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 functions
A gene 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 triplet 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 → Gln	AAA → Lys	GAA → Glu
TAG → stop	CAG → Gln	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

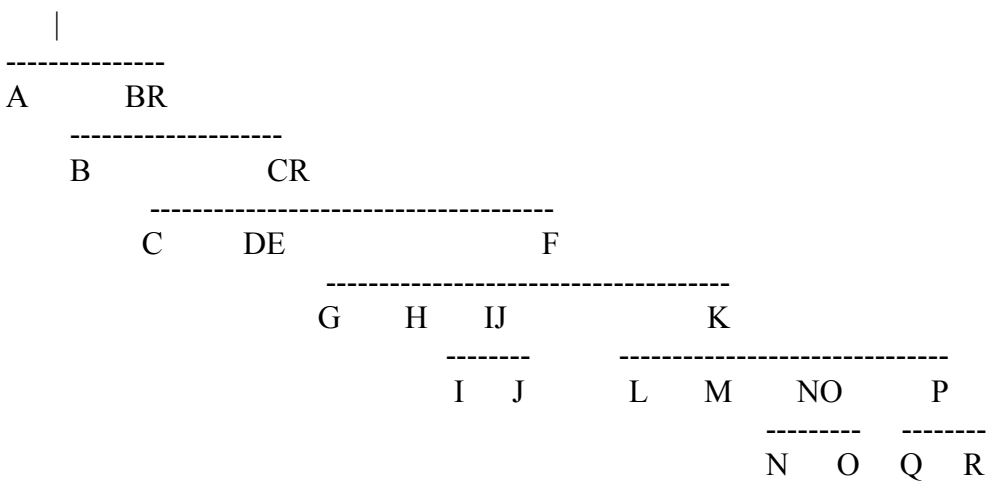
Marker: region in chromosome that is tested

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. APPLICATIONS

Genghis Khan Dynasty

800 years ago Khan conquered empire from Pacific Ocean to Caspian Sea.
Today 1 in 12 men in that region carry a common Y chromosome.
Lineage of that Y chromosome originated about 1,000 years ago.
Conclusion: Original mutation was probably Khan's great-great-grandfather or thereabout.

Anastasia Mystery

Russian royal family was murdered in 1918, including princess Anastasia.
In 1922, Anna Anderson claimed to be Anastasia and having escaped the massacre.
Anderson died in 1984, with her true identity still a mystery.
In 1991 bodies found that could be royal family, but son and one daughter were missing.
Could missing daughter be Anastasia, giving credence to Anna Anderson's claim?
DNA testing with Britain's Prince Philip matched, proving it was the royal family.
Anna Anderson's DNA was tested – it did not match the royal family's DNA.
In 2007 bodies of a young male and young female were found.
DNA testing showed they were the missing son and daughter.

Thomas Jefferson Affair

Jefferson was alleged to have fathered several children with his slave Sally Hemings.
Jefferson's grandchildren maintained that nephews Peter and Samuel Carr were the fathers.
Descendants of Jefferson, Carr, and Hemings were found and tested in 1998.
Findings showed no link between Carr's descendants and Hemings' descendants.
Findings did show a link between Jefferson's descendants and Hemings' descendants.

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. ONE-STEP WEBSITE

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