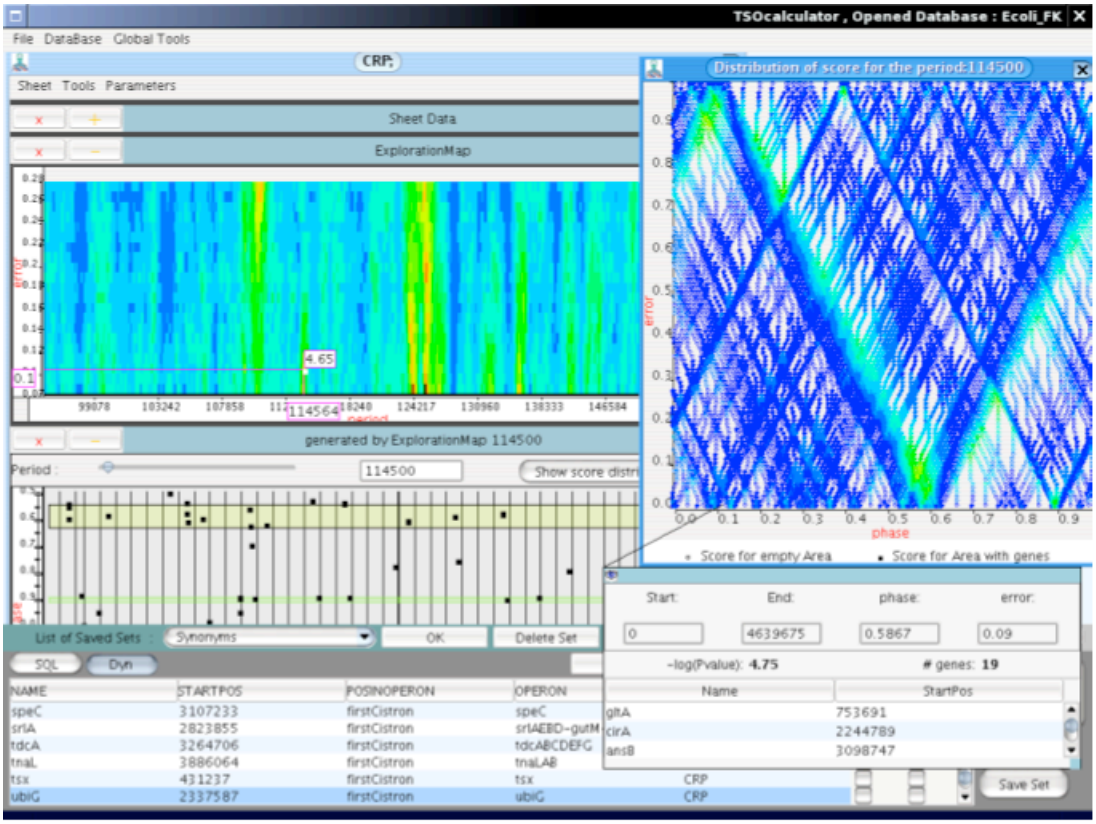
Automatic detection of regularities in the position of co-regulated genes

Recent studies showed, in prokaryote Escherichia coli and eukaryote Saccharomyces cerevisiae, that co-regulated genes tend to be positioned periodically along the chromosome (Képès 2003, Képès 2004). The observed period differs between chromosome arms in yeast, and from one step to another in bacteria. This remarquable repartition is coherent with a solenoidal epi-organization of chromosomes which result to come spatially nearer interacting partners. Thus, one period observed on the linear chromosome corresponds to a DNA loop of the solenoid. The mechanistic justification of this interpretation exists because we know that this spatial rapprochement of co-regulated genes optimize their transcription. The transcriptional organization of these genomes would be implied in the dynamical, global and spatial structuration of chromosomes.
The Epigenomics Project has developed some systematic detection techniques of regularities in co-regulated genes position along DNA.
The overall goal of the solenoid project is to describe the relation between the functional organisation of the nucleus - or bacterial nucleoid - and the linear organisation of chromosomes, in a vision that is renewed by post-genomic data. The major challenge is to characterize the prototypical DNA loop of a chromosome in different organisms. Starting from biological data, the aim is to characterize the positional and dynamical relations in a DNA loop that optimize all aspects of DNA metabolism: transcription, replication, transposition, recombination, etc.

A software tool has already been achieved within the Epigenomics Project. This software offers a range of tools through a powerfull Human-Computer Interface. This allows the user to do sophisticated studies in a intuitive way. Results might be visually or anatically analyzed within the same workflow.
This software consists of 3 parts: a database, a tools pallet, a manipulation and visualization interface.



Nowadays, the database integrates 2 organisms: C. elegans and E. coli. The software offers a view on this database in order to quickly select a part of data. It is also possible to integrate other organisms.
The algorithms already integrated essentially aim to discover regularities in the position of co-regulated genes. Based on statistical methods, these tools are quiet better than classical analyses (Fourier Transform and Wavelets).
The interface allows on one hand to manipulate tools in an efficient way through a workflow, and on other hand to visualize results. Algorithms parameters can be fit on the fly on graphs which are computed in real time (cf. figure below).

In addition, a software tool has been achieved to visualize in 3D a solenoidal structure with co-regulated genes.

The aim of this fellowship is threefold:

1. to integrate new algorithms developed under Mathematica into the platform,
2. to integrate the 3D visualization of results within the platform and,
3. to begin a bioinformatics analyze from post-genomic data.

This fellowship would allow in the same time to reinforce the idea that the genome is not radomly organized in cell, and to enhance the prediction of transcription regulation networks.