



COREMIC: a web-tool to search for a niche associated CORE MICrobiome

Richard R. Rodrigues^{1,2}, Nyle C. Rodgers³, Xiaowei Wu⁴ and Mark A. Williams^{1,5}

¹Interdisciplinary Ph.D. Program in Genetics, Bioinformatics, and Computational Biology, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America

²Department of Pharmaceutical Sciences, Oregon State University, Corvallis, OR, United States of America

³Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America

⁴Department of Statistics, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America

⁵School of Plant and Environmental Sciences, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, United States of America

ABSTRACT

Microbial diversity on earth is extraordinary, and soils alone harbor thousands of species per gram of soil. Understanding how this diversity is sorted and selected into habitat niches is a major focus of ecology and biotechnology, but remains only vaguely understood. A systems-biology approach was used to mine information from databases to show how it can be used to answer questions related to the core microbiome of habitat-microbe relationships. By making use of the burgeoning growth of information from databases, our tool “COREMIC” meets a great need in the search for understanding niche partitioning and habitat-function relationships. The work is unique, furthermore, because it provides a user-friendly statistically robust web-tool (<http://coremic2.appspot.com> or <http://core-mic.com>), developed using Google App Engine, to help in the process of database mining to identify the “core microbiome” associated with a given habitat. A case study is presented using data from 31 switchgrass rhizosphere community habitats across a diverse set of soil and sampling environments. The methodology utilizes an outgroup of 28 non-switchgrass (other grasses and forbs) to identify a core switchgrass microbiome. Even across a diverse set of soils (five environments), and conservative statistical criteria (presence in more than 90% samples and FDR q -val <0.05% for Fisher’s exact test) a core set of bacteria associated with switchgrass was observed. These included, among others, closely related taxa from *Lysobacter spp.*, *Mesorhizobium spp.*, and *Chitinophagaceae*. These bacteria have been shown to have functions related to the production of bacterial and fungal antibiotics and plant growth promotion. COREMIC can be used as a hypothesis generating or confirmatory tool that shows great potential for identifying taxa that may be important to the functioning of a habitat (e.g. host plant). The case study, in conclusion, shows that COREMIC can identify key habitat-specific microbes across diverse samples, using currently available databases and a unique freely available software.

Submitted 21 September 2017

Accepted 30 January 2018

Published 15 February 2018

Corresponding author

Richard R. Rodrigues, richrr@vt.edu

Academic editor

Thomas Rattei

Additional Information and
Declarations can be found on
page 11

DOI 10.7717/peerj.4395

 Copyright

2018 Rodrigues et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Bioinformatics, Ecology, Microbiology

Keywords Root-zone, Web-tool, Meta-analysis, Database, Rhizosphere, Data mining, App, Software, Microbiome

INTRODUCTION

Microbial diversity on earth is extraordinary, and soils alone harbor thousands of species per gram (*Hughes et al., 2001*). Understanding how this diversity is sorted and selected into habitat niches is a major focus of ecology and biotechnology, but remains only vaguely understood. The advent of next-generation sequencing technologies now allow for the potential to make great leaps in the study of microbe-habitat relationships of highly diverse microbial communities and environments. The identity and functions of this overwhelming multitude of microbes are in the beginning stages of being described, and are already providing insights into microbial impacts on plant and animal health (*Berg, 2009; Evans & Schwarz, 2011; Clemente et al., 2012*). Making use of the overwhelming amount of information on microbial taxa and habitats has enormous potential for use to further understand microbial-habitat relationships. Thus, the advent of new methods and approaches to utilize this data and describe microbiomes will benefit microbial ecology and biotechnology.

Though variations exist, a core microbiome can be defined, conceptually, using Venn diagrams, where over-lapping circles and non-overlapping areas of circles represent shared and non-shared members of a habitat, respectively (*Shade & Handelsman, 2012*). Typically, microbiomes identified in this manner are not statistically evaluated, or by nature, seek to answer specific hypothesis that are specific to an experiment. For example, studies often identify microbes associated with different plant growth stages, species, cultivars, and locations but rarely, if at all, mine databases or perform meta-analysis to statistically identify microbiomes across studies and experimental conditions (*Chaudhary et al., 2012; Liang et al., 2012; Mao et al., 2013; Mao et al., 2014; Hargreaves, Williams & Hofmockel, 2015; Rodrigues et al., 2015; Jesus et al., 2016; Rodrigues et al., 2017*). Describing differences due to treatment or habitat conditions are informative in their own right, however, extending this framework to include an easy to use, and statistically robust tool to help in the mining of data from underutilized and burgeoning databases (e.g., the National Center for Biotechnology Information (NCBI), Ribosomal Database Project) can help transform the ecological study of microbes in their natural environment. Using the vast and growing databases of organism and habitat metadata will allow for both the testing and development of hypotheses associated with habitat-microbe relationships that were not formerly possible.

To address the challenges described above, we developed COREMIC—a novel, easy to use, and freely available web tool to identify the “core microbiome”, of any well-defined habitat (e.g., plant root-zone) or niche (*Shade & Handelsman, 2012*). This straightforward approach is a novel and powerful way to complement existing analysis (e.g., indicator species analysis (ISA) *Dufrene & Legendre, 1997*) by allowing for the use of data that is now overflowing among freely available databases. It seeks to determine the core set of microbes (core microbiome) that are explicitly associated with a host system or habitat. The ability to identify core microbiomes at this scale has great potential to describe host-microbe interactions and habitat preferences of microbes.

A meta-analysis based case study was performed, combining diverse sequencing datasets derived from NCBI, to test for the occurrence of a core microbiome in the rhizosphere (root-zone) of switchgrass. Switchgrass is a US-native, perennial grass studied by many researchers, and thus has a growing database to mine for genetic information. Its widespread study is likely a result of its bioenergy potential, and the capacity of the grass to grow on marginal lands not dedicated to crops. Studies have identified different bacteria found in the root-zones of switchgrass ([Jesus et al., 2010](#); [Mao, Yannarell & Mackie, 2011](#); [Chaudhary et al., 2012](#); [Liang et al., 2012](#); [Mao et al., 2013](#); [Bahulikar et al., 2014](#); [Mao et al., 2014](#); [Werling et al., 2014](#); [Hargreaves, Williams & Hofmockel, 2015](#); [Jesus et al., 2016](#); [Rodrigues et al., 2017](#)); however, there has been no integrative study of different datasets identifying the core microbiome in switchgrass rhizospheres. It was thus proposed to identify host-habitat relationships as a proof of concept for a core microbiome. In this paper we utilize a plant host to define a habitat, but theoretically any habitat and associated organisms could make use of COREMIC and its approach to identify a core microbiome.

MATERIAL AND METHODS

Datasets used in the study

A diverse set of data composed of 61 samples from two different published datasets and collected from multiple locations ([Jesus et al., 2016](#); [Rodrigues et al., 2017](#)) were used for this study. Data were obtained from the NCBI and selected based on the availability of the raw (16S rRNA) sequence data of root-zone bacteria from switchgrass and that for an out-group of reference (native and/or other grasses) plants.

The dataset “Jesus, 2016” ([Jesus et al., 2016](#)), [PRJEB6704](#), compared the rhizosphere soil microbial communities associated with restored prairie with three grass crops, namely corn, switchgrass, and mixed prairie grasses. The grasses were grown in fields of Michigan and Wisconsin and were harvested after two and ten years. The V6–V8 region of the 16S rRNA gene was amplified and sequenced using the Roche 454 pyrosequencing. In our study, we used a total of 43 samples (three each from corn, switchgrass, mixed grasses (two years only), and restored prairie grasses grown in Wisconsin and Michigan, and sampled after two and 10 years. Switchgrass grown in Michigan, composed of 4 samples, were collected following 10 years of plant growth.

The dataset “Rodrigues, 2017” ([Rodrigues et al., 2017](#)), [PRJNA320123](#), compared the root-zone soil microbial communities associated with switchgrass cultivars: “Alamo” and “Dacotah”. The switchgrass were grown in the greenhouse using soil derived from plots growing Switchgrass (>7 years) near Blacksburg, VA. Switchgrass rhizosphere bacteria were sampled at three different growth stages. The V3–V4 region of the 16S rRNA gene was amplified and sequenced using Illumina MiSeq sequencing. In our study, we used a total of 18 switchgrass samples for Alamo (A) and Dacotah (D) from stages V2 and E3 (4 AV2, 4 DV2, 5 AE3, 5 DE3 = 18).

Overall, these datasets served as a diverse resource (relevant differences are summarized in [Fig. 1](#)) to compare the root-zone bacteria and identify core-bacteria associated with switchgrass.

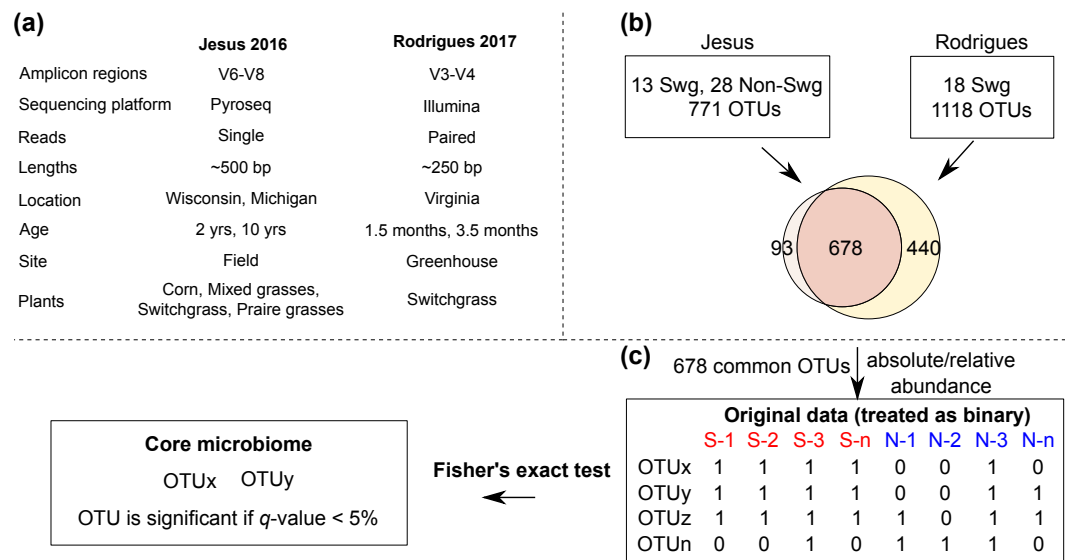


Figure 1 The COREMIC approach. (A) Differences between the Jesus, 2016 and Rodrigues, 2017 datasets. (B) Combining the two OTU tables of the two datasets and (C) the methodology used to identify core microbiome. Switchgrass and other grasses are indicated by “Swg” and “Non-Swg,” respectively.

Full-size [DOI: 10.7717/peerj.4395/fig-1](https://doi.org/10.7717/peerj.4395/fig-1)

Sequence data analysis and picking of Operational Taxonomic Units (OTU)

For the Rodrigues, 2017 dataset, the OTU table was obtained from previously performed analysis (Rodrigues *et al.*, 2017). For the Jesus, 2016 dataset, quality score (25) and read lengths (150) thresholds were enforced using cutadapt (1.8.1) (Martin, 2011) and an open reference OTU picking (enable_rev_strand_match True) was performed in QIIME v1.8.0 (Caporaso *et al.*, 2010), as previously described (Rodrigues *et al.*, 2015; Rodrigues *et al.*, 2017), to allow comparison with the other dataset. Briefly, uclust (Edgar, 2010) was used to cluster reads into OTUs (97% sequence similarity) and assign taxonomy against the Greengenes reference database version 13.8 (DeSantis *et al.*, 2006; McDonald *et al.*, 2012). Two samples from the Jesus, 2016 dataset were removed from downstream analysis due to very few sequences assigned to OTUs.

Combining two datasets

For each dataset, sequences assigned to OTUs with identical taxonomy were collapsed (summed within each sample) to create a single row for that taxonomy. This method of taxonomic summary is commonly used (e.g., QIIME) and beneficial for working with diverse datasets. Finally, the OTU tables from the two datasets were merged based on identical taxonomy (row-ids). The common OTUs (678 taxa) from the two datasets were selected, converted to biom format and used for further analyses (Fig. 1). The data table was rarefied using a sequence threshold of 1150, and the beta diversity was calculated using Bray-Curtis (Beals, 1984) distance and visualized using Principal Coordinate Analysis (Gower, 2005). Multivariate data analysis methods of MRPP (Mielke, 1984), Permanova (Anderson, 2001) and ANOSIM (Clarke, 1993) were used to identify whether the plant type

(switchgrass versus non-switchgrass) were associated with different bacterial communities. Mann–Whitney-U test from QIIME’s `group_significance.py` was used to identify OTUs that had differential abundance ($FDR < 0.05$) in switchgrass and non-switchgrass samples.

Core microbiome analysis

To find the set of core OTUs, the samples in the combined OTU table (original data) were first divided into the interest group samples (switchgrass) and out-group samples. The abundance values for each OTU in each sample are then converted to binary (present/absent) values based on whether they are zero or nonzero. For each OTU a one-tailed Fisher’s Exact Test was used to calculate a p -value testing whether an OTU was present in a significantly higher portion in the interest in-group (Switchgrass) compared to the out-group samples (numerous other grass species).

These p -values were corrected for multiple-testing using Benjamini Hochberg. The OTUs with a q -value < 0.05 were then selected to only the OTUs that are present in at least 90% of the interest group samples. Uninformative OTUs (e.g., `k_Bacteria;p_c_o_f_g_s_`) were filtered out and the remaining OTUs were candidates for the core microbiome.

Submitting data to COREMIC

COREMIC needs data and mapping files. Data file(s) provide the abundance (counts) for each OTU (taxonomy) per sample. To allow users to directly use their output of OTU picking methods, COREMIC requires the data file(s) in BIOM format with the taxonomic assignments for each OTU under the “taxonomy” key of the meta-data in the data file(s). Rows with the same taxonomic label will be combined, with their values summed per sample. If multiple data files are uploaded, they will be combined using taxonomic labels, removing any taxonomies that are not present in all the data files.

The mapping file is a tab-delimited file where each sample is represented as a row and contains group (treatment, location, age, etc.) information as columns about the samples from the data file(s). The help section of COREMIC (<http://coremic2.appspot.com/help> or <http://core-mic.com/help>) offers details about the input files and provides sample data and group files. The supplementary video (<https://oregonstate.box.com/v/coremic-tutorial>) offers a tutorial for submitting data to COREMIC.

Implementation of COREMIC

COREMIC receives the input files and parameters from web-tool. The data file is collapsed (sum the sequences of OTUs with identical taxonomy per sample) to have unique taxonomic row ids. Multiple data files, if any, are merged into a single OTU table based on common taxonomic labels. The user preferred normalization is applied and significance is calculated for each OTU followed by multiple testing. Core microbes are identified as those that pass the user-specified criteria (adjusted p -value, presence in interest and out-group). Finally, the results and output files are emailed to the user.

COREMIC and the datasets are available at <http://coremic2.appspot.com> or <http://core-mic.com>. Its code is available on github (<https://github.com/richrr/coremicro>). The web-tool was developed in Python 2.7, and is hosted on Google App Engine. Other requirements include GoogleAppEnginePipeline 1.9.22.1, pyqi 0.3.1, requests 2.10.0, requests-toolbelt

0.6.2, mailjet-rest 1.2.2, biom-format 1.1.2, ete3 3.0.0 (for tree generation—see below for details), webapp2 2.5.2, numpy 1.6.1, matplotlib 1.2.0, jinja2 2.6, ssl 2.7.11. COREMIC is accessible via any internet connected browser and emails the results to the user. The processing times with the default settings after uploading the data are provided in [Table S1](#).

A custom python script generates a phylogenetic tree using the taxonomic labels for each OTU displaying the relationship between the core OTUs obtained from the group of interest and the out-group. This tree is generated using the ete3 3.0.0 library.

RESULTS

After quality filtering, a total of 319,821 reads were obtained from the Jesus, 2016 dataset (mean 461.45 and std. dev. 69.34). Two samples with very few (48 and 75) counts were removed; each of the remaining samples had more than 1,150 sequences assigned to OTUs. A total of 1,581,679 high quality reads were used from the Rodrigues, 2017 dataset (mean 333.02 and std. dev. 57.73). The number of OTUs in the Jesus, 2016 and Rodrigues, 2017 datasets was 771 and 1118, respectively. The combined dataset had 678 OTUs, 31 switchgrass and 28 non-switchgrass (other grasses) samples. The details of sequences assigned to each sample are provided in [Table S2](#). The bacterial communities in switchgrass and grasses from the combined dataset were significantly different ([Table S3](#) shows Mann–Whitney-U test $FDR < 0.05$ and Permanova, MRPP, and ANOSIM p -values < 0.01). These differences were apparent despite significant difference across datasets ([Table S3](#) Permanova, MRPP, and ANOSIM p -values < 0.01), as can be observed using the PCoA plot using the Bray-Curtis dissimilarity metric ([Fig. 2](#)); which could be the result, for example, of the heterogeneity of the data set related to climate, soil type-condition, growth conditions, plant age and technological differences. In this regard, at the phylum level, Mann Whitney test identified Bacteroidetes and Verrucomicrobia had significantly greater (p -value < 0.05) relative abundance in switchgrass, whereas, Gemmatimonadetes were more abundant in other grasses ([Fig. S1](#)).

We used a very conservative criterion of $>90\%$ threshold i.e., an OTU has to be present in at least 90% of switchgrass samples and observed five OTUs with FDR q -values < 0.05 ([Table 1](#)). The relative abundance and a phylogenetic tree exhibiting their relationship with the core-OTUs from the non-switchgrass samples is shown in [Figs. S2](#) and [S3](#), respectively. Despite the enormous variability across the many different sampling locations, there is support for the occurrence of a core microbiome in the root-zone of switchgrass.

DISCUSSION

The case study showed how COREMIC can identify key habitat-specific microbes across diverse samples, using currently available databases and a unique freely available software. The core set of bacteria associated with switchgrass included, among others, closely related taxa from *Lysobacter spp.*, *Mesorhizobium spp.*, and *Chitinophagaceae*. The functional relevance of these bacteria related to switchgrass is currently unknown, but it is notable that these bacteria have been shown to produce bacterial and fungal antibiotics and promote

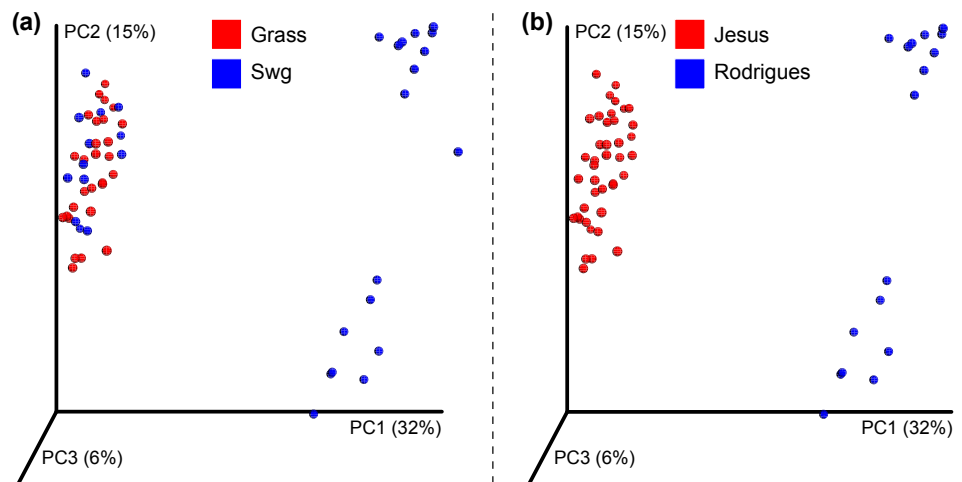


Figure 2 Beta-diversity of the combined dataset. PCoA plot showing Bray-Curtis dissimilarities for bacterial communities at the OTU level. (A) Differences in other grasses (red colored) and switchgrass (blue colored). (B) Differences in the Jesus 2016 (red colored) and Rodrigues 2017 (blue colored) datasets.

Full-size [DOI: 10.7717/peerj.4395/fig-2](https://doi.org/10.7717/peerj.4395/fig-2)

Table 1 Bacterial OTUs associated with switchgrass.

OTU	Present (%)
p_Proteobacteria;c_Gammaproteobacteria;o_Xanthomonadales;f_Xanthomonadaceae;g_Lysobacter;s_	100
p_Planctomycetes;c_Planctomycetia;o_B97;f_g_s_	96.8
p_Bacteroidetes;c_[Saprosirae];o_[Saprosirales];f_Chitinophagaceae	96.8
p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Phyllobacteriaceae;g_Mesorhizobium;s_	90.3
p_Proteobacteria;c_Gammaproteobacteria;o_Legionellales;f_g_s_	90.3

Notes.

The core bacterial OTUs those were significantly (q -value < 0.05) associated with switchgrass, calculated using presence/absence data and present in >90% switchgrass samples.

the growth of plants (Kaneko *et al.*, 2000; Kilic-Ekici & Yuen, 2004; Weir *et al.*, 2004; Islam *et al.*, 2005; Jochum, Osborne & Yuen, 2006; Ji *et al.*, 2008; Park *et al.*, 2008; Nandasena *et al.*, 2009; Yin, 2010; Bailey *et al.*, 2013; Degefu *et al.*, 2013; Guerrouj *et al.*, 2013; Madhaiyan *et al.*, 2015). The analyses from the highly diverse data sets thus provided information that helps to greatly narrow down possibilities and thus set the stage for testing, using controlled studies, how the core microbiota potentially support or antagonize the function of a native grass. This novel toolkit is simple to use and supports use by a broad range of biological scientists, and is particularly relevant to those with expertise in their field but with limited bioinformatics background. Overall, in a dataset derived from a complex and diverse set of habitats and ecosystems, this tool was shown to pinpoint microbiota of the microbiome that might have important functional implications within their habitat or host.

Methodological considerations in the use of COREMIC

COREMIC performs a complementary analysis different from that of existing methods by using presence/absence data. For two groups (A and B) it checks whether (pre-determined percentage of) samples from group A have a non-zero value for the OTU. This allows

scientists to operate without making assumptions about the PCR-based OTU relative abundances. This is considered a potential advantage of the method because it is unknown whether relative abundance of sequence data is representative of true relative differences between communities. It is well accepted that sequencing depth can affect the occurrence of rare OTUs in both relative and presence/absence data. Relative abundance analysis, however, would bias against rare OTUs, whereas presence absence equally treats abundant and rare OTUs since any count >1 (or as per the user chosen threshold) is treated as present. Using OTUs that are present in at least 90% of the interest group samples along with significance testing to define a core microbiome accounts for possibility that rare OTUs might be missed as being called present in some samples. Therefore, our proposed approach is relatively robust to sequencing depths. Further research, in this regard, will be aimed towards investigating other measures of OTU “presence”, namely the extent of exclusivity, consistency, or abundance of the group that is eventually determined to be a core microbiome.

Sampling plots used in this study were located across a range of diverse environments to help create a backdrop of heterogeneity. While this diversity of habitat conditions ignores the potential for microbe-environment interactions that might be important for the plant-microbial relationship, it has the advantage of being a conservative approach with high veracity for defining a core microbiome regardless of habitat heterogeneity. The locations from which samples were grown (Michigan, Wisconsin, Virginia) were treated as independent to help isolate the overall habitat effect of switchgrass ([Werling et al., 2014](#); [Jesus et al., 2016](#)). When the effects of habitat are thought to be habitat specific, researchers can take this into account during the design and analysis using COREMIC.

It is notable that the representation of an outgroup (multiple non-switchgrass species) is an important criteria and choice made by researchers, and is an approach that has both advantages and caveats. By definition, a habitat is defined by its differences from that of other habitats, and therefore the use of the outgroup is an important choice. A counter-argument for the current dataset might argue for exclusion of breeding lines of a cultivated grass (maize) as being unrepresentative of the grass outgroup. In our case, it was thought, *a priori*, that a diverse set of grasses would provide the best comparison; and no compelling argument was found that supported the exclusion of maize from the analysis. An implicit assumption was also made that the taxonomy of plant species (root-zone habitats) play an important role in determining root-zone microbial communities, an approach supported by extensive findings that different grass species associate with different microbial communities ([Kuske et al., 2002](#); [Kennedy et al., 2004](#); [Berendsen, Pieterse & Bakker, 2012](#); [Chaudhary et al., 2012](#); [Turner et al., 2013](#)). So although there is a need for careful consideration of the experimental questions of interest when using COREMIC, this is a common, if not ubiquitous foundation of all experimentation and hypothesis testing. The results provide a statistically valid approach using freely available software to describe and define a core microbiome of switchgrass.

The choice of the outgroup, furthermore, for determining a core microbiome is amenable to choice using deductive reasoning but ultimately limited by available data. This issue almost certainly limits inclusion of many functionally important rhizosphere microbes

that could affect the growth of switchgrass. In this study, the proof of concept utilized a conservative approach to highlight the methodology across a diversity of geographies, soil types, and plant ages. The COREMIC tool as well as the multiple methods for defining a core microbiome (e.g., QIIME [Caporaso et al., 2010](#)), ISA ([Dufrene & Legendre, 1997](#)) will always be defined by the expertise, and the nature of the hypotheses defined and defended by individual researchers.

Core microbes

The individual datasets described in this study had previously focused on identifying abundant microbes and differences due to experimental conditions. The current meta-analysis goes a step further to find common microbiota that are associated with switchgrass across the diverse experimental conditions. In fact, while *Lysobacter* would be identified as having significantly differential abundance as per the Man Whitney test, it would not be in the top 50 candidates (ranked as per FDR in [Table S3](#)) and likely missed from future testing. Clearly, our approach allows us to identify such candidates. The members of the *Lysobacter* genus, an identified core microbe of switchgrass, are known to live in soil and have been shown to be ecologically important due to their ability to produce exo-enzymes and antibiotics ([Reichenbach, 2006](#)). Their antimicrobial activities against bacteria, fungi, unicellular algae, and nematodes have been described ([Islam et al., 2005](#); [Jochum, Osborne & Yuen, 2006](#); [Park et al., 2008](#); [Yin, 2010](#)). Strains of this genus, for example, have been used for control of diseases caused by bacteria in rice ([Ji et al., 2008](#)) and tall fescue ([Kilic-Ekici & Yuen, 2004](#)). Reports of their function thus support the idea that they may play an important role in switchgrass growth and survival. The core microbiome results thus support further research into the role played by this bacterium in the switchgrass rhizosphere.

Similarly, members of the *Mesorhizobium* genus are well-known diazotrophs ([Kaneko et al., 2000](#)) and previously shown to be symbiotically associated with switchgrass ([DeAngelis et al., 2010](#); [Bahulikar et al., 2014](#)) and legumes ([Weir et al., 2004](#); [Nandasena et al., 2009](#); [Degefu et al., 2013](#); [Guerrouj et al., 2013](#)). Another identified core microbiome taxa, soil-dwelling members of the *Chitinophagaceae* family are known to have β -glucosidase ([Bailey et al., 2013](#)) and Aminocyclopropane-1-carboxylate (ACC) deaminase activities and ability to produce indole-3-acetic acid (IAA) ([Madhaiyan et al., 2015](#)). These molecules and enzymes are well known for their effects on plant growth ([Zhao, 2010](#); [Van de Poel & Van Der Straeten, 2014](#)). The capacity to degrade cellulose might provide additional and readily available options to aid survival of these bacteria near switchgrass root zones during times of environmental stress. ACC deaminase and IAA production, in contrast, are potent plant growth modulators ([Glick, 2014](#)) that could play a role in plant productivity and survival, especially under conditions of plant physiological stress. Though these examples above would need further study, they provide consistent examples describing how a core microorganism could play a role in determining plant function and growth. The power of the approach stems from the ability to identify the core microbes associated with a plant (or other habitat), and that can, with veracity, narrow down potentially important core microbes from otherwise hyperdiverse samples.

From a technological standpoint, it is important to put the current approach into context with research before the metagenomics era. The search and identification of antagonistic plant growth promoting microbes has previously been tedious and labor intensive. Screenings of hundreds of microbes were used to cultivate and identify candidate microbes that might support (or deter) plant growth. In the case of beneficial microbes, even when identified under greenhouse conditions, the beneficial effects rarely translated into plant supportive growth under field growth conditions (*Babalola, 2010; Hayat et al., 2010*). With the aid of hindsight and new knowledge suggesting the importance of the soil habitat and root-soil interactions in the development of growth promoting plant-microbial relationships, the approach used in this study reverses the focus (from top-down to bottom-up) to search for microbes that appear to already be naturally well-adapted to the root-soil habitats of interest (*Trabelsi & Mhamdi, 2013; Souza, Ambrosini & Passaglia, 2015*). This process streamlines the search for suitable microbes from a daunting pool of thousands of bacterial taxa. Bacteria and fungi with well-known partnerships with members of the core microbiome, it would be expected, to be more readily adaptable to their native environment. Indeed, the concept of adaptability to an environment has been shown to be true for many types of microbes across the environmental spectrum, and has given rise to the concept of the niche (*Lennon et al., 2012*). Existing tools (e.g., Corbata, MetaCoMET and QIIME) do not provide statistical significance for the taxa to be a core microbiome compared to the background using presence/absence data. The point of our method is to provide an approach that is definitive, rather than simply stating a set of microbe present in samples. Due to the lack of consensus in the scientific community about what a gold standard protocol/technology is for such studies (mainly due to the relatively early stages of the metagenomics field and technology still in developmental stages), one is limited (in one way or other) to sufficiently utilize the existing data. While our method of combining datasets has its limitation (like the granularity/species level information limitation from 16S, etc.), it still offers a powerful way to (i) understand the taxonomic distribution within samples (ii) mine multiple existing and often-diverse datasets (iii) generate hypothesis for future detailed experiments, and therefore, certainly a preferred alternative than not using existing data. The COREMIC tool provides an alternative and logical approach to help mine available datasets, in the search for core microbiomes associated with habitats that are ecologically and agriculturally important. Finally, each statistical test relies on different assumptions and has different strengths. COREMIC assumes that presence of a microbe, however high or low, can provide meaningful insight into potential host-microbe relationships. Therefore, it provides equal weightage for high and low abundant microbes within a sample. Since microbial abundance can be an important factor in a biological system, we recommend using COREMIC (presence/absence) in complement with other abundance-based methods (e.g., Mann Whitney test, ISA, etc.). Furthermore, while using diverse dataset has its strengths we suggest avoiding datasets where each dataset contains only a single (mutually exclusive) group. The user needs to consider similarities/differences between the datasets and the biological system while choosing an appropriate outgroup for their group of interest. It is up to the user to decide which thresholds make the most sense for their questions and hypothesis, but obviously a stricter threshold will have

higher statistical-inferential veracity. Perhaps a combinatorial approach of selecting the top candidates from the different (presence/absence and abundance based) methods, picking microbes that show significant associations by multiple methods, and using the user's biological expertise might better allow choosing candidates for future testing.

CONCLUSIONS

The COREMIC tool, by helping to mine multiple datasets fills an existing gap in the search for the core microbiome associated with a host or habitat. It allows for the development of a working hypothesis in the search for microbes well suited for a habitat or host-microbe interaction. It can also be used to confirm laboratory studies that have identified target microbes that might be important symbionts or thought to be associated with a specific habitat. In the case of plants, but not limited to them, the COREMIC approach can identify microbial targets that might be useful for plant growth promotion. An example of this would be the identification of diazotrophic bacteria that aid the growth of bioenergy grasses and help to serve the development of sustainable agricultural systems. This combined with the ongoing efforts of plant breeding and genetic modification would help to catalyze microbe-driven crop yield improvement while practicing environmental stewardship through reduced fertilizer use. Here we show the applicability of COREMIC in rhizosphere-associated microbes, but the overall concepts are translational across disciplines with interests in host-microbe and microbe-habitat relationships. The applicability of COREMIC for the identification of core genes and microbes has excellent potential to help understand the roles of microorganisms in complex and diverse microbial communities.

ACKNOWLEDGEMENTS

The authors thank Dr. Roderick Jensen and James Wrenn for their suggestions in improving the manuscript. We thank Dr. James McClure for his help in the web-tool development. We also acknowledge BIOM, Google App Engine and their developers. We thank Virginia Tech's Genetics, Bioinformatics, and Computational Biology, Department of Horticulture, College of Agriculture and Life Sciences for their support.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Virginia Tech's Genetics, Bioinformatics, and Computational Biology, Department of Horticulture provided personnel funding. Research was also partially funded by grants from USDA-NIFA (2011-03815). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
USDA-NIFA: 2011-03815.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Richard R. Rodrigues conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Nyle C. Rodgers performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Xiaowei Wu analyzed the data, wrote the paper, reviewed drafts of the paper.
- Mark A. Williams conceived and designed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.

Data Availability

The following information was supplied regarding data availability:

COREMIC and the input files are available at <http://coremic2.appspot.com> or <http://core-mic.com>. Its code is available on github (<https://github.com/richrr/coremicro>). A video tutorial is available at <https://oregonstate.box.com/v/coremic-tutorial>. Raw data used for the study are available at [PRJEB6704](#) and [PRJNA320123](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.4395#supplemental-information>.

REFERENCES

- Anderson MJ. 2001.** A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**:32–46 DOI [10.1111/j.1442-9993.2001.01070.pp.x](https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x).
- Babalola OO. 2010.** Beneficial bacteria of agricultural importance. *Biotechnology Letters* **32**:1559–1570 DOI [10.1007/s10529-010-0347-0](https://doi.org/10.1007/s10529-010-0347-0).
- Bahulikar RA, Torres-Jerez I, Worley E, Craven K, Udvardi MK. 2014.** Diversity of nitrogen-fixing bacteria associated with switchgrass in the native tallgrass prairie of northern Oklahoma. *Applied and Environmental Microbiology* **80**:5636–5643 DOI [10.1128/AEM.02091-14](https://doi.org/10.1128/AEM.02091-14).
- Bailey VL, Fansler SJ, Stegen JC, McCue LA. 2013.** Linking microbial community structure to beta-glucosidic function in soil aggregates. *The ISME Journal* **7**:2044–2053 DOI [10.1038/ismej.2013.87](https://doi.org/10.1038/ismej.2013.87).
- Beals EW. 1984.** Bray-curtis ordination: an effective strategy for analysis of multivariate ecological data. *Advances in Ecological Research* **14**:1–55 DOI [10.1016/S0065-2504\(08\)60168-3](https://doi.org/10.1016/S0065-2504(08)60168-3).
- Berendsen RL, Pieterse CM, Bakker PA. 2012.** The rhizosphere microbiome and plant health. *Trends in Plant Science* **17**:478–486 DOI [10.1016/j.tplants.2012.04.001](https://doi.org/10.1016/j.tplants.2012.04.001).

- Berg G. 2009.** Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Applied Microbiology and Biotechnology* **84**:11–18 DOI [10.1007/s00253-009-2092-7](https://doi.org/10.1007/s00253-009-2092-7).
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010.** QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**:335–336 DOI [10.1038/nmeth.f.303](https://doi.org/10.1038/nmeth.f.303).
- Chaudhary D, Saxena J, Lorenz N, Dick L, Dick R. 2012.** Microbial profiles of rhizosphere and bulk soil microbial communities of biofuel crops switchgrass (*Panicum virgatum* L.) and jatropha (*Jatropha curcas* L.). *Applied and Environmental Soil Science* **2012**:1–6 DOI [10.1155/2012/906864](https://doi.org/10.1155/2012/906864).
- Clarke KR. 1993.** Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* **18**:117–143 DOI [10.1111/j.1442-9993.1993.tb00438.x](https://doi.org/10.1111/j.1442-9993.1993.tb00438.x).
- Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012.** The impact of the gut microbiota on human health: an integrative view. *Cell* **148**:1258–1270 DOI [10.1016/j.cell.2012.01.035](https://doi.org/10.1016/j.cell.2012.01.035).
- DeAngelis KM, Gladden JM, Allgaier M, D'haeseleer P, Fortney JL, Reddy A, Hugenholtz P, Singer SW, Gheynst JSV, Silver WL, Simmons BA, Hazen TC. 2010.** Strategies for enhancing the effectiveness of metagenomic-based enzyme discovery in lignocellulolytic microbial communities. *BioEnergy Research* **3**:146–158 DOI [10.1007/s12155-010-9089-z](https://doi.org/10.1007/s12155-010-9089-z).
- Degefu T, Wolde-Meskel E, Liu B, Cleenwerck I, Willems A, Frostegard A. 2013.** *Mesorhizobium shonense* sp. nov. *Mesorhizobium hawassense* sp. nov. and *Mesorhizobium abyssinicae* sp. nov. isolated from root nodules of different agroforestry legume trees. *International Journal of Systematic and Evolutionary Microbiology* **63**:1746–1753 DOI [10.1099/ijms.0.044032-0](https://doi.org/10.1099/ijms.0.044032-0).
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006.** Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* **72**:5069–5072 DOI [10.1128/AEM.03006-05](https://doi.org/10.1128/AEM.03006-05).
- Dufrene M, Legendre P. 1997.** Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs* **67**:345–366 DOI [10.1890/0012-9615\(1997\)067\[0345:SAAST\]2.0.CO;2](https://doi.org/10.1890/0012-9615(1997)067[0345:SAAST]2.0.CO;2).
- Edgar RC. 2010.** Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**:2460–2461 DOI [10.1093/bioinformatics/btq461](https://doi.org/10.1093/bioinformatics/btq461).
- Evans JD, Schwarz RS. 2011.** Bees brought to their knees: microbes affecting honey bee health. *Trends in Microbiology* **19**:614–620 DOI [10.1016/j.tim.2011.09.003](https://doi.org/10.1016/j.tim.2011.09.003).
- Glick BR. 2014.** Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiological Research* **169**:30–39 DOI [10.1016/j.micres.2013.09.009](https://doi.org/10.1016/j.micres.2013.09.009).

- Gower JC. 2005.** *Principal coordinates analysis, encyclopedia of biostatistics*. 2nd edition. Milton Keynes: John Wiley and Sons, Ltd, The Open University.
- Guerrouj K, Perez-Valera E, Chahboune R, Abdelmoumen H, Bedmar EJ, El Idrissi MM. 2013.** Identification of the rhizobial symbiont of *Astragalus glombiformis* in Eastern Morocco as *Mesorhizobium camelthorni*. *Antonie Van Leeuwenhoek* **104**:187–198 DOI [10.1007/s10482-013-9936-y](https://doi.org/10.1007/s10482-013-9936-y).
- Hargreaves SK, Williams RJ, Hofmockel KS. 2015.** Environmental filtering of microbial communities in agricultural soil shifts with crop growth. *PLOS ONE* **10**:e0134345 DOI [10.1371/journal.pone.0134345](https://doi.org/10.1371/journal.pone.0134345).
- Hayat R, Ali S, Amara U, Khalid R, Ahmed I. 2010.** Soil beneficial bacteria and their role in plant growth promotion: a review. *Annals of Microbiology* **60**:579–598 DOI [10.1007/s13213-010-0117-1](https://doi.org/10.1007/s13213-010-0117-1).
- Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJM. 2001.** Counting the uncountable: statistical approaches to estimating microbial diversity. *Applied and Environmental Microbiology* **67**:4399–4406 DOI [10.1128/AEM.67.10.4399-4406.2001](https://doi.org/10.1128/AEM.67.10.4399-4406.2001).
- Islam MT, Hashidoko Y, Deora A, Ito T, Tahara S. 2005.** Suppression of damping-off disease in host plants by the rhizoplane bacterium *Lysobacter* sp. strain SB-K88 is linked to plant colonization and antibiosis against soilborne *Peronosporomycetes*. *Applied and Environmental Microbiology* **71**:3786–3796 DOI [10.1128/AEM.71.7.3786-3796.2005](https://doi.org/10.1128/AEM.71.7.3786-3796.2005).
- Jesus EdC, Liang C, Quensen JF, Susilawati E, Jackson RD, Balsler TC, Tiedje JM. 2016.** Influence of corn, switchgrass, and prairie cropping systems on soil microbial communities in the upper Midwest of the United States. *Global Change Biology Bioenergy* **8**:481–494 DOI [10.1111/gcbb.12289](https://doi.org/10.1111/gcbb.12289).
- Jesus EC, Susilawati E, Smith SL, Wang Q, Chai B, Farris R, Rodrigues JLM, Thelen KD, Tiedje JM. 2010.** Bacterial communities in the rhizosphere of biofuel crops grown on marginal lands as evaluated by 16S rRNA gene pyrosequences. *BioEnergy Research* **3**:20–27 DOI [10.1007/s12155-009-9073-7](https://doi.org/10.1007/s12155-009-9073-7).
- Ji G-H, Wei L-F, He Y-Q, Wu Y-P, Bai X-H. 2008.** Biological control of rice bacterial blight by *Lysobacter antibioticus* strain 13-1. *Biological Control* **45**:288–296 DOI [10.1016/j.biocontrol.2008.01.004](https://doi.org/10.1016/j.biocontrol.2008.01.004).
- Jochum CC, Osborne LE, Yuen GY. 2006.** *Fusarium* head blight biological control with *Lysobacter enzymogenes* strain C3. *Biological Control* **39**:336–344 DOI [10.1016/j.biocontrol.2006.05.004](https://doi.org/10.1016/j.biocontrol.2006.05.004).
- Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Mochizuki Y, Nakayama S, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S. 2000.** Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Research* **7**:331–338 DOI [10.1093/dnares/7.6.331](https://doi.org/10.1093/dnares/7.6.331).
- Kennedy N, Brodie E, Connolly J, Clipson N. 2004.** Impact of lime, nitrogen and plant species on bacterial community structure in grassland microcosms. *Environmental Microbiology* **6**:1070–1080 DOI [10.1111/j.1462-2920.2004.00638.x](https://doi.org/10.1111/j.1462-2920.2004.00638.x).

- Kilic-Ekici O, Yuen GY. 2004.** Comparison of strains of *Lysobacter enzymogenes* and PGPR for induction of resistance against *Bipolaris sorokiniana* in tall fescue. *Biological Control* **30**:446–455 DOI [10.1016/j.biocontrol.2004.01.014](https://doi.org/10.1016/j.biocontrol.2004.01.014).
- Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM, Belnap J. 2002.** Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. *Applied and Environmental Microbiology* **68**:1854–1863 DOI [10.1128/AEM.68.4.1854-1863.2002](https://doi.org/10.1128/AEM.68.4.1854-1863.2002).
- Lennon JT, Aanderud ZT, Lehmkuhl BK, Schoolmaster DR. 2012.** Mapping the niche space of soil microorganisms using taxonomy and traits. *Ecology* **93**:1867–1879 DOI [10.1890/11-1745.1](https://doi.org/10.1890/11-1745.1).
- Liang C, Jesus E, Duncan D, Jackson R, Tiedje J, Balsler T. 2012.** Soil microbial communities under model biofuel cropping systems in southern Wisconsin, USA: impact of crop species and soil properties. *Applied Soil Ecology* **54**:24–31 DOI [10.1016/j.apsoil.2011.11.015](https://doi.org/10.1016/j.apsoil.2011.11.015).
- Madhaiyan M, Poonguzhali S, Senthilkumar M, Pragatheswari D, Lee JS, Lee KC. 2015.** *Arachidicoccus rhizosphaerae* gen. nov. sp. nov. a plant-growth-promoting bacterium in the family *Chitinophagaceae* isolated from rhizosphere soil. *International Journal of Systematic and Evolutionary Microbiology* **65**:578–586 DOI [10.1099/ijs.0.069377-0](https://doi.org/10.1099/ijs.0.069377-0).
- Mao Y, Li X, Smyth E, Yannarell A, Mackie R. 2014.** Enrichment of specific bacterial and eukaryotic microbes in the rhizosphere of switchgrass (*Panicum virgatum* L) through root exudates. *Environmental Microbiology Reports* **6**:293–306 DOI [10.1111/1758-2229.12152](https://doi.org/10.1111/1758-2229.12152).
- Mao Y, Yannarell A, Davis S, Mackie R. 2013.** Impact of different bioenergy crops on N-cycling bacterial and archaeal communities in soil. *Environmental Microbiology* **15**:928–942 DOI [10.1111/j.1462-2920.2012.02844.x](https://doi.org/10.1111/j.1462-2920.2012.02844.x).
- Mao Y, Yannarell A, Mackie R. 2011.** Changes in N-transforming archaea and bacteria in soil during the Establishment of Bioenergy Crops. *PLOS ONE* **6**:e24750 DOI [10.1371/journal.pone.0024750](https://doi.org/10.1371/journal.pone.0024750).
- Martin M. 2011.** Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* **17**:10–12 DOI [10.14806/ej.17.1.200](https://doi.org/10.14806/ej.17.1.200).
- McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. 2012.** An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME Journal* **6**:610–618 DOI [10.1038/ismej.2011.139](https://doi.org/10.1038/ismej.2011.139).
- Mielke PW. 1984.** Meteorological applications of permutation techniques based on distance functions. In: Krishnaiah PR, Sen PK, eds. *Handbook of statistics: nonparametric methods*. Amsterdam: North-Holland. 813–830.
- Nandasena KG, O’Hara GW, Tiwari RP, Willems A, Howieson JG. 2009.** *Mesorhizobium australicum* sp. nov. and *Mesorhizobium opportunistum* sp. nov., isolated from *Biserrula pelecinus* L. in Australia. *International Journal of Systematic and Evolutionary Microbiology* **59**:2140–2147 DOI [10.1099/ijs.0.005728-0](https://doi.org/10.1099/ijs.0.005728-0).

- Park JH, Kim R, Aslam Z, Jeon CO, Chung YR. 2008.** *Lysobacter capsici* sp. nov., with antimicrobial activity, isolated from the rhizosphere of pepper, and emended description of the genus *Lysobacter*. *International Journal of Systematic and Evolutionary Microbiology* **58**:387–392 DOI [10.1099/ij.s.0.65290-0](https://doi.org/10.1099/ij.s.0.65290-0).
- Reichenbach H. 2006.** *The genus lysobacter, the prokaryotes*. New York: Springer, 939–957.
- Rodrigues RR, Moon J, Zhao B, Williams MA. 2017.** Microbial communities and diazotrophic activity differ in the root-zone of Alamo and Dacotah switchgrass feedstocks. *GCB Bioenergy* **9**:1057–1070 DOI [10.1111/gcbb.12396](https://doi.org/10.1111/gcbb.12396).
- Rodrigues RR, Pineda RP, Barney JN, Nilsen ET, Barrett JE, Williams MA. 2015.** Plant invasions associated with change in root-zone microbial community structure and diversity. *PLOS ONE* **10**:e0141424 DOI [10.1371/journal.pone.0141424](https://doi.org/10.1371/journal.pone.0141424).
- Shade A, Handelsman J. 2012.** Beyond the Venn diagram: the hunt for a core microbiome. *Environmental Microbiology* **14**:4–12 DOI [10.1111/j.1462-2920.2011.02585.x](https://doi.org/10.1111/j.1462-2920.2011.02585.x).
- Souza R, Ambrosini A, Passaglia LM. 2015.** Plant growth-promoting bacteria as inoculants in agricultural soils. *Genetics and Molecular Biology* **38**:401–419 DOI [10.1590/S1415-475738420150053](https://doi.org/10.1590/S1415-475738420150053).
- Trabelsi D, Mhamdi R. 2013.** Microbial inoculants and their impact on soil microbial communities: a review. *BioMed Research International* **2013**:Article 863240 DOI [10.1155/2013/863240](https://doi.org/10.1155/2013/863240).
- Turner T, Ramakrishnan K, Walshaw J, Heavens D, Alston M, Swarbreck D, Osbourn A, Grant A, Poole P. 2013.** Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *The ISME Journal* **7**:2248–2258 DOI [10.1038/ismej.2013.119](https://doi.org/10.1038/ismej.2013.119).
- Van de Poel B, Van Der Straeten D. 2014.** 1-aminocyclopropane-1-carboxylic acid (ACC) in plants: more than just the precursor of ethylene. *Frontiers in Plant Science* **5**:640 DOI [10.3389/fpls.2014.00640](https://doi.org/10.3389/fpls.2014.00640).
- Weir BS, Turner SJ, Silvester WB, Park DC, Young JM. 2004.** Unexpectedly diverse Mesorhizobium strains and Rhizobium leguminosarum nodulate native legume genera of New Zealand, while introduced legume weeds are nodulated by Bradyrhizobium species. *Applied and Environmental Microbiology* **70**:5980–5987 DOI [10.1128/AEM.70.10.5980-5987.2004](https://doi.org/10.1128/AEM.70.10.5980-5987.2004).
- Werling BP, Dickson TL, Isaacs R, Gaines H, Gratton C, Gross KL, Liere H, Malmstrom CM, Meehan TD, Ruan L, Robertson BA, Robertson GP, Schmidt TM, Schrotenboer AC, Teal TK, Wilson JK, Landis DA. 2014.** Perennial grasslands enhance biodiversity and multiple ecosystem services in bioenergy landscapes. *Proceedings of the National Academy of Sciences of the United States of America* **111**:1652–1657 DOI [10.1073/pnas.1309492111](https://doi.org/10.1073/pnas.1309492111).
- Yin H. 2010.** Detection methods for the genus *lysobacter* and the species *lysobacter* enzymogenes. In: *Biological sciences*. Lincoln: University of Nebraska.
- Zhao Y. 2010.** Auxin biosynthesis and its role in plant development. *Annual Review of Plant Biology* **61**:49–64 DOI [10.1146/annurev-arplant-042809-112308](https://doi.org/10.1146/annurev-arplant-042809-112308).