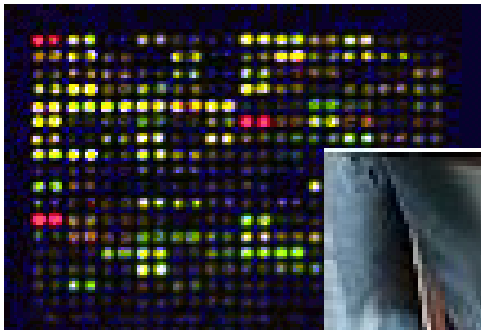


The informatics of SNPs and haplotypes



CGDN Bioinformatics Workshop
February 20, 2006



Gabor T. Marth

Department of Biology, Boston College
marth@bc.edu

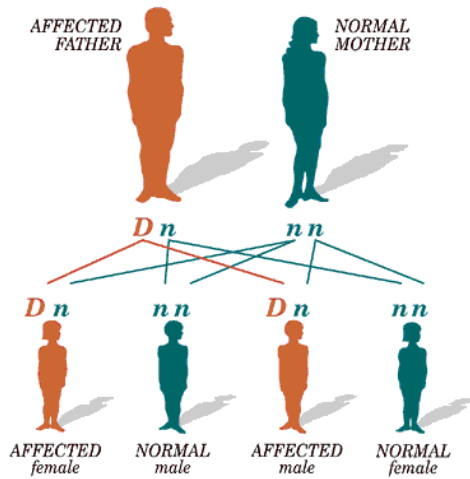


Why do we care about variations?

underlie phenotypic differences



cause inherited diseases



allow tracking ancestral human history





Where do variations come from?

• sequence variations are the result of **mutation events**

TAAAAAT



TAA**C**AAT

• mutations are propagated down through generations

MRCA

TAAAAAT

TAAAAAT

TAA**C**AAT

TAAAAAT

TAAAAAT

TAA**C**AAT

TAA**C**AAT

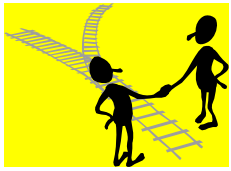
TAA**C**AAT

• and determine present-day variation patterns

```

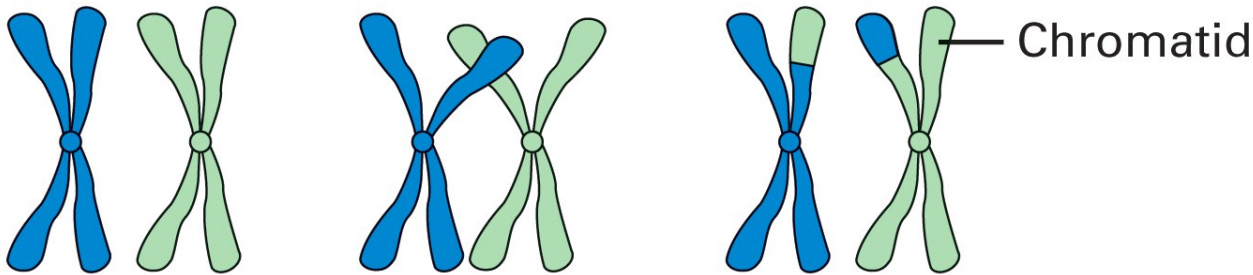
TCTGACCAATCTAAAAATACCTGTGATTAA
TCTGACCAATCTAACAATACCTGTGATTAA
TCTGACCAATCTAACAATACCTGTGATTAA
TCTGACCAATCTAAAAATACCTGTGATTAA
tctgaccaatctaacaatacctgtgattaa

```

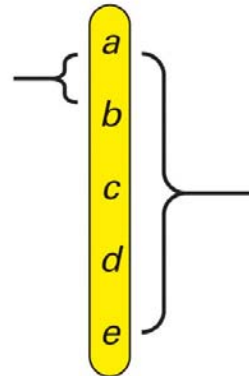



How to use markers to find disease?

Homologous chromosomes undergoing crossing over



To separate *a* and *b*,
crossover must occur
in this narrow stretch



To separate *a* and *e*,
crossover can occur
anywhere in this
stretch



<http://bioinformatics.bc.edu/marthlab>