Ruminal bacteria heritabilities: Which is the impact of different data processing ?

Guillermo Martinez Boggio, Annabelle Meynadier, Charlotte Allain, Christel Marie-Etancelin

▶ **To cite this version:**

Guillermo Martinez Boggio, Annabelle Meynadier, Charlotte Allain, Christel Marie-Etancelin. Ruminal bacteria heritabilities: Which is the impact of different data processing ?. EAAP 2021, Aug 2021, Davos, Switzerland. hal-03480858

HAL Id: hal-03480858
https://hal.inrae.fr/hal-03480858

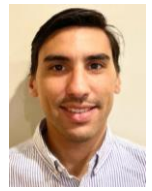
Submitted on 14 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ruminal bacteria heritabilities: Which is the impact of different data processing ?

Guillermo Martinez Boggio[1], Annabelle Meynadier[1], Charlotte Allain[2], Christel Marie-Etancelin[1]
[1]GenPhySE, Université de Toulouse, INRAE, ENVT, France. [2]INRAE, Experimental Unit of La Fage, France



guillermo.martinez-boggio@inrae.fr

Introduction and objective

Microbiome data is **compositional**: the information is found in the operational taxonomic unit (OTU) **ratios**.

As high-dimensional data with zeros, it is necessary to replace the zeros, and apply value transformation.

Conclusions

Ruminal bacteria h^2 are affected by data processing.

Zeros replacement impact the most the number of heritable OTUs.

GBM leads to less difference between the two data transformation.

To compare **data processing methods** in terms of **heritability** estimates (h^2).

Materials and methods

Rumen samples of 795 Lacaune ewes
Metabarcoding 16s rRNA gene

Table 1. *Data processing methods to be compare.*

	GBM	C1	Z01
CLR	GBM-CLR	C1-CLR	Z01-CLR
Log	GBM-Log	C1-Log	Z01-Log

Zeros replacement:
GBM= Geometric Bayesian Method
C1= Add constant equal 1
Z01= Replace only zeros with 0.001

Data transformation:
CLR= Centred Log-Ratio transformation
Log= Logarithm transformation

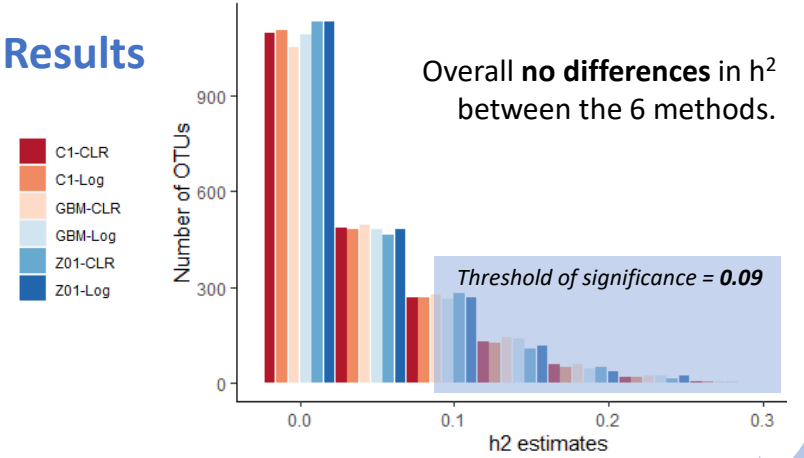
Heritability estimates*:

$$y = Xb + Za + e$$

y: OTU abundances processed with each of the 6 methods.
b: fixed effects lactation stage, litter size, year, number of lactation, sequencing run, sampling order and total nb. sequences.
a: random animal effect ~ $N(0, \sigma^2_a A)$.
e: residual ~ $N(0, \sigma^2_e I)$.

**The threshold of significance obtained by phenotype-animal permutation*

Results



Overall **no differences** in h^2 between the 6 methods.

But, the number of **heritable OTUs** defined by the methods were **variable**.

Table 2. *Number of heritable* OTUs by processing data method.*

	GBM	C1	Z01
CLR	**326**	**293**	**249**
Log	**294**	**274**	**254**

GBM > C1 > Z01

By **replacing only the zeros** we obtain lower number of heritable OTUs.

% OTUs in common between CLR and Log

78% **72%**

Using **GBM** the differences between CLR and Log are lesser than with C1.