Bioinformatics and Genomics (BZ360) Fall 2017

Study guide for exam 2. The dates reference the exercises and slides posted on the schedule.

9/28/17

1. What is the difference between conventional and next-gen sequencing?
2. What is the most common next-gen sequencing platform and what is one notable drawback to it?
3. What are two common long read sequencing platforms discussed in class?
4. What is the purpose of the adapters added to DNA or RNA fragments during next-gen sequencing library preparation?
5. What is a flow cell?
6. What is bridge amplification and cluster generation?
7. What is sequencing by synthesis and how does it work?
8. What is the paired end sequencing?
9. What is a phred quality score (know how to calculate probability that a base was called inaccurately)?
10. What is meant by coverage? Be able to calculate coverage based on read length and genome size.
11. What is library indexing used for?

10/3/17

1. What is the size of the E. coli genome?
2. What is the typical size of a mammalian genome?
3. What is the size of the human genome?
4. What was the first multicelliular eukaryote to have its genome sequenced?
5. In what year was a draft sequence of the human genome first published?
6. What is shotgun sequencing?
7. What is a contig?
8. How are contigs identified from next-gen sequencing data?

10/5/17 and 10/10/17

1. What is de novo genome assembly?
2. What software was used in class to assemble the E. coli genome?
3. What GUI program was used to view the assembly data?
4. What is read mapping?

5. What is an N50 value in assembly stats?
6. How could you determine if a file is compressed?
7. What unix command is commonly used to compress a file? What option is used to decompress a file:
8. What is tar used for?

10/12/17

1. What is the endogenous role for CRISPR and in what types of organisms does it naturally occur in?
2. What is a guide RNA (gRNA)?
3. What is Cas9 and what is its function during genome editing?
4. What is a PAM motif?
5. How are semi-random mutations introduced during CRISPR-Cas9-mediated genome editing?
6. How are specific mutations or transgenes introduced using CRISPR-Cas9?
7. Related to questions 5 and 6: know the difference between NHEJ and HDR.

10/17/17

1. What is the difference between a read and a sequence in next-gen sequencing?
2. What information in contained in each line of a fastq file and how does it relate to questions 1?
3. What GUI program was used in class to assess library quality?
4. What software was used to map reads to the E. coli genome?
5. What is SAM (you don't need to know specific details about file content)?
6. What is BAM?
7. What GUI program was used to visualized our mapped reads on the E. coli genome?

10/24/17

8. What is GWAS?
9. What is comparative genomics?
10. What is functional genomics?
11. What are stem cells?
12. What are the different types of stem cell potency?