

# COMPUTATIONAL GENOMICS – BIOL 7210 A – Spring 2018

Professor: **I. King Jordan** and T.A.: **Aroon Chande**

School of Biology, Engineered Biosciences Building (EBB), Office: 2109 Lab: 2200

[king.jordan@biology.gatech.edu](mailto:king.jordan@biology.gatech.edu)   [arch@gatech.edu](mailto:arch@gatech.edu)

**Course summary:** The science of genomics involves the intersection of experimentation and computation. Computers are quite obviously required to handle the massive amount of data produced by genome sequencing projects. More importantly however, genome sequencing efforts yield ‘information’ alone, which can only be converted into ‘knowledge’ through the use of computers. In this class, the students will convert raw genomic information (*i.e.* sequence reads) into knowledge through the use of computational genomics tools and applications. The class will be provided with unassembled genome sequence data from the Centers for Disease Control and Prevention (CDC) and will proceed through five distinct stages of analysis and interpretation of that data: 1-**genome assembly**, 2-**gene prediction**, 3-**functional annotation**, 4-**comparative genomics** and 5-production of a **predictive webserver**. This course will be entirely practical in nature. Students will learn to do the actual work of computational genomics. Expert guest lecturers will be brought in to provide information on state-of-the-art computational genomics tools. Based on this information, other class lectures and their own research, students will be solely responsible for choosing which tools (e.g. programs and/or databases) to use, how to implement them and for producing and thoroughly documenting their final results. All results will be integrated into a publicly available predictive webserver.

This class meets on Tuesdays and Thursdays from 1:30 to 2:45 pm in Klaus Advanced Computing Building 1447. There is no textbook. Required and recommended readings will be made available on the course Wiki page - <http://compgenomics2018.biosci.gatech.edu/> along with any lecture material. Students are required to use online databases and the scientific literature to inform their choice of computational tools to be used. Since there is no textbook and many of the sessions involve class discussion and lab activities rather than formal lecture, attendance and class participation are absolutely mandatory.

## **Evaluation:**

<b>Class participation</b>	<b>12.5%</b>
<b>Group presentation I</b>	<b>12.5%</b>
<b>Group presentation II &amp; Lab</b>	<b>12.5%</b>
<b>Group presentation III &amp; Demo</b>	<b>12.5%</b>
<b>Final Results &amp; Documentation</b>	<b>50%</b>

Class attendance and participation are mandatory. **Class participation** will be judged by the degree to which each student participates in class lectures and discussions (by asking questions, answering questions, offering ideas and opinions), during group presentations (by asking questions during others’ presentations, by engaging the audience during their own presentation, by connecting their presentation to previous class discussions, by working successfully in a small group), and during computer laboratory activities (by performing analyses and working with other students). Students who show up late or miss class will lose 10% of their class participation grade each time.

Each group will give a series presentations and laboratories/demos. **Group presentations and labs/demos** will be judged by the depth of analysis presented, the clarity of presentation, the utility of the exercises, the appropriateness and justification of the choices made, the validity and robustness of the results and the thoroughness of the documentation. In addition to presentations, results and documentation should be presented on the class Wiki site. All student code and analysis contributions must be shared and documented on Github – <https://github.gatech.edu/compgenomics2018/>. Specific requirements for the presentations will be provided during class sessions. Contributions of each individual student to the overall group effort must be meticulously detailed and documented.

Please see [www.honor.gatech.edu](http://www.honor.gatech.edu) for Georgia Tech’s Academic Honor Code, which you are required to uphold.

<b>Date</b>	<b>Topic</b>	<b>Presenter(s)</b>
Tues 1/9/2018	Introduction & Logistics	King Jordan & Aroon Chande Georgia Tech
Thur 1/11/2018	Class groups & Wiki	King Jordan & Aroon Chande Georgia Tech
Tues 1/16/2018	Source Code Management	Aroon Chande Georgia Tech
Thur 1/18/2018	Bioinformatics on the Cloud	King Jordan & Aroon Chande Georgia Tech
Tues 1/23/2018	Project introduction	David Weiss et al Emory
Thur 1/25/2018	Genome Based Methods for Molecular Typing	King Jordan Georgia Tech
Tues 1/30/2018	Modern approaches to genomic epidemiology	Lee Katz CDC
Thur 2/1/2018	Bioinformatics at the CDC	Scott Sammons CDC
Tues 2/6/2018	Genome Assembly Background & Strategy	Students - Genome Assembly Group
Thur 2/8/2018	Computational Phenotyping and Alignment-free methods	Luz Karime Medina and Héctor Espitia-Navarro
Tues 2/13/2018	Genome Assembly Lab & Preliminary Results	Students - Genome Assembly Group
Thur 2/15/2018	Gene Prediction with GeneMark	Mark Borodovsky Georgia Tech
Tues 2/20/2018	Functional Annotation at NCBI	Leonardo Mariño-Ramírez NCBI
Thur 2/22/2018	Gene Prediction Background & Strategy	Students - Gene Prediction Group
Tues 2/27/2018	Functional Annotation & Comparative Genomics	Aroon Chande Georgia Tech
Thur 3/1/2018	Gene Prediction Lab & Preliminary Results	Students - Gene Prediction Group
Tues 3/6/2018	Genome Assembly Results, Protocol & Demo	Students - Genome Assembly Group
Thur 3/8/2018	Functional Annotation Background & Strategy	Students - Functional Annotation Group
Tues 3/13/2018	Antibiotic resistance in <i>Klebsiella spp.</i>	David Weiss et al Emory
Thur 3/15/2018	Functional Annotation Lab & Preliminary Results	Students - Functional Annotation Group
Tues 3/20/2018		
Thur 3/22/2018		
	<b>Spring break</b>	
Tues 3/27/2018	Gene Prediction Results, Protocol & Demo	Students - Gene Prediction Group
Thur 3/29/2018	Comparative Genomics Background & Strategy	Students - Comparative Genomics Group
Tues 4/3/2018	Comparative Genomics Lab & Preliminary Results	Students - Comparative Genomics Group
Thur 4/5/2018	Predictive web servers	Aroon Chande Georgia Tech
Tues 4/10/2018	Functional Annotation Results, Protocol & Demo	Students - Functional Annotation Group
Thur 4/12/2018	Database Background & Strategy	Students - Genome Database Group
Tues 4/17/2018	Database Lab & Preliminary Results	Students - Genome Database Group
Thur 4/19/2018	Comparative Genomics Results, Protocol & Demo	Students - Comparative Genomics Group
Tues 4/24/2018	Final presentation of the Project / Database	Students - Group representatives