Challenges of 16S rRNA gene analysis in Chinese Holstein cows under heat stress conditions

B. Czech[1], K. Wang[2], S. Chen[2], Y. Wang[2], **J. Szyda**[1,3]

1. Biostatistics Group, Dep. of Genetics, Wroclaw University of Environmental and Life Sci, Wroclaw, Poland
2. College of Animal Science and Technology, China Agricultural University, Beijing, China
3. National Research Institute of Animal Production, Balice, Poland

Goal

- Assessment of the impact of normalisation and estimation methods on the significance of microbiome on heat stress

Data & Methods

Microbiome data

- 16S rRNA gene → v3 & v4 regions → faecal samples → 136 Holstein-Friesian cows
- Quantitative outcome → rectal temperature
- Analysis on genus level

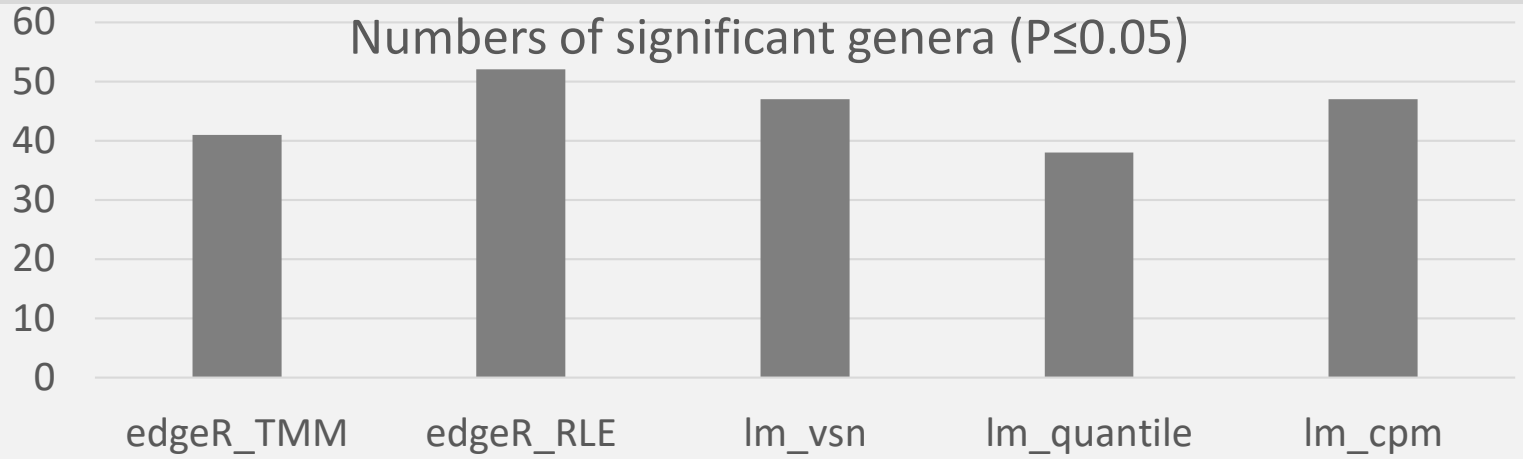
Normalisation

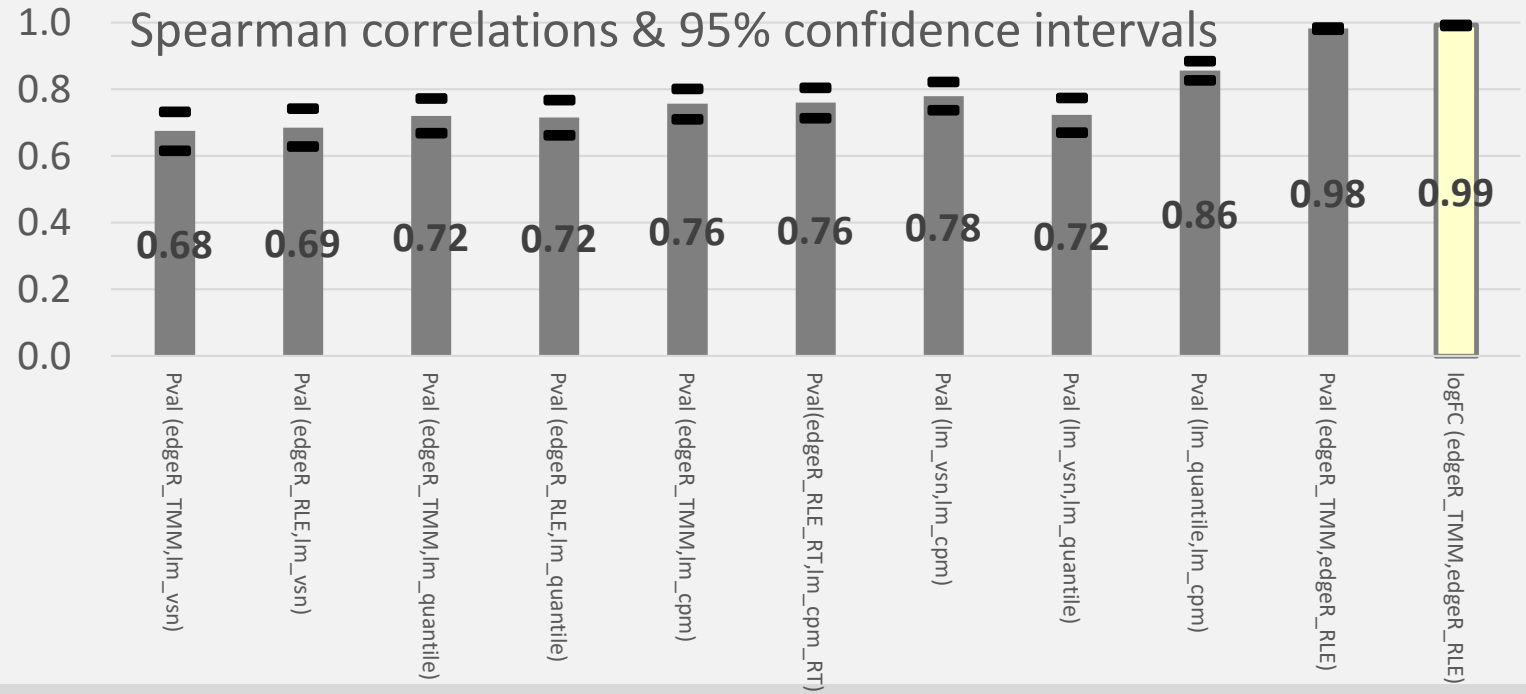
- Trimmed Mean of M values (TMM)
- Relative Log Expression (RLE)
- Quantile-based (quantile)
- Variance stabilizing (VSN)
- Counts per million (CPM)

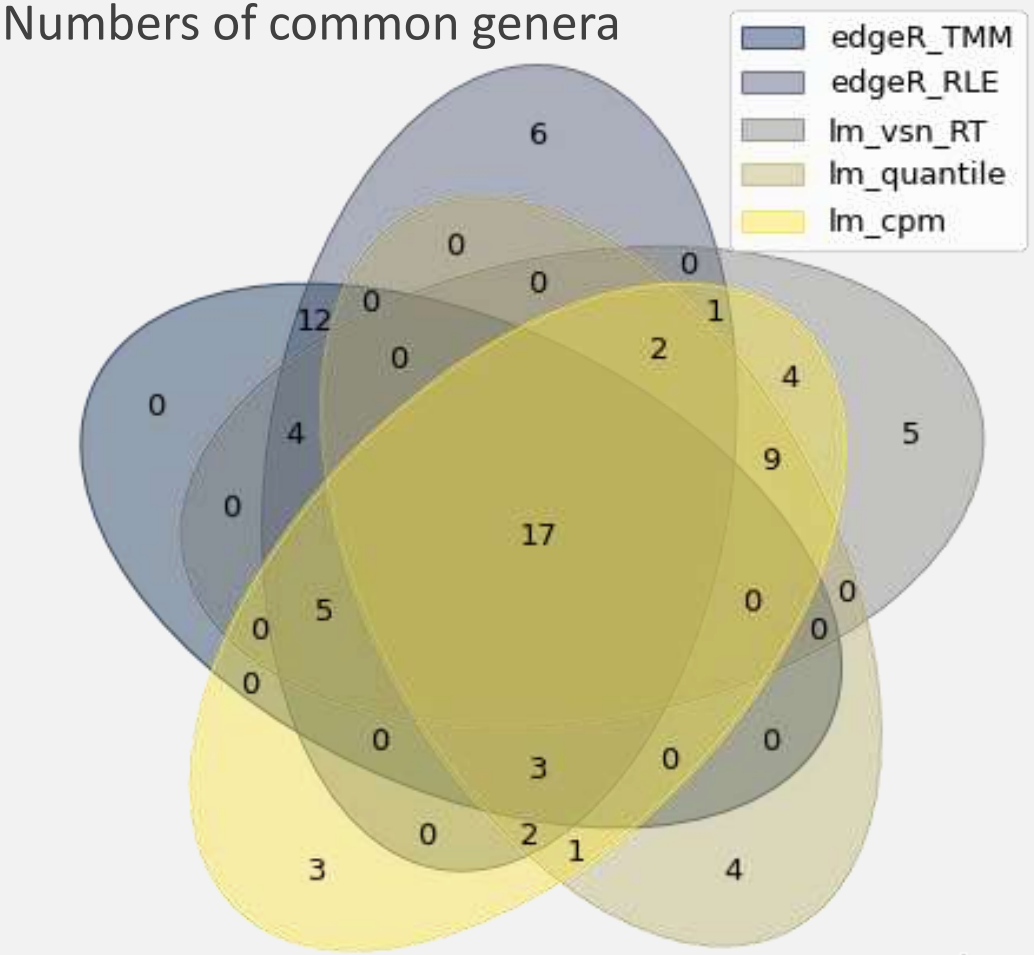
Effect estimation and testing

- Negative binomial regression → implemented via edgeR
- Linear Gaussian regression → implemented in via R lm function

Results


Numbers of significant genera (P≤0.05)
60
50
40
30
20
10
0
edgeR_TMM
edgeR_RLE
lm_vsn
lm_quantile
lm_cpm


Spearman correlations & 95% confidence intervals
1.0
0.8
0.6
0.4
0.2
0.0
0.68
0.69
0.72
0.72
0.76
0.76
0.78
0.72
0.86
0.98
0.99
Pval (edgeR_TMM,lm_vsn)
Pval (edgeR_RLE,lm_vsn)
Pval (edgeR_TMM,lm_quantile)
Pval (edgeR_RLE,lm_quantile)
Pval (edgeR_TMM,lm_cpm)
Pval(edgeR_RLE_RT,lm_cpm_RT)
Pval (lm_vsn,lm_cpm)
Pval (lm_vsn,lm_quantile)
Pval (lm_quantile,lm_cpm)
Pval (edgeR_TMM,edgeR_RLE)
logFC (edgeR_TMM,edgeR_RLE)


Numbers of common genera
edgeR_TMM
edgeR_RLE
lm_vsn_RT
lm_quantile
lm_cpm
6
0
0
0
12
0
1
2
4
0
4
5
9
0
17
0
5
0
0
0
0
0
0
0
3
0
0
0
2
1
3
4

Conclusions

- Unfortunately → Different results → different analytical approaches
- Fortunately → some overlap
- ≈ 50% of significant genera common to all analytical constellations
- Negative binomial model (edge R) more robust than linear model