**Title**

# Genome-centric analyses of seasonal phyllosphere microbiome activities in perennial crops

**Authors**

Adina Howe[1,2], Nejc Stopnisek[3,4], Shane K. Dooley[1], Fan M. Yang[1], Keara L. Grady[3,4], Ashley Shade[3,4,5,6]

1. Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011

2. Department of Bioinformatics and Computational Biology, Iowa State University, Ames, IA 50011

3. The Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, MI 48824

4. Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824

5. Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing MI 48824

6. Program in Ecology, Evolution, and Behavior, Michigan State University, East Lansing Mi 48824

**Abstract**

Plants and microbes form beneficial associations. It is expected that understanding these interactions will allow for microbiome management to enhance crop productivity and resilience to stress. Here, we apply a genome-centric approach to identify key leaf microbiome members and quantify their activities on field-grown switchgrass and miscanthus. We integrate metagenome and metatranscriptome sequencing from 237 leaf samples collected over key time points in crop phenology. We curate metagenome-assembled-genomes (MAGs), and conservatively focus analysis on the highest quality MAGs that were <15.5% contaminated, >70% complete, and detected in a minimum of 10% of samples for each crop. Populations represented by these MAGs were actively transcribing genes, and exhibited seasonal dynamics in key functions, including pyruvate metabolism, threonine, homoserine and serine biosynthesis, and stress response. Notably, we detected enrichment in transcripts annotated to nonmavalonate isoprene biosynthesis in the late season, prior to and during host senescence, concurrent with when plants are expected to decrease their isoprene biosynthesis. We also detected groups of MAGs that had coherent transcript dynamics in thioredoxin reductase, which suggests a response to reactive oxygen species, potentially released by plant hosts experiencing abiotic stress. Overall, this study overcame laboratory and bioinformatic challenges associated with field-based leaf metatranscriptomes analysis to direct to some of the key activities of phyllosphere bacteria. These activities collectively support that leaf-associated bacterial populations can be seasonally dynamic, responsive to host cues, and, likely, interactively engage in feedbacks with the plant.

42 **Keywords**

43 Metagenome, metatranscriptome, leaf, bacteria, archaea, bioenergy, biofuel feedstocks,

44 phytobiome, switchgrass, miscanthus, seasonal, time-series

45

46 **Introduction**

47 Perennial plants are a crucial target for the sustainable development of biofuels
48 (Robertson *et al.*, 2017). In addition to yielding high biomass that can be converted to biofuels
49 and bioproducts, perennial crops offer a broad range of ecosystem services that support efforts
50 to mediate climate change, including greenhouse gas mitigation and promotion of soil nutrient
51 cycling (Heaton *et al.*, 2008; Langholtz *et al.*, 2016; Robertson *et al.*, 2017; Roley *et al.*, 2018).
52 Like all plants, perennials harbor diverse microbiota, and many of these microbes are either
53 known or expected to benefit their hosts. For example, plant-associated microbes can increase
54 productivity and protect against environmental stressors (Bulgarelli *et al.*, 2013; Laforest-
55 Lapointe *et al.*, 2017; Gupta *et al.*, 2019). Because of the intimate engagement of many plant-
56 associated microbiota with the host, management of the plant microbiome is one tool
57 proposed to promote crop vigor and support crop resilience to global climate changes (Busby *et*
58 *al.*, 2017; Toju *et al.*, 2018; Haskett *et al.*, 2020; Wang and Haney, 2020). Therefore, along with
59 selective breeding and data-informed field management, regulating the plant microbiome is
60 expected to be key for the sustainable production of perennial biofuel feedstocks.

61 Like animal bodies, plants have anatomical compartments that each are inhabited by
62 distinctive microbial consortia. Generally, the diversity and composition of the plant
63 microbiome narrows from external compartments to internal, and the plant plays an active role
64 in filtering the microbiome composition inward (Hacquard *et al.*, 2015; Hardoim *et al.*, 2015;
65 Gopal and Gupta, 2016). External plant compartments include the root zone, rhizosphere and
66 rhizoplane below ground, and the phyllosphere above ground (Andrews and Harris, 2000).
67 External compartments have relatively higher representation of transient or commensal

68 microbial taxa, and these compartments engage with and recruit microbes from the immediate
69 environment. Internal compartments include the endosphere of above- and below-ground
70 tissues, and these have relatively low richness and harbor the most selected microbiota
71 (Bulgarelli *et al.*, 2013; Müller *et al.*, 2016). Of these compartments, the rhizosphere has
72 received the most attention as a key site of microbial-plant interactions that are important for
73 nutrient and water acquisition (e.g., Kuzyakov and Razavi, 2019). However, members of the
74 microbiota that inhabit the phyllosphere also provide important plant functions, such as
75 pathogen exclusion and immune priming (Bell *et al.*, 2019; Chen *et al.*, 2020). Phyllosphere
76 microorganisms have specialized adaptations to their exposed lifestyle (Lindow and Brandl,
77 2003; Vorholt, 2012; Müller *et al.*, 2016; Koskella, 2020) and contribute to global carbon and
78 other biogeochemical cycling, including transformations relevant for climate change (Bringel
79 and Couée, 2015; Dorokhov *et al.*, 2018; Cavicchioli *et al.*, 2019). Because perennial biofuel
80 feedstocks are selected to maximize foliage surface area, understanding the phyllosphere
81 microbiome is expected to provide insights into key microbial engagements that benefit the
82 plant to support productivity and stress resilience.

83 There are two general challenges in regulating the microbiome to promote crop vigor
84 and resilience to environmental stress. The first challenge is to distinguish the beneficial
85 members of the plant microbiome from transient or commensal members, with recognition
86 that some members that provide host benefit likely change situationally, either over plant
87 development or given environmental stress (Edwards *et al.*, 2015; Xu *et al.*, 2018; Zhalnina *et*
88 *al.*, 2018), while others are stable (Shade and Stopnisek, 2019; Stopnisek and Shade, 2021) . The
89 second challenge is that plants and their agroecosystems are temporally dynamic over the

90 growing season, and their associated microbiota are also dynamic. It is currently unclear what
91 functions may be associated with phyllosphere microbial dynamics and their potential
92 interactions with plant hosts.

93 Previously, we used 16S rRNA gene amplicon analysis to identify a "core" cohort of
94 bacterial and archaeal taxa that were persistently associated with the phyllosphere
95 microbiomes' of two perennial biofuel feedstocks, miscanthus and switchgrass. Persistent
96 membership was established by collecting leaf samples over replicated field plots, over a
97 temperate seasonal cycle, and across two annual growing seasons for switchgrass (Grady *et al.*,
98 2019). Other studies in switchgrass have similarly reported that the leaf, and other plant
99 compartments, can be distinguished by microbiome compositions (e.g.,Bulgarelli *et al.*, 2012;
100 Lundberg *et al.*, 2012; Bahulikar *et al.*, 2014; Roley *et al.*, 2019; Singer *et al.*, 2019), suggesting
101 selection to the leaf compartment. In the current study, we aimed to understand the functional
102 attributes of persistent phyllosphere taxa, with an interest in specialized adaptations to the leaf
103 and interactions with the host plant that may inform the mechanisms and nature of plant-
104 microbe engagements. Therefore, we performed a seasonal analysis of phyllosphere
105 metagenomes for both miscanthus and switchgrass. We paired our metagenome longitudinal
106 series with metatranscriptome analyses at select time points in switchgrass phenology to
107 determine which functions were active seasonally. We performed genome-centric analyses of
108 the metagenome and focused on understanding the dynamics and functions of the highest
109 quality and most persistent metagenome-assembled-genomes (MAGs). Our results reveal
110 specialized functions, genomic attributes, and seasonal activities of persistent phyllosphere
111 members. Finally, we provide evidence that these genomes were detected in various sites and

112 years beyond our original study plots, suggesting that they are general, consistent inhabitants
113 of these bioenergy grasses.

114 **Materials and Methods**

115 *Site description and sampling scheme*

116 Switchgrass and miscanthus leaves and corresponding contextual data were collected at
117 the Great Lakes Bioenergy Research Center (GLBRC) located at the Kellogg Biological Station in
118 Hickory Corners, MI, USA (42°23'41.6" N, 85°22'23.1" W). We sampled switchgrass (*Panicum*
119 *virgatum L.* cultivar Cave-in-rock) and miscanthus (*Miscanthus x giganteus*), from the Biofuel
120 Cropping System Experiment (BCSE) sites, plot replicates 1-4, as previously described (Grady *et*
121 *al.*, 2019). This included collecting leaves from switchgrass and miscanthus at 8 and 9 time
122 points, respectively, in 2016 and switchgrass at 7 time points in 2017 seasons, for a total of 237
123 leaf samples. Additional switchgrass leaves were collected for RNA isolation at three time
124 points in 2016 to capture host plant phenology at emergence, peak growth, and senescence
125 stages (as categorized by GLBRC according to their standard phenology methods, see
126 https://data.sustainability.glbrc.org/protocols/165 ), and at all 7 sampling time points in 2017.
127 Leaves for RNA isolation from 80 samples (24 from 2016, and 56 from 2017) were flash-frozen
128 in liquid nitrogen immediately and frozen at -80°C until processing.

129 *Phyllosphere RNA and DNA isolation*

130 Phyllosphere epiphyte RNA was isolated using a benzyl chloride liquid:liquid extraction,
131 based on (Grady *et al.*, 2019) that was adapted for RNA isolation based on the methods of
132 (Suzuki *et al.*, 2001). Approximately 5 g of intact, frozen leaf material was added to a 50 ml

133 polypropylene conical tube (Corning #430290) and kept frozen on liquid nitrogen while samples
134 were weighed and transferred. Denaturing Solution (DS) was prepared with 4.2 M guanidine
135 thiocyanate, 25 mM sodium citrate dihydrate pH 7.0, 0.5% (v/v) sodium n-laroyl sarcosine in
136 Milli-q water and was filter sterilized through 0.22 mm filters. Immediately prior to extraction, a
137 working stock of DS was prepared fresh by adding 2-mercaptoethanol to a final concentration
138 of 5% (v/v, DS/2-ME). To each leaf tube, 5 ml of benzyl chloride, 2.5 ml of 3M sodium acetate
139 (pH 5.2), and 5 ml of the working stock of DS/2-ME was added. The tube was incubated in a
140 60$^{o}$C water bath for 20 minutes with vortexing every 1 min. The leaves were removed from the
141 conical tube using ethanol-sterilized forceps and discarded.

142 Five ml of chloroform:isoamyl alcohol (24:1) were added to the remaining solution in
143 each conical tube, which were then shaken by hand for 15 seconds and incubated on ice for 15
144 minutes. The tubes were then centrifuged at 12,000 x g for 15 min at 4$^{o}$C to separate aqueous
145 and organic phases. Up to 5 ml of the upper, aqueous phase was transferred to a clean 15 ml
146 polypropylene tube, without disrupting the white interface. Two and a half ml of sodium
147 citrate buffer (1.2 M sodium chloride, 0.8 M sodium citrate dihydrate in Milli-q water, filter
148 sterilized at 0.22 mm) and ice-cold isopropanol were added to achieve a final volume of 12.5
149 ml. Next, the tubes were centrifuged at 12,000 x g for 15 minutes to pellet the RNA, and the
150 remaining supernatant aspirated using a pipette. The RNA pellets were resuspended in 0.3 ml
151 of working DS/2-ME solution, and afterwards 0.3 ml of ice-cold isopropanol was added and
152 mixed by pipetting gently. The solution was incubated for 30 minutes at -20$^{o}$C, and then the full
153 volume was transferred to a clean nuclease-free 1.7 ml tube and centrifuged at 16,000 x g for
154 15 minutes at 4 $^{o}$C. The supernatant was removed, and the pellet washed in 1 ml of nuclease-

155 free 75% ethanol. Tubes were then centrifuged at 16,000 x g for 15 minutes at 4 $^{o}$C, and
156 supernatant again removed using a pipette. The remaining pellet was air dried to completely
157 remove residual ethanol, then resuspended in 30 ml of nuclease-free Tris-EDTA buffer, pH 8.0.

158 Genomic DNA (gDNA) was removed using RNase-free DNase I (Thermo Fisher #AM2222)
159 per manufacturer's instructions. The RNA was then purified using the RNeasy MinElute Cleanup
160 Kit (Qiagen Germantown, MD, USA) according to manufacturer's instructions. The absence of
161 contaminating gDNA was confirmed by lack of amplification of the 16S rRNA gene V4 region by
162 PCR (Caporaso *et al.*, 2011) with positive and negative controls.

163 Corresponding environmental DNA (mgDNA) was isolated and processed as in (Grady *et*
164 *al.*, 2019) from each of the 488 leaf samples. DNA concentrations were normalized to 4 ng/ml.
165 DNA and RNA samples were then shipped to and sequenced at the Joint Genome Institute (JGI)
166 according to their standard protocols.

167 *Metagenome library preparation*

168 At JGI, plate-based DNA library preparation for Illumina sequencing was performed on
169 the PerkinElmer Sciclone NGS robotic liquid handling system using Kapa Biosystems library
170 preparation kit. 1.82 ng of sample DNA was sheared to 436 bp using a Covaris LE220 focused-
171 ultrasonicator. The sheared DNA fragments were size selected by double-SPRI and then the
172 selected fragments were end-repaired, A-tailed, and ligated with Illumina compatible
173 sequencing adaptors from IDT containing a unique molecular index barcode for each sample
174 library. The prepared libraries were quantified using KAPA Biosystems' next-generation
175 sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument.

176 Sequencing of the flowcell was performed on the Illumina HiSeq sequencer following a 2x151
177 indexed run recipe.

178 *Metatranscriptome library preparation*

179 At JGI, plate-based RNA sample prep was performed on the PerkinElmer Sciclone NGS
180 robotic liquid handling system using Illuminas Ribo-Zero rRNA Removal Kit (Bacteria) and the
181 TruSeq Stranded Total RNA HT sample prep kit following the protocol outlined by Illumina in
182 their user guide: https://support.illumina.com/sequencing/sequencing_kits/truseqstranded-
183 total-rna.html, and with the following conditions: total RNA starting material of 100 ng per
184 sample and 10 cycles of PCR for library amplification. The prepared libraries were quantified
185 using KAPA Biosystems' next-generation sequencing library qPCR kit and run on a Roche
186 LightCycler 480 real-time PCR instrument. Sequencing of the flowcell was performed on the
187 Illumina NovaSeq sequencer using NovaSeq XP V1 reagent kits, S4 flowcell, following a 2x151
188 indexed run recipe.

189 *Analysis of metagenomes and metatranscriptomes*

190 There were 136 metagenome and 74 metranscriptome samples that successfully
191 sequenced. We used Trimmomatic (v0.39) (Bolger *et al.*, 2014) to remove adaptors and filter
192 low-quality reads from fastq files using the following arguments: PE -phred33
193 ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:8:TRUE LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15
194 MINLEN:36. Next, plant host reads were filtered out (for both metagenomes and
195 metatranscriptomes) by removing all reads that mapped to the switchgrass genome (*Panicum*
196 *virgatum* v1.0, DOE-JGI, http://phytozome.jgi.doe.gov/) or miscanthus genome (*Miscanthus*

197 *sinensis* V7.1 http://phytozome.jgi.doe.gov/) using bowtie2 (v 7.5.0), samtools (v 1.9), and
198 bedtools (v2.29.0) (Li *et al.*, 2009; Quinlan and Hall, 2010; Langmead and Salzberg, 2012). To
199 remove the fungal reads and improve the prokaryotic signal in metagenome samples, we also
200 filtered reads against 6 fungal genomes that represent close relatives of the most abundant
201 fungal species in this system (Bowsher *et al.*, 2020) (**Table S1**). The genomes of these fungal
202 species were retrieved from the JGI Genome Portal. Filtering against fungal genomes removed
203 an average of 2% of reads in the switchgrass metagenome samples (with a bias towards reads
204 in mid-season samples, ~9%), but fewer reads for the miscanthus metagenome samples
205 (average 1.1% and evenly detected across the growing season). Nonpareil (v3) was used to
206 estimate the metagenomic coverage and sample diversity (Rodriguez-R *et al.*, 2018). Filtered
207 metagenome reads were then combined and used for co-assembly with MEGHIT (v 1.2.9) using-
208 -kmin-1pass (low sequencing depth) and--presets meta-large (complex metagenome) (Li *et al.*,
209 2015).

210 Because of the uneven distribution of sequencing depth after eukaryotic read filtering,
211 we curated metagenome assembled genomes (MAGs) from the switchgrass and miscanthus
212 metagenome libraries using Metabat (v2.2.15) (Kang *et al.*, 2019). CheckM with the lineage_wf
213 option (Parks *et al.*, 2015) was used to determine completeness and contamination of produced
214 MAGs. MAGs were assigned taxonomy using CheckM (v.1.13) based on the presence and
215 consistent assignment of phylogenetic marker genes. Among 58 MAGs assembled from the
216 switchgrass or miscanthus phyllosphere metagenomes (< 50% contamintaiton and > 50%
217 completeness, **Dataset 1**), we selected a subset of 16 highest quality MAGs based on:
218 completeness greater than 70%, contamination less than 15.5%, and detection in greater than

219 10% of both switchgrass and miscanthus metagenomes (**Table S2**). Finally, assembled MAG
220 scaffolds were aligned to the chloroplast genome of *Panicum virgatum* to confirm that these
221 genes were absent.

222 *Abundances of MAGs and associated functions*

223 The abundance of each MAG in each sample was estimated based on the average
224 median coverage of metagenome reads associated to each MAG scaffolds. Specifically, Bowtie2
225 (v2.2.2) was used to align reads to scaffolds and Bedtools (v2.28) was used to estimate the
226 coverage of each basepair within the scaffold. For each MAG scaffold, open reading frames
227 (ORFs) and functional genes were identified using Prodigal (v2.6.3 , default). The functional
228 hierarchy of the ORFs was annotated by aligning sequences to a manually-curated soil genome
229 database, RefSoil+ (Choi *et al.*, 2017; Dunivin *et al.*, 2019). Refsoil ORFs were aligned RAST
230 Subsystems (Aziz *et al.*, 2008)
231 (https://figshare.com/articles/dataset/RefSoil_Database/4362812) using BLAST (v2.10.1) to
232 identify the gene with the best scoring alignment with an expected value score < 1e-5. Similar
233 to genome abundances, the abundances of these functional genes were estimated by mapping
234 metagenome and metatranscriptome reads to ORFs and estimating the median coverage per
235 ORF in each sample.

236 Transcripts were mapped to ORFs associated with each MAG to estimate median base
237 coverage of each ORF, which was subsequently normalized by total number of reads in the
238 sample to provide comparisons between samples. Pairwise comparisons of functional roles (the
239 cumulative sum of all associated ORFs ) between the early (May – July) and late (Aug – Sep)

240 season were performed with the Kruskal-Wallis test by ranks. The distribution and dynamics of
241 MAGs' transcripts were compared with hierchical clustering using Bray-Curtis dissimilarity
242 distances of the estimated metatranscriptome abundances. Additionally, pairwise comparisons
243 of transcript abundances associated with DNA metabolism and RNA metabolism were
244 compared for each sample date with the Kruskal-Wallis test by ranks.

245 *Data and code availability*

246 Metagenome and metatranscriptome data are available through the Joint Genome Institute
247 Genome Portal under project ID 503249. Code is available on GitHub
248 (https://github.com/ShadeLab/PAPER_Howe_2021_switchgrass_MetaT).

249

250 **Results**

251 *Key leaf bacterial populations inhabit both grasses and exhibit seasonal trends*

252 We focused analysis on sixteen highest-quality MAGs assembled from the leaf time
253 series that were detected in at least 10% of both switchgrass and miscanthus metagenomes,
254 which we refer to as our "gold" leaf MAGs. This set of MAGs was dominated by representatives
255 from the orders Actinomycetales, Burkholderiales, Pseudomonadales, and Rhizobiales (**Figure**
256 **1**, **Table S2**). Among identified MAGs, one MAG, MAG13_Unkn, is a member of bacteria but
257 could not be classified further with confidence, likely due to its relatively high contamination.
258 The consistent detection of these 16 MAGs in both miscanthus and switchgrass samples (**Figure**
259 **2A**) suggests that their originating bacterial populations were not host-specific but rather
260 distributed among these perennial grasses; this confirms previous results from the 16S rRNA

261 gene amplicon survey which revealed substantial overlap in the major leaf bacterial taxa across
262 the two crops (Grady *et al.*, 2019).

263 These MAGs exhibited clear seasonality, and most were enriched in the phyllosphere
264 metagenomes late in the growing season for both grasses (**Figure 2A, Figure S1A**). The majority
265 of MAGs were not detected early in the growing season but were first detected in August (late
266 season/peak biomass). Indeed, maximum enrichment of these MAGs occurred in the latest
267 months of sampling, from August onward. August samples corresponded with (late) peak
268 biomass and fruiting for switchgrass, and with a closed canopy for both switchgrass and
269 miscanthus. Senescence occurred as early as mid-September for switchgrass and as late as
270 November for miscanthus. Exceptions to this trend were the MAGs annotated as
271 Actinomycetales MAG11, Pseduomonadales MAG8, and MAG7, Rhizobiales MAG5, and
272 Rickettsiales MAG1, which were observed more consistently throughout the growing season in
273 miscanthus or switchgrass.

274 *MAGs harbor functional genes that support a leaf-associated lifestyle*

275 The most dominant functional genes identified among the 16 highest-quality MAGs were
276 associated with bacterial growth and survival, such the production of carbohydrates, amino
277 acids, respiration, and protein metabolism (**Figure 3**). We previously showed that many
278 abundant switchgrass and miscanthus leaf taxa could be detected in low abundance in their
279 associated soils (Grady *et al.*, 2019), suggesting that soils can serve as a reservoir for leaf taxa.
280 To initially evaluate if observed functions could be unique to the phyllosphere and contrast to
281 soils, we compared functions between phyllosphere MAGs and those of 1,015 well-curated soil

282 bacterial and archaeal genomes within the RefSoil database. Relative to RefSoil genomes,
283 phyllosphere MAGs were most enriched in the proportion of genes from functional subsystems
284 associated with secondary metabolism (Kruskal-Wallis test, $p<0.01$), phage and plasmid activity
285 ($p < 0.001$), photosynthesis ($p<0.05$), virulence, disease, and defense ($p< 0.001$), and stress
286 response ($p<0.001$) (**Table S3**).

287 Next, we analyzed leaf metatranscriptomes to evaluate the functional activity of the
288 MAGs over two switchgrass growing seasons. Despite substantial signal from this contaminant
289 RNA (**Figure S2)**, we were able to filter host reads against the plant genome and effectively
290 identify microbial transcripts that could be aligned to the leaf MAGs (**Figure 2B**, **Figure S1B**).
291 The bacterial signatures in the transcriptome data confirmed that these MAG members of the
292 leaf microbiome were alive and actively transcribing genes. We observed patterns of seasonal
293 transcription of MAG genes and therefore divided the time series into early (May-July) and late
294 (Aug-Sep) groups to further investigate seasonal activity differences. Transcripts associated
295 with protein metabolism, respiration, stress response, carbohydrate metabolism, and amino
296 acid production were the most abundant in the phyllosphere in August and September as
297 compared to May to July (**Figure 3**, Kruskall Wallis $p < 2e-3$). Importantly, this enrichment was
298 observed in both late 2016 and late 2017 (**Figure S3**). In total, there were 1,034 functional roles
299 (e.g., DNA gyrase subunit A (EC 5.99.1.3)) that were enriched in the late time points relative to
300 early ones with greater than the average difference (>35-fold). For each sample date, the
301 abundance of DNA metabolism and RNA metabolism were also compared. Transcript
302 abundances were significantly different for DNA and RNA metabolism only during late season

303 sample dates (Sept 2016 and July to September 2017, Kruskall Wallis p-value < 0.1 **Table S4**),
304 further supporting seasonal transcription of these MAG genes.

305 We selected the most prominent transcripts based on their shared presence among the
306 16 highest-quality MAGs, assuming that functions shared between bacteria represent key
307 functions for survival on the leaf. A total of 21 functional roles were identified in at least 13 of
308 the 16 MAGs (**Figure S4, Table S5**). Of these, most functions were associated with protein
309 metabolism (n=4 functional roles), carbohydrates (n=4), and amino acids and derivatives (n=3).
310 Consistent with actively growing cells, many of these functions were associated with the
311 regulation of transcription and included functions for DNA and RNA repair, synthesis, and
312 production such as DNA helicases and gyrases and tRNA synthetases**.** Additionally, transcripts
313 annotated as enzymes involved in acetogenesis from pyruvate (i.e., aldehyde dehydrogenase
314 and acetyl-coenzyme A synthetase), threonine biosynthesis and degradation (i.e., aspartate
315 aminotransferase, threonine synthase, and threonine dehydrogenase), and biosynthesis of
316 isoprenes and trehalose (i.e. 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase and 1,4-
317 alpha-glucan (glycogen) branching enzyme) were consistently enriched, particularly later in the
318 season. Among detected transcripts, genes encoding for proteins associated with pyruvate
319 metabolism and with threonine synthesis and degradation comprised 4 of the 8 functional
320 groups that were consistently identified in at least 14 of the MAGs and enriched from 40 to 65-
321 fold in August to September relative to May to July (**Table S5**). Among transcripts associated
322 with central metabolism and core to our MAGs, threonine dehydrogenase was the most
323 abundant (**Table S5**).

324 *Coherent groups of MAGs shared seasonal transcript enrichments and annotations*

325 To compare the transcriptional patterns observed among MAGs and over the crop
326 growing season, we used hierarchical clustering to identify coherent MAG activity dynamics. Six
327 coherent groups (clusters) of MAGs were identified (**Figure S5, Table S2**). The first group was
328 composed of MAG1, which had the lowest estimated genome completeness. Group 2 included
329 three Rhizobiales (MAG3, MAG4, and MAG5) and a Burkholderiales (MAG2); these MAGs were
330 generally active late in the season but detected consistently throughout the metatranscriptome
331 time series, including at some of the earliest May sampling dates (**Figure 4**). Group 2 MAGs had
332 transcripts enriched in respiration, protein metabolism, stress response, and RNA metabolism,
333 and carbohydrates.

334 Group 3 MAGs included a Burkholderiales (MAG6) and two Pseudomonadales (MAG7
335 and MAG8). Group 3 MAGs were the most persistent and were detected in all but three
336 sampling dates (**Figure 4**). Their late-enriched transcripts were annotated as associated with
337 stress response, protein metabolism, and RNA metabolism. This group also had modest
338 enrichment in sulfur metabolism in August. Group 4 MAGs included two Actinomycetales
339 (MAG9 and MAG11) and a Burkholderiales (MAG10); these populations were inconsistently
340 detected in May and June but with more consistency from July onward to September. Their
341 transcripts were very modestly enriched in August and September and were annotated as
342 protein metabolism, RNA metabolism, and carbohydrates. Group 5 included the unidentified
343 bacteria MAG13 and Sphingomonadales MAG12; these populations were inconsistently
344 detected until August, and their transcripts were not appreciably enriched until September.
345 However, in September, these MAGs appeared to become highly active, with relatively
346 abundant transcripts annotated to stress response, protein metabolism, RNA metabolism,

347 virulence, membrane transport, carbohydrates, amino acids and derivatives, DNA metabolism,
348 respiration, and sulfur metabolism. In particular, in September transcripts associated with
349 thioredoxin-disulfide reduction within sulfur metabolism, Group 5 MAGs appeared to become
350 highly active while Group 2 and 3 MAG activities were diminished from August abundances
351 (**Figure S6**). Finally, Group 6 included two Actinomycetales (MAG14 and MAG15) that were
352 rarely detected until August and had September-enriched transcripts belonging to similar
353 annotations as Group 5, including protein metabolism, RNA metabolism, amino acids and
354 derivatives, and respiration (**Figure 4**). While phylogenetic signal in the MAG dynamics cannot
355 be assessed directly due to insufficient statistical power, we note that many groups contained
356 MAGs annotated to the same Order. The exception to this was MAGs annotated to the Order
357 Burkholderiales, which were distributed among 4 out of the 6 clusters.

358 *MAGs were detected in metagenomes from different field sites, crops, and years*

359 We next asked if these high-quality MAGs were detected more broadly in other crop
360 metagenomes. We curated metagenomes from local cropping systems in Michigan and also
361 from the publicly available data within the Integrated Microbial Genomes (IMG) database,
362 which serves as a repository for the Joint Genome Institute's sequencing efforts through the US
363 Department of Energy. We were surprised to find that, for each phyllopshere metagenome
364 investigated, reads could be mapped to this set of MAGs. This included metagenomes from
365 both switchgrass and miscanthus, three locations (Iowa, Michigan, and Wisconsin) and across
366 five years (**Figure S7)**. These results suggest that the detected and analyzed MAGs can be
367 widely distributed in midwestern agroecosystems, and potentially of general importance for
368 perennial crop phyllosphere environments.

369

370 **Discussion**

371 Here, we report the first seasonal metagenome and metatranscriptome assessment of
372 plant phyllosphere in agricultural field conditions, focusing on the bacterial functions associated
373 with two promising biofuel feedstocks. We expect these findings to have relevance for other
374 grasses or systems with substantial aerial biomass, including native prairie. Furthermore, our
375 collection of high-quality MAGs included phyllosphere members previously reported as
376 abundant, persistent, or key for microbiome assembly in other plants, including the model
377 *Arabidopsis* (e.g., (Carlström *et al.*, 2019)), specifically: Sphingomonadales, Pseudomonadales,
378 Actinomycetales, Burkholderiales, and Rhizobiales. Therefore, the patterns and consistently
379 detected functions among this curated collection offer general insights as well as seasonal
380 phyllosphere functions.

381 The functions identified in our MAGs suggest general, key requirements for a leaf-
382 associated lifestyle. Investigating deeper into functional subsystems, there were major
383 enrichments in pyruvate metabolism (specifically, acetyl-CoA response), amino acid
384 biosynthesis (threonine, homoserine, and serine), isoprenoid biosynthesis, and stress responses
385 (trehalose biosynthesis, ABC transporters, and DNA repair), for which we offer a detailed
386 discussion.

387 *Central metabolism and amino acid biosynthesis*

388 As expected, many transcripts shared among leaf MAGs were associated with central
389 pathways of metabolism and the production of amino acids and carbohydrates. This result is

390 consistent with observations in tropical tree phyllospheres, where 45% of all functional
391 annotated sequences in metagenomes were associated with central metabolism (Lajoie *et al.*,
392 2020). Transcripts annotated to pyruvate metabolism and with threonine synthesis and
393 degradation were among the most prevalent. Pyruvate can regulate central metabolism
394 because it can be converted to acetyl coenzyme A (actetyl-CoA) and either enter the TCA cycle
395 or initialize the synthesis of long chain fatty acids, steroids, and ketone bodies. Threonine can
396 also be converted into pyruvate as an intermediate product in eventual conversion to acetyl-
397 CoA.

398 Previous observations have suggested that both microbial- and plant-derived pyruvate
399 availability plays a central role in leaf metabolism. Alanine dehydrogenase, which converts
400 alanine into pyruvate, was the most strongly upregulated protein during growth on plants in
401 the *A. thaliana* leaf commensal *Sphingomonas meloni* (Muller *et al.*, 2016). In addition,
402 pyruvate is a key intermediate for directing nutrient allocation in plants. In barley, pyruvate
403 was shown to regulate plant gene expression levels in response to nitrogen limitation, with
404 pyruvate concentrations decreasing during nitrogen deficiency (Fataftah *et al.*, 2018). Similarly,
405 in *A. thaliana*, over-expression of pyruvate was associated with increased orthophosphate
406 dikinase (which converts pyruvate to phosphoenolpyruvate), and, ultimately, increased
407 nitrogen remobilization from leaves to seeds (Taylor *et al.*, 2010). These previous observations
408 combined with our results provide more evidence that threonine and pyruvate regulation (and,
409 potentially, exchange) may play a role in microbe-leaf dynamics; however, it is unclear whether
410 leaf bacteria are directly interacting with hosts plants or indirectly responding to available
411 metabolites released by the host. Furthermore, as pyruvate intersects several central

412 metabolic pathways, further research is needed to determine whether its regulation in the host
413 is connected to microbial survival in the phyllosphere.

414 *Isoprene biosynthesis*

415 Isoprenoids are a class of volatile terpenes that are generally abundant and reactive,
416 and they engage in indirect and complex feedbacks with methane and nitrous oxide
417 greenhouse gases (McGenity *et al.*, 2018). Isoprene is one of the simplest isoprenoids. It is
418 released by many plant species, and much of it is synthesized within the methylerythritol
419 phosphate pathway of the chloroplast (MEP, aka:nonmavalonate) (Sharkey *et al.*, 2008).
420 Isoprene is thought to act as a signaling molecule in stress response (Zuo *et al.*, 2019). Studies
421 have also found that isoprene emission protects leaf photosynthesis against short episodes of
422 high-temperature (Sharkey and Yeh, 2001). Both switchgrass and miscanthus have been
423 reported to emit relatively low basal levels of isoprene (Eller *et al.*, 2011; Morrison *et al.*, 2016).
424 Plants emit isoprene from matured, photosynthetically active leaves, and emissions are light
425 responsive (Sharkey *et al.*, 2008). However, senescing leaves have been reported to decrease
426 in their isoprene emissions relative to leaves at peak growth (Sharkey *et al.*, 1991).

427 Among the high-quality MAGs we identified, we observed a 53-fold increase in bacterial
428 transcripts for MEP isoprene synthesis in the late season. From these data, we hypothesize that
429 that members of the phyllosphere bacterial community may be either compensating for the
430 loss of plant-derived isoprene, engaging in interspecies isoprene signaling with the host,
431 protecting plant photosynthesis from thermal damage, or possibly producing isoprene as an
432 overflow metabolite (as hypothesized for *Bacillus subtilus* (Sivy *et al.*, 2002)). Bacterial isoprene

433 degraders and synthesizers are widespread in nature (McGenity *et al.*, 2018) and have been
434 previously investigated in phyllosphere communities of the relatively high isoprene emitter
435 *Populus* spp (Crombie *et al.*, 2018), as well as in soils (El Khawand *et al.*, 2016), which can serve
436 as an isoprene sink. Stable isotope assays have been used to determine that a subset of
437 bacteria community members degrade isoprene, including several Actinobacteria (Rhodococcus
438 spp.) and Variopvorax (a Proteobacteria) (El Khawand *et al.*, 2016; Crombie *et al.*, 2018). Our
439 MAG collection contains several Actinobacteria (as well as Rhizobiales) and adds support for
440 their involvement in isoprene feedbacks in the phyllosphere. Our observation that bacterial
441 isoprene synthesis increases seasonally, and concurrently with when plant isoprene emissions
442 also is expected to decrease, is a strong justification for future work to be directed to
443 understanding these isoprene dynamics and any isoprene-mediated bacterial-host
444 engagement.

445 *Stress responses*

446 Trehalose is a disaccharide that protects cells against salt, water, and osmolyte stress by
447 serving as a stabilizing chemical chaperone, either by displacing water from protein surfaces or
448 by vitrifiing around protein structures to shield them (Laskowska *et al.*, 2020). In bacteria,
449 trehalose metabolism prevents cellular overflow metabolism and carbon stress by redirecting
450 glucose-6-phosphate from conversion to pyruvate (Laskowska *et al.*, 2020). Trehalose
451 biosynthesis is common among bacteria and archaea that live in arid, saline, thermal, or
452 seasonally dry environments (e.g., (Urrejola *et al.*, 2019), (Dai *et al.*, 2015; Lacerda-Júnior *et al.*,
453 2019), and it has also been reported to be induced in a pseudomonad by ethanol (Harty *et al.*,
454 2019), which can originate inside plant cells (Kimmerer and MacDonald, 1987) or more

455 generally in the roots (e.g., Ferner *et al.*, 2012), especially during stress, fruit ripening, or
456 senescence (Kimmerer and Kozlowski, 1982) . In plants, trehalose concentration is increased in
457 response to drought conditions, and its precursor, trehalose-6-phosphate, induces senescence
458 when carbon is readily available (Wingler *et al.*, 2012). Given this, it makes sense that the
459 majority of MAGs had enrichment of transcripts related to trehalose biosynthesis, which would
460 support a plant-associated lifestyle during drought and host senescence. The *A. thaliana*
461 phyllosphere member *Sphingomonas melonis* was reported to regulate trehalose biosynthesis
462 during growth conditions that promoted mild stress (Gottschlich *et al.*, 2019).
463 The enrichment of transcripts associated with ABC transporters and DNA repair also
464 aligns with expectations of adaptations to the harsh leaf habitat, and, more specifically, to
465 substrate transport given nutrient limitation and to nucleic acid maintenance given exposure to
466 ultraviolet light radiation, respectively. Carbohydrate metabolism and transport was previously
467 reported as enriched among genomes of plant-associated bacteria as compared to soil-
468 associated microbes (Levy, Salas Gonzalez, *et al.*, 2018), though this genome collection
469 contained both below- and above-ground plant-associated microbes.
470

471 *Differential transcript dynamics among MAG clusters*

472 We describe coherent transcript enrichment from groups of MAGs that had seasonally similar
473 dynamics. While there were many functions shared across MAG groups, like RNA and protein
474 metabolism, there were a few notable functions of particular groups. For example, transcripts
475 relating to sulfur metabolism, and more specifically, thioredoxin-disulfide reductase, were
476 noted in Groups 2, 3, and 5. The relative abundance of these transcripts oscillated among

477 Group 2, 3, and 5 from August to September, suggesting a hypothesis of complementarity or
478 functional redundancy.

479 Thioredoxin reductase, together with thioredoxin and the reductant NADPH, reduces
480 disulfide bonds to defend against cellular oxidative stress. Reactive oxygen species (ROS) serve
481 as signals for various developmental and cellular processes in plants. Though the precise
482 mechanisms are unclear, homeostasis of ROS are expected to be involved in senescence
483 (Kerchev *et al.*, 2015), which is relevant to our study given that at least some to a majority of
484 scenescing plants were observed per plot in August and September sampling dates,
485 respectively. Additionally, ROS accumulate in plants that are exposed to abiotic stress, to
486 negative effects. Thioredoxins (and their reductases) contribute to a general stress response
487 strategy that phyllosphere microorganisms employ to contend with expected fluctuations in
488 ROS on the leaf surface (though, these fluxes are difficult to measure,(Kerchev *et al.*, 2015)).
489 Thioredoxins were detected in abundance in a proteomic analysis of the leaf microbiomes of
490 four Brazilian tree species (Lambais *et al.*, 2017). Also, *Methylobacteria* spp, the well-studied
491 and common phyllosphere member, has thioredoxin functionality (Vorholt, 2012; Kwak *et al.*,
492 2014). Among several other genes relevant for oxidative stress response, thioredoxin reductase
493 was differentially regulated in a wild-type phyllosphere bacteria *Sphingomonas melonis* strain
494 Fr1 as compared to a knock-out mutant for regulation of general stress response, when both
495 were grown in a medium expected to induce low-levels of stress. (Gottschlich *et al.*, 2019).
496 Given that managing plant ROS is a target for reducing crop stress and regulating plant
497 development (Considine and Foyer, 2014; Kerchev *et al.*, 2015), it is possible that manipulating
498 microbial thioredoxin-related functions could be applied as one tool to achieve such efforts, but

499 much more work is needed to understand the microbial-host interaction given ROS exposure or
500 accumulation, and any possible ROS signaling between them.

501

502 *Conclusions*

503 Many recent review, perspective, and opinion pieces have urged integration of multi-
504 omics approaches to improve understanding of the microbiome and its relationship to the host
505 plant (Rastogi *et al.*, 2013; Levy, Conway, *et al.*, 2018; Remus-Emsermann and Schlechter, 2018;
506 Beilsmith *et al.*, 2019; Trivedi *et al.*, 2020). However, most integrative studies have focused
507 almost exclusively on the rhizosphere as the compartment of soil-plant feedbacks, and nutrient
508 and water acquisition for the host. Though leaves are readily accessible for sampling, the
509 phyllosphere microbiome is challenging to investigate using throughput, cultivation-
510 independent approaches like metagenomics and metatranscriptomics. There are high levels of
511 host and chloroplast contamination in leaf samples, and relatively low microbial biomass per
512 leaf that that must be first dislodged from tightly-adhered biofilms. Signal from messenger RNA
513 in metatranscriptome analysis is masked by abundant ribosomal RNA signal, leading to further
514 challenge. Because of the combination of all of these challenges, much of our understanding of
515 the phyllosphere, as the largest surface area of microbial habitation on Earth (Peñuelas and
516 Terradas, 2014), has been learned from studies that employ model hosts and synthetic or
517 model microbial communities in controlled settings, or from description of the community
518 structure by sequencing of marker genes, amplified and bioinformatically depleted of
519 chloroplast genes to overcome the challenges of low signal and host contamination.

520 Here, we report optimized laboratory protocols (to minimize host and chloroplast
521 signals) combined with a genome-centric bioinformatic approach to perform focused functional
522 gene and transcript analysis of seasonally dynamic yet persistent phyllosphere microbiome
523 members. To our knowledge, this is the first bacterial/archaeal metatranscriptomic work
524 performed on the leaf phyllosphere of field-grown crops. Other recent leaf metatranscriptome
525 studies have focused the viral communities of tomato and pepper (Choi *et al.*, 2020), soybean
526 (Marzano and Domier, 2016), and rice (Chao *et al.*, 2020 Preprint). A key strength of this work
527 is the integration of phyllosphere metagenome and metatranscriptome data, leveraging the
528 higher coverage of the metagenomes with the activity information available from the
529 metatranscriptomes. Despite that the coverage of the MAGs was relatively limited (due to
530 substantial host and ribosomal DNA contamination), the analysis proved successful by
531 integrating both datasets and focusing on genome-centric interpretation. Thus, there are likely
532 many more prevalent and functionally active populations of the phyllosphere that were not
533 captured in this study, including those players previously known to be key in the phyllosphere,
534 such as *Methylobacterium* spp. (e.g. Delmotte *et al.*, 2009; Vorholt, 2012). Substantial
535 additional sequencing effort or an enrichment strategy would improve signal for a cultivation-
536 independent approach to target those players. While the use of genome-centric approaches
537 has the obvious shortcoming that we have obviously not captured every microbiome member,
538 our approach does allow us to link actively transcribed functions to specific microbial
539 membership. Furthemore, the functional genes and activities documented here are logical
540 given current understanding of microbial adaptation to the host and phyllosphere environment.

541 Overall, this work provides evidence of specific phyllosphere functions that are
542 seasonally activated in a temperate agroecosystem and suggests several hypotheses of
543 important host-microbe interactions in the phyllosphere, for example via central metabolism,
544 isoprene, and stress response engagements. This research contributes to our broad
545 understanding of the dynamics and activities of phyllosphere microbial communities, and
546 points to specific microbial functions to target that could prove useful for plant-microbiome
547 management.

548

549 **References**

550 Andrews, J.H. and Harris, R.F. (2000) The ecology and biogeography of microorganisms on plant
551 surfaces. *Annu Rev Phytopathol* **38**: 145–180.
552 Aziz, R.K., Bartels, D., Best, A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008) The RAST
553 Server: Rapid annotations using subsystems technology. *BMC Genomics* **9**: 1–15.
554 Bahulikar, R.A., Torres-Jerez, I., Worley, E., Craven, K., and Udvardi, M.K. (2014) Diversity of
555 nitrogen-fixing bacteria associated with switchgrass in the native tallgrass prairie of
556 Northern Oklahoma. *Appl Environ Microbiol* **80**: 5636–5643.
557 Beilsmith, K., Thoen, M.P.M., Brachi, B., Gloss, A.D., Khan, M.H., and Bergelson, J. (2019)
558 Genome-wide association studies on the phyllosphere microbiome: Embracing complexity
559 in host–microbe interactions. *Plant J* **97**: 164–181.
560 Bell, T.H., Hockett, K.L., Alcalá-Briseño, R.I., Barbercheck, M., Beattie, G.A., Bruns, M.A., et al.
561 (2019) Manipulating wild and tamed phytobiomes: Challenges and opportunities.
562 *Phytobiomes J* **3**: 3–21.
563 Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina
564 sequence data. *Bioinformatics* **30**: 2114–2120.
565 Bowsher, A.W., Benucci, G.M.N., Bonito, G., and Shade, A. (2020) Seasonal Dynamics of Core
566 Fungi in the Switchgrass Phyllosphere, and Co-Occurrence with Leaf Bacteria. *Phytobiomes*
567 *J* PBIOMES-07-20-0.
568 Bringel, F. and Couée, I. (2015) Pivotal roles of phyllosphere microorganisms at the interface
569 between plant functioning and atmospheric trace gas dynamics. *Front Microbiol* **6**:.
570 Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F.,
571 et al. (2012) Revealing structure and assembly cues for *Arabidopsis* root-inhabiting
572 bacterial microbiota. *Nature* **488**: 91–95.
573 Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E.V.L., and Schulze-Lefert, P. (2013)
574 Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* **64**: 807–
575 838.
576 Busby, P.E., Soman, C., Wagner, M.R., Friesen, M.L., Kremer, J., Bennett, A., et al. (2017)
577 Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol*
578 **15**: 1–14.
579 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., et
580 al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per
581 sample. *Proc Natl Acad Sci* **108**: 4516–4522.
582 Carlström, C.I., Field, C.M., Bortfeld-Miller, M., Müller, B., Sunagawa, S., and Vorholt, J.A. (2019)
583 Synthetic microbiota reveal priority effects and keystone strains in the *Arabidopsis*
584 phyllosphere. *Nat Ecol Evol* **3**: 1445–1454.
585 Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., et al. (2019)
586 Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Microbiol*
587 **17**: 569–586.
588 Chao, S., Wang, H., Yan, Q., Chen, L., Chen, G., Wu, L., et al. (2020) Metatranscriptomics and
589 small RNA analysis revealed that viral covert coinfection resulted in disease symptoms
590 reminiscent of rice sterility. *Prepr available Res Sq*.
591 Chen, T., Nomura, K., Wang, X., Sohrabi, R., Xu, J., Yao, L., et al. (2020) A plant genetic network
592 for preventing dysbiosis in the phyllosphere. *Nature* **580**: 653–657.

593 Choi, H., Jo, Y., Cho, W.K., Yu, J., Tran, P.T., Salaipeth, L., et al. (2020) Identification of viruses
594 and viroids infecting tomato and pepper plants in vietnam by metatranscriptomics. *Int J*
595 *Mol Sci* **21**: 1–16.
596 Choi, J., Yang, F., Stepanauskas, R., Cardenas, E., Garoutte, A., Williams, R., et al. (2017)
597 Strategies to improve reference databases for soil microbiomes. *ISME J* **11**: 829–834.
598 Considine, M.J. and Foyer, C.H. (2014) Redox regulation of plant development. *Antioxid Redox*
599 *Signal* **21**: 1305–1326.
600 Crombie, A.T., Larke-Mejia, N.L., Emery, H., Dawson, R., Pratscher, J., Murphy, G.P., et al. (2018)
601 Poplar phyllosphere harbors disparate isoprene-degrading bacteria. *Proc Natl Acad Sci U S*
602 *A* **115**: 13081–13086.
603 Dai, J., Dai, W., Qiu, C., Yang, Z., Zhang, Y., Zhou, M., et al. (2015) Unraveling adaptation of
604 *Pontibacter korlensis* to radiation and infertility in desert through complete genome and
605 comparative transcriptomic analysis. *Sci Rep* **5**: 1–9.
606 Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., et al. (2009)
607 Community proteogenomics reveals insights into the physiology of phyllosphere bacteria.
608 *Proc Natl Acad Sci* **106**: 16428–16433.
609 Dorokhov, Y.L., Sheshukova, E. V., and Komarova, T. V. (2018) Methanol in plant life. *Front Plant*
610 *Sci* **871**: 1–6.
611 Dunivin, T.K., Choi, J., Howe, A., and Shade, A. (2019) RefSoil+: a Reference Database for Genes
612 and Traits of Soil Plasmids. *mSystems* **4**: 1–10.
613 Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., et al.
614 (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc*
615 *Natl Acad Sci U S A* **112**: E911-20.
616 Eller, A.S.D., Sekimoto, K., Gilman, J.B., Kuster, W.C., de Gouw, J.A., Monson, R.K., et al. (2011)
617 Volatile organic compound emissions from switchgrass cultivars used as biofuel crops.
618 *Atmos Environ* **45**: 3333–3337.
619 Fataftah, N., Mohr, C., Hajirezaei, M.R., Wirén, N. von, and Humbeck, K. (2018) Changes in
620 nitrogen availability lead to a reprogramming of pyruvate metabolism. *BMC Plant Biol* **18**:
621 1–15.
622 Ferner, E., Rennenberg, H., and Kreuzwieser, J. (2012) Effect of flooding on C metabolism of
623 flood-tolerant (*Quercus robur*) and non-tolerant (*Fagus sylvatica*) tree species. *Tree Physiol*
624 **32**: 135–145.
625 Gopal, M. and Gupta, A. (2016) Microbiome selection could spur next-generation plant
626 breeding strategies. *Front Microbiol* **7**: 1971.
627 Gottschlich, L., Geiser, P., Bortfeld-Miller, M., Field, C.M., and Vorholt, J.A. (2019) Complex
628 general stress response regulation in *Sphingomonas melonis* Fr1 revealed by
629 transcriptional analyses. *Sci Rep* **9**: 1–13.
630 Grady, K.L., Sorensen, J.W., Stopnisek, N., Guittar, J., and Shade, A. (2019) Assembly and
631 seasonality of core phyllosphere microbiota on perennial biofuel crops. *Nat Commun* **10**:.
632 Gupta, V.V.S.R., Zhang, B., Penton, C.R., Yu, J., and Tiedje, J.M. (2019) Diazotroph Diversity and
633 nitrogen fixation in summer active perennial grasses in a mediterranean region agricultural
634 soil. *Front Mol Biosci* **6**: 115.
635 Hacquard, S., Garrido-Oter, R., González, A., Spaepen, S., Ackermann, G., Lebeis, S., et al. (2015)
636 Microbiota and host nutrition across plant and animal kingdoms. *Cell Host Microbe* **17**:

603–616.
Hardoim, P.R., van Overbeek, L.S., Berg, G., Pirttilä, A.M., Compant, S., Campisano, A., et al. (2015) The hidden world within plants: Ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiol Mol Biol Rev* **79**: 293–320.
Harty, C.E., Martins, D., Doing, G., Mould, D.L., Clay, M.E., Occhipinti, P., et al. (2019) Ethanol stimulates trehalose production through a SpoT-DksA-AlgU-dependent pathway in *Pseudomonas aeruginosa*. 1–21.
Haskett, T.L., Tkacz, A., and Poole, P.S. (2020) Engineering rhizobacteria for sustainable agriculture. *ISME J* **15**: 949–964.
Heaton, E.A., Dohleman, F.G., and Long, S.P. (2008) Meeting US biofuel goals with less land: The potential of Miscanthus. *Glob Chang Biol* **14**: 2000–2014.
Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., and Wang, Z. (2019) MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **2019**: 1–13.
Kerchev, P., De Smet, B., Waszczak, C., Messens, J., and Van Breusegem, F. (2015) Redox strategies for crop improvement. *Antioxidants Redox Signal* **23**: 1186–1205.
El Khawand, M., Crombie, A.T., Johnston, A., Vavlline, D. V., McAuliffe, J.C., Latone, J.A., et al. (2016) Isolation of isoprene degrading bacteria from soils, development of isoA gene probes and identification of the active isoprene-degrading soil community using DNA-stable isotope probing. *Environ Microbiol* **18**: 2743–2753.
Kimmerer, T.W. and Kozlowski, T.T. (1982) Ethylene, ethane, acetaldehyde, and ethanol production by plants under stress. *Plant Physiol* **69**: 840–847.
Kimmerer, T.W. and MacDonald, R.C. (1987) Acetaldehyde and ethanol biosynthesis in leaves of plants. *Plant Physiol* **84**: 1204–1209.
Koskella, B. (2020) The phyllosphere. *Curr Biol* **30**: R1143–R1146.
Kuzyakov, Y. and Razavi, B.S. (2019) Rhizosphere size and shape: Temporal dynamics and spatial stationarity. *Soil Biol Biochem* **135**: 343–360.
Kwak, M.J., Jeong, H., Madhaiyan, M., Lee, Y., Sa, T.M., Oh, T.K., and Kim, J.F. (2014) Genome information of *Methylobacterium oryzae*, a plant-probiotic methylotroph in the phyllosphere. *PLoS One* **9**:.
Lacerda-Júnior, G. V., Noronha, M.F., Cabral, L., Delforno, T.P., De Sousa, S.T.P., Fernandes-Júnior, P.I., et al. (2019) Land use and seasonal effects on the soil microbiome of a Brazilian dry forest. *Front Microbiol* **10**: 1–14.
Laforest-Lapointe, I., Paquette, A., Messier, C., and Kembel, S.W. (2017) Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature* **546**: 145–147.
Lajoie, G., Maglione, R., and Kembel, S.W. (2020) Adaptive matching between phyllosphere bacteria and their tree hosts in a neotropical forest. *Microbiome* **8**: 1–10.
Lambais, M.R., Barrera, S.E., Santos, E.C., Crowley, D.E., and Jumpponen, A. (2017) Phyllosphere metaproteomes of trees from the Brazillian Atlantic forest show high levels of functional redundancy. *Microb Ecol* **73**: 123–134.
Langholtz, M., Stokes, B., and Eaton, L. (2016) 2016 billion-ton report: Advancing domestic resources for a thriving bioeconomy (Executive Summary). *Ind Biotechnol* **12**: 282–289.
Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.

Laskowska, E., Dorota Kuczyńska-Wiśnik, ·, and Kuczyńska-Wiśnik, D. (2020) New insight into the mechanisms protecting bacteria during desiccation. *Curr Genet* **66**: 313–318.
Levy, A., Conway, J.M., Dangl, J.L., and Woyke, T. (2018) Elucidating bacterial gene functions in the plant microbiome. *Cell Host Microbe* **24**: 475–485.
Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., Clingenpeel, S., Herrera Paredes, S., Miao, J., et al. (2018) Genomic features of bacterial adaptation to plants. *Nat Genet* **50**: 138–150.
Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment / Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
Lindow, S.E. and Brandl, M.T. (2003) Microbiology of the phyllosphere. *Appl Environ Microbiol* **69**: 1875–1883.
Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86–90.
Marzano, S.Y.L. and Domier, L.L. (2016) Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. *Virus Res* **213**: 332–342.
McGenity, T.J., Crombie, A.T., and Murrell, J.C. (2018) Microbial cycling of isoprene, the most abundantly produced biological volatile organic compound on Earth. *ISME J* **12**: 931–941.
Morrison, E.C., Drewer, J., and Heal, M.R. (2016) A comparison of isoprene and monoterpene emission rates from the perennial bioenergy crops short-rotation coppice willow and Miscanthus and the annual arable crops wheat and oilseed rape. *GCB Bioenergy* **8**: 211–225.
Muller, D.B., Schubert, O.T., Rost, H., Aebersold, R., and Vorholt, J.A. (2016) Systems-level proteomics of two ubiquitous leaf commensals reveals complementary adaptive traits for phyllosphere colonization. *Mol Cell Proteomics* **15**: 3256–3269.
Müller, D.B., Vogel, C., Bai, Y., and Vorholt, J.A. (2016) The Plant Microbiota: Systems-Level Insights and Perspectives. *Annu Rev Genet* **50**: 120215–034952.
Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.
Peñuelas, J. and Terradas, J. (2014) The foliar microbiome. *Trends Plant Sci* **19**: 278–280.
Quinlan, A.R. and Hall, I.M. (2010) BEDTools : a flexible suite of utilities for comparing genomic features. **26**: 841–842.
Rastogi, G., Coaker, G.L., and Leveau, J.H.J.J. (2013) New insights into the structure and function of phyllosphere microbiota through high-throughput molecular approaches. *FEMS Microbiol Lett* **348**: 1–10.
Remus-Emsermann, M.N.P. and Schlechter, R.O. (2018) Phyllosphere microbiology: at the interface between microbial individuals and the plant host. *New Phytol* **218**: 1327–1333.
Robertson, G.P., Hamilton, S.K., Barham, B.L., Dale, B.E., Izaurralde, R.C., Jackson, R.D., et al. (2017) Cellulosic biofuel contributions to a sustainable energy future: Choices and outcomes. *Science (80- )* **356**: 1–9.
Rodriguez-R, L.M., Gunturu, S., Tiedje, J.M., Cole, J.R., and Konstantinidis, K.T. (2018) Nonpareil

3: Fast estimation of metagenomic coverage and sequence diversity. *mSystems* **3**: 1–9.

Roley, S.S., Duncan, D.S., Liang, D., Garoutte, A., Jackson, R.D., Tiedje, J.M., and Robertson, G.P. (2018) Associative nitrogen fixation (ANF) across a nitrogen input gradient. *PLoS One* **13**: 1–19.

Roley, S.S., Xue, C., Hamilton, S.K., Tiedje, J.M., and Robertson, G.P. (2019) Isotopic evidence for episodic nitrogen fixation in switchgrass (*Panicum virgatum* L.). *Soil Biol Biochem* **129**: 90–98.

Shade, A. and Stopnisek, N. (2019) Abundance-occupancy distributions to prioritize plant core microbiome membership. *Curr Opin Microbiol* **49**: 50–58.

Sharkey, T.D., Loreto, F., and Delwiche, C. (1991) High carbon dioxide and sun/shade effects on isoprene emission from oak and aspen tree leaves. *Plant, Cell Environ* **14**: 333–338.

Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from plants: Why and how. *Ann Bot* **101**: 5–18.

Sharkey, T.D. and Yeh, S. (2001) Isoprene emission from plants. *Plant Mol Biol* **52**: 407–436.

Singer, E., Bonnette, J., Kenaley, S.C., Woyke, T., and Juenger, T.E. (2019) Plant compartment and genetic variation drive microbiome composition in switchgrass roots. *Environ Microbiol Rep* **11**: 185–195.

Sivy, T.L., Shirk, M.C., and Fall, R. (2002) Isoprene synthase activity parallels fluctuations of isoprene release during growth of *Bacillus subtilis*. *Biochem Biophys Res Commun* **294**: 71–75.

Stopnisek, N. and Shade, A. (2021) Persistent microbiome members in the common bean rhizosphere: an integrated analysis of space, time, and plant genotype. *ISME J*.

Suzuki, Y., Makino, A., and Mae, T. (2001) An efficient method for extraction of RNA from rice leaves at different ages using benzyl chloride. *J Exp Bot* **52**: 1575–1579.

Taylor, L., Nunes-Nesi, A., Parsley, K., Leiss, A., Leach, G., Coates, S., et al. (2010) Cytosolic pyruvate, orthophosphate dikinase functions in nitrogen remobilization during leaf senescence and limits individual seed growth and nitrogen content. *Plant J* **62**: 641–652.

Toju, H., Peay, K.G., Yamamichi, M., Narisawa, K., Hiruma, K., Naito, K., et al. (2018) Core microbiomes for sustainable agroecosystems. *Nat Plants* **4**: 247–257.

Trivedi, P., Leach, J.E., Tringe, S.G., Sa, T., and Singh, B.K. (2020) Plant–microbiome interactions: from community assembly to plant health. *Nat Rev Microbiol* **18**: 607–621.

Urrejola, C., Alcorta, J., Salas, L., Vásquez, M., Polz, M.F., Vicuña, R., and Díez, B. (2019) Genomic features for desiccation tolerance and sugar biosynthesis in the extremophile *gloeocapsopsis* sp. UTEX B3054. *Front Microbiol* **10**: 1–11.

Vorholt, J.A. (2012) Microbial life in the phyllosphere. *Nat Rev Microbiol* **10**: 828–840.

Wang, N.R. and Haney, C.H. (2020) Harnessing the genetic potential of the plant microbiome. *Biochem (Lond)* **42**: 20–25.

Wingler, A., Delatte, T.L., O'Hara, L.E., Primavesi, L.F., Jhurreea, D., Paul, M.J., and Schluepmann, H. (2012) Trehalose 6-phosphate is required for the onset of leaf senescence associated with high carbon availability. *Plant Physiol* **158**: 1241–1251.

Xu, L., Naylor, D., Dong, Z., Simmons, T., Pierroz, G., Hixson, K.K., et al. (2018) Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proc Natl Acad Sci* **115**: E4284–E4293.

Zhalnina, K., Louie, K.B., Hao, Z., Mansoori, N., da Rocha, U.N., Shi, S., et al. (2018) Dynamic root

769 exudate chemistry and microbial substrate preferences drive patterns in rhizosphere
770 microbial community assembly. *Nat Microbiol* **3**: 470.
771 Zuo, Z., Weraduwage, S.M., Lantz, A.T., Sanchez, L.M., Weise, S.E., Wang, J., et al. (2019)
772 Isoprene acts as a signaling molecule in gene networks important for stress responses and
773 plant growth. *Plant Physiol* **180**: 124–152.
774
775

**Acknowledgements**

Support for this research was provided by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Awards DE-SC0018409 and DE-FC02-07ER64494), by the National Science Foundation Long-term Ecological Research Program (DEB 1637653 and 1832042) at the Kellogg Biological Station, Michigan State University AgBioResearch, and by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (US Department of Energy, Office of Science, Office of Biological and Environmental Research under award number DE-SC0018420). This work was also supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research and in part by the University of Wisconsin-Madison Wisconsin Energy Institute as supported by GLBRC Information Services. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. AS acknowledges support by the USDA National Institute of Food and Agriculture and Michigan State University AgBioResearch. NS acknowledges support from the Michigan State University Plant Resilience Institute.

**Author contributions**

KLG nd AS conceived and designed experiments; NS, KLG and AS performed the experiments; ACH, NS, SKD, FMY and AS analyzed the data; ACH, NS, SKD, FMY, KLG and AS contributed materials/analysis tools; and ACH, NS, SKD, FMY, KLG and AS wrote the paper.

**Competing Interests statement**

798 The authors declair no competing interests.

799

800 **Figures**

801 **Figure 1**. Summary of 16 high-quality genomes assembled and binned from switchgrass and
802 miscanthus phyllosphere metagenomes. Leaf samples were collected over two growing
803 seasons in 2016 and 2017 at Kellogg Biological Station, Michigan US. Taxonomic annotation,
804 quality/contamination assessment, and genome size of MAGs were determined using the
805 software checkM.

806

807 **Figure 2**. Estimated abundances of phyllosphere MAGs in (A) metagenomes of miscanthus and
808 switchgrass metagenomes sampled across the 2016 growing season and (B)
809 metatranscriptomes of switchgrass in 2016 and 2017 growing season. (A) Abundances are
810 median basepair read coverage for each MAG in metagenomes or (B) cumulative median
811 basepair coverage of all ORFs in each MAG in metatrancriptomes divided by total number of
812 reads in each sample.

813

814 **Figure 3**. Estimated abundances of functions associated with phyllosphere MAGs across the
815 2016 and 2017 growing season. Functional roles that were enriched and "core" in phyllosphere
816 metatranscriptomes. Enriched was defined as ratios of late to early abundances greater than
817 mean ($\mu_{ratio}$=35). Core was defined as the detection of the functional role in greater than half of
818 MAGs, where count is the number of MAGs where function was detected. The ratio of
819 functions which were not detected in the early sampling dates are shown on the top of the
820 graph.

821

822 **Figure 4**. Transcript abundances of MAGs with coherent seasonal activity dynamics (see **Figure**
823 **S5** for cluster memberships**)**, annotated by functional subsystem. Data are from switchgrass
824 metatranscriptomes sampled in 2016 and 2017. Transcripts from MAGs that fell within the
825 same group (i.e., had similar activity dynamics and annotations over the season) were
826 combined by subsystem. Rows are the groups of coherent MAGs (labeled as 1 through 6) based
827 on seasonal activities, and columns are the cumulative subsystem transcripts observed for each
828 sampling month. Open circles indicate that the function was not detected in a specific month
829 for a group cluster but is detected during the season.

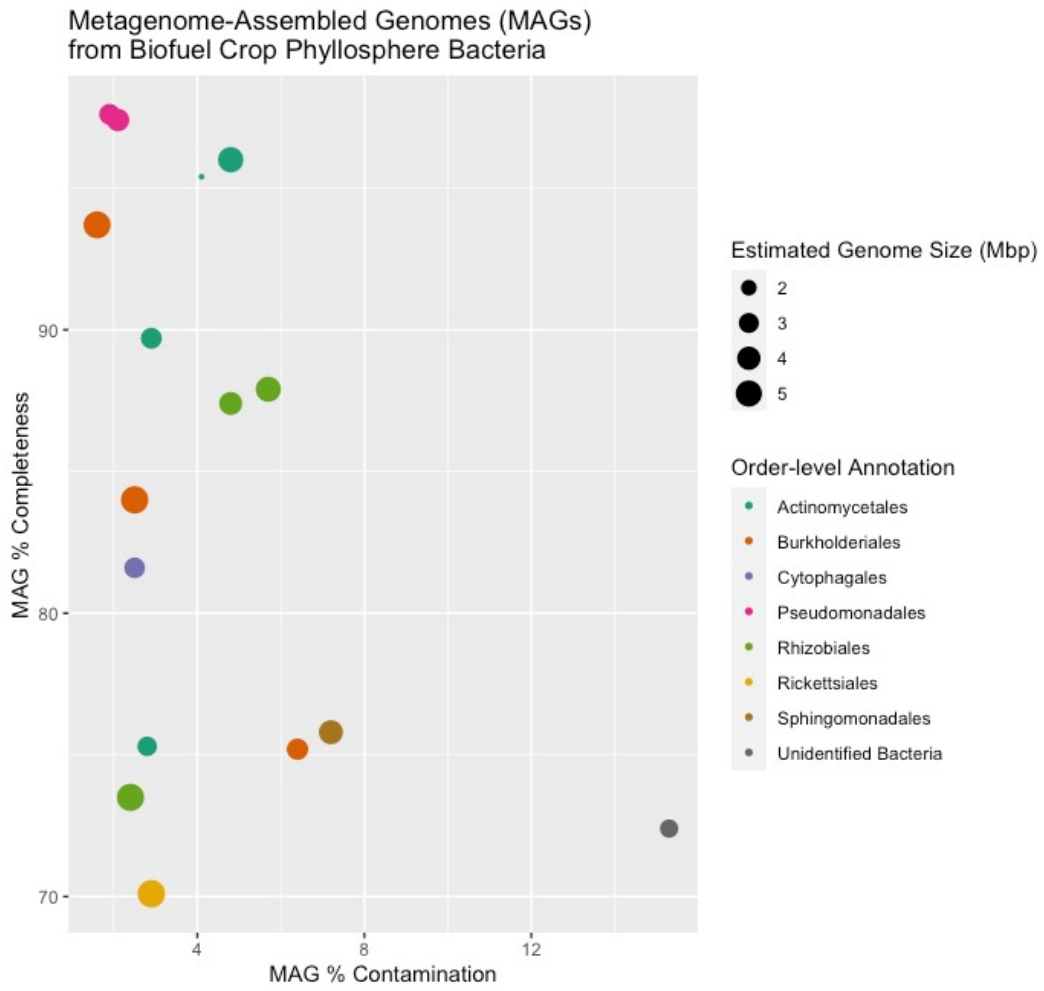
830

831 **Datasets**
832 **Dataset 1**. List of metagenome-assembled genomes from the dataset that were >50%
833 complete and <50% contamination. Excel file.
834
835

836 **Supplementary Figures**
837 **Figure S1**. Estimated abundances of phyllosphere MAGs in (A) metagenomes of miscanthus
838 and switchgrass across the 2016 growing season and (B) metatranscriptomes of switchgrass in
839 the 2016 and 2017 season. Relative abundances are (A) median basepair read coverage for
840 each MAG in metagenomes and (B) cumulative median basepair coverage of all ORFs in each
841 MAG in metatranscriptomes divided by total number of reads in sample.
842
843 **Figure S2**. Contaminating reads in switchgrass leaf metatranscriptomes from (A) plant host RNA
844 reads and (B) microbial rRNA. An average of 52% of plant host RNA and less than 10% of
845 microbial rRNA were identified as contaminant transcripts in the metatranscriptomes.
846
847 **Figure S3**. Estimated transcript abundances of functions associated with phyllosphere MAGs
848 across the 2016 and 2017 growing seasons, color coded by month of collection.
849
850 **Figure S4**. The 1,034 functional roles identified in phyllosphere MAGs that were enriched in
851 August to September relative to May to July, plotted against and the total number of highest-
852 quality MAGs in which each functional role was detected. We used a threshold of detection
853 within at least 13 out of 16 total MAGs to highlight consistent activities of the leaf-associated
854 lifestyle (blue points, see **Table S5**).
855
856 **Figure S5**. Hierarchical clustering identified six clusters of MAGS. Clustering was based on
857 metatranscriptome diversity and abundances as well as their seasonal dynamics.
858
859 **Figure S6**. Estimated transcript abundances of functions related to sulfur metabolism
860 associated with phyllosphere MAGs from 2016 and 2017 growing seasons. Open circles
861 indicate that the function was not detected in a specific month for a group cluster but is
862 detected during the season.
863
864 **Figure S7**. Detection of MAGs in publicly available metagenomes and metatranscriptomes of
865 bioenergy grasses.
866
867

868 **Supplementary Tables**
869
870 **Table S1**. Fungal genomes used for filtering metagenome reads to remove eukaryotic
871 contamination. Genomes were selected to use for filtering based on the taxonomic identities
872 of prevalent fungal taxa detected in our previous ITS2 amplicon survey that was conducted at
873 the same location (Bowsher *et al.*, 2020).
874
875 **Table S2**. Collection of 16 highest-quality ("gold") metagenome-assembled genomes (MAGs)
876 from switchgrass and miscanthus phyllosphere. These focal MAGs were in >10% of samples,
877 had >70% completeness, and <15.5% contamination.
878
879 **Table S3**. Comparison of distributions of functions in phyllosphere MAGs and Refsoil+ genomes
880 to identify leaf-enriched functions.
881
882 **Table S4**. Comparison of transcript abundances associated with DNA and RNA metabolism by
883 sampling date for switchgrass phyllosphere using the Kruskal-Wallis tests by ranks.
884
885 **Table S5**. Functions that were detected in at least 13 out of the total 16 highest-quality
886 phyllosphere MAGs, and at abundances greater than average enrichment (>35x) in late season
887 (August to September) relative to early season (May to July) for switchgrass.
888

**Figure 1**. Summary of 16 high-quality genomes assembled and binned from switchgrass and miscanthus phyllosphere metagenomes. Leaf samples were collected over two growing seasons in 2016 and 2017 at Kellogg Biological Station, Michigan US. Taxonomic annotation, quality/contamination assessment, and genome size of MAGs were determined using the software checkM.



A

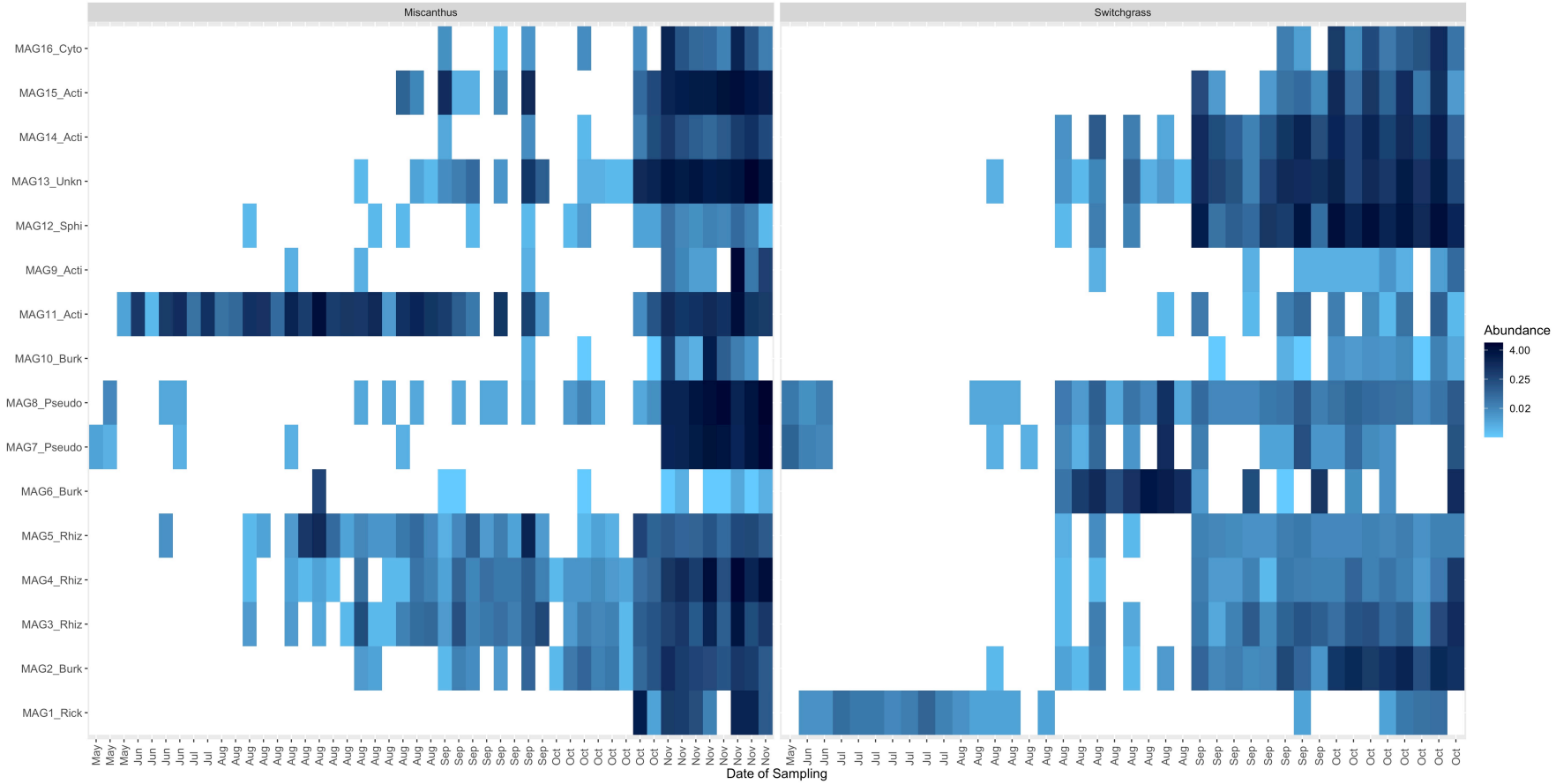



**Figure 2A**. Estimated abundances of phyllosphere MAGs in (A) metagenomes of miscanthus and switchgrass metagenomes sampled across the 2016 growing season and (B) metatranscriptomes of switchgrass in 2016 and 2017 growing season. (A) Abundances are median basepair read coverage for each MAG in metagenomes or (B) cumulative median basepair coverage of all ORFs in each MAG in metatrancriptomes divided by total number of reads in each sample.

B

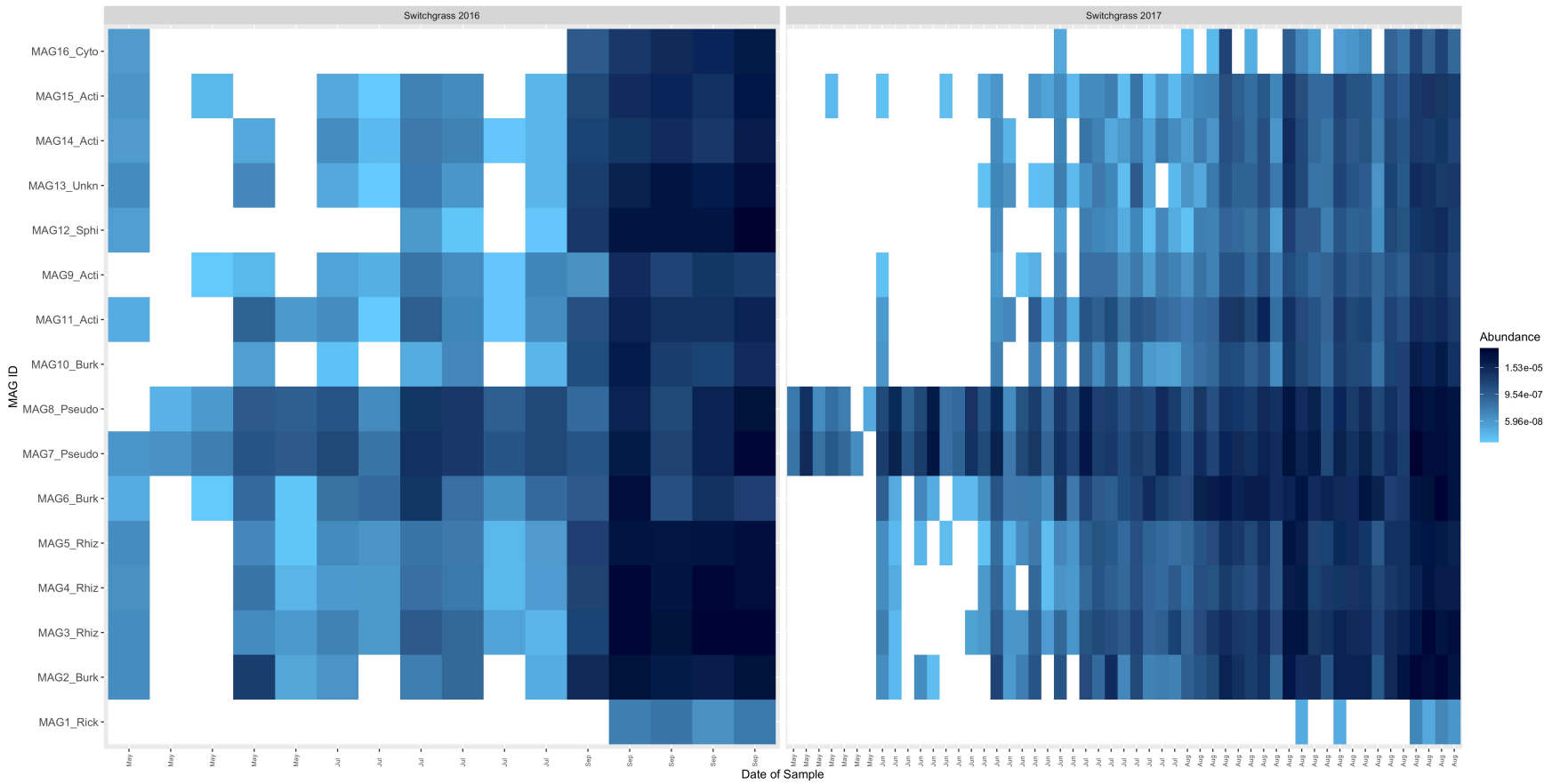



**Figure 2B**. Estimated abundances of phyllosphere MAGs in (A) metagenomes of miscanthus and switchgrass metagenomes sampled across the 2016 growing season and (B) metatranscriptomes of switchgrass in 2016 and 2017 growing season. (A) Abundances are median basepair read coverage for each MAG in metagenomes or (B) cumulative median basepair coverage of all ORFs in each MAG in metatrancriptomes divided by total number of reads in each sample.

**Figure 3**. Estimated abundances of functions associated with phyllosphere MAGs across the 2016 and 2017 growing season. Functional roles that were enriched and "core" in phyllosphere metatranscriptomes. Enriched was defined as ratios of late to early abundances greater than mean ($\mu_{ratio}$=35). Core was defined as the detection of the functional role in greater than half of MAGs, where count is the number of MAGs where function was detected. The ratio of functions which were not detected in the early sampling dates are shown on the top of the graph.

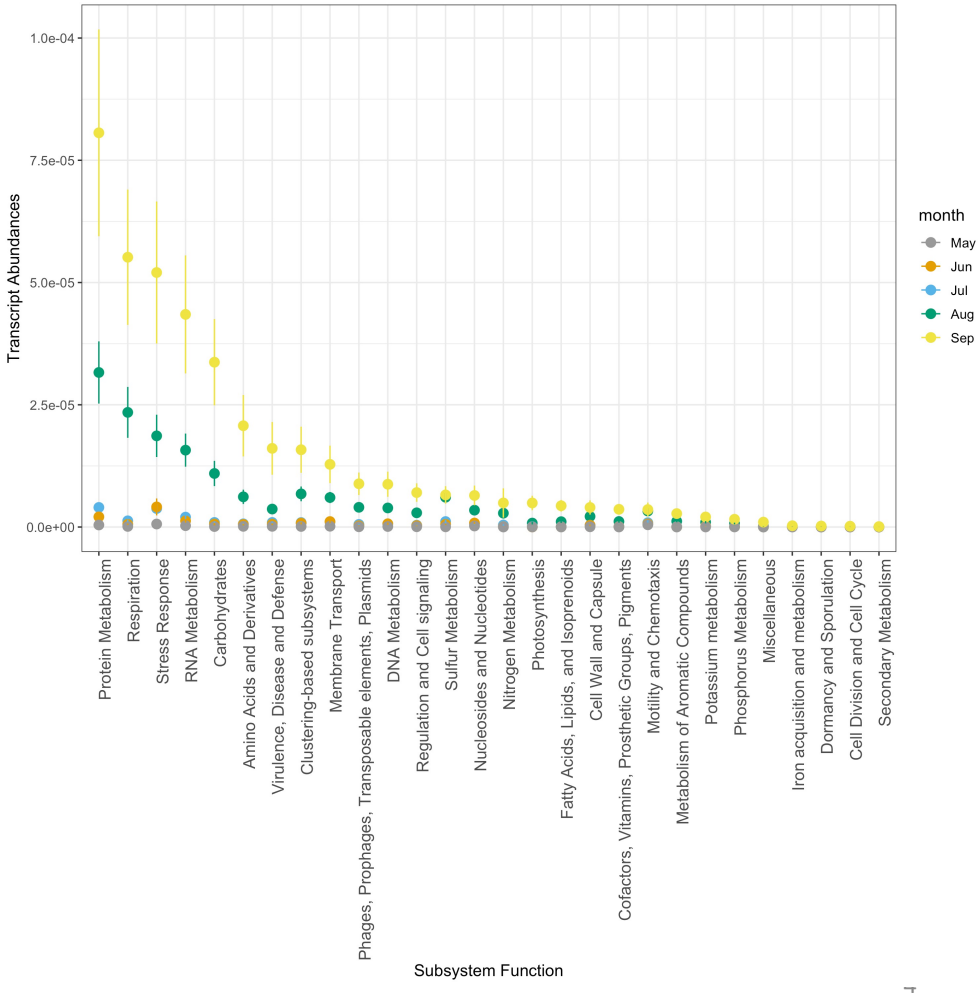


**Figure 4**. Transcript abundances of MAGs with coherent seasonal activity dynamics (see **Figure S5** for cluster memberships**)**, annotated by functional subsystem. Data are from switchgrass metatranscriptomes sampled in 2016 and 2017. Transcripts from MAGs that fell within the same group (i.e., had similar activity dynamics and annotations over the season) were combined by subsystem. Rows are the groups of coherent MAGs (labeled as 1 through 6) based on seasonal activities, and columns are the cumulative subsystem transcripts observed for each sampling month. Open circles indicate that the function was not detected in a specific month for a group cluster but is detected during the season.

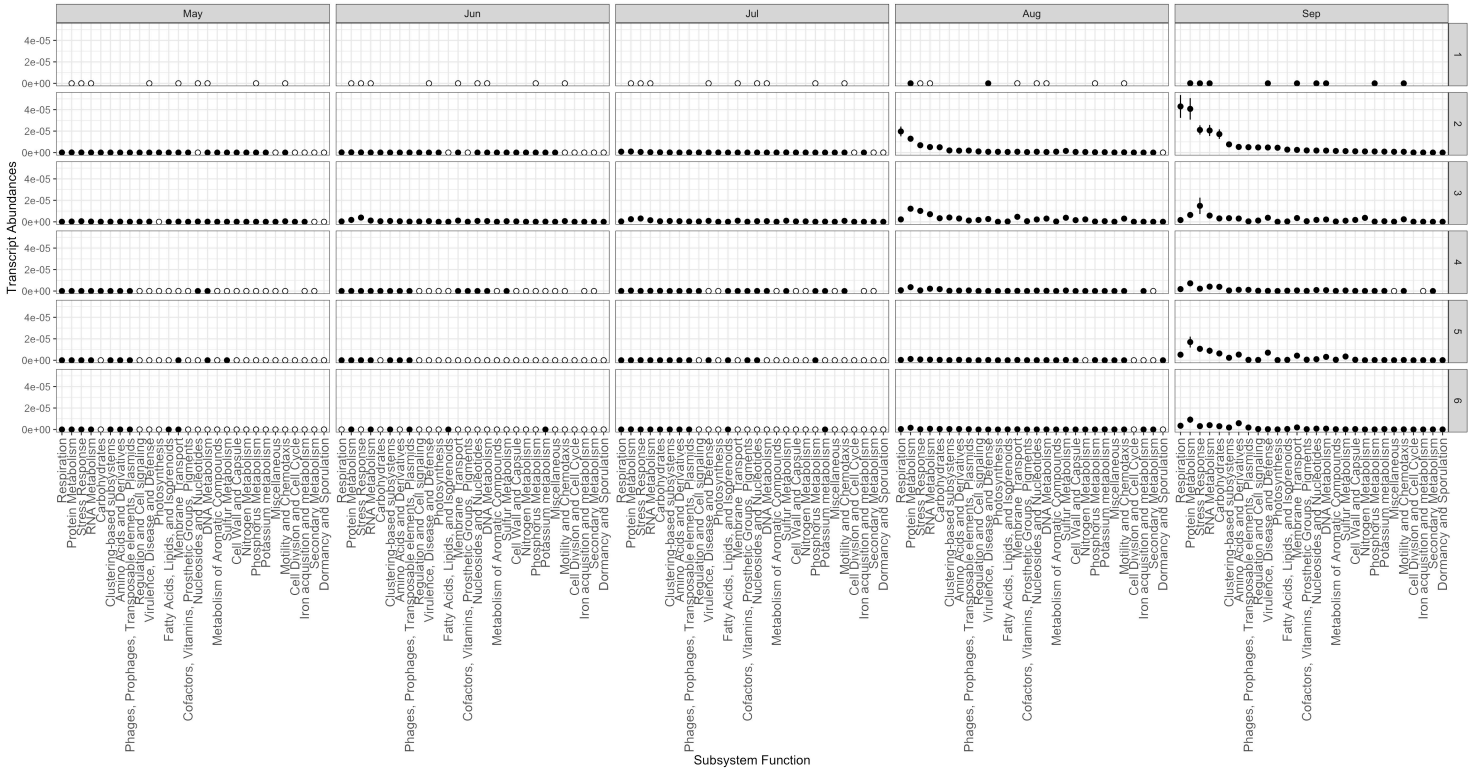


**Figure S1**. Estimated abundances of phyllosphere MAGs in (A) metagenomes of miscanthus and switchgrass across the 2016 growing season and (B) metatranscriptomes of switchgrass in the 2016 and 2017 season. Relative abundances are (A) median basepair read coverage for each MAG in metagenomes and (B) cumulative median basepair coverage of all ORFs in each MAG in metatranscriptomes divided by total number of reads in sample. Open circles indicate that the MAG was not detected in a specific month for a group cluster but is detected during the season.

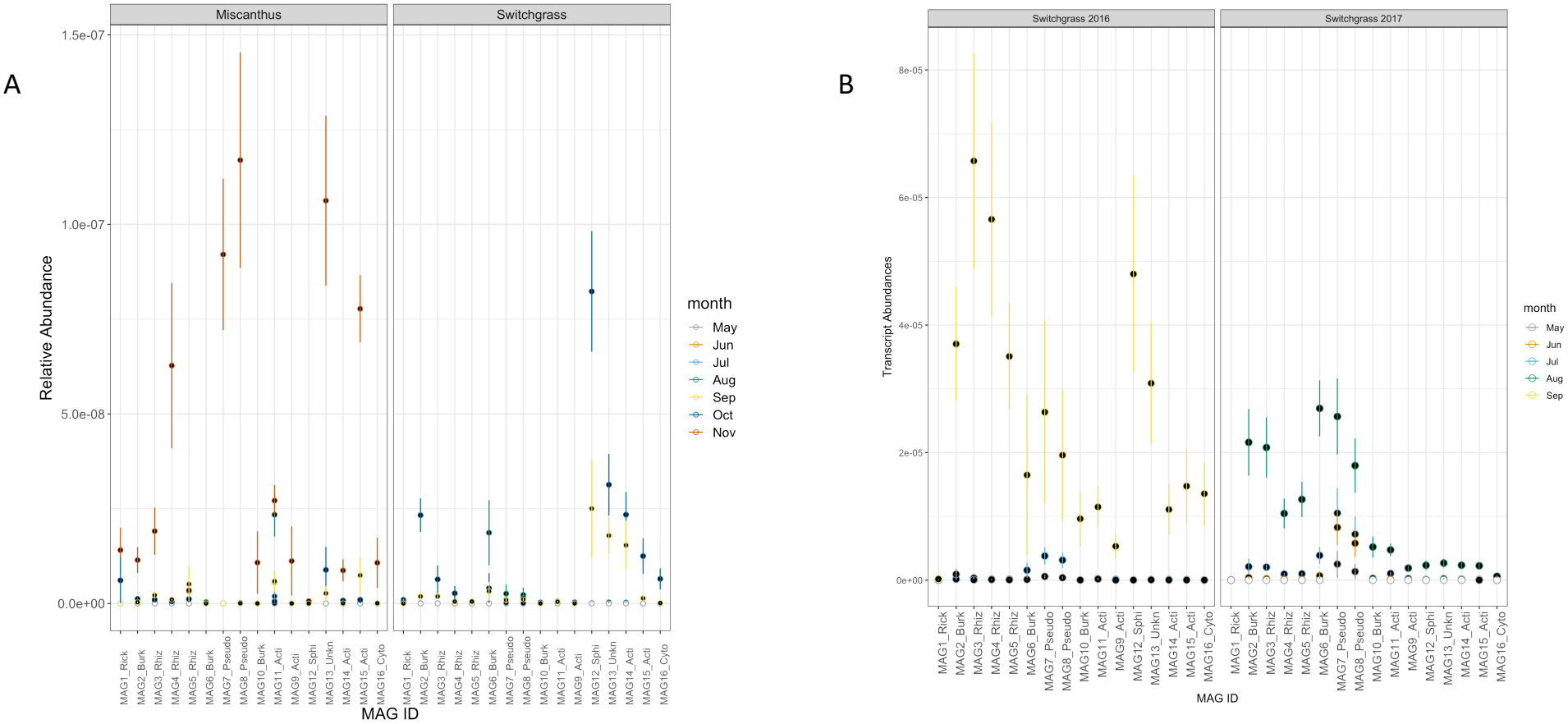


**Figure S2**. Contaminating reads in switchgrass leaf metatranscriptomes from (A) plant host RNA reads and (B) microbial rRNA. An average of 52% of plant host RNA and less than 10% of microbial rRNA were identified as contaminant transcripts in the metatranscriptomes

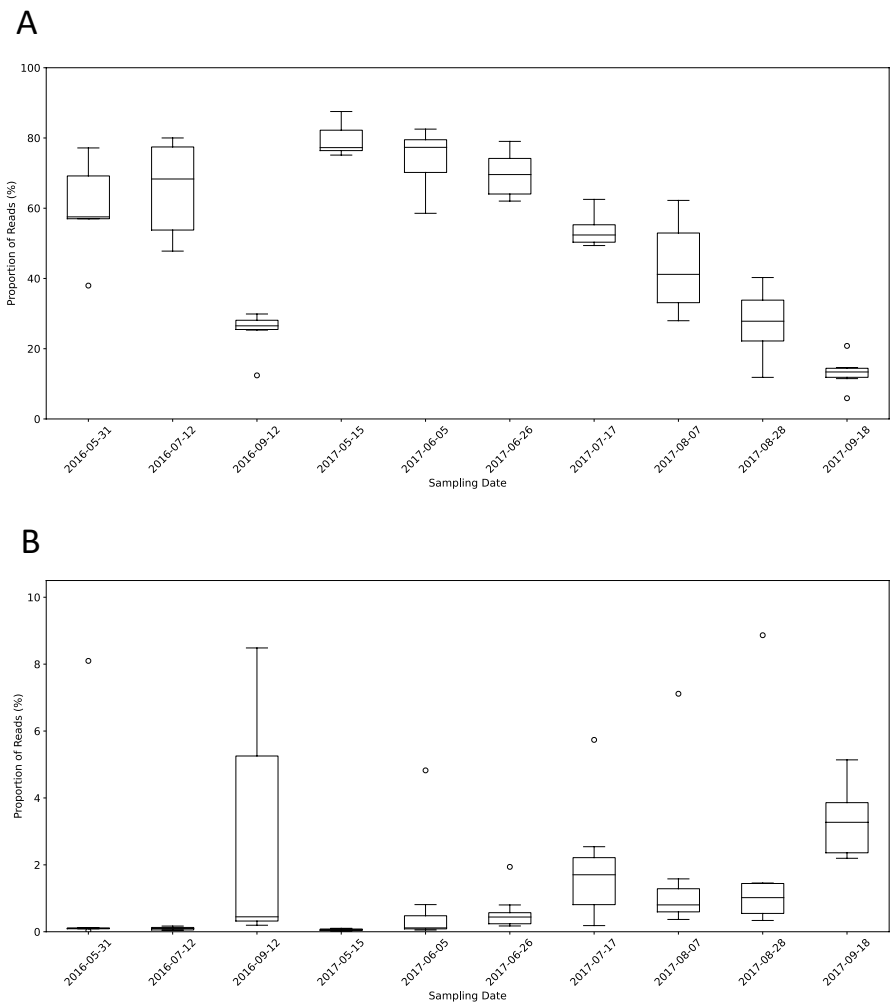


**Figure S3**. Estimated transcript abundances of functions associated with phyllosphere MAGs across the 2016 and 2017 growing seasons, color coded by month of collection.

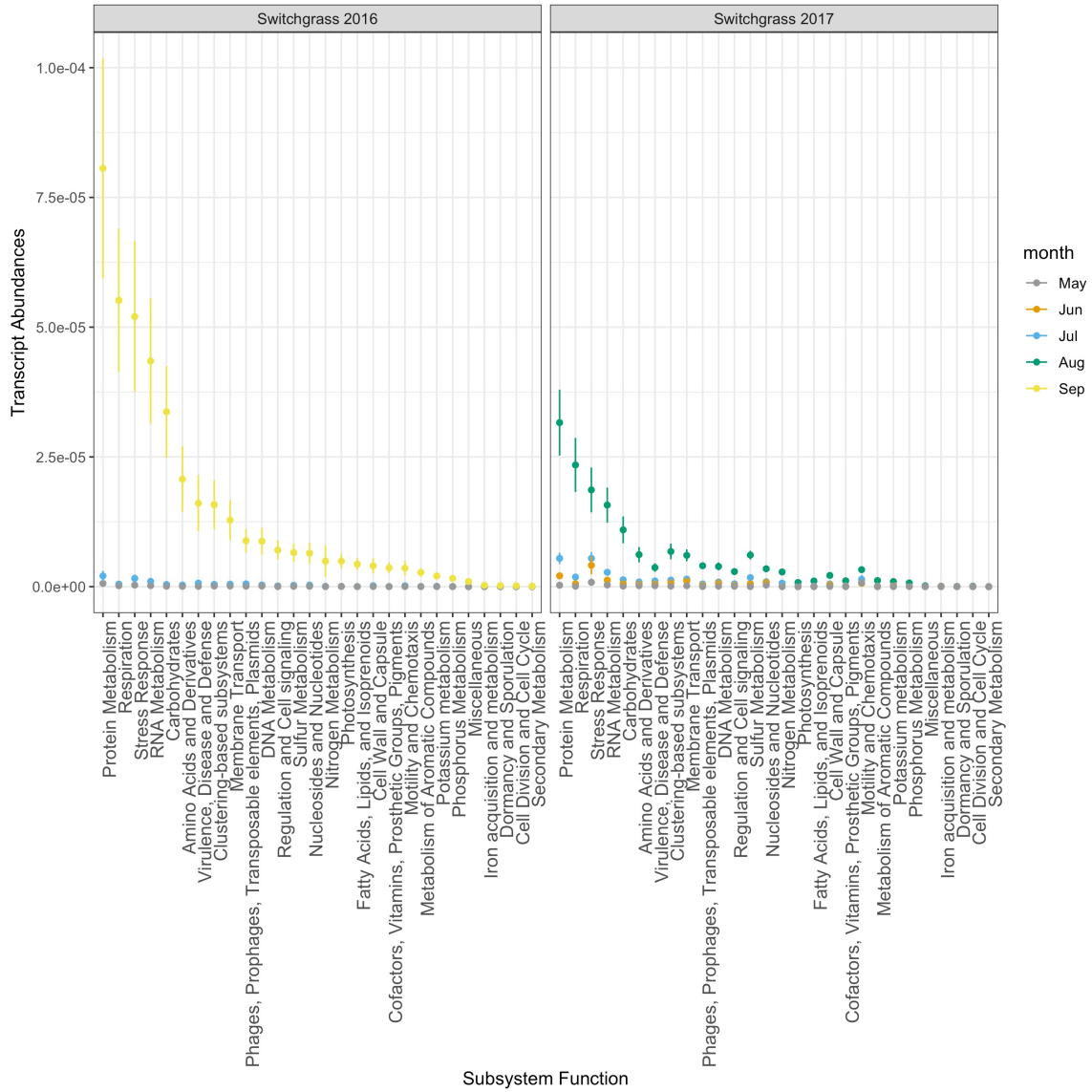


**Figure S4**. The 1,034 functional roles identified in phyllosphere MAGs that were enriched in August to September relative to May to July, plotted against and the total number of highest-quality MAGs in which each functional role was detected. We used a threshold of detection within at least 13 out of 16 total MAGs to highlight consistent activities of the leaf-associated lifestyle (blue points, see **Table S5**).

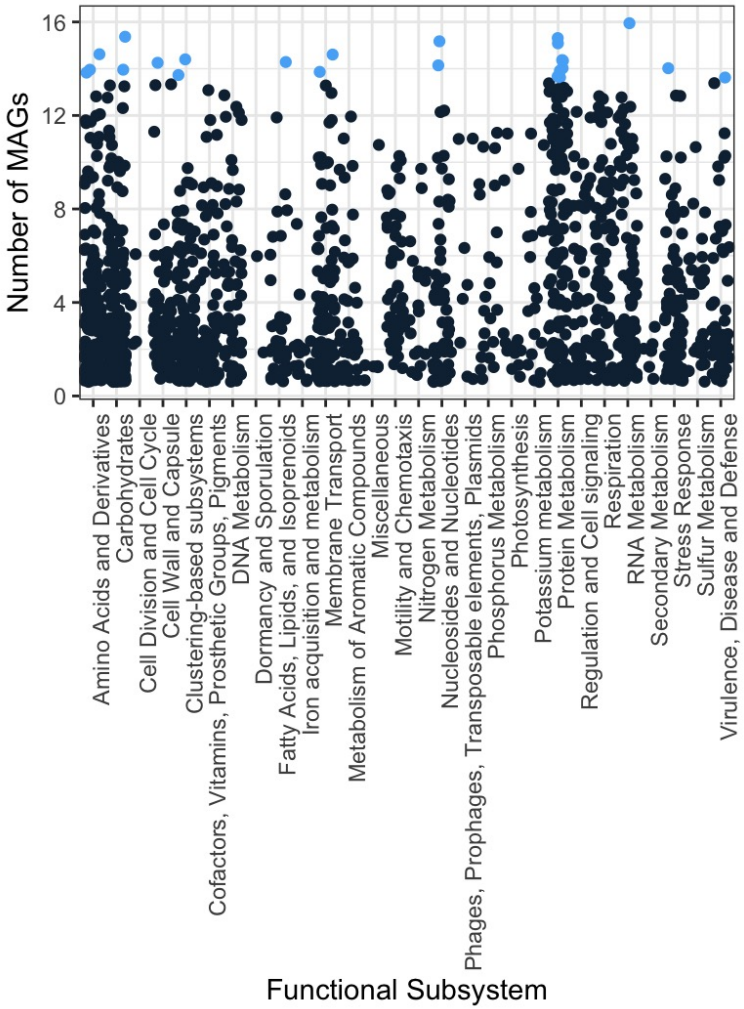


**Figure S5**. Hierarchical clustering identified six clusters of MAGS. Clustering was based on metatranscriptome diversity and abundances as well as their seasonal dynamics.

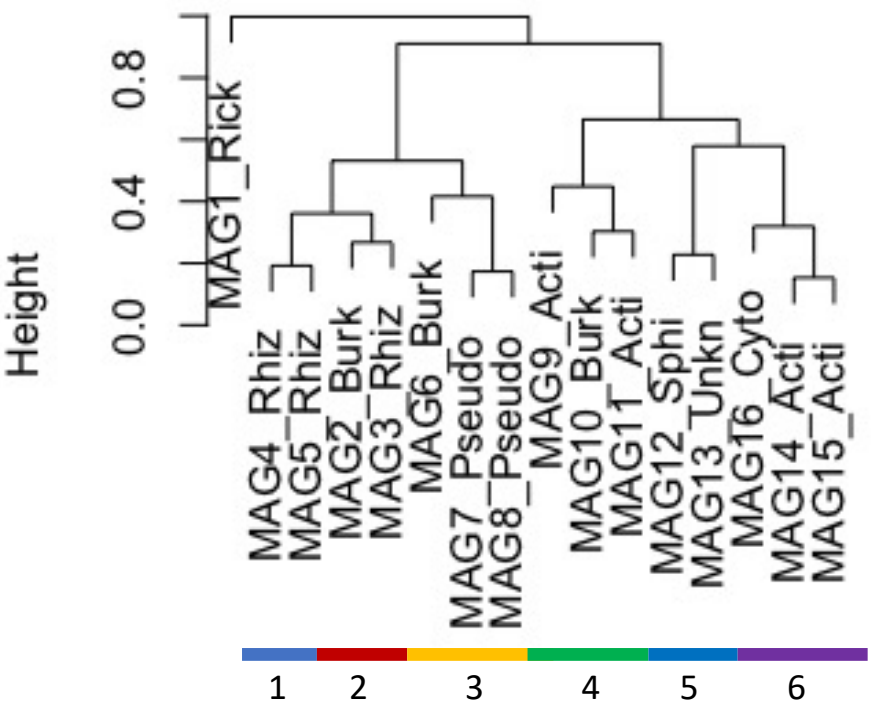


**Figure S6**. Estimated transcript abundances of functions related to sulfur metabolism associated with phyllosphere MAGs from 2016 and 2017 growing seasons. Open circles indicate that the function was not detected in a specific month for a group cluster but is detected during the season.

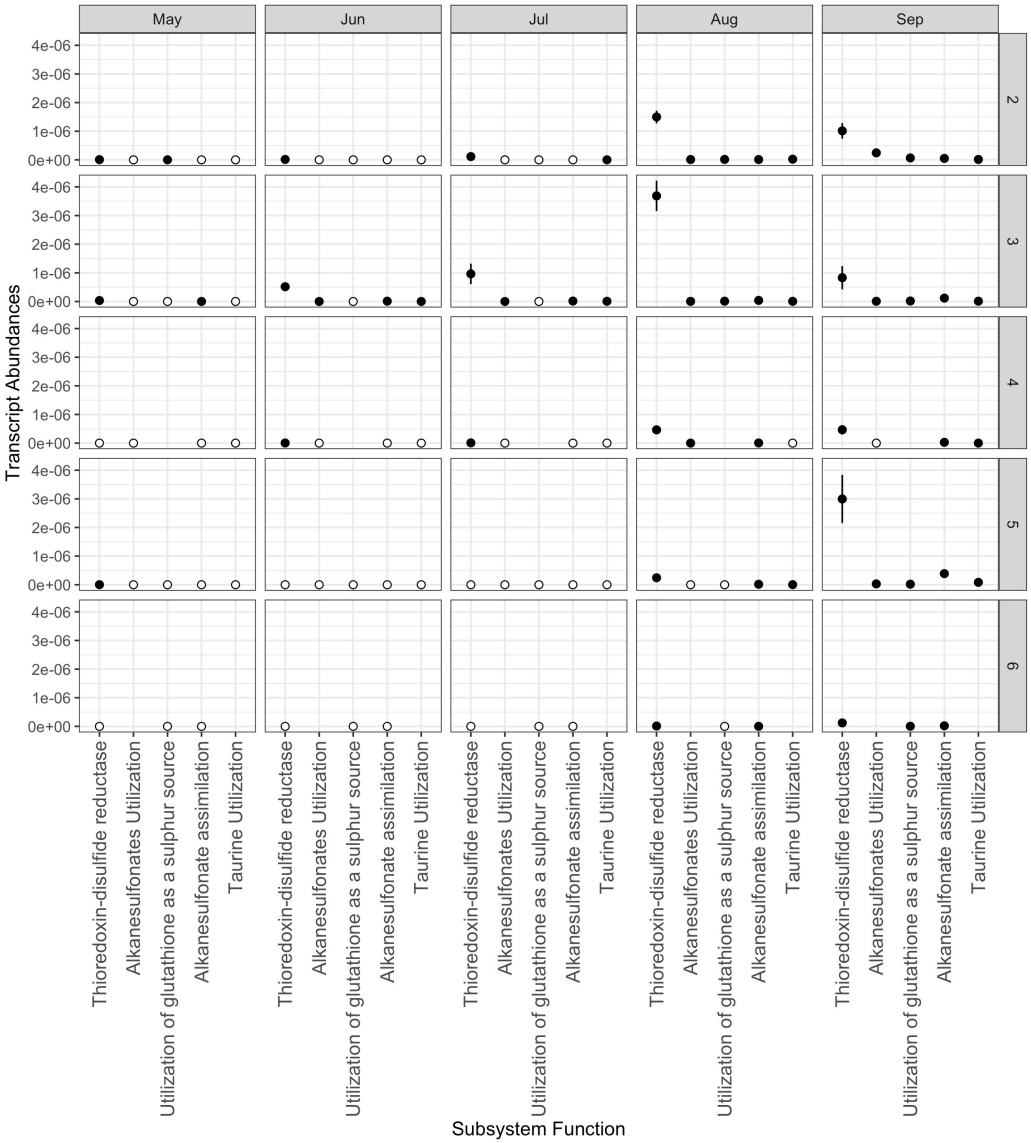


**Figure S7**. Detection of MAGs in publicly available metagenomes and metatranscriptomes of bioenergy grasses.

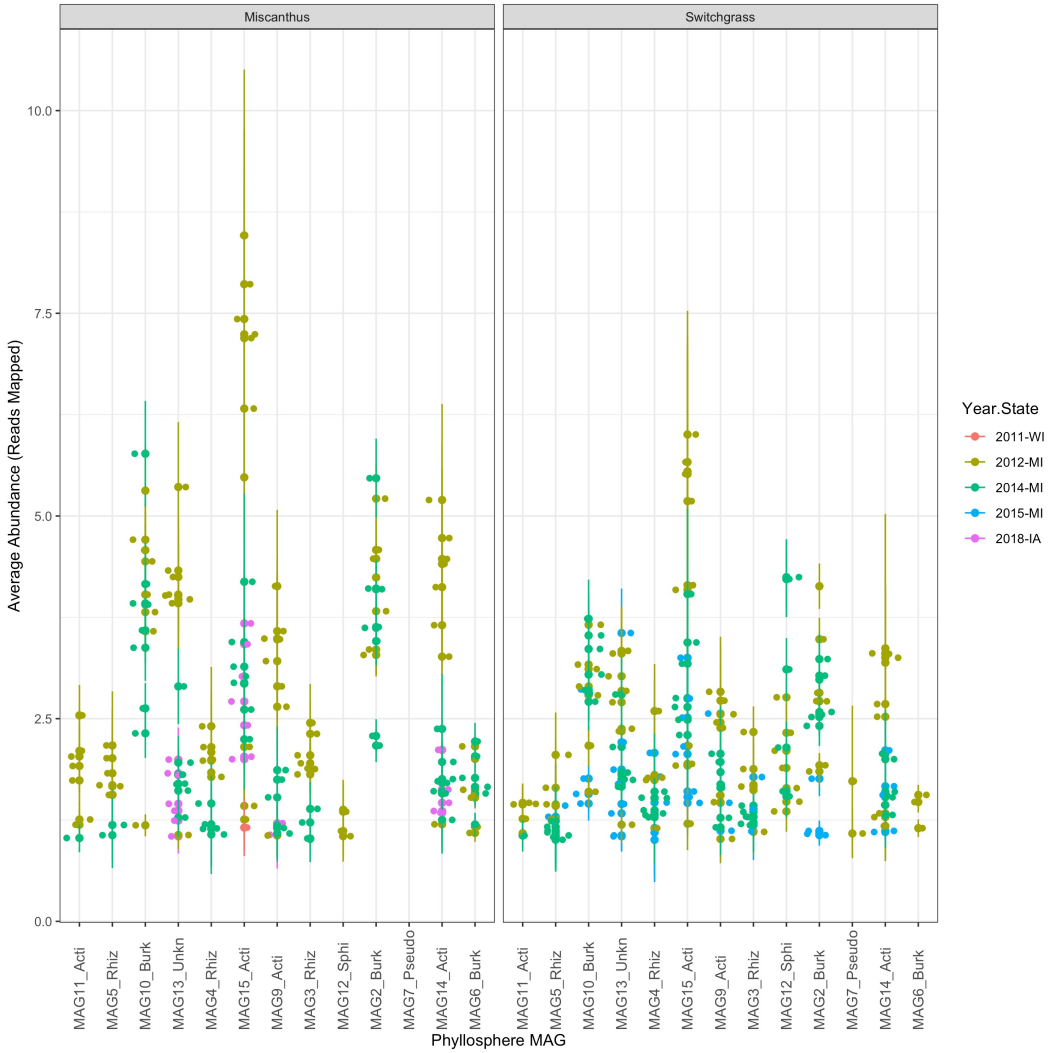


**Table S1**. Fungal genomes used for filtering metagenome reads to remove eukaryotic contamination. Genomes were selected to use for filtering based on the taxonomic identities of prevalent fungal taxa detected in our previous ITS2 amplicon survey that was conducted at the same location {Bowsher et al. 2020}.

| Fungal Genome | JGI Project ID (*) or GenBank accession (**) | Reference |
|---|---|---|
| *Trichoderma harzianum* | 403727* | (Druzhinina et al., 2018) |
| *Pleomassaria siparia* | 1011309* | (Haridas et al., 2020) |
| *Aureobasidium pullulans* | 403628* | (Gostinčar et al., 2014) |
| *Alternaria alternata* | 1103683* | NA |
| *Didymella zeae-maydis* | https://genome.jgi.doe.gov/portal/pages/dynamicOrganismDownload.jsf?organism=Didma1 | NA |
| *Sporobolomyces roseus* | 16892* | NA |
| *Puccinia novopanici* | GCA_004348175.1** | (Gill et al., 2019) |

**Table S2**. Collection of 16 highest-quality ("gold") metagenome-assembled genomes (MAGs) from switchgrass and miscanthus phyllosphere. These focal MAGs were in >10% of samples, had >70% completeness, and <15.5% contamination.

| MAG Identifier | Annotation Based on CheckM | Estimation of MAG Completeness (%) | Estimation of MAG Contamination (%) | Miscanthus Metagenomes | Switchgrass Metagenomes | Group | Estimated Genome Size (bp) |
|---|---|---|---|---|---|---|---|
| MAG1_Rick | Rickettsiales | 70.1 | 2.9 | 9 | 19 | 1 | 3361440 |
| MAG2_Burk | Burkholderiales | 75.2 | 6.4 | 23 | 22 | 2 | 1066761 |
| MAG3_Rhiz | Rhizobiales | 87.9 | 5.7 | 33 | 19 | 2 | 3240342 |
| MAG4_Rhiz | Rhizobiales | 87.4 | 4.8 | 34 | 18 | 2 | 3170845 |
| MAG5_Rhiz | Rhizobiales | 73.5 | 2.4 | 35 | 19 | 2 | 2555357 |
| MAG6_Burk | Burkholderiales | 84.0 | 2.5 | 11 | 15 | 3 | 3190093 |
| MAG7_Pseudo | Pseudomonadales | 97.6 | 1.9 | 13 | 20 | 3 | 4681335 |
| MAG8_Pseudo | Pseudomonadales | 97.4 | 2.1 | 23 | 30 | 3 | 3844481 |
| MAG9_Acti | Actinomycetales | 95.4 | 4.1 | 10 | 11 | 4 | 4589124 |
| MAG10_Burk | Burkholderiales | 93.7 | 1.6 | 10 | 11 | 4 | 5391567 |
| MAG11_Acti | Actinomycetales | 89.7 | 2.9 | 39 | 11 | 4 | 5478725 |
| MAG12_Sphi | Sphingomonadales | 75.8 | 7.2 | 17 | 19 | 5 | 5342470 |
| MAG13_Unkn | Bacteria | 72.4 | 15.3 | 23 | 24 | 5 | 4182019 |
| MAG14_Acti | Actinomycetales | 75.3 | 2.8 | 13 | 20 | 6 | 3691806 |
| MAG15_Acti | Actinomycetales | 96.0 | 4.8 | 17 | 14 | 6 | 5456568 |
| MAG16_Cyto | Cytophagales | 81.6 | 2.5 | 13 | 10 | 6 | 2848569 |

**Table S3**. Comparison of distributions of functions in phyllosphere MAGs and Refsoil+ genomes to identify leaf-enriched functions using the Kruskal-Wallis tests by ranks, with p-values less than 0.05 considered significant. Ratio of proportions in phyllosphere MAGs and RefSoil+ genomes are also shown.

| Functional Subsystem | Phyllopshere MAGs | RefSoil Genomes | p-value | Ratio of Phyllosphere:RefSoil | Significant |
|---|---|---|---|---|---|
| Secondary Metabolism | 0.49% | 0.22% | 1.6E-03 | 2.19 | * |
| Phages, Prophages, Transposable Elements, and Plasmids | 1.10% | 0.60% | 6.5E-08 | 1.83 | * |
| Respiration | 4.93% | 2.88% | 3.9E-01 | 1.71 | n.s. |
| No annotation | 38.90% | 25.18% | 2.0E-85 | 1.54 | * |
| Photosynthesis | 0.21% | 0.15% | 3.3E-02 | 1.46 | * |
| Virulence, Disease and Defense | 2.87% | 1.97% | 7.3E-05 | 1.45 | * |
| Regulation and Cell signaling | 1.77% | 1.59% | 1.7E-01 | 1.11 | n.s. |
| Sulfur Metabolism | 0.80% | 0.77% | 8.3E-02 | 1.05 | n.s. |
| Stress Response | 3.26% | 3.13% | 8.8E-04 | 1.04 | * |
| Potassium metabolism | 0.47% | 0.45% | 2.3E-14 | 1.04 | * |
| RNA Metabolism | 2.88% | 2.92% | 1.4E-32 | 0.99 | * |
| Miscellaneous | 0.31% | 0.31% | 9.0E-01 | 0.98 | n.s. |
| Nucleosides and Nucleotides | 2.44% | 2.55% | 3.6E-15 | 0.96 | * |
| Metabolism of Aromatic Compounds | 1.48% | 1.56% | 9.2E-01 | 0.95 | n.s. |
| Cell Wall and Capsule | 3.17% | 3.57% | 9.2E-04 | 0.89 | * |
| Membrane Transport | 2.44% | 2.76% | 1.5E-15 | 0.88 | * |
| Clustering based subsystems | 7.32% | 8.31% | 1.2E-31 | 0.88 | * |
| Iron acquisition and metabolism | 0.51% | 0.61% | 4.1E-05 | 0.84 | * |
| Nitrogen Metabolism | 0.72% | 1.04% | 1.5E-27 | 0.69 | * |
| Carbohydrates | 7.57% | 10.95% | 2.0E-35 | 0.69 | * |
| Cofactors, Vitamins, Prosthetic Groups, and Pigments | 2.51% | 3.78% | 4.7E-64 | 0.67 | * |
| Cell Division and Cell Cycle | 0.58% | 0.89% | 3.5E-43 | 0.65 | * |
| Fatty Acids, Lipids, and Isoprenoids | 1.97% | 3.23% | 1.5E-66 | 0.61 | * |
| Protein Metabolism | 3.50% | 5.78% | 8.0E-59 | 0.61 | * |
| Amino Acids and Derivatives | 5.00% | 8.82% | 1.4E-85 | 0.57 | * |
| DNA Metabolism | 1.38% | 2.60% | 3.0E-79 | 0.53 | * |
| Motility and Chemotaxis | 0.86% | 1.72% | 5.5E-23 | 0.50 | * |
| Phosphorus Metabolism | 0.43% | 1.14% | 7.8E-82 | 0.38 | * |
| Dormancy and Sporulation | 0.12% | 0.51% | 3.8E-23 | 0.24 | * |

**Table S4**. Comparison of transcript abundances associated with DNA and RNA metabolism by sampling date for switchgrass phyllosphere using the Kruskal-Wallis tests by ranks.

| Sample Date | Kruskal Statistic | Kruskal.test p-value |
|---|---|---|
| 31-May-16 | 1.352795031 | 0.244790079 |
| 12-Jul-16 | 1.641025641 | 0.200184804 |
| **12-Sep-16** | **3.152727273** | **0.075800175** |
| 15-May-17 | 0.926512512 | 0.335770789 |
| 5-Jun-17 | 0.540441176 | 0.462249946 |
| 26-Jun-17 | 1.334558824 | 0.247995815 |
| **17-Jul-17** | **8.647058824** | **0.003275897** |
| **7-Aug-17** | **8.040441176** | **0.004574439** |
| **28-Aug-17** | **7.455882353** | **0.006322948** |
| **18-Sep-17** | **5.025641026** | **0.024974679** |

**Table S5**. Functions that were detected in at least 13 out of the total 16 highest-quality phyllosphere MAGs, and at abundances greater than average enrichment (μ>35x) in late season (August to September) relative to early season (May to July) for switchgrass metatranscriptomes. Ratio is the average abundance observed for functions in late season divided by early season.

| Functional Subsystem Level 1 | Functional Subsystem Level 2 | Functional Subsystem Level 3 | Aug - Sep | May - July | Ratio | No. of MAGs |
|---|---|---|---|---|---|---|
| Carbohydrates | Pyruvate metabolism II: acetyl-CoA, a | Aldehyde dehydrogenase (EC 1.2.1.3) | 8.39E-07 | 1.28E-08 | 65.36791 | 15 |
| Amino Acids and Derivatives | Threonine and Homoserine Biosynthe | Aspartate aminotransferase (EC 2.6.1.1) | 2.48E-07 | 3.89E-09 | 63.71466 | 14 |
| Amino Acids and Derivatives | Threonine degradation | Threonine dehydrogenase and related Zn-dependent dehydrogenases | 2.00E-06 | 5.03E-08 | 39.81433 | 14 |
| Carbohydrates | Pyruvate metabolism II: acetyl-CoA, a | Acetyl-coenzyme A synthetase (EC 6.2.1.1) | 1.85E-07 | 4.41E-09 | 41.88832 | 14 |
| Fatty Acids, Lipids, and Isoprenoids | Nonmevalonate Branch of Isoprenoid | 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (EC 1.17.7.1) | 3.48E-07 | 6.55E-09 | 53.14039 | 14 |
| Protein Metabolism | Universal GTPases | GTP-binding protein Obg | 1.55E-07 | 1.32E-09 | 117.8691 | 14 |
| Protein Metabolism | tRNA aminoacylation, Met | Methionyl-tRNA synthetase (EC 6.1.1.10) | 3.75E-08 | 3.21E-10 | 116.5575 | 14 |
| Virulence, Disease and Defense | Resistance to fluoroquinolones | DNA gyrase subunit A (EC 5.99.1.3) | 2.14E-07 | 4.70E-09 | 45.54809 | 14 |
| Amino Acids and Derivatives | Serine Biosynthesis | D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) | 1.46E-07 | 3.11E-09 | 46.99134 | 13 |
| Carbohydrates | Trehalose Biosynthesis | 1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18) | 4.07E-07 | 7.09E-09 | 57.40183 | 13 |
| Carbohydrates | Unknown carbohydrate utilization ( c | Peptide methionine sulfoxide reductase MsrB (EC 1.8.4.12) | 5.30E-07 | 1.30E-08 | 40.61968 | 13 |
| Cell Wall and Capsule | UDP-N-acetylmuramate from Fructos | Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16) | 4.58E-08 | 0 | Inf | 13 |
| Cell Wall and Capsule | Peptidoglycan Biosynthesis | Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) | 6.67E-08 | 1.62E-09 | 41.26708 | 13 |
| DNA Metabolism | DNA repair, bacterial UvrD and relate | ATP-dependent DNA helicase UvrD/PcrA | 8.51E-08 | 1.12E-09 | 75.8473 | 13 |
| Membrane Transport | ABC transporter branched-chain amir | Branched-chain amino acid transport ATP-binding protein LivF (TC 3.A.1.4.1) | 1.20E-07 | 1.95E-09 | 61.57398 | 13 |
| Membrane Transport | ABC transporter branched-chain amir | Branched-chain amino acid transport ATP-binding protein LivG (TC 3.A.1.4.1) | 1.28E-07 | 2.56E-09 | 49.94828 | 13 |
| Protein Metabolism | Translation termination factors bacte | Methionine aminopeptidase (EC 3.4.11.18) | 2.02E-07 | 4.71E-09 | 42.80427 | 13 |
| Protein Metabolism | Translation termination factors bacte | Peptide chain release factor 2 | 2.61E-07 | 6.66E-09 | 39.15736 | 13 |
| Respiration | Ubiquinone Menaquinone-cytochror | Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2) | 2.36E-07 | 1.36E-09 | 174.2727 | 13 |
| RNA Metabolism | Polyadenylation bacterial | Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8) | 1.01E-06 | 2.15E-08 | 47.14752 | 13 |
| Virulence, Disease and Defense | Resistance to fluoroquinolones | DNA gyrase subunit B (EC 5.99.1.3) | 2.03E-07 | 5.76E-09 | 35.33268 | 13 |