Investigate clusters of co-expressed and co-located genes at a genomic scale using CoCoMap

Marion Ouedraogo [1]**, Frédéric Lecerf** [2]**, Sébastien Lê** [3]

1. INRA, Agrocampus Ouest UMR 0598 Génétique Animale, Rennes, France
2. Agrocampus Ouest, INRA UMR 0598 Génétique Animale, Rennes, France
3. Agrocampus Ouest, CNRS UMR 6625 Mathématiques Appliquées, Rennes, France
*Contact author: marion.ouedraogo@rennes.inra.fr

The study of the genome structure and its role in the gene function regulation had revealed new insights in the regulation of genes expressions by their chromosomal locations. The regulation of the genes expression is in some part due to the genome architecture. The organization of the chromosomes within the nucleus allows interactions between distant regions. Those interacting regions might present some co-expressed genes. The biological interest is to identify the co-located genes within a region which present a co-expression. In our data, genes are considered as statistical variables and individuals as statistical units; we will say that genes are co-expressed if they induce the same structure on the individuals.

The aim of the **CoCoMap** package is to identify co-expression between genomic regions. It uses transcriptomic data and the corresponding chromosomal locations of the genes and is mainly based on multivariate exploratory approaches:

i) A sequence of Principal Component Analyses among the genome is used to identify the co-located genes which present a co-expression.

We introduce the autovariogram representation to identify and discriminate regions of co-expressed genes.

ii) The Multiple Factor Analyses and its features are used to describes the co-expression between the chromosomal regions.

The final output is the groups of co-located genes which present a co-expression. The method has proven its efficiency on simulated data. Preliminary analyses on two different experimental sets identified some common clusters in interaction.

The **CoCoMap** package proposes an interface to the method and its features to investigate transcriptomic data.

References

Baker M. (2011). Genome in three dimensions. *Nature* 470: 289–294

De S. and Babu M. (2010). Genomic neighbourhood and the regulation of gene expression. *Current Opinions in Cell Biology*, 22:326–333.

Sexton T. et al. (2009). Genomic interactions: Chromatin loops and gene meeting points in transcriptional regulation. *Seminars in Cell & Developmental Biology* 20: 849–855