Bioinformatics

University of Cambridge

Ashwin Ahuja

Computer Science Tripos
Part II

January 2020

Contents

1 Purpose

Main purpose is to help doctors and help to understand the biology and computing with DNA and other biological molecules. Can use DNA as storage information. Today, through *Oxford Nanopore* and *Bento Lab* sequencing has got much cheaper and quicker. Idea of **Garage Genomics** where it often only takes an hour to get a portion of the DNA.

1.1 Definitions

- DNA made of four-letter alphabet: Adenine, Thymine, Cytosine and Guanine
- RNA is the same with Uracil instead of Thymine.
- **Amino Acid**: Made of three base pairs of DNA
- Genome: organism's genetic material - for humans = 46 strings (chromosomes) with length 3 x 10e9

2 Alignment

3 Folding

4 Trees and Phylogeny

5 Parsimony

6 Neighbour Joining

7 Genome Sequencing

8 Clustering

9 Genome Assembly and Pattern Matching

10 Hidden Markov Models