Course Description: This course is designed for students interested in learning current techniques for the analysis of large-scale genomic data sets. High-throughput sequencing has become widespread in biology and medicine over the past decade due to both rapid technological advances and decreases in overall cost. The class will discuss study design, choice of methods, including practical issues of sequencing facilities, cost and computing resources, and then proceed to hands-on data analyses used in whole genome (re)sequencing, transcriptome analysis, and reduced-representation sequencing (e.g., RAD-seq, GBS). The schedule below gives some introductory topics. Additional topics will be covered depending on the interests of students and time available.

This course is designed to build competence in the computing and statistical methods for analyzing high-throughput genomic data. The only background assumed is a basic knowledge of statistics and genetics, familiarity with your computer and interest in learning current genomic methods. Knowledge of Linux operating systems is desirable, but not necessary. All that is required is a willingness to work hard (ie, not quit after the first [or second] error message).

Course Objectives: The primary objectives of this course are:
• to learn how genomic data are being used in biology, particularly evolutionary biology.
• to become familiar with the software and databases available for bioinformatics
• to develop the ability to formulate and investigate genomic research questions, and to effectively communicate your questions, methods, and results.

Prerequisites: Genetics (BioSci 325 or equivalent) and Biostatistics (465 or equivalent) or consent of instructor.

Computers: Computers are not required as we will be using the computer lab in Lapham, but students can also use their own laptops if desired. Note, however, I will not be spending a lot of time troubleshooting installations on personal computers. Much of the software we use is also available in a (relatively) easy to install version through the BioStars Handbook, which is also the source of some exercises. So if you want to use your own computer, I would strongly recommend purchasing the license (see below) and following the directions under “2. Getting Started. How to set up your computer.”

Credits and Evaluation: This is a 2-credit course. Grades are based on in-class assignments and participation (which includes simple things like asking questions!). There is NO Final Exam!

Graduate students will also receive part of their grade based on a presentation on the last day of class. Final assessment is based on the cumulative grades, as follows:

Undergraduates: Nine in-class assignments 90% (10 pts each), and class participation (10%).

Graduate students: Eight in-class assignments 80% (10 pts each), class presentation 20%. Graduate students will make a 15-20 min presentation to the class on a topic of interest (chosen after consulting the instructor). This could be their own research or a paper they read.

Attendance and Assignment requirements: Attendance is required at each class unless excused for documented reasons, such as sickness or other serious reasons. The penalty is a zero grade for that class (ie, 8.3% per class; 12 classes), unless excused for documented reasons. Assignments must be completed in class, unless approved by the instructor.

Time investment for this course: Students should plan to spend an average of 4 hours outside of class per week reading in preparation for in-class assignments. This amount of time is based on the campus credit hour policy (Faculty Document # 2838); ie, two hours out-of-class work for each credit hour per week of class.

Letter grades will be assigned based on the final total points listed below.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Minimum Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>92 – 100%</td>
</tr>
<tr>
<td>A-</td>
<td>89 – 91%</td>
</tr>
<tr>
<td>B+</td>
<td>86 – 88%</td>
</tr>
<tr>
<td>B</td>
<td>82 – 85%</td>
</tr>
<tr>
<td>B-</td>
<td>79 – 81%</td>
</tr>
<tr>
<td>C+</td>
<td>76 – 78%</td>
</tr>
<tr>
<td>C</td>
<td>71 – 75%</td>
</tr>
<tr>
<td>C-</td>
<td>68 – 70%</td>
</tr>
<tr>
<td>D+</td>
<td>65 – 67%</td>
</tr>
<tr>
<td>D</td>
<td>61 – 64%</td>
</tr>
<tr>
<td>D-</td>
<td>56 – 60%</td>
</tr>
<tr>
<td>F</td>
<td>0 – 55%</td>
</tr>
</tbody>
</table>

Need for Special Accommodation: Students who require note-taking or test-taking accommodations in order to meet any of the requirements of this course, please contact me as soon as possible to make suitable arrangements.
Schedule of Topics:

Week # and topic (order subject to change)

1. **Sep 10.** Introduction. What is your question? DNA or RNA? What technique should I use? Methods to analyze genomes, transcriptomes, meta-genomes & SNPs. What sequencers, data and software are available?

2. **Sep 17.** Computing resources and capabilities (stand-alone, UWM cluster, Amazon and other cloud resources). How to install software without going crazy. How to use Linux. **Exercise #1** – using SciProg, Filezilla and Linux

3. **Sep 24.** How do I get sequences and align them? NCBI, Ensembl etc. FASTA and other files. **Exercise #2.** – using genome browsers and sequence alignment software.

4. **Oct 1 – NO CLASS.**

5. **Oct 8.** Short-read sequence alignment to reference genomes. Structural and functional annotation. **Exercise #3.** – using FASTQC & BWA to align Ebola virus sequences

6. **Oct 15.** SNP calling / genotyping. **Exercise #4.** – using VCFtools (& dDocent pipeline?).


8. **Oct 29.** Genome wide association studies (GWAS). **Exercise #5.** – using PLINK.

9. **Nov. 5.** Transcriptome analysis overview. de novo or reference-guided. Sampling strategies.

10. **Nov. 12.** Differential gene expression. **Exercise #6.** – using the new 'Tuxedo' Suite on Zika data

11. **Nov. 19.** Gene ontology. **Exercise #7.** – using EdgeR (Chen et al. tutorial)

12. **Nov. 26.** Network analysis with KEGG, DAVID, Cytoscape etc. **Exercise #8.**

13. **Dec 3.** Grad student presentations. **Exercise #9.**

References for corresponding weeks

Some of the class material is based on: “The Biostar Handbook: A Beginner’s Guide to Bioinformatics” (2017) by Istvan Albert (available online at: https://read.biostarhandbook.com/)


2. Unix primer for Biologists: http://korflab.ucdavis.edu/unix_and_Perl/
4. No class this week.
Chapters 19-21 in the Biostars handbook.

University Guidelines of Interest

1. **Students with disabilities.** Notice to these students should appear prominently in the syllabus so that special accommodations are provided in a timely manner.
http://www4.uwm.edu/arc

2. **Religious observances.** Accommodations for absences due to religious observance should be noted. http://www4.uwm.edu/secu/docs/other/S1.5.htm

3. **Students called to active military duty.** Accommodations for absences due to call-up of reserves to active military duty should be noted.
Students: http://www4.uwm.edu/academics/military.cfm
4. **Incomplete**. A notation of "incomplete" may be given in lieu of a final grade to a student who has carried a subject successfully until the end of a semester but who, because of illness or other unusual and substantiated cause beyond the student's control, has been unable to take or complete the final examination or to complete some limited amount of term work. [https://www4.uwm.edu/secu/docs/other/S_31_INCOMPLETE_GRADES.pdf](https://www4.uwm.edu/secu/docs/other/S_31_INCOMPLETE_GRADES.pdf)

5. **Discriminatory conduct** *(such as sexual harassment)*. Discriminatory conduct will not be tolerated by the University. It poisons the work and learning environment of the University and threatens the careers, educational experience, and well-being of students, faculty, and staff. [https://www4.uwm.edu/secu/docs/other/S_47_Discrimina_duct_Policy.pdf](https://www4.uwm.edu/secu/docs/other/S_47_Discrimina_duct_Policy.pdf)

6. **Academic misconduct**. Cheating on exams or plagiarism are violations of the academic honor code and carry severe sanctions, including failing a course or even suspension or dismissal from the University. [http://uwm.edu/academicaffairs/facultystaff/policies/academic-misconduct/](http://uwm.edu/academicaffairs/facultystaff/policies/academic-misconduct/)

7. **Complaint procedures**. Students may direct complaints to the head of the academic unit or department in which the complaint occurs. If the complaint allegedly violates a specific university policy, it may be directed to the head of the department or academic unit in which the complaint occurred or to the appropriate university office responsible for enforcing the policy. [https://www4.uwm.edu/secu/docs/other/S_47_Discrimina_duct_Policy.pdf](https://www4.uwm.edu/secu/docs/other/S_47_Discrimina_duct_Policy.pdf)

8. **Grade appeal procedures**. A student may appeal a grade on the grounds that it is based on a capricious or arbitrary decision of the course instructor. Such an appeal shall follow the established procedures adopted by the department, college, or school in which the course resides or in the case of graduate students, the Graduate School. These procedures are available in writing from the respective department chairperson or the Academic Dean of the College/School. [http://www4.uwm.edu/secu/docs/other/S28.htm](http://www4.uwm.edu/secu/docs/other/S28.htm)

9. **Other** The final exam requirement, the final exam date requirement, etc. [http://www4.uwm.edu/secu/docs/other/S22.htm](http://www4.uwm.edu/secu/docs/other/S22.htm)