

# butyrate producing genes annotation

Chunyu Zhao

12 December, 2018

## Contents

<b>1</b>	<b>butyrate-producing bacterial</b>	<b>1</b>
1.1	butyrate meta . . . . .	1
1.2	read in contigs . . . . .	2
1.3	pident and scov . . . . .	2
1.4	per base coverage . . . . .	4
1.5	merga data . . . . .	4
1.6	a complete butyrate producing pathway . . . . .	4
1.7	metadata . . . . .	4
<b>2</b>	<b>bysample: total pathway abundance</b>	<b>4</b>
2.0.1	gene versus pathway . . . . .	6
2.0.2	formula effect at 4.5 month . . . . .	6
<b>3</b>	<b>lysine pathway</b>	<b>7</b>
<b>4</b>	<b>TCA cycle pathway</b>	<b>11</b>
<b>5</b>	<b>Glutarate pathway</b>	<b>14</b>
<b>6</b>	<b>Pyruvate</b>	<b>16</b>

## 1 butyrate-producing bacterial

Butyrate-producing bacteria have recently gained attention, since they are important for a healthy colon and when altered contribute to emerging disease, such as ulcerative colitis and type II diabetes.

- a butyrate-producing taxonomic core in healthy colons
- two *final* genes in the two primary bacterial butyrate synthesis pathways, butyryl-CoA: acetate CoA-transferase (but) and butyrate kinase (buk).
- butyrate-producing community
- important things let's say it three times:

Eubacterium spp. and Roseburia spp (Clostridium cluster XIVa) and Faecalibacterium prausnitzii (Clostridium cluster IV).

### 1.1 butyrate meta

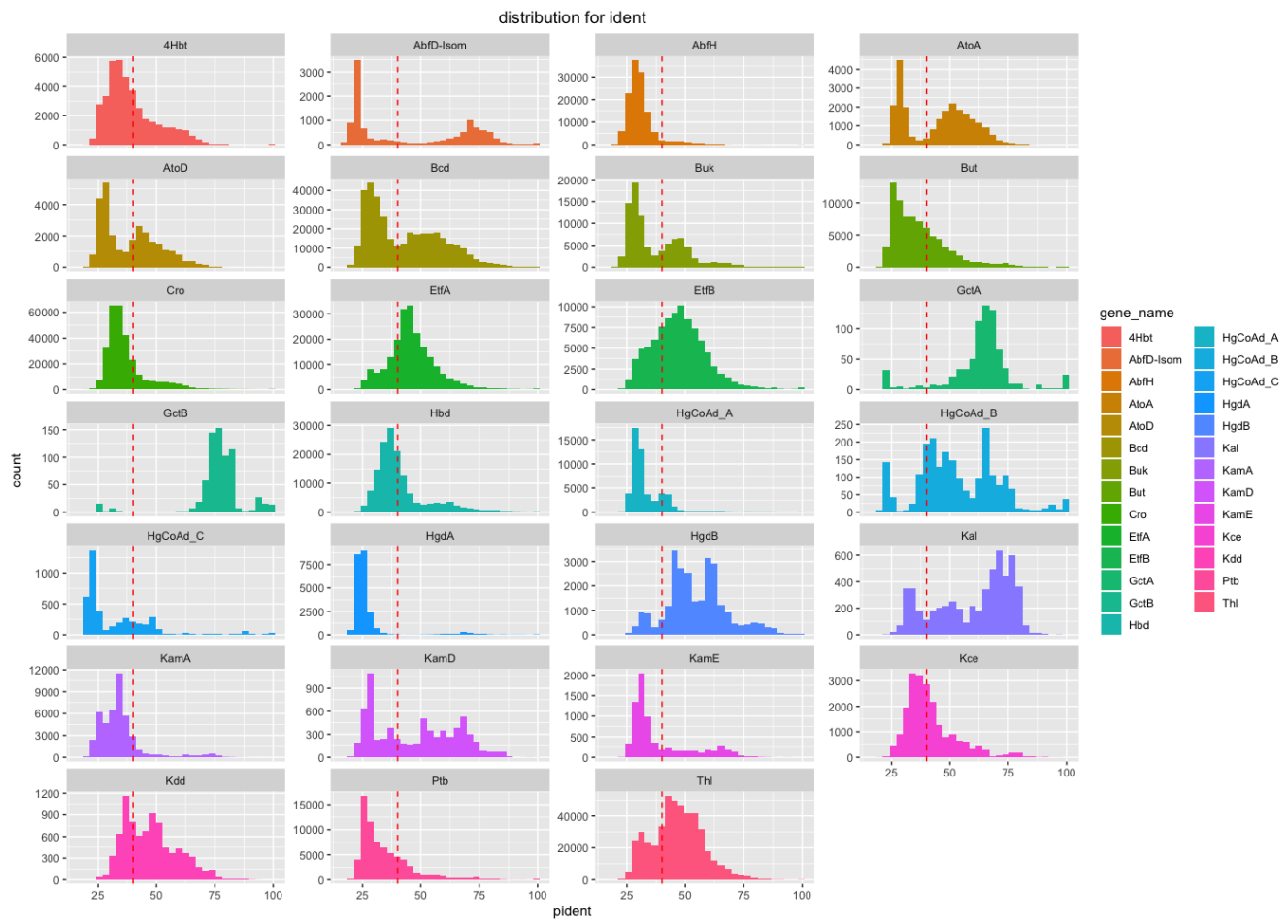
Table 1: butyrate producing genes database first look

pathway_name	gene_name	n
4aminobutyrate	4Hbt	68
4aminobutyrate	AbfD-Isom	44
4aminobutyrate	AbfH	38
Glutarate	GctA	25

pathway_name	gene_name	n
Glutarate	GctB	27
Glutarate	HgCoAd_A	27
Glutarate	HgCoAd_B	27
Glutarate	HgCoAd_C	27
Glutarate	HgdA	27
Glutarate	HgdB	28
Lysine	AtoA	62
Lysine	AtoD	62
Lysine	Kal	64
Lysine	KamA	74
Lysine	KamD	66
Lysine	KamE	66
Lysine	Kce	67
Lysine	Kdd	67
Pyruvate	Bcd	475
Pyruvate	Buk	106
Pyruvate	But	79
Pyruvate	Cro	209
Pyruvate	EtfA	254
Pyruvate	EtfB	250
Pyruvate	Hbd	205
Pyruvate	Ptb	92
Pyruvate	Thl	227

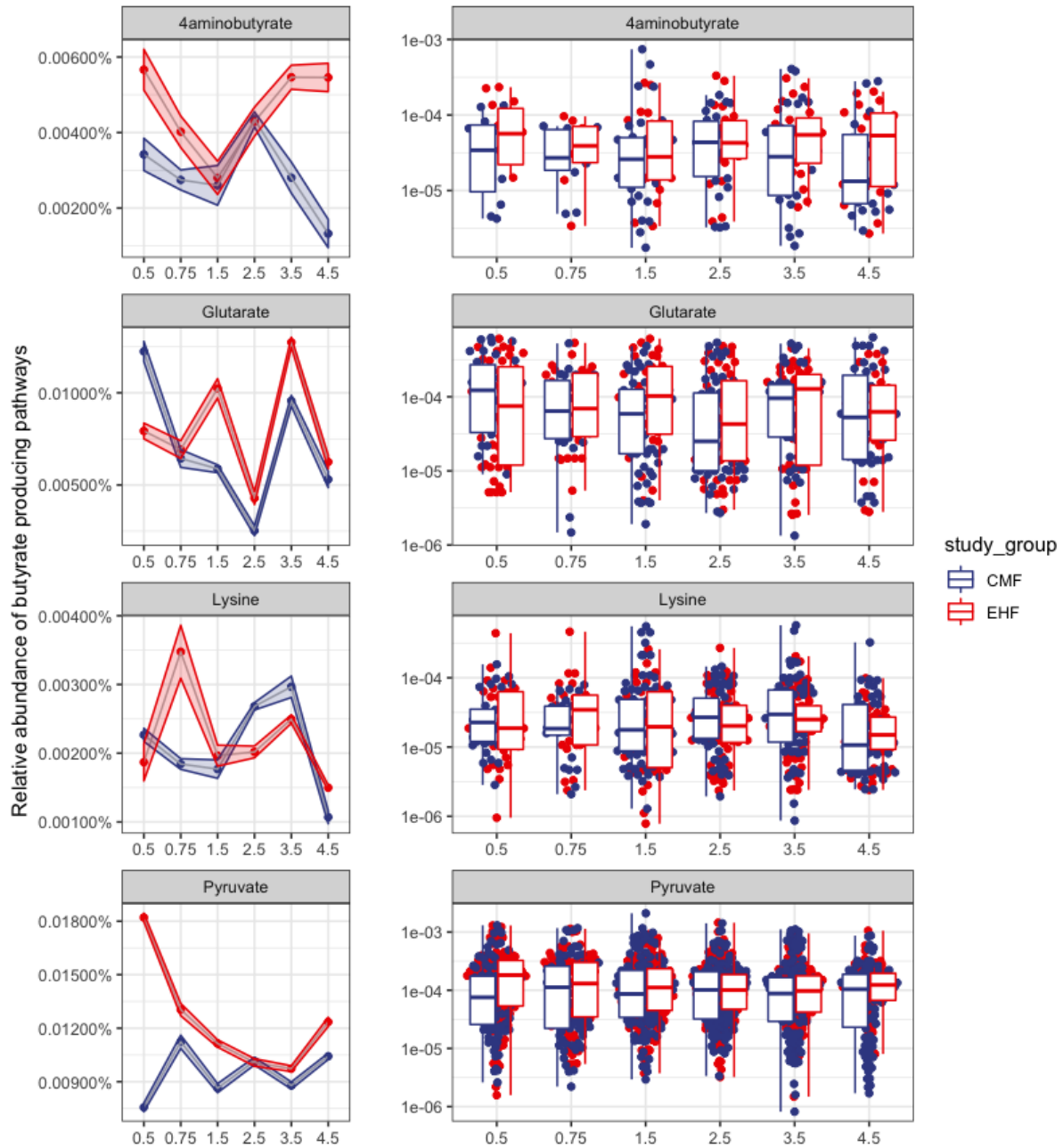
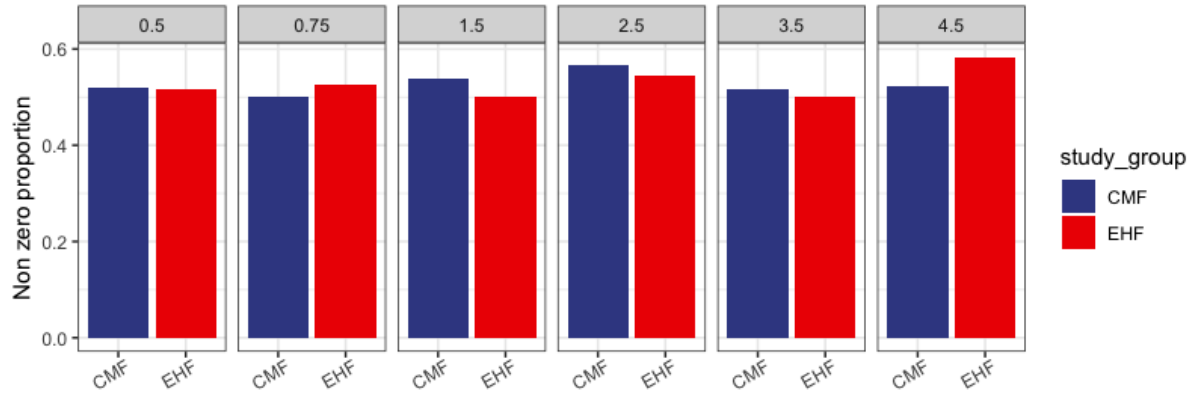
## 1.2 read in contigs

## 1.3 pident and scov





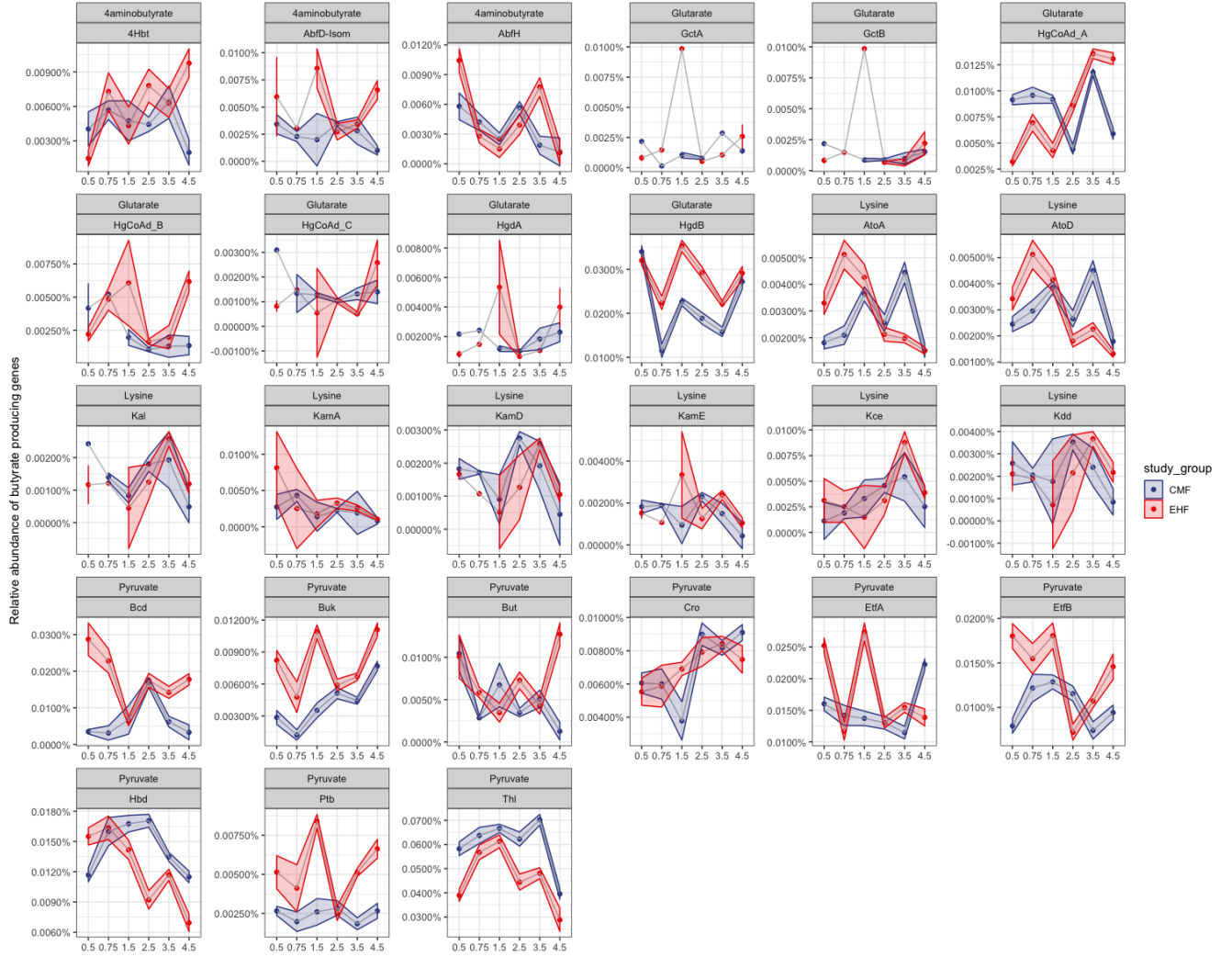
Raw butyrate producing pathway abundance without any proper filtering



## 2.0.1 gene versus pathway

Table 2: lme result

gene_name	term	Value	Std.Error	DF	t.value	p.value	fdr
Bcd	EHF	0.8196	0.2575	28	3.183	0.003558	0.08183
HgdB	study_month0.75	-0.5106	0.1125	108	-4.537	1.488e-05	0.0003423



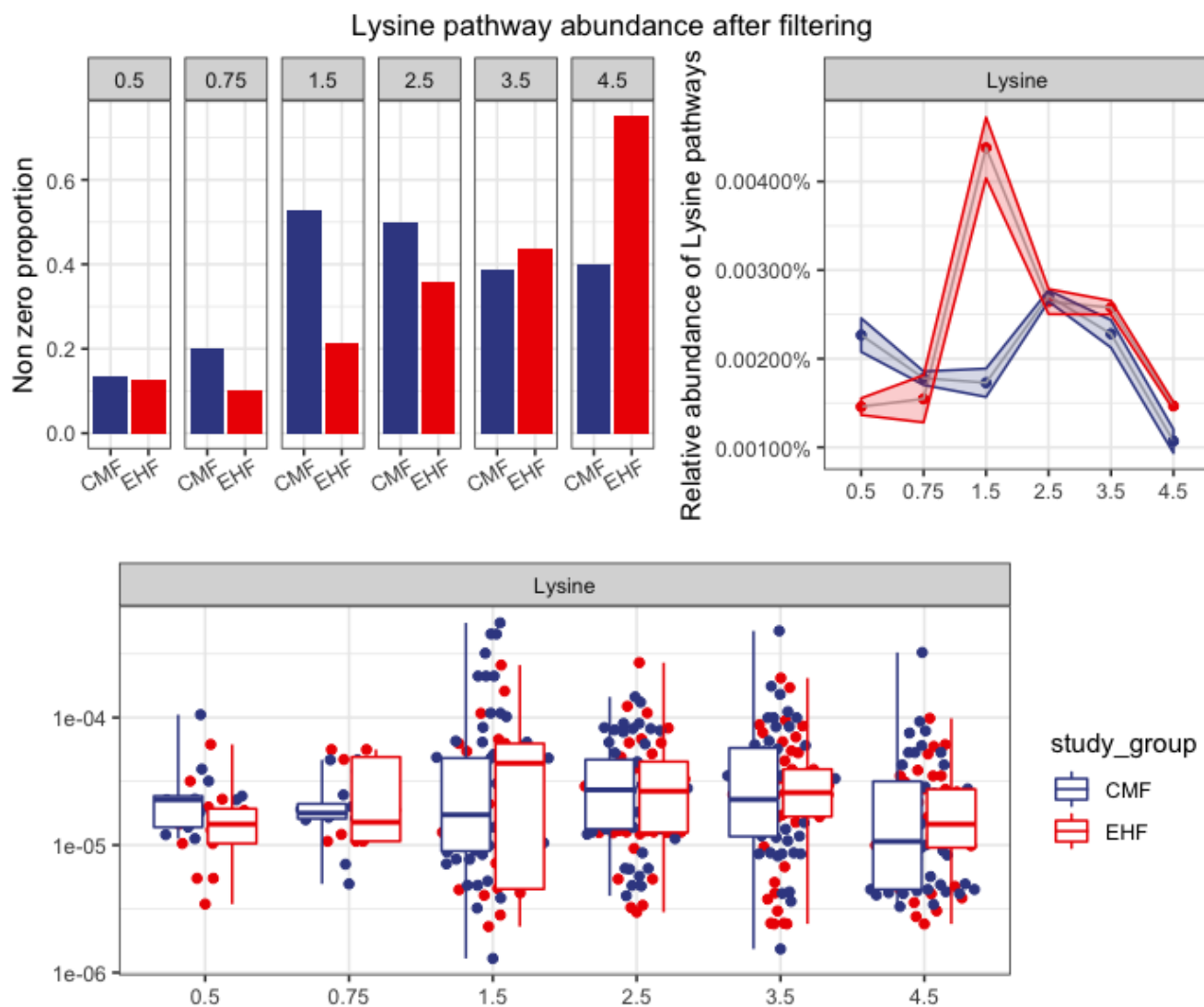
## 2.0.2 formula effect at 4.5 month

Table 3: lm result

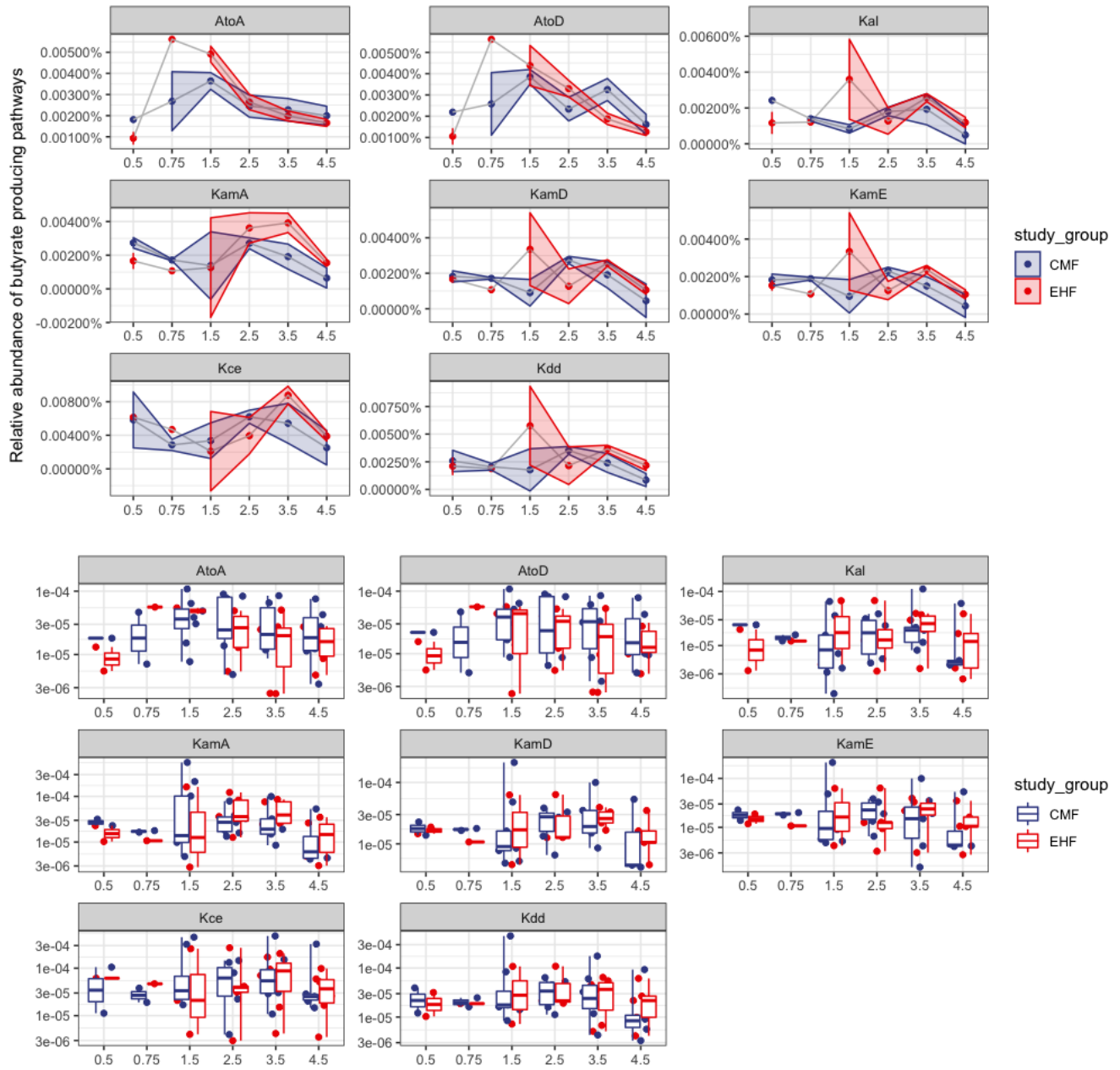
gene_name	term	estimate	std.error	statistic	p.value	fdr
Bcd	CMF-EHF	0.9482	0.4094	2.316	0.03332	0.2999
But	CMF-EHF	0.7213	0.4037	1.787	0.09182	0.3882
KamA	CMF-EHF	0.6796	0.2442	2.783	0.01275	0.2999
KamD	CMF-EHF	0.5419	0.2678	2.023	0.05908	0.3882
Ptb	CMF-EHF	0.7656	0.3145	2.434	0.02623	0.2999

### 3 lysine pathway

There are eight genes in the lysine pathway and we call **present** only when at least four genes were detected from the metagenomics data.



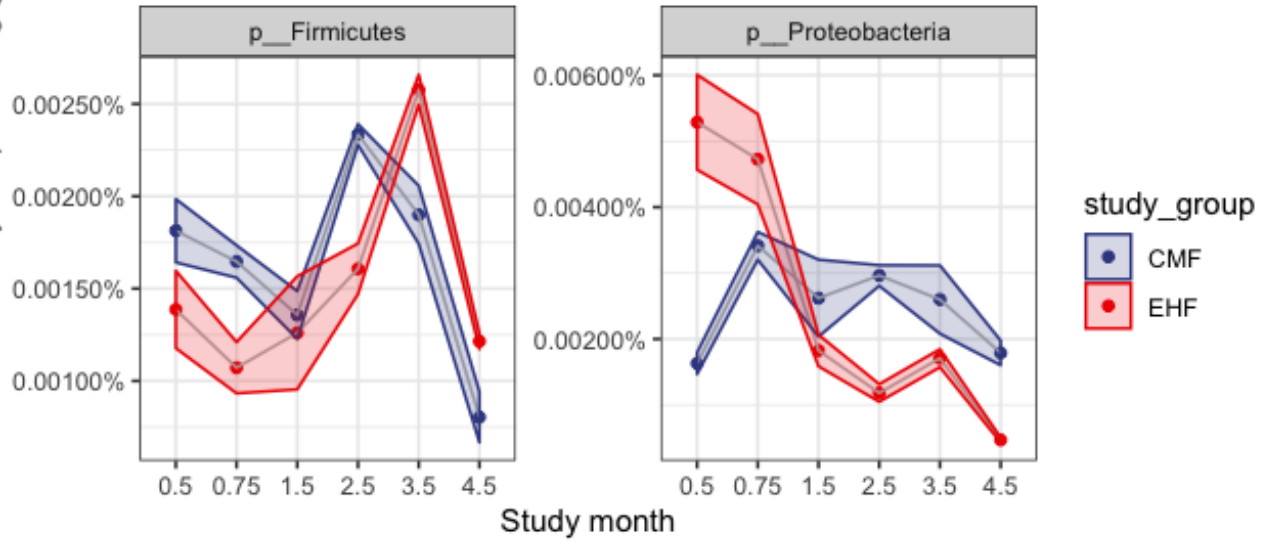
butyrate producing gene abundance in Lysine pathway





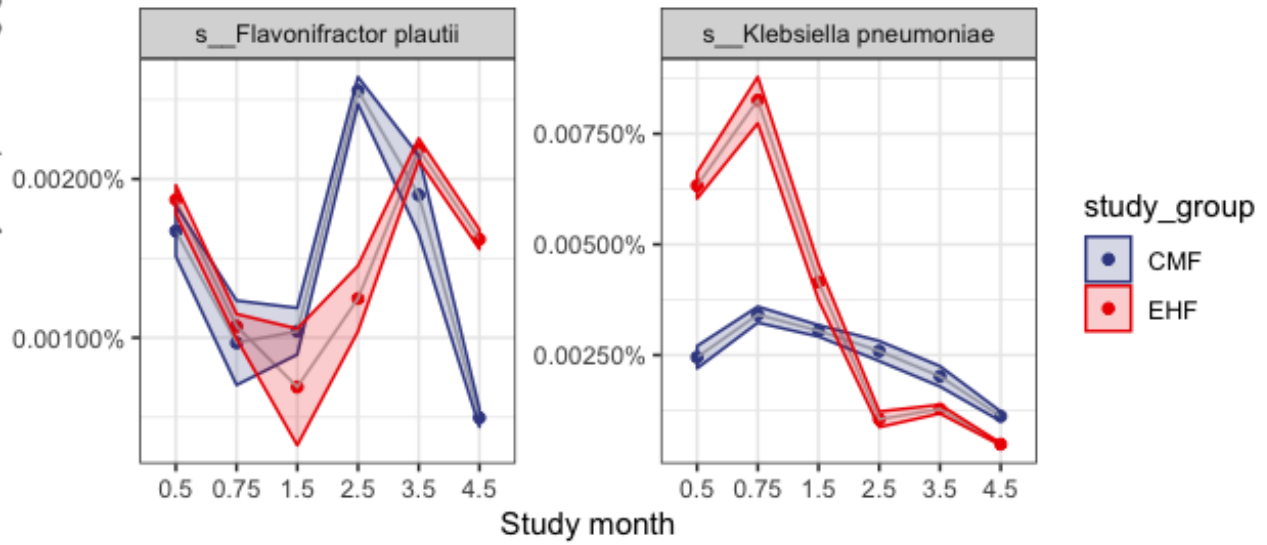
relative abundance of butyrate producing gene

### butyrate producing pathway abundance



relative abundance of butyrate producing gene

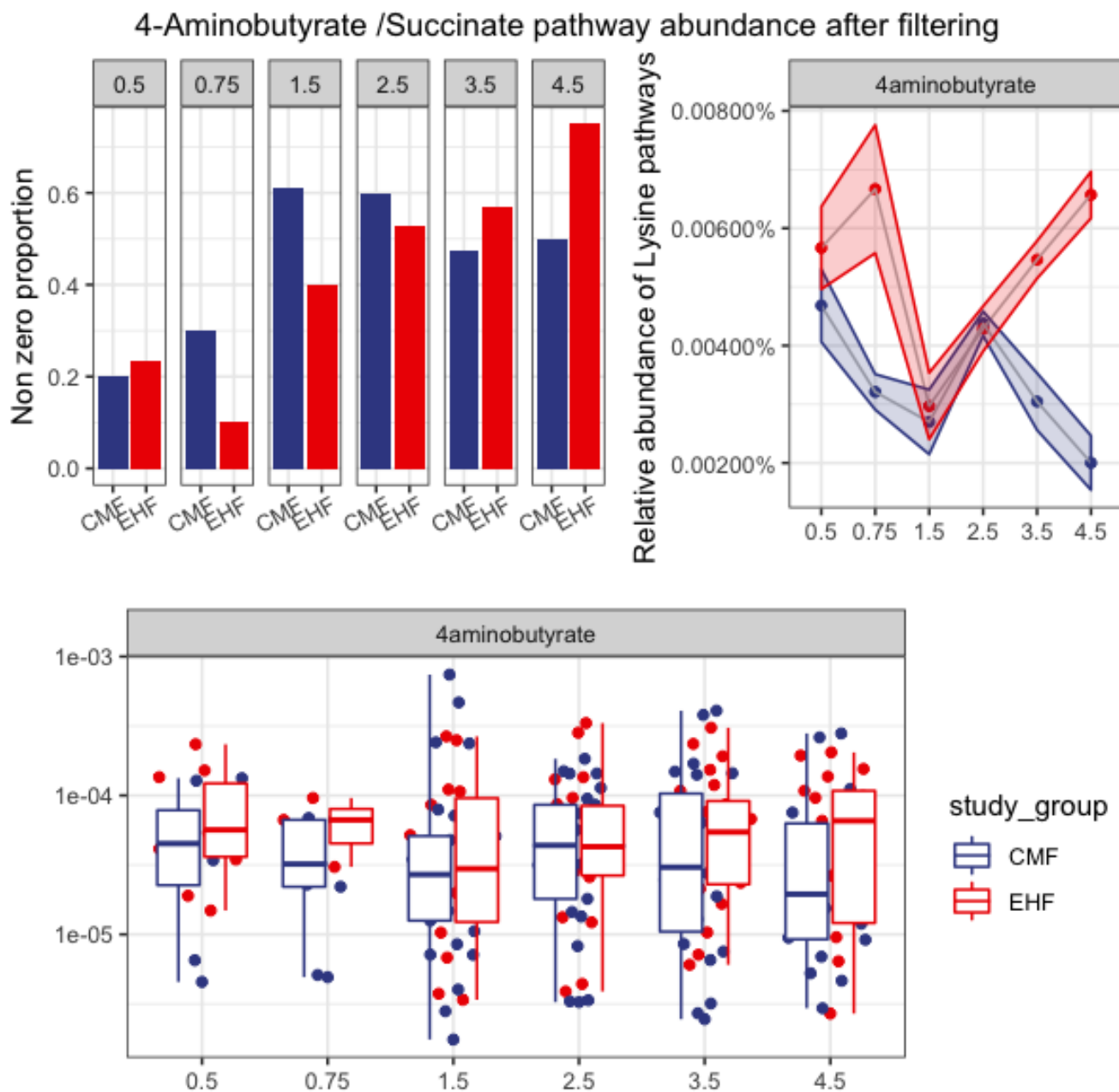
### butyrate producing pathway abundance





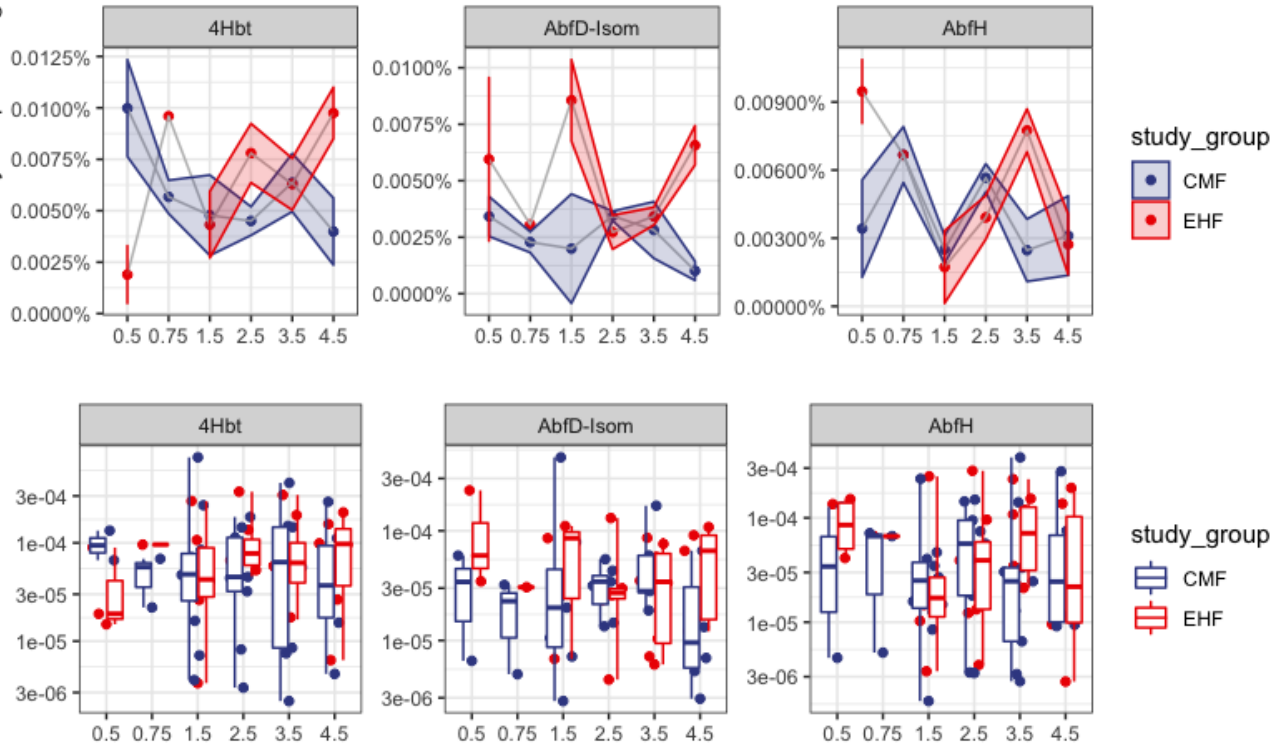
## 4 TCA cycle pathway

There are three genes in the lysine pathway and we call **present** only when at least **two** genes were detected from the metagenomics data.

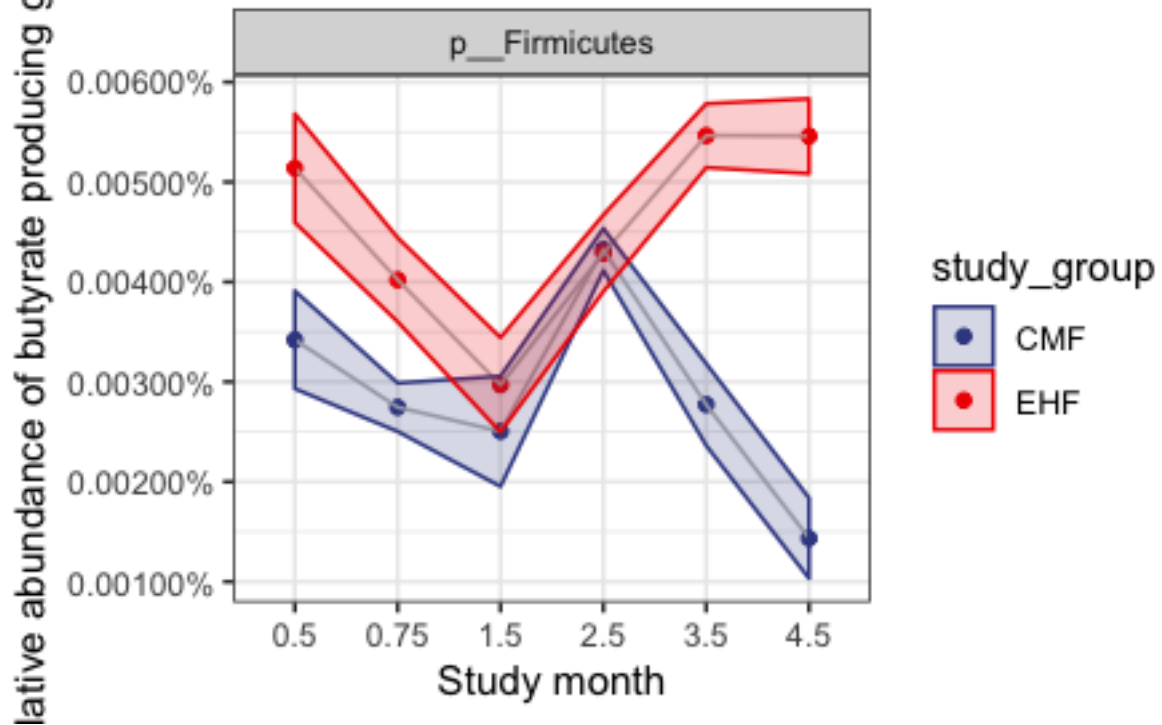


abundance of butyrate producing pat

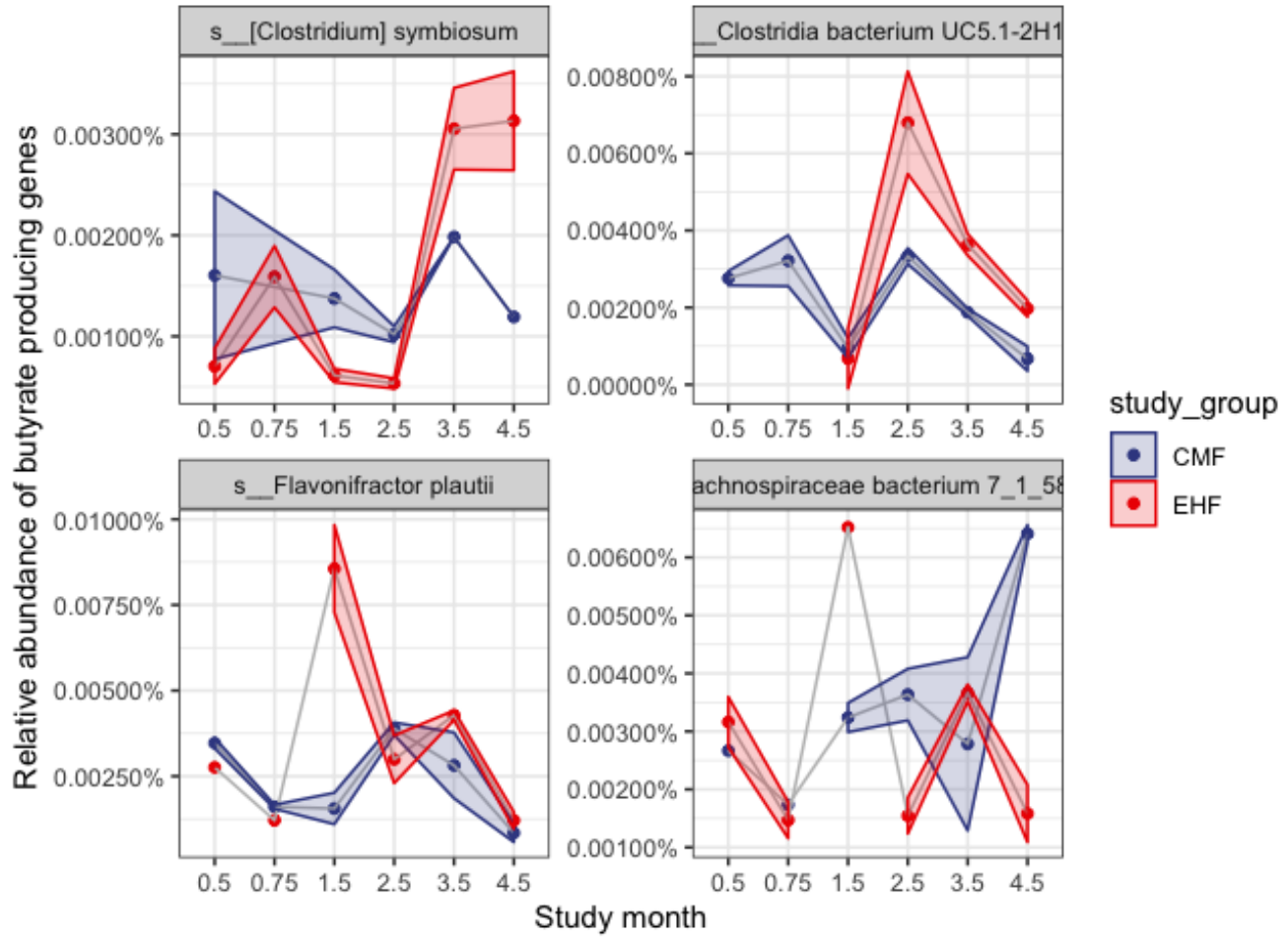
### butyrate producing gene abundance in Succinate pathway



### butyrate producing Succinate pathway abundance



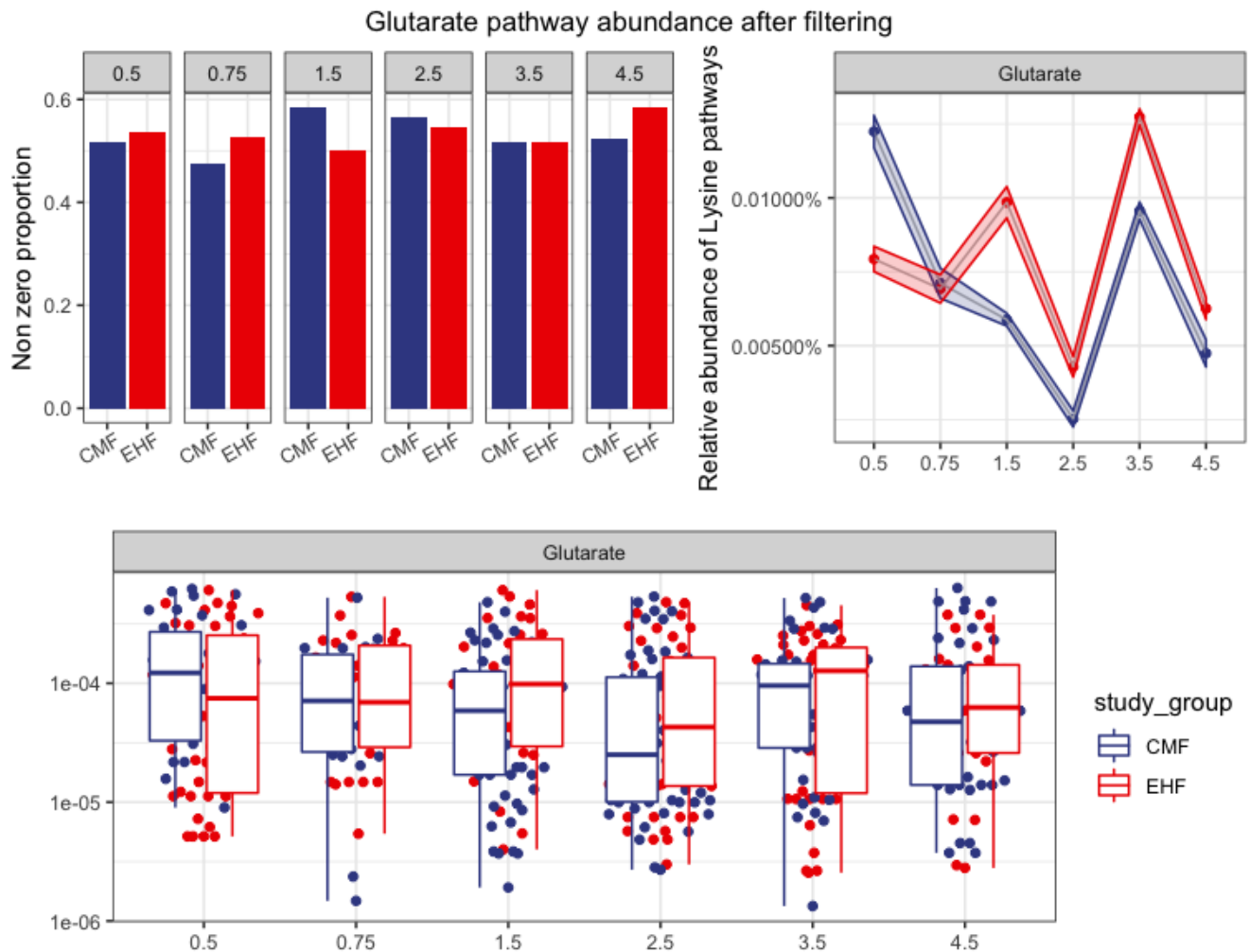
## butyrate producing Succinate pathway abundance



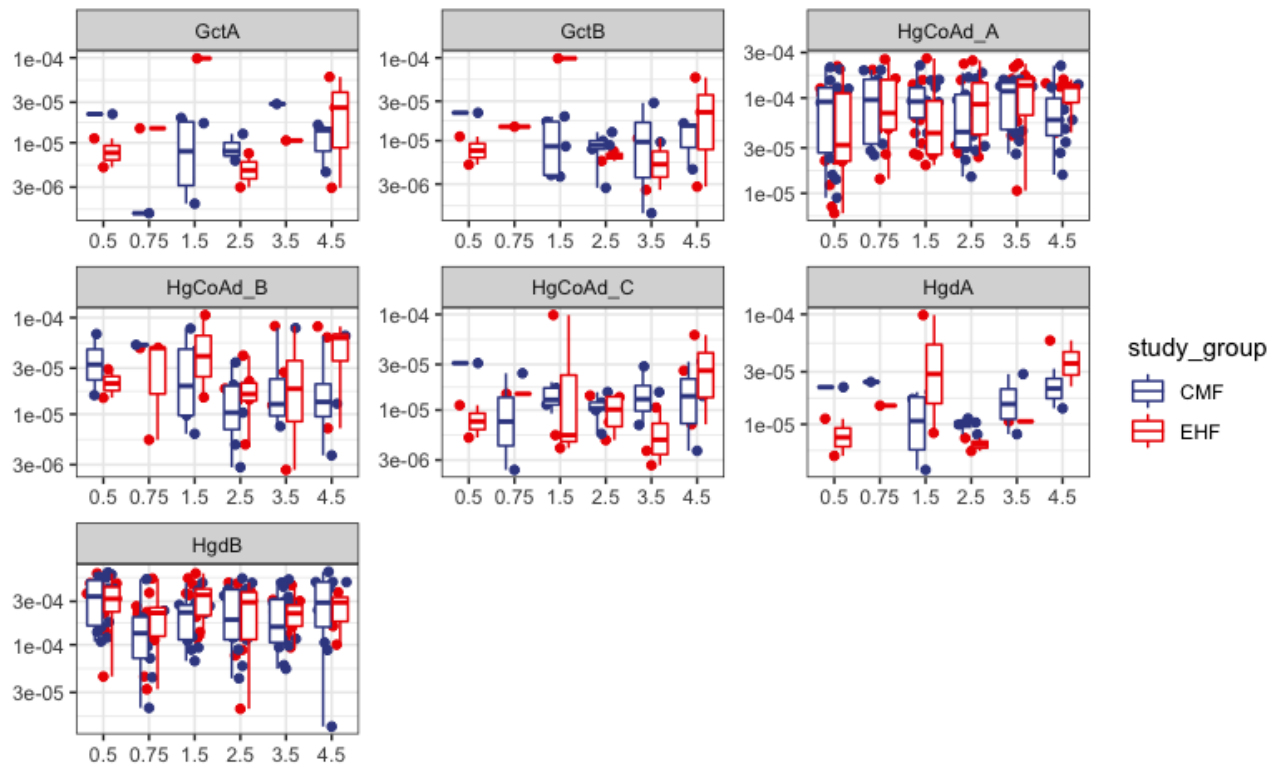
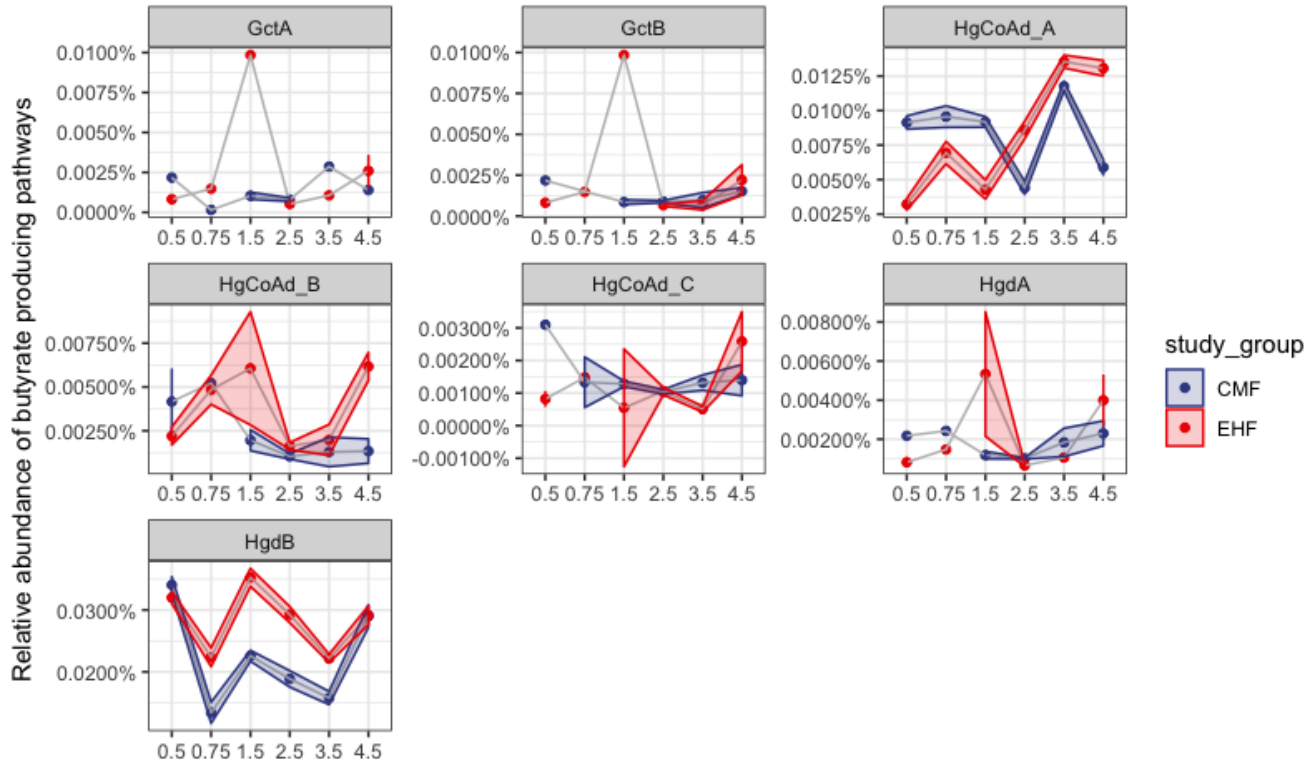
## 5 Glutarate pathway

There are eight genes in the lysine pathway and we call **present** only when at least **four** genes were detected from the metagenomics data.

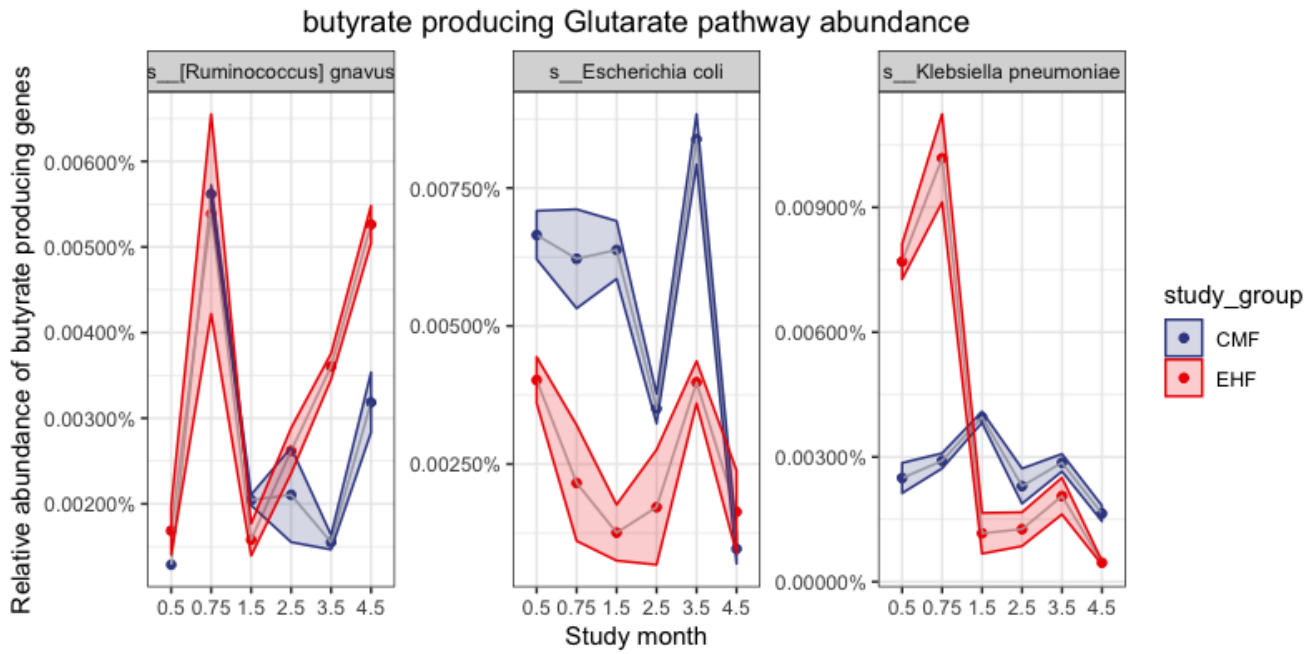
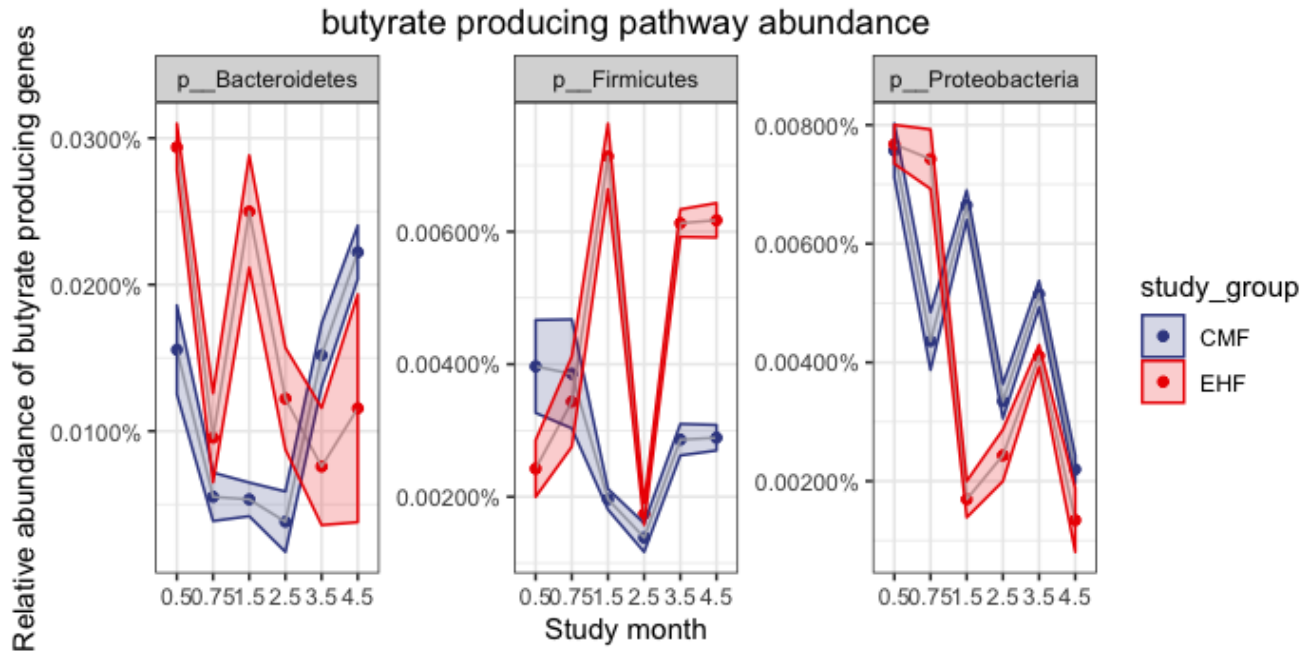
```
## # A tibble: 7 x 2
##   pathway_name gene_name
##   <chr>        <chr>
## 1 Glutarate    GctA
## 2 Glutarate    GctB
## 3 Glutarate    HgCoAd_A
## 4 Glutarate    HgCoAd_B
## 5 Glutarate    HgCoAd_C
## 6 Glutarate    HgdA
## 7 Glutarate    HgdB
```



# butyrate producing gene abundance in Glutarate pathway



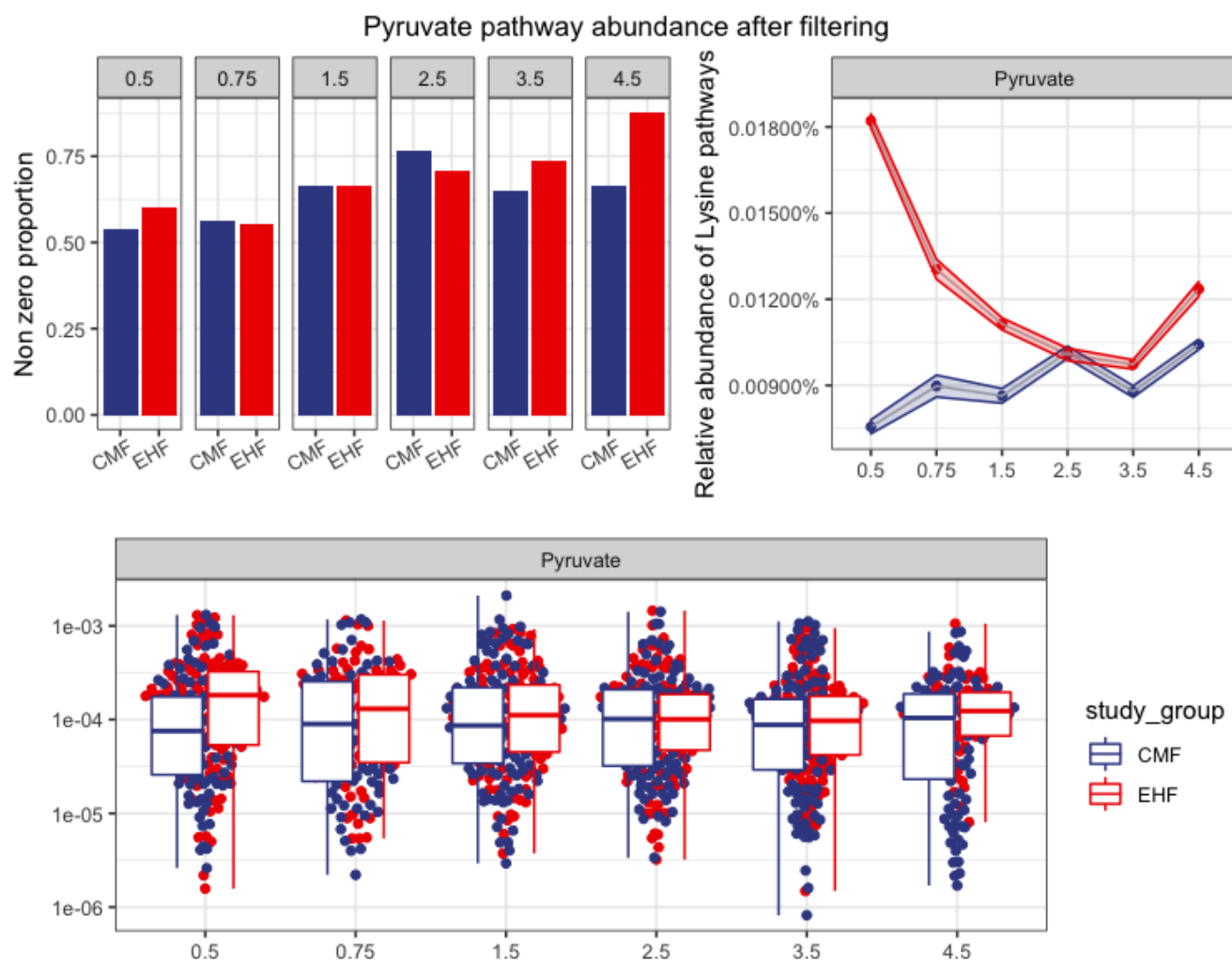




## 6 Pyruvate

There are eight genes in the Acetyl-CoA pathway and we call **present** only when at least **four** genes were detected from the metagenomics data.





# butyrate producing gene abundance in Pyruvate pathway

Relative abundance of butyrate producing pathways

