Genetic Variation II

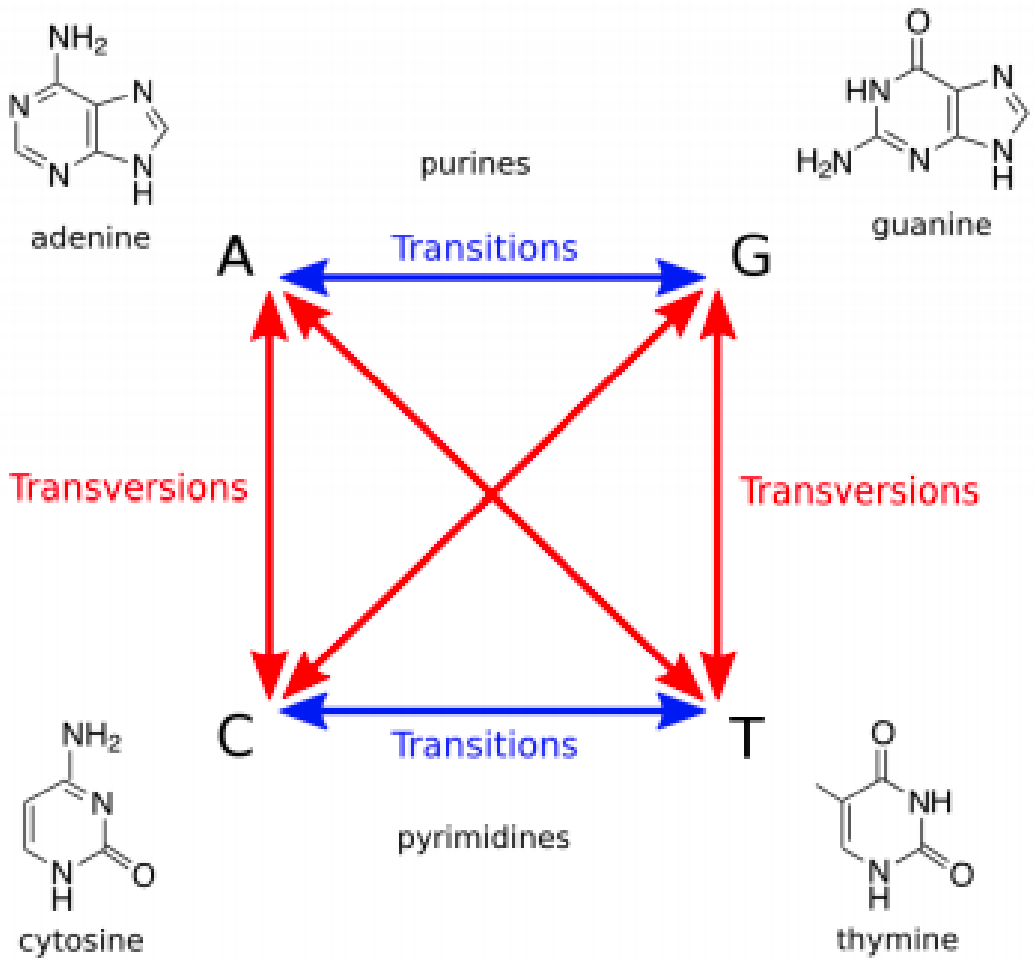
Frequency of mutations in human disease

Type of mutation	% disease causing mutations
Nucleotide substitutions	
Missense (amino acid substitution)	50%
Nonsense (premature termination codon)	10%
RNA processing (splice, polyadenylation, etc)	20%
Gene expression regulation (TF binding site, etc)	rare
Deletions & insertions	
Small indels	25%
Large rearrangements (deletion, duplication, inversion, etc)	5%
Insertion of LINE or Alu (interrupting regulation or coding)	rare
Repeat expansion	rare

Note: These data are changing!

Genetics in Medicine, 8th ed, Thompson&Thompson

Point mutations

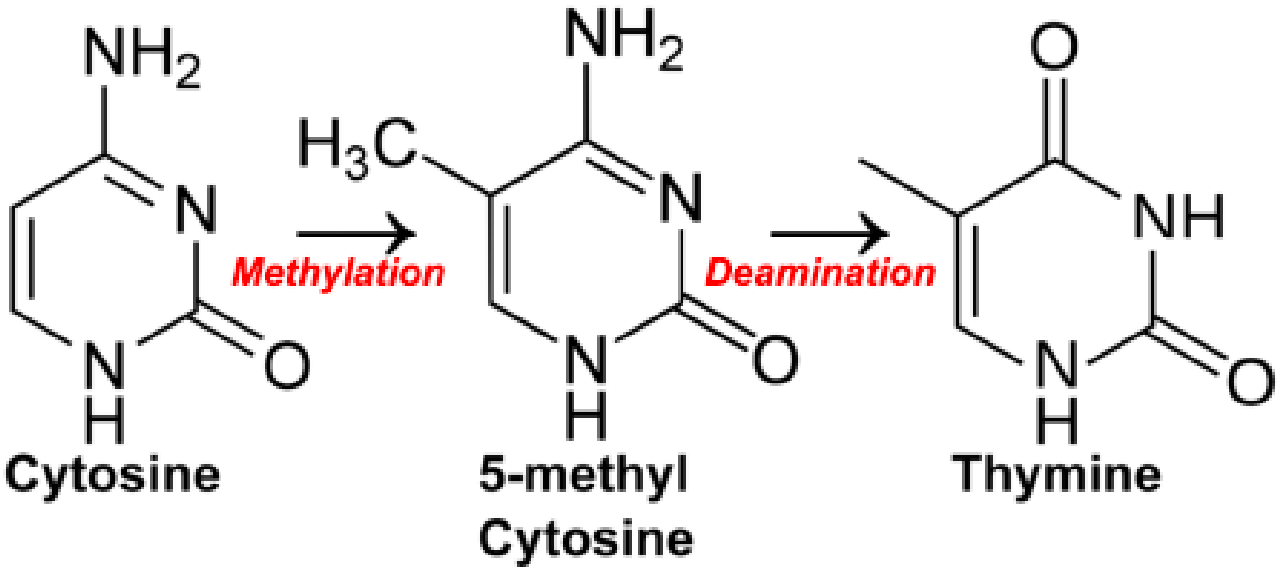


Transition:
purine to purine or
pyrimidine to pyrimidine

Transversion:
purine to pyrimidine or
pyrimidine to purine

Most common mutation: C>T transitions

- Most common type of mutation in human genome
- Due to spontaneous deamination of 5-methylcytosine to thymine



Silent (synonymous) mutations

- Do not change the amino acid (p.Ala123Ala)
- Mostly benign, but may impact splicing or RNA secondary structure!

mRNA	CAU	CAA	ACG	GGT	GCC	AAC	GGC
Protein	His	Gln	Thr	Gly	Ala	Asn	Gly

mRNA	CAU	CAA	ACG	GGT	GCU	AAC	GGC
Protein	His	Gln	Thr	Gly	Ala	Asn	Gly

May alter pre-RNA splicing

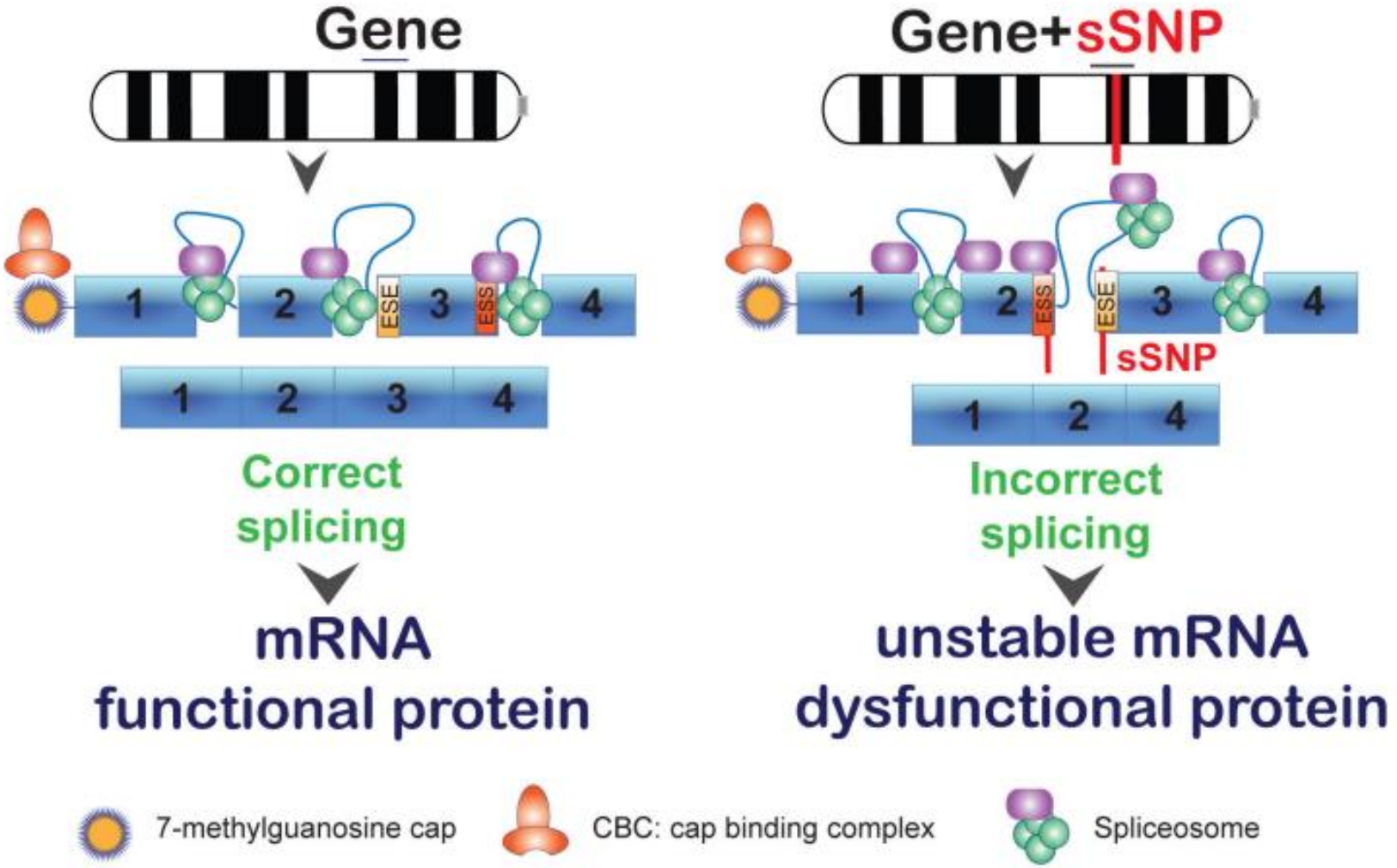
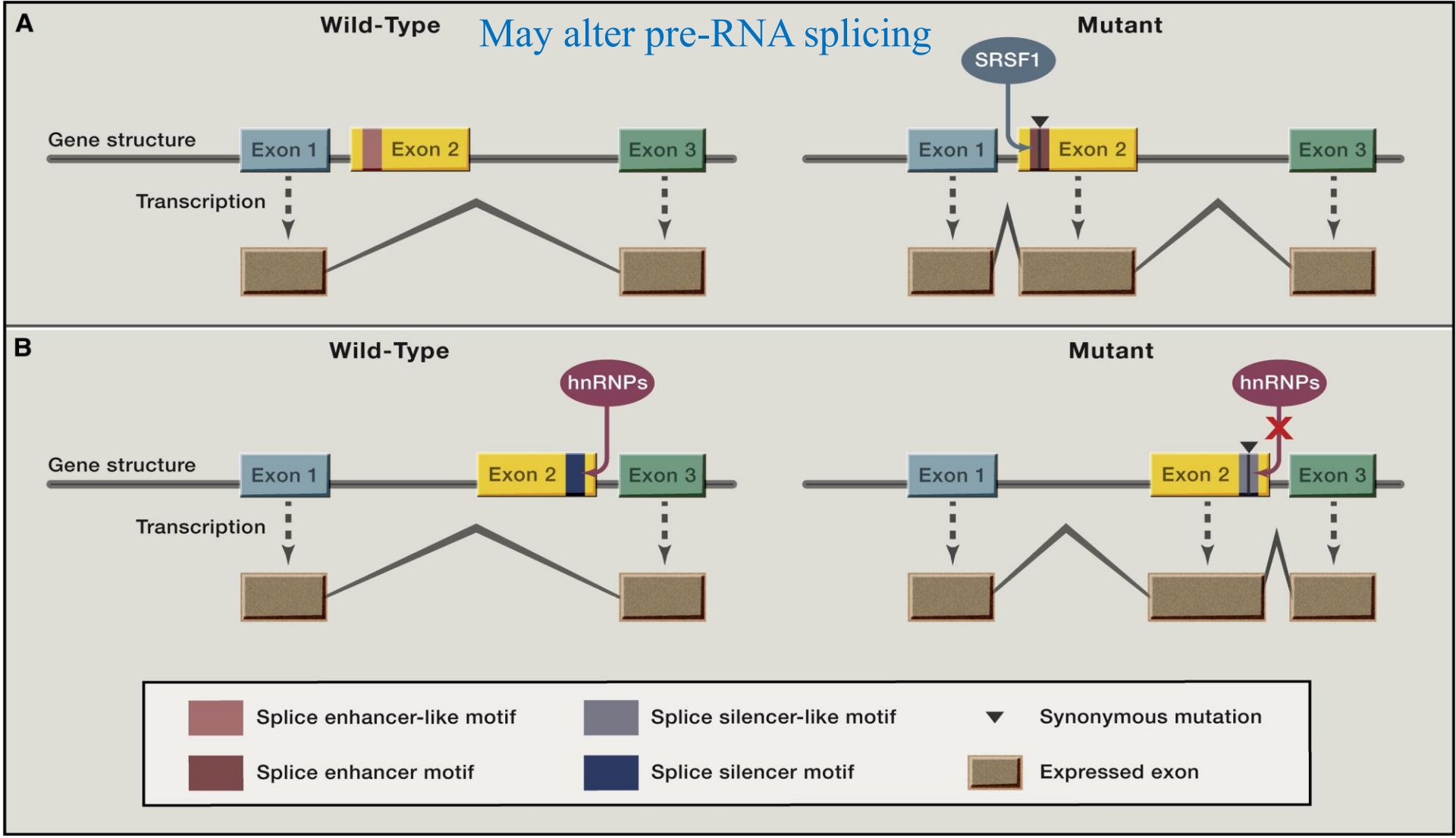


Figure 2. The consequence of synonymous mutations in exonic splice regulatory sites
sSNPs may disrupt critical elements necessary for splicing. In the example shown, this results in exon skipping. ESE: exonic splicing enhancer; ESS: exonic splicing suppressor. (For a review concerning pre-mRNA splicing refer to: (Muller-McNicoll and Neugebauer, 2013).



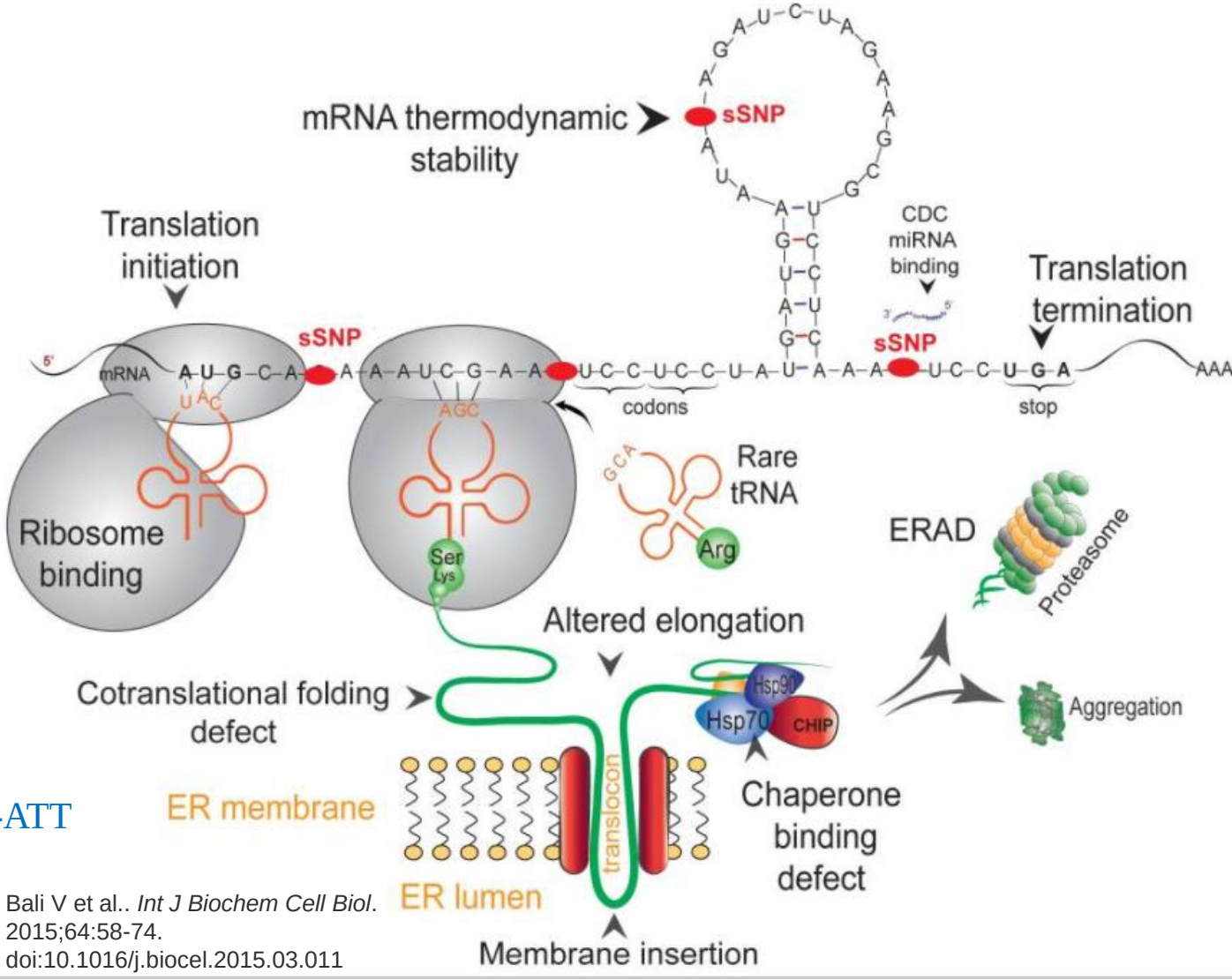
(A) A synonymous mutation leads to the gain of an exonic splicing enhancer motif. Consequently, binding of the splicing regulator SRSF1 is enhanced, resulting in the inclusion of an otherwise skipped exon.

(B) A synonymous mutation deactivates an exonic splicing silencer motif, thereby abolishing the binding of hnRNP splicing regulators

Cell 2014 1561129-1131DOI: (10.1016/j.cell.2014.02.037)

May alter mRNA secondary structure

may alter translation initiation efficiency, translation elongation rate,
ribosomal pause rhythm, cotranslational folding or the overall fate of the protein



Ile507-ATC and Ile507-ATT

Bali V et al.. *Int J Biochem Cell Biol*. 2015;64:58-74. doi:10.1016/j.biocel.2015.03.011

The consequences of a synonymous single nucleotide change on the predicted structure of the mRNA (mfold)
The predicted (mfold) structures of the Ile507-ATC and Ile507-ATT ΔF508 *CFTR* mRNAs.
The sequences represent human CFTR mRNA fragments encoding the region of NBD1 near the ΔF508 mutation. The locations of the altered nucleotides (C and U) are highlighted in red.

Missense (Non-synonymous) mutations

- Change the amino acid (substitution)
- Conservative: new amino acid has similar properties as the original (polar to polar, hydrophobic to hydrophobic, etc)
- Non-conservative: new amino acid has different properties than the original (polar to nonpolar, hydrophobic to hydrophilic, etc)
- May be benign or pathogenic

Example: *HBB* c.17A>T (p.Glu6Val)

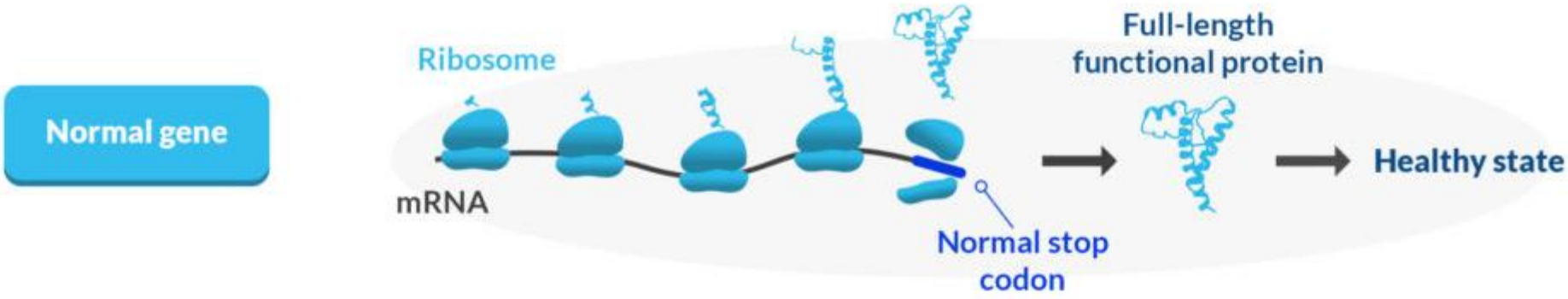
	1	2	3	4	5	6	7	8	9
	Val	His	Leu	Thr	Pro	Glu	Glu	Lys	Ser
NORMAL	GTG	CAT	CTG	ACT	CCT	GAG	GAG	AAG	TCT
SICKLE	GTG	CAT	CTG	ACT	CCT	GTG	GAG	AAG	TCT
	Val	His	Leu	Thr	Pro	Val	Glu	Lys	Ser

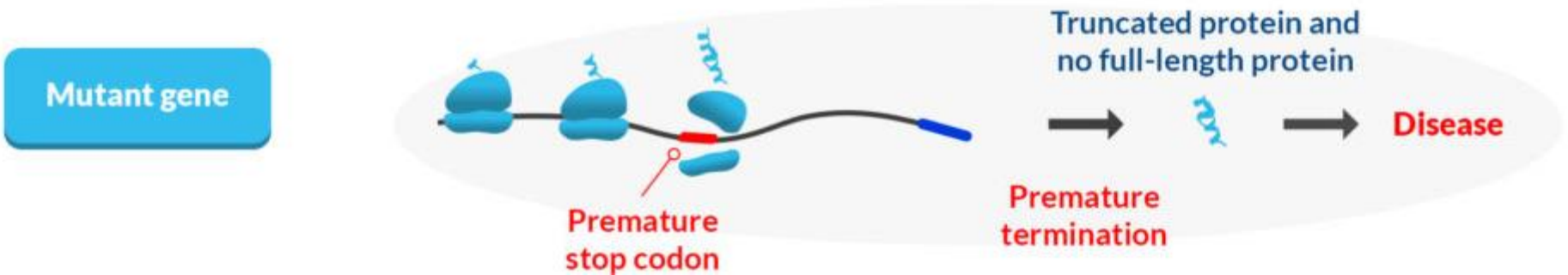
Glutamate change to Valine at position 6 of *HBB* gene encoding β-globin

Leads to β-globin protein aggregates ⟶ Causes Sickle cell anemia

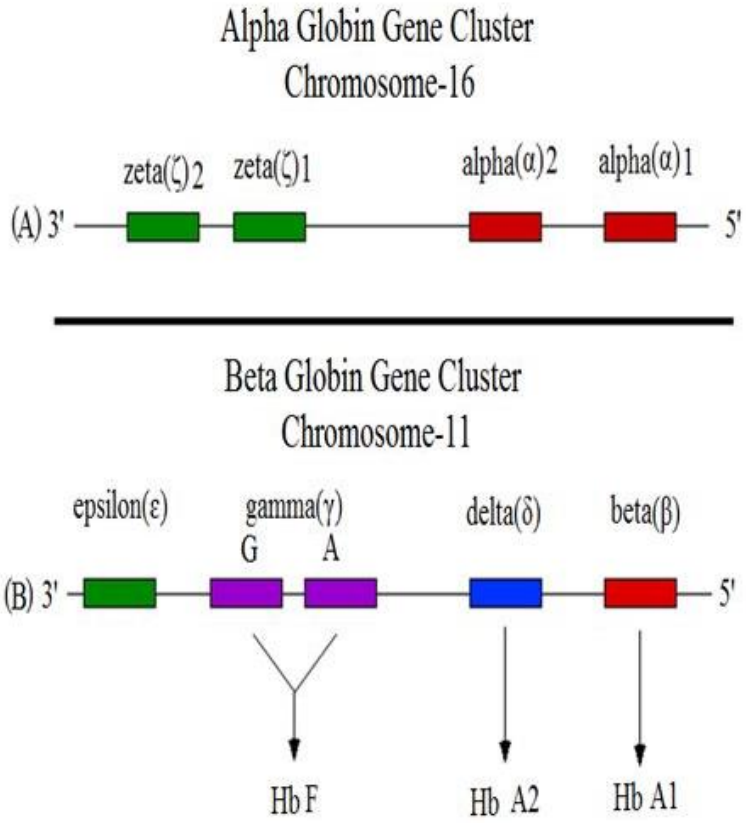
Nonsense mutations

- Cause errors in translation
- Change a codon to a termination codon (UAA, UAG, UGA)
- May result in nonsense mediated decay (NMD), truncated protein, or splicing impact
- Not always pathogenic!





- Type-1: Mutant Alpha(α) globin genes responsible for Alpha(α) thalassemia and
- Type-2: Mutant Beta(β) genes responsible for Beta(β) thalassemia.



Example: *HBB* c.118C>T (p.Gln39*)

	31	32	33	34	35	36	37	38	39
	Leu	Leu	Val	Val	Tyr	Pro	Trp	Thr	Gln
NORMAL	CTG	CTG	GTG	GTC	TAC	CCT	TGG	ACC	CAG
β^0	CTG	CTG	GTG	GTC	TAC	CCT	TGG	ACC	TAG
	Leu	Leu	Val	Val	Pro	Pro	Trp	Thr	STOP

Creates premature termination codon and leads to NMD

Homozygotes: No β-globin protein ⟶ β-thalassemia

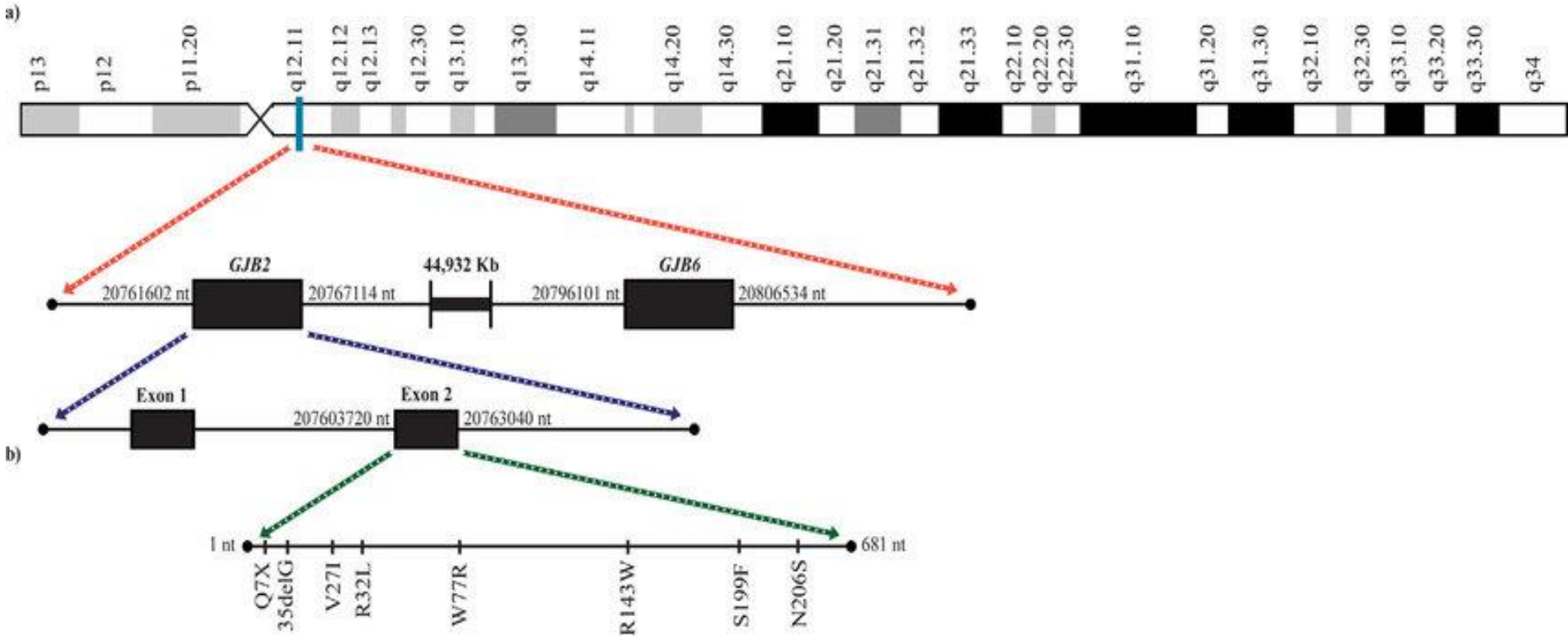
Frameshift mutations

- Cause errors in translation
- Alters the mRNA reading frame
- Often lead to a premature termination codon downstream
- Not always pathogenic!

Example: *GJB2* c.35delG (p.Gly12fs)

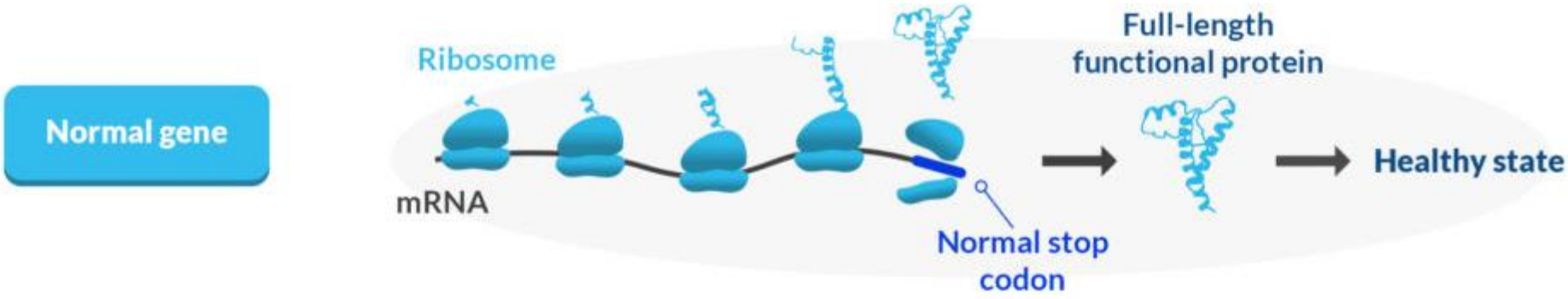
Changes Glycine at position 12 to a Valine and leads to premature termination codon downstream

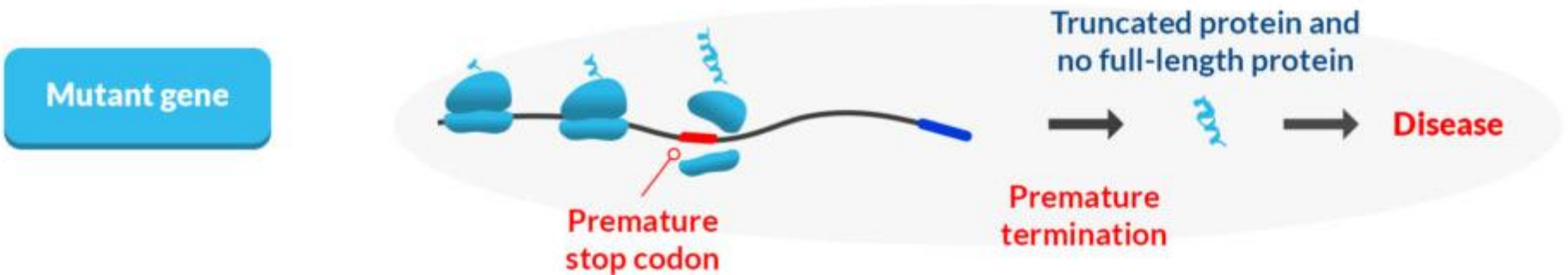
Homozygotes: Non-syndromic hearing loss


Chromosome 13
a)
p13
p12
p11.20
q12.11
q12.12
q12.13
q12.30
q13.10
q13.30
q14.11
q14.20
q14.30
q21.10
q21.20
q21.31
q21.32
q21.33
q22.10
q22.20
q22.30
q31.10
q31.20
q31.30
q32.10
q32.30
q33.10
q33.20
q33.30
q34
GJB2
44,932 Kb
GJB6
20761602 nt
20767114 nt
20796101 nt
20806534 nt
Exon 1
Exon 2
207603720 nt
20763040 nt
b)
1 nt
681 nt
Q7X
35delG
V27I
R32L
W77R
R143W
S199F
N206S

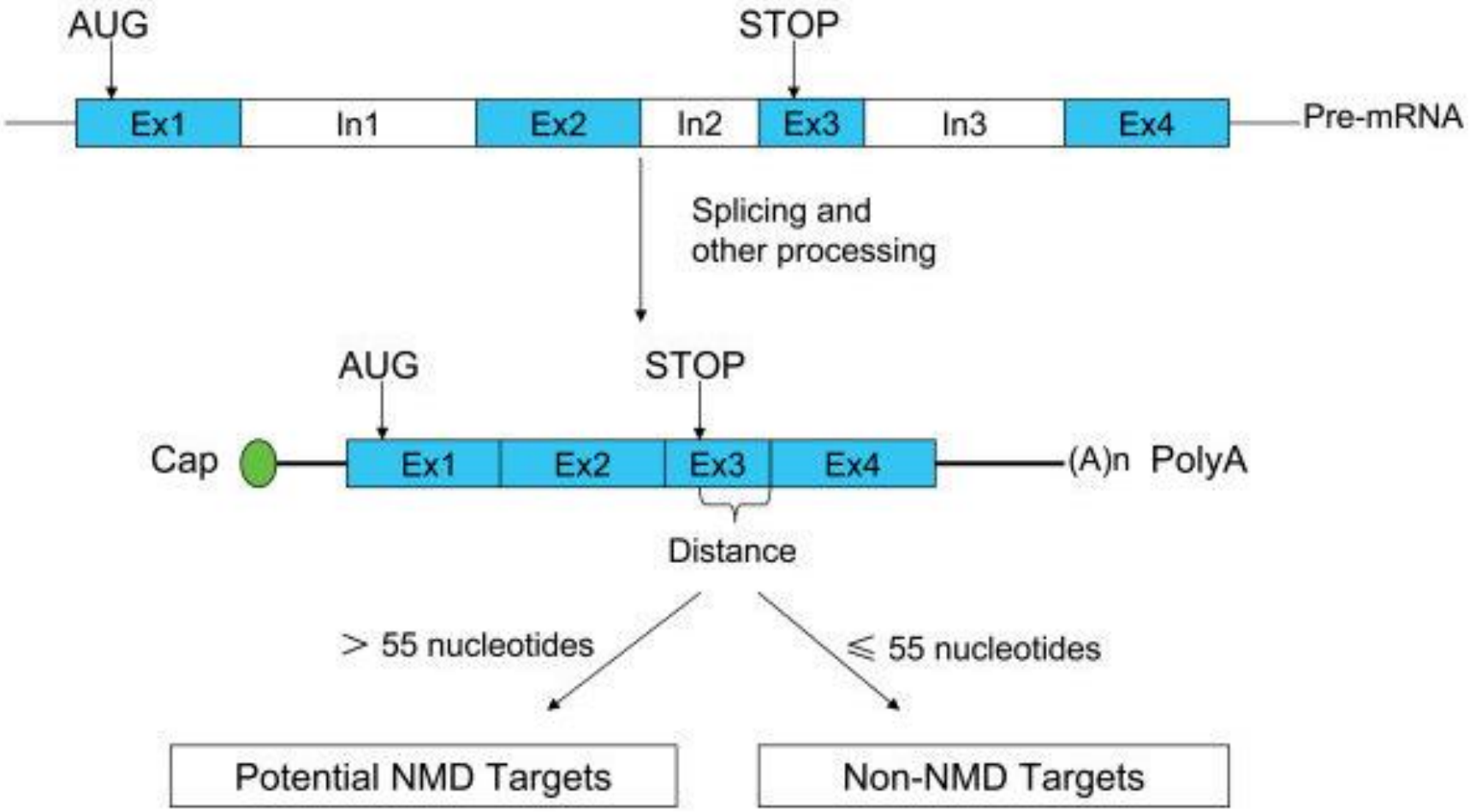
Nonsense mutations

- Cause errors in translation
- Change a codon to a termination codon (UAA, UAG, UGA)
- May result in nonsense mediated decay (NMD), truncated protein, or splicing impact
- Not always pathogenic!



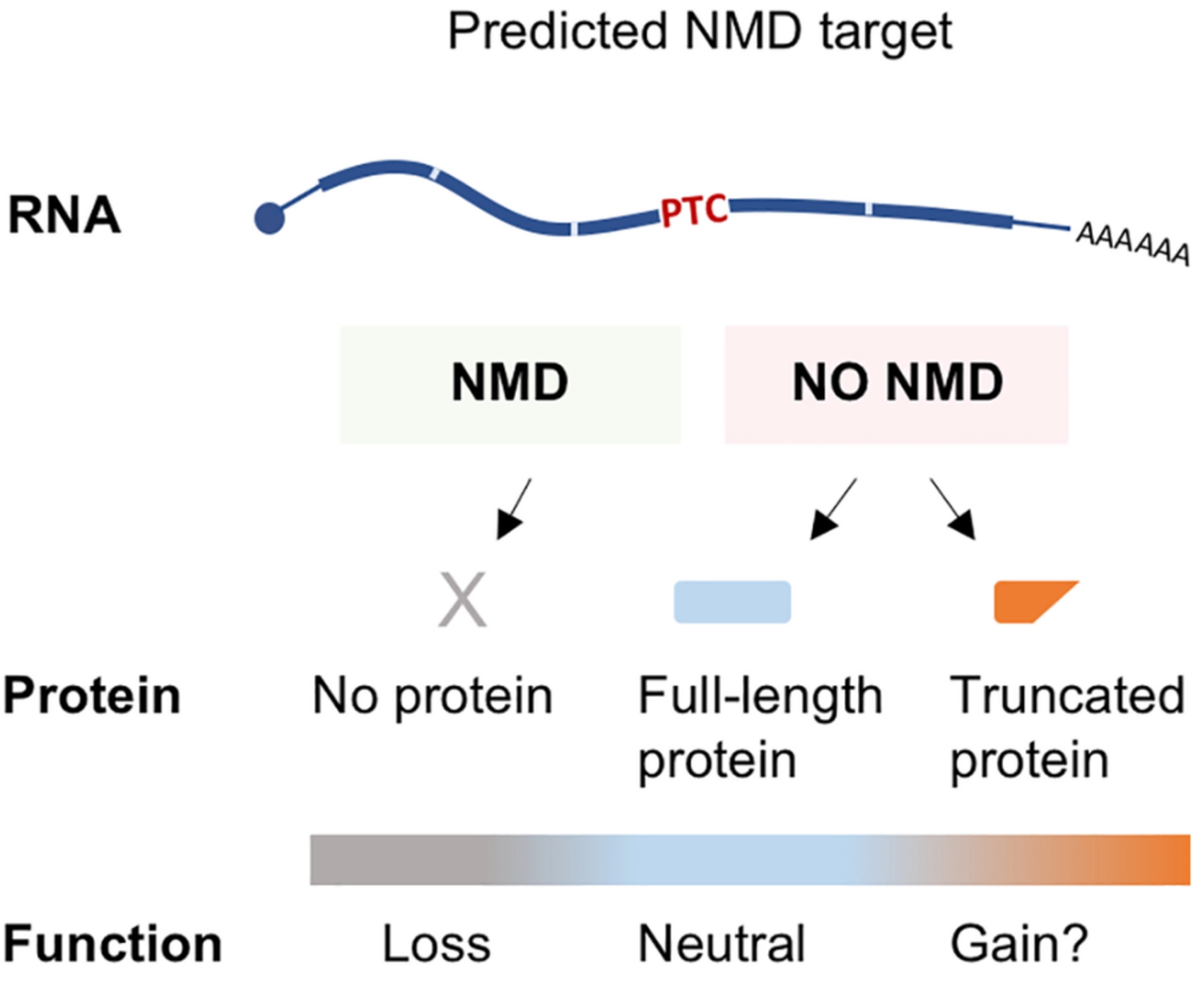


nonsense-mediated mRNA decay (NMD)



Ex*: Exons **In*:** Introns **AUG**: start codon **STOP**: termination codon

50 to 55 nucleotides upstream of the 3' most splice-generated exon-exon junction


Predicted NMD target
RNA
PTC
AAAAAA
NMD
NO NMD
Protein
No protein
Full-length protein
Truncated protein
Function
Loss
Neutral
Gain?

In-frame deletions and insertions

- Deletions or insertions of bases in multiples of 3 (3,6,9,…)
- Lead to deletions or insertions of amino acids without altering the reading frame
- May be benign or pathogenic

Example: CFTR c. (p.Phe508del – ΔF508)

Normal	ATC	ATC	TTT	GGT	GTT
	Ile	Ile	**Phe**	Gly	Val
ΔF508	ATC	ATT	GGT	GTT	
	Ile	Ile	Gly	Val	

Leads to deletion of phenylalanine at position 508 of CFTR protein

Block in processing of the protein ⟶ Cystic fibrosis

Figure 17.12-

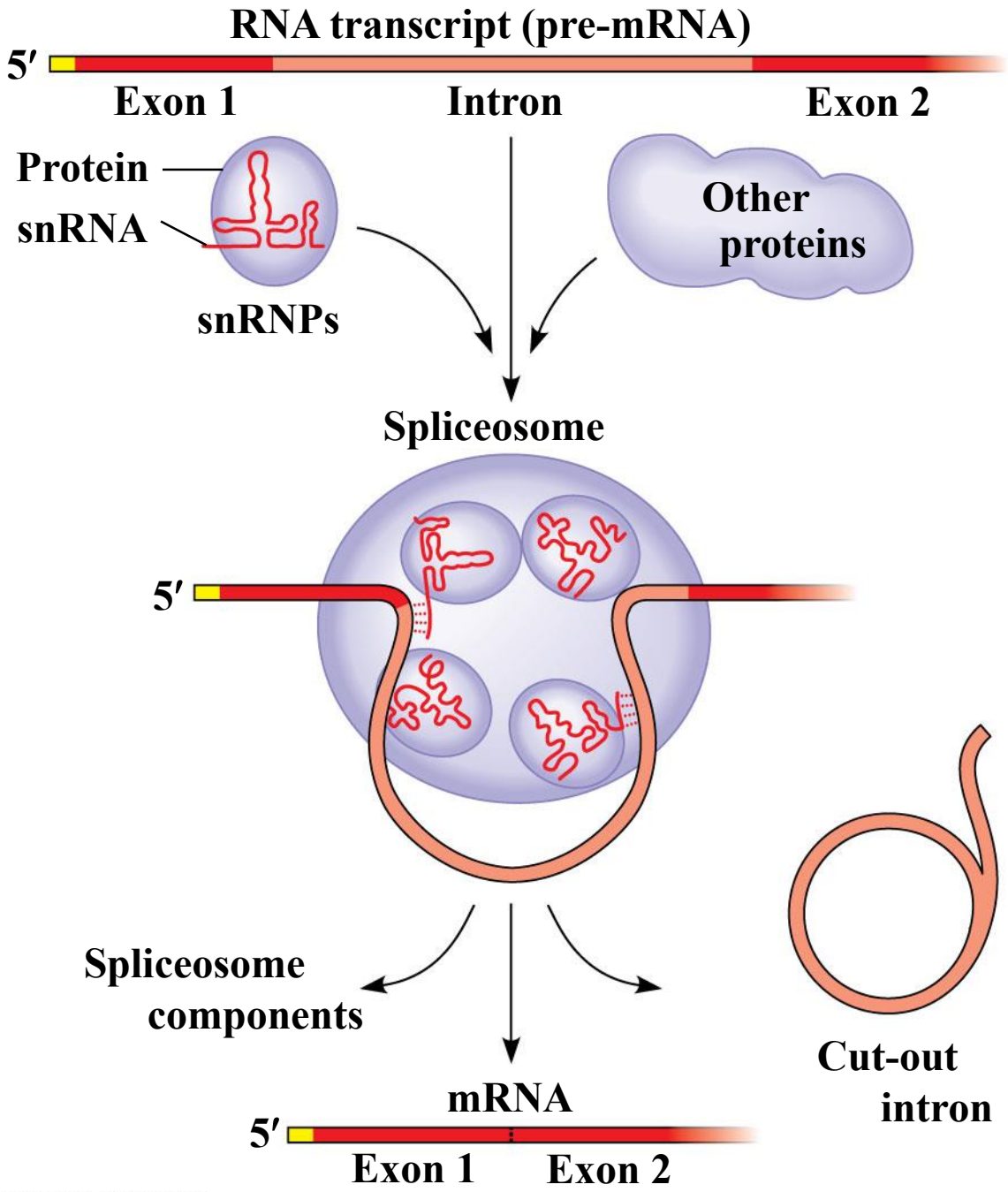


Figure 1.16 The process of RNA splicing

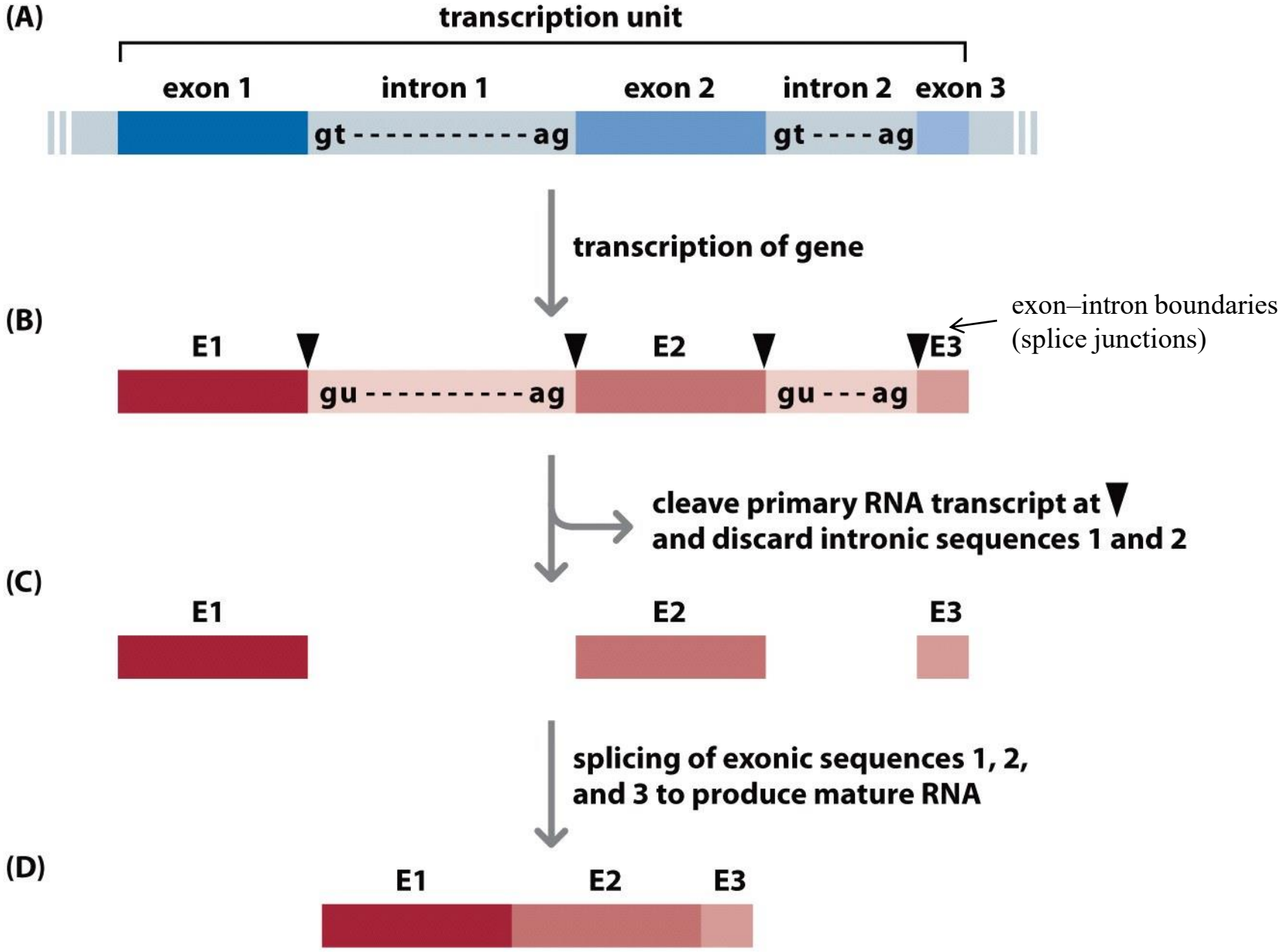


Figure 1.16 Human Molecular Genetics, 4ed. (© Garland Science)

Fig 1.17 3 splice junction **consensus DNA sequences** in introns of complex eukaryotes

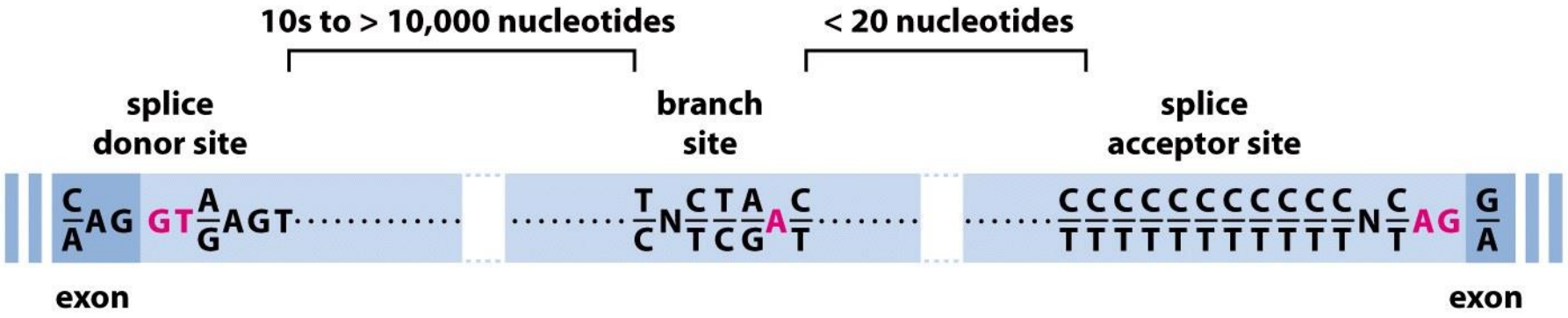


Figure 1.17 Human Molecular Genetics, 4ed. (© Garland Science)

Most introns in eukaryotic genes contain **conserved sequences** that correspond to three functionally important regions:

Two of the regions, the **splice donor** site and the **splice acceptor** site, span the 5′ and 3′ boundaries of the intron

The **branch site** is an additional important region that typically occurs less than **20 nts** upstream of the splice acceptor site

The nucleotides shown in red in these three consensus sequences are almost **invariant**.
The other nucleotides detailed in both the intron and the exons are those **most commonly** found at each position.

In some instances, two nucleotides may be **equally common**, as in the case of **C and T** near the 3′ end of the intron. Where N appears, any of the four nucleotides may occur.

Figure 1.18 The mechanism of RNA splicing

(A) The unprocessed primary RNA transcript with intronic RNA separating sequences E1 and E2 that correspond to exons in DNA

(B) The splicing mechanism involves a **nucleophilic attack** on the **G of the 5′ GU** dinucleotide. This is carried out by the **2′ OH** group on the conserved **A of the branch** site and results in the formation of a **lariat** structure and **cleavage of the splice donor** site

(C) The **3′ OH** at the 3′ end of the **E1** sequence performs a **nucleophilic attack** on the **splice acceptor** site, causing release of the intronic RNA (as a lariat-shaped structure) and fusion (splicing) of E1 and E2.

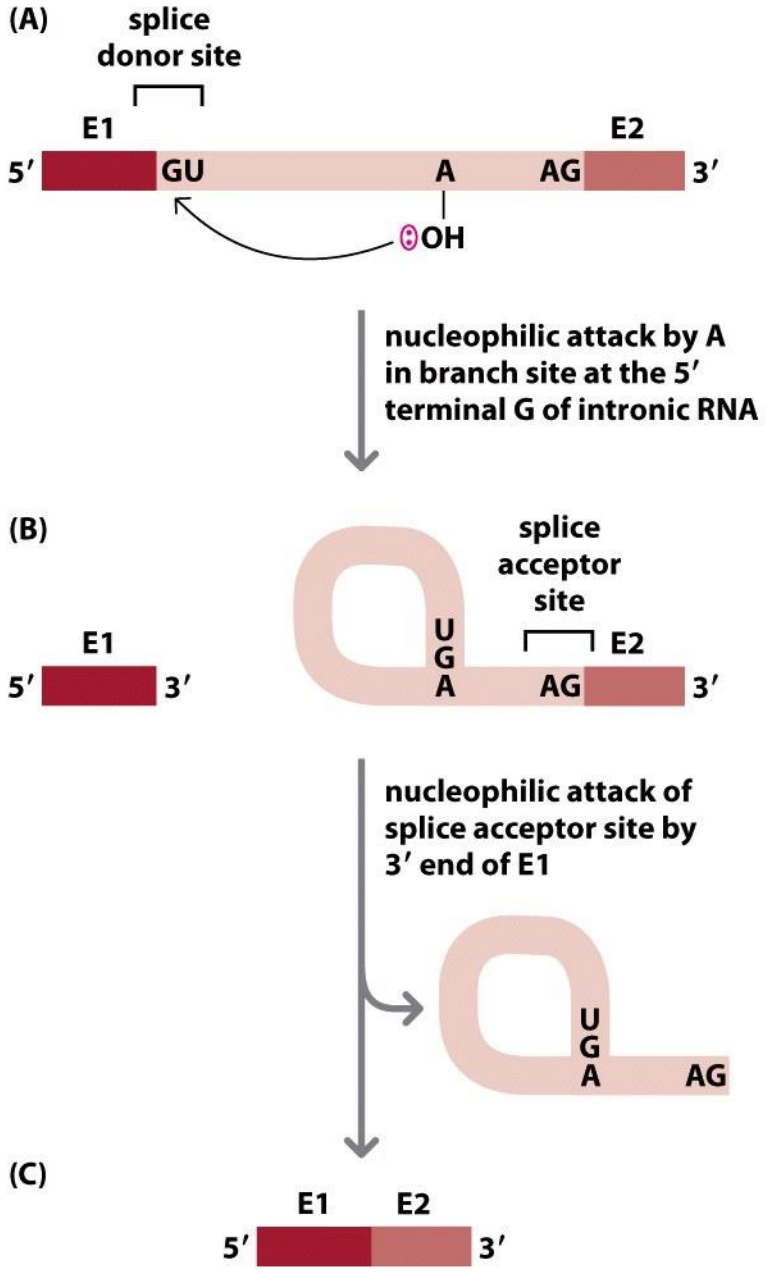


Figure 1.18 Human Molecular Genetics, 4ed. (© Garland Science)

Fig 1.19 Role of small nuclear ribonucleoprotein (snRNPs) in RNA splicing

A) The unprocessed primary RNA transcript

B) Within the spliceosome, part of the **U1** snRNA is **complementary** in sequence to the **splice donor site consensus sequence**. As a result, the U1 snRNA-protein complex (U1 snRNP) binds to the splice junction by **RNA–RNA base pairing**. The **U2** snRNP complex similarly binds to the **branch site** by RNA–RNA base pairing.

C) Interaction between the splice donor and splice acceptor sites is **stabilized** by the binding of a **multi-snRNP** particle that contains the **U4, U5, and U6** snRNAs.

- The **U5** snRNP binds simultaneously to both the splice donor and splice acceptor sites.

- Their cleavage releases the intronic sequence and allows (D) E1 and E2 to be spliced together.

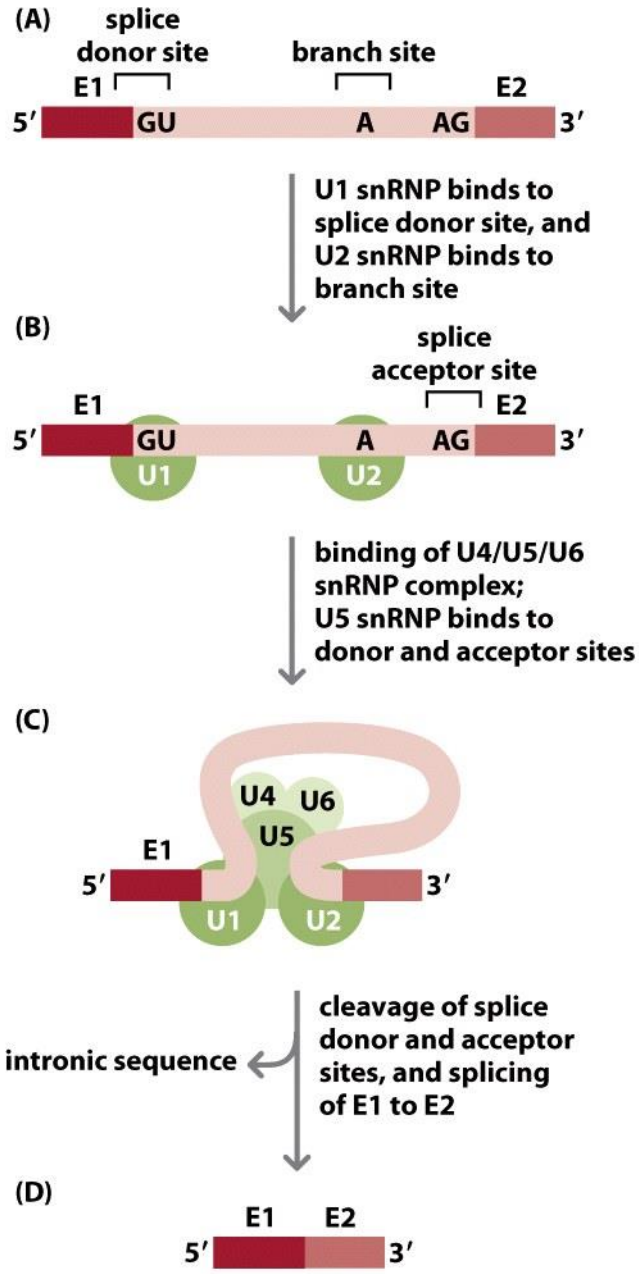


Figure 1.19 Human Molecular Genetics, 4ed. (© Garland Science)

Splice mutations

- Variants that likely impact splicing :

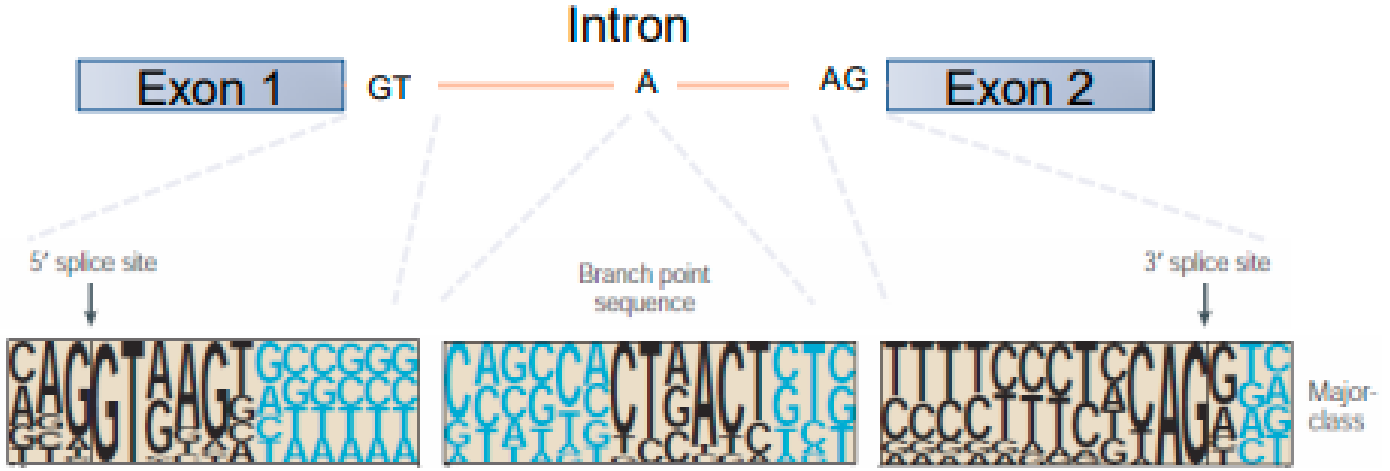
  Splice donor & acceptor positions (+/- 1,2) → destruction of 5'/3' splice consensus sequence, typically leads to exon skipping

- Variants that may impact splicing:

  Other positions in splice consensus sequence (+/- 15)

  Variants affecting 1st and last 3 bases of an exon

- Other point mutations also have potential to impact splicing



Regulatory mutations

- May be in promoter, enhancer, or UTRs
- Result in altered protein expression

Examples:

HBB c.-101C>T

- In promoter region of β-globin gene
- Leads to decreased expression
- Compound heterozygotes with a severe mutation → Mild β-thalassemia