Escape from bacterial diversity: potential enemy release in invading yellow starthistle (*Centaurea solstitialis*) microbiomes

Patricia Lu-Irving[1], Julia Harenčár[1,2], Hailey Sounart[1,3], Shana R Welles[1], Sarah M Swope[3], David A Baltrus[4,5], and Katrina M Dlugosch[1]

[1] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA; [2] Current address: Biological Sciences Department, California Polytechnic State University, San Luis Obispo, California, USA; [3] Department of Biology, Mills College, Oakland, California, USA; [4] School of Plant Sciences, University of Arizona, Tucson, Arizona, USA; [5] School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, Arizona, USA.

Author for correspondence:
Patricia Lu-Irving
Email: luirving@email.arizona.edu

Total word count (excluding summary, references and legends):	5753	No. of figures:	7
Summary:	195	No. of Tables:	0
Introduction:	1147	No of Supporting Information files:	1
Materials and Methods:	1908		
Results:	803		
Discussion:	1762		
Acknowledgements:	133		

SUMMARY

- Invasive species may benefit from introduction to new regions where they can escape their natural enemies. Here we examined whether geographic patterns of microbial community composition support a role for enemy escape in the invasion of California, USA by yellow starthistle, a highly invasive plant in western North America.
- We used high-throughput sequencing of the 16S V4 region to characterize bacterial community composition in the phyllosphere, rhizosphere, leaves, and roots of plants from seven populations in California and eight populations in the native European range. We compared bacterial diversity between the native and invaded ranges, and with previously published estimates of plant genetic diversity within each population.
- Bacterial communities differed significantly among plant compartments, and between native and invaded ranges within compartments, with consistently lower diversity in the invaded range. Plant genetic diversity did not explain this pattern in bacterial diversity, but a positive relationship was found within ranges between bacterial diversity in roots and plant genetic diversity within populations.
- Our observation of lower bacterial diversity in the invaded relative to the native range of yellow starthistle is consistent with potential enemy escape, providing some of the first evidence for this scenario in plant microbiomes.

KEYWORDS

bacteria, *Centaurea solstitialis*, endophyte, genetic diversity, invasive species, microbiome, phyllosphere, rhizosphere

INTRODUCTION

Humans continue to transport plant species around the globe, and increasing numbers of these translocations result in the invasive expansion of non-native species into recipient communities (Lonsdale, 1999; Butchart *et al.*, 2010; Essl *et*

al., 2011; Ellis *et al.*, 2012). There is a longstanding hypothesis that many species become invasive after escaping from enemies that reduce invader fitness and limit their populations in their native ranges (Darwin, 1859; Williamson, 1996). Known as the 'Enemy Release' hypothesis, this idea is highly intuitive and forms a basis for the biological control of invasive species (Keane & Crawley, 2002). Initial tests of enemy release focused on quantifying visible changes in above-ground herbivore damage (Keane & Crawley, 2002), but there has been increasing recognition that microbial enemies above- and below-ground can have large effects on plant fitness, and could thus determine whether invasive plants benefit from escaping negative species interactions (Callaway *et al.*, 2004; Colautti *et al.*, 2004; Torchin & Mitchell, 2004; Agrawal *et al.*, 2005; Mitchell *et al.*, 2006; Kulmatiski *et al.*, 2008; van der Putten *et al.*, 2013; Dawson & Schrama, 2016; Faillace *et al.*, 2017).

In recent years, microbial communities have emerged as particularly likely candidates for generating enemy release. Although many interactions between plants and microbes can be beneficial, microbial communities often appear to have negative net effects on plant fitness which may become more negative over time, e.g., via plant-soil feedbacks (Bever, 2003; Reinhart & Callaway, 2006; Kulmatiski *et al.*, 2008; Petermann *et al.*, 2008). It is now apparent that interactions between plants and their microbiomes can vary over space and environment (Nemergut *et al.*, 2013; van der Putten *et al.*, 2013; terHorst & Zee, 2016), creating opportunities for introduced plants to escape the microbial communities that characterize their native ranges. Moreover, evidence is building that reductions in microbial diversity are occurring in response to environmental change and human disturbances, and these reductions in diversity may reduce the resistance of ecosystems to invasion (Schnitzer *et al.*, 2011; Wagg *et al.*, 2014; Dawson & Schrama, 2016; van der Putten *et al.*, 2016).

Invasive plant species have provided some of the best evidence to date that microbial interactions can be locally evolved, and can vary considerably over

geographic regions (Rout & Callaway, 2012). Invaders have been shown to vary in their response to soil communities from their native and invaded ranges, and there are now many examples of more favorable interactions between plants and soil from the invaded range, consistent with escape from enemies (Reinhart *et al.*, 2003; Callaway *et al.*, 2004; Mitchell *et al.*, 2006; Engelkes *et al.*, 2008; Kulmatiski *et al.*, 2008; Maron *et al.*, 2014; van der Putten *et al.*, 2016). Plant-microbe interactions which provide net benefits to invasive species can be explained by reduced negative effects of key microbial pathogens, increased direct beneficial effects of mutualistic taxa, or increased indirect benefits from taxa that affect competitors more negatively than they do the invader (Dawson & Schrama, 2016). These mechanisms should manifest as differences in the microbial communities associated with invading vs. native plants, specifically as divergence in taxonomic composition, reduction in diversity, and/or the loss or gain of groups known to have pathogenic or mutualistic effects, where taxonomic resolution permits inference of function (Herrera Paredes & Lebeis, 2016).

Release from enemies is expected to be beneficial in and of itself, but it may further promote invasion by changing the pattern of natural selection on resource allocation by the invader (Sakai *et al.*, 2001). Plants that require reduced defenses against negative enemy interactions have the potential to adapt to invest a larger proportion of resources in traits that increase competitiveness, reproduction, and/or spread. This idea, known as the Evolution of Increased Competitive Ability (EICA) hypothesis (Blossey & Notzold, 1995) has received a great deal of attention but mixed empirical support (Maron *et al.*, 2004; Bossdorf *et al.*, 2005; Mitchell *et al.*, 2006; Felker-Quinn *et al.*, 2013). Potential contributors to evolutionary responses to enemy release are likely to become better resolved as our understanding of microbial community interactions increases, particularly since adaptive responses to microbial disease are known to be among the most rapid evolutionary changes that occur in any organism (Tiffin & Moeller, 2006; Bomblies *et al.*, 2007; Salvaudon *et al.*, 2008).

Here we conduct one of the first comparisons of plant microbiomes between invading populations and their native source region, explicitly testing for patterns consistent with enemy release (see also Gundale *et al.*, 2016). We ask whether changes in plant-associated microbial communities have the potential to generate enemy escape in the highly invasive plant yellow starthistle (*Centaurea solstitialis*). Yellow starthistle is native to a wide region of Eurasia and was introduced to South America in the 1600's and North America in the 1800's as a contaminant of alfalfa seed (Gerlach, 1997). This herbaceous annual is a colonizer of grassland ecosystems, and is often called one of the '10 Worst Weeds of the West' in North America (DiTomaso & Healy, 2007). Its extensive invasion of California in the USA (>14 million acres; Pitcairn *et al.*, 2006) is well-studied, and invading genotypes in this region have been shown to grow larger and produce more flowers than plants in the native range, suggesting an adaptive shift in resource allocation and an increase in invasiveness (Widmer *et al.*, 2007; Eriksen *et al.*, 2012; Dlugosch *et al.*, 2015). Previous research has demonstrated that yellow starthistle throughout all of its native and invaded ranges experiences net fitness reductions when grown with its local soil communities (Andonian *et al.*, 2011, 2012; Andonian & Hierro, 2011). However, these studies have also indicated that this negative interaction is weaker (more favorable) in California, raising the possibility that escape from microbial enemies has promoted this aggressive invasion.

We sample microbial communities associated with leaves (phyllosphere and endosphere) and roots (rhizosphere and endosphere) of yellow starthistle plants in both the California invasion and its source region in Europe. Previous experiments with fungicide treatments have shown that plant-soil interactions between yellow starthistle and fungi in California are *more* negative (less favorable) than those in the native range, inconsistent with a role for fungi in escape from microbial enemies (Hierro *et al.*, 2016). Here, we focus on bacterial communities as candidates for a potential role in enemy escape in this system. We use high-throughput sequencing of prokaryotic ribosomal 16S sequences to

quantify diversity and relative abundance of taxa in yellow starthistle microbiomes, designing a novel modified peptide nucleic acid clamp (Lundberg *et al.*, 2013) to reduce non-target sequencing of host plastids. We ask whether there are patterns of reduced taxonomic diversity and/or potential loss of pathogens in the invaded range, and whether patterns of diversity in plant-associated bacteria can be explained by geographic patterns of plant genetic diversity. Loss of potential pathogens would be consistent with opportunities for enemy escape that could contribute to the evolution of increased invasiveness in yellow starthistle's highly successful invasion of California.

MATERIALS AND METHODS

Study species

Yellow starthistle (*Centaurea solstitialis* L., Asteraceae) is an obligately outcrossing annual plant, diploid throughout its range (Heiser & Whitaker, 1948; Widmer *et al.*, 2007; Öztürk *et al.*, 2009). Plants form a taproot and grow as a rosette through mild winter and/or spring conditions, bolting and producing flowering heads (capitula) throughout the summer. The species is native to Eurasia, where distinct genetic subpopulations have been identified in Mediterranean western Europe, central-eastern Europe, Asia (including the Middle East), and the Balkan-Apennine peninsulas (Barker *et al.*, 2017). The invasion of California as well as those in South America appear to be derived almost entirely from western European genotypes (Fig. 1; Barker *et al.*, 2017).

Sample collection

Fifteen populations of yellow starthistle were sampled for microbial communities: seven populations across the invasion of California, six in western Europe, and two in eastern Europe (Fig. 1; Supporting Information Table S1). At each population, plants were sampled every meter (or to the nearest meter mark) along a 25 meter transect, to yield 25 individuals per population. Individuals in rosette or early bolting stages were preferentially selected. In one population (HU29), low plant density yielded 20 individuals along the 25 meter transect.

Using sterile technique, plants were manually pulled and each individual sampled for phyllosphere, rhizosphere, leaves, and roots using modified versions of protocols by Lundberg *et al.* (2013) and Lebeis *et al.* (2015) as described below. A control (blank) sample was collected for each population. Plants were pressed and dried after sampling, and submitted to the University of Arizona Herbarium (ARIZ; Supporting Information Table S1).

Phyllosphere and rhizosphere — one to three basal, non-senescent leaves were collected from each plant, as well as the upper 2-5 cm of the taproot, together with accompanying lateral roots (excess soil was brushed or shaken off). Leaf and root samples were placed in individual 50 ml tubes containing 25 ml of sterile wash solution (45.9 mM NaH_2PO_4, 61.6 mM Na_2HPO_4, 0.1% Tween 20). Tubes were shaken by hand for one minute (timed). Leaf and root samples were then removed and stored on ice in separate tubes (leaves in empty tubes, roots in tubes containing 10 ml of wash solution) until further processing. Wash samples were stored on ice during transport, then refrigerated at 4^{o}C. Phyllosphere and rhizosphere washes were pooled per population, then centrifuged at 2,200 g at 4^{o}C for 15 minutes. Supernatants were discarded, and pellets were air-dried and stored at -20^{o}C until DNA extraction.

Leaf endosphere — leaves were surface sterilized by submerging in bleach solution (10% commercial bleach, 0.1% tween 20) for two minutes. Leaves were then rinsed in distilled water, patted dry using clean kimwipe, and sealed in individual sterile surgical envelopes (Fisherbrand #01-812-50). Envelopes were kept in silica gel desiccant until leaf tissue was completely dry, then stored at room temperature until DNA extraction.

Root surface and endosphere (hereafter 'whole root') — roots were further washed by shaking in 10 ml of wash solution until visible residual soil was removed. Washed roots were stored and dried as described above for leaves.

Controls — at each collection site, a tube of sterile wash solution was left uncapped while sampling plants. Disinfected tools were periodically swished in the blank wash tube before sterilization and use for the next sample collection. For each population, rinse water and wipes used to process tissue samples were represented in controls by rinsing and wiping flame-sterilized forceps, then swishing the forceps in the blank wash tube. Controls were stored and processed in the same manner as phyllosphere and rhizosphere samples.

DNA extraction

Extractions were carried out using sterile technique in a laminar flow hood. Leaf and root DNA was extracted as bulk samples from tissue pooled by population (15 total populations), and as individual samples from 8 plants from each of 10 populations (80 total individuals). For pooled tissue extractions, equal sections of leaf tissue (50 mm^2) and root tissue (12.5 mm^3 plus 10 mm of lateral roots) were collected from each individual sample per population and pooled prior to extraction. Control (blank) samples were collected for each batch of extractions by swabbing tools and surfaces, then extracting DNA from the swab head.

All DNA samples were extracted using the MO BIO PowerSoil kit (MO BIO Laboratories, Inc.). Phyllosphere and rhizosphere DNA was extracted from up to 0.25 g of wash pellets following the standard kit protocol. Leaf and root tissues were ground to powder or sawdust consistency in liquid nitrogen using sterile mortars and pestles. Leaf and root DNA was extracted from 20 mg (leaf) or 100 mg (root) of ground tissue with the following modification to the standard protocol: tissue was incubated at 65^{o}C for 10 minutes in extraction buffer, then vortexed for 1 minute, followed by a second 10 minute incubation (as described under “alternative lysis methods” in the kit protocol). Control DNA was extracted by placing whole swab heads directly into extraction tubes. Extracted DNA was eluted in PCR-grade water and stored at -20^{o}C pending library preparation.

Library preparation and sequencing

To remove secondary compounds inhibiting PCR, DNA extracted from root and leaf tissue (together with corresponding blanks) was purified using a ZR-96 genomic DNA clean-up kit (Zymo Research). All DNA concentrations were quantified using a Qubit fluorometer high-sensitivity assay for double-stranded DNA (Invitrogen), and standardized to equimolar amounts.

Library preparation followed a dual barcoded two-step PCR protocol. In the first step (target-specific PCR), the V4 region of the 16S rRNA gene was amplified using target specific primers (515F and 806R) appended with common sequence (CS) tags through a linker sequence which varied from two to five nucleotides in length. Target-specific PCR was carried out using Phusion Flash master mix (Thermo Scientific) in 25 µl reaction volume in a Mastercycler pro thermocycler (Eppendorf) under the following conditions: 25 cycles of 1 s at 98^{o}C, 5 s at 78^{o}C, 5 s at 57^{o}C, 15 s at 72^{o}C. Products were visualized on an agarose gel and diluted by up to 1:15 (depending on yield); 1 µl of diluted product was then used as template in the second step (barcode-adapter attachment PCR). Using reagents and equipment as described above, barcoded primer pairs incorporating Illumina P5 and P7 adapters were used to amplify products from target-specific PCR in 25 µl reaction volumes under the following conditions: 10 cycles of 1 s at 98^{o}C, 5 s at 78^{o}C, 5 s at 51^{o}C, 15 s at 72^{o}C. Barcoded amplicons were quantified by fluorometry, pooled in equimolar amounts, cleaned, and submitted to the University of Idaho's IBEST Genomic Resources Core for QC and sequencing. Amplicons were multiplexed to use half the capacity of one 2 × 300 bp run on an Illumina MiSeq platform. Raw sequence data are deposited in the NCBI Short Read Archive under accession number XXXXXX [pending submission].

Peptide nucleic acid clamps (PNAs) were included in both PCR steps of library preparation to block amplification of plant chloroplast and mitochondrial 16S as recommended by Lundberg *et al.* (2013). Clamp sequences published by Lundberg *et al.* (2013) were compared with chloroplast and mitochondrial 16S sequences from yellow starthistle and three other species of Asteraceae with

published organellar genomes (*Centaurea diffusa*, *Helianthus annuus*, *Lactuca sativa*). We found a single nucleotide mismatch between the Asteraceae chloroplast 16S and the plastid PNA sequences, and designed an alternative plastid PNA specific to the Asteraceae sequence (5'—GGCTCAACTCTGGACAG—3'). All samples for this study were amplified using the plastid PNA of our design, together with the mitochondrial PNA published by Lundberg *et al.* (2013). To gauge the effectiveness of our alternative PNA, two duplicate samples were processed using both PNAs published by Lundberg *et al.* (2013).

Identification of operational taxa

Demultiplexed paired reads were merged and quality filtered using tools from the USEARCH package version 9.0.2132 (Edgar & Flyvbjerg, 2015). Merged reads were truncated to uniform length and primer sequences were removed using a combination of the seqtk toolkit version 1.2 (github.com/lh3/seqtk) and a custom script. The UPARSE pipeline (Edgar, 2013) implemented in the USEARCH package was used for further data processing and analysis: unique sequences were identified, and those represented only once or twice in the processed read set were discarded as likely PCR or sequencing errors. Remaining sequences were clustered into operational taxonomic units (OTUs) at a 97% threshold, chimeras were filtered out, and per-sample OTU read counts were tabulated using the UPARSE-OTU algorithm. Assignment of OTUs to nearest taxonomic match in the Greengenes database (McDonald *et al.*, 2012) was carried out using the UCLUST algorithm implemented in QIIME version 1.9.1 (Caporaso *et al.*, 2010; Edgar, 2010). Data were further processed using tools from the QIIME package: reads mapping to chloroplast and mitochondrial OTUs were removed, and samples were rarefied by plant compartment. Rarefaction levels were chosen to reflect the distribution of read counts per sample within plant compartments, subsampling to the minimum number of reads necessary to include all samples except those that were outliers for low read count.

297 **Microbial community analyses**

298 All statistical analyses were performed in R (R Core Team, 2015). We evaluated
299 overall differences in bacterial community composition between plant
300 compartments, and between native and invaded ranges within plant
301 compartments, by performing non-metric multidimensional scaling (NMDS) using
302 the R packages vegan (Oksanen *et al.*, 2016) and MASS (Venables & Ripley,
303 2002). Ordinations were based on Bray-Curtis distances, and were performed
304 using a two-dimensional configuration to minimize stress, using Wisconsin
305 double standardized and square root transformed data, with expanded weighted
306 averages of species scores added to the final NMDS solution. Significant
307 differences among plant compartments and between native and invaded samples
308 were assessed using the envfit function in vegan. Ellipses were drawn on NMDS
309 plots using the vegan function ordiellipse, representing 95% confidence limits of
310 the standard error of the weighted average of scores.
311
312 We further explored the underlying correlates of bacterial community variation
313 using Principal Components Analysis (PCA; using R function prcomp) for
314 samples from native and invaded ranges within each plant compartment. We
315 identified the OTUs with the highest loading on the dominant PC axis of variation
316 by examining the matrix of variable loadings produced by prcomp. The OTU
317 composition of samples pooled by population (phyllosphere, rhizosphere, and
318 bulk root samples; hereafter 'bulk samples') was visualized using a heatmap
319 generated in ggplot2 (Wickham, 2009). Bulk samples were hierarchically
320 clustered by Bray-Curtis dissimilarity (hclust function in R) using McQuitty's
321 method (McQuitty, 1966).
322
323 We compared the diversity of OTUs between the native and invaded range for
324 each plant compartment using both richness (R) and the Shannon diversity index
325 (H'; Shannon, 1948), which reflects the contributions of both taxonomic richness
326 and evenness to diversity. These values were calculated using the vegan
327 package, and compared between native and invaded ranges using a

nonparametric Kruskal-Wallis rank sum test on rarefied read counts. For plant tissue samples that included multiple individuals per site, we compared the diversity among sites using a Kruskal-Wallis test within regions.

Finally, we asked whether the geographic distribution of plant-associated bacterial diversity could be explained by the geographic distribution of genetic diversity in the plants. Measurements of plant genetic diversity at each of our sampling sites were obtained from previously published genome-wide marker analyses by Barker and colleagues (Barker *et al.*, 2017), calculated as the average proportion of pairwise nucleotide differences between alleles ($\boldsymbol{\pi}$) at variable sites across the yellow starthistle genome. Diversity estimates (H') for each plant compartment were predicted using linear models that included fixed effects of plant genetic diversity, region (native vs. invaded), and the interaction between these two effects.

RESULTS

Sequencing and data processing

Sequencing yielded 9,672,898 read pairs, of which 6,217,852 remained after merging and quality control; these were 253 bp in length after removing artificial and primer sequences. The number of raw read counts per sample ranged from 16 to 306,200 with a median of 21,964.

Analysis of the merged and processed reads resulted in 4,014 OTUs, of which 60 were identified as plastid or mitochondrial. Sequences representing yellow starthistle chloroplast and mitochondrial 16S accounted for 40% and 1% of all reads, respectively. Amplification of host chloroplast in samples using the Asteraceae-specific plastid PNA was reduced by up to 51% compared with the Lundberg *et al.* (2013) PNA (Supporting Information Table S2). This is consistent with results from a broader comparison of the two PNAs, using five Asteraceae species, which also found that blocking of host chloroplast amplification was improved by using the Asteraceae-specific PNA (FitzPatrick *et al.*, unpublished).

Despite PNA blocking activity, 83% of the total reads from leaf endosphere samples were yellow starthistle chloroplast sequences. This might potentially be attributable to high chloroplast DNA concentrations relative to endophyte DNA in leaf tissue total DNA extracts. After removal of chloroplast and mitochondrial reads, remaining read counts for most leaf endosphere samples were low (Fig. 2), so no further analysis of leaf endosphere bacterial communities was performed.

Rarefaction levels (chosen to reflect the minimum number of reads per sample by compartment, not including outliers) were 18,000 reads per sample for phyllosphere, 17,000 for rhizosphere, and 5,000 for whole root samples. These levels resulted in the exclusion of six samples which were outliers for low read count: one phyllosphere (DIA), one rhizosphere (SAZ), one bulk root (CAN), and three individual root samples (two from SAZ; one from SIE).

Microbial community analyses

Results from NMDS ordinations indicated that bacterial communities differed overall among the phyllosphere, rhizosphere, and whole root compartments (Fig. 3a; $P = 0.001$). Within compartments, NMDS further revealed significant differences between native and invaded range whole root samples (Fig. 3b; $P = 0.001$) and rhizosphere samples ($P = 0.001$). Native and invaded range phyllosphere samples differed with marginal significance ($P = 0.05$). The dominant phyla among bacterial communities were Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes (Fig. 4), which is consistent with the findings of previous characterizations of plant-associated bacterial communities (reviewed by Bulgarelli *et al.*, 2013). Principal component analyses suggested that the strongest contributions to changes in bacterial community composition between the native and invaded ranges were made by shifts in the representation of *Pseudomonas*, *Erwinia*, *Chryseobacterium*, Xanthomonadaceae, and *Bacillus* taxa (Supporting Information Fig. S1; Table S3). Clustering analyses within the phyllosphere and rhizosphere compartments

consistently grouped invaded range samples together, as well as samples from the source region in western Europe (Supporting Information Fig. S2). Native range samples from eastern Europe (HU01 and HU29) clustered together in these compartments but were variable in their relationship to the other regions. Bulk root samples showed less consistent clustering by range.

Bacterial OTU diversity (H': Fig. 5) was significantly lower in the invaded range in the phyllosphere (X^2_1 = 5.36, P = 0.02) and rhizosphere (X^2_1 = 6.21, P = 0.01), and marginally lower in bulk whole root samples (X^2_1 = 3.01, P = 0.08). Diversity in whole root individual samples (Fig. 5d) did not vary significantly among populations within the native range (X^2_4 = 1.82, P = 0.77), but did vary significantly within the invaded range (X^2_4= 15.30, P = 0.004), such that the two most extreme populations (TRI, SIE) were significantly different from one another but not the remaining sites. Bacterial OTU richness (R) values showed patterns similar to H' in general, but differences between native and invading regions were much weaker overall (Supporting Information Fig. S3), indicating that both richness and evenness of OTU representation contributed to differences in diversity between the ranges. For whole roots, our most extensively sampled plant compartment, an analysis of OTUs observed across all bulk and individual samples combined indicated that the native and invaded range shared 51% of observed OTUs, with 31% fewer unique OTUs observed in the invaded relative to the native range (Fig. 6a). These patterns were reflected across both major groups of Proteobacteria and Actinobacteria (Fig. 6b,c).

A linear model predicting microbial diversity (H') from plant diversity was significant for bulk whole root samples (Fig. 7; $F_{(2,12)}$ = 4.99; P = 0.02; r^2_{adj} = 0.36), with significant main effects of both plant genetic diversity (P = 0.04) and region (native vs. invaded; P = 0.009). The interaction between these two effects was not significant (P = 0.69) and was removed from the final model. Similar linear models did not identify significant effects of plant diversity when predicting

phyllosphere (P = 0.42) or rhizosphere (P = 0.11) diversity, nor the median diversity of individual whole root samples at a site (P = 0.95).

DISCUSSION

Our study revealed strikingly lower diversity of bacterial communities in yellow starthistle's invasion of California (USA), relative to native European populations within and beyond the source region for the invasion. Reduced bacterial diversity was apparent across phyllosphere, rhizosphere, and whole root communities, and across dominant bacterial phyla. These patterns are consistent with opportunities for enemy escape in this invasion and could explain why soils have more favorable (less negative) effects on yellow starthistle fitness in California relative to other parts of its range (Andonian *et al.*, 2011, 2012; Andonian & Hierro, 2011).

In line with other surveys of plant microbiomes, we found that differences in bacterial communities were greatest among plant compartments (i.e., phyllosphere, rhizosphere, and roots; (Bulgarelli *et al.*, 2013; Vandenkoornhuyse *et al.*, 2015). The numbers and diversity of taxa (OTUs) that we recovered in samples from each compartment were generally similar in magnitude to those reported in other studies of prokaryotic 16S sequences, from angiosperm groups as diverse as e.g. Agavaceae (Coleman-Derr *et al.*, 2016), Brassicaceae (Bodenhausen *et al.*, 2013), Cactaceae (Fonseca-García *et al.*, 2016), and other Asteraceae (Leff *et al.*, 2016). Notably, we found that diversity was approximately twice as high in the roots than the rhizosphere. Higher root endosphere diversity relative to the rhizosphere is also reported in some other recent studies (Fonseca-García *et al.*, 2016; Leff *et al.*, 2016), but previous reviews have concluded that root endosphere communities are typically less diverse than those in the rhizosphere (Bulgarelli *et al.*, 2013; Vandenkoornhuyse *et al.*, 2015). Our root collections were not surface sterilized and may represent some of the rhizoplane/rhizosphere in addition to the endosphere, elevating our estimates of diversity, though it is also possible that yellow starthistle deviates from previous

trends. In contrast to phyllosphere, rhizosphere, and root compartments, we recovered few microbial sequences for leaf bacterial endophytes, even when compared to control (blank) samples, suggesting low sequence coverage due to persistent chloroplast contamination, and potentially low overall bacterial loads within yellow starthistle leaves.

Within compartments, differences in community composition between ranges were substantial and dominated by differences in OTU diversity, which ranged from 18-40% lower in the invaded range. This variation in diversity is similar in scale to other studies that have sampled distant geographic locations (e.g. across regions of North America: (Bodenhausen *et al.*, 2013; Peiffer *et al.*, 2013). A variety of factors may explain this pattern, including environmental differences across sites (Fierer & Jackson, 2006; Bulgarelli *et al.*, 2013; Nemergut *et al.*, 2013; Vandenkoornhuyse *et al.*, 2015). Soil type appears to have a particularly strong influence on microbial communities (e.g. Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012), and is known to differ broadly across yellow starthistle's range (Hierro *et al.*, 2016). In addition, populations in California include temperature and precipitation environments that are on the warm and dry extreme of yellow starthistle's climatic niche (Dlugosch *et al.*, 2015), and our sampling was conducted at the end of a period of severe drought in California (Griffin & Anchukaitis, 2014; Diffenbaugh *et al.*, 2015), which could have amplified microbial differences related to climate (Schrama & Bardgett, 2016). Interestingly, a recent study of grassland plants found that microbial diversity increased under drought, whereas we found reduced diversity in our drought-affected invaded range (terHorst *et al.*, 2014).

We also observed an effect of plant genotypic diversity on microbial diversity in bulk samples of roots. We note that bacterial diversity at the level of the individual plant did not covary with plant population genetic diversity, indicating that it was only when samples from different plants were combined that an effect of plant genotype variation was apparent. Such within-species plant genotype

effects have been observed in other studies and may interact with the effect of environment to shape microbial communities (Peiffer *et al.*, 2013; terHorst & Zee, 2016). In many cases, it appears that genotype effects are minor relative to site effects (e.g. (Lundberg *et al.*, 2012; Peiffer *et al.*, 2013; Bodenhausen *et al.*, 2014; Bulgarelli *et al.*, 2015), and we also found that genotypic effects were not stronger than between-region variation in yellow starthistle. Moreover, plant genotype effects were only significant in the roots, the only endophytic compartment analyzed: consistent with plant genotype having the strongest influence on microbial taxa colonizing within the plant itself (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012; Lebeis *et al.*, 2015). The ability of plants to shape the composition and quantity of endophytic microbial taxa in particular has been suggested to contribute to a 'core' microbiome shared across environments (Lundberg *et al.*, 2012; Shade & Handelsman, 2012; Reisberg *et al.*, 2013), and indeed we found that the majority of OTUs observed in yellow starthistle roots were shared across invading and native populations on different continents.

Importantly, yellow starthistle's invasion to high density could be a cause rather than an effect of low microbial diversity. Species invasions and range expansions have been shown to change microbial composition over short timescales (Collins *et al.*, 2016; Gibbons *et al.*, 2017). Yellow starthistle invasions are denser than populations surveyed in the native range by an order of magnitude or more (Uygur *et al.*, 2004; Andonian *et al.*, 2011). Invaded communities that are dominated by yellow starthistle include lower diversity of plant species overall (Seabloom *et al.*, 2003; Zavaleta & Hulvey, 2004; D'Antonio *et al.*, 2007), and low plant diversity may depress the diversity of plant-associated microbes in the environment (Garbeva *et al.*, 2004; Schnitzer *et al.*, 2011; Coleman-Derr *et al.*, 2016). Such an effect of plant density could provide an explanation for a general pattern of weaker plant-soil interactions for invasive species in their introduced ranges (Kulmatiski *et al.*, 2008; Dawson & Schrama, 2016). These results reinforce a growing need for explicit observational and experimental tests of the association between microbial diversity and the potential influences of plant

density, plant community diversity, and environmental gradients (Dawson & Schrama, 2016).

This study is among the first to examine differences in microbial taxa between the native and introduced ranges of an invasive species. Gundale and colleagues (Gundale *et al.*, 2014) also identified more favorable soil interactions in invasions of lodgepole pine (*Pinus contorta*), and explored potential enemy escape in its fungal endophyte community (Gundale *et al.*, 2016). For lodgepole pine, microbial communities differed among regions, but there was no consistent pattern of loss of potential fungal pathogens or gain of mutualists in the invaded range, and it remains unclear what part of the soil community is responsible for observed differences in interactions across ranges (Gundale *et al.*, 2016). Finkel and colleagues (Finkel *et al.*, 2011; 2016) similarly explored the phyllosphere community of multiple species of *Tamarix* in native and introduced parts of their range, finding that communities are most strongly structured by geographic region. Our study reveals that this type of comparative microbiome approach can be fruitful for identifying changes in species interactions that might be contributing to invasion success.

One of the central challenges in testing the hypothesis that invaders are benefitting from enemy release is quantifying the impact of all types of enemies, with the microbial community being historically the hardest to observe (Keane & Crawley, 2002; Beckstead & Parker, 2003; Dawson & Schrama, 2016; Müller *et al.*, 2016; van der Putten *et al.*, 2016; Crawford & Knight, 2017). For yellow starthistle, a great deal of effort has gone into the identification of potential native herbivores/seed predators that could be used as biocontrol in California. Six specialist biocontrol insect species and one fungal foliar pathogen have been released into this area without resulting in effective control (DiTomaso *et al.*, 2006; Swope & Parker, 2012), suggesting that escape from these species has not facilitated the invasion. Our finding that invaded populations have not only

543 unique, but also less diverse bacterial communities, suggests particularly strong
544 opportunities for pathogen escape in this system.
545
546 We have previously argued that yellow starthistle has invaded into a low
547 competition environment in California, benefitting from the historical loss of plant
548 competitors for water in this system (Dlugosch *et al.*, 2015). Disturbance of the
549 native community is critical for yellow starthistle establishment, and functionally
550 similar native species compete well against it in experiments; however, key
551 competitors have been lost from the ecosystem due to a variety of perturbations
552 prior to the yellow starthistle invasion (Zavaleta & Hulvey, 2004; Hooper &
553 Dukes, 2010; Hierro *et al.*, 2011, 2016; Hulvey & Zavaleta, 2012). Any benefits to
554 yellow starthistle of reduced bacterial diversity could be independent of these
555 interactions with native plant species, but there are clear opportunities for these
556 factors to be related. If a lack of competition allowed yellow starthistle to increase
557 in density, then this could have reduced plant-associated microbial diversity in
558 the environment, as noted above. However, while this scenario could explain
559 lower diversity among bacteria, it appears that density is unlikely to explain
560 reduced negative interactions with the soil community in California. Yellow
561 starthistle experiences some of its strongest negative plant-soil feedbacks across
562 generations in California soils (Andonian *et al.*, 2011), suggesting that the build
563 up of high plant densities is unlikely to explain patterns of enemy release.
564 Alternatively, the historical loss of native species diversity in California (D'Antonio
565 *et al.*, 2007) could have resulted in the loss of associated microbial diversity,
566 generating particularly strong opportunities for invasion into a system with both
567 reduced competition and reduced pathogen diversity. Microbial surveys of
568 remnant native communities, as well as across densities of yellow starthistle
569 would help to clarify alternative interacting effects of plant and microbial diversity,
570 and it may be particularly enlightening to explore microbial communities
571 preserved on native plant specimens pre-dating the extensive invasion of yellow
572 starthistle into this region.
573

Conclusions

To our knowledge, our study is the first to find evidence consistent with opportunities for release from microbial enemies during invasion. We find lower overall bacterial diversity in invading plant populations, similar in scale to geographic variation in bacterial diversity that has been observed in other studies. These patterns suggest that yellow starthistle may have benefitted from introduction into disturbed plant communities with relatively low microbial diversity. Microbial interactions appear to be important for plant fitness in this system, but may interact with other factors shaping invasiveness, including disturbance and lack of effective competition from native plant species. In particular, escape from both microbial enemies and plant competitors might have created an opportunity for adaptive allocation of resources away from defensive functions and towards reproduction and the evolution of increased invasiveness in yellow starthistle. Comparative surveys of the microbiome in invading and native populations, as presented here, can reveal important variation in the species interactions that are shaping patterns of invasion.

ACKNOWLEDGMENTS

We thank D. Lundberg and S. Lebeis for helpful discussions regarding sampling design; G. Reardon and C. Chandeyson for assistance with field collections; E. Arnold, J. Aspinwall, J. Braasch, E. Carlson, K. Hockett, T. O'Connor, M. Schneider, J. U'Ren, and N. Zimmerman for 16S library preparation assistance and discussion; A. Gerritsen, D. New and staff at iBEST for assistance with sequencing; K. Andonian, J. Hierro, and # reviewers for helpful feedback on the manuscript. We are particularly indebted to C.E. Morris for hosting the European sample processing at INRA Station de Pathologie Végétale, Montfavet, France. Data collection and analyses performed by the IBEST Genomics Resources Core at the University of Idaho were supported in part by NIH COBRE grant P30GM103324. This study was supported by USDA grant 2015-67013-23000 to K.M.D., D.A.B., and S.M.S.

AUTHOR CONTRIBUTIONS

P.L-I., D.A.B., and K.M.D. designed the study. P.L-I. and J.H. collected the samples with assistance from H.S., S.M.S., and S.R.W. P.L-I conducted the microbial sequencing and bioinformatics. P.L-I, S.R.W., and K.M.D. analyzed the data. P.L-I and K.M.D. wrote the manuscript, which was edited by all authors.

REFERENCES

Agrawal AA, Kotanen PM, Mitchell CE, Power AG, Godsoe W, Klironomos J. **2005**. Enemy release? an experiment with congeneric plant pairs and diverse above- and belowground enemies. *Ecology* **86**: 2979–2989.

Andonian K, Hierro JL. **2011**. Species interactions contribute to the success of a global plant invader. *Biological Invasions* **13**: 2957–2965.

Andonian K, Hierro JL, Khetsuriani L, Becerra PI, Janoyan G, Villareal D, Cavieres LA, Fox LR, Callaway RM. **2012**. Geographic mosaics of plant-soil microbe interactions in a global plant invasion. *Journal of Biogeography* **39**: 600–608.

Andonian K, Hierro JL, Khetsuriani L, Becerra P, Janoyan G, Villarreal D, Cavieres L, Fox LR, Callaway RM. **2011**. Range-expanding populations of a globally introduced weed experience negative plant-soil feedbacks. *PLoS One* **6**: e20117.

Barker BS, Andonian K, Swope SM, Luster DG, Dlugosch KM. **2017**. Population genomic analyses reveal a history of range expansion and trait evolution across the native and invaded range of yellow starthistle (Centaurea solstitialis). *Molecular Ecology* **26**: 1131–1147.

Beckstead J, Parker IM. **2003**. Invasiveness of Ammophila arenaria: release from soil-borne pathogens? *Ecology* **84**: 2824–2831.

Bever JD. **2003**. Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. *The New Phytologist* **157**: 465–473.

634 **Blossey B, Notzold R**. **1995**. Evolution of increased competitive ability in
635 invasive nonindigenous plants: a hypothesis. *The Journal of Ecology* **83**:
636 887–889.
637 **Bodenhausen N, Bortfeld-Miller M, Ackermann M, Vorholt JA**. **2014**. A
638 synthetic community approach reveals plant genotypes affecting the
639 phyllosphere microbiota. *PLoS Genetics* **10**: e1004283.
640 **Bodenhausen N, Horton MW, Bergelson J**. **2013**. Bacterial communities
641 associated with the leaves and the roots of Arabidopsis thaliana (AM
642 Ibekwe, Ed.). *PloS one* **8**: e56329.
643 **Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D**.
644 **2007**. Autoimmune response as a mechanism for a Dobzhansky-Muller-
645 type incompatibility syndrome in plants. *PLoS Biology* **5**: e236.
646 **Bossdorf O, Auge H, Lafuma L, Rogers WE, Siemann E, Prati D**. **2005**.
647 Phenotypic and genetic differentiation between native and introduced
648 plant populations. *Oecologia* **144**: 1–11.
649 **Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y,**
650 **McHardy AC, Schulze-Lefert P**. **2015**. Structure and function of the
651 bacterial root microbiota in wild and domesticated barley. *Cell Host &*
652 *Microbe* **17**: 392–403.
653 **Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad**
654 **N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E, *et al.***
655 **2012**. Revealing structure and assembly cues for Arabidopsis root-
656 inhabiting bacterial microbiota. *Nature* **488**: 91–95.
657 **Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-**
658 **Lefert P**. **2013**. Structure and functions of the bacterial microbiota of
659 plants. *Annual Review of Plant Biology* **64**: 807–838.
660 **Butchart SHM, Walpole M, Collen B, van Strien A, Scharlemann JPW,**
661 **Almond REA, Baillie JEM, Bomhard B, Brown C, Bruno J, *et al.* 2010**.
662 Global biodiversity: indicators of recent declines. *Science* **328**: 1164–
663 1168.

664 **Callaway RM, Thelen GC, Rodriguez A, Holben WE**. **2004**. Soil biota and
665 exotic plant invasion. *Nature* **427**: 731–733.
666 **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD,**
667 **Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, *et al.* 2010**.
668 QIIME allows analysis of high-throughput community sequencing data.
669 *Nature Methods* **7**: 335–336.
670 **Colautti RI, Ricciardi A, Grigorovich IA, MacIsaac HJ**. **2004**. Is invasion
671 success explained by the enemy release hypothesis? *Ecology Letters* **7**:
672 721–733.
673 **Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel**
674 **S, Woyke T, North G, Visel A, Partida-Martinez LP, Tringe SG**. **2016**.
675 Plant compartment and biogeography affect microbiome composition in
676 cultivated and native Agave species. *The New Phytologist* **209**: 798–811.
677 **Collins CG, Carey CJ, Aronson EL, Kopp CW, Diez JM**. **2016**. Direct and
678 indirect effects of native range expansion on soil microbial community
679 structure and function. *The Journal of Ecology* **104**: 1271–1283.
680 **Crawford KM, Knight TM**. **2017**. Competition overwhelms the positive plant-soil
681 feedback generated by an invasive plant. *Oecologia* **183**: 211–220.
682 **D'Antonio CM, Malmstrom C, Reynolds SA, Gerlach J**. **2007**. Ecology of
683 invasive non-native species in California grassland. In: Stromberg MR,, In:
684 Corbin JD,, In: D'Antonio CM, eds. California grasslands: ecology and
685 management. Berkeley, California, USA: University of California Press,
686 67–83.
687 **Darwin C**. **1859**. *The Origin of Species*. New York: Random House.
688 **Dawson W, Schrama M**. **2016**. Identifying the role of soil microbes in plant
689 invasions. *The Journal of Ecology* **104**: 1211–1218.
690 **Diffenbaugh NS, Swain DL, Touma D**. **2015**. Anthropogenic warming has
691 increased drought risk in California. *Proceedings of the National Academy*
692 *of Sciences* **112**: 3931–3936.

693 **DiTomaso JM, Healy EA**. **2007**. *Weeds of California and other western states*.
694 Oakland, CA: University of California Department of Agriculture and
695 Natural Resources.
696 **DiTomaso JM, Kyser GB, Pitcairn MJ**. **2006**. *Yellow starthistle management*
697 *guide*. Berkeley, CA: California Invasive Plant Council.
698 **Dlugosch KM, Alice Cang F, Barker BS, Andonian K, Swope SM, Rieseberg**
699 **LH**. **2015**. Evolution of invasiveness through increased resource use in a
700 vacant niche. *Nature Plants* **1**: 15066.
701 **Edgar RC**. **2010**. Search and clustering orders of magnitude faster than BLAST.
702 *Bioinformatics* **26**: 2460–2461.
703 **Edgar RC**. **2013**. UPARSE: highly accurate OTU sequences from microbial
704 amplicon reads. *Nature Methods* **10**: 996–998.
705 **Edgar RC, Flyvbjerg H**. **2015**. Error filtering, pair assembly and error correction
706 for next-generation sequencing reads. *Bioinformatics* **31**: 3476–3482.
707 **Ellis EC, Antill EC, Kreft H**. **2012**. All is not loss: plant biodiversity in the
708 anthropocene. *PLoS One* **7**: e30535.
709 **Engelkes T, Morriën E, Verhoeven KJF, Bezemer TM, Biere A, Harvey JA,**
710 **McIntyre LM, Tamis WLM, van der Putten WH**. **2008**. Successful range-
711 expanding plants experience less above-ground and below-ground enemy
712 impact. *Nature* **456**: 946–948.
713 **Eriksen RL, Desronvil T, Hierro JL, Kesseli R**. **2012**. Morphological
714 differentiation in a common garden experiment among native and non-
715 native specimens of the invasive weed yellow starthistle (Centaurea
716 solstitialis). *Biological Invasions* **14**: 1459–1467.
717 **Essl F, Dullinger S, Rabitsch W, Hulme PE, Hülber K, Jarošík V, Kleinbauer**
718 **I, Krausmann F, Kühn I, Nentwig W, *et al.* 2011**. Socioeconomic legacy
719 yields an invasion debt. *Proceedings of the National Academy of Sciences*
720 **108**: 203–207.
721 **Faillace CA, Lorusso NS, Duffy S**. **2017**. Overlooking the smallest matter:
722 viruses impact biological invasions. *Ecology Letters* **EarlyView**.

Felker-Quinn E, Schweitzer JA, Bailey JK. **2013**. Meta-analysis reveals evolution in invasive plant species but little support for Evolution of Increased Competitive Ability (EICA). *Ecology and Evolution* **3**: 739–751.

Fierer N, Jackson RB. **2006**. The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences* **103**: 626–631.

Finkel OM, Burch AY, Lindow SE, Post AF, Belkin S. **2011**. Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree. *Applied and Environmental Microbiology* **77**: 7647–7655.

Finkel OM, Delmont TO, Post AF, Belkin S. **2016**. Metagenomic signatures of bacterial adaptation to life in the phyllosphere of a salt-secreting desert tree. *Applied and Environmental Microbiology* **82**: 2854–2861.

Fonseca-García C, Coleman-Derr D, Garrido E, Visel A, Tringe SG, Partida-Martínez LP. **2016**. The cacti microbiome: interplay between habitat-filtering and host-specificity. *Frontiers in Microbiology* **7**: 150.

Garbeva P, van Veen JA, van Elsas JD. **2004**. Microbial diversity in soil: selection microbial populations by plant and soil type and implications for disease suppressiveness. *Annual Review of Phytopathology* **42**: 243–270.

Gerlach JD. **1997**. How the West was lost: reconstructing the invasion dynamics of yellow starthistle and other plant invaders of western rangelands and natural areas. *California Exotic Pest Plant Council Symposium Proceedings* **3**: 67–72.

Gibbons SM, Lekberg Y, Mummey DL, Sangwan N, Ramsey PW, Gilbert JA, Shade A. **2017**. Invasive plants rapidly reshape soil properties in a grassland ecosystem. *mSystems* **2**: e00178–16.

Griffin D, Anchukaitis KJ. **2014**. How unusual is the 2012–2014 California drought? *Geophysical Research Letters* **41**: 2014GL062433.

Gundale MJ, Almeida JP, Wallander H, Wardle DA, Kardol P, Nilsson M-C, Fajardo A, Pauchard A, Peltzer DA, Ruotsalainen S, *et al.* 2016. Differences in endophyte communities of introduced trees depend on the

754 phylogenetic relatedness of the receiving forest. *The Journal of Ecology*
755 **104**: 1219–1232.
756 **Gundale MJ, Kardol P, Nilsson M-C, Nilsson U, Lucas RW, Wardle DA**. **2014**.
757 Interactions with soil biota shift from negative to positive when a tree
758 species is moved outside its native range. *The New Phytologist* **202**: 415–
759 421.
760 **Heiser CB Jr, Whitaker TW**. **1948**. Chromosome number, polyploidy, and
761 growth habit in California weeds. *American Journal of Botany* **35**: 179–
762 186.
763 **Herrera Paredes S, Lebeis SL**. **2016**. Giving back to the community: microbial
764 mechanisms of plant–soil interactions. *Functional Ecology* **30**: 1043–1052.
765 **Hierro JL, Khetsuriani L, Andonian K, Eren Ö, Villarreal D, Janoian G,**
766 **Reinhart KO, Callaway RM**. **2016**. The importance of factors controlling
767 species abundance and distribution varies in native and non-native
768 ranges. *Ecography* **EarlyView**.
769 **Hierro JL, Lortie CJ, Villarreal D, Estanga-Mollica ME, Callaway RM**. **2011**.
770 Resistance to Centaurea solstitialis invasion from annual and perennial
771 grasses in California and Argentina. *Biological Invasions* **13**: 2249–2259.
772 **Hooper DU, Dukes JS**. **2010**. Functional composition controls invasion success
773 in a California serpentine grassland. *The Journal of Ecology* **98**: 764–777.
774 **Hulvey KB, Zavaleta ES**. **2012**. Abundance declines of a native forb have
775 nonlinear impacts on grassland invasion resistance. *Ecology* **93**: 378–388.
776 **Keane R, Crawley MJ**. **2002**. Exotic plant invasions and the enemy release
777 hypothesis. *Trends in Ecology & Evolution* **17**: 164–170.
778 **Kulmatiski A, Beard KH, Stevens JR, Cobbold SM**. **2008**. Plant–soil
779 feedbacks: a meta-analytical review. *Ecology Letters* **11**: 980–992.
780 **Lebeis SL, Paredes SH, Lundberg DS, Breakfield N, Gehring J, McDonald**
781 **M, Malfatti S, Glavina del Rio T, Jones CD, Tringe SG, *et al.* 2015**.
782 Salicylic acid modulates colonization of the root microbiome by specific
783 bacterial taxa. *Science* **349**: 860–864.

784 **Leff JW, Lynch RC, Kane NC, Fierer N**. **2016**. Plant domestication and the
785 assembly of bacterial and fungal communities associated with strains of
786 the common sunflower, Helianthus annuus. *The New Phytologist*.
787 **Lonsdale WM**. **1999**. Global patterns of plant invasions and the concept of
788 invasibility. *Ecology* **80**: 1522–1536.
789 **Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S,**
790 **Tremblay J, Engelbrektson A, Kunin V, del Rio TG, *et al.* 2012**.
791 Defining the core Arabidopsis thaliana root microbiome. *Nature* **488**: 86–
792 90.
793 **Lundberg DS, Yourstone S, Mieczkowski P, Jones CD, Dangl JL**. **2013**.
794 Practical innovations for high-throughput amplicon sequencing. *Nature*
795 *Methods* **10**: 999–1002.
796 **Maron JL, Klironomos J, Waller L, Callaway RM**. **2014**. Invasive plants escape
797 from suppressive soil biota at regional scales. *The Journal of Ecology* **102**:
798 19–27.
799 **Maron JL, Vilà M, Arnason J**. **2004**. Loss of enemy resistance among
800 introduced populations of St. John's Wort (Hypericum perforatum).
801 *Ecology* **85**: 3243–3253.
802 **McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A,**
803 **Andersen GL, Knight R, Hugenholtz P**. **2012**. An improved Greengenes
804 taxonomy with explicit ranks for ecological and evolutionary analyses of
805 bacteria and archaea. *The ISME Journal* **6**: 610–618.
806 **McQuitty LL**. **1966**. Similarity Analysis by Reciprocal Pairs for Discrete and
807 Continuous Data. *Educational and psychological measurement* **26**: 825–
808 831.
809 **Mitchell CE, Agrawal AA, Bever JD, Gilbert GS, Hufbauer RA, Klironomos**
810 **JN, Maron JL, Morris WF, Parker IM, Power AG, *et al.* 2006**. Biotic
811 interactions and plant invasions. *Ecology Letters* **9**: 726–740.
812 **Müller G, Horstmeyer L, Rönneburg T, van Kleunen M, Dawson W**. **2016**.
813 Alien and native plant establishment in grassland communities is more

strongly affected by disturbance than above- and below-ground enemies. *The Journal of Ecology* **104**: 1233–1242.

Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, *et al.* 2013. Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews* **77**: 342–356.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, *et al.* 2016. vegan: Community Ecology Package.

Öztürk M, Martin E, Dinç M, Duran A, Özdemir A, Çetin Ö. **2009**. A cytogenetical study on some plants taxa in Nizip region (Aksaray, Turkey). *Turkish Journal of Biology* **33**: 35–44.

Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL, Buckler ES, Ley RE. **2013**. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences* **110**: 6548–6553.

Petermann JS, Fergus AJF, Turnbull LA, Schmid B. **2008**. Janzen-Connell effects are widespread and strong enough to maintain diversity in grasslands. *Ecology* **89**: 2399–2406.

Pitcairn M, Schoenig S, Yacoub R, Gendron J. **2006**. Yellow starthistle continues its spread in California. *California Agriculture* **60**: 83–90.

van der Putten WH, Bardgett RD, Bever JD, Bezemer TM, Casper BB, Fukami T, Kardol P, Klironomos JN, Kulmatiski A, Schweitzer JA, *et al.* 2013. Plant–soil feedbacks: the past, the present and future challenges. *The Journal of Ecology* **101**: 265–276.

van der Putten WH, Bradford MA, Pernilla Brinkman E, van de Voorde TFJ, Veen GF. **2016**. Where, when and how plant–soil feedback matters in a changing world. *Functional Ecology* **30**: 1109–1121.

R Core Team. **2015**. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

844 **Reinhart KO, Callaway RM**. **2006**. Soil biota and invasive plants. *The New*
845 *Phytologist* **170**: 445–457.
846 **Reinhart KO, Packer A, Van der Putten WH, Clay K**. **2003**. Plant–soil biota
847 interactions and spatial distribution of black cherry in its native and
848 invasive ranges. *Ecology letters* **6**: 1046–1050.
849 **Reisberg EE, Hildebrandt U, Riederer M, Hentschel U**. **2013**. Distinct
850 phyllosphere bacterial communities on Arabidopsis wax mutant leaves.
851 *PLoS One* **8**: e78613.
852 **Rout ME, Callaway RM**. **2012**. Interactions between exotic invasive plants and
853 soil microbes in the rhizosphere suggest that 'everything is not
854 everywhere'. *Annals of Botany* **110**: 213–222.
855 **Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA,**
856 **Baughman S, Cabin RJ, Cohen JE, Ellstrand NC, *et al.* 2001**. The
857 population biology of invasive species. *Annual Review of Ecology and*
858 *Systematics* **32**: 305–332.
859 **Salvaudon L, Giraud T, Shykoff JA**. **2008/4**. Genetic diversity in natural
860 populations: a fundamental component of plant–microbe interactions.
861 *Current Opinion in Plant Biology* **11**: 135–143.
862 **Schnitzer SA, Klironomos JN, HilleRisLambers J, Kinkel LL, Reich PB, Xiao**
863 **K, Rillig MC, Sikes BA, Callaway RM, Mangan SA, *et al.* 2011**. Soil
864 microbes drive the classic plant diversity–productivity pattern. *Ecology* **92**:
865 296–303.
866 **Schrama M, Bardgett RD**. **2016**. Grassland invasibility varies with drought
867 effects on soil functioning. *The Journal of Ecology* **104**: 1250–1258.
868 **Seabloom EW, Harpole WS, Reichman OJ, Tilman D**. **2003**. Invasion,
869 competitive dominance, and resource use by exotic and native California
870 grassland species. *Proceedings of the National Academy of Sciences* **100**:
871 13384–13389.
872 **Shade A, Handelsman J**. **2012**. Beyond the Venn diagram: the hunt for a core
873 microbiome. *Environmental Microbiology* **14**: 4–12.

874 **Shannon CE**. **1948**. A Mathematical Theory of Communication. *Bell System*
875 *Technical Journal* **27**: 379–423.
876 **Swope SM, Parker IM**. **2012**. Complex interactions among biocontrol agents,
877 pollinators, and an invasive weed: a structural equation modeling
878 approach. *Ecological Applications* **22**: 2122–2134.
879 **terHorst CP, Lennon JT, Lau JA**. **2014**. The relative importance of rapid
880 evolution for plant-microbe interactions depends on ecological context.
881 *Proceedings of the Royal Society: Biological Sciences* **281**: 20140028.
882 **terHorst CP, Zee PC**. **2016**. Eco-evolutionary dynamics in plant–soil feedbacks.
883 *Functional Ecology* **30**: 1062–1072.
884 **Tiffin P, Moeller DA**. **2006**. Molecular evolution of plant immune system genes.
885 *Trends in Genetics* **22**: 662–670.
886 **Torchin ME, Mitchell CE**. **2004**. Parasites, pathogens, and invasions by plants
887 and animals. *Frontiers in Ecology and the Environment* **2**: 183–190.
888 **Uygur S, Smith L, Uygur FN, Cristofaro M, Balciunas J**. **2004**. Population
889 densities of yellow starthistle (Centaurea solstitialis) in Turkey. *Weed*
890 *Science* **52**: 746–753.
891 **Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A**. **2015**.
892 The importance of the microbiome of the plant holobiont. *The New*
893 *Phytologist* **206**: 1196–1206.
894 **Venables WN, Ripley BD**. **2002**. Random and Mixed Effects. Statistics and
895 Computing. Modern Applied Statistics with S. New York: Springer, 271–
896 300.
897 **Wagg C, Bender SF, Widmer F, van der Heijden MGA**. **2014**. Soil biodiversity
898 and soil community composition determine ecosystem multifunctionality.
899 *Proceedings of the National Academy of Sciences* **111**: 5266–5270.
900 **Wickham H**. **2009**. *ggplot2: elegant graphics for data analysis*. New York:
901 Springer-Verlag.
902 **Widmer TL, Guermache F, Dolgovskaia MY, Reznik SY**. **2007**. Enhanced
903 growth and seed properties in introduced vs. native populations of yellow
904 starthistle (Centaurea solstitialis). *Weed Science* **55**: 465–473.

Williamson M. **1996**. *Biological Invasions*. London: Chapman & Hall.

Zavaleta ES, Hulvey KB. **2004**. Realistic species losses disproportionately reduce grassland resistance to biological invaders. *Science* **306**: 1175–1177.

Fig. 1. The distribution (gray) of yellow starthistle and sampling sites for this study. Maps detail the native range in Eurasia (a) and the invasion of western North America (b). Previous work has indicated that western Europe is the source for the severe invasion of California, USA (both in dark shading; Barker et al. 2017). Sampling included seven locations in California (b, filled circles), six locations in western Europe and an additional two locations in eastern Europe (a, open circles).

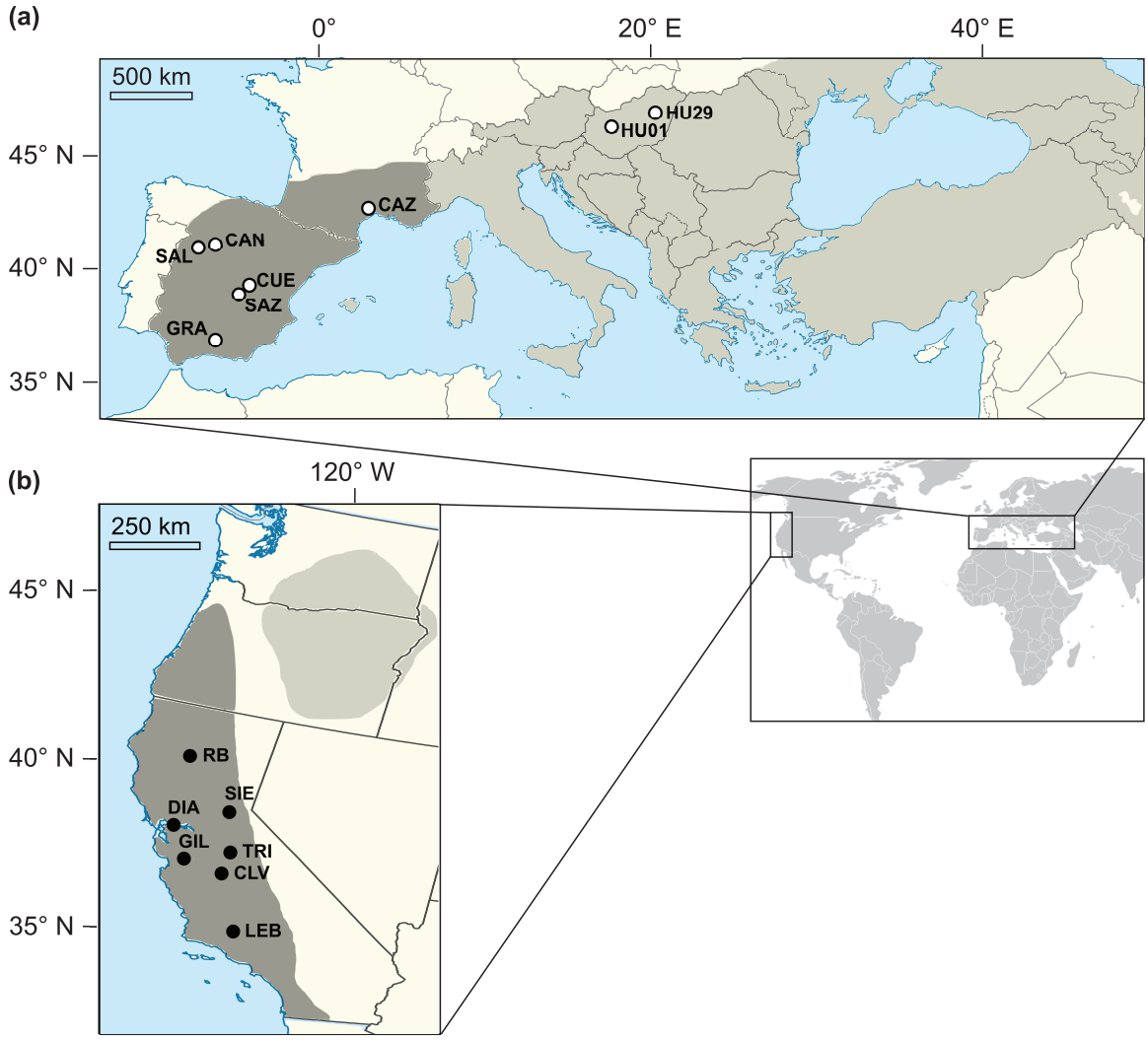

Fig. 2. Distribution of read counts for bulk samples from all four compartments (native and invading population samples combined), as well as control (blank) samples.

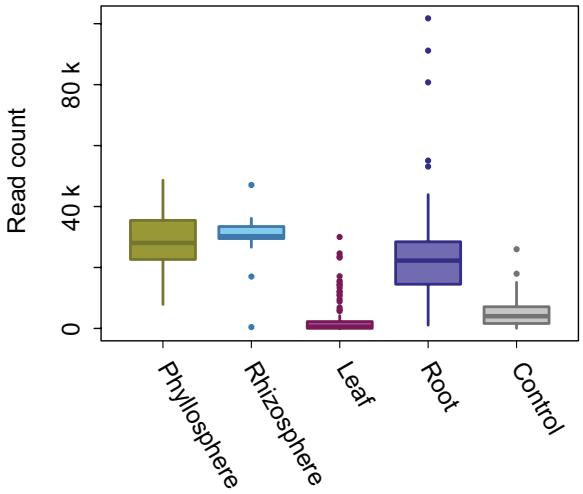

Fig. 3. NMDS plots of bacterial OTU composition in phyllosphere (green), rhizosphere (light blue), and whole root (dark blue) samples for native (open symbols) and invading (closed symbols) populations. Plotted are a) bulk samples for each population, showing overall separation by compartment and by range within compartment, and b) individual whole root samples within native and invading populations. Ellipses indicate 95% confidence intervals for samples grouped by range (native range: dashed line; invaded range: solid line).

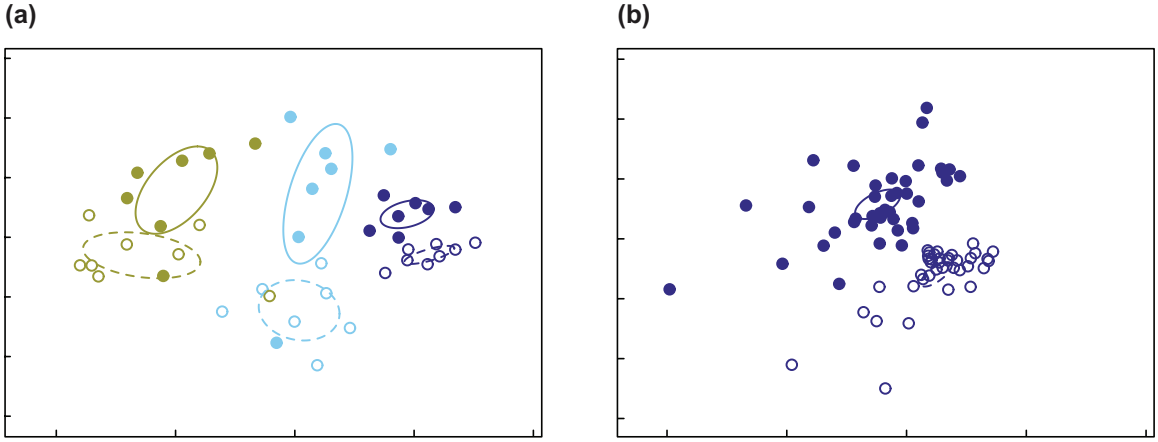

Fig 4. Relative abundance of (proportion of reads mapping to) dominant phyla in (a) phyllosphere, (b) rhizosphere, and (c) whole root bulk samples from native and invaded ranges.

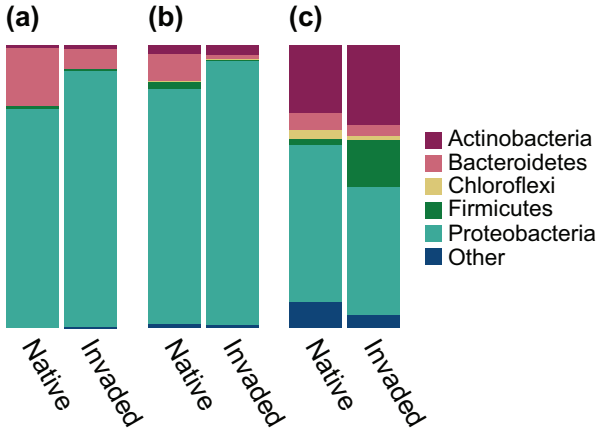

Fig. 5. Comparison of diversity (H') between samples from native and invaded ranges for (a) phyllosphere, (b) rhizosphere, (c) bulk whole roots by population, (d) individual whole roots. Significance levels from Kruskal-Wallis tests indicated with asterisks: *P < 0.05, $^{(*)}P$ < 0.1.

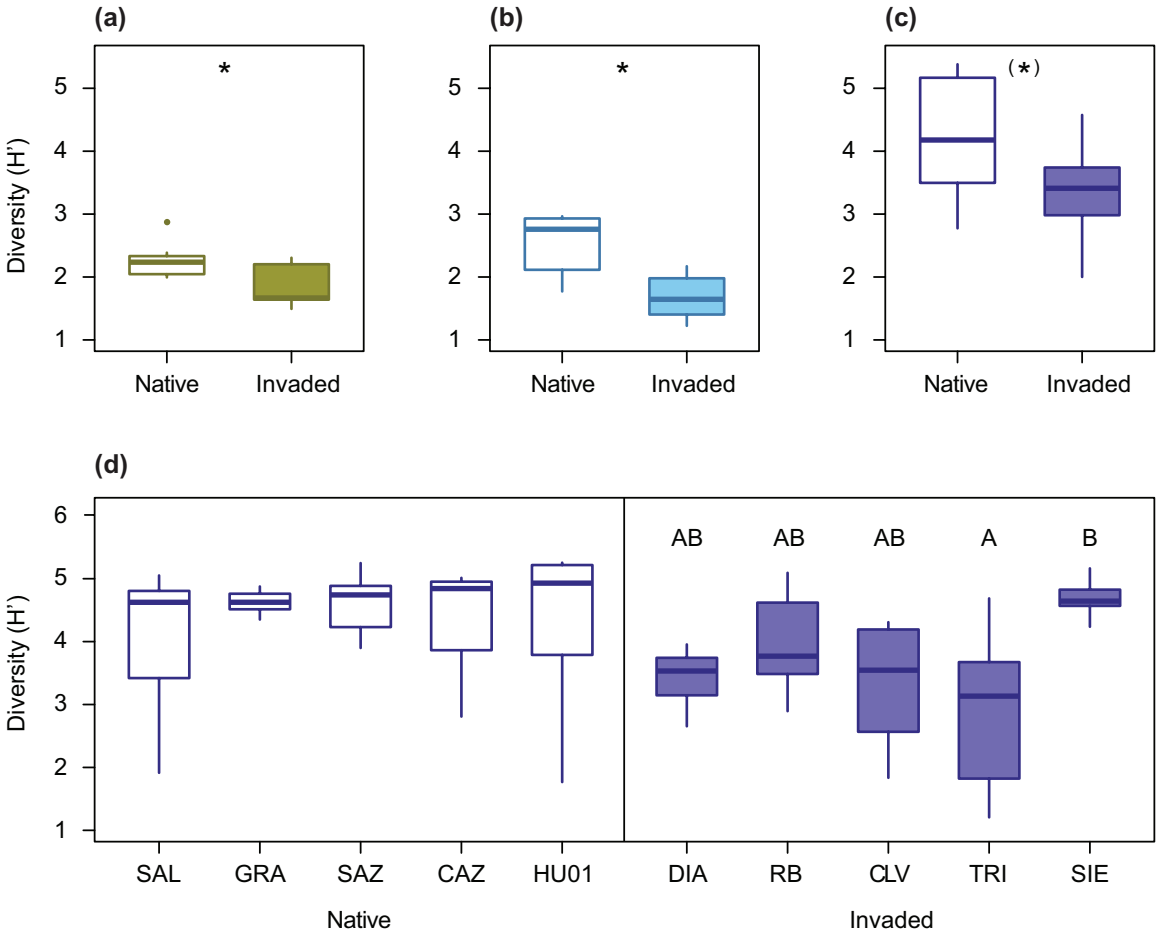

Fig 6. Venn diagrams indicating the number of OTUs shared between native and invaded ranges, and unique to each range, for whole root samples. Shown are OTUs from bulk and individual samples combined, for (a) all OTUs, and for the dominant phyla Proteobacteria (b) and Actinobacteria (c).

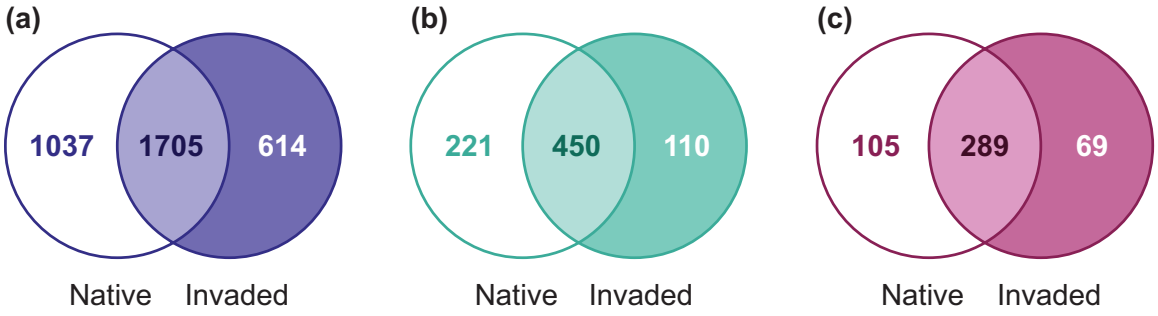

Fig 7. Bacterial diversity (H') in bulk whole root samples for each population as a function of the genetic diversity among plants in those populations (calculated as the average proportion of pairwise nucleotide differences between alleles (**π**) at variable sites across the genome; from Barker *et al.*, 2017). Lines show significant positive relationships (linear model: $P < 0.02$) between microbial and plant diversity in both the native range (open symbols, dashed line) and invaded range (closed symbols, solid line).