体现组间差异OTU/模块的微生物网络图

李雨泽

2020-08-20

### 252体现组间差异OTU/模块的微生物网络图

本节作者：李雨泽 西北农林科技大学

版本1.0.2，更新日期：2020年8月20日

本项目永久地址： <https://github.com/YongxinLiu/MicrobiomeStatPlot> ，本节目录 252NetworkModule，包含R markdown(\*.Rmd)、Word(\*.docx)文档、测试数据和结果图表，欢迎广大同行帮忙审核校对、并提修改意见。提交反馈的三种方式：1. 公众号文章下方留言；2. 下载Word文档使用审阅模式修改和批注后，发送至微信(meta-genomics)或邮件([metagenome@126.com](mailto:metagenome@126.com))；3. 在Github中的Rmd文档直接修改并提交Issue。审稿人请在创作者登记表 <https://www.kdocs.cn/l/c7CGfv9Xc> 中记录个人信息、时间和贡献，以免专著发表时遗漏。

#### 网络分析的概念、组成及特征

基于图论，网络分析将系统中的个体看作节点，利用节点间两两相关性形成的边构建起整体网络，以呈现系统的整体、个别拓扑性质。作为微生物生态学数据分析的一个重要手段受到了广泛关注，近年来从社会学、信息科学等学科向生态学、医学扩展。一般地，网络由节点、边作为基本的组成成分，点可以根据映射的属性设置颜色、大小、形状，边也如此；在微生物生态学分析中，网络节点一般为OTU/其它分类水平中的成员，着色则是门、纲水平的分类学信息、或反映当前处理/环境下的微生物群落属性。然而，部分学科，尤其是农业相关的学科往往具有较多处理，双因素试验较常见，使用常规的网络分析表现力有限，因此，从节点丰度在处理间/组间的差异入手，将来自不同处理/环境的OTU融入同一个网络中进行分析是一项具有新意的微生物组数据分析可视化方法；此外，受到传统农学“3重复”思维的影响，试验计划开始时设计重复数不足，导致较少的样本数往往不足以进行稳定可信的网络分析，将多个样本合并可以在很大程度上解决这一问题，为数据的进一步挖掘分析提供更多选择。

在这类网络分析中，微生物共现网络（co-occurrence network）或被按照所有模块着色，寻找这些模块间的丰度差异；或被按照差异OTU着色，寻找这些OTU所在模块的丰度差异。最终的组间差异模块还可以和测定的其它土壤、植物性状进行相关性/回归分析，进一步体现微生物网络分析的实际意义。

#### 体现组间差异OTU/模块的网络图实例

##### 例1.不同处理间差异OTU及模块在网络中展示

此文章是瑞士苏黎世农业研究中心Klaus Schlaeppi团队于2018年发表于Microbiome (IF = 11.607)的文章，以基于成员丰度比较的网络分析揭示了不同的智慧农业管理方式（有机农业，保护性耕作）对土壤、根系微生物网络的影响，尤其是加深了对种植方式敏感微生物的理解，为开发智慧农业的微生物组管理提供了基础。本文是一篇制作精良非常值得学习的文章，发表两年被引过百次。这里以原文中的图4a，b为例。

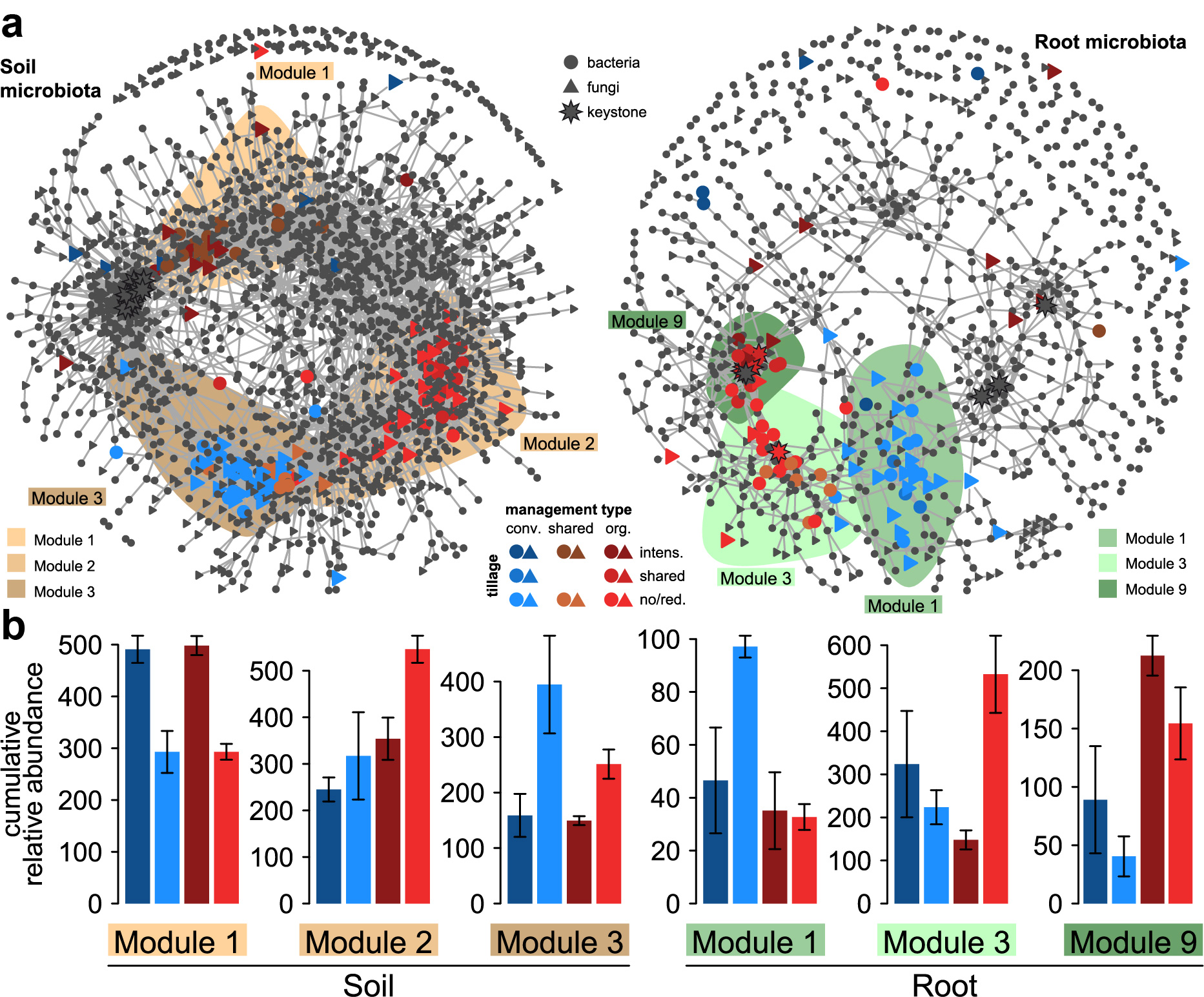


图4. 种植方式敏感OTUs的共丰度网络。a, 土壤（左图）、根系（右图）中的细菌和真菌共现性网络。b, 土壤、根系微生物**网络中包含有组间差异OTU的主要模块的相对丰度**。

Fig. 4 Co-occurrence patterns of cropping sensitive OTUs. a, Co-occurrence networks visualizing significant correlations (ρ > 0.7, p < 0.001; indicated with gray lines) between bacteria and fungi OTUs in soil and root communities. b, Cumulative relative abundance (as counts per million, CPM; y-axis in ×1000) of all bacteria and fungi of the cropping sensitive modules in soil and root networks.

* 结果

我们探索了土壤和根系细菌和真菌的亚共存模式中种植方式敏感OTUs(csOTU)的分布模式（图4a，表1）。微生物群落成员的互作网络分析表明，微生物成员的丰度模式响应种植方式（处理）。

We explored the distribution patterns of csOTUs in meta co-occurrence patterns of bacteria and fungi in soil and root communities (Fig. 4a, Table 1). We found that the abundance patterns of inter-kingdom microbial associations also responded to cropping practices.

在土壤微生物网络中，耕作对群落的影响很明显，模块M1明显在网络中独立聚类存在，其中包含集约耕作的差异OTU，与其他两个模块（M2和M3）分开；M2和M3主要包含O-RT（有机管理少耕耕作）和C-NT（常规管理免耕耕作）处理下的差异OTU（图4a，b）。农业管理也是根微生物网络的主要驱动力（图2），属于有机管理的大量差异OTU位于模块M3和M9中，并与包含常规管理差异OTU的模块M1分开（图4b）。

The effect of tillage intensity in the soil communities was apparent with a discrete module (M1) in the soil network, containing csOTUs specific to intensive tillage practices. M1 was separated from two other modules (M2 and M3) that primarily contained csOTUs specific to the O-RT and C-NT cropping systems (Fig. 4a, b). Similarly, management type presented the main driver in root communities (Fig. 2), and the numerous csOTUs assigned to organic management were predominantly located in modules M3 and M9 and separated from module M1 containing primarily conventional management specific OTUs (Fig. 4b).

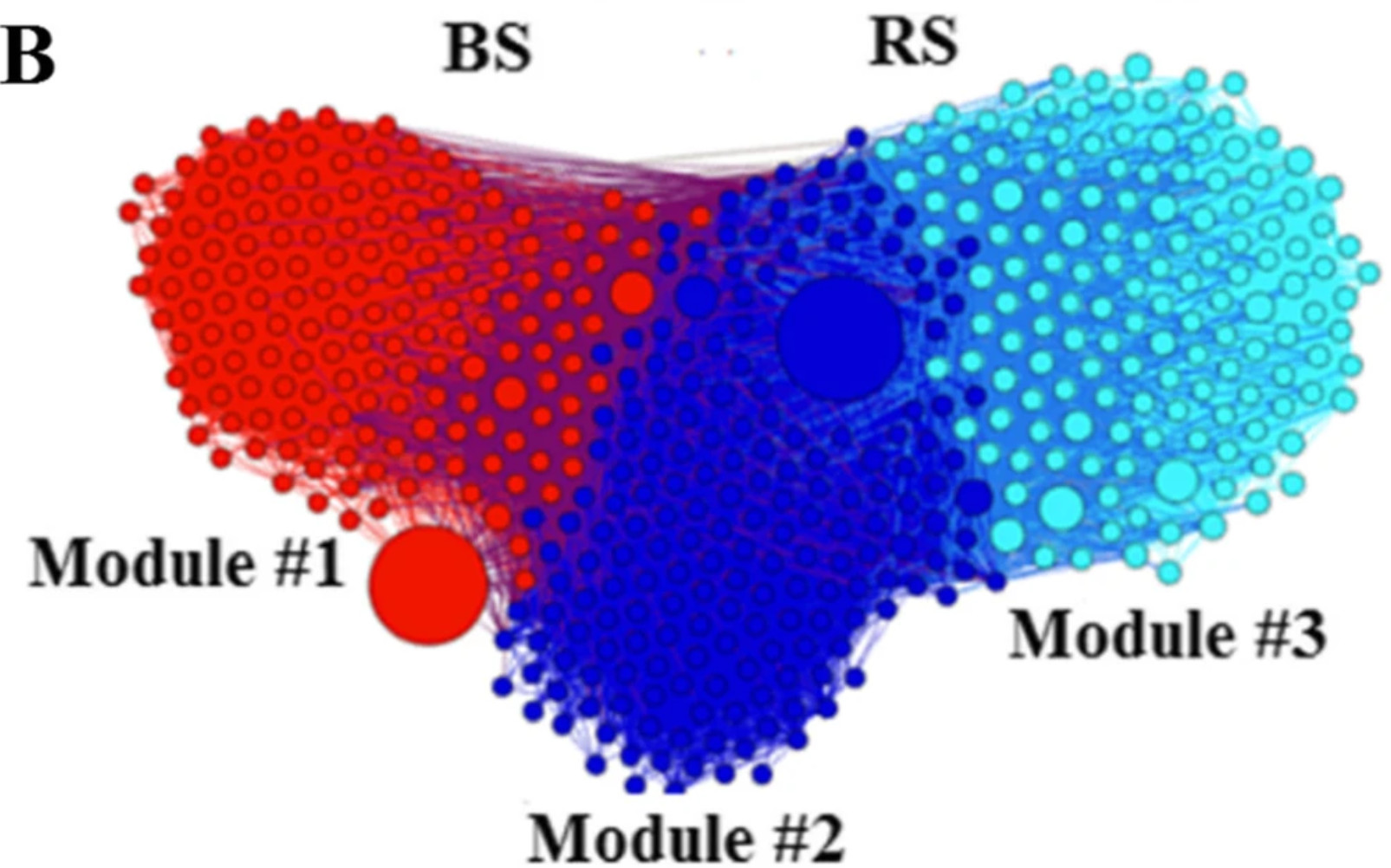
* 讨论

在土壤和根微生物网络中都观察到了包含大量组件差异OTU及其所在的模块，这些模块对不同的耕作方式有相似的响应（图4，附加文件1：图S14）。差异OTU被聚集在不同的模块中，这些模块的丰度差异反映在不同的种植系统下。

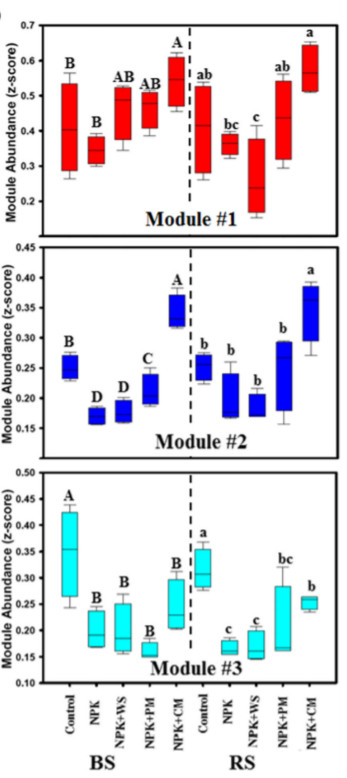
In both soil and root meta-networks, we identified modules containing high proportions of OTUs responding similarly to different cropping practices (Fig. 4, Additional file 1: Fig. S14). We observed that csOTUs grouped in distinct modules that reflected the different cropping systems.

##### 例2.不同处理和生态位间OTU网络的模块

此文章是中国科学院南京土壤研究所褚海燕团队于2019年发表于Microbiome (IF = 11.607)的文章，该文章发现长期的施肥抑制土壤固氮及固氮菌群。其中将所有样本（不同生态位，不同处理）纳入一个网络中进行模块划分及丰度比较，这里以原文中的图1b，d，图3为例。详细解读见：[Microbiome：40年施肥处理后固氮菌及氮固定受抑制](https://mp.weixin.qq.com/s/jQ21NFLVPLha6JdmzbGSZg) [新闻稿:南土所褚海燕组揭示长期施肥抑制根际微生物固氮的作用机制](https://mp.weixin.qq.com/s/XcUp6Nt14ELES9cpgir-yg)。

 图1b. 来自所有的样本进行网络分析，其节点根据三个主要生态簇进行着色（模块1-3）。

1. Network diagram with nodes colored according to each of the three main ecological clusters (modules #1–3).

 图1d. 不同施肥方式（处理）下、不同生态位间的模块相对丰度的比较。不同字母代表不同组间存在显著差异。

Relative abundance of the ecological clusters in different fertilization treatments. Different letters indicate the values that differ significantly among treatments at *P* < 0.05 (Duncan’s test) in bulk soil (A, B, C) and rhizosphere soil (a, b, c).

* 结果

我们建立了一个包含所有固氮菌的相关网络，并发现了3个固氮菌生态簇（模块）彼此强烈共生（模块＃1，＃2和＃3;图1b）。

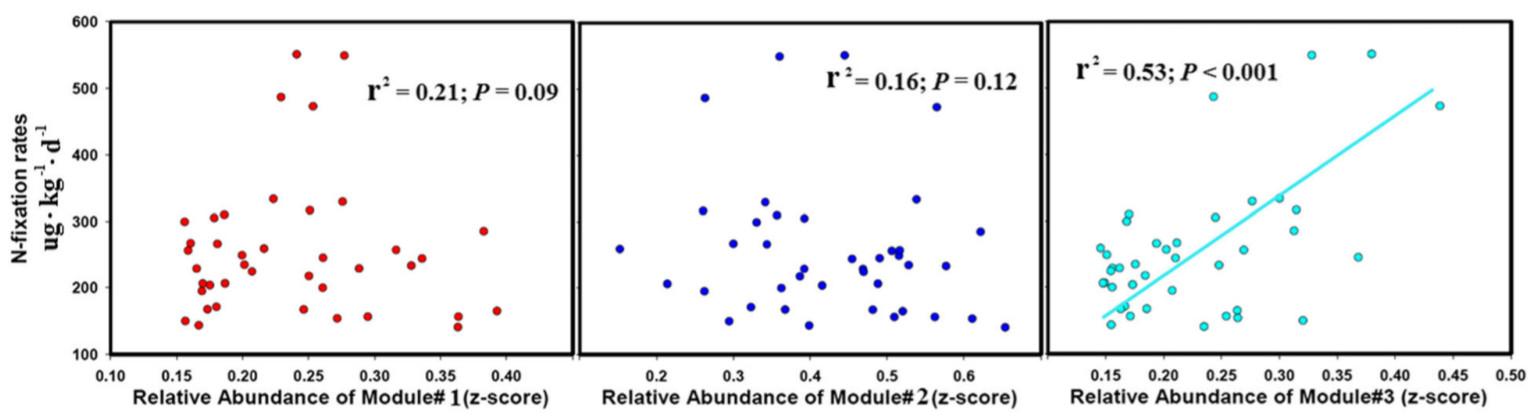
We built a correlation network incorporating the detected dominant diazotrophic phylotypes and found three ecological clusters of N fixers strongly co-occurring with each other (modules #1, #2, and #3; Fig. 1b).

长期施肥导致生态簇（模块）相对丰度发生剧烈变化，模块3的相对丰度大大降低，模块1和2的相对丰度有所提高（图1d）。

Long-term fertilization resulted in drastic changes in the relative abundance of ecological clusters; the relative abundance of module #3 was strongly reduced and that of modules #1 and #2 was somewhat increased—particularly under NPK + CM additions (Fig. 1d).

##### 例3. 观测指标与组间差异模块丰度进行联合分析

同样是例2中的文献，为了理解固氮菌的共丰度网络中模块丰度及固氮速率的关系，将二者进行了回归分析。



image

图3. 不同的固氮菌模块与固氮速率的线性回归关系。

Regressions between the nitrogen fixation rates and the relative abundance of the main diazotrophic ecological clusters. From left to right are modules #1, #2, and #3, represented by red, blue, and cyan plots, respectively）

* 结果

模块3的相对丰度变化与固氮速率之间存在显著正相关关系（图3）；模块1，模块2相对丰度和固氮速率之间无显著关联（图3）。

A strong and significant positive association between the relative abundance of module #3 and N fixation rates was found (Fig. 3). However, no significant association was detected between the relative abundance of module #1, module #2, and N fixation rates (Fig. 3).

* 讨论

物种（及所属模块）相对丰度的降低也可能对固氮速率产生负面影响。事实上，通过降低模块3中固氮菌的相对丰度，发现长期施肥与固氮间接地负相关。这表明在长期施肥下，模块3中的固氮菌受抑制，对固氮速率产生负面影响。

Reductions in the relative abundance of these taxa could also negatively influence N fixation rates. In fact, long-term fertilizations were found to be indirectly negatively associated with N fixation, by decreasing the relative abundance of N fixers within module #3. This indicates that taxa within module #3 will be suppressed under long-term fertilization with negative consequences for N fixation rates.

#### 组间差异OTU/模块网络图的

测试数据及代码

模拟的测试数据设定处理为三种耕作措施CPT，PT，ZT下的作物根际土壤，每个处理三次重复，于T，J，F三个时期采样，这里利用以下分析流程将存在显著耕作间差异的OTU及其在网络中所在模块凸显出来，并绘制模块相对丰度比较图，最后与植株性状进行关联分析。示例仅选取T时期下的样本进行操作。

##### 包和数据的准备

#####所需的包  
# Bioconductor包安装器  
if (!requireNamespace("BiocManager", quietly=TRUE))  
 install.packages("BiocManager")  
library(BiocManager)  
# 检查bioconductor来源包是否需要安装  
if (!requireNamespace("edgeR", quietly=TRUE))  
 BiocManager::install("edgeR")  
  
library(edgeR)#用于基于edgeR的组间差异OTU识别  
library(indicspecies)#用于基于indicator species的组间差异OTU识别  
library(igraph)#用于共现性网络的绘制  
library(Hmisc)#进行共现性网络构建前的OTU两两相关性计算  
library(sciplot)#用于基于模块丰度的绘图  
library(reshape2)#用于长宽数据转换  
library(ggpmisc)#用于其它观测指标与组间差异模块丰度关系分析  
  
#####输入OTU表格数据 #####  
otu\_its <- read.table("d1\_otu\_table.txt",row.names=1,sep="\t",header=T, blank.lines.skip=F,check.names=F)  
otu\_its <- as.matrix(otu\_its)  
  
#####输入OTU物种分类数据 #####  
tax\_its <- read.table("d2\_otu\_taxa.txt",row.names=1, sep="\t", header=F,stringsAsFactors=F,quote="")  
colnames(tax\_its) <- c("Kingdom","Phylum","Class","Order", "Family", "Genus", "Species")  
tax\_its[tax\_its==""] <- "unassigned"  
  
#####输入metadata数据 #####  
design\_its <- read.table("d3\_metadata.txt", header=T, row.names=1, na.strings="NA")  
design\_its$tillage <- factor(design\_its $tillage,c("CPT","PT","ZT"))  
design\_its$stage <- factor(design\_its $stage ,c("T","J","F"))  
design\_its$tilstag <- factor(design\_its $tilstag,c("CPTT","CPTJ","CPTF","PTT","PTJ", "PTF","ZTT","ZTJ","ZTF"))  
  
#####输入其它性状（plant\_trait）数据 #####  
plant\_trait <- read.table("d4\_plant\_trait.txt", header=T, row.names=1, na.strings="NA")

##### 数据筛选

由于示例涉及较多水平的样本，因此计划采用合并耕作处理，将3个生育时期分别展示的策略，这里以T时期下的网络为例

#####以其中的T时期为例进行network的绘制#####  
tsamples <- rownames(design\_its)[which(design\_its$stage == "T")]  
otu\_its\_t <- otu\_its[,tsamples]  
#去除一些低丰度OTU，在至少i个样本中具有至少j个序列  
#otu\_its\_t <- otu\_its\_t[which(rowSums(otu\_its\_t >= j) >= i),]  
design\_its\_t <- droplevels(design\_its[tsamples,])  
tax\_t\_its <- tax\_its[rownames(otu\_its\_t),]

##### 差异分析

在绘制网络之前需要找出具有组间差异的OTU，往往取2-3种差异判别方法的交集，如indicator species分析、edgeR分析、limma分析

#####寻找基于indicator species的组间显著差异OTU并保存结果#####  
edgeR\_its\_t <- DGEList(counts=otu\_its\_t,   
 group=design\_its\_t$tillage,  
 genes=tax\_t\_its)  
#CPM标准化  
otu\_norm\_its\_t <- cpm(edgeR\_its\_t, normalized.lib.sizes=T, log=F)  
# 准备数据：T时期下组间的indicator species  
indic\_t\_its <- as.data.frame(t(otu\_norm\_its\_t))  
indic\_t\_groups\_its <- design\_its\_t$tillage  
# 设置随机数种子，保证置换检验可重复  
set.seed(8046)  
# 鉴定各组指示种  
indicatorsp\_t\_its <- multipatt(indic\_t\_its,indic\_t\_groups\_its,func = "r.g",control=how(nperm=1000))  
indic\_t\_df\_its <- indicatorsp\_t\_its$sign  
  
#按照阈值p.value < 0.05筛选各组显著的指示OTU  
CPT\_t\_its <- as.matrix(indic\_t\_df\_its[which(indic\_t\_df\_its$s.CPT == 1 & indic\_t\_df\_its$p.value < 0.05),])  
PT\_t\_its <- as.matrix(indic\_t\_df\_its[which(indic\_t\_df\_its$s.PT == 1 & indic\_t\_df\_its$p.value < 0.05),])  
ZT\_t\_its <- as.matrix(indic\_t\_df\_its[which(indic\_t\_df\_its$s.ZT == 1 & indic\_t\_df\_its$p.value < 0.05),])  
#合并  
t\_r\_values\_its <- rbind(CPT\_t\_its,PT\_t\_its,ZT\_t\_its)  
# 组名修正，删除多余的"s."  
# colnames(t\_r\_values\_its)[1:3] <- c("CPT","PT","ZT")  
colnames(t\_r\_values\_its)[1:3] <- gsub("s.","",colnames(t\_r\_values\_its)[1:3])  
  
#####寻找基于edgeR的组间显著差异OTU#####  
model\_matt\_its <- model.matrix(~tillage, data=design\_its\_t)  
edgeR\_its\_t\_tillage <- DGEList(counts=otu\_its\_t, group=design\_its\_t$tillage, genes=tax\_t\_its)  
edgeR\_its\_t\_tillage <- calcNormFactors(edgeR\_its\_t\_tillage)  
dge\_ttillage\_its <- estimateGLMRobustDisp(edgeR\_its\_t\_tillage, design=model\_matt\_its)  
fit\_ttillage\_its <- glmFit(dge\_ttillage\_its, design=model\_matt\_its)  
# 2，3组分别与1组比较  
lrt\_ttillage\_its <- glmLRT(fit\_ttillage\_its, coef=2:3)  
tillage\_t\_its <- topTags(lrt\_ttillage\_its, n=Inf, p.value=0.05)  
tillage\_t\_its <- tillage\_t\_its$table  
  
##### (可选)limma分析，取indicator species、edgeR、limma的交集，但是这样可能太严格导致OTU极少  
#limma\_voom\_t\_its <- voom(edgeR\_its\_t\_tillage, model\_matt\_its)  
#fit\_t\_its <- lmFit(limma\_voom\_t\_its, model\_matt\_its)  
#fit\_t\_its <- eBayes(fit\_t\_its)  
#limma\_t\_its<-topTable(fit, coef = 2:3)  
#indic\_edge\_its\_t <- Reduce(intersect,list(rownames(t\_r\_values\_its),rownames(tillage\_t\_its),rownames(limma\_t\_its)))

##### 网络分析

进行上述分析即能够将具有组间差异的OTU“标记”，接下来就进行共现网络(co-occurrence network)的绘制及在网络图上标记这些OTU和所在的主要模块

######绘制能够体现组件丰度差异OTU/模块的共现性网络图co-occurence network #####  
#取基于indicator species分析和edgeR分析得到的显著组间差异OTU的交集  
indic\_edge\_its\_t <- intersect(rownames(t\_r\_values\_its),rownames(tillage\_t\_its))  
#基于TMM标准化的OTU表格进行OTU的两两Spearman相关计算  
t\_its\_otu\_cor <- rcorr(t(otu\_norm\_its\_t),type=c("spearman"))  
## 邻接矩阵转化为边表  
CorrDF <- function(cormat, pmat) {  
 ut <- upper.tri(cormat)  
 data.frame(  
 from = rownames(cormat)[col(cormat)[ut]],  
 to = rownames(cormat)[row(cormat)[ut]],  
 cor =(cormat)[ut],  
 p = pmat[ut]  
 )  
}  
t\_its\_cor\_df <- CorrDF(t\_its\_otu\_cor$r,t\_its\_otu\_cor$P)  
# p值校正  
t\_its\_cor\_df$padj <- p.adjust(t\_its\_cor\_df$p, method = "none") #method可选c("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none")  
#取Spearman's rho > 0.7且p-value < 0.001的关系作为入选共现性网络co-occurence network的边  
t\_its\_cor\_df\_padj <- t\_its\_cor\_df[which(t\_its\_cor\_df$cor > 0.7),]  
t\_its\_cor\_df\_padj <- t\_its\_cor\_df\_padj[which(t\_its\_cor\_df\_padj$padj < 0.001),]  
  
#生成node属性表  
# 边两列合并为结点  
nodeattrib\_t\_its <- data.frame(node = union(t\_its\_cor\_df\_padj$from,t\_its\_cor\_df\_padj$to))  
# 显著的添加标签  
nodeattrib\_t\_its$indicgroup <- 0  
for (i in as.character(nodeattrib\_t\_its$node))  
{  
 if (i %in% indic\_edge\_its\_t == TRUE)  
 {nodeattrib\_t\_its[nodeattrib\_t\_its$node==i,"indicgroup"] <- paste(colnames(t\_r\_values\_its)[which(t\_r\_values\_its[i,1:3]==1)],collapse = "\_")}  
 else  
 { nodeattrib\_t\_its[nodeattrib\_t\_its$node==i,"indicgroup"] <- "NA"}  
}  
#将OTU，即节点分类信息添加到node属性表  
nodeattrib\_t\_its <- cbind(nodeattrib\_t\_its,tax\_its[as.character(nodeattrib\_t\_its$node),])  
#用igraph绘制共现性网络图co-occurence network  
t\_net\_its <- graph\_from\_data\_frame(t\_its\_cor\_df\_padj,direct=F, vertices = nodeattrib\_t\_its)  
## 网络中的节点相对丰度  
t\_ra\_its <- apply(otu\_norm\_its\_t,1,mean)  
t\_ra\_its <- t\_ra\_its[V(t\_net\_its)$name]  
#将上述显著组间差异的OTU着色，这里有CPT,ZT,PT三种处理，因此理论上有6种可能的差异丰度分布情况  
cs <- c("CPT","CPT\_ZT","ZT","CPT\_PT","ZT\_PT","PT")  
unique(V(t\_net\_its)$indicgroup)

[1] "NA" "CPT" "PT" "CPT\_ZT" "ZT"

V(t\_net\_its)$color <- V(t\_net\_its)$indicgroup  
V(t\_net\_its)$color[!V(t\_net\_its)$color %in% cs] <- "gray30"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "CPT"] <- "#F8766D"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "CPT\_ZT"] <- "#00FFFF"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "ZT"] <- "#619CFF"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "CPT\_PT"] <- "#800080"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "ZT\_PT"] <- "#FFFF00"  
V(t\_net\_its)$color[V(t\_net\_its)$color == "PT"] <- "#00BA38"  
V(t\_net\_its)$frame.color <- V(t\_net\_its)$color  
#上述着色信息映射到node属性表中  
tits\_nodes <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup %in% cs,])  
#设置节点形状  
V(t\_net\_its)$shape <- "circle"  
##设置节点大小，非显著组间差异OTU设置为"3"，显著的为"6"  
V(t\_net\_its)$size <- V(t\_net\_its)$name  
V(t\_net\_its)$size[!V(t\_net\_its)$size %in% tits\_nodes] <- 3  
V(t\_net\_its)$size[V(t\_net\_its)$size %in% tits\_nodes] <- 6  
tits\_nodesizes <- as.numeric(V(t\_net\_its)$size)  
#向量化各类型的显著组间差异OTU以便后续计算  
CPT\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="CPT",])  
CZ\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="CPT\_ZT",])  
ZT\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="ZT",])  
CP\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="CPT\_PT",])  
ZP\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="ZT\_PT",])  
PT\_nodes\_t <- rownames(nodeattrib\_t\_its[nodeattrib\_t\_its$indicgroup=="PT",])  
cs\_nodes\_t\_all <- c(CPT\_nodes\_t,CZ\_nodes\_t,ZT\_nodes\_t,CP\_nodes\_t,ZP\_nodes\_t,PT\_nodes\_t)  
#将网络中的节点/OTU进行聚类，这里采用fast greedy法  
cfg\_t <- cluster\_fast\_greedy(as.undirected(t\_net\_its))  
#查看包含OTU数量最多的10个模块，以进行后续的着色  
t\_modules <- sort(table(membership(cfg\_t)),decr=T)  
t\_modules\_10 <- t\_modules[1:10]  
sm10\_plot <- t\_modules\_10  
names(sm10\_plot) <- as.factor(1:10)  
#寻找包含OTU数量最多的10个模块中具有显著组间差异OTU的模块  
t\_modules\_cs <- table(factor(membership(cfg\_t)[cs\_nodes\_t\_all],levels=names(t\_modules)))  
t\_modules\_cs\_10 <- t\_modules\_cs[names(t\_modules\_10)]  
smcs10\_plot <- t\_modules\_cs\_10  
names(smcs10\_plot) <- as.factor(1:10)  
#将OTU数量最多的10个模块中的OTU向量化  
t\_modules\_points <- membership(cfg\_t)[membership(cfg\_t) %in% names(t\_modules\_10)]  
t\_points <- NULL  
for(i in t\_modules\_points){  
 tx <- which(names(t\_modules\_10)==i)  
 t\_points <- c(t\_points, tx)  
}  
names(t\_points) <- names(t\_modules\_points)  
#按照组间差异OTU的类型着色这些OTU  
t\_all\_cols <- sort(t\_points)  
t\_all\_cols[!names(t\_all\_cols) %in% cs] <- "gray30"  
t\_all\_cols[names(t\_all\_cols) %in% CPT\_nodes\_t] <- "#F8766D"  
t\_all\_cols[names(t\_all\_cols) %in% CZ\_nodes\_t] <- "#00FFFF"  
t\_all\_cols[names(t\_all\_cols) %in% ZT\_nodes\_t] <- "#619CFF"  
t\_all\_cols[names(t\_all\_cols) %in% CP\_nodes\_t] <- "#800080"  
t\_all\_cols[names(t\_all\_cols) %in% ZP\_nodes\_t] <- "#FFFF00"  
t\_all\_cols[names(t\_all\_cols) %in% PT\_nodes\_t] <- "#00BA38"  
#设置节点样式，1为空心圆代表普通节点，16为实心圆代表组间差异OTU  
t\_all\_pch <- sort(t\_points)  
t\_all\_pch[names(t\_all\_pch) %in% rownames(otu\_norm\_its\_t)] <- 1  
t\_all\_pch[names(t\_all\_pch) %in% intersect(rownames(otu\_norm\_its\_t),cs\_nodes\_t\_all)] <- 16  
#设置节点缩放倍数，1为普通节点1倍，2为组间差异OTU2倍  
t\_all\_cex <- sort(t\_points)  
t\_all\_cex[!names(t\_all\_cex) %in% cs\_nodes\_t\_all] <- 1  
t\_all\_cex[names(t\_all\_cex) %in% cs\_nodes\_t\_all] <- 2  
#哪些模块包含有组间差异OTU  
t\_mods\_list\_cs <- list()  
for (i in names(t\_modules\_cs\_10)){  
 x1 <- names(membership(cfg\_t)[membership(cfg\_t)==i])  
 x2 <- x1[x1 %in% cs\_nodes\_t\_all]  
 t\_mods\_list\_cs[[i]] <- as.numeric(V(t\_net\_its)[x2])  
}  
t\_mods\_list\_cs

$`1`  
 [1] 39 129 140 142 144 167 169 170 175 179 180 193 194 210 227 249 270 283 308  
[20] 317 331 336 340 345 362 372 380 382 384 389 392 515 519 569 631 648 649 700  
  
$`2`  
 [1] 136 138 146 151 159 160 216 235 238 244 250 294 302 309 595  
  
$`3`  
[1] 113 516 561 724  
  
$`4`  
 [1] 9 10 56 80 85 98 120 124 130 152 507 590 627  
  
$`5`  
numeric(0)  
  
$`6`  
numeric(0)  
  
$`7`  
numeric(0)  
  
$`8`  
[1] 299 427 466 573  
  
$`9`  
numeric(0)  
  
$`10`  
numeric(0)

#设定layout出图，这里选择Fruchterman & Reingold  
set.seed(8051)  
coords\_t\_its <- layout\_(t\_net\_its,with\_fr(niter=9999, grid="nogrid"))  
#每次运行耗费时间，可以把文件储存  
#write.table(coords\_t\_its,paste0(output,"coords\_t\_its.txt"),sep="\t",row.names=F,col.names=F,quote=F)  
#coords\_t\_its <- as.matrix(read.table("coords\_t\_its.txt"))  
#dimnames(coords\_t\_its) <- NULL  
#出图  
pdf(paste0("p1\_network.pdf"),width=7,height=3.5)  
par(mfrow=c(1,3), mar=c(0,0,0,0))  
t\_cols <- c("antiquewhite","antiquewhite1","antiquewhite2","antiquewhite3","antiquewhite4")  
plot(t\_net\_its,vertex.label=NA,vertex.size=tits\_nodesizes, layout=coords\_t\_its,  
 mark.groups=list(t\_mods\_list\_cs$`1`,t\_mods\_list\_cs$`2`,t\_mods\_list\_cs$`3`,t\_mods\_list\_cs$`4`,t\_mods\_list\_cs$`8`),  
 mark.col=t\_cols, mark.border=t\_cols)  
legend("bottomleft",legend=c("Module 1","Module 2", "Module 3", "Module 4", "Module 8"),col=t\_cols,  
 bty="n",fill=t\_cols,border=t\_cols)  
dev.off()

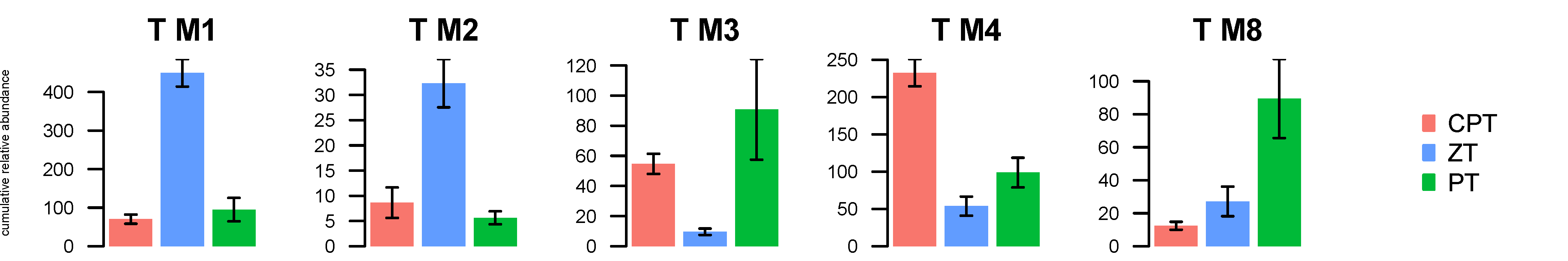
png   
 2

得到体现组间差异OTU及所在主要模块的co-ouccurrence network，原plot中该图仅占1/3的位置，为J, F两个时期的plot留了空间  图1. 共现网络展示了真菌OTU之间存在的显著相关关系。在特定耕作处理下大量富集的组间差异OTU被着色。CPT红色，ZT蓝色，PT绿色；组间差异不显著的OTU呈灰色。凸出显示了前10个网络模块中包含这些OTU的模块。

##### 模块中组内OTU相对丰度比较

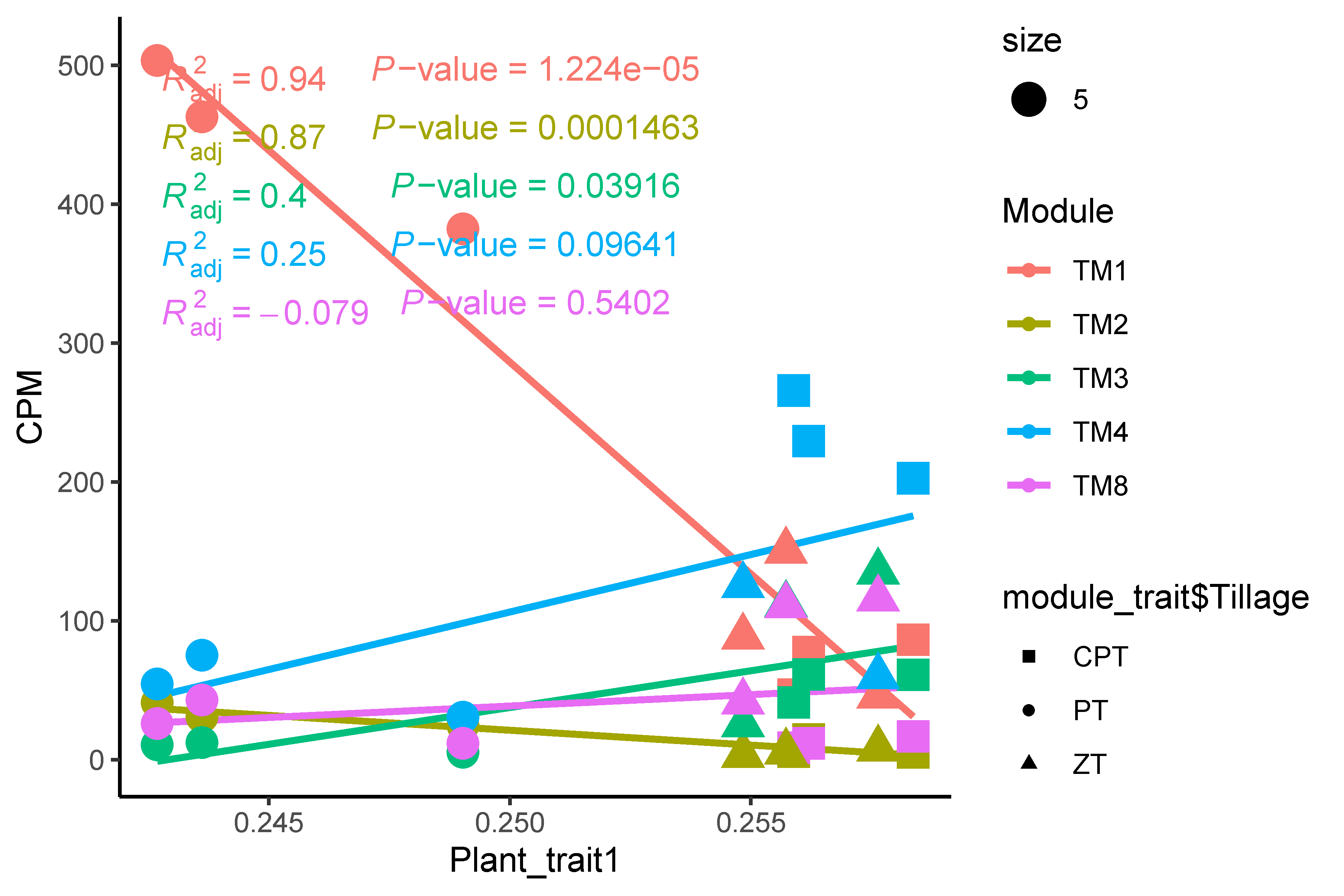
pdf(paste0("p2\_module\_abundance.pdf"),width=7,height=7/6)  
par(mfrow=c(1,6), mar=c(0.5,3.5,2,0))#根据模块数调整mfrow  
CS\_cols <- c("#F8766D","#619CFF","#00BA38")  
names(CS\_cols) <- c("CPT","ZT", "PT")  
#T module 1  
bargraph.CI(design\_its\_t$tillage, colSums(otu\_norm\_its\_t[cfg\_t[[1]],])/1000,   
 las=2, ylab="cumulative relative abundance", cex.lab=.5, cex.axis=.7, cex.names=.7,  
 err.width=.025, main="T M1", col=CS\_cols, border=F)  
#T module 2  
bargraph.CI(design\_its\_t$tillage, colSums(otu\_norm\_its\_t[cfg\_t[[2]],])/1000,   
 las=2, ylab="", cex.lab=.5, cex.axis=.7, cex.names=.7,  
 err.width=.025, main="T M2", col=CS\_cols, border=F)  
#T module 3  
bargraph.CI(design\_its\_t$tillage, colSums(otu\_norm\_its\_t[cfg\_t[[3]],])/1000,   
 las=2, ylab="", cex.lab=.5, cex.axis=.7, cex.names=.7,  
 err.width=.025, main="T M3", col=CS\_cols, border=F)  
#T module 4  
bargraph.CI(design\_its\_t$tillage, colSums(otu\_norm\_its\_t[cfg\_t[[4]],])/1000,   
 las=2, ylab="", cex.lab=.5, cex.axis=.7, cex.names=.7,  
 err.width=.025, main="T M4", col=CS\_cols, border=F)  
#T module 8  
bargraph.CI(design\_its\_t$tillage, colSums(otu\_norm\_its\_t[cfg\_t[[8]],])/1000,   
 las=2, ylab="", cex.lab=.5, cex.axis=.7, cex.names=.7,  
 err.width=.025, main="T M8", col=CS\_cols, border=F)  
plot.new()  
par(mar=c(0.5,0,2,0))  
legend("left", bty="n", cex=1, #x.intersp=0.1, y.intersp=1,  
 legend=names(CS\_cols),   
 fill=CS\_cols,   
 border=CS\_cols , xpd=T)  
dev.off()

png   
 2

得到组间差异OTU所在主要模块在各个处理下的丰度  图2. 在当前共发生网络下，图6中所示包含组间差异OTU模块的累积相对丰度。

##### 模块丰度与其它观测指标回归

#####将包含显著差异OTU的主要模块丰度与其它观测指标进行回归#####  
#提取这些模块的丰度  
csmodule<-rbind(colSums(otu\_norm\_its\_t[cfg\_t[[1]],])/1000,colSums(otu\_norm\_its\_t[cfg\_t[[2]],])/1000,colSums(otu\_norm\_its\_t[cfg\_t[[3]],])/1000,colSums(otu\_norm\_its\_t[cfg\_t[[4]],])/1000,colSums(otu\_norm\_its\_t[cfg\_t[[8]],])/1000)  
row.names(csmodule)<-c("TM1","TM2","TM3","TM4","TM8")  
#宽数据转化为长数据  
csmodule\_long<-melt(csmodule,measure.vars = c('CPT1','ZT3'))  
names(csmodule\_long)<-c("Module","Tillage","CPM")  
#将性状数据与模块丰度数据对应起来，这里有5个模块所以each=5  
plant\_trait1 <- plant\_trait[rep(1:nrow(plant\_trait),each=5),]  
module\_trait<-cbind(csmodule\_long,plant\_trait1)  
#去除重复的数字标号（i.e., CPT1→CPT）  
module\_trait$Tillage<-gsub("\\d","",module\_trait$Tillage)  
#绘制每个模块丰度-植物性状关系  
t1<-ggplot(module\_trait, aes(x = Plant\_trait1, y = CPM, color = Module)) +  
 geom\_smooth(method = 'lm', formula = y ~ x, se = F) +   
 stat\_poly\_eq(aes(label = paste(..adj.rr.label.., sep = '~~~~')), formula = y ~ x, parse = T) +   
 scale\_linetype\_manual() +   
 geom\_point(aes(size=5,shape=module\_trait$Tillage)) +  
 scale\_shape\_manual(values=c(15,16,17)) +  
 theme\_classic()+   
 stat\_fit\_glance(method = "lm",   
 label.x = "centre",  
 label.y = "top",  
 aes(label = paste("italic(P)\*\"-value = \"\*",   
 signif(..p.value.., digits = 4), sep = "")),  
 parse = TRUE)  
ggsave(t1, file=paste0("p3\_regression.pdf"), width=6, height=4)

得到组间差异OTU所在主要模块丰度与植株性状指标的关系  图3. 植物性状与网络模块的累积相对丰度之间的回归关系

#### 参考文献

* Kyle Hartman, Marcel G. A. van der Heijden, Raphaël A. Wittwer, Samiran Banerjee, Jean-Claude Walser & Klaus Schlaeppi. (2018). Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. Microbiome 6, 14, doi: <https://doi.org/10.1186/s40168-017-0389-9>
* Kunkun Fan, Manuel Delgado-Baquerizo, Xisheng Guo, Daozhong Wang, Yanying Wu, Mo Zhu, Wei Yu, Huaiying Yao, Yong-guan Zhu & Haiyan Chu. (2019). Suppressed N fixation and diazotrophs after four decades of fertilization. Microbiome 7, 143, doi: <https://doi.org/10.1186/s40168-019-0757-8>
* Natacha Bodenhausen, Vincent Somerville, Alessandro Desirò, Jean-Claude Walser, Lorenzo Borghi, Marcel G. A. van der Heijden & Klaus Schlaeppi. (2019). Petunia- and Arabidopsis-Specific Root Microbiota Responses to Phosphate Supplementation. Phytobiomes Journal 3, 112-124, doi: <https://doi.org/10.1094/pbiomes-12-18-0057-r>

责编：刘永鑫 中科院遗传发育所

版本更新历史

1.0.0，2020/7/7，李雨泽，初稿

1.0.1，2020/7/8，李雨泽，大修

1.0.2，2020/8/20，刘永鑫，全文修改