

Course Outline

1. **Course overview and biological background** (1 lecture)

- a. Problems in computational biology including genome assembly in different flavours, RNA-seq, Chip-seq, ATAC-seq, and other assays
- b. Single cell versions of various assays
- c. A discussion about the statistical and algorithmic challenges that are faced in these problems
- d. Brief discussion of biological background

2. **High-throughput sequencing** (1 lecture)

- a. Sequencing technologies (short read technologies like Illumina, long read technologies like Pacific Biosciences and Oxford Nanopore, linked read technologies like 10x)
- b. Base calling

3. **De novo Genome Assembly** (3-4 lectures)

- a. Dense read formulation: Necessary and sufficient conditions (informational view)
- b. Algorithms for assembly: de Bruijn graph based algorithms, Overlap graph based algorithms
- c. Errors and biases

4. **Read alignment** (3 lectures)

- a. Dynamic programming
- b. Hash-based seed-and-extend
- c. FM-index and Burrows-Wheeler transform
- d. Suffix arrays
- e. Minhash
- f. Applications such as spliced alignment, and alignments used in practical cases like DAligner, and Minimap.

5. **Variant calling** (1 lecture)

- a. SNV calling
- b. Structural variant calling

6. **Phasing and Imputation** (2 lectures)

- a. Imputation algorithms
- b. Phasing algorithms

7. **RNA-Seq assembly** (2 lectures)

a. Formulation

b. Algorithms

8. **RNA-Seq quantification** (2 lectures)

a. EM algorithm

9. **Single-cell RNA-Seq analysis** (3 lectures)

a. Differential expression

b. Cell Differentiation

c. Visualisation

d. Trend Analysis

10. **Genome Compression** (1 lecture)

Guest lecture by [Stephen Turner](#), Co-founder and Chief Technology Officer, [Pacific Biosciences](#) on 13 April 2016.

Useful Resources

1. Lawrence Hunter, [Molecular Biology for Computer Scientists](#) - A crisp write-up on the basics of biology which motivate, and provide insights into problems we discuss in class. This is written in a non-biologist friendly manner.
 2. Eric Lander, [Fundamentals of Biology](#), [MIT Open Course Ware](#) - Lectures covering the basics of biology. Very friendly to non-biologists.
 3. [Ben Langmead's lecture notes](#) - Covers many topics that we cover in class. Some very nice video lectures and example code in ipython notebooks.
 4. [Bioinformatics algorithms by Compeau and Pevzner](#) - Covers many topics that we cover in this class. Video lectures are also available on the book site.
-