



The Present-Day Meaning Of The Word Bioinformatics

Raúl Isea

Fundación Instituto de Estudios
Avanzados - IDEA
Hoyo de la Puerta, Baruta,
Venezuela

ABSTRACT

The objective is to reevaluate the term of Bioinformatics and we show that Bioinformatics should be considered to be an interdisciplinary science, where different disciplines converge to solve biological problems.

Keywords

Algorithms Bioinformatics, Bioinformatics, Concept, Computational Biology, Grid Computing, Translational Bioinformatics.

1. INTRODUCTION

The According to the Collins' dictionary definition, Bioinformatics is defined as "the branch of information science concerned with large databases of biochemical or pharmaceutical information" [1]. According to this definition, we can deduce that the term Bioinformatics should be the result of combining 'bio' from biochemical or pharmaceutical data, and 'informatics' from the management of data through computer science.

However, this term goes back to 1965 with the work of Margaret Dayhoff (1925–1983) where she was a pioneer in using the computer to understand biochemical processes with her famous work entitled "Atlas of Protein Sequence and Structure" [2]. Thanks to this seminal work, an explosion of scientific publications in several different journals has occurred with the general commonality that the computer was used to describe biological processes.

On the other hand, some publications attribute the origin of the term of Bioinformatics to Hogeweg and Hesper in 1978 [3] where they referred to Bioinformatics as "the study of information processes in biological systems", including as such all living organisms. Therefore Bioinformatics focus on biological processes in order to explain living organisms using an informatics infrastructure. This definition is in accordance with the definition of Nilges and Linge who indicated that the term Bioinformatics is knowledge gained from a computer analysis of biological data [4].

Up to this point, we can think that the terms Bioinformatics or Computational Biology would be interchangeable terms. However, the Information Science and Technology Initiative Consortium of the National Institutes of Health in 2000 (led by Dr. Michael Huerta), redefined Bioinformatics and Computational Biology. They stated that Bioinformatics was the "research, development or application of computational tools and approaches to leverage maximum biological, medical, behavioral or health, including those tools that allow us to acquire, store, organize, archive, analyze, or visualize such



data” and that computational biology was “the development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, behavioral, and social systems” [5].

According to the last definition, we can see that the main difference between the two terms is that Bioinformatics focuses on the management of data while computational biology provides a validity using computational techniques.

At the present time, the volume of biological data is growing exponentially since the Haemophilus influenzae data were published in 1995 [6] and it is necessary to develop tools for the analysis of an increasing amount of information. This increase is due to new generation sequencing technologies and new proteomic or microarray equipment and so on. For example, Kumari and Kumar recently published that a simple human sequence is almost 140 gigabyte in size and the European Molecular Biology Laboratory stores approximately 20 petabytes of information [7].

At this point, we wonder if it makes sense to separate the concepts of Bioinformatics and computational biology because both are necessary to manage the accumulated volume of actual data to come to valid conclusions. By now, it is impossible to deny that Bioinformatics should reach beyond the data deluge with technologies independent of the concept of Bioinformatics.

At the present time, several techniques have been employed to manage the volume of information. For example, Grid Computing, allowing the analysis of computational models incorporating hundreds of thousands of processors, is a useful technique to analyze the data. For example, the European initiative called Wisdom (Wide In Silico Docking On Malaria) is dedicated to finding a new malaria drug using Grid Computing [8]. With this example, one could suggest a new emergent discipline called e.g. Bioinformatics Grid Computing or Parallel Computing on Bioinformatics. However, let me remember what is the reason in using Grid Computing? The answer is to solve a biological problem, and this is just the goal of Bioinformatics. For this reason, we believe that we cannot define a new scientific field for each technique, and only say that this last example, is just Bioinformatics using Grid Computing.

Let us present another illustrative example. Consider the diffusion process of genetic material through a virus capsid plant. Perhaps, it is possible to do this study using nuclear magnetic resonance (NMR) which registers the dynamics of atoms. It is difficult to isolate and elucidate the dynamic process in view of the tremendous number of atoms in this system. To do that, we consider a computational model obtained with Cowpea Chlorotic Mottle Virus that was resolved by X-ray diffraction data, and find that there was a spontaneous release of genetic material through the CCMV capsid without any energy and that the infectious mechanism arising from genetic material corresponds to a free release [9]. This example belongs to Structural Bioinformatics, which is just the analysis and prediction of three dimensional structures of macromolecules.

Therefore, we can conclude that the goal of Bioinformatics at the present time is to provide answers to biological problems. Bioinformatics is a rapidly developing branch of biology using techniques and concepts from informatics, but also statistics, mathematics, chemistry, biochemistry, physics and linguistics. It has many practical applications in different areas of biology and medicine.

Just at this point, it is necessary to indicate that Bioinformatics is an interdisciplinary discipline. For to do that, it is important to understand what is meant by an interdisciplinary science. Let us cite again the dictionary concerning the definition of Interdisciplinary which defines it as: "the quality or state of involving more than one discipline" [1]. So if we say that Bioinformatics is an interdisciplinary domain since it provides an integrated view of various sciences to explain a specific problem based on available biological information. In this context, Bioinformatics is not a multidisciplinary science, since the solution of the team will be given according to the field of action in each discipline, and in addition, the problem is not focused on solving a same problem. But as indicated above, we think that Bioinformatics is really an interdisciplinary science focused to solution a biological problem.

In fact, we find another branch of Bioinformatics called Translational Bioinformatics. This concept was introduced by Jake Chen and Maricel Kann (see for example 10), and is the interception between Bioinformatics and clinical and



medical application. With this example we see that the Bioinformatics generate a rich branch of emergent disciplines, a property absent in computational biology.

What happened with the term computational biology? In the past it often described the process of inserting and storing biological information in a database. However, this definition limits this term and possibly condemns it to oblivion. We know that it is necessary to store biological information, but the technical detail are resolved by computer science, and is not to be mixed with Bioinformatics dedicated to resolve biological problems. We believe that the computational biology should be called Algorithms Bioinformatics. This term appearing in the scientific literature is the correct name for computational biology, which is dedicated to the development of algorithms in order to integrate and analyze biological data.

2. CONCLUSION

Bioinformatics is an interdisciplinary science that allows one to understand and evaluate knowledge in the field of biology, developing a hypothesis using biological information. It is true that Bioinformatics began as a tool to just manage and analyze experimental data. With recent technologies, there has been a large deluge of data that must be analyzed in order to suggest a correct model for a biological episode.

We can say that Bioinformatics is interdisciplinary and not multidisciplinary, and is dedicated to the management of obtaining a valid solution to a biological problem because it encompasses various theories and methodologies from scientists from several different disciplines analyzing the same phenomenon.

The limitations that exist in Bioinformatics can be associated with the limitations of human knowledge of biological problems instead of the consideration of physical limitations of the information technology. However, it is not a problem of resources thanks to Grid Computing and other computational techniques. Therefore its most important goal is that it enables us to formulate experimentally verifiable hypotheses that may help to understand these biological processes and to tackle important problems.

3. ACKNOWLEDGMENTS

The author wishes to express his sincere thanks to Prof. Karl E. Lonngren and Johan Hoebeke for their unconditional help and the comments concerning of this manuscript.

4. REFERENCES

- [1] Collins' Dictionary available at <http://www.collinsdictionary.com/> Accesed 1 Dic 2014
- [2] Dayhoff, M.O., Eck, R.V., Chang, M.A., and Sochard, M.R. 1965. Atlas of Protein Sequence and Structure, Vol. 1. National Biomedical Research Foundation, Silver Spring, MD.
- [3] Hogeweg, P., and Hesper, B. 1978. Interactive instruction on population interactions. *Comput Biol Med.* 8, 319–327.
- [4] Nilges, M., and Linge, J.P. Unité de Bio-Informatique Structurale, Institut Pasteur, 25–28 rue du Docteur Roux, F–75015 Paris, France
- [5] Huerta, M. 2000. NIH working definition of bioinformatics and computational biology. Available from <http://www.bisti.nih.gov/docs/compubiodef.pdf> [cited: 1 Dec 2014].
- [6] Fleischmann, R.D., Adams, M.D., White, O., et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.* 269(5223), 496-512.



GLOBAL JOURNAL OF ADVANCED RESEARCH
(Scholarly Peer Review Publishing System)

- [7] Kumari, D., and Kumar, R. 2014. Impact of Biological Big Data in Bioinformatics. IJCA. 101(11), 22-24.
- [8] Kasam, V., Salzemann, J., Botha, M., et al. 2009. WISDOM-II: Screening against multiple targets implicated in malaria using computational grid infrastructures. Malar J. 8, 88.
- [9] Isea, R., Aponte, C., and Cipriani, R. 2004. Can the RNA of the Cowpea Chlorotic Mottle Virus be released through a channel by means of free diffusion? A Test in silico". Biophysical Chemistry. 107(1-2), 101-106.
- [10] Kann, M.G. 2010. Advances in translational bioinformatics: computational approaches for the hunting of disease genes. Brief Bioinform. 11(1), 96-110.