

# Integration and Analysis of Metabolomic Dairy Cows Data Gaussian Graphical Modelling and Systems Biology Approaches

**Supervisors:** Edoardo Saccenti (SSB) & Ariette van Knegsel (Animal Physiology)

**Contacts:** edoardo.saccenti@wur.nl & ariette.vanknegsel@wur.nl

**Type of thesis:** Computational

**Required competences:** Basic statistics, Basic knowledge of R and/or Matlab and Python

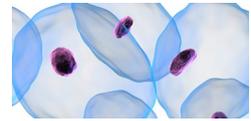
**Acquired competences:** Network analysis, Multivariate statistics, analysis of (metabol)omics data, phenotyping of human metabolism, basic of dairy cow metabolism)

**Date start of the project:** Open

## Description

Negative energy balance in dairy cows in early lactation is related to alteration of metabolic status. Classically, the energy balance has been related with a relative standard series of metabolites (e.g. glucose, ketone bodies) and metabolic hormones (e.g. insulin, growth hormone), The relationships among energy balance and a complete metabolomic profile in plasma or milk have been seldom investigated (Xu, Vervoort et al. 2018, Xu, Vervoort et al. 2020) .

Scope of the thesis is to integrate existing results from nuclear magnetic resonance (NMR) with data from liquid chromatography triple quadrupole mass spectrometry (LC-MS) obtained on blood samples of dairy cows with phenotypic data to investigate the relationship between energy balance and the metabolomics profiles of plasma and milk



and to gain understanding about the biomolecular mechanism underlying the development of negative energy balance using systems biology approaches and in network biology in particular (Rosato, Tenori et al. 2018).

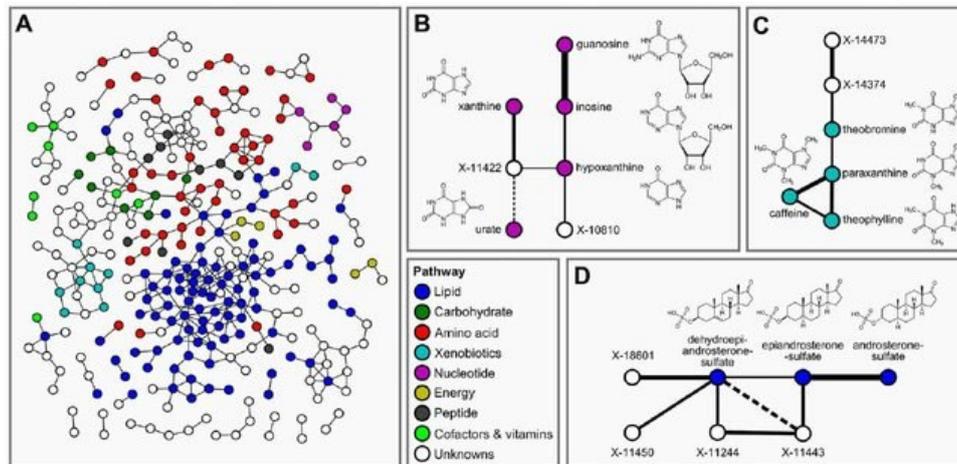
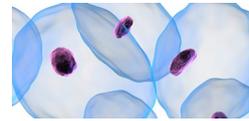
The computational/statistical work of this thesis will focus on the so called Gaussian graphical models.

Graphical models (Altenbuchinger, Weihs et al. 2020) are a unifying framework for describing the statistical relationships between large collections of random variables. In graphical models, the dependencies between different variables, the so-called nodes or vertices, are represented as edges in a network. Thus, two nodes are connected by an edge only if their association or interaction cannot be explained by any other node in the graphical model.

Consequently, probabilistic graphical models eliminate spurious associations between variables and can potentially reveal new associations adjusted for all other variables in the dataset.

At the end of this project the student will have gain knowledge in (bio)logical network inference and analysis, integration of different data type, (advanced) multivariate statistics, NMR and Mass spectroscopy, analysis of large metabolomics data sets and dairy cow science.

This thesis project is a collaboration between the Laboratory of Systems and Synthetic Biology (Dr. Edoardo Saccenti) and the Adaptation Physiology Group (Dr. Ariette van Knegsel). The student will split the thesis time between the two groups.



**Figure 1 Example of Gaussian Graphical Modelling:** GGMs embed unknown metabolites into their biochemical context. A: Complete network presentation of partial correlations that are significantly different from zero at  $\alpha = 0.05$  after Bonferroni correction. The unknown metabolites are spread over the entire network and are involved in various metabolic pathways. B–D: Selected high-scoring sub-networks. We observe that GGM edges directly correspond to chemical reactions which alter specific chemical groups (e.g. carbonyl groups and methyl groups). Solid lines denote positive partial correlation. Dashed lines indicate negative partial correlations. Line widths represent partial correlation strengths. From: doi:10.1371/journal.pgen.1003005.g003

## References

- Altenbuchinger, M., A. Weihs, J. Quackenbush, H. J. Grabe and H. U. Zacharias (2020). "Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools." *Biochim Biophys Acta Gene Regul Mech* **1863**(6): 194418.
- Rosato, A., L. Tenori, M. Cascante, P. R. De Atauri Carulla, V. A. P. Martins dos Santos and E. Saccenti (2018). "From correlation to causation: analysis of metabolomics data using systems biology approaches." *Metabolomics* **14**(4): 37.
- Xu, W., J. Vervoort, E. Saccenti, B. Kemp, R. J. van Hoeij and A. T. M. van Knegsel (2020). "Relationship between energy balance and metabolic profiles in plasma and milk of dairy cows in early lactation." *Journal of Dairy Science*.
- Xu, W., J. Vervoort, E. Saccenti, R. van Hoeij, B. Kemp and A. van Knegsel (2018). "Milk Metabolomics Data Reveal the Energy Balance of Individual Dairy Cows in Early Lactation." *Scientific Reports* **8**(1): 15828.