FROM DNA TO GENETIC GENEALOGY

Stephen P. Morse (steve@stevemorse.org)

1. GENES, CHROMOSOMES, AND DNA

Chromosomes

Every human cell = 46 chromosomes (1 to 22 in pairs, 2 sex chromosomes)
Male: sex chromosomes = X plus Y
Female: sex chromosomes = X plus X

DNA

Chromosome = long DNA molecule (double helix) with two strands
Each strand consists of 4 repeating bases (A, C, G, T)
"A" pairs with "T", "C" pairs with "G"

Genes

Genes are portions of chromosomes with identifiable function
It is a subset of the DNA sequence of the chromosome

The Numbers

Base Pairs per chromosome: between 50 to 250 million
Total Base Pairs: 3 billion
Base Pairs per gene: 27 thousand (average), 2.4 million (largest)
Genes per chromosome = 200 to 3,000
Total genes = 30 thousand

2. CRACKING THE DNA CODE

Every function in a living cell depends on proteins
Each gene is a DNA program that makes one protein
A protein is a sequence of amino acids

Amino acids

Ala/A Alanine	Gly/G Glycine	Pro/P Proline
Arg/R Arginine	His/H Histidine	Ser/S Serine
Asn/N Asparagine	Ile/I Isoleucine	Thr/T Threonine
Asp/D Aspartic Acid	Leu/L Leucine	Trp/W Tryptophan
Cys/C Cysteine	Lys/K Lysine	Tyr/Y Tyrosine
Glu/E Glutamic Acid	Met/M Methionine	Val/V Valine
Gln/Q Glutamine	Phe/F Phenylalanine	

Each DNA triplets specifies one amino acid

TTT → Phe	CTT → Leu	ATT → Ile	GTT → Val
TTC → Phe	CTC → Leu	ATC → Ile	GTC → Val
TTA → Leu	CTA → Leu	ATA → Ile	GTA → Val
TTG → Leu	CTG → Leu	ATG → Met/start	GTG → Val
TCT → Ser	CCT → Pro	ACT → Thr	GCT → Ala
TCC → Ser	CCC → Pro	ACC → Thr	GCC → Ala
TCA → Ser	CCA → Pro	ACA → Thr	GCA → Ala
TCG → Ser	CCG → Pro	ACG → Thr	GCG → Ala
TAT → Tyr	CAT → His	AAT → Asn	GAT → Asp
TAC → Tyr	CAC → His	AAC → Asn	GAC → Asp
TAA → stop	CAA → Gin	AAA → Lys	GAA → Glu
TAG → stop	CAG → Gin	AAG → Lys	GAG → Glu
TGT → Cys	CGT → Arg	AGT → Ser	GGT → Gly
TGC → Cys	CGC → Arg	AGC → Ser	GGC → Gly
TGA → stop	CGA → Arg	AGA → Arg	GGA → Gly
TGG → Trp	CGG → Arg	AGG → Arg	GGG → Gly

3. HOW WE INHERIT OUR DNA

Chromosome inheritance

Chromosomes 1 to 22 (autosomes): 1 shuffled chromosome per parent
X chromosome: shuffled chromosome from mother
second X chromosome (daughter): intact chromosome from father
Y chromosome (son): intact chromosome from father

MtDNA inheritance

passed from mother to all children

Mistakes (mutations)

SNiP: **Single** Nucleotide Polymorphism – rare event, never gets undone
Can be used to trace early migration pattern

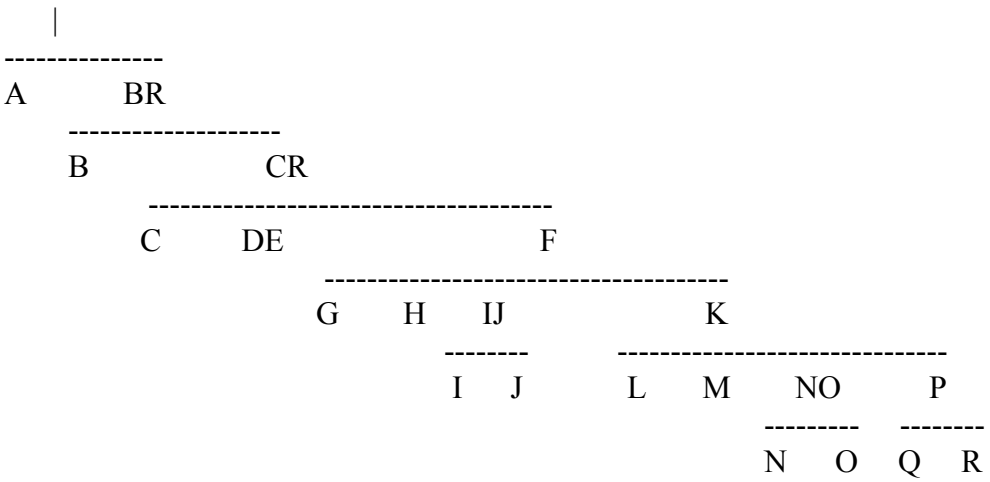
STiR: Short Tandem **Repeat** – once every 500 events, can increase or decrease
Can be used to estimate time to common ancestor

Marker: region in chromosome that is tested
Allele: value of DNA at a marker

4. OUT OF AFRICA

Each time a SNiP mutation occurred, we can identify a different "branch" of mankind
By seeing where the branches are indigenous today, we can determine migration patterns

Branches



SNiPs define the branches
STiRs are what most genealogists have tested (to find common ancestors)
From large databases, frequencies of particular STiRs in each branch have been obtained
From this, you can obtain the most probable branch corresponding to your STiRs

5. KOHANIM (JEWISH HIGH PRIEST)

Aaron was the brother of Moses
All direct male descendants of Aaron are the Kohanim
Therefore all Kohanim should have Aaron's Y chromosome (with some mutations)

We don't have Aaron's DNA to test
But we can test DNA of all people claiming to be Kohanim and see if they are similar

Cohen Modal Haplotype – those markers that are identical for many claimed Kohanim

DYS393 = 12	DYS390 = 23
DYS19 = 14	DYS391 = 10
DYS388 = 16	DYS392 = 11

But…

- Only six markers were tested
- Not enough to reliably estimate time to common ancestor
- CMH is not uncommon in general population of non Kohanim

6. GENETIC DISEASES

Down Syndrome

Found in all populations
Extra copy of chromosome 21
Two from one parent (usually mother), one from the other parent

Sickle Cell Anemia

Mostly found in sub-Saharan African populations
Chromosome 11, β-globin gene (recessive)
SNiP: GAG → GTG, changes the amino acid from glutamate to valine

Tay-Sachs

Ashkenazi Jewish, Louisiana Cajuns, French Canadian
Chromosome 15, HEXA gene (recessive)
Over 90 different mutations identified (SNiPs, STiRs, etc)
Most prevalent Jewish one is STiR: extra TATC, alters framing

Ashkenazi Jewish and French Canadian are different mutations – no relation
Louisiana Cajun is same mutation as Ashkenazi Jewish

Hemophilia

More prevalent in men than women
X chromosome
Women need two defective genes to be infected, men only one

7. COMMON APPLICATIONS FOR DNA TESTING

Human migration patterns, population studies
Genealogy
Genetic testing for characteristics (diseases)
Forensic identity checking
Paternity testing

8. ONE-STEP WEBSITE

Some useful DNA utilities can be found in the DNA section of the One-Step website at http://stevemorse.org