

LD Scans for Ongoing Selective Sweeps

Alan R. Rogers

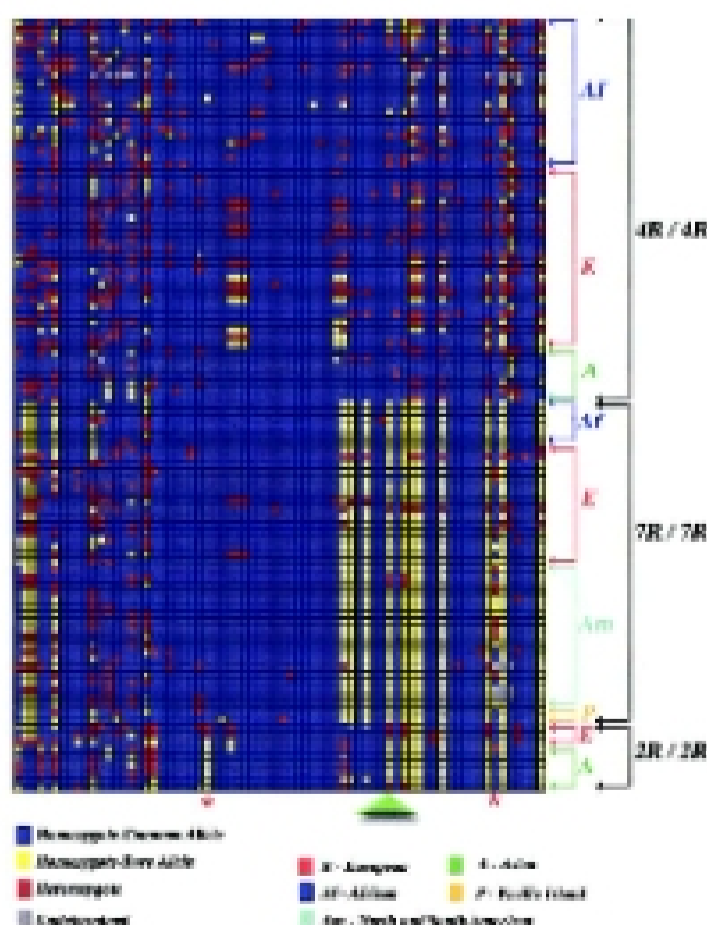
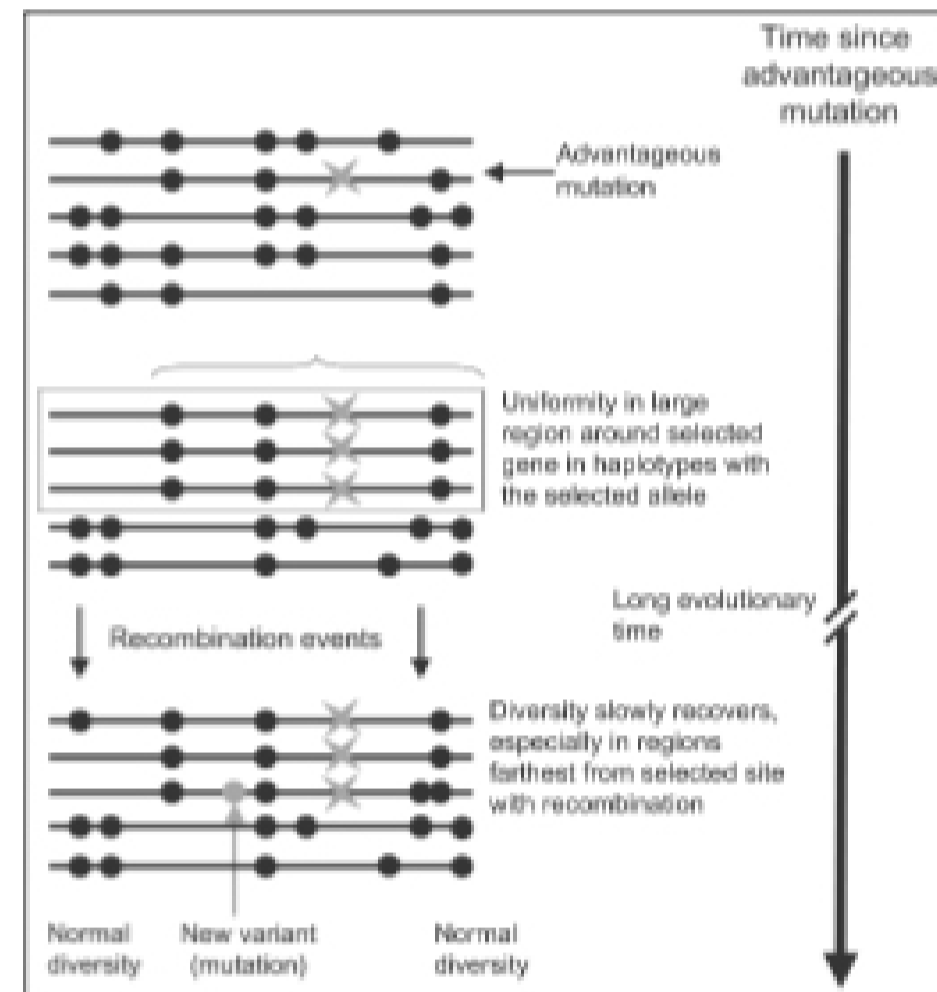
December 3, 2007

Outline

- ▶ Questions
 - ▶ Have humans evolved rapidly or slowly during the past 40 kyr?
 - ▶ What functional categories of gene have evolved most?
- ▶ Selection and recombination
- ▶ Data
- ▶ Results

Are we still evolving?

- ▶ Since split with chimps, there has been rapid evolution in proteins expressed in brain and in sperm.
- ▶ Recent selection at various loci: lactase, DRD4, etc
- ▶ How common are such loci in the human genome?
- ▶ How can we tell?



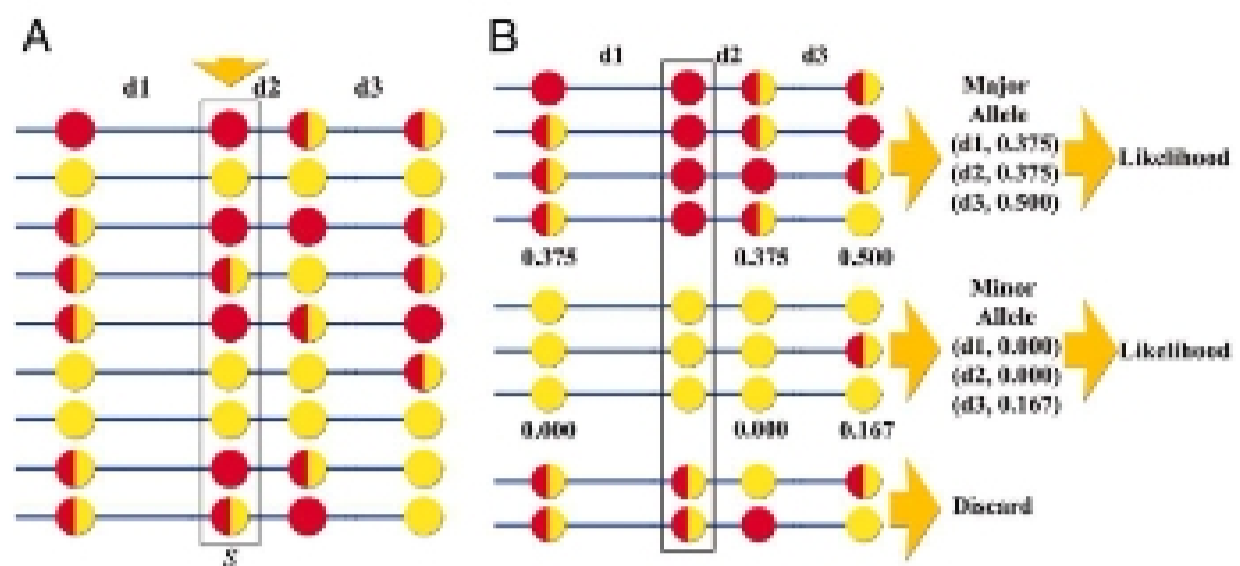
Signature of an ongoing selective sweep at DRD4

- ▶ Sweeping allele is
 - ▶ common
 - ▶ has low diversity over large region
- ▶ High LD over large region

Two methods for finding sweeping alleles

- ▶ Frequency of recombinant chromosomes (FRC): Wang et al 2006
- ▶ Extended haplotype homozygosity (EHH): Voigt et al 2006

Fraction of Recombinant Chromosomes (FRC)



At a given site

- ▶ FRC increases with time.
- ▶ Rate of increase depends on recombination rate.

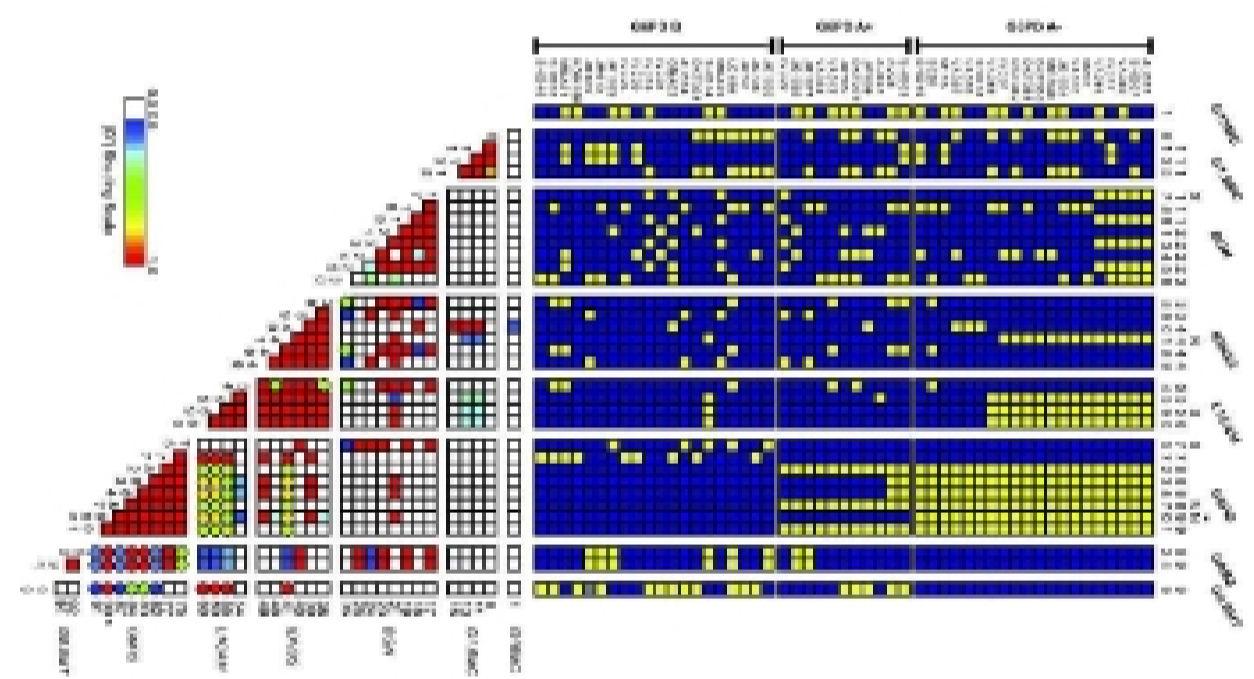
At a given time

- ▶ FRC increases with distance along the chromosome.
- ▶ Rate of increase is fast near neutral sites.
(Neutral \Rightarrow old \Rightarrow lots of recombination.)
- ▶ Rate is slow near ongoing selective sweeps.
(Selected \Rightarrow young \Rightarrow little recombination.)

Recipe
Look for regions where FRC is low in big sections of chromosome.

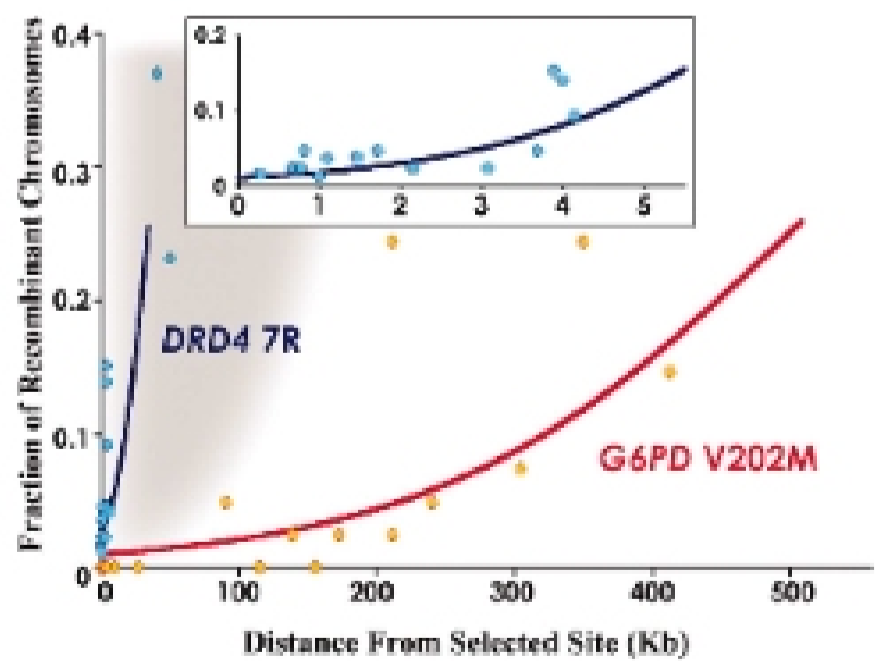
Linkage disequilibrium at G6PD

Left: LD plot; right: haplotype plot



Results for G6PD

- ▶ Favored allele is about 100 generations old.
- ▶ Selective advantage: $s \approx 0.1$.



- ▶ fraction recombinants increases with distance
- ▶ G6PD: recent selection
- ▶ DRD4: older, or in a region with more recombination

Study of Wang et al (2006)

- ▶ 1.6 million polymorphic nucleotide sites from throughout the genome
- ▶ most typed in worldwide sample of 71 people

Vertical lines: selected loci (Wang et al 2006)



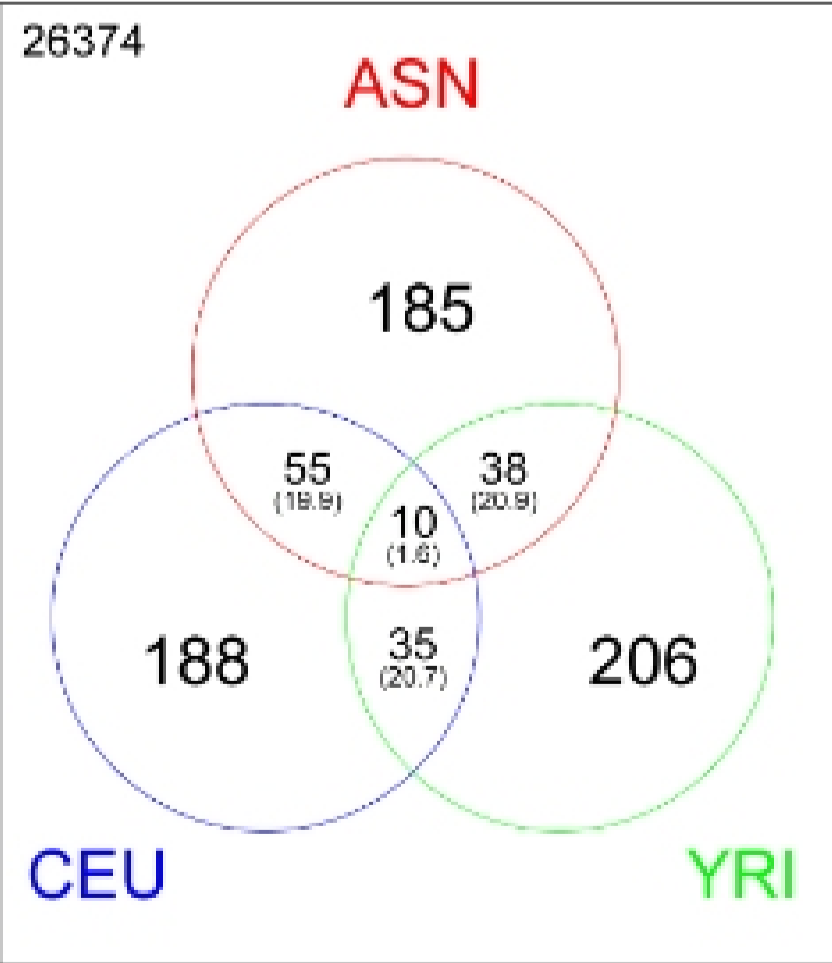
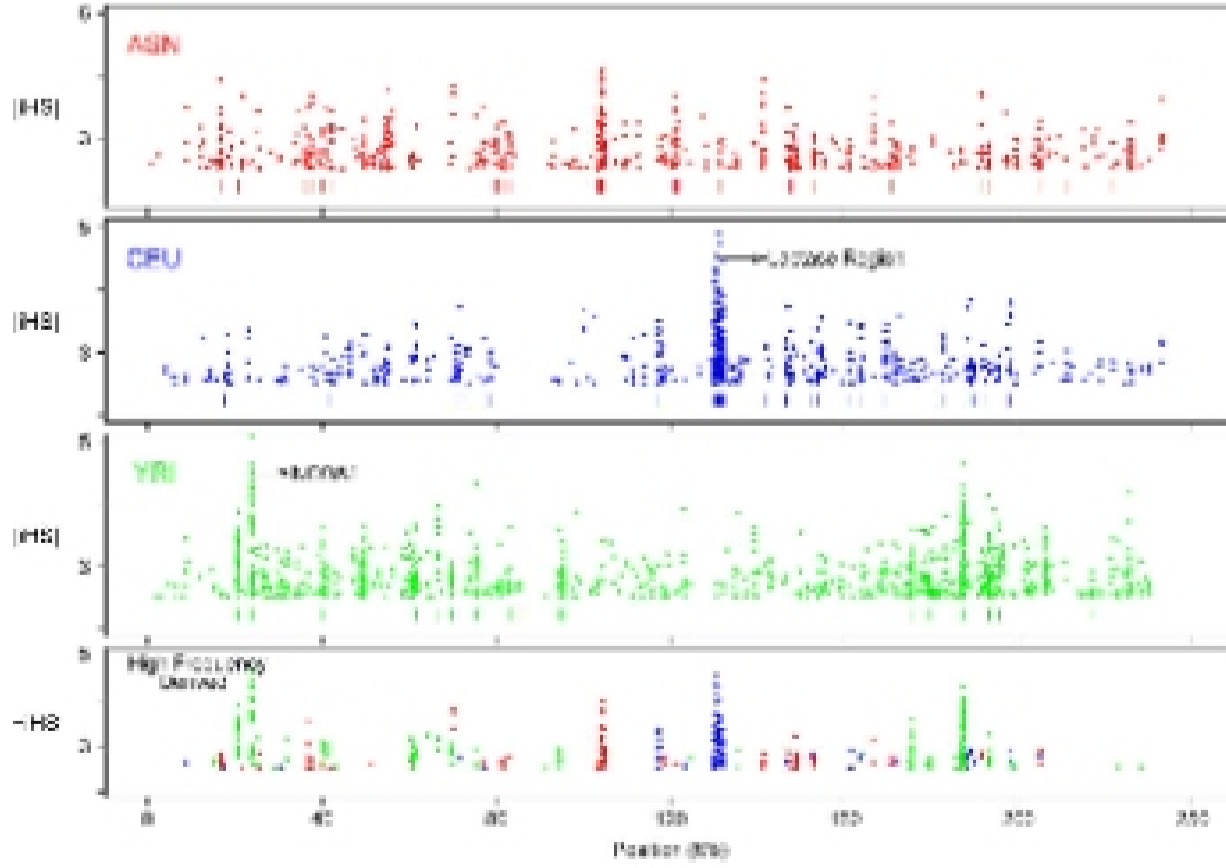
Entire genome (Wang et al 2006)



Study of Voigt et al (2006)

- ▶ 800,000 SNPs in 309 people
- ▶ 431 sweeping loci
- ▶ Most sweeps started w/i past 10,000 years

LD on human chromosome 2 (Voigt et al 2006)



Voigt et al (2006):
431 sweeping loci.
ASN: Asia
YRI: Africa
CEU: Europe.
Most are sweeping w/i only one continent.
Also true of Wang et al data.