MBB 441

SFU SIMON FRASER UNIVERSITY

BIOINFORMATICS

#Genomics #Bioinformatics

Examining the latest developments in bioinformatics, which is loosely defined as the intersection between the fields of molecular biology and computer science. Topics covered may include sequence searching and alignment, bioinformatics databases and ontologies, sequence variant analysis, sequence motif identification, metagenomics analysis, protein analysis, evolutionary/phylogenetic analysis, network/systems biology, machine learning/software development for predictive methods, and ethics in method/database development and analysis. Instruction will include lecture material and in-depth consideration of selected papers in the field.

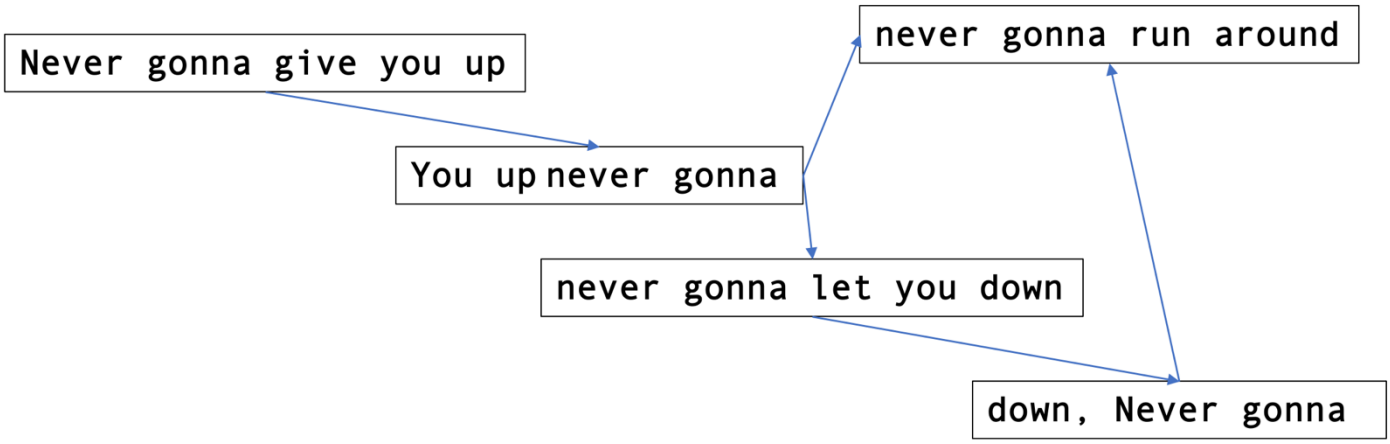
TOPICS

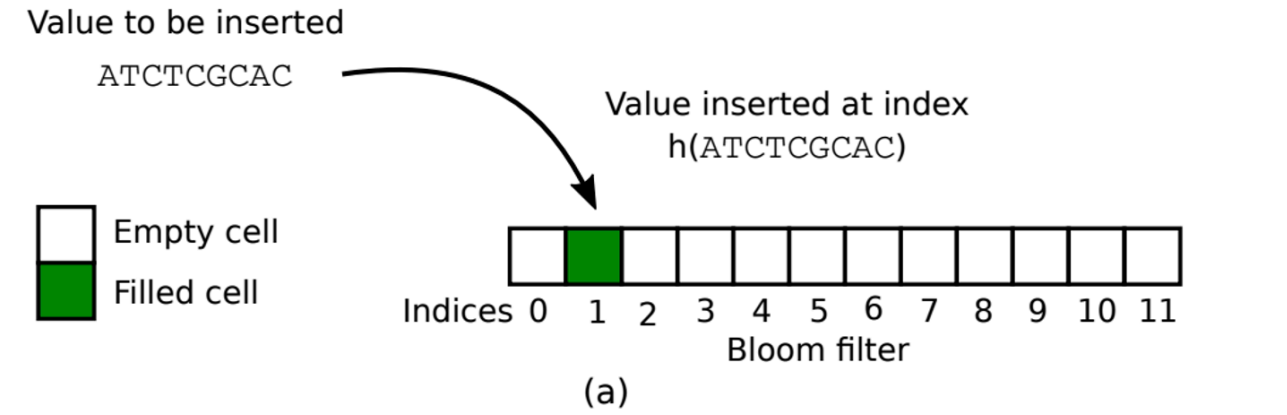
- Big data and semantics
- Ontology
- Orthology
- Sequencing and analysis
- Genome assembly
- SNV and SNP calling algorithms
- Clustering and high dimensional data analysis
- Network analysis
- Metagenomics

INSTRUCTORS

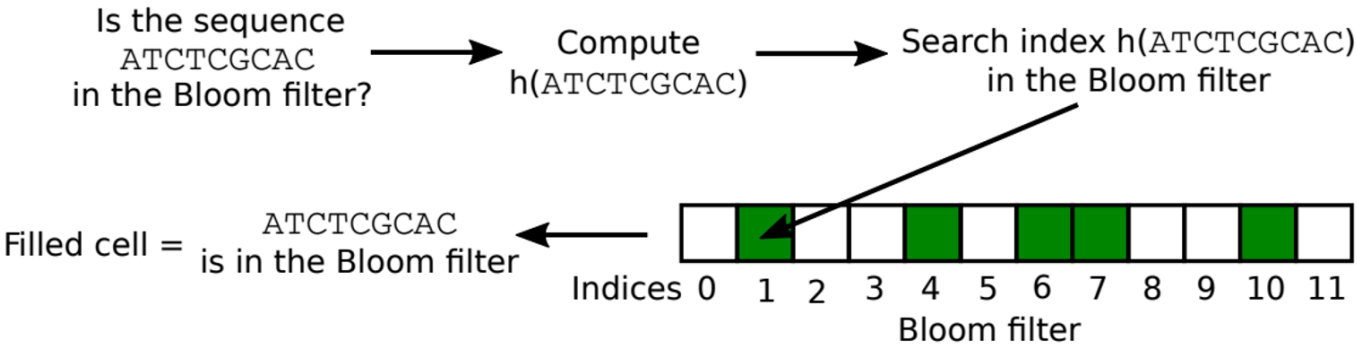
Fiona Brinkman

Sophie Sneddon ~ Bioinformatician, video game enthusiast, cat person

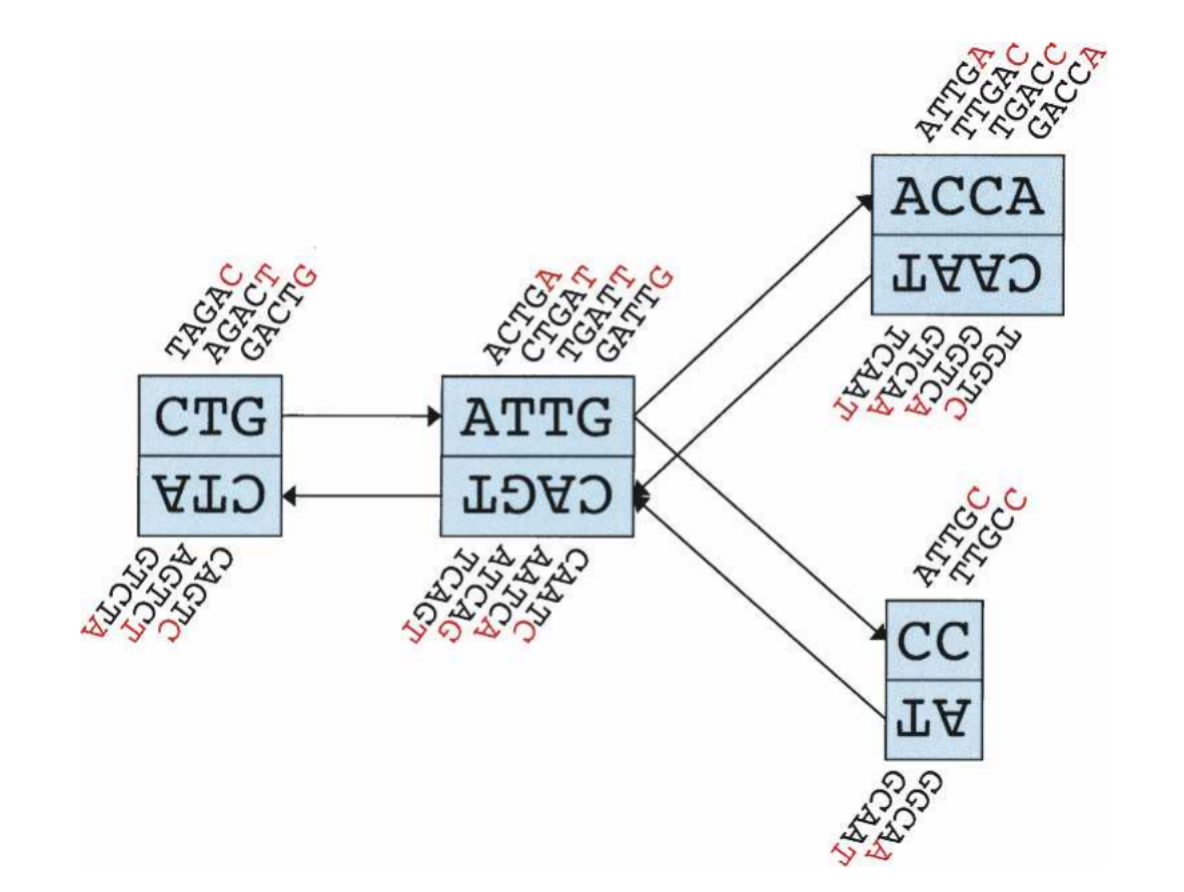

Never gonna give you up
never gonna run around
You up never gonna
never gonna let you down
down, Never gonna

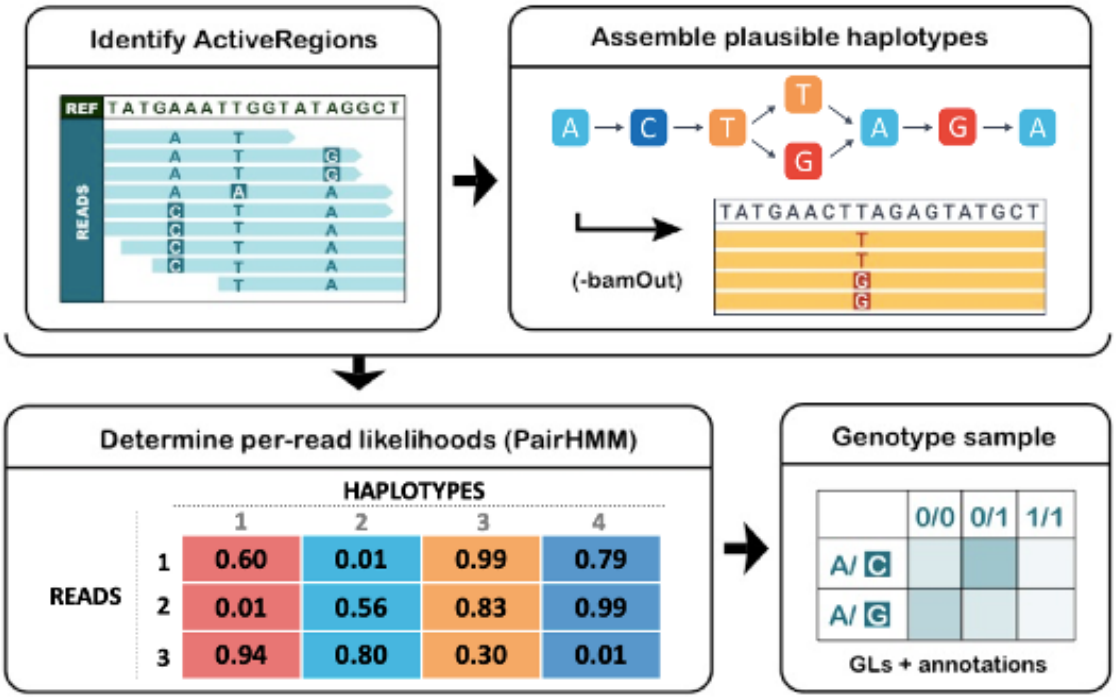

Value to be inserted
ATCTCGCAC
Value inserted at index
h(ATCTCGCAC)
Empty cell
Filled cell
Indices 0 1 2 3 4 5 6 7 8 9 10 11
Bloom filter

(a)


Is the sequence
ATCTCGCAC
in the Bloom filter?
Compute
h(ATCTCGCAC)
Search index h(ATCTCGCAC)
in the Bloom filter
Filled cell =
ATCTCGCAC
is in the Bloom filter
Indices 0 1 2 3 4 5 6 7 8 9 10 11
Bloom filter

(b)


ACCA
CAAT
ATTG
CAGT
CTG
CTA
CC
AT
TAGAC
AGACT
GACTG
ACTGA
CTGAT
TGATT
GATTG
ATTGA
TTGAC
TGACC
GACCA
ATTGC
TTGCC
TGGTC
GGTCA
GTCAA
TCAAT
CAGTC
AGTCT
GTCTA
CATC
ATCA
TCAG
CAGT
GGCAA
GCAAT


Identify ActiveRegions
REF
TATGAAATTGGTATAGGCT
READS
Assemble plausible haplotypes
A
C
T
T
G
A
G
A
(-bamOut)
TATGAACTTAGAGTATGCT
Determine per-read likelihoods (PairHMM)
HAPLOTYPES
1
2
3
4
READS
1
0.60
0.01
0.99
0.79
2
0.01
0.56
0.83
0.99
3
0.94
0.80
0.30
0.01
Genotype sample
0/0
0/1
1/1
A/C
A/G
GLs + annotations