Evolución de secuencias de DNA

4 transiciones (α)

8 transversiones (β)

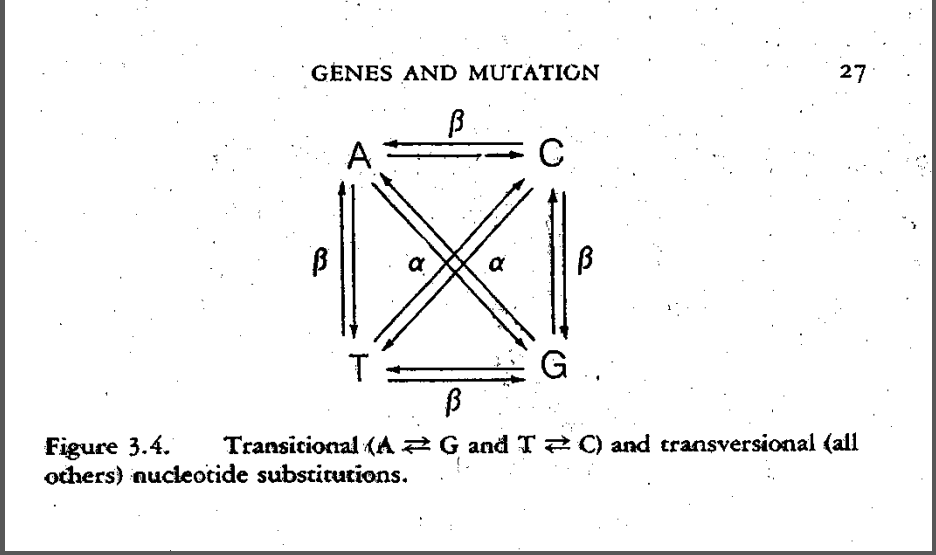


Figure 3.4. Transitional (A ⇄ G and T ⇄ C) and transversional (all others) nucleotide substitutions.

Proporción esperada Transiciones/Transversiones = 1/2

DESDE **ATCTAGATCTAGTGCATAGCATGCA**

|*|||*||*|||||*|||*|||*||

HASTA **ACCTAAATTTAGTGAATATCATCCA**

Desde/Hasta	A	T	C	G
A	--	A → T	A → C	A → G
T	T → A	--	T → C	T → G
C	C → A	C → T	--	C → G
G	G → A	G → T	G → C	--

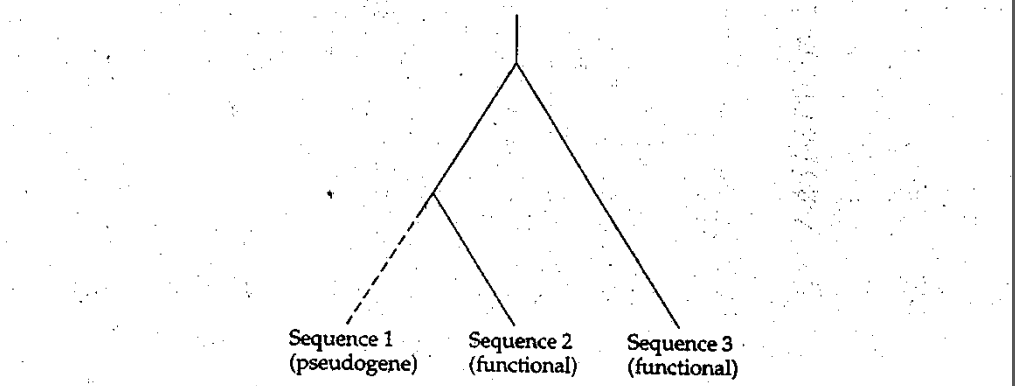


Figure 1.17 A tree for inferring the pattern of nucleotide substitution in a pseudogene sequence. The dashed line implies "nonfunctional." From Li and Graur (1991).

A	G	G	G →A (en 1)
G	T	T	T →G (en 1)
A	G	A	Se excluye
A	G	T	Se excluye
A	G	C	Se excluye

TABLE 1.5 Pattern of nucleotide substitution in pseudogenes[a]

From	To A	To T	To C	To G	Row totals
A	— —	4.7±1.3 (5.3±1.4)	5.0±0.7 (5.6±0.8)	9.4±1.3 (10.3±1.4)	19.1 (21.2)
T	4.4±1.1 (4.8±1.1)	— —	8.2±1.3 (9.2±1.3)	3.3±1.2 (3.6±1.3)	15.9 (17.6)
C	6.5±1.1 (7.1±1.3)	21.0±2.1 (18.2±2.3)	— —	4.2±0.5 (4.2±0.6)	31.7 (29.5)
G	20.7±2.2 (18.6±1.9)	7.2±1.1 (7.7±1.3)	5.3±1.0 (5.5±1.3)	— —	33.2 (31.8)
Column totals	31.6 (30.5)	32.9 (31.2)	18.5 (20.3)	16.9 (18.1)	

From Gojobori et al. (1982b) and Li et al. (1984).

[a]Table entries are the inferred percentages (f_{ij}) of base changes from i to j based on 13 mammalian pseudogene sequences. Values in parentheses were obtained by excluding all CG dinucleotides from comparison.

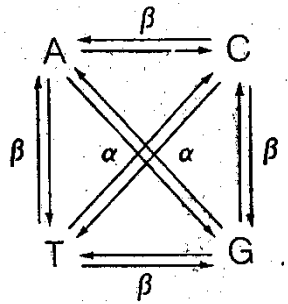


Figure 3.4. Transitional (A ⇄ G and T ⇄ C) and transversional (all others) nucleotide substitutions.

- Las transiciones suponen el 59.3%, frente al 33% (4/12) esperado
- C y G son los nucleótidos más mutables (la mutacion C → T es la más frecuente, salvo en las islas CpG)
- La mayoría de las mutaciones son hacia A o T

➔Mutación de base en el genoma: aumento de A+T

Pero los genes son ricos en G+C !!

TABLE 1.6 Pattern of nucleotide substitution in the control region of mtDNA[a]

From	To A	T	C	G	Row totals
A	—	0.4	1.1	14.1	15.6
T	0.3	—	33.8	0.3	34.4
C	1.1	25.8	—	0.5	27.4
G	20.0	1.1	1.6	—	22.7
Column totals	21.4	27.3	36.5	14.9	

From Tamura and Nei (1993).

[a]Table entries are the inferred percentages (f_{ij}) of base changes from i to j based on 95 sequences of the control region of human mtDNA.

Transition/Transversion ratio = 15.7 !!

Genes funcionales:

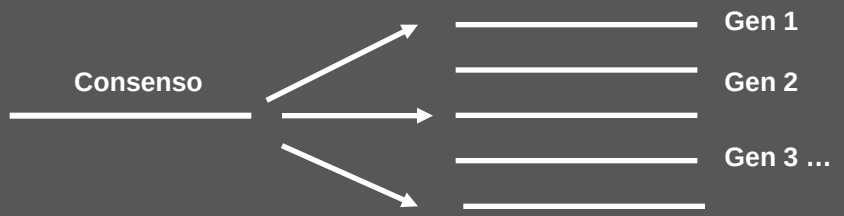


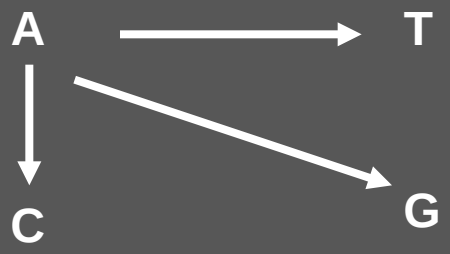
Table 5.5 Relative substitution rates among the four nucleotides A, T, C, and G in functional globin and ACTH genes. From Gojobori et al. (1982a).

Mutant nucleotide	Original nucleotide A	T	C	G
A		5.0±1.2	8.1±0.6	20.9±3.1
T	4.2±2.0		9.4±3.3	4.8±0.8
C	6.3±2.8	3.7±1.6		11.5±3.8
G	11.5±1.4	1.7±1.4	13.0±1.5	

```
ATCTAGATCTAGTGCATAGCATGCA
|*|||*||*|||||*|||*|||*||
ACCTAAATTTAGTGAATATCATCCA
```

$p = n_d / n$

n_d número de posiciones con nucleótidos diferentes



probabilidad = 1/3

No vale Poisson

Método de Jukes y Cantor

Transition matrix for the JC method:

O\S	A	T	C	G
A	1-3α	α	α	α
T	α	1-3α	α	α
C	α	α	1-3α	α
G	α	α	α	1-3α

→ ∑ = 1
→ ∑ = 1
→ ∑ = 1
→ ∑ = 1

$$K = -3/4 \ln(1 - 4/3\ p)$$

siendo $p = n_d / n$

Si la distancia entre ambas secuencias es baja, K es una buena estima de la divergencia.

Pero para distancias mayores, K sigue subestimando la divergencia

Método de Kimura con 2 parámetros (K2P)

Transition matrix for K2P method:

O\S	A	T	C	G
A	1-α-2β	β	β	α
T	β	1-α-2β	α	β
C	β	α	1-α-2β	β
G	α	β	β	1-α-2β

→ ∑ = 1
→ ∑ = 1
→ ∑ = 1
→ ∑ = 1

Table 5.1 Sixteen different types of nucleotide pairs between sequences X and Y

Identical nucleotides	AA	TT	CC	GG	Total
	R_1	R_2	R_3	R_4	**R**
Transitions	AG	GA	TC	CT	Total
	P_1	P_1	P_2	P_2	**P**
Transversions	AT	TA	AC	CA	Total
	Q_1	Q_1	Q_2	Q_2	
	TG	GT	CG	GC	
Frequency	Q_3	Q_3	Q_4	Q_4	**Q**

$$R + P + Q = 1$$

$$K = -\tfrac{1}{2} \ln [(1 - 2P - Q) \sqrt{1 - 2Q}]$$

Table 3.2 Models of nucleotide substitution.

(A) Jukes-Cantor model

	A	T	C	G
A	-	α	α	α
T	α	-	α	α
C	α	α	-	α
G	α	α	α	-

(B) Kimura model

	A	T	C	G
A	-	β	β	α
T	β	-	α	β
C	β	α	-	β
G	α	β	β	-

(C) Equal-input model

	A	T	C	G
A	-	αg_T	αg_C	αg_G
T	αg_A	-	αg_C	αg_G
C	αg_A	αg_T	-	αg_G
G	αg_A	αg_T	αg_C	-

(D) Tamura model

	A	T	C	G
A	-	$\beta\theta_2$	$\beta\theta_1$	$\alpha\theta_1$
T	$\beta\theta_2$	-	$\alpha\theta_1$	$\beta\theta_1$
C	$\beta\theta_2$	$\alpha\theta_2$	-	$\beta\theta_1$
G	$\alpha\theta_2$	$\beta\theta_2$	$\beta\theta_1$	-

(E) HKY model

	A	T	C	G
A	-	βg_T	βg_C	αg_G
T	βg_A	-	αg_C	βg_G
C	βg_A	αg_T	-	βg_G
G	αg_A	βg_T	βg_C	-

(F) Tamura-Nei model

	A	T	C	G
A	-	βg_T	βg_C	$\alpha_1 g_G$
T	βg_A	-	$\alpha_2 g_C$	βg_C
C	βg_A	$\alpha_2 g_T$	-	βg_G
G	$\alpha_1 g_A$	βg_T	βg_C	-

(G) General reversible model

	A	T	C	G
A	-	ag_T	bg_C	cg_G
T	ag_A	-	dg_C	eg_G
C	bg_A	dg_T	-	fg_G
G	cg_A	eg_T	fg_C	-

(H) Unrestricted model

	A	T	C	G
A	-	a_{12}	a_{13}	a_{14}
T	a_{21}	-	a_{23}	a_{24}
C	a_{31}	a_{32}	-	a_{34}
G	a_{41}	a_{42}	a_{43}	-

Note: An element (e_{ij}) of the above substitution matrices stands for the substitution rate from the nucleotide in the i-th row to the nucleotide in the j-th column. g_A, g_T, g_C, and g_G are the nucleotide frequencies. $\theta_1 = g_G + g_C$, $\theta_2 = g_A + g_T$.

Método de Tamura

El contenido en G+C (θ) de las secuencias puede ser ≠ 0.5

(D) Tamura model

A	-	$\beta\theta_2$	$\beta\theta_1$	$\alpha\theta_1$
T	$\beta\theta_2$	-	$\alpha\theta_1$	$\beta\theta_1$
C	$\beta\theta_2$	$\alpha\theta_2$	-	$\beta\theta_1$
G	$\alpha\theta_2$	$\beta\theta_2$	$\beta\theta_1$	-

$$d = -h\ln(1 - P/h - Q) - (1/2)(1 - h)\ln(1 - 2Q)$$

where $h = 2\theta(1 - \theta)$, and θ is the GC content.

Método de Tamura y Nei

Toma en cuenta tanto el sesgo en G+C como en la proporción de transiciones / transversiones

(F) Tamura-Nei model

-	βg_T	βg_C	$\alpha_1 g_G$
βg_A	-	$\alpha_2 g_C$	βg_C
βg_A	$\alpha_2 g_T$	-	βg_G
$\alpha_1 g_A$	βg_T	βg_C	-

$$d = -\frac{2g_A g_G}{g_R}\ln\left[1 - \frac{g_R}{2g_A g_G}P_1 - \frac{1}{2g_R}Q\right] - \frac{2g_T g_C}{g_Y}\ln\left[1 - \frac{g_Y}{2g_T g_C}P_2 - \frac{1}{2g_Y}Q\right] - 2\left[g_R g_Y - \frac{g_A g_G g_R}{g_R} - \frac{g_T g_C g_R}{g_Y}\right]\ln\left[1 - \frac{1}{2g_R g_Y}Q\right]$$

Modelo general de sustitución de nucleótidos (12 parámetros)

J. theor. Biol. (1990) **142,** 485–501

The General Stochastic Model of Nucleotide Substitution

F. Rodríguez, J. L. Oliver†, A. Marín and J. R. Medina‡†

G4H

	A	T	C	G
A	—	α	β	χ
T	δ	—	ε	ϕ
C	γ	η	—	ι
G	ψ	κ	λ	—

$$\delta = -\mathrm{tr}\,[(X(0))\log(X(0)^{-1}X(t))].$$

PERO no da estimas mejores que los de 6 parámetros

Comparación de los diferentes métodos

- En principio, se podría pensar que los métodos con más parámetros serán mejores, ya que son más realistas
- Pero en la práctica no es así debido a que hay que estimar más parámetros → menos estadística → más errores de muestreo (las secuencias son finitas...)
- Otro problema es que la fórmula de los métodos con más parámetros a menudo es inaplicable (logaritmos con argumento negativo, etc.)

Ejemplo: estimas de K entre la α-globina de conejo y ratón

Posición del codón	JC	K2P	GIN (6 parámetros)
I	0.64	0.64	0.68
II	0.42	0.42	0.52
III	0.88	0.92	Inaplicable

Divergencia

II

I

III

Table 3.3 Observed numbers of the 10 pairs of nucleotides between the DNA sequences for the human and Rhesus monkey mitochondrial cytochrome *b* genes.

	Transition		Transversion				Identical Pair					
Codon Position	TC	AG	TA	TG	CA	CG	TT	CC	AA	GG	n_d	Total (n)
First	21	22	5	1	5	4	68	93	100	56	58	375
Second	20	3	6	1	0	2	140	87	71	45	32	375
Third	60	16	6	5	49	2	11	122	102	2	138	375
All	101	41	17	7	54	8	219	302	273	103	228	1125

Note: The numbers at the first, second, and third codon positions are shown separately.

Table 3.4 Estimates ($\hat{d}$) of the number of nucleotide substitutions per site between the human and Rhesus monkey mitochondrial cytochrome *b* genes for the first, second, and third codon positions ($\hat{d} \times 100$).

Position in Codon	p	Jukes-Cantor	Kimura	Tajima-Nei	Tamura-Nei
First	15.5 ± 1.9	17.3 ± 2.4	17.8 ± 2.5	18.0 ± 2.6	17.9 ± 2.5
Second	8.5 ± 1.4	9.1 ± 1.6	9.2 ± 1.7	9.2 ± 1.7	9.3 ± 1.7
Third	36.8 ± 2.5	50.6 ± 4.9	52.3 ± 5.4	66.5 ± 9.4	87.9 ± 39.0

Divergencia

II

I

III

II: Valores bajos → Corrección poco importante (8.5 → 9.3)

III: " altos → " muy " (36.8 → 87.9)

Simulación

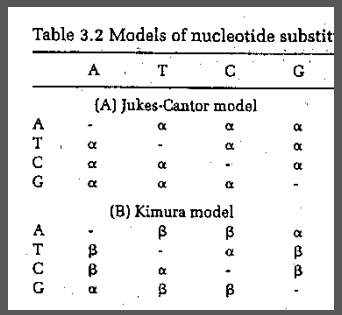
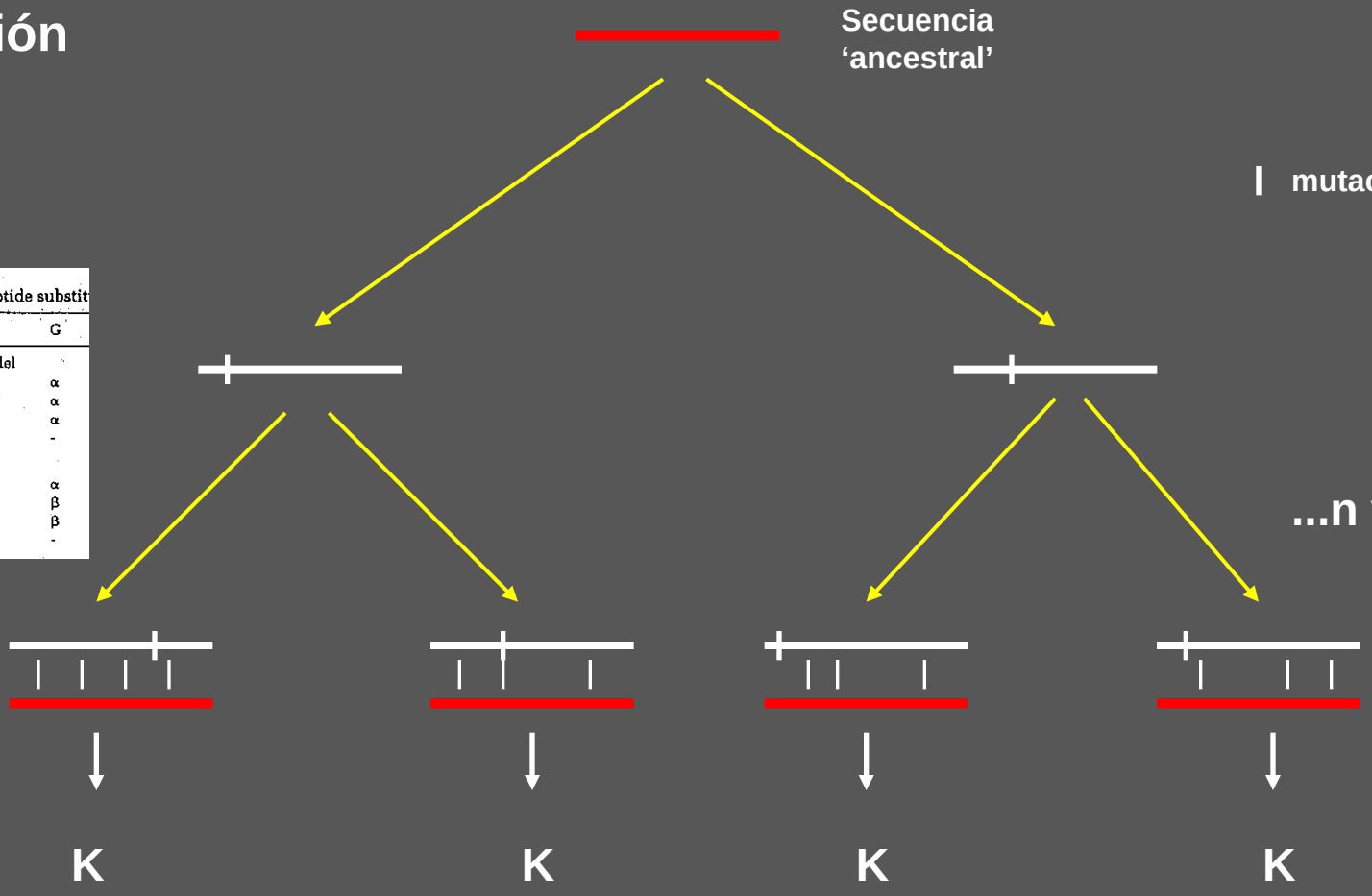


Table 3.2 Models of nucleotide substit

	A	T	C	G
(A) Jukes-Cantor model				
A	-	α	α	α
T	α	-	α	α
C	α	α	-	α
G	α	α	α	-
(B) Kimura model				
A	-	β	β	α
T	β	-	α	β
C	β	α	-	β
G	α	β	β	-



Simulación

TABLE 4.1 Estimates of the number of nucleotide substitutions per site (K) by different methods[a] from simulation data with sequence length L and the true value of K = 1

	L = 100 bp					L = 1000 bp				
Models[b]	JC	K2	TN	GIN	LA	JC	K2	TN	GIN	LA
JC	1.025	1.044	1.085	1.113	1.125	1.003	1.004	1.008	1.011	1.012
K2	0.935	1.050	1.091	1.094	1.107	0.901	1.006	1.011	1.016	1.017
EI	0.904	0.943	1.114	1.045	1.122	0.887	0.907	1.009	0.941	1.016
EIr	0.829	0.864	1.039	1.014	1.089	0.813	0.832	0.949	0.931	1.019
GIN	0.878	0.925	1.078	1.097	1.107	0.843	0.870	0.975	1.018	1.021
RAND	0.902	0.920	1.052	1.069	1.089	0.869	0.877	0.966	1.009	1.026
NA[c] (%)	0.0	1.1	2.2	24.4	33.9	0.0	0.0	0.3	3.7	4.8

From Zharkikh (1994).

[a]Methods: JC, Jukes and Cantor (1969); K2, Kimura's (1980) two-parameter method; TN, Tajima and Nei (1984); GIN, Gojobori, Ishii, and Nei (1982a); LA, Lanave et al. (1984).

[b]Models corresponded to the assumptions made in the simulation, so that each row simulated substitutions using the same model: JC, Jukes-Cantor (one-parameter); K2, Kimura's two-parameter model; EI, equal-input model (Tajima and Nei 1982); EIr, equal-input related models; GIN, Gojobori et al.'s (1982a) six-parameter model; RAND, random substitution parameters.

[c]NA: the proportion of inapplicable cases averaged over the six substitution models used.

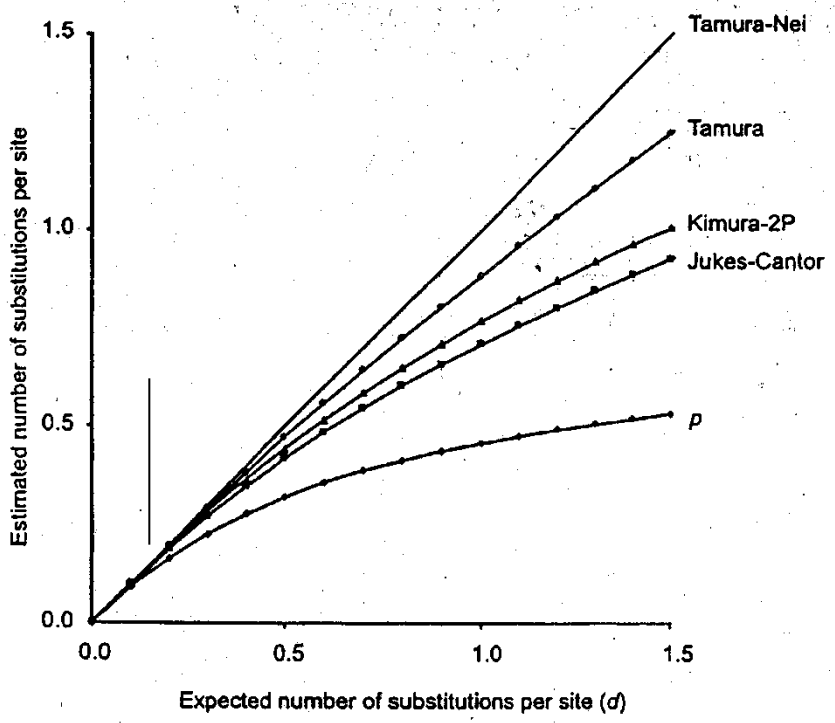


FIGURE 3.1. Estimates of the number of nucleotide substitutions obtained by different distance measures when actual nucleotide substitution follows the Tamura-Nei model. The nucleotide frequencies assumed are $g_A = 0.3$, $g_T = 0.4$, $g_C = 0.2$, and $g_G = 0.1$; and the two transition/transversion rate ratios assumed are $\alpha_1/\beta = 4$ and $\alpha_2/\beta = 8$.

Las distancias más sofisticadas (con más parámetros) son imprescindibles para estimar correctamente la longitud de las ramas en una filogenia, pero importan menos para estimar la topologia (orden de ramificación) correcta

Tasa de sustitución

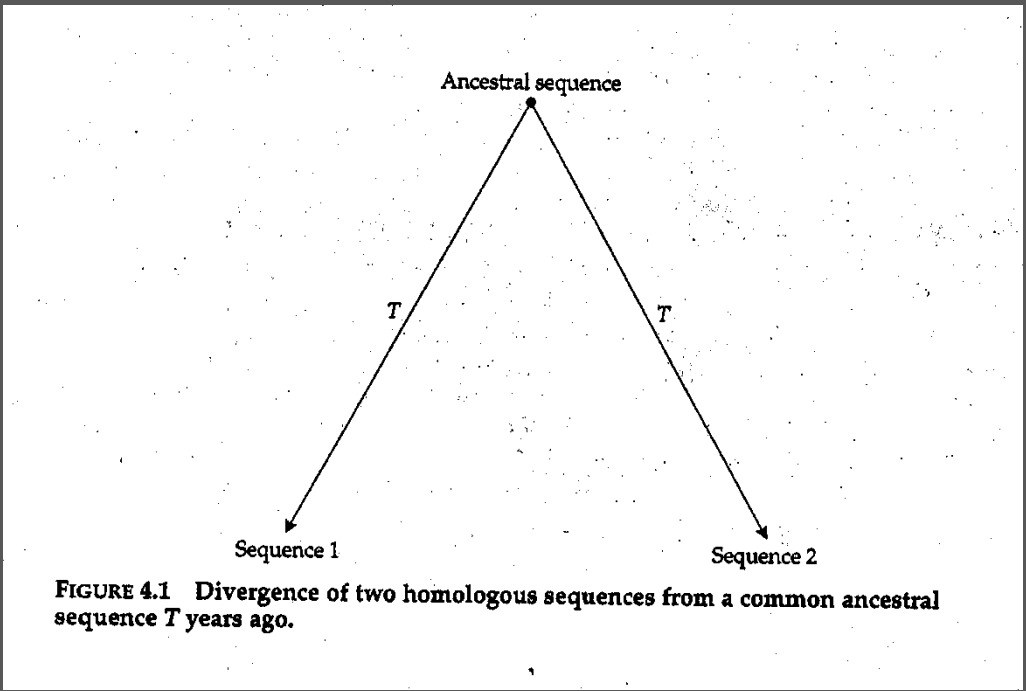


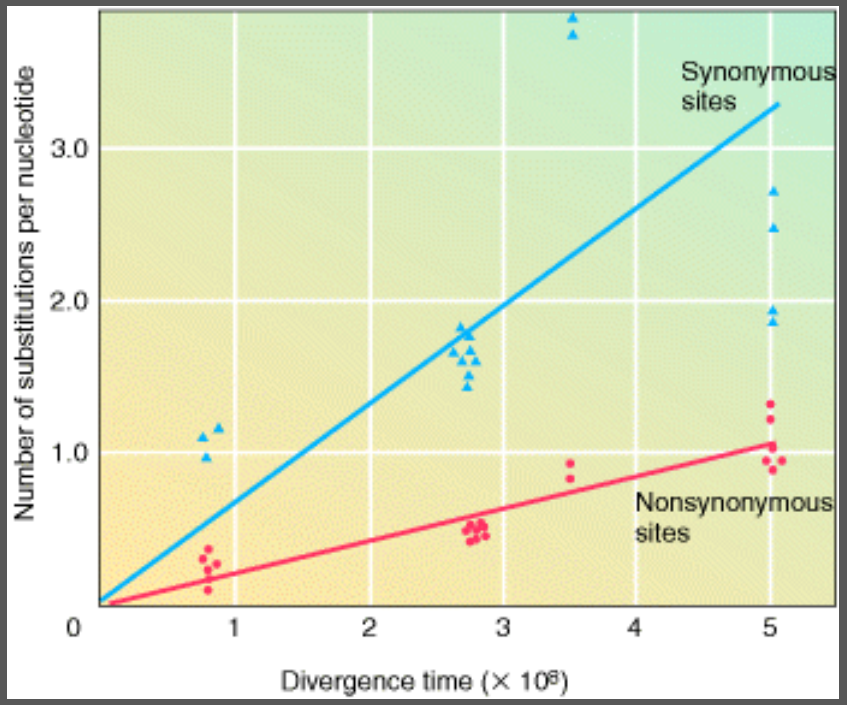
FIGURE 4.1 Divergence of two homologous sequences from a common ancestral sequence T years ago.

K

Tasa de sustitución de nucleótidos:

$$r = K / 2T$$

β-globina



- **SS > NS**
- **Mayor dispersión en SS**

The amount of nucleotide divergence at synonymous and nonsynonymous sites of the *β*-globin gene as a function of time since divergence.

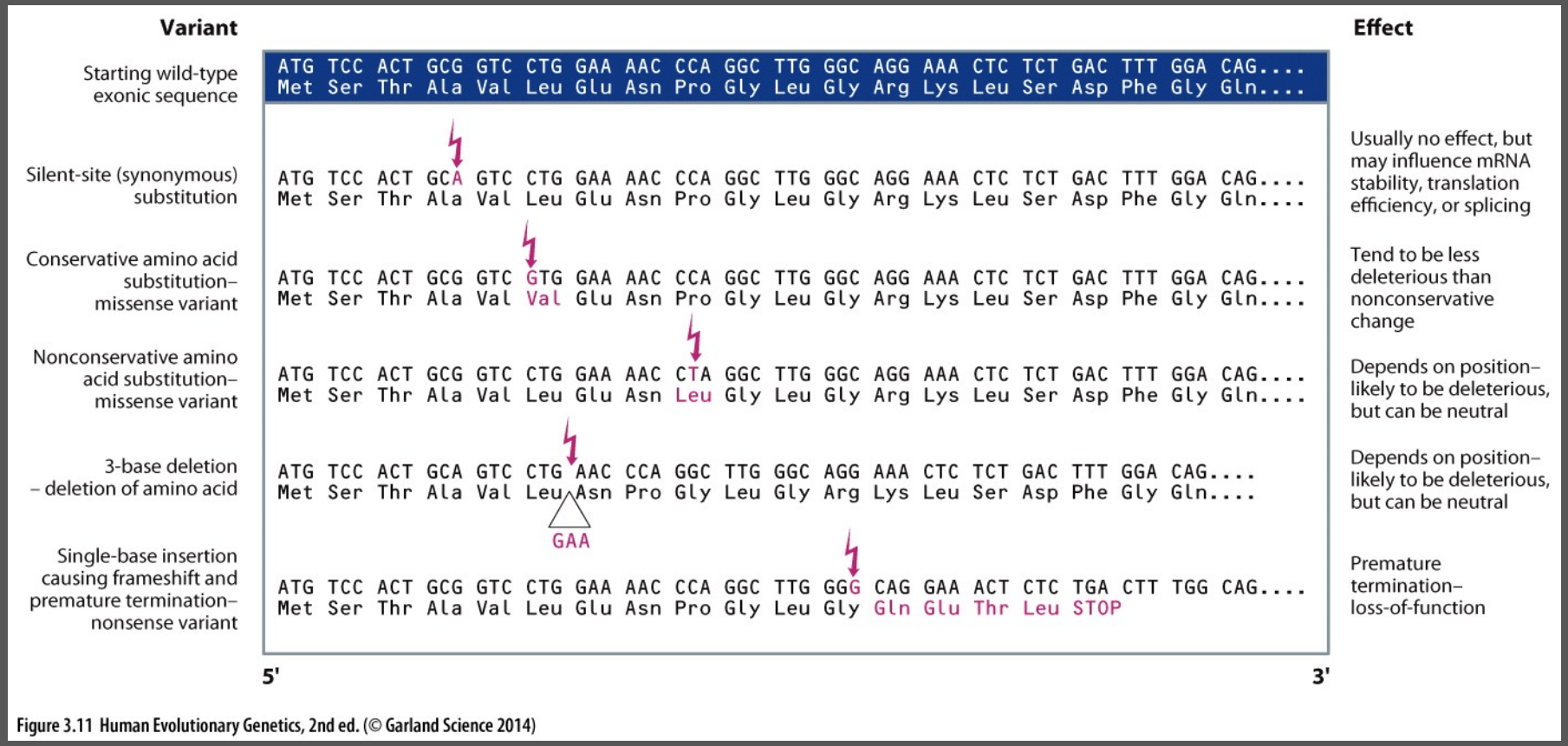
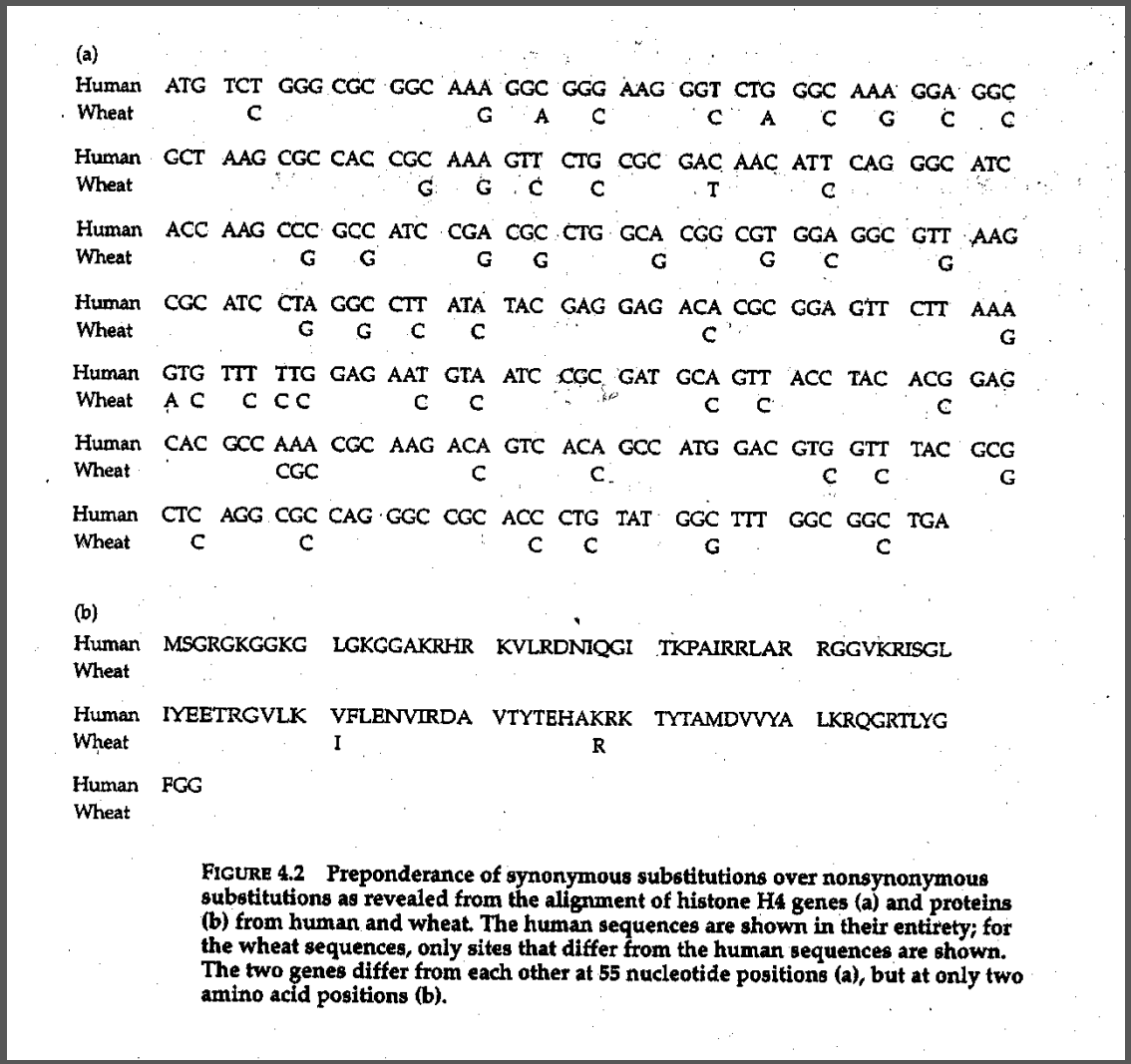


Figure 3.11 Human Evolutionary Genetics, 2nd ed. (© Garland Science 2014)

Histona H4



(a)

Human	ATG	TCT	GGG	CGC	GGC	AAA	GGC	GGG	AAG	GGT	CTG	GGC	AAA	GGA	GGC
Wheat		C				G	A	C		C	A	C	G	C	C
Human	GCT	AAG	CGC	CAC	CGC	AAA	GTT	CTG	CGC	GAC	AAC	ATT	CAG	GGC	ATC
Wheat					G	G	C	C		T		C			
Human	ACC	AAG	CCC	GCC	ATC	CGA	CGC	CTG	GCA	CGG	CGT	GGA	GGC	GTT	AAG
Wheat			G	G		G	G		G		G	C		G	
Human	CGC	ATC	CTA	GGC	CTT	ATA	TAC	GAG	GAG	ACA	CGC	GGA	GTT	CTT	AAA
Wheat			G	G	C	C				C					G
Human	GTG	TTT	TTG	GAG	AAT	GTA	ATC	CGC	GAT	GCA	GTT	ACC	TAC	ACG	GAG
Wheat	A C	C	C C		C	C				C	C			C	
Human	CAC	GCC	AAA	CGC	AAG	ACA	GTC	ACA	GCC	ATG	GAC	GTG	GTT	TAC	GCG
Wheat			CGC			C		C				C	C		G
Human	CTC	AGG	CGC	CAG	GGC	CGC	ACC	CTG	TAT	GGC	TTT	GGC	GGC	TGA	
Wheat	C		C				C	C		G			C		

(b)

Human	MSGRGKGGKG	LGKGGAKRHR	KVLRDNIQGI	TKPAIRRLAR	RGGVKRISGL
Wheat					
Human	IYEETRGVLK	VFLENVIRDA	VTYTEHAKRK	TYTAMDVVYA	LKRQGRTLYG
Wheat		I		R	
Human	FGG				
Wheat					

FIGURE 4.2 **Preponderance of synonymous substitutions over nonsynonymous substitutions as revealed from the alignment of histone H4 genes (a) and proteins (b) from human and wheat. The human sequences are shown in their entirety; for the wheat sequences, only sites that differ from the human sequences are shown. The two genes differ from each other at 55 nucleotide positions (a), but at only two amino acid positions (b).**

Gen:

- **55 sustituciones**
- **SS > NS**

Proteína:

- **2 sustituciones**

Humanos vs. roedores (ratón o rata)

Table 1. Rates of synonymous and nonsynonymous substitutions in various mammalian protein-coding genes.[a]

Gene	L^b	Nonsynonymous rate ($\times 10^9$)	Synonymous rate ($\times 10^9$)
HISTONES			
Histone 3	135	0.00 ± 0.00	6.38 ± 1.19
Histone 4	101	0.00 ± 0.00	6.12 ± 1.32
CONTRACTILE SYSTEM PROTEINS			
Actin α	376	0.01 ± 0.01	3.68 ± 0.43
Actin β	349	0.03 ± 0.02	3.13 ± 0.39
HORMONES, NEUROPEPTIDES, AND OTHER ACTIVE PEPTIDES			
Somatostatin-28	28	0.00 ± 0.00	3.97 ± 2.66
Insulin	51	0.13 ± 0.13	4.02 ± 2.29
Thyrotropin	118	0.33 ± 0.08	4.66 ± 1.12
Insulin-like growth factor II	179	0.52 ± 0.09	2.32 ± 0.40
Erythropoietin	191	0.72 ± 0.11	4.34 ± 0.65
Insulin C-peptide	35	0.91 ± 0.30	6.77 ± 3.49
Parathyroid hormone	90	0.94 ± 0.18	4.18 ± 0.98
Luteinizing hormone	141	1.02 ± 0.16	3.29 ± 0.60
Growth hormone	189	1.23 ± 0.15	4.95 ± 0.77
Urokinase-plasminogen activator	435	1.28 ± 0.10	3.92 ± 0.44
Interleukin I	265	1.42 ± 0.14	4.60 ± 0.65
Relaxin	54	2.51 ± 0.37	7.49 ± 6.10
HEMOGLOBINS AND MYOGLOBIN			
α-globin	141	0.55 ± 0.11	5.14 ± 0.90
Myoglobin	153	0.56 ± 0.10	4.44 ± 0.82
β-globin	144	0.80 ± 0.13	3.05 ± 0.56

(Continued on next page)

Humanos vs. roedores (ratón o rata)

Gene	L^b	Nonsynonymous rate ($\times 10^9$)	Synonymous rate ($\times 10^9$)
APOLIPOPROTEINS			
E	283	0.98 ± 0.10	4.04 ± 0.53
A-I	243	1.57 ± 0.16	4.47 ± 0.66
A-IV	371	1.58 ± 0.12	4.15 ± 0.47
IMMUNOGLOBULINS			
Ig V_H	100	1.07 ± 0.19	5.66 ± 1.36
Ig γ1	321	1.46 ± 0.13	5.11 ± 0.64
Ig k	106	1.87 ± 0.26	5.90 ± 1.27
INTERFERONS			
α1	166	1.41 ± 0.13	3.53 ± 0.61
β1	159	2.21 ± 0.24	5.88 ± 1.08
γ	136	2.79 ± 0.31	8.59 ± 2.56
OTHER PROTEINS			
Aldolase A	363	0.07 ± 0.03	3.59 ± 0.51
Hydroxanthine phosphoribosyltransferase	217	0.13 ± 0.04	2.13 ± 0.35
Creatine kinase M	380	0.15 ± 0.03	3.08 ± 0.37
Glyceradehyde-3-phosphate dehydrogenase	331	0.20 ± 0.05	2.84 ± 0.37
Lactate dehydrogenase A	331	0.20 ± 0.04	5.03 ± 0.61
Acetylcholine receptor γ subunit	540	0.29 ± 0.04	3.23 ± 0.31
Fibrinogen γ	411	0.55 ± 0.06	5.82 ± 0.67
Albumin	590	0.91 ± 0.07	6.63 ± 0.61
Average[c]		0.85 (0.73)	4.61 (1.44)

[a] All rates are based on comparisons between human and rodent genes and the time of divergence was set at 80 million years ago. Rates are in units of substitutions per site per 10^9 years.
[b] L = number of codons compared.
[c] Average is the arithmetic mean, and values in parentheses are the standard deviations, computed over all genes.

Variación en distintas regiones de los genes

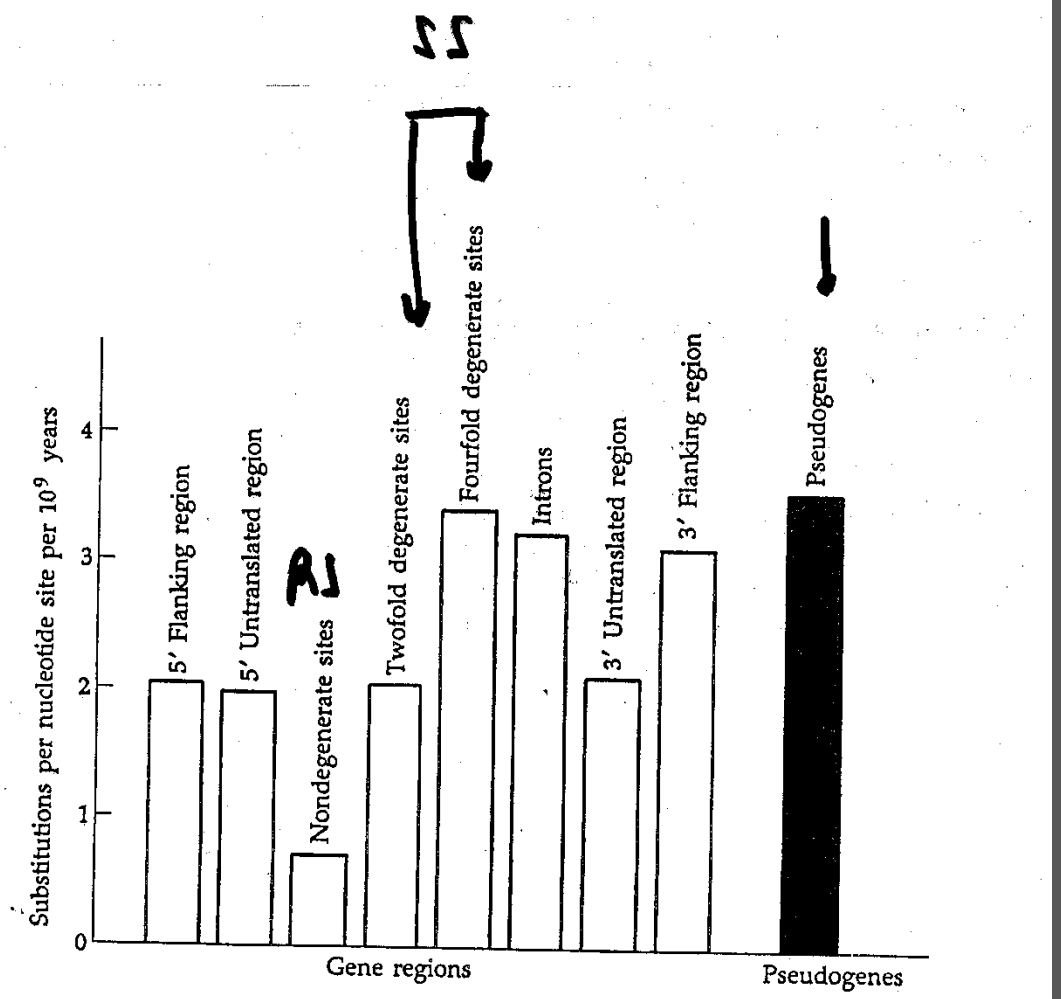


Figure 2. Average rates of substitution in different parts of genes and in pseudogenes.

Divergencia en exones vs. intrones

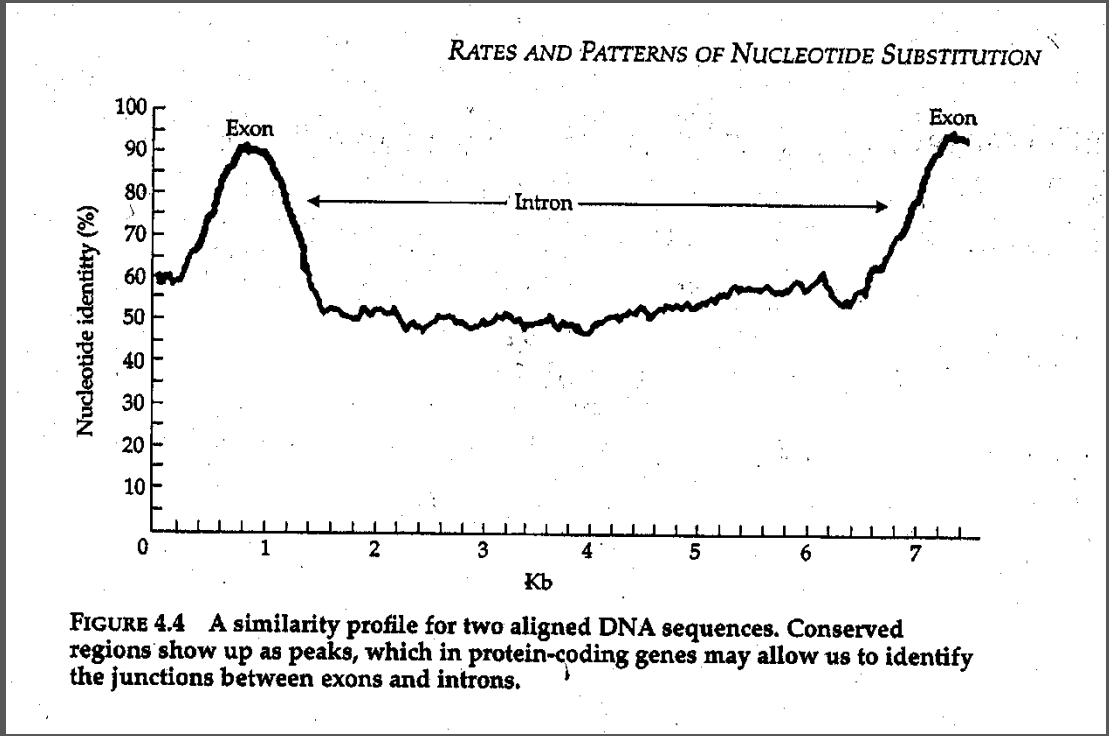
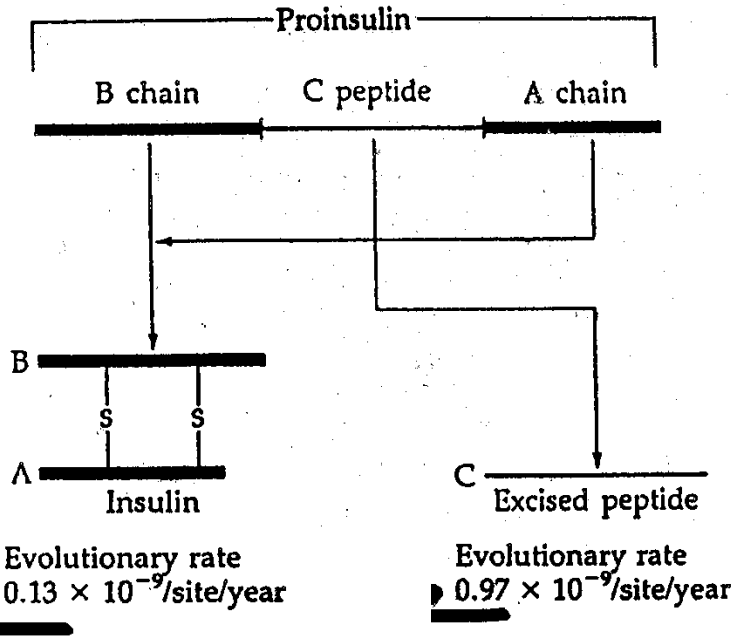


FIGURE 4.4 A similarity profile for two aligned DNA sequences. Conserved regions show up as peaks, which in protein-coding genes may allow us to identify the junctions between exons and introns.

Variación de NS en el gen de la insulina



NS

Figure 3. Comparison between the rates of nucleotide substitution for DNA regions coding for functional insulin (A and B chains) and the C peptide. A mature insulin molecule consists of one A and one B chain, linked by disulfide (S) bonds. Modifed from Kimura (1983).

Genoma mitocondrial de mamíferos

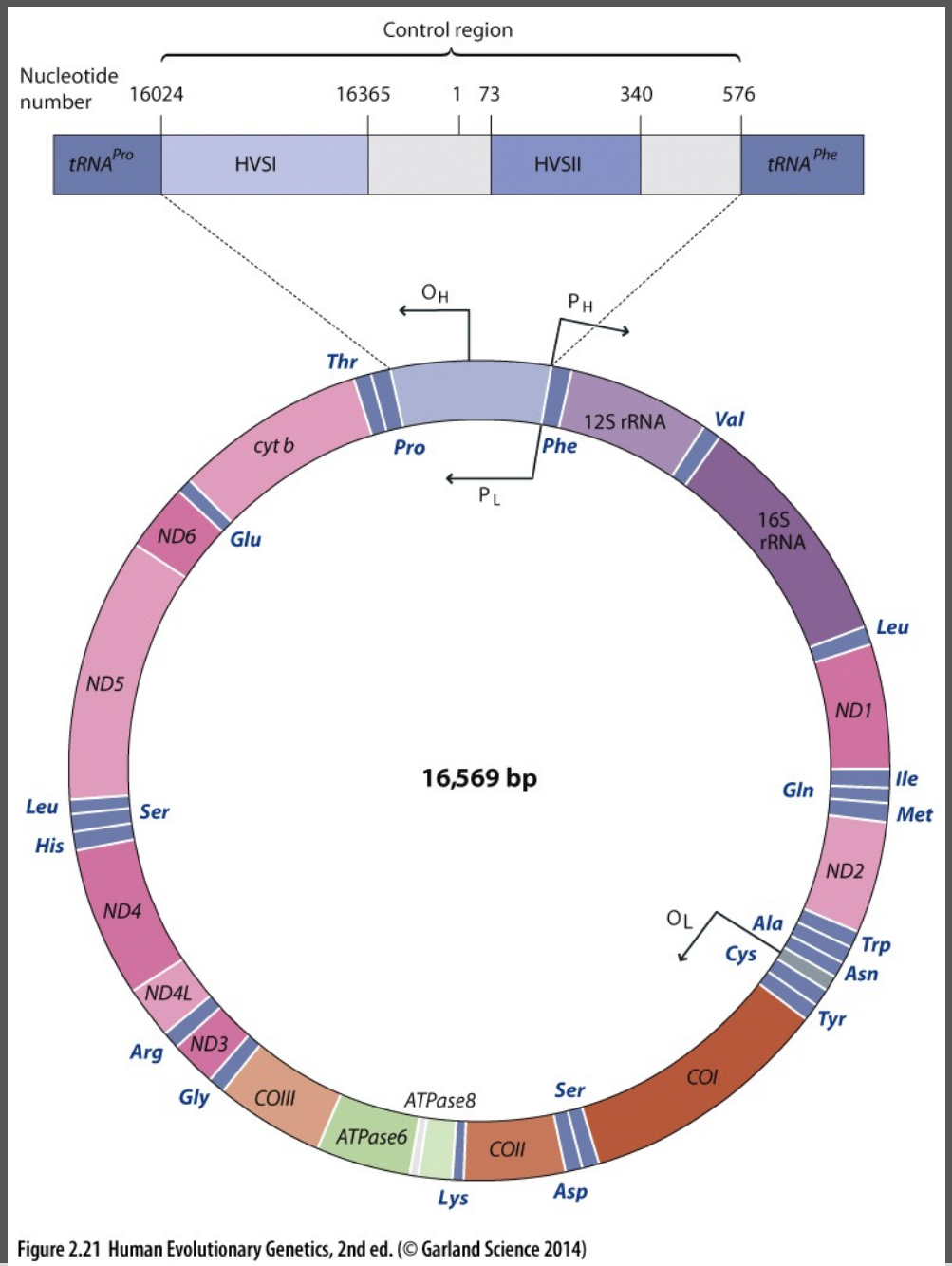


Figure 2.21 Human Evolutionary Genetics, 2nd ed. (© Garland Science 2014)

TABLE 2.3:
DIFFERENCES BETWEEN THE NUCLEAR AND mtDNA GENOMES

	Nuclear genome	mtDNA
Genome	linear chromosomes, ~50–250 Mb in size	circular molecule, ~16.5 kb
Packaging	packaged with histones as chromatin	not packaged
DNA replication origins	1000s of origins, sometimes diffuse	two discrete origins, one on each strand
Genes	~25,000 genes, mostly containing introns; much intergenic DNA	37 genes with no introns; almost no intergenic DNA
Transcripts	each gene has its own transcript(s)	only one from each strand; each mRNA covers many genes
Genetic code differences: UGA AGA, AGG AUA, AUU	 STOP Arg Ile	 Trp STOP Met

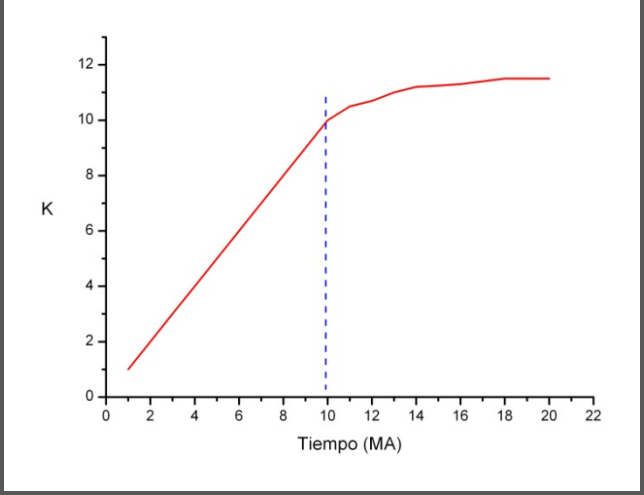
Table 2.3 Human Evolutionary Genetics, 2nd ed. (© Garland Science 2014)

Tasa evolutiva en el genoma mitocondrial de mamíferos

- Tasa de sustitución de nucleótidos muy alta: SS diez veces mayor que en el núcleo

- Causas:
  - Alta concentración de radicales superóxido O2-
  - Replicación poco fiel

- Saturación rápida: linealidad hasta sólo 10 MA → Reloj rápido



Filogenias entre especies próximas: primates, etc.

Tasa evolutiva en orgánulos de plantas

TABLE 7.7 Comparison of the rates of nucleotide substitution in plant chloroplast, mitochondrial, and nuclear genes[a]

Genomes	K_S	L_S	K_A	L_A
A. Comparison between monocot and dicot species				
Chloroplast genes	0.58 ± 0.02	4,177	0.05 ± 0.00	14,421
Mitochondrial genes	0.21 ± 0.01	1,219	0.04 ± 0.00	4,380
B. Comparison between maize and rice				
Nuclear genes	0.57 ± 0.01	8,898	0.07 ± 0.00	30,702
Chloroplast genes	0.12 ± 0.00	7,872	0.02 ± 0.00	28,518
Mitochondrial genes	0.05 ± 0.01	1,845	0.02 ± 0.00	6,357

From Wolfe et al. (1987, 1989a) and K. H. Wolfe (unpublished).

[a] K_S, number of substitutions per synonymous site. K_A, number of substitutions per nonsynonymous site between genes. L_S, number of synonymous sites. L_A, number of nonsynonymous sites.

Tasa evolutiva en virus

TABLE 7.9 Rates of synonymous and nonsynonymous substitutions of viral genes and nuclear genes

Organism	Gene	Substitutions per site per year		Ref[b]
		Synonymous	Nonsynonymous	
A. RNA viral genes				
Influenza A virus	Hemagglutinin	13.10×10^{-3}	3.59×10^{-3}	1
Hepatitis C virus	E	6.29×10^{-3}	0.32×10^{-3}	2
HIV-1	*gag*	9.70×10^{-3}	1.70×10^{-3}	3
Moloney murine sarcoma virus	v-*mos*	2.75×10^{-3}	0.82×10^{-3}	1
Moloney murine leukemia virus	*gag*	1.16×10^{-3}	0.54×10^{-3}	4
B. DNA viral genes				
Hepatitis B virus	P	4.57×10^{-5}	1.45×10^{-5}	5
Herpes simplex virus type 1	Genome[a]	3.50×10^{-8}		6
C. Nuclear genes				
Mammals	c-*mos*	5.23×10^{-9}	0.93×10^{-9}	1
Mammals	α globin	3.94×10^{-9}	0.56×10^{-9}	7

[a]Computed from restriction enzyme analysis of the genome.

[b]1, Gojobori et al. (1990); 2, Ina et al. (1994); 3, Li et al. (1988a); 4, Gojobori and Yokoyama (1985); 5, Orito et al. (1989); 6, Sakaoka et al. (1994); 7, Li et al. (1985b).

Tasa evolutiva en el virus de la gripe

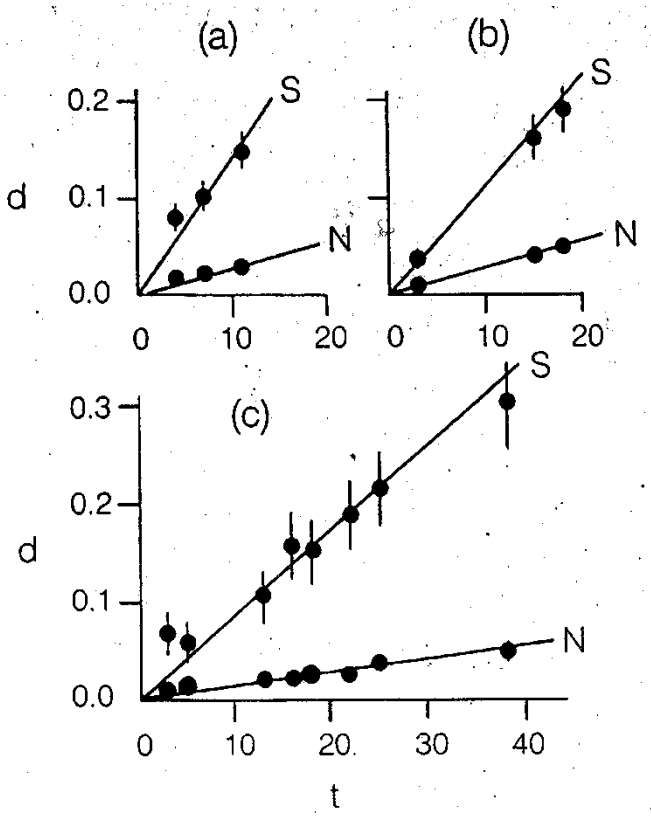


Figure 5.2. Relationships between the number of nucleotide substitutions *(d)* and evolutionary time *(t)* in influenza virus genes. Evolutionary time represents the difference in the year of isolation between the two strains compared. S: Synonymous substitution. N: Nonsynonymous substitution. (a) Hemagglutinin. (b) Neuraminidase. (c) Nonstructural protein. The rates of synonymous substitutions per site per year for the hemagglutinin, neuraminidase, and nonstructural genes are 0.014, 0.011, and 0.009, respectively, whereas the corresponding rates for nonsynonymous substitution are 0.0029, 0.0028, and 0.0015. Vertical lines represent standard errors. From Hayashida et al. (1985).