Andy Clark - Human variation -

-15 candidate genes in 24 individuals in 3 populations

-select most informative sites

-Score in 1000+ individuals (each w/ epidemiology)

① Sequence w/ genomic PCR
② Haplotype inference
Clark 1990
③ Patterning of variation

Expected # of Segregating Sites

$E(S) = \Theta \sum_{i=1}^{n-1} \frac{1}{i}$ ← sample size infinite sites (Kimura) $\Theta = 4nu$

E reaches peak quickly w/ more individuals

Four gamete tests

site 1: C A / T C
site 2: G T / A C

- Haplotypes a b c d
- Do not expect all 4 a b c d unless ① back mutations ② recombination

Observation -- more haplotypes than expect

Conclusion -- recombination common

two major clades of alleles

ACE

- 18 sites in Int. 16 seq. b/w Alu polymorphism
- Clustered together in SNPs
  (∴ Alu likely mutagenic)

Are there any conserved features of the Δ's in Alu region

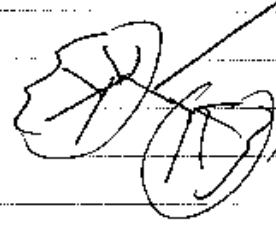
Tajima's D statistic

- most of these are positive
- this suggests some complicated history
  (e.g. expansion + bottleneck)

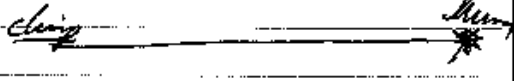
Chimpanzee
- 112 fixed differences
- 4 were coding
- All were silent

} For each varying site in humans chimp had one or other.

YAC



MtDNA



Linkage Diseq

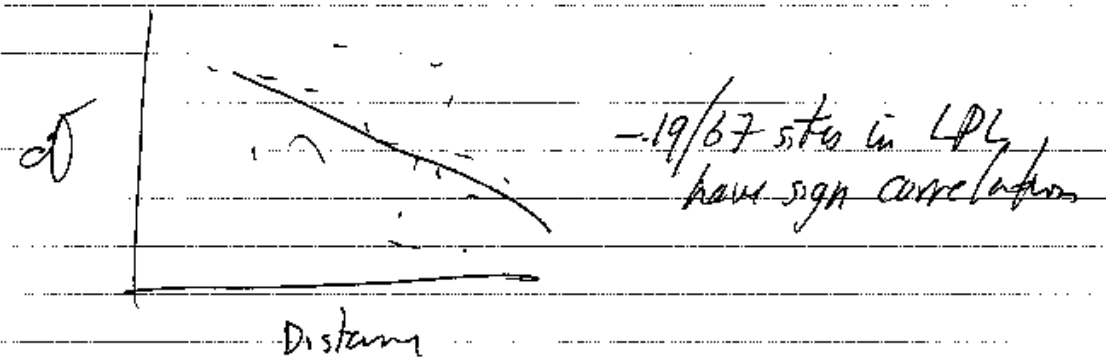
- if select site randomly from LPL ...
  most likely it is not in linkage diseq

Allele Freq + Linkage Diseq should be associated

- Rare = recent
- Common = old

Older should have more change for recombination

CAN YOU ESTIMATE AGE OF CHIMP/HUMAN
AND SEE IF CORRELATE W/ FREQ



-19/67 sites in LPL have sign correlation

Is there something about those site w/o LD

Bad News

① Intragenic recombination happens

② Gene conversion + repeated mutation generate homoplasy

③ Site pairs w/in a gene may have zero linkage disequ.

④ If disease causing mutations are recurrent or highly homoplasious, flanking sites may provide little prediction power

Good news

① Not all genes as variable as LPL

② Recomb hot pots may be common

③ recomb does not erase history

④ LD decays on avg w/ distance and w/ allele frequency

⑤ The pattern of LD is not grossly divergent across populations

But how much do to distance?