جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Course Syllabus: Contemporary Topics in Bioscience - B 394

Division	Biological and Environmental Sciences & Engineering Division
Course Number	B 394
Course Title	Contemporary Topics in Bioscience
Academic Semester	Spring
Academic Year	2017/2018
Semester Start Date	01/28/2018
Semester End Date	05/24/2018
Class Schedule (Days & Time)	02:30 PM - 04:00 PM | Sun Wed

Instructor(s)

Name	Email	Phone	Office Location	Office Hours
Timothy Ravasi	timothy.ravasi@kaust.edu.sa	+966128082387	4219, 2, Ibn Al-Haytham (bldg. 2)	
Jesper Nils Tegner	jesper.tegner@kaust.edu.sa	+9661280823888082388	4232, 2, Ibn Al-Haytham (bldg. 2)	Thursdays 1-2 PM

Teaching Assistant(s)

Name	Email
Manjula Thimma	Manjula.Thimma@KAUST.EDU.SA

Course Information

Comprehensive Course Description	**COMPUTATIONAL BIOSCIENCE AND MACHINE LEARNING** **Comprehensive Course Description:** The course includes the following modules: **(0)** Historical Overview and Conceptual Positioning. Major events in Biology, Genomics, and Computing: Turing, the Action Potential, Neumann, DNA, Cybernetics, Different modes of Computing, rise of Molecular Biology, Complex Systems, sequencing of the Human Genome, sequence based technology explosion and production of multiple data-types, Systems Biology, high-performance Computing, CRISP/CAS editing, Synthetic Biology, machine intelligence. **(1)** Transcriptomics and Epigenomics. Bioinformatics data processing and pipelines for RNA-seq, DNA methylation, microRNA, ATAC-seq, and Chip-Seq. The lectures cover the conceptual steps in data-processing as well as their respective practical implementation. This type of analysis ends with clean, normalized, and statistical relevant data. (*From Data to Information/Signals*) **(2)** Bioinformatics downstream analysis. This includes differential analysis (time, samples), enrichment, pathways, clustering and PCA. Illustrated using transcriptomics and epigenomics. (*From Information to Knowledge/biological insight*). **(3)** Overview of Available Databases, Resources, and Software. Bioconductor, IPA, cytoscape, NAR databases, tools, and software. (*Practical Resources*) **(4)** Transcriptional Networks. Definition, how to find them (prior knowledge, reverse-engineering), how to analyse them (network analysis), using case studies including E-Coli, Yeast, Cell lines, and primary cells. (*Beyond lists of molecular entities*) **(5)** Concepts, Ideas, and Tools. Multiple testing, cross validation, bootstrapping, normalization, uni/multi variate statistics, networks, supervised and unsupervised learning, Dimension reduction, clustering techniques, PCA, SVM, ICA, MDS, and tSNE. **(***Conceptual Resources and Challenges*) **(6)** Modelling, Learning, and Inference. Motivation, models, inverse problems, ODE/Boolean models, dynamical systems, and uncertainty. Machine Learning & Deep Neural Networks. **(***From Associations to Causality*) **(7)** Model Systems – Examples of integrative systems analysis. Development (Sea Urchin, Davidson), Eukaryotic Cell Cycle (Tyson/Novak), Action Potential (Hodgkin/Huxley), C-Elegance, Genomic Circuits in Stem Cells, Systems Immunology, Systems Medicine (P4 Medicine), Drug Development, Multi-scale Heart Modelling (Noble, VPH). **(***Real world examples and research programmes)* **(8)** Metabolic Modelling and Analysis. Metabolite and protein networks/charts, network analysis, pathways, modelling and flux analysis. **(***Interface between the inside and outside of cells)* **(9)** Single Cell Biology. The emerging field of single cell genomics, single cell sequencing (sc-seq), scRNAseq, CyTof, scATACseq, and analytics of data. **(***Towards precision biology and data-driven predictive analytics)* **(10)** Summary and Outlook. Summary of current and emerging concepts and trends. Data to Models, Data Integration, Multi-scale Biology, Biomedical and Industrial applications, Synthetic Biology, and Artificial Intelligence.
Course Description from Program Guide	
Goals and Objectives	**Goals and Objectives:** *The course provides a broad and practical overview of available techniques and concepts in the area of* ***computational bioscience and machine learning.*** This includes areas referred to as bioinformatics, computational biology, systems biology, systems medicine, mathematical biology, network biology, synthetic biology, data analytics, predictive modelling, computer simulations, learning algorithms, machine learning, and machine intelligence. Topics are selected to be of relevance for the computer scientist, working biologist, computational scientist, and applied investigator (biotechnology and engineering). From the standpoint of biology, the course addresses the analysis of complex biological systems at difference scales, ranging from molecules, cells, organs, to organisms. The **computer scientist/data scientist** will acquire a working knowledge on which techniques from machine learning, network theory, modelling, and statistics are useful and how to apply them in the analysis of biological systems also in the context of drug development and biomedicine. The **biologist** is provided with concepts and practical tools how advance from data to information about significant correlated features hidden in the data and advancing to insights into biological mechanisms. The scope of the course is intentionally broad since the aim is to provide the student with a comprehensive conceptual and in part practical overview of this inter-disciplinary area. The course is relevant for the researchers in life-science in academia or industry, and data-scientists developing new analytic techniques in research or industry. Upon the completion of the course the student can readily advance to further studies in selected sub-areas, which in part are offered in the BESE or CEMSE programs. Alternatively, the student develops the background and ability to collaborate and identify the relevant specialists in the appropriate sub-area. Computational techniques are necessary, not yet sufficient, to understand and analyse data produced from living systems. Eo ipso, or by the same token, complex living systems are a rich source of inspiration for developing new techniques to decipher patterns and causal mechanisms in big data.
Required Knowledge	Knowledge and skills *corresponding to* course B204 Genomics and basic mathematics and statistics facilitates the understanding of the course. In case you dont have these courses please contact the Instructor (Tegner) for an individual assessment. Entry level includes MSc or PhD students. The material will be presented at a basic level, thus setting the stage for later advanced in depths studies of selected topics.

Reference Texts	**Examples of Recommended Books (for reference only):** Selected chapters will be recommended - -**"Handbook of Systems Biology: Concepts and Insights"** Edited by Walhour, Vidal, and Dekker. Academic Press 2013 -**"A first Course in Systems Biology"** Berhard O. Voit. Garland Science, 2013. -**"Computational and Statistical Epigenomics"** Edited by Andrew E. Teschendorff, Springer, 2015. -**"Uncertainty in Biology"** Edited by Liesbet Geris & David Gomez-Cabrero, Springer, 2016. -**"Systems Medicine"** Edited by Ulf Schmitz & Olaf Wolkenhauer, Methods in Molecular Biology, Springer Protocols, Humana Press, 2016. -**"Computational Biomedicine"** Edited by Coveney, Diaz-Zuccarini, Hunter, & Viceconti. Oxford University Press, 2014. -**"Handbook of Statistical Systems Biology"** Edited by Stumpf, Balding, Girolami. Wiley, 2011.
Method of evaluation	**40.00**% - Final exam **30.00**% - Research Project **30.00**% - Midterm exam
Nature of the assignments	**Method of Evaluation:** (grading scheme) 30 % - MidTerm exam 40 % - Final exam 30 % - Research Project **Nature of Assignments:** **MidTerm exam:** Includes the topics covered. **Final exam:** Includes all topics covered in the course. **Research Projects:** Each student will be assigned a hot topic and required to give a 20-minute presentation at the end of the course. Furthermore, a 3-pages word-processed report on the same topic will be required from each student. The report and presentation reviews the area, pose one or several critical topics suitable for research. The student selects one aspect for analysis of based on their critical review & assessment of topics. The analysis is expected to suggest how to make progress on the suggested problem.
Course Policies	**Attendance Policy:** Attendance to the classes is mandatory.
Additional Information	**Webpage:** The course webpage is on the Blackboard system. Any course announcements will be posted on this site. In addition, the lecture notes, problem sets, solutions, and any extra material will be available on the Blackboard.

Tentative Course Schedule *(Time, topic/emphasis & resources)*		
Week	**Lectures**	**Topic**
1	Sun 01/28/2018 Wed 01/31/2018	Historical Overview and Conceptual Positioning
2	Sun 02/04/2018 Wed 02/07/2018	Transcriptomics and Epigenomics
3	Sun 02/11/2018 Wed 02/14/2018	Transcriptomics and Epigenomics
4	Sun 02/18/2018 Wed 02/21/2018	Bioinformatics downstream analysis
5	Sun 02/25/2018 Wed 02/28/2018	Bioinformatics downstream analysis
6	Sun 03/04/2018 Wed 03/07/2018	Overview of Available Databases, Resources, and Software
7	Sun 03/11/2018 Wed 03/14/2018	Transcriptional Networks
8	Sun 03/18/2018 Wed 03/21/2018	Concepts, Ideas, and Tools
9	Sun 03/25/2018 Wed 03/28/2018	Extra Time *Research Project Assignment* **MidTerm Exam**
10	Sun 04/01/2018 Wed 04/04/2018	SPRING BREAK
11	Sun 04/08/2018 Wed 04/11/2018	Modelling, Learning, and Inference
12	Sun 04/15/2018 Wed 04/18/2018	Model Systems – Examples of integrative systems analysis
13	Sun 04/22/2018 Wed 04/25/2018	Metabolic Modelling and Analysis
14	Sun 04/29/2018 Wed 05/02/2018	Single Cell Biology
15	Sun 05/06/2018 Wed 05/09/2018	**Research Project Presentations**
16	Sun 05/13/2018 Wed 05/16/2018	Reserve
17	Sun 05/20/2018 Wed 05/23/2018	Summary and Outlook **Final Exam**
18		

Note

The instructor reserves the right to make changes to this syllabus as necessary.