

ANÁLISIS ESTRUCTURAL DE PROTEÍNAS DE *Methanopyrus kandleri* AV19, MICROORGANISMO RELACIONADO CON LUCA.

ANGIE MILENA BARRERA CORREA

ASESOR EXTERNO: SILVIO ALEJANDRO LOPEZ PAZOS MSc. PhD.
Facultad de Ciencias, Universidad Antonio Nariño

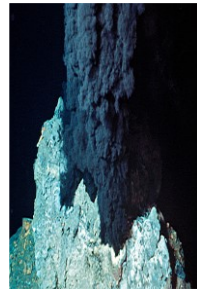
ASESOR INTERNO: SANDRA MONICA ESTUPIÑAN M. Sc.
Facultad de Ciencias de la Salud, Universidad Colegio Mayor
de Cundinamarca

UNIVERSIDAD COLEGIO MAYOR DE CUNDINAMARCA
FACULTAD DE CIENCIAS DE LA SALUD
PROGRAMA DE BACTERIOLOGIA Y LABORATORIO CLINICO
TRABAJO DE GRADO
Bogotá D.C., 2018

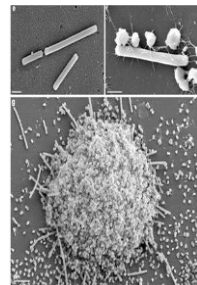
INTRODUCCIÓN



ME: condiciones severas como aguas hidrotermales, lagos alcalinos, salinas, desiertos, nevados y glaciares, plantas de energía nuclear y espacios con metales pesados



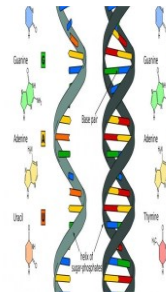
Temperatura de origen : fuentes hidrotermales de profundidades marinas.



M kandleri, es un microorganismo cercano a LUCA ya que comparte características genéticas y metabólicas, la hipertermofilia y metanogenesis.



Inicio de la vida: metabolismo a partir del código genético, o formación de moléculas informativas replicantes.



Código genético en todos los organismos indica que todos descienden de un ancestro común.

Phylogenetic Tree of Life



Bioquímica que llevo al origen de la primera forma de vida y las relaciones evolutivas de los tres dominios Bacteria, Archaea y Eukarya.

<https://cienciaybiologia.com/arqueas/>

<https://www.who.edu/main/topic/hydrothermal-vents>

[https://microbewiki.kenyon.edu/index.php/Methanopyrus_kandleri: evolutionary significance](https://microbewiki.kenyon.edu/index.php/Methanopyrus_kandleri:_evolutionary_significance)

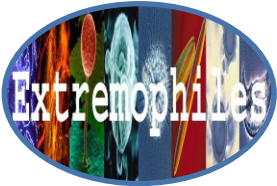
<http://biogeo.esy.es/BG4ESO/celulaunidad.htm>

<http://calealactee.planetapamant.com/2018/01/05/definitia-structurilor-de-adn-ce-sunt-si-la-ce-ne-ajuta/>

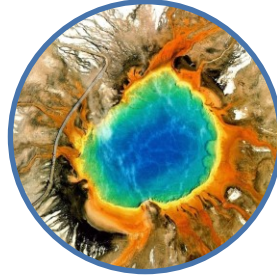
http://www.wikiwand.com/es/Sistema_de_tres_dominios

MARCO DE REFERENCIA

EXTREMOFILOS



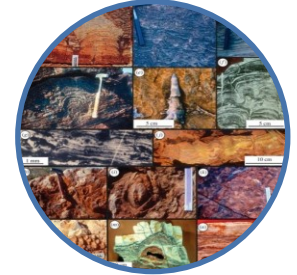
Los ME sobreviven a condiciones ambientales geofísicas y geoquímicas.



Termófilos ,
hipertermófilos,
psicrófilos , barófilos,
acidófilos, alcalófilos.



Halófilos ,
metanógenos,
metalófilos y
radiófilos .



Microfósiles indican que MO del periodo arqueano, desarrollaron fuentes metabólicas similares a MO vivos actuales.



ORIGEN DE LA VIDA

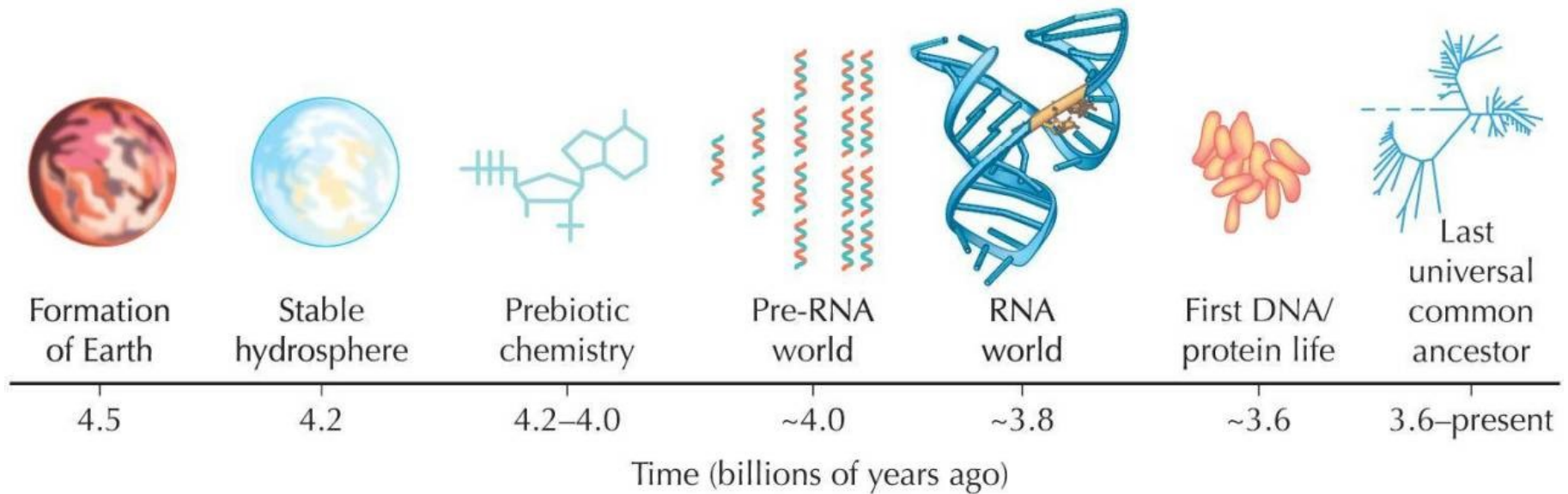
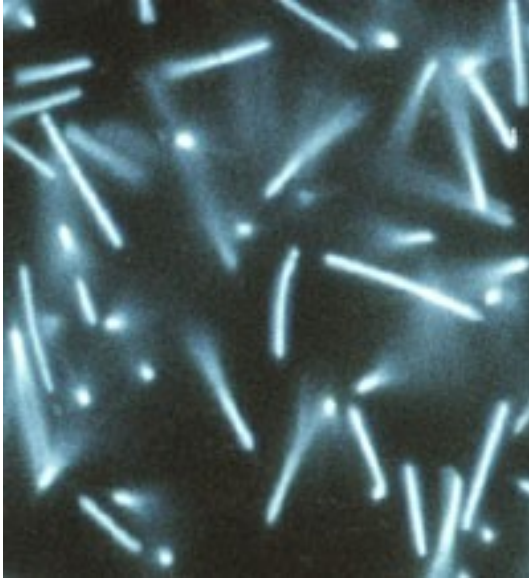


FIGURE 4.4. Steps in the origin of life.

4.4, modified from Joyce G.F., *Nature* **418**: 214–221, © 2002 Macmillan, www.nature.com

Evolution © 2007 Cold Spring Harbor Laboratory Press

Methanopyrus kandleri



Dominio: Archaea
Phylum: Euryarchaeota
Clase: Methanopyri
Orden: Methanopyrales
Familia: Methanopyraceae
Género: Methanopyrus
Especie: kandleri

Bacilo Gram positivo, hipertermofilo, halofilo moderado (0.2 y 4%, con una salinidad óptima de 2%) , metanogeno.

Aislado en una chimenea volcánica de 2,000 m de profundidad en la Cuenca Guaymas en el Golfo de California.

Puede soportar temperaturas de 130°C.

Los linajes vivos más primitivos se encuentran en la cordillera central del Océano Índico.

Methanopyrus kandleri

Los genes que codifican proteínas huérfanas se producen en dos grandes regiones. Estos genes no tienen homólogos conocidos excepto de otros genes de *M. kandleri*.

Se identificó un nuevo lípido de éter 2,3-di-O-geranilgeranil-sn-glicerol y un geranylgeraniol. Este lípido terpenoide, es una característica primitiva.

Su histona, con dos motivos de pliegues repetidos en un polipéptido indican que la N-terminal está relacionada con las histonas eucarióticas, mientras que la C-terminal se asemeja a procariontas.

Girasa inversa: única topoisomerasa heterodimérica tipo I. Formada por una subunidad RgyA (50 kDa) que interactúa covalentemente con el ADN y una subunidad RgyB (150 kDa) involucrada en la hidrólisis de ATP.

Cromosoma circular, 1.694.969 pares de bases, 1.691 genes codificadores de proteínas y 39 genes de ARNs estructurales, 73% de los productos de los genes son proteínas conservadas que pertenecen a los COG.

OBJETIVOS

General

- Estudiar la estructura del proteoma de *M. kandleri AV19* in silico en relación al ancestro común universal.

Específicos

- Clasificar las proteínas codificadas por el genoma de *M. kandleri AV19*.
- Establecer características funcionales de residuos en las proteínas de *M. kandleri AV19* importantes para el ancestro común universal.
- Comparar la estructura secundaria del proteoma de *M. kandleri AV19* con proteínas de referencia.

DISEÑO METODOLÓGICO

Hipótesis: La estructura bioquímica de las proteínas de *M. kandleri* está relacionada con la función.

Población: proteínas de *M. kandleri*.

Muestra: 597 proteínas del genoma de *M. kandleri* AV19 curadas en la base de datos de SwissProt.

Variables:

Independientes: Función de las proteínas de *M. kandleri* AV19.

Dependientes: Estructura de las proteínas de *M. kandleri* AV19.

Indicadores:

Estructurales

Alfa-hélices

Beta-laminas

Proteínas transmembrana con el programa

Parámetros bioquímicos

Tipo de aminoácido

Número total de aminoácidos

Peso molecular

Punto isoeléctrico

Residuos cargados negativamente (Asp + Glu),

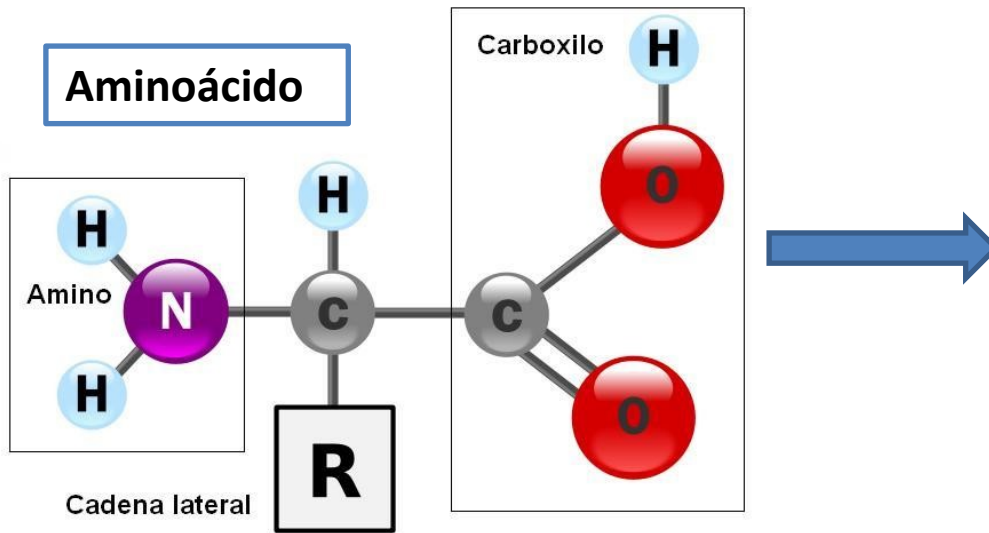
Residuos cargados positivamente (Arg + Lys),

Índice de inestabilidad

Índice alifático

Media de hidropatídicidad (GRAVY)

Parámetros bioquímicos y estructurales



20 tipos de aminoácidos, con diferente cadena lateral unida al carbono alfa.

La importancia es la formación de proteínas necesarias para cumplir las funciones necesarias en una célula.

Peso molecular (MW)

Suma de masas isotópicas medias de los aa y la masa isotópica media de una molécula de agua. Dalton (Da)

Punto isoeléctrico, pI

pH en el que una molécula tiene igual número de cargas positivas y negativas siendo eléctricamente neutra.

Índice de inestabilidad

Estabilidad de la proteína en un tubo de ensayo. <40: estable; >40: inestable

Aa Cargados positivamente

- A pH 7.0 son Arg + Lys, que contienen uno o más -NH₂ en el radical.

Aa cargados negativamente

- Asp + Glu son ácidos y tienen grupos -R por la presencia del grupo -COOH en el radical.

Índice alifático

- Volumen relativo ocupado por las cadenas laterales alifáticas (Ala, Val, Ile y Leu).

gran media de hidropatía (GRAVY)

- suma de los valores de hidropatía de todos los aa / número de estos en la secuencia

Estructura primaria



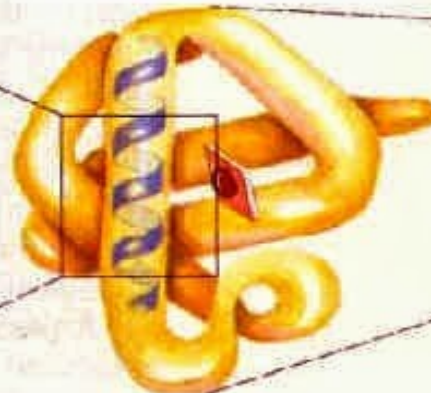
Aminoácidos

Estructura Secundaria



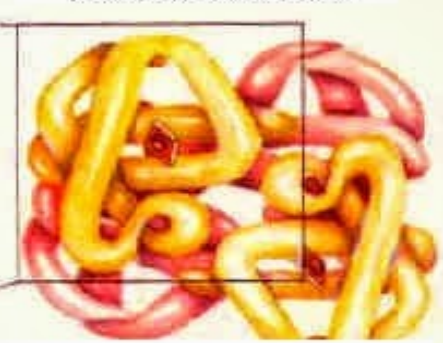
Hélice α

Estructura terciaria



Cadena polipeptídica

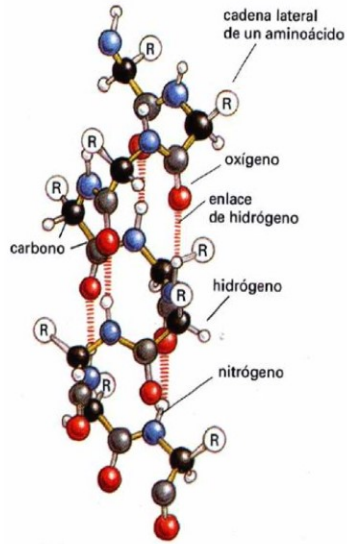
Estructura cuaternaria



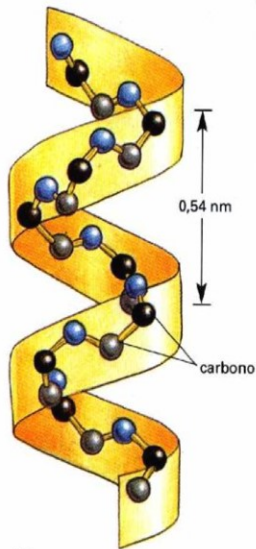
Subunidades ensambladas

Transporte y recepción

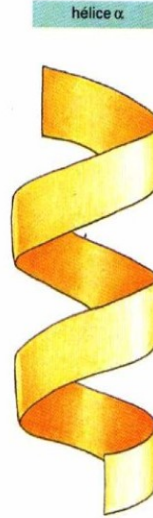
α -Hélice



(A)



(B)



(C)

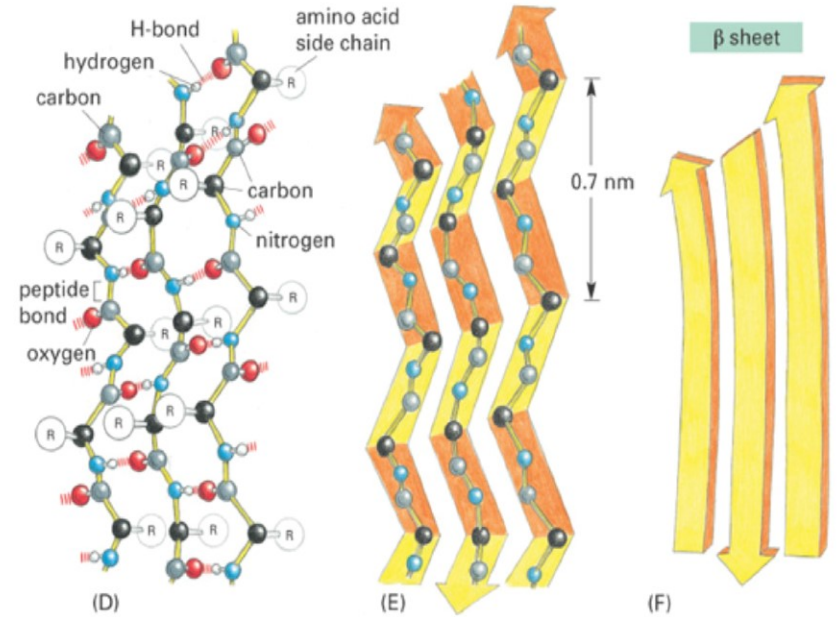
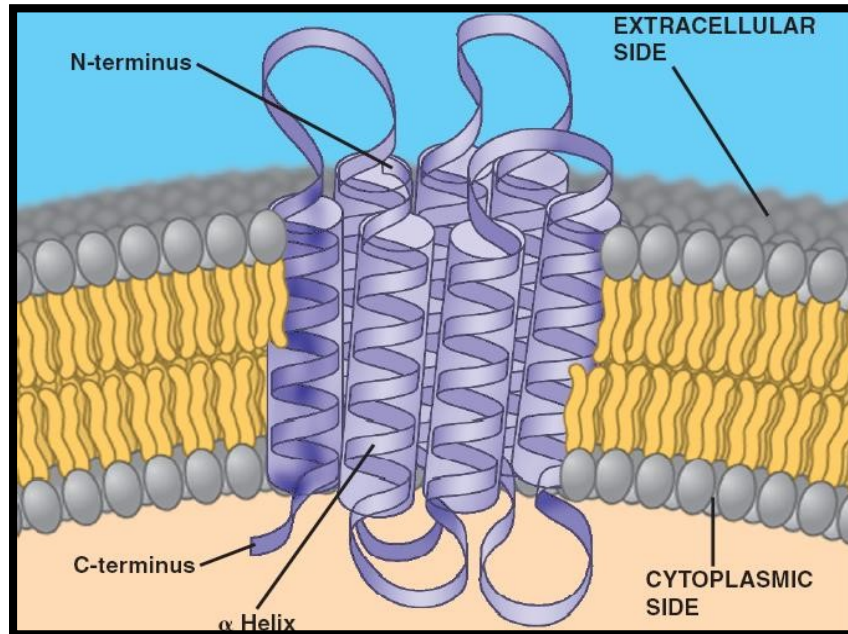


Figure 4-10 part 2 of 2 Essential Cell Biology, 2/e. (© 2004 Garland Science)



Hélices transmembrana

Procesos bioquímicos en la bicapa lipídica

Técnicas y procedimientos

Clasificación funcional

561 genes de LUCA, se clasificaron de acuerdo a las funciones de cada proteína codificada por estos (597 proteínas)

Basados en los COGs de los 1614 genes de *M. kandleri*, se obtuvo la secuencia primaria de cada una de las 597 proteínas usando la base de datos curada SwissProt/UniProt

- (A) Traducción, estructura ribosomal y biogénesis.
- (B) Modificación y procesamiento de ARN.
- (C) Transcripción
- (D) Replicación, recombinación y reparación.
- (E) Estructura y dinámica de cromatina
- (F) Control del ciclo celular, división celular, y partición cromosómica
- (G) Mecanismos de transducción de señales.
- (H) Pared celular, membrana y biogénesis.
- (I) Tráfico intracelular, secreción y transporte vesicular
- (J) Modificación postraduccional, recambio proteico, chaperones
- (K) Producción y conversión de energía.
- (L) Transporte y metabolismo de carbohidratos
- (M) Transporte y metabolismo de aminoácidos
- (N) Transporte y metabolismo de nucleótidos
- (O) Transporte y metabolismo de coenzimas
- (P) Transporte y metabolismo de lípidos
- (Q) Transporte y metabolismo de iones inorgánicos.
- (R) Biosíntesis, transporte y catabolismo de metabolitos secundarios.
- (S) Predicción de funciones generales

Análisis estructural y bioquímico

Ddatos de lparámetros bioquímicos: programa ProtParam de la suite ExPasy (<https://www.expasy.org/>).

Modelamiento proteico

La estructura secundaria de las 597 proteínas, y la estructura terciaria de 546 secuencias de aminoácidos: Phyre2.

Los modelos con cobertura <70% y de confianza <90%, se enviaron al programa I-Tasser.

Análisis estadístico

* Shapiro Test umbral de $p=0.01$.

* ANOVA de una vía, umbral de $p=0.01$.

* Kruskal Wallis, umbral de $p=0.01$.

* Prueba de Dunn con un umbral de $p=0.01$.

RESULTADOS

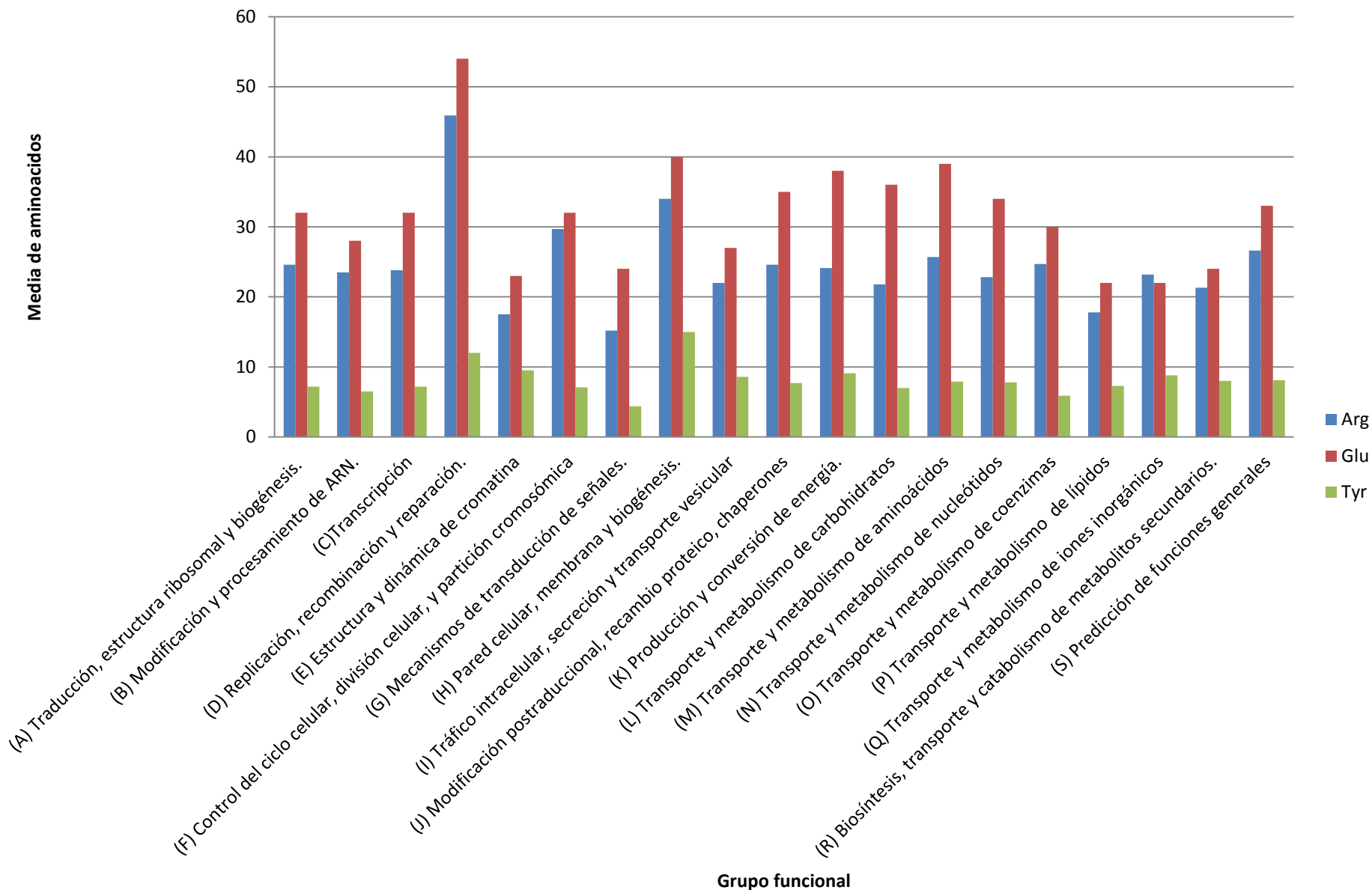
UniProt/SwissProt : 1719 entradas, 397 revisadas (Swissprot), y 1322 no curadas.

- 561 genes probables de LUCA, se les asignó una estructura secundaria basada en los COGs de 1614 genes de *M. kandleri* AV19
- Proteoma anotado teóricamente de LUCA de 597 secuencias de aminoácidos para este trabajo (182997 aminoácidos).

Tabla 1. Media de aminoácidos en cada grupo funcional. Aln: Alanina, Arg: Arginina, Asn: Asparagina, Asp: Aspartato, Cys: Cisteína, Gln: Glutamina, Glu, Ácido glutámico, Gly: Glicina, His: Histidina, Ile: Isoleucina, Leu: Leucina, Lys: Lisina, Met: Metionina, Phe: Fenilalanina, Pro: Prolina, Ser: Serina, Thr: Treonina, Trp: Triptófano, Tyr: Tirosina, Val: Valina. Azul: Grupo de datos no paramétricos.

Numero de secuencias	Nombre del grupo funcional	Aln	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
128	(A) Traducción, estructura ribosomal y biogénesis.	19	24,6	4,9	16	3,1	4,4	32	20	5,5	13,4	22,3	14,7	5,23	7,7	15	9	11	3,3	7,2	25,1
2	(B) Modificación y procesamiento de ARN.	22	23,5	4,5	13	2,5	5	28	29	7	13,5	28,5	11	4,5	6	17	12	12	2	6,5	25,5
31	(C) Transcripción	16	23,8	4,8	14	4	3,6	32	16	5,1	13,1	22,7	14,8	5	5,6	12	10	9,7	1,8	7,2	21,1
36	(D) Replicación, recombinación y reparación.	34	45,9	8,5	29	4,3	7,4	54	32	8,4	20,4	42,4	24,2	7,36	13	22	19	17	4,1	12	44,6
2	(E) Estructura y dinámica de cromatina	23	17,5	3	17	1,5	4	23	21	7	9,5	18,5	4,5	5	5	13	11	9,5	1,5	9,5	20
14	(F) Control del ciclo celular, división celular, y partición cromosómica	24	29,7	6,3	19	4,9	3,7	32	23	4,4	15,1	29,9	13,6	5,93	8,9	17	13	13	2,7	7,1	28,5
5	(G) Mecanismos de transducción de señales.	14	15,2	2,8	12	1,2	1,6	24	13	4,8	11	20,4	7,6	4	4,2	10	9	6,8	1,2	4,4	19,2
8	(H) Pared celular, membrana y biogénesis.	36	34	10	20	5,3	6,6	40	31	9,5	23,6	38,6	16,6	8,63	13	21	18	19	5,3	15	44,8
12	(I) Tráfico intracelular, secreción y transporte vesicular	28	22	6,8	15	1,5	5,7	27	26	3,3	24,5	38,2	17	10	9,6	15	15	16	2,8	8,6	31
28	(J) Modificación postraduccional, recambio proteico, chaperones	28	24,6	5,8	20	5,2	7,1	35	25	5,4	16,8	29,7	15,1	6,82	7,7	16	12	14	2,5	7,7	30,9
92	(K) Producción y conversión de energía.	31	24,1	7,2	21	8,2	5,4	38	29	6,6	19,5	30,6	14,3	7,56	9,4	19	13	16	3,1	9,1	34,6
22	(L) Transporte y metabolismo de carbohidratos	26	21,8	6,5	22	3,3	3,5	36	28	7	16	28,5	11,9	5,45	8,4	16	14	15	2	7	34,6
31	(M) Transporte y metabolismo de aminoácidos	32	25,7	7,8	22	4,2	4,5	39	29	7,9	17	30,2	11,9	7,16	9,5	17	15	15	2,5	7,9	40,5
36	(N) Transporte y metabolismo de nucleótidos	24	22,8	6,1	20	3,8	3,7	34	24	6,4	15,4	26,6	11,3	5,61	7,8	15	10	12	1,7	7,8	30,9
31	(O) Transporte y metabolismo de coenzimas	27	24,7	4,5	18	4,4	3,1	30	25	4,8	12,4	28,2	9,84	6,03	8,4	16	12	12	2,9	5,9	33,7
12	(P) Transporte y metabolismo de lípidos	28	17,8	4,8	12	2,6	3,1	22	25	5,6	13,3	25,1	10,7	5,08	9,3	11	12	14	4,2	7,3	30,5
18	(Q) Transporte y metabolismo de iones inorgánicos	37	23,2	6,4	15	3,3	4,3	22	28	6,3	17,3	43,2	10,8	6,39	11	15	18	18	3,8	8,8	39,6
6	(R) Biosíntesis, transporte y catabolismo de metabolitos secundarios.	16	21,3	5,2	17	3,2	1,7	24	16	5,7	10,8	25	7,17	4,67	5,7	12	10	8,7	2	8	25,8
83	(S) Predicción de funciones generales	23	26,6	5,2	18	4	4	33	24	7	14,6	28,6	11,9	5,67	8,6	17	13	14	2,7	8,1	32,3
597	Valor P	0,398	0,000011	0,32	0,337	0,495	0,30100	0,000002	0,331	0,24	0,306	0,374	0,362	0,281	0,301	0,27	0,42	0,3	0,3	0,000293	0,4

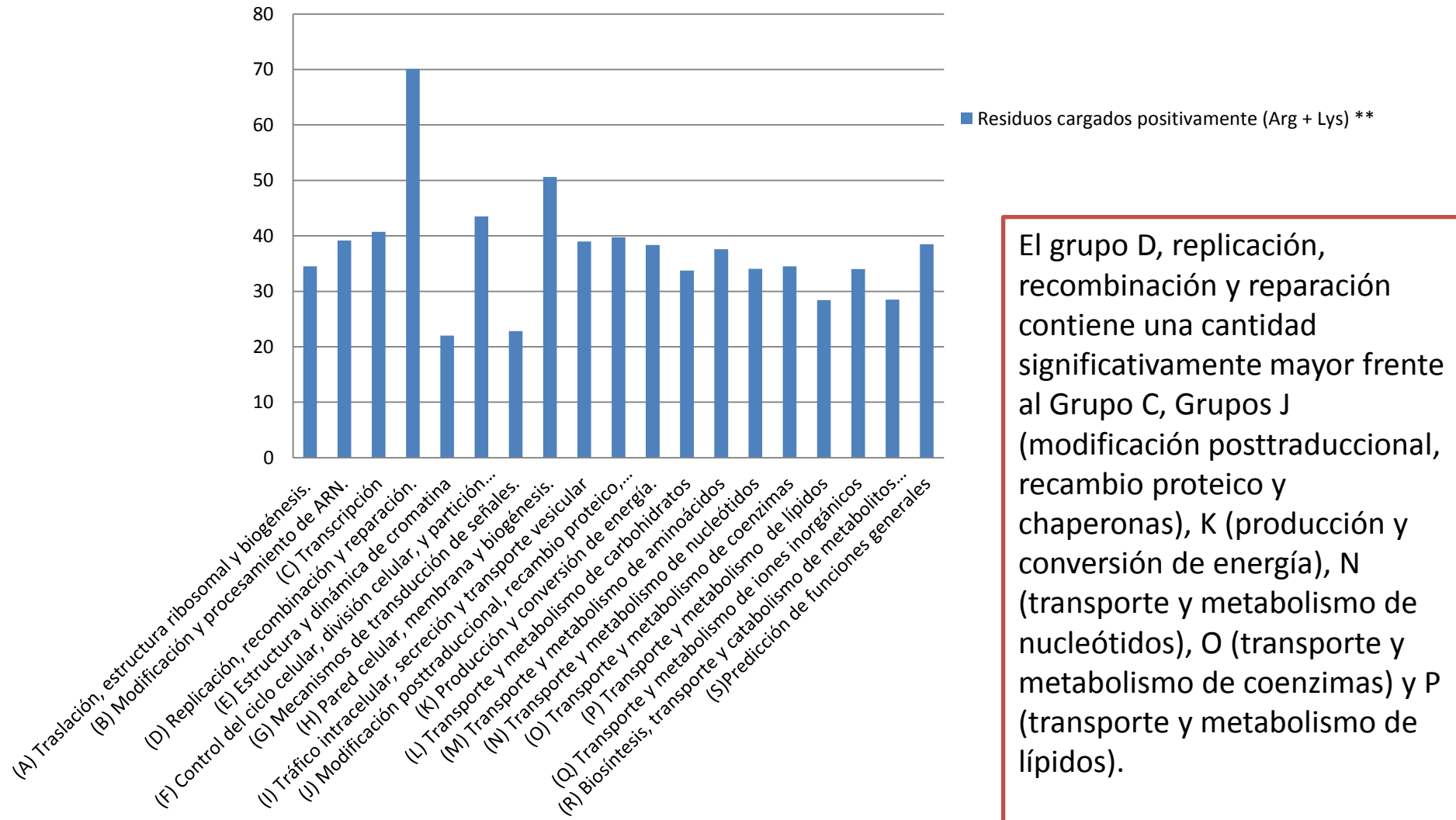
Media de aminoacidos por grupo funcional



**Tabla 4.
Parámetros
bioquímicos.**

Nombre Grupo funcional	Aminoácidos	Peso molecular	Punto isoelectrico	Residuo cargados negativos	Residuos cargados positivos (Arg + Lys) **	Índice de inestabilidad	Índice alifático.	Índice de hidropaticidad ad GRAY
(A) Traslación, estructura ribosomal y biogénesis.	262,8	29810	7,1050	40,5000	34,5000	41,3000	96,2850	-0,2325
(B) Modificación y procesamiento de ARN.	272	29817	6,9342	48,2344	39,1719	46,5202	84,8095	-0,5371
(C) Transcripción	242,14	27633	6,0579	45,8276	40,7241	46,7890	93,4503	-0,4499
(D) Replicación, recombinación y reparación.	454,4	50896	6,3272	82,9722	70,1111	43,9692	91,0897	-0,4238
(E) Estructura y dinámica de cromatina	221,5	24380	6,8500	39,0000	22,0000	39,1800	92,0650	-0,2750
(F) Control del ciclo celular, división celular, y partición cromosómica	301,21	33743	6,5529	50,3571	43,5000	45,2657	93,6579	-0,2959
(G) Mecanismos de transducción de señales.	186,2	20862	4,8860	36,2000	22,8000	46,1180	102,5440	-0,2010
(H) Pared celular, membrana y biogénesis.	413	46166	6,4700	59,3750	50,6250	39,7775	99,1488	-0,0879
(I) Tráfico intracelular, secreción y transporte vesicular	321,92	35506	7,8292	41,1667	39,0000	38,8017	114,1175	0,1569
(J) Modificación posttraduccional, recambio proteico, chaperones	314,86	34934	5,1582	55,5714	39,7500	41,4582	95,4400	-0,2277
(K) Producción y conversión de energía.	347,58	38274	5,2395	58,9474	38,3579	41,4308	92,5461	-0,1624
(L) Transporte y metabolismo de carbohidratos	312,59	34273	4,9595	58,6818	33,7273	36,7391	96,2786	-0,1846
(M) Transporte y metabolismo de aminoácidos	347,03	38060	4,7748	61,4516	37,6129	37,9065	96,4926	-0,1445
(N) Transporte y metabolismo de nucleótidos	288,61	32024	4,9336	53,4167	34,0556	37,2861	97,1533	-0,2251
(O) Transporte y metabolismo de coenzimas	288,23	31618	5,6797	47,1613	34,5161	39,1026	98,2484	-0,0741
(P) Transporte y metabolismo de lípidos	263,25	28684	7,3050	34,3333	28,4167	32,0242	106,5375	0,2093
(Q) Transporte y metabolismo de iones inorgánicos	336,44	36413	7,5439	37,0556	34,0000	31,8717	114,5422	0,3266
(R) Biosíntesis, transporte y catabolismo de metabolitos secundarios.	228,17	25775	5,6767	40,3333	28,5000	43,4667	100,8550	-0,1917
(S) Predicción de funciones generales	301,37	33510	5,7880	50,6747	38,4819	42,0047	95,7841	-0,2262
P valué <0,01	0,02	0,336	0,1175	0,36	0,000014	0,35	0,42	0,52

Residuos cargados positivamente (Arg + Lys) **



El grupo D, replicación, recombinación y reparación contiene una cantidad significativamente mayor frente al Grupo C, Grupos J (modificación posttraduccional, recambio proteico y chaperonas), K (producción y conversión de energía), N (transporte y metabolismo de nucleótidos), O (transporte y metabolismo de coenzimas) y P (transporte y metabolismo de lípidos).

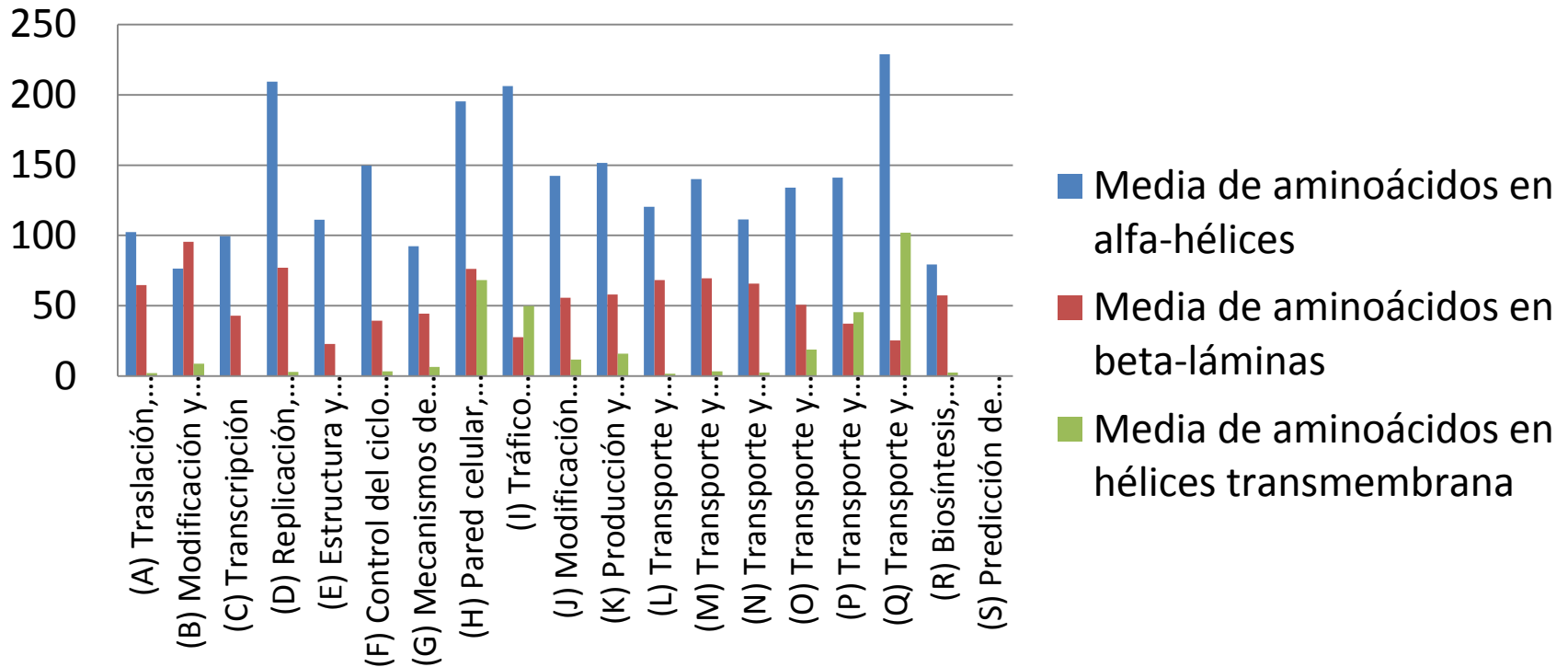
Grupo funcional	Media de aminoácidos en alfa-hélices	Media de aminoácidos en beta-láminas	Media de aminoácidos en hélices transmembrana
(A) Traslación, estructura ribosomal y biogénesis	102,29	64,71	2,12
(B) Modificación y procesamiento de ARN	76,48	95,52	8,8
(C) Transcripción	99,54	42,86	0
(D) Replicación, recombinación y reparación	209,33	76,99	2,9
(E) Estructura y dinámica de cromatina	111,23	22,88	0
(F) Control del ciclo celular, división celular, y partición cromosómica	149,69	39,32	3,4
(G) Mecanismos de transducción de señales	92,31	44,34	6,44
(H) Pared celular, membrana y biogénesis	195,29	76,1	68,14
(I) Tráfico intracelular, secreción y transporte vesicular	206,15	27,61	49,87
(J) Modificación postraduccional, recambio proteico, chaperones	142,44	55,61	11,8
(K) Producción y conversión de energía.	151,53	57,94	15,8
(L) Transporte y metabolismo de carbohidratos	120,47	68,14	1,69
(M) Transporte y metabolismo de aminoácidos	139,97	69,59	3,23
(N) Transporte y metabolismo de nucleótidos	111,44	65,69	2,58
(O) Transporte y metabolismo de coenzimas	133,99	50,73	18,88
(P) Transporte y metabolismo de lípidos	141,06	37,22	45,48
(Q) Transporte y metabolismo de iones inorgánicos	228,83	25,36	101,96
(R) Biosíntesis, transporte y catabolismo de metabolitos secundarios.	79,44	57,41	2,58
(S) Predicción de funciones generales	123.1	67.16	7.94
F de tabla=1.607	F=0.4291	F=0.3179	p=3.718478e-13*

Tabla 6. Composición estructural por grupos funcionales. El F calculado es menor que el F de tabla para todos los grupos en el ANOVA. El asterisco indica el valor p significativo en el análisis de Kruskal-Wallis ($p < 0.01$)

Tabla 7. Prueba de Dunn correspondiente a hélices transmembrana. En rojo se indican los resultados que corresponde a diferencias significativas ($p < 0.01$).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
B	1.0																		
C	1.0	1.0																	
D	1.0	1.0	1.0																
E	1.0	1.0	1.0	1.0															
F	1.0	1.0	1.0	1.0	1.0														
G	1.0	1.0	1.0	1.0	1.0	1.0													
H	3.418069e-10	1.0	3.481030e-10	7.807620e-08	0.063063	0.000034	0.15901												
I	2.018060e-01	1.0	6.977313e-02	1.0	1.0	1.0	1.0	6.819631e-02											
J	1.0	1.0	1.0	1.0	1.0	1.0	1.0	9.757609e-06	1.0										
K	1.0	1.0	8.211705e-01	1.0	1.0	1.0	1.0	2.929840e-07	1.0	1.0									
L	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.014656e-08	0.446412	1.0	1.0								
M	1.0	1.0	1.0	1.0	1.0	1.0	1.0	2.462116e-07	1.0	1.0	1.0	1.0							
N	1.0	1.0	1.0	1.0	1.0	1.0	1.0	6.956114e-08	1.0	1.0	1.0	1.0	1.0						
O	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.325054e-05	1.0	1.0	1.0	1.0	1.0	1.0					
P	1.522562e-01	1.0	5.361589e-02	1.0	1.0	1.0	1.0	8.284749e-02	1.0	1.0	1.0	0.356434	1.0	1.0	1.0				
Q	7.576780e-05	1.0	7.576780e-05	1.228030e-02	1.0	0.674545	1.0	3.279545e-01	1.0	0.5187	0.052873	0.002828	0.030813	0.011098	0.695458	1.0			
R	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.275572e-03	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
S	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.726329e-08	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.003816	1.0

Composición estructural por grupos funcionales



Se encontró que los grupos funcionales que tienen mayor contenido de aminoácidos asociados a hélices transmembrana son el Grupo H (pared celular, membrana y biogénesis), el Grupo I (tráfico intracelular, secreción y transporte vesicular), y el Grupo Q (transporte y metabolismo de iones inorgánicos).

Tabla 8. Estructura 3D de la proteína de mayor número de aminoácidos por función.



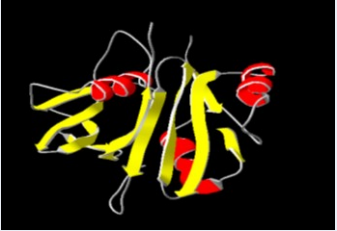
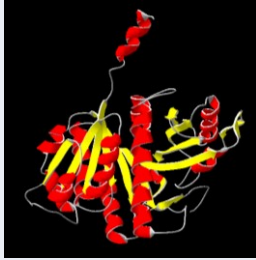



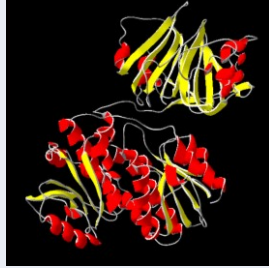
<p>(A) Traslación, estructura ribosomal y biogénesis.</p>		<p>(E) Estructura y dinámica de cromatina</p>	
	<p>(B) Modificación y procesamiento de ARN.</p>		<p>(F) Control del ciclo celular, división celular, y partición cromosómica</p>
<p>(C) Transcripción</p>		<p>(G) Mecanismos de transducción de señales.</p>	
	<p>(D) Replicación, recombinación y reparación.</p>		<p>(H) Pared celular, membrana y biogénesis.</p>

Tabla 8. Estructura 3D de la proteína de mayor número de aminoácidos por función.




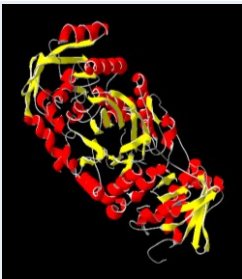


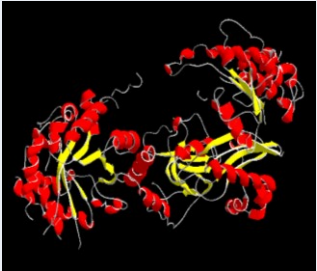

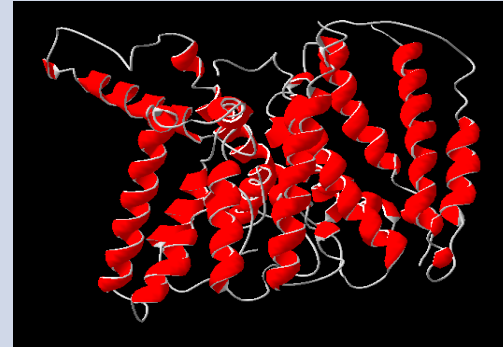
<p>(I) Tráfico intracelular, secreción y transporte vesicular</p>		<p>(M) Transporte y metabolismo de aminoácidos</p>	
	<p>(J) Modificación postraducciona, recambio proteico, chaperones</p>		<p>(N) Transporte y metabolismo de nucleótidos</p>
<p>(K) Producción y conversión de energía.</p>		<p>(O) Transporte y metabolismo de coenzimas</p>	
	<p>(L) Transporte y metabolismo de carbohidratos</p>		<p>(P) Transporte y metabolismo de lípidos</p>

Tabla 8. Estructura 3D de la proteína de mayor número de aminoácidos por función.

(Q) Transporte y metabolismo de iones inorgánicos



(R) Biosíntesis, transporte y catabolismo de metabolitos secundarios.

(S) Predicción de funciones generales



DISCUSIÓN

- Relación con LUCA
- Análisis filogenético.
- Origen de la vida.

M.
kandleri

- Regulación potenciales redox.
- Transporte de iones H⁺.

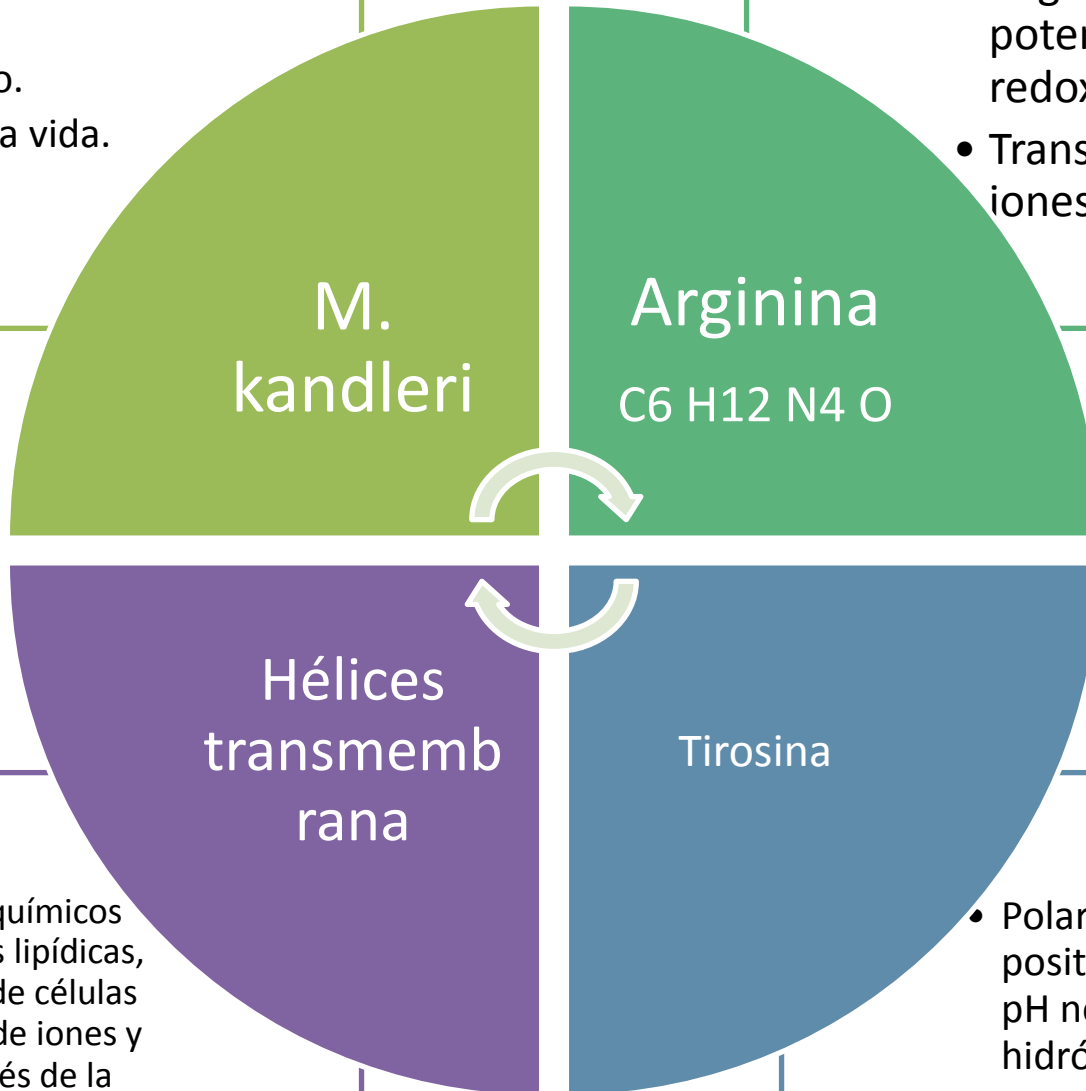
Arginina
C₆ H₁₂ N₄ O

Hélices
transmembrana

- Procesos bioquímicos en las bicapas lipídicas, señalización de células y transporte de iones y solutos a través de la membrana

Tirosina

- Polar cargado positivamente a pH neutro, hidrófilo.

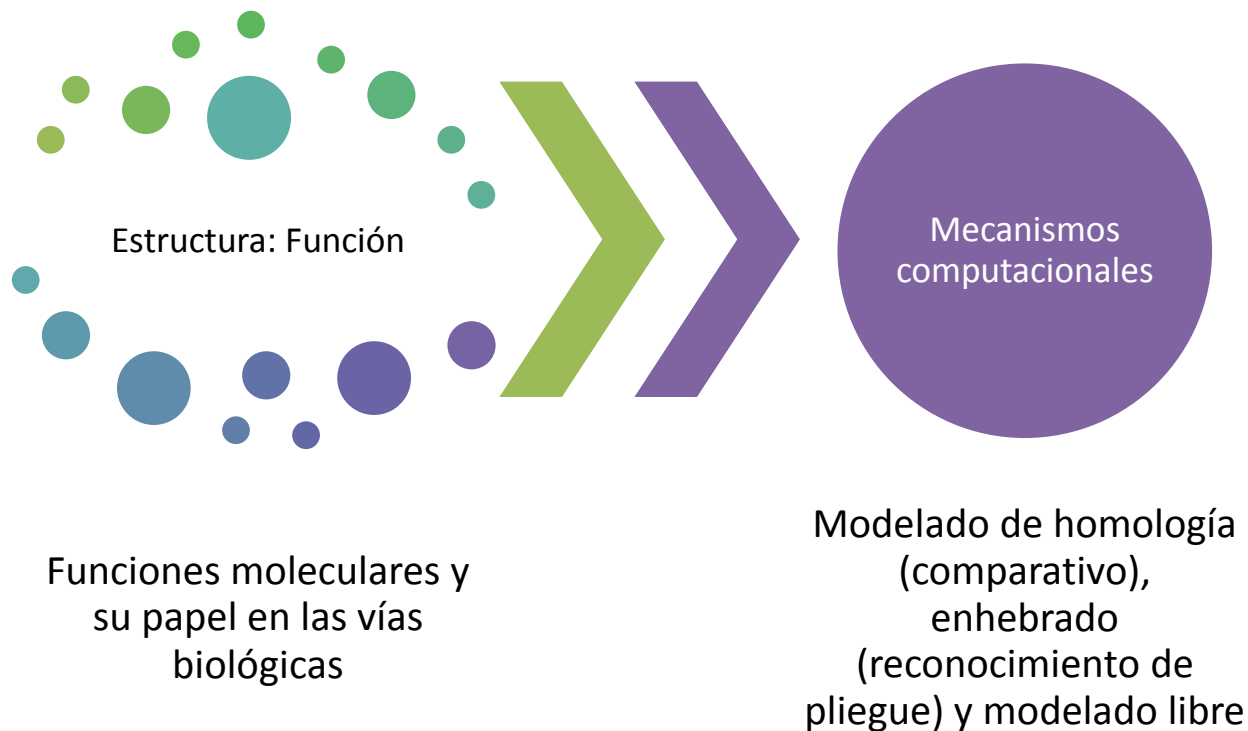


Alto de aminoácidos cargados negativamente, y proteínas con pI de aprox 5, para adaptación a alta salinidad intracelular (>3 M K⁺)

La media de pI es de 6.1. Estos resultados pueden deberse al número de proteínas usadas en este análisis (35.3%).

Escasez de proteínas implicadas en la señalización y regulación de la expresión génica.

Un aproximado de 135 de secuencias proteicas (22.6%) cumplen esta función, lo que podría concordar con dicha afirmación.



CONCLUSIONES

- En este trabajo se determinó que existían diferencias significativas para los aminoácidos Arg en las proteínas implicadas en la función de replicación, recombinación y reparación (Grupo D) y Tyr en el Grupo D respecto a las proteínas implicadas en la transcripción (Grupo C).
- Se determinó que algunos Grupos funcionales tienen alta abundancia de aminoácidos asociados a hélices transmembrana, los cuales son el Grupo H (pared celular, membrana y biogénesis), el Grupo I (tráfico intracelular, secreción y transporte vesicular), y el Grupo Q (transporte y metabolismo de iones inorgánicos).
- Se estableció las estructuras 3D de las proteínas de *M. kandleri* AV19 por cada grupo funcional, indicando una probable estructura estándar para cada función.
- Este trabajo es una aproximación teórica a un análisis bioquímico del proteoma de *M. kandleri* AV19, para inferir su posible relación con LUCA.

REFERENCIAS

- 1. Alberts B, Johnson A, Lewis J, et al. The Shape and Structure of Proteins. *Molecular Biology of the Cell*. 4th edition. New York : Garland Science, 2002.
- 2. J., Gittleman. Phylogeny - biology. [En línea] Encyclopedia Britannica, 2016. <https://www.britannica.com/science/phylogeny#ref281267> .
- 3. Koonin EV, Galperin MY.. Chapter 2, Evolutionary Concept in Genetics and Genomics. *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics*. Boston : Kluwer Academic, 2003.
- 4. *Biotechnological applications of extremophiles, extremozymes and extremolytes*. 99: Raddadi, N., Cherif, A., Daffonchio, D., Neifar, M. & Fava, F. 2015, Appl. Microbiol. Biotechnol., Vol. 99, págs. 7907–7913.
- 5. *The genomics of LUCA*. W., Mat. 2008, Frontiers in Bioscience, Vol. 13, pág. 5605.
- 6. *The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens*. Slesarev A, Mezhevaya K, Makarova K, Polushin N, Shcherbinina O, Shakhova V et al. 2002, Proceedings of the National Academy of Sciences, Vol. 7, págs. 46, 99.
- 7. *Archaeal phylogeny based on proteins of the transcription and translation machineries: tackling the Methanopyrus kandleri paradox*. . Brochier C, Forterre P, Gribaldo S. 3, 2004, Genome Biology, Vol. 5, pág. R17.
- 8. *In silico functional characterization of a double histone fold domain from the Heliothis zea virus 1*. . Greco C, Fantucci P, De Gioia L. 2005, Italian Society of Bioinformatics (BITS): Annual Meeting , Vol. 6(Suppl 4), pág. S15.
- 9. *Many non-universal archaeal ribosomal proteins are found in conserved gene clusters*. . Wang J, Dasgupta I, Fox G. 4, 2009, Archaea, Vol. 2, págs. 241-251.
- 10. *Evidence for early life in Earth's oldest hydrothermal vent precipitates*. . Dodd M, Papineau D, Grenne T, Slack J, Rittner M, Pirajno F et al. 543, 2017, Nature, Vol. 7643, págs. 60-64.
- 11. LEY DE BENFORD. [En línea] Estadística para todos, 2008. <http://www.estadisticaparatodos.es/taller/benford/benford.html>.
- 12. *Genome Sizes and the Benford Distribution*. Friar J, Goldman T, Pérez–Mercader J. 5, 2012, PLoS ONE, Vol. 7, pág. e36624.
- 13. Madigan M, Martinko J, Parker J, Gacto Fernández M. *Biología de los microorganismos*. Brock. . 10th ed. Madrid, España : Pearson, 2004.
- 14. Casanova V. Clasificación de los extremófilos. (Curso de Astrobiología, capítulo 2) . [En línea] Astrofiscayfisica.com, 2017. <http://www.astrofiscayfisica.com/2013/05/clasificacion-de-los-extremofilos-curso.html> .
- 15. *Extremophiles: Sustainable Resource of Natural Compounds-Extremolytes*. Kumas R, Patel D, Bansal D, Mishra S, Mohammed A, Arora R et al., 2009, Sustainable Biotechnology, págs. 279-294.
- 16. Darwin, C. *El origen de las especies por medio de la selección natural*. s.l. : Editorial CSIC-CSIC Press., 2009. pág. 464.
- 17. *Looking for the Last Universal Common Ancestor (LUCA)*. Koskela M, Annala A. 4, 2012, Genes, Vol. 3, págs. 81-87.
- 18. Egel, Richard. Integrative Perspectives: In Quest of a Coherent Framework for Origins of Life on Earth. [aut. libro] Dirk-Henner Lankenau, Armen Y. Mulikdjanian Richard Egel. *Origins of Life: The Primal Self-Organization*. s.l. : Springer Science & Business Media, 2011.
- 19. *The universal tree of life: an update*. . P, Forterre. 2015, Frontiers in Microbiology, Vol. 6, pág. 717.
- 20. Arunas L. . We've been wrong about the origins of life for 90 years. [En línea] EDT. The Conversation US, Inc., 2016. <http://theconversation.com/weve-been-wrong-about-the-origins-of-life-for-90-years-63744>.
- 21. *Emergence of life: from functional RNA selection to natural selection and beyond*. . J., Tze-Fei Wong. 7, 2014, Frontiers in Bioscience. , Vol. 19, pág. 1117.
- 22. *Coevolution Theory of the Genetic Code at Age Forty: Pathway to Translation and Synthetic*. Wong J, Ng S, Mat W, Hu T, Xue H. 1, 2016, Life. MDPI-Life, Vol. 6, pág. 12.
- 23. *Protein folding at extreme temperatures: Current issues*. Feller, Georges. 2017, Seminars in Cell & Developmental Biology, pág. 9.
- 24. *Search for Primitive Methanopyrus Based on Genetic Distance Between Val- and Ile-tRNA Synthetases*. Yu, Z., Takai, K., Slesarev, A. et al. J Mol Evol. 2009, Journal of Molecular Evolution, Vol. 69, págs. 386–394 .
- 25. *Analysis of two large functionally uncharacterized regions in the Methanopyrus kandleri AV19 genome*. Jensen LJ, Skovgaard M, Sicheritz-Pontén T, et al. 12, 2003, BMC Genomics , Vol. 4.
- 26. *A Novel Unsaturated Archaeal Ether Core Lipid from the Hyperthermophile Methanopyrus kandleri*. Hafenbradl D., Keller M., Thierick R., Stetter k. 2, 1993, Elsevier, Vol. 16, págs. 165-169.

- 27. *A two-subunit type I DNA topoisomerase (reverse gyrase) from an extreme hyperthermophile.* . Krah R, Kozyavkin SA, Slesarev AI, Gellert M. 1, 1996, Proceedings of the National Academy of Sciences of the United States of America, Vol. 93, págs. 106-110.
- 28. *Evidence for an early prokaryotic origin of histones H2A and H4 prior to the emergence of eukaryotes.* Slesarev AI, Belova GI, Kozyavkin SA, Lake JA. 2, 1998, Nucleic Acids Research, Vol. 26, págs. 427-430.
- 29. The Chemical Components of a Cell. [aut. libro] Johnson A, Lewis J, et al. Alberts B. *Molecular Biology of the Cell.* 4th edition. . New York : Garland Science., 2002.
- 30. Rodwell V, Murray R, Granner D, Mayes P. 26th ed. United States of America : McGraw-Hill Publishing, 2003.
- 31. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, Appel R et al. Protein Identification and Analysis Tools on the ExPASy Server . [En línea] Expasy, 2018. https://web.expasy.org/docs/expasy_tools05.pd.
- 32. Alberts B, Johnson A, Lewis J, et al. The Shape and Structure of Proteins. *Molecular Biology of the Cell.* 4th edition. New York : Garland Science, 2002.
- 33. *Structural features of transmembrane helices.* Werner P., Preissner R., Frömmel C. 2002, Elsevier, Vol. 559, págs. 145-151.
- 34. *Protein Databases on the Internet.* . Xu D, Xu Y. 2004, Current Protocols in Molecular Biology, págs. Unit–19.4.
- 35. *The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999.* Bairoch A, Apweiler R. 1, 1999, Nucleic Acids Res, Vol. 27, págs. 49-54.
- 36. Home, ExpASY: SIB Bioinformatics Resource Portal -. ExpASY. [En línea] 2017. <https://www.expasy.org/>.
- 37. *The Phyre2 web portal for protein modeling, prediction and analysis.* Kelley L, Mezulis S, Yates C, Wass M, Sternberg M. 6, 2015, Nature Protocols, Vol. 10, págs. 845-858.
- 38. *The RCSB Protein Data Bank: views of structural biology for basic and applied research and education.* Rose P, Prli A, Bi C, Bluhm W, Christie C, Dutta S et al. 1, 2014, Nucleic Acids Research, Vol. 43, págs. 345 - 356.
- 39. *The I-TASSER Suite: Protein structure and function prediction.* . J Yang, R Yan, A Roy, D Xu, J Poisson, Y Zhang. 7 - 8, 2015, Nature Methods, Vol. 12.
- 40. *Coevolution Theory of the Genetic Code at Age Forty: Pathway to Translation and Synthetic Life.* Wong, J. T.-F., Ng, S.-K., Mat, W.-K., Hu, T., & Xue, H. 1, 2016, LIFE, Vol. 6, pág. 12.
- 41. Scwald N, Dieter H. *Peptides: Chemistry and Biology.* s.l. : Wilye VCH, 2002.
- 42. *Arginine residues at internal positions in a protein are always charged.* . Harms MJ, Schlessman JL, Sue GR, García-Moreno E. B. 47, 2011, Proceedings of the National Academy of Sciences of the United States of America., Vol. 108, págs. 18954-18959.
- 43. *Transmembrane helix predictions revisited.* Chen CP, Kernytsky A, Rost B. 12, 2002, Protein Science : A Publication of the Protein Society, Vol. 11, págs. 2774-2791.
- 44. *A Universal Trend among Proteomes Indicates an Oily Last Common Ancestor.* Ranjan V. Mannige, Charles L. Brooks, Eugene I. Shakhnovich. 12, 2012, PLoS Comput Biology, Vol. 8.
- 45. *Automatic Prediction of Protein 3D Structures by Probabilistic Multi-template Homology Modeling.* Meier A, Söding J. 10, 2015, PLOS Computational Biology, Vol. 11, pág. e1004343.
- 46. *Mass Spectrometry Coupled Experiments and Protein Structure Modeling Methods.* Pi J, Sael L. 12, 2013, International Journal of Molecular Sciences., Vol. 14, págs. 20635-20657.
- 47. *Extremophiles and their application to veterinary medicine.* Irwin J, Baird A. 6, 2004, Irish Veterinary Journal., Vol. 57, pág. 348.
- 48. *The COG database: a tool for genome-scale analysis of protein functions and evolution.* . Tatusov R, Galperin M, Natale D, Koonin E. 1, 2000, Nucleic Acids Research., Vol. 28, págs. 33-36.
- 49. Koonin EV, Galperin MY. Chapter 2, Evolutionary Concept in Genetics and Genomics. . *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics.* Boston : Kluwer Academic , 2003. .
- 50. *Emergence of life: from functional RNA selection to natural selection and beyond.* J., Tze-Fei Wong. 7, 2014, Frontiers in Bioscience., Vol. 19, pág. 1117.
- 51. *Database resources of the National Center for Biotechnology Information.* Coordinators, NCBI Resource. 1, 2015, Nucleic Acids Research, Vol. 44, págs. 7 - 19.
- 52. Tool, NCBI N. BLAST: Basic Local Alignment Search. [En línea] 2017. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>.
- 53. *BLAST: a more efficient report with usability improvements.* Boratyn G, Camacho C, Cooper P, Coulouris G, Fong A, Ma N et al. 1, 2013, Nucleic Acids Research, Vol. 41, págs. 29 - 33.
- 54. *The Phyre2 web portal for protein modeling, prediction and analysis.* al., Kelley LA et. 2015, Nature Protocols, Vol. 10, págs. 845-858 .

