



Behind the Count'her

SCIENTIFIC
NETWORK BEHIND
THE COUNT HER

2022-2023

Coordination

Bertrand Servin
bertrand.servin@inrae.fr

Key words

Count data
Poisson-lognormal model
Heritability
Statistics
Genetics

INRAE units involved

[MIA Paris Saclay](#)

[MaIAGE](#)

[GenPhySE](#)

Partnerships

University of Paris

Modelling the heritability of traits from count data

Context and challenges

Understanding the genetic determinism of a trait, i.e. the set of genes involved in the development and expression of this trait, is a major challenge for better understanding biological processes and supporting genetic improvement programmes.

The recent development of genotyping and massive sequencing tools, which allow the rapid sequencing of several thousand to millions of DNA or RNA molecules simultaneously, has considerably increased the power of experimental devices in this field, leading to new challenges in analysing massive sequencing data.

In this context, one of the emerging issues is the analysis of data that correspond to a number of observed events (count data). Indeed, the analysis of this type of data by hierarchical generalized linear models is notoriously difficult, in particular when it comes to estimating the heritability of traits.

The Behind the Count'her consortium proposes to use recent statistical developments to adapt a model for estimating genetic parameters. It will be based on two cases of application in the field of breeding: the distribution of recombinations along the genome and the diversity of the ruminal meta-genome.



© pikisuperstar - Freepik



Goals

Consortium partners have recently proposed a flexible statistical model for count data, based on the Poisson-lognormal distribution, which allows complex effects to be modelled and estimated in reasonable time: covariance structure, clustering and dimension reduction, network inference, etc. However, adapting this model to the context of inferring genetic parameters requires specific developments.

The Behind the Count'her consortium therefore proposes to bring together teams from quantitative genetics and statistics to jointly develop new statistical models for the analysis of count data.

The developments will be based on two study contexts that allow a wide range of applications to be covered:

- The modelling of data on the distribution of crossings on the genome during meiosis (for which the genes involved and their variations remain unknown in many species).
- Analysis of the diversity of the meta-genome between individuals and use of these results to measure the effect of the microbiota on other traits.

The ambitions of the Behind the Count'her consortium are both methodological and genetic. The consortium will both remove methodological barriers to the efficient exploitation of count data and also provide new knowledge on the genetic effects of recombination phenotypes and microbiota.

Research units involved and partners

INRAE scientific division	INRAE research units	Expertise
<u>Mathematics, computer and data sciences, digital technologies</u>	<u>MIA Paris Saclay</u>	Statistical Modelling
	<u>MaIAGE</u>	Statistical Modelling, Evolutionary Genomics
<u>Animal genetics</u>	<u>GenPhySE</u>	Metagenomics, Genetics, Statistics
External partners		Expertise
University of Paris	<u>LPSM</u>	Statistical Modelling

Reference

- **J. Chiquet, M. Mariadassou and S. Robin:** The Poisson-lognormal model as a versatile framework for the joint analysis of species abundances, *Frontiers in Ecology and Evolution*, 2021
- **S. Fresco, C. Marie-Etancelin, A. Meynadier, G. Martinez-Boggio:** Variation in Rumen Bacteria of Lacaune Dairy Ewes From One Week to the Next, *Frontiers in Microbiology*, 2022
- **H. Vassilief, M. Id Bella, D. Hazard, F. Tortereau, T. Faraut, S.E. Johnston, B. Servin:** Sex differences in recombination maps are associated with differential hotspot usage in Sheep. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, Rotterdam, 2022

