

Advanced Bioinformatics Workshop

Date: Monday, August 20th – Friday, August 24th, 2018

Venue: Adriatico Guest House - Denardo Lecture Hall

International Centre for Theoretical Physics

Trieste, Italy

Course URL: <http://www.codata.org/datatrieste2018>

Material: <https://codata-rda-advanced-bioinformatics-2018.readthedocs.io>

Description

This advanced bioinformatics course will provide an overview of the current status of different NGS workflows (variant calling, RNA-Seq, ChIP-Seq, etc), and combine them with the appropriate Machine Learning and Data Mining approaches. The course will heavily rely on hand-on exercises and tutorials, and attempt to provide a strong foundation on the underlying theory.

Instructors

- **Fotis Psomopoulos**, Institute of Applied Biosciences (INAB | CERTH) / ELIXIR GR
Contact details: [email](#), [website](#), [twitter](#)
- **Amel Ghouila**, Institut Pasteur de Tunis / H3Bionet
Contact details: [email](#), [website](#), [twitter](#)
- **Gabriele Schweikert**, Cyber Valley Initiative, University of Tuebingen, DE / Computational Biology, University of Dundee, UK
Contact details: email, website, twitter

Course Schedule

	Topic
<i>Day 1</i>	
11:30 – 12:30	Experiments: Design and Analysis
14:00 – 15:00	Components of an Experiment. What is a good experiment design?
15:00 – 16:00	Data Distributions and Multiple Hypotheses Adjustment Methods
16:15 – 18:00	Introduction to basic NGS pipelines
<i>Day 2</i>	
09:00 – 10:00	Introduction to basic NGS pipelines
10:00 – 11:00	Short read quality and trimming (part 1)
11:30 – 12:30	Short read quality and trimming (part 2)
14:00 – 15:00	Mapping
15:00 – 16:00	Variant calling (part 1)
16:15 – 18:00	Variant calling (part 2)

<i>Day 3</i>	
09:00 – 10:00	Introduction to DM and ML, Machine Learning basic concepts
10:00 – 11:00	Taxonomy of ML and examples of algorithms
11:30 – 12:30	Applications of ML in Bioinformatics
14:00 – 15:00	Practicing using the built-in R data set iris
15:00 – 16:00	RNASeq analysis using clustering in R
16:15 – 18:00	RNASeq analysis in R to be continued
<i>Day 4</i>	
09:00 – 10:00	Introduction to ChIP-Seq, ATAC-Seq, BS-Seq
10:00 – 11:00	Retrieving and Preprocessing ChIP-Seq Data
11:30 – 12:30	Defining Regions of Interests (part 1)
14:00 – 15:00	Defining Regions of Interests (part 2)
15:00 – 16:00	Differential Shape Analysis (MMDiff2) (part 1)
16:15 – 18:00	Differential Shape Analysis (MMDiff2) (part 2)
<i>Day 5</i>	
09:00 – 10:00	Introduction to Regression
10:00 – 11:00	Hands-on application: regression algorithms - pros and cons
11:30 – 12:30	Closing, Final Remarks, Post-workshop survey