Computational Genomics

Week7 Tutorial

CpG Markov Chain Example

- Assumption:
  - Transition probability of X -> Y in a CpG island is: Pxy;
  - Transition probability of X -> Y in a Non CpG island is Qxy;
- Then, what is the probability of string ACGT for a CpG island?
- Is it a CpG island?

CpG Markov Chain Example

- Then, what is the probability of string ACGT for a CpG island?

CpG Markov Chain Example

- Then, what is the probability of string ACGT for a CpG island?

- Answer: $P_{AC} * P_{CG} * P_{GT}$

CpG Island Detector

- How is transition probability matrix obtained?

- Does it require prior knowledge of CpG islands?

CpG Island Detector

- What if the task is to locate CpG island in the genome?

- Transition probabilities between the states (in island/ outside island).

CpG Island Detector

- Build a state transition matrix for a HMM-based CpG island detector (CpG islands in underline)

- GGTTCCGCTCCCACCGCGCCGGCGTTCGGCCACGTT

Hidden Markov Model

- Use the above estimates and the brute-force algorithm from slide 17 to annotate the sequence CGCGGCACGC of:

  - All in CpG island;

  - All not in CpG island.

Hidden Markov Model

- Extend the above calculation to all possible paths. Find the annotation with the highest possibility.