

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS

ANALES  
DE LA  
ESTACION EXPERIMENTAL  
DE AULA DEI

SEPARATA



## Evolución de genes cloroplásticos emigrados al núcleo en trigo y guisante

por J.L. OLIVER<sup>1</sup> Y J.M. MARTINEZ ZAPATER<sup>2</sup>

1) Unidad de Genética, Facultad de Ciencias, Universidad de Granada, E-18071 Granada

2) D.O.E. Plant Research Lab. Plant Biology Building, Michigan State University,  
East Lansing MI - 48824-1312, U.S.A.

**Palabras Clave:** Teoría endosimbiótica, genes emigrantes, contenido en G + C, uso de codones, sitios de metilación.

### ABSTRACT

J.L. Oliver y J.M. Martínez Zapater. 1988. Evolution of chloroplast migrant genes in the nucleus of wheat and pea. *An. Aula Dei* 19 (1-2): 115-124.

Strong differences in both base composition and codon usage exist between nuclear and chloroplast genes of two plant species: wheat and pea. The migrant genes—old chloroplast genes now integrated into the nucleus, but coding for plastid-specific proteins—seem to have completed their integration process, since they are now virtually indistinguishable from the remaining nuclear genes, having adopted the base composition and codon usage typical of nuclear DNA. The nucleotide changes to reach these features have taken place mainly at silent sites. The distribution of possible methylation sites (CG doublets and C(A/T)G trinucleotides) in chloroplast, migrant and nuclear genes suggests a differential incidence of distinct methylation mechanisms in the two species. A shortage of the doublet CG was observed in nuclear and migrant genes of pea, which can be explained by the existence of DNA methylation in the nucleus but not in the chloroplast. However, in migrant genes of wheat such a shortage was not present. The integration of these genes in G + C rich chromosome regions with a decreased discrimination against CG doublets could explain this result.

### INTRODUCCION

El origen de los cloroplastos es objeto aún de controversia, si bien desde que en 1970 Lynn Margulis propusiera su derivación de algún procariota simbiote, probablemente una cianobacteria, la teoría endosimbiótica ha recibido un fuerte apoyo experimental (Woese, 1977; Schwartz y Dayhoff, 1978; Weeden, 1981; Curtis y Clegg, 1984; Shih et al., 1986). Para explicar el pequeño tamaño del genoma de los plastos, esta teoría postula la transferencia de la mayor parte del genoma del endosimbiote al núcleo del huésped, aunque las proteínas codificadas por estos genes seguirían actuando en el cloroplasto; de esta manera, los genes nucleares que codifican proteínas cloroplásticas podrían ser considerados genes desplazados o “emigrantes”, puesto que provienen de la emigración de antiguos genes del primitivo procariota simbiote. Comparando

la composición de nucleótidos, el uso de codones y la distribución de los posibles sitios de metilación, y mediante distintos índices de similitud genética, analizamos aquí la evolución seguida por los genes cloroplásticos "emigrantes" tras integrarse en el genoma nuclear.

## MATERIAL Y METODOS

Para nuestro estudio hemos empleado los genomas nuclear y cloroplástico de una especie dicotiledónea –el guisante (*Pisum sativum*)– y de una monocotiledónea –el trigo (*Triticum aestivum*)–. Las secuencias de nucleótidos de los distintos genes (Tablas 1A y 1B) se obtuvieron de GenBank (Bilofsky et al., 1986; R. 48.0, Febrero, 1987). Tres genes nucleares de guisante (*cab15*, *cab80* y *rubp15*) y dos de trigo (*cab* y *rbca*), pueden identificarse en principio como genes "emigrantes", al codificar todos ellos para proteínas típicamente cloroplásticas.

La composición de nucleótidos y la distribución de posibles sitios de metilación se han determinado utilizando el paquete de programas "DNA/Protein Sequence Analysis" de James M. Pustell (Pustell y Kafatos, 1984).

Las distancias en el uso de codones entre los genes se han estimado mediante una métrica Manhattan:

$$D(A,B) = \sum |x(i,A) - x(i,B)|$$

donde  $x(i,A)$  y  $x(i,B)$  son las frecuencias relativas por grupo sinónimo del codón  $i$  en los genes A y B, respectivamente. Para ello se ha utilizado el programa "CLAS", escrito en Fortran 77 por el primer autor.

## RESULTADOS Y DISCUSION

### A. Medidas de similitud genética

La comparación de dos secuencias de nucleótidos homólogas no ofrece mayor dificultad, existiendo toda una gama de algoritmos que permiten optimizar el alineamiento entre ellas (Nei, 1987). Es más difícil, sin embargo, intentar comparar secuencias no-homólogas, como resulta necesario en el caso de los genes "emigrantes" si se los quiere comparar con los cloroplásticos o los demás genes nucleares. Como una primera aproximación, hemos utilizado cuatro índices que fueron propuestos por Grantham en 1978 para medir la similitud entre genes de origen muy diverso (virus, bacterias y eucariotas). El primero de ellos es la proporción de codones terminados en G, NCG/NMG, donde N es cualquier base y M es cualquier base excepto C; el segundo índice es la proporción de dinucleótidos CG/GC; el tercero es la proporción de codones para arginina, CGQ/AGX, donde CGQ es el cuarteto que codifica para este aminoácido y AGX el dueto; el cuarto índice es el  $\%(G+C)$  en la tercera posición del codón. Una vez calculados estos índices, se ordenan los genes de menor a mayor valor para cada índice y se halla la suma de los rangos para cada gen. Los genes de cada especie se ordenan de nuevo de menor a mayor de acuerdo con el valor de dicha suma. Según Grantham (1978), esto suministra una medida aproximada de la similitud entre los distintos genes considerados. En la Tabla 2A puede verse que casi todos los genes nucleares del guisante ocupan uno de los extremos de la distribución, mientras que los genes cloroplásticos ocupan el otro. Esta separación no ocurre, sin embargo, en el trigo, donde los genes nucleares aparecen entremezclados con los cloroplásticos. En cualquier caso, lo que es de destacar es que en ambas especies, al menos hasta donde alcanzan estos índices, los genes emigrantes no pueden distinguirse del resto de los genes nucleares.

Tabla 1A. Genes nucleares y cloroplásticos de guisante utilizados en este estudio. Las secuencias se extrajeron de GenBank.

Entrada	Gen	Total de Codones	Proteína
Núcleo:			
PEAABN1	abn1	130	Albumina 1, exones 1,2
PEAABN2	abn2	232	Albumina 2
PEALEGA	lega	495	Legumina, exones 1-4
PEALGN	lgn	185	Legumina
PEACAB15	cab15 (*)	229	Clorofila a/b (proteína de unión)
PEACAB80	cab80 (*)	270	AB80
PEARUBP15	rubp15 (*)	157	Ribulosa 1,5-bifosfato carboxilasa (subunidad pequeña)
Cloroplasto:			
PEACPATPG	atpg	247	ATP sintetasa (subunidades a y c)
PEACPCYF	cyf	343	Citocromo f
PEACPD2	psbD	354	PII proteína D2
PEACPD2	psbC	123	PII proteína 44kDa (centro de reacción)

(\*) Genes cloroplásticos emigrados al núcleo.

Tabla 1B.- Genes nucleares y cloroplásticos de trigo utilizados en este estudio. Las secuencias se extrajeron de GenBank.

Entrada	Gen	Total de Codones	Proteína
Núcleo:			
WHTGLGB	glgb	292	Gamma-gliadina B
WHTGLIA	glia	319	Alfa gliadina
WHTGLIAA	gliaa	287	Alfa-1Y gliadina
WHTGLIABA	gliaba	292	Alfa-/beta-gliadina A-II
WHTGLIABC	gliabc	320	Alfa-/beta-gliadina A-V
WHTGLIABG	gliabg	298	Alfa-/beta-gliadina A-IV
WHTGLIABH	gliabh	283	Alfa-/beta-gliadina A-III
WHTGLUMRA	glumra	102	Gluten
WHTH3	h3	137	Histona H3
WHTH4	h4	104	Histona H4
WHTCAB	cab (*)	267	Clorofila a/b (proteína de unión)
WHTRBCA	rbca (*)	139	Ribulosa bifosfato carboxilasa (sub. pequeña)
Cloroplasto:			
WHTPATP	atp	82	ATP sintetasa
WHTCPATPS	atps	183	ATP sintetasa (subunidad CF-0 I)
WHTCPCYF	cyf	321	Citocromo f
WHTCPCYTB	cytb	84	Citocromo b-559

(\*) Genes cloroplásticos emigrados al núcleo.

Tabla 2A. Distintos índices génicos en guisante. Los genes están ordenados de acuerdo con la suma de los rangos para cada índice.

Gen	Genoma (**)	NCG/NMG	CG/GC	CGQ/AGX	%(G+C) III	Rangos	Suma
abn1	NUC	1/14	7/16	1/3	33	5,3,1,4,5	13.5
pealgn	NUC	2/32	13/47	4/9	42	4,1,2,7	14
cab80(*)	NUC	0/39	20/44	4/4	48	1,4,5,5,10	20.5
lega	NUC	10/98	39/103	14/31	47	7,2,3,9	21
cab15(*)	NUC	0/28	18/36	3/2	45	2,5,8,8	23
rubp15(*)	NUC	1/28	11/19	3/5	49	3,6,4,11	24
cyf	CP	6/41	25/34	9/6	27	8,5,8,8,1.5	26
atpg	CP	13/13	19/25	2/2	27	11,10,5,5,1.5	28
psbC	CP	1/12	12/16	3/2	40	6,9,8,6	29
psbD	CP	4/27	34/50	14/1	32	8,5,7,11,3	29.5
abn2	NUC	4/17	22/25	7/3	33	10,11,10,4,5	35.5

(\*) Genes cloroplásticos emigrados al núcleo.

(\*\*) NUC = Nuclear; CP = Cloroplástico.

Tabla 2B. Distintos índices génicos en trigo. Los genes están ordenados de acuerdo con la suma de los rangos para cada índice.

Gen	Genoma (**)	NCG/NMG	CG/GC	CGQ/AGX	%(G+C) III	Rangos	Suma
glia	NUC	17/52	20/72	0/3	36	12,5,1,5,5,5	24
gliaa	NUC	12/56	18/68	2/3	41	7,2,5,5,5,9	24
gliabc	NUC	17/52	20/72	0/3	36	12,5,1,5,5,5	24
glumra	NUC	0/0	13/36	0/0	54	1,8,3,12	24
atp	CP	3/5	5/21	1/1	15	15,1,8,1	25
gliaba	NUC	14/59	19/72	2/3	42	9,2,5,5,5,10	27
glgb	NUC	5/49	17/51	2/2	48	3,7,8,11	29
cyf	CP	7/40	33/58	11/3	31	5,10,13,2	30
gliabg	NUC	15/56	21/76	3/3	40	10,5,8,7,5	30.5
atps	CP	2/30	27/24	11/4	32	2,15,11,3,5	31.5
gliabh	NUC	19/44	25/61	1/3	40	14,9,4,7,5	34.5
rbca (*)	NUC	4/38	32/40	6/2	86	4,11,5,12,14	41.5
cytb	CP	5/3	13/7	4/2	32	16,16,10,3,5	45.5
h3	NUC	9/39	48/56	14/3	98	8,13,14,16	51
h4	NUC	5,26	44/44	15/0	95	6,14,16,15	51
cab (*)	NUC	15/46	60/75	9/0	75	12,11,5,15,13	51.5

(\*) Genes cloroplásticos emigrados al núcleo.

(\*\*) NUC = Nuclear; CP = Cloroplástico.

Tabla 3A. Distancias de uso de codones ( $\pm$  S.E.) entre genes nucleares (NUC), emigrantes (MIG) y cloroplásticos (CP) en guisante.

NUC	$10.81 \pm 1.11$		
MIG	$12.72 \pm 0.60$	$9.54 \pm 1.56$	
CP	$13.86 \pm 0.55$	$17.20 \pm 0.41$	$12.08 \pm 1.09$
	NUC	MIG	CP

Tabla 3B. Distancias de uso de codones ( $\pm$  S.E.) entre genes nucleares (NUC), emigrantes (MIG) y cloroplásticos (CP) en trigo.

NUC	$13.57 \pm 1.13$		
MIG	$16.70 \pm 0.92$	9.75	
CP	$19.49 \pm 0.93$	$25.12 \pm 0.98$	$15.17 \pm 0.96$
	NUC	MIG	CP

Tabla 4. Índice de sesgo P2 promediado para genes cloroplásticos (CP), emigrantes (MIG), y nucleares (NUC) en guisante y trigo.

Genoma	P2
Guisante:	
CP	0.48
MIG	0.69
NUC	0.59
Trigo:	
CP	0.43
MIG	0.52
NUC	0.42

## B. Uso de codones

El uso de codones se tabuló para cada gen en ambas especies. Con objeto de poder cuantificar su grado de diferenciación, hemos calculado la frecuencia relativa de cada codón en cada gen. Para ello hemos agrupado los codones en grupos de codones sinónimos, considerando como tales los que difieren sólo en la tercera posición. Hay pues un grupo de codones para cada aminoácido, excepto para arginina, leucina y serina que tienen dos. No se consideraron los codones de terminación ni los grupos constituidos por un sólo codón (metionina y triptófano). De esta manera, calculamos las frecuencias de 59 codones distribuidos en 21 grupos. Dichas frecuencias se normalizaron, de manera que para cada grupo las frecuencias de los distintos codones sinónimos sumasen 1. La diferenciación global en el uso de codones entre dos genes cualesquiera puede medirse entonces utilizando el algoritmo de distancia descrito más arriba.

Las Tablas 3A y 3B muestran las distancias medias en el uso de codones entre los genes nucleares, emigrantes y cloroplásticos. Tanto en el trigo como en el guisante, las distancias entre los genes nucleares no son significativamente distintas de las que existen entre los genes cloroplásticos, lo que indica que la tasa de evolución en el uso de codones es similar en ambos genomas. Este resultado está de acuerdo con la observación de que las secuencias de ADN cloroplástico evolucionan a la misma velocidad que las del núcleo (Curtis y Clegg, 1984) y contrasta con la alta tasa de evolución (de 5 a 10 veces más rápida) que se observa en el ADN (Wolstenholme y Clary, 1985) y en el uso de codones (resultados sin publicar del primer autor) del otro orgánulo celular de origen simbiótico, la mitocondria.

En las dos especies consideradas en este estudio, el uso de codones de los genes emigrantes es virtualmente el mismo que el del resto de los genes del núcleo. Las Tablas 3A y 3B muestran claramente que el uso de codones de los genes emigrantes se parece más al del núcleo en el que están ahora integrados que al del cloroplasto del que derivan.

Otro aspecto que hemos analizado ha sido el sesgo en el uso de codones. Hemos utilizado para ello el índice P2, que es una medida de la elección entre pirimidinas en la tercera posición del codón (Gouy y Gautier, 1982). La Tabla 4 muestra los valores de este índice promediados para los genes nucleares, emigrantes y cloroplásticos. En ambas especies, puede observarse un sesgo considerablemente mayor en los genes emigrantes que en los genes cloroplásticos o nucleares. Se ha observado una estrecha correlación entre el valor de este índice y una mayor actividad del gen correspondiente, tanto en bacterias (Gouy y Gautier, 1982) como en levaduras (Sharp et al., 1986). Si éste fuera también el caso en las plantas, el mayor sesgo de los genes emigrantes podría indicar una expresión más activa de estos genes.

## C. Composición de nucleótidos en posiciones silenciosas y de remplazamiento

La Tabla 5 muestra las frecuencias de los distintos nucleótidos en las posiciones silenciosas y de remplazamiento del codón para los genes cloroplásticos, emigrantes y nucleares de cada especie. Los sitios silenciosos son aquellos en los que los cambios de nucleótidos en el gen no conducen a cambios en la secuencia de aminoácidos de la proteína correspondiente.

En las dos especies, el contenido en G+C es más alto en el núcleo que en el cloroplasto. Puesto que se ha demostrado la existencia de una correlación entre el contenido total (genómico) en G+C y el nivel de G+C en las tres posiciones del codón (Bernardi y Bernardi, 1986; Jukes y Bhushan, 1986), cabe esperar, y así se observa, que los genes emigrantes hayan sufrido un incremento en su nivel de G+C al integrarse en el núcleo. La Tabla 5 muestra además que casi todo este incremento se ha absorbido por las posiciones silenciosas del codón y en menor proporción por las de remplazamiento.

En los genes emigrantes del guisante, cuando se los compara con los cloroplásticos, puede observarse que el alto contenido en G+C de los sitios silenciosos se debe principalmente al incremento de citosina, mientras que los niveles bajos de A+T son debidos sobre todo a un des-

Tabla 5. Composición de nucleótidos y contenido en G+C en sitios de remplazamiento (RS) y silenciosos (SS) de genes cloroplásticos (CP), emigrantes (MIG) y nucleares (NUC) de guisante y trigo.

Genoma	G+C Total	RS					SS				
		A	T	C	G	G+C	A	T	C	G	G+C
Guisante:											
CP	.40	.25	.27	.20	.27	.47	.28	.44	.17	.11	.28
MIG	.49	.24	.24	.21	.31	.52	.24	.34	.28	.14	.42
NUC	.44	.32	.21	.20	.27	.47	.30	.31	.24	.16	.40
Trigo:											
CP	.40	.30	.24	.19	.27	.46	.31	.40	.16	.13	.29
MIG	.61	.24	.23	.22	.31	.53	.05	.17	.51	.26	.77
NUC	.50	.29	.19	.38	.14	.52	.36	.17	.24	.22	.46

Tabla 6.- Distribución de sitios de metilación en genes cloroplásticos (CP), emigrantes (MIG) y nucleares (NUC) de guisante y trigo. I, II y III son la primera, segunda y tercera posición del codón.

Genoma	CG/GC				CTG	CAG
	I-II	II-III	III-I	Total	(Obs/Esp)	(Obs/Esp)
Guisante:						
CP	.33	1.26	1.81	.72	.76	.96
MIG	.18	.03	2.53	.49	.89	1.05
NUC	.32	.35	.63	.42	1.16	1.54
Trigo:						
CP	.55	1.42	.69	.71	1.34	1.10
MIG	.38	.40	2.07	.80	.74	1.14
NUC	.31	.95	.25	.40	.99	1.94



censo de timina. Así pues, las transiciones parecen haber predominado en el proceso de integración de los genes emigrantes de esta especie. En el trigo, el incremento en G+C de los genes emigrantes se debe también sobre todo al incremento de citosina, pero los niveles bajos de A+T pueden achacarse tanto a descensos de adenina como de timina. De esta manera, en el proceso de integración de los genes emigrantes del trigo parecen haber ocurrido transiciones y transversiones con frecuencia similar.

#### D. Distribución de los sitios de metilación

El ADN de las plantas tiene un alto contenido en 5-metilcitosina, distribuida tanto en los dinucleótidos CG como en los trinucleótidos C(A/T)G (Gruenbaum et al., 1981). En las dos especies que estamos considerando aquí, los genes emigrantes han pasado de un genoma con menor contenido en G+C y ausencia de metilación (cloroplasto) a otro genoma con un contenido más alto en G+C y en el que existe metilación (el núcleo). Cabe esperar, por tanto, que estos genes incrementen su contenido en G+C, reduciendo al mismo tiempo las posibles dianas para las metilasas nucleares, con objeto de evitar las alteraciones que la metilación tendría sobre su expresión.

La Tabla 6 muestra la distribución de posibles sitios de metilación en los genes cloroplásticos, emigrantes y nucleares de ambas especies. En los genes emigrantes la mayoría de los dinucleótidos CG se sitúan en las posiciones intercodón, es decir, la C en la tercera posición de un codón y la G en la primera posición del codón siguiente.

Los genes emigrantes del guisante y el trigo parecen haber seguido una estrategia diferente a la hora de "adaptarse" a su nuevo ambiente genómico. El incremento en G+C de los genes emigrantes del trigo se ve acompañado por un incremento de las proporciones de CG y CAG, así como por un descenso de la proporción de CTG. En el guisante, sin embargo, el incremento en G+C de los genes emigrantes va seguido de un descenso de la proporción de CG y un incremento de la proporción de ambos trinucleótidos. Estos datos parecen reflejar una incidencia diferencial de distintos mecanismos de metilación en ambas especies.

En los genes nucleares y emigrantes del guisante se observa una reducción importante de la proporción de dinucleótidos CG, en comparación con los genes cloroplásticos; esto probablemente se debe a la existencia de metilación en el núcleo pero no en el cloroplasto. Los genes emigrantes de esta especie, aún habiendo incrementado su contenido en G+C, habrían reducido su proporción de posibles dianas para metilación. Sin embargo, en los genes emigrantes del trigo no se observa tal reducción. Esto podría deberse a que en esta especie las metilasas utilizaran poco esta diana, pero ello no está de acuerdo con el hecho de que el resto de los genes nucleares tengan una proporción tan baja de estos dinucleótidos (tan sólo el 40% de los que cabría esperar de acuerdo con su composición de nucleótidos). Otra manera de explicar ese resultado sería suponer que los genes emigrantes del trigo se han integrado en zonas cromosómicas particularmente ricas en G+C, y que presentan, por tanto, una menor discriminación en contra de los dinucleótidos CG. Bernardi et al. (1985) han demostrado, en los vertebrados y sus virus, que la reducción en CG va siendo menor a medida que aumenta el contenido genómico en G+C. Esto parece ocurrir también en el núcleo del trigo, donde la proporción CG/GC está altamente correlacionada ( $r = 0.94$ ) con el contenido total en G+C de los distintos genes. A modo de ejemplo, los genes de histonas *h3* y *h4*, muy ricos en G+C, muestran unas proporciones CG/GC de 48/56 y 44/44, respectivamente.

#### CONCLUSIONES

En ambas especies los genes cloroplásticos emigrados al núcleo han sufrido a lo largo de su proceso evolutivo una serie de cambios tendentes, por un lado, a equiparar su composición de

nucleótidos con el contenido en G+C más alto que se encuentra en el núcleo y, por otro, a evitar las alteraciones que para la expresión génica podría tener la acción de las metilasas nucleares. Los cambios necesarios para conseguir el primer objetivo han sido absorbidos en su mayor parte por las posiciones silenciosas del codón. Los caminos seguidos para alcanzar el segundo no han sido los mismos en el guisante y en el trigo, probablemente debido a la distinta importancia que los diferentes mecanismos de metilación tienen en ambas especies. En cualquier caso, el proceso de integración de los genes emigrantes en el núcleo celular parece haberse completado en las dos especies, ya que los cambios experimentados los han hecho ya virtualmente indistinguibles, en cuanto a su composición de nucleótidos y a su uso de codones, del resto de los genes nucleares.

Este trabajo ha sido financiado en parte por la CAICYT (Pr. 84-1211 CO2-02).

### RESUMEN

Los genomas nuclear y cloroplástico de trigo y guisante están claramente diferenciados en cuanto a su composición de nucleótidos y uso de codones. En ambas especies, los genes emigrantes—antiguos genes cloroplásticos ahora integrados en el núcleo celular y que codifican para proteínas del cloroplasto—parecen haber completado su proceso de integración, ya que son virtualmente indistinguibles del resto de los genes nucleares, habiendo adoptado la composición de nucleótidos y el uso de codones típicos de este genoma. Los cambios necesarios para ello han ocurrido principalmente en las posiciones silenciosas del codón. La distribución de posibles sitios de metilación—dinucleótidos CG y trinucleótidos C(A/T)G—en los genes cloroplásticos, emigrantes y nucleares es distinta en ambas especies, lo que podría deberse a una incidencia diferencial de los distintos mecanismos de metilación. En los genes nucleares y emigrantes del guisante, se observa una reducción significativa de la frecuencia de dinucleótidos CG, lo que presumiblemente se debe a la existencia de metilación en el núcleo pero no en el cloroplasto. Dicha reducción no se observa, sin embargo, en los genes emigrantes del trigo; la integración de estos genes en zonas cromosómicas particularmente ricas en G+C, y, por tanto, con una menor discriminación en contra de los dinucleótidos CG, podría explicar este resultado.

### REFERENCIAS

- Bernardi, G.; G. Bernardi. (1986). Compositional constraints and genome evolution. **J. Mol. Evol.** 24: 1-11.
- Bernardi, G.; B. Olofsson; J. Filipiski; M. Zerial; J. Salinas; G. Cuny; M. Meunier-Rotival; F. Rodier. (1985). The mosaic genome of warm-blooded vertebrates. **Science** 228: 953-958.
- Bilofsky, H.S.; C. Burks; J.W. Fickett; W.B. Goad; F.L. Lewitter; W.P. Rindone; C.D. Swindell; C.S. Tung. (1986). The GenBank genetic sequence data bank. **Nucleic Acids Res.** 14: 1-4.
- Curtis, S.E.; M.T. Clegg. (1984). Molecular evolution of chloroplast DNA sequence. **Mol. Biol. Evol.** 1: 291-301.
- Gouy, M.; C. Gautier. (1982). Codon usage in bacteria: correlation with gene expressivity. **Nucleic Acids Res.** 10: 7055-7074.
- Grantham, R. (1978). Viral, prokaryote and eukaryote genes contrasted by mRNA sequence indexes. **FEBS Letters** 95(1): 1-11.
- Gruenbaum, Y.; T. Naveh-Manly; H. Cedar; A. Razin. (1981). Sequence specificity of methylation in higher plant DNA. **Nature.** 292: 860-862.
- Jukes, T.H.; V. Bhushan. (1986). Silent nucleotide substitutions and G+C content of some mitochondrial and bacterial genes. **J. Mol. Evol.** 24: 39-44.

- Margulis, L. (1970). Origin of eukaryotic cells. Yale University Press, New Haven.
- Nei, M. (1987). Molecular Evolutionary Genetics. Columbia University Press, New York.
- Pustell, J.M.; F.C. Kafatos. (1984). A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. **Nucleic Acids Res.** 12: 643-655.
- Schwartz, R.M.; M.O. Dayhoff. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. **Science** 199: 395-403.
- Sharp, P.M.; T.M.F. Tuohy; K.R. Mosurski. (1986). Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. **Nucleic Acids Res.** 14: 5125-5143.
- Shih, M.C.; G. Lazar; H.M. Goodman. (1986). Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. **Cell** 47: 73-80.
- Weeden, N. (1981). Genetic and biochemical implications of the endosymbiotic origin of the chloroplast. **J. Mol. Evol.** 17: 133-139.
- Woese, C.R. (1977). **J. Mol. Evol.** 10: 93-96.
- Wolstenholme, D.R.; D.O. Clary. (1985). Sequence evolution of *Drosophila* mitochondrial DNA. **Genetics** 109: 725-744.