

Cascade: a R-package to study, predict and simulate the diffusion of a signal through a temporal gene network.

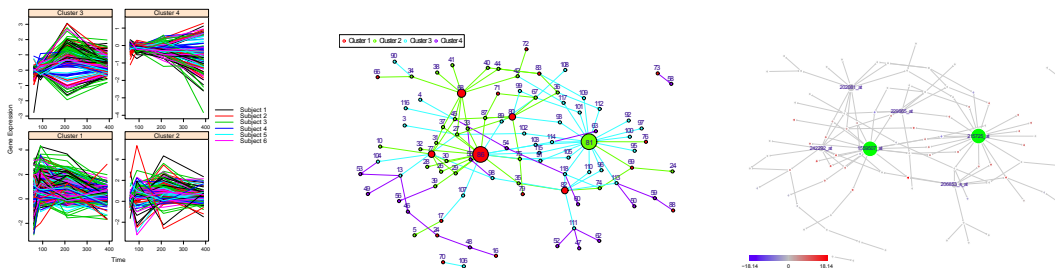
Nicolas Jung^{1,2,3}, Frédéric Bertrand^{1,2,*}, Seiamak Bahram^{1,3}, Laurent Valla^{1,3}, Myriam Bertrand^{1,2}

1. Université of Strasbourg
2. Centre national de la recherche scientifique
3. Institut national de la santé et de la recherche médicale
*Contact author: fbertran@math.unistra.fr

Keywords: Partial least squares regression, generalized linear models, bootstrap, high dimensional data, R software package

Temporal gene interactions, in response to environmental stress, is a complex system that can be efficiently described using gene regulatory networks (GRN): a GRN allows to highlight the more influential genes and to spot some targets for biological intervention experiments. Despite that many reverse-engineering tools have been designed, the **Cascade** package is an integrated solution adding several new and original key features such as the ability to predict changes in gene expressions after a biological perturbation in the network and graphical outputs that allow monitoring the spread of a signal through the network.

Since the emergence of high-throughput technologies, many tools have been developed to learn gene expression profiles and reverse-engineer their underlying GRN [1,2]. These tools are either based on static co-expression methods or, if the biological phenomenon shows any temporality, time dependent methods. While the former relies on the assumption that co-expressed genes share some biological characteristics, the latter infers a directed network with temporal dependencies. In this last case, another important distinction should be made between exogenous stress (e.g., growth response) and endogenous phenomenon (e.g., cell cycle) [3,4]. This leads to different network topologies: in exogenous stress, networks' topologies seem to have larger hubs and shorter paths through temporal dependent transcriptional waves [3]. This results in a quick response to environmental modifications [3]. The **Cascade** package is designed to model such “cascade networks” taking advantage of the assignment of genes to temporal clusters which are then used to enforce temporal causality in the network.



The three steps : gene selection with assignment to a time cluster (left), reverse-engineering of the network (center), predicted perturbations in the network after gene expression modulation at an early time (right).

- [1] Bar-Joseph, Z. *et al.* (2012). Studying and modelling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics*, **13**(8), 552–564.
- [2] Bansal, M., *et al.* (2007). Hecker, M. *et al.* (2009). Gene regulatory network inference: data integration in dynamic models—a review. *Biosystems*, **96**, 86–103.
- [3] Luscombe, N. M., *et al.* (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, **431**(7006), 308–312.
- [4] Zhu, X., Gerstein, M., and Snyder, M. (2007). Getting connected: analysis and principles of biological networks. *Genes and development*, **21**(9), 1010–1024.