Introduction to Computational Biology
Lecture # 28: Evolution Of Networks - Transcription and Protein-Protein Networks
Ya’ar Reuveni
16/01/2007

1 Brief Review:
In the previous lecture we talked about transcription network design, an interesting question would be what is the evolution of these kind of networks. how would we look on networks like this in different creatures? People have made comparisons between human and mouse networks. and there were a lot of attempts with Drosophila, insertions of genes from some of them to others etc. and also there were a few studies of this sort on yeast. So today we will elaborate some more on the transcription networks and their evolution, and later on talk a bit about protein-protein interaction networks.

2 Evolution Of Transcription Networks:
Now consider proteins go through a much more slower evolution than biological networks. and the intuition is that it is much more simple to make a few nucleotides change than the whole protein, a change that will cause a structural and functional change. Nevertheless how do we still see changes in transcription networks?
The use of transcription networks is to govern the transcription process, respond to environmental change. say we have transcription networks of a few individuals groups, now we can check the network in other species. we would like to create a module of transcription networks, so that we could find dependencies that might have evolutionary significance.
assume we see a specific protein that regulates a group of genes and that it is preserved over evolution, we would like to notice such a fact using our module. to make such an observation we could check for the binding of a transcription factor in the different species. we could also compare with results of gene expression assays to get a better combination of the information and so make a decision on the correctness of the preservation we found.
These assumptions are not groundless, for example ribosomal genes are expressed together in most species we know. a nice story on this subject, people have looked on a group of yeast genes ( 400) in view of Fermentation V.S. Respiration yeast can use both Fermentation and Respiration for creating energy, Fermentation - without the presence oxygen and Respiration - with the presence oxygen.
When they looked at the more primitive species there was only use of Respiration, and in the younger species a new control that causes Fermentation was added. so we got:

**Theorem 2.1** Once there was only a control for yeast hunger for energy that governed a few processes and yeast could only use Respiration
Later on in evolution a new species arrived that could also use Fermentation as an energy source using a newly developed control

what could we look at?
A) conservation of the binding site  
B) conservation of the binding  
C) sequence conservation

it is obvious $B$ has much more affect on functionality so probably $A$ affects on $B$ cause we can assume $B$ has a grater evolutionary pressure, cause the binding itself affects more on the transcription than location or character of the binding now if we combine with the evolutionary tree we could conclude if it is a loss or gain of function.

2.1 summary:

Evolutionary it is much more simple to gain a binding site than to lose one, but if we wait enough time even mutations in preserved sites might occur leading to changes in function, that maybe gained functions could provide a biological back up on. we could ask on conservation of binding site using methods like $PSSM$ of recognition site, and this is less stiff than sequence conservation.

now in the level of sequence conservation ($C$) we would want to see conservation of sequence using motif conservation methods like we saw a lot before. anyhow sequence conservation will always be “playing for sure” you can miss a lot, but what you find you can be sure it probably has an evolutionary significance for a reason.

3 Protein-Protein interaction Networks:

What kinds of P-P networks can we find? there are indirect networks, physical interactions lots of ways like “yest two hybrid” etc. but there is a bit of a problem this subject is still not so well defined, in view of bind strength, affinity or time. what interactions? if there are two proteins that seldom bind does this mean they should have an edge in the network? but what if they bind very strong? there are a lot of these kind of questions that can arise. thus usually this subject is not well defined in literature.

one obvious characteristic of Networks of this sort is that the distribution of the vertices ranks is not random, and so they usually have $HUBS$ - vertices that are central in many processes (e.g. actin,kinases etc...). we would expect theses $HUBS$ will be essential.

what more could we learn from these interactions? how can we really check if a $HUB$ is essential? maybe we could use knockout? what more can we learn? we can locate pathways, if two proteins have a physical connection we can guess they are involved in a process with the same function. we can also implicate on cell location, or time, when do they express? we would expect for overlapping in time.

two similar problems:

1. clustering - how to cluster proteins?  
2. how do we progress when we have information on only part of the proteins or their relations or know only part of the relations?

4 Summery:

Today we talked some more on transcription networks, and protein-protein interactions networks. next lesson we will proceed and advance some more on the protein-protein interactions networks, and see if we could find some statistical modules to investigate the subject.