

Short Tandem Repeats

Mischa Li

Biochem 118Q

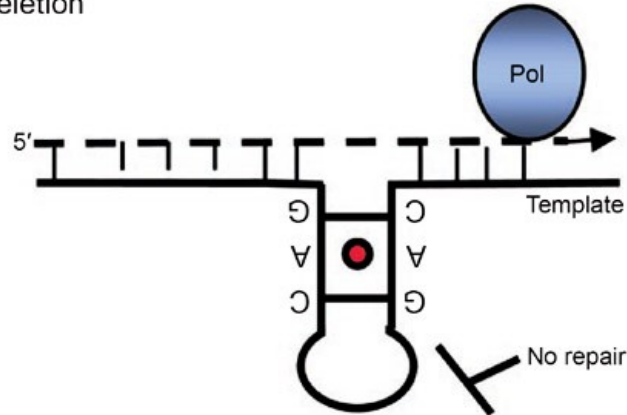


What's an STR?

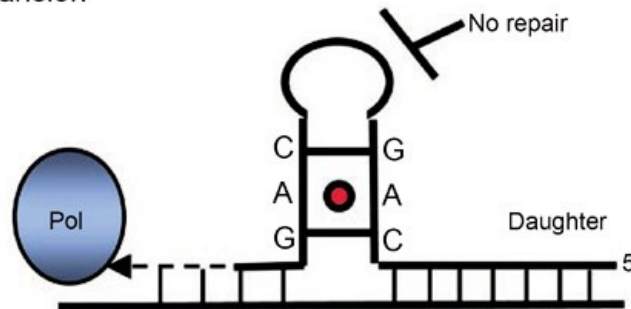
- Also called microsatellites
- Class of genetic polymorphism (like SNPs or minisatellites)
- 2 - 10 nt that are repeated in tandem
- Usually in the intron regions (junk DNA)
- Increased rate of mutation due to slippage in DNA replication
 - Rate especially high in tumor cells, where replication controls are damaged (a tumor cell can show a different profile from normal host cells)

Slip in DNA Replication

Deletion



Expansion



- <http://www.nature.com/cr/journal/v18/n1/full/cr20085a.html>



Uses

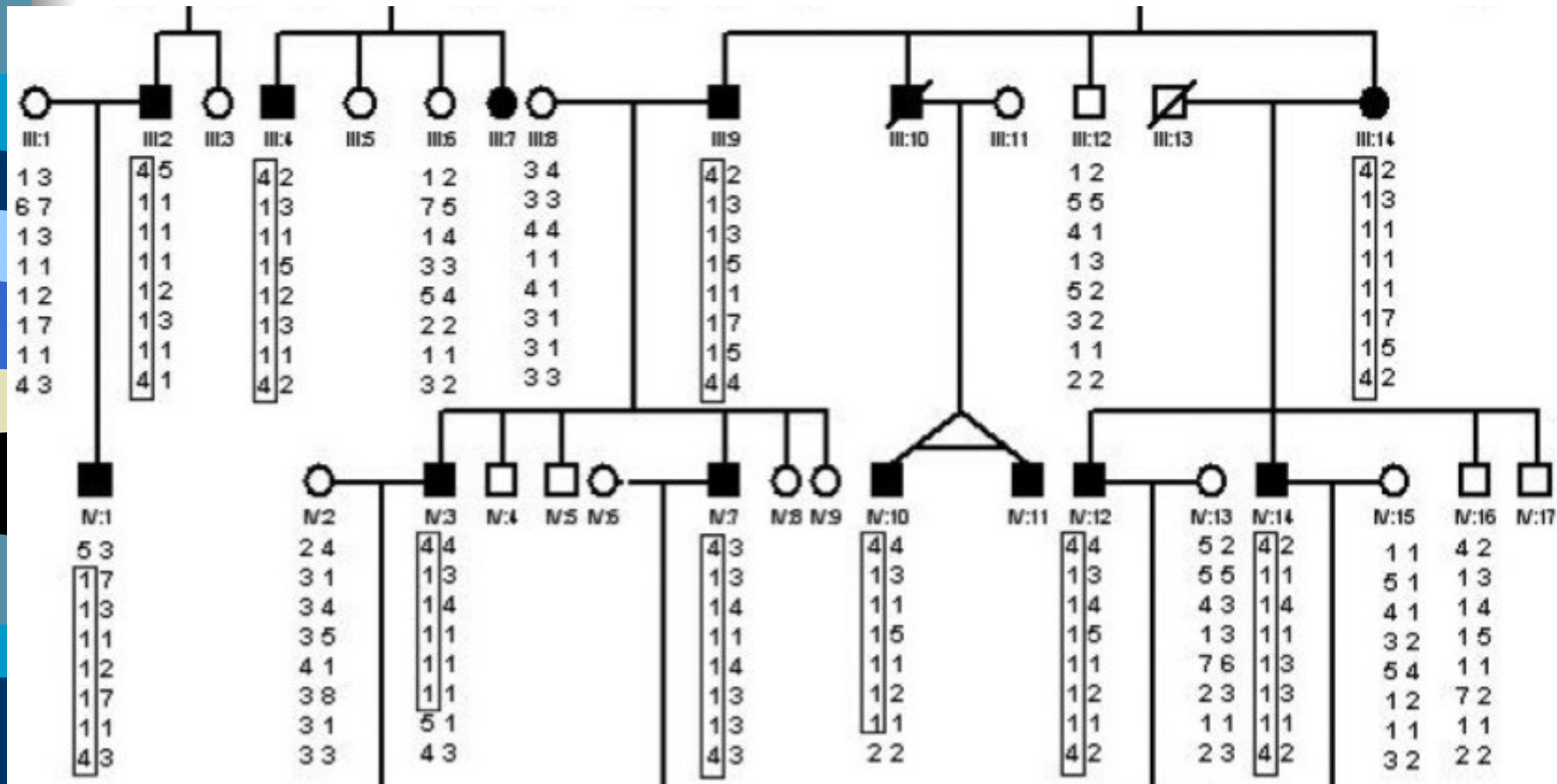
- Determining paternity
 - Does one of your alleles at a particular STR locus come from your father?
- Reconstructing genealogy
 - Tracing how a particular allele is passed down through the generations
- Population genetics
 - Particular alleles are more frequent among certain subpopulations
- Recombination mapping
 - Which alleles tend to travel together?



Pedigree

- SNPs mutate at a rate of 10^{-9} /site, whereas STRs mutate at a rate of 10^{-4} /locus.
- Allows for a more refined search

Sample STR Pedigree



- <http://www.molvis.org/molvis/v14/a88/linfig1.html>



Forensics

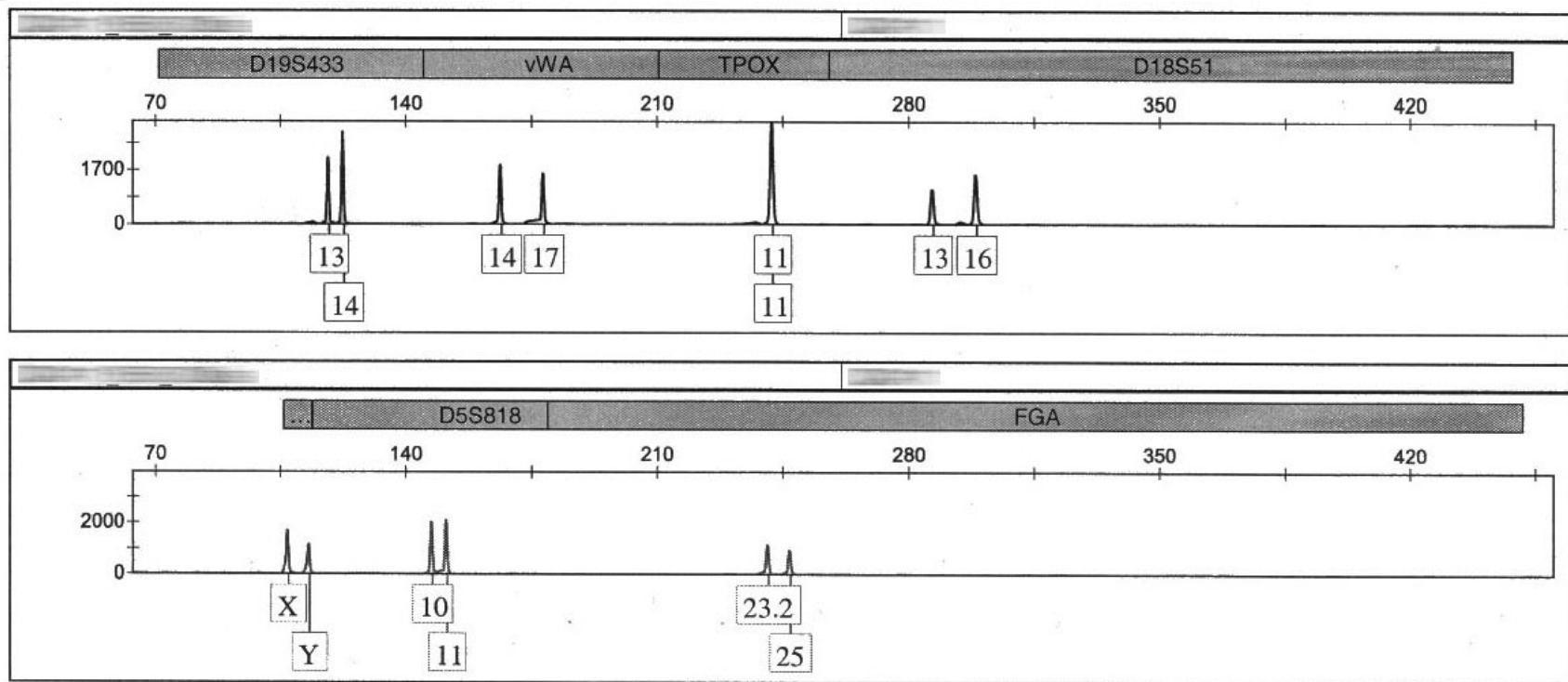
- http://news.bbc.co.uk/2/hi/uk_news/england/36
- Started in the mid-to-late 90's
- Microsatellite profiling
- Uses 4 - 5 nt repeats
 - Tri-nt repeats typically associated w/ disease
 - Long repeats not robust enough to survive degradation and harder to amplify
- In the US, 13 STR loci are used for genetic profiles. Profiles stored in databases like CODIS (Combined DNA Index System).



How microsatellite profiling works

- DNA is extracted from site samples
- Specific polymorphic STR loci amplified using PCR
- Amplified products applied to gel or capillary electrophoresis
- Determine how many repeats for each STR locus
- Most popular system: Applied Biosystems Identifiler

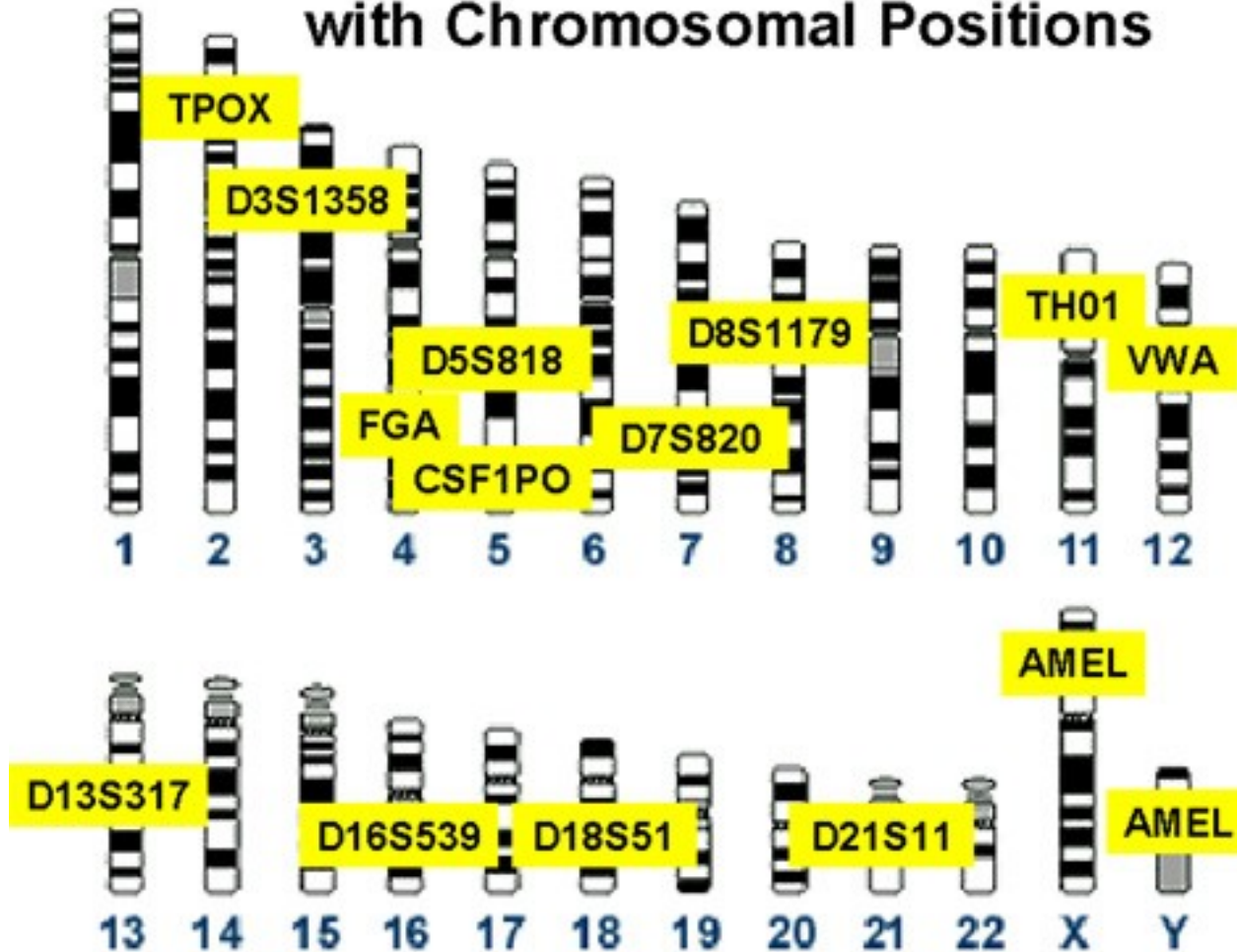
Sample test result



■ Wikipedia

The 13 CODIS Loci

13 CODIS Core STR Loci with Chromosomal Positions



- <http://cstl.nist.gov/div831/strbase/fbicore.htm>



CODIS

- In California, as a result of proposition 69 in 2004, within 5 years all suspects arrested for felony + some suspects of misdemeanors will have their DNA collected
- Nationwide, through Feb. 2005 22,122 cases were aided by CODIS hits



Prosecutor's Fallacy and Population Substructure

- Match probability of “1 in a billion” does not always mean that the guy whose DNA matches the sample is necessarily the culprit
- Because humans do not mate randomly (there is population substructure)
- Members of the same subpopulation are much more likely to share microsatellite lengths at particular loci
- Need other evidence as well



Sources

- Wikipedia: STR and microsatellite
- BBC World News
- Short Tandem Repeat Internet Database
<http://cstl.nist.gov/div831/strbase/index.r>
- Lecture slides, Ansi 14: Introduction to Anthropological Genetics, winter 2008