An outlook on SJ-INFOGEN working group

M. Jorio, V. De Fonzo, F. Aluffi-Pentini, V. Parisi

SJ-InfoGen, via Appennini, 24a - 00198 Roma

BACKGROUND

SJ-InfoGen skill and experiences come from Bio-informatics e Dynamical Genetics Group, headed by Valerio Parisi, and from Essegei, a Consulting Partnership (Strategic Marketing/ Technology and Management).

Dynamical Genetics

Recent research has shown a correlation between some genetic diseases and genomic sequences tandemly repeated a variable and excessive number of times, suggesting a genetic origin for a steadily increasing number of diseases (often neurodegenerative) showing some features not explainable by means of mendelian genetics.

The Group provides an interpretative framework (dynamical genetics) which appears to explain in a simple and unified manner many phenomena. Accordingly, a mechanism called *Tandem Repeat Lenght Manager* (TRLM) controls the repeats length in the genomic sequence and is responsible, when it goes awry, of the disease onset. This hypothesis has been supported by many plausibility arguments and by a great number of experimental data from the literature.

Within the above framework the Group is now proposing ñ as a contribution to a molecular study of the above mechanisms, and based on various supporting evidences and plausible arguments - the special importance of the DNA quadruplexes, as well as of the involved enzymes. This TRLM mechanism appears to act almost always by monitoring a DNA tract that has the characteristics of being a Variable Number Tandem Repeat (VNTR) and/or forming a DNA quadruplex; and at least one of them drives the TRLM.

The interpretation tools from Dynamical Genetics, as well as the information on RNA and DNA genomic tracts, can suggest, in certain cases, to design oligonucleotides ìactiveî for the pathology under examination.

BIO-INFORMATICS

The Group has studied the time evolution of protein networks made of kinases and phosphatases. These protein networks formally resemble attractor neural networks (ANNs), and, like ANN, their steady states can be thought of as encoding the input.

Inputs will be recognised as different, if they lead the network into different steady states. Equations for the kinase networks were derived, based on simple enzyme kinetics considerations, and the general properties of the solutions were studied with the tools of qualitative analysis of dynamical systems.

The results obtained from the simulations were in striking agreement with many experimental data, even with most parameters chosen over a wide interval, showing a great robustness of the property of the network to discriminate between to stimuli.

Our idea was of modelling signalling pathways which are used to build up complex networks of interacting proteins.

SCIENTIFIC SERVICES AND PRODUCTS

SJ INFOGEN is qualifying, developing and producing tools (from informatics) and scientific services to be supplied to genetics, biology and chemistry as well as for applications to medical therapy and for identification of active principles (constituents) for pharmaceuticals treatment of neurodegenerative or immune diseases.

PHARMA CONSTITUENTS IDENTIFICATION PROCEDURE

We are qualifying a Procedure to design oligonucleotides active in allergy treatment. The Group has developed an algorithm to identify oligonucleotides candidate - expected to influence the switch of classes of the immunoglobulines ruling the human being immune system - and suggested experiments to verify, with an iterated procedure, the biologic effectiveness. The experiments will be run in the University of Rome ìTor Vergataî.

SW PRODUCTS DEVELOPMENT

The Product n.1 is aimed to reveal and analyse special oligonucleotides repeated sequences in the framework of the Dynamical Genetics.

The Product n.2 is aimed to the numerical integration of systems of differential equations of great interest for biology.

Others SW Products are based on a new global discrete optimisation algorithm, analogous to simulating

annealing and genetic algorithms.
Our basic algorithm will be specialised to fit diversified applications: design of chemical libraries (Product n.3), identification of oligonucleotides or oligopeptides promising as pharmaceutical constituents (Product n.4), optimised modelling to describe genetic, immunology and combinatorial chemistry results (Product n.5). In more detail:
SW PRODUCT N. 1: Packages SW to identify and study the Variable Number of Tandem Repeats (VNTRs) in the genoma, of paramount importance for human health, as when they belong to human genoma (neudegenerative diseases), as well as when they belong to parasites genoma (tuberculosis, malaria, etc.).
SW PRODUCT N. 2: Packages SW to integrate ordinary differential equations systems, simulating metabolic and protein network Models.
The understanding of the biology usual processes, as well as of their pathological alterations and of their interactions with drugs requires to built and simulate suitable models.
SW PRODUCT N.3: Packages SW aimed to the design of chemical libraries. That to correspond to a need of combinatorial chemistry for the availability of intelligent libraries to make possible a fast and effective screening.
The task is therefore to create a suitable algorithm capable of optimised design of libraries as intelligent as possible; in other word to create a strong support for the Laboratories which run a great number of experiences in Chemistry and Biology. The computer processing elaboration, in fact, precedes and prepares every experimental campaign.
SW PRODUCT N.4: Packages SW aimed to identify and design oligonucleotides o oligopeptides promising as drug constituents for the treatment of several diseases.
A number of oligonucleotides e di oligopeptides shows some effects in pharmacology. Being not viable an exhaustive experimentation, heuristic methodologies become of paramount importance, in order to optimise the therapeutical efficacy with a limited number of experiments.
SW PRODUCT N. 5: Packages SW aimed to optimise descriptive models for genetics, immunology or combinatorial chemistry. The problem consists in identifying rules common to the enormous amount of data in the disciplines above, like the Quantitative Structure-Activity Relationships, describing molecules active. Our purpose - the approach being complementary to artificial intelligence, neuronal networks, Hidden Markov or Bayesian Models - is to find rational rules to supply the best interpretations of known results, and then, to generate the new ones.