersed between colony sites. For example, lineage A exhibited greater genetic differentiation by site, and haplotype diversity of lineage A at a site was more affected by the extent of bird movement into the site (Brown et al. 2008). In contrast, isolates from lineage B showed less genetic differentiation between neighboring colonies and thus presumably exhibit more gene flow. Earlier work revealed no relationship between rates of host-vector movement into a site and lineage B haplotype diversity (Brown et al. 2008). The results presented here suggest that this might be because overall rates of gene flow of lineage B between sites are higher than for lineage A.

The difference in apparent gene flow between sites for the two lineages is perplexing, given what we understand about the ecology of these virus lineages (Brown et al. 2009b). Cliff swallows do not effectively amplify BCRV of either lineage (O'Brien et al. 2010a), and thus it is unlikely that preferential movement of one lineage by selectively infected viremic birds occurs. If birds move virus primarily by transporting infected bugs, we might expect them to be equally likely to carry bugs with virus of either lineage; however, bugs infected with lineage B seem to be less likely to disperse, being more often found hidden in the recesses of the nests (Brown et al. 2009b). This should lead to greater genetic isolation of lineage B virus, the opposite of what we found. Possibly, the more restricted gene flow suggested for lineage A reflects simply this lineage's association with sites containing only house sparrows (see below), which do not attract heavy transient swallow traffic from nearby sites and are thus likely to contain more genetically isolated virus.

If genetic differentiation in BCRV among colonies results from local adaptation to environmental conditions, we would expect the signature of positive selection in the E2 gene

(Real et al. 2005). However, we failed to detect evidence of such selection in our Nebraska BCRV populations. This also suggests that geographic isolation (isolation by distance) is the best explanation for the strong spatial genetic differentiation among BCRV isolates. Possibly, other parts of the BCRV genome might be undergoing selection in these lineages, although given the importance of the E2 gene in mediating host cell receptor binding in alphaviruses (Strauss and Strauss 1994; Navaratnarajah and Kuhn 2007), selection if occurring would seem most likely to be detected in the region of the genome we studied.

Effect of house sparrows

Genetic structure of BCRV could not be predicted entirely by the geographic distance between sites. Despite their close geographic proximity, the relatively high genetic differentiation of lineage A virus between two colonies (BC and CH) that contained large numbers of house sparrows indicates that presence of sparrows also can affect BCRV genetic variation. Why this happens is unknown; we do know that sparrows preferentially amplify lineage A (rather than lineage B; Brown et al. 2010a). At sites where sparrows perennially occur, their providing of continual blood meals for bugs prevents periodic bug (and thus virus) extinction, and the consequence is that sites with sparrows maintain more lineage A throughout the year than do sites without sparrows (Brown et al. 2010b). More stable populations of lineage A should presumably lead to greater genetic homogeneity over time. However, with frequent nest failures due to sparrow nestlings succumbing to virus (O'Brien et al. 2010b), turnover of sparrow hosts at these sites may be high, with the same adult house sparrows rarely occupying a colony in successive summers (O'Brien 2009). Perhaps the genetic variability of BCRV at such sites reflects its evolutionary response to the continual replacement of host individuals with varying susceptibility to this virus. Invasive house sparrows are relatively new hosts for BCRV and may have contributed to the initial divergence of the two lineages by their effective amplifying of lineage A (Brown et al. 2009b, 2010a).

The potential for isolates to diverge significantly at sites perennially used by large numbers of house sparrows (even when such sites are close together, as two of ours were) suggests that such sites may represent predictable foci for evolutionary change in BCRV (and places to monitor for potential escape of this virus to novel hosts or vectors). Especially if cliff swallows do not use these sites for several years and thus no transient swallows enter and exit the colonies, little virus haplotype immigration or emigration among the sparrow sites is likely to occur, because adult sparrows apparently do not move bugs between colonies (we have not found any bugs on adult sparrows caught in nets). Because the wingless bugs are sedentary, little virus exchange among sites can happen in the absence of swallow traffic that circulates the hitchhiking bugs, in turn promoting genetic isolation and potential in situ evolution of BCRV at sites with only sparrows. This may account in part for the generally lower level of gene flow (number of effective migrants) for virus of lineage A (fig. 3).

Population genetics of other arboviruses

BCRV differs from the mosquito-associated arboviruses in two important ways. First, the spatial array of cliff swallow colony sites in western Nebraska, with sites largely along the

North and South Platte river valleys (fig. 1a) and few in prairie habitats away from these rivers, may constrain virus dispersal to a mostly linear east-west axis. Second, the flightless swallow bugs cannot be dispersed between sites unless a swallow carries them; passive dispersal by wind is not possible for bugs, unlike for mosquitoes, which genetic data suggest can be dispersed over wide areas (Venkatesan and Rasgon 2010). For these reasons, we might expect stronger spatial structuring in BCRV than in arboviruses transmitted by mosquitoes or other vagile arthropods. Although no similar studies on mosquito-borne alphaviruses are available to compare with our results, two North American flaviviruses seem to show genetic structure broadly similar to that of BCRV.

For St. Louis encephalitis virus (SLEV; Flaviviridae, *Flavivirus*), multiple genotypes were found to circulate within Harris County, Texas, an area about half the size of our study area in Nebraska (Chandler et al. 2001). Although formal analyses of SLEV spatial structure were not done, it appeared that some genotypes were restricted to local areas and others found more widely throughout the county (Chandler et al. 2001). SLEV is thought to be vectored only by mosquitoes.

Genetic analyses of West Nile virus (WNV; Flaviviridae, *Flavivirus*) showed similar spatial patterns. For the state of Connecticut, a region roughly equivalent in size to our Nebraska study area, WNV showed geographic-based clustering of haplotypes, although there was also evidence of some haplotypes occasionally occurring at sites up to 86 km apart (Anderson et al. 2001). Over a much smaller region than our study area, Bertolotti et al. (2008) found evidence for distance-limited WNV transmission in the city of Chicago, Illinois. WNV isolates were increasingly genetically diverse as the distance between sampling sites increased within an area of only 11 × 14 km. This suggests WNV is disseminated on a strongly local scale by short-range movements of *Culex* mosquitoes and resident bird species (Bertolotti et al. 2008), and is similar to the local transmission of BCRV among predominately closely spaced colony sites reported here.

In conclusion, while the spatial structure of BCRV is consistent with isolation by distance, possibly brought about by area-restricted movement of cliff swallows, local environmental conditions at certain sites (i.e., presence of sparrows) also contribute to the virus's genetic structure. This study indicates that arboviruses that are strongly associated with vagile birds (and that are even dispersed by birds) nevertheless can be genetically dissimilar over relatively short distances. Given our results for BCRV and those of Bertolotti et al. (2008) for WNV, isolation by distance may be a more common form of genetic structure in the bird-associated viruses than generally thought. Furthermore, in light of the restricted movement capacity of swallow bugs, the ecology and scale of vector movement could in some cases be as important as host movement in determining the genetic structure of arboviruses.

Acknowledgments – 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. House sparrow nest counts were provided by Valerie O'Brien. For field, laboratory, or technical assistance, we thank Susan Beckett, Jillian Blackwell, Ron Bonett, Eric Edwards, Allison Johnson, Jennifer Klaus, Valerie O'Brien, Cheryl Ormston, Nicholas Panella, Sunita Quick, Sara Robinson, Rajni Sethi, Stephanie Strickler, 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. Dany Garant and two anonymous reviewers provided helpful comments on the manuscript. This work was supported by grants from the National Institutes of Health (AI057569) to C.R.B. and the National Science Foundation (IBN-9974733, DEB-0075199, DEB-0514824) to C.R.B.

References

Anderson JF, Vossbrinck CR, Andreadis TG, Iton A, Beckwith WH III, Mayo DR (2001) A phylogenetic approach to following West Nile virus in Connecticut. Proc Natl Acad Sci USA 98:12885–12889.

Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge, Mass.

Bertolotti L, Kitron UD, Walker ED, Ruiz MO, Brawn JD, Loss SR, Hamer GL, Goldberg TL (2008) Fine-scale genetic variation and evolution of West Nile virus in a transmission "hot spot" in suburban Chicago, USA. Virology 374:381–389.

Brault AC, Armijos MV, Wheeler S, Wright S, Fang Y, Langevin S, Reisen WK (2009) Stone Lakes virus (family Togaviridae, genus *Alphavirus*), a variant of Fort Morgan virus isolated from swallow bugs (Hemiptera: Cimicidae) west of the Continental Divide. J Med Entomol 46:1203–1209.

Brown CR, Brown MB (1995) Cliff swallow (*Hirundo pyrrhonota*). In: Poole A, Gill F (eds) The birds of North America. Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, DC, no. 149.

Brown CR, Brown MB (1996) Coloniality in the cliff swallow: the effect of group size on social behavior. University of Chicago Press, Chicago, Illinois.

Brown CR, Brown MB (2004) Empirical measurement of parasite transmission between groups in a colonial bird. Ecology 85:1619–1626.

Brown CR, Brown MB (2005) Between-group transmission dynamics of the swallow bug, *Oeciacus vicarius*. J Vector Ecol 30:137–143

Brown CR, Komar N, Quick SB, Sethi RA, Panella NA, Brown MB, Pfeffer M (2001) Arbovirus infection increases with group size. Proc R Soc Lond B 268:1833–1840.

Brown CR, Brown MB, Moore AT, Komar N (2007) Bird movement predicts Buggy Creek virus infection in insect vectors. Vector Borne Zoo Dis 7:304–314.

Brown CR, Brown MB, Padhi A, Foster JE, Pfeffer M, Moore AT, Komar N (2008) Host and vector movement affects genetic diversity and spatial structure of Buggy Creek virus (Togaviridae). Mol Ecol 17:2164–2173.

Brown CR, Moore AT, Knutie SA, Komar N (2009a) Overwintering of infectious Buggy Creek virus (Togaviridae: *Alphavirus*) in *Oeciacus vicarius* (Hemiptera: Cimicidae) in North Dakota. J Med Entomol 46:391–394.

Brown CR, Padhi A, Moore AT, Brown MB, Foster JE, Pfeffer M, O'Brien VA, Komar N (2009b) Ecological divergence of two sympatric lineages of Buggy Creek virus, an arbovirus associated with birds. Ecology 90:3168–3179.

Brown CR, Moore AT, O'Brien VA, Padhi A, Knutie SA, Young GR, Komar N (2010a) Natural infection of vertebrate hosts by different lineages of Buggy Creek virus (family Togaviridae, genus *Alphavirus*). Arch Virol 155:745–749.

Brown CR, Strickler SA, Moore AT, Knutie SA, Padhi A, Brown MB, Young GR, O'Brien VA, Foster JE, Komar N (2010b) Winter ecology of Buggy Creek virus (Togaviridae, *Alphavirus*) in the central Great Plains. Vector Borne Zoo Dis 10:355–363.

Calisher CH (1994) Medically important arboviruses of the United States and Canada. Clin Microbiol Rev 7:89–116.

Calisher CH, Monath TP, Muth DJ, Lazuick JS, Trent DW, Francy DB, Kemp GE, Chandler FW (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.

Chandler LJ, Parsons R, Randle Y (2001) Multiple genotypes of St. Louis encephalitis virus (Flaviviridae: *Flavivirus*) circulate in Harris County, Texas. Am J Trop Med Hyg 64:12–19.

Cilnis MJ, Kang W, Weaver SC (1996) Genetic conservation of Highlands J viruses. Virology 218:343–351.

Crans WJ, Caccamise DF, McNelly JR (1994) Eastern equine encephalomyelitis virus in relation to the avian community of a coastal cedar swamp. J Med Entomol 31:711–728.

Davis CT, Ebel GD, Lanciotti RS, Brault AC, Guzman H, Siirin M, Lambert A, Parsons RE, Beasley DWC, Novak RJ, Elizondo-Quiroga D, Green EN, Young DS, Stark LM, Drebot MA, Artsob H, Tesh RB, Kramer LD, Barrett ADT (2005) Phylogenetic analysis of North American West Nile virus isolates, 2001–2004: evidence for the emergence of a dominant genotype. Virology 342:252–265.

Dusek RJ, McLean RG, Kramer LD, Ubico SR, Dupuis AP II, Ebel GD, Guptill SC (2009) Prevalence of West Nile virus in migratory birds during spring and fall migration. Am J Trop Med Hyg 81:1151–1158.

Forrester NL, Kenney JL, Deardorff E, Wang E, Weaver SC (2008) Western equine encephalitis submergence: lack of evidence for a decline in virus virulence. Virology 380:170–172.

Gubler DJ (2002) The global emergence/resurgence of arboviral diseases as public health problems. Arch Med Res 33:330–342.

Holmes EC (2009) The evolution and emergence of RNA viruses. Oxford University Press, Oxford, UK.

Hopla CE, Francy DB, Calisher CH, Lazuick JS (1993) Relationship of cliff swallows, ectoparasites, and an alphavirus in west-central Oklahoma. J Med Entomol 30:267–272.

Howard JJ, Oliver J, Grayson MA (2004) Antibody response of wild birds to natural infection with Alphaviruses. J Med Entomol 41:1090–1103.

Kondig JP, Turell MJ, Lee JS, O'Guinn ML, Wasieloski LP 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.

Kosakovsky Pond SL, Frost SDW (2005) DATAMONKEY: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21:2531–2533.

Kramer LD, Fallah HM (1999) Genetic variation among isolates of western equine encephalomyelitis virus from California. Am J Trop Med Hyg 60:708–713.

Kramer LD, Presser SB, Hardy JL, Jackson AO (1997) Genotypic and phenotypic variation of selected Saint Louis encephalitis viral strains isolated in California. Am J Trop Med Hyg 57:222–229.

Kramer LD, Styer LM, Ebel GD (2008) A global perspective on the epidemiology of West Nile virus. Ann Rev Entomol 53:61–81.

Lord RD, Calisher CH (1970) Further evidence of southward transport of arboviruses by migratory birds. Am J Epidemiol 92:73–78.

Lowther PE, Cink CL (1992) House sparrow (*Passer domesticus*). In: Poole A, Stettenheim P, Gill F (eds) The birds of North America. Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington, DC, no. 12.

Loye JE, Carroll SP (1991) Nest ectoparasite abundance and cliff swallow colony site selection, nestling development, and departure time. In: Loye JE, Zuk M (eds) Bird-parasite interactions: ecology, evolution and behaviour. Oxford University Press, Oxford, UK, pp 222–241.

Monath TP (1980) St. Louis encephalitis. American Public Health Association, Washington, DC.

Moore AT, Edwards EA, Brown MB, Komar N, Brown CR (2007) Ecological correlates of Buggy Creek virus infection in *Oeciacus vicarius*, southwestern Nebraska, 2004. J Med Entomol 44:42–49.

Moya A, Holmes EC, Gonzalez-Candelas F (2004) The population genetics and evolutionary epidemiology of RNA viruses. Nature Rev 2:1–10.

Navaratnarajah CK, Kuhn RJ (2007) Functional characterization of the Sindbis virus E2 glycoprotein by transposon linker-insertion mutagenesis. Virology 363:134–147.

Norder H, Lundström JO, Kozuch O, Magnius LO (1996) Genetic relatedness of Sindbis virus strains from Europe, Middle East, and Africa. Virology 222:440–445.

O'Brien VA (2009) Ecological interactions between arboviruses and their avian hosts. Ph.D. thesis. University of Tulsa, Tulsa, Oklahoma.

O'Brien VA, Moore AT, Young GR, Komar N, Reisen WK, Brown CR (2010a) An enzootic vector-borne virus is amplified at epizootic levels by an invasive avian host. Proc R Soc B. doi: 10.1098/rspb.2010.1098 (in press).

O'Brien VA, Meteyer CU, Ip HS, Long RR, Brown CR (2010b) Pathology and virus detection in tissues of nestling house sparrows naturally infected with Buggy Creek virus (Togaviridae). J Wildl Dis 46:23–32.

Oberste MS, Weaver SC, Watts DM, Smith JF (1998) Identification and genetic analysis of Panama-genotype Venezuelan equine encephalitis virus subtype ID in Peru. Am J Trop Med Hyg 58:41–46.

Padhi A, Moore AT, Brown MB, Foster JE, Pfeffer M, Gaines KP, O'Brien VA, Strickler SA, Johnson AE, Brown CR (2008) Phylogeographic structure and evolutionary history of two Buggy Creek virus lineages in the western Great Plains of North America. J Gen Virol 89:2122–2131.

Pfeffer M, Foster JE, Edwards EA, Brown MB, Komar N, Brown CR (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.

Posada D (2004) COLLAPSE: describing haplotypes from sequence alignments, version 1.2. University of Vigo, Vigo, Spain.

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

Powers AM, Brault AC, Shirako Y, Strauss EG, Kang W, Strauss JH, Weaver SC (2001) Evolutionary relationships and systematics of the alphaviruses. J Virol 75:10118–10131.

Rannala BH (1995) Demography and genetic structure in island populations. Ph.D. thesis. Yale University, New Haven, Connecticut.

Real LA, Henderson JC, Biek R, Snaman J, Jack TL, Childs JE, Stahl E, Waller L, Tinline R, Nadin-Davis S (2005) Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. Proc Natl Acad Sci USA 102:12107–12111.

Reeves WC (1990) Epidemiology and control of mosquito-borne arboviruses in California, 1943–1987. California Mosquito Vector Control Association, Sacramento, California.

Reisen WK, Lothrop HD, Chiles RE, Cusack R, Green EGN, Fang Y, Kensington M (2002) Persistence and amplification of St. Louis encephalitis virus in the Coachella Valley of California, 2000–2001. J Med Entomol 39:793–805.

Reisen WK, Fang Y, Brault AC (2008) Limited interdecadal variation in mosquito (Diptera: Culicidae) and avian host competence for western equine encephalomyelitis virus (Togaviridae: *Alphavirus*). Am J Trop Med Hyg 78:681–686.

Robbins CS (1973) Introduction, spread, and present abundance of the house sparrow in North America. Ornithol Monogr 14:3–9.

Sammels LM, Lindsay MD, Poidinger M, Coelen RJ, Mackenzie JS (1999) Geographic distribution and evolution of Sindbis virus in Australia. J Gen Virol 80:739–748.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: a software for population genetics data analysis, version 3.1. Genetics and Biometry Laboratory. University of Geneva, Geneva, Switzerland.

Scott TW, Bowen GS, Monath TP (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.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264–279.

Smith GC, Eads RB (1978) Field observations on the cliff swallow, *Petrochelidon pyrrhonota* (Vieillot), and the swallow bug, *Oeciacus vicarius* Horvath. J Wash Acad Sci 68:23–26.

Stamm DD, Newman RJ (1963) Evidence of southward transport of arboviruses from the US by migratory birds. Ann Microbiol 11:123–133.

Strauss JH, Strauss EG (1994) The alphaviruses: gene expression, replication, and evolution. Microbiol Rev 58:491–562.

Swofford DL (2002) PAUP*: phylogenetic analyses using parsimony (*and other methods) 4.0 b. Sinauer, Sunderlund, MA.

van den Hurk AF, Ritchie SA, Mackenzie JS (2009) Ecology and geographical expansion of Japanese encephalitis virus. Ann Rev Entomol 54:17–35.

Venkatesan M, Rasgon JL (2010) Population genetic data suggest a role for mosquito-mediated dispersal of West Nile virus across the western United States. Mol Ecol 19:1573–1584.

Weaver SC, Hagenbaugh A, Bellew LA, Gousset L, Mallampalli V, Holland JJ, Scott TW (1994) Evolution of alphaviruses in the eastern equine encephalomyelitis complex. J Virol 68:158–169.

Weaver SC, Kang W, Shirako Y, Rumenapf T, Strauss EG, Strauss JH (1997) Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. J Virol 71:613–623.

Wright S (1943) Isolation by distance. Genetics 28:139–156.