Soil heterogeneity increases plant diversity after 20 years of manipulation during grassland restoration

SARA G. BAER,1,5 TIANJIAO ADAMS,2 DREW A. SCOTT,2 JOHN M. BLAIR,3 AND SCOTT L. COLLINS4

1 Kansas Biological Survey and Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, Kansas 66047 USA
2 Department of Plant Biology and Center for Ecology, Southern Illinois University, Carbondale, Illinois 62901 USA
3 Division of Biology, Kansas State University, Manhattan, Kansas 66506 USA
4 Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131 USA


Abstract. The “environmental heterogeneity hypothesis” predicts that variability in resources promotes species coexistence, but few experiments support this hypothesis in plant communities. A previous 15-yr test of this hypothesis in a prairie restoration experiment demonstrated a weak effect of manipulated soil resource heterogeneity on plant diversity. This response was attributed to a transient increase in richness following a post-restoration supplemental propagule addition, occasionally higher diversity under nutrient enrichment, and reduced cover of a dominant species in a subset of soil treatments. Here, we report community dynamics under continuous propagule addition in the same experiment, corresponding to 16–20 yr of restoration, in response to altered availability and heterogeneity of soil resources. We also quantified traits of newly added species to determine if heterogeneity increases the amount and variety of niches available for new species to exploit. The heterogeneous treatment contained a factorial combination of altered nutrient availability and soil depth; control plots had no manipulations. Total diversity and richness were higher in the heterogeneous treatment during this 5-yr study due to higher cover, diversity, and richness of previously established forbs, particularly in the N-enriched subplots. All new species added to the experiment exhibited unique trait spaces, but there was no evidence that heterogeneous plots contained a greater variety of new species representing a wider range of trait spaces relative to the control treatment. The richness and cover of new species was higher in N-enriched soil, but the magnitude of this response was small. Communities assembling under long-term N addition were dominated by different species among subplots receiving added N, leading to greater dispersion of communities among the heterogeneous relative to control plots. Contrary to the deterministic mechanism by which heterogeneity was expected to increase diversity (greater variability in resources for new species to exploit), higher diversity in the heterogeneous plots resulted from destabilization of formerly grass-dominated communities in N-enriched subplots. While we do not advocate increasing available soil N at large scales, we conclude that the positive effect of environmental heterogeneity on diversity can take decades to materialize and depend on development of stochastic processes in communities with strong establishment limitation.

Key words: assembly; community; functional traits; niche; nitrogen; richness; tallgrass prairie.

INTRODUCTION

There is an escalating need to manage processes that maintain and promote species richness as human-driven environmental change continues to degrade ecosystems and reduce biodiversity (Vitousek et al. 1997, MacDougall et al. 2013). Global conversion of grasslands to agriculture reduces native species richness, landscape heterogeneity (Ellis and Ramankutty 2008), and propagule supply for community regeneration (Willand et al. 2013). Temperate grassland ranks highest in the ratio of area converted to protected lands, and as such, this biome is considered to be “in crisis” and a high conservation priority (Hoekstra et al. 2005). Human intervention is generally needed to restore grasslands from agricultural conditions. Grassland restoration has been practiced for over a century, with the first effort aimed at reconstructing North American tallgrass prairie (Mlot 1990, Sperry 1994). Despite the long practice of prairie restoration, many restored prairies (particularly on formerly cultivated land) suffer from diversity that declines over time to levels less than extant prairies (Kindscher and Tieszen 1998, Camill et al. 2004, Martin et al. 2005, Polley et al. 2005, Willand et al. 2013, Hansen and Gibson 2014, Barak et al. 2017, Bauer et al. 2017).
Declining diversity during grassland restoration is generally attributed to environmental filtering, increasing dominance of grasses, and limited niche space for new species to exploit (Polley et al. 2007, McCain et al. 2010, Wilsey 2010, Klopf et al. 2014, Baer et al. 2016, Scott and Baer 2018).

Numerous studies that show species richness increases with environmental heterogeneity (reviewed by Wilson 2014) and heterogeneity is one of many factors (and processes) operating simultaneously to connect richness and productivity (Grace et al. 2016). Available niche space acts as an environmental filter in community assembly and is influenced by deterministic drivers and stochastic processes acting at multiple scales. For example, fine- and intermediate-scale spatial heterogeneity have been shown to increase species density and diversity in plant communities (Richardson et al. 2012, Williams and Houseman 2014). In tallgrass prairie, varied ecological and geographical drivers promote large-scale heterogeneity in plant communities (tilman 1984, Briggs and Knapp 1995, Turner et al. 1997, Collins et al. 2018). At local scales, plant diversity responds negatively to drivers that promote grass dominance and cause local extinction of subordinate species. For example, plant diversity in native prairie is generally higher in shallow soils with limited rooting depth, lower water availability, and lower grass dominance (Gibson and Hulbert 1987, Bush and Van Auken 2010, Collins and Calabrese 2012). Alternatively, nitrogen enrichment initially increases grass dominance and reduces plant diversity (Turner et al. 1997, Collins et al. 1998). Therefore, increasing heterogeneity of soil resources (i.e., soil depth and N availability) could be key to reconstructing more diverse and heterogeneous communities under initially homogeneous agricultural conditions, provided established or colonizing species contain sufficient functional variation to exploit the range of niches that environmental heterogeneity creates.

Plant functional traits can indicate resource availability and heterogeneity under various environmental conditions (Lavorel and Garnier 2002, Harpole and Tilman 2007). The variation in plant functional traits in the pool of dispersing species is a regional process influencing community assembly and local diversity (Schellberg and Pontes 2012). Following arrival, functional traits determine the ability of a species to pass through multiple interacting environmental filters (e.g., soil nutrient status, moisture conditions, competitors, mutualists, etc.) that influence the establishment, survival, and reproduction of species in a specific environment (Reich et al. 2003). For example, species with thin leaves, tall stature, and fast growth rates establish better under conditions of high resource availability (Dyer et al. 2001). Seed traits can also influence species establishment, predation, and germination success (Westoby 1998, Franzen 2004, Larson et al. 2015). Ultimately, there must be sufficient resource availability in the environment and sufficient functional variation in the pool of dispersing species for new species to become recruited into established communities (Conradi and Kollmann 2016).

In 1997, an experimental prairie restoration was established to test the applicability of the “environmental heterogeneity hypothesis” (EHH; Ricklefs 1977, Huston 1979). The field experiment contains replicated plots of prairie restored with no soil manipulations and prairie restored under a factorial combination of altered nutrient availability and soil depth (Baer et al. 1999, 2003, 2016). Plant community dynamics over time showed no effect of manipulated soil heterogeneity on species diversity during the first five years of the experiment (Baer et al. 2004). A weak effect of heterogeneity developed over the longer term (15 yr) and was attributed to a transient increase in richness following a one-time propagule addition of new species 8 yr post-restoration, which resulted in a lower rate of species loss over time in more heterogeneous plots with manipulated N availability and soil depth (Baer et al. 2016). These results led us to hypothesize that spatial heterogeneity in belowground resources promotes greater community heterogeneity that will reduce local extirpation of previously established species and increase openness for new species to colonize over the longer term. More succinctly, we hypothesized that heterogeneity begets heterogeneity. This hypothesis can be tested by adding propagules of new species to communities, a prerequisite to increasing richness (Eriksson and Ehrlen 1992a, b, Tilman 1993, 1997, Zobel 1997, Zobel et al. 2000, Myers and Harms 2009). Because we were not testing dispersal limitation per se, we added propagules of the same suite of new species to all of the experimental plots. Plant community composition was quantified for 5 yr (corresponding to 16–20 yr post-restoration) under continuous propagule addition to (1) reveal whether heterogeneous soil conditions increase diversity through its effect on the established local community or openness to the recruitment of new species; and (2) identify which soil treatments are most influential on the diversity response to heterogeneity, if any we also quantified traits of new species supplied as propagules to reveal whether heterogeneous plots contain more niches for new species to exploit (the key mechanism underlying the EHH), indicated by establishment of species containing unique trait spaces (the collection of traits that help describe a species’ niche).

**Methods**

**Study site**

The experiment was established in a former agricultural field at the Konza Prairie Biological Station (KPBS), a 3,457-ha Long-Term Ecological Research (LTER) site in the Flint Hills region of northeastern Kansas (39°05’ N, 96°35’ W), USA. Average annual precipitation at KPBS is 835 mm. Annual precipitation
received from 2013 through 2017 was 783, 701, 998, 984, and 721 mm, respectively (precipitation data available online). The native vegetation at Konza Prairie is dominated by C₄ grass species (*Andropogon gerardii*, *Sorghastrum nutans*, *Panicum virgatum*, and *Schizachyrium scoparium*), but >300 subordinate forb species occur throughout the site and contribute most to diversity. Soil where the experiment was conducted was classified as Reading silt loam (mesic Typic Arguidoll).

**Experimental design and restoration approach**

The experiment was a randomized complete block design containing four blocks. Each block contained two 6 × 8 m plots randomly assigned to maximum heterogeneity and control treatments (Baer et al. 2016). The maximum heterogeneity treatment contained a 2 × 3 factorial combination of soil depth and nutrient availability arranged according to a split block design. Soil depth (two levels: reduced and deep) was assigned to two of four alternating 2 × 6 m strips. Soil nutrient availability (three levels: reduced-N, ambient-N and enriched-N) was assigned to one of three 2 × 8 m strips (perpendicular to the depth treatment). Reduced soil depth was achieved by excavating the soil to a depth of 20–25 cm prior to the initial planting and burying pieces of rough-cut limestone slabs (Baer et al. 1999). The average depth to limestone was 24 ± 3 cm. All plots were excavated to control for this disturbance. Sawdust was initially added to reduce N availability to plants and effectively immobilized N (Baer et al. 2003). Starting in 2005, the reduced N treatment has been maintained by applying 84.2 g C/m² (sucrose-C) three times each growing season, which sustained reduced-N conditions (Baer and Blair 2008, Baer et al. 2016). Elevated-N strips have received 5 g N/m² in the form of NH₄-N in early June every year since 1998. The factorial combination of N and soil depth treatments resulted in 12 subplots that varied in soil N level and soil depth in each maximum heterogeneity plot. The control plots contained no subplot treatments.

In 1998, all plots were sown with 42 native prairie species using a log-normal distribution of dominant grasses and subordinate forbs to resemble never-cultivated tallgrass prairie (Baer et al. 1999, 2003). In 2005, seeds of 15 additional forb species were sown into all plots at a rate of 25 live seeds/m² (Baer et al. 2016). Beginning in 2013, we added 17 additional species (14 forbs, 2 grasses, and 1 sedge) that were never previously recorded in this experiment, but occur at KPBS, to all plots (Appendix S1: Table S1). Seeds were purchased from Prairie Moon Nursery (Winona, Wisconsin, USA). Percent live seed was either provided by the seed supplier or determined by the Illinois Seed Testing Lab. Seeds were hand-broadcasted in spring each year at a rate of 20 live seeds/m² per species for a total seeding rate of 300 live seeds/m². The experimental area was fenced in 1998 to prevent deer browsing and has been burned almost annually in the early spring (a common management practice in this region) since 1998, with the exception of 2000 and 2003.

**Plant community measurements**

Each plot was divided into 12 2 × 2 m subplots based on the orthogonally crossed assignment of treatments in the maximum heterogeneity plots. The percent cover of each species was visually estimated in two permanent 0.5 × 0.5 m quadrats in each subplot in late spring (late May–early June) and late summer (late August–early September) from 2013 through 2017. The maximum cover of each species across the two seasonal samplings was used as the cover value for each species in a quadrat. Species composition surveys conducted prior to 2013 contained few seedlings, so their presence and frequency was attributed to the most recent propagule addition. Maximum cover of each species was then averaged across the two quadrats in each subplot. Plot-level richness (S) was calculated by summing the total number of species from all subplots within a plot. Plot-level diversity (Shannon’s diversity index, \( H' \)) and evenness (Pielou’s evenness, \( J = H'/\ln[S] \)) were calculated from the average cover of each species across all 12 subplots within a plot.

**Resource heterogeneity**

Nitrate availability (NO₃-N) was quantified in all subplots using buried ion exchange resin bags (Binkley and Matson 1983). The collection of ions onto exchange resins occurs passively through movement of soil water. We chose to measure nitrate because it is more mobile in soil with a high cation exchange capacity relative to ammonium. Bags were constructed with nylon and contained 5 g of anion exchange resin (Dowex 1X8-50; Dow Chemical, Midland, Michigan, USA) preloaded with Cl⁻. Two bags were buried ~10 cm deep in each subplot (opposite of species composition quadrats) in June and retrieved in September each year (Baer et al. 2003, Baer and Blair 2008, Baer et al. 2016). In the laboratory, resin bags were rinsed with deionized water then extracted using a 5:1 ratio of 2 mol/L KCl:resin on an orbital shaker (200 rpm) for one hour. Solutions were filtered through 0.4-μm polycarbonate membrane filters and extracts were analyzed for NO₃-N using an OI Analytical Flow Solution IV autoanalyzer (OI Analytical, College Station, Texas, USA).

Light availability at the soil surface was determined by measuring photosynthetically active radiation (PAR) above and below the canopy with a 50-cm ceptometer (Decagon Devices, Pullman, Washington, USA) at the same time species composition was collected. Measurements (\( n = 5 \) above and below the canopy) were taken during the growing season, which sustained reduced-N conditions (Baer et al. 2003). Starting in 2005, the reduced N treatment has been maintained by applying 84.2 g C/m² (sucrose-C) three times each growing season, which sustained reduced-N conditions (Baer and Blair 2008, Baer et al. 2016). Elevated-N strips have received 5 g N/m² in the form of NH₄-N in early June every year since 1998. The factorial combination of N and soil depth treatments resulted in 12 subplots that varied in soil N level and soil depth in each maximum heterogeneity plot. The control plots contained no subplot treatments.

In 1998, all plots were sown with 42 native prairie species using a log-normal distribution of dominant grasses and subordinate forbs to resemble never-cultivated tallgrass prairie (Baer et al. 1999, 2003). In 2005, seeds of 15 additional forb species were sown into all plots at a rate of 25 live seeds/m² (Baer et al. 2016). Beginning in 2013, we added 17 additional species (14 forbs, 2 grasses, and 1 sedge) that were never previously recorded in this experiment, but occur at KPBS, to all plots (Appendix S1: Table S1). Seeds were purchased from Prairie Moon Nursery (Winona, Wisconsin, USA). Percent live seed was either provided by the seed supplier or determined by the Illinois Seed Testing Lab. Seeds were hand-broadcasted in spring each year at a rate of 20 live seeds/m² per species for a total seeding rate of 300 live seeds/m². The experimental area was fenced in 1998 to prevent deer browsing and has been burned almost annually in the early spring (a common management practice in this region) since 1998, with the exception of 2000 and 2003.
and averaged in each species composition sampling quadrat in two perpendicular directions. Percent available PAR at the soil surface was calculated as \( \frac{\text{PAR}_{\text{soil surface}}}{\text{PAR}_{\text{above canopy}}} \times 100 \).

**Plant functional traits (PFTs)**

The same species added to the field experiment starting in 2013 were grown in Conviron CMP 6050 growth chambers (Conviron, North Branch, Minnesota, USA) in the Southern Illinois University Horticulture Research Greenhouse (Southern Illinois University, Carbondale, Illinois, USA). We quantified plant functional traits following the methods used by Tucker et al. (2011). Growth chamber conditions were 16 h day length from 06:00 to 22:00 at 25°C with 1,200 μmol light intensity and 20°C at night from 22:00 to 06:00. Plants were watered twice daily, once at 07:30 and once at 16:30 and treated biweekly with commercial fertilizer (Scotts-Miracle-Gro 24-8-16 [N-P-K] All Purpose Fertilizer; ScottsMiracle-Gro, Marysville, Ohio, USA) to eliminate nutrient stress. We used soil collected from the restoration site, but outside of the experimental plots. Twelve replicates of each species were planted in plastic cone-tainers (D-40; Stuewe and Sons, Corvallis, Oregon, USA). Several seeds of the same species were planted in each cone-tainer, and the first emerging seedling was used for PFT measurements (the rest were discarded). We used 8 of the 12 replicates of each species and PFTs were measured 8–9 weeks after the seedlings emerged.

We measured a suite of plant traits indicative of growth rate and resource use (Weiher et al. 1999, Funk et al. 2008, Tucker et al. 2011). Traits measured included stem mass fraction (stem dry mass per module dry mass), shoot length, longest internode length, maximum leaf length, and average specific leaf area (SLA), and nitrogen acquisition indicated by leaf nitrogen concentration and nitrogen use efficiency (NUE; Gubsch et al. 2011). Gas exchange was measured on the youngest fully expanded leaf using a Li-Cor LI-6400 Portable Photosynthesis System (LICOR Biosciences, Lincoln, Nebraska, USA) in the growth chamber. Plant height was used for shoot length. The upper three internodes were measured to determine the longest internode length on plants with measurable internodes. Specific leaf area (SLA) was measured by scanning fresh fully expanded leaves (n = 3 leaves per individual) and leaf area was then calculated using leaf area measurement software v.1.3 (A.P. Askew 2003, The University of Sheffield, Sheffield, UK). Leaf area was divided by the dry mass of the leaf to determine SLA. Leaf nitrogen (N) concentration was measured on dried ground leaves using a Thermo Scientific Flash CNHOS Elemental Analyzer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Percent N was then multiplied by the total dry leaf mass to obtain the leaf N mass. Nitrogen use efficiency (NUE) was obtained for healthy completely expanded leaves by dividing rate of gas exchange measured using the Li-COR 6400 by leaf N. Aboveground biomass from each cone-tainer was clipped and dried at 60°C for 5 d before weighing.

We also included seed traits of dry seed mass and seed moisture content because they are important for germination and survival. Seed moisture content was determined by drying 30 crushed seeds at 105°C for 24 h, cooling in a desiccator for 40 min, weighing, and applying the International Seed Testing Association formula (ISTA 2006).

**Statistical analysis**

Plot-level heterogeneity effects (HETTRT) on variation in resources (coefficients of variation [CV] in NO₃-N and light availability), Shannon’s diversity, species richness, evenness, grass and forb cover, forb richness, and forb diversity from 2013 to 2017 were analyzed according to a randomized complete block design with repeated measures using the mixed model procedure in SAS 9.4 (SAS Institute, Cary, North Carolina, USA). Block was assigned as a random effect and year as the repeated measure. The least squares means procedure was used to compare main effect means and contrast statements were used to test for differences between the heterogeneity and control treatments within each year (α = 0.05).

Separate (subplot-level) mixed model analyses were performed in SAS (SAS 2014) using only the maximum heterogeneity plots to examine the main effects and interaction between nitrogen (NUT) and soil depth (DEPTH) on relative NO₃-N availability, percent light availability at the soil surface, total diversity, total richness, and cover and richness of new species added from 2013 to 2017. These data were analyzed according to a split-block design with year as a repeated measure. Subplots within each NUT and DEPTH level were assigned to vertical (VS) and horizontal strips (HS), respectively (Baer et al. 2003, 2016). Block, VS(NUT), and HS (DEPTH) were assigned as random effects in the model. Because there were no significant three-way interactions (DEPTH × NUT × YEAR), or interactions between DEPTH and YEAR, we used contrast and estimate statements to compare nutrient means within a year if there was an interaction between NUT and YEAR. The least squares means procedure was used to compare main effect means (α = 0.05). Light availability at the soil surface was analyzed by year.

For all repeated-measures analyses, we used the Kenwood-Rogers method to estimate degrees of freedom. We ran each analysis with compound symmetry (CS), autoregressive (AR), and unstructured (UN) covariance structures, and selected the analysis with the lowest Akaike information criterion (AIC; Littell et al. 2006). The covariance structure accompanies each F value presented in the results, followed by numerator and denominator degrees of freedom. For many response variables, there was a significant main effect of year across plot or...
subplot treatments. Year main effects are only described in the results if they showed a clear pattern (increasing or decreasing over time).

We used PERMDISP (Anderson et al. 2006) on composition data to determine if spatial heterogeneity of vegetation varied between the control and maximum heterogeneity treatments during the 5-yr study. PERMDISP uses Bray-Curtis dissimilarity to measure the spatial dispersion of samples around the group centroid. Larger values reflect increasing dissimilarity (higher heterogeneity) among samples. We conducted PERMDISP analyses at two levels of resolution. At the first level, we compared dissimilarity among all subplots across all replicate control or maximum heterogeneity plots to determine an overall treatment effect each year using a t test. At the second level, we calculated dissimilarity among all subplots within each replicate of each treatment and compared differences between control and maximum heterogeneity plots (n = 4) using ANOSIM. PERMDISP analyses were run using PRIMER-6 (Clarke and Gorley 2006).

Comparison of added species that colonized between the heterogeneity treatments was performed with the model-based multivariate approach of the mvabund package (Wang et al. 2012) in R (R Core Team 2018). A presence–absence (binomial) response to heterogeneity (whole-plot) treatment was fit for each species individually and all species jointly. Significance was tested with likelihood ratio tests with a null model (100 bootstraps).

Plant functional and seed traits were used to create a trait space, with each species assigned to a functional group (C3 grass, C4 grass, forb, and sedge). The degree of overlap of trait space among species was assessed using nonmetric multidimensional scaling (NMDS) based on Bray-Curtis dissimilarity measures. NMDS is a robust ordination technique that provides a visual representation of the similarities of the variables being compared (Minchin 1989). Trait spaces were compared using analysis of similarity (ANOSIM). The NMDS and ANOSIM analyses were conducted using DECODA 3.01 software (Minchin 1989) on transformed data (log [x + 1]). We used α = 0.004 to determine significance for multiple comparisons accounting for 12 species; if species were significantly different, they were considered to occupy different trait spaces. Fitted vectors represent Pearson correlation coefficients between the measured plant functional trait variables and the NMDS axes.

**RESULTS**

*Treatment effects on resource variability and availability*

The soil treatments increased heterogeneity in N availability (Table 1). The CV of resin-collected NO3-N was one to three orders of magnitude higher in the maximum heterogeneity than control plots, with some years exhibiting more disparity between the treatments than others, as reflected by an interaction between HETTRT and YEAR (F4.24 (AR1) = 15.5, P < 0.001). Higher variability in NO3-N availability in the maximum heterogeneity plots was due to lower NO3-N availability in the reduced-N (C-amended) treatment and highest NO3-N availability in the enriched-N treatment in all years, with the exception of similar NO3-N availability between ambient-N and reduced-N soil in 2015 (NUT × YEAR interaction: F8.139 (AR1) = 12.5, P < 0.001). There was no effect of the soil depth treatment on nitrate availability.

The heterogeneity treatment effect on the CV of light availability was marginally significant (F1,10.1 = 3.79; P = 0.080), and the difference between the CVs over all years (26.4 vs 21.8 in the maximum heterogeneity and control plots, respectively) was small (Table 1). Despite this, the depth and nutrient treatments interacted to affect light availability in 2013 (F2,19.6 = 7.8; P = 0.003) and 2014 (F2,18.7 = 5.14; P = 0.017), and there was a main effect of nutrient treatment on light availability in 2015 (F2,24.9 = 7.4; P = 0.003) and 2017 (F2,4.7 = 11.2; P = 0.016). The only consistent pattern among all years when the subplot treatments significantly affected light availability was higher light availability in deep soil under reduced-N conditions compared to deep soil under enriched-N conditions (Table 1).

**Heterogeneity effects on plant community structure**

Following 20 yr of community assembly, we recorded a total of 49 species in the maximum heterogeneity and control plots, with 40 and 34 species occurring in each treatment, respectively. There were 15 species found only in the maximum heterogeneity plots and 9 species found only in the control plots, with 25 species occurring in both treatments. Total diversity and richness began to diverge between the maximum heterogeneity and control treatments 16 to 20 yr post-restoration (Fig. 1), concurrent with propagule addition. Despite variation in diversity among years during this period (YEAR main effect: F4.23.1 (AR1) = 7.57, P < 0.001), there was no interaction between HETTRT and YEAR (F4.23.1 (AR1) = 0.63, P = 0.649). Diversity was higher in the maximum heterogeneity plots across years 16 through 20 (HETTRT: F1.5.38 (AR1) = 8.16, P = 0.033) (Fig. 1A,B). Richness showed a similar response, with significant variation among years (YEAR: F4.21.8 (AR1) = 7.94, P < 0.001), no interaction between HETTRT and YEAR (F4.21.8 (AR1) = 0.42, P = 0.792), and higher richness in the maximum heterogeneity treatment relative to the control over the last 5 yr (HETTRT: F1.11.8 (AR1) = 4.87, P = 0.048; Fig. 1B,C). Evenness also varied among years (YEAR: F4.24 (CS) = 12.67, P < 0.001), ranging from 0.41 in 2013 to 0.48 in 2017 and oscillating between these years (data not presented). There was no effect of heterogeneity on evenness (HETTRT: F1.6 (CS) = 2.15, P = 0.193) or interaction between HETTRT and YEAR (F4.24 (CS) = 2.22, P = 0.100).
Cover of established grasses and forbs varied between the heterogeneity treatments. From 2013 to 2017, grass cover was 12.5% lower in the maximum heterogeneity treatment (56.9% ± 1.2%) relative to the control (65.1% ± 2.9%) treatment across all years (HETTRT: \( F_{1,3} \) (CS) = 15.94, \( P = 0.023 \)). Forb cover exhibited an interaction between HETTRT and YEAR (\( F_{4,24} \) (CS) = 2.82, \( P = 0.047 \)) resulting from a slow development of higher forb cover in the maximum heterogeneity treatment relative to the control over time (P values corresponding to maximum heterogeneity vs. control contrasts: 2013 = 0.541; 2014 = 0.404; 2015 = 0.248; 2016 = 0.209; and 2017 = 0.023 (Fig. 2A). Forb richness and diversity also began to respond consistently to the heterogeneity treatment. Over the last 5 yr, forb richness and diversity were 17% and 26% higher in the maximum heterogeneity treatment compared to the control (HETTRT forb richness: \( F_{1,9.2} \) (AR1) = 9.96, \( P = 0.011 \); HETTRT forb diversity: \( F_{1,6} \) (CS) = 6.76, \( P = 0.041 \); Fig. 2B–E).

The maximum heterogeneity plots became more compositionally different from one another over time relative to the control plots (Appendix S1: Table S2). Greater dispersion of the community among the maximum heterogeneity plots was due to less dominance by \textit{Andropogon gerardii} Vitman (big bluestem) and more variation in dominant and codominant species from various sources (sown in 1998, added in 2005, and natural colonization from the regional species pool) in the N-enriched subplots. In 2017, for example, different species (\textit{Ambrosia psilostachya} DC., \textit{Eupatorium altissimum} L., and \textit{Asclepias verticillata} L.) dominated or codominated different species in three of the four N-enriched subplots in Block 1. In Block 2, the codominant species, based on average percent cover in the N-enriched strip, were \textit{Salvia azurea} Michx. ex Lam. (26.5% ± 7.4%) and \textit{A. gerardii} (19.4% ± 5.0%). In Block 3, the N-enriched strip was dominated by \textit{Teucrium canadense} L. (30.4% ± 9.7%) and \textit{A. gerardii} (26.9% ± 8.1%). In Block 4, different species (\textit{Ambrosia psilostachya} DC.}

---

**Table 1.** Coefficients of variation (CV) in resin-collected NO\(_3\)-N and light availability in the maximum and control heterogeneity treatments each year and availability of NO\(_3\)-N and light in response to the soil treatments in the maximum heterogeneity plots each year.

<table>
<thead>
<tr>
<th>CV in resin-collected NO(_3)-N</th>
<th>Plot-level treatment responses</th>
<th>Subplot-level treatment responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous</td>
<td>Control</td>
<td>Reduced N</td>
</tr>
<tr>
<td>2013</td>
<td>1,247.0% ± 250.8%</td>
<td>7.40 ± 3.21</td>
</tr>
<tr>
<td>2014</td>
<td>1,499.0% ± 473.0%</td>
<td>6.41 ± 1.90</td>
</tr>
<tr>
<td>2015</td>
<td>4,027.7% ± 1,052.2%</td>
<td>0.97 ± 0.67</td>
</tr>
<tr>
<td>2016</td>
<td>3,537.3% ± 910.1%</td>
<td>0.67 ± 0.58</td>
</tr>
<tr>
<td>2017</td>
<td>2,109.5% ± 722.5%</td>
<td>0.93 ± 0.29</td>
</tr>
<tr>
<td>CV of light availability†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>34.0% ± 4.38%</td>
<td>27.5% ± 6.82%</td>
</tr>
<tr>
<td>2014</td>
<td>20.7% ± 1.58%</td>
<td>17.3% ± 3.63%</td>
</tr>
<tr>
<td>2015</td>
<td>35.0% ± 3.78%</td>
<td>39.5% ± 2.80%</td>
</tr>
<tr>
<td>2016</td>
<td>14.5% ± 1.72%</td>
<td>36.1% ± 3.07%</td>
</tr>
<tr>
<td>2017</td>
<td>27.9% ± 2.70%</td>
<td>38.5% ± 1.92%</td>
</tr>
</tbody>
</table>

**Notes:** Within a year, means accompanied by the same letter were not significantly different (\( \alpha = 0.05 \)). All values are mean ± SE.

† At the soil surface.
and Panicum virgatum L.) were equivalent or second in maximum cover to A. gerardii, respectively. In contrast, A. gerardii was the dominant species in 100% of the subplots within the control plots. Soil treatment effects on the plant community

Diversity was affected by an interaction between DEPTH and NUT ($F_{2,18.9}$ (UN) = 6.23, $P = 0.008$) resulting from higher diversity in shallow soil relative to deep soil under ambient N conditions and higher diversity in enriched-N soil relative to ambient and reduced-N conditions in deep soil (Fig. 3A). Over the last 5 yr of study, A. gerardii had 31–40% less cover in deep N-amended soil relative to all other treatments (NUT × DEPTH interaction: $F_{2,84}$ (AR1) = 29.5, $P = 0.001$) and total richness was 20% and 51% higher in the enriched-N soil relative to the ambient-N and reduced-N soil treatments, respectively (NUT: $F_{2,27.1(CS)} = 4.90$, $P < 0.015$; Fig. 3B).

Colonization of new species

Twelve of the 17 species added starting in 2013 were recorded during the 5-yr period. Not counting unidentifiable seedlings, eight new species were recorded from control plots and seven new species were recorded from the maximum heterogeneity plots. Of the eight species that established from the 2005 propagule addition, four occurred only in the maximum heterogeneity plots and the maximum heterogeneity plots contained twice as many species from this supplemental seed addition than

**Soil treatment effects on the plant community**

Diversity was affected by an interaction between DEPTH and NUT ($F_{2,18.9}$ (UN) = 6.23, $P = 0.008$) resulting from higher diversity in shallow soil relative to deep soil under ambient N conditions and higher diversity in enriched-N soil relative to ambient and reduced-N conditions in deep soil (Fig. 3A). Over the last 5 yr of study, A. gerardii had 31–40% less cover in deep N-amended soil relative to all other treatments (NUT × DEPTH interaction: $F_{2,84}$ (AR1) = 29.5, $P = 0.001$) and total richness was 20% and 51% higher in the enriched-N soil relative to the ambient-N and reduced-N soil treatments, respectively (NUT: $F_{2,27.1(CS)} = 4.90$, $P < 0.015$; Fig. 3B).

**Colonization of new species**

Twelve of the 17 species added starting in 2013 were recorded during the 5-yr period. Not counting unidentifiable seedlings, eight new species were recorded from control plots and seven new species were recorded from the maximum heterogeneity plots. Of the eight species that established from the 2005 propagule addition, four occurred only in the maximum heterogeneity plots and the maximum heterogeneity plots contained twice as many species from this supplemental seed addition than
FIG. 2. Plot-level (A) forb cover each year, (B) forb richness each year, (C) forb richness over the last 5 yr, (D) forb diversity each year, and (E) forb diversity the last 5 yr in the maximum heterogeneity and control treatments. Asterisks indicate significant differences ($P \leq 0.05$) between control and maximum. All values are mean and SE. *, **, and *** for $P \leq 0.05$, 0.01, and 0.001.
the control plots (6 vs. 3, respectively). There was a low frequency of occurrence (<2.2% of the subplots recorded over all 5 yr) of all but one species, Mirabilis nyctaginea (Michx.) MacMill. (15% of subplots). Colonization of any or all new species was not influenced by heterogeneity (bootstrap likelihood ratio test, $P > 0.05$), but the number of plots colonized varied by species.

Although new species comprised < 2% of total plant cover, the cover of species sown (starting in 2013 through 2017) in the plot-level heterogeneity treatments exhibited an interaction with time (HETTRT × YEAR: $F_{4,17.7}^{(AR1)} = 2.95$, $P = 0.049$). There was no difference in cover of new species between the heterogeneity treatments from 2013 through 2015, but their cover became higher in the maximum heterogeneity treatment compared to control in 2016 and 2017. Although the cover of new species added in the enriched-N soil was more than two times higher than the ambient-N and reduced-N treatments, the number of species/subplot averaged across years was <1 in all nutrient treatments.

**Plant functional traits**

Of the 17 sown species grown in the growth chamber, only 12 grew to maturity. A two-dimensional ordination using plant functional traits was generated with NMDS (2D stress = 0.13; Fig. 4). All species occupied a different trait space in the ordination (ANOSIM, $P < 0.002$). There was no evidence that species occupying a wider range of trait spaces colonized the maximum heterogeneity plots (data not presented). In other words, seedlings that matured enough to be confidently identified were not clustered or more dispersed in the control and maximum heterogeneity treatments, respectively.

**DISCUSSION**


Because landscape heterogeneity corresponds with high floristic diversity of tallgrass prairie (Seastedt et al. 1991, Collins et al. 1998, Collins and Calabrese 2012), we predicted that communities established under greater variation in soil resource availability would develop more heterogeneous communities that, in turn, would lead to divergence in species richness and diversity over time and create more niche space for new species to exploit. Although Williams and Houseman (2014) showed a positive short-term effect of soil heterogeneity on species richness in restored prairie, many experimental tests of the EHH in plant communities, including this experiment (Baer et al. 2005, 2016), provide underwhelming support for the EHH (reviewed by Lundholm 2009). After 20 yr of soil manipulation in this experiment, plant diversity and richness began to diverge, with higher richness and diversity emerging consistently in the maximum heterogeneity treatment. Higher diversity and richness in the heterogeneous treatment resulted from (1) higher cover, richness, and diversity of established forbs averaged across all subplots; (2) higher richness in the N-enrichment subplots within the maximum heterogeneity plots; (3) higher diversity in shallow soil relative to deep soil under ambient N and reduced-N conditions; and (4) lower cover of a dominant grass species in deep fertilized soil.

Nutrient enrichment and soil depth are known to increase and reduce the cover of C4 grasses in tallgrass prairie, respectively, and in this way indirectly determine diversity via more and less competition with subordinate species. Nitrogen is a limiting nutrient in tallgrass prairie (Blair 1997), demonstrated by positive productivity responses to added N (Tilman 1987, Seastedt et al. 1991) that typically correspond initially with a reduction in species richness and diversity (Jacquemyn et al. 2003, Baer et al. 2004, Clark and Tilman 2008, Hautier et al. 2009) owing to a reduced variety of niches (niche dimensionality; Harpole and Tilman 2007). Species loss in response to nutrient enrichment is a common phenomenon (Suding et al. 2005) that may be difficult to reverse if alternative stable states develop (Isbell et al.

![Nonmetric multidimensional scaling ordination of 12 prairie species added to the field experiment starting in 2013. Different ordination (trait) spaces were determined using ANOSIM, all species had different trait spaces (\(\alpha = 0.004\)). Vector lengths correspond to variance in axis scores explained by each environmental variable.](image_url)
Our multi-decadal study of community response to nutrient enrichment thus far demonstrates dynamic transient community state changes. For example, during the first 3 yr of restoration, there was a precipitous decline in species richness and an increase in grass dominance while richness remained higher and grass cover lower in reduced-N soil (Baer et al. 2003). Over the next 10 yr, a strong negative correlation between cover of *Andropogon gerardii* and species richness developed (Baer et al. 2016). After 15 yr of community assembly, higher richness in reduced-N soil proved to be short lived and species richness was not consistently lower in the N-enriched soil relative to the other soil treatments in later years because grass cover declined with persistent N addition (Baer et al. 2016).

Diversity became consistently higher under nutrient enrichment in this developing prairie only after 15 yr of N addition. This response was due to higher forb cover and richness under nutrient addition as grass cover declined. Avolio et al. (2014) also found that long-term N addition promoted fast-growing forbs with low N-use efficiency that replaced the dominant C₄ grasses under high N availability over the long-term in tallgrass prairie. Other studies have shown long-term N addition has potential to decrease temporal stability of grassland communities as dominance decreases (Hautier et al. 2014, Zhang et al. 2016). In our experiment, nutrient enrichment produced strong shifts in community structure over time, starting with early dominance by *Panicum virgatum* L. (Baer et al. 2003), followed by increasing dominance of *Andropogon gerardii* (Baer et al. 2016), and eventually different codominant species among N-enriched subplots as grass cover declined. This suggests that community assembly processes shift from strongly deterministic to stochastic (more randomly driven by the regional species pool) over the long-term under nutrient enrichment, supported by greater variation (dispersion) in community composition among the maximum heterogeneity plots relative to each other than composition in the controls plots relative to each other.

One possible explanation for dynamic community change and increase in species richness under long-term nutrient enrichment may be a corresponding change in the community composition and richness of arbuscular mycorrhizal fungi (AMF). The rationale for this proposed mechanism is based on a previous study at Konza Prairie that showed higher richness of AMF in long-term fertilized vs. unfertilized native grassland soil (Egerton-Warburton et al. 2007) coupled with the generally positive relationship between AMF and plant biodiversity (van der Heijden et al. 1998) and the key role AMF play in promoting plant species coexistence (Crawford et al. 2019).

High species richness in native tallgrass prairie is primarily a function of forb richness and cover, and as such, ecological drivers that reduce grass dominance in native prairie support higher plant diversity (Gibson and Hulbert 1987, Collins et al. 1998, Collins and Calabrese 2012, Manning et al. 2017). Diversity and forb richness are also inversely related to the cover of dominant grasses in restored prairie (McCain et al. 2010, Baer et al. 2016). Because fire and grazing are difficult to manipulate on small scales, we imparted the shallow soil treatment, based on the observation that there is less cover of C₄ grasses and higher richness in shallow upland soils in surrounding native prairie (Gibson and Hulbert 1987). We expected the community response to this manipulation to develop over many years. The interaction that occurred between nutrient and soil depth (i.e., higher diversity in shallow soil in only one nutrient treatment) during the first 15 yr of community assembly (Baer et al. 2016) persisted for five additional years. Interestingly, diversity was similar in shallow soil under all nutrient regimes, as was the cover of *A. gerardii*. Baer et al. (2016) previously documented that differences in the cover of *A. gerardii* among nutrient treatments only occurred in deep soil. Thus, less effect of N-availability on species composition in shallow soil appears to be persistent and suggests that stochastic processes have less influence on community assembly in shallow soil under a potentially different limiting resource (e.g., soil moisture).

In the absence of continuous propagule supply, restored communities can develop distorted species composition relative to target assemblages (Howe 1999, Maina and Howe 2000). Limited propagules or bud banks of subordinate species can constrain diversity and richness in grassland (Foster et al. 2004, Dalgleish and Hartnett 2009, Willand et al. 2013). Further, several studies indicate that plant diversity response to heterogeneity interacts with dispersal, demonstrated by higher diversity in more heterogeneous environments when propagules are supplied (Coulson et al. 2001, Foster et al. 2004, Questad and Foster 2008, Baer et al. 2016). As such, we added propagules of new species to all treatments, not to test dispersal limitation, but to determine if resource heterogeneity increased niche availability for new species to exploit.

We quantified trait spaces of newly added species to reveal community assembly processes (van der Plas et al. 2015). More specifically, we used trait analyses to elucidate the relative strength of environmental filtering in homogeneous and heterogeneous conditions. We expected greater clustering of species recruited into more homogeneous communities established under more homogeneous environmental conditions in the trait ordination space, indicative of niche availability for a narrower range of traits relative to heterogeneous conditions. Despite higher recruitment of newly sown species in the maximum heterogeneity plots, the magnitude of this response was low and these species did not collectively occupy a larger trait space, suggesting a strong filtering process (Foster et al. 2004, Ackerly and Cornell 2007, Dickson and Foster 2008, Grman et al. 2015). The most frequent establishment by *Mirabilis nytoginea*, suggests traits associated with light capture (e.g., longest internode,
shoot length, and stem mass fraction) are important traits for new species to recruit into established communities. The very low occurrence of all but one species demonstrates establishment limitation for a wide range of species and niches regardless of environmental heterogeneity. Recruitment of new species could also take longer than 5 yr, as it took >10 yr for many species supplied as propagules in 2005 to occupy more than 5% of the total cover (see Data Availability).

**CONCLUSION**

Although the “environmental heterogeneity hypothesis” is a widely accepted mechanism for species coexistence (Kolasa and Pickett 1991), this study demonstrated that long-term manipulation may be required for environmental variation in resources to increase plant diversity, particularly during ecosystem recovery following long-term disturbance. We attribute the slow-to-emerge diversity response to environmental heterogeneity to (1) initial dominance of clonal grasses in this experiment (Baer et al. 2016), (2) time required for community destabilization to occur in nutrient enriched patches, and (3) time required for roots and competition for belowground resources to develop in the shallow soil treatment. Long-term community response to N-addition in this experiment runs counter to the general phenomenon of lower species diversity with nutrient enrichment, particularly in grasslands (Bobbink et al. 2010, Harpole et al. 2016), but supports the proposition that global change drivers, such as N enrichment, can alter spatial heterogeneity in ecological communities (Avolio et al. 2015). This agrees with spatially variable change in community structure in nutrient-amended native grassland (Koerner et al. 2016). Change from deterministic to stochastic community assembly processes has also been demonstrated in successional sequences along a multi-century arable-to-grassland chronosequence (Purschke et al. 2013). Our multi-decadal study suggests that heterogeneity that includes nutrient enrichment might hasten this natural phenomenon, but we do not advocate adding nutrients to restorations at large scales. Means to increase patchy resource heterogeneity at large scales and consequences for achieving the composition of species practitioners strive to restore deserves further investigation.

**ACKNOWLEDGMENTS**

Funding for this research was provided by the National Science Foundation (IBN 9603118; DEB-1922915) with support from the Konza Prairie Long-Term Ecological Research Program. Laboratory and field assistance were provided by G. Manning, S. Black, J. Willand, P. Harris, J. Fitzgibbon, and A. Rothert.

**LITERATURE CITED**


Minchin, P. R. 1989. DECODA user’s manual: Research School of Pacific Studies. Australian National University, Canberra, Australian Capital Territory, Australia.


**Supporting Information**

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/eap.2014/full

**Data Availability**

Data sets generated and analyzed for development of this manuscript are available in the Dryad Digital Repository at https://doi.org/10.5061/dryad.63n980b and the Konza Prairie LTER data portal at https://doi.org/10.6073/pasta/28ce07278347b504fbb956a9011ac70.