Bioinformatics and its Applications in Health, Biodiversity and the Environment

Hands-on workshops on 5/9: Proteomics and RNA · Seq/ChIP-Seq

Tutorials on 6/9: Metagenomics · Phylogenomics
CONTENTS

Pre-Conference Workshops

Pre-Conference Tutorials

HB10 Scientific Agenda

Invited Speakers

Organizers

Oral Presentations List

Poster Presentation List

Oral Presentations Abstracts

Poster Presentations Abstracts

Author Index

Participant List

Useful Info

Sponsors List

HB10 Poster
# HELLENIC BIOINFORMATICS 10

## Pre-Conference Workshops

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.30-09.00</td>
<td><strong>Registration</strong></td>
</tr>
<tr>
<td>09.00-09.15</td>
<td><strong>Workshop 1: Proteomics Workshop – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Chair: Michalis Aivaliotis</td>
</tr>
<tr>
<td>09.15-10.30</td>
<td><strong>Workshop 1: Proteomics Workshop – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Michalis Aivaliotis, Nikos Kountourakis, Konstantina Psatha, Chara Seitanidou: Bioinformatics I – MS-Raw Data Processing</td>
</tr>
<tr>
<td></td>
<td>- Proteome Identification and Quantitation using Commercial available and Free Bioinformatic Platform Functional Proteome Annotation</td>
</tr>
<tr>
<td></td>
<td>- Quality Control and Filtering of Proteomics Data</td>
</tr>
<tr>
<td>10.30-13.00</td>
<td><strong>Workshop 1: Proteomics Workshop – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Michalis Aivaliotis, Nikos Kountourakis, Konstantina Psatha, Chara Seitanidou: Bioinformatics II – Meta-data analysis and visualization</td>
</tr>
<tr>
<td></td>
<td>- Comparative Proteomics Analysis</td>
</tr>
<tr>
<td></td>
<td>- Functional Proteome Annotation</td>
</tr>
<tr>
<td></td>
<td>- GO Terms and Pathways Enrichment Analysis</td>
</tr>
<tr>
<td></td>
<td>- Protein-protein Interaction Networks</td>
</tr>
<tr>
<td></td>
<td>- Visualization of Proteomics Data</td>
</tr>
<tr>
<td>14.00-14.15</td>
<td><strong>Workshop 2: RNA-Seq and ChIP-Seq Data Analysis – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Chair: Pantelis Topalis</td>
</tr>
<tr>
<td>14.15-16.00</td>
<td><strong>Workshop 2: RNA-Seq and ChIP-Seq Data Analysis – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Pantelis Topalis, Emmanuel Dialynas, George Papagiannakis: RNA-Seq Analysis</td>
</tr>
<tr>
<td></td>
<td>- De novo transcriptome assembly</td>
</tr>
<tr>
<td></td>
<td>- Differential Gene Expression</td>
</tr>
<tr>
<td></td>
<td>- Everything you Always Wanted to Know about Batch Effects</td>
</tr>
<tr>
<td>14.15-16.00</td>
<td><strong>Workshop 2: RNA-Seq and ChIP-Seq Data Analysis – Hands-on</strong></td>
</tr>
<tr>
<td></td>
<td>Pantelis Topalis, Emmanuel Dialynas: ChIP-Seq Data Analysis</td>
</tr>
<tr>
<td></td>
<td>- Peak Finding (Transcription Factors, Histone Modifications, Super Enhancers)</td>
</tr>
<tr>
<td></td>
<td>- Irreproducible Discovery Rate (IDR) Pipeline</td>
</tr>
<tr>
<td></td>
<td>- Peak Annotation</td>
</tr>
</tbody>
</table>

**Important Notes:**
- A digital textbook and references will be provided to all participants in advance
- All the participants need to have their own laptops. No computers will be provided by the organizers

* Coffee will be available inside the room during the whole duration of the workshops
**HELLENIC BIOINFORMATICS 10**  
**Pre-Conference Tutorials & Opening Session**

### Pre-Conference Tutorials – Wednesday 6/9 – “S. Orfanoudakis” Seminar Room

<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>08.30-09.00</td>
<td><strong>Registration</strong></td>
</tr>
</tbody>
</table>
| 09.00-10.30 | **Tutorial 1: Metagenomics**  
Chair: Folker Meyer  
- Alex Mitchel: *Introduction to EBI Metagenomics*  
- Folker Meyer: *MG-RAST version 4.0*  
  - This tutorial will demonstrate the features of MG-RAST v4.0 for data management and process as well as analysis. Depending on the time available the tutorial can include cmd-line as well as web interface components. The tutorial will cover metagenome, metatranscriptome and amplicon metagenome analysis. |
| 13.30-15.00 | **Tutorial 2: Phylogenomics**  
Chair: Alexandros Stamatakis  
- Alexandros Stamatakis: *Phylogenomic Inference Methods: Inferring Species Trees from Whole-genome Alignments*  
- Bastien Bousseau: *Reconciliation Methods: Interpreting Gene Trees Given Species Trees* |

* Coffee will be available inside the room during the whole duration of the workshops

### Opening Session – Wednesday 6 September  
Amphitheater “G. Lianis”  
All HB-10 and DEFORM Participants

<table>
<thead>
<tr>
<th>Time</th>
<th>Speaker</th>
</tr>
</thead>
</table>
| 18.00-18.30 | Ioannis Ioannidis, Stanford University, USA  
*Predating Data: Integrity of Methods and Statistics*  
| 18.30-19.00 | Gemma Galdon Clavell, Eticas Research and Consulting, Spain  
*Misconduct and Data: Privacy, Accountability and Responsibility in (Big) Data in Social Science*  
| 19.00-19.30 | Nektarios Tavernarakis, Univ. of Crete and IMBB-FORTH, Greece  
*The European Research Council Position on Data Integrity and Scientific Misconduct*  
| 19.30-20.00 | Discussion                                                              |
| 20.00     | 🍷olkata Welcome Reception [Sponsored by Project DEFORM]                |

**BACK TO CONTENTS**
## Scientific Agenda

### Wednesday 6/9 – FORTH Amphitheater “G: Lianis”

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>17:00-18:00</td>
<td><strong>Registration</strong></td>
</tr>
<tr>
<td><strong>Open Session – Data Integrity - Co-organized by project DEFORM (H2020)</strong></td>
<td></td>
</tr>
<tr>
<td>Chair: Vasiliki Petousi</td>
<td></td>
</tr>
<tr>
<td>18:00-18:30</td>
<td><strong>Ioannis Ioannidis</strong>, Stanford University, USA</td>
</tr>
<tr>
<td></td>
<td><em>Predating Data: Integrity of Methods and Statistics</em></td>
</tr>
<tr>
<td>18:30-19:00</td>
<td><strong>Gemma Galdon Clavell</strong>, Eticas Research and Consulting, Spain</td>
</tr>
<tr>
<td></td>
<td><em>Misconduct and Data: Privacy, Accountability and Responsibility in (Big) Data in Social Science</em></td>
</tr>
<tr>
<td>19:00-19:30</td>
<td><strong>Nektarios Tavernarakis</strong>, Univ. of Crete and IMBB-FORTH, Greece</td>
</tr>
<tr>
<td></td>
<td><em>The European Research Council Position on Data Integrity and Scientific Misconduct</em></td>
</tr>
<tr>
<td>19:30-20:00</td>
<td>Discussion</td>
</tr>
</tbody>
</table>

**Welcome Reception to all HB10 and project DEFORM participants** [Sponsored by Project DEFORM]

### Thursday 7/9 – FORTH Amphitheater “G: Lianis”

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00-12:00</td>
<td><strong>Registration</strong></td>
</tr>
<tr>
<td>09:00-09:40</td>
<td><strong>HB-10 Welcome notes and Conference outline</strong></td>
</tr>
<tr>
<td></td>
<td>Chairs: Ioannis Iliopoulos and Ilias Kappas</td>
</tr>
<tr>
<td></td>
<td>o Ilias Kappas, HB-President, AUTh</td>
</tr>
<tr>
<td></td>
<td>o Nektarios Tavernarakis, Chairman of BoD, FORTH</td>
</tr>
<tr>
<td></td>
<td>o Representative from Region of Crete</td>
</tr>
<tr>
<td></td>
<td>o Giorgos Hourdakis, Ministry of Education, Research and Religion Affairs</td>
</tr>
<tr>
<td></td>
<td>o Michalis Aivaliotis, HB-10 Co-Chair, IMBB</td>
</tr>
<tr>
<td></td>
<td>o Ioannis Iliopoulos, HB-10 Chair, University of Crete</td>
</tr>
<tr>
<td>09:40-10:00</td>
<td><strong>Anna Tramontano (In Memoriam)</strong></td>
</tr>
<tr>
<td></td>
<td>o Alfonso Valencia</td>
</tr>
<tr>
<td></td>
<td>o Christos Ouzounis</td>
</tr>
<tr>
<td>10:00-10:45</td>
<td><strong>Keynote Talk</strong> – Chair: Nektarios Tavernarakis</td>
</tr>
<tr>
<td></td>
<td><strong>John IOANNIDIS</strong>, Stanford University, USA</td>
</tr>
<tr>
<td></td>
<td><em>Credibility and Utility of Big Data</em></td>
</tr>
</tbody>
</table>

**Coffee break**

**Theme: HEALTH**

**Session 1: Medical Genomics**
<table>
<thead>
<tr>
<th>Time</th>
<th>Speaker</th>
<th>Institution</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:15-11:45</td>
<td><strong>Eleftheria ZEGGINI</strong>, Sanger Center, UK</td>
<td></td>
<td>Progress in Understanding the Genomic Aetiology of Osteoarthritis</td>
</tr>
<tr>
<td>11:45-12:15</td>
<td><strong>Evangelia PETSalaki</strong>, EBI, UK</td>
<td></td>
<td>Integrative Study of the Role of RhoGAP/GEFs in the Rho Protein Signaling Network</td>
</tr>
<tr>
<td>12:15-12:30</td>
<td><strong>Despoina KALFAKAkOU</strong>, Demokritos, Greece</td>
<td></td>
<td>CanVaS: A Database of Greek Cancer Patient Genetic Variation</td>
</tr>
<tr>
<td>12:30-12:45</td>
<td><strong>Anastasis OULAS</strong>, Cyprus Institute of Neurology, Cyprus</td>
<td></td>
<td>Network Based Approaches for Classifying Variants of Unknown Significance</td>
</tr>
<tr>
<td>12:45-13:00</td>
<td><strong>Maria NIKOLOUDAKI</strong>, CSD, University of Crete, Greece</td>
<td></td>
<td>Reconstructing Ikaros Regulatory Network by Applying Causal Discovery Methods on a Compendium of Gene Expression Profiles from Acute Lymphoblastic Leukemia Patients</td>
</tr>
<tr>
<td>13:00-13:15</td>
<td><strong>Michalis AIVALIOTIS</strong>, IMBB-FORTH, Greece</td>
<td></td>
<td>Understanding Lymphomas’ Pathobiology Drivers through an Integrative Network Analysis</td>
</tr>
<tr>
<td>13:15-13:30</td>
<td><strong>Eleni KATSANTONI</strong>, Biomedical Research Foundation, Greece</td>
<td></td>
<td>Erythropoietic Transcription Factors Networks in Beta-Thalassemia</td>
</tr>
<tr>
<td>14:30-15:00</td>
<td><strong>Nektarios TAVERNARAKIS</strong>, IMBB-FORTH and University of Crete, Greece</td>
<td></td>
<td>Mitochondrial Homeostasis in Neurodegeneration and Ageing</td>
</tr>
<tr>
<td>15:00-15:30</td>
<td><strong>Viota POIRAZI</strong>, IMBB-FORTH, Greece</td>
<td></td>
<td>Modelling Dendritic and Neuronal Contributions to Cognitive Functions</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td><strong>Pavlos PAVLIDIS</strong>, SHS, University of Crete, Greece</td>
<td></td>
<td>Modular Agent-Based Modeling Platform for Bridging Interdisciplinary Research in Ecology, Evolution, Behavioral Biology and Computational Neuroscience</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td><strong>Theodoros TAMIOIiAKIS</strong>, IMBB-FORTH, Greece</td>
<td></td>
<td>Elucidating the Role of Different CA3 Hippocampal Neurons in Pattern Completion Using Computational Modelling</td>
</tr>
<tr>
<td>16:00-17:30</td>
<td><strong>Ewa SZCZUREK</strong>, University of Warsaw, Poland</td>
<td></td>
<td>Statistical and Mathematical Models in Computational Oncology</td>
</tr>
<tr>
<td>17:30-18:00</td>
<td><strong>Kleanthi LAKIOTAki</strong>, CSD, University of Crete, Greece</td>
<td></td>
<td>Automated Machine Learning Methods to Predict Phenotype in Microarray and RNAseq Gene Expression Data</td>
</tr>
</tbody>
</table>
**Session 4: Roundtable**  
Chair: Ioannis Iliopoulos and Nektarios Tavernarakis

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>18:15-19:15</td>
<td><strong>Role of Greek Bioinformatics in Elixir</strong></td>
<td>Babis SAVAKIS - Greece, Alfonso VALENCIA - Spain, Nikos KYRPIDES - USA, Christos OUZOUNIS - Greece</td>
</tr>
<tr>
<td>19:15-20:00</td>
<td><strong>Keynote Talk</strong> – Chair: Ilias Kappas</td>
<td>Alfonso VALENCIA, Spanish National Cancer Research Centre, Spain</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Networks Based Approaches in Epigenomics, Evolution and Biomedicine</em></td>
</tr>
</tbody>
</table>

---

**Friday 8/9 – FORTH Amphitheater “G: Lianis”**

**Theme: ENVIRONMENT**

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30-10:15</td>
<td><strong>Keynote Talk</strong> – Chair: Christos Ouzounis</td>
<td>Peter KARP, SRI, USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>The Omics Dashboard for Interactive Exploration of High-Throughput Data</em></td>
</tr>
</tbody>
</table>

**Session 5: Human Microbiome**  
Chair: George Kotoulas and George Tsiamis

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:15-10:45</td>
<td><strong>GSC Metadata Standards in Action: MetaSUB and ISD</strong></td>
<td>Lynn SCHRIML, University of Maryland, USA</td>
</tr>
<tr>
<td>10:45-11:15</td>
<td><strong>Microbiota Development in the Gut of Infants, Children and Adolescents</strong></td>
<td>Maria Pilar FRANCINO, CSISP, Spain</td>
</tr>
<tr>
<td>11:15-11:30</td>
<td><strong>Bioenergetic Diversity of the Human Gut Microbiome</strong></td>
<td>Vassiliki Lila KOUMANDOU, Agricultural University of Athens, Greece</td>
</tr>
</tbody>
</table>

**Coffee Break**

**Session 6: Computational Metagenomics**  
Chair: Nikos Kyrpides and Christos Ouzounis

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00-12:45</td>
<td><strong>A Novel Pipeline for High Quality Amplicon Analysis Integrated Into MG-RAST</strong></td>
<td>Folker MEYER, ANL, USA</td>
</tr>
<tr>
<td>12:45-13:15</td>
<td><strong>Expanding the Scope and Reproducibility of Analysis in EBI Metagenomics</strong></td>
<td>Rob FINN, EBI, UK</td>
</tr>
<tr>
<td>13:15-13:30</td>
<td><strong>Updates and Applications of the Integrated Microbial NGS Platform (IMNGS): A Step towards Estimation of Global Diversity</strong></td>
<td>Ilias LAGKOUVARDOS, Technical University of Munich, Germany</td>
</tr>
</tbody>
</table>

**Lunch**

**Session 7: Environmental Microbiome**  
Chair: Michalis Aivaliotis and Despina Alexandraki

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30-15:00</td>
<td><strong>Soil Development and Rhizosphere Effect in a High Arctic Desert Chronosequence</strong></td>
<td>Daniele DAFFONCHIO, KAUST, Saudi Arabia</td>
</tr>
<tr>
<td>15:00-15:30</td>
<td><strong>Nitrogen Fixation by the Foliar Conifer Microbiome</strong></td>
<td>Carolin FRANK, UC Merced, USA</td>
</tr>
<tr>
<td>Time</td>
<td>Speaker</td>
<td>Title</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Anastasia GIOTI</td>
<td><em>The Effects of a Redox Cline and Historical Processes on Bacterial Community Structure in Aegean Marine Sediments</em></td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>George TSIAMIS</td>
<td><em>Gut and Gonadal Microflora of Insects: Challenges and Practical Applications</em></td>
</tr>
<tr>
<td>16:00-16:15</td>
<td>Georgios KOTOLAS</td>
<td><em>On Genomics Observatories; Towards Real Time Assessment of Ecosystem Structure and Functioning</em></td>
</tr>
<tr>
<td>16:15-16:30</td>
<td>Katerina VASILEIADOU</td>
<td><em>On Genomics Observatories; Towards Real Time Assessment of Ecosystem Structure and Functioning</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Coffee Break &amp; POSTER SESSION</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 8: HB Society Matters</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>20:00-22:00 Society's Annual Overview Meeting</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Saturday 9/9 – FORTH Amphitheater “G: Lianis”</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Theme: BIODIVERSITY</strong></td>
</tr>
<tr>
<td>09:30-10:15</td>
<td>Keynote Talk – Chair: Pavlos Pavlidis</td>
<td><em>Bayesian Matrix Factorization with Side Information and Application to Drug-Target Activity Prediction</em></td>
</tr>
<tr>
<td></td>
<td>Yves MOREAU</td>
<td>Leuven, Belgium</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 9: Phylogenomics</strong></td>
</tr>
<tr>
<td></td>
<td>Chair: Christoforos Nikolaou and Pantelis Topalis</td>
<td></td>
</tr>
<tr>
<td>10:15-10:45</td>
<td>Daniel LUNDIN</td>
<td><em>The Evolution of Ribonucleotide Reduction</em></td>
</tr>
<tr>
<td>10:45-11:00</td>
<td>Stella TAMANA</td>
<td><em>An Updated View of the Oligosaccharyltransferase Complex in Plasmodium</em></td>
</tr>
<tr>
<td>11:00-11:15</td>
<td>Konstantinos KYRITIS</td>
<td><em>Investigating the Specialized Roles of Ribosomal Proteins and Associated Ribosomopathies by Means of Literature, Protein Sequence and Phenotype Information Resources</em></td>
</tr>
<tr>
<td>11:15-11:30</td>
<td>Elias PRIMETIS</td>
<td><em>Evolutionary Models of Amino Acid Substitutions Based on their Neighborhood Tertiary Structure</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Coffee Break</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Session 10: Comparative Genomics</strong></td>
</tr>
<tr>
<td></td>
<td>Chair: Christos Arvanitidis and Vasilis Promponas</td>
<td></td>
</tr>
<tr>
<td>12:00-12:30</td>
<td>David USSERY</td>
<td><em>What is Life? Conserved Functional Domains in a Hundred Thousand Bacterial Genomes</em></td>
</tr>
<tr>
<td>12:30-12:45</td>
<td>Ioannis STERGIOPOULOS</td>
<td><em>Comparative Genomics of the Sigatoka Disease Complex on Banana Indicates a Direct Link between Pathogen Emergence and Nutritional Virulence</em></td>
</tr>
<tr>
<td>12:45-13:00</td>
<td>Christos NIKOLAOU</td>
<td><em>Invisible Cities: Segregated Domains in the Yeast Genome with Distinct Structural and Functional Attributes</em></td>
</tr>
<tr>
<td>Time</td>
<td>Speaker and Affiliation</td>
<td>Topic</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>13:00-13:15</td>
<td>Christos Ouzounis, CERTH, Greece</td>
<td>Ancestral Pathway Reconstruction</td>
</tr>
<tr>
<td>13:15-13:30</td>
<td>Nikos Kyrpides, JGI, USA</td>
<td>Genomic Encyclopedia of Bacteria and Archaea</td>
</tr>
<tr>
<td></td>
<td><strong>Lunch</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Session 11: Computational Technologies</strong></td>
<td>Chair: Ioannis Iliopoulos and Christos Karapiperis</td>
</tr>
<tr>
<td>14:30-15:00</td>
<td>Lars Jensen, University of Copenhagen, Denmark</td>
<td>One Tagger, Many Uses: Simple Text-Mining Strategies for Biomedicine</td>
</tr>
<tr>
<td>15:00-15:15</td>
<td>Andreas Antonakis, University of Crete, Greece</td>
<td>Latest Improvements in ProteoSign, a Web Application for Differential Proteomics Analysis</td>
</tr>
<tr>
<td>15:15-15:30</td>
<td>Lefteris Koumakis, ICS-FORTH, Greece</td>
<td>Calchas: A Semantic Search Engine for Bioinformatics Resources</td>
</tr>
<tr>
<td>15:30-15:45</td>
<td>Yannis Pantazis, University of Crete, Greece</td>
<td>Large-Scale Study on Latent Feature Construction for Gene Expression with Improved Predictive Power on Newly-Seen Datasets</td>
</tr>
<tr>
<td>15:45-16:00</td>
<td>Alexandros Kanterakis, ICS-FORTH, Greece</td>
<td>Arkalos: An open collaborative Workflow Management System for reproducible bioinformatics research</td>
</tr>
<tr>
<td></td>
<td><strong>Coffee Break</strong></td>
<td></td>
</tr>
<tr>
<td>17:00-17:45</td>
<td>Keynote Talk – Chair: Nikos Kyrpides</td>
<td>Periklis Papadopoulos, San Jose State University, USA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Space Technology and Space Biology</td>
</tr>
<tr>
<td>Speaker Name</td>
<td>Institution</td>
<td>Email</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Daniele Daffonchio</td>
<td>KAUST, Saudi Arabia</td>
<td><a href="mailto:Daniele.daffonchio@kaust.edu.sa">Daniele.daffonchio@kaust.edu.sa</a></td>
</tr>
<tr>
<td>Maria Pilar Francino</td>
<td>CSISP, Spain</td>
<td><a href="mailto:mpfrancino@gmail.com">mpfrancino@gmail.com</a></td>
</tr>
<tr>
<td>Rob Finn</td>
<td>EBI, UK</td>
<td><a href="mailto:rdf@ebi.ac.uk">rdf@ebi.ac.uk</a></td>
</tr>
<tr>
<td>John Ioannidis</td>
<td>Stanford University, USA</td>
<td><a href="mailto:jioannid@stanford.edu">jioannid@stanford.edu</a></td>
</tr>
<tr>
<td>Carolin Frank</td>
<td>UC Merced, USA</td>
<td><a href="mailto:cfrank3@ucmerced.edu">cfrank3@ucmerced.edu</a></td>
</tr>
<tr>
<td>Maria Pilar Francino</td>
<td>CSISP, Spain</td>
<td><a href="mailto:mpfrancino@gmail.com">mpfrancino@gmail.com</a></td>
</tr>
<tr>
<td>Peter Karp</td>
<td>SRI, USA</td>
<td><a href="mailto:pkarp@ai.sri.com">pkarp@ai.sri.com</a></td>
</tr>
<tr>
<td>Yiota Poirazi</td>
<td>IMBB, Greece</td>
<td><a href="mailto:poirazi@imbb.forth.gr">poirazi@imbb.forth.gr</a></td>
</tr>
<tr>
<td>Daniel Lundin</td>
<td>SciLifeLab, Sweden</td>
<td><a href="mailto:daniel.lundin@scilifelab.se">daniel.lundin@scilifelab.se</a></td>
</tr>
<tr>
<td>Hector-Garcia Martin</td>
<td>Joint Bioenergy Institute, USA</td>
<td><a href="mailto:hgmartin@lbl.gov">hgmartin@lbl.gov</a></td>
</tr>
<tr>
<td>Folker Meyer</td>
<td>Argonne National Laboratory, USA</td>
<td><a href="mailto:folker@anl.gov">folker@anl.gov</a></td>
</tr>
<tr>
<td>Yves Moreau</td>
<td>University of Leuven, Belgium</td>
<td><a href="mailto:moreau@esat.kuleuven.be">moreau@esat.kuleuven.be</a></td>
</tr>
<tr>
<td>Periklis Papadopoulos</td>
<td>San Jose State University, USA</td>
<td><a href="mailto:peripapas@gmail.com">peripapas@gmail.com</a></td>
</tr>
<tr>
<td>Evangelia Petsalaki</td>
<td>EBI, UK</td>
<td><a href="mailto:petsalaki@ebi.ac.uk">petsalaki@ebi.ac.uk</a></td>
</tr>
<tr>
<td>Lynn Schriml</td>
<td>University of Maryland, USA</td>
<td><a href="mailto:lschriml@som.umaryland.edu">lschriml@som.umaryland.edu</a></td>
</tr>
<tr>
<td>Ewa Szczurek</td>
<td>University of Warsaw, Poland</td>
<td><a href="mailto:Szczurek@mimuw.edu.pl">Szczurek@mimuw.edu.pl</a></td>
</tr>
<tr>
<td>Nektarios Tavernarakis</td>
<td>IMBB-FORTH &amp; U of Crete, Greece</td>
<td><a href="mailto:tavernarakis@imbb.forth.gr">tavernarakis@imbb.forth.gr</a></td>
</tr>
<tr>
<td>David Ussery</td>
<td>University of Arkansas, USA</td>
<td><a href="mailto:SBVoight@uams.edu">SBVoight@uams.edu</a></td>
</tr>
<tr>
<td>Alfonso Valencia</td>
<td>Spanish National Cancer Research Centre, Spain</td>
<td><a href="mailto:alfonso.valencia@bsc.es">alfonso.valencia@bsc.es</a></td>
</tr>
<tr>
<td>Eleftheria Zeggini</td>
<td>Sanger Center, UK</td>
<td><a href="mailto:Eleftheria@sanger.ac.uk">Eleftheria@sanger.ac.uk</a></td>
</tr>
</tbody>
</table>
# HB-10 ORGANIZERS

## PRESIDENT
- Ioannis Iliopoulos  
  Faculty of Medicine, University of Crete

## VICE PRESIDENT
- Michalis Aivaliotis  
  IMBB, FORTH

## SCIENTIFIC PROGRAM COMMITTEE
- Ioannis Iliopoulos  
  University of Crete
- Michalis Aivaliotis  
  IMBB, FORTH
- Vasilis Promponas  
  University of Cyprus
- Nikos Kyrpides  
  DOE Joint Genome Institute, USA
- Ilias Kappas  
  Aristotle University of Thessaloniki
- Christos Ouzounis  
  CPERI, CERTH

## ORGANIZING COMMITTEE
- Michalis Aivaliotis, Greece
- Sophia Ananiadou, UK
- Christos Arvanitidis, Greece
- Ioannis Iliopoulos, Greece
- Dimitris Kafetzopoulos, Greece
- Ilias Kappas, Greece
- Christos Karapiperis, Greece
- George Kotoulas, Greece
- Yiannis Kourmpetis, Switzerland
- Nikos Kyrpides, USA
- Tereza Manousaki, Greece
- Christophoros Nikolaou, Greece
- Christos Arvanitidis
- Anastasias Chalanaris
- Emmanouil Dialinas
- Ioannis Grigoriadis
- Nikolaos Grigoriadis
- John Kouklinos
- Anastasias Oulas
- Evaggelos Pafilis
- Eleftherios Panteris
- Nikolas Papanikolaou
- Charikleia Seitanidou
- Theodosis Theodosiou

## LOCAL ORGANIZING COMMITTEE
- Christos Arvanitidis
- Anastasias Chalanaris
- Emmanouil Dialinas
- Ioannis Grigoriadis
- Nikolaos Grigoriadis
- John Kouklinos
Theme: HEALTH

Keynote Talk

Credibility and Utility of Big Data
Ioannidis JPA

Session 1: Medical Genomics

Invited Lectures

Progress in Understanding the Genomic Aetiology of Osteoarthritis
Jeggini E

Integrative Study of the Role of RhoGAP/GEFs in the Rho Protein Signaling Network
Petsalaki E

Short Talks

CanVaS: A Database of Greek Cancer Patient Genetic Variation

Network Based Approaches for Classifying Variants of Unknown Significance
Oulas A, Minadakis G and Spyrou GM

Reconstructing Ikaros Regulatory Network by Applying Causal Discovery Methods on a Compendium of Gene Expression Profiles from Acute Lymphoblastic Leukemia Patients
Nikoloudaki M, Lagani V and Tsamardinos I

Understanding Lymphomas’ Pathobiology Drivers through an Integrative Network Analysis
Seitanidou C, Psatha K and Aivaliotis M
Erythropoietic Transcription Factors Networks in Beta-Thalassemia

Session 2: Neuro-Genomics

Invited Lectures

Mitochondrial Homeostasis in Neurodegeneration and Ageing
Tavernarakis N

Modelling Dendritic and Neuronal Contributions to Cognitive Functions
Poirazi P

Short Talks

Modular Agent-Based Modeling Platform for Bridging Interdisciplinary Research in Ecology, Evolution, Behavioral Biology and Computational Neuroscience
Sakagiannis P, Pavlidis P, Promponas VJ and Iliopoulos I

Elucidating the Role of Different CA3 Hippocampal Neurons in Pattern Completion Using Computational Modelling
Tamiolakis T, Chavlis S and Poirazi P

Session 3: Cellular Networks and Modeling

Invited Talks

Mathematical Modeling of -omics Data for Biofuel Production through Synthetic Biology
Martin H-G

Statistical and Mathematical Models in Computational Oncology
Szczurek E

Short Talk

Automated Machine Learning Methods to Predict Phenotype in Microarray and RNAseq Gene Expression Data
Lakiotaki K, Charonyktakis P and Tsamardinos I
### Session 4: Roundtable – Role of Greek Bioinformatics in Elixir

**Keynote Talk**

Networks Based Approaches in Epigenomics, Evolution and Biomedicine  
Valencia A

#### Theme: ENVIRONMENT

**Keynote Talk**

The Omics Dashboard for Interactive Exploration of High-Throughput Data  
Karp P

### Session 5: Human Microbiome

**Invited Talks**

GSC Metadata Standards in Action: MetaSUB and ISD  
Schriml L

Microbiota Development in the Gut of Infants, Children and Adolescents  
Francino MP

**Short Talk**

Bioenergetic Diversity of the Human Gut Microbiome  
Agioutantis P and Koumandou VL

### Session 6: Computational Metagenomics

**Invited Lectures**

A Novel Pipeline for High Quality Amplicon Analysis Integrated Into MG-RAST  
Meyer F

Expanding the Scope and Reproducibility of Analysis in EBI Metagenomics  
Finn R
Short Talk

**Updates and Applications of the Integrated Microbial NGS Platform (IMNGS). A Step towards Estimation of Global Diversity**
Lagkouvardos I and Clavel T

**Session 7: Environmental Microbiome**

**Invited Lectures**

**Soil Development and Rhizosphere Effect in a High Arctic Desert Chronosequence**
Daffonchio D

**Nitrogen Fixation by the Foliar Conifer Microbiome**
Frank C

**Short Talks**

**The Effects of a Redox Cline and Historical Processes on Bacterial Community Structure in Aegean Marine Sediments**

**Gut and Gonadal Microflora of Insects: Challenges and Practical Applications**
Tsiamis G

**On Genomics Observatories; Towards Real Time Assessment of Ecosystem Structure and Functioning**

**On Genomics Observatories; Towards Real Time Assessment of Ecosystem Structure and Functioning**
Vasileiadou K, Manousaki T, Tsakogiannis A, Tsigenopoulos C and Arvanitidis C

**Session 8: HB Society Matters – Society’s Annual Overview Meeting**
**Theme: BIODIVERSITY**

**Keynote Talk**
Bayesian Matrix Factorization with Side Information and Application to Drug-Target Activity Prediction
Moreau Y

**Session 9: Phylogenomics**

**Invited Talk**
The Evolution of Ribonucleotide Reduction
Lundin D

**Short Talks**
An Updated View of the Oligosaccharyltransferase Complex in *Plasmodium*
Tamana S and Promponas VJ

Investigating the Specialized Roles of Ribosomal Proteins and Associated Ribosomopathies by Means of Literature, Protein Sequence and Phenotype Information Resources
Kyritsis K, Angelis L, Ouzounis C and Vizirianakis I

Evolutionary Models of Amino Acid Substitutions Based on their Neighborhood Tertiary Structure
Primetis E and Pavlidis P

**Session 10: Comparative Genomics**

**Invited Lecture**
What is Life? Conserved Functional Domains in a Hundred Thousand Bacterial Genomes
Ussery D

**Short Talks**
Comparative Genomics of the Sigatoka Disease Complex on Banana Indicates a Direct Link between Pathogen Emergence and Nutritional Virulence
Stergiopoulos I, Chang T-C, Salvucci A and Crous P
Invisible Cities. Segregated Domains in the Yeast Genome with Distinct Structural and Functional Attributes
Nikolaou C

Ancestral Pathway Reconstruction
Ouzounis C

Genomic Encyclopedia of Bacteria and Archaea
Kyprides N

Session 11: Computational Technologies

Invited Lecture
One Tagger, Many Uses: Simple Text-Mining Strategies for Biomedicine
Jensen L

Short Talks
Latest Improvements in ProteoSign, a Web Application for Differential Proteomics Analysis

Calchas: A Semantic Search Engine for Bioinformatics Resources
Koumakis L, Giannoulis M, Kondylakis H, Kanterakis A and Potamias G

Large-Scale Study on Latent Feature Construction for Gene Expression with Improved Predictive Power on Newly-Seen Datasets
Tselas C, Pantazis Y, Lakiotaki K and Tsamardinos I

Arkalos: An open collaborative Workflow Management System for reproducible bioinformatics research
Kanterakis A, Koumakis L, Karacapilidis N and Potamias G

Keynote Talk
Space Technology and Space Biology
Papadopoulos P
1. **Genome-Wide Localization and Inter-Dependence of Aft1 and Mac1 Transcription Factors In S. Cerevisiae**

2. **Investigating the Role Of VIP+ Interneurons In Learning-Related Place Cell Dynamics In Hippocampal Area CA1**
   Chavlis S, Bozelos P, Li W, Turi G, Losonczy A and Poirazi P

3. **The Centrality Lethality Rule in Signed Protein Interaction Networks**
   Paragamian S, Nikolaou C, Sgardelis S and Antoniou I

4. **A Tale of Two Trees: Modeling Apical and Basal Tree Contribution to L2/3 V1 Pyramidal Cell Orientation Selectivity**
   Petousakis K-E, Papoutsi A and Poirazi P

5. **Forward-Backward Spatial Simulator for Genetic Data**
   Koropoulis A and Pavlidis P

6. **Evolution of Gene Regulatory Networks by Means of Genetic Drift and Natural Selection**
   Kioukis A and Pavlidis P

7. **Text Mining the Scientific Literature on Research Integrity and Misconduct**
   Papanikolaou N, Gioti A, Sifaki E, Iliopoulos I and Petoussi V

8. **Transcription Factors Binding, Gene Expression and Positive Selection detection on Topologically associated domains and Lamina associated domains in Humans and Mice**
   Vagiaki D, Primetis E, Kollias A and Pavlidis P

9. **Clinical Relevance of Single-cell Immune Signatures in Cellular Subpopulations**
   Markaki M, Tsagris M, Limaci M, Papoutsoglou G and Tsamardinos I

10. **DNA Damage Response and Zeocin-specific Transcriptional Profiles in S. cerevisiae**
    Dialynaki D, Fragiadakis G S, Stratidaki I, Gounalaki N, Topalis P and Alexandraki D
11. **Gene Co-expression Network Analysis on Mantle Cell Lymphoma Transcriptomics Data Sheds Light on Novel Diagnostic, Prognostic and Therapeutic Strategies**
   Seitanidou C, Skrempou G, Psatha K, Oulas A, Drakos E, Spyrou GM and Aivaliotis M

12. **Data Integration of Single-Cell Mass Cytometry Measurements using Advanced Imputation Methods**
   Khadka S, Pantazis Y and Tsamardinos I

13. **Phylogenomic Analyses Following Whole Genome Sequencing Resolves the Phylogenetic Position of Gilthead Seabream (Sparus aurata) within the Tree of Teleost Fishes**
   Manousaki T, Ha QV, Bargelloni L and Tsigenopoulos CS

14. **Phylogeography of the Toxic Cyanobacterium Cylindrospermopsis raciborskii**
   Panou M and Gkelis S

15. **Untangling the Microbial Dark Matter through a Metagenomic Approach: The Case of Etoliko Lagoon**
   Stathopoulou S and Tsiamis G
ORAL PRESENTATIONS
Credibility and Utility of Big Data

John P.A. Ioannidis

C.F. Rehnborg Chair in Disease Prevention; Professor of Medicine, of Health Research and Policy, of Biomedical Data Science, and of Statistics, Stanford University; Co-director, Meta-Research Innovation Center at Stanford (METRICS)

Research using "big data" is becoming more common across many biomedical domains and other areas of science. Despite its high popularity, the term lacks a proper, widely accepted definition. There is great hope and hype about big data research. The high dimensionality of the data and its increasing volume, as well as the new research practices that are associated with big data applications create several challenges and ethical, organizational, and technical/methodological concerns. The lecture will try to navigate through these opportunities and challenges and suggest some recommendations that would help maximize the credibility and utility of big data.
Progress in Understanding the Genomic Aetiology of Osteoarthritis

Eleftheria Zeggini
Sanger Center, UK

Osteoarthritis is a common complex disease with a substantial public health burden. The genetic causes and molecular pathways underpinning disease development remain largely unidentified. In this talk I will highlight recent successes in understanding the genetic aetiology of osteoarthritis through large scale genome-wide association studies and promising functional genomics approaches for deciphering the processes that lead to disease development and progression.
Rho-GTPases are molecular switches that control critical aspects of cell biology. With the help of guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs), they cycle between an active (GTP-bound) state and an inactive (GDP-bound) state. While the human genome encodes ~18 RhoGTPases, there are over 140 RhoGAP/GEFs. These have diverse sets of interaction domains, which may target them to different subcellular locations, and allow them to connect upstream and downstream molecules to form transient localized Rho signaling platforms. Despite extensive research in the field, many of the RhoGAPs/GEFs are poorly characterized, and their contributions to signaling pathways and cell dynamics are not well understood. Thus, we have undertaken a systematic approach to study these proteins, integrating interactome, imaging and functional data, with the aim to understand their role in controlling Rho signaling specificity.
**CanVaS: A Database of Greek Cancer Patient Genetic Variation**

**Despoina Kalfakakou**, Florentia Fostira, Paraskevi Apostolou, Myrto Papamentzelopoulou, Andromahi Vagena, Irene Konstanta, Angeliki Delimitsou, Ioannis Vlachos, Drakoulis Yannoukakos and Irene Konstantopoulou

*Molecular Diagnostics Laboratory, INRaSTES, National Center for Scientific Research “Demokritos”, Athens, Greece*

Following recent advances in sequencing and cancer genetics that have led to the generation of extremely large amounts of data, the need for the development of national registries and specific-to-gene or disease databases is greater than ever. A population-specific mutation and variation database can be an important resource for researchers, clinicians and laboratory scientists as inter-population variant frequencies can be recorded and evaluated for possible pathogenicity. Herein we report the construction of the first Greek database for hereditary cancer. Data derive from the genetic analyses of more than 5000 cancer patients that have been analyzed in Molecular Diagnostics Laboratory by Sanger or Next Generation Sequencing (panel testing) as well as through collaborations with other Research Institutes, namely University of Washington and Mayo Clinic. The dataset consists of ~3000 variants in coding regions (which will be sub-divided in loss-of-function mutations and nucleotide variations depending on their effect) and ~3800 variants in non-coding regions of approximately 100 known cancer susceptibility genes. The database will provide minor allele frequencies for the Greek population, possible genotype-phenotype associations, impact on clinical significance and links to other public databases. Most importantly, all data will be available for download in multiple formats via anonymous FTP. Hopefully, this database will lead the way to a coordinated effort along with other public and private diagnostic laboratories in order to summarize all data and clarify Greek cancer mutation spectrum.
Network Based Approaches for Classifying Variants of Unknown Significance

Anastasis Oulas, George Minadakis and George M Spyrou

The Cyprus Institute of Neurology & Genetics, Cyprus

Human genetic variation analysis has recently been influenced by next generation sequencing (NGS) technologies in the form whole-genome sequencing (WGS), whole-exome sequencing (WES) and multigene panels. These methodologies, as well as genome wide association studies (GWAS), have paved the way for analysing genetic variation by means of global, high-throughput methods. Currently there is a growing trend to incorporate these technologies in clinical diagnosis, prognosis and even choice of underlying treatment, as well as overall patient consultation and advising. The notion of utilizing these methods regularly in medical care, however, poses multiple challenges one of which is how to address variants of uncertain significance (VUS). These VUS, such as single nucleotide polymorphisms (SNPs) or insertions/deletions (indels), cause dilemmas for clinicians and uncertainty on how to advise patients, hence, MDs often refrain from making patients aware of information that might cause further discomfort and unnecessary concern.

In this study we describe computational approaches, that make use of a network gene association information coupled with a polygenic risk model approach to detect groups (>=2) of synergistically acting variants required to accurately predict disease outcome. The approach makes use of real data in the form of genotypes derived from WGS, WES, multigene panels or GWAS data and provides for a disease related classification for VUS while also assessing their synergistic effect on risk model prediction accuracy for the disease under investigation. Synergy can be defined as the phenomenon where a combination of variants provides additive value than the sum of individual variants. Our methodology aims to detect such synergistic sets of variants and provides sound statistical validation to support findings. Moreover, biological significance and interpretation of the results generated by our methodology further substantiates the overall methodology and provides concrete evidence of its practical application and usability.
Reconstructing Ikaros Regulatory Network by Applying Causal Discovery Methods on a Compendium of Gene Expression Profiles from Acute Lymphoblastic Leukemia Patients

Maria Nikoloudaki, Vincenzo Lagani and Ioannis Tsamardinos

Computer Science Department, University of Crete

Acute Lymphoblastic Leukemia (ALL) is one among the most common types of cancer, especially in children. The transcript factor Ikaros (IKZF1) has been proven to have principal role in B cell differentiation and is associated to leukemogenesis. A better understanding of IKZF1 interaction network might provide a better understanding of leukemia mechanisms at a molecular basis and a clearer view of Ikaros role in the disease.

The goal of this study is reconstructing the regulatory causal network of IKZF1 by applying (causal) network reconstruction methods on a compendium of microarray gene expression data. Data from five different studies on B Cell Precursor BCP ALL [1-5] were analyzed together after undergoing a uniform pre-processing and batch-effect removal process. For each study, subjects with IKZF1-impairing genetic aberrations were identified, leading to a total number of 151 Ikaros mutated subjects out of 597 ALL patients. Two different approaches were used for investigating the Ikaros network: (a) feature selection algorithms able to identify genes whose expression is directly affected by Ikaros mutations and (b) network reconstruction algorithms able to identify probesets (causally) connected to IKZF1 expression variations.

The results were contrasted against the information contained in the STRING online biological database and were visualized through the Cytoscape application. Both approaches identified several putative interactions between Ikaros and other genes, some of which were already known in the literature.

Understanding Lymphomas’ Pathobiology Drivers through an Integrative Network Analysis

Charikleia Seitanidou\textsuperscript{1,2}, Konstantina Psatha\textsuperscript{1} and Michalis Aivaliotis\textsuperscript{1}

\textsuperscript{1} IMBB-FORTH
\textsuperscript{2} University of Crete

Lymphomas, a haematological disorder of the lymphocytes, are well known for their high heterogeneity in terms of etiology, diagnosis and clinical outcome \cite{Psatha_2017}. One of the major characteristics of this blood cancer is the inactivation of the wild type (wt) p53 protein by the overexpression of its main negative regulator, MDM2. This dysfunction affects directly the p53-dependent signaling pathways, resulting in loss of DNA damage control and uncontrolled proliferation \cite{Fahraeus_2014}. Nutlin-3a (N3a), an MDM2 antagonist which has recently entered clinical trials, has proven to efficiently re-activate non-functional wt p53 in lymphoma as, exhibited by cell cycle arrest and apoptosis \cite{Drakos_2009, Drakos_2011}. Our work aims on decoding lymphomas’ pathobiology and progression by using a systemic approach towards the identification of the key mal-functioning p53-centered network of interactions. Geared towards this goal, microarray transcriptomic and mass-spectrometry-based proteomic analyses on three model lymphoma cell lines (ALCL, HL, MCL) treated with N3a were integrated and interrogated using specific network inference methods and algorithms \cite{Shannon_2003}. Development and application of two distinct approaches (Hypothesis Free and Targeted) led us to the identification of novel protein-protein interactions, delineating specific and overlapping molecular signatures among the different lymphoma subtypes. Using available plugins and tools from Cytoscape revealed differentially regulated transcript-protein correlations specific to the different lymphoma phenotypes. To conclude, our systems biology point of view investigated the regulatory pathways and protein interaction networks involved in a wide range of cellular processes of lymphoma pathophysiology. Such findings may delve into the molecular identification of unique and common characteristics between different lymphoma entities and other diseases, leading future research to new drug development and/or drug repurposing.

Erythropoietic Transcription Factors Networks in Beta-Thalassemia

Aikaterini Nanou¹,#, Chrisavgi Toumpeki¹,#, Konstantinos Vougas¹, Giorgos Giagkas¹, Pavlos Fanis², Nicoletta Bianchi³, Lucia Carmela Cosenza³, Roberto Gambari³, Marina Kleanthous² and Eleni Katsantoni¹

¹ Basic Research Center, Biomedical Research Foundation, Academy of Athens, Athens, Greece
² The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus
³ Department of Life Sciences and Biotechnology, Ferrara University, Ferrara, Italy
# Equal contribution

Beta-thalassemia is an autosomal recessive disease characterized by severe anemia. More than 200 mutations have been identified, the majority of which are point mutations in the beta-globin gene or in regulatory regions that control the expression of beta-globin gene, causing abnormal production of beta-globin. There are three types of beta-thalassemia: major, intermedia and minor. Beta-thalassemia major and intermedia differ in their clinical characteristics and the necessity for blood transfusions. Here, we applied a transcriptomics and proteomics approach in erythroid progenitors cell cultures from thalassemia major and intermedia patients, and healthy donors. Identified differentially expressed proteins / genes included transcription factors, proteasome subunits, factors with acetyltransferase activity, kinase activity etc and highlighted the changes in erythropoietic circuits between the different groups used. Pathway analysis provided further information on the molecular differences between the two different phenotypes of beta-thalassemia. The findings will delineate the mechanisms by which erythropoietic transcription factors exert key roles in erythropoiesis and will lead to identification of potential biomarkers or therapeutic targets for the therapeutic management of beta-thalassemia patients.
Mitochondrial Homeostasis in Neurodegeneration and Ageing

Kostas Palikaras\textsuperscript{1} and \textit{Nektarios Taverakis}\textsuperscript{1,2}

\textit{Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-Hellas}\textsuperscript{1} and \textit{Medical School, University of Crete, Heraklion, Crete, Greece}\textsuperscript{2}

Mitochondrial function impinges on several signalling pathways modulating cellular metabolism, cell survival and healthspan. Maintenance of mitochondrial homeostasis requires both generation of new, and elimination of dysfunctional mitochondria. Impaired mitochondrial content homeostasis is a common characteristic of ageing and several human pathophysiological conditions, highlighting the pivotal role of the coordination between mitochondrial biogenesis and mitophagy. However, the cellular and molecular underpinnings of the relevant mechanisms remain obscure. We found that DCT-1, the Caenorhabditis elegans homolog of mammalian BNIP3 and BNIP3L/NIX, is a key mediator of mitophagy, promoting longevity under stress. DCT-1 acts downstream of the PINK-1/Parkin pathway and is ubiquitinated upon mitophagy-inducing conditions to mediate the removal of damaged mitochondria. Accumulation of damaged mitochondria triggers SKN-1 activation, which initiates a bipartite retrograde signaling pathway stimulating the coordinated induction of both mitochondrial biogenesis and mitophagy genes. Our results unravel a homeostatic feedback loop that allows cells to adjust their mitochondrial population in response to environmental and intracellular cues. Age-dependent decline of mitophagy both inhibits removal of dysfunctional or superfluous mitochondria and impairs mitochondrial biogenesis resulting in progressive mitochondrial accretion and consequently, deterioration of cell function.
**Modelling Dendritic and Neuronal Contributions to Cognitive Functions**

Yiota Poirazi

*Institute of Molecular Biology and Biotechnology (IMBB), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, GREECE*

The goal of this presentation is to provide a set of predictions generated by biophysical and/or abstract mathematical models regarding the role of dendrites in information processing, learning and memory across different brain regions. Towards this goal I will present modelling studies from our lab –along with supporting experimental evidence- that investigate how dendrites may be used to facilitate the learning and coding of both spatial and temporal information at the single cell, the microcircuit and the neuronal network level. I will briefly present early work on how the dendrites of individual CA1 pyramidal neurons may allow a single cell to act as a 2-stage neural network classifier\(^1\), thus massively increasing the storage capacity of the neural tissue\(^2\). I will then discuss how such dendritic nonlinearities may enable stimulus specificity in individual PFC pyramidal neurons during working memory\(^3\) and underlie the emergence of sustained activity at the single cell and the microcircuit level\(^3,4\). The role of dendrites in memory phenomena will be assessed using circuit models of the Dentate Gyrus implementing pattern separation\(^5,6\) as well as hippocampal models capable of learning associative memories and linking them across time\(^7\). This presentation aims to highlight how dendrites are likely to serve as key players in different memory functions.

References:


Modular Agent-Based Modeling Platform for Bridging Interdisciplinary Research in Ecology, Evolution, Behavioral Biology and Computational Neuroscience

Panagiotis Sakagiannis\textsuperscript{1}, Pavlos Pavlidis\textsuperscript{2}, Vasilis J Promponas\textsuperscript{3} and Ioannis Iliopoulos\textsuperscript{1}

\textsuperscript{1} School of Health Sciences, Faculty of Medicine, University of Crete, Greece
\textsuperscript{2} ICS, FORTH, Greece
\textsuperscript{3} Bioinformatics Research Laboratory, Department of Biological Sciences, University of Cyprus, Cyprus

Agent-based modeling (ABM) is a paradigm widely used in computational ecology for modeling populations of living organisms. It provides a framework for conducting interdisciplinary research by allowing findings from experimental biology to be incorporated in the single organism's model. Population features emerge without being explicitly modeled by simulating multiple agents’ interactions in the same environment. In our model, elementary interactions comprise external interactions such as feeding, mating, sensing touch or odor and moving, and internal drives such as hunger and sexual drive. These are coupled within functional modules such as metabolism, sensorimotor, appraisal memory, decision making and genetic inheritance during agent’s life within a complex spatial environment. In contrast to previous ABM or Artificial Life implementations, the platform has been constructed in a highly abstract and generic way in order to be extendable towards implementations in (i) computational neuroscience, (ii) population genetics and (iii) behavioral biology. As an alternative to rule-based modeling, the open-source platform supports the use of biologically plausible large-scale neural networks shaped by reinforcement learning in distinct modules of the organism, e.g. sensorimotor coupling in a simple motor module to approach or avoid odor gradients. Evolutionary and population genetics studies are supported by the realistic genome implementation incorporating distinct phenotypically functional loci, recombination and mutation process, differential survival and reproduction rates as a function of agent’s genotype, learning and random genetic drift. Finally, bridging to experimental behavioral biology is supported by allowing the storage of the state of a certain agent (genetic and acquired attributes). Then, by creating a population of its clones and putting them in a specific-task test simulation environment we can compare their performance against a control population.
Elucidating the Role of Different CA3 Hippocampal Neurons in Pattern Completion Using Computational Modelling

Theodoros Tamiolakis, Spyridon Chavlis and Panayiota Poirazi

IMBB-FORTH, Greece

Hippocampus, one of the basic components of the limbic system, is crucial for the efficient long-term memory storage in the brain. Damage to the hippocampus has been associated with neurodegenerative diseases such as Alzheimer's disease and Schizophrenia. Therefore, it is utterly important to understand how it works. Here, we study CA3, an area of the hippocampus, which has been associated with pattern completion. This process occurs when a network recovers a complete, stored memory, based on a partial, noisy or degraded input. Towards this goal, we have built an integrate-and-fire model of CA3 with adaptation using the BRIAN simulator. The network consists of 300 excitatory pyramidal cells and 76 inhibitory interneurons from three different classes: Basket (38), OLM(19) and RLM(19) cells. Pyramidal cells receive input from the Dentate Gyrus (DG) and Entorhinal Cortex (EC), and have a high level of recurrency. All neurons are modelled as simplified point neurons without dendrites. All neuron models as well as the network connectivity are validated against experimental data to ensure biological relevance. The network reproduces the spiking profiles of CA3 neurons in response to random Poisson input. The next goal is to demonstrate the network’s ability to perform pattern completion when presented with partial input.
Mathematical Modeling of -omics Data for Biofuel Production through Synthetic Biology

Hector-Garcia Martin

Joint Bioenergy Institute, USA

The 21st century has seen biology radically change its nature by the capability to synthesize DNA and express it inside a cell. At the Joint BioEnergy Institute (JBEI, http://www.jbei.org), we are leveraging synthetic biology to produce a variety of renewable biofuels. However, while DNA sequencing, DNA synthesis and functional genomics (-omics) data generation capabilities improve rapidly, our capability to predict the outcomes of bioengineering efforts remains nascent.

The goal of the Quantitative Metabolic Modeling directorate at JBEI is to develop models of metabolism which are both quantitative and predictive, in order to improve biofuel production in a rationally directed fashion. We use experimental, computational and mathematical tools to achieve this goal.

Our efforts are divided into three main areas: 1) flux-based mechanistic models that leverage $^{13}$C carbon labeling experiments, 2) data mining and machine learning of -omics data; 3) development of web-based tools for -omics data visualization and storage. In this talk, I will give examples of each area and I will showcase how they are used to improve biofuel production in engineered cells.
Statistical and Mathematical Models in Computational Oncology

Ewa Szczurek

University of Warsaw, Poland

The immense complexity of alterations that affect tumor genomes calls for computational methods that will improve our understanding of this disease. Here, I will present the set of tools we develop to model mutational patterns in tumor genomic data. I will focus on our previous and current methods for detecting synthetic lethal interactions among human cancer genes. Synthetic lethality occurs when the co-inactivation of two genes results in cellular death, while inactivation of each individual gene is viable. Synthetic lethality can be exploited in cancer therapy. For cancer patients, one inactivation already occurs via the endogenous mutation of a specific gene in the tumor cells, and not in the normal cells. Thus, applying a drug that targets the synthetic lethal partner of that gene will selectively kill cancer cells, leaving the rest viable. Our recent efforts, summarized at http://www.mimuw.edu.pl/~Szczurek/sl, aim at broad search of clues for synthetic lethal interactions in various data sources, from patient survival to cell line data.
Automated Machine Learning Methods to Predict Phenotype in Microarray and RNAseq Gene Expression Data

Kleanthi Lakiotaki\(^1\), Paulos Charonyktakis\(^2\) and Ioannis Tsamardinos\(^1,2\)

\(^1\)Computer Science Department, University of Crete, Greece
\(^2\)Gnosis Data Analysis, PC, Greece

High throughput transcriptomics became possible with microarrays and RNA sequencing that produce a vast amount of publicly available data. Gene expression analysis has become routine nowadays helping researchers to study the behavior of cells, the functionality of tissues, and a wide range of biological processes. Knowledge gained through gene expression data analysis is increasingly important as it is useful for phenotype classification of diseases. Identification of differentially expressed genes has mainly focused on the most significant changes and may not allow discovery of more subtle patterns in the data. Great potential exists for computational methods to analyze these data for the study of cellular regulation mechanisms, disease diagnosis and drug development. The high dimension and noise associated with these data though, require advanced data analysis methods. In this work, we use tools and algorithms from machine learning to build data-driven predictive models for both microarray and RNASeq gene expression data sets. The prediction models were trained with the Just Add Data Bio (JADBio; Gnosis Data Analysis; http://www.gnosisda.gr). JADBio performs multiple feature selection, hyper-parameter tuning and automated prediction model selection and ensures unbiased estimation of the mean performance and the confidence intervals of the final selected model. The pipeline generates several configurations (combinations of feature selection algorithm with classification algorithm for specific values of their hyper-parameters) and subsequently, estimates their performance using stratified K-fold cross-validation. Subsequently, it selects the best configuration and trains with it the final model using all available training data. The last step is to provide an estimate of the performance of the returned model. Because, the model was produced with the best configuration, its cross-validated performance is optimistic. Therefore, the tool estimates the bias of the performance using a bootstrap method, and removes it to return the final performance estimate and confidence intervals. We applied JADBio in three gene expression microarray and RNASeq datasets from Gene Expression Omnibus, studying two chronic inflammatory diseases, psoriasis and systemic lupus erythematosus. We were able to discover a small set of genes that achieve high phenotype prediction accuracy measured in AUC in all datasets. We also used Enrichr (http://amp.pharm.mssm.edu/Enrichr/) for a functional analysis of those gene sets.

Acknowledgment. Part of research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement n. 617393]
**Networks Based Approaches in Epigenomics, Evolution and Biomedicine**

**Alfonso Valencia**  
*ICREA professor, Barcelona Supercomputing Centre, Spain*

In the first study, we processed heterogeneous ChIP-Seq information to build a comprehensive genome co-localization network of Chromatin Related Proteins (CRPs), histone marks and DNA modifications in mouse embryonic stems cells. In this network, co-localization preferences can be translated into specific of “mESC Chromatin States”, such as active regions or enhancers. The study of the properties of the “co-localization” network points to the 5hmC DNA modifications, as the key component in the organization of the mouseESC network.

In a second network based study, the importance of 5hmC, as organizer of the epigenetic network, is reinforced by the evolutionary analysis of the protein components of the network. There, 5hmC acts as a mediator in the co-evolution of the CRPs protein components of the mESC network.

The third network-based approach explores the functional significance of the mESC Epigenetic Properties and Chromatin States, by analysing them in the context of the structure of the chromatin in the cell nucleus. The results revealed interesting properties of the organization of the mESC epigenetic control system, in line with the emerging models of gene expression control and chromatin organization, and again support the role of 5hmC as a factor present in a significant number of interactions related with active transcription in mouse embryonic stems cells.

One additional network approach shows how the same network properties can help to understand the complex relations between expression patterns related with human diseases.


*The direct and invers comorbidity network.* Sanchez-Valle et al., 2017 in preparation.

*Parts of this work were developed in collaboration with: Vingron’s (MPIMG, Berlin), Fraser’s (Babraham Institute), and Baudot’s labs (CNRS, Marseille).*
The Omics Dashboard for Interactive Exploration of High-Throughput Data

Peter Karp
SRI, USA

The Omics Dashboard is a novel software tool for interactive exploration and analysis of gene-expression and metabolomics datasets. The dashboard is organized as a hierarchy of cellular systems. At its highest level the dashboard contains panels for cellular systems such as biosynthesis, energy metabolism, regulation, and central dogma. Each of those panels contains a series of X—Y plots depicting the expression levels of genes within subsystems of that panel, e.g., subsystems within the central dogma panel include transcription, translation, and protein maturation and folding. The dashboard presents a visual read-out of the expression status of cellular systems to facilitate a rapid user survey of how all cellular systems are responding to a given stimulus, and to enable the user to quickly find and understand the response of genes within one or more specific systems of interest. The dashboard enables an investigator with a limited understanding of biochemistry to quickly observe and analyze the functioning of the entire metabolic system without analyzing the expression of single genes. The dashboard is complementary to traditional statistical methods for analysis of gene-expression data. We present the capabilities of the dashboard using case studies of analysis of lipid production for the marine alga Thalassiosira pseudonana, and during a shift from anaerobic to aerobic growth for the bacterium Escherichia coli.
The Genomic Standards Consortium (GSC) has developed and implemented a suite of metadata standards to capture contextual metadata for genomic sampling and sequencing projects. In 2016, the GSC MIxS standards were utilized in the study of microbiome communities in Crete soil samples and subway surface samples as part of the MetaSUB project.
Microbiota Development in the Gut of Infants, Children and Adolescents

Maria Pilar Francino
CSISP, Spain

The human gut microbiota develops rapidly throughout the first year of life, but little is known regarding the continuation of this process during childhood. We are applying a variety of “omics” approaches to understand the changes that occur in the gut microbiota at different stages from birth to childhood, adolescence and adulthood. Biodiversity analyses based on 16S rRNA gene amplicons, as well as metagenomic and metatranscriptomic analyses, are revealing some general trends in this process and allowing us to address how the temporal dynamics of the gut microbiota change with age. Network analyses indicate likely interactions among gut microbiota taxa and suggest how these interactions appear and evolve through time.
Bioenergetic Diversity of the Human Gut Microbiome

Panagiotis Agioutantis¹ and Vassiliki Lila Koumandou²

¹ National Technical University of Athens, Greece
² Agricultural University of Athens, Greece

The human microbiome has lately emerged as an important factor in health and disease. Metagenomics studies of the microbial diversity in the human gut of healthy adults, children and infants, show that (a) the microbiome is less diverse early in life, and gets enriched after infancy, and (b) that it is dominated by organisms from two bacterial phyla: bacteroidetes and firmicutes. However, there is still considerable diversity between individuals, based on age, habitat and health status, while a lot of questions on how the diversity of the microbiome affects our health still remain unanswered. One aspect that has been largely overlooked is the diversity of the microbiome with regards to bioenergetic pathways. We therefore set out to address this by re-examining freely available data from three major studies of human gut microbiota, as deposited in MG-RAST, based on the available complete genome sequences of a set of species which represent the full bioenergetic diversity across all the major bacterial and archaeal lineages. Our results indicate that several species present in the human gut of both adults and infants normally derive their energy from methanogenesis, iron oxidation, iron reduction, sulfate and arsenate reduction, and even anoxygenic photosynthesis. We discuss the possible effects of the presence of these bioenergetic pathways to the dynamics of the human gut microbial community, as well as future directions to better characterize this previously unsuspected diversity.
A Novel Pipeline for High Quality Amplicon Analysis Integrated Into MG-RAST

Folker Meyer

ANL, USA

We are expanding MG-RAST to incorporate additional workflows. A novel high-precision workflow for 16s and ITS amplicon data analysis is the first additional workflow that was added to MG-RAST.

I will briefly describe the MG-RAST system and then provide an overview of the novel workflow and compare its performance to existing workflows.
Expanding the Scope and Reproducibility of Analysis in EBI Metagenomics

Rob Finn

European Bioinformatics Institute, UK

The field of metagenomics has changed radically over the past decade. Initially, the SSU marker gene was used to probe which bacteria were present in a particular sample. However, with the advent of modern sequencing technologies and evolving experimental design, datasets have radically changed, growing to large scale shot-gun, multi-sample datasets looking at the whole genomic composition of all taxa. Such changes have lead the research community to produce a myriad of tools and reference databases that can perform a range of analyses, from specific to generic. The EBI metagenomics portal (http://www.ebi.ac.uk/metagenomics) provides a free-to-use analysis platform, which is agnostic of biome or sequencing methodology. However, due to the diversity of metagenomics experiments the one-pipeline-does-all approach no longer applies. I will describe the latest developments, which includes the analysis of eukaryotes and assembled data. How we communicate the different analysis to the user and ensure reproducibility and facilitate the rapid extension of the pipelines will be outlined.
Updates and Applications of the Integrated Microbial NGS Platform (IMNGS). A Step towards Estimation of Global Diversity

Ilias Lagkouvardos and Thomas Clavel
Technical University of Munich, Germany

The Integrated Microbial Next Generation Sequencing (IMNGS) platform is a useful resource of thousands of uniformly processed microbial profiles based on 16S rRNA gene amplicon sequences. The ability to identify natural ecosystems of target bacteria of interest and to estimate the global diversity of selected taxonomic groups are exclusive features of the system. IMNGS also offers a de novo clustering functionality for processing users’ own raw amplicon data. We are currently expanding functionality of the platform to allow streamlined submission of users data to the European Nucleotide Archive after processing, which will facilitate the inflow of dataset submission to public repositories. We recently used IMNGS to bring clarity to estimates of the cultured fraction of bacterial populations in the human and mouse gut. Comparison of thousands of amplicon datasets to an in-house built catalogue of approximately 15,000 full-length 16S rRNA gene sequences from isolates revealed that, on average, 35 to 65% of microbial species detected by high-throughput sequencing in the mouse and human intestine have representative strains in culture. IMNGS is also being used to assign environmental preferences to species of the enigmatic family Muribaculaceae (formerly family S24-7 within the Bacteroidales). In the latter work, we map all amplicons in IMNGS to almost 600 candidate species of this family to obtain an overview of their environmental and host prevalence and abundance. Overall, the platform keeps expanding both in terms of the number of available samples, users, and functionalities, forming a valuable resource dedicated to microbiology research freely available at http://www.imngs.org.
Soil Development and Rhizosphere Effect in a High Arctic Desert Chronosequence

Daniele Daffonchio

KAUST, King Abdullah University of Science & Technology, BESE, Thuwal, Saudi Arabia

In desert ecosystems, either cold or warm, or in lands undergoing desertification, soil is poorly structured and has not yet or it is loosing the properties of mature soils and the capacity of supporting plant growth and crop production. While the effect of soil vs plant-induced selection of rhizosphere microbiomes has been extensively studied in mature natural and agricultural soils, the role of soil developmental state on the assembly of rhizosphere microbiomes is poorly understood. To respond to this question I discuss a study of bacterial diversity changes and OTU co-occurrence network in unvegetated and rhizosphere soils of a pioneer plant across a glacier foreland chronosequence in the High Arctic (Svalbard Islands, 78°53’N), where the barren substrate of the moraine is progressively undergoing a soil developmental process.
Nitrogen Fixation by the Foliar Conifer Microbiome

Carolin Frank1, Andrew Moyes2, Lara M Kueppers2,3, Jennifer Pett-Ridge4, Dana Carper1, Dianne Quiroz2, James Kupihea1, Paola Saldierna1

1 Life and Environmental Sciences, UC Merced, USA
2 Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, USA
3 Energy and Resources group, UC Berkeley, USA
4 Physical and Life Sciences Directorate, Lawrence Livermore National Lab, USA

Mature temperate and boreal forest are nitrogen (N) limited, yet N budgets indicate unknown sources of N in these ecosystems. Symbiotic N-fixing plants are notably absent from coniferous forests, and sources to overcome N limitation are not well understood, but include epiphytic N2-fixation in mosses, free-living N2, fixation in litter and soil, bedrock nitrogen where sedimentary rocks occur, and deposition of nitrogen pollution.

Via Illumina sequencing of the 16S RNA gene, we have found a consistent association between pines growing in nutrient limited ecosystems and specific bacteria, most notably potential N2-fixing acetic acid bacteria and Rhizobiales spp. These associations appear to be conserved across host species, time, and geographic distance, suggesting selection on the part of the tree, the bacteria, or both, potentially reflecting a functional partnership based on N2-fixation. Using the acetylene reduction assay on surface sterilized foliar samples, we have confirmed nitrogenase activity in the subalpine conifer limber pine growing at high elevation in Colorado, as well as in bishop pine and lodgepole pine growing along a gradient in soil age and associated variation in soil fertility at the “ecological staircase” in Mendocino, California.

Not surprisingly, N2-fixation rates of needle endophytes are much lower than those of nodulating N2-fixers, and comparable to rates of free-living fixation in soils in temperate and boreal ecosystems. So far, we have found no evidence that local differences in soil N availability affect rates of foliar N fixation. Together, these results suggest that foliar endophytes represent a low-cost, evolutionarily stable N2-fixing strategy for long-lived conifers that never fully alleviates N limitation in temperate and boreal ecosystems. These results open up the possibility that ‘hidden’ symbiotic N2-fixers hide in other N-poor ecosystems as well.
The Effects of a Redox Cline and Historical Processes on Bacterial Community Structure in Aegean Marine Sediments

Anastasia Gioti, Vassiliki Kalogeropoulou, Ioannis Karakassis and Emmanuel D. Ladoukakis
University of Crete, Greece

The structure of bacterial communities is determined by two groups of ecological processes: Environmental filtering, such as changes in temperature and other parameters, and historical processes, i.e. past selection and ecological drift in combination with some degree of dispersal limitation. The relative contribution of these processes in shaping the biogeographical patterns of prokaryotic communities is a field of active research. In this study, we examined the parameters that determine the bacterial community structure in marine sediments coming from seven sites of the Central and Northern Aegean Sea, sampled at different depths and at a high degree of technical and biological replication. We combined a meta-barcoding approach (Illumina sequencing of the 16S gene) with multivariate statistical analyses to determine and compare the community structure of these sites, and examine the contribution of selective and historical processes in the distribution of taxa. For this purpose, a number of factors was considered, such as the redox potential (Eh), temperature and chlorophyll α concentration, as well as geographical location.

We identified a surprisingly high number of low-abundance novel taxa, indicating that the diversity of the Aegean Sea sediment remains largely uncharacterized. Our results show that both environmental filtering and historical processes affect bacterial community structure. However Eh, an indicator of oxygen availability in the sea sediment, appears as the most important driver of community structure in the studied marine sediments. Interestingly, we show that low oxygen availability increases phylogenetic diversity, while we observe benthic stratification of communities, even at small depth differences.

This study was funded by the EU/GSRT (Αριστεία ΙΙ) grant “Benthic pelagic coupling: hypoxia and regime shifts (HYPOXIA)”.

PROGRAMME
LIST
Insects are by far the most diverse and most abundant animal group with respect to the number of species globally, in ecological habitats and in biomass. The ecological and evolutionary success of insects depends on their countless relationships with beneficial microorganisms, which are known to influence all aspects of their biology, physiology, ecology and evolution. As for essentially all animals, microbial communities are particularly prominent in the digestive tract, where they may be key mediators of the varied lifestyles of insect hosts. The contribution of microorganisms, particularly gut microorganisms, to insect function is highly relevant from several perspectives, linking to applications in medicine, agriculture, and ecology. Gut-associated bacteria can influence: (a) vectoring efficiency, (b) developmental time, (c) decomposition of plant biomass and carbon cycle, (d) nitrogen fixation and nitrogen cycle, (e) mating incompatibilities, and (f) detoxification of pesticides leading to the acquisition of insecticide resistance.

We have deployed an amplicon sequencing approach which we couple with an in situ hybridization approach in order to decipher and characterize symbiotic bacterial communities of important insects. I will be using case studies from the Mediterranean fruit fly, tsetse fly and other fruit flies in order to highlight mutualistic, and applied aspects of these symbiotic communities. In more detail, Spiroplasma has been identified as a new symbiont in tsetse flies and a putative mutualistic role has been attributed in Glossina fuscipes fuscipes. The Mediterranean fruit fly, Ceratitis capitata, has been used in order to provide more direct evidence of gut symbiotic structuring during a laboratory adaptation process.
Genomics observatories address the molecular and biological component of ecosystems by time-series assessment of the genomic and transcriptomic makeup of biological communities across food webs, captured and analyzed together with high-quality environmental data. This implies standardized collection, management, analysis and dissemination of BIG COMPLEX DATA (BCD). The ambition of GOs is no less than assessing the dynamics of the structure and function of ecosystems and understanding the evolutionary processes that have shaped the observed patterns. Modeling of such complex systems will allow hypothesis testing, theory building and developing prognostic tools. The technical challenges for operational and evolving GOs are rivaled by social type challenges such as participatory processes and community building necessary for networks of GOs transcending ecosystem types and scales. GOs have been capitalizing on new technologies of the DNA/RNA sequencing revolution and the subsequent developments in bioinformatics and data science. The integration of additional data types (high-throughput bioimaging, proteomics, metabolomics) in combination with fast improvements and automation of sampling and sequencing methods, are expected to render BCD and machine intelligence the major drivers of GOs towards a real time assessment of ecosystems and the knowledge production for well-informed societies. However, at this point GOs are still in their infancy, currently demonstrating both their huge potential but also important challenges from sampling, sample and data processing to data analysis and interpretation. For instance, integrating all relevant available data reveals emerging needs to be addressed: (a) the harmonization of the protocols and standards currently used to allow machine-to-machine readable data to become discoverable and reusable and (b) the building of top level ontologies (TLOs) in order to establish organic and functional links between different types of data. Although these challenges may take time to be met, in the long run, they will substantially reduce current biodiversity and environmental health assessment and monitoring costs and will increase our effectiveness on knowledge production. IMBBC has initiated the establishment of a Genomic Observatory in the Cretan Sea representing a coastal oligotrophic and highly diverse ecosystem. This is located close to the newly developed underwater biotechnological park of Crete which is a large scale facility of HCMR to support basic and applied research equipped with state of the art sensors for continuous monitoring (ubpcrete.hcmr.gr). Appropriate equipment for atmospheric sampling and meteorological parameters has been also installed to support the Cretan Sea GO.
Environmental Influence on Genetic Diversity Patterns Revealed through ddRAD Sequencing

Katerina Vasileiadou, Tereza Manousaki, Alexandros Tsakogiannis, Costas Tsigenopoulos and Christos Arvanitidis

Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC), Hellenic Centre for Marine Research (HCMR), Gournes Pediados, 71500 Heraklion Crete, Greece

The patterns of SNP clusters can offer a good estimate on the fluctuations of the populations demography and elucidate the role of the environment on these shifts. Lagoons are transitional water ecosystems with variable environmental conditions, spatially and temporally. The Mediterranean lagoons often undergo anoxic conditions followed by mass mortality of the populations inhabiting them. The macrobenthic polychaete Nephtys hombergii is a common species of the Mediterranean lagoons with genetic diversity and great dispersal ability, used in many studies as an indicator of the ecological ecosystem status. The genome-wide genetic diversity of the polychaete N. hombergii from three lagoons of Amvrakikos Gulf (W Greece), was investigated with the ddRAD sequencing method. The study aims to compare the SNP cluster patterns against the patterns of environmental attributes in an attempt to elucidate the mechanisms that influence the demography of lagoonal polychaete populations and contribute to the understanding on the role of the environmental variables to the populations structuring, providing insights for biodiversity and conservation biology.
Bayesian Matrix Factorization with Side Information and Application to Drug-Target Activity Prediction

Yves Moreau
University of Leuven, Belgium

Matrix factorization/completion methods provide an attractive framework to handle sparsely observed data, such as the prediction of biological activity of chemical compounds against drug targets, where only 0.1% to 1% of all compound-target pairs are measured. Matrix factorization searches for latent representations of compounds and targets that allow an optimal reconstruction of the observed measurements. These methods can be further combined with linear regression models to create multitask prediction models. In our case, fingerprints of chemical compounds are used as “side information” to predict target activity. By contrast with classical Quantitative Structure-Activity Relationship (QSAR) models, matrix factorization with side information naturally accommodates the multitask character of compound-target activity prediction. This methodology can be further extended to a fully Bayesian setting to handle uncertainty optimally, which is of great value in this pharmaceutical setting where experiments are costly. We have developed a significant innovation in this setting, which consists in the reformulation of the Gibbs sampler for the Markov Chain Monte Carlo Bayesian inference of the multilinear model of matrix factorization with side information. This reformulation shows that executing the Gibbs sampler only requires performing a sequence of linear regressions with a specific noise injection scheme. This reformulation thus allows scaling up this MCMC scheme to millions of compounds, thousands of targets, and tens of millions of measurements, as demonstrated on a large industrial data set from a pharmaceutical company. We have developed a Python/C++ library, called Macau, implementing this method and which can be applied to many modeling tasks, well beyond our pharmaceutical setting. We discuss the application of our method to drug-target activity prediction using compound structure fingerprints as side information. We also discuss the application of this method to drug-target activity prediction using high-content imaging assays as side information. Our results suggest that high-content imaging assays can be broadly repurposed for drug-target activity prediction and the broad exploration of chemical space.
The Evolution of Ribonucleotide Reduction

Daniel Lundin
SciLifeLab, Sweden

Reduction of ribonucleotides is the only known biological pathway for de novo synthesis of deoxyribonucleotides, the building blocks of DNA. The enzyme catalyzing the reaction – ribonucleotide reductase (RNR) – is hence encoded by virtually all cellular organisms plus many dsDNA viruses. RNRs have diversified into three distinct classes, with distinct mechanisms for generation of the radical that is essential for the reaction. The three classes have different oxygen requirements: strict anoxy, independence and full dependency on oxygen respectively. Whether organisms are obligately or facultatively aerobic or obligately anaerobic is hence reflected in which RNR classes they encode.

We study the evolution of ribonucleotide reduction from several perspectives:

- The origin of the biochemically challenging reaction
- Evolutionary divergence of the family, including origin of the classes with their differences in cofactors as well as later divergence of specialized subclasses
- Ecological distribution of classes and subclasses, including identification of horizontal gene transfer events
- Structural evolution, in particular the dual allosteric regulation of the enzyme

The research is performed as a combination of computational analyses and experiments. The primary type of computational analysis is phylogenetics and sequence searches with HMM profiles. The former poses challenges due to the high degree of divergence between classes (cannot be aligned to each other over a sufficient number of residues for phylogenetics) and within classes for the two most ancient classes. In addition to sequence-based approaches we have therefore explored structural phylogenetics based on distances – with some success – and algorithms combining sequence alignment, structure alignment and tree estimation. Results and hypotheses arrived at computationally are used to identify project ideas for experimental approaches as well as to investigate the generality and distribution of patterns detected in experiments. Recently, several projects have concerned the evolution of overall activity of the enzyme. This is controlled by a separate domain – the ATP cone – that is gained and lost over evolutionary time and appears to be able to add activity regulation without requiring much in terms of adaptation in the rest of the enzyme, in contrast to what might be expected from allosteric regulation.

Sequence databases are regularly searched with a hierarchical, manually curated, set of HMM profiles and collected in a publically available database: [http://rnrdb.dbb.su.se/pfitmap-apps/pfitmap/](http://rnrdb.dbb.su.se/pfitmap-apps/pfitmap/).
An Updated View of the Oligosaccharyltransferase Complex in Plasmodium

Stella Tamana and Vasilis J Promponas

Bioinformatics Research Laboratory, Department of Biological Sciences, University of Cyprus, Cyprus

Despite the controversy regarding the importance of protein N-linked glycosylation in species of the genus Plasmodium, genes potentially encoding core subunits of the oligosaccharyltransferase (OST) complex have already been characterized in completely sequenced genomes of malaria parasites. Nevertheless, the currently established notion is that only 4 out of 8 subunits of the OST complex – which is considered conserved across eukaryotes – are present in Plasmodium species.

In this study, we provide unequivocal evidence that all components of the OST complex (except of Swp1/Ribophorin II) can be reliably identified within completely sequenced plasmodial genomes, based on carefully conducted bioinformatics analysis. In fact, most of the subunits currently considered as absent from Plasmodium refer to uncharacterized protein sequences already existing in the sequence databases. Interestingly, the main reason why the unusually short Ost4 subunit (36 residues long in yeast) has not been identified so far is the failure of gene-prediction pipelines to detect such a short coding sequence.
Investigating the Specialized Roles of Ribosomal Proteins and Associated Ribosomopathies by Means of Literature, Protein Sequence and Phenotype Information Resources

Konstantinos Kyritsis¹, Lefteris Angelis², Christos Ouzounis³ and Ioannis Vizirianakis¹

¹ Department of Pharmacology, School of Pharmacy, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece
² Department of Informatics, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece
³ Biological Computation & Process Laboratory, Chemical Process & Energy Resources Institute, CERTH, GR-57001 Thessaloniki, Greece

Ribosome biogenesis is a highly coordinated cellular process that involves a multitude of macromolecular components, leading to the stoichiometric assembly of the ribosome. In humans, 4 rRNAs and 80 ribosomal proteins (RPs) are expressed, processed and assembled into the large (60S) and small (40S) ribosomal subunits, each of which possess specific functions for the translation of mRNA into protein. Mild disruption of this well-orchestrated process is known to contribute to certain pathological phenotypes. Interestingly, many heterozygous mutations have been recorded in specific RP genes, which result in a reduction of their expression levels (haploinsufficiency). These mutations constitute the cause for the rise of rare congenital diseases, known as ribosomopathies. While most RPs are ubiquitously expressed across all tissues, ribosomopathies present highly specific symptoms in patients, including hematological disorders, bone marrow aplasia and cancer predisposition. The unexpected specificity of phenotypes arising from mis-regulation of RP gene expression strongly implies that these proteins exhibit roles beyond those of structural constituents of the ribosome. Indeed, many recent studies have revealed the involvement of RPs in the regulation of the translational efficiency of distinct mRNAs. These regulatory functions depend on several factors, including the presence of specific cis-regulatory elements within mRNAs (IRES elements), post-translational modifications of RPs, as well as the expression levels and the stoichiometry of structural constituents of ribosome and their paralogs. The present study reports on preparatory work, conducted to study the reported multifunctional roles of RPs in the biomedical literature and their association with disease phenotypes. Specifically, we employed ‘reutils’ R package to perform multiple queries on NCBI’s PubMed database, aiming to analyze the literature coverage of 82 RPs. Notably, the level of literature coverage for RPs ranges widely, as reflected by the number of returned abstract entries, with some RPs having been extensively studied, while others are apparently less well characterized by targeted experimental analysis. Furthermore, pairwise alignments were performed using NCBI’s BLASTP software with default settings, to estimate identity levels between human and mouse ortholog RP sequences. Based on our analysis, a varying degree of conservation was observed for RPs, ranging from 75% to 100% sequence identity. Remarkably, the level of literature coverage for RPs appears to be unrelated to the level of their conservation. Finally, we used NCBI’s PubTator software to retrieve disease-related bioterms, in the form of MeSH IDs, from titles and abstracts of all RP PubMed articles. Next, we associated all RPs with MeSH IDs and PubMed results into a complex network, using the ‘BioLayout’ software suite. Cancer, developmental defects and hematological disorders are the predominant disease phenotypes with many RP connections. Nevertheless, a considerable number of unrelated disease phenotypes was observed, for which the involvement of RP functionality is less clear. Our results constitute a first step towards a deeper understanding of the functionality of RPs and how its disruption results in a range of pathological conditions.
**Evolutionary Models of Amino Acid Substitutions Based on their Neighborhood Tertiary Structure**

**Elias Primetis**\(^1\) and Pavlos Pavlidis\(^3\)

1 Department of Biology, University of Crete, Greece
2 Institute of Computer Science (ICS), Foundation of Research-Technology Hellas (FORTH), Greece

Intra-protein interactions depend on the involved amino acids. It is assumed that the energetically favorable interactions have been preserved, while the unfavorable have been eliminated during evolution. We have used the Protein Interaction Statistics (PrInS) algorithm to statistically describe interactions between amino acids using protein structures. PrInS produces a scoring matrix to describe the frequency of amino acid interactions in the protein structures. In this project, we used structures of alpha helical membrane proteins from the RCSB PDB database (Berman et al. 2000).

The resulting scoring matrix was converted to an amino acids distance (Euclidean, Manhattan or Pearson) matrix \(M\), where \(M_{ij}\) value denotes the distance between the neighborhoods of amino acid \(i\) and \(j\). To test the validity of our methodology, we counted the observed number of amino acid changes in 224 alignments of homologous proteins and we correlated them with the \(M\) matrix. Assuming human protein sequences as a reference, we calculated the distance between human and 19 other species (16 primates and 3 other mammals) for all 224 proteins of our dataset, using our approach, BLOSUM62 and PAM120 amino acid substitution matrices. Outcomes were comparable, suggesting that our approach captures information about protein evolution process in a similar fashion as BLOSUM62 and PAM120. Finally, distance matrices were converted to rate matrices to calculate the likelihood of multiple alignments and the likelihood of each site in alignments. For the sites that our approach produces a better likelihood we described their location on the tertiary structure of the protein.

References
What is Life? Conserved Functional Domains in a Hundred Thousand Bacterial Genomes

David Ussery

University of Arkansas, USA

We have examined the conservation of more than 16,000 PfamA domains across more than a hundred-thousand bacterial genomes, and find a set of about 450 domains are found in nearly all (~99%) of the genomes. These functional domains appear to contain the basic essential functions for life as a bacterium, including basic processes such as protein synthesis, DNA and RNA replication, as well as basic metabolism.
Comparative Genomics of the Sigatoka Disease Complex on Banana Indicates a Direct Link between Pathogen Emergence and Nutritional Virulence

Ioannis Stergiopoulos¹, Ti-Cheng Chang¹, Anthony Salvucci¹ and Pedro Crous²

¹ Department of Plant Pathology, University of California Davis, USA.
² Westerdijk Fungal Biodiversity Institute, Netherlands

Understanding the evolutionary and genomic changes involved in the emergence of new pathogens and shifts in their virulence spectra after speciation is critical. Such knowledge is vital for deciphering the biological process of disease and for designing new disease control methods. In order to understand the evolutionary trends and genomic modifications associated with speciation and virulence-jumps in fungi, we have sequenced the 82.8 Mb and 53.8 Mb genomes of Pseudocercospora musae and P. eumusae, respectively and compared them with the 74.1 Mb genome sequence of P. fijiensis. The three species constitute the Sigatoka disease complex of banana, currently the most destructive disease in this crop. However, despite their common ancestry and shared host-specificity, they show clear differences in virulence with P. fijiensis and P. eumusae being the most aggressive. Our comparative genomic and evolutionary analyses revealed that changes in gene family sizes among the three species are not selectively neutral but more respectful of the species virulence profiles rather than their evolutionary relationships. Specifically, P. eumusae and P. fijiensis share convergent patterns of expansions and contractions in core gene families related to metabolism and degradation of plant cell walls, suggesting that virulence-jumps and speciation in these fungi are to a certain extent linked to recurrent genomic changes in molecular pathways associated with nutrient acquisition and assimilation.
Invisible Cities. Segregated Domains in the Yeast Genome with Distinct Structural and Functional Attributes

Christoforos Nikolaou

Department of Biology, University of Crete, Greece

In a recent work we argued for the existence of a functional compartmentalization of the yeast genome, according to which genes occupying the chromosomal regions at the nuclear periphery have distinct structural, functional and evolutionary characteristics compared to their centromeric-proximal counterparts. At the same time, it was recently shown that even the small and gene dense genome of *S. cerevisiae* is organized in Topologically-associated Domains (TADs), which are associated to the replication timing of different genomic regions. In this work we went on to investigate whether such units of three-dimensional genomic architecture reflect a particular organization in terms of structure and function. Through the application of a simple boundary-calling criterion in genome-wide 3C data we define ~100 TAD-like domains which can be clustered in 6 different classes with radically different nucleosome organization, transcription factor binding site enrichment and chromosomal distribution. Around ~20% of the genome is found to be confined in regions with “closed” chromatin structures around the gene promoters. Most interestingly we find these regions to be segregated in the sense that they tend to avoid inter-chromosomal interactions. Our data further enforce the notion of a marked compartmentalization of the yeast genome in isolated territories, with implications in its function and evolution.
Ancestral Pathway Reconstruction

Christos Ouzounis

Biological Computation & Process Laboratory, Chemical Process & Energy Resources Institute, CERTH, GR-57001 Thessaloniki, Greece

Ancestral reconstruction is an essential element for modern genomics, bioinformatics and synthetic biology. While sophisticated methods for ancestral sequence inference exist and methods for estimating the ancestral gene content of entire taxa have also been developed, the detection and subsequent computation of ancestral metabolic pathways is still in its infancy. We describe PathTrace, an algorithm for the reconstruction of the evolutionary history of individual pathways. The algorithm is a five-step process through which pathways are represented as fuzzy vectors, reminiscent of gene phylogenetic profiles for the corresponding enzymes in a pathway. The method is validated with a carefully selected benchmark set of pathways against collections of genome sequences from key data resources and is expected to find uses in the analysis of pathway evolution and the design of novel, synthetic pathway modules.
Genomic Encyclopedia of Bacteria and Archaea

Nikos Kyrpides

DOE Joint Genome Institute, USA

Sequencing of bacterial and archaeal genomes has revolutionized our understanding of the many roles played by microorganisms. So far, bacterial and archaeal genomes available, most of which were chosen for sequencing on the basis of their physiology. As a result, the perspective provided by the available genomes is limited by a highly biased phylogenetic distribution. To explore the value added by choosing microbial genomes for sequencing based on their evolutionary relationships, we have sequenced and analysed the genomes of 1003 culturable species of Bacteria and Archaea selected to maximize phylogenetic coverage. These genomes double the number of existing type strains and expand their overall phylogenetic diversity by 25%. Comparative analyses with previously available finished and draft genomes reveal a 10.5% increase in novel protein families as a function of phylogenetic diversity. The GEBA genomes recruit 25 million previously unassigned metagenomic proteins from 4,650 samples, improving their phylogenetic and functional interpretation. We identify numerous biosynthetic clusters and experimentally validate a divergent phenazine cluster with potential new chemical structure and antimicrobial activity. This Resource is the largest single release of reference genomes to date. Bacterial and archaeal isolate sequence space is still far from saturated, and future endeavors in this direction will continue to be a valuable resource for scientific discovery.

Reference:

PROGRAMME
LIST
Automatic annotation of text is an important complement to manual annotation, because the latter is highly labour intensive. We have developed a fast dictionary-based named entity recognition (NER) system and addressed a wide variety of biomedical problems by applying it to text from many different sources. We have used this tagger both in real-time tools to support curation efforts and in pipelines for populating databases through bulk processing of entire Medline, the open-access subset of PubMed Central, NIH grant abstracts, FDA drug labels, electronic health records, and the Encyclopedia of Life. Despite the simplicity of the approach, it typically achieves 80–90% precision and 70–80% recall. The tagger software and dictionaries are available under open licenses, and pre-computed results for entire Medline are provided through web resources, REST APIs, and bulk download files.
Latest Improvements in ProteoSign, a Web Application for Differential Proteomics Analysis

Georgios Efstatthiou\textsuperscript{1,2}, \textbf{Andreas N. Antonakis}\textsuperscript{1}, Georgios A. Pavlopoulos\textsuperscript{3}, Theodosios Theodosiou\textsuperscript{1}, Peter Divanach\textsuperscript{4}, David C. Trudgian\textsuperscript{5}, Benjamin Thomas\textsuperscript{2}, Nikolas Papanikolaou\textsuperscript{1}, Michalis Aivaliotis\textsuperscript{4}, Oreste Acuto\textsuperscript{2} and Ioannis Iliopoulos\textsuperscript{1}

\textsuperscript{1} Medical School, Division of Basic Sciences, University of Crete, Greece
\textsuperscript{2} Sir William Dunn School of Pathology, University of Oxford, UK
\textsuperscript{3} Lawrence Berkeley National Labs, DOE Joint Genome Institute, USA
\textsuperscript{4} Institute of Molecular Biology and Biotechnology, FORTH, Greece
\textsuperscript{5} Department of Bioinformatics, UT Southwestern Medical Center, Dallas, USA

Bottom up mass spectrometry has been proven to be the most efficient tool to identify proteins in large scale, however this method produces large data files that are difficult to process and analyse. Our team recently developed ProteoSign, a post-quantification differential proteomics analysis program which targets the end-user. ProteoSign accepts as input quantified data from two commonly used proteomics quantification programs (MaxQuant and Proteome Discoverer) from both labelled and label-free experiments, statistically analyses them using the Linear Models For Microarray Data methodology and generates data tables and high quality data plots. In comparison to other existing tools that require advanced computer or even programming skills, ProteoSign allows users to perform abundance analysis requiring minimal user interaction. It is accessible through a user-friendly web application at \url{http://bioinformatics.med.uoc.gr/ProteoSign}. By utilizing a client-server architecture, the analyses are performed in a resource-efficient manner. Our team is currently working on improving ProteoSign by adding support to time-series proteomics experiments and complicated experimental structures. Additionally, we are creating a REST API to make our functions accessible to other programs.
Calchas: A Semantic Search Engine for Bioinformatics Resources

**Lefteris Koumakis**, Mihalis Giannoulis, Haridimos Kondylakis, Alexandros Kanterakis and George Potamias

*Computational BioMedicine Laboratory (CBML), Institute of Computer Science (ICS), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Greece*

Over the last years we are experiencing an explosion of biomedical research tools, data formats and analysis methods. While the global scientific output is doubling every nine years, researchers still use generic Google-like searches in order to locate useful tools or rely on specialized web forums to seek technical and analytical advices. It is not an exaggeration to state that science on that matter has not changed over the last 20 years. This is surprising given that the results of a plethora of published papers in experimental science are generated through an analysis pipeline of some kind.

We introduce Calchas (calchas.ics.forth.gr), a web based framework that takes advantage of domain specific ontologies (e.g. the EDAM ontology is utilized edamontology.org), and Natural Language Processing (NLP), aiming to empower exploration of biomedical resources via semantic-based querying and search. The NLP engine analyzes the input free-text query (description of the facts/data, the research question and the desired output) and translates it into targeted queries with terms from the underlying ontology. Each query is passed to the semantically-annotated tools repository, and based on similarity matches, it ranks the available resources. The current version of Calchas is an improvement, both in terms of the included resources and functionality, of an older version of it (Sfakianaki et al., 2015).

In addition, Calchas features a rich GUI to present search results as a network of tools and input/output entries. The recommended tools are mapped on a directed graph where each node represents a data type or data format and each edge a tool that supports as input and output the source and target node, respectively. Different edge (tool) colors indicate whether the tool supports both the requested input and output (green), only the input (blue) or only the output (orange). Such a visualization provides a direct overview of putative pipelines that could serve the research quest, based on the identified tools and the engaged input/output formats. Other features include interaction with the network allowing to explore descriptions of the tools and matching scores of the respective nodes, and edges as well.

Calchas tries to link the gap between research question and efficient dynamic biomedical resources discovery. Currently, Calchas supports more than 6000 tools from the bioinformatics domain, as managed by the Elixir tools and data services registry (https://bio.tools). Work is also in progress to include resources from the FAIRDOM repository (https://fair-dom.org) aiming to offer a FAIR-compliant resource discovery service in the life sciences domain.
Large-Scale Study on Latent Feature Construction for Gene Expression with Improved Predictive Power on Newly-Seen Datasets

Christos Tselas, Yannis Pantazis, Kleanthi Lakiotaki and Ioannis Tsamardinos

University of Crete, Greece

Gene expression analysis aims to improve the understanding of the intrinsic cellular processes and contribute towards the successful implementation of personalized medicine. The advent of high-throughput gene expression technologies such as microarrays and RNA-sequencing (RNAseq) as well as the recent reduction of cost resulted in an explosion of publicly-available datasets. The generated datasets are inevitably high-dimensional with typically small sample size that severely limits the potential for developing reproducible prognostic models. Being able to increase the predictive power without losing the information of the measured genome on a newly-produced dataset is of paramount importance. Despite the fact that various studies attempt to perform dimensionality reduction and dataset integration so as to increase classification performance and robustness, there are still challenging issues due to the limited number of data as well as the technological diversity and heterogeneity across the datasets. Exploiting the redundancy of genomics data, we constructed low-dimensional, universal latent feature spaces of the genome utilizing several dimensionality reduction approaches and a diverse set of curated datasets. Standard PCA, kernel PCA and Neural Network Autoencoder were applied on datasets from four different platforms. While linear techniques show better reconstruction performance, nonlinear approaches capture more complex gene interactions, and thus enjoy stronger classification performance. When newly-seen microarray datasets projected to a latent space of only 200 dimensions, the classification power was improved. Moreover, we performed a large-scale experiment where we trained the dimensionality reduction methods on an integrated set of 59864 unique samples. The classification power was further improved especially with Autoencoder showing, rather surprisingly, that the statistical variability of the additional datasets increases the performance implying that more intricate biological features were better learnt. We additionally test the possibility of cross-platform data augmentation by constructing an intermediate feature space showing that when platforms share common characteristics (such as GLP570 and GLP96) the predictive performance is also improved.
Arkalos: An open collaborative Workflow Management System for reproducible bioinformatics research

Alexandros Kanterakis¹, Lefteris Koumakis¹, Nikos Karacapilidis² and George Potamias¹

¹Computational BioMedicine Laboratory (CBML), Institute of Computers Science (ICS), Foundation for Research and Technology-Hellas (FORTH), Heraklion, Crete, Greece
²University of Patras, Department of Mechanical Engineering, Rio, Patras, Greece

Today there are more than 50 Workflow Management Systems (WMS) in the bioinformatics domain [1]. The landmark objective of these systems is to enable and serve reproducible science. Although the system requirements, design and objectives of each system differ, we can pinpoint a general set of drawbacks that hinder significantly their adoption mainly from young and IT inexperienced researchers. The most important is that none of these environments includes an active, well-indexed and well-annotated repository of tools and pipelines. Despite the existence of such repositories [2], the rich information that they include has not been incorporated properly in modern WMSs. Namely, modern WMS environments provide auxiliary gluing functions for bioinformatics tools but the process of locating, downloading, installing and configuring of these tools is still left to the researcher. Additionally, we are experiencing an explosion of online services that promise to enhance scientific collaboration through various ways [3]. Examples are sites that offer social media-like services (i.e. www.academia.edu), sites for collaborative scientific writing (i.e. www.overleaf.com) and sites for sharing of data analysis (i.e. www.myexperiment.org). Yet the process of locating and contacting a potential skilled collaborator, building strong research teams, sharing research protocols and debating scientific ideas, hypotheses and findings is still based on traditional person to person communication. Here we present the initial design and prototype of the Arkalos platform. Arkalos is an online Research Object management service. A Research Object [4] can be any of a scientific tool, data, workflow, citation, user, result, hypothesis and ontology. Each Research Object can be executable, citable, sharable, findable, commentable, rateable, import/export-able and linkable. These features not only support but also extend the prominent FAIR principles for open and reproducible science [5]. Execution can happen to any computational environment ranging from a local computer to user defined High Performance Computational environments (i.e. cluster, grid, cloud) through the Docker virtualization technology. Arkalos also includes a graphic UI environment for the composition of workflows and a citation management system. Additionally, Arkalos will be equipped with an AI assisted collaborative environment that consults users on proper tool selection, manages expert opinions and suggests potential collaborators according to user’s profiles and activities, enabling thus a set of microattribution services [6]. Initially Arkalos is populated with a collection of tools, data and pipelines in the area of population genetics research.


References

PROGRAMME LIST
Space Technology and Space Biology

Periklis Papadopoulos
San Jose State University, USA

The US space program and NASA have established space biology as key for human space flight and exploration of our solar system. The main goals in support of this objective include (a) understanding the adaptation of the basic biological processes in micro-gravity environments, (b) biological effects of humans exposed to the harsh space environments as it affects space exploration, and (c) application of related technologies to improve our day-to-day lives on Earth.

The future for Human Space Exploration and the proposed Life and Physical Sciences research is documented in NASA’s New Era decadal survey. The survey establishes an integrated framework of biological and physical sciences research in the decade of 2010-2020. The current presentation will provide specific examples on how mission objectives are tightly integrated with the biological decadal scope. Representative space missions and designs will be presented to showcase biological payloads and their retrieval from the international space station using micro reentry systems, and space biological experiments performed on small form factor small sat systems.

In addition, examples on how the space community is addressing micro-gravity effects on biology and human physiological effects will be presented. Finally the role of space biology for the planned manned mission to Mars will be presented.
POSTER PRESENTATIONS
**Genome-Wide Localization and Inter-Dependence of Aft1 and Mac1 Transcription Factors In S. Cerevisiae**

Konstantinos Barsakis\(^1\), Konstantina Koukourikou\(^2\), Antonis Klonizakis\(^2\), George Fragiadakis\(^3\), Farbod Babrzadeh\(^4\), Michael Mindrinos\(^4\), Despina Alexandraki\(^3\) and Christoforos Nikolaou\(^2\)

\(^1\) Stanford Blood Center, Stanford School of Medicine, USA
\(^2\) Department of Biology, University of Crete, Greece
\(^3\) Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology
\(^4\) Stanford Genome Technology Center, Stanford University

The yeast Aft1 and Mac1 proteins are iron and copper responding respectively. They bind to DNA and activate the transcription of genes, responsible for Fe/Cu homeostasis. They are additionally localized on various genomic loci, performing known and unknown functions in chromatin. Aft1 is involved in the proper chromosome segregation, interacting with centromeric proteins, and in genome surveillance of fragile sites interacting with the ScRad9 DNA damage response mediator.

We have previously noted Aft1 and Mac1 co-localization on several promoters specific for either one. That led us to examine their possible interdependence by analyzing the genome-wide localization of each one in the presence and at the absence of the other. We performed ChIPSeq analysis for Aft1-9Myc and Mac1-9Myc in both wild-type and mac1\(^\Delta\)/aft1\(^\Delta\) mutants respectively, under Fe/Cu starvation conditions, in rich medium, to achieve maximum Aft1 and Mac1 functions in the nucleus.

We found a large number of genomic site being bound by Aft1 (>700) and Mac1 (>1500) and a highly significant degree of co-localization (>480 sites). More interestingly, we observed marked differences in the binding patterns of both Aft1 and Mac1 in the mutant strains, with a general shift towards the 3' end of genes, that was more pronounced in the case of Aft1/mac1\(^\Delta\). Our findings suggest a co-operative function for these two factors, that goes beyond their standard gene regulatory role.
Investigating the Role Of VIP+ Interneurons In Learning-Related Place Cell Dynamics In Hippocampal Area CA1

Spyridon Chavlis¹, Panagiotis Bozelos¹, Wenke Li², Gergely Turi², Attila Losonczy² and Panayiota Poirazi¹

¹ Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, Greece
² Department of Neuroscience, Columbia University Medical Center, USA

Decades of accruing evidence have landed strong credence to the idea of the hippocampus being an indispensable part of the episodic memory formation, storage and retrieval processes. Specifically, in the CA1 subregion a substantial percentage of the pyramidal neurons, now called place cells, exhibit a firing pattern strongly modulated by environment location, thus acting as neural anchors of spatial memories. However, in the brain, no player is acting alone, and it is no surprise that pyramidal cells are tuned by a wide variety of local-network interneurons. Among those, interneuron-targeting, disinhibitory interneurons, expressing vasoactive intestinal polipeptide (VIP) have recently been proposed to play an important role in mice spatial learning behavior during a goal oriented task. It is hypothesized that one mechanism by which VIP+ cells affect spatial learning is the alterations in place field properties.

To examine this hypothesis we developed a biophysically constrained network model of the CA1 region that consists of 100 cells. More specific, the network includes 80 pyramidal cells and 20 interneurons. Specifically, there are five classes of interneurons, namely basket, bistratified, axoaxonic, OLM and VIP+. All neuron models were validated against experimental data regarding basic electrophysiological, connectivity and input properties. To simulate place cell formation in the network model, we generated grid cell input from the Entorhinal Cortex (EC) and the CA3 regions, activated in a realistic manner as observed when an animal transverses a linear track. Some of the data used to simulate the grid-like inputs are taken from in vivo experiment, such as the animal’s speed an the path it follows. Realistic place fields emerged in a subpopulation of pyramidal neurons (10-20%), in which similar EC and CA3 grid cell inputs converged onto distal/proximal apical and basal dendrites. The tuning properties of these cells are very similar to the ones observed experimentally in awake, behaving animals.

Ongoing work aims to assess the role of VIP+ interneurons in the formation and/or tuning properties of place fields. Towards this goal, we will selectively remove connections for VIP+ cells onto basket cells and VIP+ cells onto OLM cells, as well as we will lesion the whole VIP+ population. Given the lack of experimental data on the precise role of VIP+ cells in spatial memory, our modeling manipulations will provide new predictions as to the mechanistic effects of these neurons at the cellular level. These predictions can in turn guide experimental testing that will ultimately reveal whether VIP+ cells contribute to the formation and learning related reorganization of place cells via their disinhibitory effects on somatic and/or dendritic inhibition.
**The Centrality Lethality Rule in Signed Protein Interaction Networks**

Savas Paragamian¹, Christoforos Nikolaou², Stefanos Sgardelis³ and Ioannis Antoniou¹

¹ Department of Mathematics, Aristotle University of Thessaloniki, Greece
² Department of Biology, University of Crete, Greece
³ Department of Ecology, Aristotle University of Thessaloniki

Essential are the genes/proteins which are indispensable for the organisms. A lot of research has focused on the identification of essential genes/proteins because they are considered part of the minimal gene set; they are possible drug targets for pathogens and more knowledge about them will contribute to the improvement of therapeutic strategies for human diseases. The experimental procedures for the detection of essential genes are expensive, laborious and in most cases unfeasible. Hence scientists have created tools for their prediction from other data using computational approaches. The most important results have come from centrality indices in protein interaction networks which formed the centrality - lethality rule. According to centrality - lethality rule the higher the interactions of a protein the more likely for it to be essential. Since its introduction, this rule has been expanded to other centralities and many novel methods have been developed that integrate a variety of data. Despite all these advancements, the protein - protein interaction network has largely remained the same. For a better representation of protein interactions, additional information should be taken into consideration like activation/inhibition, direction and molecular function. In this work, we used the first large scale signed protein interaction network, which was constructed using protein interaction and RNAi screen data for *D. melanogaster*, to predict essential protein using centrality indices. This revealed that when a protein has many activation interactions it is more likely to be essential. In addition, we found that the interactions between essential proteins are only activation interactions.
A Tale of Two Trees: Modeling Apical and Basal Tree Contribution to L2/3 V1 Pyramidal Cell Orientation Selectivity

Konstantinos-Evangelos Petousakis\textsuperscript{1,2}, Athanasia Papoutsi\textsuperscript{1} and Panayiota Poirazi\textsuperscript{1}

\textsuperscript{1} IMBB, FORTH, Heraklion, Crete, 70013, Greece
\textsuperscript{2} Brain and Mind Postgraduate Programme, Heraklion, Crete, 70013, Greece

Pyramidal neurons, a mainstay of cortical regions, receive a plethora of inputs from various areas. Afferent synapses are received by either the apical or basal dendritic trees, which are morphologically distinct. Both trees contribute in different ways to the somatic response, although their exact roles remain unclear. Inputs to apical dendrites are integrated en-masse at the apical trunk and propagate to the soma. Basal dendrites, on the other hand, branch out from the soma, with inputs being integrated semi-independently. Thus, these trees define distinct anatomical and possibly functional sub-units. To assess the latter, we modeled the complex response pattern of the L2/3 V1 pyramidal neuron to spatially tuned synaptic input. Our goal was to elucidate the contribution of each tree to the response pattern of the neuron, namely its orientation tuning curve. Towards this goal, we created a morphologically detailed computational model of one such cell in the NEURON simulation environment. The model was validated using electrophysiological data recorded in vitro and in vivo. We investigated the role of dendritic integration at the basal and apical trees, and how it shaped the response of the cell. Results show that the basal tree exhibits extensive dendritic spiking, which is primarily caused by background synapses, rather than stimulus-driven ones. In contrast, apical dendritic spiking activity is solely stimulus-driven. Finally, somatic action potentials are generated only when input coincides bilaterally, with single unilateral stimuli unable to evoke spiking activity at the soma. This model provides evidence for distinct computations occurring in the basal and apical trees of the neuron.
Forward-Backward Spatial Simulator for Genetic Data

Aggelos Koropoulis\textsuperscript{1,2} and Pavlos Pavlidis\textsuperscript{1}

\textsuperscript{1} FORTH-ICS, Greece
\textsuperscript{2} University of Crete, Greece

We have implemented an open source software, named sps, which can be used for obtaining samples from a population that lives in a complex environment, and enable us to study the statistical properties of samples and perform hypothesis testing (e.g. neutrality vs selection) or parameter inference. Different regions of the environment may be characterized by varying carrying capacity, inflow and outflow migration rates. sps efficiently simulates genetic polymorphisms using a forward-backward in time approach. First, forward in time we generate the evolutionary history of the population, that is all birth and migration events. After initial colonization of a region with one population, stochastic birth and migration takes place in each generation. Generations are assumed to be discrete. The population size in a region follows a discrete logistic function. Migration may take place between neighboring demes. In addition to the independent migration model, we assume a mass-migration model, where the majority of a population inhabiting a deme, or a number of demes, will migrate as a flock to another region, thus implementing a “follow-the-leader” model. A leader, which we consider as the first migrant to leave a region, will set a migration direction for the population. After the forward-in-time process ends, we obtain a sample from the present day population and we track its genealogy backwards in time until its most recent common ancestor. During the backward process we also generate recombination events. Finally, mutations are placed on the branches of the genealogy following a Poisson process. The output of sps comprises a polymorphic table to present all mutations inherited to the present.
Evolution of Gene Regulatory Networks by Means of Genetic Drift and Natural Selection

Antonios Kioukis and Pavlos Pavlidis

FORTH-ICS, Greece

The evolution of a population by means of genetic drift and natural selection that act on its gene regulatory networks (GRNs) has not been studied in depth. Thus, the relative importance of genetic drift and natural selection on shaping genetic variability has not been examined. Furthermore, it is not known if existing tools used to detect strong and positive selection from genomic sequences in simple models of evolution can be used to detect recent selection when it operates on GRNs.

Here, we propose a simulation framework, called EVONET, that simulates forward-in-time the evolution of GRNs in a population. Since the population size is finite, random genetic drift is explicitly modelled. The fitness of a mutation is not constant. In contrast, we evaluate the fitness of each individual by measuring its genotypic distance from an optimal genotype. Each generation is initialized from the previous generation. Mutations and recombination may take place modifying the genotypic composition of the population. Each individual goes through a maturation period, where its GRN reaches an equilibrium. At the next step individuals compete to produce the next generation. As time progresses, the beneficial genotypes push the population higher on the fitness landscape. We examine properties of the GRN evolution such as robustness against the deleterious effect of mutations and the role of genetic drift. We confirm classical results from Andreas Wagner’s work that GRNs show robustness against mutations and we provide new results regarding the interplay between random genetic drift and natural selection.
Text Mining the Scientific Literature on Research Integrity and Misconduct

Nikolaos Papanikolaou¹, Anastasia Gioti¹, Eirini Sifaki²,³, Ioannis Iliopoulos¹ and Vassiliki Petoussi²

¹ Department of Medicine, University of Crete, Voutes University Campus, 70013 Heraklion, Greece.
² Department of Sociology, University of Crete, Gallos Campus, 74100 Rethymnon, Greece.
³ Hellenic Open University, Greece.

Research integrity and research misconduct are emerging themes in academia, with increasing awareness on their societal repercussions. The H2020 project DEFORM aims to define the global and financial impact of research misconduct. In this context, it is necessary to quantitatively depict the extent and trends of the scientific discussion on issues of research misconduct and research integrity. Scholarly publications treating these issues are dispersed in several databases, each with its own limitations in accessibility and search options.

Herein, we describe the methodology and first results from a systematic collection of the publicly available literature on research misconduct and research integrity. We automatically gathered articles, reviews and editorial letters from three public databases: Medline, Web of Science and Scopus. The data was collected by data mining tools adapted to each database’s particularities. An array of ~100 keywords or their combination was searched within the titles, abstracts and keywords of all publications, and the relevant publications from the three databases were combined into a unique dataset of ~20,000 publications. Further filtering of this dataset was based on feedback provided by a team of experts in sociology, and led to a more restricted set of ~3,700 most relevant publications. Each publication is associated with full information (abstract, year, scientific field, authors, geographical location of publisher etc.), allowing dynamic investigation of the major trends in scientific discourse on research integrity and misconduct. These data are currently formatted for integration into a comprehensive database, expected to represent a valuable online resource for social scientists.

This work is based upon work supported by the DEFORM research project financed by the European Commission within the H2020 framework program. Project ID: 710246.
Transcription Factors Binding, Gene Expression and Positive Selection detection on Topologically associated domains and Lamina associated domains in Humans and Mice

Danai Vagiaki\textsuperscript{1}, Elias Primetis\textsuperscript{1}, Anastasios Kollias\textsuperscript{1} and Pavlos Pavlidis\textsuperscript{2}

\textsuperscript{1} University of Crete, Greece
\textsuperscript{2} FORTH-ICS, Greece

Lamina-Associated Domains (LADs) and Topologically-Associated Domains (TADs) represent distinct 3D organizations of the genetic material inside the nucleus. TADs can be described as genomic regions or chromosomal neighborhoods used to summarize the three dimensional nuclear organization of mammalian genomes. In other words, a TAD is a DNA region within which physical interactions between DNA residues occur more frequently. LADs represent another organization of DNA inside the nucleus, consisting of the parts of the chromatin that heavily interacts with lamina at the inner membrane of the nucleus. In this work, we examine the preferences of various transcription factors (TFs) to bind within, outside or in the border proximity of TADs and LADs. Specifically, we compare the binding sites of TFs from the ENCODE Consortium against random points and we perform hypothesis testing and statistical tests for the location of TFs and Histone Modifiers relative to the boundaries of LADs and TADs. Furthermore, we examine the distribution of Transcription Start Sites (TSS) from the GENCODE project within or outside TADs’ and LADs’ boundaries, and we measure the distance of the former from the latter, as well as from the center of each 3D DNA conformation. Simultaneously, we address the type of transcript that arises from each TSS, with our focus being placed on protein coding genes. In addition, we identify the recent evolutionary forces that shape the patterns of polymorphisms in relation to LAD and TAD boundaries within Human and Mice populations. Thus, we provide insights into whether recent, within populations evolution acts on the 3D structures of DNA within the nucleus. Finally, using Dn/Ds measurements we examine whether older selective forces have been applied on genes located within or outside TADs and LADs.
Clinical Relevance of Single-cell Immune Signatures in Cellular Subpopulations

Maria Markaki, Michael Tsagris, Mira Limaci, Georgios Papoutsoglou and Ioannis Tsamardinos

Department of Computer Science, University of Crete, Greece

Algorithms that robustly partition high-parameter single-cell data into phenotypically distinct subpopulations, are key to discovering novel associations with experimental endpoints of interest. Specifically, the abundance and/or functional activity of some cellular subpopulations has been shown to serve as a surrogate marker of disease status, to predict clinical outcome, or enable a better understanding of disease mechanism. Citrus and Phenograph are two recent data-driven approaches for the identification of cell subsets in mass cytometry datasets, without the subjectivity and bias inherent in manual analysis. In this work, we compare clusters from Citrus and Phenograph with a given set of hand-gated populations, using F-measure; performance is compared on a large collection of public mass cytometry datasets, where the network state of intra-cellular signaling proteins of the human immune system is perturbed. We characterize these signaling networks using Single CEll NEtwork Reconstruction sYstem (SCENERY), a web server featuring several standard and advanced network reconstruction implementations specific for cytometry data. Sparse regression models that explicitly account for correlated features are incorporated into the workflow, to reveal the most clinically relevant differences in signaling patterns between subpopulations.

Acknowledgement: Part of research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 617393]
**DNA Damage Response and Zeocin-specific Transcriptional Profiles in S. cerevisiae**

Dimitra Dialynaki\(^1\), George S. Fragiadakis\(^2\), Irini Stratidaki\(^2\), Niki Gounalaki\(^2\), Pantelis Topalis\(^2\) and Despina Alexandraki\(^1,2\)

\(^1\) Department of Biology, University of Crete, Greece  
\(^2\) Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology-HELLAS, Greece

We have recently identified a specific connection between the radiomimetic, antibiotic and antitumor drug Zeocin and the copper/iron homeostasis in yeast. Zeocin, that we used to induce DNA damage response in the cells, additionally induced copper starvation disrupting copper-regulated transcription. We have evidence for a functional interference of Zeocin with the copper-modulated transcription factor Mac1.

In order to examine the overall effects of Zeocin in the cells, we have analyzed the profiles of accumulated RNA using RNA sequencing technology. In parallel to Zeocin, we have examined RNAs from cells grown in other DNA damaging conditions (UV, 4NQO) for comparison. RNA-Seq data have been analyzed with edgeR and deseq2 algorithms (using the chipster platform/tool) and only the genes commonly predicted as differentially expressed (adjusted p-value or FDR < 0.05) by both algorithms survived and kept for downstream analysis.

Our data revealed that, while a number of genes were upregulated in all three DNA damaging conditions (mostly stress related ensuring cellular survival) and some were down regulated (DNA and cell division specific), Zeocin affected distinct biological processes. Mac1/copper deregulation is indeed a Zeocin-specific effect since the most transcriptionally repressed genes are Mac1-regulated and responsible for copper ion reduction and cellular import. Other downregulated genes include those of ribosome biogenesis. Zeocin specific upregulated genes are mostly involved in energy production. More analyses will be presented.
Gene Co-expression Network Analysis on Mantle Cell Lymphoma Transcriptomics Data Sheds Light on Novel Diagnostic, Prognostic and Therapeutic Strategies

Charikleia Seitanidou1,2,#, Georgia Skrempou3,#, Konstantina Psatha1, Anastasios Oulas3, Elias Drakos4, George M. Spyrou3 and Michalis Aivaliotis1

1. IMBB FORTH, Heraklion, Greece
2. Department of Biology, University of Crete, Heraklion, Greece
3. Bioinformatics ERA Chair, Bioinformatics Group, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus
4. School of Medicine, University of Crete, Heraklion, Greece
# Equal contribution

Mantle cell lymphoma (MCL), a rare, currently incurable and aggressive type of B cell lymphoma comprises 6-8% of all non-Hodgkin's lymphomas (NHL) [1]. The t(11;14)(q13;q32) translocation leading to overexpression of Cyclin D1 is the molecular hallmark of MCL, along with other distinctive clinical, biologic, and molecular characteristics [2]. Recent high-throughput approaches applying transcriptome profiling of MCL support the disease genetic landscape exploration involved in lymphoma initiation and progression, and provide a valuable source of information for a better understanding of its pathobiology on molecular level. In this context, the present study is a systemic approach aiming at shedding light on complex and heterogenous MCL’s pathobiology, collecting all the available comparative transcriptomics data on MCL deposited in both ArrayExpress and GEO databases [3, 4]. MCL gene expression datasets were statistically analyzed by proper R scripts in order to find the top differentially expressed genes, for each case under study. Derived gene lists were used as input for co-expression network reconstruction based on well-established network inference methods[5]. Application of network analysis approaches retrieved significantly affected protein network patterns unveiling the most significant gene-gene links related to different stages and grading, revealing potential disease- and stage-related molecular pathways so as to propose potential (repurposed) drug treatments. Taking into consideration MCL high diversity in different levels (diagnosis, treatment, clinical outcome), application of a network-based bioinformatics approachis of high importance for shedding light on the pathobiology of the disease.

References
Data Integration of Single-Cell Mass Cytometry Measurements using Advanced Imputation Methods

Shyam Khadka, Yannis Pantazis and Ioannis Tsamardinos

University of Crete, Greece

Recent technological advances in mass cytometry enable biologists to measure dozens of surface and inner proteins of a cell. In mass as well as in flow single-cell cytometry, major objectives are to measure and then analyze the sub-populations of cells through cluster analysis, infer the network of interactions between proteins and ultimately diagnose health disorders. However, there are inside the cell as well as on the surface of the cell hundreds of proteins that are impossible to measure simultaneously. In this study, we aim to integrate single-cell measurements that contain a partially overlapping set of proteins but came from different experiments. We adopt the missing data perspective and treat the datasets integration problem as an imputation problem. Imputation methods try to estimate both the forward mapping from the complete space of the data to the intrinsic space as well as the inverse mapping; hence, the key assumption for successful imputation is that the intrinsic dimension of the data should be smaller compared to the full dimension. There are several algorithms that perform imputation and we deploy probabilistic PCA as well as neural network autoencoder. To the best of our knowledge, this is the first time deep learning is used for imputation in cytometry. We perform several imputation experiments using the single-cell mass cytometry datasets of Bendall et al. and Bodenmiller et al. We quantify the results using maximum mean discrepancy distance which evaluates the similarity between two sample distributions as well as we employ SPADE which is a clustering analysis tool specifically designed for single-cell cytometry. Preliminary results indicate that the neural network approach outperforms probabilistic PCA when the number of samples in the datasets is large or when the distribution of the samples is multimodal making it a perfect candidate for single-cell cytometry data. Moreover, the initial results from SPADE clustering shows that autoencoder provides clusters that are qualitatively closer to the clusters produced by probabilistic PCA.
Phylogenomic Analyses Following Whole Genome Sequencing Resolves the Phylogenetic Position of Gilthead Seabream (Sparus aurata) within the Tree of Teleost Fishes

Tereza Manousaki\textsuperscript{1}, Quoc Viet Ha\textsuperscript{1}, Luca Bargelloni\textsuperscript{2} and Costas S. Tsigenopoulos\textsuperscript{1}

\textsuperscript{1} Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research, Greece
\textsuperscript{2} Department of Comparative Biomedicine and Food Science, University of Padova

The phylogenetic relationship of Sparidae and other teleost families is a long-standing question with controversial findings coming from multiple small-scale studies. In the genomics era, phylogenomic approaches are being employed to resolve questions on evolutionary relationships among organisms that remain unanswered. The aim of this study was to conduct a phylogenomic analysis among Sparid species and other teleosts and assess robustly the phylogenetic position of the family within the tree of teleosts.

To that end, we sequenced and functionally characterized the genome of Gilthead seabream (Sparus aurata), a Sparid of high economical value and only member of the family with cumbersome information on genetics and general biology. The high quality predicted gene set has been used to establish one-to-one orthology relationships among Gilthead seabream and all model teleost species available in Ensembl to date. The genes with high confidence orthology relationships among all fishes included in the analysis were recruited to build a supermatrix where rigorous phylogenomic analyses were performed. The outcome of the analyses provides a definite answer regarding the position of Sparidae within the tree of teleosts and lays the ground for future comparative genomic analyses on a solid phylogenetic framework.
Phylogeography of the Toxic Cyanobacterium Cylindrospermopsis raciborskii

Manthos Panou and Spyros Gkelis

Department of Botany, School of Biology, Aristotle University of Thessaloniki, GR 541-24 Thessaloniki, Greece

Planktonic Nostocales cyanobacteria represent a challenge for researchers because of the wide range of cyanotoxins they synthesize and their invasive behaviour, which is presumably enhanced by global warming [1]. *Cylindrospermopsis raciborskii* is a cyanobacterial species extensively studied for its toxicity, bloom formation, and invasiveness potential, which may have consequences to public and environmental health [2]. Its current geographical distribution, spanning different climates, suggests that *C. raciborskii* has acquired the status of a cosmopolitan species. Greece is considered as the origin of the expansion of *Cylindrospermopsis raciborskii* in Europe [3]. In the context of the present study, the phylogeography of the strain *Cylindrospermopsis raciborskii* TAU-MAC 14144 isolated from Lake Karla and representing the first *C. raciborskii* strain isolated from Greek freshwaters, was assessed based on five molecular markers (16S rRNA, 16S-23S rRNA ITS, cpcBA-IGS, nifH, rpoC1). *Cylindrospermopsis raciborskii* TAU-MAC 1414 clustered with strains from Mediterranean regions (Italy, Spain), Northern Africa (Tunisia), Northern America, and South America, in contrast to the rest of European strains which were clustered with strains from Asia, Australia, and Central Africa (Uganda). All the markers examined revealed the same pattern, which led us to a dispersal hypothesis with primary center in South America. The hypothesis describes two dispersion routes: one to Northern America then to Northern Africa and from there to Mediterranean regions, and the second one to Central Africa then to Australia and Asia and from there to the rest of Europe. Our findings suggest an intriguing phylogeographic pattern in *C. raciborskii* species, providing new prospects for future microbiological research.

References:
Untangling the Microbial Dark Matter through a Metagenomic Approach: The Case of Etoliko Lagoon

Panagiota Stathopoulou and George Tsiamis

Department of Environmental and Natural Resources Management, University of Patras, Agrinio, Greece, e-mails: gtsiamis@upatras.gr, panayotastathopoulou@gmail.com

Etoliko lagoon is part of a complex wetland in Western Greece that covers an area of 1,700 ha. It has been formed tectonically and is extremely rich in biodiversity. The physical characteristics of the Etoliko lagoon include a permanent thermocline, halocline and the anoxic conditions in the hypolimnion. This anoxic zone extends from 10 m to 28 m and a steep oxycline is evident between 5 m and 10 m. The temperature increases with depth in the anoxic zone from 13°C at 10 m to 14°C at 28 m. Salinity also increases with depth in the anoxic zone from 18.5 ppt at 10m to 26 ppt at 28m. Etoliko lagoon is characterized by high concentrations of methane and H$_2$S and is also unique regarding the distribution of sulfates, which increases towards the bottom of the lagoon while it decreases in all other well-known anoxic basins. In this study, an advanced “omic” approach was utilized to reveal the microbial diversity in this unusual environment. For this purpose a sampling station was chosen (38.4844 , 21.3169) and a surface (2m depth) water sample was collected with a 5L Niskin bottle. The in situ temperature of water was 26.29 °C and the salinity was 15.09 psu.

Water sample was filtered through 0.2-μm membranes under vacuum and total DNA extraction was carried out. The extracted metagenome was analyzed by MiSeq sequencing. The obtained metagenome indicated the dominance of Bacteria and within this domain, a little over than 50 % of the microbial community was comprised of the phylum Proteobacteria. The second most abundant phylum was Bacteroidetes which represented more than 8 % of the sample. The phylum Actinobacteria was also detected with a significant abundance. Lower abundance of Planctomycetes, Firmicutes, Cyanobacteria, Verrucomicrobia, Chlamydiae, Acidobacteria, and Chloroflexi were also detected. Moreover, a large number of sequences (23.91 %) remained taxonomically unassigned. Regarding the functional content of the sequences, 448,568 annotations were related to biological processes, 381,642 annotations to molecular function and 66,003 annotations to cellular components. This study highlights the unique ecology of Etoliko lagoon and allows a better understanding of the Etoliko microbial community.

P.S. was supported by IKY fellowship of excellence for postgraduate studies in Greece-Siemens Program.
AUTHOR INDEX
PARTICIPANT LIST
Michalis AIVALIOTIS • IMBB-FORTH • Heraklion • Greece • aivaliot@imbb.forth.gr
Alexia AKALESTOU • University of Crete • Heraklion • Greece • jecycraft@yahoo.gr
Despina ALEXANDRAKI • University of Crete and IMBB-FORTH • Heraklion • Greece • alexandr@imbb.forth.gr
Andreas ANTONAKIS • University Of Crete • Heraklion • Greece • antonakisandreas@yahoo.gr
Christos ARVANITIDIS • HCMR • Heraklion • Greece • arvanitidis@hcmr.gr
Maria ASTRINAKI • University of Crete • Heraklion • Greece • maryastrinaki@yahoo.gr
Zein Mersini BESHARAT • Sapienza University of Rome • Rome • Italy • zeinmersini.besharat@uniroma1.it
Dimitra BOUROU • University of Crete • Heraklion • Greece • annabd24@yahoo.com
Anastasios CHANALARIS • University of Oxford • Oxford • UK • anastasios.chanalaris@kennedy.ox.ac.uk
Spyridoula CHAROVA • IMBB-FORTH • Heraklion • Greece • charova@imbb.forth.gr
Christina CHATZIPANTSIOU • University of Crete • Heraklion • Greece • chatzipantsiou@gmail.com
Syridon CHAVLIS • IMBB-FORTH • Heraklion • Greece • chavlis.spiros@gmail.com
Daniele DAFFONCHIO • KAUST • • SA • daniele.daffonchio@kaust.edu.sa
Elena DELIGIANNI • IMBB-FORTH • Heraklion • Greece • delena@imbb.forth.gr
Dimitra DIALYNAKI • University of Crete • Heraklion • Greece • d.dialinaki@gmail.com
Peter DIVANACH • University of Crete • Heraklion • Greece • tomodachiami@gmail.com
Andreas DOULIS • HAO-DEMETER • Heraklion • Greece • andreasntoulis@yahoo.gr
Stelios FANOURAKIS • FORTH • Heraklion • Greece • stelios_fanourakis@yahoo.gr
Rob FINN • EBI • Cambridge • UK • rdf@ebi.ac.uk
Maria Pilar FRANCINO • CSISP • Valencia • Spain • mpfrancino@gmail.com
Carolin FRANK • UC Merced • Catheys Valley • USA • cfrank3@ucmerced.edu
Alexandros GALARAS • National and Kapodistrian University of Athens • Athens • Greece • alexgalaras@gmail.com
Anastasia GIOTI • University of Crete • Heraklion • Greece • natassagioti@googlemail.com

Fotios GKOUNTROMICHOS • University of Crete • Heraklion • Greece • fotis.gk@windowlive.com

Ioannis ILIPOULOS • University of Crete • Heraklion • Greece • ioannis@uoc.gr

John IOANNIDIS • Stanford University • Stanford • USA • jioannid@stanford.edu

Lars JENSEN • University of Copenhagen • Copenhagen • Denmark • Lars.Juhl.Jensen@cpr.ku.dk

Dimitris KAFETZOPOULOS • IMBB-FORTH • Heraklion • Greece • kafetzo@imbb.forth.gr

Ilias KAGKOUVARDOS • TUM • Freising • Germany • ilias.lagkouvardos@tum.de

Maria Lydia KALAITZAKI • University of Crete • Heraklion • Greece • lydiakalaitzaki@gmail.com

Despoina KALFAKAKOU • NCSR "Demokritos" • Athens • Greece • dkalfakakou@gmail.com

Maria KALYVA • University of Crete • Heraklion • Greece • mariakalyva96@gmail.com

Alexandros KANTERAKIS • FORTH • Heraklion • Greece • kantale@ics.forth.gr

Ilias KAPPAS • Aristotle University of Thessaloniki • Thessaloniki • Greece • ikappas@bio.auth.gr

Manouela KAPSETAKI • University of Crete • Heraklion • Greece • kapseta@imbb.forth.gr

Christos KARAPIPERIS • Aristotle University of Thessaloniki • Thessaloniki • Greece • ckarapiperis@bio.auth.gr

Peter KARP • SRI • • USA • pkarp@ai.sri.com

Chrysovalantou KATAITZIDOU • University of Crete • Heraklion • Greece • chrysova@gmail.com

Eleni KATSANTONI • BRFAA • Athens • Greece • ekatsantoni@bioacademy.gr

Antionios KIOUKIS • ICS-FORTH • Athens • Greece • antokioukis@gmail.com

Antonis KLONIZAKIS • University of Crete • Heraklion • Greece • bio2265@edu.biology.uoc.gr
Aggelos KOROPOULIS • FORTH-ICS UOC • Heraklion • Greece • koropoulis.a@gmail.com

Georgios KOTOULAS • HCMR • Heraklion • Greece • kotoulas@hcmr.gr

Margarita-Ioanna KOUFAKI • University of Crete • Heraklion • Greece • koufakimargianna@gmail.com

Konstantina KOUKOURIKOU • University of Crete • Heraklion • Greece • bio2126@edu.biology.uoc.gr

Vassiliki Lila KOUMANDOU • Agricultural University of Athens • Athens • Greece • koumandou@aua.gr

Konstantinos KOUNAKIS • IMBB-FORTH • Heraklion • Greece • kkounakis@gmail.com

Yiannis KOURMPETIS • Nestlé Institute of Health Sciences • Lausanne • Switzerland • yiannis.kourmpetis@gmail.com

Jon Bent KRISTOFFERSEN • HCMR • Heraklion • Greece • jonbent@hcmr.gr

Dimitrios KYRIAKIS • University of Crete • Heraklion • Greece • kyriakds@gmail.com

Athina KYRIAZI • University of Crete • Heraklion • Greece • bio2443@edu.biology.uoc.gr

Konstantinos KYRITSIS • Aristotle University of Thessaloniki • Thessaloniki • Greece • kkyritsis@pharm.auth.gr

Nikos KYRPIDES • DOE Joint Genome Institute • Walnut Creek • USA • nckyrpides@lbl.gov

Vincenzo LAGANI • University of Crete • Heraklion • Greece • vlagani@csd.uoc.gr

Kleanthi LAKIOTAKI • University of Crete • Heraklion • Greece • kliolak@gmail.com

Alexios LOUKAS • Aristotle University of Thessaloniki • Thessaloniki • Greece • aloukas@bio.auth.gr

Daniel LUNDIN • SciLifeLab • Stockholm • Sweden • daniel.lundin@scilifelab.se

Maria MALLIAROU • University of Crete • Heraklion • Greece • bio1870@edu.biology.uoc.gr

Tereza MANOUSAKI • HCMR • Heraklion • Greece • tereza@hcmr.gr

Maria MARKAKI • University of Crete • Heraklion • Greece • maria.g.markaki@gmail.com

Hector-Garcia MARTIN • Joint Bioenergy Institute • Emeryville • USA • hgmartin@lbl.gov
Anna MATHIOUDAKI • University of Crete • Heraklion • Greece • annamathioudaki94@gmail.com
Alexios-Fotios MENTIS • Hellenic Pasteur Institute • Athens • Greece • mentisaf@gmail.com
Folker MEYER • Argonne National Laboratory • Argonne • USA • folker@anl.gov
Eleni MINA • Leiden University • Leiden • Netherlands • e.min@lumc.nl
George MINADAKIS • University of Cyprus • Nicosia • Cyprus • georgem@cing.ac.cy
Yves MOREAU • University of Leuven • Leuven • Belgium • moreau@esat.kuleuven.be
Christoforos NIKOLAOU • University of Crete • Heraklion • Greece • nikolaou@uoc.gr
Maria NIKOLOUDAISKI • University of Crete • Heraklion • Greece • marianikoloud@gmail.com
Anastasis OULAS • University of Cyprus • Nicosia • Cyprus • anastasioso@cing.ac.cy
Christos OUZOUNIS • CERTH • Thessaloniki • Greece • ouzounis@certh.gr
Evangelos PAFILIS • HCMR/IMBBC • Heraklion • Crete • pafilis@hcmr.gr
Charalampos PAGONIS • University of Crete • Heraklion • Greece • pagonisbabis@gmail.com
Manthos PANOU • Aristotle University of Thessaloniki • Thessaloniki • Greece • mattrpol@bio.auth.gr
Yannis PANTAZIS • University of Crete • Heraklion • Greece • yannis.pantazis@gmail.com
Panteleimon PANTELIDIS • Aristotle University of Thessaloniki • Thessaloniki • Greece • pan.g.pantelidis@gmail.com
Stefanos PAPADANTONAKIS • University of Crete • Heraklion • Greece • spapadonakis@gmail.com
Periklis PAPADOPOULOS • San Jose State University • San Jose • USA • peripapas@gmail.com
Zacharias PAPADOVASILAKIS • University of Crete • Heraklion • Greece • spapadonakis@gmail.com
Nikolas PAPANIKOLAOU • University of Crete • Heraklion • Greece • papnikol@gmail.com
Savas PARAGAMIAN • Aristotle University of Thessaloniki • Thessaloniki • Greece • s.paragamian@gmail.com

Pavlos PAVLIDIS • University of Crete • Heraklion • Greece • pavlidisp@gmail.com

Konstantinos-Evangelos PTOUSAKIS • IMBB-FORTH • Heraklion • Greece • kepetousakis@gmail.com

Evangelia PETALAKI • EBI • Cambridge • UK • petsalaki@ebi.ac.uk

Panayiotis POIRAZI • IMBB-FORTH • Heraklion • Greece • poirazi@imbb.forth.gr

George POTAMIAS • ICS-FORTH • Heraklion • Greece • potamias@ics.forth.gr

Elias PRIMETIS • University of Crete and IMBB-FORTH • Heraklion • Greece • eliasprimetis93@gmail.com

Andrea PRINCE • IMBB-FORTH • Heraklion • Greece • princ@imbb.forth.gr

Vasilis PROMPONAS • University of Cyprus • Nicosia • Cyprus • vasilis.promponas@gmail.com

Konstantina PSATHA • IMBB-FORTH • Heraklion • Greece • konstantina_psatha@imbb.forth.gr

Nikolaos PSONIS • IMBB-FORTH • Heraklion • Greece • nikos.psonis@gmail.com

Panagiotis SAKAGIANNIS • University of Crete • Heraklion • Greece • bagjohn0@gmail.com

Ioulia SANTI • HCMR • Heraklion • Greece • isanti@hcmr.gr

Ele SARROPOULOU • HCMR • Heraklion • Greece • sarris@hcmr.gr

Lynn SCHRIML • University of Maryland • Heraklion • Greece • lschriml@som.umaryland.edu

Charikleia SEITANIDOU • IMBB-FORTH • Heraklion • Greece • seit.chara@gmail.com

Lefteris SPANOS • IMBB-FORTH • Heraklion • Greece • spanos@imbb.forth.gr

Alexandros STAMATAKIS • HITS gGmbH • Heidelberg • Germany • Alexandros.Stamatakis@h-its.org

Panagiota STATHOPOULOU • University of Patras • Agrinio • Greece • panayotastathopoulou@gmail.com

Ioannis STERGIOPoulos • University of California Davis • Davis • USA • istergiopoulos@ucdavis.edu
Ewa SZCZYREK • University of Warsaw • Warsaw • Poland • eszczurek@gmail.com

Stella TAMANA • University of Cyprus • Nicosia • Cyprus • stellatamana@gmail.com

Theodoros TAMIOLAKIS • University of Crete • Heraklion • Greece • ttamiolakis@gmail.com

Nektarios TAVERNARAKIS • IMBB-FORTH • Heraklion • Greece • tavernarakis@imbb.forth.gr

Evangelos THEODORAKIS • University of Crete • Heraklion • Greece • van.theodorakis@gmail.com

Pantelis TOPALIS • IMBB-FORTH • Heraklion • Greece • topalis@imbb.forth.gr

Ioannis TSAMARDINOS • University of Crete • Heraklion • Greece • tsamard.it@gmail.com

George TSIAMIS • University of Patras • Agrinio • Greece • gtsiamis1@gmail.com

Antastasia TSIOLA • HCMR • Heraklion • Greece • atsiola@hcmr.gr

Maria TSONTAKI • University of Crete • Heraklion • Greece • mariatsontaki@gmail.com

David USSERY • University of Arkansas • Oak Ridge • USA • DWUssery@uams.edu

Danai VAGIAKI • University of Crete • Heraklion • Greece • danaivag@outlook.com

Alfonso VALENCIA • Spanish National Cancer Research Centre • Barcelona • Spain • alfonso.valencia@bsc.es

Kleio Maria VERROU • University of Crete • Heraklion • Greece • kleioverrou@yahoo.com

Ioannis VIZIRIANAKIS • Aristotle University of Thessaloniki • Thessaloniki • Greece • ivizir@pharm.auth.gr

George VLACHAKIS • University of Crete • Heraklion • Greece • ph3310@edu.physics.uoc.gr

Trudy WASSENAAR • UAMS • Arkansas • USA • ph3310@edu.physics.uoc.gr

Drakoulis YANNOUKAKOS • NCSR "Demokritos" • Athens • Greece • yannouka@gmail.com

Sofia ZAFEIROPOULOU • University of Crete • Heraklion • Greece • bio2470@edu.biology.uoc.gr

Eleftheria ZEGGINI • Sanger Center • Cambridge • UK • Eleftheria@sanger.ac.uk
Interlab Ltd. based in Athens has been providing laboratories with consumables and equipment since 1995. We represent a number of different companies such as the ones (but not limited to) seen below.

For more information please contact us on:

info@interlab.gr
21095 32 220
Survivor’s Guide...

❖ A small piece of advice: Better take your POSTER onboard rather than leave it in your luggage; the effort putting your roll under your seat or above your head is worth avoiding the risk of your presentation traveling to some exotic destination on its own...

❖ How to reach FORTH

❖ BY BUS

There are two bus routes to and from FORTH (ITE in Greek) several times per day:

a. From Heraklion city center to FORTH: Line 8, the sign on the front of the bus reads ITE: [http://astiko-irakleiou.gr/](http://astiko-irakleiou.gr/) (select route No 8)

<table>
<thead>
<tr>
<th>Timetable:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Airport - FORTH</td>
<td>FORTH - Airport</td>
</tr>
<tr>
<td>07:20</td>
<td>08:00</td>
</tr>
<tr>
<td>08:30</td>
<td>09:15</td>
</tr>
<tr>
<td>11:30</td>
<td>12:15</td>
</tr>
<tr>
<td>14:00</td>
<td>14:40</td>
</tr>
<tr>
<td>15:30</td>
<td>16:10</td>
</tr>
<tr>
<td>17:20</td>
<td>18:00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Port - FORTH</th>
<th>FORTH - Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>19:40</td>
<td>20:00</td>
</tr>
<tr>
<td>21:15</td>
<td>21:30</td>
</tr>
</tbody>
</table>

* The bus arrives at Astoria bus station (Platia Eleftherias) about 10 minutes later. The bus drops you off at FORTH premises.

b. From Heraklion city center to PAGNI (ΠΑΓΝΗ in Greek, the University Hospital): [http://astiko-irakleiou.gr](http://astiko-irakleiou.gr) (select route No 11)

Line 11, the sign on the front of the bus reads ΠΑΓΝΗ.

This line has a stop approximately 300 meters away from FORTH's entrance. In order to avoid missing the stop, please consult the bus driver by asking for the stop to FORTH. The frequency of this service is approximately every 15 minutes.

One-way tickets can be purchased from most kiosks around the downtown area. The duration of each trip is approximately 40 minutes.
BY TAXI

From Heraklion city center to FORTH:
The most convenient taxi queue is at "Eleftheria's Square" ("Platia Eleftherias" in Greek), very close to the Astoria Hotel. A taxi ride takes about 20-30 minutes and costs approximately €15,00. We suggest that groups of 3-4 share a taxi, as this is the fastest and most convenient option.

If you wish to call for a TAXI, there are two TAXI companies: Ikaros (2810210102) and Candia (2810361362 or 18300)

Arriving at FORTH

On arrival (workshops, tutorials and HB-10), posted labels will guide you to the seminar room and the amphitheater.

PRESENTATIONS

If you are a speaker:
There will be a data projector connected to a PC so kindly prepare your presentation file accordingly. There will be assistance inside the conference room; we prefer everyone to have their presentation on a USB memory stick.

The time dedicated for your speech includes discussion.

You are strongly advised to keep to the time frames strictly, as the Agenda is tightly packed.

If you are presenting a poster:
Poster boards will be located at the ground floor of the building, outside the restaurant area
You are not allowed to use push-pins or any other mount material which could damage boards. Therefore, you should ask for proper mounting material at the secretariat desk. Remember to consult the detailed poster presentation guidelines (see here). POSTERS SHOULD BE PORTRAIT ORIENTED.

Posters should be up for display by Friday, 8 September the latest by 13.30. All posters will remain mounted for the whole duration of the conference.

There is one Poster session: Friday, 8 September @ 16:30. There will be a poster list where you can check your poster number/board.

ASSISTANCE

People from the local laboratory will be of your assistance during the conference for any help you might need. You will be able to locate them very easily since their name badges will be printed in light yellow background.

- There will be Free WiFi everywhere inside the building.

We are all looking forward to welcoming you in Crete!

The Organizers
Central bus stops for FORTH where you can also find tickets to buy. Usually you can also buy tickets at the kiosks close to bus stops and the restaurant of FORTH.
POSTER PREPARATION GUIDELINES

➤ Each author will have a board space of (HxW) 180cm x 96cm (5ft 10in x 3ft 2in), on which to mount the poster. The poster should be designed to summarize current research in graphic forms. Presentations should be self-explanatory so that the author is free to supplement and discuss particular points. For easy identification, provide a poster heading, listing its title and author(s), identical to that on the official programme.

➤ The poster board is double-sided with one presentation on each side. Your assigned number will be posted on the poster board. The boards will be arranged in numerical order outside the conference hall.

➤ Mounting materials will be provided by the Conference Secretariat.

➤ Do not use push-pins or glue

➤ Do not paint or write on the board

➤ Do not print your poster on heavy or tick backing, as it may be difficult to fasten to the board. If you require assistance with mounting or removing your poster, please notify the Conference Secretariat.

➤ Each author is responsible for assembly and removal of his/her own presentation.

➤ Please remove your poster promptly. Materials left on the poster boards after the removal deadline will be taken down. The organizers or the hotel staff has no responsibility of materials that may be lost or damaged.

The poster sessions have a designated time in which the poster presenters are requested to be available at their poster to discuss their research with the conference attendees.
With the kind support of

FORTH
FOUNDATION FOR RESEARCH AND TECHNOLOGY-HELIAΣ

ΠΕΡΙΦΕΡΕΙΑ ΚΡΗΤΗΣ
REGION OF CRETE

hcmr

UNIVERSITY OF CRETE
SCHOOL OF MEDICINE

SafeBlood
BioAnalytica

biogenea™

PHARMACEUTICALS

Interlab

DEFORM

Crete
University
Press

www.cup.gr
CONFERENCE: HELLENIC BIOINFORMATICS 10
6-9 SEPTEMBER 2017 • FORTH | HERAKLION | GREECE

FOUNDATION FOR RESEARCH AND TECHNOLOGY - HELLAS

Bioinformatics and its applications in Health, Biodiversity and the Environment
about Elixir • biodiversity • bioinformatics • bioprospecting • computational biology • enzyme discovery • genomics • metabolic engineering • microbe • pharmacogenomics • phylogeny • structural genomics • synthetic biology • systems biology • text mining • functional proteomics

hands-on workshops on 5/9: Proteomics • RNA-Seq/Dip-Seq
tutorials on 6/9: Metagenomics • Phylogenomics • Data Integrity

Keynote Speakers
• John Ioannidou, USA
• Peter Karp, USA
• Yves Moreau, Belgium
• Periklis Papadopoulos, USA
• Alfonso Valencia, Spain

Invited Speakers
• Daniele Delferro, Saudi Arabia
• Rob Flein, UK
• Maria Pilar Francisco, Spain
• Carola Frank, USA
• Lars Jensen, Denmark
• Daniel Lundin, Sweden
• Hector Garcia Martin, USA
• Periklis Papadopoulos, USA
• Evangelia Petlassi, UK
• Volta Piron, Greece
• Peter Roosma, USA
• Lynn Schindl, USA
• Ewa Szczerba, Poland
• David Uney, USA
• Eleftheria Zergioti, UK
• Ilia Kappas, Greece
• Christos Kamparris, Greece
• Georgiou Goulianas, Greece
• Yiannis Kourmpetis, Switzerland
• Nikos Kypides, USA
• Jurgen Manousakis, Greece
• Christopher Nikolaou, Greece
• Christos Ouzounakis, Greece
• Evangelos Polli, Greece
• Ioannis Pavlidis, Greece
• George Petasis, Greece
• Vasili Pappou, Cyprus
• Alexandros Stamatakis, Germany
• Nikos Tzanis, France
• Paraskevi Tsalapouli, Greece
• Ioannis Tsamardinos, Greece
• George Tzanis, Greece
• Ioannis Viotsas, Greece

Local Co-Organizing Committees
Arvanitis C (GR), Billis K (UK), Chaniotakis A (UK), Elias E (GR), Grigoriadis I (GR), Konstantinou F (GR), Kourtesis G (GR), Oulas A (CY), Papadopoulos E (GR), Pappou A (GR), Petasis G (GR), Theodorou T (GR)

Organizing Committee
• Michalis Arvanitis, Greece (co-chair)
• Sophia Anagnostou, UK
• Christos Arvanitis, Greece
• Ioannis Vlachou, Greece (chair)
• Dimitris Katsikopoulos, Greece