Poster E-9

Dynamical System Modeling of the Estrogen Transcription Control Network

Authors:
Marcelo Ris (*IME-USP*)
Junior Barrera (*IME-USP*)
Helena Brentani (*Hospital do Câncer*)

Short Abstract: Identify the estrogen transcription control network, that is, modeling the relationships of the estrogen genes regulated. We adopt the PGN model and estimate its parameters from gene expression profiles, using the U-curve algorithm. Based on the relationships, we can group the genes in the different mechanisms of transcription control.

Long Abstract:

Estrogen is not only one hormone but a group of hormones, which is found mainly in 3 forms in human body: estrone (E1), estradiol (E2) and estriol (E3). Estrogen has an important role in the reproductive tissues: the growing of the mammary gland and endometry during pregnancy and menstrual cicle are estrogen dependent. The tumoral cells growing stimulated by estrogen may require the presence of estrogen receptors. The tumoral cells can be clssified in two main groups: cells that express estrogen receptors (ER+) and that not express them (ER-). The last one (ER-) is more agressive and poor treatment success. Therapies using anti-estrogen drugs (as Tamoxifen) is used with some success in the ER+ tumors.

There are more than 300 genes positive and negative regulated by estrogen. Between the positive regulation we have: IGFBP4, GREB1, PGR; and negative regulation: NMA, BPM7. Our goal, in our work, is to identify the estrogen transcription control network. More precisely we are interested in modeling the relationships of these genes between each other, using for this Probabilistic Genetic Network (PGN). Then we can group this genes based on their relatioships and correlate these groups with different mechanisms of transcription control.

Probabilistic Genetic Network is a Markov Chain with some aditional parameters: transition function is conditional independent and almost deterministic. Each state of the Markov Chain is a vector where each component charcaterized a gene. The conditional independent property means that the probability of a new state is obtained from the product of the probability of each component of this state, given a previous state. Our desire is to find which genes influence the expression of a given gene i, and to find this we should find which genes minimize the entropy of the gene i given these genes. There are several algorithms that estimate good candidates, but don't guarantee the better result.

We have developed a stochastic algorithm called U-curve algorithm that can compute the best candidate (or candidates). The algorithm uses the premise that the entropy function has an U-curve form and from this it can achieve the best result without having to compute values for each candidate. As a first approach, we have computed the 89 genes in the Ed Liu et al. Work using the U-curve algorithm, and build the PGN (characterized by a directed graph), using a developed pipeline of algorithms. The graph constructed can be viewed in an html browser, and each node (representing a gene) has a link to the prediction table and links to GeneBank.