

# **Molecular Basis of Genetic Variation**

**L. Elliot Hong, MD**

**Professor and Director  
UM Center for Brain Imaging Research  
Maryland Psychiatric Research Center  
University of Maryland School of Medicine, USA  
[www.mdbrain.org](http://www.mdbrain.org)**

# Acknowledgements

**Trygve Bakken, PhD**

University of California at San Diego

– Taught this course at HBM 2010

**Thomas Nichols, PhD**

University of Warwick

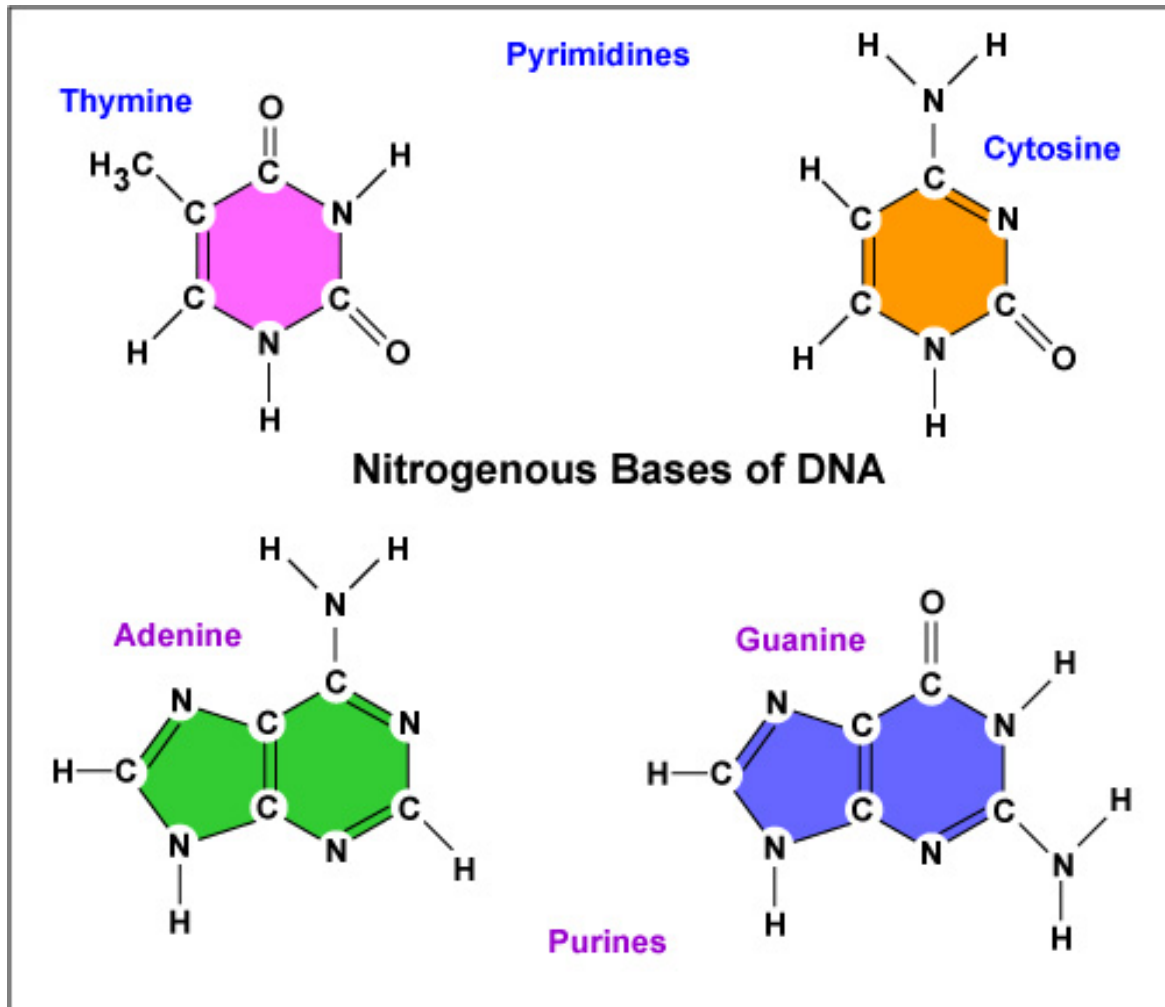
– Taught this course at HBM 2011

Drs. Bakken and Nichols provided most of the structure, content and many slides for this lecture

# Outline

- Commonly used Genetic Concepts and Terms in Imaging Genetics
- Transcription & Translation
- Recombination & Linkage Disequilibrium
- Polymorphism & Mutation types

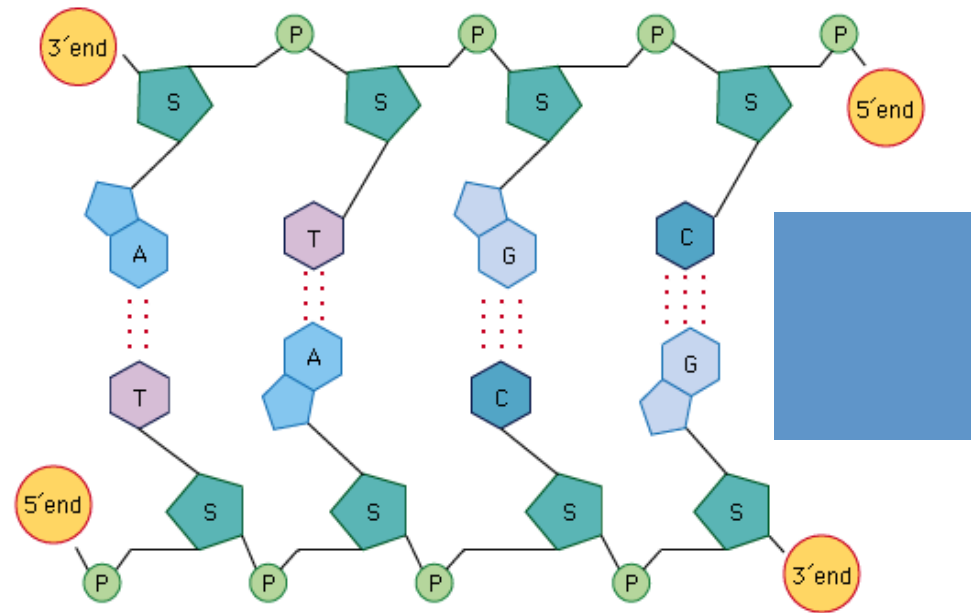
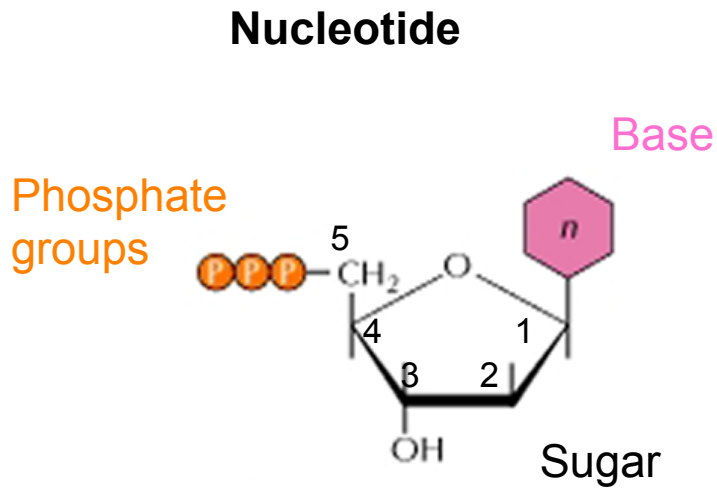
# The Bases – T C A G



T and C

A and G

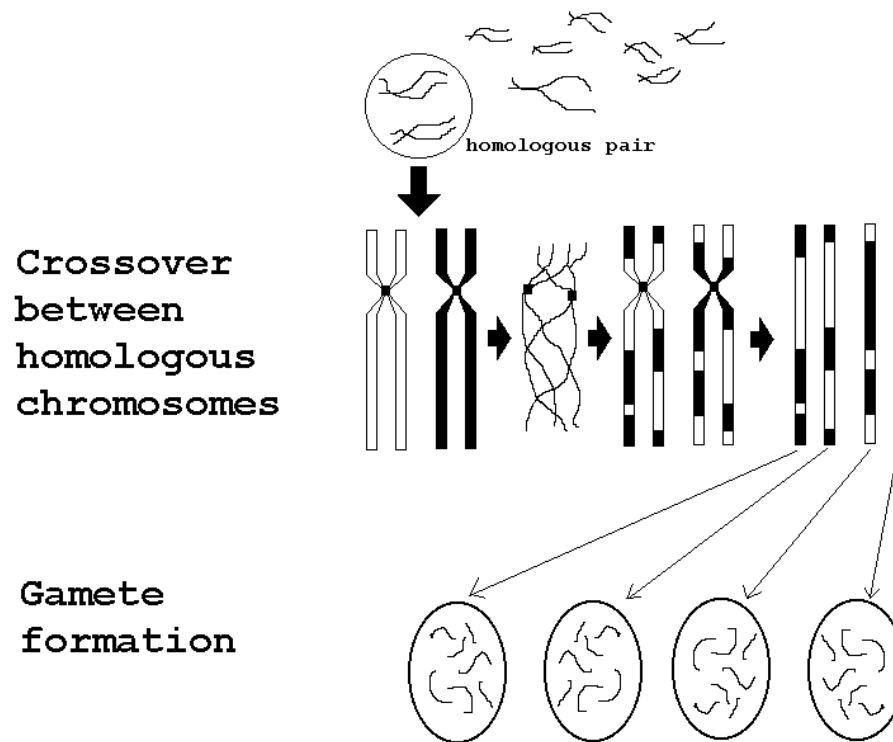
# DNA sugar-phosphate backbone



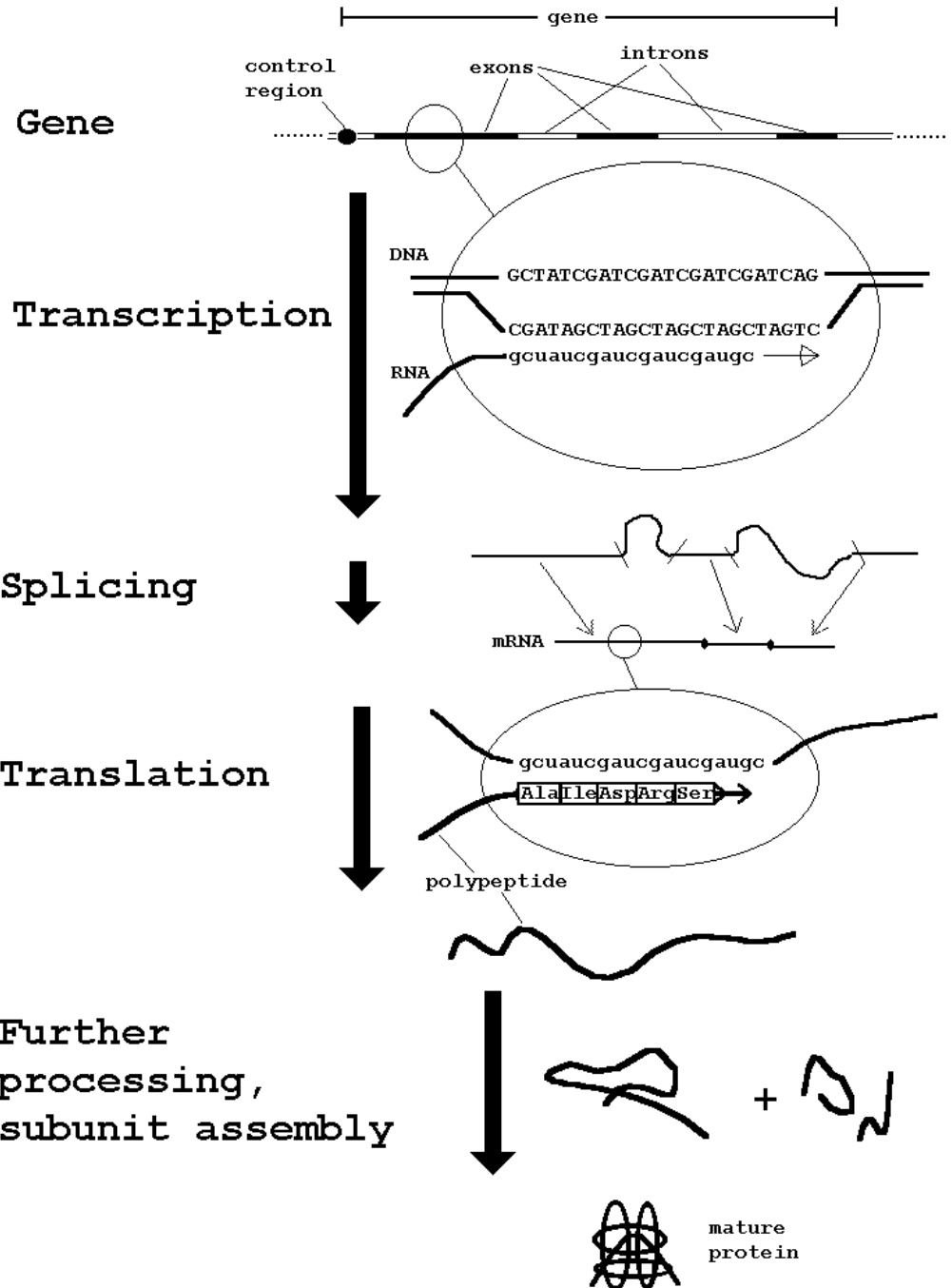
©1998 Encyclopaedia Britannica, Inc.

# Cross-over and Recombination

- During **meiosis**, two chromosomes, one from mom and one from dad, twist around each other
- Large segments of DNA are exchanged and **recombined**
- **Gamete** is formed, each carries copies from both sides of the parents



# Protein synthesis



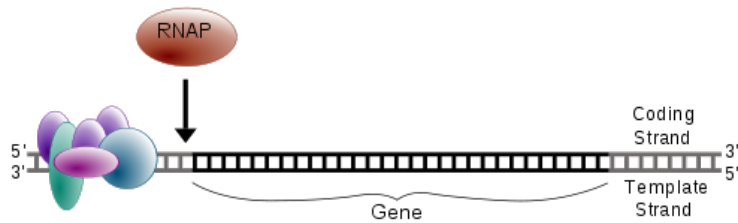
## Key concepts

- Promoter region
- Exon
- Intron
- Splicing

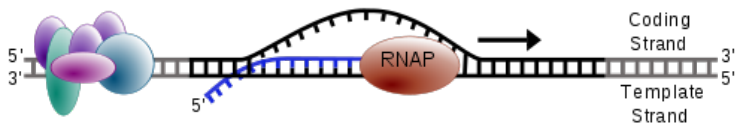
# Transcription of DNA into RNA

Ribonucleic acid (RNA), double helix like DNA, except

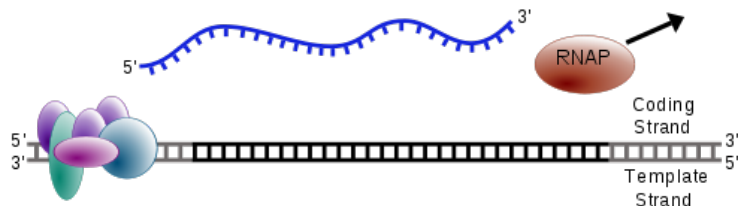
DNA bases: Adenine (A) Thymine (T) Guanine (G) Cytosine (C)    A – T    G – C  
RNA bases: Adenine (A) Uracil (U)    Guanine (G) Cytosine (C)    A – U    G – C



1. Initiation: transcription factors mediate RNA polymerase (RNAP) binding



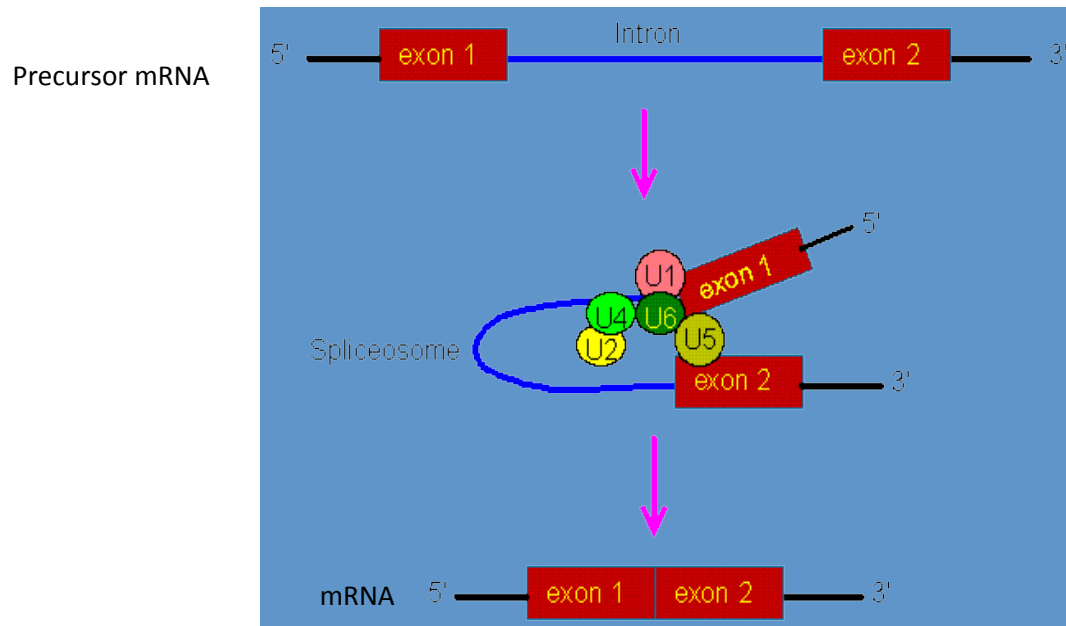
2. Elongation: RNAP reads template DNA strand



3. Termination: single-stranded pre-mRNA is released



# Splicing messenger RNA (mRNA)



- Exons contain information to make protein
- Spliceosome removes introns from pre-mRNA
- Alternate splicing generates protein diversity

# Genetic code specifies rules to make proteins

## Codon - Amino Acid mapping

Amino Acid	Codons					
Isoleucine			AUU	AUC	AUA	
Phenylalanine			UUU	UUC		
Valine			GUU	GUC	GUA	GUG
Leucine	UUA	UUG	CUU	CUC	CUA	CUG
Methionine						AUG
Tryptophan						UGG
Alanine			GCU	GCC	GCA	GCG
Glycine			GGU	GGC	GGA	GGG
Cysteine			UGU	UGC		
Tyrosine			UAU	UAC		
Proline			CCU	CCC	CCA	CCG
Threonine			ACU	ACC	ACA	ACG
Serine	AGU	AGC	UCU	UCC	UCA	UCG
Histidine			CAU	CAC		
Glutamate					GAA	GAG
Asparagine			AAU	AAC		
Glutamine					CAA	CAG
Aspartate			GAU	GAC		
Lysine					AAA	AAG
Arginine	AGA	AGG	CGU	CGC	CGA	CGG
STOP	UGA				UAA	UAG

## Second bp

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

64 codons, 20 amino acids

- Genetic code is unambiguous and redundant
- Code is highly conserved across all organisms

# 3 Types of Genetic Codes

1. 1 codon (AUG) encodes methionine **and** starts translation of all protein
2. 61 codons encode 20 amino acids
3. 3 codons stop protein translation

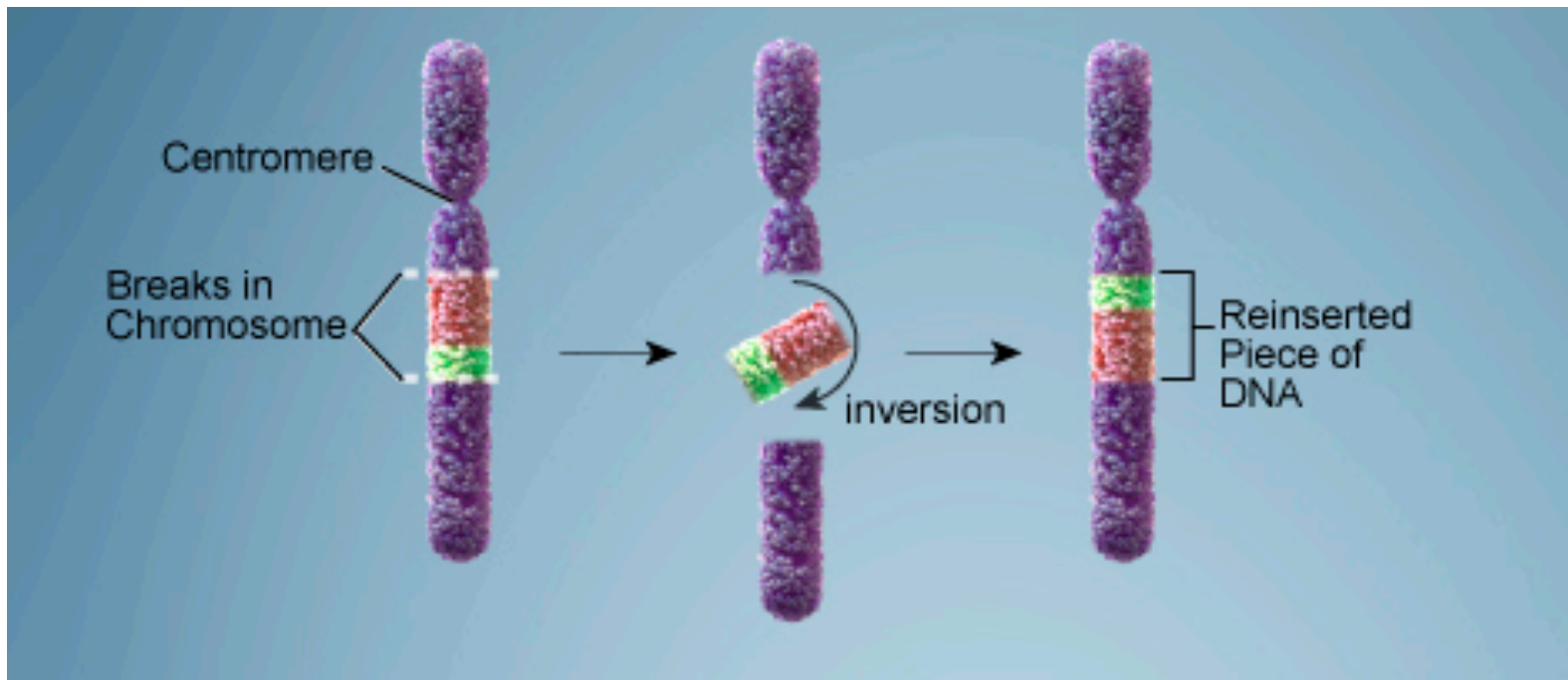
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

# Mutation at the Chromosome Level

- Inversion
- Deletion
- Translocation
- Duplication
- Nondisjunction

# Inversion

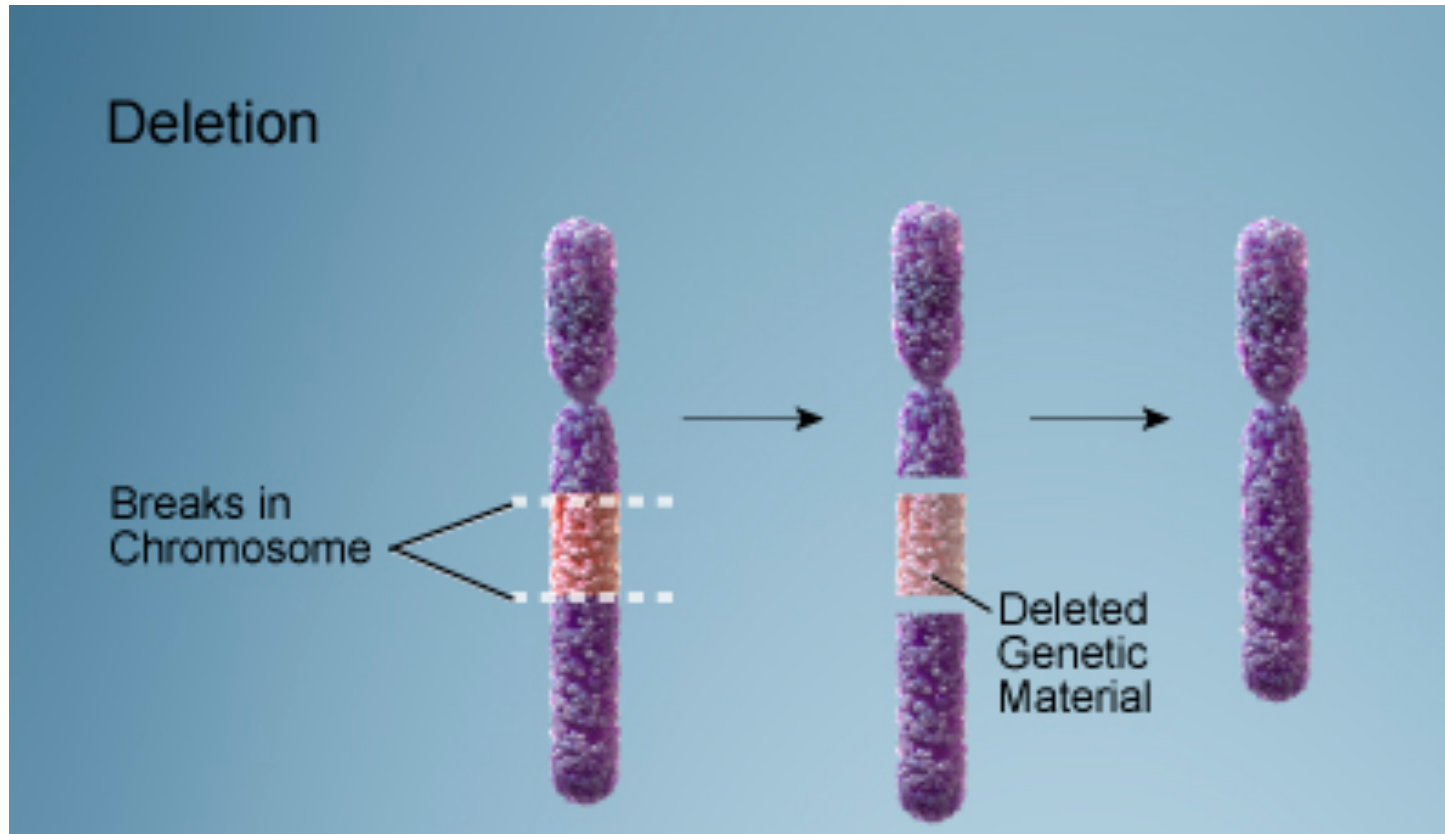
- Chromosome segment **breaks off**
- Segment flips around **backwards**
- Segment **reattaches**



From

# Deletion

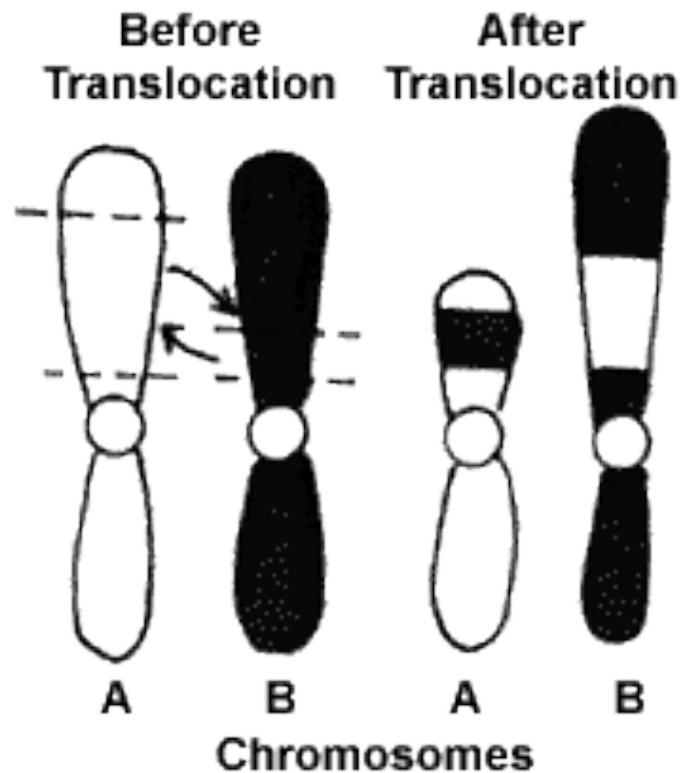
A **piece** of a chromosome is **lost**



From

# Translocation

- Part of one chromosome is transferred to **another chromosome** that is **not identical**



From

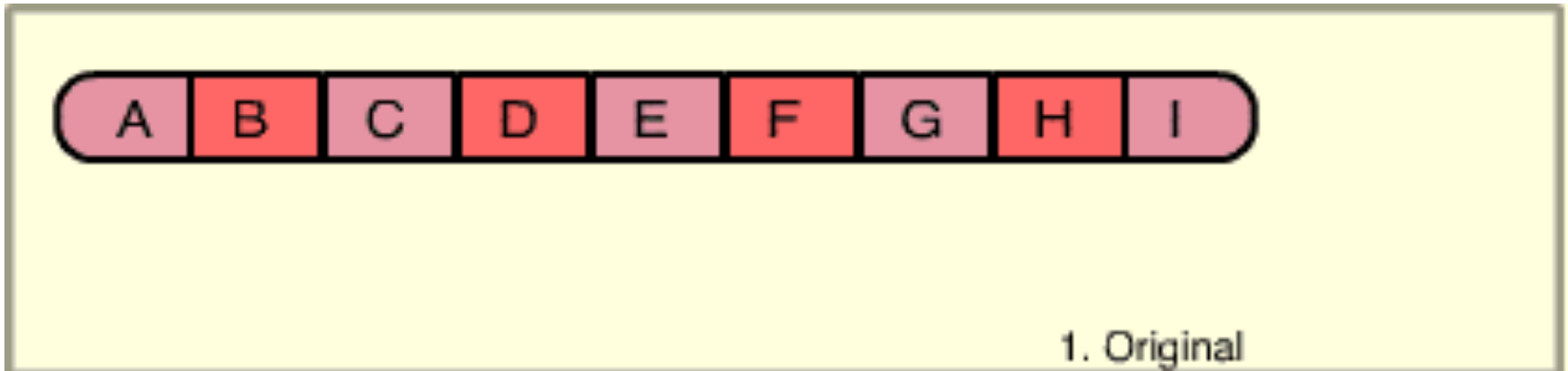




# Nondisjunction

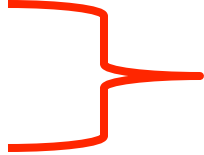
- Failure of chromosomes to separate during **meiosis**
- Causes gamete to have too many or too few chromosomes
  - **Down Syndrome**

# Chromosome Mutation Animation

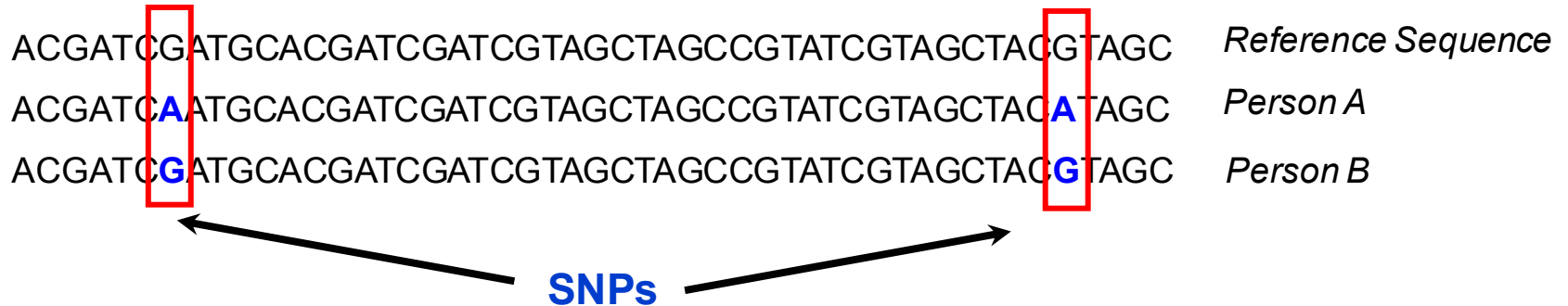


- They are often **copy number variants** (CNV)
- About 10 % of all genetic variants are CNV
- Fully discovered by human genome project
- 0.4% of humans differ in copy numbers

# Mutation at the Gene Level

- Insertion
  - Deletion
  - Frameshift
  - Substitution
-  **Indel:** CNV in smaller scale

# Single nucleotide polymorphisms (SNPs)



- Variation at a single nucleotide
- Purine to purine (A - G)
- Pyrimidine and pyrimidine (C - T)
- Purine – pyrimidine: less often

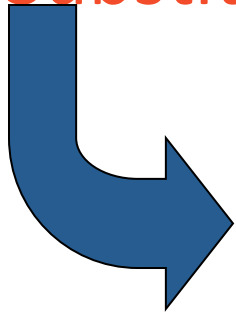
# Gene Mutation Animation



- 2 humans differ about 1 out of 1000 bases = 3–4 million
- About 90% of all genetic variants are SNP
- Much more abundant sources for the variance in brain and behavior

# Mutation at the Gene Level continues...

- Substitution

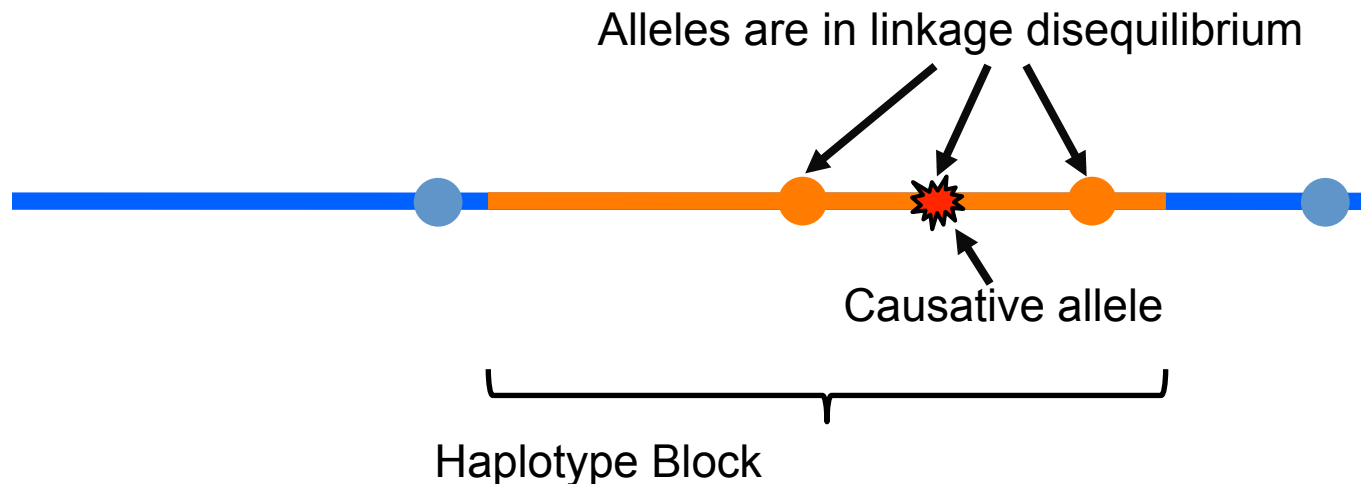


- Intronic SNP
- Synonymous SNP – silent
- Nonsynonymous SNP – missense, nonsense
- Nonsense SNP – premature stop
- SNP at 5' , 3' , promotor, stop codon
- SNP at splicing site

# SNP for Imaging Genetics

SNP is a proxy vs. SNP is **causative**

- SNP as mere marker for another functional variant in **linkage disequilibrium** – still need to find that variant
- SNP itself is functional and causative



# Gene Structure and SNP Alleles

- **Locus**: the position of a gene is called a locus
- **Allele**: the exact form of the gene is called allele
- Two copies of the same chromosome in a cell
- Therefore, two physical copies of each gene in a cell
- Each gene exist in the form of 0, 1, or 2 alleles

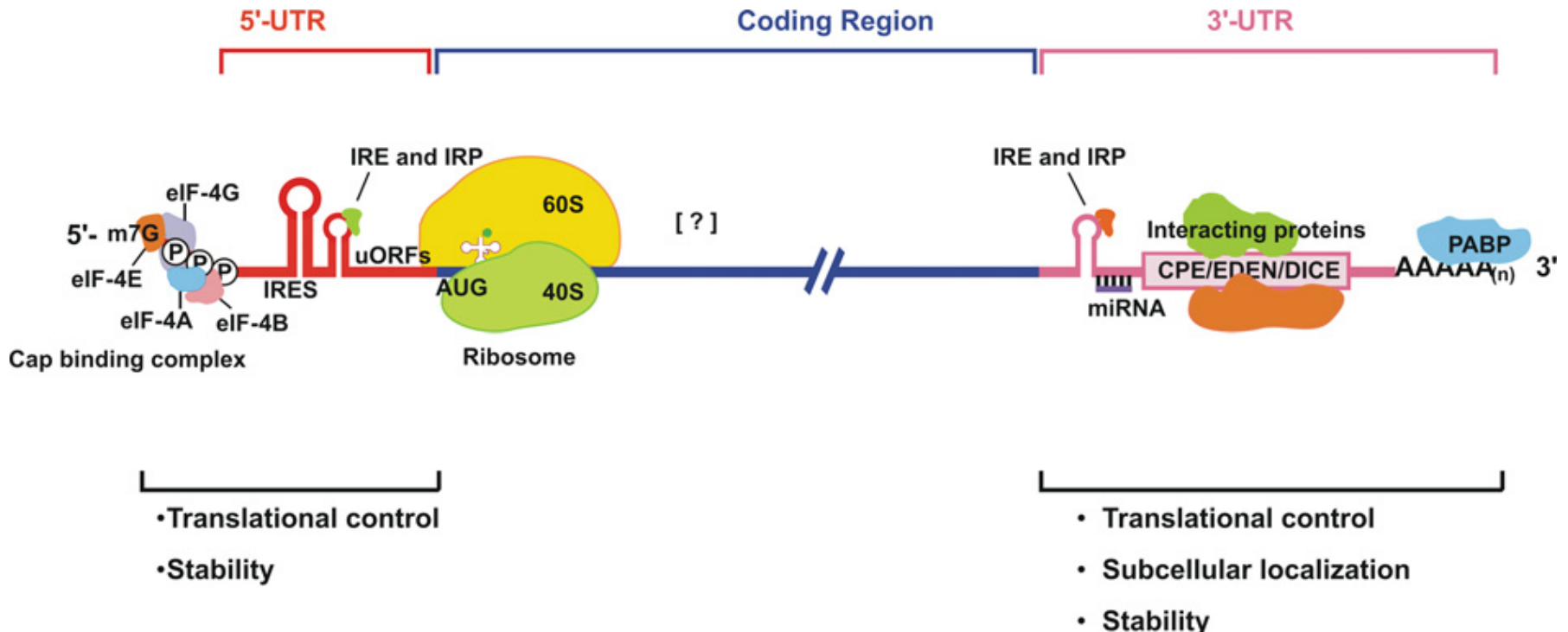


# SNP Allele Count

- One allele is the **minor allele** – the one less frequent in a population. Not always, but often is the **risk allele**
- The other is the **major allele**
- An individual can have 0, 1, or 2 copies of this risk allele
- **Additive**: each risk allele contributes some to phenotype
- **Recessive**: one risk allele has no effect; needs **homozygote**
- **Dominant**: **heterozygote** risk allele has the same effect as **homozygote** on the phenotype

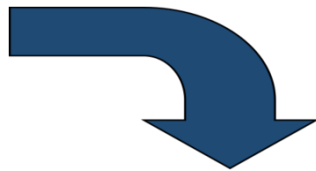
# 5' UTR and 3' UTR SNPs

- Multi-pronged machinery for protein translational control
- Not part of the protein
- Mutation leads to reduction or acceleration of translation

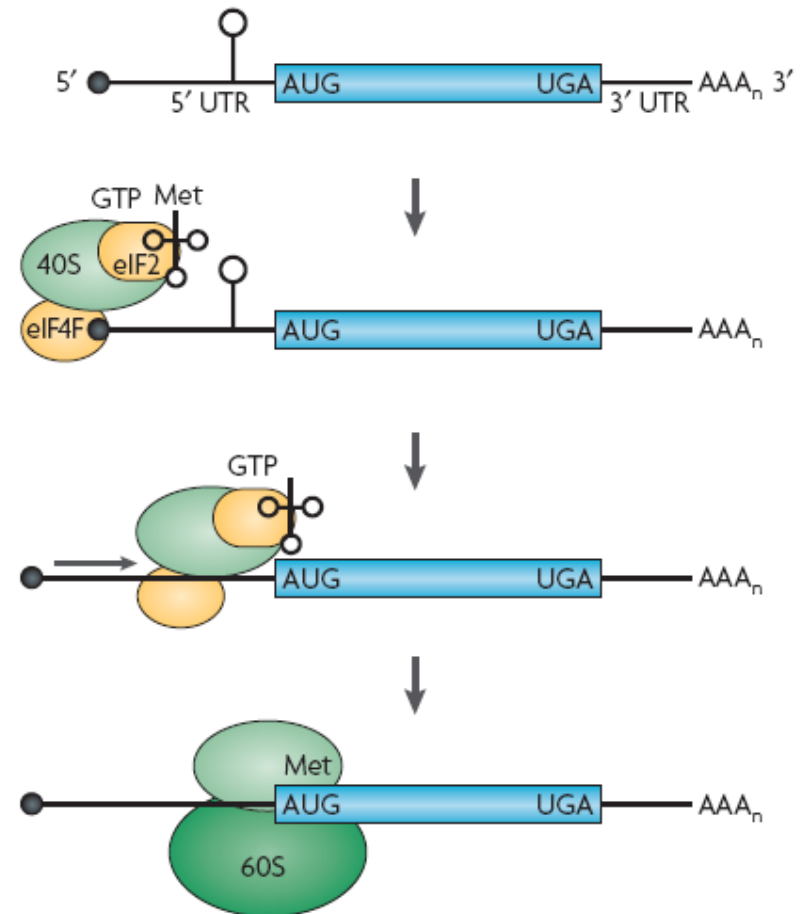


# Mutation at 5' UTR Affects Protein Translation

- Ribosome scanning
- Upstream open reading frame (uORF)
- Internal ribosome entry site (IRES)
- eIF2
- A number of other factors
- All influence the rate of mRNA translation



- Example: eIF2 variants - childhood ataxia with central hypomyelination



# Synonymous versus Nonsynonymous SNP

- SNP occurs in exons
- Synonymous substitution is **silent** and does not cause amino acid change; but **not always silent**, still may be functional
- **Nonsense** mutation: cause the formation of a **stop codon**, produce a truncated protein
- Nonsynonymous is **missense** and cause change of amino acid. Not necessarily functional, but much more likely than synonymous substitution

		Second bp				
		U	C	A	G	
U	Phe	Ser	Tyr	Cys	U	
	Phe	Ser	Tyr	Cys	C	
	Leu	Ser	STOP	STOP	A	
	Leu	Ser	STOP	Trp	G	
C	Leu	Pro	His	Arg	U	
	Leu	Pro	His	Arg	C	
	Leu	Pro	Gln	Arg	A	
	Leu	Pro	Gln	Arg	G	
A	Ile	Thr	Asn	Ser	U	
	Ile	Thr	Asn	Ser	C	
	Ile	Thr	Lys	Arg	A	
	Met	Thr	Lys	Arg	G	
G	Val	Ala	Asp	Gly	U	
	Val	Ala	Asp	Gly	C	
	Val	Ala	Glu	Gly	A	
	Val	Ala	Glu	Gly	G	

First bp Third bp

# Intronic SNPs

- Do not participate in protein coding
- Frequently identified using GWAS chips – norm rather than exception?
- Still poorly understood, but often **replicable**
- **Assume to be in LD** with another functional mutation
- Or, could be directly functional, for example, affecting mRNA folding, altering splicing site