Multi-Omic

Christian Maltecca

Multiomics

- **Definition:**
  Multiomics is an integrative approach that combines data from multiple "omics" fields (genomics, transcriptomics, proteomics, metabolomics, etc.) to provide a comprehensive understanding of biological systems.

- **Goal:**
- To gain understanding of molecular mechanisms and interactions within a cell or organism.
- In animal breeding context also an attempt to improve prediction accuracy or better account for additional variation

1. **Genomics:** Study of an organism's complete set of DNA (genome).
2. **Transcriptomics:** Study of RNA transcripts produced by the genome.
3. **Proteomics:** Study of the full set of proteins encoded by the genome.
4. **Metabolomics:** Study of the chemical processes involving metabolites.
5. **Epigenomics:** Study of chemical modifications on DNA and histone proteins affecting gene expression without altering the DNA sequence.
6. **Lipidomics:** Study of cellular lipids and their role in metabolism and cell signaling.
7. **Metagenomics:** Is the study of genetic material from entire microbial communities directly from environmental samples, revealing their diversity and functions.

- Can be seen as either additional source of information in modeling phenotypic variation
- Or can be seen as intermediate phenotypes closing the gap between genome and phenome
- Or in some cases can be seen as a direct target of selection

- Most of them have some common data-structure because they arise from NGS

Illumina Sequencing: Uses sequencing by synthesis with reversible terminator bases, offering high accuracy and throughput for a wide range of applications.
Ion Torrent Sequencing: Detects DNA sequence by measuring changes in pH as nucleotides are incorporated, suitable for targeted sequencing and smaller genomes.
PacBio (Pacific Biosciences) Single Molecule Real-Time (SMRT) Sequencing: Utilizes real-time observation of DNA polymerase activity, enabling long-read sequencing with high consensus accuracy.
Oxford Nanopore Sequencing: Reads DNA or RNA molecules as they pass through a nanopore, providing long reads and real-time data for real-time analysis and flexibility in read length.
SOLiD (Sequencing by Oligonucleotide Ligation and Detection) Sequencing: Uses ligation-based sequencing, known for high accuracy in sequencing by repeated rounds of ligation and imaging.
BGI/MGI Sequencing: Employs DNA nanoballs and combinatorial probe-anchor synthesis, providing high throughput and cost-effectiveness for a variety of sequencing applications.

- After some (lengthy and sometime perilous bioinformatics) are normally represent as a table of counts per feature (akin to what we have seen for the SNP)

- The advantage of this structure is that it is straghtforward to extend the machinery we have devolped for BLUP and GBLUP to other technologies, since we can always obtain a matrix product $\mathbf{WW}^{\mathbf{T}}$ based on a matrix $\mathbf{W}$ of n individuals by m features

$$\begin{bmatrix} \mathbf{X}^\top\mathbf{X} & \mathbf{X}^\top\mathbf{Z} \\ \mathbf{Z}^\top\mathbf{X} & \mathbf{Z}^\top\mathbf{Z} + \text{Etc.}^{-1}\lambda \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top\mathbf{y} \\ \mathbf{Z}^\top\mathbf{y} \end{bmatrix}$$

JOURNAL ARTICLE

Leveraging Multiple Layers of Data To Predict *Drosophila* Complex Traits

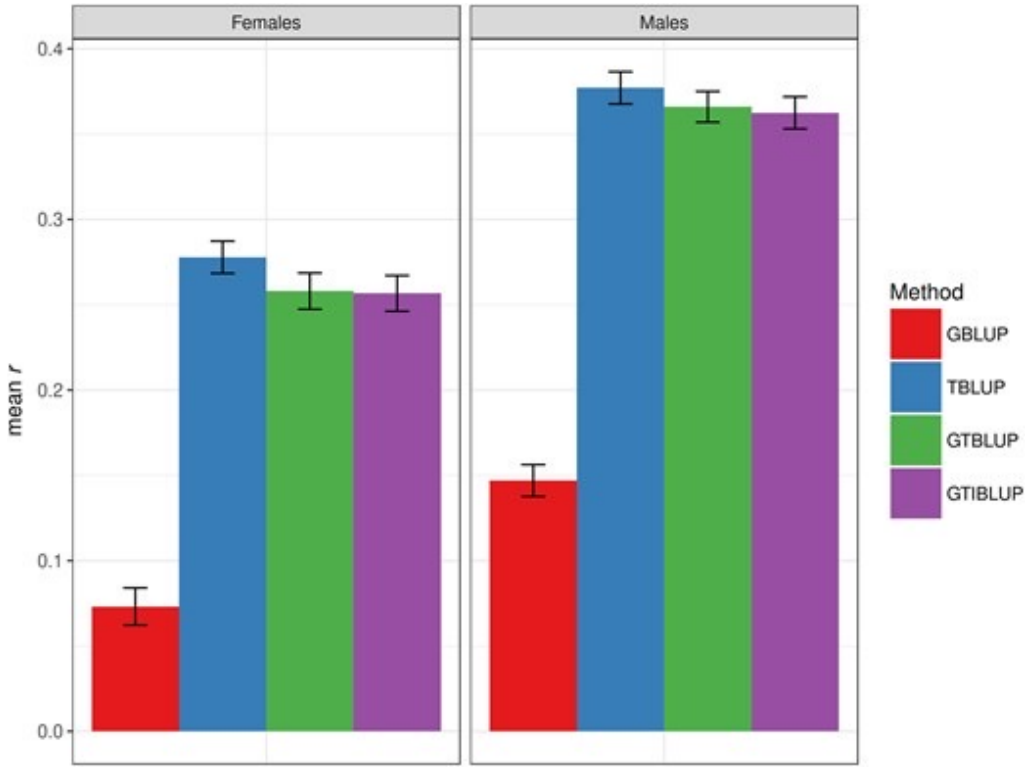
Fabio Morgante ✉, Wen Huang, Peter Sørensen, Christian Maltecca, Trudy F C Mackay

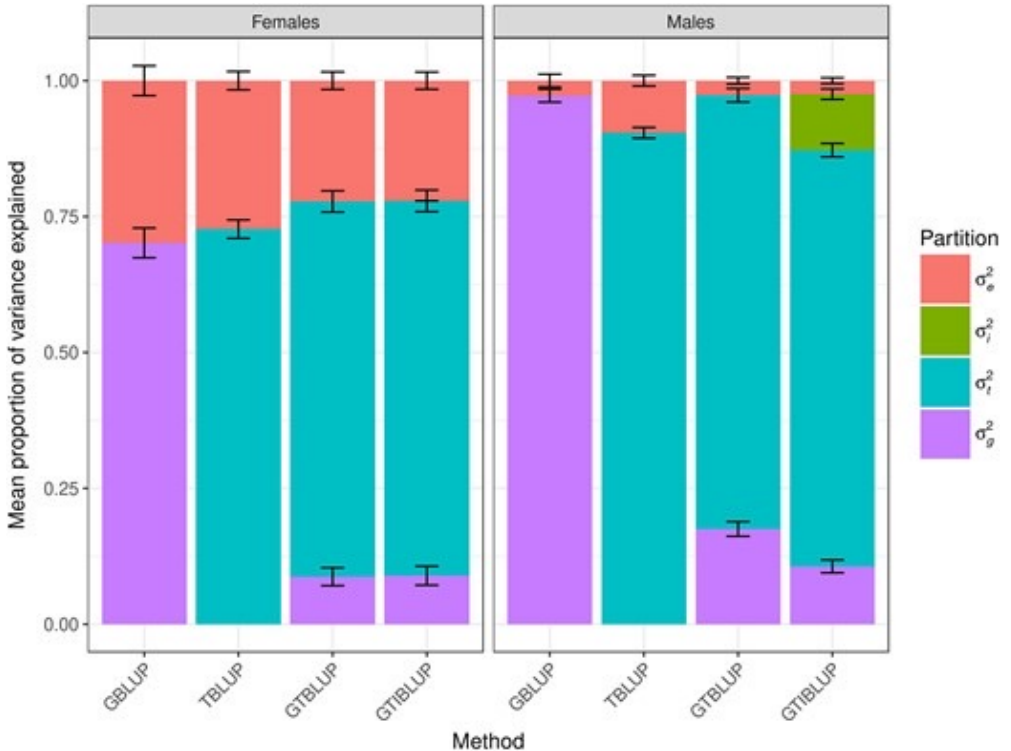
Author Notes

G3 Genes|Genomes|Genetics, Volume 10, Issue 12, 1 December 2020, Pages 4599–4613,

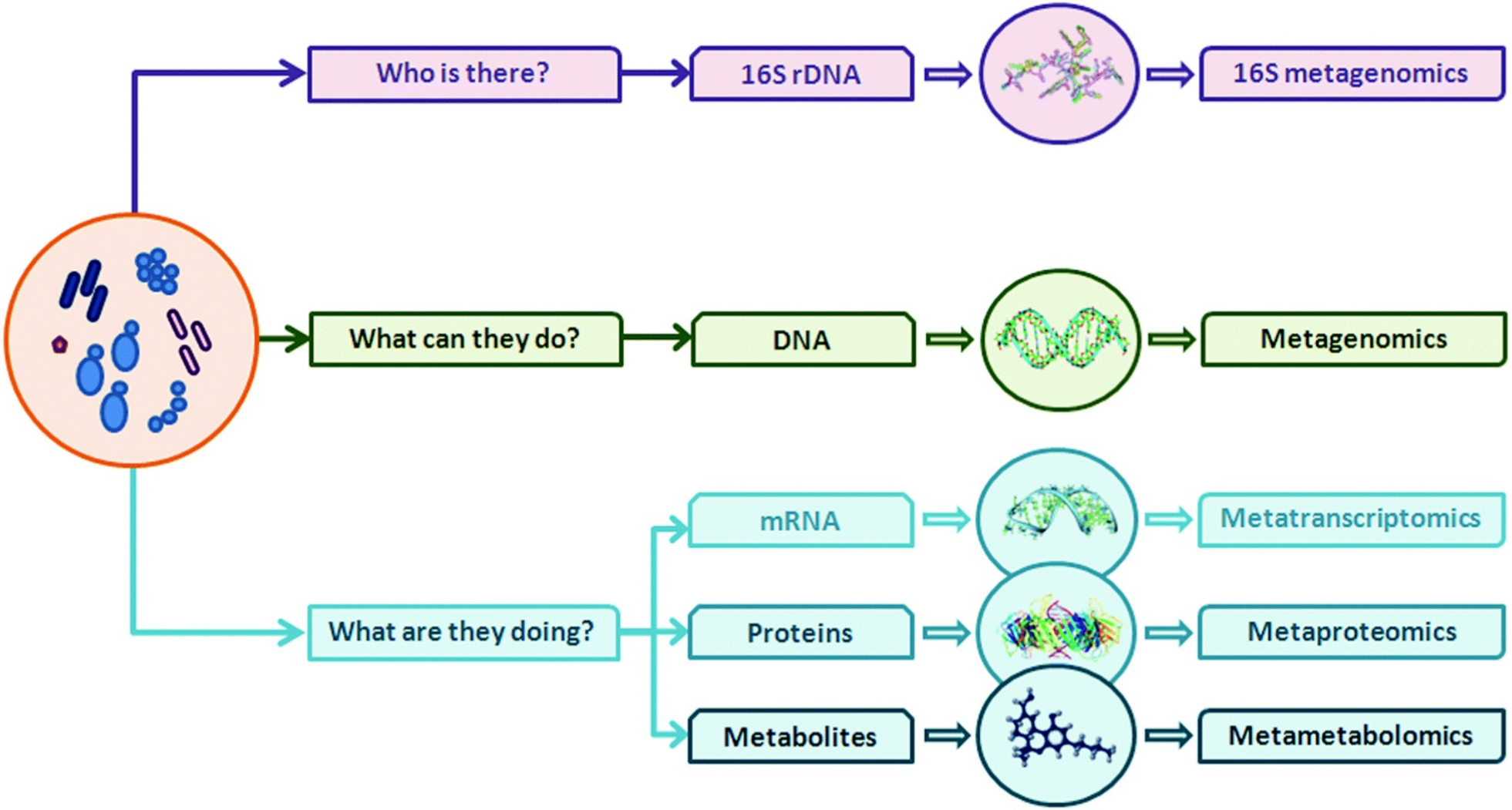
https://doi.org/10.1534/g3.120.401847

Published: 01 December 2020 **Article history** ▾





- Microbiome as a case example



Livestock Science
Volume 269, March 2023, 105171

Invited Review

Invited review: Novel methods and perspectives for modulating the rumen microbiome through selective breeding as a means to improve complex traits: Implications for methane emissions in cattle

O. González-Recio [a], M. Martínez-Álvaro [b], Francesco Tiezzi [c], A. Saborío-Montero [d e], C. Maltecca [f], R. Roehe [b]

INVITED REVIEW

The interaction between microbiome and pig efficiency: A review

Christian Maltecca, Matteo Bergamaschi, Francesco Tiezzi

First published: 01 October 2019 | https://doi.org/10.1111/jbg.12443 | Citations: 36

Chapter 18 Manipulation of gut microbiome composition by genetic selection of the host

In: Environmental effects on gut health in production animals

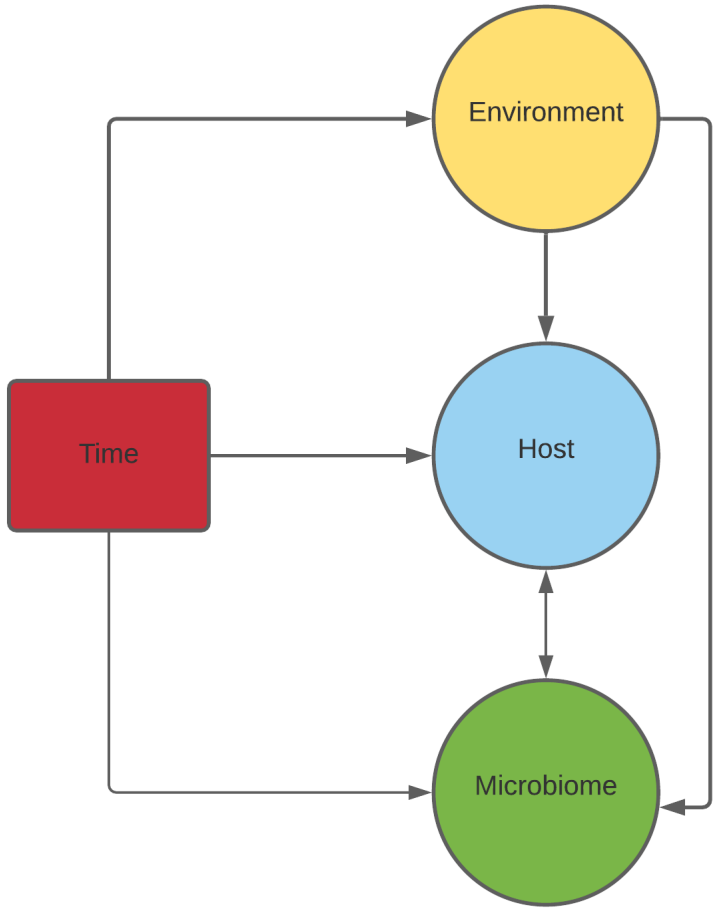
Authors: Francesco Tiezzi and Christian Maltecca

Type: Chapter
Pages: 459–487
DOI: https://doi.org/10.3920/9789004695467_019

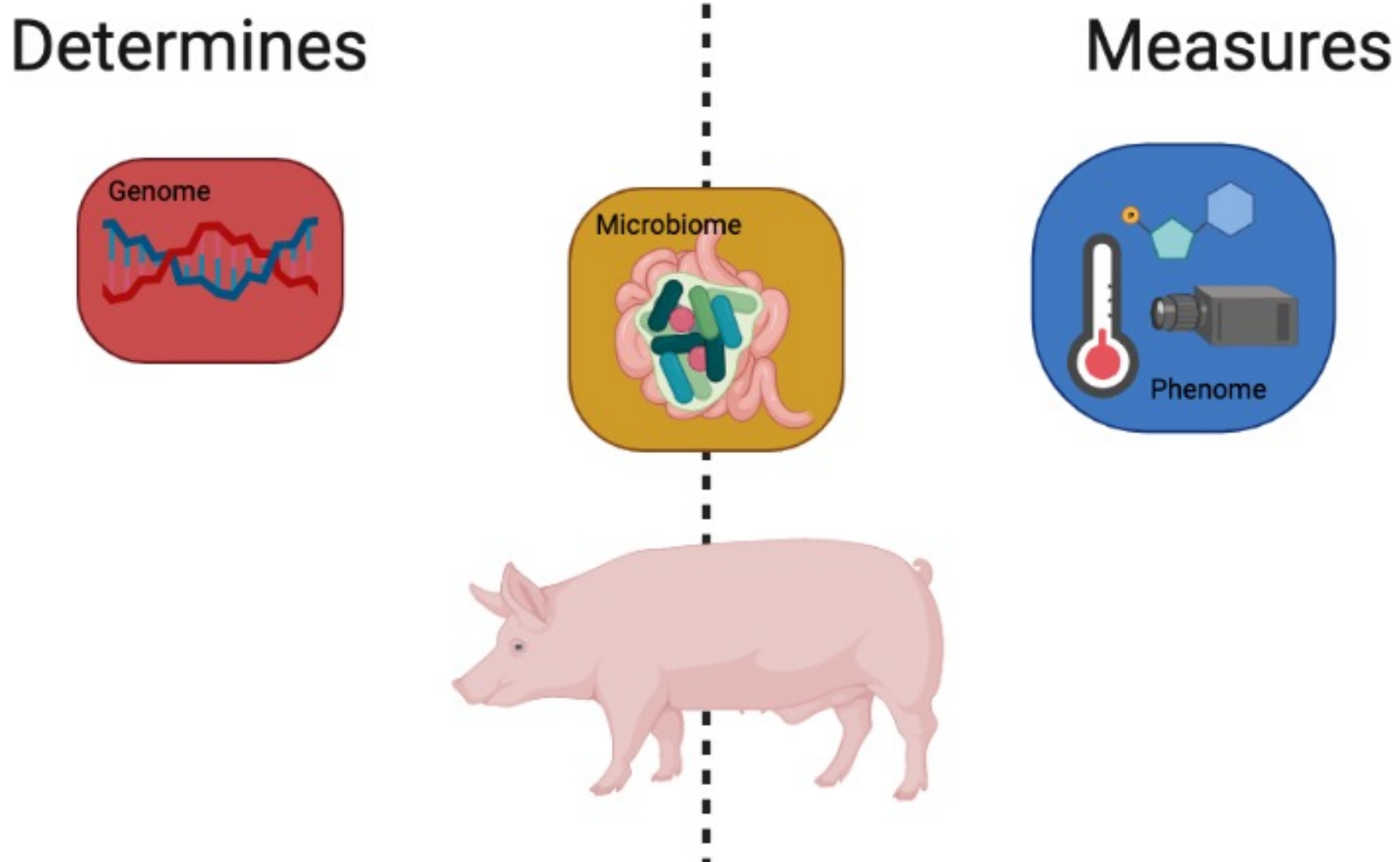
If the only tool you have is a hammer, it is tempting to treat everything as if it were a nail

Breeding while accounting for the interaction between Host Genome Microbiome Environment

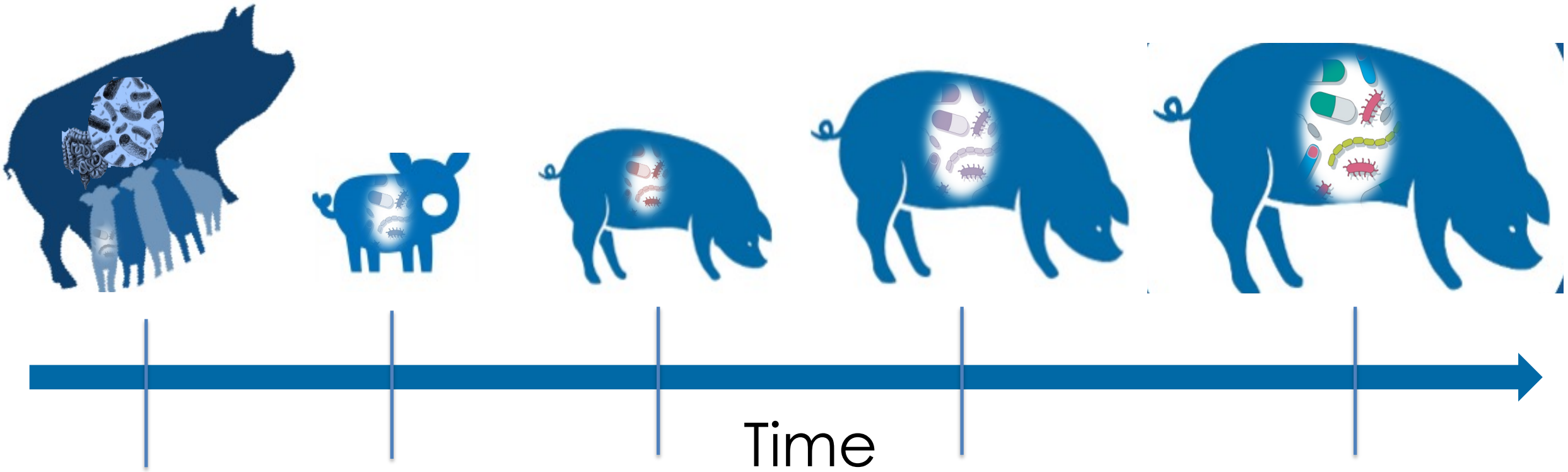
$$G + E + (G\ E) + M + (G\ M) + (E\ M) + (G\ E\ M)$$



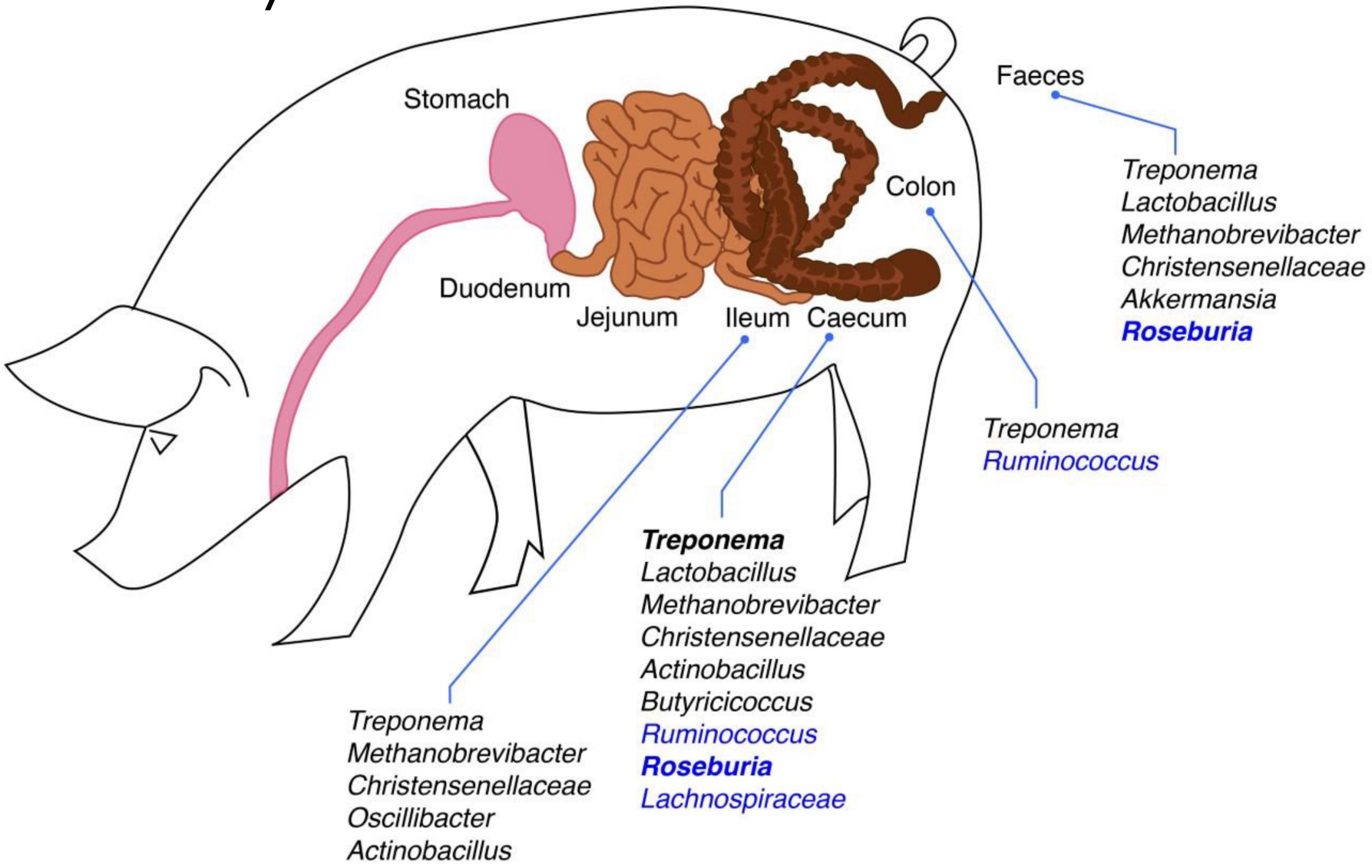
Microbiome is "Signal Dense"



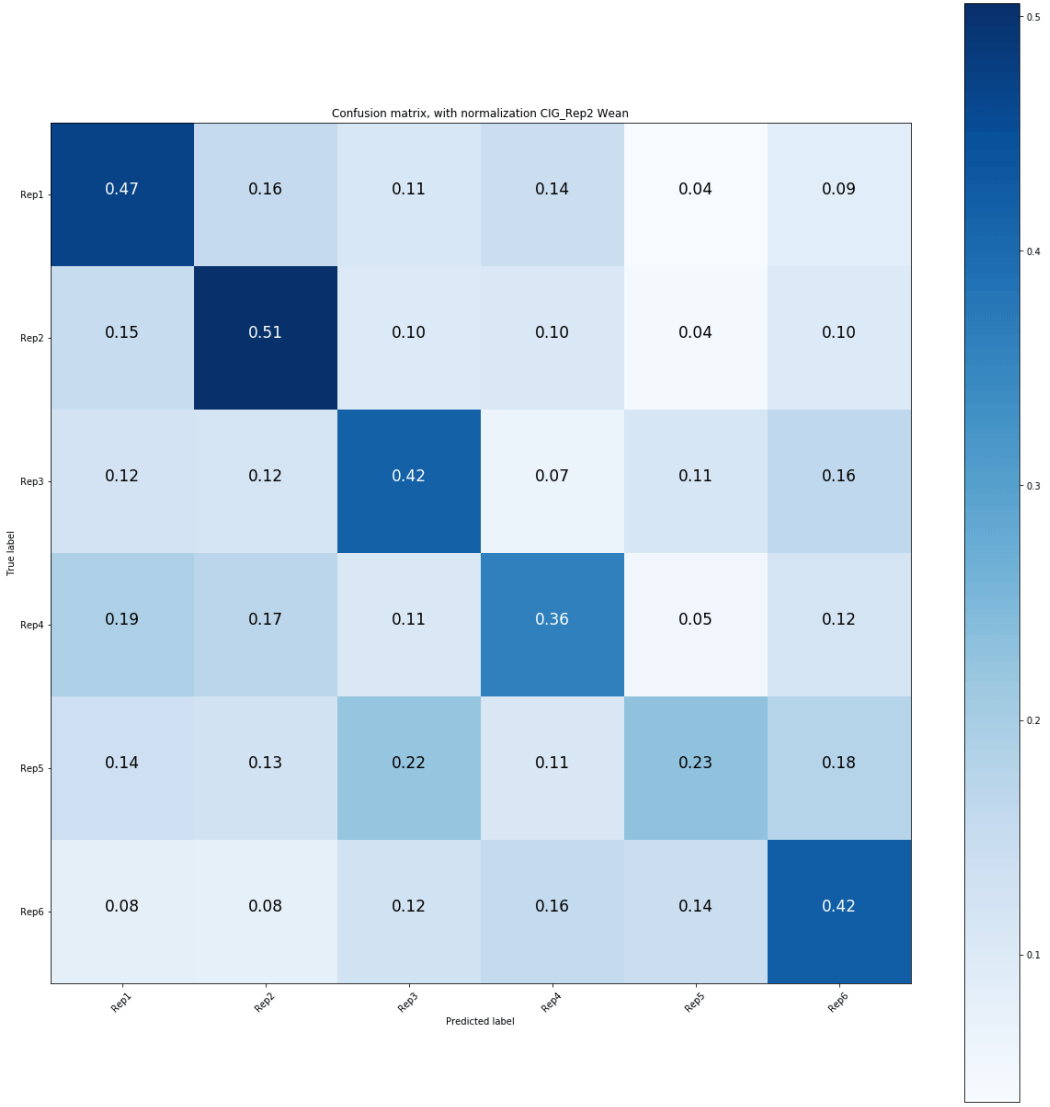
Longitudinal variability

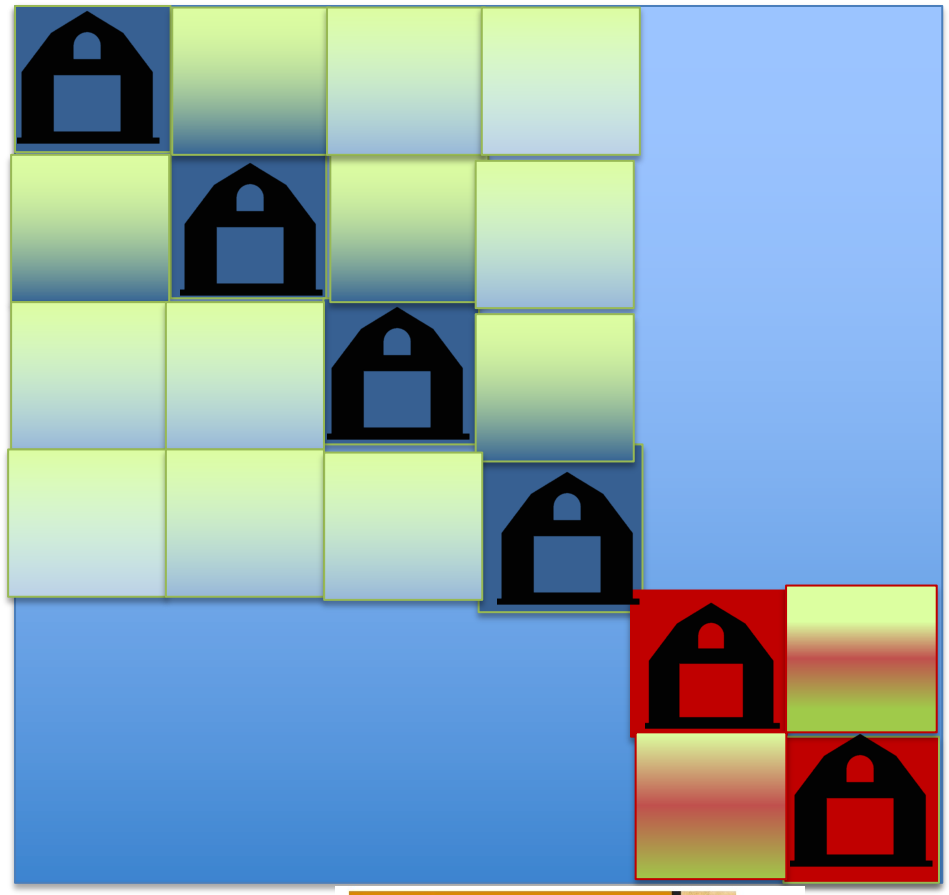


Spatial Variability



Spatial and temporal Variability





Journal of Animal Breeding and Genetics

INVITED REVIEW

The interaction between microbiome and pig efficiency: A review

Christian Maltecca, Matteo Bergamaschi, Francesco Tiezzi

First published: 01 October 2019 | https://doi.org/10.1111/jbg.12443 |

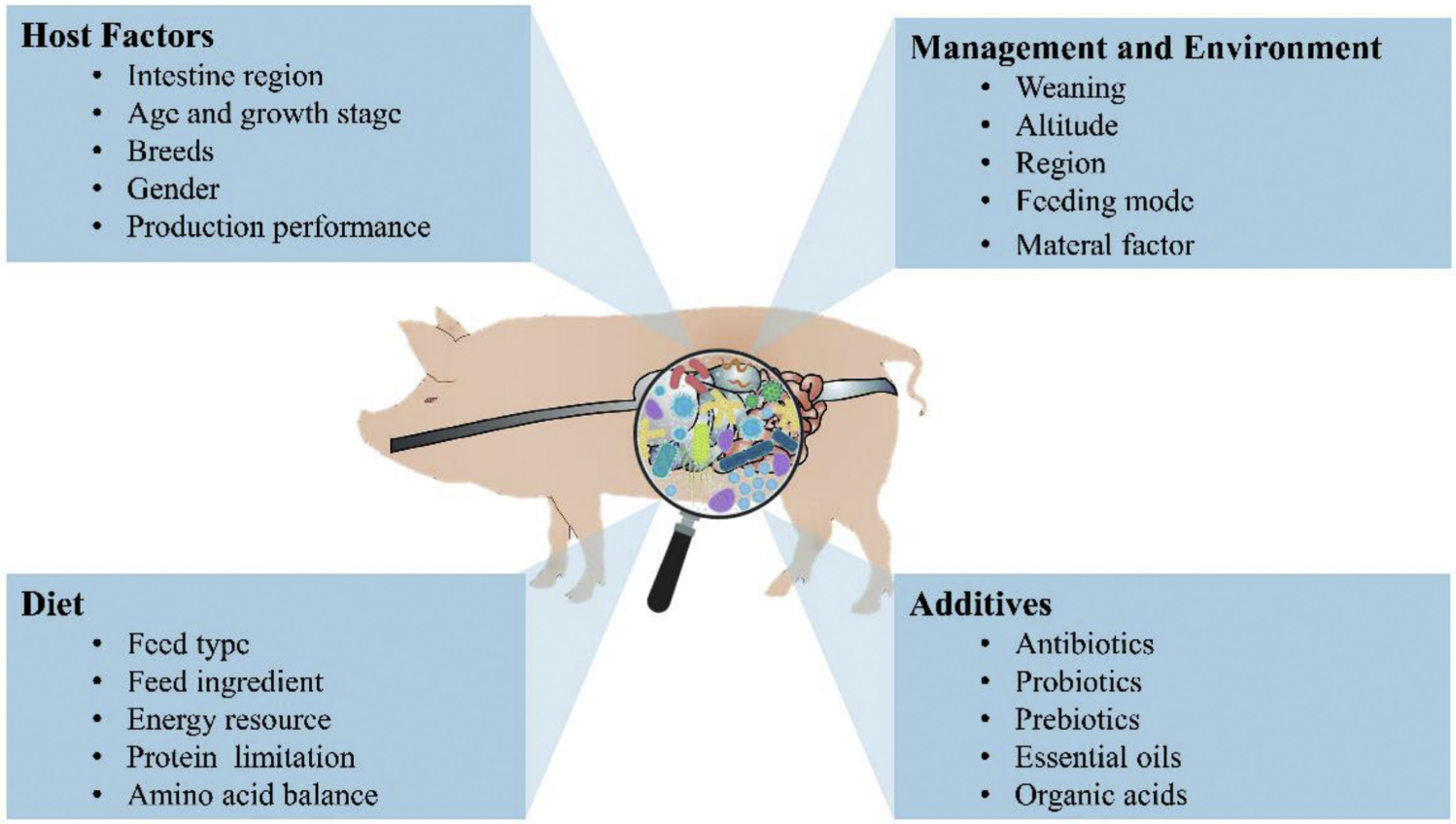
Research | Open Access | Published: 04 January 2018

Host contributes to longitudinal diversity of fecal microbiota in swine selected for lean growth

Duc Lu, Francesco Tiezzi, Constantino Schillebeeckx, Nathan P. McNulty, Clint Schwab, Caleb Shull & Christian Maltecca

Microbiome 6, Article number: 4 (2018) | Cite this article

3463 Accesses | 25 Citations | 16 Altmetric | Metrics



Hongyu Wang, Rongying Xu, He Zhang, Yong Su, Weiyun Zhu, Swine gut microbiota and its interaction with host nutrient metabolism, Animal Nutrition,

- What Microbiome measured with RNA 16S looks like

```
> M[1:10,1:10]
              S1_m37981    S1_m2462    S1_m2464    S1_m2466    S1_m2469    S1_m3732    S1_m3730     S1_m3735   S1_m12017    S1_m2929
GLDS_246469  1.58678767  1.58678767  0.74591747  1.30763592  0.3150390  0.60775331  1.40205477  0.015087119  0.63158829  0.37459862
GLDS_246470  1.50030975  1.50030975  0.70191963  1.23568450  0.2907513  0.57132880 -0.04347259  0.004010025  0.59398139  0.34869026
GLDS_246471 -0.04958071 -0.04958071 -0.08662655 -0.05385650 -0.1445418 -0.08148515  1.09311844 -0.194517910 -0.08002377 -0.11564941
GLDS_246472 -0.07408433 -0.07408433 -0.09909339 -0.07424403 -0.1514238 -0.09180608 -0.35172545 -0.197656618 -0.09067973 -0.12299057
GLDS_246473 -0.74811544 -0.74811544 -0.44202386 -0.63505189 -0.3407282 -0.37570802 -0.48369496 -0.283994336 -0.38379751 -0.32492703
GLDS_246479 -0.84201460 -0.84201460 -0.48979744 -0.71317793 -0.3671002 -0.41525835 -0.50207961 -0.296022028 -0.42463170 -0.35305876
GLDS_246482  0.06082816  0.06082816 -0.03045323  0.03800597 -0.1135330 -0.03498093 -0.32531074 -0.180375462 -0.03200996 -0.08257144
GLDS_246485 -1.08151278 -1.08151278 -0.61164823 -0.91244539 -0.4343642 -0.51613500 -0.54897130 -0.326699731 -0.52878294 -0.42481125
GLDS_246486 -0.61548582 -0.61548582 -0.37454516 -0.52470129 -0.3034786 -0.31984442 -0.45772722 -0.267005596 -0.32612050 -0.28519183
GLDS_246487 -1.23919744 -1.23919744 -0.69187431 -1.04364230 -0.4786506 -0.58255179 -0.57984461 -0.346897810 -0.59735571 -0.47205281
> |
```

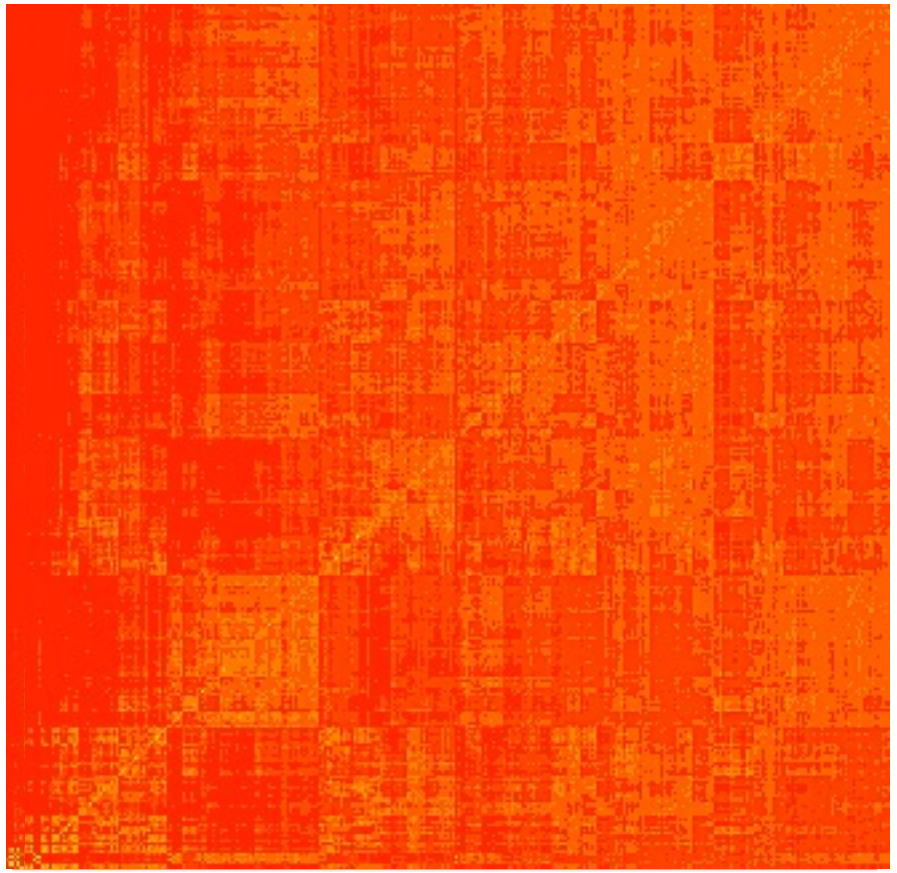
- After some massaging needed due to the nature of the datapoints
  - Compositional
  - 0-Inflated
  - …

Will just pretend everything is fine here

Exploring methods to summarize gut microbiota composition for microbiability estimation and phenotypic prediction in swine

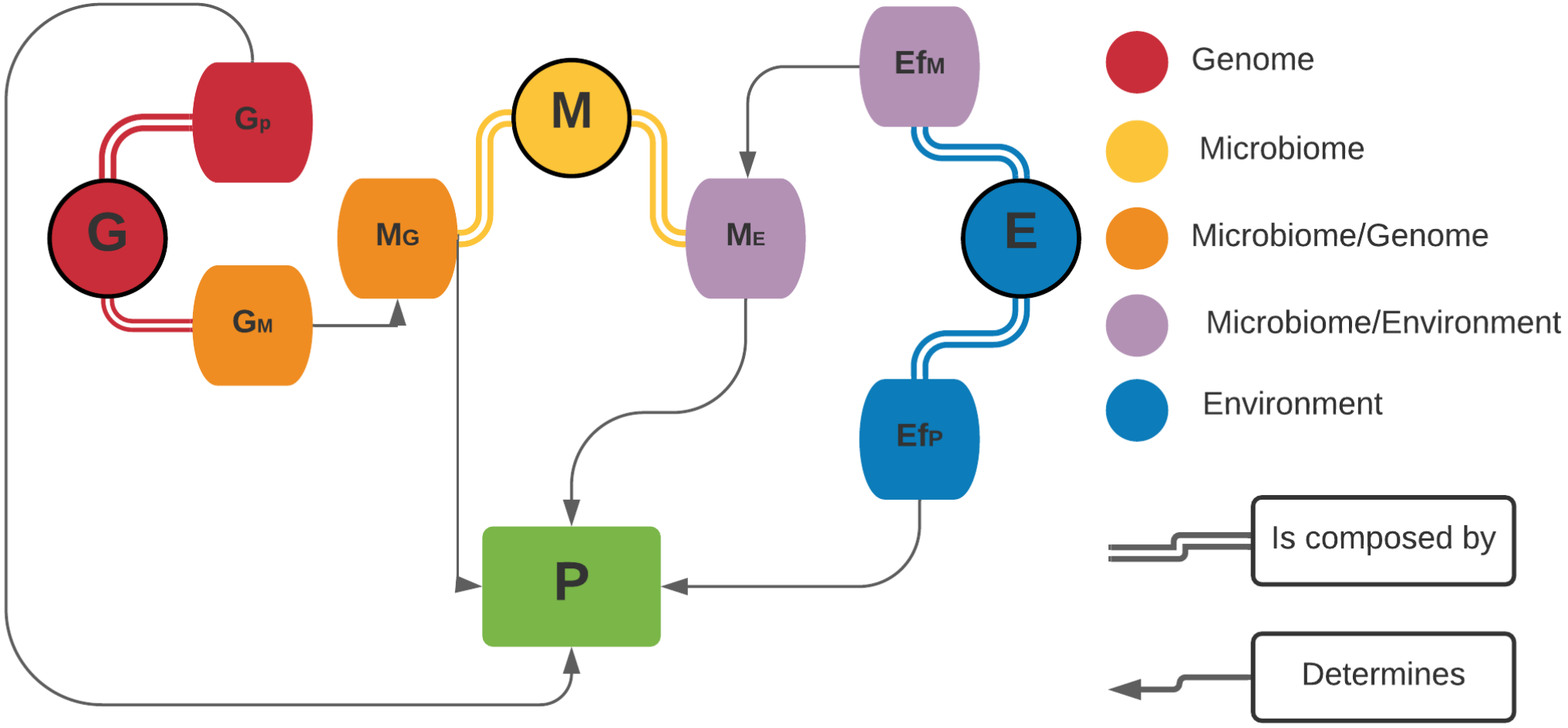
Yuqing He ✉, Francesco Tiezzi, Jicai Jiang, Jeremy Howard, Yijian Huang, Kent Gray, Jung-Woo Choi, Christian Maltecca

Journal of Animal Science, Volume 100, Issue 9, September 2022, skac231,
https://doi.org/10.1093/jas/skac231
Published: 01 July 2022 **Article history** ▾

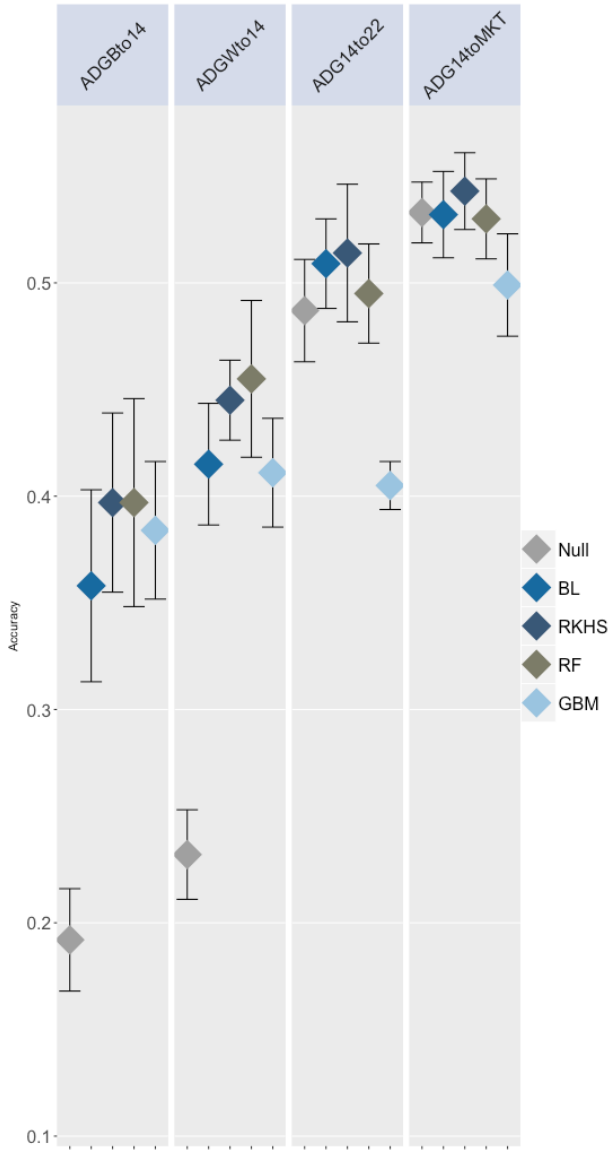


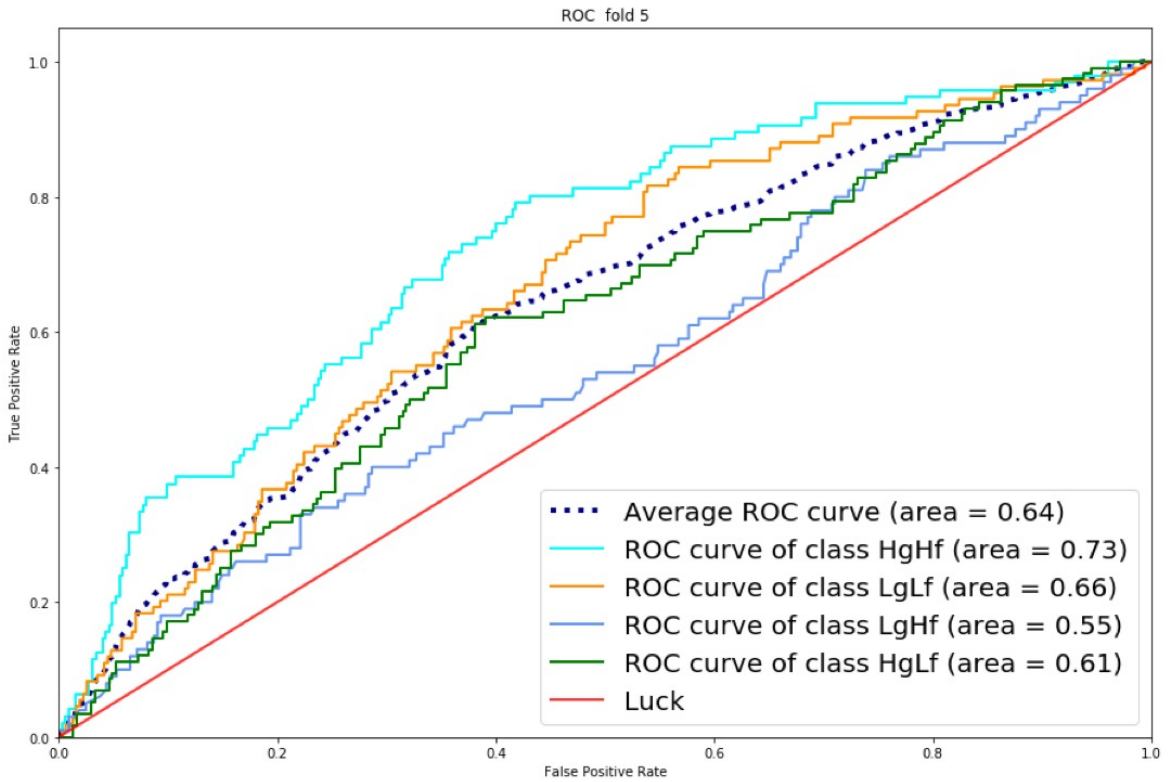
Group	Method[1]	Input Data	Function[2]	Construction of *M* matrix
General Kernel	LK	Centred and scaled log(count + 1)	$LK = XX^T(\frac{1}{p})$	LK
	PK		$PK = (XX^T(\frac{1}{p}))^3$	PK
	GK		$GK = e^{-\frac{1}{p}[X^TX-2X^TX+X^TX]}$	GK
	AK1		$\theta = \cos^{-1}(\frac{X^TX}{\|X\|\|X\|})$ $AK1 = \frac{1}{\pi}\|X\|\|X\|[\sin(\theta)+(\pi-\theta)\cos(\theta)]$	AK1
Dissimilarity	BC	log(count + 1)	$BC = \frac{\sum\|a_i - b_i\|}{\sum(a_i + b_i)}$	1 – BC
	JA		$JA = \frac{y+z}{x+y+z}$	1 – JA
Ordination	MDS	log(count + 1)	$BC = \frac{\sum\|a_i - b_i\|}{\sum(a_i + b_i)}$ X = Vectors	$XX^T(\frac{1}{p})$
	DCA		$BC = \frac{\sum\|a_i - b_i\|}{\sum(a_i + b_i)}$ X = Projections	$XX^T(\frac{1}{p})$

A- Biological Structure



Predicting Performance with the use of microbiome information





Research Article | Open Access | Published: 29 July 2020

Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine

Piush Khanal, Christian Maltecca, Clint Schwab, Justin Fix, Matteo Bergamaschi & Francesco Tiezzi

Genetics Selection Evolution **52**, Article number: 41 (2020) | Cite this article

1122 Accesses | **8** Altmetric | Metrics

Article | Open Access | Published: 25 April 2019

Predicting Growth and Carcass Traits in Swine Using Microbiome Data and Machine Learning Algorithms

Christian Maltecca, Duc Lu, Constantino Schillebeeckx, Nathan P. McNulty, Clint Schwab, Caleb Shull & Francesco Tiezzi

Scientific Reports **9**, Article number: 6574 (2019) | Cite this article

Journal of Animal Breeding and Genetics

ORIGINAL ARTICLE | Open Access

Microbiability of meat quality and carcass composition traits in swine

Piush Khanal, Christian Maltecca, Clint Schwab, Justin Fix, Francesco Tiezzi

First published: 26 September 2020 | https://doi.org/10.1111/jbg.12504

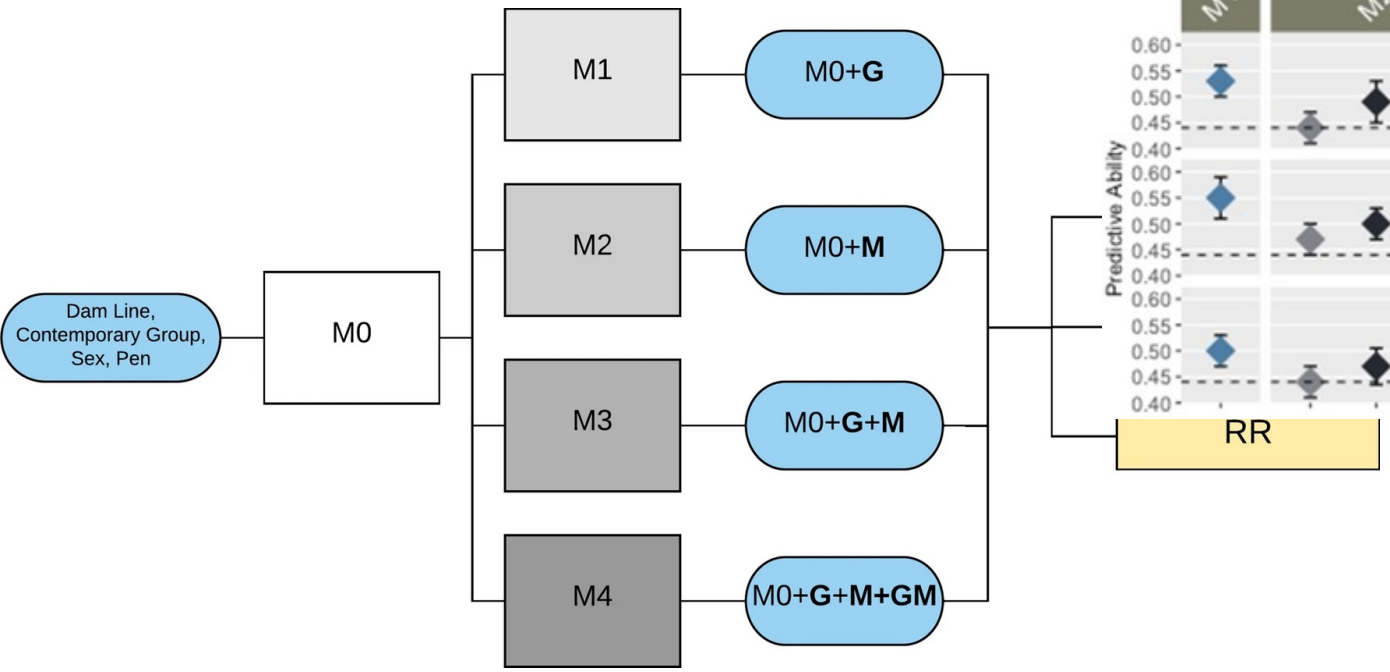
Research Article | Open Access | Published: 29 July 2020

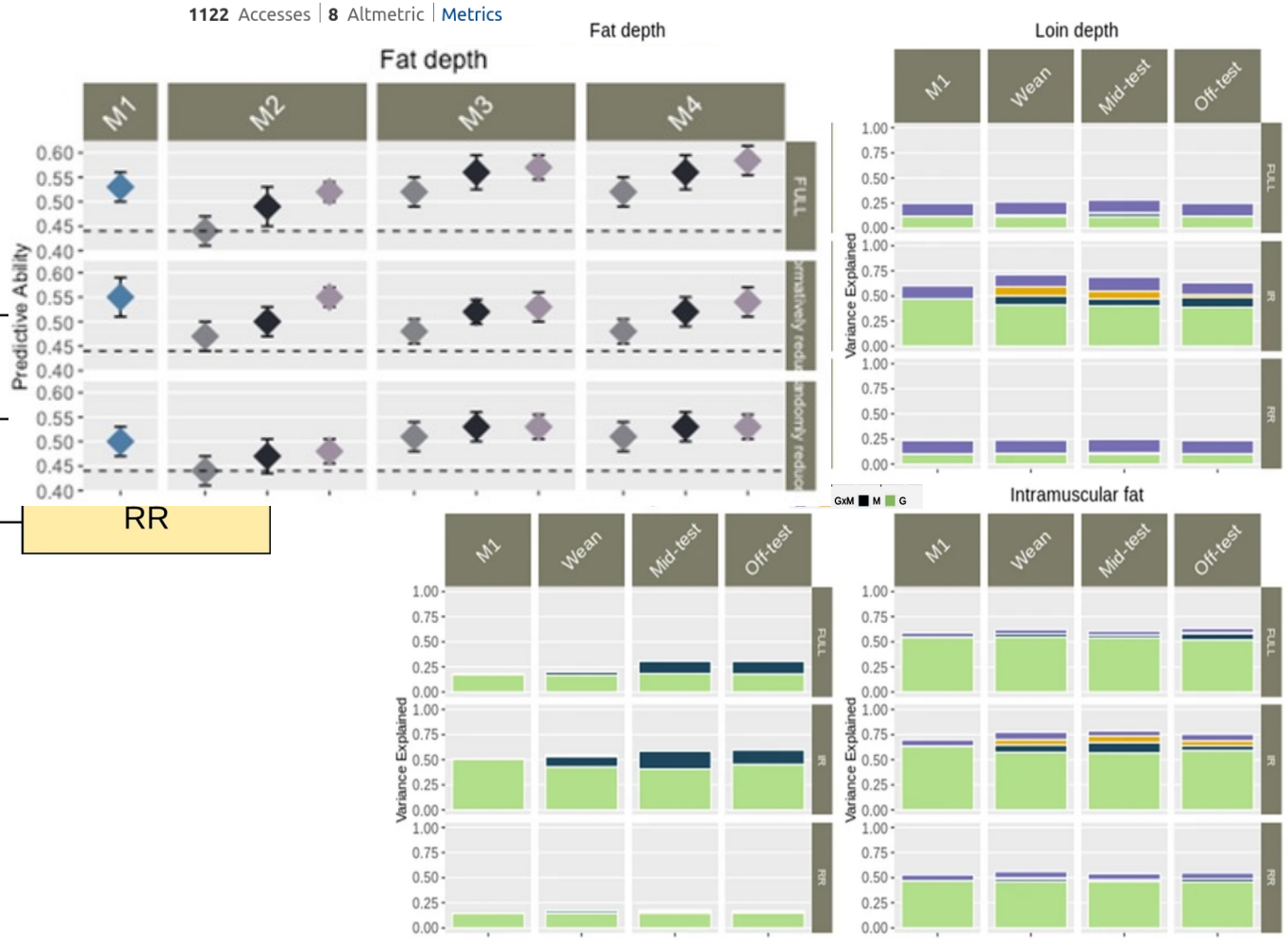
Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine

Piush Khanal, Christian Maltecca, Clint Schwab, Justin Fix, Matteo Bergamaschi & Francesco Tiezzi

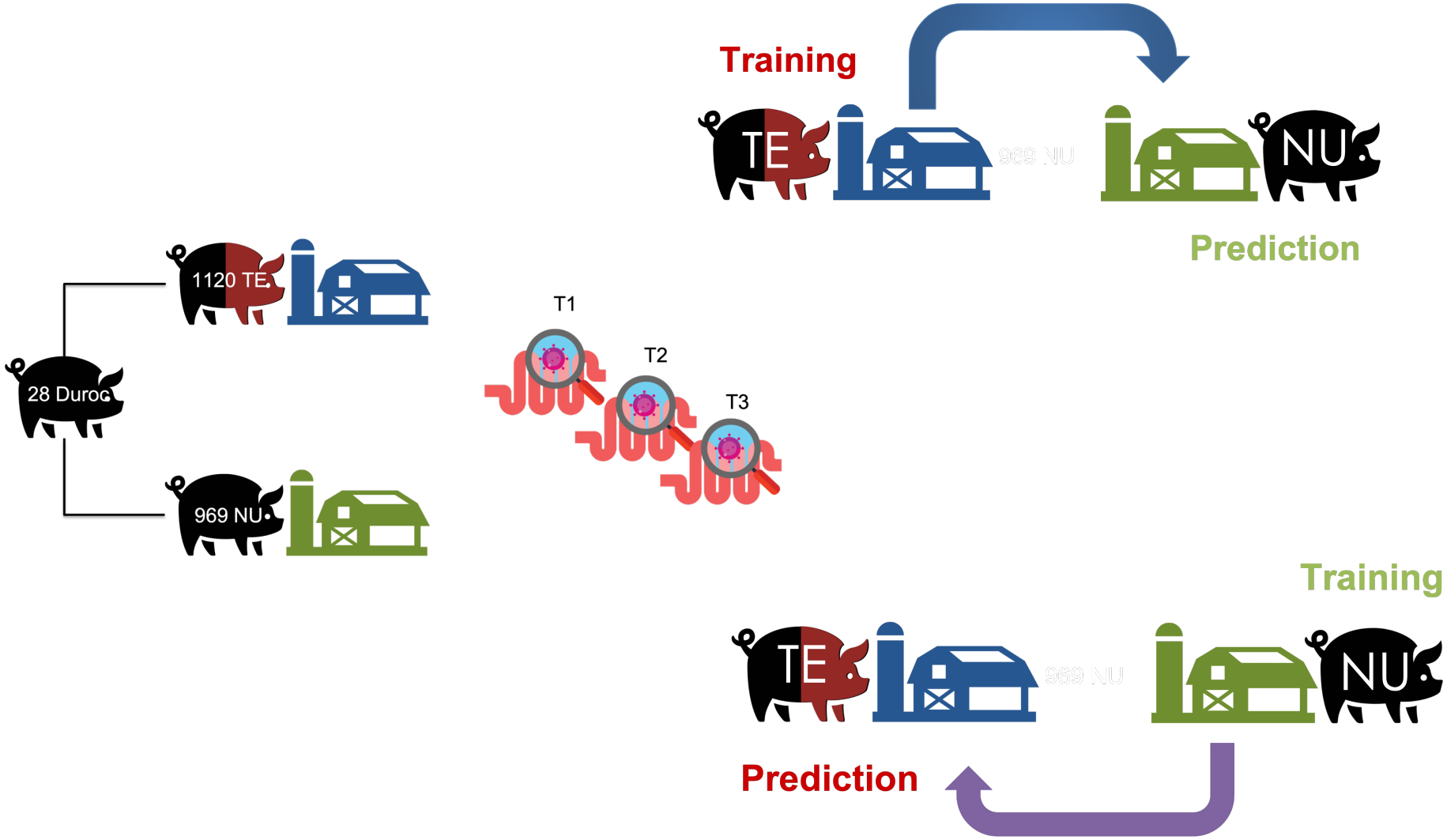
Genetics Selection Evolution **52**, Article number: 41 (2020) | Cite this article

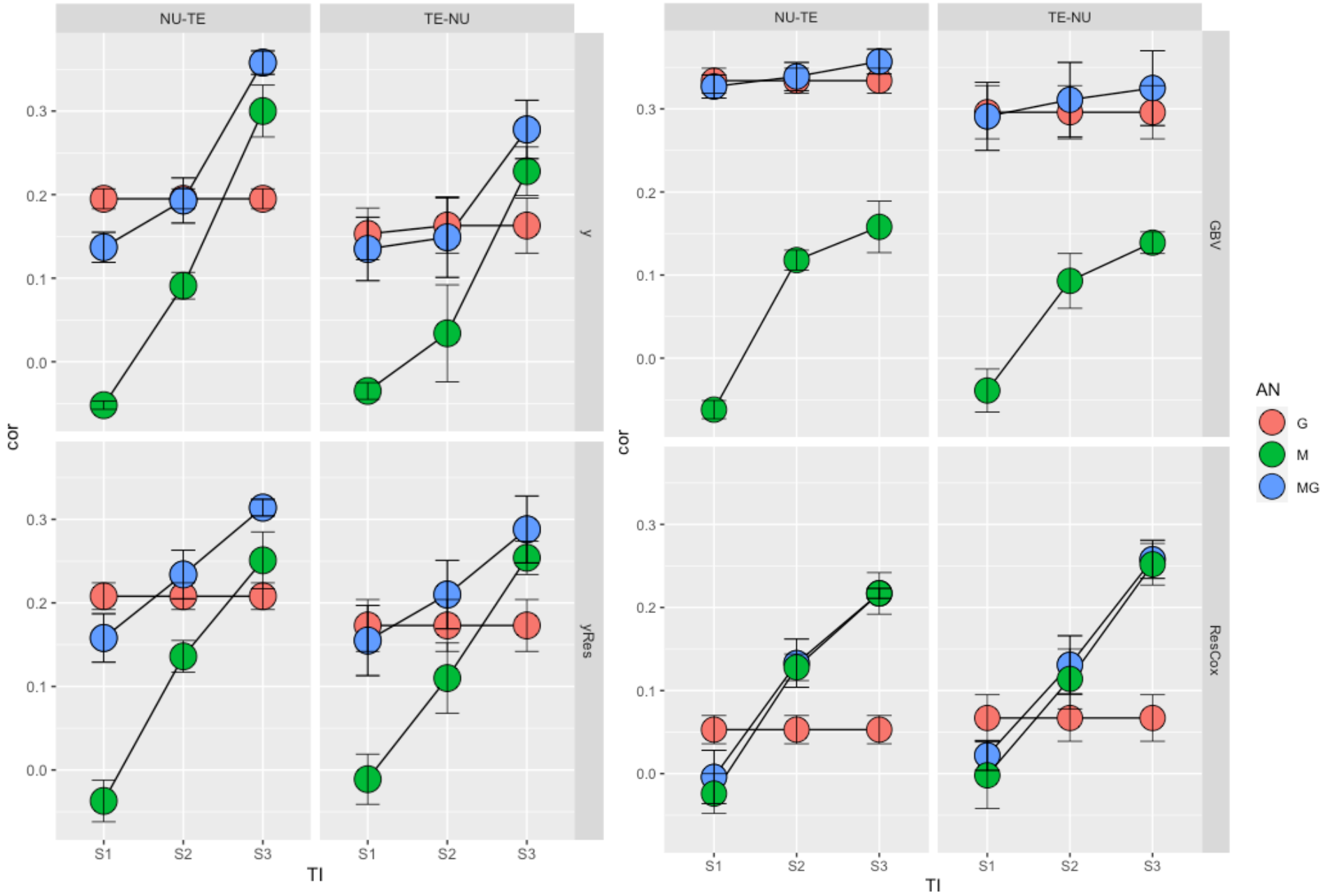
1122 Accesses | **8** Altmetric | Metrics



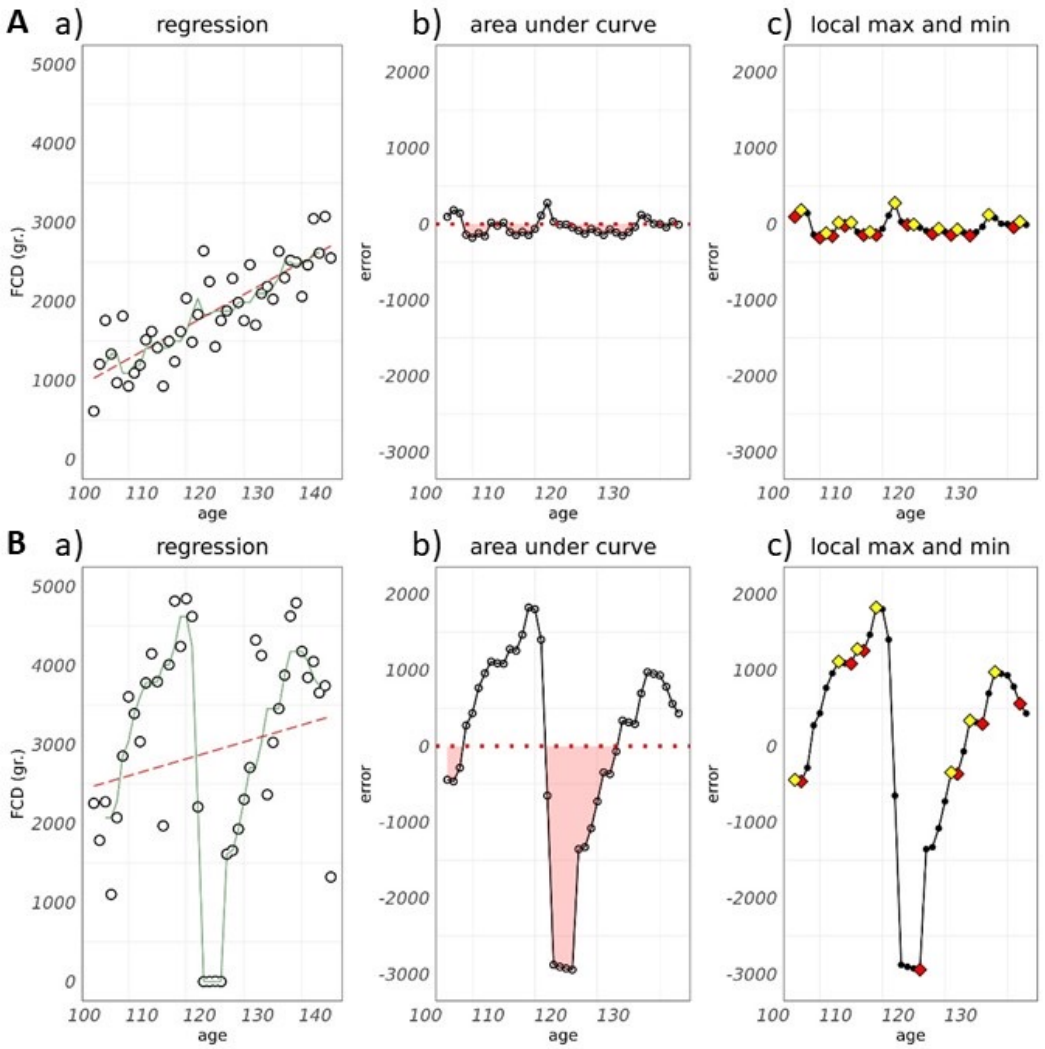


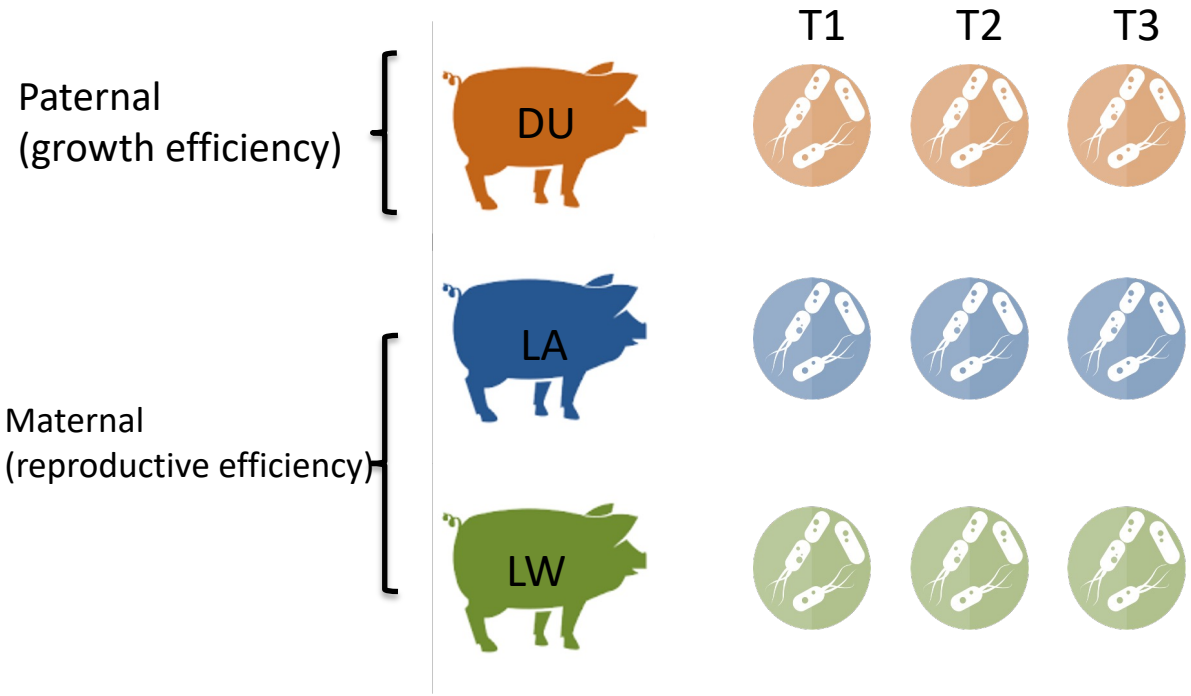
Predicting Performance across systems




NU-TE
TE-NU
y
yRes
GBV
ResCox
cor
TI
S1
S2
S3
0.0
0.1
0.2
0.3
AN
G
M
MG

A first characterization of the Microbiota-Resilience Link in Swine

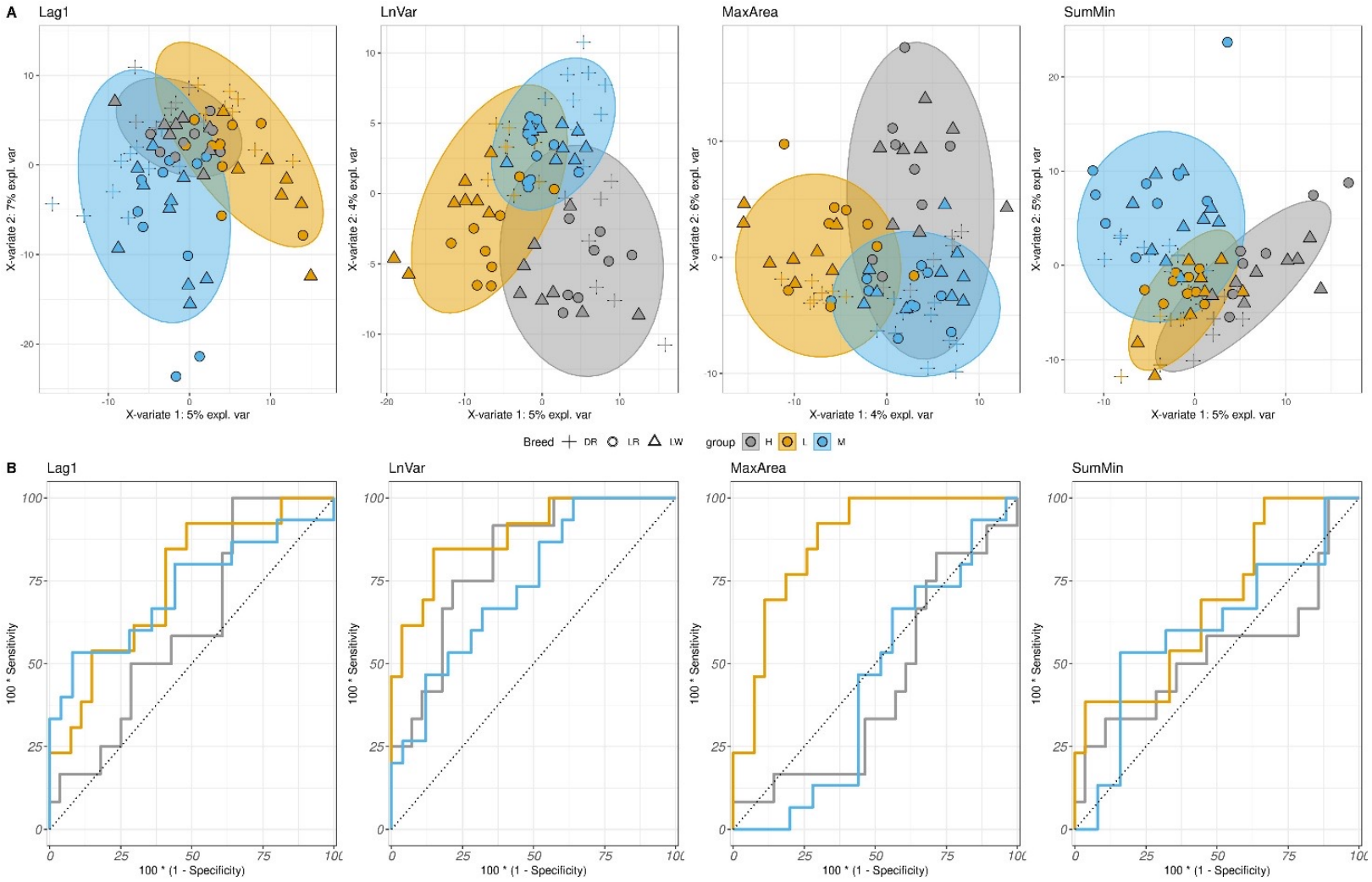




- Does microbiome contribute to broad-sense resilience in pigs?

A first characterization of the microbiota–resilience link in swine

Research | Open access | Published: 15 March 2024
Volume 12, article number 53, (2024) Cite this article



Microbiome as a source of phenotypic variability

$$\mathbf{y=Xb+Mu+Za+e}$$

y is the selection trait of interest,

X and **b** incidence matrix and vector of solutions for the environmental effects,

M is a matrix that contains the information on the microbial features (e.g. species abundance, microbial diversity)

uis the vector of microbial effects

Z and **a** are the incidence matrix and vector of solutions for the additive genetic effects of the host

$$\mathbf{y=Xb+Mu+Za+e}$$

Estimates for the variance components for the two random effects allows calculate the ratio of each variance component to the total phenotypic variance. h^2 and m^2

m^2 "microbiability" consideres m independent of g (problem)

Covariance between m and g can be considered in the model (akin to maternal effects)

Disentangling the covariance between the two terms might be challenging in practice

- EBV of an individual is determined by the genetic architecture of the trait and the known genotype of the individual

- EMV of an individual is determined by the effect of each microbial feature on the trait and the (relative) abundance of microbial features in the individual.
  - Time dependent
  - Genetic and Environment determine microbial composition.

- EMV of an individual could be determined by an environmental component that is not found in the EBV.
  - Hard to estimate covariance
    - Holds in statistical terms
    - Might lack biological rationale

Host genetic and microbial effects can also be fitted in interaction. Hadamard product of G and M.

- Holobiont concept, used to describe the system composed by a host and its associated communities of microorganisms.

- Estimate of holobiability (Saborio Montero 2019)

  - $h^2_0=(\sigma^2_g+\sigma^2_m+\sigma^2_{gm})/(\sigma^2_g+\sigma^2_m+\sigma^2_{gm}+\sigma^2_e)$

- Genetic variation in microbial composition can cause partial genotype by genotype interaction.

Microbiome as a trait

$$\begin{cases} \mathbf{m_i} = \mathbf{Xb}_{\mathbf{m}_i} + \mathbf{Za}_{\mathbf{m}_i} + \mathbf{e}_{\mathbf{m}_i} \\ \mathbf{y} = \mathbf{Xb}_{\mathbf{y}} + \mathbf{Za}_{\mathbf{y}} + \mathbf{e}_{\mathbf{y}} \end{cases}$$

$$\begin{bmatrix} \mathbf{a}_{\mathbf{m}_i} \\ \mathbf{a}_{\mathbf{y}} \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{G})$$

m (with i = 1,2, … , k) could be a
- Individual feature
- principal component
- ecological measurement of richness and diversity of the microbiota itself

$$\begin{bmatrix} \mathbf{e}_{\mathbf{m}_i} \\ \mathbf{e}_{\mathbf{y}} \end{bmatrix} \sim N(0, \mathbf{R} \otimes \mathbf{I})$$

- This model allows the estimation of the host genetic effects on both the microbiome and the phenotype
- Note that estimation of the VCV is now feasible since both am and ay are now free of environmental effects

A microbial trait can be considered as a selection criterion if:

- It is present in a large part of the population
- Shows considerable phenotypic variation across animals
- Is heritable and genetically correlated with traits of interest

Today taxonomical composition is considered as the obvious selection target

- Even within the same genus, different species might have very different metabolic pathways
  - Heritability of a microbial trait is most likely to occur in terms of functional pathways
  - Mixed results with core bacteria and core genes showing similar h2 estimates in cattle (and in pigs)
- The magnitude of r_g depends strongly on the trait and the "microbic" effect on the objective trait.

A few options:

- Selection index based on the top XX % alr(clr)-transformed features
- Latent component of SVD (or alternative dimensionality reductions)
  - Loss of interpretability but facilitate breeding

Caveats:

- Microbial metabolic pathways might be shared by several objective traits
  - Also likely to interact with other microbial activities that affect host metabolism.
  - Before microbial activity is targeted by genetic selection, the expected correlated response on other productive traits and overall animal fitness must be examined.

- Potential advantage of breeding on microbiome profiles:
  - Unfavorable genomic may no be reflected in functional microbiome.
  - Some microbial activities may be found with positive effect on more than one trait offsetting negative genetic covariances.

Predicting Breeding Values with including Omics

Genetics. 2021 Oct; 219(2): iyab130.
Published online 2021 Aug 7. doi: 10.1093/genetics/iyab130

Genetic evaluation including intermediate omics features

Ole F Christensen,[1] Vinzent Börner,[1] Luis Varona,[2] and Andres Legarra[3]

Author information · Article notes · Copyright and License information PMC Disclaimer

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{ZM}\alpha + \mathbf{Z}_r\mathbf{a}_r + \epsilon,$$

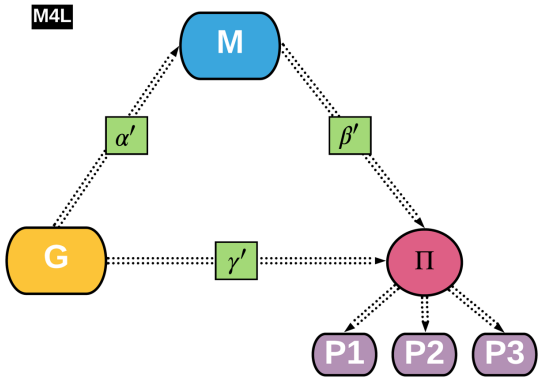
$$\mathbf{m}_i = \tilde{\mathbf{X}}\tilde{\beta}_i + \tilde{\mathbf{Z}}\mathbf{g}_i + \mathbf{e}_i, \quad i = 1, \ldots, k.$$

Weishaar R, Wellmann R, Camarinha-Silva A, Rodehutscord M, Bennewitz J.. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs—a novel selection index. *J Anim Breed Genet*. **137**:14–22. [PubMed] [Google Scholar] [Ref list]

ELSEVIER
Computational and Structural Biotechnology Journal
Volume 19, 2021, Pages 530-544

Gut microbiome mediates host genomic effects on phenotypes: a case study with fat deposition in pigs

Francesco Tiezzi [a], Justin Fix [b], Clint Schwab [b, c], Caleb Shull [c], Christian Maltecca [a]



$$\mathbf{a} = \mathbf{G}\alpha + \mathbf{a}_r. \qquad h^2 = c_m^2 h_m^2 + h_r^2,$$

BLUP of regression effects of omics expression levels and residual genetic effects are obtained as solutions to the mixed model equations (MME)

$$\begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathbf{Z}\mathbf{M} & \mathbf{X}^{\mathrm{T}}\mathbf{Z}_r \\ (\mathbf{Z}\mathbf{M})^{\mathrm{T}}\mathbf{X} & (\mathbf{Z}\mathbf{M})^{\mathrm{T}}\mathbf{Z}\mathbf{M} + \xi_1\mathbf{I} & (\mathbf{Z}\mathbf{M})^{\mathrm{T}}\mathbf{Z}_r \\ \mathbf{Z}_r^{\mathrm{T}}\mathbf{X} & \mathbf{Z}_r^{\mathrm{T}}\mathbf{Z}\mathbf{M} & \mathbf{Z}_r^{\mathrm{T}}\mathbf{Z}_r + \xi_2\mathbf{H}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \\ \hat{\mathbf{a}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{y} \\ (\mathbf{Z}\mathbf{M})^{\mathrm{T}}\mathbf{y} \\ \mathbf{Z}_r^{\mathrm{T}}\mathbf{y} \end{bmatrix},$$

$$\mathbf{M}\hat{\alpha} + \tilde{\mathbf{Z}}\hat{\mathbf{a}}_r,$$

$$\begin{bmatrix} \tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}} & \tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{Z}} \\ \tilde{\mathbf{Z}}^{\mathrm{T}}\tilde{\mathbf{X}} & \tilde{\mathbf{Z}}^{\mathrm{T}}\tilde{\mathbf{Z}} + \zeta\mathbf{H}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\tilde{\mathbf{B}}} \\ \hat{\mathbf{G}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}^{\mathrm{T}}\mathbf{M} \\ \tilde{\mathbf{Z}}^{\mathrm{T}}\mathbf{M} \end{bmatrix},$$

$$\hat{\mathbf{a}} = \mathrm{E}[\mathbf{a}|\mathbf{y},\mathbf{M}] = \mathrm{E}[\mathbf{G}\alpha + \mathbf{a}_r|\mathbf{y},\mathbf{M}] = \mathrm{E}[\mathbf{G}|\mathbf{M}]\mathrm{E}[\alpha|\mathbf{y},\mathbf{M}] + \hat{\mathbf{a}}_r$$

$$= \hat{\mathbf{G}}\hat{\alpha} + \hat{\mathbf{a}}_r,$$

directly predicting genetic effects on individuals

$$\begin{aligned}\hat{\mathbf{a}} &= \mathrm{E}[\mathbf{G}\alpha|\mathbf{y},\mathbf{M}] + \mathrm{E}[\mathbf{a}_r|\mathbf{y},\mathbf{M}] \\ &= \mathrm{E}[\mathrm{E}[\mathbf{G}\alpha|\mathbf{y},\mathbf{M},\alpha]|\mathbf{y},\mathbf{M}] + \mathrm{E}[\mathbf{a}_r|\mathbf{y},\mathbf{M}] \\ &= \mathrm{E}[\mathbf{G}\alpha|\mathbf{M}\alpha]|_{\mathbf{M}\alpha=\mathrm{E}[\mathbf{M}\alpha|\mathbf{y},\mathbf{M}]} + \mathrm{E}[\mathbf{a}_r|\mathbf{y},\mathbf{M}],\end{aligned}$$

$$\begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathbf{Z} & \mathbf{X}^{\mathrm{T}}\mathbf{Z}_r \\ \mathbf{Z}^{\mathrm{T}}\mathbf{X} & \mathbf{Z}^{\mathrm{T}}\mathbf{Z} + \xi_1(\mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1} & \mathbf{Z}^{\mathrm{T}}\mathbf{Z}_r \\ \mathbf{Z}_r^{\mathrm{T}}\mathbf{X} & \mathbf{Z}_r^{\mathrm{T}}\mathbf{Z} & \mathbf{Z}_r^{\mathrm{T}}\mathbf{Z}_r + \xi_2\mathbf{H}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{a}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{y} \\ \mathbf{Z}^{\mathrm{T}}\mathbf{y} \\ \mathbf{Z}_r^{\mathrm{T}}\mathbf{y} \end{bmatrix},$$

$$\begin{bmatrix} \tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}} & \tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{Z}} \\ \tilde{\mathbf{Z}}^{\mathrm{T}}\tilde{\mathbf{X}} & \tilde{\mathbf{Z}}^{\mathrm{T}}\tilde{\mathbf{Z}} + \zeta\mathbf{H}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\tilde{\beta}}\alpha \\ \hat{\mathbf{a}}_m \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}^{\mathrm{T}}\hat{\mathbf{u}} \\ \tilde{\mathbf{Z}}^{\mathrm{T}}\hat{\mathbf{u}} \end{bmatrix}.$$