Leveraging aquatic-terrestrial interfaces to capture putative habitat generalists

Sarah C Richards[1,2,3,4], William L King[1,5], Jeremy L Sutherland[6], Terrence H Bell[1,3,4,7]

[1]Department of Plant Pathology and Environmental Microbiology, The Pennsylvania State University, University Park, PA, 16802, USA

[2]Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, PA, 16802, USA

[3]Intercollege Graduate Degree Program in Ecology, The Pennsylvania State University, University Park, PA, 16802, USA

[4]Department of International Agriculture and Development, The Pennsylvania State University, University Park, PA, 16802, USA

[5]School of Biological Sciences, University of Southampton, SO17 1BJ, UK

[6]Plant Sciences Department, The Pennsylvania State University, University Park, PA, 16802, USA

[7]Department of Physical & Environmental Sciences, University of Toronto Scarborough, Toronto, ON, M1C1A4, Canada

*Correspondence:

Postal address: Department of Physical & Environmental Sciences, University of Toronto Scarborough, Toronto, ON, M1C1A4, Canada

Telephone number: +1 647-985-2017

Email: terrence.bell@utoronto.ca

Key Words: Microbial colonization, habitat generalism, trait conservation, environmental interfaces, abiotic constraints, microbiome manipulation

Abstract

Habitat type is a strong determinant of microbial composition. Habitat interfaces, such as the boundary between aquatic and terrestrial systems, present unique combinations of abiotic factors for microorganisms to contend with. Aside from the spillover of certain harmful microorganisms from agricultural soils into water (e.g. fecal coliform bacteria), we know little about the extent of soil-water habitat switching across microbial taxa. In this study, we developed a proof-of-concept system to facilitate the capture of putatively generalist microorganisms that can colonize and persist in both soil and river water. We aimed to examine the phylogenetic breadth of putative habitat switchers and how this varies across different source environments. Microbial composition was primarily driven by recipient environment type, with the strongest phylogenetic signal seen at the order level for river water colonizers. We also identified more microorganisms colonizing river water when soil was collected from a habitat interface (i.e. soil at the side of an intermittently flooded river, compared to soil collected further from water sources), suggesting that environmental interfaces could be important reservoirs of microbial habitat generalists. Continued development of experimental systems that actively capture microorganisms that thrive in divergent habitats could serve as a powerful tool for identifying and assessing the ecological distribution of microbial generalists.

Introduction

Microorganisms are found in almost every environment that we can access on Earth, but individual taxa differ widely in their ubiquity. While some can grow across a broad environmental range, many are restricted to relatively narrow niche space (i.e. habitat generalists and specialists, respectively) (Barberán *et al.*, 2014). In some cases, microbial biogeography may be limited by environmental opportunity, as microorganisms are less likely to be well-adapted to conditions they have not been exposed to. For instance, a lineage that is endemic to isolated cave sediments may have never had the opportunity to colonize a human gut; whether it would develop this ability given the opportunity would then depend on its specific suite of traits. Similarly, there is evidence that the ability to biodegrade plastics among microorganisms is increasing as this opportunity becomes more widespread in the environment (Zrimec *et al.*, 2021). Due to their direct relevance to public health, certain disease-causing microorganisms are well-known for likely habitat switching between animals and humans (e.g. *E.coli* (Rwego *et al.*, 2008); SARS-CoV-2 -19 (Cohen, 2022)), soils and humans (e.g. *Bacillus anthracis* (Steffan et al., 2020)), or soil and water (e.g. fecal coliform bacteria (Pachepsky & Shelton, 2011)). However, we do not have a general understanding of the frequency or extent of habitat switching among non-pathogenic microorganisms, despite acknowledgement of the forces that shape microbial biogeography (Chu *et al.*, Larkin & Martiny, 2017).

Environmental heterogeneity and stability play an important role in shaping the potential habitat range of microorganisms (Hotaling *et al.*, 2019, Bell & Bell, 2021, Xu *et al.*, 2022). In spatially or temporally heterogeneous environments, some organisms may undergo niche expansion by developing broadly adapted phenotypes to cope with the range of conditions they

encounter. For instance, in frequently perturbed environments, organisms can rely on compensatory adaptations (e.g. copiotrophy; dormancy; phenotypic plasticity (Shoemaker & Lennon, 2018, Kurm *et al.*, 2019)) to cope with periods of abiotic or biotic stress. Alternatively, microorganisms may respond to environmental heterogeneity by confining themselves to a narrow portion of their fundamental niche and increasing their competitive advantage within that space (e.g. oligotrophy; antibiotic compound production), particularly when conditions remain stable (Barrett *et al.*, 2005, Hibbing *et al.*, 2010, Stone *et al.*, 2023). The costs associated with generalist and specialist lifestyles have long been debated, with some finding generalism to be linked with higher metabolic burden (Lipson *et al.*, 2008, Hall *et al.*, 2010) and others finding no clear evidence of fitness or functional trade-offs (Bennett & Lenski, 2007). Ultimately, this likely depends on the type of generalism and the traits required to allow microorganisms to function across a broad range of conditions.

There are several practical implications to understanding the factors that promote habitat switching. For instance, repurposing soils for agriculture will inherently reshape the microbial composition and function of those soils but may also have spillover effects on the microbiomes of adjacent land or water (Bell & Tylianakis, 2016, Bell & Bell, 2021). In addition, identifying taxonomic groups or traits that promote a broad survival range could benefit the development of microbial inoculants, which are notorious for their unpredictable survival across field conditions (Kaminsky *et al.*, 2019). In this study, we developed a proof-of-concept system for capturing putative habitat generalists. Such systems are of interest, because they could be used to 1) clearly identify microorganisms based on observed ecological traits and 2) test the functional capacity of different microbiome subsets post-capture. Here, we used microcosms to interface a non-sterile

source (water or soil) with a biotically-cleared recipient environment (soil or water) to capture microorganisms that can colonize and persist between divergent habitat types. We assayed two distinct types of soil: one from a riverbed that is intermittently flooded through the year, though not exposed to water at the time of sampling (representing a natural soil-water interface), and one from a nearby farm that is well above the water level of the river. We hypothesized that: 1) putative habitat switching would be restricted to specific lineages, as we expected traits that allow for habitat switching to be complex and therefore phylogenetically conserved, and 2) we would identify more putative habitat switchers from soil collected at a water-soil interface than soil collected further from the river, as microbiomes conditioned to fluctuating water and soil environments are likely to have undergone niche expansion into both. This proof-of-concept study demonstrates the potential for using active microbial capture systems to move beyond survey-based studies and collect microbiome subsets based on ecologically relevant traits.

Materials and Methods

Soil and water collection and processing

On 14 December 2020, river water and two distinct soil samples were collected from Cole Farm, a cultivated land site within the Susquehanna Shale Hills in Pennsylvania containing a sub-catchment of Shaver's Creek. This site was formerly part of the NSF-funded Critical Zone Observatory Network (Li *et al.*, 2018). River water was collected in bleach-sterilized 15 L buckets and the containers were pre-rinsed six times with river water downstream of the final collection site to minimize capture of disturbed sediments. Collected water samples were stored in a cold room at 4°C or frozen at -20°C for short- or long-term use, respectively, and a portion immediately filter-sterilized with 0.2 um cellulose nitrate membranes (Majaneva *et al.*, 2018). Water quality

110 before and after sterilizing was analyzed by the Agricultural Analytical Services Laboratory at
111 Pennsylvania State University (See Supplementary Table 1).
112
113 Soils were collected from two locations within the study site to obtain microbiomes
114 historically conditioned to distinct environmental pressures: 1) farm soil: directly from an
115 agriculturally-managed field not susceptible to river flooding (40°38'07.6"N 77°56'38.0"W), and
116 2) riverbed soil: from un-managed river-adjacent sediments (40°38'00.6"N 77°56'29.4"W), where
117 intermittent flooding was apparent from soil compaction and vegetation patterns. Each soil was
118 homogenized and divided into batches of "natural" (non-sterile) or biotically-cleared soil (further
119 referred to as sterilized soil). Natural field soils from each collection site were immediately frozen
120 until experimental setup to preserve their microbiomes. To standardize comparisons of microbial
121 colonization between two different soil types, soils from each location intended for sterilization
122 were sieved through a 12 mm wire mesh to minimize rocks and soil aggregates that could generate
123 spurious noise in colonization between replicates. They were then autoclaved three individual
124 times (3 x 60 min @ 121ºC) with 24 h between each cycle (Howard *et al.*, 2017, King *et al.*, 2022).
125 While all soil sterilization methods will alter soil properties to some extent (McNamara *et al.*,
126 2003) we have found autoclaving to be most effective in removing microorganisms (Trexler &
127 Bell, 2019, Kaminsky *et al.*, 2021, Yates *et al.*, 2022). While no sterilized version of water or soil
128 can fully replicate the original environment, our goal was to assess microbial movement into
129 complex soil and water analogues in the absence of biotic pressures. Soils before and after
130 sterilizing were analyzed for nutrient concentrations and particle size by the Agricultural
131 Analytical Services Laboratory at Pennsylvania State University (Supplementary Table 2 & 3).
132

Microcosm development

Examples of fully constructed microcosms are provided in Fig. 1, along with a visual representation of the full experimental design. To evaluate microbial movement between divergent habitats, we designed our microcosms to interface soil and water using 565 ml Microbox containers with filter membrane lids that allow sterile gas exchange (Sac O_2 company, Deinze, Belgium, catalogue: O118/80+OD118). Microcosms were filled partially with 450 ml of river water, leaving an air buffer between the meniscus and lid. Soil was aliquoted (200 ml) and encased in sterile nylon stockings, to allow microbial movement across the barrier while maintaining a soil-water interface within the microcosm. Nylon stockings were sterilized by immersing in absolute ethanol for 20 min followed by a 20 min UV-sterilization and drying in a biological safety cabinet prior to use. Control treatments included soil-only and water-only microcosms, which were constructed as either a unit containing a nylon stocking filled with soil (200 ml) or containing 450 ml of river water, respectively. All soil-only microcosms were normalized to 70 % water holding capacity prior to setup, estimated by measuring the volume of water needed to saturate each soil type and back-calculating the volume needed to perform the adjustment.

Microcosm experiment setup - Initial colonization phase

One of our goals was to assess this microcosm design as a means to reproducibly capture microorganisms that can colonize and persist between divergent habitats. Experimental treatments are described throughout this paper as a ***source*** microbiome colonizing a sterile ***recipient*** environment. Treatments consisted of: 1) source farm soil to recipient river water, 2) source riverbed soil to recipient river water, 3) source river water to recipient farm soil, and 4) source river water to recipient riverbed soil. Our baseline controls include both a natural and sterile

version of soil-only and water-only treatments, which are defined further in the next section. All treatments and controls were replicated six times yielding 60 microcosms (and a total of 120 microcosms in the full experiment). Microcosms were randomized on a bench and incubated at a constant room temperature (24 °C) until harvest. The initial colonization phase of our experiment includes the day of microcosm setup (19 December 2020) to the end of the 31-day incubation period (19 January 2021). Our chosen timeframe has previously been shown to be sufficient for observing microbial recolonization of sterile soil in prior experiments (King *et al.*, 2022).

Microcosm experiment setup - Secondary transfer phase

Because there is a high degree of microbial dormancy and varying abundances of relic DNA across habitats (e.g. up to 99% in soils; (Lebaron *et al.*, 2001, Carini *et al.*, 2016)), we sought to maximize our capture of the active microbial pool by initiating a "secondary transfer phase". A portion of the colonized recipient at the end of the initial colonization phase was transplanted into a secondary sterile recipient of the same habitat type. For example, a recolonized recipient water in the initial phase would be transferred into another sterile recipient water. The additional transfer would allow putative habitat switchers to recolonize in an environment that is, at least initially, free of dormant cells and relic DNA. Water inoculum at the end of the initial colonization phase was transplanted to recipient sterile river water in a 1:10 dilution (45 ml in 405 ml), whereas a 5% direct transfer of colonized soil (5 g) was transplanted into sterile recipient soil (95 g), as outlined in Trexler & Bell (2019). Constructed microcosms were again randomized on a bench and allowed to incubate for 21 d at room temperature, starting on day of transplant into the secondary sterile recipient (19 January 2021) and ending the day of sample harvest (9 February 2021). All setup and harvest procedures in both initial and secondary transfer phases were completed within 24 h.

Our baseline controls, which were incubated in parallel to experimental treatments, include both a natural and sterilized version of soil-only and water-only treatments. Natural baseline controls serve as a benchmark for comparing colonization dynamics in recipient environments across habitats, while sterile baseline controls were used to identify any regrowth of sterilization-resistant taxa through the timeline of the experiment (sterile baseline control data is presented exclusively in Supplementary Fig. 2). We define a natural baseline control as the regrowth of a natural source environment into its sterilized habitat analogue throughout initial colonization and secondary transfer phases of the experiment (for instance, source river water > recipient river water (Initial colonization phase) > recipient river water (Secondary transfer phase)). This allows us to compare the regrowth dynamics of a natural source microbiome in its recipient analogue, with that of recipients from a divergent habitat source. By also including sterile baseline controls, we can assess the regrowth of bacteria that were not removed during sterilization, though sequencing depth in half of the sterile soil baseline controls was poor and the remainder were compositionally distinct from experimental samples (Supplementary Table 5 and Supplementary Fig. 2A & 2B). For sterile water baseline controls, the dominant genera found after regrowth were aquatic microorganisms that have previously been observed to pass through 0.22 um filtration systems (Supplementary Table 5 and Supplementary Fig. 2C & 2D). A total of 60 microcosms were constructed at the secondary transfer phase of this experiment. Soil and water used to construct secondary transfer microcosms were sterilized and handled using the same methodology outlined in preparation for the initial colonization phase, and again analyzed for nutrient concentrations and particle size post-sterilization by the Agricultural Analytical Services Laboratory at Pennsylvania State University (Supplementary Tables 1 & 2).

DNA extraction and amplicon sequencing

All microcosm samples were collected within a UV-sterilized biological safety cabinet. At the end of the initial colonization phase, after sacrificing a portion of water or soil to seed secondary transfer phase microcosms, DNA was extracted using the NucleoSpin 96 Soil DNA extraction kit (Macherey-Nagel; catalogue: 740787.2) directly from soil as per the manufacturer's instructions and from water with the following modifications. Water samples at the end of secondary transfer phase were collected by passing the remaining 405 ml of microcosm water through 0.2 um cellulose nitrate filter membranes, which were then cut in half using sterile scissors to fit into Type A bead tubes. Many water samples from water-soil microcosms at the end of the initial colonization phase were un-filterable due to clogging, so to minimize extraction bias, DNA was extracted from a pellet via centrifugation for all samples at this phase. These methods are detailed (Supplementary Table 4) and represent data that is presented exclusively in supplementary material. Bacterial composition was characterized with amplicon sequencing of partial 16S rRNA gene fragments (515F and 806R) (Apprill *et al.*, 2015, Parada *et al.*, 2016). Raw data files in FASTQ format were deposited in the NCBI sequence read archive under BioProject number PRJNA798746.

PCR

Bacterial composition was characterized with amplicon sequencing of partial 16S rRNA gene fragments (515F and 806R) (Apprill *et al.*, 2015, Parada *et al.*, 2016). The PCR mix for reactions was as follows: 12 µL of 5Prime HotMasterMix, 1.5 µL of each primer (10 µM), 1.5 µL template DNA and 13.5 µL molecular grade water for a final PCR volume of 30 µL. Bacterial 16S

rRNA gene PCR cycling conditions were as follows: 3 min at 94°C, 25 cycles of: 45 sec at 94°C, 60 sec at 50°C, and 90 sec at 72°C, and a final elongation step of 10 min at 72°C. The resulting amplicons were cleaned using Mag-Bind TotalPure NGS magnetic beads (Omega Bio-tek; catalogue: M1378-01). Illumina indexes were added to the cleaned amplicons with the following PCR ingredients: 12.5 µL of Platinum II Hot-Start PCR Master Mix, 2.5 µL of each index (10 µM) and 2.5 µL of sterile water for a final volume of 20 µL. The indexing PCR cycling conditions were as follows: 1 min at 98°C, 8 cycles of: 15 sec at 98°C, 30 sec at 55°C, and 20 sec at 72°C, and a final elongation step of 5 min at 72°C. Indexed amplicons were then normalized using the SequalPrep normalization plate kit (ThermoFisher; catalogue: A1051001), pooled, concentrated with a Savant SpeedVac (Thermo Scientific, Waltham, MA, USA) at 50 °C for 3 h, and purified with a gel extraction using the PureLink Quick Gel Extraction Kit (ThermoFisher; catalogue: K210012). The pooled library was sequenced on the Illumina V3 600 cycle MiSeq sequencing platform (2 x 250bp yielding ~45,000 pairs of reads per sample) at the Pennsylvania State University Genomics Core Facility (Huck Institutes for the Life Sciences).

Sequence analysis

Raw demultiplexed 16S rRNA gene data were processed using the Quantitative Insights into Microbial Ecology (QIIME 2 version 2021.4) pipeline (Bolyen *et al.*, 2019). Briefly, paired-ended 16S rRNA gene sequences were imported and trimmed, and denoised using DADA2, which also removes chimeric sequences (Callahan *et al.*, 2016). The classify-sklearn qiime feature classifier was used to assign taxonomy against the Silva v.138 database (Quast *et al.*, 2013) at the single nucleotide threshold (amplicon sequencing variants; ASVs). The dataset was further cleaned by removing sequences identified as chloroplasts or mitochondria, and by removing ASVs with

248 less than 48 (0.001 %) sequences for the 16S rRNA gene dataset (Bokulich *et al.*, 2013). The
249 cleaned 16S rRNA gene data were then rarefied at 3,670 sequences per sample.
250
251 *Statistical analysis*

252 Processed sequencing data were imported into the R statistical environment v.4.2.1 (R Core
253 Team, 2012) and used to create a *Phyloseq* object (McMurdie & Holmes, 2013). Chao1 species
254 richness and Shannon's index (as a measure of richness and evenness) were used to estimate alpha
255 diversity in recipient environments colonized by divergent sources. Recipient river water
256 colonized by either source riverbed or farm soil, and the two recipient soil types colonized by river
257 water were tested for statistical differences with a Kruskal-Wallis test from the *stats* package,
258 followed by a Dunnett's post hoc test. To compare bacterial composition between microcosm
259 treatments, a Principal Coordinates Analysis (PCoA) was performed based on Bray-Curtis
260 dissimilarities using the "ordinate()" function in the *Phyloseq* package. Permutational multivariate
261 analysis of variance (PERMANOVA) using the "adonis2" function from the *vegan* package
262 (Oksanen *et al.*, 2013) was used to determine the impact of recipient environment type
263 (explanatory variables) on bacterial composition of the colonizer communities (response variable).
264 P values were adjusted using the Bonferroni method (Jafari & Ansari-Pour, 2019) correcting for 6
265 comparisons. To identify taxonomic colonization trends across treatments, the total number of
266 unique ASVs in each recipient environment was quantified and summarized in a 4D Venn diagram.
267 Further, the relative abundance of ASVs at least 80% prevalent across all replicates of
268 experimental water and soil recipients was plotted to determine and visualize the taxonomic
269 classification of core colonizers.
270

271 **Results and Discussion**

Understanding the factors that shape the breadth of microbial habitat colonization has been a paramount focus for decades from global (Lozupone & Knight, 2007, Nemergut *et al.*, 2011) to local scales (Wu *et al.*, 2019), including how these patterns interface with ecological theory, for example dispersal limitation and functional tradeoffs. It is therefore of interest to improve our ability to predict which taxa are likely to be found in a particular niche, their potential functional contributions under different environmental conditions, and the likelihood of microbial spillover between divergent environments. In this proof-of-concept study, we develop a new system to perform a reciprocal transplant between two natural soil types and river water, to assess community-level patterns in microbial habitat switching. To comprehensively assess the performance of this system for the first time, we focused on higher within-treatment replication and testing of microbial composition at multiple experimental stages, rather than increasing the number of source and recipient environments, which would be a target for future studies. We define putative habitat-switching as a certain taxon being detected in a recipient environment (colonized by a divergent source environment type) at the end of the secondary transfer phase incubation period. Previous work has attempted to categorize microorganisms as specialists or generalists based on the number of habitats they occupy in publicly available sequence data (Monard *et al.*, 2016, Delgado-Baquerizo *et al.*, 2018); however, widespread variation in genomic surveying methodologies and a lack of controlled environmental manipulations have made it difficult to assess under which conditions microbial habitat switching between source and recipient environments is phylogenetically restricted (exclusive to certain microbial lineages) or a broadly acquired ability.

Diversity of putative habitat generalists may be greater along natural habitat interfaces

295 To explore the impact of soil origin on bacterial colonizer composition in a receiving
296 aquatic environment, we contrasted alpha diversity in the river water recipients colonized by two
297 soils, which were either in direct contact (riverbed soil) or farther away (farm soil) from a natural
298 river. We found higher species richness and evenness in water recipients colonized by a riverbed
299 soil (Shannon diversity; $p = 0.003$, Chao1; $p = 0.002$; Fig. 2A), suggesting the capacity for
300 microbial habitat switching could be greater when microorganisms exist at a contextual habitat
301 interface. Along a riverbed, the resident soil microbiome is forced to interact with river water, as
302 it introduces frequent but irregular fluxes in environmental conditions such as aqueous submersion,
303 oxygen gradients, dynamic or transient biodiversity deposits, nutrient economy, etc. Over time,
304 frequent environmental fluctuations may select for more generalist lifestyles due to heterogenous
305 conditions. In support of our data, Evans & Wallenstein (2012) found that microbiomes exposed
306 to a decade of experimentally-intensified precipitation in a grassland soil led to improved moisture
307 adaptation in laboratory microcosms and Porter & Rice (2013) identified differences in nickel
308 tolerance in *Rhizobium* species along a serpentine soil gradient. Improved community fitness or
309 plasticity is also observed in numerous adaptive laboratory evolution experiments under varying
310 environmental conditions (Dragosits & Mattanovich, 2013). In contrast, our chosen farm soil was
311 atop a hill and never subject to interactions with the river water. The historical differences in
312 environmental conditions between our soils could be a factor driving the lower diversity of
313 organisms able to colonize the recipient water when the source was a farm soil.
314
315 A central goal of this experiment was to assess the extent to which our microcosm approach
316 can be used to capture putative habitat switchers across divergent environments, so we limited our
317 diversity of source material (soil & water) to improve replication potential and validate our

318 microcosm system. As only one type of source river water was assayed, we were not able to
319 compare potentially similar patterns in the opposite direction of colonization (i.e., distinct source
320 water types on a common recipient soil). However, we examined alpha diversity in our divergent
321 soil recipients when colonized by a common source river water and found no significant difference
322 in species diversity or richness (Shannon diversity; $p = 0.775$, Chao1; $p = 0.438$; Fig. 2B). River
323 water microbiomes are likely to be more transient than a localized riverbed soil community
324 (Teachey *et al.*, 2019) and therefore may not experience the same degree of environmental
325 conditioning to be able to switch efficiently between the water and various soils they encounter in
326 a stream. As an alternative hypothesis, there is some evidence of greater niche space and
327 phylogenetic diversity being found in sediments compared to soils, thus, the pool of organisms
328 with the potential to colonize water may have been larger in our riverbed soil, leading to higher
329 diversity in colonized water compared to when farm soil was the source (Lozupone & Knight,
330 2007, Thompson *et al.*, 2017). However, we observed relatively similar starting species richness
331 and significant compositional differences between our natural source soil environments,
332 suggesting source diversity was not a confounding factor (Supplementary Fig. 3).
333
334 ***Recipient environment type structures microbial colonizer communities***
335 We sought to identify compositional overlap between endpoint communities, to examine
336 the strength of environmental filtering when source microbiomes were introduced into recipient
337 environments (Fig. 3; Table 1); (see Supplementary Fig. 1 for comparisons including
338 "intermediate" communities at the end of initial colonization phase. Many microbial traits are
339 conserved at different taxonomic levels based on their complexity (Martiny *et al.*, 2015). For
340 example, pH tolerance, which can require cell membrane modification, and photosynthesis, which

involves numerous genes that are deeply conserved, while others can develop rapidly at shallow taxonomic levels, such as viral resistance (Martiny *et al.*, 2013, Martiny *et al.*, 2015). Therefore, we explored the phylogenetic levels at which recipient communities of water or soil were statistically indistinguishable from each other, to explore the potential complexity of traits that allow habitat switching. For source river water colonizing recipient farm or riverbed soils, we identified significant differences between the community composition in recipient soil types up to the Phylum level, which was the broadest taxonomic level that we assessed (Fig. 3A; Table 1). These data suggest that edaphic properties are strong environmental filters for microbial composition (soil property analyses in Supplementary Table 2) as identified in numerous previous studies (Lauber *et al.*, 2009, Leff *et al.*, 2015, Delgado-Baquerizo *et al.*, 2018), and that there does not appear to be trait conservation. This is further supported by the patterns exhibited by the natural baseline controls, which serve as a reference point for what the community would look like under a self-colonization scenario. The deviation from the natural baseline indicates that the introduced microbes from divergent source environments are responding to and interacting with the novel environment (i.e., environmental filtering). For our two source soils colonizing recipient river water, we identified significant compositional differences at finer taxonomic resolutions (ASV-Family), but not at the Order-Phylum levels (Table 1). These data suggest that the traits required for water colonization, from two disparate soil sources, were conserved roughly at the Order level, with the top three most abundant being *Burkholderiales*, *Rhizobiales*, and *Micrococcales*. Despite the limitations of using phylogeny to identify trait-based conservation along environmental axes (such as microbial habitat breadth), our findings show that microbial transfer between our aquatic-terrestrial environments is not evolutionarily trivial, as it was restricted to specific taxonomic groups. The ecological attributes that enable microbial “habitat switching” (Supplementary Table

364 6) are not well-characterized, likely vary across types of environmental transitions, and could
365 consist of multiple traits rather than one alone, therefore, analyses at the community-level using
366 controlled manipulations and differentiating between active and dormant microbial pools is needed
367 to improve our modeling and prediction capability of where we might expect to find microbial
368 taxa across spatial and temporal scales.
369
370 ***Core microbes were shared across environments***
371 We surveyed taxa that were consistently found across both soil and water recipient
372 environments at the end of the secondary transfer phase as an indicator for microbial habitat
373 switching ability (i.e. growth or persistence in a novel environment). The most dominant taxa
374 across all soil or water recipients being *Actinobacter* or an unclassified genus in the
375 *Comamonadaceae* family respectively, (see Supplementary Figure 4 for top ten most abundant
376 taxa plots). Relative to the total richness of ASVs found in each recipient environment, a mere 18
377 (< 1 %) of them were shared (Fig. 4A). Of those, only two were at least 80 % prevalent in replicates
378 of each recipient, an unclassified bacterial species in the *Micrococcaceae* family
379 (*Actinobacteriota*) and an unclassified bacterial species assigned to the *Devosia* genus
380 (*Proteobacteria*). The *Devosia* are a motile genus of bacteria widely reported in soil habitats. The
381 *Devosia* are known for their ability to break down an array of hydrocarbons and have a high degree
382 of genetic promiscuity via horizontal gene transfer (Rivas *et al.*, 2002, Talwar *et al.*, 2020), which
383 would support the potential proclivity toward a generalist lifestyle. However, it represented an
384 extremely low percent relative abundance in each sample (Fig. 4B) and may be bacterial
385 persistence rather than adaptation and active growth, despite our efforts to differentiate between
386 these fractions by initiating a secondary transfer phase. While the exact mechanisms that govern
387 bacterial genome size are not comprehensively understood, larger or smaller genome sizes have

been associated with generalist or specialist lifestyles respectively (Gweon *et al.*, 2017, Sriswasdi *et al.*, 2017, Bell & Bell, 2021). Bacterial and Archaeal genomes can range from 100 kbp to 16 Mbp (Garcia *et al.*, 2014, Rodríguez-Gijón *et al.*, 2022), and genome size has also been found to vary according to habitat type (Beier *et al.*, 2022). While categorizing generalist or specialist lifestyles is not as straightforward as establishing a size threshold, when BLASTing the representative sequences for our two universal ASVs they had estimated genome sizes of 4.5 Mbp (*Actinobacteriota;* CP066362.1; E-value of $3e^{-127}$) and 5.8 Mbp (*Devosia;* MT239494.1; E-value of $3e^{-127}$), aligning with those of other microbial genera broadly linked with generalist life histories, suggesting functional potential for our core microbes to adapt to environmental novelty. For instance, Barberán *et al.,* (2014) found that the ubiquity and habitat breadth of soil dwelling bacteria was positively correlated with genome size.

Conclusions

A growing body of work attempts to map microbiome composition to ecosystem functions such as nitrogen fixation or decomposition rates, which can be particularly useful when assessing the health and productivity of managed systems (e.g. agriculture) or predicting ecosystem response under global change conditions (Wang & Xue, 2021). But while the relatively low cost of high-throughput sequencing has made it a common tool in estimating microbial presence in natural environments, it remains challenging to predict ecosystem function from this information alone (Morris *et al.*, 2020, Wemheuer *et al.*, 2020).

While our study focuses mainly on microbial presence and observable habitat switching, it lays the foundation for this microcosm approach to be coupled with additional techniques, such as enzyme assays, stable isotope probing (SIP), or multi-omics to enhance our understanding of active

412 microbial communities. These techniques can be specifically tailored to assess the functional
413 attributes of active microbial communities, but they often face challenges due to signals from relic
414 DNA or dormant cells. Our method of active capture has the potential to limit some of these
415 challenges by enriching for active colonizing microbes, and in combination with other techniques,
416 could provide additional insights between microbiome composition, microbial activity, and
417 ecosystem functionality. Further, it could provide a means to cross-validate the results generated
418 by other methods intended to distinguish between active, dormant, and dead cells. Bioorthogonal
419 noncanonical amino acid tagging (BONCAT), for instance, is a cutting edge method that uses
420 translational incorporation of a noncanonical amino acid probe into cellular proteins, followed by
421 azide-alkyne click-chemistry to detect metabolically active bacteria (Couradeau *et al.*, 2019).
422 However, it lacks a reliable benchmark for quantifying microbiome constituents that are not
423 captured by the method (e.g., microorganisms that might be metabolically active but not
424 undergoing protein synthesis or are fundamentally incompatible with the chosen amino acid
425 probe).
426
427 Our method provides a potentially scaleable approach to physically capture subsets of a
428 microbiome with an ecological trait of interest, allowing for active microbial pools to be cross
429 compared with other existing methods. Together, this research and future studies can further our
430 understanding of the likelihood that soil microbial composition in one environment could impact
431 function and composition in connected systems. Future studies may also consider including
432 multiple reciprocal inoculations to apply stronger selection for taxa that can consistently switch
433 between and actively grow in divergent source and recipient environments, while further excluding
434 those that persist alone.

Acknowledgements

We thank Brandon Forsythe, who coordinated field sampling efforts with Dr. Herbert Cole, the landowner of this study site. Logistical support was provided by the NSF-supported Susquehanna Shale Hills Critical Zone Observatory. Fig. 1 created with BioRender.com.

Funding

This work was supported by the United States Department of Agriculture National Institute of Food and Agriculture (grant no. 2020-67013-30865), under Hatch Appropriations PEN04651 and Accession 1016233.

Author Contributions

SCR and THB designed the experiment, analyzed the data, and wrote the manuscript, incorporating edits and feedback from contributing authors. SCR and WLK set up microcosms, and prepared and submitted samples for sequencing. SCR and JLS generated the microbial prevalence analyses.

Conflict of Interest

None declared.

References

Apprill A, McNally SP, Parsons RJ & Weber L (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**: 129-137.

459 Barberán A, Ramirez KS, Leff JW, Bradford MA, Wall DH & Fierer N (2014) Why are some
460 microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecology*
461 *Letters* **17**: 794-802.

462 Barrett Rowan DH, MacLean RC & Bell G (2005) Experimental Evolution of Pseudomonas
463 fluorescens in Simple and Complex Environments. *The American Naturalist* **166**: 470-480.

464 Beier S, Werner J, Bouvier T, Mouquet N & Violle C (2022) Trait-trait relationships and tradeoffs
465 vary with genome size in prokaryotes. *Frontiers in Microbiology* **13**.

466 Bell T & Tylianakis JM (2016) Microbes in the Anthropocene: spillover of agriculturally selected
467 bacteria and their impact on natural ecosystems. *Proc Biol Sci* **283**.

468 Bell TH & Bell T (2021) Many roads to bacterial generalism. *FEMS Microbiology Ecology* **97**:
469 fiaa240.

470 Bennett AF & Lenski RE (2007) An experimental test of evolutionary trade-offs during
471 temperature adaptation. *Proc Natl Acad Sci U S A* **104 Suppl 1**: 8649-8654.

472 Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA & Caporaso JG
473 (2013) Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing.
474 *Nature Methods* **10**: 57-59.

475 Bolyen E & Rideout JR & Dillon MR*, et al.* (2019) Reproducible, interactive, scalable and
476 extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**: 852-857.

477 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA & Holmes SP (2016) DADA2:
478 High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581-583.

479 Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS & Fierer N (2016) Relic DNA is
480 abundant in soil and obscures estimates of soil microbial diversity. *Nature Microbiology* **2**: 16242.

481 Chu H, Gao G-F, Ma Y, Fan K, Delgado-Baquerizo M & Shade A Soil Microbial Biogeography
482 in a Changing World: Recent Advances and Future Perspectives. *mSystems* **5**: e00803-00819.

483 Cohen J (2022) Do three new studies add up to proof of COVID-19's origin in a Wuhan animal
484 market? p.^pp.

485 Couradeau E, Sasse J, Goudeau D, Nath N, Hazen TC, Bowen BP, Chakraborty R, Malmstrom
486 RR & Northen TR (2019) Probing the active fraction of soil microbiomes using BONCAT-FACS.
487 *Nature Communications* **10**: 2770.

488 Delgado-Baquerizo M, Oliverio Angela M, Brewer Tess E, Benavent-González A, Eldridge David
489 J, Bardgett Richard D, Maestre Fernando T, Singh Brajesh K & Fierer N (2018) A global atlas of
490 the dominant bacteria found in soil. *Science* **359**: 320-325.

491 Delgado-Baquerizo M, Oliverio AM, Brewer TE, Benavent-González A, Eldridge DJ, Bardgett
492 RD, Maestre FT, Singh BK & Fierer N (2018) A global atlas of the dominant bacteria found in
493 soil. *Science* **359**: 320-325.

494 Dragosits M & Mattanovich D (2013) Adaptive laboratory evolution – principles and applications
495 for biotechnology. *Microbial Cell Factories* **12**: 64.

496 Evans SE & Wallenstein MD (2012) Soil microbial community response to drying and rewetting
497 stress: does historical precipitation regime matter? *Biogeochemistry* **109**: 101-116.

498 Garcia R, Gemperlein K & Müller R (2014) Minicystis rosea gen. nov., sp. nov., a polyunsaturated
499 fatty acid-rich and steroid-producing soil myxobacterium. *International journal of systematic and*
500 *evolutionary microbiology* **64**: 3733-3742.

501 Gweon HS, Bailey MJ & Read DS (2017) Assessment of the bimodality in the distribution of
502 bacterial genome sizes. *The ISME Journal* **11**: 821-824.

503 Hall EK, Singer GA, Kainz MJ & Lennon JT (2010) Evidence for a temperature acclimation
504 mechanism in bacteria: an empirical test of a membrane-mediated trade-off. *Functional Ecology*
505 **24**: 898-908.

506 Hibbing ME, Fuqua C, Parsek MR & Peterson SB (2010) Bacterial competition: surviving and
507 thriving in the microbial jungle. *Nat Rev Microbiol* **8**: 15-25.

508 Hotaling S, Foley ME, Zeglin LH, Finn DS, Tronstad LM, Giersch JJ, Muhlfeld CC & Weisrock
509 DW (2019) Microbial assemblages reflect environmental heterogeneity in alpine streams. *Global*
510 *Change Biology* **25**: 2576-2590.

511 Howard MM, Bell TH & Kao-Kniffin J (2017) Soil microbiome transfer method affects
512 microbiome composition, including dominant microorganisms, in a novel environment. *FEMS*
513 *Microbiology Letters* **364**.

514 Jafari M & Ansari-Pour N (2019) Why, When and How to Adjust Your P Values? *Cell J* **20**: 604-
515 607.

516 Kaminsky LM, Esker PD & Bell TH (2021) Abiotic conditions outweigh microbial origin during
517 bacterial assembly in soils. *Environ Microbiol* **23**: 358-371.

518 Kaminsky LM, Trexler RV, Malik RJ, Hockett KL & Bell TH (2019) The Inherent Conflicts in
519 Developing Soil Microbial Inoculants. *Trends Biotechnol* **37**: 140-151.

520 King WL, Kaminsky LM, Richards SC, Bradley BA, Kaye JP & Bell TH (2022) Farm-scale
521 differentiation of active microbial colonizers. *ISME Communications* **2**: 39.

522 Kurm V, Geisen S & Gera Hol WH (2019) A low proportion of rare bacterial taxa responds to
523 abiotic changes compared with dominant taxa. *Environmental Microbiology* **21**: 750-758.

524 Larkin AA & Martiny AC (2017) Microdiversity shapes the traits, niche space, and biogeography
525 of microbial taxa. *Environmental Microbiology Reports* **9**: 55-70.

526 Lauber CL, Hamady M, Knight R & Fierer N (2009) Pyrosequencing-based assessment of soil pH
527 as a predictor of soil bacterial community structure at the continental scale. *Appl Environ*
528 *Microbiol* **75**: 5111-5120.

529 Lebaron P, Servais P, Agogué H, Courties C & Joux F (2001) Does the High Nucleic Acid Content
530 of Individual Bacterial Cells Allow Us To Discriminate between Active Cells and Inactive Cells
531 in Aquatic Systems? *Applied and Environmental Microbiology* **67**: 1775-1782.

532 Leff JW, Jones SE, Prober SM*, et al.* (2015) Consistent responses of soil microbial communities
533 to elevated nutrient inputs in grasslands across the globe. *Proc Natl Acad Sci U S A* **112**: 10967-
534 10972.

535 Li L, DiBiase RA, Del Vecchio J*, et al.* (2018) The Effect of Lithology and Agriculture at the
536 Susquehanna Shale Hills Critical Zone Observatory. *Vadose Zone Journal* **17**: 180063.

537 Lipson D, Monson R, Schmidt S & Weintraub M (2008) The Trade-Off Between Growth Rate and
538 Yield in Microbial Communities and the Consequences for Under-Snow Soil Respiration in a High
539 Elevation Coniferous Forest. *Biogeochemistry* **95**: 23-35.

540 Lozupone CA & Knight R (2007) Global patterns in bacterial diversity. *Proceedings of the*
541 *National Academy of Sciences* **104**: 11436-11440.

542 Majaneva M, Diserud OH, Eagle SHC, Boström E, Hajibabaei M & Ekrem T (2018)
543 Environmental DNA filtration techniques affect recovered biodiversity. *Scientific Reports* **8**: 4682.

544 Martiny AC, Treseder K & Pusch G (2013) Phylogenetic conservatism of functional traits in
545 microorganisms. *The ISME Journal* **7**: 830-838.

546 Martiny JB, Jones SE, Lennon JT & Martiny AC (2015) Microbiomes in light of traits: A
547 phylogenetic perspective. *Science* **350**: aac9323.

548 McMurdie PJ & Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis
549 and Graphics of Microbiome Census Data. *PLOS ONE* **8**: e61217.

550 McNamara NP, Black HIJ, Beresford NA & Parekh NR (2003) Effects of acute gamma irradiation
551 on chemical, physical and biological properties of soils. *Applied Soil Ecology* **24**: 117-132.

552 Mizoguchi H, Mori H & Fujio T (2007) Escherichia coli minimum genome factory. *Biotechnology*
553 *and Applied Biochemistry* **46**: 157-167.

554 Monard C, Gantner S, Bertilsson S, Hallin S & Stenlid J (2016) Habitat generalists and specialists
555 in microbial communities across a terrestrial-freshwater gradient. *Scientific Reports* **6**: 37719.

556 Morris A, Meyer K & Bohannan B (2020) Linking microbial communities to ecosystem functions:
557 what we can learn from genotype–phenotype mapping in organisms. *Philosophical Transactions*
558 *of the Royal Society B: Biological Sciences* **375**: 20190244.

559 Nemergut DR, Costello EK, Hamady M*, et al.* (2011) Global patterns in the biogeography of
560 bacterial taxa. *Environ Microbiol* **13**: 135-144.

561 Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P & O'Hara RB (2013) vegan: Community
562 Ecology, Package. R Package Version 2.2–1.

563 Pachepsky YA & Shelton DR (2011) Escherichia Coli and Fecal Coliforms in Freshwater and
564 Estuarine Sediments. *Critical Reviews in Environmental Science and Technology* **41**: 1067-1110.

565 Parada AE, Needham DM & Fuhrman JA (2016) Every base matters: assessing small subunit
566 rRNA primers for marine microbiomes with mock communities, time series and global field
567 samples. *Environ Microbiol* **18**: 1403-1414.

568 Porter SS & Rice KJ (2013) TRADE-OFFS, SPATIAL HETEROGENEITY, AND THE
569 MAINTENANCE OF MICROBIAL DIVERSITY. *Evolution* **67**: 599-608.

570 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J & Glöckner FO (2013)
571 The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.
572 *Nucleic acids research* **41**: D590-D596.

573 R Core Team (2012) R: A language and environment for statistical computing.

574 Rivas R, Velázquez E, Willems A, Vizcaíno N, Subba-Rao NS, Mateos PF, Gillis M, Dazzo FB
575 & Martínez-Molina E (2002) A new species of Devosia that forms a unique nitrogen-fixing root-
576 nodule symbiosis with the aquatic legume Neptunia natans (L.f.) druce. *Appl Environ Microbiol*
577 **68**: 5217-5222.

578 Rodríguez-Gijón A, Nuy JK, Mehrshad M, Buck M, Schulz F, Woyke T & Garcia SL (2022) A
579 Genomic Perspective Across Earth's Microbiomes Reveals That Genome Size in Archaea and
580 Bacteria Is Linked to Ecosystem Type and Trophic Strategy. *Frontiers in Microbiology* **12**.

581 Rwego IB, Gillespie TR, Isabirye-Basuta G & Goldberg TL (2008) High rates of Escherichia coli
582 transmission between livestock and humans in rural Uganda. *J Clin Microbiol* **46**: 3187-3191.

583 Shoemaker WR & Lennon JT (2018) Evolution with a seed bank: The population genetic
584 consequences of microbial dormancy. *Evol Appl* **11**: 60-75.

585 Sriswasdi S, Yang C-c & Iwasaki W (2017) Generalist species drive microbial dispersion and
586 evolution. *Nature Communications* **8**: 1162.

587 Steffan JJ, Derby JA & Brevik EC (2020) Soil pathogens that may potentially cause pandemics,
588 including severe acute respiratory syndrome (SARS) coronaviruses. *Curr Opin Environ Sci Health*
589 **17**: 35-40.

590 Stone BWG, Dijkstra P, Finley BK*, et al.* (2023) Life history strategies among soil bacteria—
591 dichotomy for few, continuum for many. *The ISME Journal*.

592 Talwar C, Nagar S, Kumar R, Scaria J, Lal R & Negi RK (2020) Defining the Environmental
593 Adaptations of Genus Devosia: Insights into its Expansive Short Peptide Transport System and
594 Positively Selected Genes. *Scientific Reports* **10**: 1151.

595 Teachey ME, McDonald JM & Ottesen EA (2019) Rapid and Stable Microbial Community
596 Assembly in the Headwaters of a Third-Order Stream. *Appl Environ Microbiol* **85**.

597 Thompson LR & Sanders JG & McDonald D*, et al.* (2017) A communal catalogue reveals Earth's
598 multiscale microbial diversity. *Nature* **551**: 457-463.

599 Trexler R & Bell T (2019) Testing sustained soil-to-soil contact as an approach for limiting the
600 abiotic influence of source soils during experimental microbiome transfer. *FEMS microbiology*
601 *letters* **366**.

602 Wang Y & Xue K (2021) Linkage between microbial shift and ecosystem functionality. *Global*
603 *Change Biology* **27**: 3197-3199.

604 Wemheuer F, Taylor JA, Daniel R, Johnston E, Meinicke P, Thomas T & Wemheuer B (2020)
605 Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on
606 16S rRNA gene sequences. *Environmental Microbiome* **15**: 11.

607 Wu L & Ning D & Zhang B*, et al.* (2019) Global diversity and biogeography of bacterial
608 communities in wastewater treatment plants. *Nature Microbiology* **4**: 1183-1195.

609 Xu Q, Luo G, Guo J, Xiao Y, Zhang F, Guo S, Ling N & Shen Q (2022) Microbial generalist or
610 specialist: Intraspecific variation and dormancy potential matter. *Molecular Ecology* **31**: 161-173.

611 Yates CF, Trexler RV, Bonet I, King WL, Hockett KL & Bell TH (2022) Rapid niche shifts in
612 bacteria following conditioning in novel soil environments. *Functional Ecology* **36**: 3085-3095.

613 Zrimec J, Kokina M, Jonasson S, Zorrilla F & Zelezniak A (2021) Plastic-Degrading Potential
614 across the Global Microbiome Correlates with Recent Pollution Trends. *mBio* **12**: e02155-02121.
615

616