COMBIOMED: A Cooperative Thematic Research Network on COMputational BIOMEDicine in Spain

Fernando Martin-Sanchez, Victoria Lopez-Alonso, Isabel Hermosilla-Gimeno, Guillermo Lopez-Campos, and the COMBIOMED Network.

Medical Bioinformatics Department, Institute of Health "Carlos III", Madrid, Spain

fms@isciii.es

Abstract

The Cooperative Thematic Research Network on Computational Biomedicine COMBIOMED was approved in the last call for Thematic Networks in Health Research within the Spanish National Plan for Scientific Research, Development and Technological Innovation and it is funded for the period 2008-2011. The COMBIOMED Network is currently addressing various aspects that range from basic to applied research in science for the development of methods and tools to solve problems in biomedical science in the context of personalized medicine. This paper describes and analyses the organizational aspects and scientific areas in which this network has been focused (gene-disease association, pharmainformatics and decision support systems at the point of care). At the same time, COMBIOMED aims to play a central role in the education of researchers and in the training of health professionals in techniques for the processing of biomedical information.

1. Introduction

The COMBIOMED Network continues the work initiated by INBIOMED, the Cooperative Thematic Research Network on Biomedical Informatics (2004-2007) that developed a platform for the storage, integration and analysis of clinical, genetic, and epidemiological data and images focused on the investigation of complex diseases [1]. Computational biomedicine represents the interface between biomedical and computer sciences. It provides an inclusive environment for a better understanding of the biological processes that take place in each of the levels of organization of living organisms and the intricate network interactions between

them. One of the objectives of COMBIOMED is to establish contacts and to collaborate with the most relevant international initiatives in the field, such as the National Centers for Biomedical Computing [2], or the Biomedical Informatics Cores of the Clinical Science and Translational Awards [3] funded by the National Institutes of Health (NIH).

2. COMBIOMED: description and organization

Several of the 12 groups participating in COMBIOMED have previously participated and led some of the European Networks of Excellence (NoE) on biomedical informatics, bioinformatics and systems biology, such as INFOBIOMED [4] and BIOSAPIENS [5]. Previous experience in Spanish initiatives (INBIOMED [6], INB [7]) has been crucial to the creation and development of those European initiatives.

The design of the network, shown in Figure 1, consists of the following levels:

- Coordination and management.
- Computational aspects, which serve as instrumental support to the network including software and middleware, hardware, GRID, algorithms, and programming.
- Coordination of the work carried out by the research groups. This level addresses aspects such as data and text mining, clinical decision-making, electronic health records, image processing, disease simulation, and biomedical ontologies that help manage and integrate chemical, genetic, environmental, clinical and imaging data.
- Horizontal activities affecting all groups and lines of work. Particular attention is paid to the connection of the network with the scientific community and society (integrated Knowledge Management, education and training, communication, dissemination, Quality and Safety) (Figure 2).

3. COMBIOMED: scientific areas

COMBIOMED focuses on three scientific research areas: gene-disease association (Disease-omics), pharmainformatics and decision support systems at the point of care (Info-POC).

3.1 Gene disease association (Disease-omics)

The study of the molecular causes of disease and individual genetic variations allows deepening into personalized medicine [8] by developing safer and more efficient preventive, diagnostic and therapeutic solutions. The scientific community needs more advanced computational resources (functional analysis of genes and proteins in the context of genetic variability, alternative splicing…) [9], access to specific comparative genomic information (genomic data visualization) and prediction of the effects of individual mutations (SNPs) in the pathways and macro-molecular complexes with the consequent implications in the associated diseases.
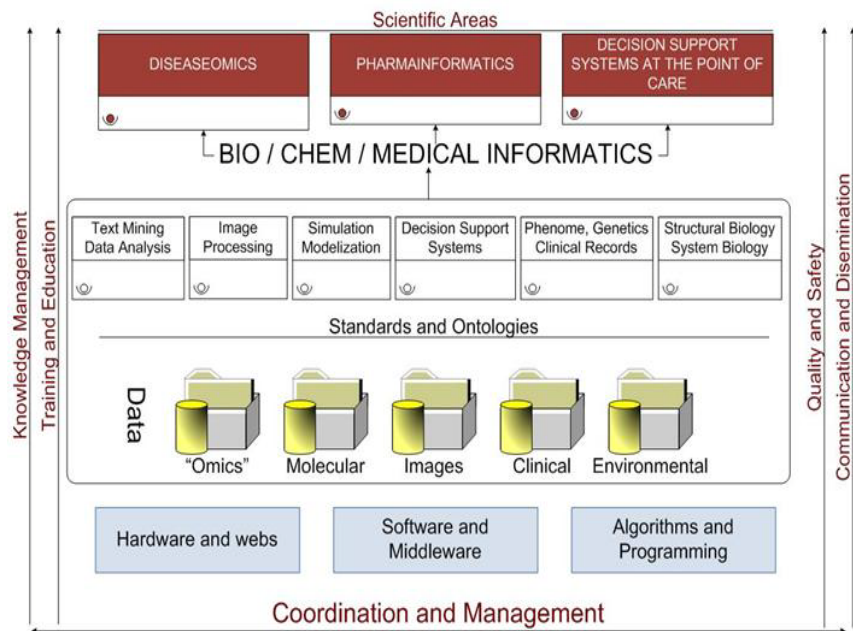

Figure 1. Graphical representation of the structure of the COMBIOMED Network

COMBIOMED works on these computational challenges in genotype-phenotype association and genomic epidemiology studies, to advance the understanding and modeling of the influence of environmental and genetic factors in the development of diseases. The network is using modules already developed by the National Institute of Bioinformatics (INB) to connect new methods that will be made available as Web services. This will help develop specific solutions for the analysis of genomic and clinical data.

The network is also developing systems that facilitate access to textual information about gene-disease relationships using automated information extrac-

tion methods and natural language processing with specific applications to problems of biomedical importance [10].
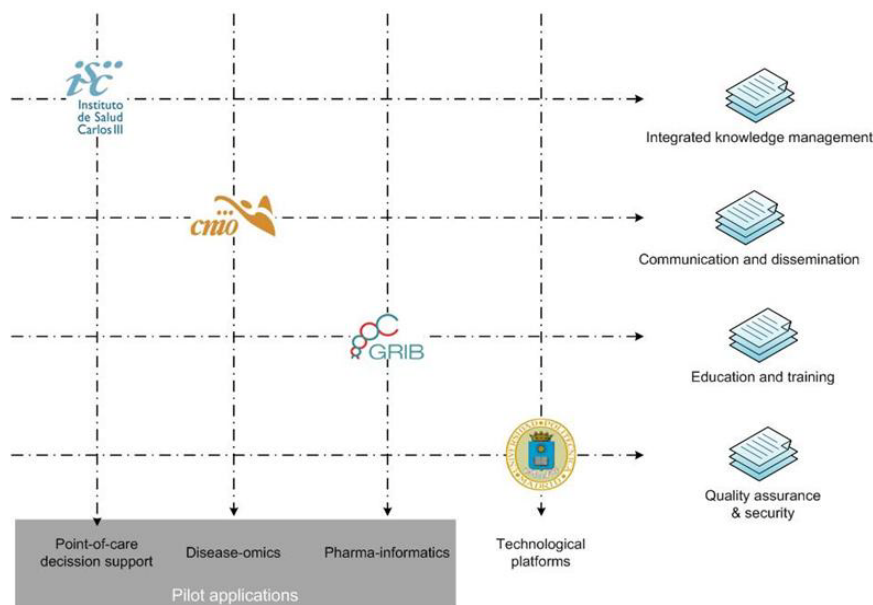

Figure 2. Organisation of scientific and horizontal activities within COMBIOMED

3.2 Pharma-informatics

The discovery and development of drugs is an area of great significance for human health, and at the same time it is an area of great socio-economic importance because it gives its "raison d'etre" to an industry which business is highly knowledge-intensive. Biomedical research in general and the R & D of drugs in particular, generate enormous amounts of data that require sophisticated computational tools for their management and analysis in order to extract the knowledge they need. This is one of the main reasons for the emergence of a new field of scientific activity that includes disciplines such as Computational Biology, and Biomedical Informatics.

Pharmaceutical research labs were pioneers in identifying the need and usefulness of computational approaches for the management and exploitation of the data generated in pre-clinical and clinical research. They are aware that certain

computational methods and their associated software can perform simulations and predictions that save time and investment in the development of drugs [11]. Computational approaches in systems biology are facilitating the management, visualization and development of predictive and descriptive mathematical models on interaction networks between biomolecular entities. This information is generated in the experimental laboratory largely based on the use of microarrays technologies [12].

Virtual screening and computer simulation techniques are very useful for the selection and testing of compounds to be considered in the initial stages of the design of a new drug. Moreover, the pharmacological and toxicological knowledge accumulated on the different groups of compounds allows for the development of quantitative models that can be used to perform *in-silico* prediction studies of the pharmacological and toxicological behavior of compounds not synthesized or tested.

Information technology also plays an important role in areas such as the management and exploitation of data from clinical trials. In addition, physiological advanced simulation techniques may allow the study of the behavior of organs of different individuals when exposed to drugs with different properties.

In coordination with the INB and the Spanish Technological Platform of Innovative Medicines [13], the COMBIOMED Network is developing technological solutions to facilitate the advancement of biomedical knowledge management geared towards the development of pharmaceutical R & D in all its stages.

3.3 Decision support systems at the point of care (INFO-POC)

In recent decades medical practice has sought greater integration of scientific knowledge in its routine. The tremendous growth of scientific knowledge and technological innovation requires the development of solutions that allow the use of a large amount of information in the clinical decision-making process. Within this context, Computational Biomedicine promotes the combination of techniques such as Medical Informatics (MI), bioinformatics (BI) and computing in the development of new methods and standards for clinical and biomolecular data integration and analysis [14]. At the same time, they facilitate a new approach that has as its overall objective to create a new integrated research framework for the development of diagnostic methods within the context of genomic medicine in the so-called "point of care".

The COMBIOMED network proposes the common research line of INFO-POC to carry out computational developments to represent and analyze clinical and biomedical knowledge at the point of patient care (POC). The collaboration between the diverse groups of the COMBIOMED network makes possible a continuous exchange of information and tools.

The network will support decision-making processes in a context of miniaturization of diagnostic systems and accessibility to information about molecular causes of diseases. This context is in line with recent trends on Convergent Technologies NBIC (Nano, Bio, Info and Cogno) with the objective of contributing to the development of a line of intelligent and miniaturized systems to be used at the point of care.

The availability and applicability of new technologies at the point of care could be a key incentive for translational research which may also imply a reduc-
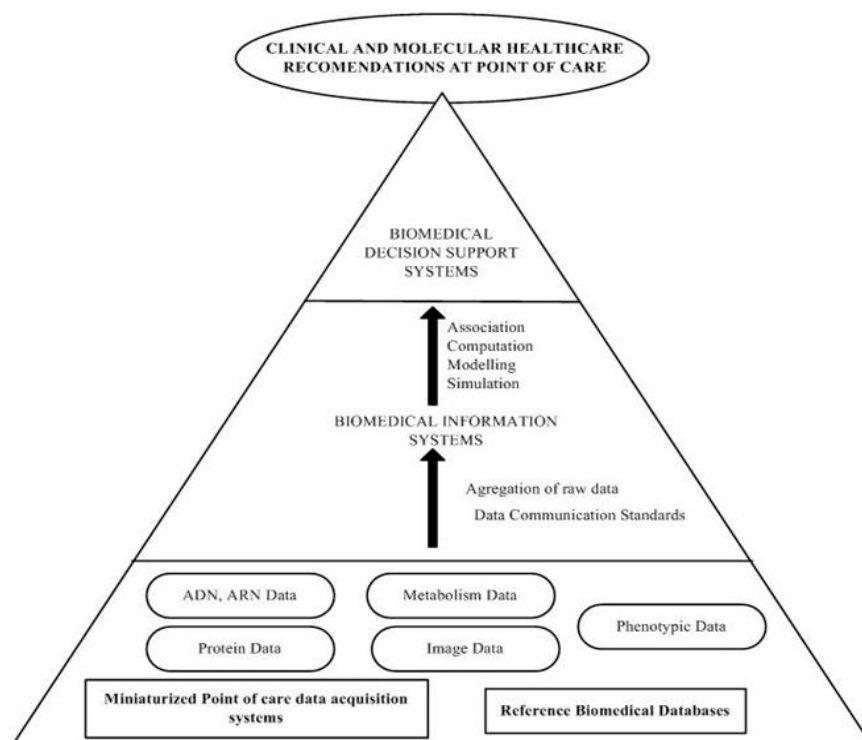

Figure 3. A conceptual diagram of the design of the INFO-POC pilot.

tion in the time devoted to decision-making.

The DNA microarray technology and the Bioinformatics tools that allow microarray data storage, management and analysis have enabled the development of diagnostic tests for complex diseases [15]. In addition to the biomolecular results obtained through these miniaturized point-of-care test systems there exists the requirement of placing molecular data (i.e. mutations in a gene, sequences of DNA, proteins…) in context, through the recovery of relevant information from reference databases (in silico), and its interpretation by implementing systems to support the diagnosis process (in info). The enormous complexity of cellular processes (metabolism, signal transduction, gene expression, and so on) needs the

development of new computational models and simulations to understand their behavior overall. The recent boost of systems biology and computational cell biology reflects this fact. The design of new computer-based methods in the Semantic Web for data recovery can contribute to the representation and computational analysis of biological knowledge at the POC. The knowledge generated will be integrated into computerized protocols for the diagnosis, treatment and management of patients (Figure 3).

The combination of bioinformatics and biomedical computing tools will facilitate the development of diagnostic models, supported by new standards. These tools need to be linked by using standard medical terminologies and coding with clear semantics to facilitate the effective implementation within clinical information systems.

Conclusions

The creation of the COMBIOMED Network represents a national and international reference in biomedical computing, which aims to provide solutions to the computational challenges posed by basic and translational research, and clinical practice in the context of the new personalized medicine. The most relevant research groups in Spain are cooperating to develop methods, systems, applications and pilot projects and to yield educational recommendations to promote biomedical computing research in the next years.

More specifically, computational developments within the COMBIOMED Network allow advancing in the representation and analysis of clinical and biomolecular knowledge, and the joint research will enable the new generation of miniaturized systems to support decision making with obvious clinical applications in health at the point of care.

References

[1] López V, et al. INBIOMED: a platform for the integration and sharing of genetic, clinical and epidemiological data oriented to biomedical research. 4th IEEE Intern. Symp. on BioInf.. and BioEng. 2004;222-226.

[2] NCRR webpage available at:
http://www.ncrr.nih.gov/clinical_research_resources/clinical_and_translational_science_awards/ Accessed on 02/26/2009.

[3] NIH BISTI webpage available at: http://www.bisti.nih.gov/ncbc/ accessed on 02/26/2009.

[4] The INFOBIOMED Network of Excellence webpage available at http:// www.infobiomed.org Accessed on 02/26/2009.

[5] BIOSAPIENS Network of Excellence webpage available at http://www.biosapiens.info Accessed on 02/26/2009.

[6] Cooperative Thematic Research Network of Biomedical Informatics INBIOMED web page available at www.inbiomed.retics.net/ Accessed on 02/26/2009

[7] Bioinformatics National Institute INB webpage available at http://www.inab.org/ Accessed on 02/26/2009.

[8] Sadee W, Dai Z. Pharmacogenetics/ genomics and personalized medicine. Hum Mol Genet. 2005;14: 207-214.

[9] Lopez-Bigas N, Audit B, Ouzounis C, Parra G, Guigo G. Are splicing mutations the most frequent cause of hereditary disease?. FEBS Lett. 2005:28:1900-1903.

[10] Krallinger M, Valencia A. Text-mining and information-retrieval services for molecular biology. Genome Biol. 2005;6(7):224.

[11] Jorgensen WL. The many roles of computation in drug discovery. Science. 2004;303:1813-1818.

[12] Butcher EC, Berg EL, Kunkel EJ. Systems Biology in drug discovery. Nat. Biotechnol. 2004;22:1253-1259.

[13] The Spanish Technological Platform for Innovative Medicines. http://www.medicamentos-innovadores.org/. Accessed on 02/26/2009.

[14] Alonso-Calvo R, Maojo V, Billhardt H, Martin-Sanchez F, García-Remesal M, and Pérez-Rey D. An Agent- and Ontology-based System for Integrating Public Gene, Protein and Disease Databases. J. Biomed. Inform. 2007;40(1):17-29.

[15] Vissers LE, Veltman JA, van Kessel AG, Brunner HG. Identification of disease genes by whole genome CGH arrays. Hum Mol Genet. 2005. 2: 215-223.