Network mapping of root-microbe interactions in *Arabidopsis thaliana*

Xiaoqing He$^{1,2,}$, Qi Zhang$^2$, Yi Jin$^2$, Libo Jiang$^{1,2}$, and Rongling Wu$^{1,2,3,,*}$

$^1$Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing 100083, China

$^2$Center for Computational Biology, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China

$^3$Center for Statistical Genetics, Departments of Public Health Sciences and Statistics, The Pennsylvania State University, Hershey, PA 17033, USA

All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

*Corresponding author: R.W. (rwu@bjfu.edu.cn or rwu@phs.psu.edu)

*Corresponding author: Rongling Wu, Center for Computational Biology, Beijing Forestry University, Beijing 100083, China; Tel: +86-10-6233-6283; rwu@bjfu.edu.cn, or Center for Statistical Genetics, The Pennsylvania State University, Hershey, PA 17033, USA; Tel: +1-(717)531-2037; Email: rwu@phs.psu.edu
Abstract

Understanding how plants interact with their colonizing microbiota to determine plant phenotypes is a fundamental question in modern plant science. Existing approaches for genome-wide association studies (GWAS) are based on the association analysis between host genes and the abundance of individual microbes, failing to characterize the genetic architecture of microbial interactions that are thought to a determinant of microbiota structure, organization, and function. Here, we implement a behavioral model to quantify various patterns of microbe-microbe interactions, i.e., mutualism, antagonism, aggression, and altruism, and map host genes that modulate microbial networks constituted by these interaction types. We reanalyze a root-microbiome data involving 179 accessions of Arabidopsis thaliana and find that the four networks differ structurally in the pattern of bacterial-fungal interactions and microbiome complexity. We identify several fungus and bacterial hubs that play a central role in mediating microbial community assembly surrounding A. thaliana root systems. We detect 1142 significant host genetic variants throughout the plant genome and then implement Bayesian networks (BN) to reconstruct epistatic networks involving all significant SNPs and find 91 hub QTLs. Gene annotation shows that a number of the hub genes detected are biologically relevant, playing roles in plant growth and development, resilience against pathogens, root development, and improving resistance against abiotic stress conditions. The new model allows us to better understand the underlying mechanisms that govern the relationships between plants and their entire microbiota and harness soil microbes for plant production.

Keywords: Arabidopsis thaliana; root microbe interaction; hub microbes; network mapping; QTL network; hub genes

INTRODUCTION

The microbiota has been widely thought to be an important determinant of various natural
processes ranging from biogeographical cycling to human health. Many studies have characterized strong associations between the microbiota and a variety of human disorders (1-4), but research on how the microbiota impacts plant growth has not been conducted until recently (5, 6). Increasing evidence shows that the microbiota plays a pivotal role in promoting plants’ stress tolerance, determining plant productivity, improving the bioavailability of nutrients, and preventing invasion by bacterial pathogens (5, 7-12). Some bacteria can fix and preserve nitrogen in root nodules for plants (13, 14), whereas others even can modulate the timing of flowering of plants (15, 16) and save hosts from extinction (17). Under drought stress, root microbiomes can help crop plants maintain more sustainable production (10, 18).

While the microbiota affects the phenotypes of the hosts they colonize, the hosts can also shape the structure and function of the microbial communities (2, 4, 19, 20) in a way that is determined by environmental factors (2, 21). It has been commonly recognized that the microbiota and their hosts form complex but well-orchestrated interaction networks (22, 23). For example, the crop microbiome assembly is shaped by the host, even to a greater extent environmental factors (24). There is a greater variability among plant species or genotypes in their ability to recruit specific microbial communities (10, 25). Plant genes affect root metabolism, immune system functioning, and root exudate composition, which in turn influence the activity and structure of the root microbiome (26, 27). Recent studies provide a ‘cry-for-help’ hypothesis to explain that stressed plants assemble health-promoting soil microbiomes by changing their root exudation chemistry (28). It was still arguable that, to what extent, the overall microbiome is shaped by host genes.

Roots of healthy plants are colonized by multi-kingdom microbial consortia (22, 29, 30). The whole microbiome structure and function are determined by the pattern and strength of how the constituent microbes interact with each other through cooperation or competition (31). F. Getzke et al. (32) found that interactions between microbiota members, particularly bacterial-fungal interactions, contribute to plant health. Given that fungi have a strong influence on the structure of the root microbiome, characterizing both bacteria and fungi can enhance our understanding of the root microbiome (33). Several studies have identified highly interconnected ‘hub species’ in microbial networks that act as mediators between a host and its associated microbiome (2, 34). Yet, we are still unclear in which way microbes interact with each other within to shape polymicrobial
communities (35). We know little about how microbiota members contribute to the establishment, stability, and resilience of microbial communities essential for the maintenance of plant health.

Understanding the fundamental questions described above requires integrated systems approaches (1, 13). Recently, with the application of next-generation sequencing, the microbiome data and host genetic data measured at unprecedented resolution have been increasingly available (30, 36-41). From these data, genome-wide association studies (GWAS) have been developed to systematically characterize the genetic underpinnings of microbiota-host associations in plants (33, 42, 43). However, traditional GWAS models can only detect the host QTLs responsible for the abundance of individual microbes, failing to disentangle the relationships of diverse microbial species and microbe–host interactions (23, 35, 44). In a previous study, we have integrated ecology theory and network science to derive a series of mathematical descriptors for measuring all possible types of microbe-microbe interactions, and quantify how each microbe interacts with every other through a web of cooperation and competition (45). These descriptors show their utility to to map the genetic architecture of the gut microbiota (45).

In this article, we report the application of our ecology-based network model to root-microbiota interactions in Arabidopsis. As a model system, Arabidopsis has been extensively studied, aimed to explore the interactions between microbial communities and hosts. In a GWAS including 179 accessions of A. thaliana, Bergelson et al (33, 42) identified associations between the abundance of individual microbes within root microbiomes and plant genotypes. By reanalyzing this dataset, we further reveal the intricate relationship among A. thaliana and its colonizing microorganisms. We identify hub microbes within the root microbiome, characterize how microbes interact across kingdoms, and illustrate to what extent those interactions affect plant health and how this process is governed by the host genes.

**RESULTS**

**Co-occurrence networks of the root microbiota**

We developed a behavioral model to define the strengths of mutualism, antagonism, aggression, and altruism between each pair of microbes, quantitatively described by $Z_{mu}$, $Z_{an}$, $Z_{ag}$, and $Z_{al}$.
respectively (see Experimental Procedures). These mathematical descriptors have been biologically validated (45, 46), allowing us to reconstruct mutualism, antagonism, aggression, and altruism networks for the root microbiota of the *A. thaliana*. To reduce the complexity of the networks, we chose the most abundant 100 OTUs in bacteria and fungi, respectively, for the reconstruction of four types of bacterial-fungal networks (OTU1-100 are listed as bacteria and OTU100-200 as fungi). We calculated node-level topological properties (i.e., degree, betweenness, closeness and eigenvector) using the “igraph” R package. Bacterial and fungal co-occurrence network characteristics are listed in Table S1.

Interkingdom functional diversity among fungi and bacteria is important for maintaining ecosystem functioning (30) and microbial interkingdom interactions in roots can promote Arabidopsis survival (22). We calculated degree-centrality parameters to determine the relative importance of bacteria and fungi in each network. It indicates that bacteria are more central to the structure of the mutualism and altruism networks than fungi (Fig. 1A and 1D), as bacteria tend to have a higher number (i.e., degree) of network connections than fungi (P=0.00038; P=0.0001029). In contrast, fungi in the antagonism and aggression networks appears to have a higher number of network connections than bacteria (Fig. 1B and 1C; P = 0.0002128; P= 0.01961). We also calculated and compared interkingdom microbial OTU relationships (the number of links; edges information) among bacterial and fungal taxonomic groups in four interaction networks (Fig.1; Table S2).

Bacterial OTUs belonging to classes Betaproteobacteria, Flavobacteriia, Actinobacteria, Gammaproteobacteria, and Alphaproteobacteria displayed a strong mutualistic relationship with fungal OTUs belonging to classes Leotiomycetes, Dothideomycetes, Sordariomycetes, Agaricomycetes, and others, respectively (Fig. 1A). Bacterial classes, such as Actinobacteria, Alphaproteobacteria, and Deltaproteobacteria and fungal classes, such as Leotiomycetes, Dothideomycetes and Sordariomycetes were antagonistic to each other (Fig. 1B). In the aggression network, there were three bacterial classes (Betaproteobacteria, Actinobacteria, and Flavobacteriia) which were aggressive to fungal classes (Leotiomycetes, Sordariomycetes, Dothideomycetes, etc.) (Fig. 1C). Bacterial classes including Actinobacteria, Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria were altruistic to fungal classes, Dothideomycetes, Leotiomycetes, Sordariomycetes, Mortierellomycetes, etc. (Fig. 1D).

**Hubs of the co-occurrence network identification**
Hub microbes are important in shaping microbial communities due to their critical roles in maintaining network function (2). The four networks differed structurally in the pattern of social links and the number of hub microbes. Fungal and bacterial OTUs that display the highest degree and the highest closeness centrality scores may serve as hub taxa to drive fungal-bacterial interaction equilibrium in *A. thaliana* roots (24) (Fig. 2A; Table S1).

We identified six hub microbes, four bacteria and two fungi (nodes with degree >11 and closeness centrality values >0.08 in the network; P<0.01), which dominate the mutualism network. The four hub bacteria are classes Betaproteobacteria (2 OTUs), Sphingobacteriia, and Actinobacteria, and the two fungal hubs are phylum Ascomycota (2 OTUs). In the antagonism network, 26 hub microbes were found to act as ‘public enemies’, which were antagonistic to many more microbes than other microbes (degree >139 and closeness centrality values >0.78). The ‘agonists’ were observed to be more abundant than the ‘antagonists’ that are more combative (P<0.01; Fig.2B). In the aggression network, three OTUs belonging to the bacterial class Betaproteobacteria (2 OTUs) and Actinobacteria might represent the hub taxa (degree >100 and closeness centrality values >0.60; P<0.5). The hawks which are considered to aggressively repress others are abundant than the doves (those inhibited by others). The altruism network includes some ‘altruists’ (24 hub microbes; Fig. 2A; Table S1) that sacrificed their own growth by providing resources to beneficiaries (degree >90 and closeness centrality values >0.64). The hub microbes (altruists) are less abundant than the beneficiaries (Fig. 2B; P<0.01).

A total of 59 OTUs were identified as hub species, which were mainly from bacterial phyla proteobacteria (21 OTUs), Actinobacteria (12 OTUs), Chloroflexi (1 OTU), Bacteroidetes (3 OTUs) and fungal phyla Ascomycota (20 OTUs), Mortierellomycota (2 OTUs) (Table S1). The altruism network shares 4, 3, and 2 phyla with the mutualism, antagonism, and aggression networks, respectively (Fig. 2C).

**Mapping Root-microbe interactions**

We calculated six centrality indices including connectivity (Con), closeness (C(u)), betweenness (B(u)), eccentricity (E(u)), eigencentrality (G(u)), and PageRank (P(u)) (Fig. 3) using the formulas...
given in Jiang et al. (45). In the same network type, these indices exhibit differences among hosts and, also, the same index varies among network types. By treating each network index as a phenotype, we performed association mapping for the interaction networks (Fig. S1). Our model identified 1142 significant host genetic variants throughout the plant genome, which are responsible for centrality indices of each network, including 225 acting through mutualism, 845 through antagonism, 49 through aggression, and 23 through altruism (19.70% for mutualism, 73.99% for antagonism, 4.29% for aggression, and 2.01% for altruism) (Table S3). It appears that more variants control mutualism and antagonism than aggression and altruism.

**QTL networks**

We implemented Jiang et al.’s (46) procedure to reconstruct Bayesian QTL networks among the significant SNPs detected to affect each type of microbial network and identified 91 hub QTLs (Fig. 4; Table 1). Through QTL network analysis, we can better characterize how a QTL mediates microbial cooperation or competition through its epistatic interactions with other QTLs. In the QTL network for the microbial mutualism network, we identify a hub QTL that affects connectivity QTL (Fig. 4A), annotated to gene *TMK3* (*AT2G01820*) that orchestrates plant growth by regulation of both cell expansion and cell proliferation and as a component of auxin signaling (47). A hub QTL for the betweenness of microbial mutualism network is located in gene *IBR1* (*AT4G05530*), encoding indole-3-butyric acid response 1 (IBR1). IBR1 are involved in root hair elongation (48). A hub QTL for the PageRank of the mutualism network represents gene *NTRB* (*AT4G35460*) encoding NADPH-dependent thioredoxin reductase. Thioredoxin is a key regulator of intracellular redox status that determine plant development in response to biotic and abiotic stress. Thioredoxin reductase (ntra ntrb) mutant alters both auxin transport and metabolism, causing a loss of apical dominance and reduced secondary root production, etc., largely regulated by auxin (49).

In the QTL network for the microbial antagonism network, a closeness hub QTL acts like gene *UBP22* (*AT5G10790*). *UBP22* encodes a ubiquitin-specific protease, which plays roles in regulating plant development and stress responses (50). A hub QTL for the betweenness of the microbial antagonism network is located in gene *Hrd1A* which may be an important regulator of heat stress response in Arabidopsis (51). A hub QTL for the closeness of the microbial antagonism network acts like gene *PEPR2* (*AT1G17750*) encoding PEP1 receptor 2, which is transcriptionally induced by wounding and pathogen-associated molecular patterns and contributes to defense responses in...
Arabidopsis (52). CHL1(AT5G40090) is the hub QTL for the eigenvector of the antagonism network, which encodes disease resistance protein (TIR-NBS class). TIR-NBS protein is involved in disease resistance in Arabidopsis (53).

There are 12 pleiotropic genes including UPL4(AT5G02880), which are detected to influence multiple types of microbial networks or properties (Table 1). UPL4 encodes a ubiquitin-protein ligase, function additively in the regulation of plant growth and development, and positively modulate immune hormone salicylic acid (SA)-mediated basal and induced resistance responses (54).

**DISCUSSION**

Plant rhizosphere is considered as the second genome of plants. As such, research on rhizosphere interactions has now become one of the hottest topics in modern biology and agriculture (55). However, the intrinsic principles governing the assembly of the root microbial community remain unclear (56). The potentially beneficial bacteria and fungi may serve as a valuable foundation for bio-fertilizer development in agriculture and forestry. Our understanding of the underlying mechanisms that govern the relationships between plants and their entire microbiota will be crucial for the benefit of sustainable plant growth (57).

**Co-occurrence network**

Currently, network analysis has emerged as an extremely promising approach for modeling complex biological systems and can potentially provide deep and unique perspectives on microbial interactions and ecological assembly rules beyond those of simple richness and composition (58, 59). Properties of co-occurrence networks can reveal the intrinsic mechanisms of microbial interactions in response to environmental disturbance (38, 60). The connection and strength of the network even are crucial for the resistance to the pathogens. In this study, based on four descriptors, we successfully reconstructed corresponding 200-node networks accordingly. Each described root microbiome interactions according to a different ecological interaction metric. The four networks differ structurally in the pattern of bacterial-fungal interactions and microbiome complexity and the number of hub microbes.
Microbial interkingdom interactions

In our study, bacteria have a higher degree in the mutualism and altruism networks and are more central to the structure of networks. Fungi have more connections in the antagonism and aggression networks, however, based on the directions of edges, bacteria still are more antagonistic and aggressive to fungi (Table S2). Understanding the interaction among different species within a community is one of the central goals in ecology (61). Microbial interkingdom interactions in roots could promote Arabidopsis survival. Bacterial communities aid in maintaining the microbial balance and protect host plants against the detrimental effects of filamentous eukaryotic microbes (22). A previous research showed that microbes tended to be positively related within kingdoms but negatively related between kingdoms (2). Besides bacteria and fungi, rhizosphere bacteriophages and protists also play roles in plant health (62, 63), which should be included in further research.

Hub microbes

The identification of network hubs and their importance in microbial community structure has crucial implications for studying microbe–microbe interactions and can facilitate the design of strategies for future targeted biocontrol (2). Hub microorganisms have a regulatory influence on the network of microbial interactions, which can exert strong effects on microbiome assembly and serve as mediators between the plant and microbiome (2). According to centrality measurements, such as degree, closeness centrality and betweenness centrality, hub microorganisms are tightly connected within a co-occurrence network of (34). We identified hub microbes in four types of networks, mutualism, antagonism, aggression, and altruism. The most dominant taxa as hub microbes belong to bacterial phyla Proteobacteria (21 OTUs), Actinobacteria (12 OTUs) and fungal phyla Ascomycota (20 OTUs). This is consistent with the finding that Proteobacteria, Actinobacteria and Ascomycota are the most abundant phyla in plants and soil (24, 64). Actinobacteria is one of the bacteria whose dysbiosis in abundance in tomato rhizosphere causes the incidence of bacterial wilt disease (12).

Host QTLs

Plant phenotypes are inextricably shaped by their interactions with microbes (1). In a well-designed GWAS study, J. Bergelson et al. (33) found a few significant QTLs that are associated with root
microbial species richness and community structure, which are involved in plant immunity, cell-wall integrity, root and root-hair development. In this study, we used a newly developed network mapping model (45) to reanalyze Bergelson et al.'s (33) data, characterizing previously undetected QTLs that mediate microbial interactions. We found that most of the QTLs detected by the new model can be annotated to candidate genes with known biological functions including plant growth and development, resilience against pathogens, root development, and improved resistance against abiotic stress conditions (Table 1; Table S4). Understanding how microbes improve plant stress resistance will enhance our understanding of how plants survive in stress conditions. In the near future, it will be crucial to unravel the complex network of genetic, microbial, and metabolic interactions, including the signaling events mediating microbe–host interactions.

Future prospects

Scientists has also linked the phyllosphere microbiome to plant Health (65) and found host genes could affect bacterial communities in the phyllosphere (20). Building synthetic microbiomes in plants has been proved to be useful for future research on plant–microbe interactions (66–71). Studies have showed the potential of microbiome adjustment tailored to bring benefits for plant growth and resistance to biotic and abiotic challenges (72). Bioorganic fertilizers promote indigenous soil plant-beneficial consortium to enhance plant disease suppression (73). The design of more efficient biofertilizers to update soil function has important implications for the manipulation of crop microbiomes for sustainable agriculture. Our work provides a comprehensive exploration of microbial interkingdom interactions, hub microbes, and plant genes for the structure of the root microbiome. The results obtained could help design synthetic microbiomes beneficial for plant growth.

EXPERIMENTAL PROCEDURES

Root microbiome experiment

Bergelson et al (33, 42) conducted a genome-wide association study (GWAS) for the root microbiome in Arabidopsis thaliana. The study included 179 accessions of A. thaliana, each measured for the bacterial and fungal abundance of the root microbiota using a 16S/ITS rRNA gene
sequencing technique and genotyped for Arabidopsis SNPs by a high-throughput sequencing technology (Table S5).

**Microbial interactions analysis**

We chose the 100 most abundant bacterial OTUs and the 100 most abundant fungal OTUs (Table S6) to reconstruct microbial interaction networks using a microbial behavioral network model (45). This model is based on mathematical descriptors of four types of microbe-microbe interactions, mutualism, antagonism, aggression, and altruism, expressed as

\[
Z_{\text{mu}} = \frac{x_u x_v}{x_u - x_v} \quad \text{Mutualism}
\]

\[
Z_{\text{an}} = \frac{1}{(x_u x_v)(x_u - x_v)} \quad \text{Antagonism}
\]

\[
Z_{\text{ag}} = \frac{x_u}{x_v} \quad \text{Aggression}
\]

\[
Z_{\text{al}} = 1 - \frac{x_v}{x_u} \quad \text{Altruism}
\]

where \(x_u\) and \(x_v\) (\(x_u > x_v\), \(u \neq v\), \(u, v = 1, \ldots, m\)) are the abundance of two microbes \(u\) and \(v\) and \(m\) is the number of microbes. Based on the above equations, we use the corrected microbial abundance to quantify four interactional relationships of two microbes \(u\) and \(v\). The descriptor, \(Z_{\text{mu}}\), can be used to quantify a cooperative relationship (mutualism) between two microbes. The descriptor, \(Z_{\text{an}}\), can be used to quantify a competitive relationship (antagonism) between two microbes. The descriptor, \(Z_{\text{ag}}\), can represent the utilization extent (aggression) of a more abundant microbe to a less abundant microbe. The descriptor, \(Z_{\text{al}}\), can represent the sacrifice extent (altruism) of a more abundant microbe to a less abundant microbe.

**Microbial networks analysis**

By microbial interactions analysis, we can reveal internal workings within the root microbial community. The interaction networks can be visualized using Gephi (https://gephi.org/). We constructed the corresponding network for mutualism, antagonism, aggression, and altruism, respectively. Emergent properties of each network can be calculated in the “igraph” R package (74). We calculated six network indices to describe the features of various networks, including connectivity (Con), closeness (C(u)), betweenness(B(u)), eccentricity (E(u)), eigenvector (G(u)) and Pagerank (P(u)). The specific calculation method was described by Jiang et al. (45). The heat maps of each network index were generated by package `pheatmap` in R (https://CRAN.R-project.org/package=pheatmap). Meanwhile, microbial networks can be used to
statistically identify hub taxa. We calculated the degree of each node for every network using the “igraph” R package (74). It is generally believed that hub microbes with a high degree and closeness centrality value play crucial roles in microbial networks (2, 24).

**Mapping microbial network properties**

To study how host genes influence root microbiomes, we consider six network property parameters as phenotypic traits that are associated with host SNPs (Single Nucleotide Polymorphisms). We chose those SNPs with MAF > 5% for association analysis. A regression model of log-transformed phenotypes at a SNP is expressed as

\[ y_i = \mu + \sum_{j=1}^{2} x_i \beta + e_i \]

where \( y_i \) represents the phenotype of the \( i \)th host, \( \mu \) is the mean of the phenotypes over all hosts, \( x_i \) is the genotype indicator of the \( i \)th host which is 0 for high-frequency allele and 1 for low-frequency allele, \( \beta \) is the geneic effect of the SNP and \( e_i \) is a random error value. Then, we used \( \text{lm} \) function in R for association analysis from which to get the \( P \)-value of each SNP. Package \textit{qqman} (https://CRAN.R-project.org/package=qqman) was used to draw the Manhattan plot. By statistical testing, we can find significant SNPs that are associated with each network property.

**QTL networks**

Many existing approaches attempt to reveal the genetic architecture of complex traits through identifying key individual genes underlying the traits. However, epistatic interactions among different genes have been increasingly recognized to play an important role in genetic control. Several approaches have been developed to map epistatic interactions based on gene pairwise analysis, failing to systematically chart a network of epistasis involving all genes. More recently, Jiang et al. (46) proposed an analytical procedure of reconstructing epistatic networks from mapping data. This procedure was used to infer QTL networks of the significant SNPs that mediate the emergent properties of microbial networks. At each SNP, we calculated the mean value of each genotype for a network parameter and assigned this value to each \textit{Arabidopsis} accession, transforming the GWAS data structure from its SNP-phenotype illustration to SNP-based genotype representation. We implemented Bayesian networks (BN) to reconstruct genetic networks involving
all significant SNPs for each network parameter. The BN-based QTL networks are directed acyclic graphs, encoded by casual SNP-SNP interactions. We identified hub QTLs that play a crucial role in the genetic architecture of plant microbiomes assembly.

**Data availability**

The data used can be downloaded at https://doi.org/10.1038/s41598-018-37208-z and the computer code can be freely downloaded at https://github.com/lenahe2006/mSystem for any purpose of research. Also, all questions in data and computation can be addressed to the corresponding author.

**Acknowledgments**

We thank Dr. Bergelson and Dr. Horton for supplying their Arabidopsis microbiome and SNP data to us. This work was funded by Natural Science Foundation of China (31971398, 31700633), the Fundamental Research Funds for the Central Universities (2017JC05, 2015ZCQ-SW-06), and Science and Technology Service Network Initiative (KFJ-STS-ZDTP-036).

**Author contributions**

X.H. performed data analysis and wrote the manuscript. Q.Z. wrote the R software. L.J. and Y.J. critically revised the manuscript. R.W. conceived of the idea and formulated the model. All authors discussed results.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**


Table 1 Numbers of mutualism, antagonism, aggression, and altruism QTLs that affect the emergent properties of ecological networks. The names of genes to which QTLs are annotated are given below.

<table>
<thead>
<tr>
<th>Emergent property</th>
<th>Con</th>
<th>C(u)</th>
<th>B(u)</th>
<th>E(u)</th>
<th>G(u)</th>
<th>P(u)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mutualism QTLs</strong></td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>AT5G35602</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G05594</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMK3</td>
<td>-</td>
<td>ARPC3</td>
<td>RAD23A</td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>AT1G11700</td>
<td>AT4G10201</td>
<td>CYP721A1</td>
<td>IBR1</td>
<td>AT1G04867</td>
<td>AT1G04867</td>
<td></td>
</tr>
<tr>
<td>AT1G04867</td>
<td>AT4G05594</td>
<td>AT4G05594</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMB2296</td>
<td>TNPO3</td>
<td>AT3G22421</td>
<td>WSS1/WSS1B</td>
<td>VIII A</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Antagonism QTLs</strong></td>
<td>0</td>
<td>25</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hrd1A</td>
<td>AT3G01850</td>
<td>SS5</td>
<td>AT5G11430</td>
<td></td>
</tr>
</tbody>
</table>
Gene

AT5G35555 - AT1G72230

UBP22 PIN8 CSTF77

- CHL1

AT3G12350 PXMT1

DELTA-VPE DFL2

DEL2 DFL2

MAIN AT4G03410

CAP1

DEL2

- AT4G31360

AT4G25610

NAC058

AT5G20660

- AT2G34185

AT5G19480

AT4G00320
<table>
<thead>
<tr>
<th>Gene</th>
<th>Aggression QTLs</th>
<th>Altruism QTLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPL4</td>
<td>AT3G13662</td>
<td>2</td>
</tr>
<tr>
<td>-</td>
<td>PEPR2</td>
<td>1</td>
</tr>
<tr>
<td>AT5G11750</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>-</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>-</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>GA2OX5</td>
<td>AT1G12760</td>
<td>1</td>
</tr>
<tr>
<td>AT5G19097</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>AT5G35495</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>PDE320</td>
<td>AT3G29175</td>
<td>1</td>
</tr>
<tr>
<td>AT3G30405</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>-</td>
<td>AT3G14800</td>
<td>2</td>
</tr>
<tr>
<td>AT1G14800</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>UPL4</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>-</td>
<td>AT4G08280</td>
<td></td>
</tr>
<tr>
<td>EGRET</td>
<td>RGLG5</td>
<td>2</td>
</tr>
<tr>
<td>AT5G60470</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>PUM11</td>
<td></td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1 Microbial $Z_{mu}$-based mutualism network (A), $Z_{an}$-based antagonism network (B), $Z_{ag}$-based aggression network (C), and $Z_{al}$-based altruism network (D) at the OTU level of the root microbiome in *Arabidopsis thaliana*. In each network, bacteria and fungi are distinguished by different colors. The network analysis was performed in the “igraph” R package and visualized in Cytoscape v3.7.1. The number of links between root inter-kingdom microbes was given at the right. Bacterial and fungal OTUs were grouped at the class level and sorted according to number of edges between bacteria and fungi within each network. In boxplots, the fungal and bacterial degrees were calculated to determine the relative importance of bacteria and fungi in each network.

Figure 2 (A) The identity of each OTU is labeled by a number, 1 to 100 for bacteria and 101 to 200 for fungi. In each network, hub microbes are highlighted in dark circles. The distribution of ‘Hub microbes’ in four different microbial networks were based on degree and closeness centrality values. These two values of each OTU within each network were given at the right. The red dotted line represents the screening cutoffs of ‘Hub microbes’ corresponding to each network. (B) The abundance of hub microbes within each network. (C) The shared hub microbes within each microbial network. This ‘shared network’ was represented in Cytoscape. Circle shapes represented hub microbes from each microbial network and irregular shapes represented different microbial network types. The edges were connected to hub microbes and microbial network. The distribution of ‘Hub microbes’ in four different microbial networks were based on degree and closeness centrality values. The red dotted line represents the screening cutoffs of ‘Hub microbes’ corresponding to each network. Visualization were done with Gephi for four microbial networks.

Figure 3 Heatmap of six emergent property indices constituting (A) Mutualism, (B) Antagonism, (C) Aggression, and (D) Altruism among 200 OTUs for network properties for each host.

Figure 4. Bayesian QTL networks of significant SNPs mediating interaction networks of microbial mutualism (A), antagonism (B), aggression (C), and altruism (D). Hub QTLs within each genetic network are highlighted in green circles. The emergent properties of each microbial network are
described by connectivity (Con), closeness (C(u)), betweenness (B(u)), eccentricity (E(u)), eigencentrality (G(u)), and PageRank (P(u)).

SUPPLEMENTAL MATERIAL

Table S1 Bacterial and fungal co-occurrence network characteristics and hub microbes.
Table S2 The edge information in the four networks.
Table S3 Significant QTLs associated with the microbial interaction network.
Table S4 The gene enrichment analysis of Hub QTLs.
Table S5 The list of accessions.
Table S6 The relative abundance of top 100 OTUs in bacteria and fungi.

Figure S1 Manhattan plots of GWAS results for Arabidopsis thaliana genomic regions associated with emergent properties of root microbial networks. The x-axis showed SNP positions (Mb) and the y-axis was the –log(P-value) resulting from the association test. Each dot in the plot represented an SNP, and a reference line was used on the y-axis to reflect genome-wide significance. The red dotted lines correspond to the chosen significant threshold –log_{10}(P)≥5. (A) Mutualism, (B) Antagonism, (C) Aggression, and (D) Altruism.