



HAL
open science

ProteINSIDE: a web service to computerize an in-depth analysis of functions, sequences, secretions and interactions for proteins

Nicolas Kaspric, Matthieu Reichstadt, Brigitte B. Picard, Muriel Bonnet, Jérémy Tournayre

► **To cite this version:**

Nicolas Kaspric, Matthieu Reichstadt, Brigitte B. Picard, Muriel Bonnet, Jérémy Tournayre. ProteINSIDE: a web service to computerize an in-depth analysis of functions, sequences, secretions and interactions for proteins. 11. Meeting on Research in Computational Molecular Biology (RECOMB) Comparative Genomics, Oct 2013, Lyon, France. pp.32, 2013, Proceedings of Eleventh Annual Research in Computational Molecular Biology (RECOMB) Satellite Workshop on Comparative Genomics, Lyon, France, 17-19 October 2013. ⟨hal-02746269⟩

HAL Id: hal-02746269

<https://hal.inrae.fr/hal-02746269v1>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

IWBBIO 2014

**INTERNATIONAL WORK-CONFERENCE ON
BIOINFORMATICS AND BIOMEDICAL ENGINEERING**

**April 7-9
Granada (Spain)**

Proceedings

IWBBIO 2014

**International Work-Conference
on Bioinformatics and
Biomedical Engineering**

**Granada
April, 7-9 2014**

volumen 1

**Proceedings
IWBBIO 2014
International Work-Conference on Bioinformatics
and Biomedical Engineering
volumen 1**

Editors and Chairs

Francisco Ortuño
Ignacio Rojas

Deposito Legal: 978- 84-15814-84-9

I.S.B.N: GR 738/2014

Edita e imprime: Copicentro Granada S.L

Reservados todos los derechos a los autores. Queda rigurosamente prohibida, sin la autorización escrita de los titulares del Copyright, bajo las sanciones establecidas en las leyes, la reproducción total o parcial de esta obra.

Preface

We are proud to present the set of final accepted papers for the second edition of the IWBBIO conference "International Work-Conference on Bioinformatics and Biomedical Engineering" held in Granada (Spain) during April 7-9, 2014.

The IWBBIO 2014 (International Work-Conference on Bioinformatics and Biomedical Engineering) seeks to provide a discussion forum for scientists, engineers, educators and students about the latest ideas and realizations in the foundations, theory, models and applications for interdisciplinary and multidisciplinary research encompassing disciplines of computer science, mathematics, statistics, biology, bioinformatics, and biomedicine.

The aims of IWBBIO 2014 is to create a friendly environment that could lead to the establishment or strengthening of scientific collaborations and exchanges among attendees, and therefore, IWBBIO 2014 solicited high-quality original research papers (including significant work-in-progress) on any aspect of Bioinformatics, Biomedicine and Biomedical Engineering.

New computational techniques and methods in machine learning; data mining; text analysis; pattern recognition; data integration; genomics and evolution; next generation sequencing data; protein and RNA structure; protein function and proteomics; medical informatics and translational bioinformatics; computational systems biology; modelling and simulation and their application in life science domain, biomedicine and biomedical engineering were especially encouraged. The list of topics in the successive Call for Papers has also evolved, resulting in the following list for the present edition:

1. **Computational proteomics.** Analysis of protein-protein interactions. Protein structure modelling. Analysis of protein functionality. Quantitative proteomics and PTMs. Clinical proteomics. Protein annotation. Data mining in proteomics.
2. **Next generation sequencing and sequence analysis.** De novo sequencing, re-sequencing and assembly. Expression estimation. Alternative splicing discovery. Pathway Analysis. Chip-seq and RNA-Seq analysis. Metagenomics. SNPs prediction.
3. **High performance in Bioinformatics.** Parallelization for biomedical analysis. Biomedical and biological databases. Data mining and biological text processing. Large scale biomedical data integration. Biological and medical ontologies. Novel architecture and technologies (GPU, P2P, Grid,...) for Bioinformatics.
4. **Biomedicine.** Biomedical Computing. Personalized medicine. Nanomedicine. Medical education. Collaborative medicine. Biomedical signal analysis. Biomedicine in industry and society. Electrotherapy and radiotherapy.
5. **Biomedical Engineering.** EComputer-assisted surgery. Therapeutic engineering. Interactive 3D modelling. Clinical engineering. Telemedicine. Biosensors and data acquisition. Intelligent instrumentation. Patient Monitoring. Biomedical robotics. Biotechnology. Genetic engineering.
6. **Computational systems for modelling biological processes.** Inference of biological networks. Machine learning in Bioinformatics. Classification for biomedical data. Microarray Data Analysis. Simulation and visualization of biological systems. Molecular evolution and phylogenetic modelling.
7. **Healthcare and diseases.** Computational support for clinical decisions. Image visualization and signal analysis. Disease control and diagnosis. Genome-phenome analysis. Biomarker identification. Drug design. Computational immunology.

8. **E-Health.** E-Health technology and devices. E-Health information processing. Telemedicine/E-Health application and services. Medical Image Processing. Video techniques for medical images. Integration of classical medicine and E-Health.

After a careful peer review and evaluation process (each submission was reviewed by at least 2, and on the average 2.8, program committee members or additional reviewer), 197 papers were accepted for oral, poster or virtual presentation, according to the recommendations of reviewers and the authors' preferences.

During IWBBIO 2014 several Special Sessions will be carried out. Special Sessions will be a very useful tool in order to complement the regular program with new and emerging topics of particular interest for the participating community. Special Sessions that emphasize on multi-disciplinary and transversal aspects, as well as cutting-edge topics are especially encouraged and welcome, and in this edition of IWBBIO 2014 are the following:

1. **SS1: Multi-biomarker and informatics in cancer diagnosis** The session will discuss the joint research between Hospital Doctors and Informatic Scientists. With the development of life science, the clinical laboratory could provide much more test for patients than before, currently this use reference interval or cut off values as clinical diagnostic standard which could not significantly improve the specificity and sensitivity simultaneously. If the informatic tools could be used to combined analysis multiple clinical test and biomarkers, it will be enhance the using of clinical information from medical laboratory and also could stimulate the translational medicine studies of "omic" technology and theory.

Organizer: Dr. Prof. Yaping Tian, Department of Clinical Biochemistry, Chinese PLA General Hospital, Beijing (China).

2. **SS2: Discovery of non-coding and structured RNAs** Non-coding (nc)RNAs are emerging as some of the most versatile and important biological molecules in the cell. They can act both in cis and in trans to mediate functions as diverse as catalysis, metabolite sensing, regulation of gene expression, epigenetics, chromatin stability, splicing, and more. The explosion of RNA sequences generated by next-gen sequencing and accumulating in various databases represents a massive, and relatively un-explored, "New World" of ncRNAs. The identification of novel ncRNAs from these data is a crucial, but challenging, problem for bioinformatics. Many ncRNAs require that they fold into thermodynamically stable RNA structures to carry out their functions. If these functions are evolutionarily conserved, then structures may also be conserved between related sequences. Thus, identifying thermodynamically stable and conserved RNA structures from sequence data can help identify putative ncRNAs. This session will cover various approaches for modeling RNA structures using thermodynamics, biochemistry, and sequence comparison, identifying homologous RNA sequences/structures, and using these methods to suggest which sequences have likely non-coding functions. This session will also cover the application of these methods to identify ncRNAs in target species.

Organizer: Dr. Walter N. Moss, Yale University and Howard Hughes Medical Institute, New Haven (USA).

3. **SS3: Biological Knowledge Visualization** Biotechnology produces large amounts of data, to be handled and analyzed by computational methods. The scale and complexity of such input data, but also of the analysis results, is large, and can be tackled by visual techniques for representation and interpretation. The focus of this special session is the discussion of approaches that integrate visual intelligence into the analytical process of such biological data. Contributions may range from the development of novel visual

techniques to the successful application of existing ones in biological or biomedical studies.

Organizer: Dr. Rodrigo Santamaria, University of Salamanca (Spain).

4. **SS4: High Performance Computing in Bioinformatics** The goal of this special session is to explore the use of emerging parallel computing architectures as well as High Performance Computing systems (Supercomputers, Clusters, Grids) for the simulation of relevant biological systems and for applications in Bioinformatics, Computational Biology and Computational Chemistry. We welcome papers, not submitted elsewhere for review, with a focus in topics of interest ranging from but not limited to: -Parallel stochastic simulation. -Biological and Numerical parallel computing. -Parallel and distributed architectures. -Emerging processing architectures (e.g. GPUs, Intel Xeon Phi, FPGAs, mixed CPU-GPU or CPU-FPGA, etc). -Parallel Model checking techniques. -Parallel algorithms for biological analysis. -Cluster and Grid Deployment for system biology. -Biologically inspired algorithms.

Organizers: Dr. Horacio Perez-Sanchez and Dr. Jose M. Cecilia, Catholic University of Murcia (UCAM), (Spain).

Dr. Ivan Merelli, Institute for Biomedical Technologies, National Research Council of Italy, Milano (Italy).

5. **SS6: ePathology - Realities and Perspectives** The session will discuss the current research and ongoing activities implemented in the field of ePathology. Digital imaging as well as virtual pathology boards acquire more and more importance. Application of medical and laboratory information management systems for pathology purposes is also important and actual. Special importance acquires implementation of eLearning technologies for continuous medical education of pathologists as well as introduction and practical application for realization of quality assurance programs in pathology and cytology.

Organizer: Dr. Ekaterine Kldiashvili, Georgian Telemedicine Union (Association), Tbilisi (Georgia).

6. **SS7: Modelling of cellular pathways and disease** This session will discuss applications of mathematical and computational techniques to the representation of biological processes or pathways, with the aim of gaining a mechanistic understanding of cellular functions. A wide range of approaches will be considered, ranging from qualitative network representation to fully quantitative kinetic models. Applications may focus on a specific metabolic, signalling and regulatory pathway, or on the contrary offer genome-scale coverage of processes implicated in a particular biological function or disease. We welcome interdisciplinary papers presenting the integration of modelling techniques with experimental analyses.

Organizers: Dr Jean-Marc Schwartz, Faculty of Life Sciences, University of Manchester (UK).

Prof Marija Krstic-Demonacos, School of Environment & Life Sciences, University of Salford (UK).

7. **SS8: Integration of data, methods and tools in biosciences** With the rise of biotechnology and bioinformatics, a number of problems regarding storing, searching and using biological data occur. The data are scattered over a large number of repositories (public or private) and stored in a large number of different formats. Furthermore, some formats are not adequate for automatic computer processing (such as text documents) and require some kind of preprocessing before they can be input into computer algorithms. This situation makes searching and analyzing the data very difficult, leaving

facts and knowledge on observed biological phenomena hidden in databases and digital collections. The above problems and other similar problems can be overcome only by an integrative bioinformatics approach. The integration can be done across different aspects and on various levels. Topics of interests include (but are not limited to): -database integration techniques, -data acquisition from heterogeneous sources, -genotype-phenotype associations researches, -integrative modeling and analysis of processes and systems in biosciences, -tool integration and workflow systems, -computational infrastructure for biological researches, including laboratory management systems, -biological ontologies and metadata, -integrative data and text mining approaches. Contributions that address any other aspect of integrating data, methods, techniques and/or tools, are also welcome. Organizer: Dr. Vesna Pajic, University of Belgrade, Faculty of Agriculture (Serbia).

8. **SS9: Biomaterials in Biomedicine: Computational approaches** The development of materials that can successfully replace biological tissues both in function and appearance, the so-called Biomaterials, is a field of growing interest in Biomedicine. The use of computational methods to determine or predict the physical properties of these materials has become one of the main focuses of the latest research in this area, allowing pre-clinical non-invasive methods for evaluation and testing of this type of materials. Also, theoretical modelling of these materials can expedite research by allowing conducting simulated experiments in order to find the solution in the real system that is being studied in response to changing conditions, and helps obtaining information on optical, mechanical and many other potentially interesting properties. The main topics of interest of this Special Session include (but are not limited to): -Computational methods for assessment and prediction of physical (including mechanical, optical,...) properties of biomaterials; -Application of biomaterials in Biomedicine; -Modelling systems of biomaterials; -Software tools and simulation packages for the evaluation of suitability of biomaterials;

Organizers: Dr. Razvan Ghinea, Department of Optics, University of Granada (Spain).

Dr. Luis Javier Herrera, Department of Computer Architecture and Computer Technology, University of Granada (Spain).

Dr. Maria del Mar Perez, Department of Optics, University of Granada (Spain).

9. **SS10: Effective Soft Computing Methods for Biomedical Signals** The goal of this special session is to elaborate applications of soft computing methods for biomedical signals such as ECG, EMG, and EEG. This session will discuss new hybrid algorithms and investigate effective as well as high performance computing techniques for the classification of biomedical signals for diagnosis or the other applications in Biomedical Engineering and Bioinformatics. We welcome papers, not submitted elsewhere for review, with a focus in topics of interest ranging from but not limited to: -Artificial Neural Networks Algorithms on Biomedical Signals. -Fuzzy Systems and Fuzzy Clustering Algorithms on Biomedical Signals. -Probabilistic Model Algorithms on Biomedical Signals. -Metaheuristic Algorithms on Biomedical Signals. -Hybrid Algorithms on Biomedical Signals.

Organizer: Prof. Dr. Bekir KARLIK, Faculty of Engineering, Selcuk University, Konya, (Turkey)

10. **SS11: Chaperone Therapy for Protein Misfolding Disorders with Brain Dysfunction** Chaperone therapy is a new concept of molecular therapeutic approach, first developed for lysosomal diseases, based on a paradoxical molecular phenomenon involving lysosomal enzyme proteins and their competitive inhibitors as intracellular enhancers

(chaperones). The misfolded mutant enzyme protein is stabilized as a molecular complex with its substrate analogue, chaperone, and transported safely to the lysosome. The complex is automatically dissociated in the lysosome under the acidic condition, the free mutant protein remains stable, and the enzyme activity is expressed. The small chaperone molecule has been confirmed to be delivered to the brain tissue through the blood-brain barrier. This new trial was targeted initially at a few lysosomal diseases, and then has been expanded to many other diseases with pathological protein dysfunction due to structural misfolding. The advantages of this molecular approach over other currently available therapies for genetic and protein misfolding diseases are summarized twofold; first, oral administration to individuals with intractable diseases; and second, delivery to the central nervous system for therapy of brain diseases currently not available with other experimental or clinical trials. In this session we discuss the present status of chaperone therapy research, focusing on chemical compounds, enzyme-chaperone interactions, and animal and human experiments for a few genetic diseases. This therapeutic approach will become a novel and revolutionary scientific and medical strategy in the near future.

Organizer: Dr. Yoshiyuki Suzuki, Tokyo Metropolitan Institute of Medical Science (Japan).

11. **SS13: Computational analysis of gene regulatory elements with next-gen sequencing data** The advent of next-generation sequencing (NGS)-based epigenetic profiling assays opens new perspectives for studying gene regulation. In particular, ChIP-Seq against histone marks and transcription factors (TFs), BS-Seq for DNA-methylation profiling and various protocols for assessing chromatin accessibility allow for a detailed characterization of the chromatin states of tens of thousands of cis-regulatory elements at once. Moreover, high-throughput protocols for sequencing RNA 5'ends allow for genome-wide monitoring of transcription initiation events at single base pair resolution. This session focuses on novel computational approaches to extract insights on gene regulatory mechanism from the vast amounts of epigenetic profiling data that have been accumulated over the last few years. We welcome contributions on topics ranging from but not limited to: -Promoter inference and classification from transcript mapping data. -Functional classification of regulatory elements based on chromatin features. -Methods for identifying differentially modified chromatin regions. -Prediction of nucleosome positioning from DNA sequence and TF binding events -Discovery of genetic variants associated with changes in chromatin state. -Usage of epigenetic profiling data for medical diagnosis and treatment decision

Organizer: Dr. Philipp Bucher, Swiss Federal Institute of Technology in Lausanne (EPFL) (Switzerland).

12. **SS14: Better Oncology Treatment and Patient Outcomes by Using Therapy-Related Symptom Checklists (TRSC/TRSC-C) and a Computerized Two-Way Communication System** This special session will describe and illustrate the development/measurement properties, current use, and future electronic uses of patient-friendly symptom checklists for adults (TRSC) and children (TRSC-C) oncology patients. These checklists can be completed in less than 5 minutes by adults and children/parents, and involve no radical alterations in or increased costs of clinical practice. The adult checklist was developed 1984-1995 and the child checklist 2004-2010. Both checklists now have versions in English, Spanish, Thai, Pilipino, Chinese, and Bahasa Indonesia, and have been used in clinic settings in the USA and other countries. The checklists were originally developed to reduce observed under-documentation of treatment symptoms of

concern to patients in medical records. The hope was that use of these checklists (25 symptoms for adults and 30 for children) would lead to better documentation through integration of the checklists into electronic medical records (EMR), enhance communications among patients and clinicians, and improve health outcomes. Completed research has found high levels of patient/clinician satisfaction with use of the TRSC, no increase in clinic costs, and strong correlations of TRSC/TRSC-C scores with the number of patient symptoms documented and managed, patient functional status, and patient quality of life. A recently published sequential cohort trial with adult outpatients at a Mayo Clinic Health System community cancer center reported that use of the TRSC produced a 7.2 higher covariate adjusted patient quality of life, 116 more symptoms documented and managed, and higher functional status. These results were statistically and substantively significant. This special session will illustrate transitions among design, measurement, application, and informatics required for good patient care. Audience participation will be solicited, and establishment of relationships among the audience for work along the lines presented and discussed in this special session will be encouraged.

Organizers: Dr. Arthur R. Williams, Research Associate, CINDRR, US Department of Veterans Affairs, and Professor, Department of Healthcare Policy and Management, College of Public Health, University of South Florida, Tampa (USA)

Dr. Phoebe D. Williams, Professor, School of Nursing, Kansas University Medical Center, Kansas City, Kansas (USA)

13. **SS15: Computational MRI: Theory, Dynamics and Applications** Computational techniques are invaluable to the continued success and development of Magnetic Resonance Imaging (MRI) and to its widespread applications. New processing methods are essential for addressing issues at each stage of MRI techniques. Magnetic Resonance Imaging simulations based on the Bloch NMR equations are of high educational value. They serve as essential tools in basic MRI method development, sequence design and protocol optimization. In this special session, the underlying physical and biomedical models of the Bloch NMR flow equations, their field of applications and possible limitations will be discussed. Magnetization preparation will be simulated in order to tailor sequence protocols for specific applications, exploiting basic spin relaxation as well as advanced Magnetic Resonance contrast mechanisms such as flow and diffusion. The main objective of this special session is to bring together the mathematicians, computer scientists, theoretical physicists, medical scientists and the engineer and apply their tools through high quality paper presentations and contribute to this fast developing and most exciting field of our time without acquiring the most sophisticated equipment. Volume I of a new book titled "Theory, Dynamics and Applications of MRI" will be released and presented at this special session. The book is intended to present basic theory of MRI and develop several fundamental equations which can be invaluable for quantitative and qualitative analysis of NMR magnetizations and signals. Based on this special session, scientists should be able to apply basic MRI methods to solve real life problems using computational methods of their choice.

Organizer: Dr. Omotayo Bamidele Awojoyogbe, Department of Physics, Federal University of Technology, Minna, Niger-State (Nigeria)

14. **SS16: Bioinformatical Approaches to Disordered Proteins** According to the "structure-function-paradigm" a stable, folded structure is a pre-requisite of protein function. However, since the turn of the century it became evident via an increasing number

of known examples that many proteins are able to serve crucial functions in vivo without adopting a well-defined 3D structure. In accordance with the typical functions these Intrinsically Unstructured/Disordered Proteins (IUPs/IDPs) serve in signalling, regulation and transcription, they were found to be at the heart of many diseases such as cancer, neurodegenerative diseases and diabetes, among others. This sparked interest in the focused research of IUPs not only at a basic science level but also in biomedicine/pharmacology. However, due to biological and technical reasons, experimental study of these proteins remains difficult and expensive and therefore the exact biological function, mode of action and biophysical/thermodynamical description of the majority of IUPs remain elusive. Because of these difficulties, bioinformatics tools that target protein disorder play an important role in the identification and characterization of IDPs. This is exemplified by the fact that the majority of our systems-, network- and evolutionary level knowledge of IUPs are based in various bioinformatics prediction methods and analyses. In this session we present the concept of IUPs and focus on the current, state-of-the-art bioinformatics approaches, tools and results of analyses concerning protein disorder.

Organizer: Prof. Istvan Simon, Institute of Enzymology, Research Centre for Natural Sciences (Hungary)

15. **SS18: Stochastic Modelling of Biological Systems** Recently there has been a significant interest in the stochastic approach for modelling biological systems, mainly because experimental data are providing evidences that random events play significant roles in determining the complex behaviours observed in living organisms. The growing amount of evidences arising from experimental observations at the single cell level are showing that fundamental biological processes such as, e.g., fate decision making, gene expression regulation, phenotypic variability are deeply conditioned by random fluctuations at the molecular level. Approaches grounding on computational stochastic modelling and simulations are proving to be useful tools for gaining insights about the role played by random events in determining the global dynamics of biological networks. In these cases deterministic descriptions fail both in predicting the observed random fluctuations and in capturing stochastic-driven phenomena such as stochastic focusing, stochastic switching and multiplicative noise effects. The most popular strategy for building stochastic models of biological phenomena consist in describing the temporal evolution of the considered system as a discrete-state, continuous time Markov process. Recently, an increasing number of works proposing non-Markovian methods is emerging, with the aim of providing more accurate representations of the random events observed experimentally. The need for non-Markovian modeling is also related to the finding of long-range memory and non-ergodicity. In this special session we aim at collecting original research works reporting on topics related to the stochastic modelling of biological systems at different levels, ranging from the inter-molecular scale to the inter-cellular scale, including neural networks. Particular attention will be devoted to those proposals highlighting how the proposed modeling approach allows to gain insights on the investigated biological phenomenon and on the emerging properties of biological networks.

Organizers: Dr. Paolo Paradisi and Dr. Davide Chiarugi, Institute of Science and Technologies of Information (ISTI) - Italian National Research Council (CNR), Pisa (Italy)

This second edition of IWBBIO was organized by the Universidad de Granada together with the Spanish Chapter of the IEEE Computational Intelligence Society. We wish to thank to our main sponsor BioMed Central, e-Health Business Development BULL (España) S.A.,

and the institutions Faculty of Science, Dept. Computer Architecture & Computer Technology and CITIC-UGR from the University of Granada for their support and grants. We wish also to thank to the Editor-in-Chief of different international journal for their interest in editing special issues from the best papers of IWBBIO.

We would also like to express our gratitude to the members of the different committees for their support, collaboration and good work.

April, 2014
Granada

IWBBIO Editors and Chairs
Francisco Ortuño
Ignacio Rojas

Program and Steering Committee

Irina Abnizova	Wellcome Trust Sanger Institute
Jesus Aguilar	University of Pablo Olavide
Carlos Alberola	University of Valladolid
Patrick Aloy	Institute for Research in Biomedicine (IRB)
Rui Alves	Universitat de Lleida
Miguel Andrade	Max Delbrck Center for Molecular Medicine (MDC)
Eduardo Andres	Spanish National Cancer Research Centre (CNIO)
Jose Andreu	Centre of Biological Research (CIB-CSIC)
Marco Antoniotti	Universit degli Studi di Milano Bicocca
Ana Aransay	Center for Cooperative Research in Biosciences (CIC bio-GUNE)
Saul Ares	Spanish National Center for Biotechnology (CNB)
Rubén Armañanzas	Technical University of Madrid (UPM)
Antonia Aránega	University of Granada
Gualberto Asencio Cortés	"Pablo Olavide" University
Bamidele Awojoyogbe	Federal University of Technology in Minna
Jaume Bacardit	University of Nottingham
Esteban Ballestar	Chromatine and Disease Group (IDIBELL)
Gianni Barcaccia	University of Padova
Ugo Bastolla	Research Center of Molecular Biology "Severo Ochoa"
Paul Bates	Cancer Research UK London Research Institute
Oresti Baños	University of Granada
Shomi Bhattacharya	Andalusian Molecular Biology and Regenerative Medicine Centre (CABIMER)
Alfredo Benso	Politechnic University of Turin
Concha Bielza	Technical University of Madrid (UPM)
Jose M. Blanca	Technical University of Valencia (UPV)
Armando Blanco	University of Granada
Ignacio Blanquer	Technical University of Valencia (UPV)
Andrs Bueno	Catholic University of Murcia (UCAM)
Philipp Bucher	Swiss Federal Institute of Technology in Lausanne
Juan Calvete	Biomedicine Institute of Valencia (IBV-CSIC)
Daniela Calvetti	Case Western Reserve University
Mario Cannataro	University Magna Gracia of Catanzaro
Carlos Cano	University of Granada
Emidio Cappriotti	University of Alabama at Birmingham
Jose Maria Carazo	Spanish National Center for Biotechnology (CNB)
Pablo Casaseca	University of Salamanca
Nazareth Castellanos	Complutense University of Madrid
Monica Chagoyen	Spanish National Center for Biotechnology (CNB)
M. Gonzalo Claros	Andalusian Bioinformatics Platform
Jean Claverie	National Centre of Scientific Research (CNRS)
Jacques Colinge	Research Center for Molecular Medicine of the Austrian Academy of Sciences
Jose M. Cecilia	Catholic University of Murcia (UCAM)
Ana Conesa	"Principe Felipe" Research Center (CIPF)

Christian Conrad	German Cancer Research Center (DKFZ)
Juan Cruz Cigudosa	Spanish National Cancer Center (CNIO)
Alfonso Marquez Chamorro	Pablo de Olavide University
Davide Chiarugi	Italian National Research Council (CNR)
Guillermo de La Calle	Technical University of Madrid (UPM)
Carmen de Mendoza	"Carlos III" Hospital of Madrid
Coral Del Val	University of Granada
Ramón Díaz Uriarte	University Autnoma of Madrid (UAM)
Jose Díaz	Centre of Biological Research (CIB-CSIC)
Hernan Dopazo	"Principe Felipe" Research Center (CIPF)
Joaquin Dopazo	"Principe Felipe" Research Center (CIPF)
Juergen Eils	German Cancer Research Center (DKFZ)
Ernesto Estrada	University of Strathclyde
Jose Jesus Fernandez	University of Almeria
Sylvain Foissac	Center for Genomic Regulation (CRG)
Jean-Fred Fontaine	Max Delbrck Center for Molecular Medicine (MDC)
Lutz Froenicke	University of California Davis
Hans Gabius	University of Munich
Juan García Gómez	Technical University of Valencia (UPV)
M. Teresa Garca Valverde	Catholic University of Murcia (UCAM)
Razvan Ghinea	University of Granada
Paulino Gmez-Puertas	"Severo Ochoa" Molecular Biology Center
Daniel Gonzalez	University of Vigo
Humberto Gonzalez	University of Santiago de Compostela
Juan Ramón Gonzalez	Center for Research in Environmental Epidemiology (CREAL)
Marina Gordaliza	University of Salamanca
Ananth Grama	Purdue University
Raik Gruenberg	Center for Genomic Regulation (CRG)
Concettina Guerra	University of Padova
Alberto Guillen	University of Granada
Roderic Guigo	Center for Genomic Regulation (CRG)
Stefan Götz	"Principe Felipe" Research Center (CIPF)
Christophe Guyeux	IUT Belfort-Montbeliard
Michael Hackenberg	University of Granada
Kim Henrick	European Bioinformat Institute (EBI)
Vicente Hernández	Politechnic University of Valencia (UPV)
Lynette Hirschman MITRE	MITRE Corporation
Martijn Huynen	Nijmegen Centre for Molecular Life Sciences
Iñaki Inza	University of Basque Country
Ana Maria Ionescu	University of Granada
Richard Jackson	University of Leeds
Andrew Jenkinson	European Bioinformatics Institute (EBI)
Bekir Karlik	Selcuk University
Guzin Kecek	Fatih University
Ekaterine Kldiashvili	Georgian Telemedicine Union
Juliane Klein	University of Tübingen
Martin Krallinger	Spanish Nacional Cancer Research Centre (CNIO)

Natalio Krasnogor	University of Nottingham
Abhay Krishan	Andalusian Molecular Biology and Regenerative Medicine Centre (CABIMER)
Marija Krstic-Demonacos	University of Salford
Sajeesh Kumar	UT Health Science Center
Vipin Kumar	University of Minnesota
Rainer König	German Cancer Research Center (DKFZ)
Pedro Larrañaga Computational Intelligence Group, UPM	Technical University of Madrid (UPM)
David Leader	University of Glasgow
Isaac Lera	University of the Balearic Islands
Miguel A. Lopez Gordo	University of Cadiz
Ernesto Lowy	Center for Genomic Regulation (CRG)
Jose Lozano	University of the Basque Country
Victor Maojo	Politechnic University of Madrid (UPM)
Marc Marti	Spanish National Center of Genomic Analysis
Diego Martín	University of Salamanca
Fernando Martín	Institute of Health Carlos III
Victoria Martín Requena	University of Malaga
Francisco Martínez-Álvarez	University "Pablo Olavide" of Seville
Victor Martínez Gómez	University of Granada
Alexandre Masselot	Swiss Institute of Bioinformatics (CMU)
Marco Masseroli	Polytechnical University of Milano
Jose Mato	Center for Cooperative Research in Biosciences (CIC bio-GUNE)
Rune Matthiesen	University of Porto
Miguel Medina	University of Malaga
Ivan Merelli	Italian National Research Council (CNR)
Jordi Mestres	University Pompeu Fabra
Iain Moal	Barcelona Supercomputing Center (BSC)
Antonio Morreale	Center of Molecular Biology "Severo Ochoa"
Walter N. Moss	Yale University and Howard Hughes Medical Institute
Enrique Muro	Max Delbrck Center for Molecular Medicine (MDC)
Carmen Navarro	University of Granada
Isabel Nepomuceno	University of Seville
Michael Ng	Hong Kong Baptist University
Baldomero Oliva	University Pompeu Fabra
José Luis Oliveira	University of Aveiro
Jose Luis Oliver	University of Granada
Salvatore Orlando	University of Venezia
Julio Ortega	University of Granada
Francisco Ortuño	University of Granada
Christos Ouzounis	Centre for Research and Technology Hellas
Erola Pairó	Institute for Bioengineering of Catalonia (IBEC)
Akhilesh Pandey	Johns Hopkins University
Lorena Pantano NA	Institute of Predictive and Personalized Medicine of Cancer (IMPPC)
Vesna Pajic	University of Belgrade

Paolo Paradisi	Italian National Research Council (CNR)
Alberto Pascual	University "Complutense" of Madrid (UCM)
David Pelta	University of Granada
Alexandre Perera	Technical University of Catalonia (UPC)
Mara del Mar Prez	University of Granada
Joanna Polanska	Silesian University of Technology
Javier Perez Florido	Genomics and Bioinformatics Platform of Andalusia (GBPA)
Horacio Perez Sanchez	Catholic University of Murcia (UCAM)
Graziano Pesole	University of Bari
David Pisano CNIO	Spanish National Cancer Center (CNIO)
Alberto Policriti	University of Udine
Héctor Pomares	University of Granada
Pablo Porras	European Bioinformatics Institute (EBI)
Robert Preissner	University of Charit
Alberto Prieto	University of Granada
Carlos Puntonet	University of Granada
Omer Rana	Cardiff University
Jairo Rocha	University of the Balearic Islands
Ignacio Rojas	University of Granada
Julio Rozas	University of Barcelone
Gregorio Rubio	Politechnic University of Valencia (UPV)
Gonzalo Ruiz	University of Granada
Maria Jose Saez Lara	University of Granada
Yvan Saeys	Flanders Institute for Biotechnology (VIB)
Rodrigo Santamaria	University of Salamanca
Pilar Santisteban	Institute of Biomedical Researchs "Alberto Sols"
Javier Santoyo	Andalusian Institute of Human Sequencing
Michael Sattler	Institute of Structural Biology (Helmholtz Zentrum Mnchen)
Peter Schmidtke	University of Barcelona
Vicky Schneider	European Bioinformat Institute (EBI)
Jose Seoane	University of Bristol
Luis Serrano	Center for Genomic Regulation (CRG)
James Sharpe	Centre for Genomics Regulation (CRG)
Vladimir Shulaev	University of North Texas
Jean-Marc Schwartz	University of Manchester
Istvan Simon	Hungarian Research Centre for Natural Sciences
Richard Sinnott	University of Glasgow
Prashanth Suravajhala	Bioclues.org and Bioinformatics.org
Yoshiyuki Suzuki	Tokyo Metropolitan Institute of Medical Science
Li Teng	University of Iowa
Yaping Tian	Chinese PLA General Hospital
Carolina Torres	University of Granada
Huseyin Tombuloglu	Fatih University
Oswaldo Trelles	University of Malaga
Paolo Trunfio	University of Calabria
Shusaku Tsumoto	Shimane University
Renato Umeton	Massachusetts Institute of Technology
Jose Miguel Urquiza	Chromatin and Disease Group (IDIBELL)

Olga Valenzuela
Alfredo Vellido
Nicola Vitacolonna
Arthur R. Williams
Phoebe D. Williams
Wolfgang Wurst
Albert Zomaya

University of Granada
Technical University of Catalonia (UPC)
University of Udine
CINDRR and University of South Florida
Kansas University Medical Center
National Research Center of Environment & Health
University of Sydney

Table of Contents

SS10: Effective Soft Computing Methods for Biomedical Signals	
Comparison Machine Learning Algorithms for Recognition of Epileptic Seizures in EEG ..	1
<i>Bekir Karlik and Sengul Bayrak Hayta</i>	
A new computational measure for detection of extrapyramidal symptoms	13
<i>José Ignacio Aznarte, Sergio Iglesias-Parro, Antonio Ibañez-Molina and María Felipa Soriano</i>	
The Optimization of Breathing Signals and Ventilatory Control with Nonlinear Respiratory Mechanics under Hypercapnia and Eucapnia	23
<i>Shyan-Lung Lin, Yu-Zhe Tsai and Andy Liao</i>	
A Novel Feature Extraction Method for Heart Sounds Classification	34
<i>Yücel Koçyiğit</i>	
Task Related and Spatially Regularized Common Spatial Patterns for Brain Computer Interfaces	42
<i>Ayhan Yuksel and Tamer Olmez</i>	
Fuzzy Clustering of ECG Beats Using a New Metaheuristic Approach	54
<i>Berat Dogan and Tamer Ölmez</i>	
SS9: Biomaterials in Biomedicine: Computational approaches	
Predictive algorithms for determination of reflectance data from quantity of pigments within experimental dental resin composites	66
<i>Razvan Ghinea, Oscar E. Pecho, Luis Javier Herrera, Ana Maria Ionescu, Juan De La Cruz Cardona, Maria Purificacion Sanchez Sanchez, Rade D. Paravina and Maria Del Mar Perez Gomez</i>	
Ultrasonic monitoring of artificial tissue mechanical properties in biorreactor	77
<i>Guillermo Rus, Nicolas Bochud, Juan Melchor, Laura Peralta, Juan Chiachio, Manuel Chiachio and Antonio Gómez</i>	
Influence of the Length in Biomimetic Ion Channels Based on Derivatized γ -Self Assembled Peptide Nanotubes. A Molecular Dynamics study.	84
<i>Rebeca Garcia-Fandino, Juan Outeiral, Saulo Vazquez and Juan R. Granja</i>	
FDTD simulations for ultrasound propagation in a 2-D cervical tissue model	85
<i>Laura Peralta, Guillermo Rus, Nicolas Bochud, Juan Melchor, Juan Chiachio and Manuel Chiachio</i>	
Model-based probability of detection of pathologies in soft tissue	97
<i>Juan Melchor, Guillermo Rus, Nicolas Bochud, Laura Peralta, Juan Chiachio, Manuel Chiachio and Antonio Gómez</i>	

Information-theory approach to model class assessment for tissue-engineered cultures consistence evolution	108
<i>Juan Chiachío-Ruano, Manuel Chiachío Ruano, Guillermo Rus-Carlborg, Nicolas Bochud, Laura Peralta Pereira and Juan Manuel Melchor Rodriguez</i>	

SS1: Multi-biomarker and informatics in cancer diagnosis

The mathematical models of serum HE4 and CA125 combined application to improve the pelvic tumor differential diagnosis rate	109
<i>Yaping Tian, Chuanxin Wang, Liming Cheng, Peng Zhang, Lin Guo, Wanli Liu, Zhongying Zhang, Yanchun Huang, Qishui Ou, Xinyu Wen, Yuelei Xing, Guixi Zheng, Ziyong Sun, Huijun Li, Aimin Zhang, Ying Chen, Wen Liu, Huiming Ye, Yi Xu, Yishan Huo and Jing Chen</i>	
Exploration of ovarian cancer micro array data focus on gene expression patterns relevant to survival using artificial neural networks.....	116
<i>Clare Coveney, Dong Tong, David Boocock, Robert Rees and Graham Ball</i>	
High efficiency for activated KRAS detection from peripheral blood using weighted enzymatic gene chip array method	124
<i>Tai Feng Hsu, Ming Yii Huang, Hsueh Chiao Liu, Jia Yuan Chang, Jian Jhang Huang and Shiu Ru Lin</i>	
HER-2/neu Breast Cancer Diagnosis Procedure, Based on Histopathology Image Analysis	135
<i>Martin Tabakov, Marta Tabakov, Halina Kwasnicka, Pawel Kozak and Bartosz Pula</i>	

SS6: ePathology - Realities and Perspectives

Image Quality Assessment in Digital Pathology The Analysis of Background in Whole Slide Images	147
<i>Thomas Schrader, Anne Nadolny, Robert Piduch and Peter Hufnagl</i>	
Automatic image quality assessment in digital pathology: from idea to implementation ...	148
<i>David Ameisen, Christophe Deroulers, Valerie Perrier, Fatiha Bouhidel, Maxime Battistella, Luc Legres, Anne Janin, Philippe Bertheau and Jean-Baptiste Yunès</i>	
Web-based remote diagnosis system using virtual slide for routine pathology slides, analysis of discrepancies between virtual and real microscopic diagnosis.....	158
<i>Ichiro Mori, Takashi Ozaki, Yasuteru Muragaki and Yoshiyuki Osamura</i>	
Evaluation of cytokeratin-19 in breast cancer tissue samples: a comparison of automatic and manual evaluations of scanned tissue microarray cylinders.....	160
<i>Cristina Callau, Marylène Lejeune, Anna Korzynska, Marcial García, Gloria Bueno, Ramon Bosch, Joaquín Jaén, Guifré Orero, Teresa Salvadó and Carlos López</i>	
The medical information system and its application for quality assurance programs in cytology Georgian experience.....	173
<i>Ekaterina Kldiashvili, Archil Burduli, Gocha Ghortlishvili, Dodo Agladze and Ivan Sheklashvili</i>	

<hr/>	
SS7: Modelling of cellular pathways and disease	
<hr/>	
In silico prediction of elementary mode fluxes.....	183
<i>Jean-Marc Schwartz and Peter Neal Taylor</i>	
Applications of p53 interactome analysis to personalised drug discovery.....	192
<i>Michelle Hussain, Benjamin Stutchbury, Kun Tian, Rengul Atalay, Jean-Marc Schwartz and Marija Krstic-Demonacos</i>	
Evolutionary and functional studies on the novel Hepatitis C virus core+1/ARF protein..	204
<i>Ioly Kotta-Loizou</i>	
<hr/>	
Biomedical Data Mining	
<hr/>	
Using Biomedical Terminologies to extract Noun Phrases for managing knowledge evolution.....	205
<i>Adila Merabti, Lina F. Soualmia and Stéfan J Darmoni</i>	
Computing Pathways in Bio-Models Derived from Bio-Science Text Sources	217
<i>Troels Andreassen, Henrik Bulskov, Jørgen Fischer Nilsson and Per Anker Jensen</i>	
Comparing BioPortal and HeTOP: towards a unique biomedical ontology portal ?	227
<i>Julien Grosjean, Lina F. Soualmia, Khedidja Bouarech, Clement Jonquet and Stefan J. Darmoni</i>	
<hr/>	
SS2: Discovery of non-coding and structured RNAs	
<hr/>	
Analyses of non-coding RNAs generated from the EpsteinBarr virus W repeat region.	238
<i>Walter N. Moss</i>	
New frontiers in the investigation of structural functional RNA domains in viral genomes. Understanding the hepatitis C virus (HCV)	253
<i>Alfredo Berzal-Herranz and Cristina Romero-López</i>	
Identifying functional SNVs that map to non-coding regions of the genome and alter RNA Structure.....	272
<i>Alain Laederach</i>	
Detection of structural constraints and conformational transitions in the influenza virus RNA genome using structure predictions and mutual information calculations	273
<i>Alexander Gultyaev, Rene Olsthoorn, Monique Spronken and Ron Fouchier</i>	
In silico discovery of de novo structured RNAs in genomic and transcriptomic sequence ...	285
<i>Jan Gorodkin</i>	
Comparative Detection of Processed Small RNAs in Archaea	286
<i>Christian Hoener Zu Siederdisen, Sarah Berkemer, Fabian Amman, Axel Wintsche, Sebastian Will, Sonja J. Prohaska and Peter Stadler</i>	

<hr/>	
SS14: Better Oncology Treatment and Patient Outcomes by Using Therapy-Related Symptom Checklists (TRSC/TRSC-C) and a Computerized Two-Way Communication System	
<hr/>	
The Development and Application of an Oncology Therapy-Related Symptom Checklist for Adults (TRSC) and Children (TRSC-C)	298
<i>Arthur R. Williams, David D. Williams and Phoebe D. Williams</i>	
The TRSC-C and Childhood Leukemia in Thailand and the USA: Symptom Occurrence/Severity and Care Strategies for Symptom Relief	308
<i>Phoebe D. Williams, Ubolrat Piamjariyakul and Jenna Degennaro</i>	
The TRSC and Symptom Monitoring, Alleviation, and Self-Care among Mexican-Americans during Outpatient Cancer Treatments	315
<i>Phoebe D. Williams, Leticia S. Lantican, Julia O. Bader and Daniela Lerma</i>	
Computers that Show Recognition of Patients' Symptoms	321
<i>Farrokh A. Alemi, Hosai Hesham, Arthur R. Williams, Phoebe D. Williams, Blaine Donley and Raya E. Kheirbek</i>	
<hr/>	
High Performance for Sequence Analysis	
<hr/>	
Exploring Sequence Alignment Algorithms on FPGA-based Heterogeneous Architectures .	330
<i>Xin Chang, Fernandao A. Escobar, Carlos Valderrama and Vincent Robert</i>	
AutoFlow: an easy way to build workflows	342
<i>Pedro Seoane, Rosario Carmona, Rocío Bautista, Darío Guerrero and M. Gonzalo Claros</i>	
A Probabilistic Genome-Wide Gene Reading Frame Sequence Model	350
<i>Christian Theil Have and Søren Mørk</i>	
Inexact Sequence Mapping Study Cases: Hybrid GPU Computing and Memory Demanding Indexes	362
<i>José Salavert, Andrés Tomás, Ignacio Medina and Ignacio Blanquer</i>	
<hr/>	
SS11: Chaperone Therapy for Protein Misfolding Disorders with Brain Dysfunction	
<hr/>	
Concept and Development of Chaperone Therapy for Protein Misfolding Diseases	374
<i>Yoshiyuki Suzuki</i>	
Design and Synthesis of Bioactive Valienamine-type Chaperones	385
<i>Seiichiro Ogawa, Shinichi Kuno, Katsumi Higaki, Atsushi Takahashi and Eiji Nanba</i>	
Pharmacological Chaperones by Design	389
<i>José M. García Fernández</i>	
Enzyme Enhancement Therapy through non-competitive pharmacological chaperones	390
<i>Juan Aymami, Xavier Barril, Aida Delgado, Marc Revés, Rodolfo Lavilla, Katsumi Higaki, Ana María García-Collazo, Laura Rodríguez-Pascau, Elena Cubero, Pilar Pizcueta and Marc Martinell</i>	
Structural basis of pharmacological chaperoning for human β -galactosidase	396
<i>Toshiyuki Shimizu</i>	

Identification and characterization of chaperone compounds for human beta-galactosidase deficiency	397
<i>Eiji Nanba and Katsumi Higaki</i>	

SS8: Integration of data, methods and tools in biosciences

Integration of data in biosciences	398
<i>Vesna Pajić, Gordana Pavlović Lažetić, Dragana Dudić, Dragica Radovanović and Jelena Kozoderović</i>	
A Machine Learning Approach to Enhance Scoring Performance in Docking-Based Virtual Screening Experiments: COX-1 as a Case Study	406
<i>Cândida Silva, Pedro Carreiras, Elsa Henriques, Carlos Simões and Rui Brito</i>	
Omic Data Modelling for Information Retrieval	415
<i>Chloé Cabot, Julien Grosjean, Romain Lelong, Arnaud Lefebvre, Thierry Lecroq, Lina F. Soualmia and Stéfan J. Darmoni</i>	
Numerical Simulation of ISFET Structures for BioSensing Devices with TCAD Tools	425
<i>Daniele Passeri, Arianna Morozzi, Keida Kancheri and Andrea Scorzoni</i>	
Mining Associations for Organism Characteristics in Prokaryotes - an Integrative Approach	436
<i>Gordana Pavlović Lažetić, Vesna Pajić, Nenad Mitić, Jovana Kovačević and Miloš Beljanski</i>	
LPS: a strategy for the generation of longer DNA sequence fragments from short reads ...	451
<i>Francisco Vera Voronisky, Ansel Y. Rodriguez Gonzalez, Ivan Olmos Pineda, Patricia Sanchez Alonso, Candelario Vazquez Cruz and Jesus A. Gonzalez</i>	
In Search of Predictive Models for Inhibitors of 5-alpha Reductase 2 Based on the Integration of Bioactivity and Molecular Descriptors Data	463
<i>Joana Sousa, Rui Brito, Jorge Salvador and Cândida Silva</i>	
Assisted prescription for improving treatments in Obstetrics-Gynecology Department	473
<i>Mihaela Marcella Crian-Vida, Oana Lupe and Lacramioara Stoicu-Tivadar</i>	
Integrating Expression Data from Different Microarray Platforms in Search of Biomarkers of Radiosensitivity	484
<i>Anna Papiiez, Paul Finnon, Christophe Badie, Simon Bouffler and Joanna Polanska</i>	

SS4: High Performance Computing in Bioinformatics

The role of High Performance Computing in Bioinformatics	494
<i>Horacio Pérez-Sánchez, José M. Cecilia and Ivan Merelli</i>	
Hamming Distance based Binary PSO for Feature Selection and Classification from high dimensional Gene Expression Data	507
<i>Haider Banka and Suresh Dara</i>	
Experience with Lamport Clock Ordered Events with Intel Threading Building Blocks in a Glucose-Level Prediction Software	515
<i>Tomas Koutny</i>	

Hermite Polynomial Characterization of Heartbeats with Graphics Processing Units	527
<i>Alberto Gil, Gabriel Caffarena, David G. Marquez and Abraham Otero</i>	
Entropy-based High Performance Computation of Boolean SNP-SNP Interactions Using GPUs	539
<i>Carlos Riveros, Manuel Ujaldon and Pablo Moscato</i>	
Evaluating mixed HTC/cloud approaches for parameter sweep applications in systems biology	551
<i>Ivan Merelli, Ettore Mosca, Daniele Cesini, Elisabetta Ronchieri and Luciano Milanesi</i>	
Design of a Generic Architecture for executing Bioinformatics Workflows on Distributed Infrastructures	563
<i>Abel Carrión, Ignacio Blanquer, Miguel Caballer, Cristina Yenyxe González and Ignacio Medina</i>	
An Efficient Solvent Accessible Surface Area calculation applied in Ab Initio Protein Structure Prediction	575
<i>Daniel Bonetti, Horacio Pérez-Sánchez and Alexandre Delbem</i>	
Parallel Computation of Non-Bonded Interactions in Drug Discovery: Nvidia GPUs vs. Intel Xeon Phi	579
<i>Jianbin Fang, Ana Lucia Varbanescu, Baldomero Imbernon, Jose Cecilia and Horacio Perez-Sanchez</i>	
Accelerating Phylogenetic Inference on GPUs: an OpenACC and CUDA comparison	589
<i>Lidia Kuan, Joao Neves, Frederico Pratas, Pedro Tomas and Leonel Sousa</i>	
High Performance Computing Studies of RNA Nanotubes	601
<i>Shyam Badu, Roderik Melnik, Maxim V. Paliy, Sanjay Prabhakar and Bruce A. Shapiro</i>	
<hr/> SS3: Biological Knowledge Visualization <hr/>	
Generalized macro level models of amino acid sequences using passive electrical circuits . .	608
<i>Roger Marshall</i>	
Exploratory Visualization of Misclassified GPCRs from their transformed unaligned sequences using manifold learning techniques	623
<i>Martha Ivón Cárdenas, Alfredo Vellido, Caroline König, René Alquézar and Jesús Giraldo</i>	
Mapping regional changes in the glycerophosphocholine second messenger lipidome following brain injury using CIRCOS	631
<i>Graeme Taylor, Hongbin Xu, Andrew Syrett and Steffany Bennett</i>	
Multi-dimensional anatomical representation: A volumetric comparison of the C57BL/6 and N3 C57BL/6 x 129/Sv mouse brain modelled from serial section using Autodesk Maya	642
<i>Stephen Fai, Katie Wurts, Andrew Syrett, Brendan Trickey, Nico Valenzuela and Steffany Bennett</i>	

<hr/>	
Bioinformatics Tools and Databases	
<hr/>	
Bio4j: bigger, faster, leaner.....	660
<i>Pablo Pareja-Tobes, Alexey Alekhin, Evdokim Kovach, Marina Manrique, Eduardo Pareja, Raquel Tobes and Eduardo Pareja-Tobes</i>	
Massive Automatic Functional Annotation - MAFA	661
<i>Cristian Alejandro Rojas Quintero, Nelson Enrique Vera Parra and José Nelson Pérez Castillo</i>	
Advancing Lipidomic Bioinformatic Technologies: Visualization and Phospholipid Identification (VaLID) version 3.0	670
<i>Graeme S.V Mcdowell, Alexandre P. Blanchard, Daniel Figeys, Stephen Fai and Steffany A.L. Bennett</i>	
MG7: A fast horizontally scalable tool based on cloud computing and graph databases for microbial community profiling	681
<i>Evdokim Kovach, Alexey Alekhin, Marina Manrique, Pablo Pareja-Tobes, Eduardo Pareja, Raquel Tobes and Eduardo Pareja-Tobes</i>	
<hr/>	
SS13: Computational analysis of gene regulatory elements with next-gen sequencing data	
<hr/>	
Principles of ChIP-seq Data Analysis Illustrated with Examples	682
<i>Giovanna Ambrosini, René Dreos and Philipp Bucher</i>	
Transcription factor binding and nucleosome positioning are alternative pathways for transcription start site selection in eukaryotic promoters.....	695
<i>René Dreos, Giovanna Ambrosini and Philipp Bucher</i>	
Changes in heat shock duration influence regulatory schemes of HSF1 activity.	707
<i>Marta Iwanaszko, Patryk Janus, Tomasz Stokowy, Piotr Widlak and Marek Kimmel</i>	
Transcription Factor Binding Site Detection Algorithm Using Distance Metrics Based on a Position Frequency Matrix Concept	715
<i>Mohammad Al Bataineh, Lun Huang and Guillermo Atkin</i>	
<hr/>	
Gene Expression and Microarrays	
<hr/>	
Implications of RBBP6 in various types of cancer.....	727
<i>Firdous Khan, Mushal Allam, Marius Tincho and Ashley Pretorius</i>	
Automatic detection of outlying microarrays using multi-array quality metrics	738
<i>Michal Marczyk, Lukasz Krol and Joanna Polanska</i>	
Formal Concept Analysis and Knowledge Integration for Highlighting Statistically Enriched Functions from Microarrays Data.	747
<i>Sidahmed Benabderrahmane</i>	
<hr/>	
Biomedical Engineering and eHealth Applications	
<hr/>	
Evaluating the effects of signal segmentation on activity recognition	759
<i>Oresti Baños, Juan Manuel Galvez, Miguel Damas, Alberto Guillen, Luis Javier Herrera, Hector Pomares and Ignacio Rojas</i>	

A Supervised Cooperative Learning System for Early Detection of Language Disorders ...	766
<i>Maria Luisa Martín Ruiz, Miguel Ángel Valero Duboy, Iván Pau de La Cruz, María Peñafiel Puerto and Carmen Torcal Loriente</i>	
Improvement in the accuracy of Nuclear Magnetic Resonance spectrum analysis by automatic tuning of phase correction algorithms.	778
<i>Franciszek Binczyk, Rafal Tarnawski and Joanna Polanska</i>	
Analysis of Respiratory Flow Signals to Identify Success of Patients on Weaning Trials ...	789
<i>Hernando Gonzalez Acevedo, Carlos Arizmendi and Beatriz Giraldo</i>	
A novel framework to enhance scientific knowledge of cardiovascular MRI biomarkers and their application to pediatric cardiomyopathy classification.	798
<i>Vanathi Gopalakrishnan, Prahlad Menon and Shobhit Madan</i>	
<hr/> SS15: Computational MRI: Theory, Dynamics and Applications <hr/>	
Resolving the enhanced flow parameters for an in-depth analysis of the MRI-Neuroimaging	810
<i>Moses Emetere, Bamidele Awojoyogbe, Uno Uno, Kasim Isah and Michael M. Dada</i>	
Computational Magnetic Resonance Imaging based on Bloch NMR Flow Equation and Bessel Functions	820
<i>Bamidele Omotayo Awojoyogbe, Michael Oluwaseun Dada and Adewale Omoniyi Adesola</i>	
Computational Phase Contrast Magnetic Resonance Imaging based on Legendre Polynomials.....	842
<i>Michael Oluwaseun Dada and Omotayo Bamidele Awojoyogbe</i>	
Computational model of NMR Molecular Dynamics for the Analysis Blood Brain Barrier .	860
<i>Michael Oluwaseun Dada, Omotayo Bamidele Awojoyogbe and Simona Baroni</i>	
Multiple Sclerosis lesion segmentation using Active Contours model and adaptive outlier detection method	878
<i>Foued Derraz, Antonio Pinti, Laurent Peyrodie, Miloud Boussahla, Hechmi Toumi and Patrick Hauteceur</i>	
<hr/> Computational Proteomics and Biological Systems <hr/>	
Protein function easily investigated by genomics data mining using the ProteINSIDE web service	890
<i>Nicolas Kaspric, Brigitte Picard, Matthieu Reichstadt, Jérémy Tournayre and Muriel Bonnet</i>	
Protein Fold Classification using Kohonen's Self-Organizing Map.....	903
<i>Ozlem Ozbudak and Zümray Dokur</i>	
A framework for modelling spatially dependent interactions of biological systems in CCP	912
<i>Davide Chiarugi, Moreno Falaschi, Diana Hermith and Carlos Olarte</i>	

Rewriting Logic and Symbolic Systems Biology applied to EGF Signaling Pathway	924
<i>Gustavo Santos García, Javier De Las Rivas and Carolyn Talcott</i>	

Ethical Principles in Biotechnology and Bioengineering

Biotechnology, Biomedicine and the Precautionary Principle.....	936
<i>Robin Attfield</i>	

High Performance Bioinformatics for Healthcare and Diseases

Blood Vessel Segmentation in Retinal Images based on Local Binary Patterns and Evolutionary Neural Networks.....	941
<i>Antonio Rodriguez-Jimenez and Enrique J. Carmona</i>	

Outlier detection for single particle analysis in Electron Microscopy.....	950
<i>Carlos O.S. Sorzano, Javier Vargas, Jose Miguel de La Rosa-Trevin, Airen Zaldivar-Peraza, Joaquin Oton, Vahid Abrishami, Ignacio Foche, Roberto Marabini, Gabriel Caffarena and Jose M. Caraza</i>	

2D and 3D Alignment for Electron Microscopy via Graphics Processing Units	960
<i>Eduardo García de La Cueva, Miguel Mateo de La Puente, Alessandro Deideri, Ana Iriarte, Carlos O.S. Sorzano and Gabriel Caffarena</i>	

Application of parallel blind docking with BINDSURF for the study of platinum derived compounds as anticancer drugs.....	972
<i>José P. Cerón Carrasco, Javier Cerezo, José Zuñiga, Alberto Requena, Julia Contreras-García, Sonali Chavan, Miguel Manrubia-Cobo, Baldomero Imbernón, José M. Cecilia and Horacio Pérez-Sánchez</i>	

miRNA Regulation Networks

Model-based design of synthetic networks	977
<i>Liliana Ironi and Diana X Tran</i>	

miRNAO: An Ontology Unfolding the Domain of microRNAs	989
<i>Vicky Dritsou, Pantelis Topalis, Elvira Mitraka, Emmanuel Dialynas and Christos Louis</i>	

Inference of Circadian Regulatory Networks	1001
<i>Marco Grzegorzczuk, Andrej Aderhold, V. Anne Smith and Dirk Husmeier</i>	

miR-SEA: miRNA Seed Extension based Aligner Pipeline for NGS Expression Level Extraction	1015
<i>Gianvito Urgese, Giulia Paciello, Claudio Isella, Enzo Medico, Enrico Macii, Elisa Ficarra and Andrea Acquaviva</i>	

Computational Approaches for Genomics and NGS

An Integrated Approach to Comparative Assembly	1027
<i>John Healy and Desmond Chambers</i>	

Hkera, a human transcriptome partitioner	1039
<i>Austin Chiang, Grace Shaw and Ming-Jing Hwang</i>	

Robust Error Correction for De Novo Assembly via Spectral Partitioning and Sequence Alignment	1040
<i>Andrei Alic, Andres Tomas, José Salavert Torres, Ignacio Medina and Ignacio Blanquer</i>	
Using a Random Forest proximity measure for variable importance stratification in genotypic data	1049
<i>Jose A. Seoane, Ian N.M. Day, Colin Campbell, Juan Pablo Casas and Tom R. Gaunt</i>	
<hr/> SS16: Bioinformatical Approaches to Disordered Proteins <hr/>	
Predicting functional sites in disordered proteins - implications in disease	1061
<i>Bálint Mészáros, István Simon and Zsuzsanna Dosztányi</i>	
Dynamic approaches to structural ensembles of intrinsically disordered proteins	1062
<i>Peter Tompa</i>	
The Roles of Short Linear Motifs in Human Diseases	1063
<i>Bora Uyar and Toby Gibson</i>	
<hr/> SS18: Stochastic Modelling of Biological Systems <hr/>	
Flexible docking of the fragment of the troponin I to the troponin C	1064
<i>Jacek Wabik, Mateusz Kurcinski and Andrzej Kolinski</i>	
Analysis of risk factors of hip fracture with causal Bayesian networks	1074
<i>Alex Aussem, Pascal Caillet, Zara Klemm, Maxime Gasse, Anne-Marie Schott and Michel Ducher</i>	
Incorporating covariates in a flowgraph model for bladder carcinoma	1086
<i>Gregorio Rubio, Belen Garcia-Mora, Cristina Santamaria and Francisco Santonja</i>	
<hr/> Poster Session <hr/>	
Shape of a dilution curve as the consequence of stochasticity within microcirculation	1097
<i>Victor Kislukhin</i>	
Inhibition of Wb-iPGM using analogues of Clorsulon and co-administration with DEC for bancroftian filariasis treatment	1104
<i>Om Prakash Sharma and M. Suresh Kumar</i>	
Identification of Potent Inhibitors for Resistant Form of Chronic Myelogenous Leukaemia (CML)	1122
<i>Hemanth Naick B, Om Prakash Sharma, Muthuvel Suresh Kumar and Baskaran Rajasekaran</i>	
Comparison of pregnancy predictive models applied to women who received IVF/ICSI in Valencia (Spain) using ROC curves	1145
<i>Ana Debón, Patricia Carracedo and Inmaculada Molina</i>	
Applying Stacked and Cascade Generalizations to B-cell Epitope Prediction	1154
<i>Yuh-Jyh Hu, Shun-Chien Lin and Yu-Lung Lin</i>	
An Interactive X-Ray Image Segmentation Technique for Bone Extraction	1164
<i>Cristina Stolojescu-Crisan and Stefan Holban</i>	

Sequencing by Ligation with Double-Labeled Fluorescent Probes	1172
<i>Dan Pu, Jing Chen and Pengfeng Xiao</i>	
Fast Parallel Bayesian Networks Reconstruction with BNFinder	1179
<i>Alina Frolova and Bartek Wilczynski</i>	
Patient Trajectory Modeling and Analysis.....	1185
<i>Jalel Akaichi and Marwa Manaa</i>	
N-body Information Theory (NbIT) Analysis of Rigid-Body Dynamics in Intracellular Loop 2 of the 5-HT _{2A} Receptor	1190
<i>Michael V. Levine, Jose Manuel Perez-Aguilar and Harel Weinstein</i>	
Doctors and researchers: integrating data for bipolar disorder studies.....	1202
<i>Patricia Rodriguez Tomé, Carla Melis, Raffaella Ardau, Caterina Chillotti, Maria Del Zompo, Alessio Squassina and Giovanni Severino</i>	
An Effective 3-Dimensional Regional Myocardial Strain Computation Method with Displacement ENcoding with Stimulated Echoes (DENSE) in Dilated Cardiomyopathy Patients and Healthy Subjects.....	1203
<i>Julia Kar, Andrew Knutsen, Kevin Kulshrestha, Brian Cupps and Michael Pasque</i>	
Fully Automatic Renal Parenchyma Volumetry in LDA-based Probability Maps Using Variational Outer Cortex Edge Alignment Forces.....	1207
<i>Oliver Gloger, Volkmar Liebscher, Klaus Tönnies and Henry Völzke</i>	
Positioning Method Based on Infrared Spectrum Detection of Neurotransmitters for Electrical Nerve Stimulation after Spinal Cord Injury.....	1219
<i>Min Nie and Guang Yang</i>	
Specificities of Medical Devices Affecting Health Technology Assessment Methodology	1229
<i>Vladimir Rogalewicz and Ivana Jurickova</i>	
Supervised Retinal Vessel Segmentation Based on Neural Network Using Broader Aging Dataset	1235
<i>Mohd Zulfaezal Che Azemin, Fadilah Abd Hamid, Mohd Izzuddin Mohd Tamrin, and Anang Hdaya Muhamad Amin</i>	
Label-free detection of viruses using liquid crystals on a polymeric surface with periodic nanostructures	1243
<i>Chang-Hyun Jang and Dong Min Kim</i>	
Voice controller for Image Guided Surgery and per-sonalized Interactive Visualisation.....	1244
<i>Andoni Beristain, Alessandro De Mauro, Koen Van De Weyer and Dominique Segers</i>	
Improving Stability of Feature Selection for Brain Tumour Diagnosis Using 1H-MRS Data	1254
<i>Albert Vilamala and Lluís A. Belanche</i>	
Stress-strain analyses of the jaws with multiple keratocysts before and after surgery	1266
<i>Josef Danek, Tatjana Dostalova, Milan Hubacek and Nima Mahdian</i>	
Intelligent System for Premature Babies Healthcare at Home based on Case-based Reasoning.....	1278
<i>Beatriz López, Fco. I. Gamero, Jordi Coll, Natàlia Mordvaniuk, Judit Bassols, Eva Bargallo, Borja Guarch and Abel López-Bermejo</i>	

Quantitative Analysis of Pathological Mitochondrial Morphology in Neuronal Cells in Confocal Laser Scanning Microscopy Images.....	1290
<i>Herbert Süße, Wolfgang Ortmann, Janin Lautenschläger, Christian Lautenschläger, Marco Körner, Julian Grosskreutz and Joachim Denzler</i>	
PETRA: Multivariate analyses for neuroimaging data.....	1302
<i>Fermín Segovia Román, Ignacio Álvarez Illán, Diego Salas González, Francisco Jesús Martínez Murcia, Christophe Phillips, Carlos García Puntonet, Javier Ramírez Pérez de Inestrosa and Juan Manuel Górriz Sáez</i>	
Liposomes in polymer matrix. Stability of liposomes in PEG 400 and PEG 8000 solutions.....	1313
<i>Magdalena Bajgrowicz, Jerzy Detyna and Marek Langner</i>	
Motif discovery in speech: application to monitoring Alzheimers disease.....	1323
<i>Dragana Nikolic, Anna Barney, Vanda Nemes and Peter Garrard</i>	
Designing preterm neonatal cyanosis simulation.....	1325
<i>Peter Peters, Frank Delbressine and Loe Feijs</i>	
Computational approach for modeling and testing NF- κ B binding sites.....	1338
<i>Marcin Pacholczyk, Karolina Smolińska, Marta Iwanaszko and Marek Kimmel</i>	
Dynamic Gap Selector: A Smith Waterman Sequence Alignment Algorithm with Affine Gap Model Optimization.....	1347
<i>Gianvito Urgese, Giulia Paciello, Andrea Acquaviva, Elisa Ficarra, Mariagrazia Graziano and Maurizio Zamboni</i>	
Multiple-criteria decision making: application to medical devices.....	1359
<i>Vladimir Rogalewicz and Ivana Jurickova</i>	
Barriers to implementation of a clinical information system in an emergency department .	1373
<i>Ivana Jurickova and Pavla Hejmová</i>	
Case study: Mobile X-ray equipment selection for a traumatology department using value engineering and multi-criteria decision methods.....	1389
<i>Ivana Jurickova and Aleš Kraina</i>	
Automatically building database from biological ontology.....	1403
<i>Zina Nakhla and Kaouther Nouira</i>	
Statika: managing cloud resources, bioinformatics tools and data.....	1412
<i>Alexey Alekhin, Evdokim Kovach, Pablo Pareja-Tobes, Marina Manrique, Eduardo Pareja, Raquel Tobes and Eduardo Pareja-Tobes</i>	
Nispero: a cloud-computing based Scala tool specially suited for bioinformatics data processing.....	1414
<i>Evdokim Kovach, Alexey Alekhin, Marina Manrique, Pablo Pareja-Tobes, Eduardo Pareja, Raquel Tobes and Eduardo Pareja-Tobes</i>	
Determination of changes in plasma structure during extracorporeal circulation studies by ATR-FTIR spectroscopy and machine learning methods.....	1416
<i>Sylwia Olsztyńska-Janus, Barbara Kmiecik, Bartosz Krawczyk and Malgorzata Komorowska</i>	

Acoustic Study of a Neonatal Intensive Care Unit: Preliminary Results.....	1418
<i>Ganna Raboshchuk, Climent Nadeu, Blanca Muñoz Mahamud, Ana Riverola de Veciana and Santiago Navarro Hervas</i>	
Optimal preictal period in seizure prediction	1427
<i>Mojtaba Bandarabadi, Jalil Rasekhi, César Teixeira and Antonio Dourado</i>	
Revealing Helitron signatures in <i>Cænorhabditis elegans</i> by the Complex Morlet Analysis based on the Frequency Chaos Game Signals	1434
<i>Imen Messaoudi, Afef Elloumi Oueslati and Zied Lachiri</i>	
The impact of the quality filter for RNA-Seq data over differential expression profile.....	1445
<i>Pablo C. Gomes de Sa, Siomar de Castro Soares, Adonney A. de Oliveira Veras, Anne C. Pinto, Luis Guimaraes, Vasco Azevedo, Artur Silva and Rommel Ramos</i>	
An Automatic Wavelet Selection Scheme for Heart Sounds Denoising.....	1450
<i>Omari Tahar and Bereksi-Reguig Fethi</i>	
Digital Human Model and Motion Capture Techniques for Home Kinesitherapy.....	1463
<i>Karolina Grzechnik and Tadeusz Burczynski</i>	
Analysis of ATM signaling pathway as an activator of p53 and NF- κ B regulatory modules and the role of PPM1D	1471
<i>Krzysztof Puszynski, Katarzyna Jonak, Monika Kurpas, Patryk Janus and Katarzyna Szoltysek</i>	
High-throughput, Scalable, Quantitative, Cellular Phenotyping using X-Ray Tomographic Microscopy	1483
<i>Kevin Mader, Leah-Rae Donahue, Ralph Müller and Marco Stampanoni</i>	
Pervasive System for Searching the Appropriate Road: A Mobile Physician on Road Network Case Study	1499
<i>Hanen Faiez and Jalel Akaichi</i>	
The experimental model of lysozyme sustained release based on poly(3-hydroxybutyrate)-poly(ethylene glycol)/hydroxyapatite microparticles	1511
<i>A.L. Zernov, E.V. Ivanova and K.V. Shaitan</i>	
The experimental model of mesenchymal stem cells growth and differentiation based on application of porous scaffold from bacterial origin poly(3-hydroxybutyrate)-poly(ethylene glycol).....	1520
<i>I.I. Zharkova, E.A. Akulina, N.V. Andreeva and K.V Shaitan</i>	
Trajectory data warehouse modeling based on a Trajectory UML profile: Medical example	1527
<i>Wided Oueslati and Jalel Akaichi</i>	
A semantic cache for queries optimization of Health care services communities.....	1539
<i>Hela Limam and Jalel Akaichi</i>	
Infectious disease risk from anti-psychotic drug use: A population-based study.....	1554
<i>Usman Iqbal and Yu-Chuan Li</i>	

In-silico approaches to sequence and structure based scrutiny of nonsynonymous SNPs and synteny of ACAA2 for its implicated role in metabolomics.....	1555
<i>Sidrah Anjum</i>	
An Integrative Analysis of ncRNA-mRNA Using Co-expression Network to Discover Potential Contributions of Coding-non-coding RNA Clusters	1568
<i>Li Guo, Yang Zhao, Sheng Yang, Hui Zhang and Feng Chen</i>	
Biochemical Reactions as Renewal Processes: the case of mRNA Degradation.....	1574
<i>Paolo Paradisi and Davide Chiarugi</i>	
Multi-point accelerometric detection and principal component analysis of heart sounds ...	1577
<i>Simone De Panfilis, Carlo Moroni, Fabrizio Pompili, Giorgio Parisi and Rosario Cassone</i>	
A New Algorithm for Fetal QRS Detection in Abdominal Recordings.....	1578
<i>Bouabida Zohra, Hadj Slimane Zinne Eddine and Bereksi Requiig Fethi</i>	
Transcriptomic analysis of human liver identifies a novel class of regulatory RNAs in chronic viral hepatitis and associated cancer.....	1586
<i>Praveen Sethupathy</i>	
Docking analysis and catalytic site prediction of azoreductase in E. coli, with a wide range of industrially important azodyes	1587
<i>Samrat Adhikari, Samudra Sutradhar, Nangkyntiewbor Jungai, Bikash Thakuria and Gopesh Paul</i>	
Computational docking analysis on selective inhibition and binding affinity of synthetic inhibitors towards Matrix Metalloproteinase (MMPs) isoforms.....	1600
<i>Sudandiradoss C and Petrek Martin</i>	
CHEEK FORCE: A DEVICE FOR QUANTITATIVE EVALUATION	1601
<i>Monalise Berbert, Carlos Thomas and Rogério Marczak</i>	
A performance comparison between conventional SSVEP and Emokey based Emotiv EPOC matrix speller.....	1602
<i>S M Abdullah Al Mamun</i>	
Impact of missing genotype imputation on the power of Genome Wide Association Studies.....	1603
<i>Lukasz Król, Ghazi Alsbeih, Christophe Badie and Joanna Polanska</i>	
Automatic Peptides Selection for Targeted Proteomics.....	1615
<i>Yassene Mohammed, Dominik Domański, Angela M. Jackson, Derek S. Smith, André M. Deelder, Magnus N. Palmblad and Christoph H. Borchers</i>	
The Necessity for Improvement of the Algorithms Used for In Silico Allergenicity Assessment of Novel Proteins.....	1616
<i>Najaf Allahyari Fard</i>	
Discriminative Modeling of Cell Signaling as Bayesian Networks.....	1617
<i>Ayşe Gul Yaman, Aybar C Acar, Volkan Atalay and Rengul Cetin-Atalay</i>	

Instrumented 30-s Chair Stand Test: evaluation of an exercise program in frail nonagenarians	1618
<i>Nora Millor, Pablo Lecumberri, Alicia Martínez-Ramírez, Jon Martinikorena, Marisol Gómez and Mikel Izquierdo</i>	
An improved Fuzzy Clustering methodology applied to the study of Protein Conformational Ensembles	1630
<i>Duhu Man, Isabel María Timón-Pérez, Jesús Soto-Espinosa, Antonio Flores-Sintas, José M. Cecilia and Horacio Pérez-Sánchez</i>	
Data Recycling of Historical Records and Integration in New Information Systems in Cardiology ervice	1631
<i>R. Penafiel, Alberto Guillen and Luis Javier Herrera Maldonado</i>	
Advances in bioinformatics using soft-computing paradigms	1634
<i>Olga Valenzuela, Fernando Rojas, Francisco Ortuó, Jose Luis Bernier, M. Jose Saez, Belen San-Roman, Luis Javier Herrera, Alberto Guillen, Ignacio Rojas</i>	
Development of Soft-Computing techniques capable of diagnosing Alzheimers Disease in its pre-clinical stage combining MRI and FDG-PET images	1644
<i>Olga Valenzuela, Belen San-Roman, Francisco Ortuño, Jose Luis Bernier Villamor, Maria Jose Saez, Fernando Rojas and Ignacio Rojas</i>	
<hr/> Virtual Presentation <hr/>	
Hospital bed management support using regression data mining models	1651
<i>Sérgio Oliveira, Filipe Portela, Manuel Filipe Santos, José Machado and António Abelha</i>	
Introducing the Concept of Second Neighbours to FPNC algorithm for Improving the Functional Modules Detection	1662
<i>Mohammad Rahman and Nafisa Chowdhury</i>	
The Binding Sites of miR-619-5p, miR-5095, miR-5096 and miR-5585-3p in the Human mRNAs	1674
<i>Anatoly Ivashchenko, Olga Berillo, Anna Pyrkova, Raigul Niyazova and Shara Atambayeva</i>	
ERROR-RELATED POTENTIAL -IN BRAIN- ACTUATED WHEELCHAIR	1685
<i>Mana Taeb, Mohamad.B Shamsollahi, Farnaz Ghassemi and Behnam Asefi</i>	
Real-time True-color Volume Visualization of Multi-channel 3D CLSM Images Based on CUDA	1686
<i>Yakang Dai, Yunhai Zhang, Zhiyong Zhou, Haomin Yang and Xiaojun Xue</i>	
Computation Based Disease Associations in Disease Inference	1694
<i>Preeti Kale and Jagannath Aghav</i>	
Impact of Single amino acid Polymorphisms in Protein-Protein interactions in tumorigenic cluster A and cluster B of VHL: Computational molecular dynamics	1705
<i>George Priya Doss C, Chiranjib Chakraborty, Siddarth R, Nagasundaram N, Magesh R, Udhayakumar S, Priyanka N and Priyadharshini Christy J</i>	
Physiological Data Stream from Monitoring System in Intensive Care Unit	1717
<i>Fahmi Ben Rejab, Kaouther Noura and Bilel Amri</i>	

State and Parameter Estimation of a Neural Mass Model from Electrophysiological Signals during Induced Status Epilepticus	1729
<i>Armando López-Cuevas, Bernardino Castillo-Toledo, Laura Medina-Ceja and Consuelo Ventura-Mejía</i>	
Cardiac Arrhythmia Classification Using a combination of Quadratic Spline-Based Wavelet Transform and Artificial Neural Classification Network.....	1743
<i>Jose Antonio Gutierrez Gnechi, Rodrigo Morfin Magaña, Adriana Del Carmen Téllez Anguiano, Daniel Lorias Espinoza, Enrique Reyes-Archundia and Obeth Hernandez Diaz</i>	
Electrical Impedance Spectroscopy imaging of the thigh using current excitation frequencies in the mid- β frequency dispersion range.....	1755
<i>Jose Antonio Gutierrez Gnechi, Miguel Angel Mendoza Mendoza, Carlos Eduardo Guillen Nepita, Daniel Lorias Espinoza and Adriana Del Carmen Téllez Anguiano</i>	
Global Topology of Codon Usage Equality Networks of Escherichia Coli Essential Genes ..	1757
<i>Mohammad-Hadi Foroughmand-Araabi, Sama Goliaei and Bahram Goliaei</i>	
A parallel approach for accelerated parameter identification of Gene Regulatory Networks.....	1769
<i>Tariq Saeed and Jamil Ahmad</i>	
Enhancing Hotelling's T2 Statistic using Shrinkage Covariance Matrix for Identifying Differentially Expressed Gene Sets	1780
<i>Suryaefiza Karjanto, Rasimah Aripin, Norazan Mohamed Ramli and Nor Azura Md Ghani</i>	

Protein function easily investigated by genomics data mining using the ProteINSIDE web service

Nicolas KASPRIC^{1,2}, Brigitte PICARD^{1,2}, Matthieu REICHSTADT^{1,2}, Jérémy TOURNAYRE^{1,2} and Muriel BONNET^{1,2}

¹ INRA, UMR1213 Herbivores, F-63122 Saint-Genès-Champanelle, France

² Clermont Université, VetAgro Sup, UMR1213 Herbivores, BP 10448, F-63000, Clermont-Ferrand, France

Corresponding authors: nicolas.kaspric@clermont.inra.fr & muriel.bonnet@clermont.inra.fr

Abstract. ProteINSIDE is a new workflow to analyse lists of protein or gene identifiers from ruminant species and gather biological information provided by functional annotations, putative secretion of proteins and proteins interactions networks. ProteINSIDE gets results from several software and databases with a single query. From a unique list, ProteINSIDE uses orthologs identifiers within well studied species (Human, Rat or Mouse) to extend analyses and biological information retrieval. ProteINSIDE is freely available at: <http://www.proteinside.org>.

Keywords: web service, workflow, protein-protein interaction, protein secretion, gene ontology, networks.

1 Introduction

The current challenge for scientists working on the efficiency of ruminant (cattle, sheep or goat) and the quality of their products (meat, milk...) is to understand which genes and proteins control nutrient metabolism and partitioning between tissues or which genes and proteins control tissues growth and physiology [1]. Such questioning leads to the genome annotation, the sequencing and the quantification of gene expression or protein abundance. The quantity of data produced by these genomic and proteomic studies increases continuously [2-4]. There is a necessity to analyse, understand and generate biological information and knowledge from these data [5]. This is possible by using a panel of tools requiring different identifiers (IDs) per protein or gene and time to read and analyse the results. Moreover, most databases (DB) like UniProtKB [6] or NCBI [7] possess a large quantity of information and most of existing bioinformatic tools implemented as web services are specific to one analysis: as the annotation according to the Gene Ontology (GO) [8] or the prediction of signal peptide [9] or the molecular interactions identification [10] and visualization as networks [11, 12]. Many workflows that integrate several analyses are available [13-16]

and are specific to a species (*Drosophila*, *Arabidopsis thaliana*, *Escherichia coli*...), and thus are not suitable for the analysis of genomic data from ruminant species. The few workflows working with ruminant data are multispecies, the results are not species-specific and the data source is not available because of the privacy of databases (as the license software Pathway Studio [17]). Other workflows are specialized on the identification of candidate genes related to diseases as ToppGene [18]. Thus, to date there is no workflow dedicated to the integrative analysis of genomic data from ruminant species.

Unlike Human or model species like mouse or rat, ruminant species are less annotated and protein sequences are not always verified. Often, scientists use orthologs with the aim to increase the meaningful biological contexts for proteins. For this purpose biologists query for annotations according to Gene Ontology, the putative secretion of proteins, protein-protein interactions (PPI) and network analysis first in ruminant and then in Human or in rodents. The integration within a workflow of gateways between proteins / genes from ruminants and their orthologs from Human and models species has never been done.

Here we propose ProteINSIDE, a web service dedicated to a systematic and integrative analysis of protein's biological information. ProteINSIDE works using lists of proteins or genes IDs from 6 species (Bovine, Ovine, Caprine, Human, Rat, and Murine) to annotate functions and cellular location, predict secreted proteins, search for interactions between proteins within and/or outside a dataset and allowing cross-species analysis using orthologs.

2 Materials and methods

This section lists necessary equipment, ProteINSIDE resources and describes the dataset used to assess the functionalities of our tool.

2.1 Equipment

ProteINSIDE doesn't require an installation on a computer and the web service is available online at www.proteinside.org by using an internet browser. ProteINSIDE is completely adapted for any internet browser, but for better performances we recommend to use Firefox, Chrome, or Safari.

2.2 Implementation

ProteINSIDE is divided into three parts: the workflow, the database and the web interface. The workflow is a combination of Perl (version 5.10.1; CPAN modules (Comprehensive Perl Archive Network) used and BioPerl [19]) and R (version 3.0.1; with "tnet" package [20]) scripts to query databases, recover protein data, perform calculations and run algorithms for signal peptide predictions and network visualisation. The MySQL database aims to reduce server load and thus stores both available knowledge from major public databases and results (and settings) from queries. The

web interface is programmed in PHP, HTML, and JavaScript. It is devoted to the creation of a new analysis, the view of results and users information with updates.

2.3 ProteINSIDE structure and interface

A flow chart (Fig 1) details the type of analysis (basic or customizable) and the four main queries proposed to the user. Whatever the type of analysis, the workflow uses data from the input file and runs default scripts (basic analysis) or scripts and options selected through the settings (customs analysis). At analysis completion, results are created and uploaded on ProteINSIDE database to decrease web interface treatment duration (results have to be deleted by the user; visitors results are automatically deleted monthly).

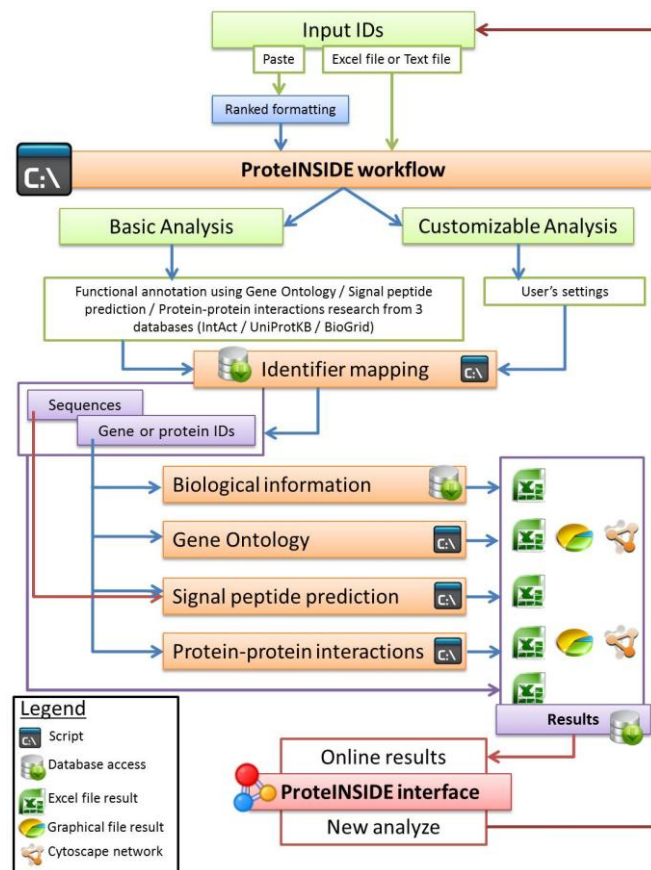


Fig. 1. Flow chart of ProteINSIDE structure. The four modules for querying the available biological information, annotations according to the gene ontology, signal peptide predictions and protein-protein interactions are either all present in the basic analysis or individually selected in the custom analysis.

ProteINSIDE is easily run by biologist through the interface. Registered user or visitor run a new analysis by using the web interface menus “Basic Analysis” (automatic settings) or “Custom Analysis” (user selects the settings):

1. Click on “Basic Analysis” menu on the homepage of ProteINSIDE
2. Fill in “the job name” box
3. Select the species for the analysis (related to the IDs that will be used on this query)
4. Upload your input file or directly paste your IDs
5. Click on the “Run the job” button to submit a new analysis

ProteINSIDE gives a link and an access code to view analysis status and get the results. The analysis status is indicated by the colour of a button: red for “analysis on the waiting list”, yellow for “the analysis is running” and green “analysis completion”. The blue globe is the link to access to the online results views:

1. Click on the blue globe button to view the results (use the trash to delete them)
2. Visualise the results summary produced by selected modules on the first default page
3. Navigate to module’s results pages by clicking on the module’s name on the toolbar menu.

2.4 The input and the output of ProteINSIDE

ProteINSIDE inputs are genes or proteins IDs (e.g. ADIPO or ADIPO_HUMAN) or UniProtKB protein accession numbers (e.g. Q15848). These IDs are uploaded as text tabulated files (extension .tab or .txt) or as Excel files (.xls or .xlsx). The input files have to be ranked as three columns (Fig. 2) because of the database format. Alternatively, the IDs are directly pasted.

	A	B	C
1	1	1	A4D1N9
2	2	2	ADIPO_HUMAN
3	3	3	C1T9A_HUMAN
4	4	4	F261_HUMAN
5	5	5	F262_HUMAN
6	6	6	PFKFB3
7	7	7	PFKFB4

Fig. 2. Example of an input files made using Excel 2010 and formatted for an upload.

The output files are Excel file (.xls), Cytoscape file (.cys or .xgmml), text or FASTA file (.txt or .fa) and pictures (.jpg or .png or .pdf). They are downloadable from the page results of each module of analysis.

2.5 The sample dataset

We created a dataset to assess ProteINSIDE performances. This dataset is composed of the UniProtKB accession numbers of 133 proteins (Table 1): 34 proteins related to the glycolysis cycle, 11 proteins from the respiratory chain, 5 proteins from the tricar-

boxylic acid cycle, 79 hormones or secreted proteins and proteins with very specific functions unrelated to the others. We also included a duplicated protein among proteins of the glycolysis to verify its recognition by ProteINSIDE.

We created this dataset on bovine species, but the numbers of annotations and PPI weren't sufficient for a clear representation of the functionalities of ProteINSIDE. Then, we used the same proteins in Human to test ProteINSIDE with the "Basic" and the "Custom Analysis" (Table 1).

Table 1. Results summary of ProteINSIDE analysis performances. The numbers are the proteins that belong to main pathways in the sample dataset, that are properly annotated by GO terms relevant to glycolysis and tricarboxylic acid (TCA) functions, and that have been predicted as secreted by SignalP for hormones.

Analyses and data	Glycolysis	Hormones	TCA	Analysis time (min)
Dataset	33+1 (duplicate)	79	5	-
Basic Analysis	29	78	3	2
Custom Analysis	33	78	5	10

3 Results and discussion

Here we present the results produced by a "Basic Analysis" and a "Custom Analysis" from our sample dataset, and we discuss the relevance of biological information extracted by ProteINSIDE. All of the 133 proteins were recognized by ProteINSIDE, the protein in duplicate was identified and excluded from the analysis (Fig. 3). Thus, 132 proteins were submitted to the analyses. The numbers of proteins / genes submitted to the analyses, numbers of annotations, PPI and predicted secreted proteins are recorded on the default page following the access to the results (Fig. 3).

General information	Analyze general results
Name: ProtHumainglycolyseChaineResTCA Type: Basic Species: HUMAN Date: 2014-01-23 Analysis parameters: ID Mapping / Gene Ontology / SignalP / Interaction research (IntAct / UniProtKB / BioGrid)	<input type="checkbox"/> Query find in ID Mapping DB: 133 <input type="checkbox"/> Blastp alignment executed: 0 <input type="checkbox"/> All GO detected: 568 for 123 annotated protein(s) <input type="checkbox"/> Signal peptides detected: 85 <input type="checkbox"/> Interactions detected on the dataset: 36
Incomplet informations <input type="checkbox"/> Incomplet query: 12 + show/hide + <input type="checkbox"/> Duplicate query: 1 + show/hide +	

Fig.3. A table is provided by the default page after the access to the results both for a basic and a custom analysis. In addition to general information, the table provide counts of results retrieved by each module that has been run, incomplete query (IDs with missing biological information) and duplicate query.

3.1 Results of the Basic Analysis

The first module of analysis has extracted and summarized, as a downloadable table, other gene or protein IDs, gene or protein names, a summary for the protein function, the gene chromosomal location, information on tissue expression and cellular location, and the species in which orthologs have been identified. Thus on the “ID resume” page of the toolbar menu, a user has access at a glance to several information for a list of genes or their products, and also to direct links with the UniProtKB and the NCBI databases.

Multiple annotations GOs

No of annotated protein by GO: 17,0 (slider) 63,0

Function: Proteins: Gene Name: Number of items: 10 Gene products detected on: HUMAN

GO	Function	Proteins	Gene Name	Ontology group	Count	GO frequency (%)	Number of Gene products detected / expected	Percent of Gene products detected / expected
GO:0006096	glycolysis	Q01813 Q16877 P09104 P16118 P00338 P00558 P15259 P08237 P04406 P14618 P30613 P07205 P18669 P52789 P60174 P36871 Q60825 P04075 P19367 Q75356 Q16875 P09972 P06733 Q14556 P13929 P05062 P17858 P06744	PFKP PFKFB4 ENO2 PFKFB1 LDHA PGK1 PGAM2 PFKM GAPDH PKM PKLR PGK2 PFAM1 HK2 TP1 PGM1 PFKFB2 ALDOA HK1 ENTPD5 PFKFB3 ALDOC ENO1 GAPDH3 ENO3 ALDOB PFKL GPI	BP	28	22.764	28 /45	62.222
GO:0005184	neuropeptide hormone activity	P15089 P10092 P01282 P06850 P11509 P01185 P22466 P06307	UCN CALCB VIP CRH ADCYAP1 AVP GAL CCK	MF	8	6.504	8 /15	53.333
GO:0005179	hormone activity	P11148 P06850 P01298 P06307 P81277 P09529 P12272 P01242 P01282 P61278 Q9Y581 P04808 P01350 Q14406 Q95399 P01225 Q43555 P01189 P10082 P09681 Q15848 P01308 P09683 P01233 P52823 P08476 P01270 P05111 P18860 Q76051 P01215 P01222	GNRH1 CRH PPY CCK PRLH INHBB PTHLH GH2 VIP SST INSL6 RLH1 GAST CSHL1 UTS2 FSHB GNRH2 POMC PPY GIP ADIPOQ INS SCT CGB8 STC1 INHBA PTH INHA NPPB STC2 CGA TSHB	MF	32	26.016	32 /62	51.613

Fig. 4. Results of the functional annotations according to the Gene Ontology are available as dynamic tables. Results can be sorted by: the GO’s identifier, the function, protein ID or gene name, the ontology group, the number of annotated proteins or the number of expected gene products.

On the “GO” page of the toolbar menu, we checked the relevance of the annotations extracted by ProteINSIDE by looking for the over-representation of annotations relative to glycolysis and hormones. First, among the 132 proteins submitted, ProteINSIDE annotates 123 proteins with 568 unique GOs (Fig 3). We classed these GO according to the number of proteins annotated by GO and the percentage of gene products detected/expected to identify the most common pathways associated to our sample dataset (Fig. 4). By this way, we retrieved as the most common pathways: glycolysis and hormone activity (about 62% and 52% of expected annotated gene products with these GO in Human, respectively). We have to note a lack of annotations for 12 proteins of the sample dataset, and a lack of annotations relative to glycolysis for 4 proteins (28 of the 33 expected proteins related to the glycolysis were annotated; Table 1). This lack of annotations is related to our choice to use only GO terms that have been agreed by review curator in the “Basic Analysis”. This means that the “Basic Analysis” doesn’t use GO annotations with IEA (Inferred by Electronic Annotation) evidence code, but the option to use IEA is provided in the custom analysis to extend the annotations.

On the “Secreted protein” page of the toolbar menu, the proteins potentially secreted are listed in a dynamic table (Fig. 3 and 5). From our sample dataset, 85 proteins were predicted as secreted by SignalP [9], among them 78 of the 79 proteins that were expected (Table 1). This lack of perfect prediction of the protein can be explained by the false positive and false negative prediction rates of SignalP, as already evaluated [21]. The prediction of secretion is then confirmed by a search for GOs related to the “secretion” function. Over the 85 predicted secreted proteins, 63 were annotated by GOs related to the “secretion” function. The double query of protein secretion both by the peptide signal prediction from protein sequence and the GO annotation is unique to ProteINSIDE.

Proteins	GO related to secretion	Gene Name	Number of rows	Signal peptides detected	Download table
			10	85 (on 133 proteins imported)	

Proteins	Protein ID	Gene Name	Peptide	GO related to secretion	Number of GO
Q9UBU3	GHRL_HUMAN	GHRL	noTM	GO_0051464 GO_0060124 GO_0005576 GO_0034774 GO_0005615 GO_0030252 GO_0051461 GO_0032024 GO_0043400	9
P01308	INS_HUMAN	INS	noTM	GO_0050796 GO_0005576 GO_0060715 GO_0034774 GO_0005615 GO_0090277 GO_0050708	7
Q15848	ADIPO_HUMAN	ADIPOQ	noTM	GO_0005576 GO_0005615 GO_0045715 GO_0034393	4
P01189	COL1_HUMAN	POMC	noTM	GO_0005576 GO_0034774 GO_0005615 GO_0030141	4
P08476	INHBA_HUMAN	INHBA	noTM	GO_0046881 GO_0005576 GO_0042701 GO_0046880	4
P16860	ANFB_HUMAN	NPPB	noTM	GO_0005576 GO_0005615 GO_0007589	3
P06850	CRF_HUMAN	CRH	noTM	GO_0051464 GO_0005576 GO_0005615	3
P09681	GIP_HUMAN	GIP	noTM	GO_0050796 GO_0005576 GO_0034774	3
P01275	GLUC_HUMAN	GCG	noTM	GO_0050796 GO_0005576 GO_0034774	3
P35318	ADML_HUMAN	ADM	noTM	GO_0005576 GO_0005615	2

prev next 1 2 3 4 5 6 7 8 9

Fig. 5. Results as a dynamic table, of the potentially secreted proteins predict by SignalP.

On the “Protein interactions” page of the toolbar menu, proteins within the dataset are linked by the “Interaction detection methods” or reviewed by a curator (clicking on node gives information about the protein and a link to UniProtKB database). We selected to query BioGrid [22], UniProtKB [6] and IntAct [23] because these PPI databases are reviewed by curators and the query of PPI in 2 or 3 PPI database delivered best results [24]. PPI are listed by a dynamic table or viewed as a network (Fig. 6). The interactions research between proteins of our sample dataset has identified 36 PPI that involved 23 different proteins. As expected, PPI within the sample dataset linked proteins known to contribute to the pyruvate dehydrogenase complex (Fig 6A), the complexes IV (Fig 6B) and I (Fig 6C) of the respiratory chain, and also some proteins linked to the glycolysis and the carbohydrate oxidation (Fig 6D and 6E).

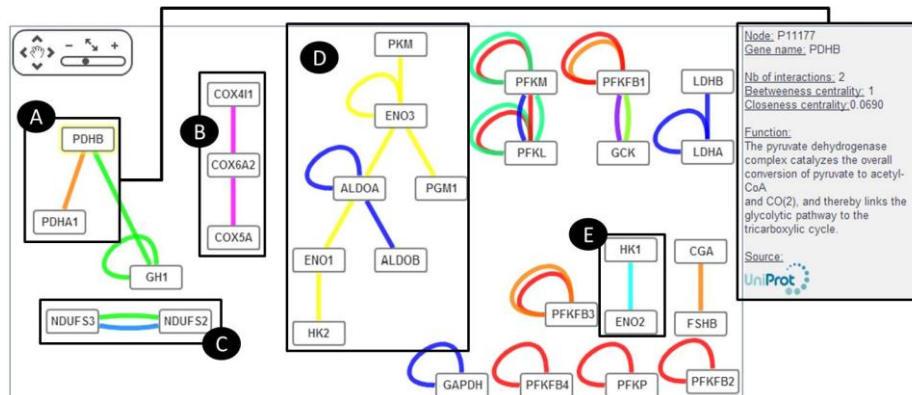


Fig. 6. Network of PPI retrieved by querying BioGrid, UniProtKB and IntAct databases. PPI are between proteins within the sample dataset. Information about a protein is obtained by clicking on a protein/gene or node. Edge colour depends on the detection method of the PPI.

3.2 The Custom Analysis: an added-value provided by the extension of the analysis

We made a “Custom Analysis” using the same major settings as the “Basic Analysis” but we used the proposed additional options:

1. The GO electronic annotation (IEA) evidence codes proposed to extend the annotation
2. GOTree network to view linked GOs
3. The search and the view of the PPI between proteins from our sample dataset and proteins outside the dataset (but still in the same species, here in Human) to extend the network and the biological information.

The use of electronic annotation has increased both the number of annotated proteins (132 rather than 123 without IEA in the basic analysis) and of annotations by around 50% since 1031 unique GOs were retrieved by ProteINSIDE. Over the 33 expected proteins related to the glycolysis, the Custom analysis of ProteINSIDE has annotated 32 proteins with the GO 0006096, glycolysis (Table 1). The GOTree network linked 236 GOs. We have chosen to visualize the GOs of the “Molecular Function” group (Fig. 7). In this visualisation, the dark red colour represents the most common GO associated to our sample dataset. As expected the GO:0005179, nominates “Hormone activity”, which is consistent with the over-representation of hormones in our sample dataset. This network has also linked more specific GOs or child terms [25] of the “Hormone activity” GO, as for example “Neuropeptide hormone activity” (GO:0005184).

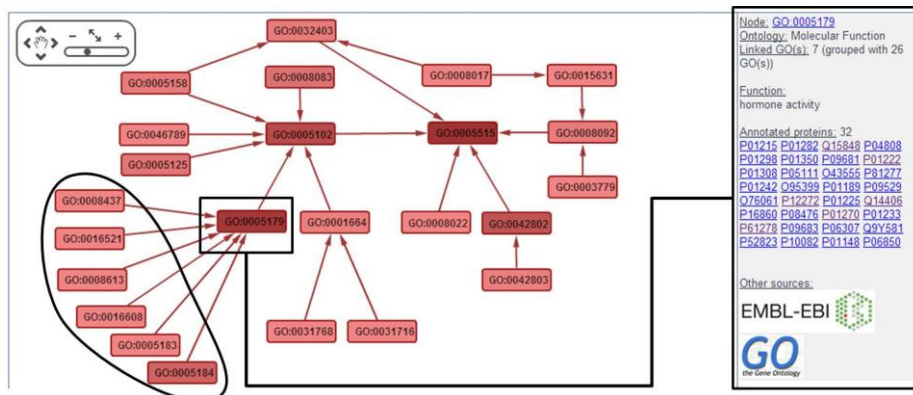


Fig. 7. A Network that links GOs used to annotate proteins of the sample dataset. Red colour is only for the GO terms relative to the Molecular Function. The degree of colour saturation represents the quantity of proteins annotated by a GO. Each edge means that a term A is a sub-type of a term B (is_a). Information about a GO is obtained by clicking on the GO or the node.

Lastly, the same proteins as the “Basic Analysis” were predicted as secreted (Table 1). Thanks to the IEA electronic annotation, 82 proteins over the 85 proteins predicted to be secreted by SignalP, were also annotated by GOs related to the “secretion” function. It’s 19 more than the 63 of the “Basic Analysis” because of the use of IEA evidence code for GO annotation.

By comparison with “Basic Analysis”, “Custom Analysis” searches for PPI between the proteins within and outside the sample dataset by querying up to 28 DB. We chose to query the same 3 DB (BioGrid, UniProtKB and IntAct) to compare with the “Basic Analysis”. ProteINSIDE retrieved 616 PPI made by 221 proteins among them 61 from the dataset. We visualized the network of the PPI (Fig. 8) and we retrieved some sub-networks relative to the respiratory chain (Fig. 8A), hormone activity such as signalization by adipokines (Fig. 8B), the growth hormone (Fig. 8C) and thyroid hormones (Fig. 8D), as well as sub-networks relative to glycolysis and carbohydrate metabolism (not highlighted).

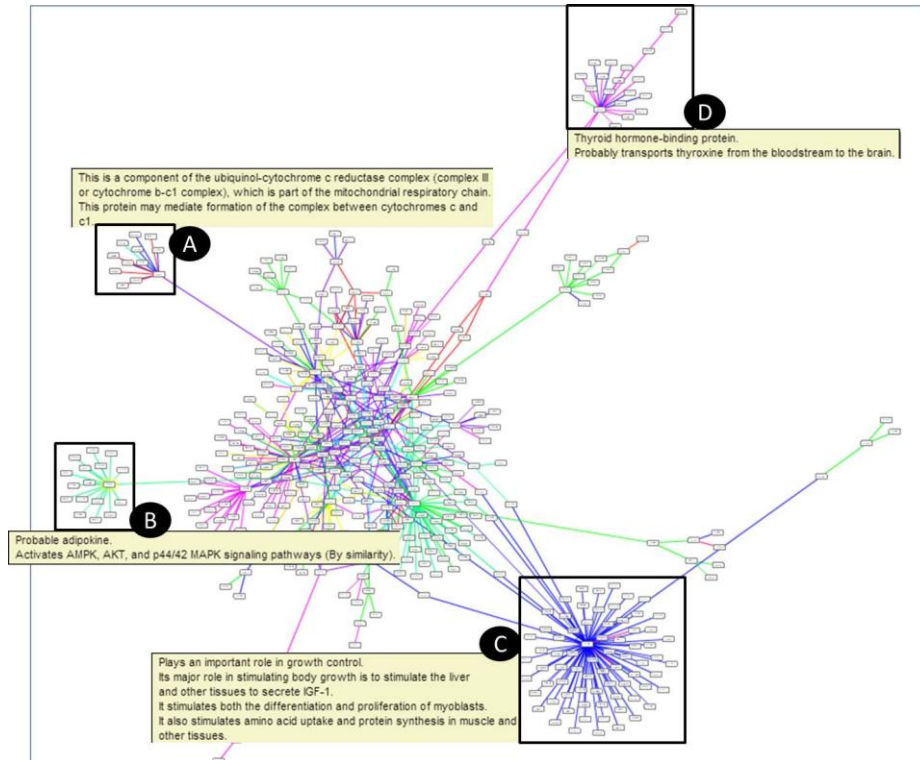


Fig. 8. Network of PPI retrieved by querying the BioGrid, UniProtKB and IntAct DB and using PPI with human proteins outside of the dataset.

4 Conclusion

In this work we present the performances of ProteINSIDE, a new powerful workflow which gathers tools and public databases to retrieve biological information of genes or proteins lists from 6 species (Bovine, Ovine, Caprine, Human, Rat, and Murine). The presented web service has correctly identified a dataset of 133 proteins, has excluded a duplicate query and has retrieved biological information for each protein. According to our dataset, ProteINSIDE properly annotates the proteins related to the glycolysis, the proteins affiliated as hormones, and the putatively secreted proteins. ProteINSIDE has revealed the most common pathways related to our dataset by creating networks from PPI interactions within and outside the dataset and from links between GOs. Each result is easily accessible and downloadable.

ProteINSIDE offers a great support to analyse a large quantity of data from genomic and proteomic studies. ProteINSIDE is also the unique web service that makes all of these analyse using ruminant IDs.

Acknowledgment

This work was supported by the Region Auvergne (FRANCE) and Apis-Gene (FRANCE).

References

1. Bonnet M, Cassar-Malek I, Chilliard Y, Picard B. "Ontogenesis of muscle and adipose tissues and their interactions in ruminants and other species". *Animal : an international journal of animal bioscience*, 4(7), pp. 1093-109. PubMed PMID: 22444612. (2010)
2. Picard B, Cassar-Malek I, Guillemain N, Bonnet M. Quest for Novel Muscle Pathway Biomarkers by Proteomics in Beef Production. In: Moo-Young M, editor. *Comprehensive Biotechnology (Second Edition)*. Burlington: Academic Press; 2011. p. 395-405.
3. Chaze T, Meunier B, Chambon C, Jurie C, Picard B. "Proteome dynamics during contractile and metabolic differentiation of bovine foetal muscle". *Animal : an international journal of animal bioscience*, 3(7), pp. 980-1000. PubMed PMID: 22444818. (2009)
4. Taga H, Chilliard Y, Meunier B, Chambon C, Picard B, Zingaretti MC, et al. "Cellular and molecular large-scale features of fetal adipose tissue: is bovine perirenal adipose tissue brown?". *Journal of cellular physiology*, 227(4), pp. 1688-700. PubMed PMID: 21678425. (2012)
5. Woelders H, Te Pas MF, Bannink A, Veerkamp RF, Smits MA. "Systems biology in animal sciences". *Animal : an international journal of animal bioscience*, 5(7), pp. 1036-47. PubMed PMID: 22440099. (2011)
6. Magrane M, Consortium U. "UniProt Knowledgebase: a hub of integrated protein data". *Database : the journal of biological databases and curation*, 2011, pp. bar009. PubMed PMID: 21447597. Pubmed Central PMCID: 3070428. (2011)
7. Coordinators NR. "Database resources of the National Center for Biotechnology Information". *Nucleic Acids Res*, 41(Database issue), pp. D8-D20. PubMed PMID: 23193264. Pubmed Central PMCID: 3531099. (2013)
8. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. "QuickGO: a web-based tool for Gene Ontology searching". *Bioinformatics*, 25(22), pp. 3045-6. PubMed PMID: 19744993. Pubmed Central PMCID: 2773257. (2009)
9. Petersen TN, Brunak S, von Heijne G, Nielsen H. "SignalP 4.0: discriminating signal peptides from transmembrane regions". *Nature methods*, 8(10), pp. 785-6. PubMed PMID: 21959131. (2011)
10. Aranda B, Blankenburg H, Kerrien S, Brinkman FS, Ceol A, Chautard E, et al. "PSICQUIC and PSISCORE: accessing and scoring molecular interactions". *Nature methods*, 8(7), pp. 528-9. PubMed PMID: 21716279. Pubmed Central PMCID: 3246345. (2011)

11. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. "Cytoscape 2.8: new features for data integration and network visualization". *Bioinformatics*, 27(3), pp. 431-2. PubMed PMID: 21149340. Pubmed Central PMCID: 3031041. (2011)
12. Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. "Cytoscape Web: an interactive web-based network browser". *Bioinformatics*, 26(18), pp. 2347-8. PubMed PMID: 20656902. Pubmed Central PMCID: 2935447. (2010)
13. Wong AK, Park CY, Greene CS, Bongo LA, Guan Y, Troyanskaya OG. "IMP: a multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks". *Nucleic Acids Res*, 40(Web Server issue), pp. W484-90. PubMed PMID: 22684505. Pubmed Central PMCID: 3394282. Epub 2012/06/12. eng. (2012)
14. Pache RA, Ceol A, Aloy P. "NetAligner--a network alignment server to compare complexes, pathways and whole interactomes". *Nucleic Acids Res*, 40(Web Server issue), pp. W157-61. PubMed PMID: 22618871. Pubmed Central PMCID: 3394252. Epub 2012/05/24. eng. (2012)
15. Renaud Y, Baillif A, Perez JB, Agier M, Mephu Nguifo E, Mirouse V. "DroPNet: a web portal for integrated analysis of Drosophila protein-protein interaction networks". *Nucleic Acids Res*, 40(Web Server issue), pp. W134-9. PubMed PMID: 22641854. Pubmed Central PMCID: 3394298. Epub 2012/05/30. eng. (2012)
16. Tuncbag N, McCallum S, Huang SS, Fraenkel E. "SteinerNet: a web server for integrating 'omic' data to discover hidden components of response pathways". *Nucleic Acids Res*, 40(Web Server issue), pp. W505-9. PubMed PMID: 22638579. Pubmed Central PMCID: 3394335. Epub 2012/05/29. eng. (2012)
17. Nikitin A, Egorov S, Daraselia N, Mazo I. "Pathway studio--the analysis and navigation of molecular networks". *Bioinformatics*, 19(16), pp. 2155-7. PubMed PMID: 14594725. (2003)
18. Chen J, Bardes EE, Aronow BJ, Jegga AG. "ToppGene Suite for gene list enrichment analysis and candidate gene prioritization". *Nucleic Acids Res*, 37(Web Server issue), pp. W305-11. PubMed PMID: 19465376. Pubmed Central PMCID: 2703978. (2009)
19. Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, Dagdigian C, et al. "The Bioperl toolkit: Perl modules for the life sciences". *Genome research*, 12(10), pp. 1611-8. PubMed PMID: 12368254. Pubmed Central PMCID: 187536. (2002)
20. T. O. "Structure and Evolution of Weighted Networks". University of London (Queen Mary College), London, UK, pp. 104-22. (2009)
21. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. "Locating proteins in the cell using TargetP, SignalP and related tools". *Nat Protoc*, 2(4), pp. 953-71. PubMed PMID: 17446895. (2007)
22. Chatr-Aryamontri A, Breitkreutz BJ, Heinicke S, Boucher L, Winter A, Stark C, et al. "The BioGRID interaction database: 2013 update". *Nucleic Acids Res*, 41(Database issue), pp. D816-23. PubMed PMID: 23203989. Pubmed Central PMCID: 3531226. (2013)
23. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, et al. "The IntAct molecular interaction database in 2012". *Nucleic Acids Res*, 40(Database

issue), pp. D841-6. PubMed PMID: 22121220. Pubmed Central PMCID: 3245075. (2012)

24. Martha VS, Liu Z, Guo L, Su Z, Ye Y, Fang H, et al. "Constructing a robust protein-protein interaction network by integrating multiple public databases". *BMC Bioinformatics*, 12 Suppl 10, pp. S7. PubMed PMID: 22165958. Pubmed Central PMCID: 3236850. (2011)

25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium". *Nature genetics*, 25(1), pp. 25-9. PubMed PMID: 10802651. Pubmed Central PMCID: 3037419. (2000)