ANALYSIS OF THE DEVELOPMENT AREA AS A BIOINFORMATICS RESEARCH- TEACHING AND EXTENSION IN THE META UNIVERSITY

Ana María Castañeda González [1], Diana Carolina Garcia Burbano [2]

[1] *Meta-University Corporation Chief Technology Research Center Engineering and Basic Sciences Raul Cuero Rengifo, Biologist Specialist Microbiology (Friedrisjh Schiller Jenia Germany), Specialist in Natural Resources and Environmental Ecology Education, Master in Rural Development (Pedagogical University Colombia and Technology), e-mail: ana.catañeda@unimeta.edu.co*
2 University Corporation Target - Exclusive Dedication Full Time Faculty, Chemistry, Specialization Candidate HSEQ (University Corporation of Meta), MBA Candidate at Integrated Quality Safety and Environment (Viña del Mar University), e-mail.: diana.garcia@unimeta.edu.co

RESUMEN

"The biological sciences are about to become the 21st century, what was unusual physical in the 20th century." Just as the discovery of DNA in the early de1950 we resulted in a profound revolution in biological understanding, we are now about to take such a leap, in which we move through computational tools that can be used to understand biological systems in all their complexity preserving and exploiting these systems in a sustainable manner "." computational biology is part of a larger revolution that affect how all science is conducted. This revolution is being driven by the generation and use of information in all forms and requires the development of intelligent systems to store and access to the information.

In the last ten years biology have been faced with new methods derived from computational area such as the development of new algorithms and mathematical applications, especially designed for integration and analysis of cluster data (Human Genome) that have generated from the application of the methodologies in biological research.

Palabras clave: Bioinformatics, Computing, Research, Genetics.

1. INTRODUCCIÓN.

The objective of this work is to establish a position on the possible development of bioinformatics at the University Target Corporation, as part of the teaching, research and extension will promote existing programs, as this program is interdisciplinary (biology, chemistry. mathematics. Computing, Informatics, Electronics. Genetics) in this area and thus ensure the survival and prosperity of this discipline at this time, is key to our University.

This paper discusses the themes and proposals for action, on the prospects of the University Corporation Meta implement and strengthen the science.

The growth of biologists data in recent years that have passed 606 DNA sequences stored in 1982 to over 182 million today, was driven by the development of technology chain reaction (PCR) in the year 1986 (Mullis, 1990), and now by the emergence of new high-throughput technologies called biological experimentation

These rapid advances have been particularly increased by the Human Genome Project is trying to completely sequence three billion nucleotides in the human <2> genome (3). The number of magazines and research reports, documents and tools for the analysis of these sequences are also increased. For this the life sciences need tools information technology and computing to prevent the degeneration of these data in an accumulation of nascent unconnected facts and figures.

Bioinformatics deals with organizing and presenting this information effectively with the globalization of the Internet and the data deluge Genome Project bioinformatics are going through a period of explosive growth and development. The WWW (World Wide Web) facilitates the exchange of this treasure, and changed the nature of learning by providing greater access to resources in a variety of media.

Bioinformatics as being interdisciplinary field where it joins the Biology, Computational Science, Chemistry and Mathematics is essential to solve many problems, such as:

• The data generated by sequencing projects that require new ways of processing and analysis.

• Requirements on better algorithms to thoroughly scan the biological databases.

• Runs vital experimentation that can be obtained by bioinformatics and biocomputacional tools saving time and efforts.

2. RESULTADOS

2.1Historia and current status
This report will discuss the history, the current state of bioinformatics, the tools developed for the integration of databases, software bioinformatics and facilitated access to diversity of information. In particular to discuss the access to:
• BioKleisli.

• Bio Portal.
• Biogenetic.
Access to international databases and Integration (BioKleisli).
The data obtained from genome sequencing projects were seen without success and users could not share their findings. With the rapid increase in the volume of data from recombinant DNA became necessary to allow researchers submit their data directly to the community and make the literature review of the sequences.
There are currently three institutions responsible for the distribution of DNA
• The National Center for Biotechnology Information (NCBI) in the U.S. with the online tool called Bankit, which provides a simple form based in the presentation of GenBank.
• The European Molecular Biology Laboratory (EMBL) system with data submission
• DNA Data Bank of Japan sequence submission tool called Sakura (S).

Besides these three main databases, there are over l00 specific databases. This generation of databases has become so varied and scattered that are no longer easy to access and use. Users of industries today are particularly uncomfortable system data because they do not want to put their data in the public domain, but also want to use these databases and if possible to integrate information from public databases and to generate new private information. Therefore, BioKleísli provides links between the complex and heterogeneous sources that are geographically dispersed data also allows the integration of databases, analysis software and visualization tools.

The Internet has become a valuable asset in the globalization of computer resources, exchange and fast communication although this new frontier has enormous potential, the power of the Internet has not yet been fully exploited. Today there are information in Internet about bioinformatics which is difficult to locate, information update on research and development {l + D).

2.2 History
When in 1953 Watson and Crick proposed the model of the double helix to explain the structure of DNA, not envisioned the tremendous volume of information exponentially be generated from the time <5> and would give origin to algorithmic problems likely to a highly organized and careful handling. In the following decades computational tools that made possible the analysis and resolution of questions already raised in the DNA structure OF the genetic information coding for proteins <6> in the structural properties of these and the factors regulating <71 3 Without clutch computing tools applied in biology began long before the emergence of the Internet. In 1960 their first biological databases and nucleotide metabolic proteins, which later with the use of computer algorithms bioinformatics and computational biology appears. The bioinformatics finds and uses patterns and structures inherent in biological data, such as gene sequences and the development of new methodologies to access and search databases <9, while the second, computational biology, refers to the physical and mathematical simulation of biological processes <10>.

The bioinformatics is an interdisciplinary science as well, whose story broke in two after it was first fully sequenced on a protein, insulin by Sanger and his colleagues at the University of Cambridge in 1955 (it). Where Sanger won the Nobel Prize for chemistry in 1958.

Subsequent sequencing methods other less wasteful and more efficient than Sanger, as Edman degradation reaction, ion exchange columns and electrophoresis, which contributed to the automation of the sequencing and development of amino acid libraries were developed <m.
But the development of the new science would not have been possible without the high-speed digital computers. In the 1970s, though with very limited availability, it appeared the Internet, where 15% of all research centers and universities in the United States have this service <13l,

That's when the diffusion of new techniques for sequencing DNA and proteins was increased and their sequences were stored in data banks. This necessitated the creation of algorithms to catalog and compare sequence, which is recognized as a pioneer Margaret Oakley Dayhoff (1925Ml983). Connoted researcher at Georgetown University Medical Center. The Dr. Dayhoff developed computational methods that allowed him to compare protein from sequence alignments between them. The first edition of the "Atlas" contained 65 protein sequences. These editions are cited more than 4,500 times and are an invaluable source of reference for scientists worldwide.

Throughout more than 40 years of existence, access to too many databases provides proteins including PSD was included. From the year 2002, PIR-PSD was associated with EBI (European Bioinformatics Institute) and SIB (Swiss Institute of Bioinformatics) to give rise to a unique database of protein sequence and function, known today as UniProt.
Currently, there are secondary bases, also called knowledge bases that they contain the accumulated biological knowledge necessary to understand the functioning and utility at all levels of organization of a living being (molecular, cellular, organism). For example, these databases include all families of proteins with functional domains and their three-dimensional structures, as well as different signaling pathways.

For the future, it is expected to have a complete computer representation of the cell and the organism in order to understand the principles that determine the high level of complexity of biological systems.

2.3 Bioinformatics as a pioneer of research multidisciplinary.
The bioinformatics is a new multidisciplinary research area within computer practice, having a strong relationship in the theoretical part of this science. Conversely, aided biology computer, either call ("computational biology") is to the analysis and interpretation of large data sets of structures and functional genomics, representation and modeling of functional networks of the cell, and the development of methods for the prediction properties of macromolecules and biological significance. These interactions are the subject of bioinformatics and are answered with the methods of computing.
From the perspective of biology, bioinformatics objectives include:

• The analysis and interpretation of information. Genomics through complex algorithmic process and organization of biological knowledge in the relevant databases. Here, methods to play confident predictions of properties of DNA and protein sequences a prominent role.
• The interpretation of the complex and dynamic processes of regulation and cell differentiation.
• Prediction of the structural properties of the models coevolution or comparative analysis of the known structures ("threading").
• The prediction of the interactions of molecules I molecule for functional analysis of networks.
• The identification and characterization of mutations as the cause cellular and genetic disorders (carcinogenesis related genetic disposition).
• Identification of objectives called differential; Genome analysis of the development of pharmaceuticals, for example to combat pathogens.
• The simulation of metabolic networks and regulation.
• Simulation of molecular interactions.
Applications include bioinformatics practically all areas of modern molecular biology and molecular medicine (Human Genome Project, population genetics and neurobiology).
From the perspective of computational sciences, the bioinformatics tasks are:
• On the one hand develop algorithmic methods to systematically explore the use of computers by getting information.

• In addition, to simulate life processes with computers to investigate faster (This approach is used among other things to solve unreached objects in the life sciences strategic objectives of the structure prediction biodiversity, molecules, and simulations using neural networks in neurobiology).

The development of complex algorithmic process requires the use of the methods of theoretical computer science. Not only in the life sciences, but also in areas such as physical and earth sciences including mathematics and social sciences, who lately have importance in computer aided simulation.

Bioinformatics methods focus on both the analysis of small genomes (bacteria, fungi), as well as eukaryotic cells • (Genomes model as C. elegans, A. thaliana, D. melanogaster) the identification and characterization of the genetic elements by sequence analysis applications and the development of algorithms and performance evaluation plays an important role in genome analysis bioinformatics (e.g. BLAST, FASTA, hidden Markov models for sequence comparison, genetic modeling, predictions secondary structure, and pattern recognition domains, etc..)

2.4 Methods Bioinformatics

The Bioinformatics application methods are varied, can be highlighted:
• Genome analysis of genomes (currently exist and are available to more than 20 genomes completely sequenced and made publicly available, including sequence analysis and functional predictions).
• The analysis of genetic and physical maps (microsatellites, "Fingerprints" call signature biometric, etc.).
• Sequencing by hybridization (SBH short) and changing sequence with biochip (Focus being researched and used by the company Affymetrix}.
• Systematic genome comparisons
• Structure prediction and classification of protein folding. (This area structural genomics, where an attempt to model the structure of proteins in a genome, either predicting or experimentally done).
• Molecular evolution (e.g., analysis of Neanderthal evolution and paleontology in general the origin of species).
• Expression analysis to study the biosynthesis, for example, gene disruption, stress, cell division, growth conditions, organ specificity, developmental stage.
• The proteome analysis to study the cellular concentrations of proteins and their modifications.
• The development based on sequence comparisons and interactions of molecules (target identification, development agents) (" drugs design ", "drug testing"} .
• Hybridization methods, for example, oligonucleotide chip making fingerprints and hybridization for genetic analysis.
• "Computing with Biomolecules" (i.e. the use of molecular biology to electrophoresis method and chain reaction or PCR polymerase {) calculating complex mathematical problems such as cryptographic decryption).
• "Profiles of expression to educate regulatory networks in RNA.
• Modeling and Simulation of neurobiological processes through neural networks.
 • Simulations with time-related ecology.

In the same way biology in bioinformatics has a molecular basis for the chemistry; finally, it should be mentioned that bioinformatics is adjacent to the cognitive sciences. Cognitive science is a science of interdisciplinary research skills and thought processes as well as the so-called bio-information, e.g. physiological.
In summary we can say that bioinformatics is set as integral part of the multidisciplinary research.

3. DISCUSION

3.1 Importance of Bioinformatics research in Colombia, in the goal and in our College.

Whereas at the present time the bioinformatics not only to the analysis of molecular data, integration of biodiversity data is one of the aspects of research and development in which bioinformatics groups can find a set of interesting problems and restricted relevant to solve, because although Colombia is one of the most bio diverse countries in the world, faces the challenge of starting various courses of action in a systematic and coordinated manner, to consolidate the full knowledge of the biodiversity. Bioinformatics provides the tools and concepts to systematize that knowledge.

Current Importance any worldwide project that seeks to do research in genomics, transcriptomics, proteomics, metabolomics, and any other omics, requires strong support bioinformatics in the development and implementation of new individual applications allow information generated profitable, especially if one considers that, based on the responses obtained using bioinformatics developments. Progress will be much faster and with greater confidence in biological experimentation of any kind. This means that not only should encourage and fund research in the area of bioinformatics in Colombia that has shown to be globally competitive regardless of biological experiments carried out in the country, but to make it more competitive generation biological knowledge must be integrated with equal rank to the lines of inquiry, to give them the competitive they need for their development.

This becomes more pressing when we consider that at this time our research are confronted necessarily the use of high-performance techniques for the generation of biological data, making it necessary to advance the implementation of a common bioinformatics platform for systematic management these data, especially those obtained in bulk by the application of technologies in areas such as genomics, transcriptomics, proteomics and metabolomics, contributing to lower development costs and implementation tools.

The development of bioinformatics in our university would have a virtual character as with the Swiss Institute of bioinformatics, but at the same time keeping them connected to their home institutions (universities and research institutes) in order that there will always share these within lines of each research institution.

In our case we could begin the coordination of human and technical resource that although limited currently exists. Strive for an interdisciplinary and inter-institutional research, some universities like the Valley, the National University of Colombia with groups like

Bioinformatics Institute of Biotechnology charge of the operation of Node Colombian European Molecular Biology Network (EMBnet) and Laboratory of Intelligent Systems Research (LISI), Faculty of Engineering Cenicafé, Universidad de los Andes. FIDIC and Xaverian University, among others to combine individual skills for the benefit of the academic community requires services, consultancy and training in bioinformatics.

The future of bioinformatics in Colombia was formed on the effort to conduct the research groups of our country together, to transform the information generated in the biological area knowledge and technological development.

3.2 Infrastructure needed:
• Training courses Bi <> information.
• Establishment of a modern electronic communications infrastructure (including videoconference).
• Availability of bioinformatics software teaching:
• Join efforts for the realization of the contents of the interdisciplinary course from the Biology, Chemistry and Pharmacy and the Institute of Information to establish a program • Perform interagency agreements with universities that project between research and teaching programs

FINAL COMMENT.

The bioinformatics is a key to the development of biology, Chemistry, pharmacy and medicine as well as for teaching research and Extension University of Meta Corporation is able to mount this given his leadership infrastructure, and the desire to delve into future disciplines doing this serious analysis has to our university offers all the conditions and opportunities for research and extension training for this discipline, opening a new field for our students in the job market.

4. REFERENCIAS BILIOGRAFICAS.

1. Dra. Mary Clutter (National Science Foundation), *mensaje al Comité del Senado de EE.UU. de Comercio, Ciencia y Transporte, 17.9.2006*

2. Hawkins TL ,McKernan KJ , Jacotot LB, MacKenzie JB, Richardson PM , Lander ES: Una atracción magnética a la genómica de alto rendimiento. Ciencia 276 (5320): 1887-1889 (20 de junio 1997)

3. Patrinos A, Drell DW : El Proyecto Genoma Humano: vista desde el Departamento de Energía. J Am MedAssocWomens 52 (1): 8-10 (1997).

4. Universidad de Berkeley Museo de Paleontología (UCMP) del sitio Web de Internet. Dirección de Internet: http://ucmp1.berkeley.edu/subway/bioinfo.html.

5. TuliMA , Flores TP , Cameron GN: Presentación de datos de secuencias de nucleótidos que EMBL / GenBank / DDBJ . Mol Biotechnol 6 (1) : 47-51 (agosto 1996).

6. El Centro Nacional de Información Biotecnológica (NCBI) Sitio Web de Internet. http://www.ncbi.nlm.nih.gov/.

7. El Laboratorio Europeo de Biología Molecular (EMBL) Sitio Web de Internet. http://www.embl-heidelberg.de/Services/index.html
8 El Banco de Datos de ADN de Japón (DDBJ) Sitio Web: http://www.ddbj.nig.ac.jp/

9. GenBank. (Fecha de acceso diciembre 3 de 2007). Disponible ttp://www.ncbi.nlm.nih.gov/GenBank/index.html
10. Gamow G, Rich A, Ycas M. The problem of information transfer from nucleic acids to proteins. Adv Biol Med Phys.1956; 4:23-68.

11. Pauling L, Corey RB, Branson HR. The structure of proteins: two hydrogen-bonded helical configurations of the polypeptidechain. Proc Natl Acad Sci USA.1951; 37:205-11.

12. Szent-Györgyi AG, Cohen C. Role of proline in polypeptide chain configuration of proteins. Science.1957; 126:697.5. Britten RJ. Davidson EH. Gene regulation for higher cells: a theory. Science.1969; 165:347-57.

13. Turing AM. The chemical basis for morphogenesis. PhilTrans R Soc London B.1952; 237:37-72.7. Horowitz NH. On the evolution of biochemical syntheses.Proc Natl Acad Sci USA.1945; 31:153-7.

14. NCBI Bioinformatics. (Fecha de acceso diciembre 3 de 2007).Disponible h:tp://www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html..