

---

# Computational Approaches to Haplotype Inference

---

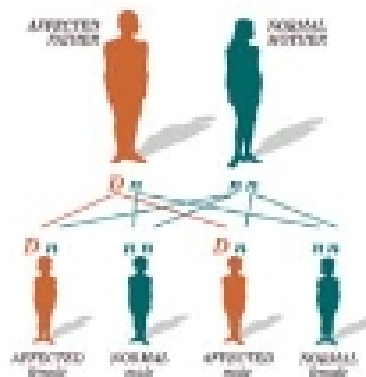
Ravi Vijaya Satya  
Amar Mukherjee

---

## Overview

- SNPs & Haplotypes
  - The HapMap Project
  - Why “Infer” Haplotypes?
  - Computational Methods
    - Maximum Resolution
    - Perfect Phylogeny Haplotyping
    - Haplotyping with Pedigree information
  - Haplotyping via sequencing
  - Direct Approach for PPH (Bafna, Gusfield, et. al.)
-

## Genetic Variations

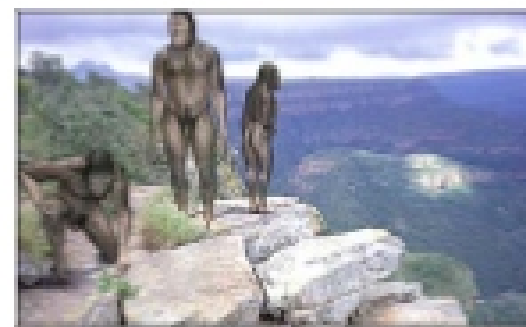


underlie phenotypic differences



cause inherited diseases

allow tracking ancestral human history



Source: Gabor T. Marth,

[www.varbug.org/talk\\_ppts/Gabor\\_2004.ppt](http://www.varbug.org/talk_ppts/Gabor_2004.ppt)

## SNP: Single Nucleotide Polymorphism

“Loci in the human genome in which a considerable percentage of the population differs from the rest.”

...CATGATCACGTCGACGATCGAT...

...CATGATCACGTCGACGATCGAT...

...CATGATCATGTCGACGATCGAT...

...CATGATCACGTCGACGGTCGAT...

Allele - One of the possible states of a given a locus

The locations, or loci, are also called 'markers'

## Types of SNPs

- Number of alleles:
  - Bi-allelic: A site is called bi-allelic if there are only two possible states for that site.
  - Multi-allelic: A site is called multi-allelic if there are more than two possible states for that site
  - Almost all the SNPs are bi-allelic
- Coding / Noncoding
  - Coding (CSNP), if the SNP occurs in an exon
  - Non-coding, if it occurs in an intron or in a non-coding region

## Types of SNPs (contd...)

- Coding SNPs can be:
    - Silent
    - Non-silent
- ```
.....aca gat cag atc atg.....  
..... T D Q I M .....
```
- ```
.....aca gaa cag atc atg.....  
..... T E Q I M .....
```