

## NORAYBIO REACHES THE END OF EUROPEAN PROJECT GENNETEC COLLABORATION

In November of 2009, the Specific Targeted Research Project (STREP) **GENNETEC** (Project no. 34952), funded by the European Commission under the Sixth Framework Programme, will be finished.

Gennetec gathers partners from 4 countries: CNRS and INRIA from France, Fondation ISI and ICTP from Italy, IST from Portugal, and NorayBio from Spain. This consortium includes teams from several subdomains of computer science, biology and physics, and even multidisciplinary participants.

Gennetec takes inspiration from two biological cases of transcriptional interaction networks that exemplify many challenges in complex systems research, including elemental heterogeneity, interconnected scales in time and space, emergence of higher level behaviour, and hierarchies.

GENNETEC has exploited its case study to build a theoretical framework for modelling/simulating complex systems and for the analysis of their emergent properties. This development has made use of methods from formal software engineering and from qualitative approaches in statistical physics. It has dealt with the two-way information flow between two or more successive levels of model description, with particular attention on consequences for aggregate behaviour. To tackle complex systems with more modelling levels, hierarchies of descriptions have been studied.

Within this theoretical framework the consortium has developed bottom-up multi-scale methodologies to analyse global aggregation behaviours. Design of multi-stable systems has been handled using multiobjective optimization algorithms, borrowing from evolutionary computation and from message-passing techniques. These studies have been extended and generalized across other, non-biological applications to understand and design complex systems with desired aggregate behaviour, having stability and versatility / adaptability.

The biological outcome of these studies has been transferred into a scalable software tool to achieve a breakthrough in the bioinformatics domain of inference in transcriptional interaction networks. GENNETEC has also delivered a software tool for modelling/simulating complex networks, as well as reports and publications on the above-described issues.

The Strategic objectives of the project have been:

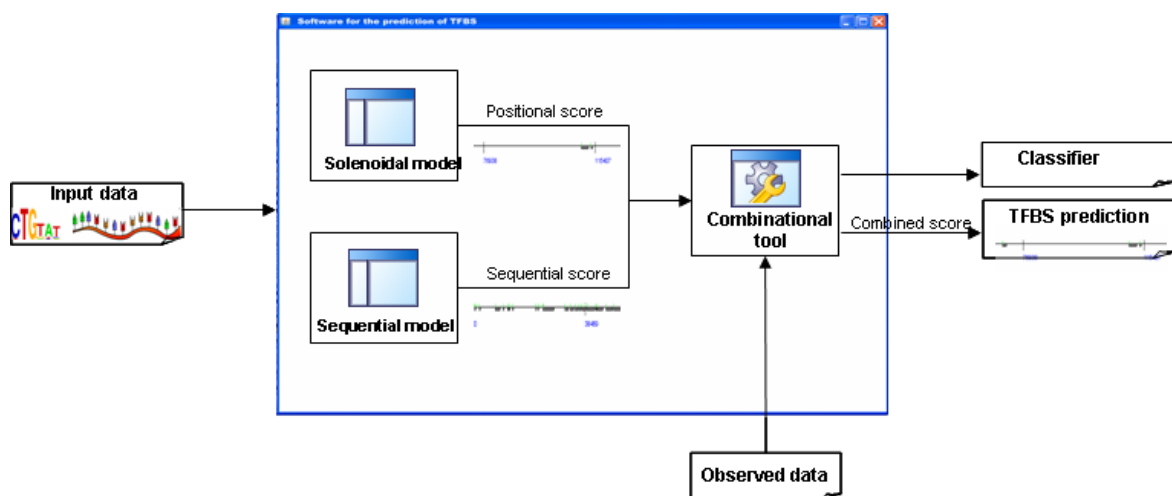
1. "Develop scalable computational modelling and inference tools and scalable simulation techniques for complex systems." This was the main strategic objective of the project, in order to succeed in developing the models of emergence of aggregate behaviour that permit the formulation of design strategies for systems with a specified aggregate behaviour.
2. "Projects should lead to breakthroughs in one of the following application areas: critical information infrastructures, bioinformatics, or systems engineering." Breakthroughs in the application area of bioinformatics are anticipated. This objective relates to technology transfer.

## NorayBio collaboration in the project

NorayBio was the leader of a work package focused on building software applications that integrate and make easy the access to the algorithms and methods emerging from all the partners involved in the project.

In that respect, the work has been focused on the development of a new software tool for the prediction of transcription factor binding sites. This application aims to improve that matter through the combination of classical methods and new methods developed under the project in a machine learning framework.

Sequential and solenoidal models are combined using observed interactions to obtain a classifier that will be used to predict new interactions.



## Final results

The main result of this Project obtained by NorayBio is a software platform for the prediction of transcription factor binding sites.

The final product is based on the combination of the algorithms based on gene positions along the chromosome, and the state-of-the-art algorithms classically based on the sequence of DNA regulatory sites combined in a machine learning framework that has evolved during the last half of the project life. Both, the algorithms obtained, and the way they are combined with the classical methods, had never before been used for transcription factor binding site prediction. Thus, this new software aims to be a new reference for this kind of studies.

November of this year, the software will be freely available at NorayBio's and Gennetec's Web site.