NETTAB 2017

Methods, tools & platforms for Personalized Medicine in the Big Data Era

October 16-18, 2017

ICAR CNR Site
Via Ugo La Malfa, 153, Palermo, Sicily, Italy

Scientific Programme

NETTAB 2017

October 16-18, 2017, Palermo, Sicily, Italy

Methods, tools & platforms for Personalized Medicine in the Big Data Era

Scientific Programme at a glance

Monday October 16, 2017	
13.30-14.30	Registration and poster hang-up
14.30-14.45	Welcome and Introduction
14.45-15.30	Keynote Lecture – Anita Grigoriadis
15.30-16.05	Scientific Session 1
16.05-16.50	Keynote Lecture – Alfonso Valencia
16.50-17.20	Coffee break
17.20-17.35	Flash Poster presentations 1
17.35-18.05	Guest Lecture – Raffaele Giancarlo
18.05-18.45	Scientific Session 2
18.45	Day closure

Tuesday October 17, 2017	
08.30-09.00	Registration and poster hang-up
09.00-09.55	Scientific Session 3
09.55-10.40	Keynote Lecture – Winston Hide
10.40-10.55	Flash Poster presentations 2
10.55-12.15	Poster session with coffee break service
12.15-12.45	Guest Lecture – Francis Ouellette
12.45-13.15	Guest Lecture – Alexander Kel
13.15-14.30	Lunch Break
14.30-15.05	Scientific Session 4
15.05-15.35	Guest Lecture – Matthias Reumann
15.35-16.30	Panel Discussion – Giovanni Vizzini and all speakers
16.30	Day closure
16.30-20.30	Social activity – Guided Tour to Cefalù (transfer by bus)
20.30-23.00	Social Dinner in Cefalù

Wednesday October 18, 2017	
08.30-09.00	Registration
09.00-09.30	Guest Lecture – Luana Licata
09.30-10.15	Keynote lecture – Inna Kuperstein
10.15-11.05	Scientific Session 5
11.05-12.15	Poster session with coffee break service
12.15-12.50	Scientific Session 6
12.50-13.20	Guest Lecture – Emanuela Merelli
13.20-13.30	NETTAB 2018 Announcement
13.30	Workshop Closure and Farewell

OSPEDALE POLICLINICO SAN MARTINO

NETTAB 2017
October 16-18, 2017, Palermo, Sicily, Italy
ICAR CNR Site, Via Ugo La Malfa, 153, Palermo

Methods, tools & platforms for Personalized Medicine in the Big Data Era

Scientific Programme

Monday October 16, 2017	
13.30-14.30	Registration and poster hang-up
14.30-14.45	***Welcome and Introduction*** G. De Pietro,HPC & Networking Institute, National Research Council, Palermo, Italy P. Romano, IRCCS Ospedale Policlinico San Martino, Genoa, Italy L. Milanesi, BITS Bioinformatics Italian Society, Italy A. M. Urso, HPC & Networking Institute, National Research Council, Palermo, Italy
14.45-15.30	***Keynote Lecture***
	Interoperability of clinical, pathological and omics data to execute personalised medicine Anita Grigoriadis, King's College, London, United Kingdom
15.30-16.05	***Scientific Session 1***
	Epistasis analysis reveals associations between gene variants and bipolar disorder C. Maj, E. Milanesi, M. Gennarelli, L. Milanesi and I. Merelli
	Identification of Mutational Signatures Active in Individual Tumors S. Krüger and R. M. Piro
16.05-16.50	***Keynote Lecture***
	Disease comorbidities and network approaches Alfonso Valencia, Barcelona Supercomputing Center BSC, Spain
16.50-17.20	***Coffee break***
17.20-17.35	***Flash Poster presentations 1***
	P1 Visualizing Mutation Occurrence using Big Data S. Albert
	P2 A Gene Set Enrichment Analysis of multiomic celiac disease data E. Del Prete, A. Facchiano and P. Liò
	P3 Network diffusion on multiple-layers: current approaches and integrative analysis of Rheumatoid Arthritis data. N. Di Nanni, M. Gnocchi, M. Moscatelli, L. Milanesi and E. Mosca
	P4 Evolutionary relationships of Microbial Transglutaminases D. Giordano and A. Facchiano
	P5 Exploiting transcriptomic data in genome scale metabolic networks: new insights into obesity I. Granata, M. Sangiovanni, E. Troiano and M. Guarracino
	P6 Implementing a multilayer framework for pathway data integration, analysis and visualization Z. Hammoud and F. Kramer
	P7 Withdrawn

OSPEDALE POLICLINICO SAN MARTINO

17.35-18.05	***Guest Lecture***
	Getting Beyond Proof of Principle for Big Data Technologies in Bioinformatics: MapReduce Algorithmic, Programming and Architectural issues *Raffaele Giancarlo, University of Palermo, Italy*
18.05-18.45	***Scientific Session 2***
	Melanoma expression analysis with Big Data technologies *A. Fernandez-Rovira, R. Lavado, M. A. Berciano Guerrero, I. Navas-Delgado and J. F Aldana Montes*
	GenHap: A Novel Computational Method Based on Genetic Algorithms for Haplotype Assembly *A. Tangherloni, S. Spolaor, L. Rundo, M. S. Nobile, I. Merelli, P. Cazzaniga, D. Besozzi, G. Mauri and P. Liò*
18.45	***Day closure***

Tuesday October 17, 2017	
08.30-09.00	*Registration and poster hang-up*
09.00-09.55	***Scientific Session 3***
	Parameters tuning boosts hyperSMURF predictions of rare deleterious non-coding genetic variants *A. Petrini, M. Schubach, M. Re, M. Frasca, M. Mesiti, G. Grossi, T. Castrignanò, P. Robinson and G. Valentini*
	Greedy motif-based approach to parsing large and diverge coiled-coil proteins into domains *H. Khakzad, J. Malmström and L. Malmström*
	From trash to treasure: detecting unexpected contamination in unmapped NGS data *I. Granata, M. Sangiovanni, M. R. Guarracino and A. Singh Thind*
09.55-10.40	***Keynote Lecture***
	Making genomics Come true: How can we achieve real acceleration of genomics into medicine? *Winston Hide, University of Sheffield, United Kingdom*
10.40-10.55	***Flash Poster presentations 2***
	P8 Querying and analyzing biological data with BioGraph *A. Messina, A. Fiannaca, L. La Paglia, M. La Rosa and A. M. Urso*
	P10 ISMARA: Completely automated inference of gene regulatory networks from high-throughput data *M. Pachkov, P. Balwierz, P. Arnold, A. Gruber, M. Zavolan and E. van Nimwegen*
	P11 An Algebraic Representation for Tree Alignment of RNA Pseudoknotted Structures *M: Quadrini, L. Tesei and E. Merelli*
	P12 Procedure to integrate i2b2 and REDCap: a case study at ICSM *V. Tibollo, M. Bucalo, D. Vella, M. Stuppia, N. Barbarini and R. Bellazzi*
	P13 Stability Analysis of MTopGO for Module Identification in PPI Networks *D. Vella, A. Tucker and R. Bellazzi*
	P14 A Quality Management System for scientific research activities and its related management software *L. Caruana, A. Pensato, L. Riccobono, G. L. Liguori, A. Lanati, A. Kisslinger, M. Di Carlo and A. Bongiovanni*
10.55-12.15	***Poster session with coffee break service***
12.15-12.45	***Guest Lecture***
	Open Data is Essential for Personalized Medicine *Francis Ouellette, Génome Québec, Canada*
12.45-13.15	***Guest Lecture***
	Walking pathways in cancer *Alexander Kel, GeneXplain, Germany*
13.15-14.30	***Lunch Break***

14.30-15.05	***Scientific Session 4***
	MTopGO: a tool for module identification in PPI Networks *D. Vella, S. Marini, F. Vitali and R. Bellazzi*
	GenotypeAnalytics: an RESTful Platform to mine multiple associations between SNPs and drug response in case-control studies *G. Agapito, P. H. Guzzi and M: Cannataro*
15.05-15.35	***Guest Lecture***
	Big data and cognitive computing in healthcare: weathering the perfect storm *Matthias Reumann, IBM Research - Zurich, Switzerland*
15.35-16.30	***Panel Discussion***
	Pairing expectations and achievements in Precision Medicine *Giovanni Vizzini, UPMC Italy, Palermo and all speakers*
16.30	***Day closure***
16.30-20.30	***Social activity – Guided visit to Cefalù (transfer by bus)***
20.30-23.00	***Social Dinner in Cefalù***

Wednesday October 18, 2017	
08.30-09.00	*Registration and poster hang-up*
09.00-09.30	***Guest Lecture***
	ELIXIR-IIB – the Italian Infrastructure for Bioinformatics: a growing support to national and international research in life sciences *Luana Licata, ELIXIR-IIB, Italy*
09.30-10.15	***Keynote lecture***
	Atlas of Cancer Signaling Network and NaviCell: Systems Biology resources for studying cancer biology *Inna Kuperstein, Institut Curie, Paris, France*
10.15-11.05	***Scientific Session 5***
	Personalised models for human – gut microbiota interaction *D. Gilbert, M. Heiner and L. Ghanbar*
	A Dynamic Bayesian Network model for simulation of disease progression in Amyotrophic Lateral Sclerosis patients *A. Zandonà, M. Francescon, M. Bronfeld, A. Calvo, A. Chiò and B. Di Camillo*
	Multi-resolution network modelling of T-cells for precisions medicine of multiple sclerosis *M. Gustafsson*
11.05-12.15	***Poster session with coffee break service***
12.15-12.50	***Scientific Session 6***
	PGxO: A very lite ontology to reconcile pharmacogenomic knowledge units *P. Monnin, C. Jonquet, J. Legrand, A. Napoli and A. Coulet*
	Interoperability of RehabRobo-Onto *Z. Dogmus, V. Patoglu and E. Erdem*
12.50-13.20	***Guest Lecture***
	Topological Field Theory of Data: a new venue for Biomedical Big Data Analysis *Emanuela Merelli, University of Camerino, Italy*
13.20-13.30	***First announcement of NETTAB 2018*** *Paolo Romano, IRCCS Ospedale Policlinico San Martino, Genoa, Italy*
13.30	***Workshop Closure and Farewell***

KEYNOTE TALKS

Interoperability of clinical, pathological and omics data to execute personalised medicine

Anita Grigoriadis, King's College, London, UK.

Translational research has seen an increasing trend towards omics techniques imaging approaches, in combination with clinical and pathological data. Multifactorial data, both large in sample size and heterogeneous in context, needs to be integrated in a standardised, cost-effective, secure manner so that it can be utilised and searched by researchers and clinicians. Small- to moderate-sized research groups need to find solutions to handle and administer enormous data volumes whilst researching for new discoveries.

Here, I represent solutions to support data management, the integration of digital microscopy and pathology, and illustrate the utility of R-shiny to make high-throughput data searchable.

Disease comorbidities and network approaches

Alfonso Valencia, Barcelona Supercomputing Center (CBS), Spain.

Abstract not available.

Making genomics Come true: How can we achieve real acceleration of genomics into medicine?

Winston Hide, University of Sheffield, UK, and Harvard TH Chan School of Public Health, Boston, USA.

We are now rapidly moving from single human genomes to deca-, centi- and even milligenome projects. With more ways to compare gene variation against a background, comes new methods to select variants and genes for their potential in prediction and impact for a disease. Gene hunting is still very much a fashion and genes represent tempting targets for drug development.

But like David Bowie we need to push the boundaries to embrace the growing realisation that genes work in cohorts and it is the interaction of these cohorts that drive the disease phenotype. Identifying and targeting pathways and processes that drive disease is the new black. To action discovery, we need to address ways in which to benchmark selection of disease genes, pathways and processes. In turn we need to develop more efficient (read less ineffective) ways to select therapeutics that are likely to be acceptable for real health interventions.

The talk will present how we address these challenges through commoning for data sharing, provenance, reproducibility and workflows, benchmarks for assessment of approaches, standardisation for pathway activity, and integrative approaches to discovering the relationships between therapeutic target prioritisation, network topology, pathway interaction, genome variation, disease modelling and drug repurposing.

Atlas of Cancer Signaling Network and NaviCell: Systems Biology resources for studying cancer biology

Inna Kuperstein, Institut Curie, Paris, France

Studying reciprocal regulations between cancer-related pathways is essential for understanding signaling rewiring during cancer evolution and in response to treatments. With this aim we have constructed the Atlas of Cancer Signaling Network (ACSN), a resource of cancer signaling maps and tools with interactive web-based environment for navigation, curation and data visualization. The content of ACSN is represented as a seamless 'geographic-like' map browsable using the Google Maps engine and semantic zooming. The associated blog provides a forum for commenting and curating the ACSN maps content. The integrated NaviCell web-based tool box allows to import and visualize heterogeneous omics data on top of the ACSN maps and to perform functional analysis of the maps. The tool contains standard heatmaps, barplots and glyphs as well as the novel map staining technique for grasping large-scale trends in numerical values projected onto a pathway map.

To demonstrate applications of ACSN and NaviCell we show a study on drug sensitivity prediction using the networks. We performed a structural analysis of Cell Cycle and DNA repair signaling network together with omics data from ovary cancer patients resistant to genotoxic treatment. Following this study we retrieved synthetic lethal gene sets and suggested intervention gene combinations to restore sensitivity to the treatment. In another example, analysis of cell lines multi-level omics data, interpreted in the context of signaling network

maps, highlighted different DNA repair molecular profiles associated with sensitivity to each one of the drugs, rationalizing combined treatment in some cases.

Analysis of multi-omics data together with cell signaling information helps finding personalized treatments. In additional study we show how epithelial to mesenchymal transition (EMT) signaling network from the ACSN collection has been used for finding metastasis inducers in colon cancer through network analysis. We performed structural analysis of EMT signaling network that allowed highlighting the network organization principles and complexity reduction up to core regulatory routs. Using the reduced network we modeled single and double mutants for achieving the metastasis phenotype. We predicted that a combination of p53 knock-out and overexpression of Notch would induce metastasis and suggested the molecular mechanism. This prediction lead to generation of colon cancer mice model with metastases in distant organs. We confirmed in invasive human colon cancer samples the modulation of Notch and p53 gene expression in similar manner as in the mice model, supporting a synergy between these genes to permit metastasis induction in colon.

GUEST TALKS

Getting Beyond Proof of Principle for Big Data Technologies in Bioinformatics: MapReduce Algorithmic, Programming and Architectural issues

Raffaele Giancarlo, University of Palermo, Italy.

High Performance Computing (HPC) in Bioinformatics has a classic architectural paradigm: shared-memory multi-processor. With the advent of Cloud Computing, such a new way of managing Big Data has been considered for Bioinformatics. Initially, with Proof of Concept results investigating the advantages of the new computational paradigm. They have been followed by an increasing number of specific Bioinformatics tasks, developed mainly with the use of the MapReduce programming paradigm, which is in turn supported by Hadoop and Spark Middleware.

A careful analysis of the State of the Art indicates that the main advantage of those Big Data Technologies is the perception of boundless scalability, at least in terms of time. However, how effectively the computing resources are used in the Cloud… is rather cloudy, as most of the software available almost entirely delegates the management of the distributed workload to the powerful primitives of Hadoop and Spark. On a private cloud, i.e., a physical computing cluster that can be configured at will by the user, one can show that carefully designed MapReduce algorithms largely outperform the ones that naively "delegate" to Hadoop and Spark. In the public cloud, e.g., virtual clusters (for instance created via OpenStack) with a dynamic and instance-dependent allocation of physical resources, issues of an architectural nature or related to configuration of the virtual cluster, largely oblivious to the end-user, may translate in a lack of data locality that results in a poor MapReduce performance with respect to the resources used.

In order to obtain resource-effective, portable, Cloud-based software for Bioinformatics pipelines, the issues mentioned earlier must be carefully studies and accounted for, in particular to have an impact for Personalized Medicine. As a matter of fact, the need is so pressing and apparently the expected demand so high that Edico genome and Amazon have started a collaboration that makes available Bioinformatics pipelines that are highly engineered to take advantage of FPGA programmability and the Cloud. The objective is to take the already highly performing shared-memory multi-processor based solutions offered by Edico genome in order to make them "real time". Fortunately, this is only the "high end" of the spectrum where a transition from the old HPC paradigm to the new of Cloud Computing one has gone beyond Proof of Concept.

Walking pathways in cancer

Alexander Kel, GeneXplain, Germany

Huge regions of non-coding DNA in genomes are the source of high adaptability of molecular genomic systems of multicellular eukaryotic organisms (such as human) to varying external conditions. We think, that such high adaptability is provided, first of all, through structural plasticity of gene regulatory networks. Binding of highly variable complexes of transcription factors to their highly fluctuating opened chromatin regions in genome (epigenomic variations) underlies the fundamental basis for such structural plasticity of gene regulatory networks. In this talk we will discuss the evolutional advantages of such structural plasticity of gene regulatory networks as well as the high price such systems have to pay for this plasticity – terrible diseases such as cancer. We think, that often non-reversible structural changes of the regulatory networks due to an epigenomic "evolution" of genome regulatory regions cause transformations in the system switching the normal state to a

disease state. We call such structural network changes as “walking pathways”. The analysis of this phenomenon helps us to understand the mechanisms of molecular switches (e.g. between programs of cell death and programs of cell survival) and to identify prospective drug targets to treat cancer. Such structural plasticity of regulatory networks observed in genomes of higher eukaryotes, in our view, is the result of an evolutional “aramorphose” towards emergence of completely new mechanism of evolution of multi-cellular organisms.

Empirical information about the interaction of transcription factors and the regulated target genes, obtained by either conventional or high-throughput methods, has been collected in the TRANSFAC database since 28 years, and statistical models inferred from this information have been included as positional weight matrices (PWMs) and made available for the prediction of regulatory sites as well. New extension includes syntax (relevant combinations) and semantics (regulated processes) of regulatory sites. Extended annotation of gene-disease associations is available in the Human Proteome Survey Database (HumanPSD), connected with signaling pathways that control the activity of TFs (TRANSPATH database). All this carefully curated information can be used in full power to analyze disease related multi-omics data using recently created geneXplain platform, which helps to decipher the molecular mechanisms of disease often on very early stages of its progression.

Finally, in this talk we will present an “upstream analysis” strategy for causal analysis of such multiple “-omics” data. It analyzes promoters using the TRANSFAC database, combines it with an analysis of the upstream signal transduction pathways and identifies master regulators as potential drug targets for a pathological process. We applied this approach to a complex multi-omics data set that contains transcriptomics, proteomics and epigenomics data. We identified the following potential drug targets against induced resistance of cancer cells towards chemotherapy by methotrexate (MTX): TGFalpha, IGFBP7, alpha9-integrin, and the following chemical compounds: zardaverine and divalproex as well as human metabolites such as nicotinamide N-oxide.

Open Data is Essential for Personalized Medicine

Francis Ouellette, Génome Québec, Canada

Abstract not available.

Big data and cognitive computing in healthcare: weathering the perfect storm

Matthias Reumann, IBM Research - Zurich, Switzerland

Big data in healthcare is experiencing the perfect storm: the volume is increasing exponentially with accelerating speed, the variety of data ranges from multi-omics information to lifestyle measures with the help of mobile devices backed by cloud infrastructures. State of the art analytical methods are generally limited by computational approaches. Furthermore, the convergence of data analytics, sophisticated modelling approaches and cognitive computing gives promise to solve the big data challenges in healthcare and lifescience. Data analytics especially in today’s omics era yield results of large volumes given computational challenges are overcome. Sieving through the results requires expert and translational knowledge. Cognitive computing can play a significant role in making transparent results. Cognitive computing tools can be used to create hypotheses to guide experimental studies but also as prior knowledge that drives data analytics. The increasing amount of data requires a larger amount of computation that can at some point only be tackled using supercomputers. In biophysical modelling we have already shown how the computational challenge can be overcome using high performance computing systems. The sophistication of computer modelling of biophysical processes has made the transition from basic research to translational science and medicine. It is feasible today that data in healthcare will be augmented by simulation of biophysical models tailored to each patient. Cognitive computing is a promising path to make the analytical results transparent. The IBM Watson™ technology allows analysis results to be represented within a global context of accumulated knowledge of published literature. To view data and analysis in that global context will not only enable verification of results, but also helps accelerate discovery and identification of, for example, new targets in drug discovery. The combination of data-driven and knowledge-based analytics in a cognitive computing environment becomes a powerful way to create hypothesis and to limit the search space so that it can efficiently be tested using traditional laboratory methods. The IBM Watson™ technology allows one to find “the needle in the haystack” of today’s big data challenge. Hence, the power of big data can only be unleashed by embracing new approaches in data-driven analysis within a cognitive computing environment. This creates a holistic view that places big data analytics into the context of the accumulated knowledge of the scientific community.

ELIXIR-IIB – the Italian Infrastructure for Bioinformatics: a growing support to national and international research in life sciences

Luana Licata, ELIXIR-IIB, Italy

ELIXIR (https://www.elixir-europe.org/) unites Europe's leading life science organisations in managing and safeguarding the increasing volume of data being generated by publicly funded research, and coordinates, integrates and sustains bioinformatics resources across its member states (ELIXIR nodes).

ELIXIR-IIB (http://elixir-italy.org/), the ELIXIR Italian node and infrastructure for bioinformatics, is coordinated by the National Research Council and currently includes 17 centres of excellence among which are research institutes, universities and technological institutions.

The infrastructure supports the exchange and development of skills, and the integration of publicly available and internationally recognised Italian bioinformatics resources within the European infrastructure.

ELIXIR-IIB, which aims to bring together all the Italian researchers working in the field of bioinformatics, is striving to assume a pivotal role for the national and international life science communities. This is reflected by the growing number of bioinformatics services, initiatives and projects supported or participated by ELIXIR-IIB, including H2020 grants, and the development of the ELIXIR-IIB Training Programme (https://elixir-iib-training.github.io/website/), which is building a thriving community who strongly believes that quality training in bioinformatics is essential to achieve excellence in life science research.

Topological Field Theory of Data: a new venue for Biomedical Big Data Analysis

Emanuela Merelli, University of Camerino, Italy

In her talk, she will challenge the current thinking in IT for the Big Data question, proposing a program aiming to construct an innovative methodology to perform data analytics that goes beyond the usual paradigms of data mining rooted in the notions of Complex Networks and Machine Learning. The method presented – at least as scheme – that returns an automaton as a recognizer of the data language, is, to all effects, a Field Theory of Data.

She will discuss, by using biomedical case studies, how to build, directly out of probing the data space, a theoretical framework enabling to extract the manifold hidden relations (patterns) that exist among data as correlations depending on the semantics generated by the mining context.

The program, that is grounded in the recent innovative ways of integrating data into a topological setting, proposes the realization of a Topological field theory of data, transferring and generalizing to the space of data notions inspired by physical (topological) field theories and harnesses the theory of formal languages to define the potential semantics necessary to understand the emerging patterns.

ORAL COMMUNICATIONS

- **Melanoma expression analysis with Big Data technologies.**
  Alicia Fernandez-Rovira, Rocio Lavado, Miguel Ángel Berciano Guerrero, Ismael Navas-Delgado and Jose F Aldana Montes.

- **Personalised models for human – gut microbiota interaction.**
  David Gilbert, Monika Heiner and Leila Ghanbar.

- **Greedy motif-based approach to parsing large and diverge coiled-coil proteins into domains.**
  Hamed Khakzad, Johan Malmström and Lars Malmström.

- **Epistasis analysis reveals associations between gene variants and bipolar disorder.**
  Carlo Maj, Elena Milanesi, Massimo Gennarelli, Luciano Milanesi and Ivan Merelli.

- **PGxO: A very lite ontology to reconcile pharmacogenomic knowledge units.**
  Pierre Monnin, Clément Jonquet, Joël Legrand, Amedeo Napoli and Adrien Coulet.

- **Parameters tuning boosts hyperSMURF predictions of rare deleterious non-coding genetic variants.**
  Alessandro Petrini, Max Schubach, Matteo Re, Marco Frasca, Marco Mesiti, Giuliano Grossi, Tiziana Castrignanò, Peter Robinson and Giorgio Valentini.

- **GenHap: A Novel Computational Method Based on Genetic Algorithms for Haplotype Assembly.**
  Andrea Tangherloni, Simone Spolaor, Leonardo Rundo, Marco S. Nobile, Ivan Merelli, Paolo Cazzaniga, Daniela Besozzi, Giancarlo Mauri and Pietro Liò.

- **MTopGO: a tool for module identification in PPI Networks.**
  Danila Vella, Simone Marini, Francesca Vitali and Riccardo Bellazzi.

SHORT ORAL COMMUNICATIONS

- **GenotypeAnalytics: an RESTful Platform to mine multiple associations between SNPs and drug response in case-control studies.**
  Giuseppe Agapito, Pietro Hiram Guzzi and Mario Cannataro.

- **Interoperability of RehabRobo-Onto.**
  Zeynep Dogmus, Volkan Patoglu and Esra Erdem.

- **From trash to treasure: detecting unexpected contamination in unmapped NGS data.**
  Ilaria Granata, Mara Sangiovanni, Mario Rosario Guarracino and Amarinder Singh Thind.

- **Multi-resolution network modelling of T-cells for precisions medicine of multiple sclerosis.**
  Mika Gustafsson.

- **Identification of Mutational Signatures Active in Individual Tumors.**
  Sandra Krüger and Rosario M. Piro.

- **A Dynamic Bayesian Network model for simulation of disease progression in Amyotrophic Lateral Sclerosis patients.**
  Alessandro Zandonà, Matilde Francescon, Maya Bronfeld, Andrea Calvo, Adriano Chiò and Barbara Di Camillo.

POSTERS

P1 **Visualizing Mutation Occurrence using Big Data**
Silvana Albert

P2 **A Gene Set Enrichment Analysis of multiomic celiac disease data**
Eugenio Del Prete, Angelo Facchiano and Pietro Liò

P3 **Network diffusion on multiple-layers: current approaches and integrative analysis of Rheumatoid Arthritis data**
Noemi Di Nanni, Matteo Gnocchi, Marco Moscatelli, Luciano Milanesi and Ettore Mosca

P4 **Evolutionary relationships of Microbial Transglutaminases**
Deborah Giordano and Angelo Facchiano

P5 **Exploiting transcriptomic data in genome scale metabolic networks: new insights into obesity**
Ilaria Granata, Mara Sangiovanni, Enrico Troiano and Mario Guarracino

P6 **Implementing a multilayer framework for pathway data integration, analysis and visualization**
Zaynab Hammoud and Frank Kramer

P7 **WITHDRAWN**

P8 **A Quality Management System for scientific research activities and its related management software.**

Luca Caruana, Alessandro Pensato, Loredana Riccobono, Giovanna Lucia Liguori, Antonella Lanati, Annamaria Kisslinger, Marta Di Carlo and Antonella Bongiovanni

P9 **Querying and analyzing biological data with BioGraph**
Antonio Messina, Antonino Fiannaca, Laura La Paglia, Massimo La Rosa and Alfonso Urso.

P10 **WITHDRAWN**

P11 **ISMARA: Completely automated inference of gene regulatory networks from high-throughput data**
Mikhail Pachkov, Piotr Balwierz, Phil Arnold, Andreas Gruber, Mihaela Zavolan and Erik van Nimwegen.

P12 **An Algebraic Representation for Tree Alignment of RNA Pseudoknotted Structures**
Michela Quadrini, Luca Tesei and Emanuela Merelli.

P13 **Procedure to integrate i2b2 and REDCap: a case study at ICSM**
Valentina Tibollo, Mauro Bucalo, Danila Vella, Morena Stuppia, Nicola Barbarini and Riccardo Bellazzi.

P14 **Stability Analysis of MTopGO for Module Identification in PPI Networks**
Danila Vella, Allan Tucker and Riccardo Bellazzi.

P15 **A novel shiny platform for the geo-spatial analysis of large amount of patient data**
Mario Alesandro Russo, Francesco Guarino, Monica Franzese, Dario Righelli, Giovanni Improta, Claudia Angelini and Maria Triassi

P16 **Detecting significant features in modeling microRNA-target interactions**
Claudia Coronnello, Giovanni Perconti, Patrizia Rubino, Flavia Contino, Serena Bivona, Salvatore Feo and Agata Giallongo

P17 **Application of network diffusion approaches to Genomics of Drug Sensitivity in Cancer (GDSC) data: a perspective on integrating multiple layers of information extracted from cell lines and drugs**
Vigneshwari Subramanian, Bence Szalai, Luis Tobalina and Julio Saez-Rodriguez

P18 **Distributed stream processing for genomics pipelines**
Francesco Versaci, Luca Pireddu and Gianluigi Zanetti

P19 **Automatic simulation of RNA editing in plants for the identification of novel putative Open Reading Frames**
Fabio Fassetti, Claudia Giallombardo, Ofelia Leone, Luigi Palopoli, Simona E. Rombo, Pierluigi Ruffolo and Adolfo Saiardi

Workshop Chairs

P. Romano

A. M. Urso

A. Valencia

Tutorial Chairs

L. Milanesi

A. M. Urso

A. Via

Under the Patronage of

BITS
Italian Society
of Bioinformatics

Città di Palermo

ELIXIR-IIB
Italian Institute of Bioinformatics
ELIXIR Italian Node

EMBnet
The Global
Bioinformatics Network

SFBI
Societé Française
de Bioinformatique

PTBI
Polish Society
of Bioinformatics

With the contribution from

INdAM – GNCS Project 2017
Efficient Algorithms and Techniques
for Biological Big Data

InterOmics
Flagship Project
National Research Council

ELIXIR-IIB
Italian Institute of Bioinformatics
ELIXIR Italian Node

ICAR / CNR
HPC & Networking Institute
National Research Council

IRCCS
Ospedale Policlinico San Martino,
Genova

OSPEDALE POLICLINICO SAN MARTINO