

RESEARCH ARTICLE

Plant community responses to bison reintroduction on the Northern Great Plains, United States: a test of the keystone species concept

Nicholas A. McMillan^{1,2,3} , Kyran E. Kunkel⁴, Donald L. Hagan¹, David S. Jachowski¹

Keystone species restoration, or the restoration of species whose effect on an ecosystem is much greater than their abundance would suggest, is a central justification for many wildlife reintroduction projects globally. Following restoration, plains bison (*Bison bison* L.) have been identified as a keystone species in the tallgrass prairie ecoregion, but we know of no research to document similar effects in the mixed-grass prairie where restoration efforts are ongoing. This study addresses whether Northern Great Plains (NGP) mixed-grass prairie plant communities exhibit traits consistent with four central keystone effects documented for bison in the tallgrass prairie. We collected species composition, diversity, abundance, bare ground cover, and plant height data in three treatments: where livestock (*Bos taurus* L.) continuously grazed, livestock were removed for 10 years, and bison have been introduced and resident for 10 years. We observed mixed support for bison acting as keystone species in this system. Supporting the keystone role of bison, we observed higher species richness and compositional heterogeneity (β -diversity) in the bison treatment than either the livestock retention or livestock removal treatments. However, we observed comparable forb, bare ground, and plant height heterogeneity between bison-restored sites and sites where livestock were retained, contradicting reported keystone effects in other systems. Our results suggest that after 10 years of being restored, bison partially fulfill their role as a keystone species in the mixed-grass prairie, and we encourage continued long-term data collection to evaluate their influence in the NGP.

Key words: Bison, keystone species restoration, livestock removal, mixed-grass prairie, Northern Great Plains, plant communities

Implications for Practice

- Short-term bison reintroduction (<10 years) may not lead to sweeping increases in plant diversity or structural heterogeneity in the Northern Great Plains of Montana, but may lead to higher plant species richness and compositional heterogeneity compared to short-term cattle-removal (<10 years) and long-term cattle-grazing (>70 years).
- If bison are to be reintroduced across the Great Plains based on expected keystone effects, it is imperative that managers utilize long-term, robust monitoring techniques to track their influence on rangeland plant communities.

Introduction

The restoration of keystone species, or those species that influence ecosystems more than their abundance would suggest (Paine 1969), has been posited as essential for the maintenance and management of heightened biodiversity and ecosystem function (Mills et al. 1993). The keystone species concept has been expanded and employed by several studies to explore how best to recognize and categorize various keystone species. Thus, keystone species can be lumped into five generally accepted categories: keystone predators, prey, plants, links, or modifiers

(Mills et al. 1993; Power et al. 1996). Herbivores operating as keystone species are typically labeled keystone modifier species; defined as a species whose behaviors alter ecosystem structure and composition, such that removal of that species results in reduced overall biodiversity through shifts in the structure and diversity of ecological communities (Mills et al. 1993).

The restoration of keystone (or perhaps foundational; see Soulé et al. 2003) interactions is a central theme guiding restoration and reintroduction projects to generate and sustain the habitat structure needed for high native species abundance and diversity (Conway 1989). For example, restoration of the gray wolf (*Canis lupus* L.)—a keystone predator—to the Greater Yellowstone Ecosystem (GYE) is hypothesized to have increased biodiversity and improved habitat structure by reducing the browsing influence of elk (*Cervus elaphus*

Author contributions: NAM, DSJ, KEK conceived and designed the research; NAM performed the research and analyses of the data; DSJ contributed equipment and funding; all authors wrote and edited the manuscript.

¹Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC 29634, U.S.A.

²Present address: 008C Agricultural Hall, Oklahoma State University, Stillwater, OK 74078

³Address correspondence to N. A. McMillan, email namcmil@g.clemson.edu

⁴American Prairie Reserve, 7 East Beall Street No. 100, Bozeman, MT 59715, U.S.A.

ssp. *canadensis* Erxleben) that would otherwise limit plant abundance and diversity (Ripple & Beschta 2012). Similarly, the reintroduction or rewilding of large herbivores, namely those considered keystone species (e.g. American plains bison [*Bison bison*]), has recently become a key rationale globally for many restoration projects attempting to increase biodiversity and ecosystem function (Seddon et al. 2014).

The American plains bison is widely viewed by ecologists to be a keystone species that historically had a major influence shaping North American grassland ecosystems (Knapp et al. 1999; Freese et al. 2007). American plains bison populations declined from nearly 60 million individuals to near extinction in the nineteenth century due to over-hunting (Hornaday 1889; Freese et al. 2007). Grazing disturbance by bison paired with other disturbance regimes (e.g. fire) have been shown to generate increased biodiversity in the tallgrass prairie ecosystem (Fuhlendorf & Engle 2001; Towne et al. 2005; Allred et al. 2011). In addition, bison rubbing and wallowing behaviors directly alter habitat structure (Knapp et al. 1999). For example, in the tallgrass prairie ecosystem (hereafter referred to as tallgrass prairie) where woody encroachment potential is high, bison typically alter forest succession through rubbing and horning behaviors on woody vegetation (Coppedge & Shaw 1997). Bison patch-grazing and wallowing behaviors can also generate high levels of structural (i.e. plant height or growth habit) and compositional heterogeneity (i.e. β -diversity, or the amount of compositional variation from community to community), collectively viewed as increased grassland heterogeneity (Fuhlendorf & Engle 2001; Toombs et al. 2010). Declines in grassland heterogeneity have been linked to declines in avian, amphibian, reptilian, insect, and botanical diversity (Knapp et al. 1999; Koerner et al. 2014). In addition, bison grazing can increase heterogeneity in plant functional group richness (Knapp et al. 1999). Bison grazing behaviors lower warm season (C_4) grass dominance in tallgrass prairie plant communities, and thus act to generate more equal proportions of cool season (C_3) grasses, forbs, and C_4 grasses (Knapp et al. 1999). However, while numerous projects attempt to restore bison to the Great Plains of North America, to date evidence that bison are acting as keystone species has been limited to the tallgrass prairies of the mid-western United States, and the Konza Prairie Biological Station in particular (Knapp et al. 1999; Towne et al. 2005). Thus, we still have little understanding about whether

the keystone effects of *B. bison* can be observed in other prairie ecosystems or how long post-reintroduction these effects become observable in those various prairie plant communities.

Our objective for this comparative study was to test whether the keystone role of bison as described in the tallgrass prairie ecosystem (Knapp et al. 1999; Fuhlendorf et al. 2010; Gates et al. 2010) can be applied to the mixed-grass prairies of the Northern Great Plains (NGP). To describe the extent to which restored bison could be acting as keystone species in mixed-grass prairies, we set out to evaluate support for four key plant community compositional and structural responses observed by Knapp et al. (1999) in their review of how bison can act as a keystone species in tallgrass ecosystems (Table 1). These hypothesized effects included: (H_1) bison restoration results in lower overall abundance of C_4 grasses compared to our livestock retention treatment, and bison restoration results in higher overall abundance of C_3 grasses and forbs compared to our livestock retention treatment, (H_2) bison grazing and disturbance generates higher overall species diversity and richness of plants than does cattle grazing and disturbance or livestock removal, (H_3) heterogeneity in bare ground cover and average plant height are higher in bison-grazed treatments compared to livestock retention or livestock removal treatments, and (H_4) greater compositional heterogeneity occurs within bison restoration treatments than occurs within livestock retained or livestock removal treatments. Accordingly, we designed this study to investigate the plant community responses that bison reintroduction, livestock removal, and cattle-grazing have on the mixed-grass prairie, and thus to compare how different grassland restoration or management strategies affect plant communities in the NGP.

Methods

Study Area

Our study took place in the NGP region of the United States, in a portion of southern Phillips County, Montana (Fig. 1). This area is part of a region of northeastern Montana known colloquially as the Missouri Breaks: a hilly, and formerly glaciated, part of the mixed-grass prairie (Manning 2009). The Missouri Breaks region is a 1.4 million hectare patchwork of public and private lands, with 23% of lands being privately held, 36% of lands publicly managed by the Bureau of Land Management

Table 1. The results of our study, where we compared the influence of bison reintroduction, livestock retention, and cattle-removal on the plant communities in the mixed-grass prairie of northeastern Montana. Our results are shown in relation to a subset of the major botanical findings by Knapp et al. (1999) comparing bison and cattle in the tallgrass prairie of Kansas.

Hypotheses		Knapp et al. (1999)	Our Study
H_1	C_4 grasses	Decreased (↓)	No significant difference
	C_3 grasses and forbs	Increased (↑)	No significant difference compared to cattle Decreased (↓) compared to livestock removal
H_2	Diversity	Increased (↑)	No significant difference
	Richness	Increased (↑)	Increased (↑)
H_3	Heterogeneity in plant height and bare ground cover	Increased (↑)	No significant difference
H_4	Compositional heterogeneity	Increased (↑)	Increased (↑)

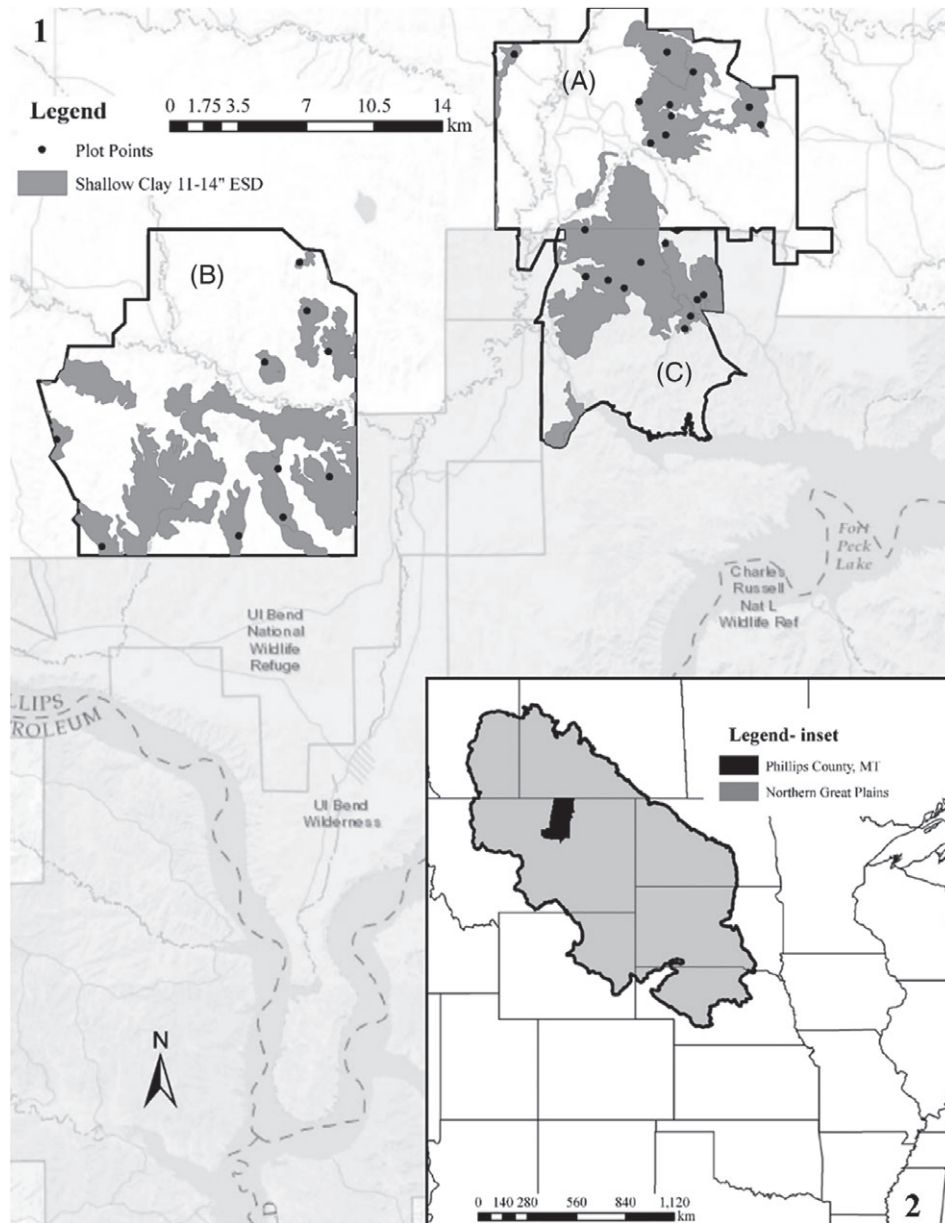


Figure 1. Map of our study area, located in Phillips County, Montana, showing the spatial arrangement of our three treatments: (A) bison-grazed, (B) livestock retention, and (C) livestock removal. Plots are represented as filled circles.

(BLM), and the remaining 41% being publicly managed by the U.S. Fish and Wildlife Service (Manning 2009). This area is dominated by a mixture of short and tallgrass prairie species, e.g. *Pascopyrum smithii* (Rydb.) Á. Löve, *Schizachyrium scoparium* (Michx.) Nash, *Hesperostipa comata* (Trin. & Rupr.) Barkworth var. *comata*, etc.. The mixed-grass prairie, as a region, generally has a moderate but highly variable climate (Savage 2011) as, for example, Phillips County averaged 34 cm in total precipitation from 2000 to 2016; ranging from 29 cm in 2015 to 61 cm in 2016.

Livestock grazing across the region was largely unregulated until 1934 when the Taylor Grazing Act was passed amid

concerns about range degradation on public lands, providing a regulatory process for allowing ranchers to lease federal lands for grazing at specific stocking rates. From 1915 to 1970, most of southern Phillips County was grazed similarly, most grazing being year-round, with unregulated grazing intensity (i.e. stocking rate), leading to a generally homogeneous distribution of extremely poor rangeland quality (high erosion, little vegetation cover, increased stream head cutting, or channelization) (W. Haglan 2017, USFWS, personal communication). Accordingly, we assumed that all our treatments historically were subjected to loosely regulated year-round grazing pre-1970, but transitioned to seasonal (i.e. growing season only), moderate

to light grazing intensity (i.e. 0.20–0.40 Animal Unit Month (AUM)/ha; the amount of land needed to support one cow and one calf for 1 month) rest-rotation cattle management from 1970 to 2005 through combined regulation and oversight by federal and state agencies (W. Haglan 2017, USFWS, personal communication; B. J. Rhodes 2015, BLM, personal communication). Thus, while we were not able to implement a classical before-after-control-impact (BACI) study design (Smith 2014), our selection of a study area and treatment types represented a natural experiment from which to assess correlations in vegetation response to grazer management.

Bison Restoration Treatment

Our bison treatment plots were located within a 12,545 ha reintroduction area that is owned and managed by the American Prairie Reserve (APR; Fig. 1). Specifically, we conducted our sampling in the 7,092 ha Telegraph and Box Elder creek drainages, referred to collectively as Box Elder, where cattle were removed in early 2004 (M. Kohl 2017, Utah State University, personal communication). In October 2005, the APR reintroduced 16 bison to Box Elder, and the area has experienced year-round bison grazing with the bison population growing to roughly 600 animals (including juveniles and subadults) by 2015 (18.75% growth per year with an average 29 imported animals per year [Reserve 2015]). Thus, bison have grazed Box Elder for 10 years by the time of our sampling (Fig. 1) during which bison grazing intensity has been maintained below a threshold of 0.39 AUM/ha.

Livestock Removal Treatment

We selected a 4,059 ha portion of the Charles M. Russell National Wildlife Refuge (CMR), where livestock were removed in 2004 as our livestock removal treatment. Since 1975, the refuge (managed by the USFWS) has been expiring grazing leases, leading to some unique areas to the NGP where grazing by large herbivores has not occurred in up to 40 years. Specifically, we focused our sampling within an allotment named Telegraph Creek Pasture Five where cattle had been absent for 10–11 years (Fig. 1). The allotted cattle grazing intensity data for the CMR (i.e. the AUM/ha permitted for grazing) are unavailable to compare with our other treatments, but the observed (or actual) grazing intensity from 1990 to 2004 fell within the range observed in our other two treatments.

Livestock Retention Treatment

Our livestock retention (control) treatment was located within the 8,303 ha Fourchette Creek grazing allotment that is managed by the BLM (Fig. 1). Cattle grazing on this allotment has been consistently managed via rest-rotation grazing since 1983, allowing for up to 2,815 AUMs of grazing pressure during a grazing season that lasted from 1 May to 30 October of each year (B. J. Rhodes 2015, BLM, personal communication). Thus, grazing intensity in our livestock retention treatment area was maintained below a threshold of 0.33 AUM/ha from 1 May to

30 October during years 1983–2016 (B. J. Rhodes 2015, BLM, personal communication).

Site Selection for Field Sampling

In addition to selecting for areas with similar historical (i.e. pre-2005) grazing pressures, to ensure comparability in site condition and potential community productivity among our three treatments, we selected sampling areas that were similar in terms of soil condition, slope, aspect, elevation, and other abiotic variables. To accomplish this, we utilized Ecological Site Description (ESD) data created by the U.S. Natural Resources Conservation Service (NRCS) that describe and outline areas with similar biotic and abiotic conditions (potential plant community, soil description, slope, elevation, etc.; <https://esis.sc.gov.usda.gov>). In addition, ESDs were developed as a tool to predict forage production potential for each ecological site (Bestelmeyer & Brown 2010). Given that we assumed vegetative community responses to grazing were likely to follow the dynamic equilibrium model, which predicts that the effects of disturbance are correlated positively with productivity (Huston 1979; Yuan et al. 2016), we focused sampling in ESDs where plant community productivity, and thus grazing pressure, was likely highest. Therefore, our use of ESDs enabled us to effectively restrict sampling to areas with high levels of forage productivity and potential for grazing to impact vegetative communities, as well as areas with homogeneous environmental conditions (e.g. soil texture, slope, elevation, etc.). The most common ESD with high productivity occurring in all three treatments was shallow clay 11–14" (SC) capable of producing from 703 to 5,878 kg of grass and sedge forage per hectare (<https://esis.sc.gov.usda.gov>). We then randomly selected 10 sample points within SC polygons for each treatment in the study area.

Vegetation Sampling and Assessment

Between June–August 2015 and May–August 2016, we established study plots and evaluated the vegetative community at 10 random sampling points within each of our treatments. We chose to sample 10 points based on a power analysis, where we utilized pilot data to generate reasonable estimates of (1) the difference in mean response among our treatments and (2) the variability among points within our treatments. We determined a statistical power estimate of ≥ 0.80 with an α of 0.05 for 10 sampling points within each treatment. At each sampling point, we established study plots following a modified Whitaker plot design employed in many long-term ecological studies across the United States (Newell & Peet 1998; Peet et al. 1998; Cavender-Bares et al. 2004; Fridley et al. 2005; Reilly et al. 2006; Carr et al. 2009). Specifically, at each point we established a 0.1 ha survey plot (20 m \times 50 m), each consisting of 10 individual 100 m² subplots (Fig. 2; referred to hereafter as modules). We sampled modules 1–10 in each plot (Fig. 2) for species incidence and directly used those data to estimate species richness for the entire plot. We restricted vegetation and bare ground cover sampling to modules 2, 3, 8, and 9 (Fig. 2; hereafter referred to as intensive modules). We estimated both the

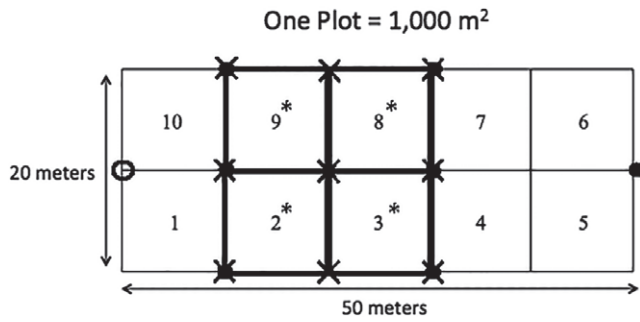


Figure 2. The spatial scale of an individual plot, consisting of 10 modules of 100 m². Modules 2, 3, 8, and 9 were sampled as intensive modules (bold boxes). All other modules (i.e. 1, 4, 5, 6, 7, 10) were sampled as an aggregate, searching only for species not recorded in the intensive modules. The origin of the plot is marked with an open circle, and other long-term plot markers are represented by filled black circles. Each bare ground cover measurement is denoted by an asterisk (*), and the individual sample points for plant height are marked as a cross (X).

cover for each species encountered and bare ground cover for each intensive module ($n = 4$ per plot) on a scale from 1 to 9, to represent a range of percent cover values for each metric (trace, 0–1, 1–2, 2–5, 5–10, 10–25, 25–50, 50–75, and 75–95%, respectively; Fig. 2). We estimated vegetation height variability across the whole 0.1 ha plot by recording standing vegetation heights (m) at each corner of a plot's intensive modules ($n = 9$ per plot; Fig. 2).

Statistical Analysis

We applied mixed-effect modeling techniques using the statistical program JMP 13.2.0 (SAS 2017) to evaluate support for the first three of our hypotheses (i.e. H_1 – H_3), testing for significance in each response variable (e.g. variance in bare ground cover) among our three treatments using mixed-effect analysis of variance (ANOVA) (sums of squares type I). In our mixed-effect models, we held our three treatment sites as fixed-effect variables, and our plots as random-effect variables nested within each treatment. When our model residuals were non-normally distributed, we transformed them as appropriate, or excluded extreme outliers when transformation did not result in normal residual distributions. We determined that treatment effects were statistically significant at an α of 0.05 for all statistical tests.

To test that C_4 grass, C_3 grass, and forb abundances differed among our three treatments (H_1), we transformed our cover class values to a midpoint percentage value (e.g. a species with 5–10% range was given a value of 7.5% for cover) for each species encountered in our four intensive modules at each plot (Peet et al. 1998). We then calculated the mean abundance for each species across the plot aggregate, and organized them into their broader functional groups prior to analysis. The residuals from our C_3 grass abundance data were normally distributed and did not require transformation. We log-transformed our non-normal C_4 grass abundances to meet normality assumptions for analysis. To meet the assumption of normality for our forb

data, we excluded an extreme outlier and log-transformed prior to testing. We used each of the aforementioned functional group data per treatment as fixed-effects in our statistical analysis.

To test our hypothesis (H_2) that bison generated higher plant species diversity than cattle or livestock removal, we first calculated plot-level diversity values using our abundance (i.e. the recorded cover) data for species encountered. We then used the “diversity” function in the R package “vegan” (Oksanen et al. 2016) to calculate the inverse Simpson ($1D^{-1}$; where $D = 1/\sum_{i=1}^S p_i^2$) and exponential Shannon (e^H ; where $H = -\sum_{i=1}^S p_i \ln p_i$) index values for each plot. We then utilized these diversity index values for each treatment as dependent variables in our statistical analysis.

To test the hypothesis that bison generated higher species richness (i.e. the total number of species found in an area) than cattle or livestock removal (H_2), we first utilized the “rarefy” function in the R package “vegan” (Oksanen et al. 2016) to standardize species richness across treatments. The “rarefy” function in vegan is based on calculations from Hurlbert (1971) and Heck et al. (1975). Rarefaction has become standard practice in analyzing species richness data, where rarefaction reduces potential influences of variability in species abundance values across samples in each treatment (Gotelli & Colwell 2011; Chao et al. 2014). Rarefied species richness values for each treatment were then tested as dependent variables in our statistical analysis.

We tested our hypothesis that bison generated higher bare ground cover heterogeneity than livestock retention or cattle removed areas (H_3) by comparing the coefficient of variation (CV) of bare ground across study plots within each of our three treatments ($n = 5$ for bison; $n = 10$ for cattle; $n = 10$ for grazer removal). Our sample size for bison treatments for this variable was reduced due to field limitations. We calculated the median of our bare ground cover range values (e.g. 5–10% would be said to equal 7.5%), and then averaged these values to obtain the mean bare ground cover value at each plot (Fig. 2). We lastly calculated the CV for mean bare ground cover for each treatment prior to testing, and used the CV values per treatment as dependent variables in our statistical analysis.

We evaluated whether differences in plant height heterogeneity occurred among treatments (H_3) by comparing the CV of plant height per plot. We used our plant height measurements per plot ($n = 270$; $n = 90$ per treatment) to calculate the CV for plant height across our three treatments ($n = 30$). To maintain assumptions of normality, we transformed the CV plant height data on a log-scale.

We applied principal coordinate analysis (PCoA) multivariate analysis techniques in the statistical program R (R Core Team 2017) to test our hypothesis that compositional heterogeneity was highest in our bison treatment, and lower in livestock retention and removal areas (H_4). Compositional heterogeneity is a measure of the amount of compositional difference between study sites, sample areas, or communities (Legendre et al. 2005; Arroyo-Rodríguez et al. 2013; Myers et al. 2013, 2015), and has been used to assess how management changes (e.g. fire or grazing) may directly affect compositional heterogeneity in plant communities (Reilly et al. 2006; Conradi et al. 2015). We

were interested in the amount of variance in species composition among plots for each treatment, and thus we utilized group dispersion ordination techniques to visually compare compositional heterogeneity between bison, cattle, and livestock removal treatments (Gardener 2014). To accomplish this, we first generated a dissimilarity index using species incidence data collected in each plot within the bison, cattle, and livestock removal treatments using Sørensen equation of dissimilarity:

$$\beta_s = \frac{(b + c)}{(2a + b + c)} \quad (1)$$

where a equals the number of species shared between two sites, b is equal to the species unique to site 1, and c is equal to the species unique to site 2 (Anderson 2006; Gardener 2014). We then used the multivariate dispersion function (betadisper) as part of the package “vegan” in R to generate an ordination of group dispersion using a PCoA (Gardener 2014; R Core Team 2017). This multidimensional ordination visually shows the heterogeneity among plots within and across treatments, where each point in the ordination represents a compositional heterogeneity value for a plot. The farther the distance that the point is from the center of the treatment (i.e. the center of each circle), the more different that sample point is from all the other samples in the whole treatment (i.e. the higher the variability) (Anderson 2006). Additionally, the bigger the polygon, the more variable the treatment is in terms of species composition, and the greater the β -diversity (Anderson 2006). Lastly, we used the analysis of variance function “ANOVA” and the pairwise comparison function “TukeyHSD” in R to test for significant differences in β -diversity among our treatments.

Results

Functional Group Abundance

We found no evidence that 10 years following bison reintroduction, C_4 grass abundance decreased (H_1 ; $F = 0.55$, $p = 0.58$, $df = 2$, 26) or forb abundance increased ($F = 1.9$, $p = 0.2$, $df = 2$, 26) compared to livestock retention or removal treatments. Although we did detect differences among our treatments in C_3 grass abundances ($F = 5.96$, $p < 0.01$, $df = 2$, 27), we did not observe a significant difference in C_3 grass abundance between bison and livestock treatments (Tukey-HSD, $p = 0.12$, $df = 2$). However, we observed that our bison restoration treatment ($\bar{x} = 4.2\%$, $SD = 2.8$, range = 0.9–9.7%) had 53.7% lower C_3 grass ($p < 0.01$, $df = 2$) abundance than was recorded in our livestock removal treatment ($\bar{x} = 9.1\%$, $SD = 3.3$, range = 5.3–16.8).

Species Diversity and Richness

Contrary to our predictions, we did not observe significant differences in plant species diversity among our three treatments using either of our diversity indices ($1/D$: $F = 0.30$, $p = 0.74$, $df = 2$, 27; or e^H : $F = 0.88$, $p = 0.43$, $df = 2$, 27).

In line with our predictions, our bison-restored treatment had significantly higher species richness than sites

where livestock were retained or removed ($F = 7.1$, $p < 0.01$, $df = 2$, 27; Fig. 3). We found that our bison-grazed treatment had 57% higher species richness ($\bar{x} = 30.7$, $SD = 6.1$, range = 18.5–39.0) compared to our cattle treatment ($\bar{x} = 24.2$, $SD = 2.4$, range = 19.6–28.4), and 56% higher richness compared to our livestock removal treatments ($\bar{x} = 23.9$, $SD = 4.2$, range = 16.2–29.5; Fig. 4). We did not detect any difference in species richness between livestock retention and removal treatments (Fig. 3).

Bare Ground and Vegetation Height

We found no support for our hypothesis that our bison-grazed treatment had significantly higher heterogeneity in bare ground ($F = 2.3$, $p = 0.13$, $df = 2$, 22) and plant height ($F = 1.27$, $p = 0.30$, $df = 2$, 27) than livestock retention or removal treatments.

Compositional Heterogeneity (β -Diversity)

Our bison-restored treatment exhibited significantly higher compositional heterogeneity than our livestock retention or removal treatments ($F = 7.7$, $p < 0.01$, $df = 2$, 27). We observed in our multivariate dispersion model that compositional heterogeneity was 56% higher in our bison treatment than in our livestock retention treatment, and 59% higher than livestock removal treatment (Fig. 4). This variation in species composition was also observed in our multivariate ordination, where the bison treatment polygon was visibly larger than our other two treatments, showing higher compositional heterogeneity within our bison treatment than within the others (Fig. 4).

Discussion

We found support for a subset (i.e. 2/4) of our hypotheses, indicating that bison may be generating plant community responses indicative of a keystone modifier species in the mixed-grass prairie of the NGP 10 years postreintroduction. Of those results, our higher observed species richness and compositional heterogeneity (β -diversity) in mixed-grass prairie plant communities between bison-grazed versus cattle-grazed rangeland follows what has been reported following bison restoration in the tall-grass prairie ecosystem (Knapp et al. 1999; Towne et al. 2005; Fuhlendorf et al. 2010). Thus, despite general floristic community differences between the mixed and tallgrass prairie ecosystems (see Towne 2002; Charboneau et al. 2013), we still observed support for evidence of keystone impacts by bison on plant diversity. However, unlike other studies in tallgrass ecosystems, we did not observe functional group or habitat structural (i.e. bare ground or standing plant height) differences among our treatments. This in turn raises the important question: At what scale and which vegetation responses are sufficient for bison to be considered a modifier keystone species in a system?

In addition to questions regarding the scale and number of measures that vegetative responses should be detected at to qualify bison as a modifier keystone species in a system, it remains

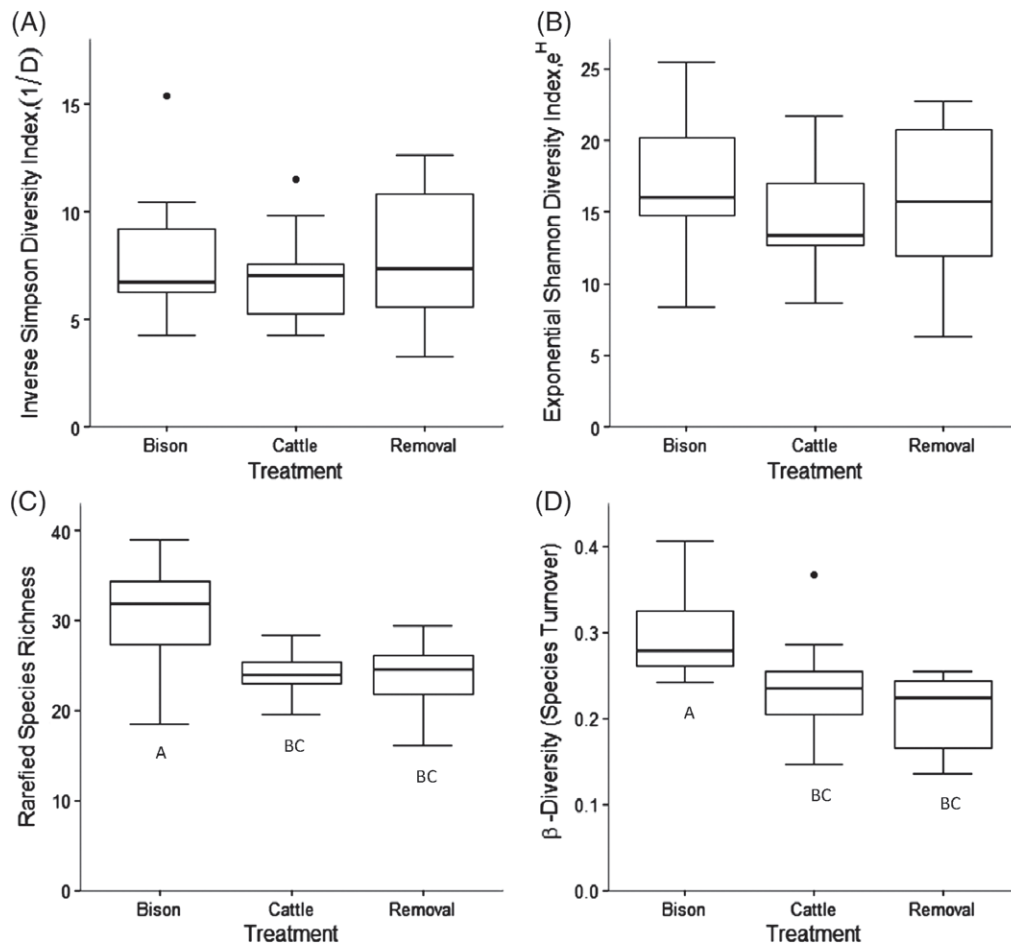


Figure 3. (A) Inverse Simpson diversity index, (B) exponential Shannon diversity index values, and (C) rarefied species richness for each of our three treatments, with 95% confidence intervals for each represented as error bars. (D) Beta-diversity, or compositional heterogeneity, is plotted as the distance of each plot from the overall centroid of the data for each treatment (see Fig. 4), representing the overall variation in species composition for the respective sites (significance is represented via pairwise notation; $\alpha = 0.05$).

unclear how much time must elapse postrelease before keystone effects become observable. We observed that sites exposed to 10 years of bison grazing showed increased forb species richness and abundance compared to livestock retention areas, which is consistent with observations in the tallgrass prairie. However, counter to what would be expected in the mesic tallgrass ecosystem, we did not observe significant differences with livestock retention areas in the dominance of C_4 grasses over C_3 grasses (Knapp et al. 1999; Towne et al. 2005; Fuhlendorf et al. 2010). Similar studies looking at cattle grazing in the shortgrass steppe showed very little functional group difference between grazing treatments in less than 20 years, with functional group shifts not occurring for almost 70 years (Augustine et al. 2017). However, more productive ecotones like the mixed-grass prairie will likely show these shifts sooner than the arid shortgrass steppe (Huston 1979), and this is likely reflected in the partial transition observed in our study. Indeed evidence we observed in relatively fine-scale measures of species richness and abundance may suggest that bison are starting to transition plant communities in our study area toward those expected shifts in functional groups

indicative of a keystone species effect (Knapp et al. 1999), suggesting more time may be needed. Therefore, if conservation organizations justify reintroduction of bison based on the belief they will exhibit keystone impacts on grassland systems, it is critical that focused and sustained long-term monitoring are supported to test these hypothesized effects (Jachowski et al. 2016). In particular, it is essential that bison reintroductions in other portions of the NGP establish robust, long-term monitoring strategies utilizing BACI experimental designs to further define the role that bison may play in the region's ecosystems.

Although our bison reintroduction, livestock retention, and livestock removal treatments did not show predicted differences using broad functional groups (e.g. forbs), we suggest that future research utilize groupings based on more than just growth-habit alone to effectively detect differences among treatments. Bison reintroduction is predicted to increase rangeland forb abundance and richness compared to livestock retention (Knapp et al. 1999); however, we reported the contrary. It should be noted, however, that when we took a finer-scale look at the forb composition of each treatment, we found that non-native (and often

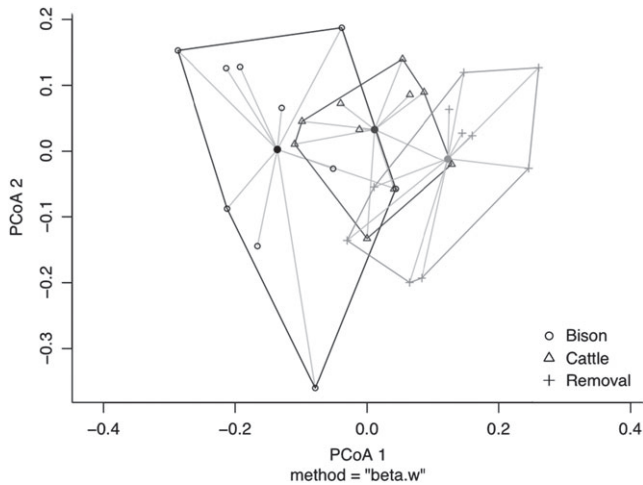


Figure 4. Multivariate dispersion (via principal coordinates analysis) for compositional heterogeneity (β -diversity) between our bison, cattle, and cattle-removal treatments, showing both the difference in composition between each plot within each treatment (i.e. distance from the middle of each circle) and the difference among treatments (i.e. the distance from the center of the plot).

invasive) species constituted 70 and 46% of the total forb cover in our livestock retention and removal treatments, respectively, compared to 20% of the total in our bison-restored treatment. While it is possible that our results could be due to habitat differences between the tallgrass and mixed-grass prairie, our contradictory results regarding the effects of bison on forb abundance and composition may be partially explained by the disproportionate cover of non-native forbs in our other treatments. Studies attempting to reveal the interactions between wildlife and plant communities often utilize broad groupings of plants (e.g. Wedin et al. 1996); however, our study suggests that such basic views of plant communities and functional groups may lead to erroneous or conflicting interpretations of plant–wildlife interactions.

Stocking rate can have varying influence on rangeland community dynamics (Olf & Ritchie 1998), and could have influenced our ability to discern vegetation responses to bison restoration. We attempted to control for the effect of stocking rate on plant community composition and species abundances, but detailed records on the precise stocking rates (i.e. AUM/hectare) that traditionally cattle-grazed lands have experienced is lacking or inconsistent between agencies (B. J. Rhodes 2016, BLM, personal communication; R. Matchett 2016, USFWS, personal communication). Stocking densities at our study sites (0.39 AUM/ha for bison; 0.33 AUM/ha for cattle) were lighter than those at Konza prairie (2.1 AUM/ha for bison; 2.4 AUM/ha for cattle), where Knapp et al. (1999) performed their study supporting keystone effects of bison (Allred et al. 2011). The higher stocking densities at Konza may have resulted in the more dramatic community shifts described there following bison restoration. Therefore, we suggest that future attempts should be made to compare vegetative community responses across a range of bison-stocking densities, including

those that are closer to the Konza prairie. At the same time, bison reintroduction at the spatial scales used in this study (from approximately 3,000 to 12,500 ha over 10 years) may not produce the effects expected from historic bison disturbance on vegetative communities. Sanderson et al. (2008) suggested that bison reintroduction projects at scales greater than 200,000 ha are the most effective, and likely represent historic foraging behaviors, with a single bison foraging patch being greater than 11,000 ha in the NGP (Kohl et al. 2013). Thus, Kohl et al. (2013) suggested that the APR bison reintroduction area (3,555 ha at the time of their study) may not be large enough to facilitate our predicted historic ecological influence, and that continued access to areas greater than 12,000 ha may both result in a return to historic disturbance patterns and vegetative community composition. Additionally, while most nonbovine ungulate species in our study area (e.g. *Odocoileus virginianus* Zimmermann, *Odocoileus hemionus* Rafinesque, *Antilocapra americana* Ord.) are predominantly browsers, elk (*Cervus elaphus*) have been shown to be graminoid grazers, potentially amplifying the effect of bovine grazers on prairie plant communities (Keller 2011). Therefore, we suggest that future studies test along a gradient of grazing regimes at a variety of spatial scales postreintroduction, and integrate the potential additive effects of nonbovine grazers to further parse out differences between bison and cattle.

Our results suggest that livestock removal may contribute to declines in overall grassland biodiversity. The plant communities of the NGP have undoubtedly evolved in the presence of large grazers (Knapp et al. 1999), and large grazer exclusion or livestock removal in the tallgrass and shortgrass prairie is expected to result in lower annual forb cover compared to grazed sites (Fuhlendorf & Engle 2001; Valone & Sauter 2005; Manier & Hobbs 2006). Our findings are comparable with studies in the shortgrass prairie where livestock removal resulted in heightened forb abundance (Augustine et al. 2017). However, our data suggest that removal may additionally lead to landscape homogenization (evidenced by lower compositional heterogeneity compared to bison treatment). Increased rangeland homogenization (or lack of heterogeneity) can be commensurate with declines in grassland biodiversity, and the decline of wildlife species such as the chestnut-collared longspur (*Calcarius ornatus* J.K. Townsend) and Baird's sparrow (*Ammodramus bairdii* Audubon) who require the influence of grazing to generate variability in plant community structure (standing biomass and bare ground cover) and composition (Fuhlendorf & Engle 2001; Toombs et al. 2010). Thus, managers implementing grazer removal may both decrease habitat for many grassland plant and wildlife species that require structurally heterogeneous plant communities, and support lower biodiversity than sites grazed by either bison or cattle (Fuhlendorf & Engle 2001).

Different ecological sites (or habitats) may respond slightly differently than our sites within the SC-ESD, as it is hypothesized that grazing's relationship to grassland plant community composition varies along environmental gradients (e.g. soil, altitude, etc.) (Zemlich et al. 2010). Discrepancies in habitat selection between bison and cattle may, in turn, result in stronger plant community responses in some ecological sites compared

to others (Kohl et al. 2013). Because our study only investigated a single rangeland community type or ESD, it is essential that more studies be conducted to detect community-level effects of cattle-grazing, removal, and bison reintroduction, across a range of environmental conditions.

Species introductions or reintroductions are often popularly advocated based on possible keystone effects (Hansen et al. 2010; Newsome et al. 2015), but there have been comparatively few studies directly evaluating the influence that these species' translocations have had on restored ecosystems. For example, in Yellowstone National Park there is evidence of keystone predator effects by gray wolves following reintroduction (Ripple & Beschta 2012; Painter et al. 2015), yet there is a lack of evidence to suggest these effects would similarly be observed in other ecosystems where wolves have been restored (Mech 2012; Ford & Goheen 2015). Similar to evidence regarding the keystone predator effect in large carnivores (Haswell et al. 2017), our findings suggest that keystone modifier effects by large herbivore are likely highly context dependent. Thus, if ecologists and managers are to continue to gain support for translocation and reintroduction projects involving large mammals based on them being a means for achieving biodiversity restoration goals, it is essential to continually evaluate and confirm the transferability of these processes across ecosystems (Hansen et al. 2010; Newsome et al. 2015). For bison restoration, while there are many ecological, cultural, and aesthetic reasons for various groups to undertake a bison reintroduction project (Freese et al. 2007), if bison are to be reintroduced across the Great Plains based on expected keystone effects, it is imperative that managers implement long-term, robust monitoring techniques to track their influence on rangeland plant communities.

Acknowledgments

The authors thank M. L. McMillan, P. D. McMillan, R. Rock, T. Coats, N. Collier, J. Segar, B. Ross, O. Wood, D. Bertuna, M. Caplice, and E. Freese for their tireless help in establishing and collecting data from our field plots, and for making this project possible. We also thank G. and P. Segars, the Alexander P. and Lydia Anderson Fellowship, and the Wade T. Batson Scholarship for financial support. We want to thank B. J. Rhodes of the Bureau of Land Management's Malta, Montana Field Office for technical and logistical assistance throughout our field work, and both R. Matchett and M. DeRosier of the U.S. Fish and Wildlife Service's Charles M. Russell National Wildlife Refuge for all the technical, logistical, and data collection assistance that made this project possible. Additionally, we wish to thank D. Austin, L. Anderson, E. Anderson, and B. Holder of the American Prairie Reserve for providing excellent logistical assistance throughout the project, along with technical information pertaining to the APR bison reintroduction project. Lastly, we would like to express our gratitude to both B. Bridges and S. Dewalt for their enthusiastic technical assistance, comments, and guidance throughout our data analysis process and study design.

LITERATURE CITED

- Allred BW, Fuhlendorf S, Hamilton R (2011) The role of herbivores in Great Plains conservation: comparative ecology of bison and cattle. *Ecosphere* 2:26
- Anderson MJ (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62:245–253
- Arroyo-Rodríguez V, Rös M, Escobar F, Melo F, Santos B, Tabarelli M, et al. (2013) Plant β -diversity in fragmented rain forests: testing floristic homogenization and differentiation hypotheses. *Journal of Ecology* 101:1449–1458
- Augustine DJ, Derner JD, Milchunas D, Blumenthal D, Porensky LM (2017) Grazing moderates increases in C3 grass abundance over seven decades across a soil texture gradient in shortgrass steppe. *Journal of Vegetation Science* 28:562–572
- Bestelmeyer BT, Brown JR (2010) An introduction to the special issue on ecological sites. *ARS USDA Submissions* 32:3–4
- Carr SC, Robertson KM, Platt WJ, Peet RK (2009) A model of geographical, environmental and regional variation in vegetation composition of pyrogenic grasslands of Florida. *Journal of Biogeography* 36:1600–1612
- Cavender-Bares J, Kitajima K, Bazzaz FA (2004) Multiple trait associations in relation to habitat differentiation among 17 Floridian oak species. *Ecological Monographs* 74:635–662
- Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological Monographs* 84:45–67
- Charboneau JLM, Nelson BE, Hartman RL (2013) A floristic inventory of Phillips and Valley Counties, Montana (U.S.A.). *Journal of the Botanical Research Institute of Texas* 7:847–878
- Conradi T, Strobl K, Wurfer AL, Kollmann J (2015) Impacts of visitor trampling on the taxonomic and functional community structure of calcareous grassland. *Applied Vegetation Science* 18:359–367
- Conway WG (1989) The prospects for sustaining species and their evolution. Pages 199–209. In: *Conservation: for the twenty-first century*. Oxford University Press, New York
- Coppedge BR, Shaw JH (1997) Effects of horning and rubbing behavior by bison (Bison bison) on woody vegetation in a tallgrass prairie landscape. *American Midland Naturalist* 138:189–196
- Ford AT, Goheen JR (2015) Trophic cascades by large carnivores: a case for strong inference and mechanism. *Trends in Ecology and Evolution* 30:725–735
- Freese CH, Aune KE, Boyd DP, Derr JN, Forrest SC, Gates C, et al. (2007) Second chance for the plains bison. *Biological Conservation* 136:175–184
- Fridley JD, Peet RK, Wentworth TR, White PS (2005) Connecting fine- and broad-scale species-area relationships of southeastern U.S. flora. *Ecology* 86:1172–1177
- Fuhlendorf SD, Engle DM (2001) Restoring heterogeneity on rangelands: ecosystem management based on evolutionary grazing patterns. *Bioscience* 51:625
- Fuhlendorf SD, Allred BW, Hamilton RG (2010) Bison as keystone herbivores on the Great Plains: can cattle serve as proxy for evolutionary grazing patterns? American Bison Society Working Paper No. 4. American Bison Society, Bronx, NY
- Gardener M (2014) *Community ecology: analytical methods using R and Excel*. Pelagic Publishing, Exeter, U.K.
- Gates C, Freese CH, Gogan P, Kotzman M (2010) *American bison: status survey and conservation guidelines*. ICUN, Gland, Switzerland
- Gotelli N, Colwell R (2011) Estimating species richness. Pages 39–54. In: Magurran AE, McGill BJ (eds) *Frontiers in measuring biodiversity*. Oxford University Press, New York
- Hansen DM, Donlan CJ, Griffiths CJ, Campbell KJ (2010) Ecological history and latent conservation potential: large and giant tortoises as a model for taxon substitutions. *Ecography* 33:272–284
- Haswell PM, Kusak J, Hayward MW (2017) Large carnivore impacts are context dependent. *Food Webs* 12:3–13

- Heck KL, van Belle G, Simberloff D (1975) Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology* 56:1459–1461
- Hornaday WT (1889) Pages 369–548. The extermination of the American Bison. Report of the National Museum 1886–87. Vol 31. US Government Printing Office, Washington, DC
- Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577–586
- Huston M (1979) A general hypothesis of species diversity. *The American Naturalist* 113:81–101
- Jachowski D, Millsbaugh JJ, Angermeier PL, Slotow RH (2016) Reintroduction of fish and wildlife populations. University of California Press, Berkeley, CA
- Keller BJ (2011) Factors affecting spatial and temporal dynamics of an ungulate assemblage in the Black Hills, South Dakota. University of Missouri, Columbia
- Knapp AK, Blair JM, Briggs JM, Collins SL, Hartnett DC, Johnson LC, Towne EG (1999) The keystone role of Bison in North American tallgrass prairie: bison increase habitat heterogeneity and alter a broad array of plant, community, and ecosystem processes. *Bioscience* 49:39–50
- Koerner SE, Burkepille DE, Fynn RWS, Burns CE, Eby S, Govender N, et al. (2014) Plant community response to loss of large herbivores differs between North American and South African savanna grasslands. *Ecology* 95:808–816
- Kohl MT, Krausman PR, Kunkel K, Williams DM (2013) Bison versus cattle: are they ecologically synonymous? *Rangeland Ecology & Management* 66:721–731
- Legendre P, Borcard D, Peres-Neto PR (2005) Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75:435–450
- Manier DJ, Hobbs NT (2006) Large herbivores influence the composition and diversity of shrub-steppe communities in the Rocky Mountains, USA. *Oecologia* 146:641–651
- Manning R (2009) Rewilding the west: restoration in a prairie landscape. University of California Press, London, United Kingdom
- Mech LD (2012) Is science in danger of sanctifying the wolf? *Biological Conservation* 150:143–149
- Mills LS, Soulé ME, Doak DF (1993) The keystone-species concept in ecology and conservation, management and policy must explicitly consider the complexity of interactions in natural systems. *Bioscience* 43:219–224
- Myers JA, Chase JM, Jiménez I, Jørgensen PM, Araujo-Murakami A, Paniagua-Aambrana N, Seidel R (2013) Beta-diversity in temperate and tropical forests reflects dissimilar mechanisms of community assembly. *Ecology Letters* 16:151–157
- Myers JA, Chase JM, Crandall RM, Jiménez I (2015) Disturbance alters beta-diversity but not the relative importance of community assembly mechanisms. *Journal of Ecology* 103:1291–1299
- Newell CL, Peet RK (1998) Vegetation of Linville gorge wilderness, North Carolina. *Castanea* 63:275–322
- Newsome TM, Ballard GA, Crowther MS, Dellinger JA, Fleming PJS, Glen AS, et al. (2015) Resolving the value of the dingo in ecological restoration. *Restoration Ecology* 23:201–208
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB et al. (2016) Vegan: community ecology package. <https://CRAN.R-project.org/package=vegan>
- Olff H, Ritchie ME (1998) Effects of herbivores on grassland plant diversity. *Trends in Ecology and Evolution* 13:261–265
- Paine RT (1969) A note on trophic complexity and community stability. *The American Naturalist* 103:91–93
- Painter LE, Beschta RL, Larsen EJ, Ripple WJ (2015) Recovering aspen follow changing elk dynamics in Yellowstone: evidence of a trophic cascade? *Ecology* 96:252–263
- Peet RK, Wentworth TR, White PS (1998) A flexible, multipurpose method for recording vegetation composition and structure. *Castanea* 63:262–274
- Power ME, Tilman D, Estes JA, Menge BA, Bond WJ, Mills LS, et al. (1996) Challenges in the quest for keystones. *Bioscience* 46:609–620
- R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Reilly MJ, Wimberly MC, Newell CL (2006) Wildfire effects on plant species richness at multiple spatial scales in forest communities of the southern Appalachians. *Journal of Ecology* 94:118–130
- Reserve AP (2015) Bison report 2015. American Prairie Reserve, Bozeman, MT
- Ripple WJ, Beschta RL (2012) Trophic cascades in Yellowstone: the first 15 years after wolf reintroduction. *Biological Conservation* 145:205–213
- Sanderson EW, Redford KH, Weber B, Aune K, Baldes D, Berger J, et al. (2008) The ecological future of the North American bison: conceiving long-term, large-scale conservation of wildlife. *Conservation Biology* 22:252–266
- SAS (2017) JMP®. SAS Institute, Inc., Cary, NC
- Savage CS (2011) *Prairie: a natural history*. Greystone Books, Vancouver, BC, Canada
- Seddon PJ, Griffiths CJ, Soorae PS, Armstrong DP (2014) Reversing defaunation: restoring species in a changing world. *Science* 345:406–412
- Smith EP (2014). BACI design. In: Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL (eds) Wiley StatsRef: statistics reference online. doi:10.1002/9781118445112.stat07659
- Soulé ME, Estes JA, Berger J, Martinez Del Rio C (2003) Ecological effectiveness: conservation goals for interactive species. *Conservation Biology* 17:1238–1250
- Toombs TP, Derner JD, Augustine DJ, Krueger B, Gallagher S (2010) Managing for biodiversity and livestock. *Rangelands* 32:10–15
- Towne EG (2002) Vascular plants of Konza Prairie Biological Station: an annotated checklist of species in a Kansas Tallgrass Prairie. *Sida: Contributions to Botany* 20:269–294
- Towne EG, Hartnett DC, Cochran RC (2005) Vegetation trends in tallgrass prairie from bison and cattle grazing. *Ecological Applications* 15:1550–1559
- Valone TJ, Sauter P (2005) Effects of long-term cattle enclosure on vegetation and rodents at a desertified arid grassland site. *Journal of Arid Environments* 61:161–170
- Wedin DA, Tilman D, Briggs JM, Blair JM, Steinauer EM (1996) Influence of nitrogen loading and species composition on the carbon balance of grasslands. *Science (New York, N.Y.)* 274:1720–1723
- Yuan ZY, Jiao F, Li YH, Kallenbach RL (2016) Anthropogenic disturbances are key to maintaining the biodiversity of grasslands. *Scientific Reports* 6:22132
- Zemlich A, Manthey M, Zerbe S, Oyunchimeg D (2010) Driving environmental factors and the role of grazing in grassland communities: a comparative study along an altitudinal gradient in western Mongolia. *Journal of Arid Environments* 74:1271–1280

Coordinating Editor: Jacob Bowman

Received: 28 November, 2017; First decision: 6 February, 2018; Revised: 30 June, 2018; Accepted: 1 July, 2018