Journal of Visualized Experiments

Divergence of root microbiota in different habitats based on weighted Co-expression networks

--Manuscript Draft--

Article Type:	Invited Methods Collection - Author Produced Video
Manuscript Number:	JoVE62205R5
Full Title:	Divergence of root microbiota in different habitats based on weighted Co-expression networks
Corresponding Author:	Jing Tang, Ph.D Key Laboratory of State Forestry Administration on Biodiversity Conservation in Karst Mountainous Areas of Southwestern China,School of Life Sciences,Guizhou Normal University Guiyang, Guizhou CHINA
Corresponding Author's Institution:	Key Laboratory of State Forestry Administration on Biodiversity Conservation in Karst Mountainous Areas of Southwestern China,School of Life Sciences,Guizhou Normal University
Corresponding Author E-Mail:	tangjing2016@gznu.edu.cn
Order of Authors:	Jing Tang, Ph.D
	Xinzhu He
	Ming Tang
	Xiaorong Xu
	Ximin Zhang
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US$1200)
Please specify the section of the submitted manuscript.	Biology
Please confirm that you have read and agree to the terms and conditions of the author license agreement that applies below:	I agree to the Author License Agreement
Please provide any comments to the journal here.	
Please indicate whether this article will be Standard Access or Open Access.	Standard Access ($1400)
Please confirm that you have read and agree to the terms and conditions of the video release that applies below:	I agree to the Video Release

1   **TITLE:**
2   **Divergence of root microbiota in different habitats based on weighted correlation networks**
3
4   **AUTHORS:**
5   Jing Tang[1, 2#], Xinzhu He[1#], Ming Tang[1], Xiaorong Xu[2], Ximin Zhang[1*]
6
7   1. Key Laboratory of State Forestry Administration on Biodiversity Conservation in Karst
8   Mountainous Areas of Southwestern China, School of Life Sciences, Guizhou Normal
9   University, Guiyang, Guizhou, P R China
10 2. State Key Laboratory of Plant Physiology and Development in Guizhou Province, Guizhou
11 Normal University, Guiyang, Guizhou, PR China
12
13 [#]: Jing Tang and Xinzhu He contributed equally to this work.
14
15 Jing Tang: tangjing2016@gznu.edu.cn
16 Xinzhu He: 1326562756@qq.com
17 Ming Tang: 8043484@qq.com
18 Xiaorong Xu: 512742911@qq.com
19 Ximin Zhang: zhxm409@163.com
20
21 * corresponding author
22 zhxm409@163.com
23

27 **Summary:**
28 Network analysis was applied to evaluate the association of various ecological microbial
29 communities, such as soil, water and rhizosphere. Presented here is a protocol on how to use
30 the WGCNA algorithm to analyze different co-occurrence networks that may occur in the
31 microbial communities due to different ecological environments.
32

33 **Abstract**
34 The root microbiome plays an important role in plant growth and environmental adaptation.
35 Network analysis is an important tool for studying communities, which can effectively explore
36 the interaction relationship or co-occurrence model of different microbial species in different
37 environments. The purpose of this manuscript is to provide details on how to use the weighted
38 correlation network algorithm to analyze different co-occurrence networks that may occur in
39 microbial communities due to different ecological environments. All analysis of the
40 experiment is performed in the WGCNA package. WGCNA is an R package for weighted
41 correlation network analysis. The experimental data used to demonstrate these methods
42 were microbial community data from the NCBI (National Center for Biotechnology
43 Information) database for three niches of the rice (Oryza sativa) root system. We used the
44 weighted correlation network algorithm to construct co-abundance networks of microbial
45 community in each of the three niches. Then, differential co-abundance networks among
46 endosphere, rhizoplane and rhizosphere soil were identified. In addition, the core genera in
47 network were obtained by the "WGCNA" package, which plays an important regulated role in

48 network functions. These methods enable researchers to analyze the response of microbial
49 network to environmental disturbance and verify different microbial ecological response
50 theories. The results of these methods show that the significant differential microbial
51 networks identified in the endosphere, rhizoplane and rhizosphere soil of rice.
52
53 **Introduction**
54 Microbiome research has important implications for understanding and manipulating
55 ecosystem processes [1,2]. Microbial populations are interconnected by interacting ecological
56 networks, whose characteristics can affect the response of microorganisms to environmental
57 changes [3,4]. Furthermore, the properties of these networks affect the stability of microbial
58 communities, and are closely associated with soil function[5]. Weighted gene correlation
59 network analysis has now been widely applied for research on the relationship between genes
60 and microbial communities[6]. Previous studies have focused mainly on the associations
61 between networks of different genes or populations and the outside world[7]. However, the
62 differences in correlation networks formed by microbial populations under different
63 environmental conditions have been scarcely investigated. The purpose of the research
64 presented in this paper is to provide insights and details on the rapid implementation of the
65 WGCNA algorithm to construct a co-occurrence network of microbiome samples collected
66 under different environmental conditions. Based on the analysis results, we assessed the
67 composition and differences of the population and further discussed the relationship between
68 different microbial populations. The following basic flow of weighted correlation network
69 algorithm[8] was applied. First, a similarity matrix needed to be constructed by calculating the
70 Pearson correlation coefficient between the Operational Taxonomic Units (OTU) expression
71 profiles. Then, the parameters of the adjacency functions (the power or the sigmoid adjacency
72 functions) were adopted with a scale-free topology criterion, the similarity matrix was
73 transformed into an adjacency matrix, and each co-occurrence network corresponded to an
74 adjacency matrix. We used average linkage hierarchical clustering coupled with the TOM-
75 based dissimilarity to group OTUs with coherent expression profiles into modules. Further, we
76 calculated the relationship between conservative statistics and the related parameter analysis
77 modules, finally identifying the hub OTU in the module. These methods are particularly
78 suitable for analysis of the differences in network structures among various microbial
79 populations under divergent environmental conditions. In this manuscript, we have described
80 in detail the method of co-expression network development, the analysis of the dissimilarities
81 between the modules, and have provided a brief overview of the steps in the procedure
82 applied to obtain the core species in different module networks.
83
84 **Protocol**
85
86 **1. Data Download**
87
88 1.1. Download the data of the accession PRJNA386367 form the NCBI database. From the data
89 of the accession PRJNA386367, select the rhizosphere, rhizoplane, and endosphere
90 microbiome data from rice plants grown for 14 weeks in a submerged rice field in Arbuckle,
91 California in 2014.
92
93 NOTE: The rhizosphere, rhizoplane, and endosphere microbiome data were presented by the
94 OTUs table in accession PRJNA386367.

95

2.   Optimal power value determination

96

97

NOTE: The WGCNA package contains all of the following functional parameters. WGCNA is an
R package for weighted correlation network analysis. The key command lines refer to the
Supplement S1.

98
99
100

101

2.1. In the R language environment, open the Rstudio software and install the WGCNA
package.

102
103

104

2.2. Load the data and use the goodSamplesGenes function to check the correctness of the
data. Execute the command lines:

105
106

```
"gsg = goodSamplesGenes(datExpr0, verbose = 3)
gsg$allOK "
```
Click **Run**.

107
108
109

110

2.3. Check for outliers and store samples that meet the requirements. When the check result
is TRUE, continue to the next step. Save the result.

111
112

113

2.4. Use the PickSoftThreshold function to calculate the scale-free index R^2 of the two groups
of the data under different power values. Execute the command line:

114
115

```
"sft = pickSoftThreshold(datExpr0, powerVector = powers, verbose = 5)"
```
Click **Run**.

116
117

118

2.5. Visualize the results (**Figure 1**). Execute the command line:

119

```
"plot(sft$fitIndices[,1], -sign(sft$fitIndices[,3])*sft$fitIndices[,2],
       xlab="Soft Threshold (power)",ylab="Scale Free Topology Model Fit,signed
R^2",type="n",
       main = paste("ES_Scale independence"));
text(sft$fitIndices[,1], -sign(sft$fitIndices[,3])*sft$fitIndices[,2],
       labels=powers,cex=cex1,col="red");
abline(h=0.9,col="red")
plot(sft$fitIndices[,1], sft$fitIndices[,5],
       xlab="Soft Threshold (power)",ylab="Mean Connectivity", type="n",
       main = paste("ES_Mean connectivity"))
text(sft$fitIndices[,1], sft$fitIndices[,5], labels=powers, cex=cex1,col="red")"
```
Click **Run**.

120
121
122
123
124
125
126
127
128
129
130
131

132

NOTE: The premise of the weighted correlation network algorithm is that the established co-
expression network structure conforms to the standards of the scale-free topology criterion,
increasing its robustness. A scale-free index closer to 1 indicates a network structure that is
closer to the scale-free network.

133
134
135
136

137

2.6. Select the power value when the scale-free index R^2 squared greater than 0.9 and proceed
to the next step of analysis.

138
139

140

141 NOTE: When the scale-free index is close to 1, the network structure is closer to the scale-free
142 network. When analyzing two or more networks, it is necessary to choose to make each
143 network close to the power value of the scale-free network to satisfy the comparability
144 between the co-expressed networks.
145
146 **3. Construction of a co-expression network and module identification**
147
148 NOTE: Based on the above calculated power value, the co-occurrence network is constructed.
149 The key command lines refer to the **Supplement S2**.
150
151 3.1. Use the adjacency function in the WGCNA package to add signed parameters for the
152 construction of a symbolic co-occurrence network. Execute the command line:
153   "adjacency = adjacency(datExpr0, power = softPower)"
154   Click **Run**.
155
156 3.2. Apply the TOM-similarity function to develop a topological overlapping network and
157 calculate the dissimilarity network. Execute the command line:
158   "TOM = TOMsimilarity(adjacency);
159   dissTOM = 1-TOM"
160   Click **Run**.
161
162 NOTE: The signed parameter was added to set the topology overlap network type.
163
164 3.3. Use the hclust function to select the average linkage hierarchical clustering method for
165 hierarchical clustering. Execute the command line:
166   "geneTree = hclust(as.dist(dissTOM), method = "average");"
167   Click **Run**.
168
169 3.4. Use the cutreeDynamic function to perform dynamic branch cutting and set the
170 minClusterSize parameter to 30. Obtain the module recognition result. Execute the command
171 line:
172   "dynamicMods = cutreeDynamic(dendro = geneTree, distM = dissTOM, deepSplit = 2,
173   pamRespectsDendro = FALSE, minClusterSize = minModuleSize);"
174   Click **Run**.
175
176 NOTE: The minimum module size could not be lower than 30.
177
178 3.5. Calculate the module eigen of each OTUs module by the moduleEigengenes function.
179 Execute the command line:
180   "MEList = moduleEigengenes(datExpr0, colors = dynamicColors)
181   MEs = MEList$eigengenes"
182   Click **Run**.
183
184 NOTE: The module eigen represented the overall OTU expression level in the module. It was
185 not a specific OTU, but the first principal component of each cluster obtained by singular
186 network value decomposition.
187

188 3.6. Perform the cluster function based on the correlation coefficient of module eigen. Use
189 the mergeCloseModules function to merge the modules with a value lower than 0.25. Execute
190 the command line:
191   "merge = mergeCloseModules(datExpr0, dynamicColors, cutHeight = MEDissThres,
192   verbose = 3)"
193   Click **Run**.
194
195 3.7. Finally, use the plotDendroAndColors function for visualization to obtain the module
196 assignment display diagram of each co-expression network (**Figure 2**). Use the table function
197 to extract the module attribution corresponding of each OTin the module assignment table.
198 Execute the command line:
199   "plotDendroAndColors(geneTree, mergedColors, "Merged dynamic",dendroLabels =
200   FALSE,
201   hang = 0.03,addGuide = TRUE, guideHang = 0.05,
202   main = "ES_Gene dendrogram and module colors")"
203   Click **Run**.
204
205 NOTE: In the module assignment diagram of the co-expressing network, different colors
206 represent different modules, and gray represents OTUs that cannot be classified into any
207 module. A greater number of OTUs in the gray module indicates that the early-stage
208 preprocessing quality of the expression matrix is poor.
209
210 **4.   Module comparison**
211
212 NOTE: This method can be used to compare the network modules of two ecological microbial
213 communities. In this article, compare the differences of microbial network modules between
214 endosphere and rhizoplane, endosphere and rhizosphere, rhizosphere and rhizoplane.
215
216 4.1. Preservation test
217
218 4.1.1.   Load the parameters and results of the two data sets saved in the previous steps.
219
220 4.1.2.   Set the network module assignment result of a group of microbial data as the
221 reference group, whereas the other group as the test group.
222
223 4.1.3.   Use the modulePreservation function to calculate the values of conservativeness
224 statistical parameters Z_summary and medianRank. Execute the command line:
225   "system.time({mp=modulePreservation(multiExpr,multiColor,referenceNetworks=1,
226   nPermutation=100, randomSeed=1,quickCor=0,verbose=3)})"
227   Click **Run**.
228
229 NOTE: This result can quantified the conservativeness between modules. Z_summary>10
230 indicates that two modules are highly preserved, whereas Z_summary<2 denotes non-
231 preserved modules. medianRank expresses the relative preservation of the module assessed
232 by ranking. Higher medianRank values denote non- preserved modules. (The key command
233 lines refer to the *Supplement S3*.)
234

235   4.1.4.   Use the plot function to visualize the results (**Figure 3**). Get the parameters
236   Z_summary and medianRank (**Table 1**).
237

238   NOTE: The network modules that satisfy both the Z_summary value less than 2 and the
239   median Rank value at the top, is the most highly non-preserved module in the two ecological
240   microbial communities.
241

242   4.1.5.   Based on the results of the aforementioned two statistical parameters to identify the
243   module with most highly non-preserved module of the two networks.
244

245   4.2. 2. Correlation analysis of the module membership
246

247   4.2.1.   Set the module assignment results of the two networks were set as the reference and
248   the test group, respectively.
249

250   NOTE: The settings need to be the same as Preservation test.
251

252   4.2.2.   Use the corPvalueStudent function to extract the kME (module membership) value
253   of each OTU in several candidate modules.
254   Execute the command line:
255   "Pvalue = as.data.frame(corPvalueStudent(as.matrix(ModuleMembership), Samples))"
256   Click **Run**.
257

258   NOTE: kME stands for the degree of module membership. ME stands for module eigen, which
259   represents the overall level of OTU expression in the module. kME is the correlation coefficient
260   between each OTU and the ME. Quantify the importance of OTU in the network by the kME
261   value of OTU. (The key command lines refer to the *Supplement S4*.)
262

263   4.2.3.   Then, use the verboseScatterplot function to calculate the correlation coefficient of
264   the kME value of the corresponding OTUs in the two networks and draw the correlation
265   analysis diagram (**Figure 4**).
266   Execute the command line:
267   "verboseScatterplot(abs(TModuleMembership[TmoduleGenes, Tcolumn]),
268   abs(NModuleMembership[NmoduleGenes, Ncolumn]),
269   xlab = paste("kME in", "ES"),
270   ylab = paste("kME in", "RP"),
271   main = paste("lightyellow"),
272   cex.main = 1.7, cex.lab = 1.6, cex.axis = 1.6, col = modulecolor)"
273   Click **Run**.
274

275   4.2.4.   Select the module with the smallest correlation coefficient of the kME value of the
276   OTU of the two networks. Consider this module to have the largest difference of the two
277   networks.
278

279   **5.   Analysis of the microbial differential network module**
280

281   5.1. Obtain data of the dominant bacteria phyla through statistical analysis of the OTU
282   sequence set of the module with the largest difference.
283
284   NOTE: The OTU sequence set of the module with the largest difference is summed by the
285   taxonomy of phyla. The dominant bacteria phyla accounted for more than 10%.
286
287   5.2. Then, use the exportNetworkToCytoscape function to obtain the file containing the
288   interaction relationship information of the OTU in the largest differential module.
289   Execute the command line:
290   "cyt = exportNetworkToCytoscape(modTOM,
291   edgeFile = paste("NEW-ES_CytoscapeInput-edges-", modules , ".txt",
292   sep=""),
293   nodeFile = paste("NEW-ES_CytoscapeInput-nodes-", modules, ".txt",
294   sep=""),
295   weighted = TRUE,threshold = 0.5, nodeNames = modProbes,
296   altNodeNames = modGenes, nodeAttr =
297   moduleColors[inModule])"
298   Click **Run**.
299
300   5.3. Import the file into Cytoscape. Set the threshold to 0.5 and adjust other parameters as
301   needed.
302
303   5.4. Construct a co-occurrence network of differential microorganisms (Figure 5).
304
305   5.5. Obtained the information of the core genus that has the most important regulatory role
306   in the network.
307
308   NOTE: According to the the kME value of OUT, the core genus can be defined.
309
310   5.6. Finally, the functions of the core genus were assessed and its influence on the entire
311   difference network was analyzed.
312
313   **Representative Results**
314   The representative results in this article were downloaded from the 2014 California Abaker
315   rice root microbiome data in the NCBI database (PRJNA386367)[9]. The data includes the
316   rhizosphere, rhizoplane, and endosphere microbiome samples from rice plants grown for 14
317   weeks in a submerged rice field. We used the WGCNA algorithm to select the power value
318   that satisfied the three networks that were close to the scale-free network (**Figure 1**) and
319   developed three co-expression networks (**Figure 2**). In the endosphere, rhizoplane, and
320   rhizosphere soil microbial co-expression network, 23, 22, and 21 modules were identified,
321   respectively. These results indicate that the numbers of microbial interaction networks in the
322   three niches were basically equal.
323
324   We further compared the differences in the microbial network modules within the
325   endosphere, rhizoplane, and rhizosphere soil. The following preservation test results of the
326   modules of the three niches groups were obtained. Three extremely non-preserved modules
327   existed between the rhizosphere soil and the rhizoplane (**Figure 3a, Table 1**). Additionally, nine

328   extremely non-preserved modules were present between the rhizosphere soil and
329   endosphere (**Figure 3b, Table 2**) and six extremely non-preserved modules between the
330   rhizoplane and endosphere (**Figure 3c, Table 3**). Furthermore, extremely non-preserved
331   modules among the three niches were found, indicating the presence of large differences in
332   the composition of microorganisms among the three niches. The results of the module
333   membership correlation analysis of the obtained non-conservative modules are illustrated in
334   **Figure 4**. From the figure, a significantly different module with the least correlation between
335   each two niches is visible among the three niches, which represents the most significant
336   difference between them.
337
338   In the rhizosphere-rhizoplane difference network (**Figure 5a**), the dominant phylum was
339   *Proteobacteria* (72.97%). In the rhizosphere-endosphere difference network (**Figure 5b**), the
340   dominant phyla were *Proteobacteria* (66.36%), *Actinobacteria* (10.1%), and *Bacteroidetes*
341   (10.9%). In the rhizoplane-endosphere difference network **(Figure 5c)**, the dominant phyla
342   were *Proteobacteria* (41.41%), *Bacteroidetes* (10.10%), *Firmicutes* (12.12%), and
343   *Verrucomicrobia.*
344
345   Three core genera (**Figure 5a**), including *Rhodobacter* and *Novosphingobium*, six core genera
346   (**Figure 5b**), including *Blvii28* and *Dechloromonas*, and five core genera (**Figure 5c**), including
347   *Cellvibrio* and *Geobacter*, exerted important regulatory functions in the three differential co-
348   occurrence networks. All core genera, except for *Dechloromonas*, had influence on only one
349   network, indicating the availability of considerable differences in the relative abundance of
350   microbial populations and species among the three niches of rice roots, which critically
351   affected the abundance and diversity of the existing root microbial communities.
352
353   The core genus *Azospirillum*, present in the rhizosphere-endosphere difference network of
354   rice, participated in nitrogen fixation and promoted plant growth[10]. Additionally, the
355   *Geobacter* genus, which was significantly enriched in the rhizoplane-endosphere difference
356   network, may be the main factor inducing the reduction of insoluble Fe and Mn oxides in many
357   soils and sediments[11]. These microorganisms interact with a range of microbial communities
358   in the root niches and actively participate in the regulation of microbial networks, which might
359   be critically important for the growth and development of rice roots.
360
361   **Figure 1. Evaluation of power β in the datasets.** (a) Evaluation of power β in the ES dataset;
362   (b) Evaluation of power β in the RS dataset; (c) Evaluation of power β in the RP dataset,
363   distribution of the scale-free index R2 (left), distribution of the mean connectivity (right) along
364   different soft power indices. The value of the best power was achieved when R2 tended to
365   saturation and was not lower than 0.8. The three co-expression networks had to be set to the
366   same power value to ensure their comparability. (RS: rhizosphere soil, RP: rhizoplane, and ES:
367   endosphere).
368
369   **Figure 2. OTU dendrogram obtained by average linkage hierarchical clustering. (**a) OTU
370   dendrogram from ES network; (b) OTU dendrogram from RS network; (c) OTU dendrogram
371   from RP network. The color row underneath the dendrogram indicates the module
372   assignment determined by the Dynamic Tree Cut algorithm. (RS: rhizosphere soil, RP:
373   rhizoplane, and ES: endosphere)
374

375 **Figure 3. Results of the preservation test.** (a) The analysis results are based on the RS co-
376 expression network module allocation as the reference group and the RP co-expression
377 network module allocation results as the test group; (b) The analysis results are based on the
378 ES co-expression network module allocation as the reference group and the RS co-expression
379 network module allocation results as the test group; (c) The analysis results are based on the
380 ES co-expression network module allocation as the reference group and the RP co-expression
381 network module allocation results as the test group. Z_summary > 10 indicates that two
382 modules are highly preserved, whereas Z_summary < 2 denotes non- preserved modules.
383 medianRank expresses the relative preservation of the module assessed by ranking. Higher
384 medianRank values denote non- preserved modules. (RS: rhizosphere soil, RP: rhizoplane, and
385 ES: endosphere)
386
387 **Figure 4. Correlation analysis of the module memberships.** (a) Module correlation of the kME
388 value between the RS and RP network; (b) Module correlation of the kME value between RS
389 and RP network; (c) Module correlation of the kME value between RP and ES network (RS:
390 rhizosphere soil, RP: rhizoplane, ES: endosphere, and KME: module membership).
391
392 **Figure 5. Co-occurrence network of differential microbial population in the root of rice. (**a)
393 Co-occurrence network of differential microbial populations in RS-RP; (b) Co-occurrence
394 network of differential microbial populations in ES-RS; (c) Co-occurrence network of
395 differential microbial populations in ES-RP. The analysis was performed using Cytoscape
396 software. Different colors represent different gates in the figure. (RS: Rhizosphere soil, RP:
397 rhizoplane, and ES: endosphere).
398
399 **Table 1. Result of Zsummary and medianRank between the rhizosphere soil and the**
400 **rhizoplane.**
401
402 **Table 2. Result of Zsummary and medianRank between the rhizosphere soil and endosphere.**
403
404 **Table 3. Result of Zsummary and medianRank between the rhizoplane and endosphere.**
405
406 **Discussion**
407 Correlation networks have been increasingly used in bioinformatics applications. WGCNA is a
408 systems biology method for descriptive analysis of the relationships between various
409 elements of a biological system[12]. R software package was used in earlier work on
410 WGCNA[13,14,15]. The package includes functions for network construction, module detection,
411 calculations of topological properties, data simulation, visualization, and capability for
412 interfacing with external software. WGCNA has been extensively employed to analyze gene
413 expression data from brain cancer[16], yeast cell cycle[17], mouse genetics[18,19], primate brain
414 tissue[20,21], diabetes[22], and plants[23]. Use the weighted gene correlation network analysis to
415 construct the network must include at least 8 sample. In this article, we focused on gene co-
416 expression networks that describe the interactions between microbial populations in different
417 environments. We obtained differential networks between microbial populations in divergent
418 environments and identified the key species in each network. The notion that key species are
419 important to community has been widely employed in food-web research[24]. Some of the
420 species in a complex microbial community can be essential to maintain the stability and
421 functionality of the community, such as *Bacteroides* in the intestinal flora[25]. The analysis of

microbial communities can be substantially simplified by targeting specific species of potential importance.

Our representative results highlight the differences in microbial communities, which can be identified using the above-described method. Here, microorganisms in different niches of the rice root system were subjected to WGCNA. The differential among the three niches were identified using conservative and module membership analyses. We determined the key species in the difference modules and obtained information on the differences in the composition of the microbial communities in the three niches. Meanwhile, the co-occurrence network revealed the presence of a significant interaction between the changing microorganisms in the roots of rice. Our findings provide direct evidence for the significance and feasibility of WGCNA in the evaluation of microbial community differences in various environments.

References
1.   Philippot, L., Raaijmakers, J. M., Lemanceau, P., van der Putten, W. H. Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology*. **11**, 789-799 (2013).
2.   Fierer, N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Review Microbiology*. **15** (10), 579-590 (2017).
3.   Jin J., Wang G. H., Liu X. B., Liu J. D., Chen X. L., Herbert S. J. Temporal and spatial dynamics of bacterial community in the rhizosphere of soybean genotypes grown in a black soil. *Pedosphere*. **19** (6), 808-816 (2009).
4.   Ma, B. et al. Genetic correlation network prediction of forest soil microbial functional organization. *ISME J*. **12** (10), 2492-2505 (2018).
5.   de Vries, F. T. et al. Soil bacterial networks are less stable under drought than fungal networks. *Nature Communications*. **9** (1), 3033 (2018).
6.   Colin, C. et al. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis.. *Carcinogenesis*. (10), 2300-2308 (2013).
7.   Ma, B., Zhao, K., Lv, X. et al. Genetic correlation network prediction of forest soil microbial functional organization. *ISME J*. **12**, 2492–2505 (2018).
8.   Zhang, B., Horvath, S. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*. **4** (1) (2005).

469   9.   Edwards, J. A. et al. Compositional shifts in root-associated bacterial and archaeal
470   microbiota track the plant life cycle in field-grown rice. *PLoS Biology*. **16**(2), e2003862 (2018).
471   10. Bashan, Y., De-Bashan, L. E. How the Plant Growth-Promoting Bacterium Azospirillum
472   Promotes Plant Growth—A Critical Assessment. *Advances in Agronomy*. **108**, 77-136 (2010).
473   11. Lovley, D. R. 1. et al. Geobacter: The Microbe Electric's Physiology, Ecology, and Practical
474   Applications. *Advances in Microbial Physiology*. **59**, 1-100 (2011).
475   12. Langfelder, P., Horvath, S. WGCNA: an R package for weighted correlation network
476   analysis. *BMC Bioinformatics*. **9**, 559 (2008).
477   13. Zhang, B., Horvath, S. A General Framework for Weighted Gene Co-expression Network
478   Analysis. *Statistical Applications in Genetics and Molecular Biology*. **4** (1), 17 (2005).
479   14. Horvath, S., Dong, J. Geometric interpretation of Gene Co-expression Network Analysis.
480   *PLoS Computational Biology*. **4** (8), e1000117 (2008).
481   15. Langfelder, P., &Horvath, S. Eigengene networks for studying the relationships between
482   co-expression modules. *BMC Systems Biology*. **1**, 54 (2007).
483   16. Horvath, S. et al. Analysis of Oncogenic Signaling Networks in Glioblastoma Identifies
484   ASPM as a Novel Molecular Target. *Proceedings of the National Academy of Sciences of the*
485   *United States of America* .**103**(46), 17402-17407 (2006).
486   17. Carlson, M. R. et al. Gene Connectivity, Function, and Sequence Conservation: Predictions
487   from Modular Yeast Co-expression Networks. *BMC Genomics*. **7** (1), 40 (2006).
488   18. Fuller, T. et al. Weighted Gene Co-expression Network Analysis Strategies Applied to
489   Mouse Weight. *Mammalian Genome*. **6** (18), 463-472 (2007).
490   19. Yin L , Wang Y , Lin Y , et al. Explorative analysis of the gene expression profile during liver
491   regeneration of mouse: a microarray-based study[J]. Artifi*cial Cells Nanomedicine &*
492   *Biotechnology*. **47** (1), 1113-1121 (2019).
493   20. Oldham, M., Horvath, S., Geschwind, D. Conservation and Evolution of Gene Co-
494   expression Networks in Human and Chimpanzee Brains. *Proceedings of the National Academy*
495   *of Sciences of the United States of America*. **103** (47), 17973-17978 (2006).
496   21. Oldham, M. C. et al. Functional organization of the transcriptome in human brain. *Nature*
497   *Neuroscience*. **11** (11), 1271-1282 (2008).
498   22. Keller, M. P. et al. A gene expression network model of type 2 diabetes links cell cycle
499   regulation in islets with diabetes susceptibility. *Genome Research*. **18** (5), 706-716 (2008).
500   23. Weston, D., Gunter, L., Rogers, A., Wullschleger, S. Connecting genes, coexpression
501   modules, and molecular signatures to environmental stress phenotypes in plants. *BMC*
502   *Systems Biology*. **2** (1), 16 (2008).
503   24. Jordaˊn, F. Keystone species and food webs. *Biological Sciences*. **364**, 1733–1741 (2009).
504   25. Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson D.A., Gordon J.I. Host-Bacterial
505   Mutualism in the Human Intestine. *Science*. **307**, 1915–1920 (2009).
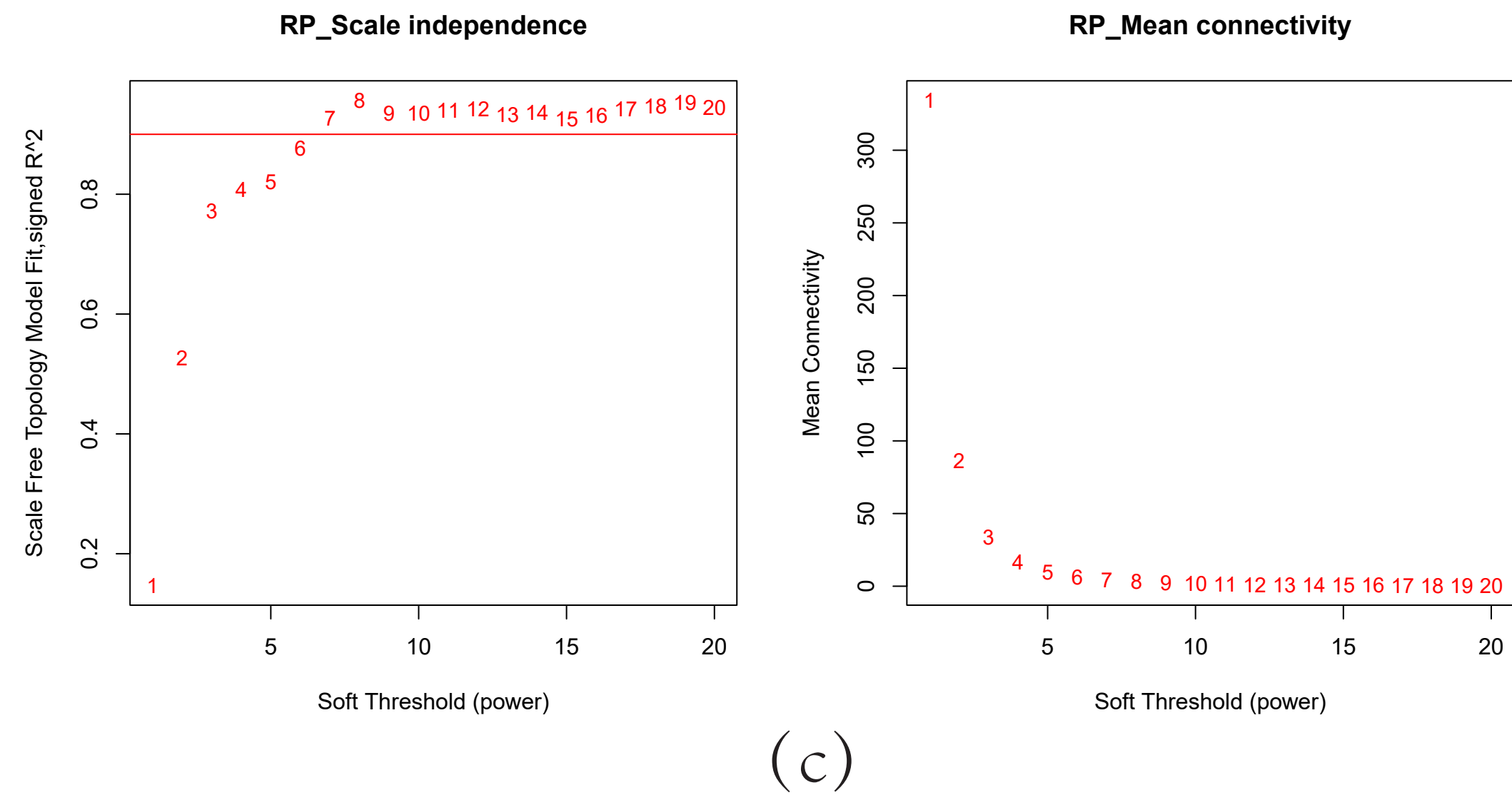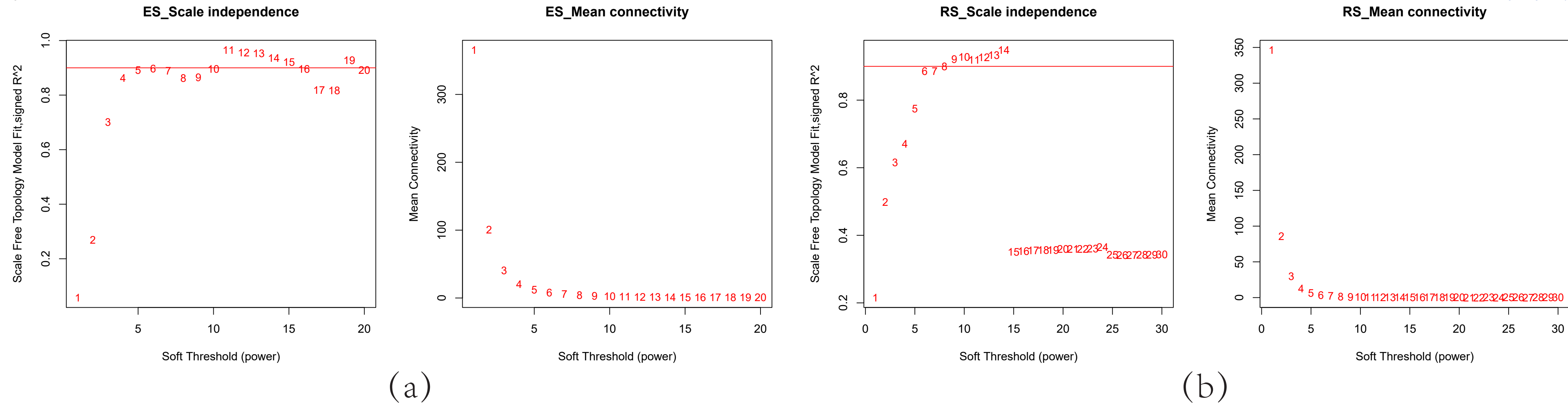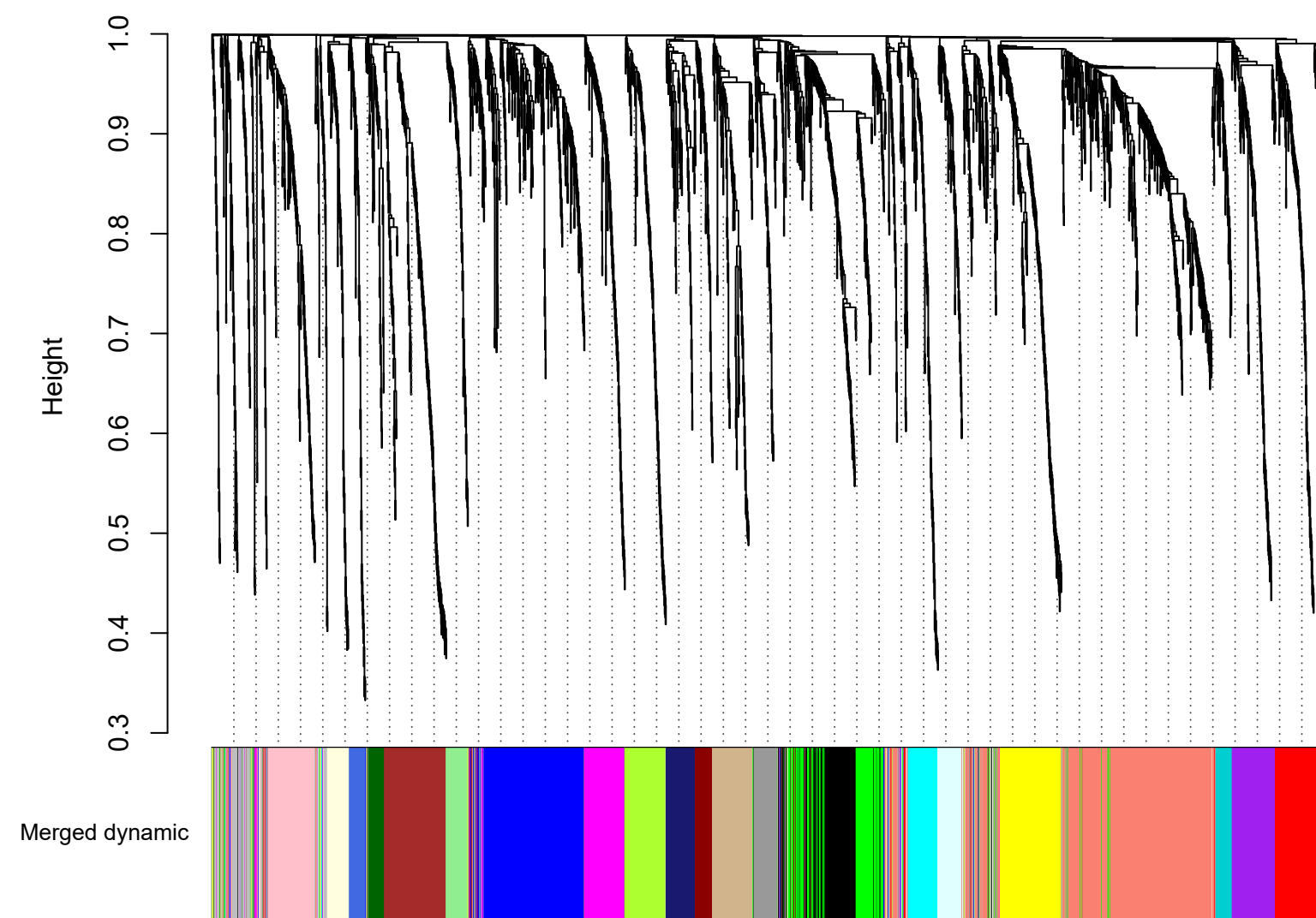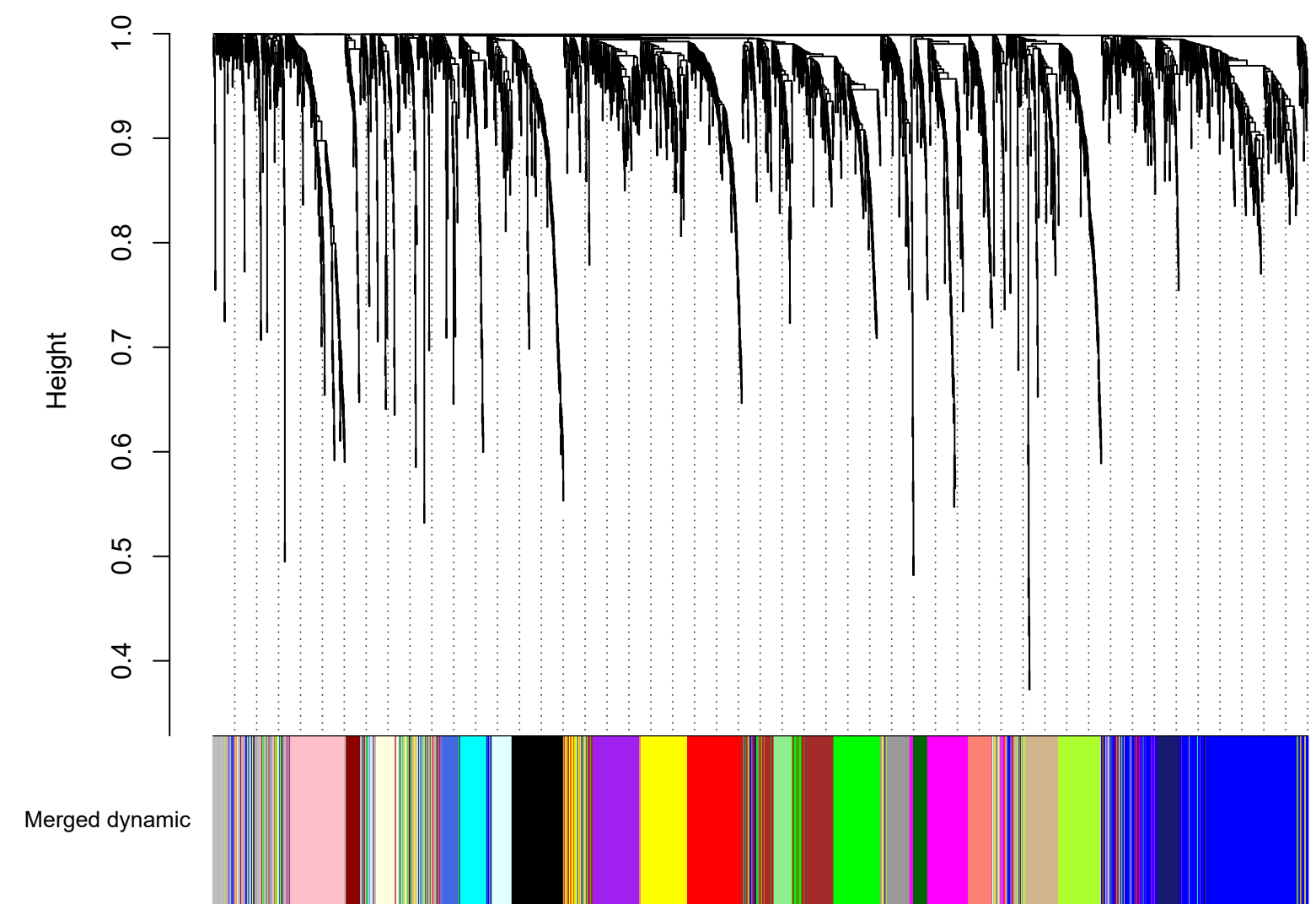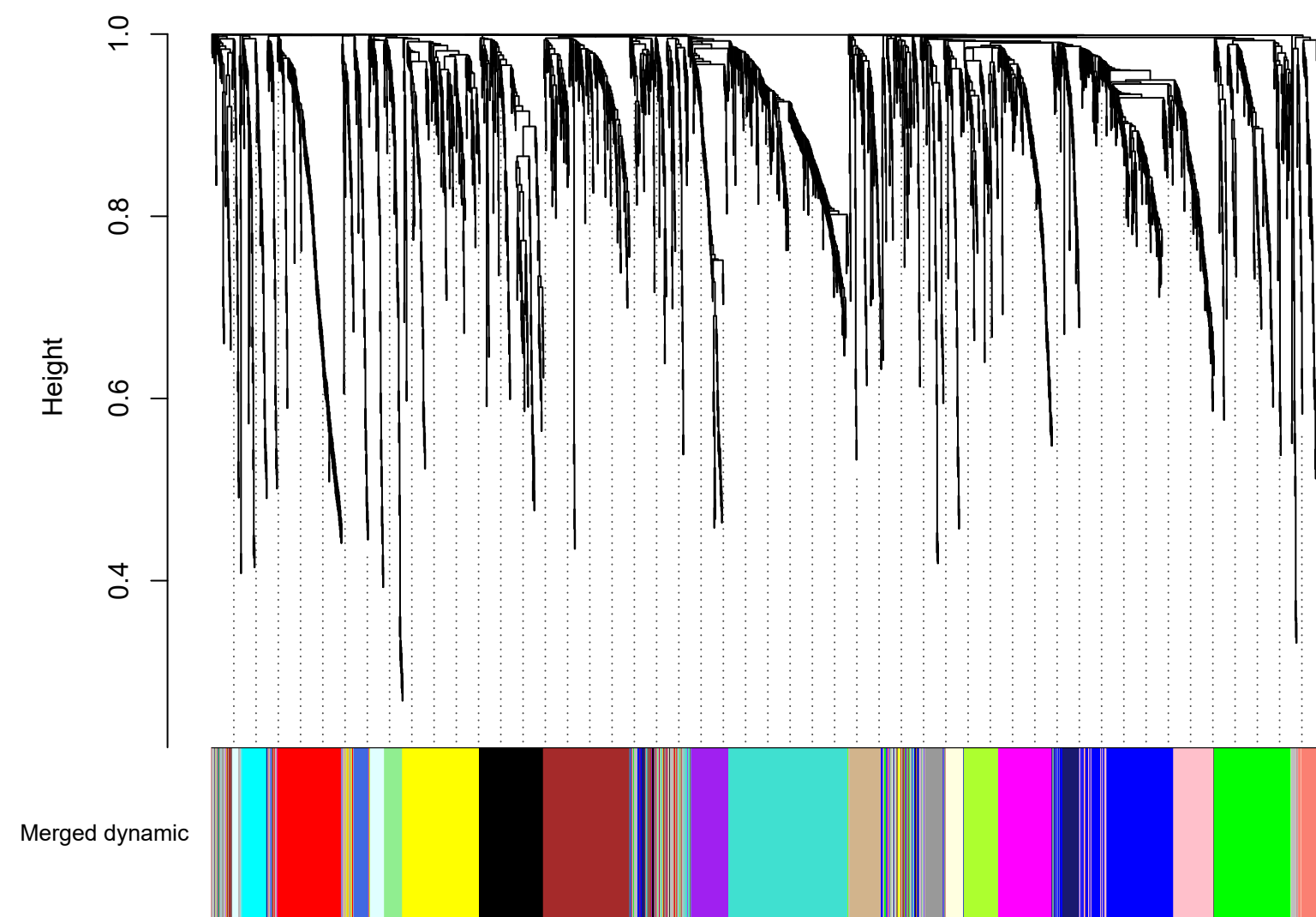
Figure 1

(a)

(b)

(c)

Figure 2

(a)

(b)

(c)

Figure 3

(a)

(b)

(c)

Figure 4

(a)

(b)
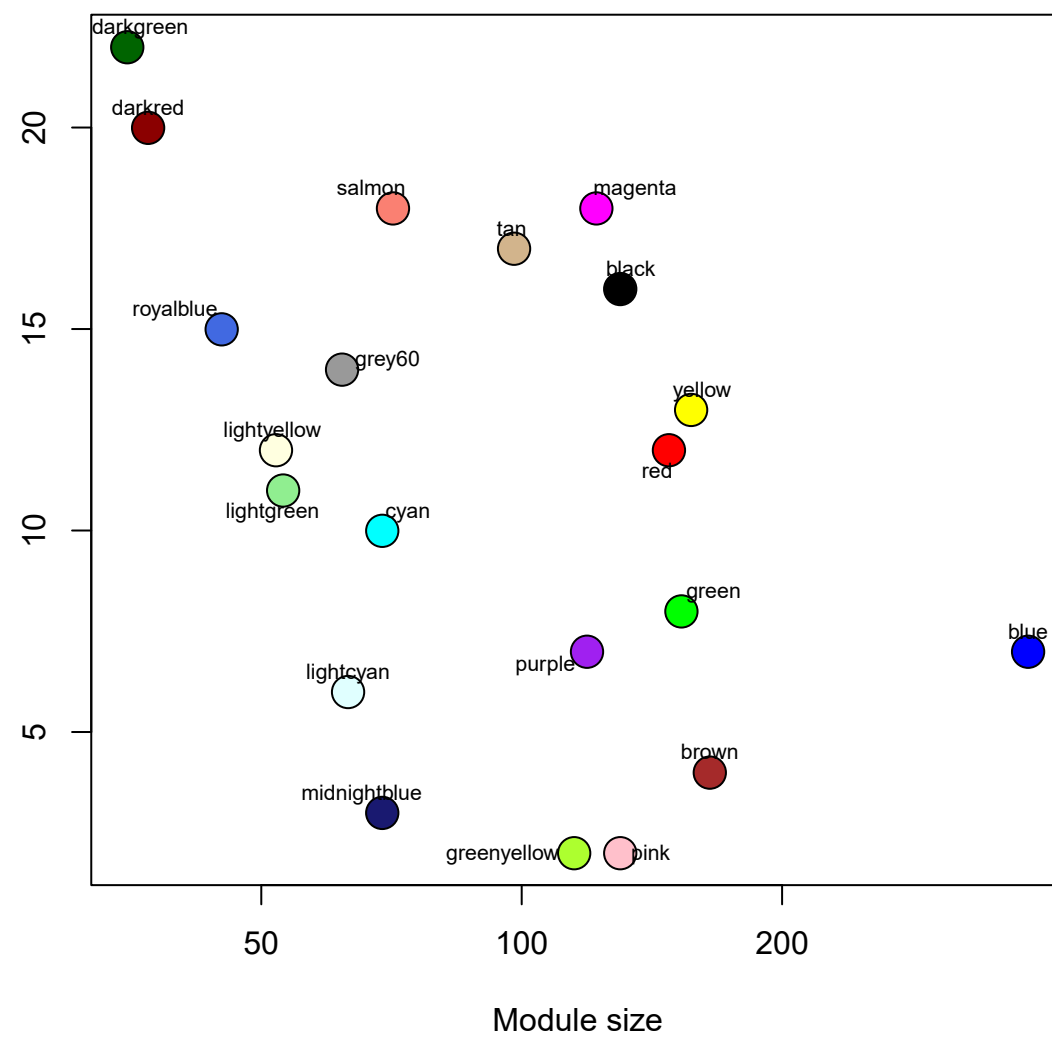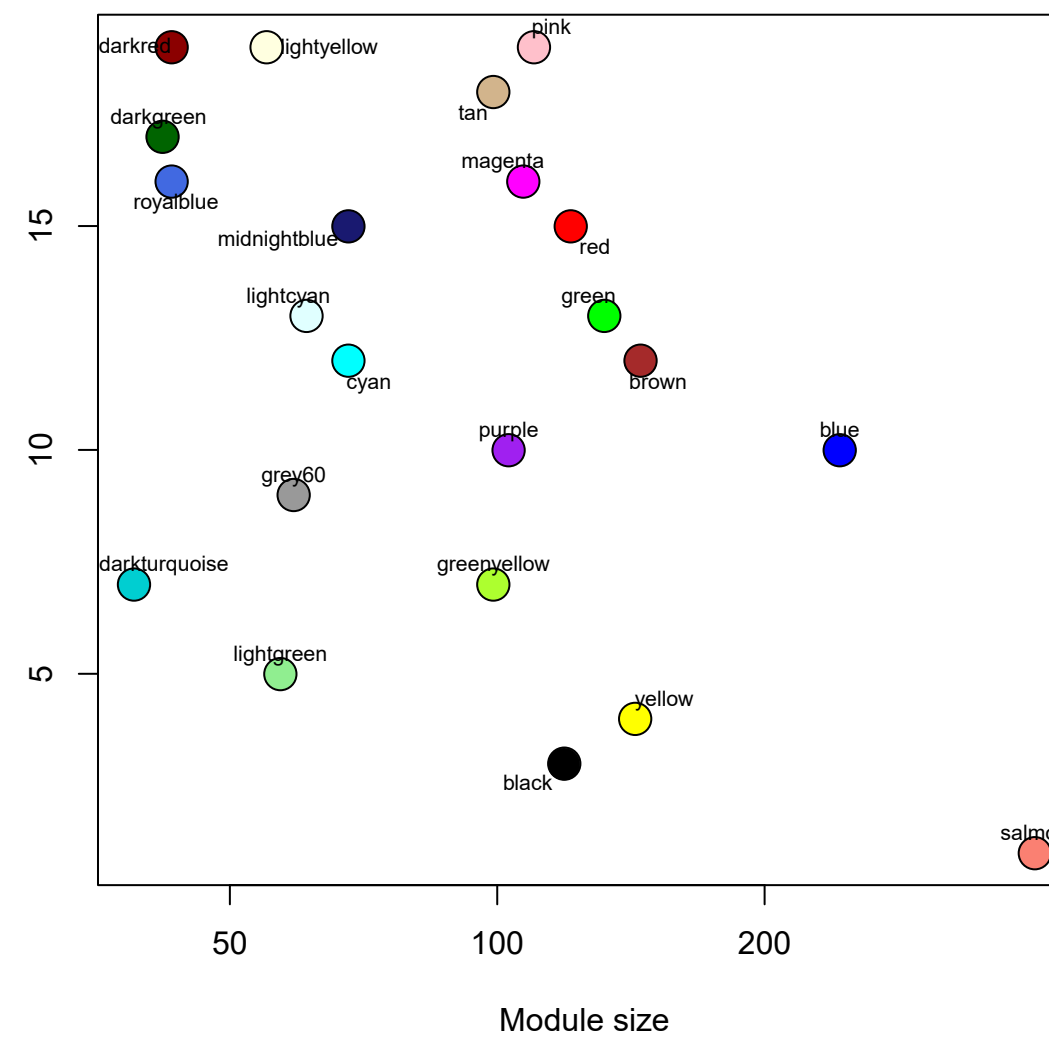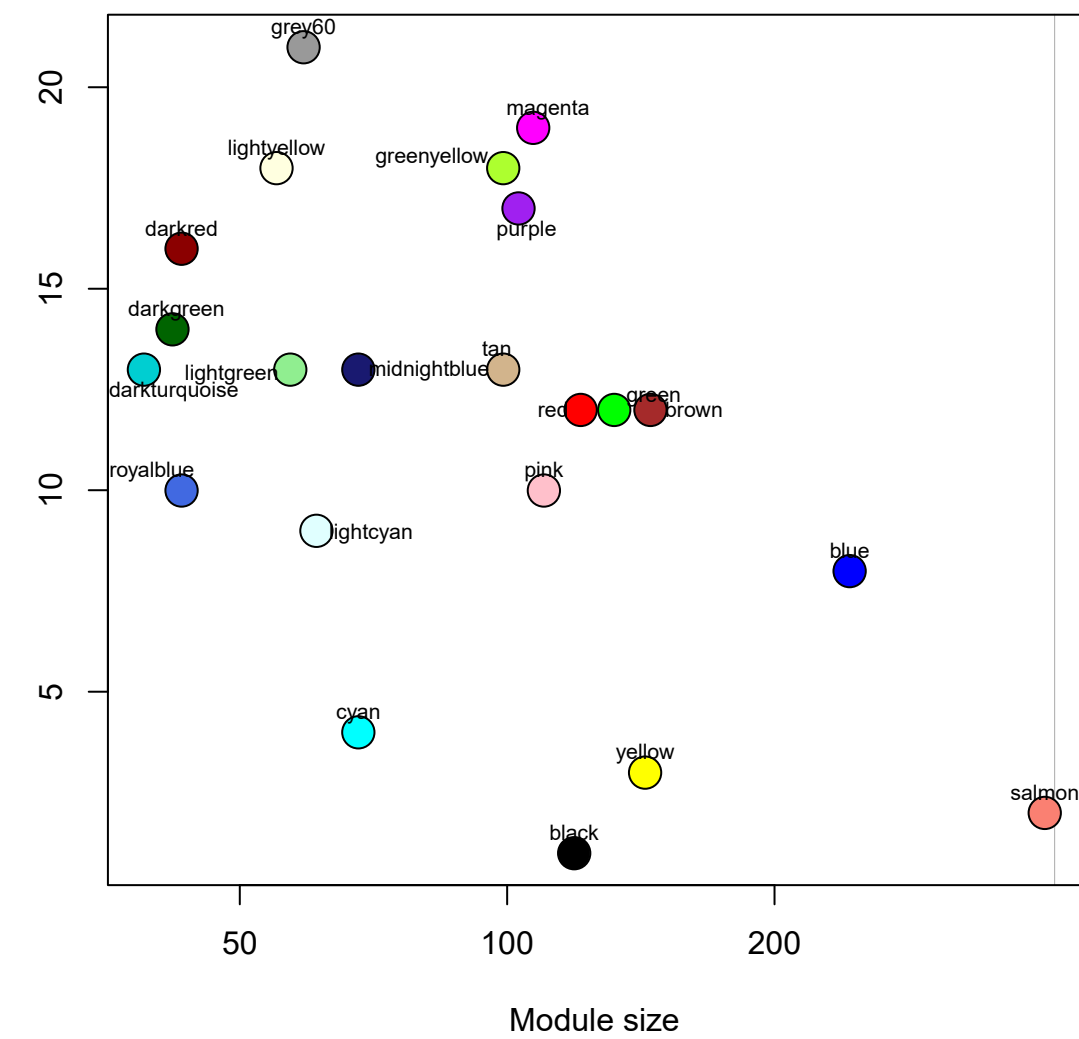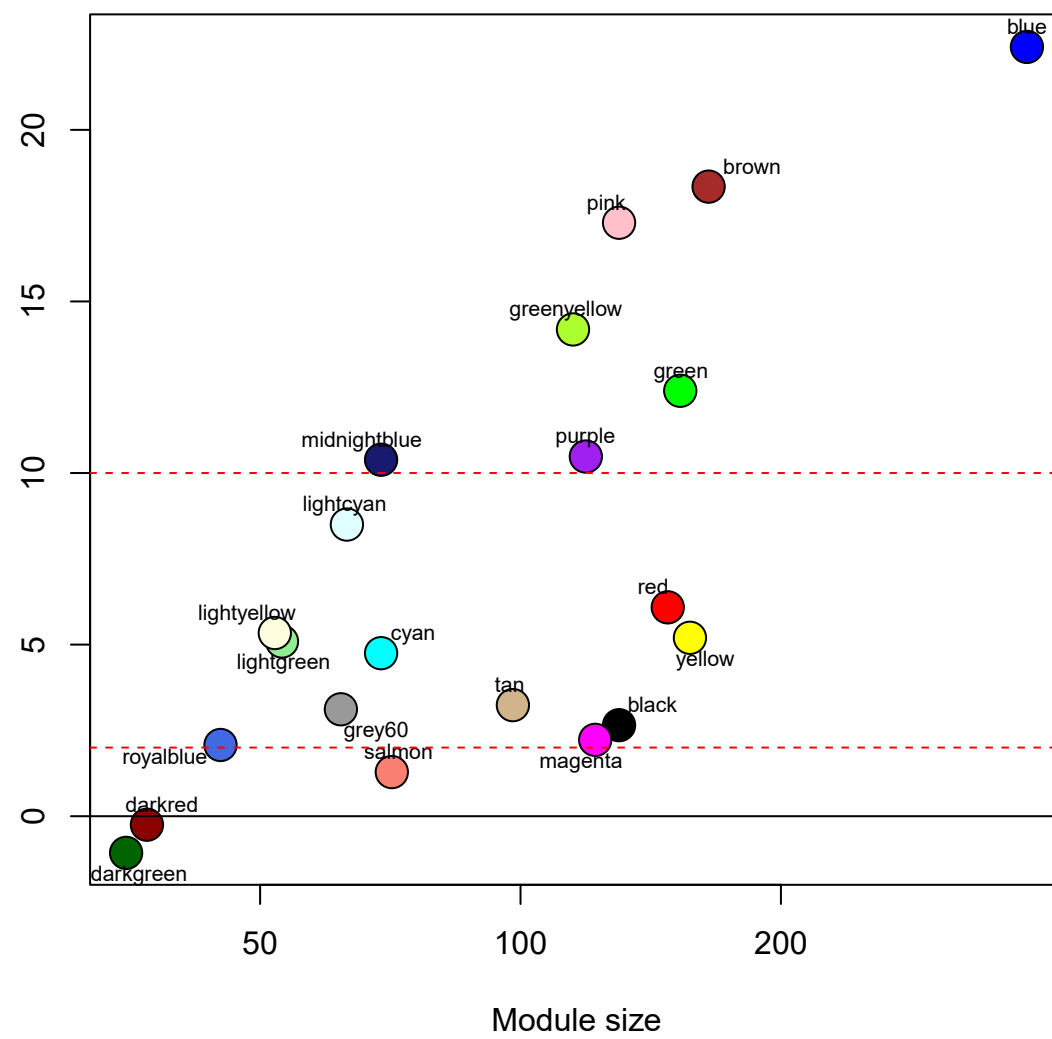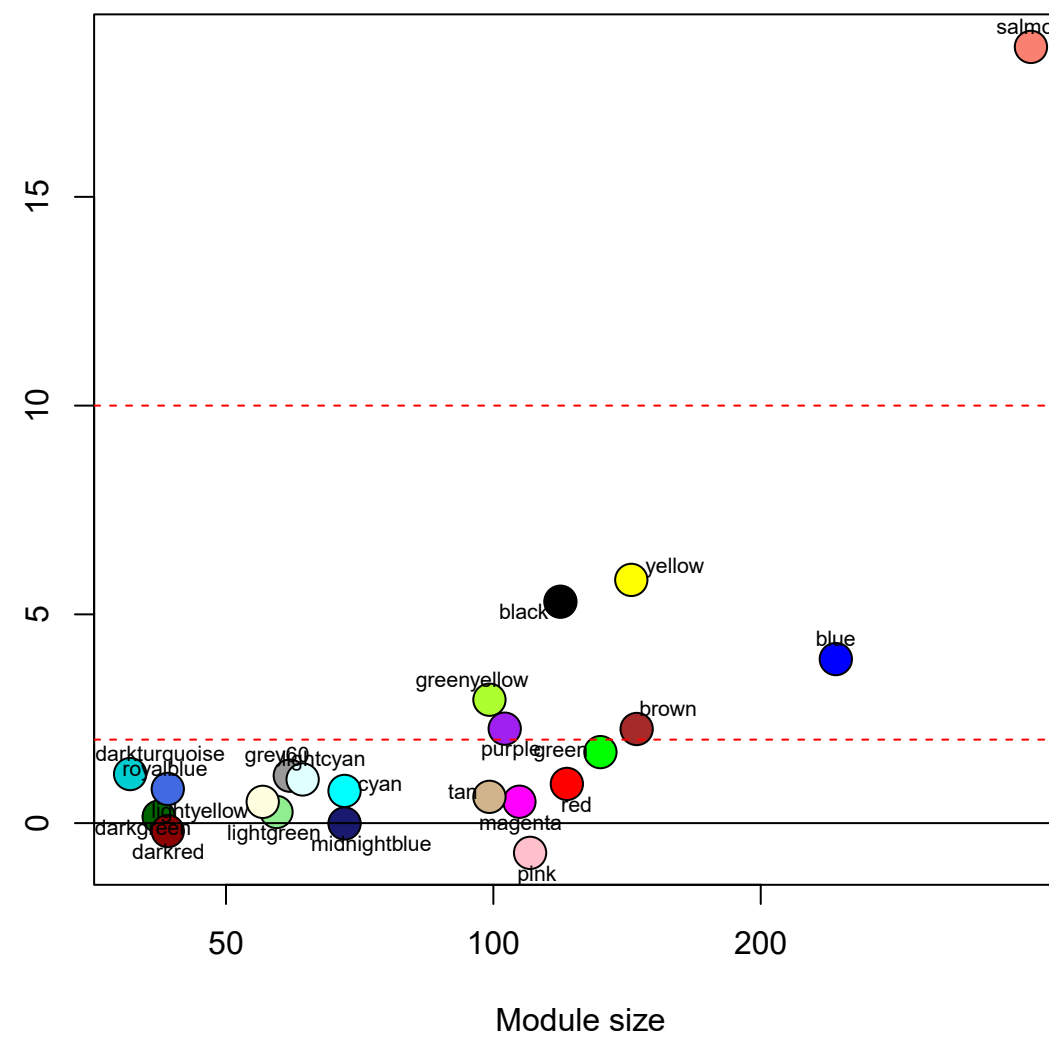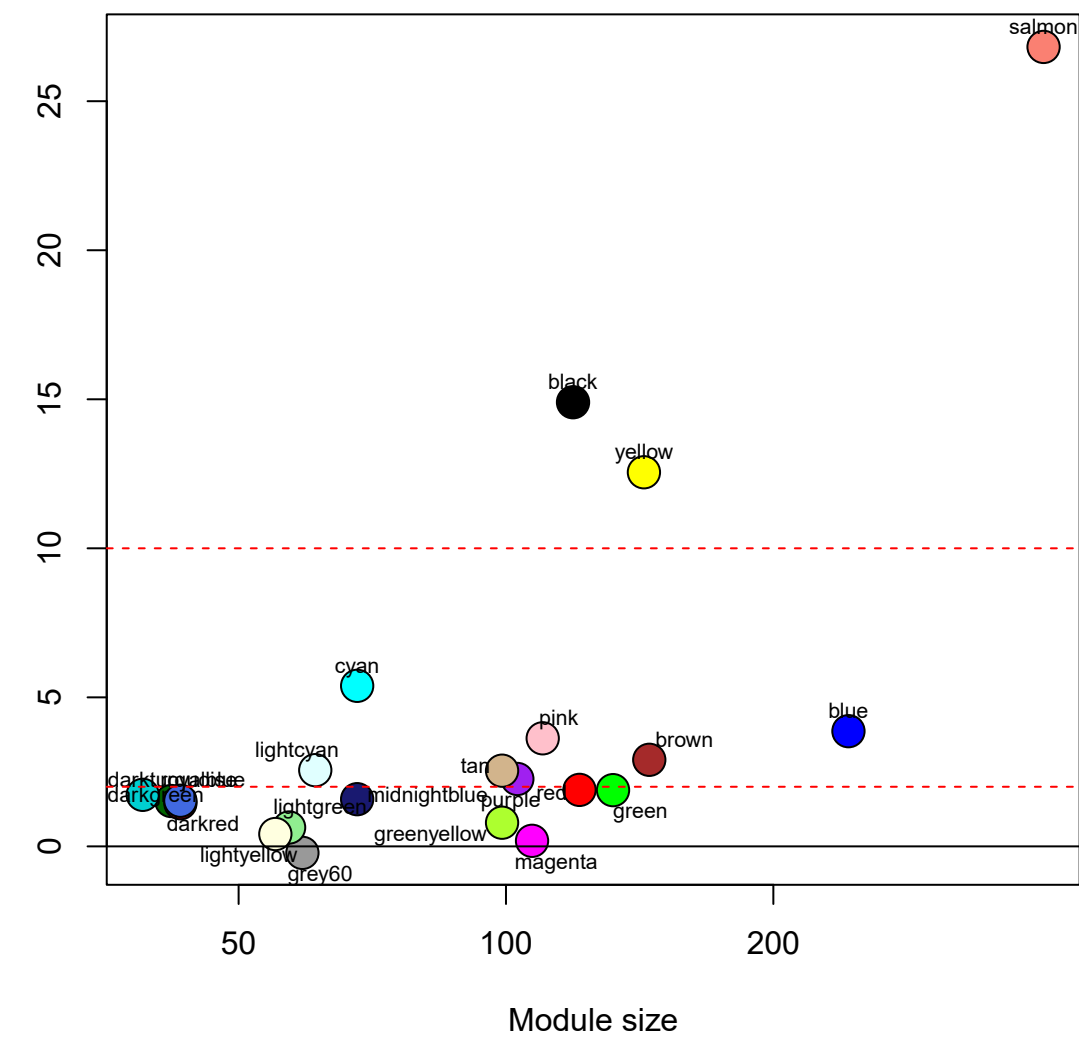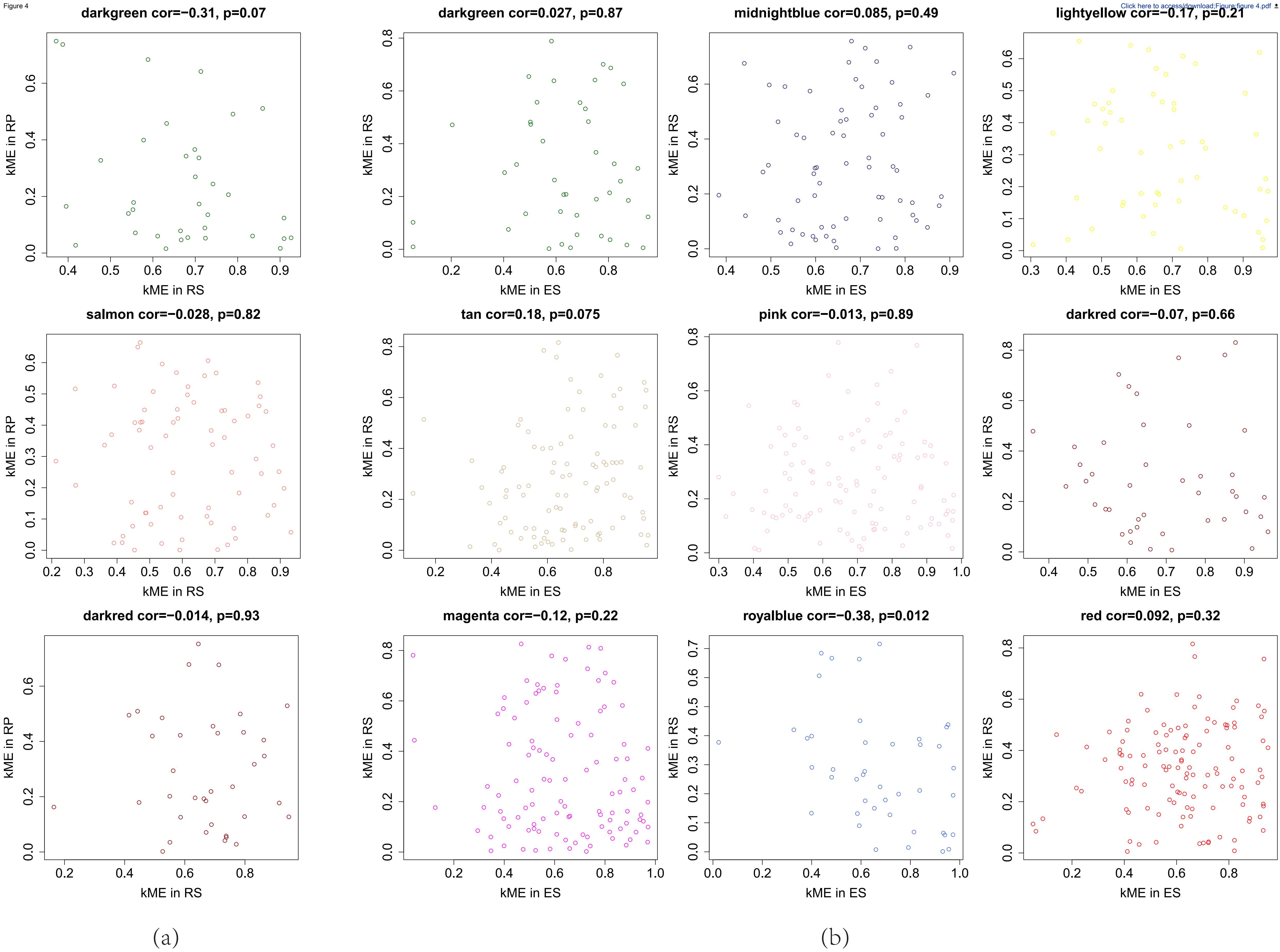
(c)

(a)

(b)

(c)

phylum

- Acidobacteria
- Actinobacteria
- Bacteroldetes
- Chloroflexi
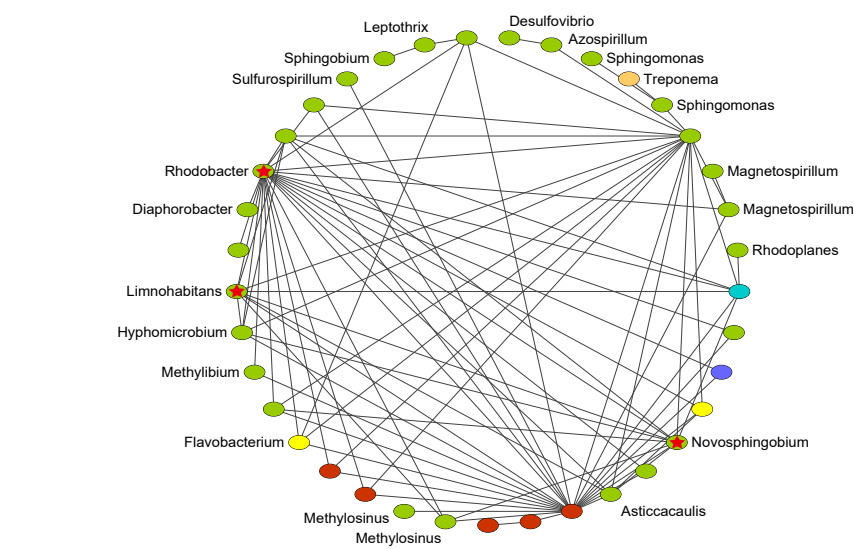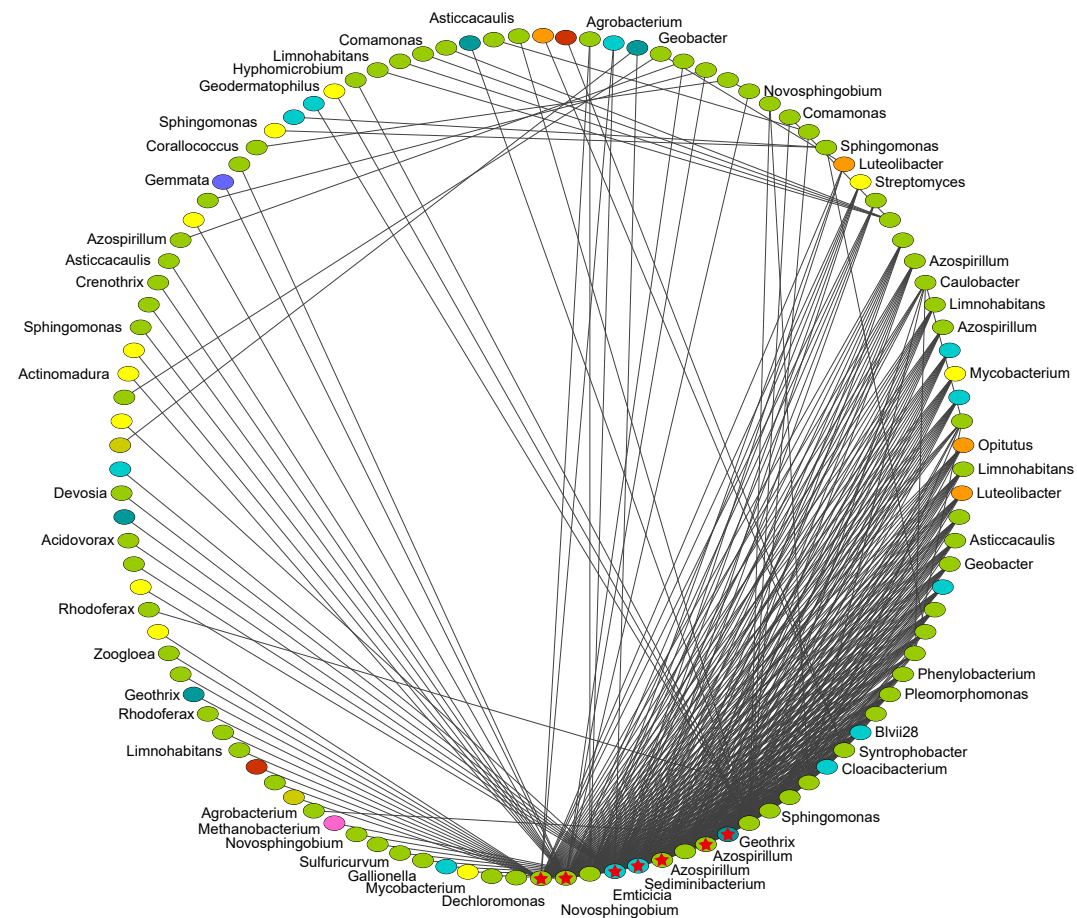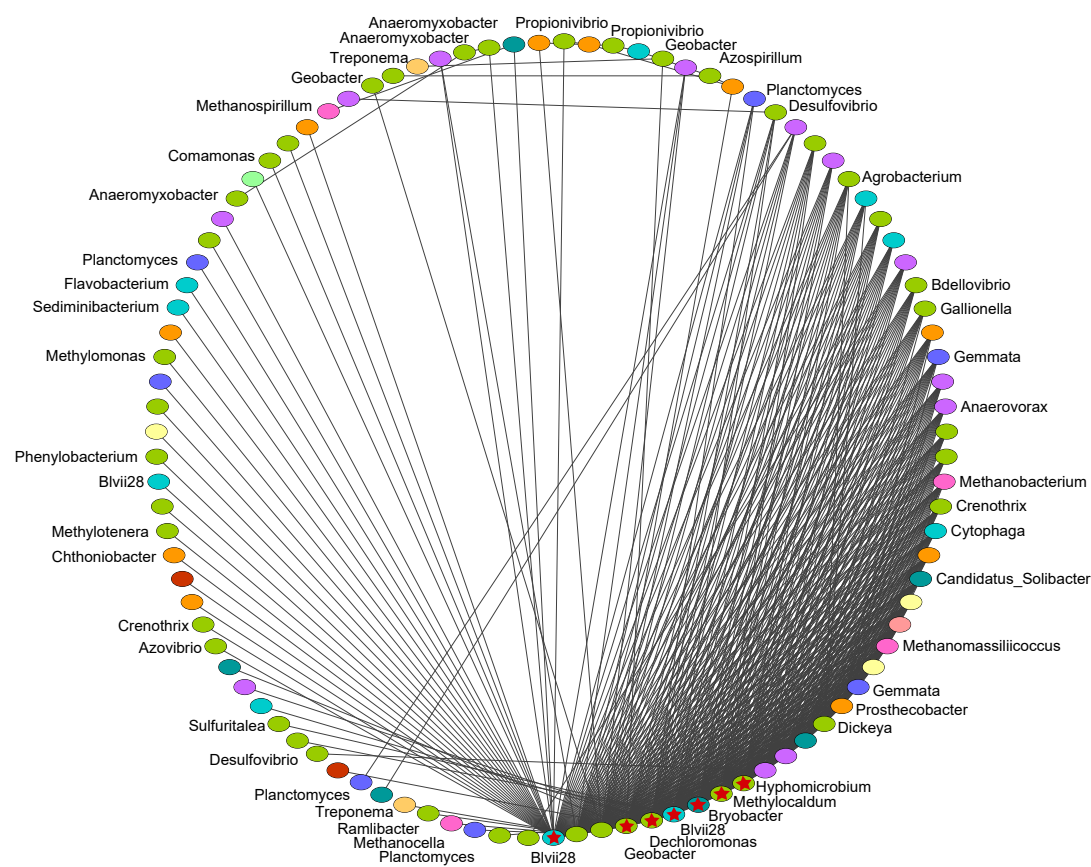- Euryarchaetes
- Firmicutes
- Spirobacteria
- Proteobacteria
- Planctomycetes
- Verrucomicrobia

Table 1

Table 1 Result of Zsummary and medianRank between the rhizosphere soil and the rhizoplane

module	medianRar	Zsummary	module	medianRar	Zsummary
greenyellow	2	14	yellow	13	5.2
pink	2	17	grey60	14	3.1
midnightblue	3	10	royalblue	15	2.1
brown	4	18	black	16	2.7
lightcyan	6	8.5	tan	17	3.2
purple	7	10	**salmon**	**18**	**1.3**
blue	7	22	magenta	18	2.2
green	8	12	**darkred**	**20**	**-0.24**
cyan	10	4.8	gold	20	14
lightgreen	11	5.1	**darkgreen**	**22**	**-1.1**
lightyellow	12	5.3	grey	22	0.21
red	12	6.1			

Table 2   Click here to access/download;Table;revised table 2.xlsx ⬇

Table 2 Result of Zsummary and medianRank between the rhizosphere soil and endosphere

module	medianRank	Zsummary	module	medianRank	Zsummary
salmon	1	19	lightcyan	13	1.1
black	3	5.3	**red**	**15**	**0.95**
yellow	4	5.8	**midnightblue**	**15**	**-0.0016**
lightgreen	5	0.27	**royalblue**	**16**	**0.83**
greenyellow	7	3	**magenta**	**16**	**0.52**
darkturquoise	7	1.2	**darkgreen**	**17**	**0.16**
grey60	9	1.1	**tan**	**18**	**0.64**
blue	10	3.9	**lightyellow**	**19**	**0.52**
purple	10	2.3	**darkred**	**19**	**-0.18**
brown	12	2.3	**pink**	**19**	**-0.71**
cyan	12	0.78	gold	21	11
green	13	1.7			

Table 3

Table 3 Result of Zsummary and medianRank between the rhizoplane and endosphere

module	medianRank	Zsummary	module	medianRank	Zsummary
black	1	15	darkturquoise	13	1.7
salmon	2	27	midnightblue	13	1.6
yellow	3	13	lightgreen	13	0.64
cyan	4	5.4	**darkgreen**	**14**	**1.5**
blue	8	3.9	darkred	16	1.5
lightcyan	9	2.6	purple	17	2.3
pink	10	3.6	**greenyellow**	**18**	**0.8**
royalblue	10	1.5	**lightyellow**	**18**	**0.42**
brown	12	2.9	**magenta**	**19**	**0.2**
green	12	1.9	gold	21	18
red	**12**	**1.9**	**grey60**	**21**	**-0.21**
tan	13	2.5			

Click here to access/download
Table of Materials
revised Table_of_Materials.xlsx

Dear Dr. Nam Nguyen and and reviewers,

Thank you for your and the reviewers' comments on our manuscript entitled "Divergence of root microbiota in different habitats based on weighted correlation networks". These comments are very helpful for us in improving our paper. According to your suggestions, we have made amendments in the revised manuscript carefully. Here, we respond to the reviewers' comments point-by-point as follows.

- 3:36-3:47: Please increase the narration to the proper volume (-12 to -6 dB).

Response：Thanks for the suggestions. We adjusted the narration in the place to the proper volume.

Supplemental Coding Files

Click here to access/download
Supplemental Coding Files
Supplement S1.txt

Click here to access/download
Supplemental Coding Files
Supplement S2.txt

Click here to access/download
Supplemental Coding Files
Supplement S3.txt

Click here to access/download

Supplemental Coding Files

Supplement S4.txt