Soil microbial community composition in tallgrass prairie restorations converge with remnants across a 27-year chronosequence

Nicholas A. Barber,1,2 Karley M. Chantos-Davidson,1 Rene Amel Peralta,3 Jared P. Sherwood4 and Wesley D. Swingley1,2*
1Department of Biological Sciences, Northern Illinois University, DeKalb, IL, USA.
2Institute for the Study of the Environment, Sustainability, and Energy, Northern Illinois University, DeKalb, IL, USA.
3Great Minds in STEM, Monterey Park, CA, USA.
4University of San Francisco, San Francisco, CA, USA.

Summary
Restoration and management of natural ecosystems is a critical strategy in mitigating global biodiversity loss. This is exemplified in the American Midwest by efforts aimed at reclaiming historical grasslands lost to high-yield agriculture. While restorations traditionally take the form of plant reintroduction and management, advances in microbial analyses suggest that soil communities could be indicators of restoration success. However, current understanding of key microbial taxa and functional activities in both natural and restored ecosystems is limited. Here, we investigated the impact of nearly 30 years of carefully managed restoration on soil microbial communities at the Nachusa Grasslands in northern Illinois, USA. We characterized bacterial and archaeal communities in a chronosequence of restored tallgrass prairies ranging from 1 to 27 years old across a growing season and compared them to communities in pre-restoration agricultural fields and remnant prairies. Results indicate that older restorations harboured communities statistically distinct from newer restorations. These communities converged toward those in local prairie remnants, suggesting that plant-focussed restoration has yielded soil bacterial communities reflective of a successful restoration.

Recovery of microbial clades within the Verrucomicrobia and Acidobacteria are an important feature of this convergence, and these groups could be targeted for future soil-focussed, bottom-up restoration studies.

Introduction
Ecosystem restoration has become an important strategy in mitigating and reversing biodiversity loss (Palmer et al., 1997; Benayas et al., 2009; Bullock et al., 2011). The goal of restoration is to establish diverse communities that carry out and contribute to desirable ecosystem functions (Hobbs and Harris, 2001; Thorpe and Stanley, 2011). Traditional restoration efforts have focussed on the introduction, maintenance and monitoring of native macro-organisms, particularly plants (Hobbs and Norton, 1996; Van Andel and Aronson, 2012; McAlpine et al., 2016), although investigations of belowground organisms have revealed that soil microbial communities can play important roles in the pace and trajectory of recovery in ecosystem restoration (Bever et al., 2003; Middleton and Bever, 2012; Wubs et al., 2016). The advent of modern sequencing technologies now allows researchers to delve deeper into the complex soil communities that drive nutrient cycling in both undisturbed and restored ecosystems (e.g., Leff et al., 2015; Calderón et al., 2016; Nelson et al., 2016). Interactions between above- and belowground ecosystem compartments are predicted to be driven largely by both direct and indirect influences of plants on nearby soils (changes in pH, organic matter, soil texture, etc.), coupled with microbially-mediated elemental cycles (N, C, S, etc.) and the direct interactions between plants and other soil organisms including both pathogens and mutualists like mycorrhizal fungi (Bardget and Wardle, 2010). However, more foundational work examining microbial community dynamics and the specific geochemical influences of microbial taxa is needed before restorations can be driven from the ground up, successfully coordinating traditional plant and animal management with soil microbiome manipulation, maintenance and monitoring (Chaparro et al.,...
We are only beginning to understand how ecological, geochemical and physicochemical parameters directly influence microbial communities in natural systems, and it is an open question whether any unifying principles or overarching themes derived in natural systems can be applied to restored ecosystems (Fierer et al., 2009), which likely involve more transient processes as they progress from degraded states toward targeted goals. The paucity of common variables in the wide variety of ecosystems under study has confounded the identification of primary ecological/geophysical drivers of microbial diversity—e.g. the effect of moisture on microbial diversity (Lauber et al., 2008; Griffiths et al., 2011; Brockett et al., 2012; Kuramae et al., 2012; Lauber et al., 2013; Pan et al., 2014; O’Brien et al., 2016). Studies constrained to systems with strong similarities (ecosystem type, climate, plant diversity, etc.) have yielded more consistent patterns, identifying, for example, specific members of the bacterial phyla Acidobacteria, Proteobacteria, Verrucomicrobia and others regularly linked with undisturbed terrestrial grasslands and other members of the Proteobacteria and phylum Gemmatimonadetes with cultivated soils (McNamara et al., 2006; Bergmann et al., 2011; Griffiths et al., 2011; Nacke et al., 2011; Van Trump et al., 2011; Fierer et al., 2012; da Rocha et al., 2013; Fierer et al., 2013; Carbonetto et al., 2014; Pan et al., 2014). The causes of these associations are still poorly understood, in part due to the dearth of knowledge on these phyla, which contain primarily uncultivable species (Bergmann et al., 2011; DeBruyn et al., 2011; da Rocha et al., 2013), highlighting the importance of continuing to characterize environmental strains in natural and laboratory settings.

North American tallgrass prairies represent one of the most widely destroyed and degraded ecosystems on Earth, with >90% of original prairie lost, and in some regions such as Illinois, >99% (Samson and Knopf, 1994; Samson and Knopf, 1996). Tallgrass prairie was largely converted to row crop agriculture to take advantage of the deep, productive soils. Starting in the late 20th century, restoration efforts targeting these lost ecosystems have become widespread, representing a rapidly growing percentage of total prairie habitat. Restoration activities typically include the re-establishment of diverse plant communities, control of exotic plant species, and implementation of regular activities that mimic historic disturbance regimes (fire, and in some cases grazing, including reintroduction of American bison (Bison bison)).

Despite the importance of microbial communities in prairie soil nutrient cycling and the knowledge that conversion to agriculture has significantly altered these communities (Lauber et al., 2008; Jiang et al., 2010; Van Trump et al., 2011; Kuramae et al., 2012; Fierer et al., 2013; Lauber et al., 2013; Suleiman et al., 2013; Carbonetto et al., 2014; Mendes et al., 2015), the soil response to prairie restoration is poorly understood. Although an implicit or explicit goal of most restoration projects is to mimic reference sites, it is not known whether restoration soil microbial community compositions converge with reference sites over time or what the pace of this convergence may be (Allison et al., 2005; Bach et al., 2012; Suleiman et al., 2013). Here we use a 27-year chronosequence of high-quality tallgrass prairie restoration sites to examine the recovery of soil bacterial communities. We examined community and soil environment changes in comparison to local pre-restoration (row crop) and remnant sites. We document these changes over time, both across the chronosequence as well as within a growing season, to understand relationships between short-term turnover of seasonal patterns and longer-term successional changes across nearly three decades.

Results

From April through September, we repeatedly sampled soil bacterial and archaeal communities from a focal set of five restorations, one remnant and a corn field currently under cultivation. In September, we also sampled eight additional restorations, one other remnant, and a soy field for a larger cross-sectional analysis to validate results from the focal sites. From each soil sample, we extracted and sequenced DNA, analysed community structure and measured edaphic variables (see Methods).

16S rRNA gene sequencing

Illumina sequencing yielded an average of ~109 000 16S rRNA genes per sample (each sample being a site/date combination) eight samples at each focal site and 1 at each annual site) after paired-end assembly and quality control filtering, with the smallest sample containing 61 909 sequences (Table 1). On average 90.6% ± 2.1% of 16S rRNA gene sequences at each site were confidently assigned at the Phylum level, with very little variation between assignments at newly restored prairies and older or remnant prairies. Rarefaction analysis on these samples suggested that the sequencing depth adequately captured community diversity at the phylum and class level, the latter of which is the primary focus of discussion in this work, but did not capture the full range of diversity at the genus level (Supporting Information Fig. S1). All sequence data can be accessed on NCBI using project number (PRJNA382854).

Community metrics and composition

Shannon diversity, Pielou’s evenness and taxonomic richness calculated using class-level taxonomy all declined
Results in bold are significant at $P < 0.05$.

© 2017 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*, **00**, 00–00

---

Table 1. Statistics of 16S rRNA gene sequencing and taxonomic classification.

<table>
<thead>
<tr>
<th></th>
<th>Agricultural field</th>
<th>1 year old</th>
<th>2 years old</th>
<th>3 years old</th>
<th>13 years old</th>
<th>27 years old</th>
<th>Remnant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total rRNA Gene Sequences</td>
<td>986 111</td>
<td>805 602</td>
<td>716 749</td>
<td>885 154</td>
<td>932 435</td>
<td>807 212</td>
<td>875 285</td>
</tr>
<tr>
<td>Archaeal Sequences</td>
<td>18 036</td>
<td>10 337</td>
<td>4377</td>
<td>13 671</td>
<td>7817</td>
<td>6317</td>
<td>1158</td>
</tr>
<tr>
<td>Bacterial Sequences</td>
<td>967 945</td>
<td>795 063</td>
<td>712 129</td>
<td>871 400</td>
<td>924 354</td>
<td>800 687</td>
<td>871 957</td>
</tr>
<tr>
<td>Classified at Phylum-level</td>
<td>890 511</td>
<td>728 794</td>
<td>624 808</td>
<td>795 785</td>
<td>849 936</td>
<td>738 279</td>
<td>807 507</td>
</tr>
<tr>
<td>Unclassified at Phylum-level</td>
<td>95 600</td>
<td>76 808</td>
<td>91 941</td>
<td>89 369</td>
<td>82 499</td>
<td>68 933</td>
<td>67 778</td>
</tr>
</tbody>
</table>

with increasing site age in the focal sites and in the full-chronosequence September validation set, indicating that older restorations and remnants tended to have lower diversity, evenness and richness than newly-planted restorations and that older restorations approach the community characteristics of remnants, based on vector-fitting with envfit (Oksanen et al., 2017) (Tables 2 and 3, Fig. 1A, C and E; Supporting Information Fig. S2). Conversely, all three of these metrics increased within a season from spring to fall in the focal sites, (Table 2, Fig. 1B, D and F). NMDS fits had low stress for both data-sets (focal sites stress = 0.148; full chronosequence stress = 0.079), and site age and soil C:N ratio were significantly correlated with community composition (Table 4, Fig. 2 and Supporting Information Fig. S3). NMDS analyses highlight, however, that agricultural bacterial community composition does not represent an ‘age = 0’ restoration site, but rather an independent soil community that does not correlate significantly with either young or old restorations (PERMANOVA, $P < 0.0001$ for both comparisons). In the full chronosequence dataset, soil pH was also significantly related to community composition. The Mantel test comparing Bray-Curtis distances to the spatial distances between sites was not significant (Mantel $r = 0.003$, $P = 0.427$), indicating that spatial proximity did not correlate with microbial community composition. PERMANOVA revealed significant differences among the focal site communities ($F_{6,49} = 11.0$, $P < 0.001$), and all pairwise comparisons were highly significant ($P < 0.0003$) except the 1- and 2-year-old sites ($P = 0.0026$), which did not meet the strict critical value cut-off of $P < 0.0024$.

Most focal taxa had significant changes in abundance across the chronosequence, with quadratic and logarithmic relationships more frequent than linear relationships (Table 5, Fig. 3). These patterns included both significant increases and decreases as restorations aged, varying among taxa. Very few focal taxa showed significant changes in abundance within a season (Supporting Information Table S3, Fig. S4). If the agricultural site was included in these analyses as age = 0, none of the major trends were affected (data not shown); however, the trends in Acidobacteria groups 3 and 6 were marginally significant, reflecting the fact that these clades were present in similar abundances in older restorations and agricultural sites.

### Edaphic variables

In both the within-season dataset (Supporting Information Table S1) and the full chronosequence (Supporting Information Table S2), carbon content increased with site age, resulting in higher C:N ratios in older sites (Tables 2 and 3, Fig. 4 and Supporting Information Fig. S5). Age also interacted with sampling date for soil C such that C content also increased slightly across the season, but only in the two oldest restorations and the remnant. Within the season, pH was marginally significantly related to day of year, with a trend of soils becoming slightly more acidic as the season progressed (Table 2, Fig. 4C).

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Site age $\chi^2$ $P$</th>
<th>Day of year $\chi^2$ $P$</th>
<th>Age $\times$ Day $\chi^2$ $P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil C</td>
<td>9.61 0.002</td>
<td>1.86 0.172</td>
<td>4.76 0.029</td>
</tr>
<tr>
<td>Soil N</td>
<td>1.16 0.282</td>
<td>0.14 0.713</td>
<td>0.61 0.436</td>
</tr>
<tr>
<td>Soil C:N ratio</td>
<td>11.17 &lt;0.001</td>
<td>0.04 0.834</td>
<td>0.28 0.592</td>
</tr>
<tr>
<td>Soil pH</td>
<td>0.39 0.533</td>
<td>3.56 0.059</td>
<td>0.94 0.331</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>0.13 0.724</td>
<td>1.14 0.285</td>
<td>0.45 0.503</td>
</tr>
<tr>
<td>Shannon diversity</td>
<td>9.89 0.002</td>
<td>13.25 &lt;0.001</td>
<td>0.00 0.998</td>
</tr>
<tr>
<td>Pielou’s evenness</td>
<td>11.78 &lt;0.001</td>
<td>14.12 &lt;0.001</td>
<td>0.09 0.767</td>
</tr>
<tr>
<td>Class richness</td>
<td>7.76 0.005</td>
<td>7.33 0.007</td>
<td>0.00 0.992</td>
</tr>
</tbody>
</table>

Table 2. Results of generalized linear mixed models analysing edaphic conditions and community metrics of within-season sampling.

Table 3. Results of generalized linear mixed models analysing edaphic conditions and community metrics during September sampling of all chronosequence sites.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Site age $F$ $P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soil C</td>
<td>9.12 0.009</td>
</tr>
<tr>
<td>Soil N</td>
<td>4.06 0.062</td>
</tr>
<tr>
<td>Soil C:N ratio</td>
<td>8.65 0.010</td>
</tr>
<tr>
<td>Soil pH</td>
<td>0.01 0.935</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>2.63 0.126</td>
</tr>
<tr>
<td>Shannon diversity</td>
<td>15.79 0.001</td>
</tr>
<tr>
<td>Pielou’s evenness</td>
<td>21.78 &lt;0.001</td>
</tr>
<tr>
<td>Class richness</td>
<td>8.47 0.011</td>
</tr>
</tbody>
</table>

Results in bold are significant at $P < 0.05$.
Table 4. Test of relationships between site characteristics and community composition based on envfit vector-fitting.

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Within-season sampling</th>
<th>Full chronosequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r^2$</td>
<td>$P$</td>
</tr>
<tr>
<td>Site age</td>
<td>0.607</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Day of year</td>
<td>0.081</td>
<td>0.111</td>
</tr>
<tr>
<td>Soil C:N ratio</td>
<td>0.138</td>
<td>0.018</td>
</tr>
<tr>
<td>Soil pH</td>
<td>0.061</td>
<td>0.133</td>
</tr>
<tr>
<td>Soil moisture</td>
<td>0.016</td>
<td>0.649</td>
</tr>
</tbody>
</table>

Results in bold are significant at $P < 0.05$.

Discussion

We documented significant changes in soil bacteria communities across a tallgrass prairie restoration chronosequence and over the course of a growing season. Richness, diversity and evenness declined as restorations became older, a pattern that was documented in both the focal subset and the full chronosequence, but these same metrics increased from spring through fall within the sampling year (Fig. 1). Bacterial community composition was also significantly related to site age, with immediate effects of restoration activities resulting in communities in young restorations (the first 1–3 years of growth) that distinctly differed both from the pre-restoration agricultural field communities (PERMANOVA, $P < 0.0001$) and from old restorations ($P < 0.0001$). This independence of agricultural field community composition from the greater successional pattern in restoration sites contradicts the trend in edaphic variables and broader diversity metrics, which supported the role of agricultural field communities as an age = 0 restoration site (Fig. 4). Older restorations (13 and 27-year-old restorations sampled over the season (Fig. 2), as well as most sites older than ~6 years in the full chronosequence (Supporting Information Fig. S3) were also compositionally similar to the local prairie remnant site, providing strong evidence that restoration management has successfully converted former agricultural areas to soil resembling reference remnants, in some cases in less than a decade. The fact that these communities approach remnants in less than three decades is noteworthy given that most ecosystem restoration projects, including sites with active restoration management such as Nachusa, achieve less than 50% recovery of communities and ecosystem functions within this timeframe (Moreno-Mateos et al., 2017).

Here, we focus our discussion on four phyla previously identified as important players in grassland restorations and cultivated soils (Acidobacteria, Gemmatimonadetes, Proteobacteria and Verrucomicrobia). These phyla also consistently showed the largest statistically significant changes in relative abundance between restoration ages in this study. We also discuss abiotic changes in the soil environment and how these relate to intra- and interannual shifts in bacterial communities.

Bacterial populations affiliated with an undisturbed prairie microbiome

Previous studies in grassland systems have suggested a link between Acidobacteria abundance and undisturbed or native ecosystems (McNamara et al., 2006; Griffiths et al., 2011; Nacke et al., 2011; Van Trump et al., 2011; Fierer et al., 2011). As an age = 0 restoration site, the microbial communities of young restorations are simpler in composition and have a lower amount of operational taxonomic units (OTUs), supporting the idea that restoration age is a strong factor driving bacterial community composition. Future monitoring will be able to identify the age at which restoration is complete, as well as the extent to which the remaining communities approach the communities of the local native prairie remnants.

Table 5. Relationships between site age and taxa relative abundance, excluding agricultural site but including remnant as age = 30, with lowest AIC model retained and evaluating the age factor with an F-test.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Trend with Age</th>
<th>Average Abundance</th>
<th>$F$</th>
<th>$P$</th>
<th>Best Predictor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidobacteria (total)</td>
<td>Increasing</td>
<td>13.8%</td>
<td>39.3</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Acidobacteria group 1</td>
<td>Decreasing</td>
<td>1.6%</td>
<td>10.13</td>
<td>0.003</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Acidobacteria group 3</td>
<td>Decreasing</td>
<td>1.4%</td>
<td>28.14</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Acidobacteria group 4</td>
<td>Increasing</td>
<td>2.5%</td>
<td>38.89</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Acidobacteria group 6</td>
<td>Increasing</td>
<td>5.1%</td>
<td>34.87</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Acidobacteria group 16</td>
<td>Increasing</td>
<td>1.6%</td>
<td>44.51</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Increasing</td>
<td>29.0%</td>
<td>7.34</td>
<td>0.010</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Gemmatimonadetes</td>
<td>Decreasing</td>
<td>1.8%</td>
<td>150.62</td>
<td>$&lt;0.001$</td>
<td>Log(age + 1)</td>
</tr>
<tr>
<td>Proteobacteria (total)</td>
<td>Decreasing</td>
<td>28.0%</td>
<td>22.93</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Decreasing</td>
<td>13.7%</td>
<td>41.99</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Decreasing</td>
<td>6.3%</td>
<td>28.77</td>
<td>$&lt;0.001$</td>
<td>Log(age + 1)</td>
</tr>
<tr>
<td>Deltaproteobacteria</td>
<td>—</td>
<td>3.8%</td>
<td>2.75</td>
<td>0.105</td>
<td>Log(age + 1)</td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Decreasing</td>
<td>3.1%</td>
<td>24.69</td>
<td>$&lt;0.001$</td>
<td>Log(age + 1)</td>
</tr>
<tr>
<td>Verrucomicrobia (total)</td>
<td>Increasing</td>
<td>5.3%</td>
<td>17.17</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Opitutae</td>
<td>Decreasing</td>
<td>0.5%</td>
<td>19.1</td>
<td>$&lt;0.001$</td>
<td>Log(age + 1)</td>
</tr>
<tr>
<td>Spartobacteria</td>
<td>Increasing</td>
<td>3.3%</td>
<td>15.63</td>
<td>$&lt;0.001$</td>
<td>Age$^2$</td>
</tr>
<tr>
<td>Subdivision 3</td>
<td>Decreasing</td>
<td>1.2%</td>
<td>7.93</td>
<td>0.008</td>
<td>Log(age + 1)</td>
</tr>
</tbody>
</table>

Results in bold are significant at $P < 0.05$. 

© 2017 Society for Applied Microbiology and John Wiley & Sons Ltd, Environmental Microbiology, 00, 00–00
et al., 2012; Naether et al., 2012; da Rocha et al., 2013; Pan et al., 2014). While an increase in total Acidobacteria abundance did correlate with restoration age, changes in abundance of several classes within the Acidobacteria (also known as Groups or Subdivisions) varied with

restoration age (Table 5). Weak trends that began in the first 3 years after planting, namely decreases in the abundance of the acidophilic Groups 1 and 3 and corresponding increases in neutrophilic/alkaliphilic Groups 4, 6 and 16, were much more apparent in old restorations and remnants. The relative ease of cultivation of members of Groups 1 and 3 (compared to 4, 6 and 16) suggests that they may be relatively more eutrophic than members of Groups 4, 6 and 16, which may drive their decrease in abundance as restored soils became more oligotrophic or more metabolically recalcitrant, and large carbon-rich plant polymers accumulate (as suggested by increases in carbon content with age) (McLauchlan et al., 2006).

Nutrient selection may be more important than the trend in acidification for any of the highlighted Acidobacteria classes, as the pH in all Nachusa soils is typically two orders of magnitude higher than the optima for Groups 1 and 3 (Jones et al., 2009). The circumneutral pH likely contributes to the relative success of members of Group 4, 6 and 16, which have been shown to be weakly alkaliphilic (Jones et al., 2009; Foesel et al., 2013; Huber et al., 2014). While little is known about members of Groups 4, 6 and 16, it is presumed that uncultivable members are obligate oligotrophs (Jones et al., 2009; Naether et al., 2012; Foesel et al., 2013), and their abundance in undisturbed grasslands may be due to decreased nutrient availability rather than the minor differences in pH. Although soil moisture content might be anticipated to correlate Group 4 Acidobacteria abundance given cultures isolated from arid environments (Foesel et al., 2013; Huber et al., 2014), there was no relationship with this environmental parameter (Pearson correlation $r = 0.15$, $P = 0.232$, $R^2 = 0.02$).

In recent years, improvements in 16S rDNA primer design have led to the discovery that members of the relatively poorly-described phylum Verrucomicrobia are major components of uncultivated grassland soils (Bergmann et al., 2011; Nacke et al., 2011; da Rocha et al., 2013; Fierer et al., 2013; Carbonetto et al., 2014; Pan et al., 2014). This realization comes despite the continuing dearth of representative cultures. The Nachusa representative remnant prairie and restorations older than 10 years agree with previous observations in their relatively high abundance of Verrucomicrobia from class Spartobacteria. The slow rise of Verrucomicrobia during early succession—rarely rising beyond 3% of total 16S rRNA reads in early restorations—is in sharp contrast to rapid changes within the Acidobacteria classes (Fig. 5), and it is unclear what specific physicochemical parameters caused this stark difference. Approximately 48% of Spartobacteria 16S rRNA genes found in 10+ year-old restorations and the remnant prairie share $>97\%$ identity with Candidatus Udaeobacter copiosus, a draft Spartobacteria genome reconstructed from a Colorado grassland terrace metagenome (Brewer et al., 2016). Conversely, only 25% of

![Fig. 1. Changes in (A) soil C content and (B) soil C:N ratio with site age in focal sites. (C) Marginally significant changes in soil pH across the season in these same sites. Open circles are the agricultural field, and triangles are the remnant. In all panels, each site was sampled on 8 days from April through September. A single line of best fit is depicted in each panel, although analyses accounted for the non-independence of repeatedly-sampled sites.](image-url)
Spartobacteria 16S rRNA genes in <3-year-old restorations shared >97% identity with this genome, suggesting that this ecotype is a key player in the rise of Verrucomicrobia in prairie soils. The putatively streamlined genome and oligotrophic metabolism of *Candidatus Udaeobacter copiosus* might explain their slow response, but ultimate success during restoration succession. As observed in previous studies, members of Verrucomicrobia Subdivision 3 and class Opitutae were more abundant in early restorations (Bergmann *et al.*, 2011), but did not increase in relative abundance over the course of restoration (Supporting Information Table S4), which is consistent with their apparent preference for cultivated soils (Sangwan *et al.*, 2005).

Proteobacteria have been found to be major members of soil microbial communities (McNamara *et al.*, 2006; Griffiths *et al.*, 2011; Nacke *et al.*, 2011; Van Trump *et al.*, 2011; Fierer *et al.*, 2012; Kuramae *et al.*, 2012; Pan *et al.*, 2014); however, their abundance in Nachusa soils is slightly lower than analyses in other grassland systems, averaging roughly 25% of the microbial community. The most abundant Proteobacterial orders are the Rhizobiales in class Alphaproteobacteria followed by Myxococcales in the Deltaproteobacteria and the Burkholderiales in the Betaproteobacteria. The general decrease in Gammaproteobacteria with increasing restoration age was driven primarily by decreases in members of the family Xanthomonadaceae, a group of common plant pathogens.

---

**Fig. 2.** Relationships between community metrics (Shannon diversity, Pielou’s evenness and Class richness) in within-season sampling sites and (A, C, E) site age and (B, D, F) day of year. Open circles are the agricultural field, and triangles are the remnant. In all panels, each site was sampled on 8 days from April through September. A single line of best fit is depicted in each panel, although analyses accounted for the non-independence of repeatedly-sampled sites.
Fig. 3. Non-metric multidimensional scaling ordination of microbial community composition in within-season sampling sites, with each site color-coded. Numbers on each dot represent sampling session (1 – 30 April, 8 – 26 September). Vectors illustrate significant correlations between environmental parameters and community composition.

Fig. 4. Box and whisker plot representing the average percent change in abundance per year between all dates using each seasonal sample site ranked according to mean annual change for the four phyla discussed in this work. Each box represents one quartile above and below the mean and the whiskers represent 5% and 95% quantiles. Box widths approximate relative abundance of each taxon.

Fig. 5. Average distribution of 16S rRNA genes for bacterial phyla (and Proteobacterial and Acidobacterial classes) with a relative abundance >1% in each of the seven full-season sampling sites. The eight seasonal sample date abundances were normalized and averaged for each site. Black columns represent all taxa <1% at each site.

© 2017 Society for Applied Microbiology and John Wiley & Sons Ltd, Environmental Microbiology, 00, 00–00
decreasing from 1.12% of total 16S rRNA genes to 0.40%; however class-level decreases were partly offset by increases in sequences affiliated with Sinobacteracaeae, the sister family to Xanthomonadaceae in order Xanthomonadales, in older restorations, increasing from 0.12% of total 16S rRNA genes to 0.40%. Similarly, while Alphaproteobacteria did not show a significant trend with age, increases in members of family Hyphomicrobiaceae, typically composing nearly 20% of all Alphaproteobacteria, in older restorations were balanced by decreases in several other families. As many Alphaproteobacteria are associated with plant roots, these changes may be caused by both shifts in plant communities and a broadening of the overall rhizosphere, as root networks become increasingly dense in prairie soils (Cook et al., 1988; Matamaia et al., 2008). Like the Proteobacteria, members of the abundant phylum Actinobacteria (averaging ~28% of 16S rRNA genes) show a relatively weak correlation with prairie age, with very few of its classes showing more than a twofold change over the course of restoration. As such, while Actinobacteria was the most abundant community member, its increase in abundance correlated with restoration age was considerably lower than in the four focal phyla (Fig. 3).

Similar to previous studies, 16S rRNA gene sequences affiliated with phylum Gemmatimonadetes were more abundant in the cultivated corn field and early restorations (Lauber et al., 2008; Jiang et al., 2010; Lauber et al., 2013; Carbonetto et al., 2014). While it has been suggested that soil moisture is a driving factor in Gemmatimonadetes abundance (DeBruyn et al., 2011; DeBruyn et al., 2013; Zeng et al., 2015), moisture did not appear to be the largest driver in the Nachusa restorations. A mixed response of Gemmatimonadetes to nutrient addition (Eilers et al., 2010; Fierer et al., 2012) highlights the need for further study of both environmental and cultivable members of this phylum to characterize why this group of microaerophilic oligotrophs appear to thrive more in cultivated than in undisturbed soils.

**Microbial community structure in agricultural soils**

Surprisingly, agricultural soil bacterial communities did not show strong similarity to the youngest (1-year) restoration site (PERMANOVA, $F_{1,16} = 6.18$, $P < 0.001$), but rather represented an intermediate composition on most sampling dates between the young (1–3-year-old) and old (13–27-year-old and remnant) communities (Fig. 2). This supports the idea that restoration impacts, namely seed planting, burning and the cessation of agricultural practices (fertilization, tilling and pesticide application), have an almost immediate impact on soil ecosystems and underscore how different the soil environment is under active rowcrop agriculture. However, more intriguing is the divergent nature of agricultural communities, containing high relative abundances of taxa affiliated with both young and old prairie restorations. Average abundances of Verrucomicrobia, Proteobacteria and Gemmatimonadetes (2.4%, 33.8% and 2.1%, respectively) were on par with the three youngest restorations, while the profile of Acidobacteria classes was nearly identical to the 27-year-old restoration (Groups 4, 6 and 16 at 2.7%, 6.8% and 2.2%, respectively).

Agricultural field communities do not appear to be specialized at deeper taxonomic levels, as assessed by family- and genus-level diversity (Supporting Information Table S4), sharing many of the same major clades with restoration plots. In fact, in the Proteobacteria, which more strongly resemble early restorations, the relatively high abundance of class Gammaproteobacteria is due to increases in the Sinobacteracaeae, a clade more abundant in older restorations and remnants.

The treatment of agricultural soil as a zero time-point for restoration appears to be ecologically relevant in terms of edaphic variables and broad community metrics, whether considering the focal sites or the full September chronosequence (Figs 1 and 2; Supporting Information Figs S2 and S3). However, community metrics such as diversity, richness and evenness are coarse-grained and may not detect compositional changes at finer taxonomic levels. Agricultural management practices, e.g. tillage, fertilization and annual monoculture, likely impact other soil characteristics such as aeration, water retention rates and soil compaction that are not reflected in the edaphic variables measured here. These differences may drive the drastic shift from agricultural to first-year restoration soil bacterial communities when finer-scale taxonomic and compositional changes are examined.

**Seasonal impacts on grassland soils**

According to long-term climate records for nearby Dixon, IL, the summer of 2014 was a relatively average year with regard to temperature and precipitation (i.e. no statistical significance in differences in monthly min/max temperature and precipitation averages). As a result, this study is reflective of an average seasonal variation, with bursts of rain and dry spells, heat waves and cold snaps that did not push the year toward abnormal drought or flooding. The impact of these short-term (weather) events on soil microbial communities is most noticeable in moisture trends, which mirror rainfall events (Supporting Information Fig. S6), as there were relatively few extremely hot days that might drive excessive evaporation. The regular rainfall may have contributed to the marginally significant declining pH trend over the season through the dissolution and removal of basic cations such as Ca$^{2+}$ and Mg$^{2+}$. Leguminous nitrogen cycling could also have played a role in acidifying
that soil microbes, yielding communities that rapidly resemble both local remnants and other undisturbed prairies throughout the American Midwest. Although within-season community patterns follow an opposite trend, with increasing richness and diversity from spring to fall possibly due to changes in soil pH or seasonal plant growth dynamics, decade-scale changes in bacterial communities mirror the declines in richness and diversity over time in restorations for other taxa like plants (Hansen and Gibson, 2014) and ground beetles (N.A. Barber, K.A. Lamagdeleine-Dent, J.E. Willand, K.W. McCravy, and H.P. Jones. Species and functional trait re-assembly of ground beetle communities in restored grasslands, in preparation).

Active restoration management in grasslands has traditionally focussed on macro-organisms. However, our results suggest that the restoration activities at Nachusa over the past 30 years have produced successes that extend beyond plant and animal biodiversity. The restoration management activities at this location, which lead to microbial communities that mimic undisturbed prairie diversity and, presumably, ecosystem functions, are carried out in other restoration projects as well. Thus, grassland restoration projects can provide a robust experimental framework to begin to understand what factors drive differences in cultivated and undisturbed soil communities. Further, results will perhaps allow managers to track restoration success through soil microbial communities and make management decisions with a goal of supporting soil-driven ecosystem functions (Heneghan et al., 2008; Harris, 2009). Continued studies at the Nachusa Grasslands and other restored sites will help to illuminate the physicochemical parameters that support ‘native grassland’ taxa, especially potential indicator taxa within the Verrucomicrobia and Acidobacteria.

Conclusions and future directions

As in many ecosystem restoration projects, the long-term trajectory of restoration at Nachusa has been monitored and managed to meet the goal of maximizing richness and diversity of native plant communities (Hobbs and Norton, 1996; Van Andel and Aronson, 2012; McAlpine et al., 2016). Yet the consistent management practices focussed on guiding plant succession has also influenced soil microbes, yielding communities that rapidly resemble both local remnants and other undisturbed prairies throughout the American Midwest. Although within-season community patterns follow an opposite trend, with increasing richness and diversity from spring to fall possibly due to changes in soil pH or seasonal plant growth dynamics, decade-scale changes in bacterial communities mirror the declines in richness and diversity over time in restorations for other taxa like plants (Hansen and Gibson, 2014) and ground beetles (N.A. Barber, K.A. Lamagdeleine-Dent, J.E. Willand, K.W. McCravy, and H.P. Jones. Species and functional trait re-assembly of ground beetle communities in restored grasslands, in preparation).

Active restoration management in grasslands has traditionally focussed on macro-organisms. However, our results suggest that the restoration activities at Nachusa over the past 30 years have produced successes that extend beyond plant and animal biodiversity. The restoration management activities at this location, which lead to microbial communities that mimic undisturbed prairie diversity and, presumably, ecosystem functions, are carried out in other restoration projects as well. Thus, grassland restoration projects can provide a robust experimental framework to begin to understand what factors drive differences in cultivated and undisturbed soil communities. Further, results will perhaps allow managers to track restoration success through soil microbial communities and make management decisions with a goal of supporting soil-driven ecosystem functions (Heneghan et al., 2008; Harris, 2009). Continued studies at the Nachusa Grasslands and other restored sites will help to illuminate the physicochemical parameters that support ‘native grassland’ taxa, especially potential indicator taxa within the Verrucomicrobia and Acidobacteria.
Table 6. Characteristics of chronosequence sites. ‘Within-season sampling’ indicates sites that were sampled on eight dates across the growing season.

<table>
<thead>
<tr>
<th>Site</th>
<th>Year planted</th>
<th>Age at sampling</th>
<th>Focal site</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG corn</td>
<td>–</td>
<td>–</td>
<td>Y</td>
</tr>
<tr>
<td>AG soy</td>
<td>–</td>
<td>–</td>
<td>N</td>
</tr>
<tr>
<td>HF</td>
<td>2013</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>HN</td>
<td>2012</td>
<td>2</td>
<td>Y</td>
</tr>
<tr>
<td>L</td>
<td>2011</td>
<td>3</td>
<td>Y</td>
</tr>
<tr>
<td>SB</td>
<td>2009</td>
<td>5</td>
<td>N</td>
</tr>
<tr>
<td>CCW</td>
<td>2008</td>
<td>6</td>
<td>N</td>
</tr>
<tr>
<td>HW</td>
<td>2008</td>
<td>6</td>
<td>N</td>
</tr>
<tr>
<td>CCE</td>
<td>2007</td>
<td>7</td>
<td>N</td>
</tr>
<tr>
<td>FC</td>
<td>2006</td>
<td>8</td>
<td>N</td>
</tr>
<tr>
<td>TC</td>
<td>2002</td>
<td>12</td>
<td>N</td>
</tr>
<tr>
<td>SF</td>
<td>2001</td>
<td>13</td>
<td>N</td>
</tr>
<tr>
<td>HLP</td>
<td>2001</td>
<td>13</td>
<td>Y</td>
</tr>
<tr>
<td>WH</td>
<td>1992</td>
<td>22</td>
<td>N</td>
</tr>
<tr>
<td>MU</td>
<td>1987</td>
<td>27</td>
<td>Y</td>
</tr>
<tr>
<td>MR remnant</td>
<td>–</td>
<td>–</td>
<td>Y</td>
</tr>
<tr>
<td>TC remnant</td>
<td>–</td>
<td>–</td>
<td>N</td>
</tr>
</tbody>
</table>

Experimental procedures

Study site

Samples were obtained from the Nachusa Grasslands, a 1400 ha tallgrass prairie restoration project owned and operated by The Nature Conservancy that encompasses several prairie remnants. Restorations of former agricultural fields to diverse prairie plant communities have followed similar methodology since the 1980s, resulting in replicated plantings that differ in age but are managed using the same techniques. In brief, restoration involves the cessation of tilling and fertilization associated with agriculture, planting of a diverse mix of locally-collected seeds of native prairie plant species, prescribed fire (generally ever 2–3 years) to control invasive plants and woody plant encroachment, and active removal of weeds and other invasive plants using chemical and mechanical removal. More details of management and restoration practices are available elsewhere (Willand et al., 2013; Hansen and Gibson, 2014; Wodika and Baer, 2015; Barber et al., 2017). Upon acquisition, prairie remnants have been managed under the same practices as restorations, including prescribed fire and invasive species removal.

We assembled a chronosequence of thirteen sites planted between 1987 and 2013, ranging from the 1st to 27th growing season during sampling in 2014 (Table 6). Each site was >2ha in size, and selection controlled for characteristics such as soil texture (all loam, or mix of silt loam and sandy loam). We avoided sites that were heavily invaded by exotic weeds. We also selected two agricultural fields, one planted in soybeans and the other corn, to represent pre-restoration conditions and two on-site remnant plots that were never ploughed but were grazed by cattle until the mid-1980s, at which point management practices mirrored those in restorations. The agricultural fields have been in soybean-corn rotation for at least the past decade. This design allowed us to examine stages of restoration development, including comparisons to pre-restoration conditions and to remnants. Although these remnants are not perfect references, given their grazing history, they represent the only large, high-quality remnant tallgrass prairies with >50 km of the restoration sites and are thus the most appropriate sites for comparison.

Soil sampling

Our sampling and analyses focussed on a subset of seven focal sites for repeated sampling across a single growing season to investigate both community changes with restoration age and detailed within-season phenological change in the soil environment and soil bacterial communities (Table 6, ‘Focal site’). These focal sites included five restorations, spanning the age range of the full chronosequence, as well as a corn field and one remnant. We collected soil samples on eight dates from 30 April to 26 September 2014. At five haphazardly-selected points within each site (different points on each sampling date), we collected approximately 10 g of soil from the top 10 cm with a sterile scoop (avoiding contact with aboveground vegetation) and homogenized them into a single ~50g sample in a Whirlpack bag. We immediately placed samples on ice for transportation and stored them at −80°C until processing (stored typically within 4 h of sampling and processed within 48 h).

To validate restoration age-related community patterns with a larger cross-sectional sample, we sampled the full chronosequence on 26 September. This collection consisted of the full thirteen restorations, both remnants and both agricultural fields, following the same procedure. Samples from the focal sites on 26 September were analysed in both the context of focal sites data and the full chronosequence validation data.

DNA sequencing and analysis

DNA was extracted from each sample using the Mo Bio PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Inc., Carlsbad, CA). 16S rRNA genes were amplified in preparation for Illumina sequencing using protocols developed for the Earth Microbiome Project with universal primers 515F and 808R (Caporaso et al., 2012) containing 5 Illumina CS1 and CS2 adaptors, respectively (Illumina, Inc., San Diego, CA). rDNA was sequenced using 250 bp paired-end amplicon sequencing on an Illumina HiSeq2500 through the University of Illinois at Chicago DNA Services facility. Paired reads were assembled and quality trimmed using QIIME (Caporaso et al., 2010).

16S rDNA sequences were assigned to taxonomic units using complete linkage clustering from RDP (Wang et al., 2007). Assignment of reads to a taxonomic level was designated at the last level to exceed the suggested RDP confidence threshold of 50% (for sequences between 50–250 bp), where sequences scoring below 50% at a given taxonomic level were labelled at the previous >50% confident level appended with ‘-Undefined’. For example, a sequence confidently assigned to phylum Firmicutes, but scoring <50% to an order within the Firmicutes was labelled as ‘Firmicutes-Undefined.’

Edaphic variables

Soil samples were removed from −80°C storage and after a brief thawing period approximately 5 g was weighed out into a
small coin envelope, which was placed in a 60°C oven to dry for 48 h. Total moisture was calculated as the percent loss in dry versus starting mass. Soil was ground to a fine powder using a mortar and pestle and remaining large particles (primarily rocks and plant fibres) were removed using a 1 mm mesh. 1 g of this dry, homogenized soil was stirred into 5 mL of nanopure water and the pH was measured after settling. A further 0.5 g of processed soil was weighed out into crucibles for quantifying total C and N in a Viromax CNS analyzer (Elementar, Hanau, Germany) calibrated to manufacturer recommendations.

Statistical analyses

To understand how edaphic conditions (C content, N content, C:N ratio, pH and moisture) change during a growing season, we used linear mixed models (LMMs), treating age of each site and the day of year on which sampling took place as continuous fixed factors. Age of the remnant was set at 30 years, the approximate time since cattle were removed and management activities matching restorations (prescribed fire, invasive plant control) began, and agricultural sites were set to age 0. Arbitrarily setting remnant age at 50 years rather than 30 did not qualitatively change results for edaphic variables or the community metrics discussed below, nor did excluding age 0 agricultural sites (data not shown). Site was a random factor to account for repeated-sampling across the season, and models assumed a Gaussian error distribution. We evaluated the age × day of year interaction, and each of these independent variables (age first, then day), using likelihood ratio tests (LRT's). The lme() function in the nlme package (Pinheiro et al., 2014) of R (Team, 2016) was used to construct LMMs. We analysed these same five edaphic variables in the full chronosequence using generalized linear models (GLMs, using glm() function) including just age as a fixed factor, because the full chronosequence was sampled on a single date. Age was evaluated using F-tests.

We examined three community metrics: Shannon diversity, Pielou's evenness and richness, at the Class taxonomic level, with Class richness only including taxa representing >1% of the community. We analysed these metrics in the focal-site samples using LMMs and LRT's and in the full chronosequence samples using GLMs and F-tests as described above.

To examine soil microbial community composition and its relationship to site variables, we calculated Bray-Curtis distances between all pairs of sites within the two datasets. We visualized community composition using non-metric multidimensional scaling (NMDS) with two axes using metaMDS() in the vegan package (Oksanen et al., 2017) and tested relationships with site characteristics (age, C:N ratio, soil pH and soil moisture) using the vegan function envfit(), and the focal sites dataset also included sampling day of year as an independent variable. Because community composition was significantly related to site age (see Results), we wanted to verify that this was not due to spatial autocorrelation. That is, if microbial communities are similar to each other simply because of geographical proximity, especially in the case of adjacent sites that are similar in age, and whether this could result in a spurious correlation with age. With the full chronosequence, we used a Mantel test to find the correlation between Bray-Curtis distances among pairs of sites and the geographic distance between site pairs as determined by GPS coordinates of sampling locations. Finally, we used permutational MANOVA (PERMANOVA) to compare community composition of the focal sites using anosim() in vegan. Because sites differed significantly (see Results), we also carried out pairwise PERMANOVA between each pair of focal sites, controlling for multiple comparisons by using a conservative Bonferroni critical-value adjustment of $P = 0.0024$.

We examined four focal phyla, Acidobacteria, Gemmatimonadetes, Proteobacteria and Verrucomicrobia, in detail to determine how their relative abundances varied across the chronosequence and within a season. For each phylum (and select classes within each phylum), we constructed a set of three LMMs treating abundance as the dependent variable, site age as a fixed independent variable and sampling date as a random factor to control for repeated sampling of each site. In the three models, the age variable was either raw age (linear model), age$^2$ (quadratic model), or ln(age + 1) (logarithmic model). These three models were selected because we could not predict a priori the shape of abundance responses to time, and the selected models were chosen to represent continuous linear change, high/low abundance at intermediate ages, or rapid increases/decreases to a stable abundance, respectively. Agricultural field samples were not utilized in these analyses, as community composition analyses did not support the hypothesis that agricultural soil would represent an age of 0. The AIC of the three models was compared, and the lowest model AIC was selected for evaluation using a Wald F-test. To determine if the abundance of these focal groups changed within a season (e.g., from spring to fall), we repeated these analyses using sets of three models that treated sampling day of year in the same manner to create linear, quadratic and logarithmic models. We treated site as a random factor in these models to control for the fact that the same sites were sampled on each sampling date, so samples from the same site on different dates are not independent of one another. Again, we selected the model with the lowest AIC and evaluated sampling date with a Wald F-test.

Acknowledgements

We thank The Nature Conservancy and the staff at Nachusa Grasslands for supporting this work, especially B. Kleiman and C. Considine, and the numerous students who provided assistance in the field and laboratory. This work was supported by the National Science Foundation SMA-1156789, the Friends of Nachusa Grasslands, and Research and Artistry Opportunity Grants from Northern Illinois University to WDS and NAB. The authors declare no conflict of interest for this work.

References


© 2017 Society for Applied Microbiology and John Wiley & Sons Ltd, *Environmental Microbiology*. **00**: 00–00


**Supporting information**

Additional Supporting Information may be found in the online version of this article at the publisher’s website:
Fig. S1. Rarefaction curves for 16S rRNA genes classified at (A) phylum, (B) class, and (C) genus taxonomic level in all 2014 full-season samples.

Fig. S2. Changes in (A) soil C content, (B) soil N content, and (C) soil C:N ratio with age in the full chronosequence, sampled on 26 September. Dashed line in panel (B) indicates that increase in soil N was marginally significant. Each point represents a different site, open circles are agricultural fields, and triangles are remnants.

Fig. S3. Relationships between community metrics (Shannon diversity, Pielou’s evenness, and Class richness) in the full chronosequence and site age. Each point represents a different site sampled on 26 September, open circles are agricultural fields, and triangles are remnants.

Fig. S4. Non-metric multidimensional scaling ordination of microbial community composition in the full chronosequence, each site sampled on 26 September, open circles represent site age in years. Vectors illustrate significant correlations between environmental parameters and community composition.

Fig. S5. Average distribution of 16S rRNA genes for bacterial phyla (and Proteobacterial and Acidobacterial classes) with a relative abundance >1% during each of the eight seasonal sample dates. The seven seasonal sample site abundances were normalized and averaged for each date. Black columns represent all taxa <1% at each site.

Fig. S6. Percent soil moisture in each focal site (gray dots) and total precipitation in the preceding week (blue bars) for each sampling date. Each of 7 focal sites was sampled on 8 dates from April through September.

Table S1. Edaphic variables for the within-season dataset. Number in parentheses after site name is the age of restorations in years.

Table S2. Edaphic variables for the full chronosequence dataset, sampled on 26 September 2014. Number in parentheses after site name is the age of restorations in years.

Table S3. Relationships between day of year and taxa relative abundance, excluding ag site but including remnant as age = 30. Three models were evaluated, either with age (linear), age (quadratic), or log age as predictor. Lowest AIC model was retained and the age variable was evaluated with an F-test. Results in bold are significant at P < 0.05.

Table S4. Taxonomic distribution of 16S rRNA gene sequences at five taxonomic levels using RDP’s Classifier with a confidence threshold of 50%.