Mecanismos de evolución de los genomas

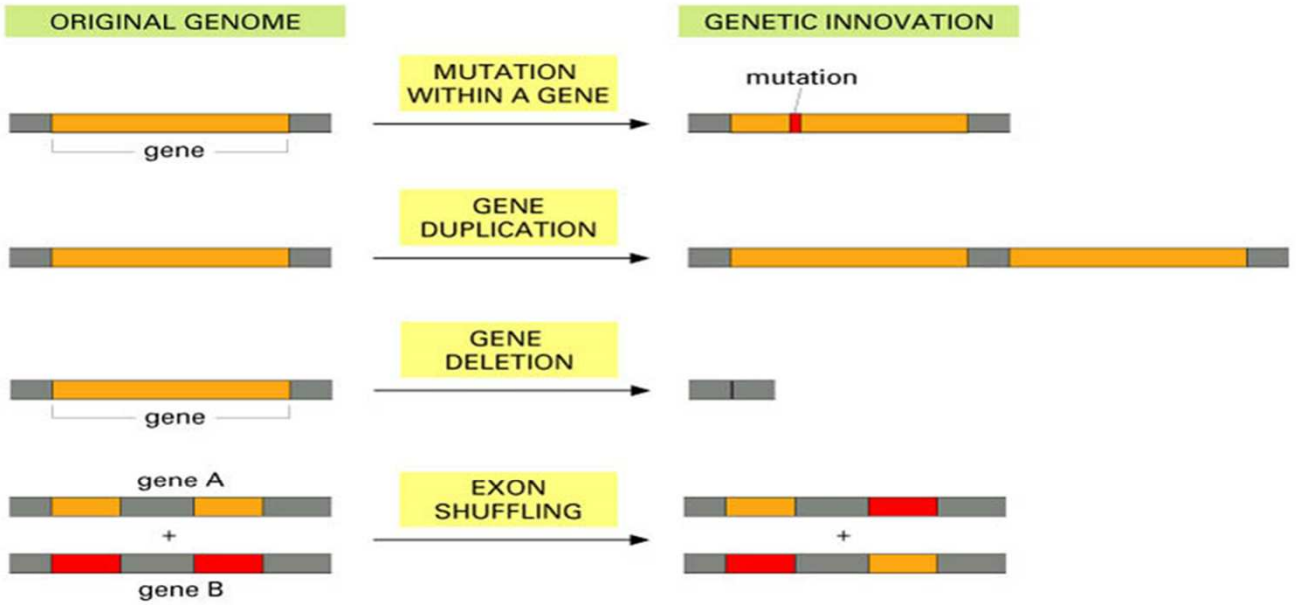
- Generación de la variación genética: Clases de mutaciones

- Duplicación de dominios y alargamientos de genes.
- Exon-shuffling
- Genes quimera

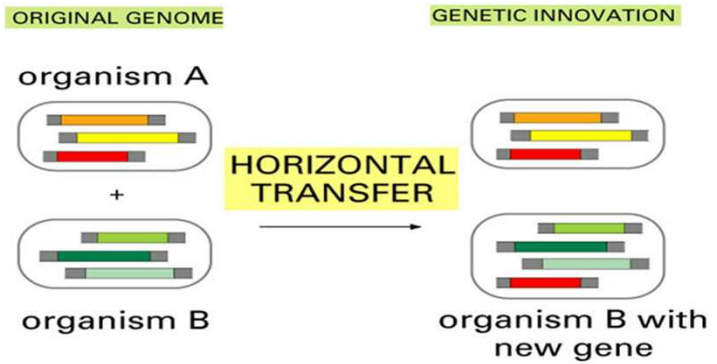
Generating Genetic Variation

Cinco tipos de cambio contribuyen a la evolución de un gen individual.

- La mutación dentro de un gen
- Duplicación de un gen
- Supresión de genes
- exón shuffling



- La transferencia horizontal - raro en eucariotas



Mutations and Molecular Evolution

- 3 different effects on fitness by mutations:
  - deleterious
  - increase efficiency or performance
  - no effect (“neutral”)

Different models of selection

- Directional selection- pushes population toward homozygosity and phenotypes toward one extreme
- Balancing selection- favors heterozygotes but maintains all phenotypes
- Disruptive selection- favors both homozygotes, eliminates heterozygotes and increases extremes of phenotypes

Red blood cells in someone with sickle-cell trait

La mutación dentro de un gen

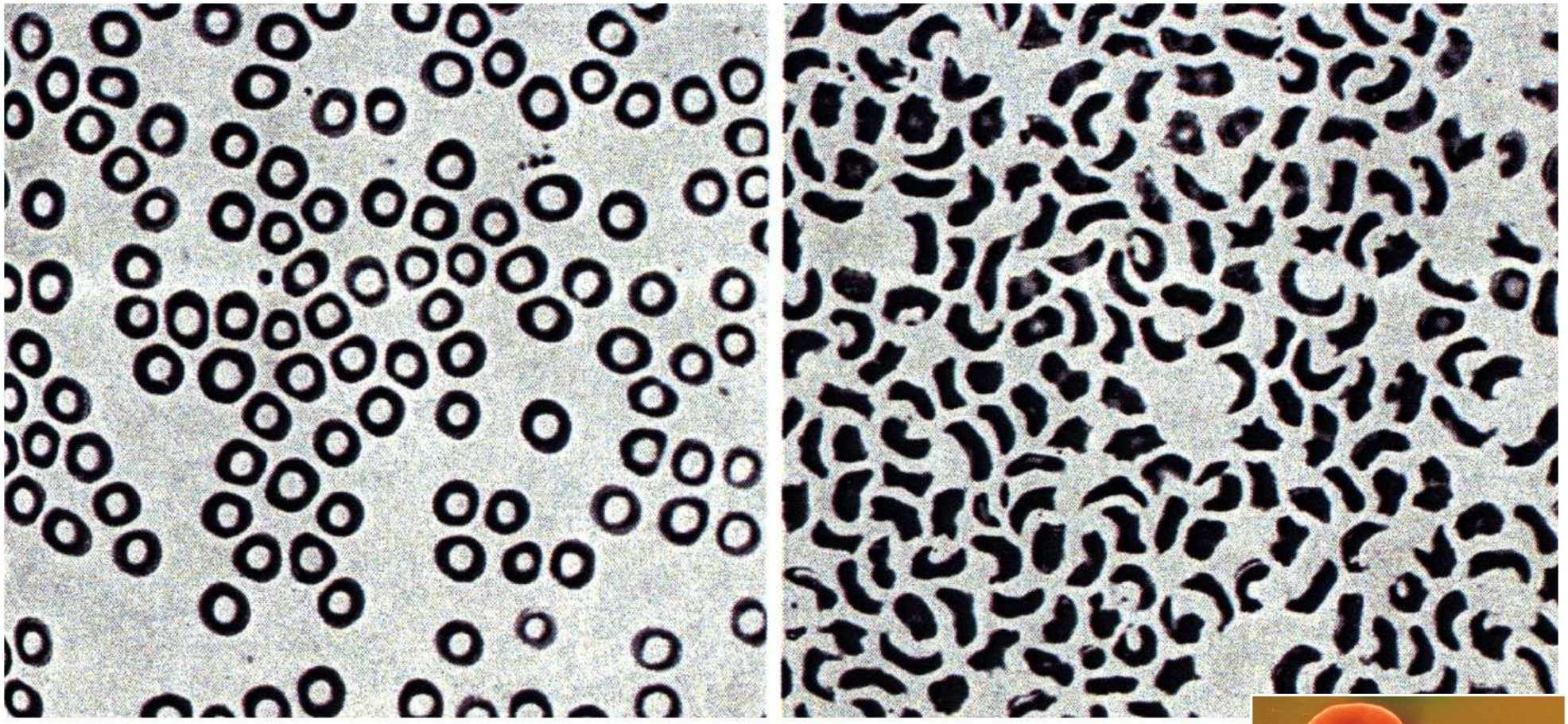


Figure 20-5
Introduction to Genetic Analysis, Tenth Edition

Malarial parasites live within red blood cells

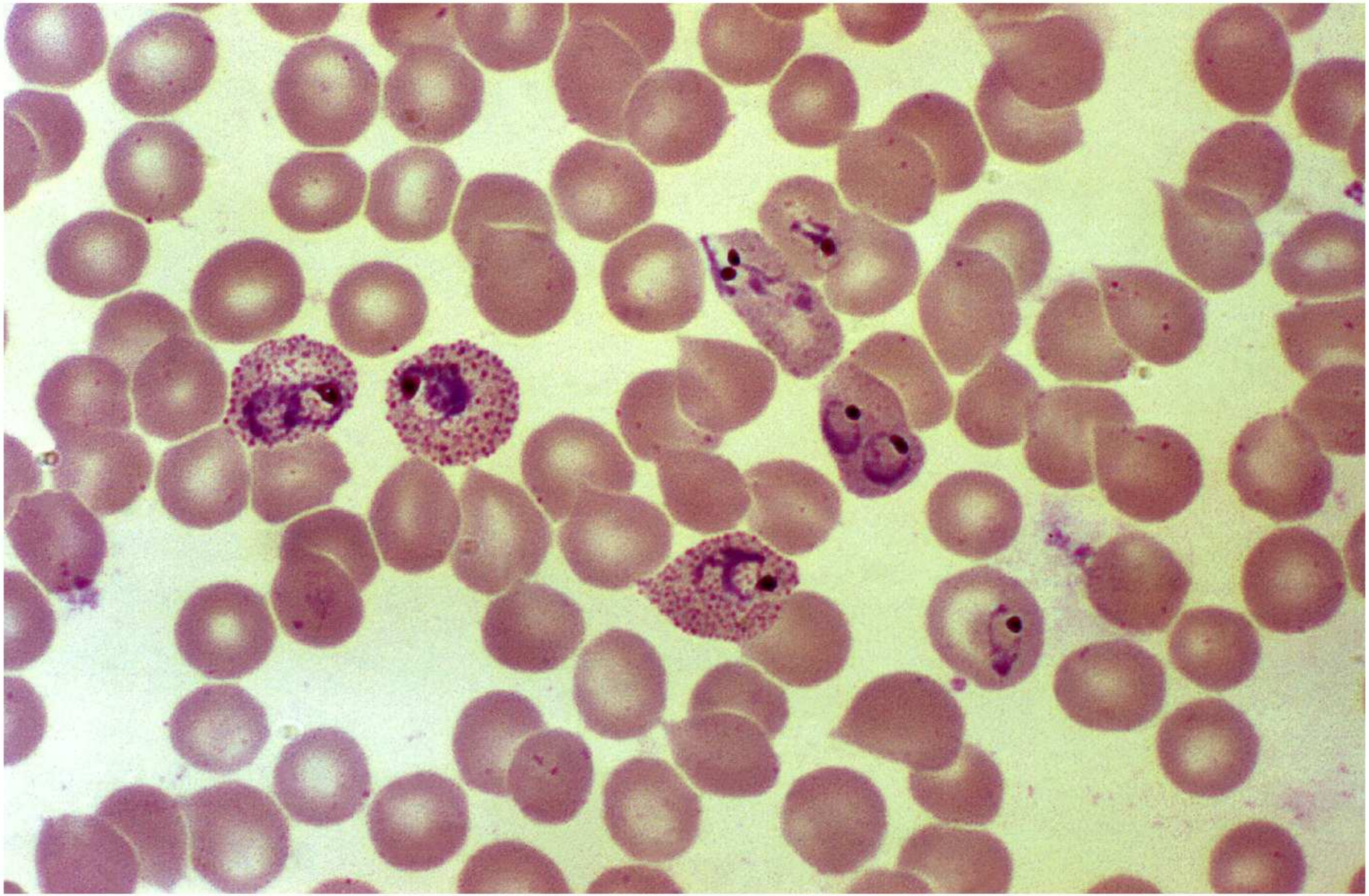
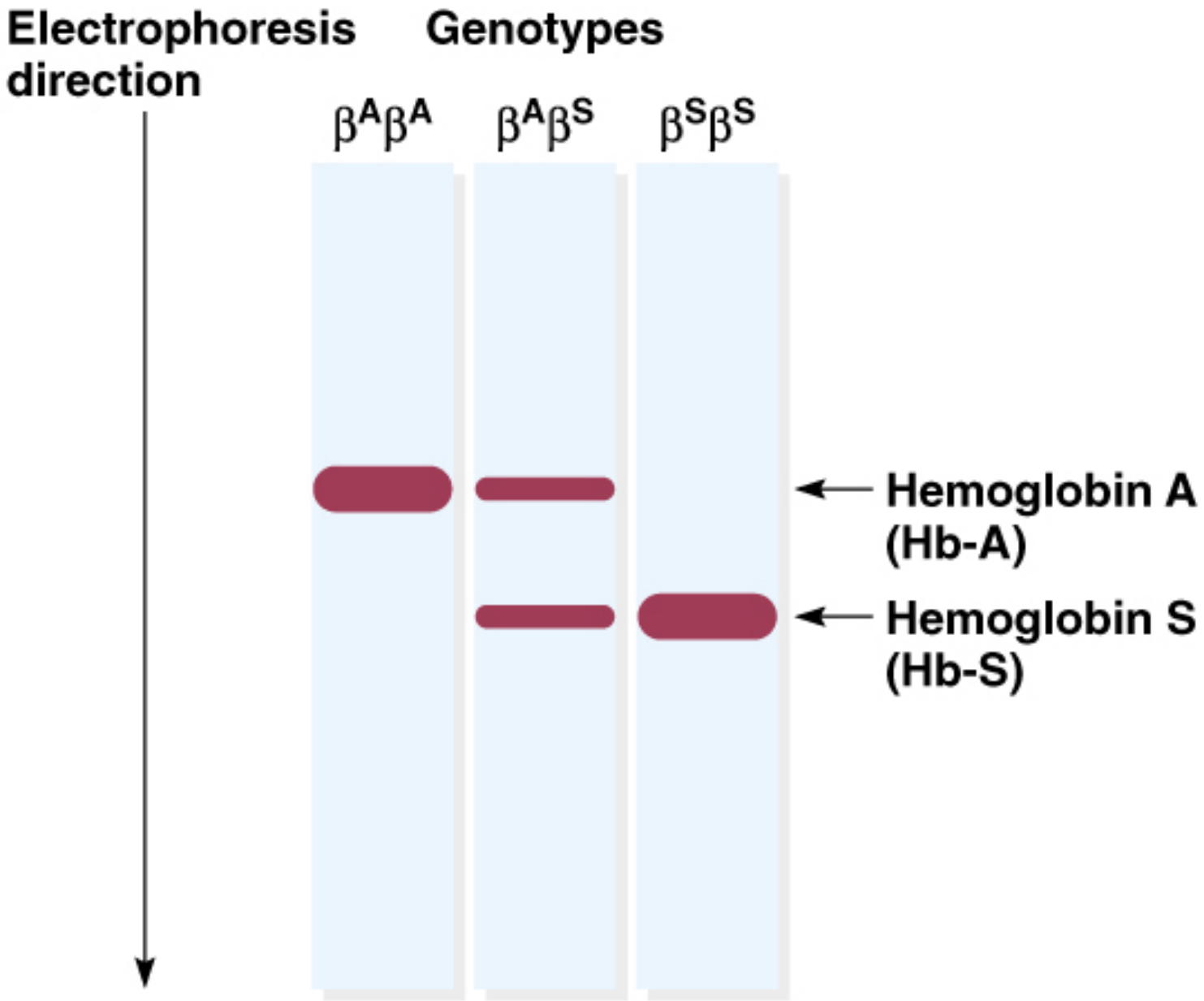
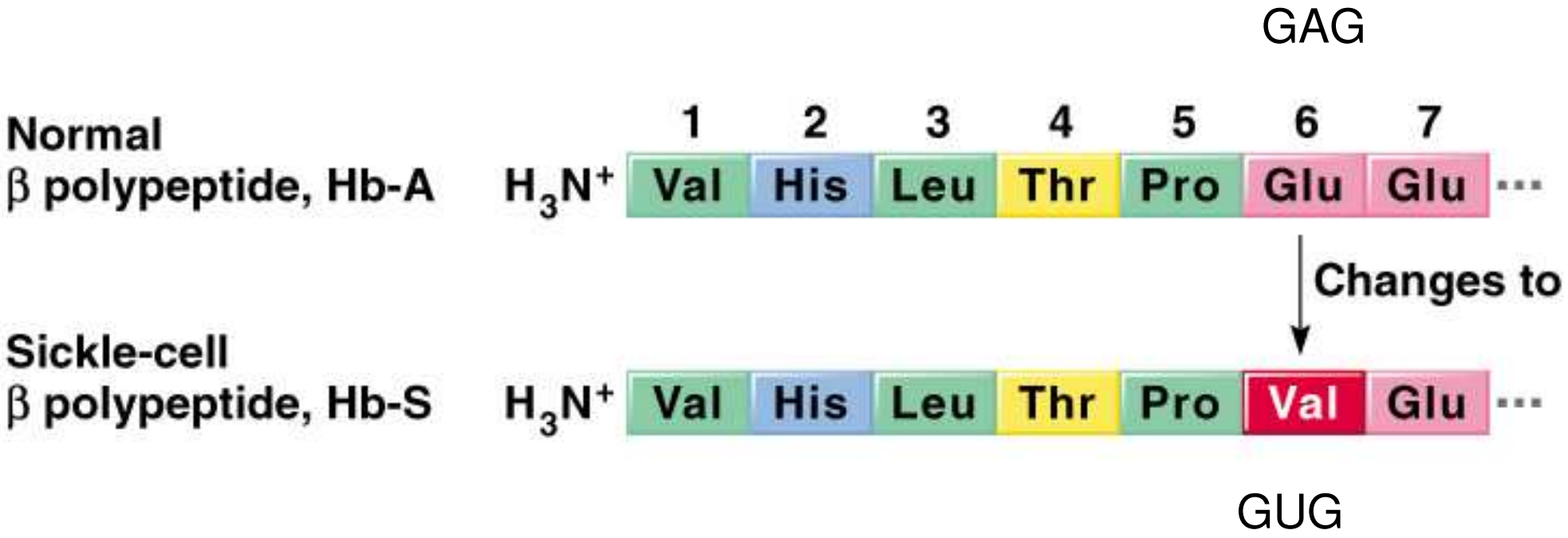


Figure 20-6
Introduction to Genetic Analysis, Tenth Edition

Electrophoresis of hemoglobin variants



The first seven N-terminal amino acids in normal and sickle cell hemoglobin β polypeptides



GLU = Glutamic acid is acidic
VAL = Valine is neutral non-polar

• Evolución del tamaño de los genes

- **mecanismos de amplificación y de aparición de nuevos genes**
  - **Duplicación Parcial de Genes**
    - **Barajamiento de exones/dominios (exon-shuffling)**
    - **Genes quimera**

Relaciones entre los distintos exones de un gen y los dominios proteícos.

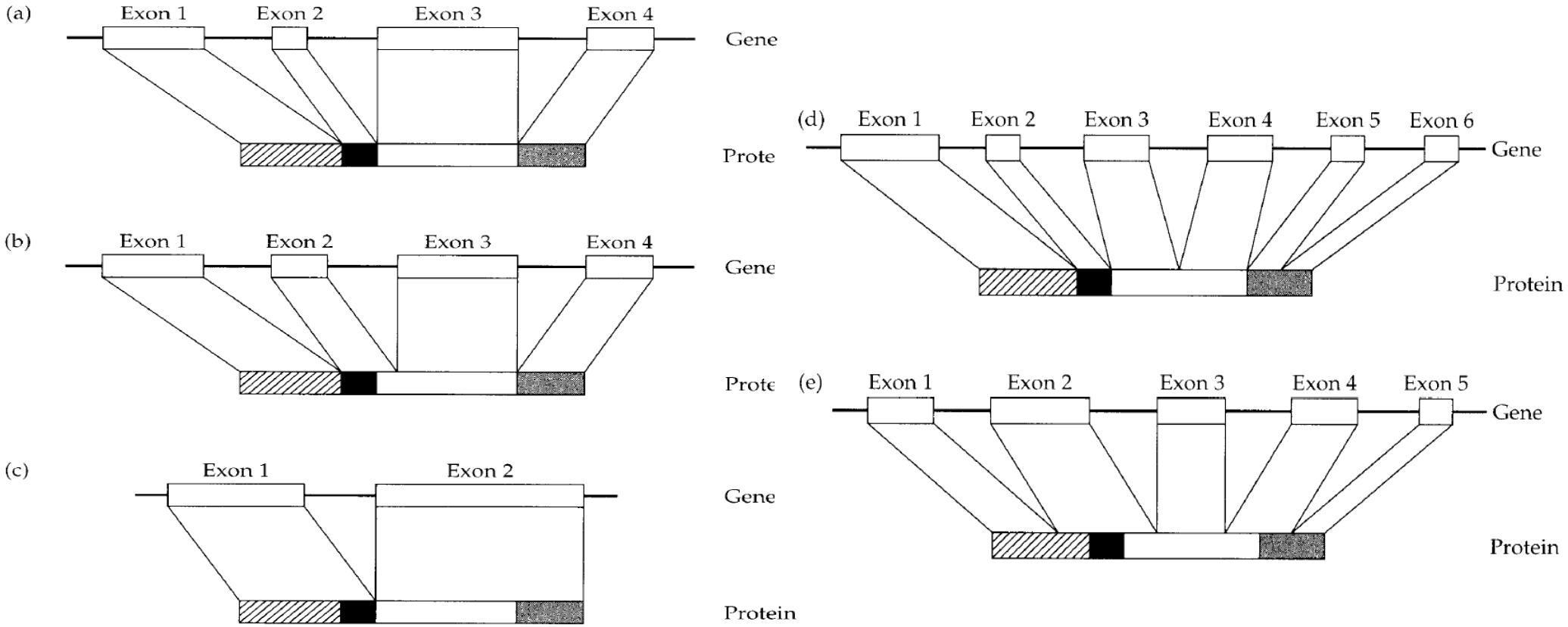
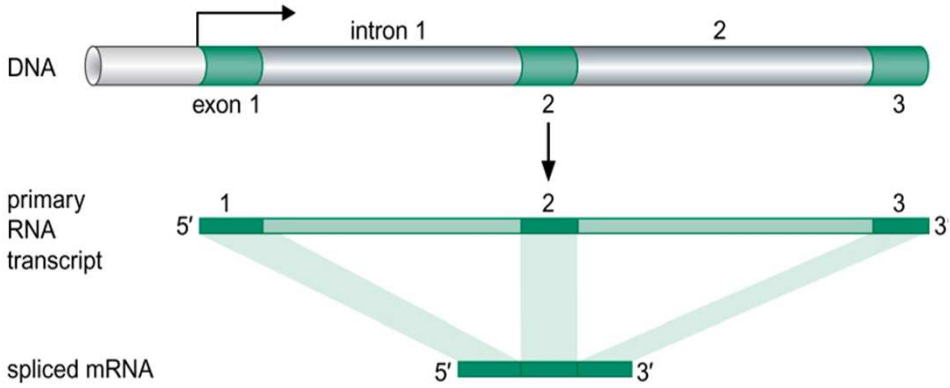


FIGURE 6.2 Five possible relationships between the arrangement of exons in a gene and the structural domains of its protein. (a) Each exon corresponds exactly to a structural domain. (b) The correspondence is only approximate. (c) An exon encodes two or more domains. (d) A single structural domain is encoded by two or more exons. (e) Lack of correspondence between exons and domains. The structural domains of the protein are designated by different boxes (hatched, black, white, and gray).

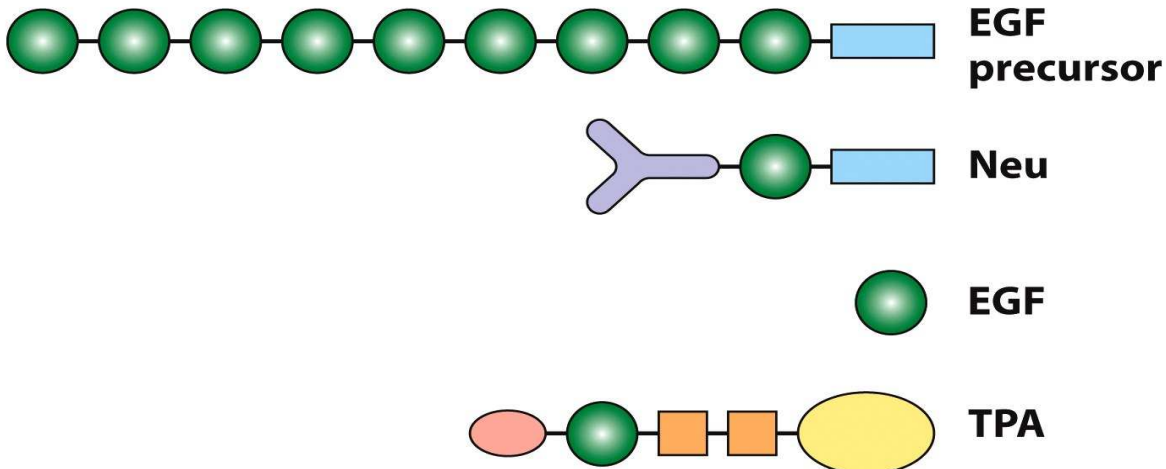
Genomas eucarióticos

Genes interrumpidos (con frecuencia)

Exones-intrones

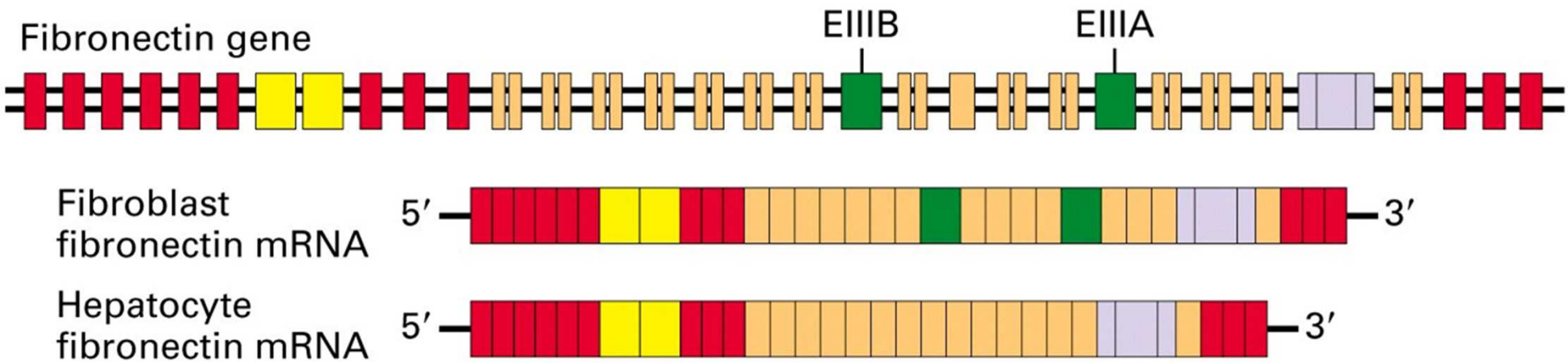


Los dominios son independientemente plegables y funcionalmente especializados unidades de estructura terciaria dentro de una proteína. La estructura de dominio modular de muchas proteínas ha resultado de la barajadura y el empalme de sus secuencias de codificación dentro de genes más largos



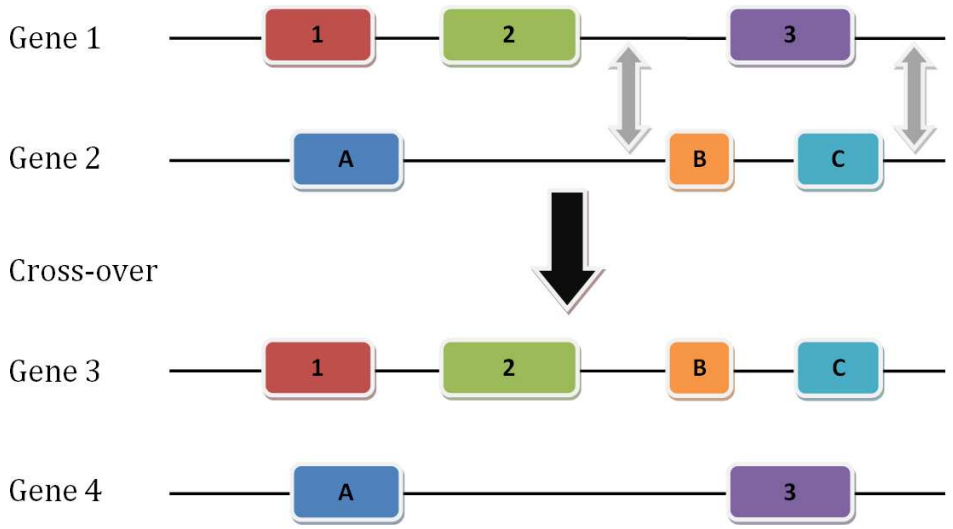
Alternative Splicing & Gene Regulation

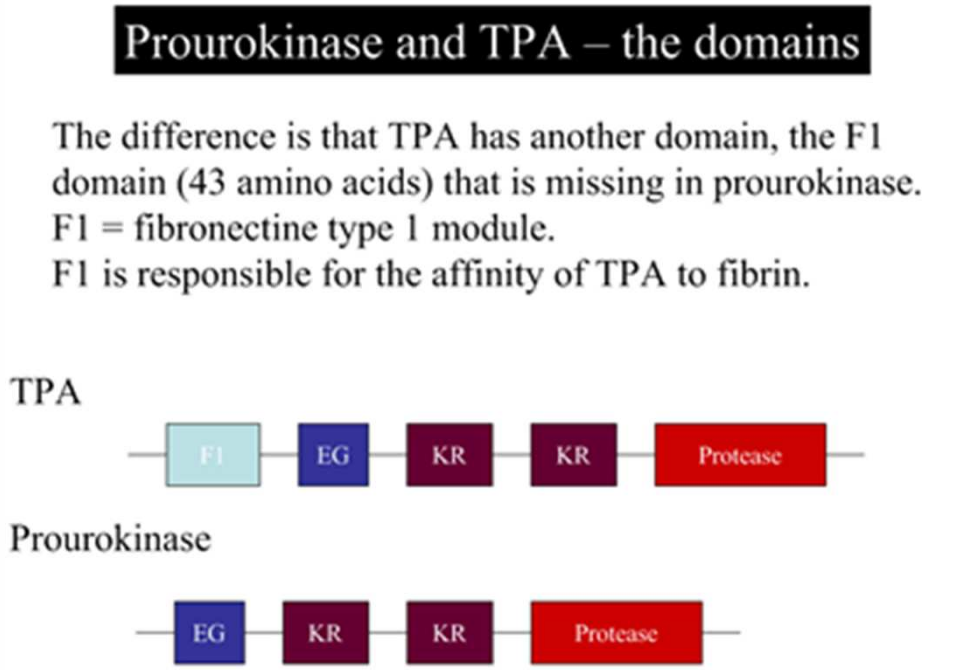
- Protein domains can be encoded by a single exon or by a small collection of exons within a larger gene. The coding regions for domains can be spliced in or out of the primary transcript by the process of alternative splicing. The resulting mRNAs encode different forms of the protein, known as isoforms. Alternative splicing is an important method for regulation of gene expression in different tissues and different physiological states. It is estimated that 60% of all human genes are expressed as alternatively spliced mRNAs. Alternative splicing is illustrated in Fig for the fibronectin gene. The fibroblast and hepatocyte isoforms differ in their content of the EIIIA and EIIIB domains which mediate cell surface binding. Twenty different isoforms of fibronectin produced by alternative splicing have been identified.



Exon-suffling. Es la idea de que la recombinación, la inserción o la deleción de un exón o intrón puede producir nuevos genes •

- **Cuatro tipos de exon-suffling**
  -Deleción
- - Recombinación
  - Duplicación
- - Inserción





Exon-shuffling: pérdida de un intrón

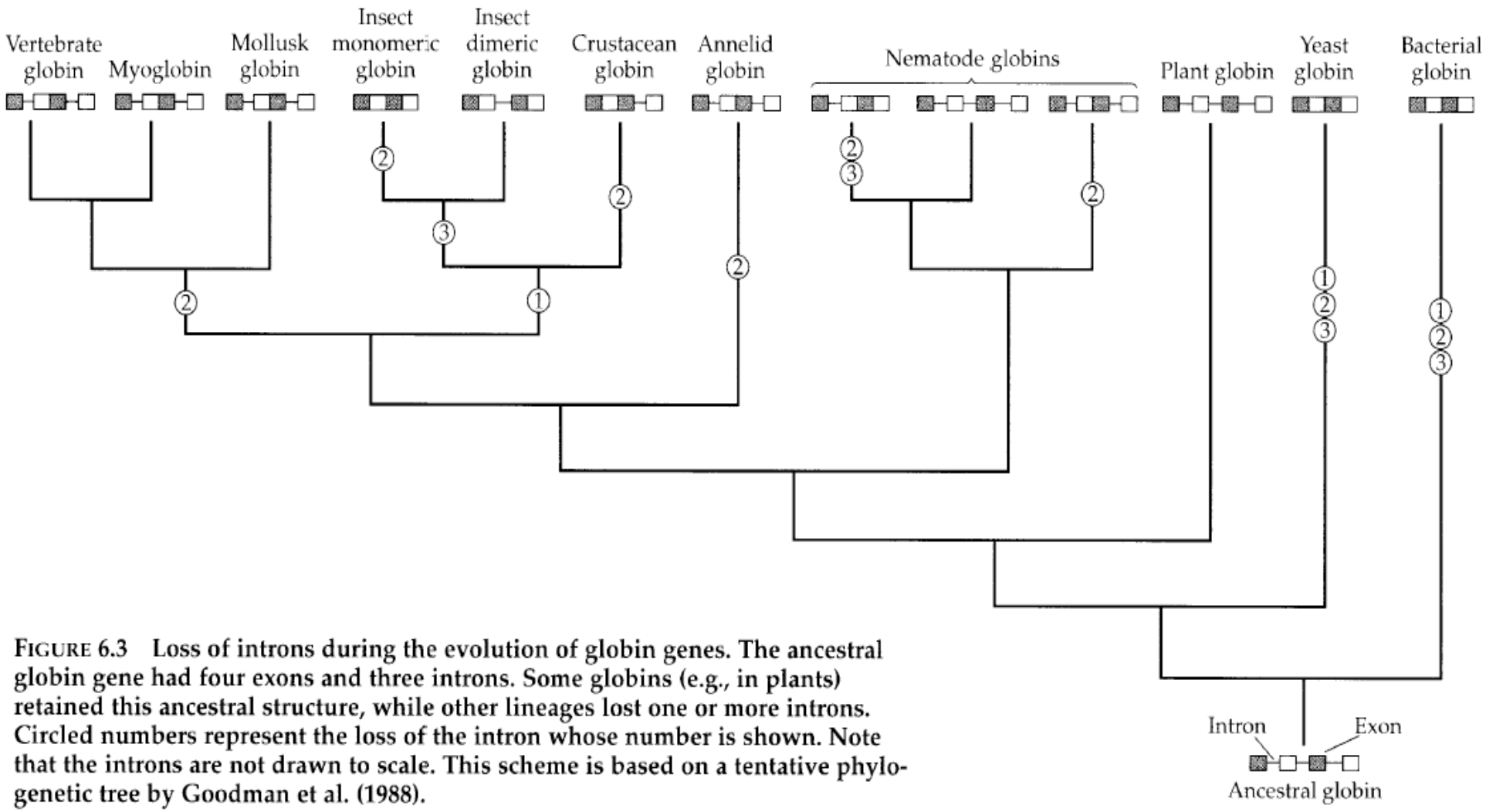
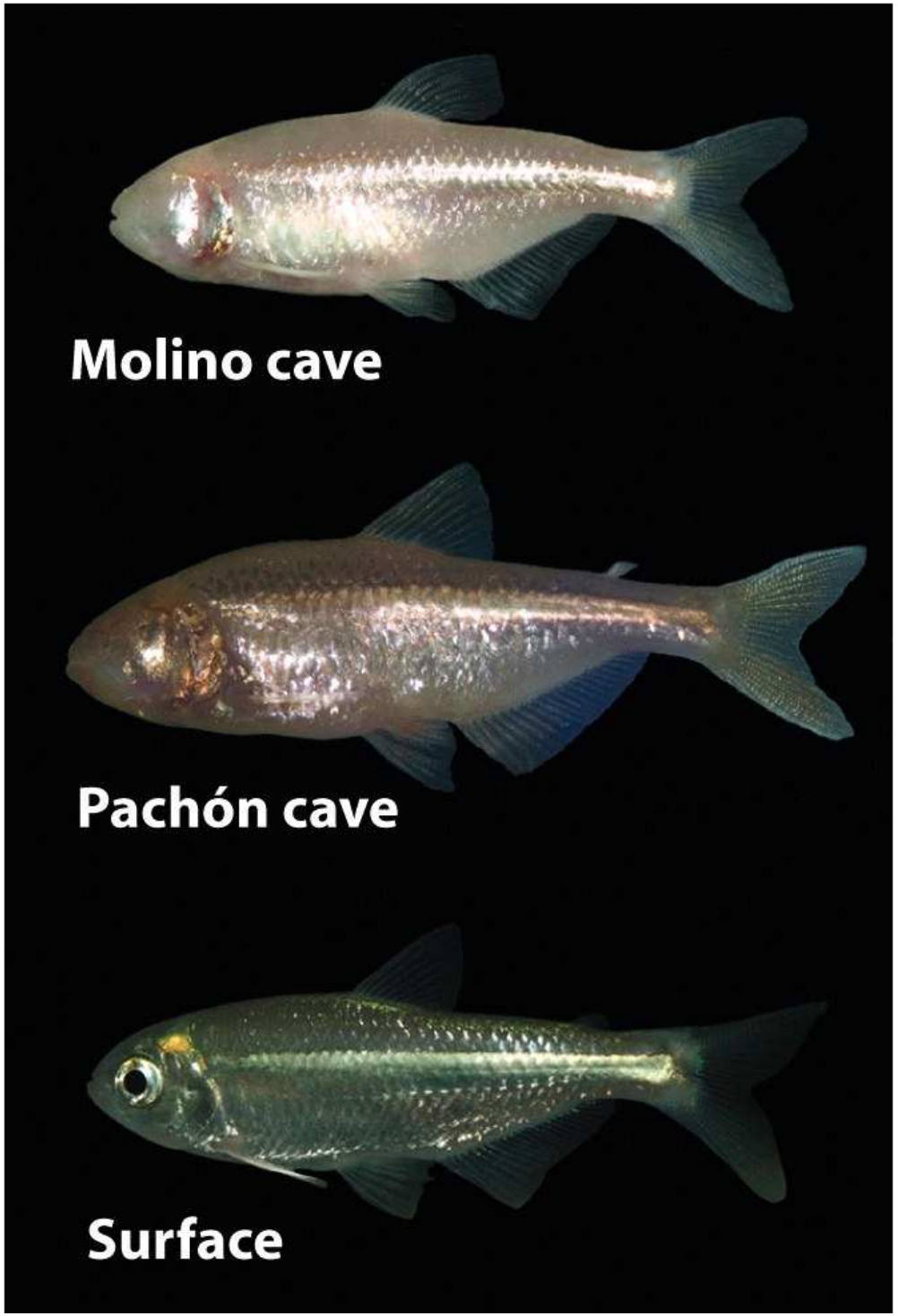


FIGURE 6.3 Loss of introns during the evolution of globin genes. The ancestral globin gene had four exons and three introns. Some globins (e.g., in plants) retained this ancestral structure, while other lineages lost one or more introns. Circled numbers represent the loss of the intron whose number is shown. Note that the introns are not drawn to scale. This scheme is based on a tentative phylogenetic tree by Goodman et al. (1988).

- **Evolución globinas**: pérdida de intrones
- en microbios (todos)
- y 1, 2 o 3 en animales

Morphological evolution-evolution of albinism in blind cave fishes (gene inactivation)



Morphological evolution—albinism in blind cave fishes

- albinism common in cave organisms (incl. fishes, crustaceans), often accompanied by eye loss
- genetic studies of Mexican blind cave fish (*Astyonax mexicanus*) in 2 different populations, Pachón and Molino, revealed different mutations in *Oca2* gene—gene inactivation
- Pachón fishes are homozygous for deletion of intron and most of exon in *Oca2* gene

Recombinación: tissue plasminogen activator evolves from four unrelated genes

protease

kringle (plasminogen)

epidermal growth factor

fibronectin type-1

(from Graur & Li 2000)

A survey of modern genes in eukaryotes shows that internal duplications have occurred frequently in evolution. This increase in gene size, or **gene elongation**, is one of the most important steps in the evolution of complex genes from simple ones. Theoretically, elongation of genes can also occur by other means. For example, a mutational change converting a stop codon into a sense codon can also elongate the gene (Chapter 1). Similarly, either insertion of a foreign DNA segment into an exon or the occurrence of a mutation obliterating a splicing site will achieve the same result. These types of molecular changes, however, would most probably disrupt the function of the elongated gene, because the added regions would consist of an almost random array of amino acids. Indeed, in the vast majority of cases, such molecular changes have been found to be associated with pathological manifestations.

Duplicación de exones

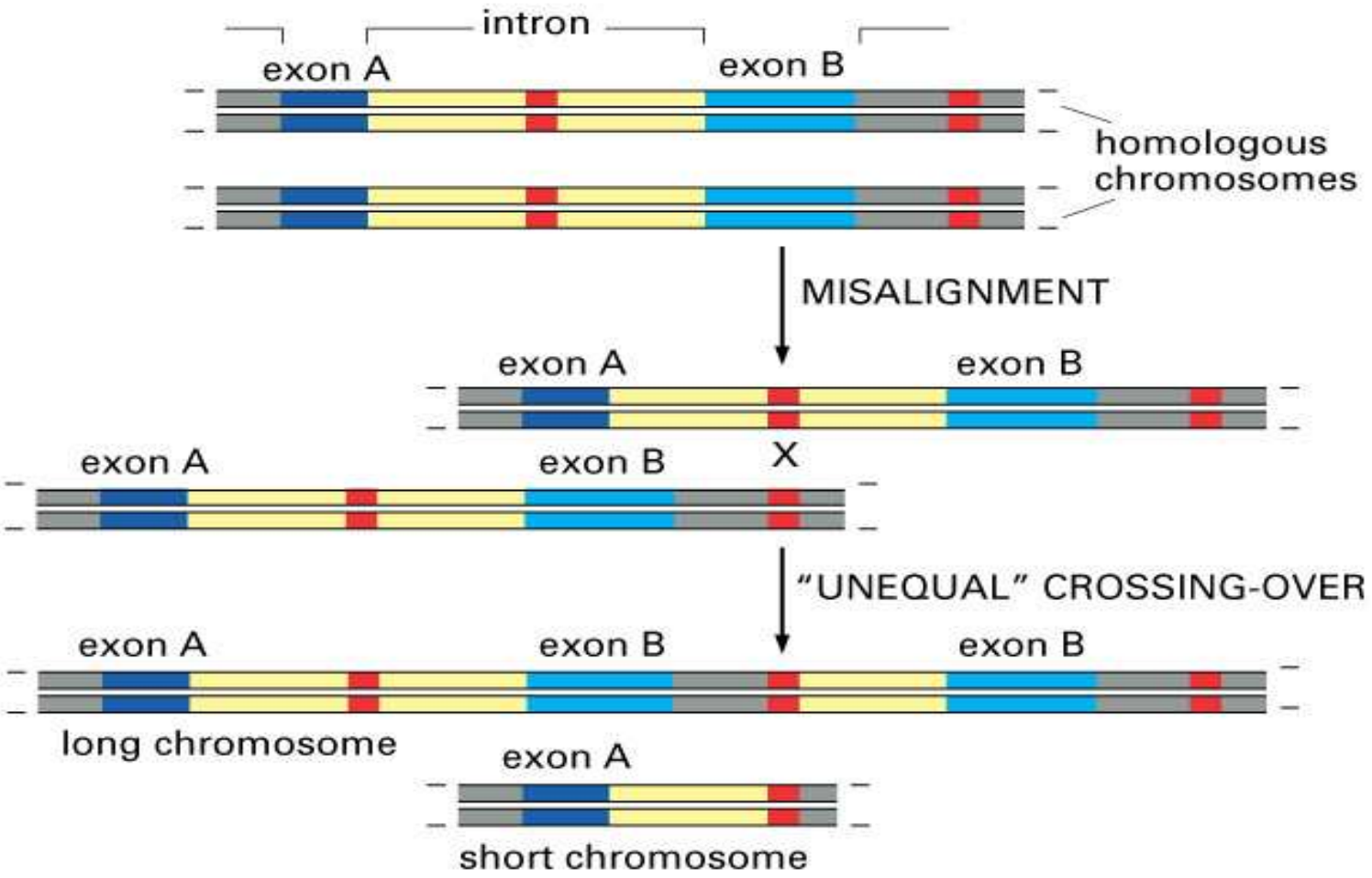


Figure 9-9 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Ejemplo de evolución por dulicación de exones de la proteína del colágeno de aves

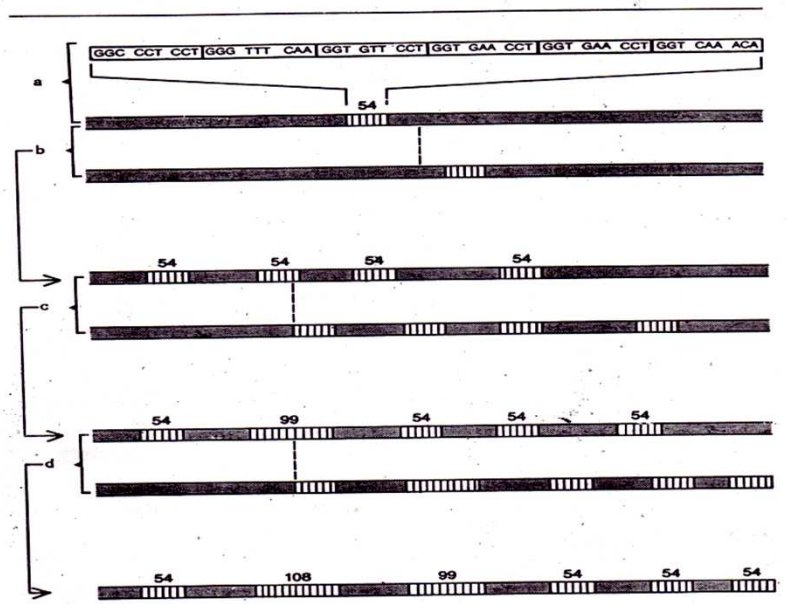
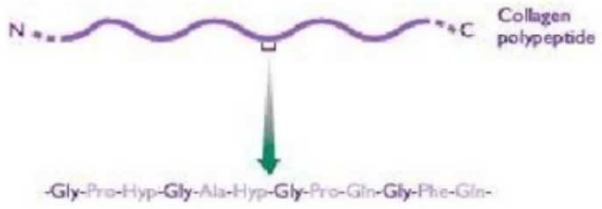


Fig. 5-2. Diagrama esquemático de la evolución del gen que codifica el colágeno en las gallinas. En primer lugar se formó un exón con una longitud de 54 nucleótidos a partir de la repetición en tándem por seis veces de una secuencia básica de 9 nucleótidos (*a*). El exón se fue duplicando por entrecruzamientos desiguales (*b*). En algunos casos (*c* y *d*), el entrecruzamiento produjo nuevos exones con un número de nucleótidos distinto de 54 (99 en el caso *c* y 108 en el *d*)[22].

Domain Duplication

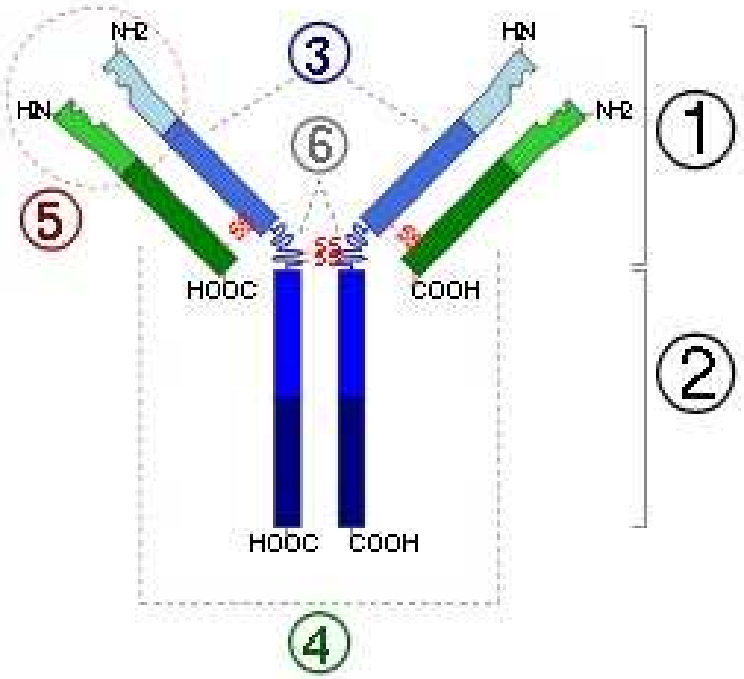
Clip slide

- The α2 Type I collagen has repetitive Gly-X-Y
- It codes for 338 of these repeats, is split into 52 exons, 42 of which cover the part of the gene coding for the glycine-X-Y repeats.
- The number of repeats per exon varies but is 5 (5 exons), 6 (23 exons), 11 (5 exons), 12 (8 exons) or 18 (1 exon).
- Gene have evolved by duplication of exons leading to repetition of the structural domains.



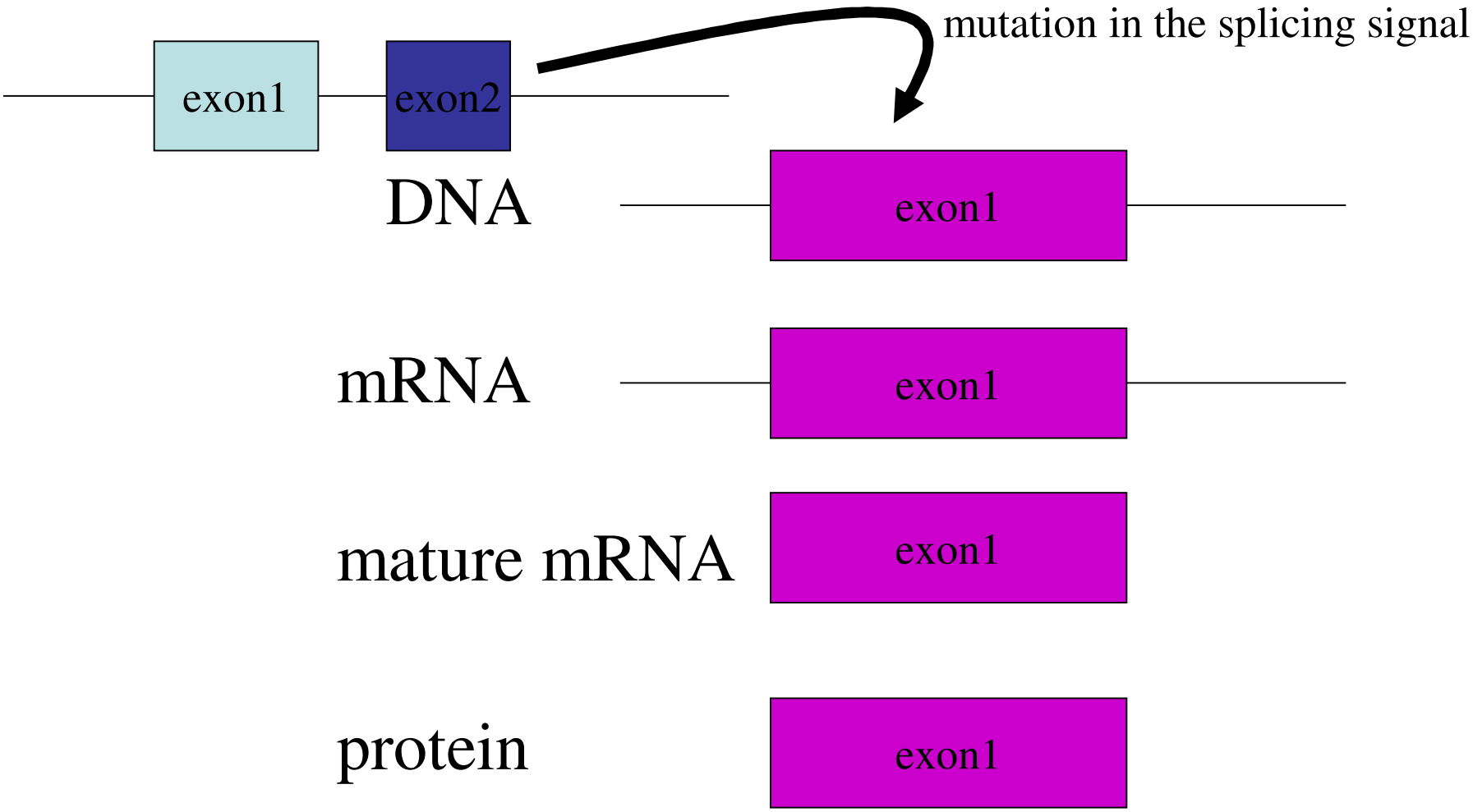
Evolución de exones duplicados: divergencia entre la partes de la duplicación

The second possibility for the emergence of a novel function following partial gene duplication is for the internal copies thus produced to diverge in sequence, ultimately resulting in each of them performing a different function. For instance, the variable and constant regions of immunoglobulin genes were probably derived from a common primordial domain, but have since acquired distinct properties (Leder 1982). Thus, despite common molecular ancestry, the variable region of immunoglobulins binds antigens, while the constant region mediates non-antigenic functions. Many complex genes might have arisen in this manner.



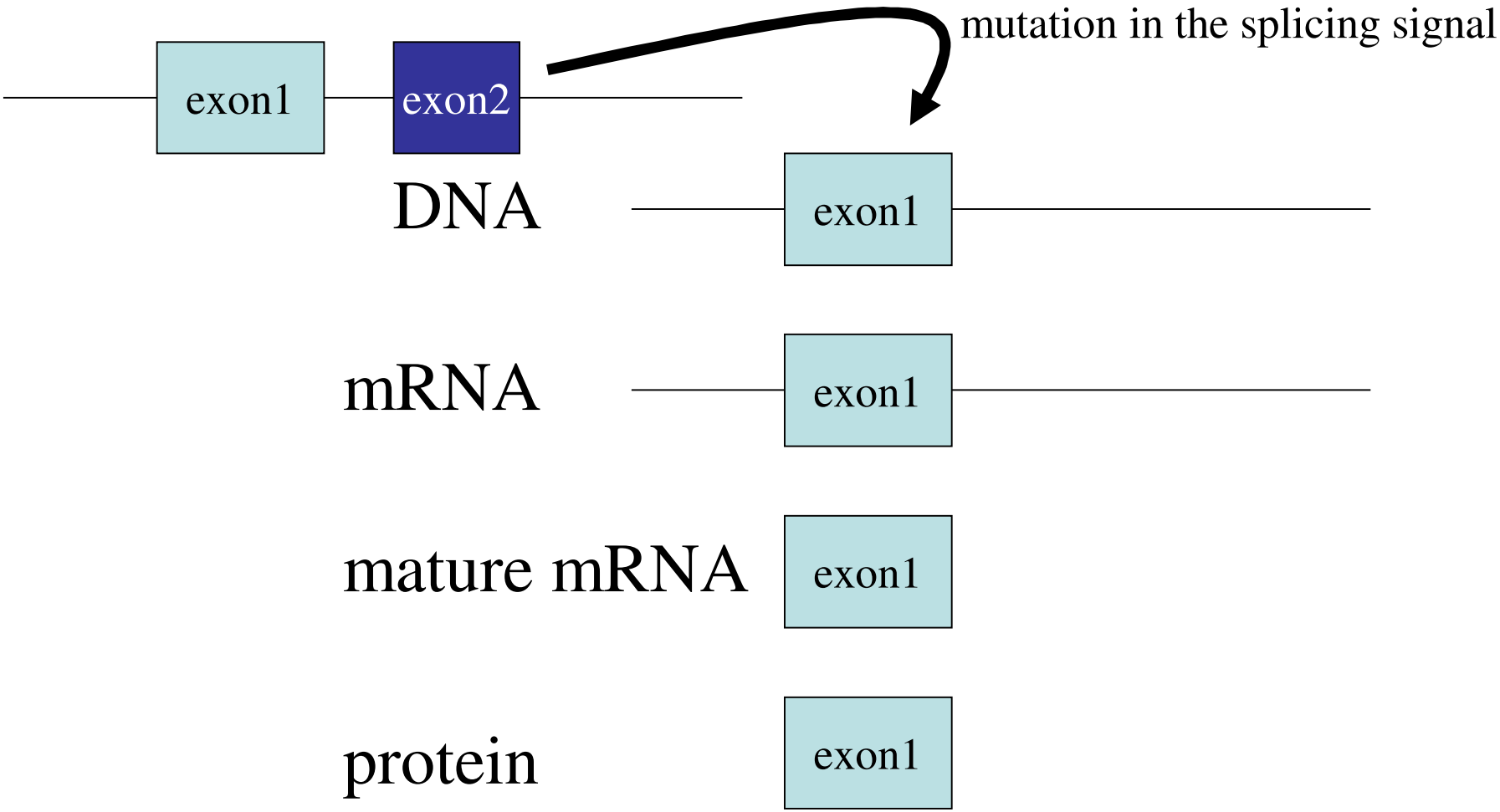
Exonization

Mutations in the DNA that encode signals for intron excision might result in exonization of the intron.



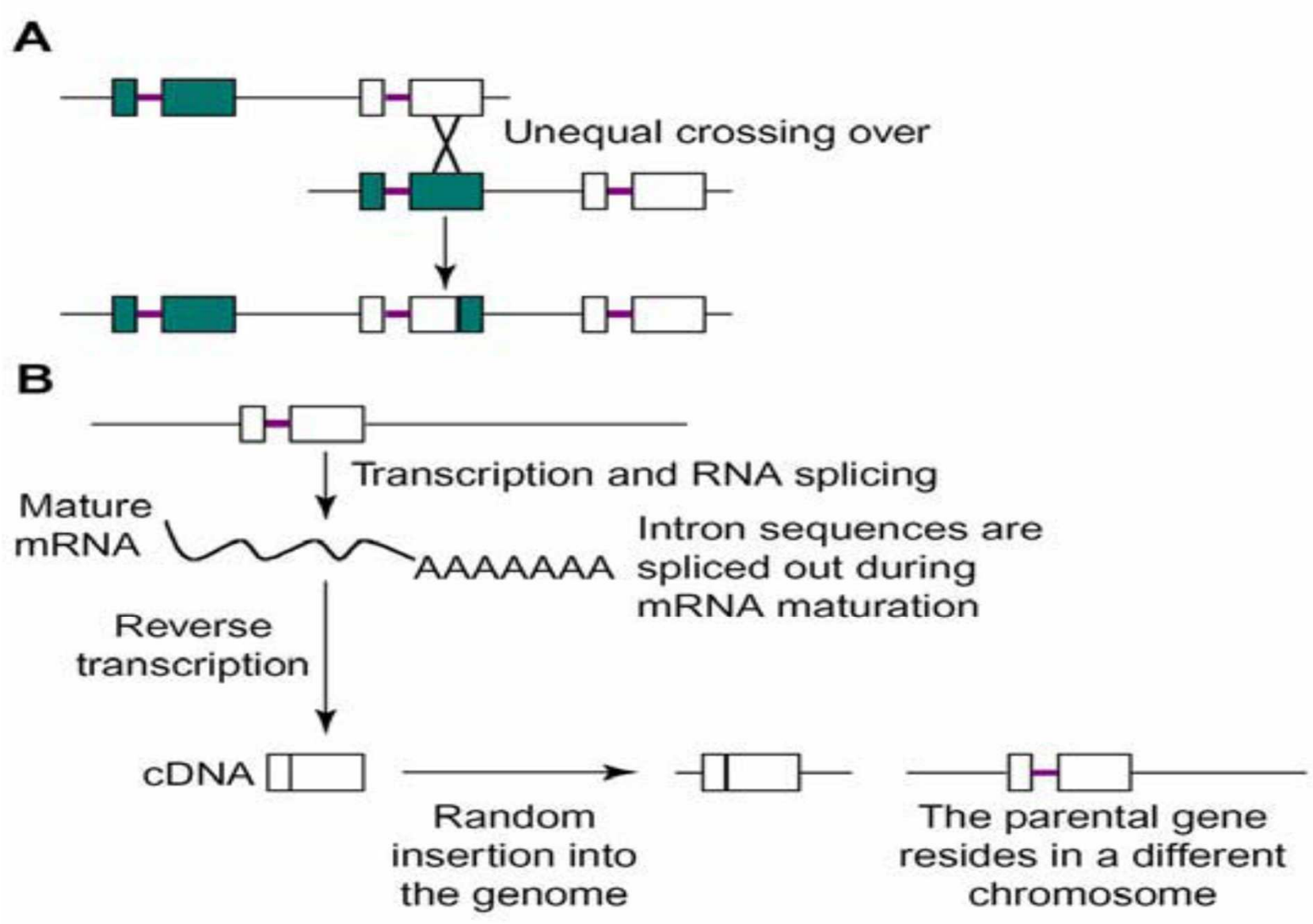
Exon lost

Of course, in a similar vain, exons can also be removed due to such mutations.



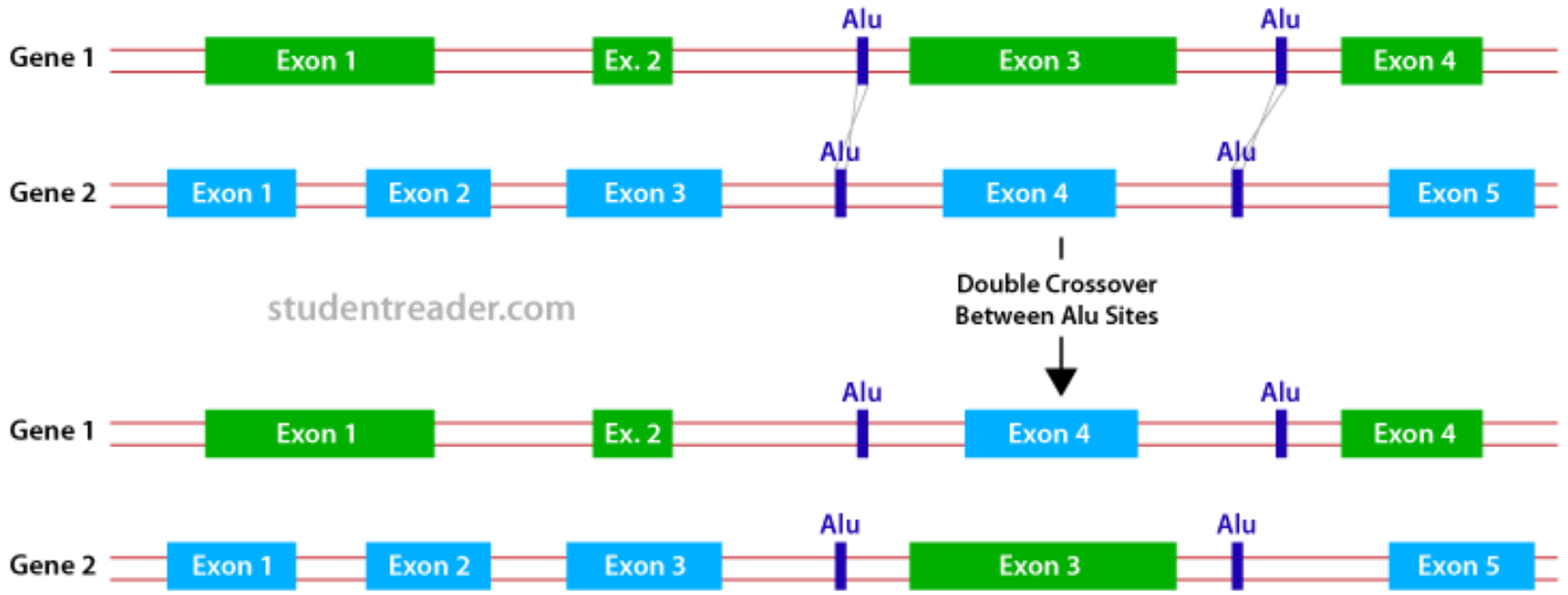
Mecanismos más importantes de formación de duplicaciones exónicas.

(A) Entrecruzamiento desigual. (B) Retrotransposición.



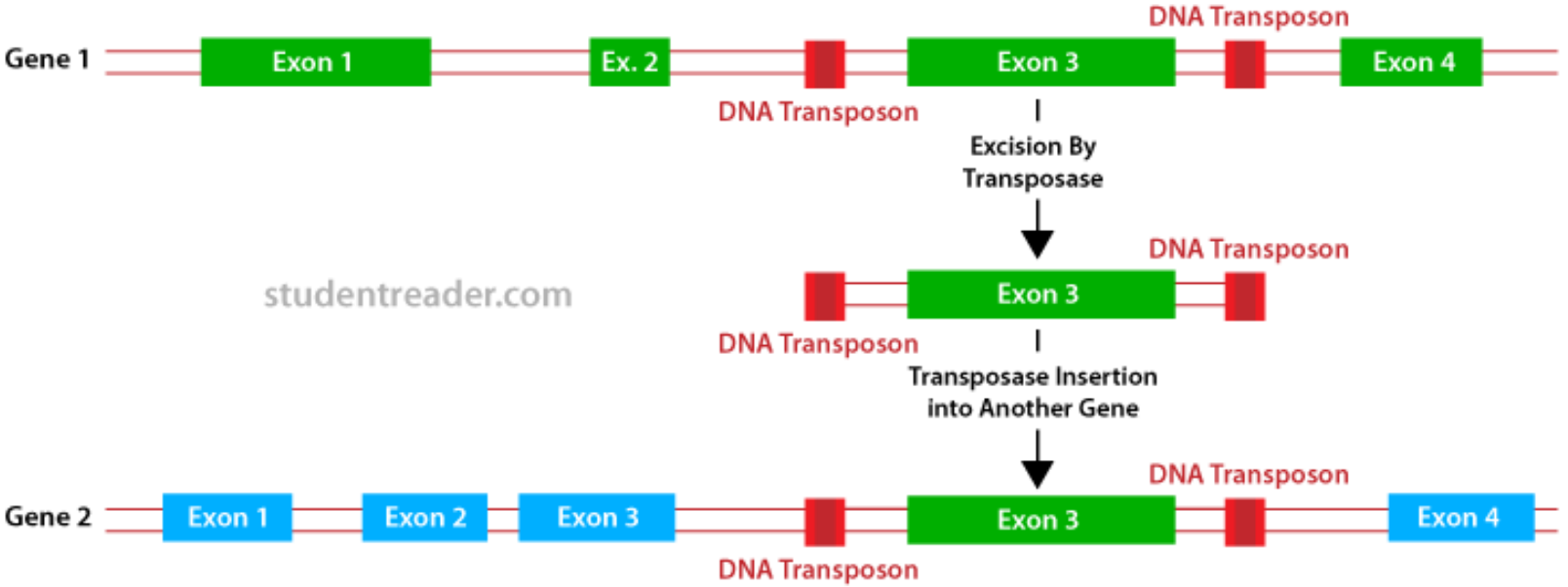
B1 **Via Entrecruzamiento entre elementos intercalados**

We previously have noted that gene evolution has involved exon shuffling between protein-coding genes in the genome. A large amount of shuffling has occurred due to the prevalence of interspersed repeats in the genome. Due to sequence conservation within these regions, crossover events can take place at these sites (Fig. 6.18). This results in exon shuffling between nonhomologous genes and the formation of new genes with new combinations of protein domains. As illustrated in Fig. 6.2, such events also have been important in exon and gene duplications.

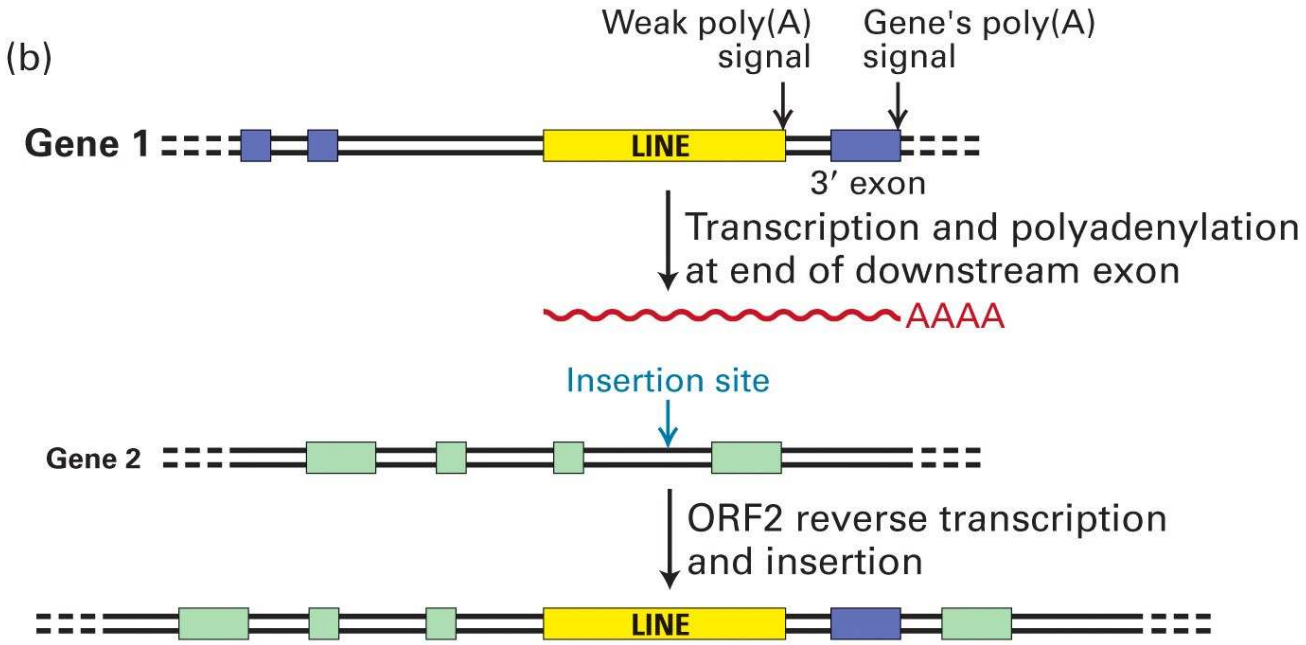


2Via DNA Transposones

Exon shuffling can also occur via cut-and-paste transpositions mediated by DNA transposons. The mechanism by which this occurs is illustrated in Fig. 6.19a. It requires that two copies of the transposon flank the target exon. Both DNA transposons and the exon will move as one piece of DNA if the transposase happens to cleave DNA at the left inverted repeat of the upstream transposon and at the right inverted repeat of the downstream transposon. Gene 1 ends up losing the exon, and Gene 2 acquires the exon

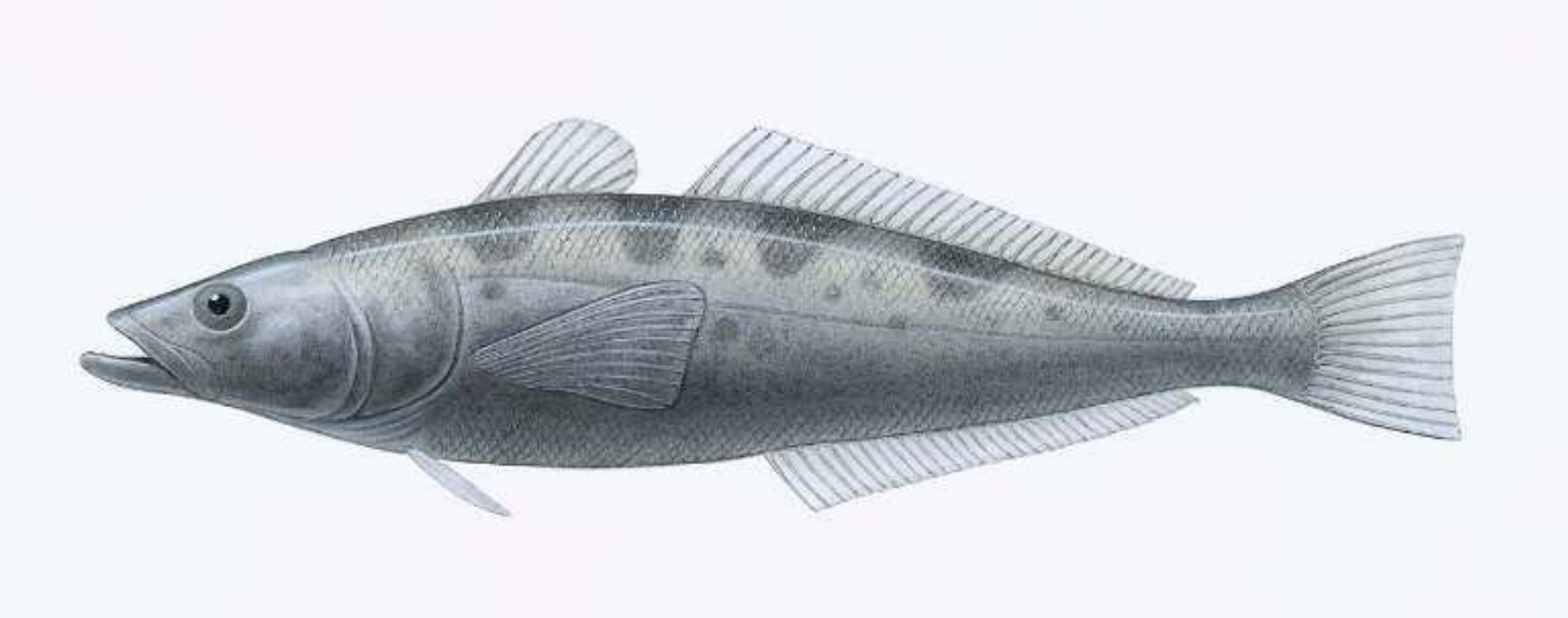


Exons can move along with a LINE element when it transposes via its copy-and-paste mechanism (Fig. 6.19b). When a LINE element has a weak poly(A) signal, RNA polymerase II continues to transcribe downstream, potentially through an exon. If this exon has a strong poly(A) signal, then transcription stops and the RNA is polyadenylated. Then following the mechanism in Fig. 6.17, DNA encoding the exon and the LINE element can be incorporated into another gene. The spliced mRNA produced from the acceptor gene may contain the newly introduced exon. Exon shuffling is supported by experimental evidence and the enormous amount of interspersed repeat DNA in genomes. Over billions of years, it has played a major role in evolution of genomes.



Cómo se originan nuevos genes

An example: the antifreeze glycoprotein (AFGP) gene in the Antarctic fish, *Dissostichus mawsoni*



Motif multiplication and exon loss

Gene function may be altered through the multiplication of motifs in an ancestral gene and/or loss of exons.

Notothenioid fish - live in very cold waters off Antarctica.

Avoid blood freezing; antifreeze glycoprotein (AFGP) genes - code for Thr-Ala-Ala repeat.

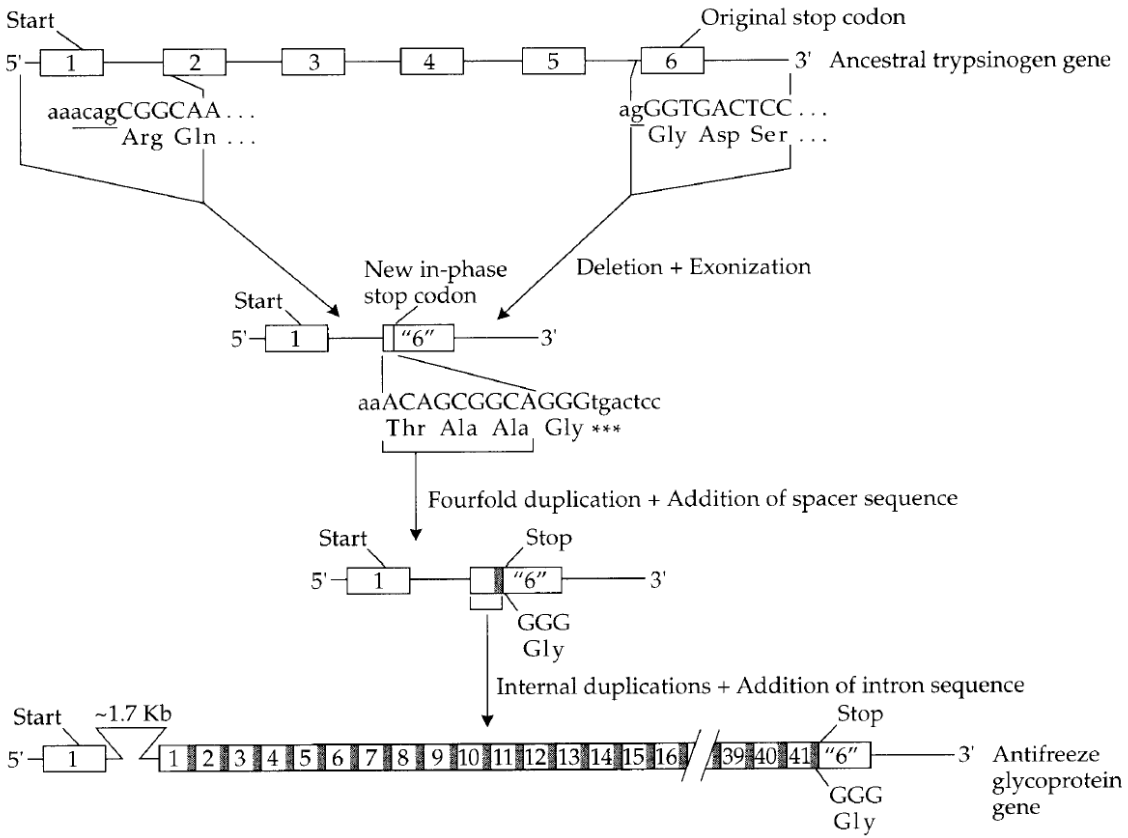
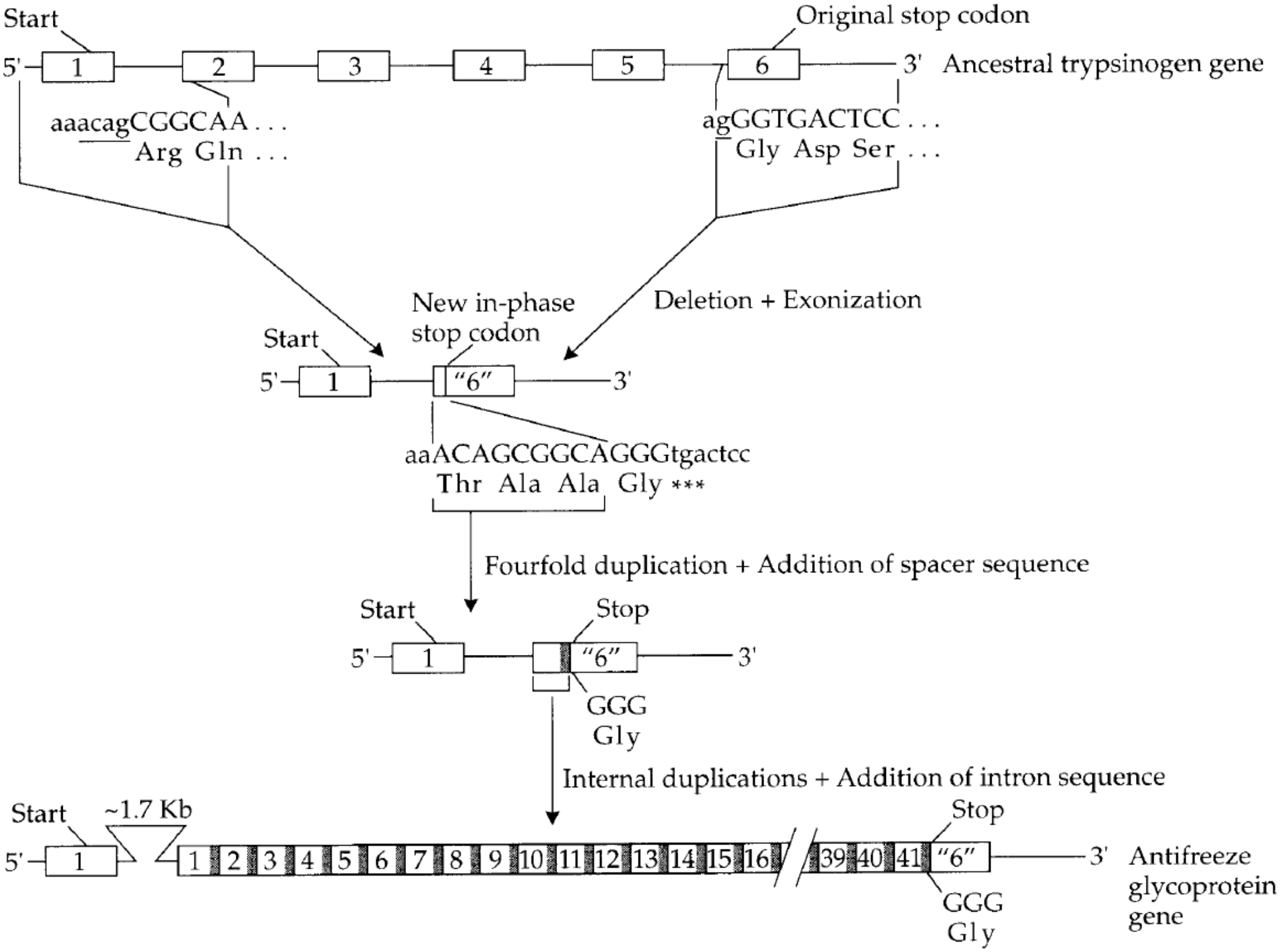
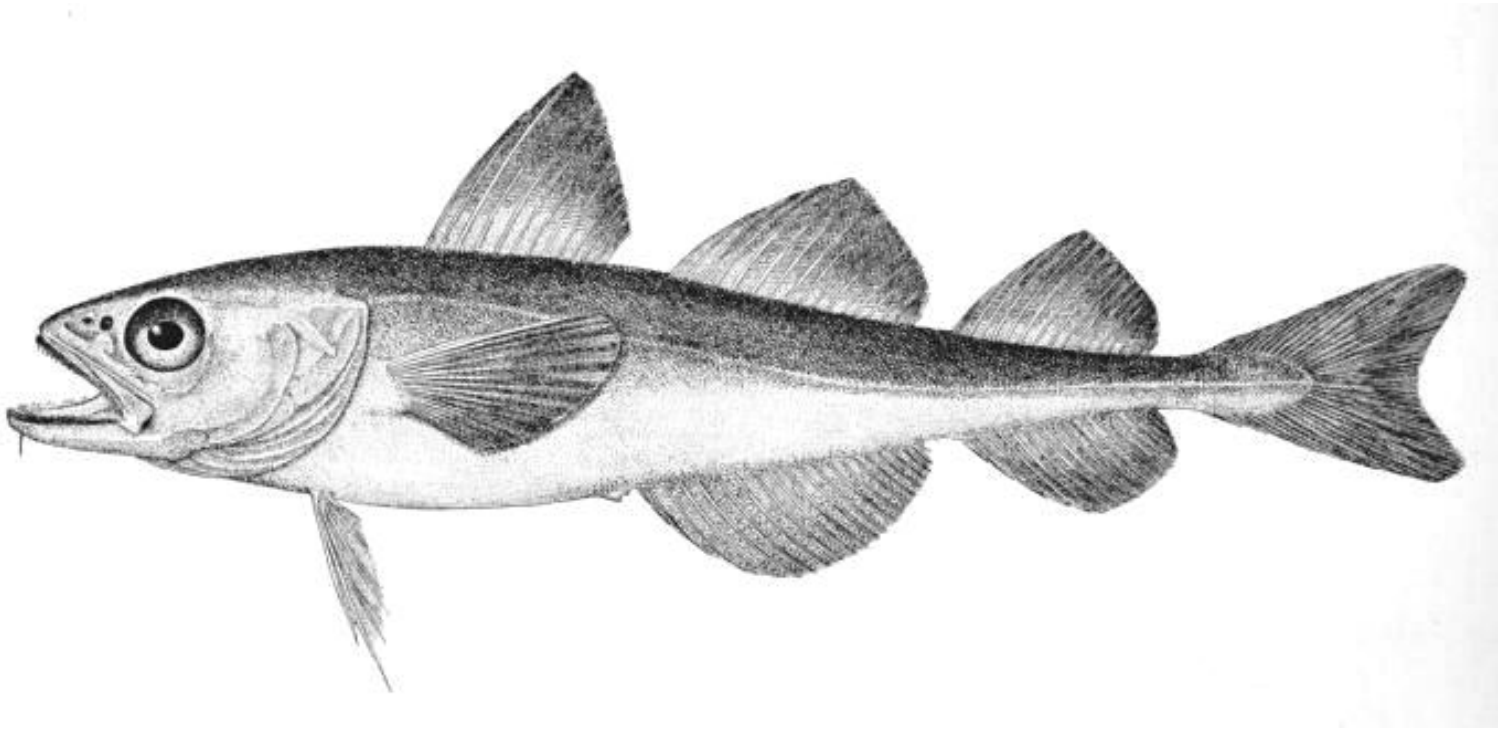


FIGURE 6.6 **A likely evolutionary pathway by which an ancestral trypsinogen gene with six exons (numbered boxes) was transformed into an antifreeze glycoprotein gene in the giant Antarctic cod (*Dissostichus mawsoni*). Following a deletion and the exonization of five intronic nucleotides (underlined lowercase letters) in the ancestral trypsinogen gene, a new gene with two exons emerged. (The second exon is marked "6" to emphasize its ancestry, and the new stop codon that was brought into frame by the exonization is marked by asterisks.) The sequence encoding Thr-Ala-Ala was duplicated to create a fourfold repetition, and a short spacer sequence of unknown origin (shaded box) was added to the repeated unit. Multiple internal gene duplications resulted in 41 repeats. The addition of the ~1.7 Kb-long sequence to the intron is indicated as a triangular loop. This addition could have occurred in any one of the previous steps in the evolution of this gene, and was added at the end for graphical convenience only.**


Start
Original stop codon
5'
1
2
3
4
5
6
3'
Ancestral trypsinogen gene
aaacagCGGCAA . . .
Arg Gln . . .
agGGTGACTCC . . .
Gly Asp Ser . . .
New in-phase
stop codon
Deletion + Exonization
Start
5'
1
"6"
3'
aaACAGCGGCAGGGtgactcc
Thr Ala Ala Gly ***
Fourfold duplication + Addition of spacer sequence
Start
Stop
5'
1
"6"
3'
GGG
Gly
Internal duplications + Addition of intron sequence
Start
~1.7 Kb
Stop
5'
1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
39 40 41 "6"
3'
Antifreeze
glycoprotein
gene
GGG
Gly

Convergent evolution of an AFGP gene in the arctic cod, *Boreogadus saida*



Convergent evolution of an AFGP gene in the arctic cod, *Boreogadus saida*

- the AFGP gene in *B. saida* also has a Thr-Ala-Ala repeating motif!
- appears to have evolved independently because:
- 1. flanking regions show no homology to trypsinogen
- 2. different number and locations of introns
- 3. codons used in repeating unit are different

Examples of proteins with internal domain duplications taking up 50% or more of the total length of the protein

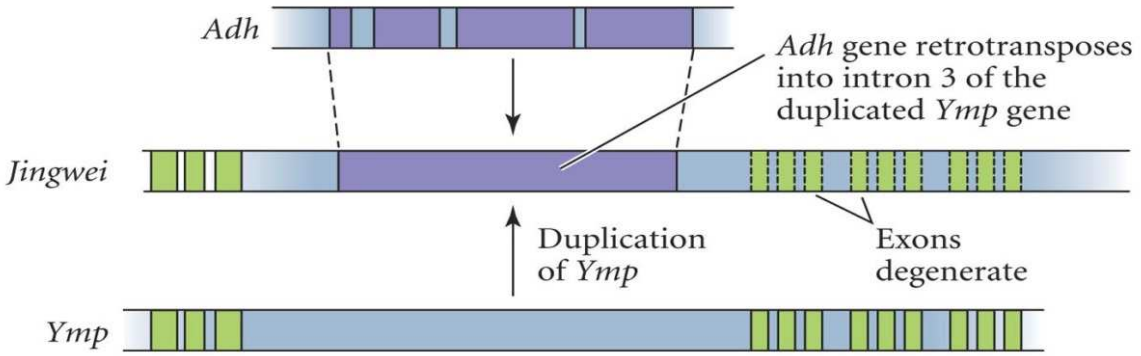
Sequence (organism)	Length of protein[a]	Length of repeat[a]	Number of repeats[b]	Percent repetition[c]
α1β-glycoprotein (human)	474	91	5	96
Angiotensin I-converting enzyme (human)	1,306	357	2	55
Calbindin (human, bovine)	260	43	6	99
Calcium-dependent regulator protein (human)	148	74	2	100
Ferredoxin (*Azobacter vinelandii*)	70	30	2	86
Ferredoxin (*Azobacter pasteurianum*)	55	28	2	100
Hemopexin (human)	439	207	2	94
Hexokinase (human)	917	447	2	97
Immunoglobulin γ chain C region (human)	329	108	3	98
Immunoglobulin ε chain C region (human)	423	108	4	100
Interleukin-2 receptor (human)	251	68	2	54
Interstitital retinol-binding protein (bovine)	1,263	302	4	96
Lactase-phlorizin hydrolase (human)	1,927	480	3	79
Lymphocyte activation gene-3 protein (human)	470	138	2	59
Multidrug resistance-1 P-glycoprotein (human)	1,280	609	2	95
Ovoinhibitor (chicken)	472	64	7	95
Parvalbumin (human)	108	39	2	72
Plasminogen (human)	790	79	5	50
Preproglucagon (rat)	180	36	3	60

Gen Quimera

Los Genes quimera se producen bien por via exon shuffling o retrotransposicion.

Mosaic (or **chimeric**) **protein** = a protein encoded by a gene that contains regions also found in other genes. The existence of such proteins provides evidence of exon shuffling.

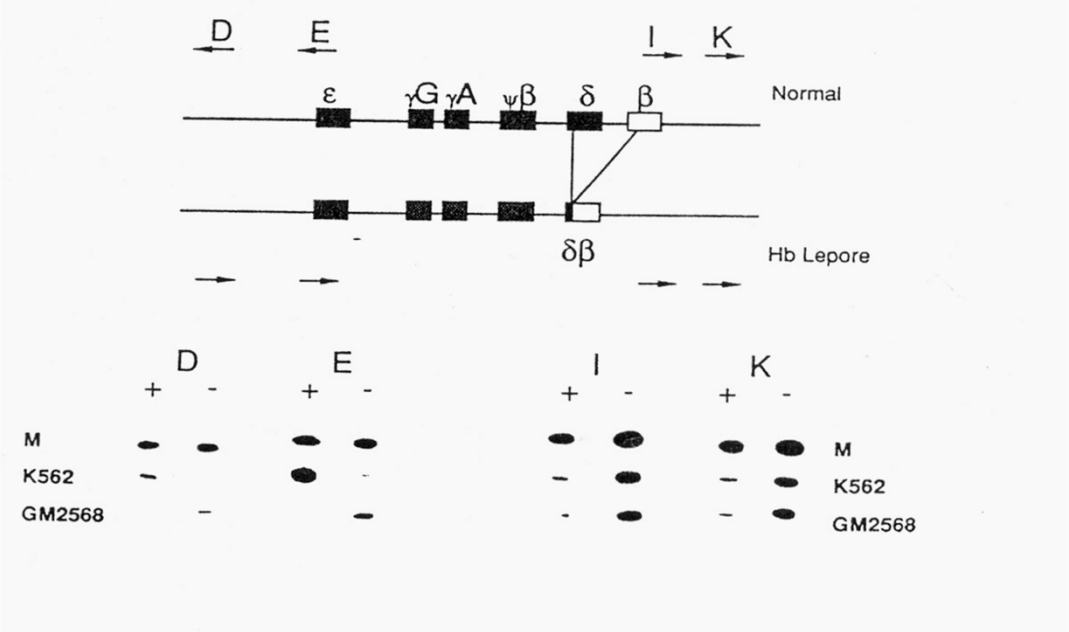
Under retrotransposition, mRNA transcribed to cDNA, which is then inserted into a gene (without introns).



EVOLUTION, Figure 19.14 © 2005 Sinauer Associates, Inc.

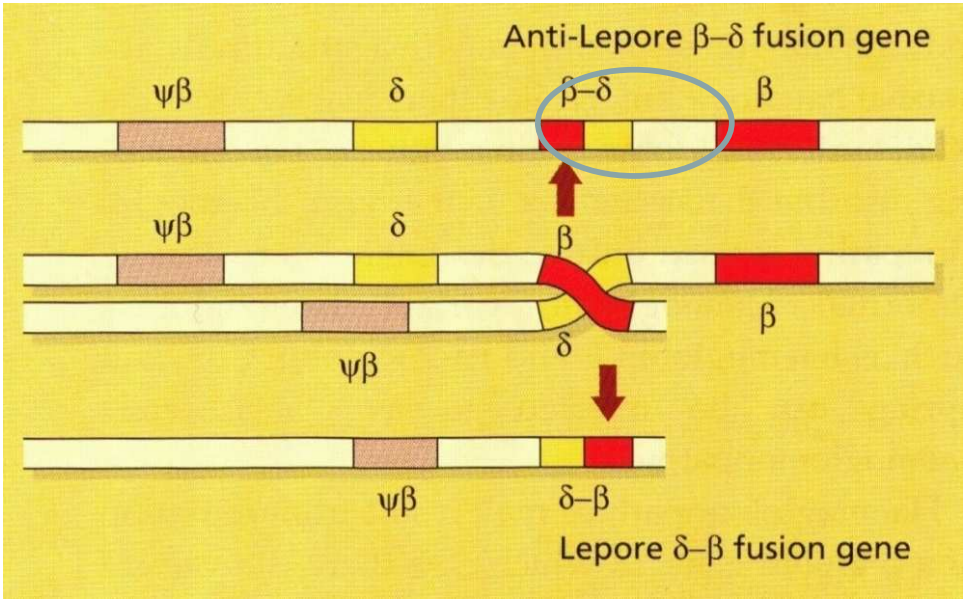
Both *Jingwei* and *Ymp* are expressed in *Drosophila* testes.

- Gen quimera: Hemoglobina Lepore (Gen quimera)



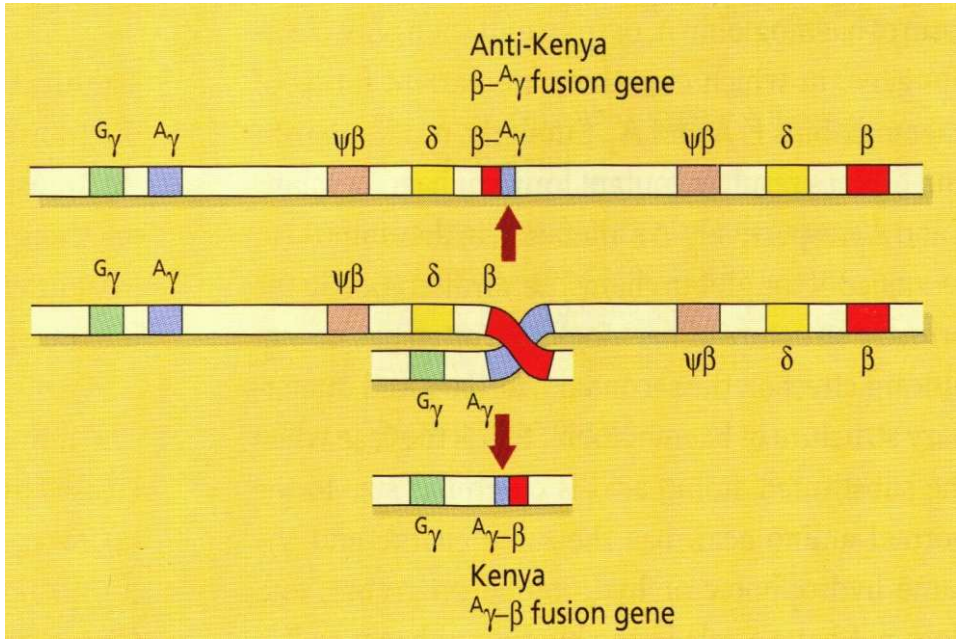
- Debido a que los locus de los genes de globina contienen grupos de genes similares existe el potencial para que haya un entrecruzamiento desigual entre las cromátides hermanas durante la meiosis. La generación de **hemoglobina Gun Hill** y **hemoglobina Lepore** son el resultado de eventos de entrecruzamientos desiguales. La Hemoglobina Gun Hill es el resultado de una deleción de 15 nucleótidos causados por el cruce desigual entre los codones 91–94 de un gen de β-globina y los codones 96–98 del otro. La generación de la hemoglobina Lepore resulta del cruce desigual entre δ-globina y los genes β-globina. El gen híbrido resultante δβ se llama Lepore y el gen híbrido βδ se llama anti-Lepore. Según lo indicado anteriormente, el promotor del gen δ-globina es ineficiente así las consecuencias de este evento de entrecruzamiento desigual son tanto cualitativas como cuantitativas

Hb Lepore



- Anti-Lepore
  - βδ fusion gene
  - Longer than the original
- Lepore
  - δβ fusion gene
  - Shorter than the original

Hb Kenya



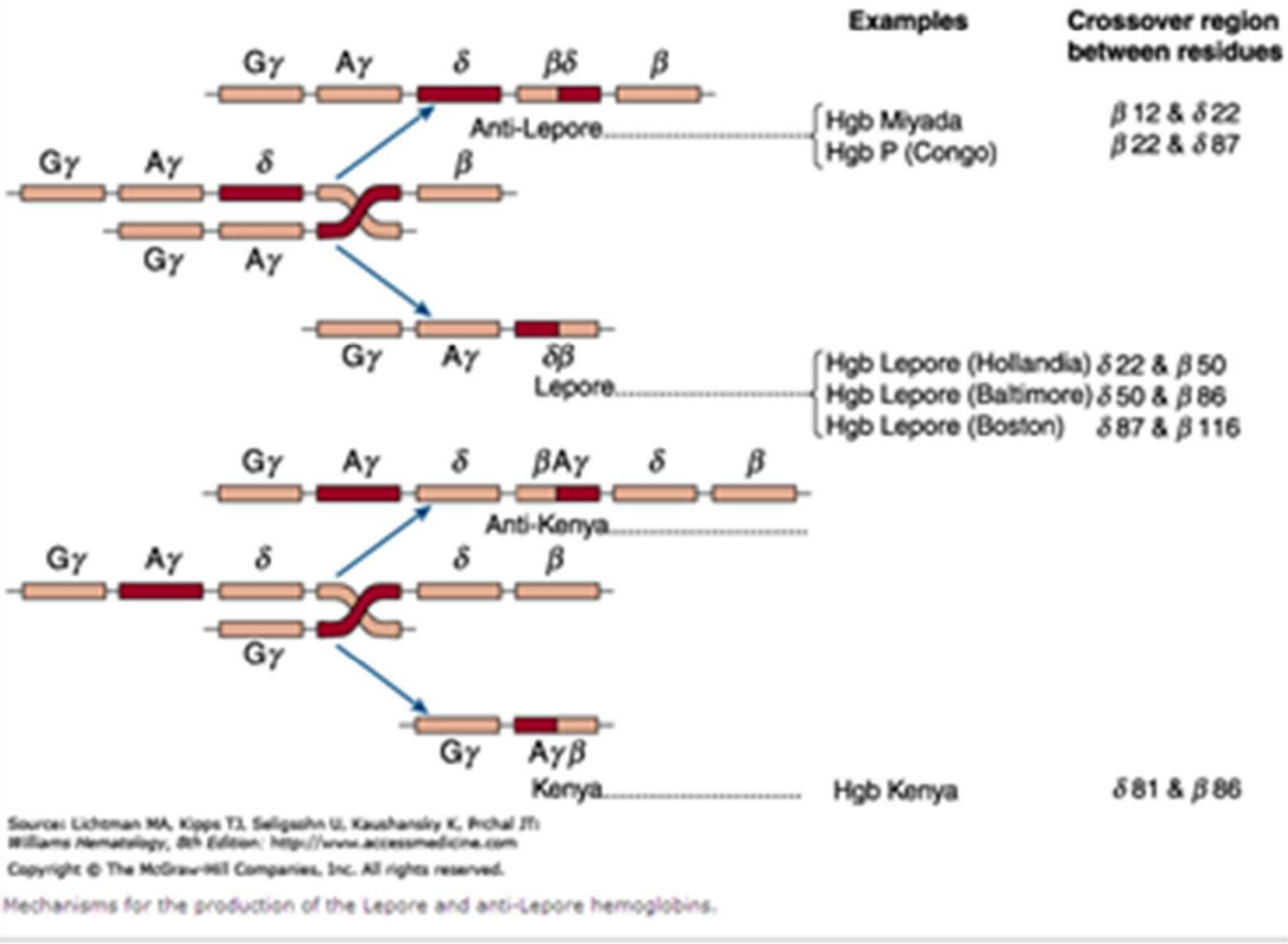
- Anti-Kenya
  - β-^{A}γ fusion
  - (much) longer than the original
- Kenya
  - ^{A}γ- β fusion
  - (much) shorter than the original

- d- Hemoglobinas de fusión

Hemoglobina Lepore

Hemoglobina Kenya

Hb Lepore tiene los primeros 20 a 80 aminoácidos de las cadenas δ y los últimos 50 a 100 aminoácidos del extremo C-terminal de la cadena β.



Source: Lichtman MA, Kipps TJ, Seligsohn U, Kaushansky K, Prchal JT: *Williams Hematology, 8th Edition*: http://www.accessmedicine.com

Mechanisms for the production of the Lepore and anti-Lepore hemoglobins.