



# Phylogenetic conservatism of thermal traits explains dispersal limitation and genomic differentiation of *Streptomyces* sister-taxa

Mallory J. Choudoir<sup>1,2</sup> · Daniel H. Buckley<sup>1</sup>

Received: 18 October 2017 / Revised: 22 March 2018 / Accepted: 26 March 2018 / Published online: 7 June 2018  
© International Society for Microbial Ecology 2018

## Abstract

The latitudinal diversity gradient is a pattern of biogeography observed broadly in plants and animals but largely undocumented in terrestrial microbial systems. Although patterns of microbial biogeography across broad taxonomic scales have been described in a range of contexts, the mechanisms that generate biogeographic patterns between closely related taxa remain incompletely characterized. Adaptive processes are a major driver of microbial biogeography, but there is less understanding of how microbial biogeography and diversification are shaped by dispersal limitation and drift. We recently described a latitudinal diversity gradient of species richness and intraspecific genetic diversity in *Streptomyces* by using a geographically explicit culture collection. Within this geographically explicit culture collection, we have identified *Streptomyces* sister-taxa whose geographic distribution is delimited by latitude. These sister-taxa differ in geographic distribution, genomic diversity, and ecological traits despite having nearly identical SSU rRNA gene sequences. Comparative genomic analysis reveals genomic differentiation of these sister-taxa consistent with restricted gene flow across latitude. Furthermore, we show phylogenetic conservatism of thermal traits between the sister-taxa suggesting that thermal trait adaptation limits dispersal and gene flow across climate regimes as defined by latitude. Such phylogenetic conservatism of thermal traits is commonly associated with latitudinal diversity gradients for plants and animals. These data provide further support for the hypothesis that the *Streptomyces* latitudinal diversity gradient was formed as a result of historical demographic processes defined by dispersal limitation and driven by paleoclimate dynamics.

## Introduction

The latitudinal diversity gradient is one of the most widely documented biogeographic patterns in all of ecology, having been described for a broad range of plant and animal systems [1]. Latitudinal diversity gradients, however, are rarely described in microbial systems, although a few examples have been documented in marine [2–5] and terrestrial environments [6–8]. Growing evidence suggests that microbes are subject to many of the same forces that drive

biogeographic patterns in ‘macro-organisms’ [9–12]. Although adaptive traits are known to promote microbial diversification and reinforce certain patterns of biogeography, the degree to which patterns of biogeography are driven by dispersal limitation and drift remain less well described (as reviewed in [13]). We struggle to discern patterns of microbial biogeography driven by dispersal limitation and drift because microbes are small, they have vast populations, and microbial species are difficult to differentiate [13–16].

Several hypotheses have been proposed to describe the formation of latitudinal diversity gradients [17, 18]. Briefly, ecological hypotheses typically associate species richness with niche breadth and productivity [19, 20], evolutionary hypotheses implicate the effects of temperature on the kinetics of molecular evolution and rates of speciation [21], and historical hypotheses account for contemporary distributions of biodiversity as a consequence of demographic processes associated with previous geological, climatological, or ecological events [22, 23]. For instance, climatic oscillations throughout the Pleistocene resulted in repeated

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1038/s41396-018-0180-3>) contains supplementary material, which is available to authorized users.

---

✉ Daniel H. Buckley  
dbuckley@cornell.edu

<sup>1</sup> School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

<sup>2</sup> University of Colorado Boulder, Boulder, CO 80309, USA

cycles of glacial and interglacial periods. On these geologic time scales, the tropics are less susceptible to glaciation, and hence ‘time for speciation effects’ permits an accumulation of greater species diversity at lower latitudes [23, 24]. In contrast, the species diversity of higher latitudes is constrained by limits on post-glacial range expansion from lower latitudes and glacial refugia (reviewed in [25]).

Ultimately, multiple mechanisms can create diversity gradients, and these causes vary in their relative importance across diverse taxa and habitats. For example, in marine microbial systems, latitudinal diversity gradients have been linked to increased metabolic rates at warmer latitudes [2]. Alternatively, latitudinal gradients of bacterial diversity in forest soils have been explained in relation to metabolic niche partitioning [8]. In addition, we have proposed that historical demography generated a latitudinal diversity gradient in terrestrial *Streptomyces* as a consequence of paleoclimate dynamics [6, 7, 26]. Glacial dynamics have also impacted the population structure of other microbial taxa in North America soils, including spore-forming bacteria such as *Bacillus* [27] and *Actinobacteria* [28]. A recent study by Delgado-Baquerizo et al. [29] shows that paleoclimate conditions explain much of the variation in the richness and composition of microbial communities across diverse soil habitats, suggesting that the genetic legacy of these historical events are discernable in extant microbes.

Bacteria of the genus *Streptomyces* (phylum *Actinobacteria*) are ubiquitous in soil [30] where they form desiccation and starvation-resistant spores, which are easily dispersed [31]. *Streptomyces* are often associated with freshly decomposing organic matter, and they play an important role in the terrestrial carbon cycle because they are prolific degraders of complex organic molecules like cellulose [32] and chitin [33]. *Streptomyces* also produce a majority of microbial-derived antibiotic, antifungal, and tumor suppressant compounds [31, 34]. We previously assembled a culture collection of >1000 *Streptomyces* isolates spanning the United States. *Streptomyces* species exhibit dispersal limitation acting at regional spatial scales and manifesting in a latitudinal diversity gradient [6, 7]. Furthermore, they exhibit patterns of phylogeography and population structure consistent with post-glacial range expansion [6, 7, 26]. We hypothesize that, as spores are small, numerous, and easily disseminated, dispersal limitation is caused not by limits to dissemination but rather by adaptive barriers and density-dependent blocking, which inhibits colonization of previously occupied sites (as described [35]). The imposition of adaptive barriers to dispersal would manifest as niche conservatism of ecological traits [36, 37].

We have identified *Streptomyces* sister-taxa that occupy distinct geographic ranges delimited by latitude [26]. We

hypothesize that lineage subdivision occurred as a result of geographic range expansion and subsequent diversification processes resulting in gene flow limitation across climate regimes defined by latitude. Here, we use comparative genomics to evaluate patterns of genetic diversity, differentiation, and gene flow between southern-derived (SDR) and northern-derived (NDR) lineages, and we evaluate physiological characteristics of strains from these lineages to identify adaptive constraints, which could account for dispersal limitation.

## Materials and methods

### *Streptomyces* isolation and DNA extraction

A *Streptomyces* culture collection was assembled from 15 sites spanning the United States, as previously described [6, 7, 26]. Briefly, sites were selected to represent a narrow range of ecological characteristics including meadow, pasture, or native grasslands dominated by perennial grasses and having moderately acidic soil (pH:  $6.0 \pm 1.0$ ; mean  $\pm$  s.d.). Climate data for each site were obtained from the National Centers for Environmental Information (NOAA) (<https://www.ncdc.noaa.gov>). Surface soils were sampled at 0–5 cm depth, and strains were isolated using uniform conditions as this will select for strains with similar physiological characteristics.

*Streptomyces* strains were isolated from air-dried soils on glycerol–arginine (pH 8.7) agar containing antifungals cycloheximide and Rose Bengal [38, 39], as previously described [40]. Pure cultures were obtained by streaking for isolation on glycerol–arginine (pH 8.7) agar. For genome sequencing, high-quality genomic DNA was extracted, using a standard phenol/chloroform/isoamyl alcohol protocol [41], from pure cultures grown in liquid yeast extract-malt extract medium (YEME) with 0.5% glycine [31] for 72 h at 30 °C with shaking.

### Identification of *Streptomyces* sister-taxa

First, we assessed the genetic diversity of the entire culture collection using the *rpoB* locus, which encodes the RNA polymerase beta-subunit. This initial screen allowed us to identify 107 phylogroups [6, 7]. We then analyzed six of the most abundant phylogroups using multilocus sequence analysis (MLSA) [26]. From this work, we identified phylogroups WA1063 and MS152, which we hypothesize to constitute recently diverged sister-taxa representing NDR and SDR lineages, respectively. We also identified additional strains from our culture collection that shared close phylogenetic affiliation across MLSA loci with phylogroups WA1063 and MS152. These sister-taxa lineages were

selected for the present work because they share a recent common ancestor, inhabit discrete geographic ranges, and sufficient representatives of each lineage were present in our culture collection to support robust comparative genomic approaches. Furthermore, the analysis of physiologically similar strains from ecologically comparable sites improves our ability to detect biogeographical patterns that result from drift by minimizing the effects of environmental selection (as reviewed in [13]).

### Genome sequencing, assembly, and annotation

Genomic libraries were prepared for 24 strains using the Nextera DNA Library Preparation Kit (Illumina, San Diego, CA, USA). Draft genome sequences were generated using paired-end reads ( $2 \times 100$  bp) sequenced on an Illumina HiSeq2500 instrument located at the Cornell University Biotechnology Resource Center (BRC). The A5 pipeline was used for quality control and genome assembly [42], and assembled genomes were annotated with the online RAST Server [43]. Orthologous genes were identified using ITEP [44] with MCL clustering parameters as follows: inflation value = 2.0, cutoff = 0.04, maxbit score. Genome sequences are available at NCBI under BioProject PRJNA401484 accession numbers SAMN07606143–SAMN07606166.

### Phylogenetic and genomic analyses

Phylogenetic relationships were reconstructed from multiple genome alignment nucleotide sequences. Multiple genome alignments were performed using Mugsy [45], and poorly aligned regions were removed using trimAl v1.2 [46]. Whole-genome pairwise average nucleotide identity (ANI) was calculated from multiple whole-genome alignments with mothur [47]. Maximum likelihood (ML) trees were constructed from nucleotide sequences of multiple genome alignments using the generalized time reversible nucleotide substitution model [48] with gamma distributed rate heterogeneity among sites (GTRGAMMA) in RAxML v7.3.0 [49]. Bootstrap support was determined following 20 ML searches with 100 inferences using the RAxML rapid bootstrapping algorithm [50].

All clade-level characteristics were evaluated using nucleotide sequences for each of 2778 single-copy orthologous genes conserved in all 24 genomes. Nucleotide sequences of orthologous genes were aligned using MAFFT v.7 [51], and poorly aligned regions were removed with Gblocks. The extent of genomic differentiation between lineages inferred from  $F_{ST}$  values calculated using arlecore [52]. ANI of orthologous genes was determined using mothur [47]. Additional diversity metrics and were calculated using the R packages pegas [53] and ape [54, 55].

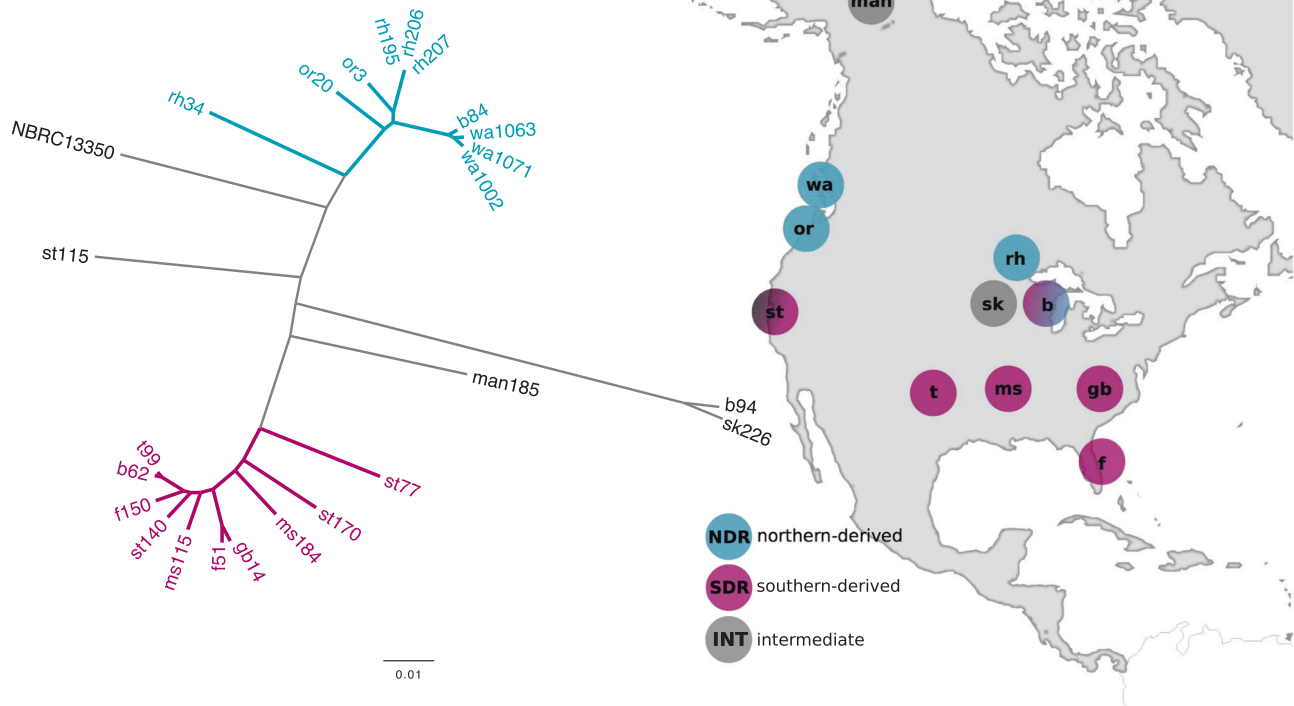
### Evaluation of physiological traits

Wild-type *Streptomyces* strains do not grow uniformly, but instead tend to aggregate, when suspended in liquid media, so growth rates cannot be determined by standard turbidity measurements made on liquid cultures. Hence, growth was evaluated using two different approaches: (1) by measuring change in absorbance of cultures grown on solid medium, or (2) by measuring biomass production (see below). Spore suspensions were generated from mature aerial hyphae by washing mycelia with Tween®80 (Sigma Aldrich, St. Louis, MO, USA), filtering the spore suspension through sterile cheese cloth, and centrifuging the suspension at 1800 g for 5 min to obtain a spore pellet. Spore pellets were suspended in 20% sterile glycerol. Spores were counted using a Petroff-Hausser Counting Chamber, and spore suspensions were stored at  $-80^{\circ}\text{C}$ .

Absorbance-based assays were performed by inoculating  $1 \times 10^{12}$ – $1 \times 10^{13}$  spores onto 1 ml of glycerol–arginine (pH 8.7) agar within the well of a 24-well CELLSTAR® culture plate (Greiner Bio-One, Monroe, NC, USA). Growth was measured by determining change in absorbance at 450 nm in a FilterMax™ F5 plate reader (Molecular Devices, Sunnyvale, CA, USA). As *Streptomyces* growth results from hyphal elongation and septation and not binary fission, we determined that the change in ABS 450 nm over time approximated a linear relationship, and hence we quantified growth rate as the slope for ABS 450 nm  $\text{hr}^{-1}$  across the entire growth curve. Biomass production assays were performed by spreading  $1 \times 10^{13}$ – $1 \times 10^{14}$  total spores onto pre-weighed 0.22  $\mu\text{M}$  GV Durapore® membrane filters (Millipore, Billerica, MA, USA) that were placed onto glycerol–arginine (pH 8.7) agar. Following growth, filters were removed, oven dried for 5–7 days, and biomass production was calculated by subtracting the mass of the filter paper.

To identify differences in thermal adaptation between lineages, we calculated the growth rate of each strain relative to the growth rate of the inferred ancestral population. In this way, relative growth rate,  $\text{GR}_{\text{rel}}$ , was defined as  $(\text{GR}_{\text{str}} - \text{GR}_{\text{ref}}) / \text{GR}_{\text{ref}}$ , where  $\text{GR}_{\text{str}}$  is the absolute growth rate of a strain and the reference growth rate,  $\text{GR}_{\text{ref}}$ , is defined as the average growth rate of the SDR population. Thus, a value of  $\text{GR}_{\text{rel}}$  equal to zero indicates no difference in growth rate relative to the average of the inferred ancestral population (i.e., SDR).

We assayed the ability to grow on a range of single carbon sources that have been shown to phenotypically differentiate strains within the *S. griseus* species complex [56], including arabinose, arginine, citrate, glycine, lactose, lactate, leucine, lysine, rhamnose, and succinate. Cultures were incubated at  $30^{\circ}\text{C}$  for 7 days on M9 (pH 7) minimal media [57] solidified with Ultrapure noble agar (Affymetrix, Santa Clara, CA, USA) and containing a final concentration



**Fig. 1** *Streptomyces* lineages NDR and SDR were partitioned across sites with respect to latitude. The phylogenetic tree was constructed from multiple whole-genome alignments using maximum likelihood and a GTRGAMMA model of evolution. Scale bar represents nucleotide substitutions per site. Bootstrap values for all internal nodes are 100%. Colored clades depict the northern-derived (NDR) and southern-derived (SDR) populations. The remaining strains belong to

the intermediate groups (INT). Strain names reflect the location they were isolated from (Table S1). Genome NBRC13350 is the type strain *Streptomyces griseus* subsp. *griseus* NBRC 13350. The geographic distribution of strains is illustrated in the right panel. Sample locations are labeled with their site codes, which are found in Table S1 and are colored to reflect population ranges (color figure online)

of 1% carbon source. Evidence of growth was scored on a four-point scale. Growth on dextrose was used as positive control, as all strains used this substrate, whereas growth without a carbon source was used as negative control.

Antibiotic resistance was determined by plating a lawn of spores on glycerol–arginine (pH 8.7) agar and measuring the zone of inhibition around discs containing antibiotics ampicillin (10 µg), chloramphenicol (30 µg), erythromycin (15 µg), kanamycin (30 µg), oxytetracycline (30 µg), penicillin (10 IU), tetracycline (30 µg), and streptomycin (10 µg) after incubation at 30 °C for 7 days.

## Results

### Genome divergence between *Streptomyces* sister-taxa

We generated 24 whole-genome sequences from *Streptomyces* strains, which originated from geographically

disparate grassland sites spanning over 6000 km (Fig. 1, Table S1). Strain names reflect the site from which they were isolated (Table S1). NDR genomes belong to the previously described phylogroup WA1063, and SDR genomes belong to phylogroup MS152 [26]. The four genomes we define as intermediate groups (INT) were isolated from both northern (isolates man185, sk226, and b94) and southern (isolate st115) sites (Fig. 1, Table S2). These four isolates were initially expected to affiliate with NDR or SDR based on housekeeping loci, but whole-genome sequencing reveals that these four genomes belong to distinct lineages (Fig. 1). Assembled draft genomes are 7.5–9.1 Mb with 6776–8078 predicted open reading frames (ORF) and have a G + C content of 71.4–72.5% (Table S2). The core genome, defined conservatively here as orthologous genes present in all genomes, contains 3234 genes (2778 single-copy genes), which encompasses 40–48% of total genes per strain. A total of 22,054 genes were identified across all genomes, with 8785 of these genes being strain specific. The core and pan-genomes of NDR



contained 4234 and 9447 genes, respectively, and of these 6565 genes were lineage specific (i.e., found in NDR isolates but not found in SDR genomes). The core and pan-genomes of SDR contained 4400 and 7859 genes, respectively, of these 5143 genes were lineage specific (i.e., found in SDR isolates but not found in NDR genomes).

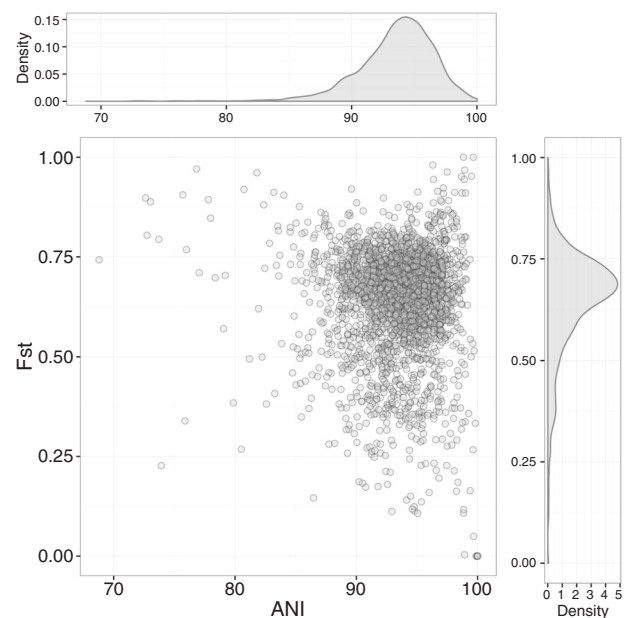
We evaluated pairwise genome-wide ANI to determine the genetic relatedness between sister-taxa lineages. Typically, strains that belong to the same genomic species have ANI > 95% [58]. NDR and SDR each have intra-lineage genome-wide ANI between 95.6–99.9%. Inter-lineage ANI between NDR and SDR genomes is 92.6–93.3%, and all 24 genomes have pairwise ANI > 91%. In addition, the NDR and SDR genomes all have pairwise ANI in the range of 91–94.6% with *Streptomyces griseus* subsp. *griseus* NBRC 13350, which is the most closely related type strain (Fig. 1). These results are consistent with the conclusion that SDR and NDR represent recently diverged yet genetically distinct species, and that NBRC 13350 and the intermediate group isolates b94 plus sk226, st115, and man185 represent species independent of SDR and NDR (Fig. 1). Furthermore, NDR and SDR genomes share nearly identical 16S ribosomal RNA (rRNA) gene sequences, with inter-lineage 16S rRNA gene distances of 0–0.21% ( $0.12 \pm 0.066\%$ , mean  $\pm$  s.d.).

### Dispersal limitation and allopatric divergence

We assessed genome-wide patterns of diversity and gene flow between lineages to identify signatures of allopatric diversification consistent with dispersal limitation. The average inter-lineage (NDR versus SDR) ANI across core genes is  $93.5 \pm 3.2\%$  (mean  $\pm$  s.d.), and the ANI distribution is skewed-left with a few rare genes demonstrating high genetic diversity (i.e., those genes with < 80% ANI) (Fig. 2). Next, we examined potential differentiation of the observed genetic diversity between NDR and SDR using  $F_{ST}$ -statistics. Inter-lineage  $F_{ST}$  values across core genes are  $0.64 \pm 0.13$  (mean  $\pm$  s.d.), with 14.9% ( $n = 415$ ) of genes demonstrating high levels of differentiation (i.e., those genes with  $F_{ST} \geq 0.75$ ) (Fig. 2). The distribution of  $F_{ST}$  values is right-skewed, indicating that a small number of genes are poorly differentiated between lineages due to current or recent gene flow (Fig. 2).

### Latitudinal partitioning of genomic diversity

We next evaluated genome-wide intra-lineage diversity for NDR and SDR as a function of ANI, proportion of segregating sites ( $S$ ), and nucleotide diversity ( $\pi$ ) for each core gene (Fig. 3). ANI is significantly greater for NDR compared with SDR (Mann–Whitney  $U$ -test:  $P < 0.0001$ ; Fig. 3a), indicating more standing diversity in SDR core



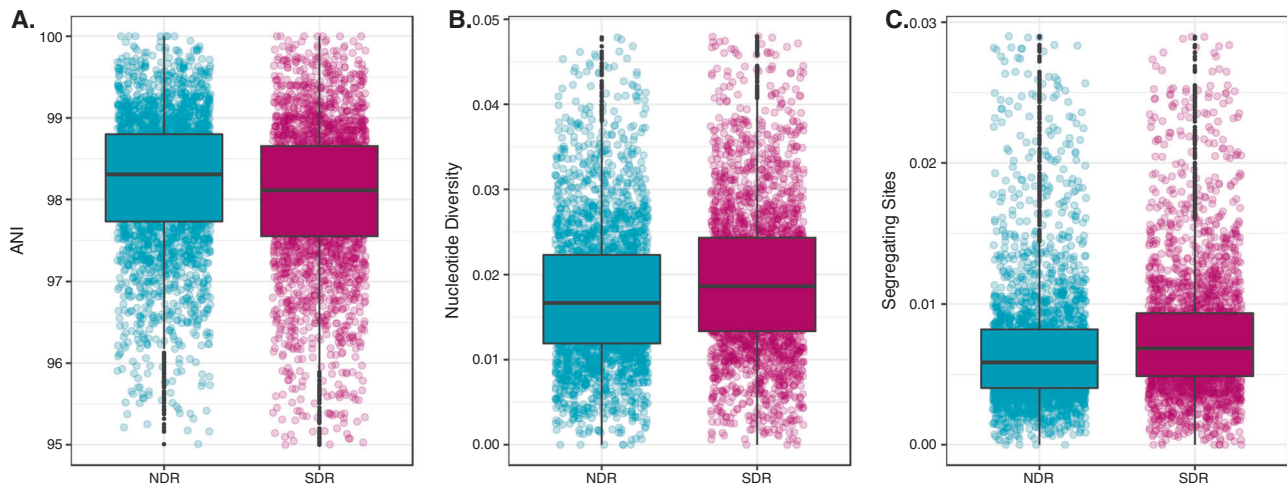
**Fig. 2** NDR and SDR lineages exhibit substantial genome-wide genetic differentiation as revealed by pairwise values of genetic differentiation ( $F_{ST}$ ) and average nucleotide identity (ANI) for 2778 single-copy core genes. Kernel density distributions for gene-by-gene ANI and  $F_{ST}$  are displayed on the corresponding axes

gene sequences. SDR genes also demonstrate significantly greater nucleotide diversity (Mann–Whitney  $U$ -test:  $P < 0.0001$ ) (Fig. 3b) and proportion of segregating sites (Mann–Whitney  $U$  test:  $P < 0.0001$ ) (Fig. 3c) compared with NDR genes.

### Temperature-dependent growth phenotypes

*Streptomyces* spores are totipotent, and relative fitness (in terms of fecundity) is directly related to growth rate. NDR grows approximately 10 times faster than SDR at 4 °C (Fig. 4) indicating greater average fitness than SDR in colder sites. In contrast, SDR grows approximately twice as fast as NDR at temperatures of 23 °C and above (Fig. 4) indicating greater average fitness than NDR in warmer sites. The mean annual temperatures of sites occupied by NDR and SDR were  $8.7 \pm 5.6$  °C (mean  $\pm$  s.d.), and  $16.3 \pm 5.5$  °C (mean  $\pm$  s.d.), respectively (Table S1).

We found a significant interaction between the effects of lineage (NDR versus SDR) and temperature (4 °C, 23 °C, 30 °C, 35 °C) when assessed either as growth rate (Fig. 4; analysis of variance (ANOVA):  $F_{3,72} = 25.6$ ,  $P < 0.0001$ ), or total biomass production (Figure S2; ANOVA:  $F_{3,144} = 13.8$ ,  $P < 0.0001$ ). SDR grows faster than NDR at high temperatures both in terms of relative growth rate (Fig. 4; Mann–Whitney  $U$ -test: 23 °C,  $P = 0.014$ ; 30 °C,  $P = 0.023$ ; 35 °C,  $P = 0.0088$ ), and biomass production (Figure S2; Mann–Whitney  $U$ -test: 23 °C,  $P < 0.0001$ ; 30 °C,  $P < 0.0001$ ;



**Fig. 3** Overall genomic diversity is lower in NDR compared with SDR as assessed by pairwise analysis of 2778 single-copy core genes with respect to average nucleotide identity (ANI) **a**, percentage of segregating sites **b**, and per site nucleotide diversity **c**. Boxplots show the interquartile range of values. The median is indicated by the central

horizontal line, whiskers indicate the 1.5 interquartile range beyond the box boundaries, and black points indicate outlying values. Colored circles further illustrate the distributions of diversity metrics (color figure online)

35 °C,  $P < 0.0001$ ). In contrast, NDR grows faster than SDR at low temperature when assessed as relative growth rate (Fig. 4; Mann–Whitney  $U$ -test: 4 °C,  $P = 0.00044$ ), but our biomass assay likely lacked sensitivity to discern differences in biomass production at the lowest growth rates. Growth at 45 °C was undetectable for all strains and undetectable at 40 °C for all strains except SDR isolates ms184 and st77 (Table S3). The use of relative growth rates (Fig. 4) standardizes the effect size across temperature, although the same conclusions can be drawn from absolute growth rates (Figure S1).

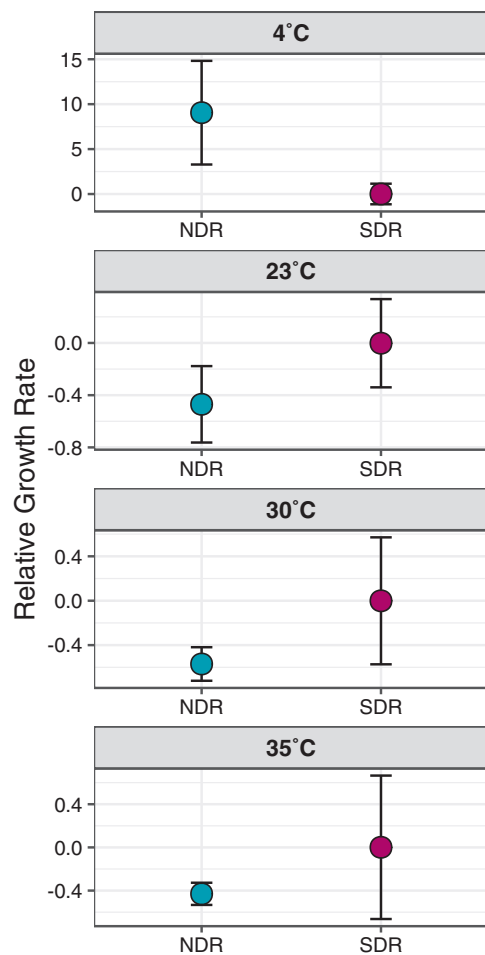
### Carbon source utilization and antibiotic resistance profiles

We found that carbon source utilization profiles weakly differentiate NDR and SDR populations (permutational multivariate analyses of variance (PERMANOVA):  $R^2 = 0.2$ ,  $P = 0.039$ ; Figure S3A). However, carbon source utilization was not strongly conserved within each lineage and the frequency of strains found to use each carbon source did not differ significantly between NDR and SDR for any of the carbon sources tested. Antibiotic resistance profiles also weakly differentiate NDR and SDR lineages (PERMANOVA:  $R^2 = 0.22$ ,  $P = 0.004$ ; Figure S3B). We found a significant difference between NDR and SDR with respect to kanamycin resistance (ANOVA:  $F_{1,18} = 12.6$ ,  $P = 0.0023$ ) and streptomycin resistance (ANOVA:  $F_{1,18} = 8.5$ ,  $P = 0.0093$ ). We also found that NDR has significantly lower growth rates than NDR when exposed to streptomycin (Mann–Whitney  $U$ -test:  $P < 0.0001$ ) (Table S3, Figure S4).

### Discussion

The historical legacy of climatic events, which occurred during the late Pleistocene persists in the biogeography of North American plants and animals [59, 60]. Similarly, the latitudinal diversity gradient we observe in our *Streptomyces* culture collection can also be explained by historical changes in climate (previously presented in [6, 7, 26]). We hypothesize that, much like North American macroorganisms [59, 60], ancestral *Streptomyces* lineages dispersed northward from glacial refugia, colonizing glacial retreat zones, and adapting to these newly colonized habitats. Furthermore, we hypothesize that subsequent limits to dispersal across climate regimes contributed to a contemporary latitudinal diversity gradient by creating barriers to gene flow that reinforced divergence between northern and southern lineages. Here, we examine the genomic attributes and physiological traits of recently diverged *Streptomyces* sister-taxa to determine whether lineage divergence could be explained by adaptive differences, which imposed limits to dispersal.

Analyses of sister-taxa provide a particularly powerful approach for determining evolutionary and ecological processes driving microbial speciation and diversification. Here, we generated 24 whole-genome sequences from sister-taxa, which are affiliated with the *Streptomyces griseus* species-cluster and which originated from geographically disparate grassland sites spanning over 6000 km (Fig. 1, Table S1). Whole-genome phylogeny confirms that NDR and SDR represent independent species-like clusters (Fig. 1). Independent bacterial species generally have genome-wide ANI  $< 95\%$  [58, 61], and we show that strains



**Fig. 4** NDR exhibits a tradeoff in thermal traits with higher relative growth rate than SDR at 4 °C and lower relative growth rate at 23 °C, 30 °C, and 35 °C. Relative growth rate is defined as  $(GR_{str} - GR_{ref}) / GR_{ref}$ . Circles depict the population-level mean values, and error bars depict the population-level standard deviations. The same trend is observed in absolute growth rate (Figure S1), but the use of relative growth rate standardizes the effect size for differences in growth rate across temperature

within NDR and SDR have pairwise genome-wide ANI > 95%. Although NDR and SDR represent distinct genomic species with distinct geographic ranges, they typify closely related sister-taxa as comparison of strains between NDR and SDR indicate pairwise genome-wide ANI of  $93 \pm 0.0014\%$  (mean  $\pm$  s.d.), just beyond the standard cutoff for species.

Historical hypotheses accounting for the latitudinal diversity gradient make the prediction that species richness [24, 62], as well as intraspecific diversity, decrease with increasing latitude [63, 64]. We previously described in *Streptomyces* an inverse relationship between latitude and community species richness [6, 7], as well as intra-population nucleotide diversity [26]. Here, we find further evidence to support this biogeographic pattern and show that genome-wide diversity partitions with latitude in sister-

taxa, with NDR genomes demonstrating significantly lower genetic and nucleotide diversity relative to SDR (Fig. 3).

Evidence for horizontal gene transfer (HGT) among bacteria is widespread [65, 66], and rates of HGT estimated for *Streptomyces* are among some of the highest reported [40, 67]. Given that rates of gene exchange are known to increase with increasing genetic similarity between lineages [68], we would predict high levels of gene flow between NDR and SDR. Although we find that NDR and SDR share 16 identical alleles, suggesting contemporary horizontal gene exchange processes, such alleles can also be acquired vertically from ancestral gene exchange and incomplete lineage sorting [6, 7, 26]. Furthermore, the spatial distribution of haplotypes within and among *Streptomyces* species indicates allopatric diversification resulting from dispersal and gene flow limitation [26], and these results suggest that most gene flow between SDR and NDR likely occurred prior to lineage subdivision.

Genome analysis of sister-taxa is consistent with these observations. The genome-wide estimate of  $F_{ST}$  for single-copy core genes between NDR and SDR is  $0.64 \pm 0.1$  (mean  $\pm$  s.d.), indicating that these are strongly differentiated lineages with minimal rates of contemporary gene exchange between NDR and SDR. Physical barriers to gene flow, including geographic isolation, generally produce elevated  $F_{ST}$  values distributed across the microbial genome as opposed to ecological barriers, which can create genomic ‘islands’ and eventually ‘continents’ of high differentiation, depending on the extent of speciation [69]. For example, high rates of gene flow are observed across a majority of the genome for ecological populations of *Vibrio cyclitrophicus*, and fixation of just a few ‘ecoSNPs’ suggest incipient speciation [70]. Alternatively, for co-occurring lineages of *Sulfolobus islandicus*, an archaeon found in geothermal hot springs, most of the genome is highly differentiated ( $F_{ST} > 0.5$ ) with few regions of low differentiation ( $F_{ST} < 0.5$ ), suggesting barriers to recombination and consistent with latter speciation processes [71]. Hence, the extent of genome-wide differentiation between NDR and SDR suggests later-stage speciation processes, and in combination with the fact that NDR and SDR inhabit discrete geographic regions (Fig. 1), this is generally consistent with restricted gene flow due to geographic isolation (Fig. 3).

Lineage divergence due to gene flow limitation can result from both adaptive and non-adaptive processes associated with colonization dynamics (as described in [72]). Phylogenetic niche conservatism (PNC) is the tendency of species to preserve ecological attributes associated with their ancestral niche [73, 74] and is often associated with adaptive traits that constrain the geographical or ecological range of a species. As a result, PNC is associated with the formation of latitudinal diversity gradients [37, 75, 76]. For example, tropical niche conservatism predicts that

phylogenetic conservatism of traits ancestral to tropical species will limit successful colonization of higher latitudes, thereby limiting the number of species that disperse to and diversify within temperate sites [37]. Alternatively, PNC can result in inverse latitudinal diversity gradients for species with temperate origins [77]. We hypothesize that the last common ancestor of our sister-taxa inhabited the contemporary range of SDR (Fig. 1). Following this logic, we evaluated a range of physiological phenotypes for NDR and SDR to determine to what extent niche-related adaptations are phylogenetically conserved. In particular, we hypothesized that temperature-related phenotypes would be important drivers of adaptive differences between lineages.

Dispersal limitation and phylogenetic trait conservation are expected to strengthen patterns of PNC [36]. Niche conservatism and phylogenetic signal is discernible for a range of bacterial phenotypic and morphological traits [78], but generally, complex poly-genic microbial traits, such as thermal adaptation, exhibit greater phylogenetic conservation than simple traits [79]. We observe significant thermal trait tradeoffs in NDR and SDR, with NDR having higher growth rate at low temperatures and SDR having higher growth rate at higher temperatures (Fig. 4). With the exception of two SDR strains (st77 and t99), which did not grow at 4 °C, all strains were able to grow across the temperature range of 4–35 °C. These strains were all isolated at the same pH from ecologically similar habitats using the exact same selective conditions, and there were no categorical differences in carbon source utilization, and so presumably, in the absence of competition, all strains should be able to colonize all sites regardless of latitude. Based on these observations, it seems likely that ancestral lineages could have expanded northward and established new populations in previously un-colonized sites. Within this new range, selection would have favored strains having greater cold tolerance resulting in adaptation over time. A similar demographic history has been documented for populations of many North American plants [80].

There is evidence that thermal adaptation in bacteria can be rapid and highly plastic in the absence of interspecific competition [81, 82]. However, differences in growth rates indicate that NDR has a competitive advantage over SDR in cooler sites, whereas the reverse is true of SDR (Fig. 4). Hence, interspecific competition likely poses a barrier to successful establishment in previously colonized sites, thereby promoting genetic isolation and divergence of northern lineages from their southern ancestors. The distributions of many soil bacteria, including *Actinobacteria*, appear to be temperature sensitive [83]. Thus, we propose that the divergence of SDR and NDR was driven by dispersal following glacial retreat, followed by thermal adaptation, and the imposition of barriers to colonization. Under this scenario, we would predict transient co-existence of

competitive species along range boundaries. We note that one SDR strain (b62) was observed in Wisconsin despite having the SDR thermal phenotype. We do not yet know the true range boundaries for any species of *Streptomyces* but the observation of both SDR and NDR in WI suggests that this region, which falls at the maximum extent of historical glaciation [84], may represent a range boundary for these lineages.

We also found evidence for PNC in both carbon utilization (Figure S3A) and antibiotic resistance (Figure S3B, Figure S4). These data support niche partitioning between lineages with respect to thermal adaptation, antibiotic resistance, and carbon use. Although we cannot definitely say that one adaptive trait is more important than another in defining adaptive barriers to colonization, no single carbon source was able to clearly differentiate these lineages. Thus, thermal adaptation and competitive antagonism are more likely to constrain the contemporary dispersal of these microbes than carbon source utilization. Whereas movement between climate regimes creates obvious selection pressure to favor thermal adaptation, it is less clear how such dispersal would favor adaptive divergence in antibiotic resistance profiles. We expect that divergence in antibiotic resistance profiles results from drift, or local adaptive process, but that these local changes further act to reinforce colonization barriers over time.

Although thermal adaptation could be sufficient to explain latitudinal barriers to dispersal across different climate regimes, non-adaptive processes can produce similar patterns of dispersal limitation. For example, density-dependent priority effects can generate non-adaptive limits to dispersal [35]. The ‘founder takes all’ hypothesis posits that early dispersers easily colonize uninhabited niches, but that late dispersers find niche access blocked by prior occupancy [25, 35]. Naturally, this density-dependent blocking would be enhanced by even slight adaptive differences between the indigenous (early-dispersing) and invading (late-dispersing) strains, and hence both adaptive and non-adaptive mechanisms should be self-reinforcing.

The sister-taxa NDR and SDR were isolated from ecologically similar habitats, yet they appear to occupy distinct geographic range as defined by latitude. Genomic analyses show genetic differentiation between these lineages, with low levels of inter-lineage gene flow, and lower genetic diversity in NDR relative to SDR. We observe evidence of thermal adaptation between lineages, with NDR having higher growth rates at cold temperatures and lower growth rates at high temperatures. These results are broadly consistent with hypotheses that link latitudinal diversity gradients to observations of PNC [37, 75, 76], and with the hypothesis that this diversity gradient was formed by the historical effects of paleoclimate dynamics during the late Pleistocene. Similar patterns of biogeography are found



widely in many North American plants and animals, but to our knowledge, this is the first documentation of comparable events shaping diversification of microbial lineages. A better understanding of the historical events leading to contemporary biogeographical patterns within *Streptomyces* may help to further illuminate the mechanisms creating and maintaining biodiversity in soils. We think it will be particularly interesting to explore the degree to which antibiotic production and antagonism contribute to dispersal limitation imposed by priority effects.

**Acknowledgements** We would like to thank Mary Jo Choudoir and Dunroven Farm for providing soil samples from northern Wisconsin. This material is based upon work supported by the National Science Foundation under grant no. DEB-1456821 awarded to Daniel H. Buckley.

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

### References

- Hillebrand H. On the generality of the latitudinal diversity gradient. *Am Nat.* 2004;163:192–211.
- Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL, et al. A latitudinal diversity gradient in planktonic marine bacteria. *Proc Natl Acad Sci USA.* 2008;105:7774–8.
- Pommier T, Canbäck B, Riemann L, Boström KH, Simu K, Lundberg P, et al. Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol.* 2006;16:867–80.
- Sul WJ, Oliver TA, Ducklow HW, Amaral-Zettler LA, Sogin ML. Marine bacteria exhibit a bipolar distribution. *Proc Natl Acad Sci USA.* 2013;110:2342–7.
- Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, Gonzalez JM, et al. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc Natl Acad Sci USA.* 2013;110:11463–8.
- Andam CP, Choudoir MJ, Nguyen VA, Park SH, Buckley DH. Contributions of ancestral inter-species recombination to the genetic diversity of extant *Streptomyces* lineages. *ISME J.* 2016a;10:1731–41.
- Andam CP, Doroghazi JR, Campbell AN, Kelly PJ, Choudoir MJ, Buckley DH. A latitudinal diversity gradient in terrestrial bacteria of the genus *Streptomyces*. *mBio.* 2016b;7:e02200–15.
- Zhou J, Deng Y, Shen L, Wen C, Yan Q, Ning D, et al. Temperature mediates continental-scale diversity of microbes in forest soils. *Nat Commun.* 2016;7:12083.
- Fierer N, Lennon JT. The generation and maintenance of diversity in microbial communities. *Am J Bot.* 2011;98:439–48.
- Green J, Bohannan BJM. Spatial scaling of microbial biodiversity. *Trends Ecol Evol.* 2006;21:501–7.
- Locey KJ, Lennon JT. Scaling laws predict global microbial diversity. *Proc Natl Acad Sci USA.* 2016;113:5970–5.
- Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al. Microbial biogeography: putting microorganisms on the map. *Nat Rev Microbiol.* 2006;4:102–12.
- Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH. Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol.* 2012;10:247–506.
- Choudoir MJ, Campbell AN, Buckley DH. Grappling with Proteus: population level approaches to understanding microbial diversity. *Front Microbiol.* 2012;3:336.
- Finlay BJ. Global dispersal of free-living microbial Eukaryote species. *Science.* 2002;296:1061–3.
- Finlay BJ, Fenchel T. Cosmopolitan metapopulations of free-living microbial eukaryotes. *Protist.* 2004;155:237–44.
- Mittelbach GG, Schemske DW, Cornell HV, Allen AP, Brown JM, Bush MB, et al. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol Lett.* 2007;10:315–31.
- Pianka EricR. Latitudinal gradients in species diversity: a review of concepts. *Am Nat.* 1966;100:33–46.
- Currie DJ, Mittelbach GG, Cornell HV, Field R, Guegan J-F, Hawkins BA, et al. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol Lett.* 2004;7:1121–34.
- Field R, Hawkins BA, Cornell HV, Currie DJ, Diniz-Filho JAF, Guégan J-F, et al. Spatial species-richness gradients across scales: a meta-analysis. *J Biogeogr.* 2009;36:132–47.
- Allen AP, Brown JH, Gillody JF. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science.* 2002;297:1545–8.
- Stevens RD. Historical processes enhance patterns of diversity along latitudinal gradients. *Proc Biol Sci.* 2006;273:2283–9.
- Wiens JJ, Donoghue MJ. Historical biogeography, ecology and species richness. *Trends Ecol Evol.* 2004;19:639–44.
- Stephens PR, Wiens JJ. Explaining species richness from continents to communities: the time-for-speciation effect in Emydid turtles. *Am Nat.* 2003;161:112–28.
- Hewitt GM. Some genetic consequences of ice ages, and their role, in divergence and speciation. *Biol J Linn Soc Lond.* 1996;58:247–76.
- Choudoir MJ, Doroghazi JR, Buckley DH. Latitude delineates patterns of biogeography in terrestrial *Streptomyces*. *Environ Microbiol.* 2016;18:4931–45.
- Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Ravel J, et al. Pre-Columbian origins for North American anthrax. *PLoS ONE.* 2009;4:e4813.
- Eisenlord SD, Zak DR, Upchurch RA. Dispersal limitation and the assembly of soil *Actinobacteria* communities in a long-term chronosequence. *Ecol Evol.* 2012;2:538–49.
- Delgado-Baquerizo M, Bissett A, Eldridge DJ, Maestre FT, He J-Z, Wang J-T, et al. Palaeoclimate explains a unique proportion of the global variation in soil bacterial communities. *Nat Ecol Evol.* 2017;1:1339–47.
- Kämpfer P. The family *Streptomycetaceae* Part I: taxonomy. *The Prokaryotes.* New York, NY: Springer; 2006. p. 538–604. Vol. 3
- Kieser T, Bibb MJ, Buttner MJ, Charter KF, Hopwood DA. *Practical Streptomyces Genetics.* Norwich, UK: John Innes Foundation; 2000.
- Ljungdhal LG, Eriksson K. The ecology of microbial cellulose degradation. In: *Advances in microbial ecology.* Vol. 8. Springer: New York, USA; 1985. p. 237–99.
- Schrempf H. Recognition and degradation of chitin by streptomycetes. *Antonie Van Leeuwenhoek.* 2001;79:285–9.
- Wavre M, Tickoo R, Jog M, Bhole B. How many antibiotics are produced by the genus *Streptomyces*? *Arch Microbiol.* 2001;176:386–90.
- Waters JM, Fraser CI, Hewitt GM. Founder takes all: density-dependent processes structure biodiversity. *Trends Ecol Evol.* 2013;28:78–85.
- Crisp MD, Cook LG. Phylogenetic niche conservatism: what are the underlying evolutionary and ecological causes? *New Phytol.* 2012;196:681–94.

37. Wiens JJ, Ackerly DD, Allen AP, Anacker BL, Buckley LB, Cornell HV, et al. Niche conservatism as an emerging principle in ecology and conservation biology. *Ecol Lett*. 2010;13:1310–24.
38. El-Nakeeb MA, Lechevalier HA. Selective isolation of aerobic *Actinomycetes*. *Appl Microbiol*. 1963;11:75–55.
39. Ottow JCG. Rose bengal as a selective aid in the isolation of fungi and *Actinomycetes* from natural sources. *Mycologia*. 1972;64:304.
40. Doroghazi JR, Buckley DH. Widespread homologous recombination within and between *Streptomyces* species. *ISME J*. 2010;4:1136–43.
41. Roberts MA, Crawford DL. Use of randomly amplified polymorphic DNA as a means of developing genus-and strain-specific *Streptomyces* DNA probes. *Appl Environ Microbiol*. 2000;66:2555–64.
42. Tritt A, Eisen JA, Facciotti MT, Darling AE. An integrated pipeline for de novo assembly of microbial genomes. *PLoS ONE*. 2012;7:e42304.
43. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom*. 2008;9:75.
44. Benedict MN, Henriksen JR, Metcalf WW, Whitaker RJ, Price ND. ITEP: an integrated toolkit for exploration of microbial pan-genomes. *BMC Genom*. 2014;15:8.
45. Angiuoli SV, Salzberg SL. Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics*. 2011;27:334–42.
46. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 2009;25:1972–3.
47. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009;75:7537–41.
48. Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, American Mathematical Society. 1986;17:57–86.
49. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22:2688–90.
50. Stamatakis A, Hoover P, Rougemont J, Renner S. A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol*. 2008;57:758–71.
51. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.
52. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10:564–7.
53. Paradis E. pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*. 2010;26:419–20.
54. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*. 2004;20:289–90.
55. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 2007;56:564–77.
56. Kämpfer P, Kroppenstedt RM, Dott W. A numerical classification of the genera *Streptomyces* and *Streptoverticillium* using miniaturized physiological tests. *Microbiology*. 1991;137:1831–91.
57. Sambrook J, Russell RW. *Molecular cloning: a laboratory manual*. 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.
58. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA*. 2005;102:2567–72.
59. Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci*. 2004;359:183–95.
60. Shafer ABA, Cullingham CI, Côté SD, Coltman DW. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Mol Ecol*. 2010;19:4589–621.
61. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol*. 2007;57:81–91.
62. Hawkins BA, Porter EE. Relative influences of current and historical factors on mammal and bird diversity patterns in deglaciated North America. *Glob Ecol Biogeogr*. 2003;12:475–81.
63. Adams RI, Hadly EA. Genetic diversity within vertebrate species is greater at lower latitudes. *Evol Ecol*. 2013;27:133–43.
64. Bernatchez L, Wilson CC. Comparative phylogeography of Nearctic and Palearctic fishes. *Mol Ecol*. 1998;7:431–52.
65. Gogarten JP, Doolittle WF, Lawrence JG. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol*. 2002;19:2226–38.
66. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature*. 2000;405:299.
67. Doroghazi JR, Buckley DH. Intraspecific comparison of *Streptomyces pratensis* genomes reveals high levels of recombination and gene conservation between strains of disparate geographic origin. *BMC Genom*. 2014;15:970.
68. Fraser C, Hanage WP, Spratt BG. Recombination and the nature of bacterial speciation. *Science*. 2007;315:476–80.
69. Krause DJ, Whitaker RJ. Inferring speciation processes from patterns of natural variation in microbial genomes. *Syst Biol*. 2015;64:926–35.
70. Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabó G, et al. Population genomics of early events in the ecological differentiation of bacteria. *Science*. 2012;336:48–51.
71. Cadillo-Quiroz H, Didelot X, Held NL, Herrera A, Darling A, Reno ML, et al. Patterns of gene flow define species of thermophilic *Archaea*. *PLoS Biol*. 2012;10:e1001265.
72. Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Mol Ecol*. 2013;22:5983–99.
73. Harvey PH, Pagel MD. *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press; 1991.
74. Losos JB. Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecol Lett*. 2008;11:995–1003.
75. Buckley LB, Davies TJ, Ackerly DD, Kraft NJB, Harrison SP, Anacker BL, et al. Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proc R Soc Lond Biol*. 2010;277:2131–8.
76. Wiens JJ, Graham CH. Niche conservatism: integrating evolution, ecology, and conservation biology. *Annu Rev Ecol Evol Syst*. 2005;36:519–39.
77. Morinière J, Van Dam MH, Hawlitschek O, Bergsten J, Michat MC, Hendrich L, et al. Phylogenetic niche conservatism explains an inverse latitudinal diversity gradient in freshwater arthropods. *Sci Rep* 2016;6:26340.
78. Barberán A, Velazques HC, Jones S, Fierer N. Hiding in plain sight: mining bacterial species records for phenotypic trait information. *mSphere*. 2017;2:e00237–17.
79. Martiny AC, Treseder K, Pusch G. Phylogenetic conservatism of functional traits in microorganisms. *ISME J*. 2013;7:830–8.
80. Davis MB, Shaw RG. Range shifts and adaptive responses to Quaternary climate change. *Science*. 2001;292:673–9.

81. Cooper VS, Bennett AF, Lenski RE. Evolution of thermal dependence of growth rate of *Escherichia coli* populations during 20,000 generations in a constant environment. *Evolution*. 2001;55:889–96.
82. Bennett AF, Lenski RE, Mittler JE. Evolutionary adaptation to temperature. I. Fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution*. 1992;46:16–30.
83. Oliverio AM, Bradford MA, Fierer N. Identifying the microbial taxa that consistently respond to soil warming across time and space. *Glob Chang Biol*. 2017;23:2117–29.
84. Clayton L, Attig JW, Mickelson DM, Johnson MD, Syverson KM. Glaciation of Wisconsin. *Wisconsin Geological and Natural History Survey*. 2006;36:1–4