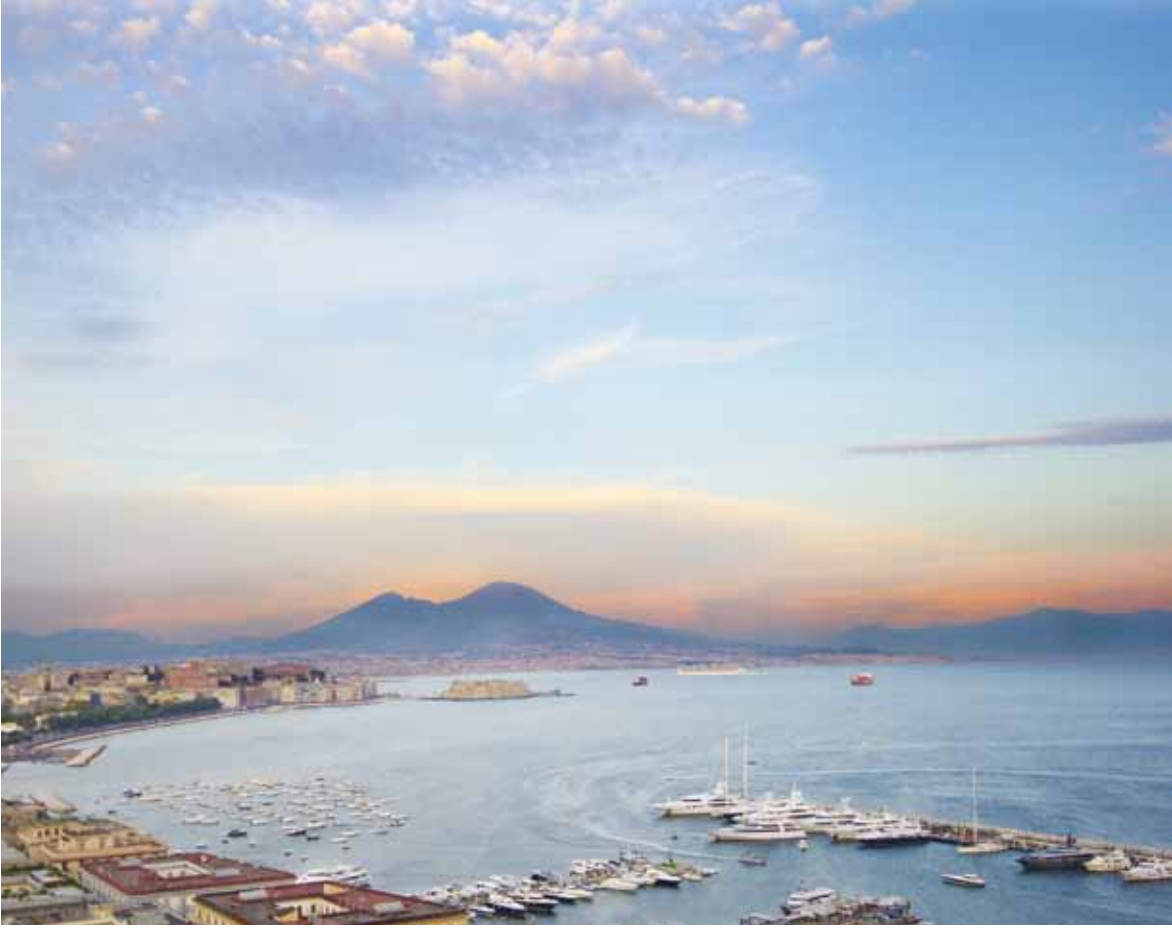

Bringing Maths to Life

Naples (Italy) October 27-29, 2014

ABSTRACT

Workshop Bringing Maths to Life

October 27-29, 2014

Aula Magna Partenope, Centro Congressi "Federico II"
Naples, Italy

SCIENTIFIC TOPICS

- **Genetic variability and differential expression: sequence data analysis**
- **Deciphering complex relationships: networks and interactions**
- **Zoom inside the cell: microscopy image processing**

INVITED SESSIONS

- **Molecular Dynamics and Modelling of Protein Structure and Function via High Performance Computing Simulations**
- **Artificial neurons and realistic simulation of neuronal functions**
- **Statistical challenges in omics research within Life Sciences**

Organizing Institutions

- University of Naples “Federico II”
- National Research Council “CNR”
- University of Rome “Sapienza”
- Vu University Amsterdam
- ISBE Infrastructure for Systems Biology Europe

Organizing Committee

Vincenza Colonna
IGB - National Research Council (CNR) – Italy
vincenza.colonna@igb.cnr.it
Maria Brigida Ferraro
DSS - Sapienza University of Rome – Italy
mariabrigida.ferraro@uniroma1.it
Mario Rosario Guarracino
ICAR - National Research Council (CNR) – Italy
mario.guarracino@cnr.it
Martijn Moné
VU University - Netherlands & Infrastructure for Systems Biology Europe (EU)
m.j.mone@vu.nl
Alessandra Rogato
IBBR - National Research Council (CNR) – Italy
alessandra.rogato@ibbr.cnr.it
Valeria Zazzu
IGB - National Research Council (CNR) – Italy
valeria.zazzu@igb.cnr.it

Scientific Committee

Scientific Advisor: Gerardo Toraldo
Department of Mathematics and Applications "*Renato Caccioppoli*", University of Naples Federico II - Italy
Guido Barbujani
University of Ferrara - Italy
Maria Vittoria Cubellis
University of Naples Federico II - Italy
Angela Blanco Fernandez
University of Oviedo – Spain
Paola Festa
University of Naples Federico II – Italy
Erik Garrison
Boston College – USA
Maria Angeles Gil Alvarez
University of Oviedo – Spain
Luca Tardella
Sapienza University of Rome – Italy
Angela Tino
CIB - National Research Council (CNR) - Italy

SPONSORS

We acknowledge the support of the BMTL Workshop by the following institutions/projects:

- University of Naples “Federico II”
- Department of Mathematics and Applications “Renato Caccioppoli”, University of Naples Federico II – Italy
- Genomics, Transcriptomics and Proteomics Laboratory (LabGTP), ICAR, CNR-Italy
- Institute of Genetics and Biophysics (IGB) - National Research Council (CNR) – Italy
- Department of Bio-Agriculture Sciences - National Research Council (CNR) – Italy
- Department of Biomedical Sciences - National Research Council (CNR) – Italy
- INDAM - Institute of Higher Mathematics “F. Severi” - Italy
- Tecnologica srl
- MIMOmics
- InterOmics

GIVEN THAT ALL GENETICISTS WANT TO KNOW ABOUT MATHS AND *VICE VERSA,* ORGANISERS OF **BRINGING MATHS TO LIFE** OUTLINE A NEW WORKSHOP DESIGNED TO ESTABLISH CONNECTIONS BETWEEN THE TWO FIELDS

It all adds up

Often, in many instances, life scientists are not always fully aware of the power that mathematical models have in both answering biological research questions and in making predictions. They have a clear view of the problem; they know the questions; they have identified ways to answer; so far so good. They then produce the data that is to be analysed. They use novel high throughput technologies that give rise to an unprecedented quantity of data, but the data is 'noisey', and the answer to every question can be well hidden under terabytes of incomprehensible text files.

It is here that the mathematicians can help: they know 'how' to do things; they love the huge, ugly text files; they foresee hundreds of statistics that could be calculated; they want to try all of them because there is always uncertainty; they see paths, trends, connections, and correlations. But soon the time will come when they need to identify the beautiful biological mechanisms that are hidden, and here, in turn, they too get stuck, lost among protein sticks, bubbles, helices, and sheets.

Workshop

Because of this, six enthusiastic researchers from different disciplines, institutions and countries have now come together and begun a discussion to improve the interface between life sciences and mathematics. From this a workshop to establish connections between the two fields, and indeed, to pinpoint needs, broaden views, exchange ideas and share knowledge, has been developed.

The workshop, entitled 'Bringing Maths to Life' (www.bmtl.it), will be held in Naples, Italy, on 27-29 October 2014, with a list of confirmed speakers coming from leading European universities. Participants from the international scientific community, especially computational biology, mathematics, and the life science disciplines are also expected.

Analysing and interpreting rich biological datasets requires more expertise than just knowledge of the biological system at hand. Extracting reliable insights from complex bodies of data calls for suitable mathematical solutions.

This workshop will therefore allow biologists and mathematicians to join forces in order to address key areas in biology that face demanding mathematical challenges. Discussing existing cases to identify gaps or sharing existing solutions should help these disciplines to successfully link up. Finding the best mathematical resolution to interpret data from a biological perspective, or – inversely – understanding the biological issue and its real-life constraints from a mathematical viewpoint, require both communities to engage closely.

Sessions

The workshop will feature three main sessions. The first will be **'Genetic variability and differential expression: sequence data analysis'**. Recent revolution in DNA sequencing technology has made the sequencing of an increasing number of genomes both feasible and cost effective. Changes in data quantity and format (large numbers of short reads or pairs of short reads *versus* relatively long reads produced by traditional Sanger sequencing) imply changes of sequence data management, storage, and visualisation, and provide a challenge for bioinformatics.

The second concerns **'Deciphering complex relationships: networks and interactions'.** Biological systems comprise of thousands of different types of components. These parts form huge networks that comprise numerous non-linearly interacting dimensions, from which, in turn, biological functions emerge. The networks are far too complex to be understood by the unassisted human mind. To analyse these complex biological systems and obtain relevant answers, biology requires quantitative models that draw from modern computer science and mathematics.

'Zoom inside the cell: microscopy images processing' will form the topic of the third session. Biological visualisation provides the means through which to place genomic and proteomic information in a cellular or tissue context. While existing software enables particular assays for particular cell types, high throughput image analysis has, to this point, been impractical unless an image analysis expert develops a customised solution, or unless commercial packages are used with their built-in algorithms for a limited set of cellular features and for a limited set of cell types. There exists a clear need for powerful, flexible tools for high throughput cell image analysis. Hopefully, computer vision researchers will contribute new algorithms to the project so that their theoretical work can be applied to practical biological problems.

Invited sessions

In addition, there will also be three invited sessions:

'Molecular Dynamics and Modelling of Protein Structure and Function via High Performance Computing Simulations' (organized by Alessandro Grottesi from CINECA, Italy). Molecular dynamics simulations are computational tools aimed at studying protein structure and dynamics as well as protein-protein interactions at the atomic level. The high performance computing of current computer architectures, as well as the developing of valid force fields for the mathematical modelling of biochemical interactions, have provided new tools to help biologists studying and testing hypotheses to understand biochemical phenomena in a new perspective. This session will thus highlight the advantages and limitations of this powerful computational technique.

In the second session, **'Statistical challenges in omics research within Life Sciences'** (organized by J.J Houwing-Duistermaat from Leiden University Medical Center, The Netherlands and Luciano Milanesi from Institute of Biomedical Technologies, CNR, Itlay), we will ad-

dress several statistical issues in omics datasets from preprocessing up to building statistical models for joint interpretation of the datasets. Examples of omics datasets are genomics, metabolomics and proteomics, and microbiome data. These datasets contain information about different aspects of the same biological processes. Therefore in many studies multiple omics datasets are nowadays available and integrated analyses of these omics datasets is the ultimate goal to understand biological mechanisms underlying traits. However integration of these datasets is not straightforward since they vary in measurement error distributions, scale, sparseness and size. In this session we will therefore address challenges in single omics datasets analysis as well as combined analysis of multiple omics datasets.

The third invited session is dedicated to **'Artificial neurons and realistic simulation of neuronal functions'** (organised by Angela Tino from Institute of Cybernetics, CNR, Italy). Modern neuroscience research has generated vast volumes of experimental data, and large scale initiatives launched in recent years will gather much more. Nonetheless, much of the knowledge needed to build multilevel atlases and unifying models of the brain is still missing. Brains are a large network composed of many neurons with their synaptic connections, each expressing different proteins on the cell membrane and each with its own complex internal structure. Despite huge advances, there is no technology that allows us to characterise more than a tiny part of this complexity. The session will shed light on novel solutions from neural-inspired artificial models and software, realistic neuronal function simulation, and functional and molecular neurobiology. The session aims to gather scientists from diverse disciplines to foster integrated approaches to unravel complex brain functions.

For more info visit www.bmtl.it

KEYNOTES

Evidence of non-Poissonian Dynamics in Cancer Mutations

Jose M Muino[1], Ercan E Kuruoglu[2,1], Peter F Arndt[1]

1. Max Planck Institute for Molecular Genetics, Berlin, Germany
2. ISTI-CNR, Pisa, Italy

It is a common assumption in various work on modelling mutation dynamics that mutations follow a Poisson dynamics; that is, in a given portion of genome the number of mutations follow a Poisson law. Equivalently, the distance between to mutations follows an exponential distribution. This can actually be verified when Human and Chimpanzee genomes are compared. It is of interest to see if this law generalizes also to somatic mutations which cause cancer. A recent survey published a catalogue of somatic mutations in cancer genome analysing 4,938,362 mutations from 7,042 cancers of 30 different cancer types. We have analysed this data to find the interoccurence time (space) distributions for different types of cancer. It has been found that specific cancer types show a power-law in interoccurrence distances, instead of the expected exponential distribution dictated with the Poisson assumption. Cancer genomes exhibiting power-law interoccurrence distances were enriched in cancer types where the main mutational process is described to be the activity of the APOBEC protein family, which produces a particular pattern of mutations called Kataegis. Therefore, the observation of a power-law in interoccurence distances could be used to identify cancer genomes with Kataegis. It is our objective to develop parametric models to differentiate between cancer and non-cancer mutations and between different cancer types.

References

1. Alexandrov et al, "Signatures of mutational processes in human cancer," *Nature,* 500, 415–421, (22 August 2013), doi:10.1038/nature12477.