

# 基于微生物成分数据的差异 zOTU 分析流程

## Workflow for discriminative zOTU analysis based on compositional data

焦金真<sup>1</sup>, 张小丽<sup>1,2</sup>, 王敏<sup>1</sup>, 谭支良<sup>1,\*</sup>

<sup>1</sup> 中国科学院亚热带农业生态研究所, 亚热带农业生态过程重点实验室, 畜禽养殖污染控制与资源化技术国家工程实验室, 动物营养生理和代谢过程湖南省重点实验室, 湖南长沙 410125

<sup>2</sup> 中国科学院大学, 北京, 100049

\*通讯作者邮箱: zltan@isa.ac.cn.

### 摘要

近年来, 得益于测序技术的革新及成本的大幅度下降, 一系列微生物组 (microbiome) 研究技术, 如扩增子、宏基因组和宏转录组等, 被用于医学、工农业和环境等各个领域的研究, 推动微生物组研究进入了黄金发展期<sup>[1,2]</sup>。面对如此海量的测序数据, 如何从中有效且准确地挖掘信息显得尤为重要。值得注意的是, 测序时核酸通常被设定到一定浓度, 结果提供的总序列数均为一有限值, 使得任一条序列均为相对比例, 称为成分数据 (compositional data)<sup>[3]</sup>。本文在 Gloor 等 (2017)<sup>[4]</sup> 开发的基于成分数据的分析流程的基础上, 以山羊空肠细菌 16S rRNA 测序数据为示例 (食糜细菌 vs. 粘膜细菌), 在获得试验设计、zOTU 丰度表、zOTU 物种注释信息的基础上, 运用 R 和 R Studio 软件, 详述基于微生物成分数据的差异 zOTU 分析流程。zOTU 指 zero-radius OTU, 是通过 unosi3 命令去噪输出的有效生物学序列, 类似于 DADA2 和 Deblur 输出的扩增序列变体 (amplicon sequence variants, ASV)。

**关键词:** 扩增子测序, 成分数据, 差异 zOTU

### 分析策略

一般而言, 微生物 16S rRNA 扩增子的分析流程是从将测序序列数归一到同一测序深度, 即 Normalization 开始的 (表 1)。主流的 Normalization 是将测序序列进行抽平<sup>[5]</sup>, 或者用 R 的 DESeq 包对序列数进行均一化<sup>[6, 7]</sup>。基于成分数据的 Normalization 是通过 centered log-ratio (CLR) 将测序数据转化为比例数据<sup>[8]</sup>, 极大地降低了下游分析受测序深度的影响程度。下一步的分析通常是计算样品之间的  $\beta$  多样性 (PCoA), 降维出图 (Ordination) 和多元比较 (Multivariate comparison)。主流的方法中这些分析是基于样品间的 Bray-Curtis、UniFrac 和 Jensen-Sha

nnon 距离等<sup>[4, 7]</sup>。成分数据的下游分析是基于通过 CLR 计算的 Aitchison 距离，之后进行 PCA（Principal Component Analysis）降维，这可以最大程度地保证成分数据的降维是通过数据集中变异程度最大的 zOTU 实现的。主流的和基于成分数据的分析均通过 ANOSIM、ADONIS 等方式进行多元比较，来统计两组的微生物之间是否存在显著差异。最后，在筛选差异微生物（如差异 zOTU）时，与主流流程通过 LEfSe、DESeq2 等不同的是，基于成分数据的差异 zOTU 分析是通过 R 包中的 ANCOM 和 ALDEx2 等来实现的<sup>[4,7,8]</sup>。本流程最大的优势是可以极大程度降低测序深度对分析的影响程度，使得分析结果更准确可信。

表 1. 微生物数据分析的主流和成分数据分析流程

步骤	主流分析	成分数据分析
均一化	Rarefaction DESeq	CLR ILR ALR
距离	Bray-Curtis UniFrac Jenson-Shannon	Aitchison
降维	PCoA (Abundance)	PCA (Variance)
多元比较	ADONIS ANOSIM	ADONIS ANOSIM
差异分析	LEfSe DESeq2	ANCOM ALDEx2

软件

R (version 3.6.2)  
R studio (<https://github.com/rstudio/rstudio/>)

输入数据

试验设计表格: conds.txt  
zOTU 表格: zotu.txt  
zOTU 注释信息表格: taxonomy.txt

实验步骤

1. R 包安装及加载

本分析需要用到 5 个 R 包：  
ALDEx2 (version 1.14.0, 差异分析)、  
CoDaSeq (version 0.99.2, 成分数据分析)、

factoextra (version 1.0.6, PCA 图)、  
ggplot2 (version 3.2.1, 作图)、  
vegan (version 2.5-6, ADONIS 分析)。

1.1. R studio 中可以通过菜单安装 R 包。选择“packages”窗口的“install”进行安装。

注：ALDEx2 和 CoDaSeq 可以分别用以下命令进行安装。

```
source("https://bioconductor.org/biocLite.R")  
biocLite("ALDEx2")  
library(devtools)  
devtools::install_github('ggloor/CoDaSeq/CoDaSeq')
```

1.2. 加载 R 包

```
library (R 包名)
```

## 2. 数据的载入

通过以下三条命令分别载入试验设计表格、zOTU 表格、zOTU 物种注释信息表格。

```
conds <- read.delim ("conds.txt", header=T, row.names= , sep="\t")  
otu = read.delim ("zotu.txt", header=T, row.names= 1, sep="\t")  
tax = read.delim ("taxonomy.txt", header=T, row.names= 1, sep="\t")
```

注：本流程数据来源于山羊空肠食糜和粘膜细菌 16S rRNA 扩增子数据，分别测定了其食糜（Digesta, n=6）和粘膜（Mucosa, n=6）的细菌组成。前期通过 usearch 的 unosie3 命令将拼接的 Tags 在 99%的相似度上聚类为 zOTU。之后执行 SINTAX 命令比对了 RDP 数据库，对 zOTU 进行了物种注释。本流程旨在筛选空肠不同生态位细菌的差异 zOTU。

## 3. 运用 ALDEx2 包的相关命令对 zOTU 表进行 CLR 变换，获得每个 zOTU 的效应量大小和E（CLR）矩阵

3.1 运用 codaSeq.filter 命令过滤 zotu 表格，之后通过 aldex.clr 进行中心对数（CLR）变换。

```
f <- codaSeq.filter(zotu, min.reads=5000, min.prop=0.001, min.occurrence = 0.1, samples.by.  
row=FALSE)  
f.x <- aldex.clr(f, as.vector(conds$Region), mc.samples=256, verbose=FALSE, useMC = TRUE)  
E)
```

注：aldex.clr 将数据存储在 s4 类型变量 f.x 对象中，包括参数 reads、conds、mc.samples、denom、verbose、useMD 和 analysisData 等（图 1）。\$reads 为 CLR 的表达矩阵，conds 为样品的分组信息。

Name	Type	Value
f.x	S4 (ALDEx2::aldex.clr)	S4 object of class aldex.clr
reads	list [382 x 12] (S3: data.frame)	A data.frame with 382 rows and 12 columns
conds	character [12]	'Digesta' 'Digesta' 'Digesta' 'Digesta' 'Digesta' 'Digesta' ...
mc.samples	double [1]	256
denom	double [382]	1 2 3 4 5 6 ...
verbose	logical	FALSE
useMC	logical	TRUE
analysisData	list [12]	List of length 12

图 1. CLR 变换后的 f.x 的数据结构

3.2 通过 `aldex.effect` 命令获得每个 zOTU 的效应量大小 (expected CLR value)。

```
f.e <- aldex.effect(f.x, as.vector(conds$Region),
  include.sample.summary=TRUE,
  verbose=FALSE,
  useMC = TRUE)
```

注: `aldex.effect` 的输出结果 `f.e` (图 2) 包括: `rab.sample.***` (zOTU 在各个样品的 CLR 值); `dif`  
`f.btw` (zOTU 在两组间的中位数差异); `diff.win`, 每个 zOTU (zOTU 在组内 Dirichlet instances 的  
 中位数差异); `effect` (zOTU 的效应量)。`effect` 值可用于后续的差异 zOTU 筛选标准之一。

	rab.all	rab.win.Digesta	rab.win.Mucosa	rab.sample.Jd101	rab.sample.Jd102	rab.sample.Md106	rab.sample.Md107	diff.btw	diff.win	effect	overlap
Zotu1	8.172914908	4.57879492	12.17528485	1.8056631	5.2701882	11.55018525	10.0456767	7.51447988	1.977831	3.99796924	9.137503e-05
Zotu144	1.893965191	-2.04426090	4.59332659	-3.0295515	-2.6379802	3.85882659	2.5301268	6.52340505	2.685136	2.45116840	1.313454e-03
Zotu158	1.942465195	-2.52278929	4.47842245	-2.9288451	-2.8410703	3.87262043	2.5471512	6.93960670	3.024335	2.16459396	9.137503e-05
Zotu197	-0.310746047	-3.02627905	3.80022222	-3.2214130	-2.9971123	3.00038494	-0.4368222	6.46330960	3.524127	1.87339854	2.604233e-02
Zotu346	-0.364915142	-3.04133306	3.09970455	-3.3206246	-2.8120473	0.77852995	-0.2732471	5.73114259	3.363307	1.60732312	1.950674e-02
Zotu168	1.268914409	-3.14501007	4.03517173	-2.9218140	-2.8959488	2.86397918	2.1156404	7.14689842	2.923363	2.43348628	9.137503e-05
Zotu268	0.648982592	-3.00012872	2.46826364	-2.5393453	-3.3725742	1.89345087	1.3515293	5.62793175	2.668757	2.11290762	1.313454e-03
Zotu1471	0.174739592	-2.90685943	2.91341044	-2.7730845	-2.8949441	2.15706254	0.5744063	5.78808837	2.889445	1.91053076	6.504638e-03
Zotu195	1.245314119	-3.12841445	4.01135666	-3.0964201	-2.7335759	2.74412861	2.1202250	6.92214108	2.746124	2.47458944	9.137503e-05
Zotu147	1.674913556	-2.13984110	4.33364864	-3.1612352	-0.6295517	3.01717220	2.5106072	6.50701463	2.686203	2.41382665	9.137503e-05
Zotu103	1.746401371	-2.01700265	4.72727341	-3.0518347	-0.7364707	2.71538902	2.8039113	6.48351578	2.573343	2.44940761	9.137503e-05
Zotu203	1.297897622	-2.98421244	3.49687966	-2.9214620	-2.7889620	2.43890258	2.0963305	6.48849830	2.636902	2.39041011	1.313454e-03
Zotu357	-0.366922550	-3.12683658	3.23769298	-2.6629272	-3.2485984	1.49586464	-0.4341450	6.02584575	3.369531	1.73129214	2.604233e-02
Zotu227	0.869364604	-2.54503026	2.94066416	-2.8939452	-0.5349344	1.76378235	1.5624595	5.45647841	2.658810	2.00906926	1.315165e-03
Zotu258	0.808667532	-3.04230246	3.08195359	-3.1485416	-2.8460423	2.36428266	1.5706194	6.27275760	2.867834	2.07989891	1.313454e-03
Zotu620	-1.125941753	-3.06952965	1.89201964	-2.9305499	-3.1099134	-0.39491494	-3.8223296	4.87017283	4.058062	1.06326228	1.209364e-01

图 2. `aldex.effect` 的输出结果 `f.e`

3.3 获得 E (CLR) 值的矩阵 (E.clr), 用于后续的 PCA 降维和作图。

```
E.E.clr <- t(f.e[,grep("rab.sample", colnames(f.e))])
rownames(E.E.clr) <- gsub("rab.sample.", "", rownames(E.E.clr))
exp <- apply(E.E.clr, 1, function(x) 2^x)
E.clr <- t(apply(exp, 2, function(x) log2(x) - mean(log2(x))) )
```

注: `E.clr` 矩阵中行是 12 个样品, 列是 382 个 zOTU 的 `E.clr` 值。

111

## 112 4. PCA 降维及作图

### 113 4.1 通过 prcomp 命令进行 PCA 分析

```
114 pcx <- prcomp(E.clr)
```

115

### 116 4.2 执行 factoextra 包的 fviz\_pca\_ind 命令进行 PCA 作图

```
117 fviz_pca_ind(pcx,
118               geom.ind = c("point", "text"),
119               fill.ind = conds$Region, col.ind = conds$Region,
120               pointshape = 21, pointsize = 2,
121               palette = "jco",
122               addEllipses = TRUE, mean.point = FALSE,
123               ellipse.level = 0.95, ellipse.alpha = 0,
124               repel = TRUE,
125               title = "Compositional PCA sample plot Digeta Mucosa",
126               labels = "Regions", labelsize = 4,
127               ggtheme = theme_minimal()) +
128   theme(plot.title = element_text(hjust = 0.5)) +
129   labs(fill = "Regions") + guides(color=FALSE)
130 ggsave("Compositional_PCA_sampleplot_Digeta_Mucosa.jpg", width = 297, height = 210, un
131 its = "mm", dpi = 600)
```

132

## 133 5. 基于 Aitchison 距离进行 ADONIS 分析

### 134 5.1 获得样品间的 Aitchison 距离矩阵

```
135 dist.clr <- dist(E.clr, method = "euclidean")
```

136 注：基于 CLR 向量的欧氏距离的样品间的比例距离，称为 Aitchison 距离。

137

### 138 5.2 利用 vegan 包的 adonis2 和 betadisper 命令进行 ADONIS 分析

```
139 adonis.At = adonis2(dist.clr ~ Region, data = conds, contr.unordered = "contr.sum")
```

```
140 adonis.At
```

```
141 adonis.At.disp = betadisper(dist.clr, conds$Region)
```

```
142 adonis.At.disp
```

```
143 permutest(adonis.At.disp, pairwise = TRUE)
```

```
144 TukeyHSD(adonis.At.disp)
```

注：函数 `adonis2` 研究两组间的 Aitchison 距离的平均值的差异，跟多元方差分析的功能相同；而函数 `betadisper` 研究组内同质性，统计两组内的 Aitchison 距离是否有差异，类似于 Levene's 等方差检验。

## 6. 运用 ALDEx2 包的相关命令进行差异 zOTU 分析

6.1 通过 `aldex.ttest` 命令将变换后的值，用参数或非参数检验进行单变量统计检验，并返回 Benjamini-Hochberg 校正后的 *P* 值。

```
f.t <- aldex.ttest(f.x, as.vector(conds$Region))
```

注：`aldex.ttest` 输出结果 `f.t`（图 4）为矩阵型变量，包括：`we.ep` (Welch's *t* 检验的 *P* 值)；`we.eBH` (Welch's *t* 检验校正后的 *P* 值)；`wi.ep` (秩和检验的 *P* 值)；`wi.eBH` (秩和检验校正后的 *P* 值)。

	we.ep	we.eBH	wi.ep	wi.eBH
Zotu1	1.430234e-05	0.0003169299	0.002164502	0.005809750
Zotu144	2.906558e-03	0.0075968382	0.002172957	0.005826966
Zotu158	3.877126e-03	0.0097829251	0.002215233	0.005912284
Zotu197	3.549797e-03	0.0089389519	0.005521172	0.011914454
Zotu346	6.127402e-03	0.0139089621	0.004895495	0.010869644
Zotu168	2.485323e-03	0.0068113629	0.002164502	0.005809750
Zotu268	4.639330e-03	0.0110395916	0.002358969	0.006194901
Zotu1471	4.246090e-03	0.0104100764	0.002815544	0.007093174
Zotu195	2.352534e-03	0.0065312056	0.002172957	0.005827728
Zotu147	1.776306e-03	0.0052953746	0.002164502	0.005809750
Zotu103	2.222163e-03	0.0060975263	0.002164502	0.005809750
Zotu203	2.666732e-03	0.0070486737	0.002181412	0.005845139
Zotu357	5.000967e-03	0.0120279651	0.005385890	0.011821236
Zotu227	5.766243e-03	0.0135166544	0.002426610	0.006328418
Zotu258	2.563412e-03	0.0071480613	0.002232143	0.005947312
Zotu620	6.730485e-02	0.0918786234	0.045090977	0.069279701
Zotu810	1.577556e-01	0.1926019867	0.203505479	0.240928471

图 3. `aldex.ttest` 的输出结果 `f.t`

6.2 基于 `effect` 效应量值和校正后的 *P* 值进行差异 zOTU 筛选，获得在 *Mucosa* 中富集的 zOTU (`high.e.Mucosa`) 和在 *Digesta* 中富集的 zOTU (`high.e.Digesta`)。

```
low.p <- which(f.t$we.eBH < 0.05)
```

```
high.e <- which(abs(f.e$effect) > 1)
```

```
high.e.Mucosa <- which(f.e$effect >= 1 & f.t$we.eBH < 0.05)
```

```
high.e.Digesta <- which(f.e$effect <= -1 & f.t$we.eBH < 0.05)
```

注：本文中差异 zOTU 要同时满足以下两个标准：Welch's *t* 检验校正后的 *P* 值<0.05；`effect` 值>1 或者 `effect` 值<-1。

167

168 6.3 差异 zOTU 的 effect 图。

169 pdf (file = "Discriminative zOTUs.pdf")

170 par(mfrow=c(1,2))

171

172 plot(f.e\$diff.win, f.e\$diff.btw,

173 pch=19, col=rgb(0,0,0,0.3),

174 cex=0.5, main="effect plot",

175 xlab="Dispersion", ylab="Difference")

176 points(f.e\$diff.win[high.e],

177 f.e\$diff.btw[high.e],

178 col=rgb(1,0,0,0.5), cex=0.8)

179 points(f.e\$diff.win[low.p],

180 f.e\$diff.btw[low.p],

181 pch=19, col=rgb(0,0,1,0.5),

182 cex=0.5)

183 abline(0,1, lty=2, lwd=2,col="grey")

184 abline(0,-1, lty=2, lwd=2,col="grey")

185

186 plot(f.e\$effect, f.t\$we.eBH, log="y",

187 pch=19, col=rgb(0,0,0,0.3),

188 main="E vs p", xlab="effect size",

189 ylab="E(p.adjust)", cex=0.5)

190 points(f.e\$effect[high.e], f.t\$we.eBH[high.e],

191 col=rgb(1,0,0,0.5), cex=0.8)

192 points(f.e\$effect[low.p], f.t\$we.eBH[low.p],

193 pch=19, col=rgb(0,0,1,0.5),cex=0.5)

194 abline(v=1, lty=2, lwd=2,col="grey")

195 abline(v=-1, lty=2, lwd=2,col="grey")

196 abline(h = 0.05, lty=2, lwd=2,col="grey")

197

198 dev.off();

199

200 6.4 基于 zOTU 注释信息，获得两组差异 zOTU 的物种信息表格，并写入 csv 文件。

201 high.e.Mucosa.tax &lt;- tax[rownames(f)[high.e.Mucosa],]

202 high.e.Digesta.tax &lt;- tax[rownames(f)[high.e.Digesta],]

```
write.table (high.e.Mucosa.tax,"Mucosa enriched OTUs.csv", row.names = T, quote = F, sep
=",")
write.table (high.e.Digesta.tax,"Digesta enriched OTUs.csv", row.names = T, quote = F, sep=
=",")
```

结果与分析

1. PCA 主成分分析图

如图 4 所示，基于 Aitchison 距离的 PCA 图表明，空肠食糜（Digesta）和粘膜（Mucosa）的细菌组成聚成两个相互独立的簇，即：食糜细菌聚成一簇，而粘膜细菌聚成另外一簇。坐标轴中 PC1/2 的数值为总体差异的解释率分别为 65.1%和 12.9%。

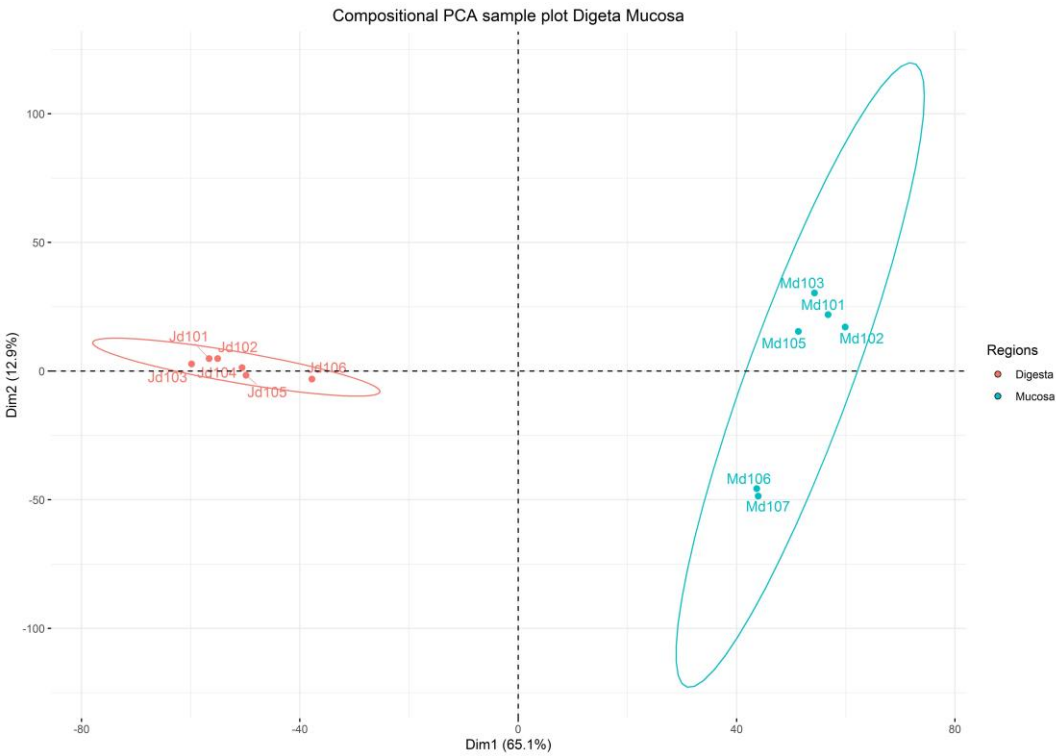


图 4. 基于 Aitchison 距离的 PCA 图

2. ADONIS 分析结果

ADONIS 分析是一种判断不同组的同质性和差异的统计方法。本例中 adonis2 函数的  $P = 0.002$ ，表明山羊空肠食糜和粘膜细菌组成显著不同。betadisper 函数的  $P = 0.026$ ，提示粘膜细菌的组间差异显著高于食糜细菌的组间差异。

3. 差异 zOTUs



### 3.1 差异 zOTUs 的 effect 图

以 Welch's t 检验校正后的  $P$  值 $<0.05$ ；effect 值 $>1$  或者 effect 值 $<-1$  为标准，本示例共筛选出 179 个差异 zOTUs，其中 70 个 zOTUs 在空肠粘膜样品中显著富集，109 个 zOTUs 在空肠食糜样品中显著富集。如图 5 所示，图中的每个点代表 1 个 zOTU。左侧的 effect 图是基于 Dispersion（x 轴，每个 zOTU 的组内中位数差异）和 Difference（Y 轴，每个 zOTU 的组间中位数差异）而做的。当 zOTUs 的 Difference 的绝对值大于 Dispersion 时（即 effect 值 $>1$  或者 effect 值 $<-1$ ），图中的点外围被标红色圆圈；当 zOTUs 的 Welch's t 检验校正后的  $P$  值 $<0.05$ ，图中的点被标蓝。由此判断，左图中的蓝色圆圈且外围标红的圆点即为差异 zOTUs。右图中的 x 轴是每个 zOTU 的 effect size（效应量大小），y 轴是 Welch's t 检验校正后的  $P$  值 $<0.05$ 。与左图类似，右图中的蓝色圆圈且外围标红的圆点即为差异 zOTUs。

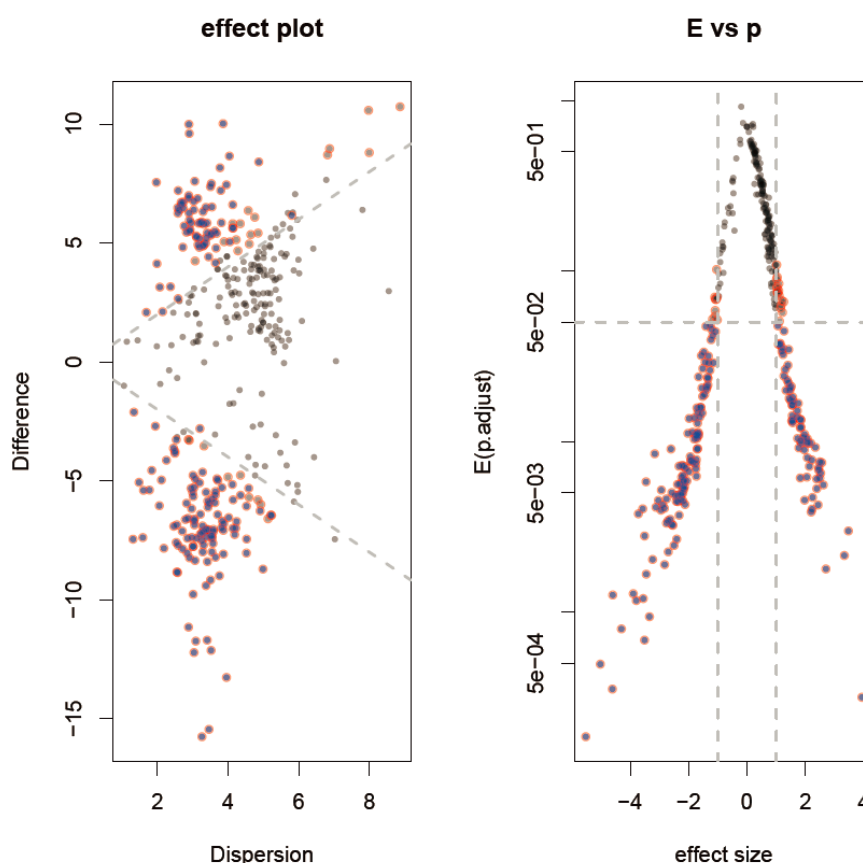


图 5. 差异 zOTUs 的 effect 图

### 3.2 差异 zOTUs 的物种注释

差异 zOTUs 的文件格式如图 6。山羊空肠食糜中（Digesta enriched OTUs.csv）富集的细菌 zOTUs 主要为 *Eubacterium*、*Ruminococcus*、*Romboutsia*、*Lactobacillus*、*Streptococcus*、*Desulfov*

*ibrio*、*Halomonas*、*Shewanella* 等，而山羊粘膜中（*Mucosa enriched OTUs.csv*）富集的细菌 zOTUs 主要为 *Porphyromonas*、*Bacteroides*、*Campylobacter* 和一些未注释到物种信息的 zOTUs。

	Kingdom	Phylum	Class	Order	Family	Genus	Species
Zotu393	Bacteria	Proteobacter	Betaproteobact	Neisseriales	Neisseriaceae	NA	NA
Zotu605	Bacteria	Proteobacter	Betaproteobact	Neisseriales	Neisseriaceae	NA	NA
Zotu302	Bacteria	Bacteroidete	NA	NA	NA	NA	NA
Zotu35	Bacteria	Bacteroidete	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	NA
Zotu76	Bacteria	Bacteroidete	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	NA
Zotu587	Bacteria	Bacteroidete	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides	NA
Zotu2	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	NA	NA
Zotu44	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	NA	NA
Zotu9	Bacteria	Firmicutes	NA	NA	NA	NA	NA
Zotu322	Bacteria	Firmicutes	Clostridia	Clostridiales	NA	NA	NA
Zotu57	Bacteria	Firmicutes	Clostridia	Clostridiales	Eubacteriaceae	Eubacterium	NA
Zotu75	Bacteria	Firmicutes	Clostridia	Clostridiales	Eubacteriaceae	Eubacterium	NA
Zotu36	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu45	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu398	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu121	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu18	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae_1	Clostridium_s	NA
Zotu161	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae_1	Clostridium_s	NA
Zotu7	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	NA	NA
Zotu22	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu53	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu112	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA
Zotu176	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcus	NA

图 6. 差异 zOTUs 的物种注释结果

致谢

感谢国家自然科学基金委青年项目（31601967）对本分析流程的资助。本试验方案内容改编于 Gloor 等（2017）开发的流程，特此致谢。本分析流程已应用于 Zhang 等（2020）的文章中，特此致谢。

参考文献

1. Claesson MJ, Clooney AG and O'Toole PW. 2017, A clinician's guide to microbiome analysis. Nat Rev Gastroenterol Hepatol. 14(10): 585-595.

2. Zimmermann M, Zimmermann-Kogadeeva M, Wegmann R, Goodman AL. 2019, Separating host and microbiome contributions to drug pharmacokinetics and toxicity. Science. 363(6427).

3. Lovell D, Pawlowsky-Glahn V, Egozcue JJ, Marguerat S, Bahler J. 2015. Proportionality: a valid alternative to correlation for relative data . PLoS Comput Biol. 11(3): e1004075.

4. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. 2017, Microbiome datasets are compositional: and this is not optional. Front Microbiol. 8: 2224.

5. Jiao J, Huang J, Zhou C, Tan Z. 2015, Taxonomic identification of ruminal epithelial bacterial diversity during rumen development in goats. Appl Environ Microbiol. 81(10): 3502-3509.

6. Love MI, Huber W and Anders S. 2014, Moderated estimation of fold change and dispersion for

RNA-seq data with DESeq2. *Genome Biol.* 15(12): 550.

7. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vazquez-Baeza Y, Birmingham A, Hyde ER, Knight R. 2017, Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome.* 5.
8. Thorsen J, Brejnrod A, Mortensen M, Rasmussen Morten A, Stokholm Jakob, Abu Al-Soud Waleed, Sorensen Soren, Bisgaard Hans, Waage Johannes. 2016, Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. *Microbiome.* 4.
9. Zhang X, Wu J, Zhou C, Tan Z, Jiao J. 2020. Spatial and temporal organization of jejunal microbiota in goats during animal development process. *J Appl Microbiol.* doi: 10.1111/jam.14961.