The microbiome as a biosensor: functional profiles elucidate hidden stress in hosts

Avihai Zolti[a,b], Stefan J. Green[c], Noa Sela[b], Yitzhak Hadar[a], and Dror Minz[b].

a. Department of Plant Pathology and Microbiology, Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel

b. Institute of Soil, Water and Environmental Sciences, Agricultural Research Organization – Volcani Center, Rishon Lezion 7528809, Israel

c. Sequencing Core, Research Resources Center, University of Illinois at Chicago, Chicago, IL, USA

Significance Statement

This study examines the potential for microbial communities to provide insight into stresses experienced by their eukaryotic host organisms, through profiling of metagenomes and metatranscriptomes. Our study uses plant host-associated microorganisms as an *in vivo* and *in situ* microsensor to identify environmental stresses experienced by the microbial community and by the plant. Transcriptionally active host-associated microbial communities are responsive in a highly specific manner to environmental conditions. Conversely, host transcriptome sequencing provides only a very general stress response. This study is a proof-of-concept for the use of microbial communities as microsensors, with a great potential for interrogation of a wide range of host systems.

Abstract:

Microbial communities are highly responsive to environmental cues, and both their structure and activity can be altered in response to changing conditions. We hypothesized that host-associated microbial communities, particularly those colonizing host surfaces, can serve as *in situ* sensors to reveal environmental conditions experienced by both microorganisms and the host. For a proof-of-concept, we studied a model plant-soil system and employed a non-deterministic gene-centric approach. A holistic analysis was performed using plants of two species and irrigation with water of low quality to induce host stress. Our analyses examined the genetic potential (DNA) and gene expression patterns (RNA) of plant-associated microbial communities, as well as transcriptional profiling of host plants. Transcriptional analysis of plants irrigated with treated wastewater revealed significant

enrichment of general stress-associated root transcripts relative to plants irrigated with fresh water. Metagenomic analysis of root-associated microbial communities in treated wastewater-irrigated plants, however, revealed enrichment of more specific stress-associated genes relating to high levels of salt, high pH and lower levels of oxygen. Meta-analysis of these differentially abundant genes obtained from other metagenome studies provided evidence of the link between environmental factors such as pH and oxygen and these genes. Analysis of microbial transcriptional response demonstrated that enriched gene content was actively expressed, which implies contemporary response to elevated levels of pH and salt. We demonstrate here that microbial profiling can elucidate stress signals that cannot be observed even through interrogation of host transcriptome, leading to an alternate mechanism for evaluating *in situ* conditions experienced by host organisms.

Introduction

Advances in sequencing have propelled the field of microbiology and shifted focus from analysis of microbial isolates or low diversity ecosystems to analysis of environments with highly diverse microbial communities. Global surveys of microbial community structure have been conducted in a wide range of natural environments (1, 2, 3), also reviewed in (4), and many of these studies have focused on host-associated microbiomes. Such host-associated microbial environments include plant-associated communities (5, 6, 7, 8, 9) though the greatest effort has been placed on the human-associated microbial communities (10, 11, 12, 13). Studies examining plant host-associated microbial communities have focused on soil microorganisms that are enriched in the rhizosphere- the soil surrounding and affected by the roots (8). In the rhizosphere, soil type (6, 8, 9) and plant host type (14, 15) have been identified as the main forces determining rhizosphere and root microbiomes. The selection of rhizosphere-competent organisms from soil has been well established, with specific plants and different growth stages of plants each selecting for different microbial communities from among the high diversity of microorganisms in soil (16, 17). Rhizosphere microorganisms are further enriched to form sub-populations colonizing root surface (9), as plants shape the soil-plant continuum in a gradient-depended manner (6, 8, 9) mainly through carbon flux to the root environment (18). Functional profiling of microbial communities associated with different plants has demonstrated that these microbiomes differ in their metabolic activities and has suggested the presence of niche conditions associated with a wide range of factors, of which oxygen concentration is one (19).

More broadly, factors influencing plant root microbiome include geographic location (5, 9), plant developmental stage (15, 20, 21), nutrient (*e.g.,* N or P) availability (7, 22) and redox status (23). Numerous agricultural practices, which modify many of the above, have been shown to have an impact on root microbiome. These include fertilization (24), compost amendment (25) and irrigation with water of lower quality (26). Each of these practices alters a wide range of environmental variables, thus confounding the ability to identify the most consequential abiotic factor influencing the plant system and modifying the root microbiome.

Microorganisms sense minor changes in environmental conditions and respond rapidly through transcriptional changes, as well as through microbial amplification- the dynamic modification of the abundance of microbial taxa; these changes occur on a time-scale that is much shorter than for the host (27). Thus, interrogation of the microbial community may be used as a means to understand environmental conditions on short to long time scales, as well as small to large physical scale. In this study, the root surface is used as a model to test the hypothesis that microorganisms can be sensitive *in situ* detectors of environmental conditions. The root zone has a number of favorable features for such interrogation, including: (a) the presence of a high percentage of microorganisms that are transcriptionally active; (b) high microbial competition for access to root exudates – and therefore likely rapid turn-over if environmental conditions change; and (c) access to high microbial diversity in the soil. Thus, both the composition of the root-associated microbial community and the transcriptional activity of the microbial community can be informative regarding root environmental conditions.

In this study, we use the root surface microbiome functional response as a micro-sensor to identify stresses imposed by irrigation with water of lower quality, such as treated wastewater (TWW). Our study employs the basic assumptions that the microbiome inhabiting the root surface is exposed to the same environmental conditions as its host, and that the response of the microbiome (*i.e.*, alteration of community structure, associated gene abundance and transcriptional profiles) to stress can identify the specific stress or stresses in the root environment. To examine these assumptions and our general hypothesis, we performed deep DNA and RNA sequencing of plant roots grown in soil irrigated with fresh water (FW) or TWW. Plant transcriptional profiling was examined together with microbial community taxonomic and functional gene content characterization (shotgun metagenome sequencing) and microbial transcriptional profiling (shotgun metatranscriptome sequencing). We observed that the host responded to TWW irrigation in

a highly general manner, whereas the microbial response was specific to stresses present in TWW, including elevated salinity and elevated pH.

Results

Here, we assess the use of root-associated microorganisms as an indicator tool to reveal environmental conditions and stresses affecting plant hosts at a micro-scale. Our model system was a long-term anthropogenic disturbance caused by soil irrigation with water of lower quality (*i.e.,* treated wastewater, TWW) as compared to irrigation with fresh water (FW). We characterized plant-host response and root microbiome composition and response using deep sequencing of RNA and DNA extracted from roots. Shotgun metagenomic (DNA-based) and metatranscriptomic (RNA-based) analyses were performed on root systems from two host plant types (tomato and lettuce) across two consecutive years and with two different water treatments (FW or TWW), to describe taxonomic shifts and functional responses associated with long-term root irrigation with water of differing quality. For shotgun metagenome analyses, 25-49 million DNA sequences (paired-end) were generated per sample (**supplementary Table S1**). In tomato roots, 13 to 28% of the reads mapped to the host genome, with the remaining reads were attributed to the microbiome. In lettuce, 55-69% mapped to the host with the remaining reads attributed to the microbiome. A *de novo* assembly was performed using all non-host data from all root metagenomes. This assembly yielded 1,760,490 contigs larger than 500bp, and this assembly had an N50 value greater than 1,600 bases. In total, there were 6,422,376 predicted genes and a non-redundant gene catalogue (based on 95% similarity) was established with 5,359,885 genes. These genes were mapped to the SEED (28) and KEGG (Kyoto Encyclopedia of Genes and Genomes) (29) databases for functional predictions, and approximately 22% of the non-redundant genes were annotated by each database. For shotgun metatranscriptome analyses, 27-33 million sequences (paired-end) were generated per sample (**Table S1**). Metatranscriptome reads were mapped to the gene catalogue established from the metagenomics analysis (microbial transcriptome, with 11-51% of the sequences mapped) or to the available plant genome (host transcriptome analysis). As the lettuce genome is not fully annotated, sequence data generated from lettuce plants were screened for orthologs of known tomato genes, when possible.

131 ***Host functions actively associated with TWW irrigation***

132 Plant host physiological response to irrigation with water of lower quality has been
133 previously reported, with significantly reduced yield of both plants under TWW irrigation
134 (*e.g.*, (26) **Table S2**). In this study, we performed deep sequencing of plant transcripts to
135 identify stresses that TWW irrigation imposes on plant roots. Across two growing seasons, a
136 total of 45 tomato genes and 645 predicted lettuce transcripts were significantly
137 differentially expressed between irrigation treatments. Of the 645 lettuce transcripts, only
138 141 could be annotated by comparison with known tomato genes (**Table S3**).

139 To identify the most robust effects of TWW treatment, tomato and lettuce differentially
140 abundant transcripts were analyzed together. A network analysis of enriched transcripts was
141 performed to predict interactions and highlight clusters of associated genes (**Figure 1**). The
142 FW-enriched gene network consisted of 97 nodes, indicating the number of enriched genes
143 that were identified by the STRING protein-protein interaction network database. Similarly,
144 the TWW-enriched gene network consisted of 86 nodes. The FW-enriched gene network,
145 however, was linked by only seven edges representing predicted protein interactions (direct
146 physical interactions, as well as predicted functional association). In contrast, the TWW-
147 enriched gene network was linked by 69 edges, with a significantly higher number of
148 interactions then expected (p value<0.0001, by Random Graph with Given Degree Sequence
149 (RGGDS)) (**Figure 1a, b**). The TWW gene network of both plants was enriched (Aggregate
150 Fold Change, permutation-based, non-parametric test) primarily with various heat-shock
151 transcripts, including Hsp20, Hsp70, and DnaJ. Heat shock proteins are prevalent in plants
152 and are active during normal growth (30) Such genes also show a stress response, and can
153 also be activated in response to many stress cues, including heat, cold, water stress, salinity,
154 osmotic stress, and oxidative stress (30, 31). In addition, tubulin and 'FKBP-type peptidyl-
155 prolyl cis-trans isomerase' genes were also significantly enriched under TWW exposure
156 (**Figure 1b,c**). Tubulin reorganization has been shown under salt stress, cold shock,
157 aluminum exposure, interaction with pathogens and more (32, 33, 34, 35). Overall, the plant
158 response to irrigation with TWW, as detected by transcriptome analysis, was largely
159 restricted to highly general stress response genes that are expressed under a wide range of
160 environmental conditions.

161

162 ***Shifts in microbiome associated with TWW irrigation (DNA-based metagenomics)***

Functional profiling demonstrates the extent to which root microbiomes respond to environmental factors.

The taxonomic affiliation of root-associated microbial communities was determined by analysis of annotated genes from the metagenomes (**Fig. 2f; Table S4, S5**). The vast majority of annotated genes were derived from bacteria (96.7% of all mapped reads), while the percentage of reads derived from Fungi (1.2%), Archaea (0.5%), and viruses (0.13%) was much lower. Despite a prior mapping step to remove host reads, 1.2% of annotated gene counts could still be mapped to plant genomes. The root microbiome was primarily composed of bacteria from the phyla Proteobacteria (44% of all mapped reads) and Actinobacteria (33%).

The relative abundance of taxa was compared across experimental conditions of plant host type (tomato *vs.* lettuce) and irrigation water quality (FW *vs.* TWW) using DESeq2 method for comparing differential abundant count data (36)(**Figure 2f**). Broadly, 34% of all taxonomic groups (with highest available taxonomic resolution, based on MEGAN6 least common ancestor) were significantly (Wald test ,FDR corrected p value<0.05) more abundant in tomato roots, as compared to 31%, significantly associated with lettuce roots. Many taxa from the phlya Actinobacteria and Bacteroidetes/Chloroflexi were significantly more abundant in tomato roots relative to lettuce, while Betaproteobacteria and Planctomycetes were strongly and significantly associated with lettuce roots. Irrigation water quality mostly affected Proteobacterial taxa (10% of microbial taxa were significantly more abundant in TWW-irrigated roots, as compared to 11% of microbial taxa enriched in FW-irrigated roots. 60% of all significantly abundant taxonomic group were identified as Proteobacteria). Acidobacteria and Betaproteobacteria were significantly more abundant in FW-irrigated roots and Gammaproteobacteria significantly more abundant in TWW-irrigated roots (all data available at supplementary **Table S4**).

Root microbial metagenomes from lettuce and tomato were annotated and mapped to the SEED database to identify functional genes significantly associated with plant host type (tomato and lettuce) and irrigation water quality (FW and TWW). A comparison of differentially abundant functional genes between tomato and lettuce root microbiomes, demonstrated strong host specificity in microbiome gene content (**Figure 2a**), consistent with our prior analyses of root microbiomes of different plant species grown in identical soils (19). In this study, greater than 50% of SEED annotated genes (from a total of 2625 "functional role"(28)) were significantly (based on Wald test, adjusted p-value<0.05) more

196 abundant in either tomato or lettuce roots (26% in tomato relative to lettuce and 27% in
197 lettuce relative to tomato; **Figure 2a**).

198 Irrigation water quality also affected root-associated microbiome functional profile (**Figure**
199 **2b, c, d**). Initially, the effect of irrigation type was examined in tomato and lettuce systems
200 independently by comparing SEED-annotated gene abundance using DESeq2 method.
201 Irrigation type determined 36% of the tomato root metagenome (15% of all annotated gene
202 list in tomatoes were significantly more abundant in FW irrigated plants compare to 21%
203 more abundant in TWW irrigated plants, Wald test, **Figure 2b**) similarly to 34% of the lettuce
204 root metagenome (FW- 16%, TWW- 18%, **Figure. 2c**). To identify commonalities in response
205 to irrigation, irrigation effects were examined in a dataset of both plant hosts combined. This
206 combined analysis identified microbiome genes that were positively associated with FW
207 irrigation (11% of all SEED annotated genes) or with TWW irrigation (15% of all SEED
208 annotated genes) (**Figure 2d**). Overall, the combined root microbiome functional profiles
209 were significantly associated with plant host (tomato *vs*. lettuce, F=10.1, p value=0.001) and
210 irrigation treatment (FW *vs*. TWW, F=6.6, p value=0.004) as determined by a PERMANOVA
211 test based on the Bray-Curtis dissimilarity index of the SEED annotated gene counts (**Figure**
212 **S1**). For further analyses, unless otherwise indicated, ecosystem comparisons of plants
213 grown in FW and TWW were performed on data combined from both plant hosts.

214 We previously measured significant increases in pH, dissolved organic matter (DOC) and
215 electrical conductivity (EC) in TWW-irrigated soils relative to FW-irrigated soils (26)(also at
216 **Table S6**). Similar patterns were observed in this study through canonical correspondence
217 analysis (CCA) (**Figure 2e**). The CCA presents the relationship of the measured soil
218 parameters and root microbiome functional gene profile (all SEED-annotated gene counts).
219 An ANOVA permutation test was significant (F= 3.2, p value=0.001 for the full model), and
220 the constrained variables (i.e., pH, DOC, EC) accounted for 54.8% of the variance. DOC
221 (F=2.4, p value=0.038) and pH (F=5.7, p value=0.002) were found to significantly explain
222 portions of the variance associated with the observed microbiome functional profile, while
223 EC was not significant (F=1.57, p value=0.163). The microbial community functional gene
224 profiles of FW- and TWW-irrigated root samples were separated along the CCA2 axis, with
225 DOC loading primarily on the CCA2 axis. Conversely, the microbial community functional
226 gene profiles of samples from different plant hosts were separated primarily along CCA1
227 axis, with pH loading primarily on the CCA1 axis.

228

SEED & KEGG functional categories enriched or depleted in metagenomes of TWW-irrigated roots relative to FW-irrigated roots.

SEED- annotated genes highly significantly ($p<0.01$) associated with irrigation water quality across both plants were examined, and in total 438 genes were identified (**Figure. 3**). Of these genes, 286 were enriched in TWW-irrigated roots and 152 enriched in FW-irrigated roots. These genes were clustered into general categories (SEED level 1, based on the hierarchical clustering available on MEGAN6): *e.g.,* carbohydrates, amino acid derivatives, membrane transports, respiration and regulation of cell signaling (**Figure 3a**). Rare categories (represented by fewer than 5 genes) were removed from the analysis. To compare category enrichment, we examined the proportion of genes enriched for each category (**Figure 3b**). For some gene categories (*e.g.,* cell division, cell cycle or carbohydrates), a similar number of genes were enriched in both TWW- or FW-irrigated roots (*i.e.*, no specific effect of irrigation treatment), while others were more strongly skewed to either FW or TWW. For example, membrane transport and transposable element genes were substantially enriched in TWW-irrigated roots. Conversely, the gene categories of nucleosides and nucleotides and sulfur metabolism, were substantially enriched in FW-irrigated roots.

An enrichment analysis was also conducted for gene subsystem enrichment and depletion by irrigation method (level 2, based on the SEED hierarchical clustering, tested using Wallenius non-central hypergeometric distribution) (**Figure 4a, Table S7**). One of the most strongly enriched categories ($log_2FC=1.4$; relative abundance=0.07%, p value<0.0001) was the Na(+)-translocating NADH-quinone reductase (NQR), a membrane complex that utilizes the respiratory chain to generate a sodium gradient in place of a proton gradient in high pH and sodium conditions (37). In addition, enrichment of multiple membrane-associated subsystems in TWW-irrigated roots was observed, including: (i) sodium-hydrogen antiporter, a common membrane transporter that supports sodium balance in exchange for proton motive force ($log_2FC=1.67$; p value=0.002) , (ii) pH adaptation potassium efflux system ($log_2FC=1.87$; p value=0.007), (iii) mannose-sensitive hemagglutinin, type 4 pilus (MSHA4) ($log_2FC=1.03$; p value=0.0009), (iv) alginate metabolism membrane complex ($log_2FC=0.23$; p value=0.01). In addition to membrane-associated subsystems, other subsystems were enriched in TWW-irrigated roots, including arginine degradation ($log_2FC=0.56$; p value=0.04). Five genes enriched within this subsystem catalyze the complete arginine to glutamate pathway (**Table S7**). Soil Na^+ concentration, K^+ concentration and pH were correlated with observed gene abundance patterns (**Figure S2, Table S8**). The relative abundance of TRAP

transporters (Pearson's R_{DOM}=0.81, P- $value_{DOM}$=0.001; R_{Na}=0.78, P_{Na}=0.003; R_{K+}=0.81, P_{K+}=0.001) and sodium hydrogen antiporters (R_{DOM}=0.86, P_{DOM}=0.0003; R_{Na}=0.78, P_{Na}=0.003; R_{K+}=0.8, P_{K+}=0.001) correlated to organic matter and Na^+/ K^+ concentrations. The relative abundance of potassium antiporter genes was correlated with Na^+ and K^+ concentrations (R_{Na}=0.9, P_{Na}<0.0001; R_{K+}=0.86, P_{K+}=0.0003), and MSHA4 gene relative abundance was correlated with pH (R_{pH}=0.88, P_{pH}=0.0001). NQR gene abundance was significantly correlated with salt concentration (R_{Na}=0.8, P_{Na}=0.002; R_{K+}=0.81, P_{K+}=0.001) and pH (R_{pH}=0.79, P_{pH}=0.002).

Analysis of enriched KEGG pathways (**Figure 4b, c, Table S9**) and modules (**Figure S3**) revealed additional biological processes enriched in TWW-irrigated root microbiomes relative to FW-irrigated root microbiomes. Two-component systems were significantly enriched (40 KEGG genes were enriched, p value<0.0001) in TWW-irrigated root microbiomes relative to FW-irrigated root microbiomes, including pathways involved in misfolded proteins, flagellar assembly, iron acquisition, and Mg^{2+} starvation (**Figure 4c**). In addition, the denitrification gene module was significantly (4 genes, p value=0.006) enriched in TWW irrigated roots (**Figure S3**). Conversely, ABC transporter gene pathways associated with sugar (maltose, galactose and oligogalacturonide), peptide (dipeptide and glutamate/ aspartate) and nutrient (cobalt or nickel) transport were enriched (43 genes, p value<0.0001) in FW-irrigated root microbiomes relative to TWW-irrigated roots. A Type 3 secretion system (T3SS) gene module was also enriched (12 genes, p value<0.0001) in FW-irrigated root communities.

Meta-analysis of selected gene counts relative to environmental variables from publicly available metagenomes.

A meta-analysis was conducted to establish a global link between metagenome functional gene content and measured environmental variables. We focused on subset of prominent genes from this study that were strongly positively correlated with pH (NQR, Na^+-H^+ antiporter) or negatively correlated with oxygen levels (periplasmic nitrate reductase, *napAB*, nitric oxide reductase- *norBC*, nitrous oxide reductase Z, *nosZ*). Metagenomes available at the Joint Genome Institute's (JGI) Integrated Microbial Genomes and Microbiomes repository (n=14,596) were screened. Environmental pH measurements were available for a subset of these metagenomes (n=1,588), and of this subset, 160

295 metagenomes had a total number of predicted genes greater than 100,000 (**Table S10**).
296 Within these 160 metagenomes, the relative abundance of genes annotated as NQR
297 (pairwise Wilcoxon rank test, Bonfferoni correction- p $value_{pH<7: pH7-8}$=0.02; p $value_{pH7-}$
298 $_{8:pH>8}$<0.0001; p $value_{pH<7:pH>8}$=0.002) and Na^+-H^+ antiporter (p $value_{pH<7: pH7-8}$<0.0001; p
299 $value_{pH<7:pH>8}$<0.0001) were strongly correlated with measured pH values (**Figure 5a, b**).
300 Using the same filtering criteria, 257 metagenomes were identified with oxygen
301 measurements and greater than 100,000 predicted genes. In these metagenomes, the
302 abundance of *napAB*, *norBC* and *nosZ*, in the denitrification pathway, were enriched in
303 samples with lower measured oxygen (**Figure 5d, e, f, g**). No such trend was observed for
304 housekeeping genes such as gyrase B (*gyrB*, **Figure 5h**). Salinity or Na^+ concentrations were
305 measured only in small subset of available metagenomes and were not analyzed further.

306

307 ***Microbial gene expression patterns associated with TWW irrigation (RNA-based***
308 ***metatranscriptomics)***

309 An enrichment analysis of the root-associated microbial metatranscriptomes was performed
310 to identify SEED-annotated genes and subsystems that were significantly differentially
311 transcribed between plants irrigated with TWW relative to those irrigated with FW (**Figure 6,**
312 **Table S11**). In total, 10.1% of SEED-annotated genes were significantly differentially
313 expressed in roots of TWW-irrigated plants relative to FW-irrigated plants (Wald test ,FDR
314 corrected p value<0.05). Specifically, 7.2% of such genes had higher expression and 2.9%
315 lower expression in TWW-irrigated roots relative to FW-irrigated roots. SEED-annotated
316 genes were clustered into 761 SEED subsystems (level 2, based on SEED hierarchical
317 clustering), and of these, 8 were over-represented in TWW-irrigated root microbial
318 communities while only a single subsystem was significantly over-represented in the FW-
319 irrigated root transcriptomes. The most highly and significantly over expressed gene sub-
320 systems in TWW-irrigated roots were NQR, TRAP transporters, sodium-hydrogen antiporters,
321 alginate metabolism genes and MSHA4 (**Figure 6a**). All of these genes were also significantly
322 enriched in metagenomic analysis of TWW-irrigated roots relative to FW-irrigated roots.
323 Genes involved in alginate metabolism were only slightly enriched in metagenomes of TWW-
324 irrigated roots (log_2FC=0.23, p value=0.01) but were strongly over-expressed in TWW-
325 irrigated metatranscriptomes relative to FW-irrigated roots (log_2FC=1.7, p value=0.0004).
326 Conversely, the overall expression level of hydrogenase subsystem genes was significantly
327 higher in FW-irrigated roots relative to TWW-irrigated roots (log_2FC=1.6, p value>0.0001),

328 though their relative abundance at the DNA level was not substantially affected by irrigation
329 treatment (**Figure 6b**).

330 Enrichment analysis of KEGG pathways (**Figure 6c**) and modules (**Figure S4**) was performed
331 (**Table S12**). The transcriptome of the microbial community of TWW-irrigated roots was
332 significantly enriched in 2.75% of all KEGG-annotated genes. In addition, 8 KEGG pathways
333 and 2 KEGG modules were significantly enriched in TWW-irrigated root metatranscriptomes
334 relative to FW-irrigated root metatranscriptomes. In contrast, the transcriptome of the
335 microbial community of FW-irrigated roots was significantly enriched in 1.8% of all KEGG-
336 annotated genes. In addition, 3 KEGG pathways and 3 KEGG modules were significantly
337 enriched in FW-irrigated root metatranscriptomes relative to TWW-irrigated root
338 metatranscriptomes. This analysis revealed higher microbial relative expression of two-
339 component systems, including the C4 dicarboxylate gene cluster, in TWW-irrigated roots
340 relative to FW-irrigated roots (**Figure 6c**). Moreover, in TWW-irrigated roots, higher relative
341 expression of arginine and proline metabolism genes, particularly those in the arginine-to-
342 spermidine pathway, was observed (**Figure S5**). The relative expression level of ABC
343 transporter genes, including glutamate and galactose transporters, were higher in the
344 metatranscriptomes of FW-irrigated roots relative to TWW-irrigated roots (10 enriched
345 genes, p value=0.0004). The abundance of type 3 secretion system (T3SS) genes was
346 significantly higher in the metagenomes of TWW-irrigated root microbial communities
347 relative to those of the FW-irrigated root metagenomes. However, the level of expression of
348 type 6 secretion system (T6SS) genes was most highly expressed under FW-irrigation
349 conditions (7 enriched genes, p value<0.0001).

350

351 **Discussion**

352 We previously studied the effect of TWW irrigation on soil and root microbial community
353 structure and composition (26). In that study, irrigation water quality and soil type were
354 major explanatory variables for the observed soil microbial community structure and were
355 of a similar magnitude. Similarly, the effect of irrigation water quality on root microbial
356 community structure was of a similar magnitude to the plant host effect (26), demonstrating
357 the responsiveness of the microbial community to both host and environmental factors. In
358 the current study, we have attempted to harness the rhizoplane microbiome – existing at
359 the interface between the plant and the surrounding soil – as a sensor for detecting *in situ*
360 environmental conditions at the plant-soil interface, including factors leading to host stress.

The main incentive in using the host-associated microbiome as a biosensor lies in the fact that high resolution is desired for accurate definition of the factors contributing to host physiological status (38). Comparing the differences in the relative abundance of microbial genetic features (*i.e.,* metagenome analysis) or expression of microbiome genes (*i.e.,* metatranscriptome analysis) can aid in the identification of long-term stressors imposed on the host under these conditions as well as short-term stressors revealed by expression of genes processing environmental cues at time of sampling. Analyses can be performed at different levels of hierarchical gene annotation and can be performed using gene level annotation (*e.g.,* SEED database) and enriched pathways or modules (*e.g.*, KEGG annotation).

Most commonly used methods for studying root- soil interface employs microelectrodes (39), or specific dyed root imaging in "rhizoboxes" (40, 41). Both methods measure only pre-defined variables, eliminating the possibility of discovering novel or unsuspected stressors. Moreover, experimental design forces manipulating natural environment by growing plants in designed cells or by removing plants from soil for further experimental procedure. Furthermore, in studies where transcriptional response is examined, plant host response is often tested under severe stress in unnatural short term experimental design (42). We sought to be able to assess environmental factors leading to plant physiological status under more natural agricultural conditions.

A secondary motivating factor for the use of the microbiome as a biosensor lies in the observed low-resolution response of the host organism. In this study, upregulation of stress response genes was identified in the transcriptome of host roots irrigated with TWW relative to those roots irrigated with FW. However, the specific nature of the stresses remained unresolved, with transcriptome analysis revealing only the differential expression of genes involved in a highly general stress response associated with heat shock proteins (43, 44). In fields, plants are expose to myriad fluctuating biotic and abiotic environmental conditions, which force plants to tailor their gene transcriptional profiles. Therefore, individual abiotic stress response cannot be extrapolated to plant experiencing multiple stress conditions. The nature of the stress cannot be predicted based on experimental profiling of individual stress response under regulated conditions (45, 46). Moreover, other types of stress regulation can also mediate plant response, including post-transcriptional regulation of RNA by micro RNAs (miRNAs) or other small noncoding RNAs (47), and through protein modification (48). Such regulation is not as easily measured as gene expression.

In contrast to the host, the genetic diversity of the host-associated microbiome is much greater (27), and the extraordinarily high microbial diversity in soils provides the plant with a wide selection of organisms competing for access to root exudates (18). While the plant host can alter gene expression profiles in response to changing environmental conditions, both the membership of the root-associated microbial community and the expression patterns of the root-associated microbial community can be altered. Thus, the microbiome provides us with a highly dynamic and sensitive target with the potential for both short-term responsiveness (*i.e.*, metatranscriptome) and long-term responsiveness (*i.e.*, metagenome). In this study, we observed that in response to long-term irrigation with TWW, both the metagenome and metatranscriptome were significantly altered. Statistical analysis of microbial features lead to the identification of significantly differently abundant genes, gene transcripts and pathways. Some gene of interest, being most significantly enriched, or with known and informative supported data, are presented in **Figure 7**. Critically, the identification of the differentially abundant or expressed microbial features was consistent with the known key stresses imposed by TWW irrigation on the microbial community and the host.

Level of pH and salinity:

Cytoplasmic pH in microorganisms must be at a range suitable for maintaining protein integrity. Bacterial cytoplasmic pH lies at a pH range of 7.4–7.8 (49). In alkaline environments, organisms deploy various mechanisms to maintain intracellular pH and preserve electrochemical gradient in the presence of low proton concentration. To prevent proton loss in alkaline environments, an increase in cytoplasmic pH is achieved by reducing the activity of the proton pumping machinery of the cell respiratory chain. An increase in intercellular proton concentration is attained by an increase in the activity of proton-cation antiporters, reducing the pH gradient along the cell membrane, but increasing the transmembrane electrical potential (50). Under such conditions, some bacteria form a transmembrane sodium gradient, alternatively or concomitantly with a proton gradient. In our plant system, the abundance and expression of Na^+-transporting NADH:ubiquinone oxidoreductase (*nqr*) genes was significantly enriched under TWW irrigation relative to FW irrigation. In many alkaline environments, NQR constructs the primary sodium efflux system through oxidation of NADH and reduction of quinone. This process creates an electrochemical gradient of net negative charge in the cytosol (51), and the gradient is used

427 by cation/proton antiporters (*e.g.,* Na^+/H^+, and K^+/H^+) to exchange non-balanced movement
428 of positive charge (H^+) to the cell (more protons enter the cytoplasm as compared to the
429 efflux of sodium or potassium ions (49, 51, 52, 53). In a global mapping of soil bacterial
430 communities, cation/proton antiporters were observed to be key genes overrepresented in
431 dryland soil, presumably due to the high levels of salt and pH in arid soils (54). Additionally,
432 we measured an increase in the relative abundance and expression of Tripartite ATP-
433 Independent Periplasmic (TRAP) transporters in TWW roots. TRAP transporters have been
434 recently demonstrated to use a membrane-associated sodium gradient to facilitate
435 transport of ligands (55, 56, 57). Similarly, the increase in abundance in flagella assembly
436 genes under TWW irrigation conditions may also suggest the use of sodium motive force for
437 flagella performance (58, 59). This is further consistent with cell motility as a critical feature
438 of rhizosphere competence (19). The signs for high pH stress obtained by the metagenome
439 and metatranscriptome analysis are consistent with soil chemical analysis, as elevated pH
440 conditions can result from long-term TWW irrigation.

441 Bacterial adaptation to alkaline conditions is frequently dependent on salt concentration,
442 and elevated salt levels are also found in TWW and TWW-irrigated soils (60). Elevated soil
443 salinity can develop through long-term TWW irrigation, and can adversely affect protein and
444 cell membrane stability. Commonly, microorganisms adapt to high levels of salts through
445 osmoregulation by synthesizing organic solutes, thereby avoid salt imbalance and the influx
446 of toxic salts. More rarely, in stable saline environments, some halophiles mitigate salinity
447 levels through adapting the cell enzymatic activity to the high ionic strength. Both strategies
448 require stabling salt concentration in the cell, mostly by regulating cation proton antiporter
449 activity (61). Efflux of sodium ions by NQR activity and the activation of cation/proton
450 transporters demonstrate that the TWW-irrigated root microbiome and the plant roots are
451 indeed exposed to elevated salinity as compared to the FW-irrigated roots. This finding is
452 consistent with the measurement in this study of higher levels of Na^+ in the leaves of TWW-
453 irrigated tomato and lettuce plants relative to FW-irrigated plants (**Table S2**) and in other
454 plants (62).

455 To the best of our knowledge, no prior study has linked environmental microbiome functions
456 to pH level or salinity. However, the association between our significantly differentiated
457 genes to processing specific environmental conditions is established *in vitro* in numerous
458 studies (51, 63). Here, we hypothesized that differentially abundant genes could be used as
459 predictive markers of environmental cue, and this hypothesis is supported by meta-analyses
460 demonstrating a link between single isolate studies and microbial communities *in vivo*. A

461 greater effort in collecting, publishing and making metadata more accessible for
462 metagenomics surveys, will better aid the interpretation and modeling of microbiomes
463 processing environmental conditions.

464

465 *Oxygen levels:*

466 Microbial gene content and expression patterns have great potential for identifying oxygen
467 conditions *in situ*. We previously demonstrated that different plants, even when grown in
468 the same soil, may produce highly different rhizoplane oxygen conditions, and that shotgun
469 metagenome and metatranscriptome analysis revealed differential expression of
470 denitrification genes and catalase genes (19). In this study, we observed an enrichment in
471 denitrification genes in TWW-irrigated root metagenomes relative to FW-irrigated
472 metagenomes, possibly suggesting a lower oxygen concentration under TWW-irrigation (63).
473 However, the expression level of denitrification genes was not significantly higher in TWW-
474 irrigated roots relative to FW-irrigated roots. This difference in enrichment between
475 metagenome and metatranscriptome could be due to enrichment in the TWW-irrigated
476 rhizoplane of facultative denitrifying microorganisms, with rhizosphere selection based on
477 other physiological capabilities (*e.g.*, motility). Conversely, the lack of enrichment in
478 denitrification genes in TWW-irrigated metatranscriptomes could be a result of time of
479 sampling. The root microbiome functional profile is expected to fluctuate by diurnal or
480 hydration-dehydration cycles (64, 65, 66). Therefore, gene abundance is indicative of the
481 chronic, long term exposure to stress imposed by TWW, while expression levels may
482 represent transient conditions. Plants in this study were subject to twice-daily irrigation and
483 irrigation conditions in TWW deliver higher levels of organic matter and this may lead to
484 localized oxygen depletion (67). However, at the time of sampling, oxygen levels may have
485 increased. Further short-term longitudinal analysis will be required to demonstrate diel- and
486 irrigation-derived shifts in denitrification gene expression patterns.

487

488 *Bacterial life style:*

489 We observed a significant enrichment of genes associated with surface attachment,
490 colonization and biofilm formation in TWW-irrigated root microbiome. These enriched
491 microbial features included genes encoding for flagella and MSHA type 4 pili; both features
492 have been previously demonstrated to facilitate near-surface motility and bacterial

493 attachment (68). Furthermore, an increase in the relative abundance and expression of
494 alginate producing genes, which catalyze the formation of extracellular polysaccharide
495 matrix in biofilms of many bacterial clades (69), was observed. These results further point to
496 the critical importance of both motility and attachment for rhizoplane microorganisms, as
497 has been previously indicated (70). The reason for significant enrichment of these genes in
498 TWW-irrigated roots relative to FW-irrigated roots is not entirely clear but may be due to
499 overall elevated organic matter in TWW (71). While the focus in this study has been largely
500 on genes enriched in TWW-irrigated systems, several key ABC transporters were depleted in
501 the metagenomes of roots of both plant types irrigated with TWW relative to those irrigated
502 with FW. These genes include transporters for oligogalacturonide, maltose, general sugar
503 and glutamate. The change in relative abundance of these transporters may indicate
504 differential pattern of root deposits, as was observed in roots from cucumber and wheat
505 (19). Plant glutamate secretion patterns have been shown to be mediated by external cues
506 such as salinity, oxidative stress and availability of nutrients (72). Glutamate ABC
507 transporters were depleted, while glutamine were enriched in TWW irrigated roots.

508

509 ***Conclusions:***

510 We hypothesized that the microbial functional gene profiles and expression patterns can
511 serve as *in vivo* sensors of environmental factors affecting hosts and host-associated
512 microbial communities. Environmental surveys or host-associated microbiome analyses
513 frequently yield contradictory or context-dependent results, making the predictive power of
514 such observations inconclusive. Studying the microbiome as a functional unit reacting to a
515 specific environment, however, constitutes a non-deterministic approach, thereby
516 eliminating the need for marker features (*e.g.*, genes, pathways or specific taxa) associated
517 with specific conditions. As the host and its microbiome are similarly exposed to
518 environmental conditions, genetic profiling and expression analysis of the microbiome may
519 be used as a predictive tool to identify stresses affecting hosts. In this study, we employ a
520 well-defined plant host-microbiome system under experimental treatment with FW- or
521 TWW-irrigation, but this approach may be used to define microscale conditions in other host
522 systems.

523

524

Materials and methods:

Experimental design; Mesocosm scale experiment:

Tomato (*Solanum lycopersicum*-Heinz 4107) and lettuce (*Lactuca sativa*-Romaine-Assaph) were grown in lysimeters (0.5 m3) for 98 and 42 days, respectively, at Kiryat-Gat-Lachish agricultural research station, northern Negev, Israel (31.605760, 34.791179). The lysimeters were filled with loamy sand soil collected from western Negev, Israel (31.351722, 34.403471). Plants were drip irrigated for 8 summers with fresh water (FW) or tertiary treated wastewater (TWW), derived from the Kiryat Gat wastewater treatment plant (WTP). At harvest, roots were collected, vigorously washed, dried, and frozen on site for further procedures. All samples were composites, consisting of 2-4 plants collected from separate lysimeters, except for FW irrigated lettuce plants which were collected from two lysimeters and were composites of 2 plants each. Detailed procedures, including soil and plant measurements, were described previously (26).

DNA and RNA extraction

Standard phenol-chloroform nucleic acid extraction protocol was employed for DNA and RNA isolation (26, 73). In brief, 0.2 gr of roots were moderately bead beaten for 45 s at low speed (4.5 m/s) by Fast Prep FP120 (Savant Instruments Inc., Holbrook, NY, USA) with phenol, phosphate buffer pH 8 (with additional 10 µl ml−1 β-mercaptoethanol -Sigma-Aldrich, St Louis, MO, USA) and 1.25% CTAB (Hexadecyltrimethylammonium bromide, Sigma Aldrich). Following phenol-chloroform wash, nucleic acids were precipitated with polyethylene glycol (PEG) and ethanol. Nucleic acids were split for DNA and RNA isolation. RNA samples were treated with RQ1- DNase (Promega, Madison, WI, USA) and complete DNA removal from RNA samples was validated by real time reverse transcription PCR. RNA integrity was evaluated with Agilent TapeStation (Santa Clara, CA, US). Ribosomal RNAs were removed using the Ribo-Zero rRNA Removal Kit (Illumina, San Diego, CA), combining bacteria and plant probes. Double- strand complementary DNA (cDNA) synthesis was conducted by Maxima H Minus Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA USA).

Library prep and sequencing

Shotgun metagenome libraries were generated using a Nextera XT library preparation kit according to the manufacturer's instructions (Illumina). Complementary DNA for transcriptome analysis was sheared using a Covaris S2 acoustic device, and libraries were generated using a Accel-NGS 1S Plus DNA Library Kit (Swift Biosciences, Ann Arbor, MI)

according to the manufacturer's instructions. Libraries were pooled sequenced using high-output flow cells with paired-end 2x150 base reads on an Illumina NextSeq500 sequencer. Library preparation and sequencing was performed at the University of Illinois at Chicago Sequencing Core (UICSQC) .

Bioinformatic analysis:

Quality control of raw double-strand FASTQ sequences was evaluated by FASTQC software (74), and adjusted by Trimmomatic (75) with customized parameters set to: SLIDINGWINDOW:4:15 MINLEN:100 CROP:145 HEADCROP:15 .

Metagenome analysis: Host sequences were removed by comparing quality checked reads to host genome (tomato or lettuce) with bowtie2 (76), and subsequently removing the reads with SAMtools (77). Metagenomics reads from all three replicates were de novo assembled together with metaSPAdes (78). Gene prediction was performed on scaffolds using the software package Prodigal (79). Predicted genes from all samples were combined, and a non-redundant gene catalog was established based on 95% similarity, using CD-HIT (80). The gene catalog was aligned to the NCBI non-redundant protein database using the software package DIAMOND in sensitive mode (81). Sequence annotation (SEED- (28) and Kyoto Encyclopedia of Genes and Genomes- KEGG- (29)) and predicted taxonomy were achieved with MEGAN V6 (82). To attain count data (number of mapped read for each gene), quality checked reads (after host read removal) were aligned to the annotated gene catalog by bowtie2, while analogous read annotated terms were summed using a custom python script .

Metatranscriptome analysis: Quality checked RNA reads were aligned to the gene catalog established from the metagenomics analysis, in a similar fashion to metagenomics count data .

Host RNAseq: An estimation of transcript abundance for tomato root samples was obtained by aligning quality checked sequences (prior to host reads removal) to the predicted *Solanum lycopersicum* transcripts with Trinity RSEM transcript quantification method (83). Lettuce transcripts first predicted by Tophat and cufflinks for transcript prediction (84), than infered to ortholog tomato genes by OrthoFinder. Transcript quantification was done following similar analysis as for tomato samples.

Statistical analysis:

588 Metagenome and metatranscriptome statistical enriched gene list (SEED or KEGG
589 annotated) or taxonomic groups were obtained by DESeq2 (36) and compared using the
590 VennDiagram R package (85). Taxonomic trees were visualized using the interactive tree of
591 life (86) and applying the least common ancestor MEGAN algorithm. SEED subsystems
592 enrichment analysis was conducted with the 'R' goseq package (87), normalizing to SEED
593 counts. KEGG pathway and module enrichment were analyzed by clusterProfiler package in
594 R (88). Statistical test (MANOVA, ANOSIM) were conducted in R package 'vegan' (89), and
595 figures were plotted with R 'ggplot2' (90) or 'pheatmap' (91). Differentially expressed host
596 transcripts were obtained using the EdgeR 'R' package (92), followed by annotation and
597 visualization using the STRING network (93, 94) Cytoscape integrated application (95) for
598 both plant hosts combined. The minimum required interaction score was customized to
599 medium confidence (0.4), and PFAM protein domain enrichment was set to a false discovery
600 rate p value of 0.05.

601

602 **Acknowledgements**

603 The authors thank Dr. Nesli Tovi and Dr. Sammy Frenk from the Agricultural Research
604 Organization – Volcani Center. The authors thank Dr. Jonathan Friedman from the faculty of
605 agriculture, Hebrew university. The authors would also like to thank Dr. Maya Ofek Lalzar
606 from Haifa University. This research was supported by research grant no. IS- 4662-13 from
607 the Binational Agricultural Research& Development Fund (BARD), research grant no. 821-
608 0142-13 from the Israel Ministry of Agriculture and Rural Development and USAID-MERC
609 research grant no. M34-011.

610

611

612

613

614

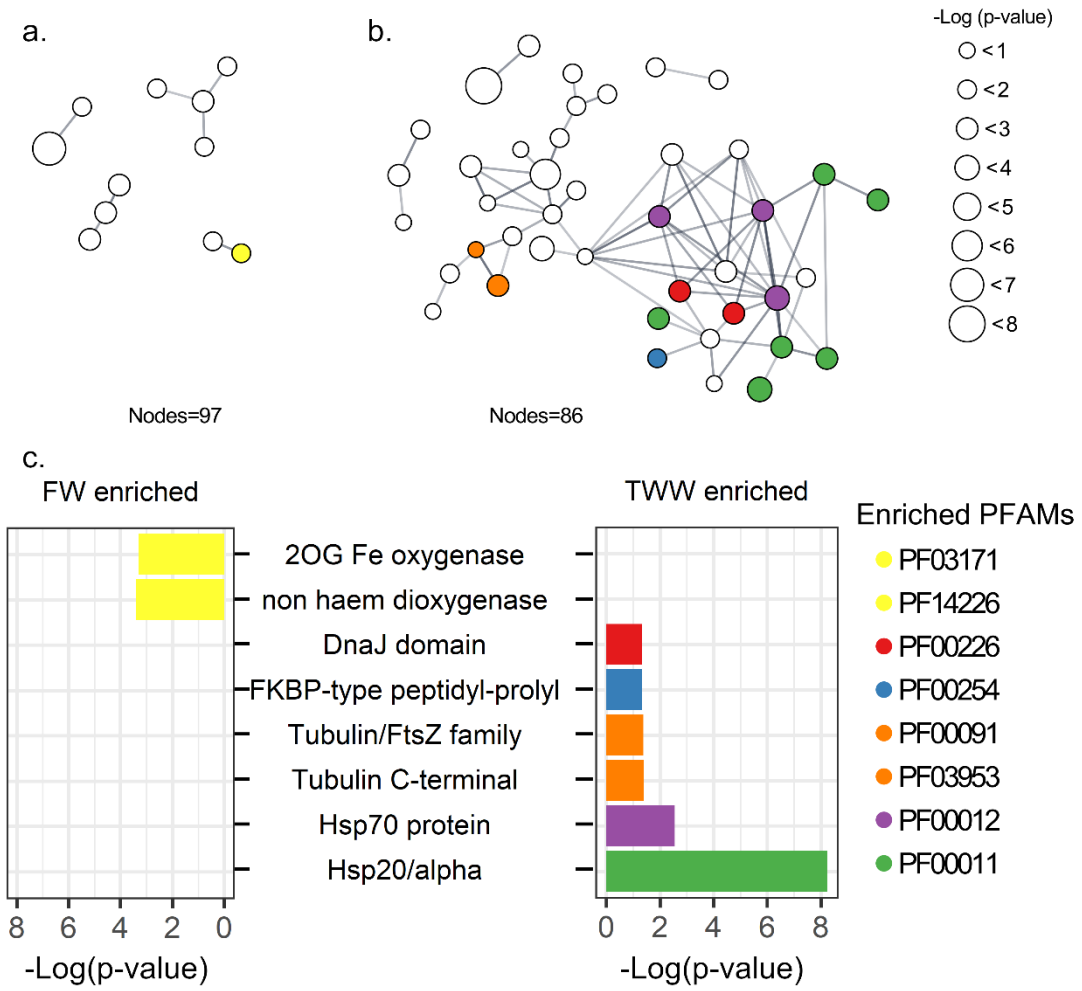
615

616

617

618

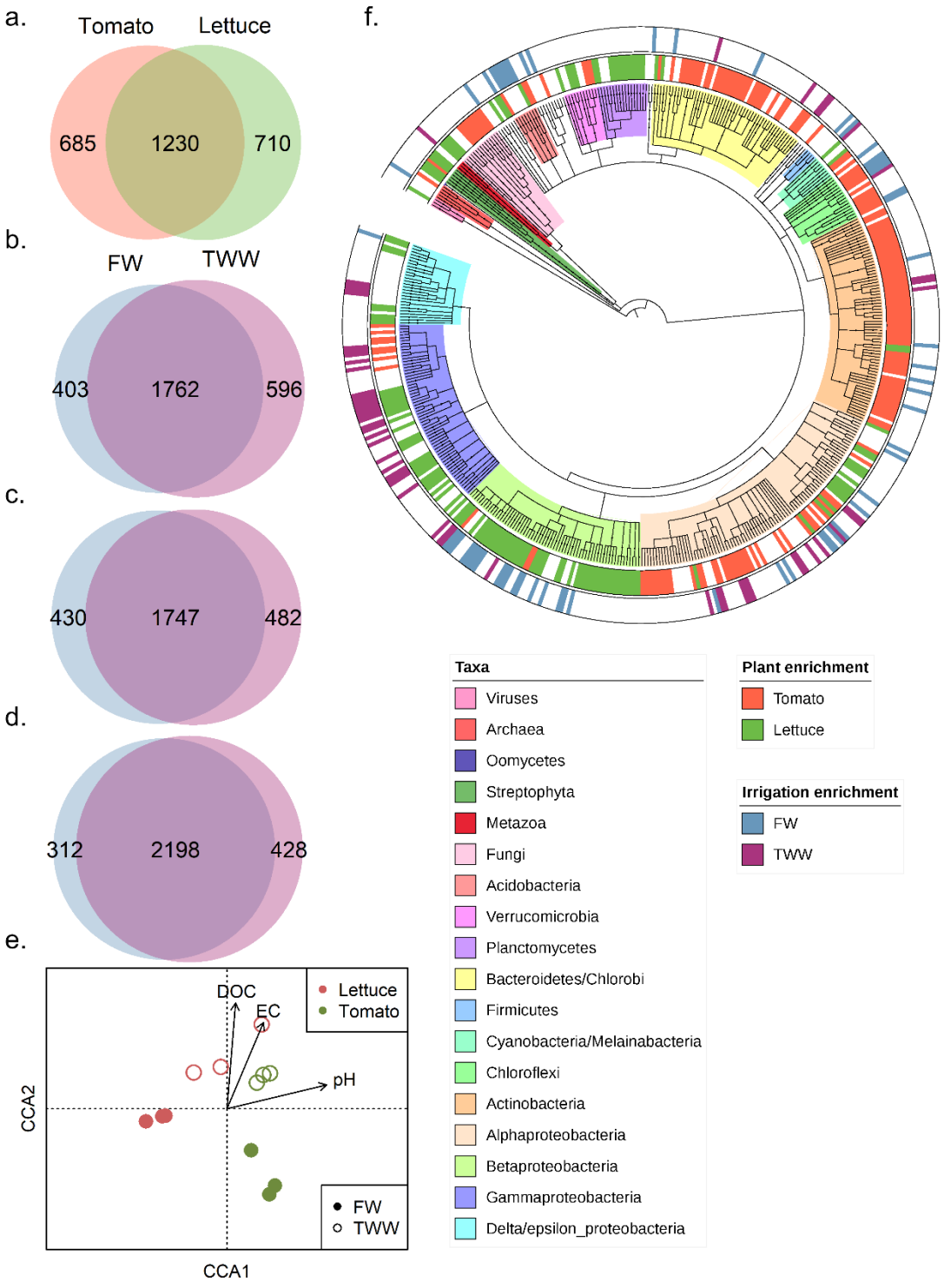
619 **Figures:**



620

621 **Figure 1:** Plant host functional response to TWW irrigation revealed by tomato and lettuce
622 transcriptome analysis. Differentially expressed plant genes between FW or TWW irrigation
623 were determined using the software package EdgeR, with significance set at FDR P<0.05.
624 Lettuce transcripts were annotated by comparison to the tomato genome. STRING protein
625 interaction networks are presented for (a) FW and (b) TWW irrigation enriched gene list.
626 Only nodes linked by edges are shown. The size of each node is proportional to the log10 (p-
627 value) of the enriched gene. Colored are PFAM protein domains significantly enriched within
628 the network landscape. The log10 (p value) of the enriched PFAM domains are presented in
629 (c) FW and (d) TWW enrichment analyses.

630

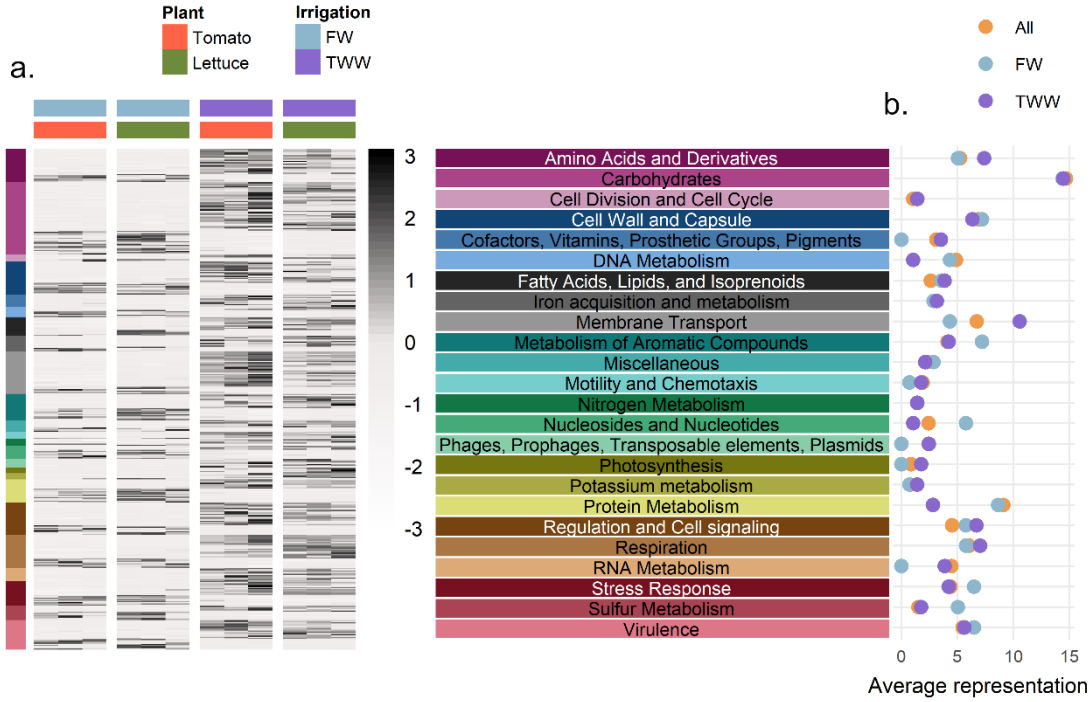


631

632 **Figure 2:** Effect of plant host type and irrigation treatment on root-associated metagenome.
633 (a) VENN diagram of differentially-abundant genes between FW-irrigated tomato and lettuce
634 (DESeq2 FDR P<0.05). VENN diagrams of differentially-abundant genes by irrigation type in
635 (b) tomato or (c) lettuce, and (d) in both hosts combined. (e) Conical correspondence
636 analysis (CCA) of all SEED-annotated genes with DOC, EC and pH as constrained variables. (f)
637 Microbial composition predicted by a least common ancestor (LCA) pipeline (MEGAN6) of

638 the predicted gene catalogue. Taxonomic groups are displayed in the inner ring, and
639 differentially-abundant taxonomic groups between the two tested plant types are
640 highlighted in the middle ring. The outer ring highlights the taxonomic groups that are
641 significantly differentially abundant between irrigation treatments.

642



643

644 **Figure 3:** Genetic profile of the 438 significantly differentially abundant genes between
645 irrigation treatments. (a) Heatmap of genes enriched or depleted (FDR P<0.01) in the
646 metagenomes of TWW-irrigated roots (displaying trimmed mean of M values- TMM). Gene
647 abundance was normalized by scaling each row separately. The gene list was clustered to
648 high hierarchy SEED categories. (b) The proportion of genes enriched or depleted in TWW-
649 irrigated root metagenomes compared to the total abundance of that category in all
650 metagenome analysis. Enriched category (TWW, magenta), deprived (FW, light blue), or the
651 proportion within the full gene catalogue (marked as "All", colored by orange), are
652 highlighted. The proportion was calculated based on the number of gene assigned to the
653 different categories with-in each data set.

654

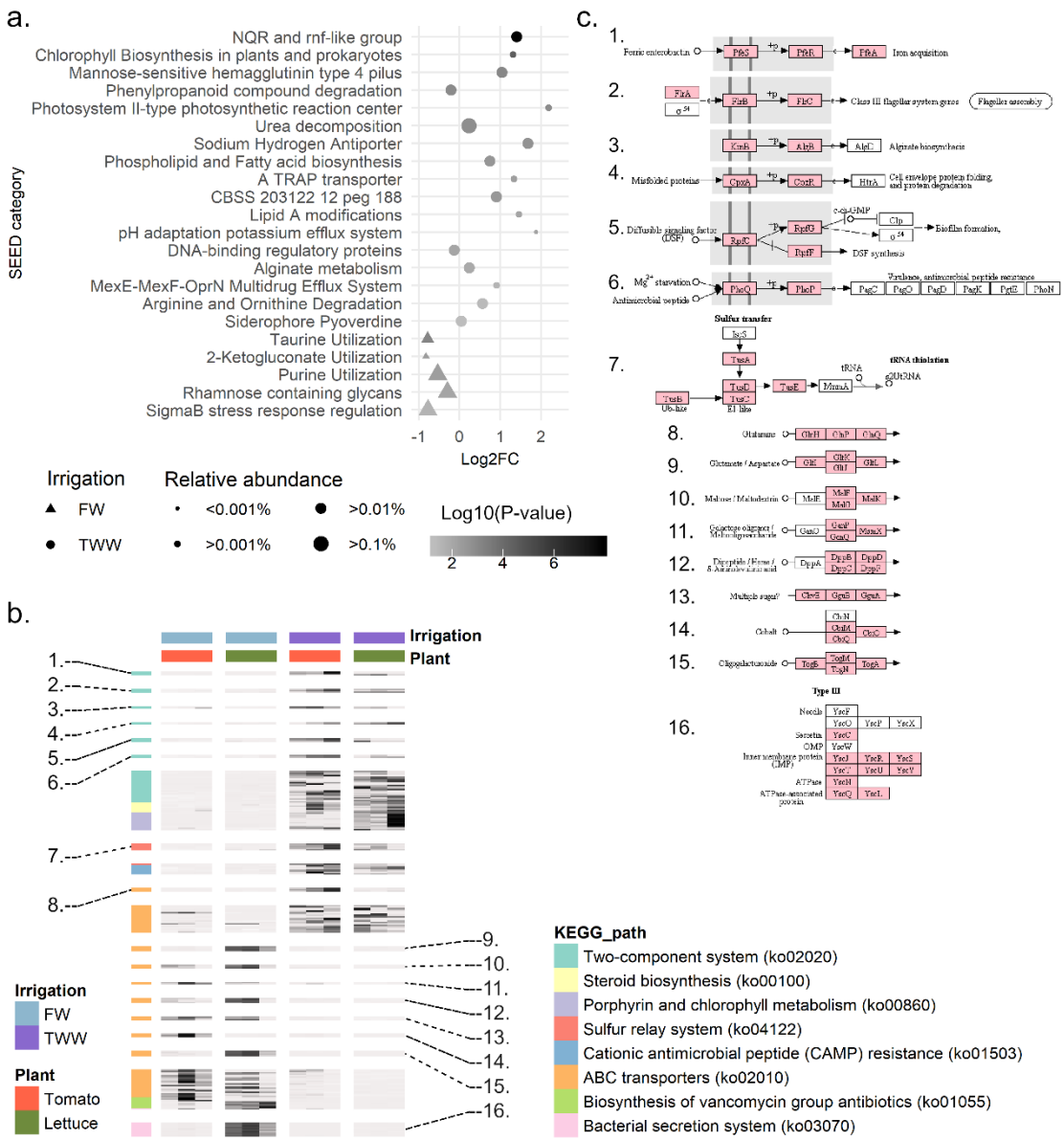


Figure 4: Analyses of SEED subsystems and KEGG pathway significantly enriched or depleted in metagenome of TWW irrigated roots. (a) Dot plot of log2(fold-change) relative abundance of enriched or depleted SEED subsystem. Significantly (FDR P<<0.05, represented by more than two gene families) enriched or depleted gene abundance was computed using the goseq software package, with correction for read counts. Symbols are proportional to the sub-system relative abundance and colored based on the enrichment/depletion log10(p-value). Circles indicate TWW-enrichment while triangles indicated TWW-depleted categories. (b) Heatmap of TWW-enriched or depleted (p-value < 0.05) KEGG pathways (characterized by keggProfiler). Genes of interest, significantly enriched or depleted in TWW-irrigated root metagenomes, are highlighted and colored in pink (c).

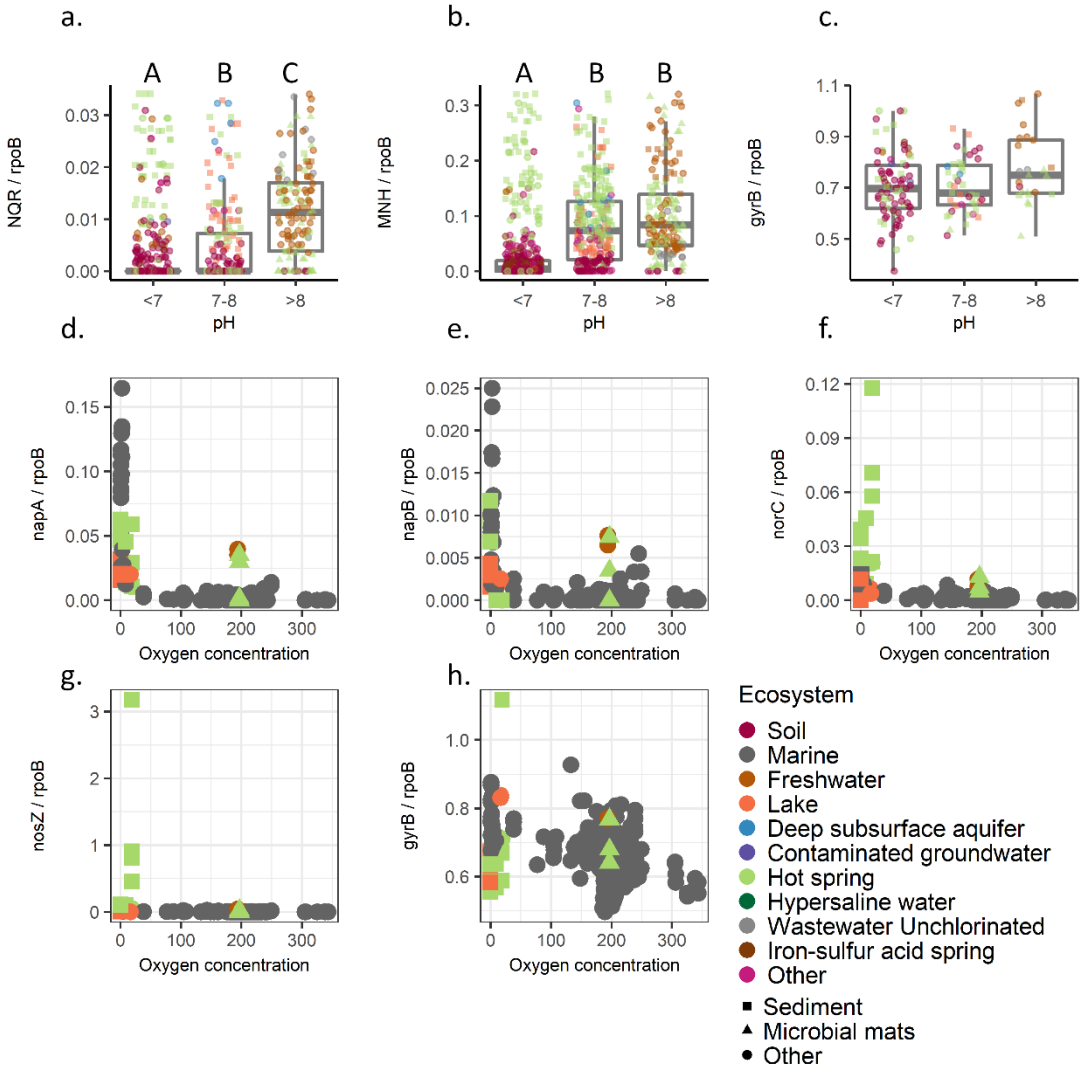


Figure 5: Meta- analysis of selected pH and oxygen responsive genes. Publicly available (by JGI) subset of 160 environmental metagenomes with pH measurements, were screened. Box plot of gene counts in acidic pH (<7), neutral (7-8), or alkaline (>8) pH for (a) NQR operon (6 subunits), (b) Na- H antiporter operon (7 subunits), and (c) *gyrB* as control. Detailed pattern for each subunit is available in Figure S6, S7. Outliers are not displayed (1.5x 0.25-0.75 quantiles). Significant differences in gene counts, by Wilcoxon rank sum test (Bonfferoni correction, P<0.05), are marked in letter report (A, B and C).

Oxygen measurements were available for subset of 257 environmental metagenomes. In these metagenomes, the abundance of *napA* (d), *napB* (e) *norBC* (f) and *nosZ* (g) was compared to oxygen levels. (h) *gyrB* was used as control. All gene counts are in proportion to *rpoB* gene.

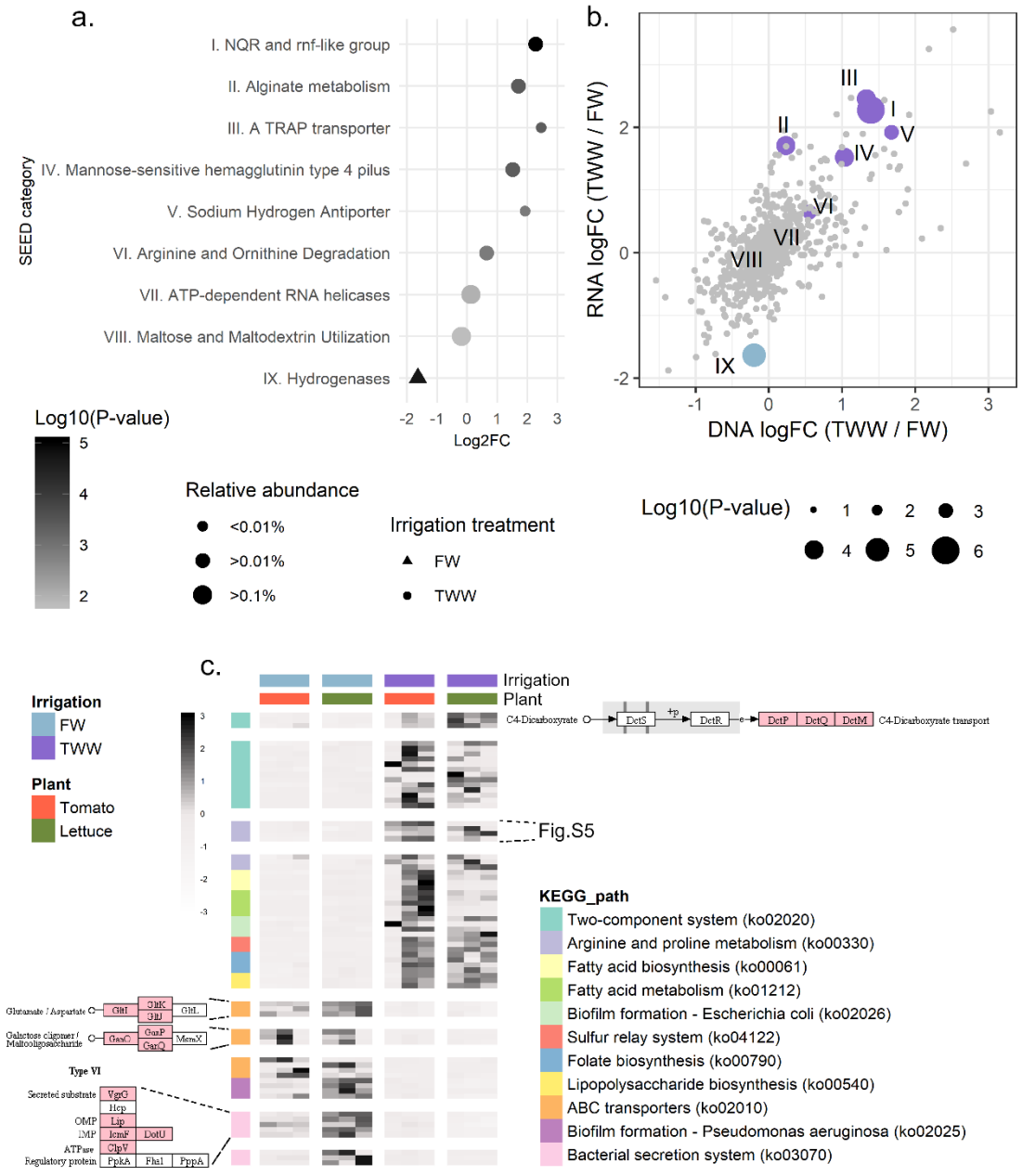
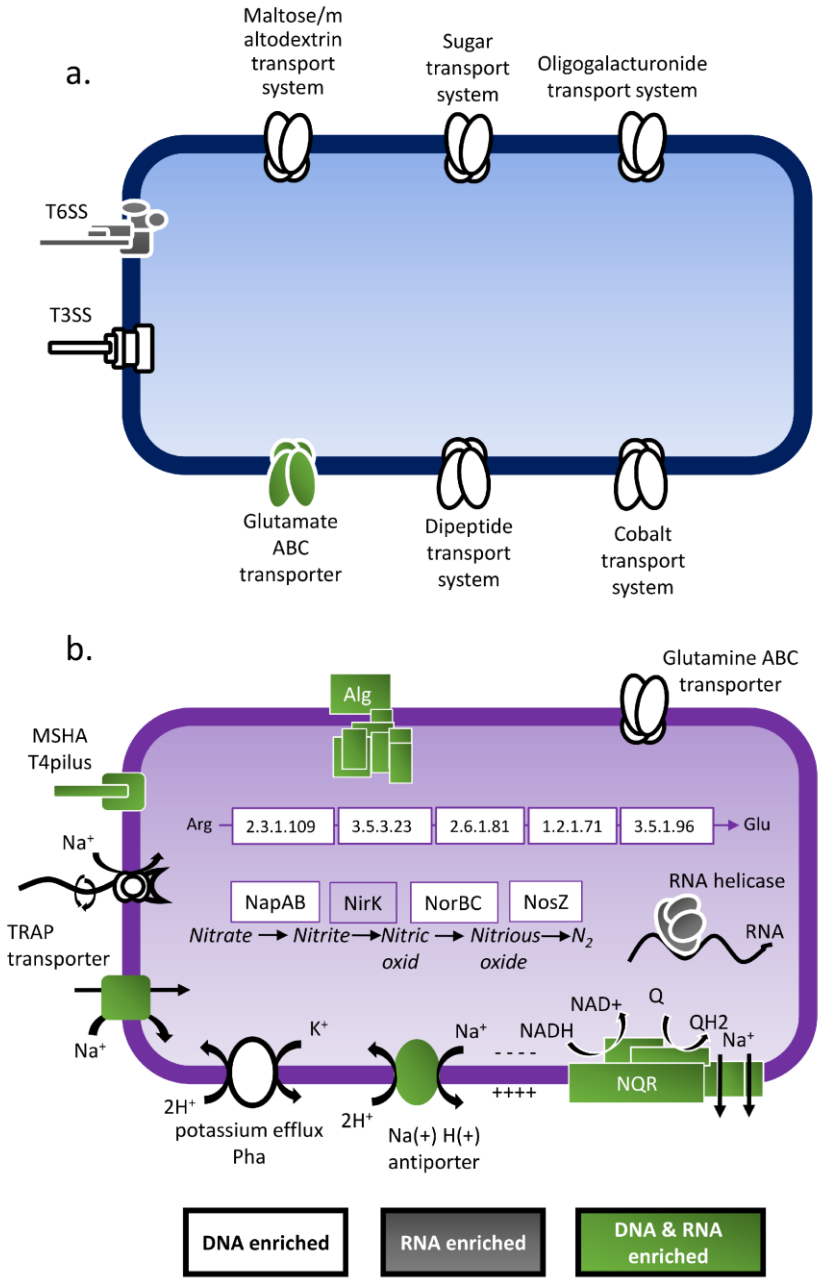


Figure 6: Expressed functions associated with irrigation treatment. (a) Dot plot of Log2(fold-change) SEED subsystems enriched or depleted in TWW irrigated root metatranscriptomes. Significantly (FDR p <0.05, represented by more than two gene families) enriched or depleted transcript abundance was computed using the goseq software package, with corrections for read abundance. Symbols are proportional to the sub-system relative abundance and colored based on the enrichment or depletion log10(p-value). Circles indicate TWW-enriched categories and triangles indicate TWW-depleted categories. (b) Differential metagenome enrichment (TWW/FW fold change) compared to differential

688 metatranscriptome expression level. Highlighted (colored) categories significantly enriched
689 or depleted in the metatranscriptome analysis. 'ns'= 'not significant'. Symbols are
690 proportional to the log10(p- value) enrichment in the metatranscriptome analysis. Numbers
691 label the enriched category, as marked in (a). (c) KEGG pathway-level enrichment or
692 depletion in TWW-irrigation root metatranscriptomes (p-value < 0.05), based on keggProfiler
693 enrichment analysis. Significantly enriched or depleted gene clusters in TWW-irrigated roots
694 are highlighted and colored in pink.



695

Figure 7: Conceptual model of hypothetical bacteria harboring physiological features (i.e., genes, pathways and modules) enriched in (a) TWW- or (b) FW-irrigated root microbiomes. White symbols indicate features that are significantly enriched at the DNA level (metagenomes), grey features are highly expressed (metatranscriptomes), and green features are significantly abundant and expressed in one treatment relative to the other.

References

1. Sunagawa S, et al. (2015) Structure and function of the global ocean microbiome. *Science* 348(6237):1261359.

2. Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 103(3):626–31.

3. Mendes LW, et al. (2015) Soil-borne microbiome: linking diversity to function. *Microb Ecol* 70(1):255–265.

4. Fierer N (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15(10):579–590.

5. Peiffer JA, et al. (2013) Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci U S A* 110(16):6548–53.

6. Lundberg DS, et al. (2012) Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488(7409):86–90.

7. Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM (2014) Taxonomical and functional microbial community selection in soybean rhizosphere. *ISME J* 8(8):1577–1587.

8. Bulgarelli D, et al. (2012) Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* 488(7409):91–95.

9. Edwards J, et al. (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* 112(8):E911-20.

10. Turnbaugh PJ, et al. (2007) The human microbiome project. *Nature* 449(7164):804–810.

11. Dewhirst FE, et al. (2010) The human oral microbiome. *J Bacteriol* 192(19):5002–17.

12. Goodrich JK, et al. (2014) Human genetics shape the gut microbiome. *Cell* 159(4):789–799.

13. Grice EA, Segre JA (2011) The skin microbiome. *Nat Rev Microbiol* 9(4):244–253.

14. Ofek M, Voronov-Goldman M, Hadar Y, Minz D (2014) Host signature effect on plant root-associated microbiomes revealed through analyses of resident *vs* . active communities. *Environ Microbiol* 16(7):2157–2167.

15. Walters WA, et al. (2018) Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proc Natl Acad Sci U S A* 115(28):7368–7373.

16. Torsvik V, Øvreås L, Thingstad TF (2002) Prokaryotic diversity--magnitude, dynamics, and controlling factors. *Science* 296(5570):1064–6.

17. Tringe SG, et al. (2005) Comparative metagenomics of microbial communities. *Science* 308(5721):554–7.

18. Sasse J, Martinoia E, Northen T (2018) Feed your friends: do plant exudates shape the root microbiome? *Trends Plant Sci* 23(1). doi:10.1016/j.tplants.2017.09.003.

19. Ofek-Lalzar M, et al. (2014) Niche and host-associated functional signatures of the root surface microbiome. *Nat Commun* 5(1):4950.

20. Edwards JA, et al. (2018) Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLOS Biol* 16(2):e2003862.

21. Chaparro JM, Badri D V, Vivanco JM (2014) Rhizosphere microbiome assemblage is affected by plant development. *ISME J* 8(4):790–803.

22. Castrillo G, et al. (2017) Root microbiota drive direct integration of phosphate stress and immunity. *Nature* 543(7646):513–518.

23. Pett-Ridge J, Firestone MK (2005) Redox fluctuation structures microbial communities in a wet tropical soil. *Appl Environ Microbiol* 71(11):6998–7007.

24. Staley C, et al. (2018) Urea amendment decreases microbial diversity and selects for specific nitrifying strains in eight contrasting agricultural soils. *Front Microbiol* 9:634.

25. Green SJ, Inbar E, Michel FC, Hadar Y, Minz D (2006) Succession of bacterial communities during early plant development: transition from seed to root and effect

754 of compost amendment. *Appl Environ Microbiol* 72(6):3975–83.

755 26. Zolti A, Green SJSJ, Ben Mordechay E, Hadar Y, Minz D (2019) Root microbiome
756 response to treated wastewater irrigation. 655:899–907.

757 27. Zilber-Rosenberg I, Rosenberg E (2008) Role of microorganisms in the evolution of
758 animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev*
759 32(5):723–735.

760 28. Overbeek R, et al. (2005) The subsystems approach to genome annotation and its use
761 in the project to annotate 1000 genomes. *Nucleic Acids Res* 33(17):5691–5702.

762 29. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic*
763 *Acids Res* 28(1):27–30.

764 30. Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins
765 and molecular chaperones in the abiotic stress response. *Trends Plant Sci* 9(5):244–
766 52.

767 31. Swindell WR, Huebner M, Weber AP (2007) Transcriptional profiling of Arabidopsis
768 heat shock proteins and transcription factors reveals extensive overlap between heat
769 and non-heat stress response pathways. *BMC Genomics* 8(1):125.

770 32. Wang C, Li J, Yuan M (2007) Salt tolerance requires cortical microtubule
771 reorganization in Arabidopsis. *Plant Cell Physiol* 48(11):1534–1547.

772 33. Abdrakhamanova A, Wang QY, Khokhlova L, Nick P (2003) Is microtubule disassembly
773 a trigger for cold acclimation? *Plant Cell Physiol* 44(7):676–686.

774 34. Sivaguru M, Pike S, Gassmann W, Baskin TI (2003) Aluminum rapidly depolymerizes
775 cortical microtubules and depolarizes the plasma membrane: evidence that these
776 responses are mediated by a glutamate receptor. *Plant Cell Physiol* 44(7):667–675.

777 35. Takemoto D, Hardham AR (2004) The cytoskeleton as a regulator and target of biotic
778 interactions in plants. *Plant Physiol* 136(4):3864–76.

779 36. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and
780 dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.

781 37. Juárez O, Barquera B (2012) Insights into the mechanism of electron transfer and
782 sodium translocation of the Na+-pumping NADH:quinone oxidoreductase. *Biochim*
783 *Biophys Acta - Bioenerg* 1817(10):1823–1832.

784 38. Jones DL, Nguyen C, Finlay RD (2009) Carbon flow in the rhizosphere: carbon trading
785 at the soil–root interface. *Plant Soil* 321(1–2):5–33.

786 39. Pang JY, Newman I, Mendham N, Zhou M, Shabala S (2006) Microelectrode ion and
787 O2 fluxes measurements reveal differential sensitivity of barley root tissues to
788 hypoxia. *Plant, Cell Environ* 29(6):1107–1121.

789 40. Blossfeld S, Schreiber CM, Liebsch G, Kuhn AJ, Hinsinger P (2013) Quantitative
790 imaging of rhizosphere pH and CO2 dynamics with planar optodes. *Ann Bot*
791 112(2):267–276.

792 41. Blossfeld S, Gansert D, Thiele B, Kuhn AJ, Lösch R (2011) The dynamics of oxygen
793 concentration, pH value, and organic acids in the rhizosphere of Juncus spp. *Soil Biol*
794 *Biochem* 43(6):1186–1197.

795 42. Munns R (2005) Genes and salt tolerance: bringing them together. *New Phytol*
796 167(3):645–663.

797 43. Nishizawa A, et al. (2006) Arabidopsis heat shock transcription factor A2 as a key
798 regulator in response to several types of environmental stress. *Plant J* 48(4):535–547.

799 44. Bhardwaj AR, et al. (2015) Global insights into high temperature and drought stress
800 regulated genes by RNA-Seq in economically important oilseed crop Brassica juncea.
801 *BMC Plant Biol* 15(1):9.

802 45. Atkinson NJ, Urwin PE (2012) The interaction of plant biotic and abiotic stresses: from
803 genes to the field. *J Exp Bot* 63(10):3523–3543.

804 46. Mittler R (2006) Abiotic stress, the field environment and stress combination. *Trends*
805 *Plant Sci* 11(1):15–19.

806 47. Sunkar R, Zhu J-K (2004) Novel and stress-regulated microRNAs and other small RNAs
807 from Arabidopsis. *Plant Cell* 16(8):2001–19.

808 48. Vlad F, et al. (2009) Protein phosphatases 2C regulate the activation of the Snf1-
809 related kinase OST1 by abscisic acid in Arabidopsis. *Plant Cell* 21(10):3170–84.

810 49. Padan E, Bibi E, Ito M, Krulwich TA (2005) Alkaline pH homeostasis in bacteria: New
811 insights. *Biochim Biophys Acta - Biomembr* 1717(2):67–88.

812 50. Krulwich TA, Sachs G, Padan E (2011) Molecular aspects of bacterial pH sensing and
813 homeostasis. *Nat Rev Microbiol* 9(5):330–343.

814 51. Vorburger T, et al. (2016) Role of the Na+-translocating NADH:quinone
815 oxidoreductase in voltage generation and Na+ extrusion in Vibrio cholerae. *Biochim*
816 *Biophys Acta - Bioenerg* 1857(4):473–482.

817 52. Ito M, Guffanti AA, Oudega B, Krulwich TA (1999) *mrp*, a multigene, multifunctional
818 locus in Bacillus subtilis with roles in resistance to cholate and to Na+ and in pH
819 homeostasis. *J Bacteriol* 181(8):2394–402.

820 53. Yamaguchi T, Tsutsumi F, Putnoky P, Fukuhara M, Nakamura T (2009) pH-dependent
821 regulation of the multi-subunit cation/proton antiporter Pha1 system from
822 *Sinorhizobium meliloti*. *Microbiology* 155(8):2750–2756.

823 54. Delgado-Baquerizo M, et al. (2018) A global atlas of the dominant bacteria found in
824 soil. *Science* 359(6373):320–325.

825 55. Mulligan C, Kelly DJ, Thomas GH (2007) Tripartite ATP-independent periplasmic
826 transporters: application of a relational database for genome-wide analysis of
827 transporter gene frequency and organization. *J Mol Microbiol Biotechnol* 12(3–
828 4):218–226.

829 56. Mulligan C, et al. (2009) The substrate-binding protein imposes directionality on an
830 electrochemical sodium gradient-driven TRAP transporter. *Proc Natl Acad Sci U S A*
831 106(6):1778–83.

832 57. Mulligan C, Fischer M, Thomas GH (2011) Tripartite ATP-independent periplasmic
833 (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* 35(1):68–86.

834 58. Atsumi T, McCartert L, Imae Y (1992) Polar and lateral flagellar motors of marine
835 Vibrio are driven by different ion-motive forces. *Nature* 355(6356):182–184.

836 59. Yorimitsu T, Homma M (2001) Na+-driven flagellar motor of Vibrio. *Biochim Biophys*
837 *Acta - Bioenerg* 1505(1):82–93.

838 60. Assouline S, Narkis K (2013) Effect of long-term irrigation with treated wastewater on
839 the root zone environment. *Vadose Zo J* 12(2):0.

840 61. Oren A (2006) Life at high salt concentrations. *The Prokaryotes* (Springer New York,
841 New York, NY), pp 263–282.

842 62. Paudel I, et al. (2018) Treated wastewater irrigation: Soil variables and grapefruit tree
843 performance. *Agric Water Manag* 204:126–137.

844 63. Dalsgaard T, et al. (2014) Oxygen at nanomolar levels reversibly suppresses process
845 rates and gene expression in anammox and denitrification in the oxygen minimum
846 zone off northern Chile. *MBio* 5(6):e01966.

847 64. Staley C, et al. (2017) Diurnal cycling of rhizosphere bacterial communities is
848 associated with shifts in carbon metabolism. *Microbiome* 5(1):65.

849 65. Naylor D, DeGraaf S, Purdom E, Coleman-Derr D (2017) Drought and host selection
850 influence bacterial community dynamics in the grass root microbiome. *ISME J*
851 11(12):2691–2704.

852 66. Šťovíček A, Kim M, Or D, Gillor O (2017) Microbial community response to hydration-
853 desiccation cycles in desert soil. *Sci Rep* 7(1):45735.

854 67. Henderson SL, et al. (2010) Changes in denitrifier abundance, denitrification gene
855 mRNA levels, nitrous oxide emissions, and denitrification in anoxic soil microcosms
856 amended with glucose and plant residues. *Appl Environ Microbiol* 76(7):2155–64.

857 68. Utada AS, et al. (2014) *Vibrio cholerae* use pili and flagella synergistically to effect
858 motility switching and conditional surface attachment. *Nat Commun* 5(1):4913.

859 69. O'Toole G, Kaplan HB, Kolter R (2000) Biofilm formation as microbial development.
860 *Annu Rev Microbiol* 54(1):49–79.

861 70. RodrÃ-guez-Navarro DN, Dardanelli MS, RuÃ-z-SaÃ-nz JE (2007) Attachment of
862 bacteria to the roots of higher plants. *FEMS Microbiol Lett* 272(2):127–136.

863 71. Jimenez-Sanchez C, Wick LY, Cantos M, Ortega-Calvo J-J (2015) Impact of dissolved
864 organic matter on bacterial tactic motility, attachment, and transport. *Environ Sci*
865 *Technol* 49(7):4498–4505.

866 72. Forde BG, Lea PJ (2007) Glutamate in plants: metabolism, regulation, and signalling. *J*
867 *Exp Bot* 58(9):2339–2358.

868 73. Angel R, Claus P, Conrad R (2012) Methanogenic archaea are globally ubiquitous in
869 aerated soils and become active under wet anoxic conditions. *ISME J* 6(4):847–862.

870 74. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data.
871 Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

872 75. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina
873 sequence data. *Bioinformatics* 30(15):2114–2120.

874 76. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat*
875 *Methods* 9(4):357–359.

876 77. Li H, et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*
877 25(16):2078–2079.

878 78. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile
879 metagenomic assembler. *Genome Res* 27(5):824–834.

880 79. Hyatt D, et al. (2010) Prodigal: prokaryotic gene recognition and translation initiation
881 site identification. *BMC Bioinformatics* 11(1):119.

882 80. Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-
883 generation sequencing data. *Bioinformatics* 28(23):3150–3152.

884 81. Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using
885 DIAMOND. *Nat Methods* 12(1):59–60.

886 82. Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data.
887 *Genome Res* 17(3):377–86.

888 83. Haas BJ, et al. (2013) De novo transcript sequence reconstruction from RNA-seq using
889 the Trinity platform for reference generation and analysis. *Nat Protoc* 8(8):1494–
890 1512.

891 84. Cantarel BL, et al. (2008) MAKER: An easy-to-use annotation pipeline designed for
892 emerging model organism genomes. *Genome Res* 18(1):188–196.

893 85. Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-
894 customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12(1):35.

895 86. Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display
896 and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44(W1):W242–
897 W245.

898 87. Young MD, Wakefield MJ, Smyth GK, Oshlack A (2010) Gene ontology analysis for
899 RNA-seq: accounting for selection bias. *Genome Biol* 11(2):R14.

900 88. Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R Package for comparing
901 biological themes among gene clusters. *Omi A J Integr Biol* 16(5):284–287.

902 89. Dixon P (2003) VEGAN, a package of R functions for community ecology. *J Veg Sci*

14(6):927–930.

90. Wickham H (2016) ggplot2: elegant graphics for data analysis (*Springer-Verlag New York*) Available at: https://ggplot2.tidyverse.org.

91. Kolde R (2015) pheatmap: Pretty heatmaps [Software].

92. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140.

93. Szklarczyk D, et al. (2017) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res* 45(D1):D362–D368.

94. Szklarczyk D, et al. (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43(D1):D447–D452.

95. Shannon P, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504.