Phylogenetic Analysis of Buggy Creek Virus: Evidence for Multiple Clades in the Western Great Plains, United States of America

Martin Pfeffer, Jerome E. Foster, Eric A. Edwards, Mary Bomberger Brown, Nicholas Komar, and Charles R. Brown*

Bundeswehr Institute of Microbiology, Neuherbergstrasse 11, 80937 Munich, Germany; Department of Biological Sciences, University of Tulsa, Tulsa, Oklahoma 74104; and Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention, P.O. Box 2087, Fort Collins, Colorado 80522

We present the first detailed phylogenetic analysis of Buggy Creek virus (BCRV), a poorly known alphavirus with transmission cycles involving a cimicid swallow bug (Oeciacus vicarius) vector and cliff swallows (Petrochelidon pyrrhonota) and house sparrows (Passer domesticus) as the principal avian hosts. Nucleotide sequences of a 2,075-bp viral envelope glycoprotein-coding region, covering the entire PE2 gene, were determined for 33 BCRV isolates taken from swallow bugs at cliff swallow colonies in Nebraska and Colorado in the summer of 2001 and were compared with the corresponding region of BCRV isolates collected from Oklahoma in the 1980s. We also analyzed isolates of the closely related Fort Morgan virus (FMV) collected from Colorado in the 1970s. Phylogenetic analysis indicated that BCRV falls into the western equine encephalomyelitis complex of alphaviruses, in agreement with antigenic results and a previous alphavirus phylogeny based on the E1 coding region. We found four distinct BCRV/FMV clades, one each unique to Nebraska, Colorado, and Oklahoma and one containing isolates from both Nebraska and Colorado. BCRV isolates within the two clades from Nebraska showed 5.7 to 6.2% nucleotide divergence and 0.7 to 1.9% amino acid divergence, and within these clades, we found multiple subclades. Nebraska subclades tended to be confined to one or a few cliff swallow colonies that were close to each other in space, although in some cases, near-identical isolates were detected at sites up to 123 km apart. Viral gene flow occurs when cliff swallows move (bugs) between colony sites, and the genetic structure of BCRV may reflect the limited dispersal abilities of its insect vector.

Buggy Creek virus (BCRV) is a poorly known alphavirus (Togaviridae) that is part of the western equine encephalomyelitis antigenic complex, which includes other alphaviruses such as Sindbis virus (SINV), western equine encephalomyelitis virus (WEEV), Highlands J virus (HVJ), and Fort Morgan virus (FMV) (9, 17, 29, 37). Like other members of the WEEV complex, BCRV is a natural recombinant virus derived from Old World SINV and New World eastern equine encephalomyelitis virus (EEEV) (29, 41). BCRV (and the closely related FMV) is apparently widely distributed in North America, having been found in Texas, Oklahoma, Nebraska, Colorado, South Dakota, and Washington state (7, 8, 17). It was first isolated in 1980 at Buggy Creek in Grady County, Oklahoma (17, 22), although the ecologically very similar FMV was discovered in the 1970s in Colorado (8, 14).

Like most alphaviruses, BCRV is a vector-borne virus with transmission cycles involving an invertebrate vector and a vertebrate host. BCRV is unusual among alphaviruses, however, in that it is routinely vectored by an invertebrate other than mosquitoes: this virus is commonly associated with the cimicid swallow bug (Hemiptera: Cimicidae: Oeciacus vicarius). The bug is an ectoparasite of the colonially nesting cliff swallow (Petrochelidon pyrrhonota) and, to a lesser extent, the introduced house sparrow (Passer domesticus), and both birds serve as hosts to BCRV (14, 17, 34). FMV is also associated with swallow bugs, cliff swallows, and house sparrows. Because the wingless swallow bugs are sedentary and confined to active and inactive cliff swallow nests during most of the year (5, 21), the spatial foci for virus occurrence are predictable and allow the study of how genetic divergence of BCRV potentially varies from site to site over relatively small spatial scales without making assumptions as to the vectors’ geographic origin. There is a close antigenic relationship between BCRV and the other WEEV complex alphaviruses (8, 9, 14, 17) that have medical and veterinary significance, and consequently, the study of how BCRV may undergo spatial structuring and evolutionary change may provide important information on this virus group that could not be obtained in studies where volant arthropods (e.g., mosquitoes) are the primary vectors.

To understand more about the evolution of BCRV, we performed a phylogenetic analysis of BCRV isolates collected from bugs in cliff swallow colonies in (primarily) southwestern Nebraska in summer 2001. Our analysis also included the first described strains of BCRV, collected from Oklahoma in the 1980s, and FMV isolates from Colorado in the 1970s. One of our objectives was to determine the phylogenetic relationship of FMV to BCRV. These viruses share ecological, antigenic, and nucleotide sequence similarities (9, 29; R. M. Kinney and M. Pfeffer, unpublished data) and may be strains of the same virus.

We based our phylogeny on the majority of the structural protein genes encoded by the viral subgenomic 26S RNA, including the entire PE2 gene (which encodes the E2 and E3 viral envelope proteins). Most of the major neutralizing
epitopes are located on the E2 glycoprotein (29, 37). Under the pressure of a vertebrate immune system, mutations on this gene may help the virus escape destruction. Thus, if there are functional differences among virus isolates, they are likely to be expressed in the E2 gene, and consequently, this region should be highly informative for the study of virus diversification and evolution. This work represents the first phylogenetic survey of BCRV within a population, and from the results, we describe four separate clades of BCRV in the swallow bug vectors of the western Great Plains. We also analyze how genetic similarity of the isolates is related to colony of origin. The results should help us understand patterns of evolutionary change in this alphavirus and how these changes may be related to ecological variables.

MATERIALS AND METHODS

Study organisms. The swallow bug is an ectoparasite primarily of cliff swallows and is found throughout the bird's wide geographic range (4, 24). 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 hematophagous, feeding on the birds mostly at night, and they travel on the adult birds relatively rarely (5, 6, 12, 21). Infestations can reach up to 2,600 bugs per nest, and the bugs affect many aspects of cliff swallow life history (2, 3, 5, 10, 23). Swallow bugs are long-lived ectoparasites that begin to reproduce as soon as they feed in the spring. The bugs have access to hosts only when cliff swallows occupy a colony site or reuse existing nests. The birds do not use all of the colony sites in a given year (5), and the bugs seem to be adapted to withstanding long periods of host absence, in some cases persisting at a site not used by cliff swallows for up to perhaps four consecutive years (21, 23, 30, 35). The bugs appear capable of parasitizing introduced house sparrows (Passer domesticus) that occupy nests in some cliff swallow colonies (7, 17; C. Brown et al., personal observation).

Cliff swallows are highly colonial passerines that breed throughout most of western North America (4). 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 outsides of cliff swallow nests. Bugs generally either clustered just inside the tubular entrances of the nests in a relatively dense mass or were distributed in typically lower densities across the bottom and sides of the nests and below the entrance. We brushed bugs from nests and sorted them into pools of 100 individuals while alive. Pools were frozen immediately at ~70°C. Multiple pools often came from bugs taken from a single nest.

Virus isolation and sequencing. Pools of 100 bugs were macerated by mortar and pestle and suspended in 2.0 ml of BA-1, a growth medium containing M-199 Hanks' salts, 1% bovine serum albumin, 0.05 M Tris-HCl (pH 7.5), 0.35 g/liter sodium bicarbonate, 100 U/ml penicillin, 100 µg/ml streptomycin, 1 µg/ml Fungizone ( Gibco-BRL, Gaithersburg, Md.), and sterile distilled water. Homogenates were clarified by centrifugation and passed through 0.45-

<table>
<thead>
<tr>
<th>Primer</th>
<th>Direction</th>
<th>Sequencea</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-8187</td>
<td>+</td>
<td>TGGCAYCAYGNGCNGTNCCARTA</td>
<td>Amplification/sequencing</td>
</tr>
<tr>
<td>cec-10262</td>
<td>–</td>
<td>RTCRCAARACANTNRCNCCNCCCACAT</td>
<td>Amplification/sequencing</td>
</tr>
<tr>
<td>8170</td>
<td>+</td>
<td>GACTTTTGGGACGAGTC</td>
<td>Sequencing</td>
</tr>
<tr>
<td>8171</td>
<td>–</td>
<td>AATGGCTATTATGACTCC</td>
<td>Sequencing</td>
</tr>
<tr>
<td>8256</td>
<td>–</td>
<td>GATGIGAGGGTIAAGGTG</td>
<td>Sequencing</td>
</tr>
<tr>
<td>8257</td>
<td>+</td>
<td>TGATTACCCCTCCTATGGC</td>
<td>Sequencing</td>
</tr>
<tr>
<td>8258</td>
<td>–</td>
<td>ATCCGACTCTGGCCCAAAC</td>
<td>Sequencing</td>
</tr>
</tbody>
</table>

a Y = C or T; R = A or G; N = A, G, C, or T.

Viral RNA was extracted from 140 µl of the infectious precleared supernatant of the second Vero cell passage using the RNeasy extraction kit as recommended by the manufacturer (Qiagen, Hilden, Germany). Five microliters of the eluted RNA suspension was used as a template in an alphavirus reverse transcription-PCR (RT-PCR) (27) to amplify a 2,075-bp region within the subgenomic 26S RNA spanning a 192-bp segment of the 3′ end of the capsid, the entire E3 (180 bp), E2 (1269 bp), and 6K (165 bp) genes, and the first 269 bp of the E1 gene. Details of the RT-PCR have been described previously (19). Primers used to amplify the four separate clades of BCRV in the swallow bug vectors in the western Great Plains are listed in Table 1. Sequences were resolved on a 3130-Avant model genetic analyzer (Applied Biosystems, Darmstadt, Germany).
Sequences were aligned against the corresponding region in a 1981 BCRV reference sequence (strain 81V1822 [GenBank accession no. AF339474]), and fragments were combined for a given isolate using SeqMan 6.1 (DNAStar; Lasergene) to obtain a continuous nucleotide sequence for each sample. Isolates from Nebraska are labeled here as “colony number-sample number-year,” and thus, the site (colony) of sampling is indicated by the first two digits of the isolate label.

Phylogenetic analysis. Sequences were aligned in MEGALIGN v. 6.1.2 (Lasergene) using the Clustal algorithm (15, 16). An initial neighbor-joining (NJ) tree was constructed with our 33 BCRV isolates and other representative alphaviruses available at PubMed (GenBank accession numbers AF339474 to AF339488, AF093102, AF093103, AF103728, AF079454, AF079456, AF214040, AF229608, AF429428, AF448539, AJ316244, AJ316246, AY112987, AY604236, DQ189086, DQ241304, K00046, M20162, M69094, M69205, and NC001924). Also included in the NJ tree was the corresponding region from 10 additional unpublished BCRV (strains 81V1823-6 and 84S217) and FMV (75V7468, 77V654, CM4-110, CM4-368, and CM4-976) isolates. A maximum-likelihood (ML) tree of all isolates that grouped with the BCRV isolates was generated using PAUP 4.0b10 (38). We based tree calculations on models of sequence evolution whose parameters were estimated by maximum likelihood using Mod...
RESULTS

We analyzed sequences from 32 BCRV isolates collected from eight different cliff swallow colonies in southwestern Nebraska in 2001 and 1 isolate from a colony in northeastern Colorado in 2001 (Fig. 1). The NJ phylogenetic tree shows that all the sequences analyzed belong to the WEEV complex and include the six FMV sequences (Fig. 1).

TABLE 2. Amino acid changes in a 2,075-bp region of the Buggy Creek virus genome

<table>
<thead>
<tr>
<th>Virus clade</th>
<th>Amino acid at different positions†</th>
<th>E3</th>
<th>E2</th>
<th>6K</th>
<th>E1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>18</td>
<td>57</td>
<td>60</td>
</tr>
<tr>
<td>Okl</td>
<td>D</td>
<td>Q</td>
<td>D</td>
<td>K</td>
<td>D</td>
</tr>
<tr>
<td>Col</td>
<td>N</td>
<td>G</td>
<td>Q</td>
<td>G</td>
<td>T</td>
</tr>
<tr>
<td>NebA-C</td>
<td>N</td>
<td>Q</td>
<td>G</td>
<td>K</td>
<td>L</td>
</tr>
<tr>
<td>NebA-N1</td>
<td>N</td>
<td>Q</td>
<td>G</td>
<td>K</td>
<td>L</td>
</tr>
<tr>
<td>NebA-N2</td>
<td>N</td>
<td>Q</td>
<td>G</td>
<td>L</td>
<td>—</td>
</tr>
<tr>
<td>NebA-N3</td>
<td>N</td>
<td>Q</td>
<td>G</td>
<td>L</td>
<td>—</td>
</tr>
<tr>
<td>NebB-N1</td>
<td>—</td>
<td>Q</td>
<td>G</td>
<td>L</td>
<td>—</td>
</tr>
<tr>
<td>NebB-N2</td>
<td>G</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>NebB-N3</td>
<td>G</td>
<td>R</td>
<td>Q</td>
<td>L</td>
<td>M</td>
</tr>
</tbody>
</table>

† All isolates of each clade (Fig. 2) are compared to the Oklahoma clade; residues identical to those present in Okl are represented by dashes.

eltest 3.07 PPC (28). We assessed the reliability of the tree topology by bootstrapping the data set 1,000 times to generate NJ trees. Internal nodes of particular interest were labeled with percent times obtained, if greater than 50.

Each BCRV isolate is actually a sample of multiple and potentially variable virus particles within the host (11, 40); in this case, a “sample” refers to a pool of bugs, and an isolate from a given sample represents the dominant genotype present. In one case, an isolate had evidence of a sequence polymorphism, of bugs, and an isolate from a given sample represents the dominant genotype present. In one case, an isolate had evidence of a sequence polymorphism, indicating the presence of multiple viral genotypes, and it was excluded from the analysis. We use the term “clade” to describe groups of similar and presumably related individuals that share a common ancestor (monophyletic) when that relationship is supported at least 75% of the time by bootstrap analysis and exhibit similar amino acid substitutions (equivalent in some usage to “lineage”).

We use the term “subclade” to describe individuals within a given clade that form subgroups that are also supported at least 75% of the time by bootstrap analysis and amino acid similarity. Both terms may refer to groups of isolates that share common silent and nonsilent nucleotide substitutions. We use the term gene flow to refer to the movement of viral genes and the movement of BCRV within the bird colony sites containing vectors and hosts.

Nucleotide sequence accession numbers. All sequences generated from this study can be accessed in GenBank (accession numbers DQ451557 to DQ451559).

We analyzed sequences from 32 BCRV isolates collected from eight different cliff swallow colonies in southwestern Nebraska in 2001 and 1 isolate from a colony in northeastern Colorado in 2001 (Fig. 1). The NJ phylogenetic tree shows that all the sequences analyzed belong to the WEEV complex and include the six FMV sequences (Fig. 1). There is close association with other alphaviruses within that group, such as HJV, SINV, WEEV, and FMV. The BCRV sequences are a distinct group within the WEEV complex and include the six FMV sequences (Fig. 1).

Among the BCRV and FMV sequences, there were (in total) 137 nucleotide substitutions and 15 amino acid residues that underwent a change relative to the earliest BCRV isolates from Oklahoma (Table 2). There were 2 amino acid changes in the E3 gene, 10 in the E2 gene, 2 in the 6K gene, and 1 in the fragment of the E1 gene examined. The capsid region had no changes in the fragment used in this analysis.

Based on the ML tree of the 2,075-bp region in the structural genes that we examined, there were four BCRV clades (Fig. 2). One is a clade composed solely of Oklahoma isolates from 1981 to 1983, designated clade Okl (Fig. 2). There were two distinct BCRV clades (designated NebA and NebB) circulating in swallow bugs during 2001 at the Nebraska study site (Fig. 2). These two clades both share two amino acid changes at residues E2-62 (K→Q) and E2-243 (I→L), which distinguish them from the Oklahoma isolates (Table 2 and Fig. 2). The substitution at E2-62 is nonconservative, with a change from a basic residue (lysine) to a neutral but polar one (glutamine), while the change at E2-243 is conservative. These substitutions are also shared with the FMV isolates (all from Colorado) and the only BCRV isolate from Colorado (COL 2001) (Table 2 and Fig. 2).

In addition to these two substitutions, the four subclades of NebA (NebA-N1, NebA-N2, NebA-N3, and NebA-C) (Fig. 2) are characterized by changes at E3-17 (D→N) and E2-119 (D→G), both of which are nonconservative. The three distinct Nebraska subclades within NebA (N1 to N3) are defined by both silent and nonsilent changes. NebA-N1 is made up solely of isolates from one colony (site NN) and has two distinct conservative changes at E2-60 (D→N) and E2-222 (R→K). NebA-N2 is widespread with respect to sampling and has a distinct conservative change at E2-378 (A→V). Isolate 88-02-01 lies outside this and the other subclades within NebA, lacking any of the distinct amino acid changes that define these clades. Based on the ML tree, it is most closely related to NebA-N2, with moderate bootstrap support of 68 (Fig. 2). Subclade NebA-N3 has a unique mutation at E2-57 (D→N) and is comprised mostly of isolates from one colony (site 03). It also has a mutation at E2-218 (M→T), which it shares with only subclade NebA-C (Table 2). Subclade NebA-C consists of FMV isolates from Colorado. It is similar to isolates taken from Nebraska, and it shares the majority of its amino acid changes with the other NebA subclades (Table 2). NebA-C is distinct only by the presence of several silent mutations and a change at E2-218 (M→T) shared with subclade NebA-N3.

NebB, the other Nebraskan clade, is the most distinct of the four. Its isolates have two mutations that separate them from the other sequences in the tree, E2-372 (T→A) and 6K-33 (L→F). The first change is a nonconservative change from a polar to a nonpolar residue, while the second is conservative. As in clade NebA, there are distinct subgroups within NebB (Fig. 2). Subclade NebB-N1 is comprised of isolates solely from one colony (site 30). NebB-N2 and NebB-N3 are widespread with respect to colonies and share two common amino acid changes at E2-270 (L→M) and 6K-53 (V→A), which are not functionally significant.
The fourth clade consisted of one BCRV isolate (COL 2001) and one FMV isolate (75V7468), both collected in Colorado, and was designated Col. This clade was more closely related to clade NebA than to any other clade. Its position in the tree suggests that the other Colorado isolates (NebA-C) and the three NebA subclades from Nebraska have more recently shared a common ancestor than any have with the Col clade (Fig. 2). In addition to a distinct amino acid change at E3-17 (D→N), which it shares with clade NebA, the Col clade also has a unique mutation at E2-60 (D→G) that distinguishes it from the other clades (Table 2).

A number of isolates also had individual mutations not seen elsewhere: E2-141 (Q→R) in FMV 75V7468, E2-143 (P→R) in isolates 82-01-01 and 82-06-01, E2-284 (T→A) and E2-344 (A→T) in isolate NN-20-01, E2-288 (T→A) in isolate 95-02-01, and E2-422 (H→R) in isolate COL 2001.

Overall, Nebraska isolates from the NebA and NebB clades showed 5.7 to 6.2% interclade divergence at the nucleotide
level and 0.7 to 1.9% divergence at the amino acid level. Both clades appear to be widespread within the Nebraska study area and coexist at some sites. Five colonies (sites 30, 82, 88, 95, and NN) had isolates in both branches of the phylogeny (Fig. 2). Within the subclades, however, there was greater spatial structuring: NebA-N1 and NebB-N1 each consisted of isolates all from the same colony, NebA-N2 consisted of isolates from two colonies that were only 4.9 km apart, NebA-N3 consisted of nine isolates from the same colony and another colony (site 31) from a site 7.2 km away, and two of the four isolates in NebB-N2 were from the same colony (Fig. 2). However, there was also similarity in virus isolates taken from the two colonies on the extreme eastern and western edges of our study area (sites 82 and 30). These colonies were approximately 123 km apart, and they each contained representatives of clades NebA-N3 and NebB-N2 (Fig. 2).

**DISCUSSION**

This report is the first detailed phylogenetic analysis of BCRV. The NJ tree constructed from a 2,075-bp region encompassing a significant portion of the E2 viral envelope glycoprotein-coding region resulted in an accurate phylogenetic categorization of the virus as a WEEV complex member, similar to previous analyses using a portion of the E1 gene (29, 41) or the E2 gene (25). This confirms the suitability of the region that we used for phylogenetic analysis and supports that all our isolates were BCRV and not another, closely related WEEV complex virus.

The placement of the FMV sequence within the BCRV branch indicates that these two virus are the most closely related to each other (compared to other WEEV complex viruses). This phylogenetic analysis (also see reference 29) demonstrates that Buggy Creek virus and Fort Morgan virus are strains of the same virus, a conclusion that is also consistent with the close ecological similarity between the two. The FMV strains available to us are more evolutionarily derived than the original BCRV strains from Oklahoma reported previously by Hopla et al. (17) (Fig. 2).

Several of the amino acid changes observed in the BCRV isolates can be mapped to regions with pathogenic and immunogenic properties in related viruses. The change at E2-62 (K→Q), common to clades NebA, NebB, and Col, is a change from a basic to a polar residue, and research with monoclonal antibodies has shown this residue to be the component of a pathogenesis site in SINV associated with virulence and host cell penetration (26, 32). E2-243 (I→L) is also common to all three non-Oklahoma clades and is found within an antigenic determinant of the E2 ectodomain in Ross River virus (39) and Semliki Forest virus (13). The changes at E2-372 (T→A) and 6K-33 (L→F) are unique to clade NebB, and the former mutation falls within the transmembrane domain of the E2 glycoprotein (31, 36). Whether any of these changes influence the fitness of the virus is unknown. However, the finding that they are linked with distinct clades found in nature indicates that some form of purifying or positive selection might be operating, and these data would be appropriate for comparative functional genome analyses. The nucleotide and amino acid divergences between the clades are within or close to the range of differences seen among recognized subtypes of other alphaviruses (e.g., SINV and EEEV) (25, 41).

Clades NebA and NebB were widely distributed in our Nebraska study area, with near-identical isolates detected in cliff swallow colonies up to 123 km apart. However, the BCRV isolates also showed a relatively high degree of spatial structure, with some of the subclades confined to one or a few colonies that were quite close to each other in space. Interpreting these spatial patterns requires a knowledge of the extent to which viral gene flow occurs between sites. Although the transmission dynamics of BCRV are poorly understood, the virus presumably is moved from site to site whenever an infected bug moves and possibly also when an infected cliff swallow or house sparrow transfers between colony sites (assuming that birds maintain viremias of sufficient titers to infect bugs that feed on them). Bug movement correlates strongly with cliff swallow movement, because bugs disperse only by clinging to the feet of a swallow that is moving between sites (6). The relatively high degree of spatial structure among BCRV subclades probably reflects the birds’ greater likelihood of moving between closely spaced colonies (5) and the fact that successful dispersal by a bug is more likely when the distance moved (and thus the time spent clinging to a bird’s feet) is short. Formal analyses of the genetic similarity of isolates in relation to colony spatial position await a larger sample of sequences from additional years.

If viral gene flow is largely dependent on cliff swallows’ moving of bugs, one would predict some genetic divergence of the virus (i) over distances greater than the typical movement radius of cliff swallows selecting colonies and (ii) when gaps in the distribution of swallow colony sites, due to unsuitability of habitat, serve to isolate colonies in one area from bird movement from another area. Consistent with this, we found that the FMV isolates collected in Colorado approximately 250 km from the western edge of the Nebraska study area occupied basal positions to the NebA branch (Fig. 2) and showed some amino acid substitutions not seen in the Nebraska samples (Table 2). Similarly, the Oklahoma BCRV isolates from 1981 to 1983, collected approximately 750 km south of the Nebraska study area, showed some amino acid divergence from both the Nebraska and Colorado clades (Table 2). Although the Oklahoma and (most of the) Colorado results are potentially confounded by different sampling years, the distances from Nebraska to both the Colorado and Oklahoma sites are well beyond the typical movement radius of a cliff swallow when assessing and choosing colony sites to occupy (and thus when potentially moving infected bugs) (5). This is supported by the more recently isolated Colorado isolate (COL 2001), which, although collected 25 years after the “Fort Morgan” viruses from the CDC’s reference collection, was more similar to those isolates than to recent ones from Nebraska. Because cliff swallow breeding distribution is typically concentrated along east-west-flowing rivers in most of the Great Plains (C. Brown, personal observation), the distribution of colonies is not continuous from the Nebraska study area to the Colorado or Oklahoma sampling sites. Thus, the genetic divergence of these isolates is probably attributable in part to reduced viral gene flow between these areas.

An unresolved question is whether the two Nebraska BCRV clades are functionally different in terms of effects on hosts,
vectors, or ecology (e.g., overwintering persistence). More samples and additional studies will be necessary to address these questions. However, we note that clade NebB was not found in the two largest cliff swallow colonies sampled (660 and 240 nests from sites 03 and 88, respectively). The other colonies, including all the ones with clade NebB, varied in size between 1 and 90 nests. Colony size could affect virus prevalence from year to year (and thus whether a particular clade is detected) because the number of active nests correlates directly with bug population size (7), which may affect the likelihood of local virus persistence over the winter. Perhaps one clade is more susceptible to extreme seasonal fluctuations in bug population size that are brought about by small colonies being less likely to be occupied by birds in consecutive years (5).

The relatively high degree of spatial structure among the BCRV subclades detected in this study suggests that WEEV complex alphaviruses have the capacity to show sequence variation that can be associated with particular geographic locations. If so, this may make it possible to identify a given isolate's probable geographic origin simply by sequencing and may assist in determining the movement characteristics and the source of epizootics in the medically important WEEV complex viruses. Documentation of nucleotide variation among the spatially distinct BCRV clades could be especially important, given that mutations at only one or few nucleotide positions in phylogenetically predicted viral mutation. Proc. Natl. Acad. Sci. USA 103: 4994–4999.


