Genotyping:
Determining our Differences

Michelle Amaral, PhD

UAB Heflin Center

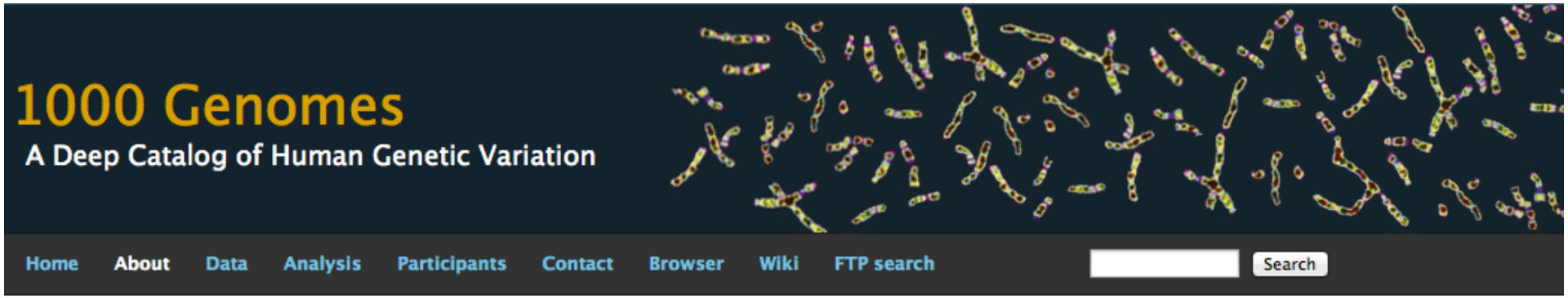
Genotyping

- Determine the DNA sequence of any organism
- Compare that sequence to a reference sequence
- Looking for variation between the sequences
  - Find causes of species diversity
  - Understand cause of disease

Reference Genomes

- Human Reference Genome
  - Combination of anonymous donors
  - >70% from one male in Buffalo, NY

www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/



www.1000genomes.org

British in England and Scotland (**GBR**)	no	yes	0	89	92	94
Finnish in Finland (**FIN**)	no	no	0	93	99	100
Iberian populations in Spain (**IBS**)	no	yes	0	14	107	107
Toscani in Italy (**TSI**)	no	no	66	98	108	110
Utah residents with Northern and Western European ancestry (**CEU**)	no	yes	94	85	99	103
Total European Ancestry (EUR)			**160**	**379**	**505**	**514**
Colombian in Medellin, Colombia (**CLM**)	no	yes	0	60	94	95
Mexican Ancestry in Los Angeles, California (**MXL**)	no	yes	0	66	67	69
Peruvian in Lima, Peru (**PEL**)	yes	yes	0	0	86	86
Puerto Rican in Puerto Rico (**PUR**)	yes	yes	0	55	105	105
Total Americas Ancestry (AMR)				**181**	**352**	**355**
Total			**553**	**1092**	**2535**	**2577**

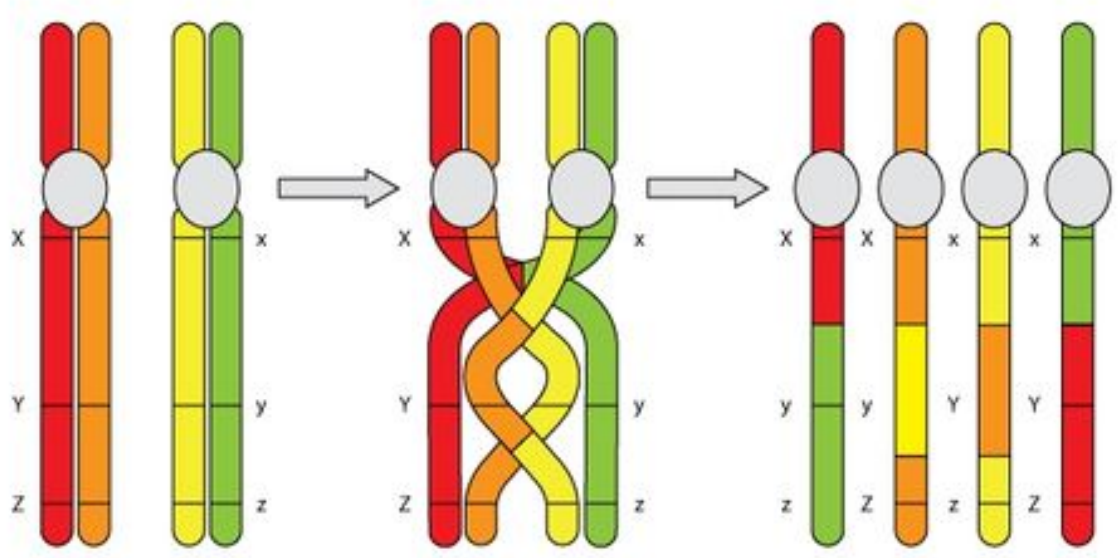
1000 Genomes Samples						
Population	**DNA sequenced from blood**	**Offspring Samples from Trios Available**	**Pilot Samples**	**Phase 1 Samples**	**Final Phase Sample**	**Total**
Chinese Dai in Xishuangbanna, China(**CDX**)	no	yes	0	0	99	99
Han Chinese in Bejing, China (**CHB**)	no	no	91	97	103	106
Japanese in Tokyo, Japan (**JPT**)	no	no	94	89	104	105
Kinh in Ho Chi Minh City, Vietnam (**KHV**)	yes	yes	0	0	101	101
Southern Han Chinese, China (**CHS**)	no	yes	0	100	108	112
Total East Asian Ancestry (ASN)			**185**	**286**	**515**	**523**
Bengali in Bangladesh (**BEB**)	no	yes	0	0	86	86
Gujarati Indian in Houston,TX (**GIH**)	no	yes	0	0	106	106
Indian Telugu in the UK (**ITU**)	yes	yes	0	0	103	103
Punjabi in Lahore,Pakistan (**PJL**)	yes	yes	0	0	96	96
Sri Lankan Tamil in the UK (**STU**)	yes	yes	0	0	103	103
Total South Asian Ancestry (SAN)			**0**	**0**	**494**	**494**
African Ancestry in Southwest US (**ASW**)	no	yes	0	61	66	66
African Caribbean in Barbados (**ACB**)	yes	yes	0	0	96	96
Esan in Nigeria (**ESN**)	no	yes	0	0	99	99
Gambian in Western Division, The Gambia (**GWD**)	no	yes	0	0	113	113
Luhya in Webuye, Kenya (**LWK**)	no	yes	102	97	101	116
Mende in Sierra Leone (**MSL**)	no	yes	0		85	85
Yoruba in Ibadan, Nigeria (**YRI**)	no	yes	106	88	109	116
Total African Ancestry (AFR)			**208**	**246**	**669**	**691**

FROM WHERE DO THESE VARIATIONS ARISE?

DNA Mutations

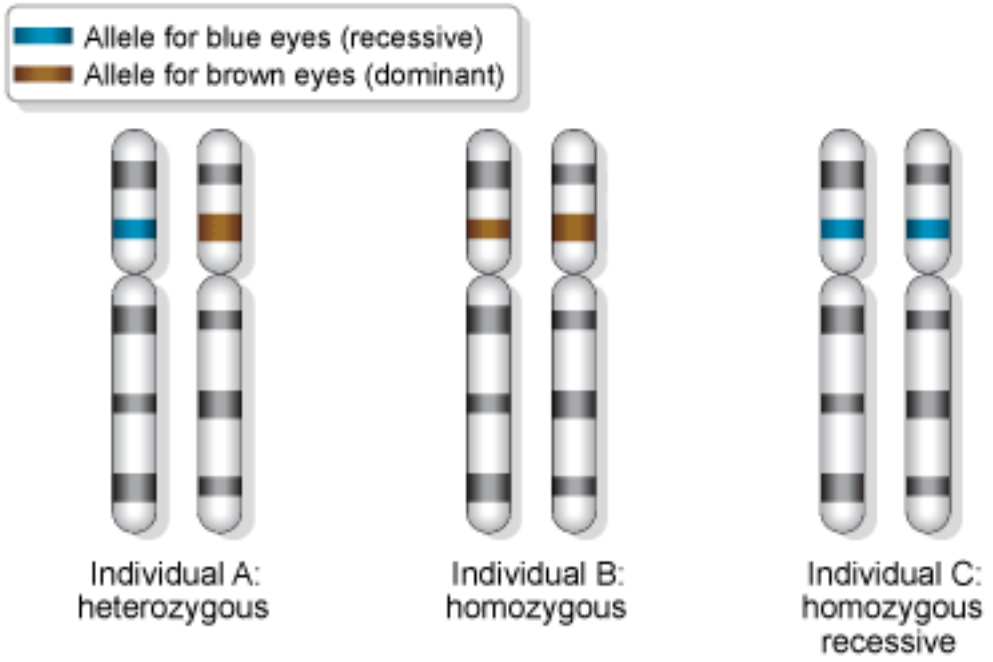
- Spontaneous errors in replication
- Aging
- Mutagenic agents
- Irradiation
- Viruses

Crossing Over in Meiosis



Definition: Allele

- Allele: alternative forms of the same gene
- Homozygote: 2 copies of same allele
- Heterozygote: 2 different alleles



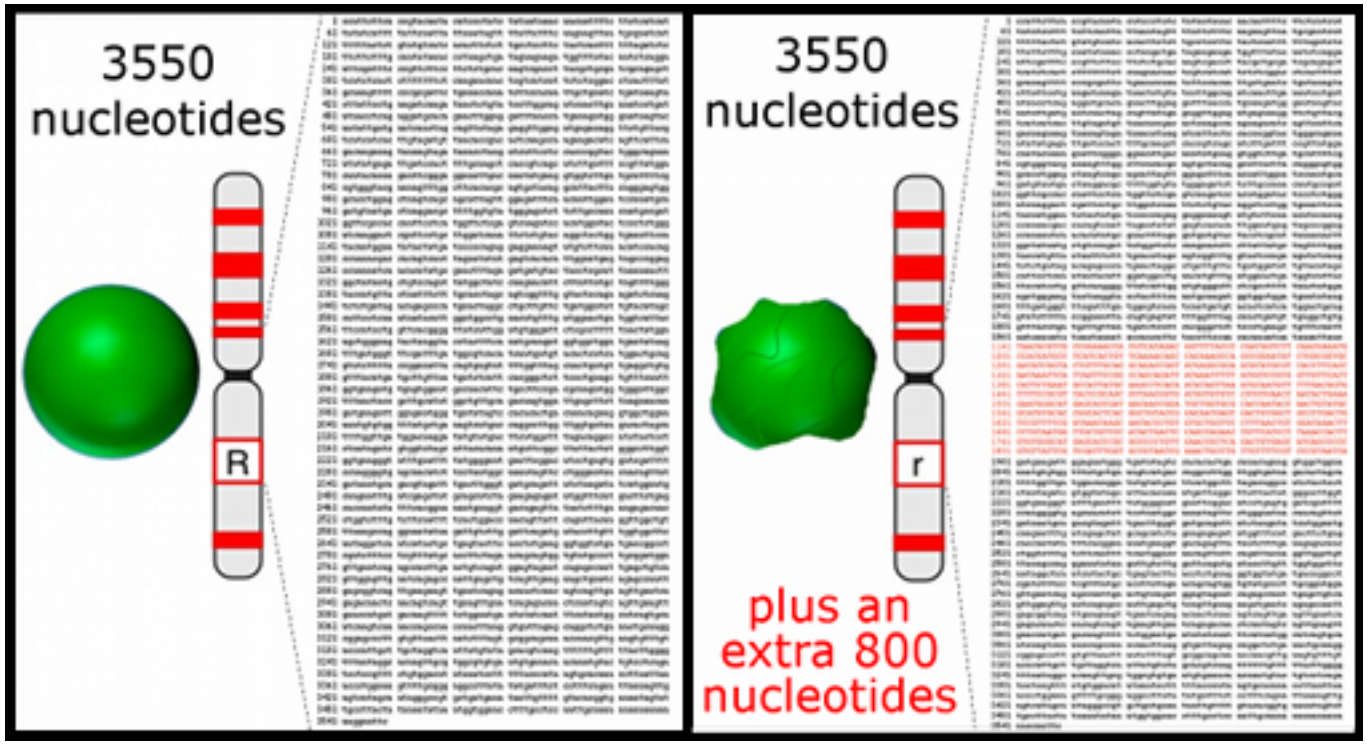
“Ginger Gene”

- Melanocortin 1 receptor (*MC1R*)
- V60L allele caused ginger gene
- It lightened the skin and allowed people to get more vitamin D from weaker sunlight
- But increased their vulnerability to melanomas
- Common in Europe
- Recessive

EXAMPLES

Round & Wrinkled Peas

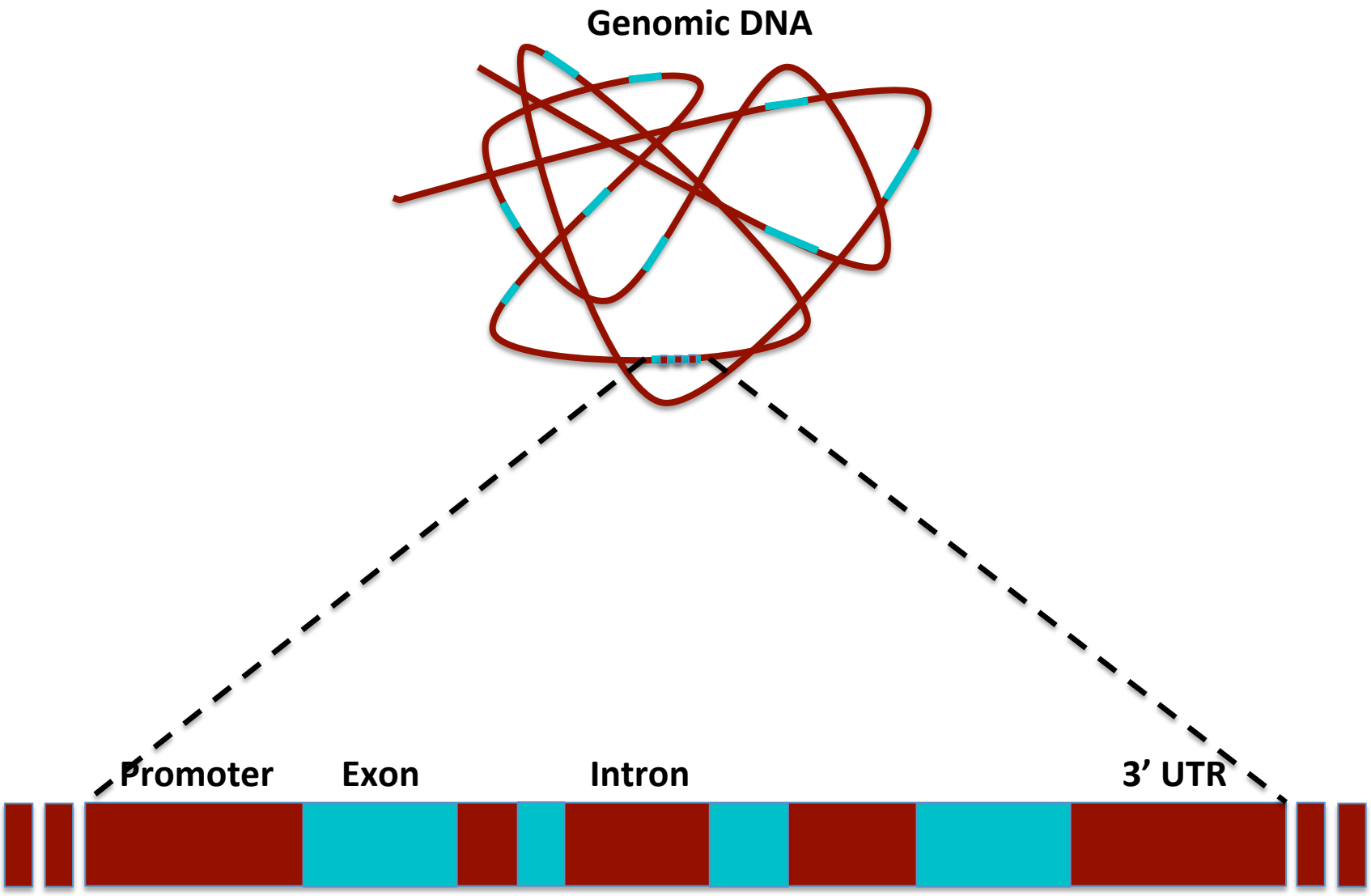
- *Sbe1* – starch branching enzyme I
- *rr* peas taste sweeter



Practical Uses

- Investigating disease
- Paternity testing
- Genotyping transgenic organisms

GENOMIC LOCATION OF VARIABILITY


Genomic DNA
Promoter
Exon
Intron
3’ UTR

Variation in Promoter Region

- Possibilities
  - prevent or promote transcription factor binding
  - alter RNA polymerase binding
  - produce unstable RNA polymerase/transcription factor binding
  - **all of these could alter gene expression**
- Could also have no effect

Variation in Exon Region

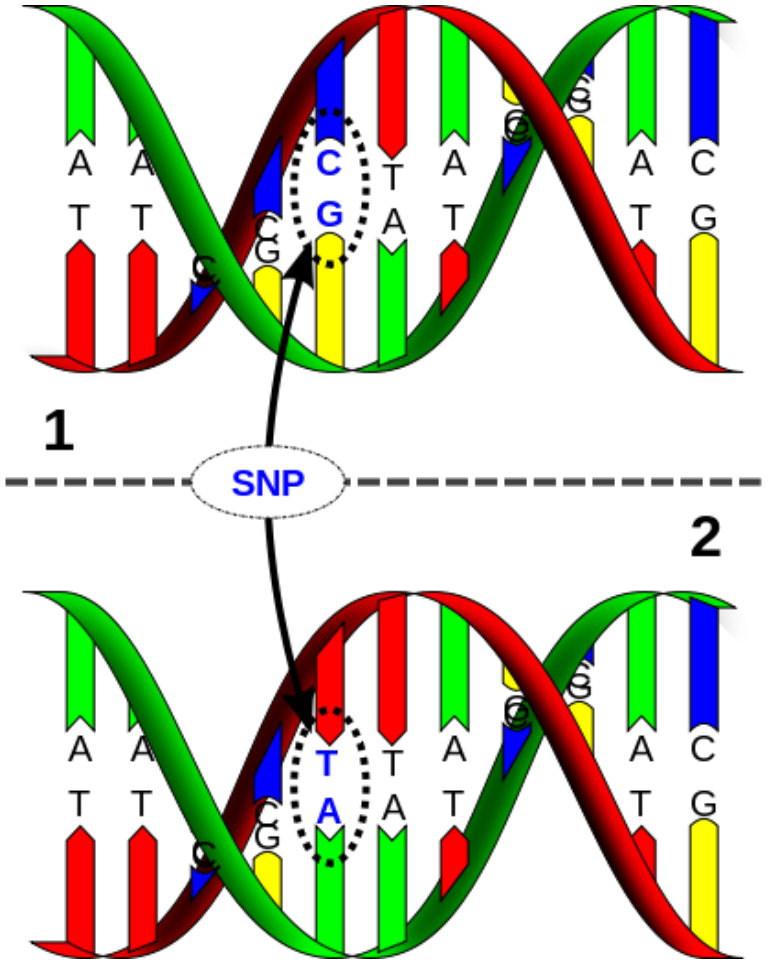
- Possibilities
  - alter amino acid sequence
  - produce truncated protein product
  - produce unstable or non-functional protein product
- Could have no effect

Variation in Introns & 3' UTR

- Possibilities
  - Alter RNA splicing and processing
  - Affect stability of mRNA
  - Alter protein product
  - Promote degradation of the mRNA
- Could have no effect

TYPES OF DNA VARIATION

Single Nucleotide Polymorphism (SNP)



Example: Cold Sore Susceptibility

- SNP in the promoter region for IL28B
- IL28B consistently mutated in people w/sores
- Therefore, can't make a functional protein and mount adequate immune response
- Same people w/ IL28B mutation – less likely to respond to hepatitis C treatment
- Therapies: acyclovir but perhaps in the future giving IL28B

Insertion/Deletion

Pathogenic indel

CARD15: 3020Cins → Frameshift → Crohn's disease

Wild type

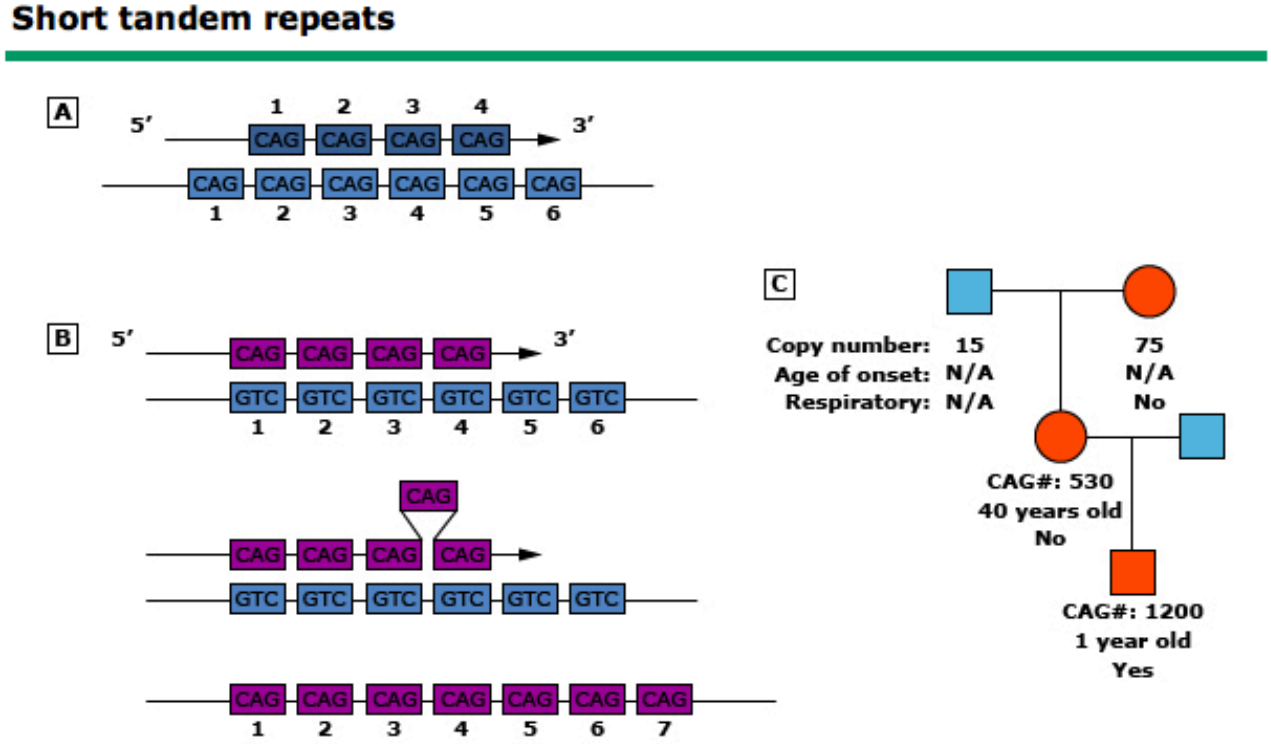
GCC	CTC	CTG	CAG	GCC	CTT	GAA	AGG	AAT	GAC
A	L	L	Q	A	L	E	R	N	D

Mutation

GCC	CTC	CTG	CAG	GCC	CCT	TGA	AAG	GAA	TGA
A	L	L	Q	A	P	*			

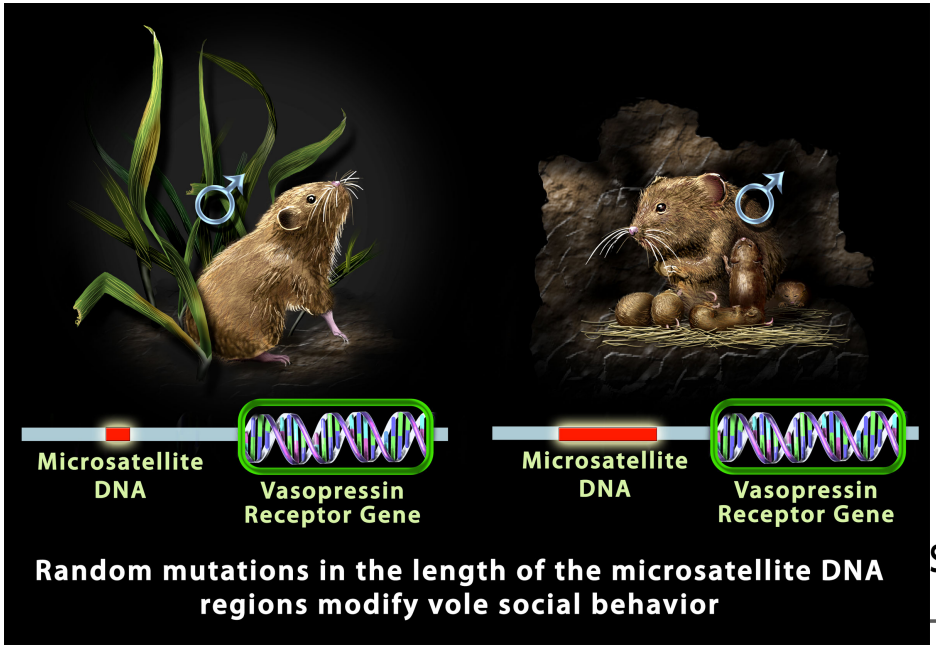
Insertion of Cytosine at nucleotide position 3020 of the CARD15 gene results in frameshift mutation inducing a stop codon sequence in the following codon (denoted by *).

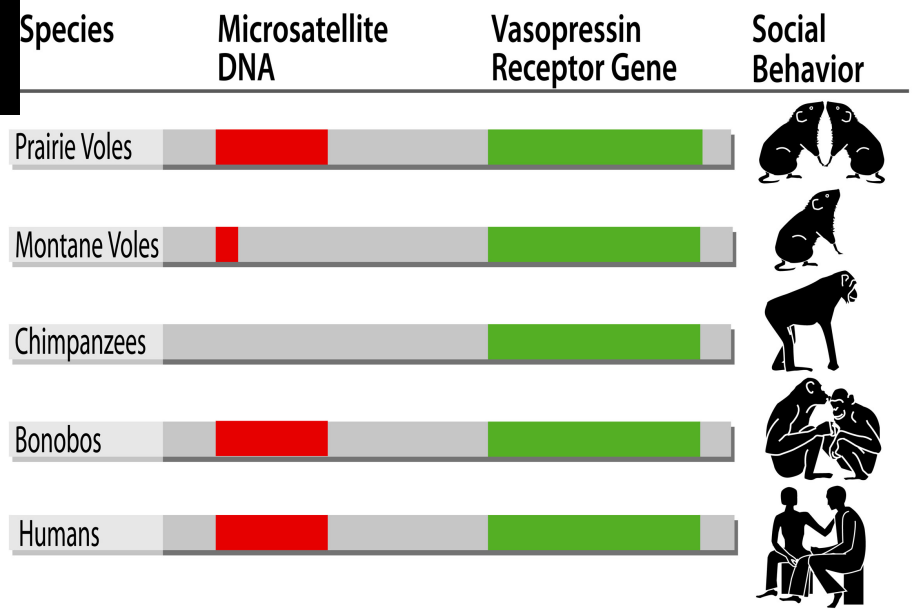
Short Tandem Repeats



(A) Example of a CAG short tandem repeat with genotype 4/6. (B) Slip mispairing: During DNA replication, nascent DNA copy (in purple) slips back one tandem repeat, resulting in duplicate replication of tandem repeat #5. Results in additional copy (total 7) on new strand. (C) Pedigree of Myotonic Dystrophy in three generations with anticipation. Grandmother has 75 tandem repeats at the DMPK gene, no symptoms. Mother has expansion of 530 repeats, with mild symptoms (no respiratory distress) at age 40. Grandson has progressive expansion to 1200 repeats, manifesting with myotonia and respiratory failure at one year of age.

Microsatellites





Copy Number Variation

- Fady Mikhail, MD, PhD