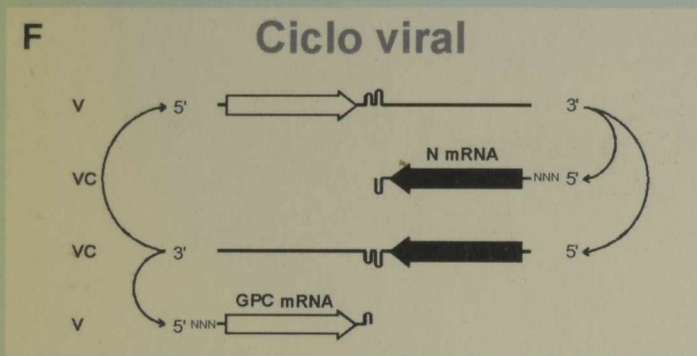
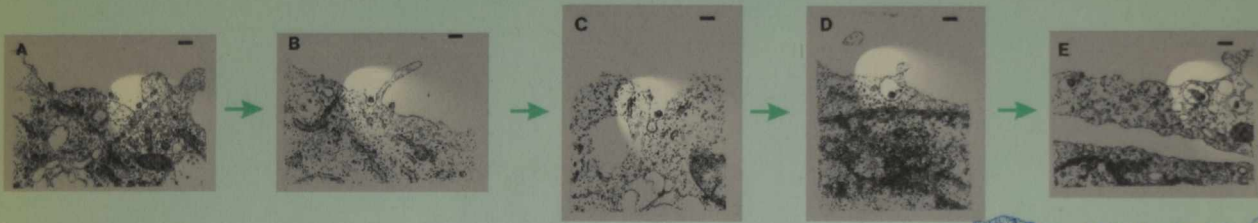




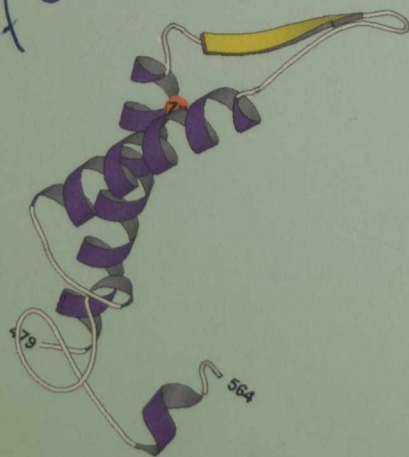
Universidad Nacional de La Plata
 Facultad de Ciencias Exactas
 Departamento de Ciencias Biológicas

Virus Junín: clonado molecular y análisis estructural y funcional del RNA S y sus productos génicos



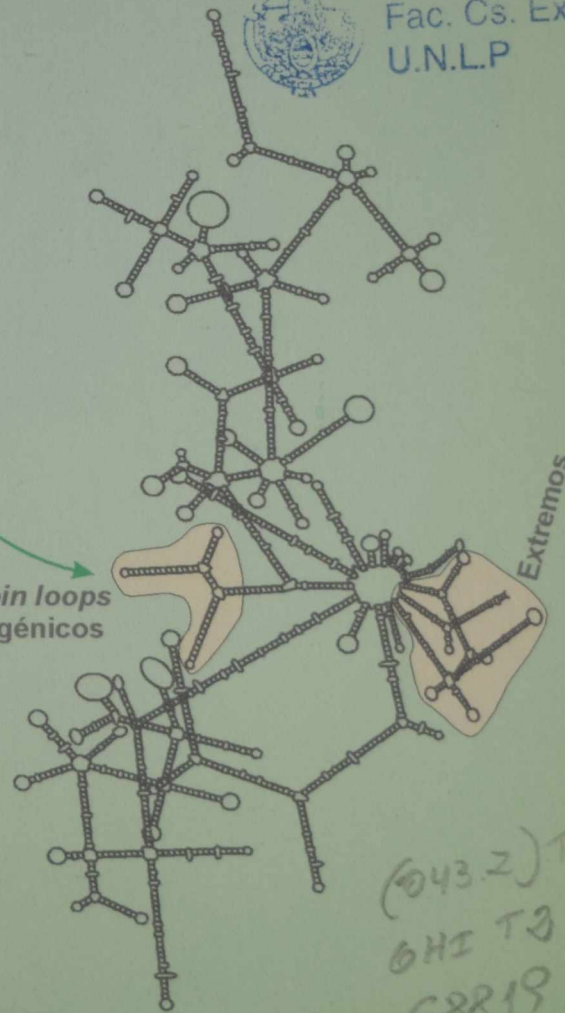
Biblioteca Central
 Fac. Cs. Exactas
 U.N.L.P

TESIS
 No 957 (2)



Dominio 3
 Proteína N del virus Junín

Hairpin loops
 intergénicos



(943.2) Tesis
 GHI T3
 58819

Pablo Daniel Ghiringhelli
 Tesis Doctoral
 La Plata, 2002

DONACION.....

Fecha: 21-09-2009

Inv. B..... Inv..... B 58819



TOMO II

Partes VI a IX



Biblioteca Central
Fac. Cs. Exactas
U.N.L.P.

PARTE V

MATERIALES Y METODOS

DONACION
A.....
Fecha *12/11/02*.....
Inv. E.....**Inv.**.....**B**.....

Capítulo 15 Materiales

REACTIVOS COMERCIALES

PRODUCTOS QUÍMICOS Y ENZIMAS

En todos los casos, se usaron reactivos de grado analítico de Merck (Darmstadt, Alemania), Sigma (St. Louis, EE.UU.), Fluka (Buchs, Suiza), BDH (Poole, Inglaterra), Carlo Erba (Milán, Italia) o Mallinkrodt (Nueva York, EE.UU.)

Los reactivos y enzimas usados en los experimentos de manipulación de genes fueron de grado Biología Molecular de Sigma (St. Louis, EE.UU.), Boehringer Mannheim (Mannheim, Alemania), IBI (New Haven, EE.UU.), U.S.B. (Cleveland, EE.UU.), Bethesda Research Laboratories (Gaithersburg, EE. UU.), Stratagene (La Jolla, EE.UU.), Promega (Madison, EE.UU.), New England Biolabs (Beverly, EE.UU.) o Cetus (Norwalk, EE.UU.). Los componentes de los medios de cultivo para bacterias se adquirieron en Difco (Detroit, EE.UU.).

Los antibióticos, ampicilina, tetraciclina y cloranfenicol fueron gentilmente donados por laboratorios Bagó.

Los medios de cultivo de células eucarióticas se adquirieron en Gibco (Grand Island, EE.UU.) o Flow (Mc Lean, EE.UU.). El suero fetal bovino fue provisto por Gibco (Grand Island, EE.UU.) o Bioser (Buenos Aires, Argentina).

MATERIAL RADIOACTIVO

El material utilizado fue de New England Nuclear (NEN) (Boston, EE.UU.). Los nucleótidos radiactivos, utilizados en las reacciones de marcación de ácidos nucleicos, se sintetizaron a partir del $^{32}\text{P}[\text{H}_3\text{PO}_4]$. El dosaje del material radiactivo se realizó por medición con un contador Geiger (para ^{32}P) o por medio

de un contador de centelleo líquido por efecto Cerenkov.

Para las autorradiografías se usó película X-OmatTM K de Kodak (Rochester, EE.UU.) revelándose con revelador manual para placas radiográficas.

CELULAS, VIRUS Y MUESTRAS CLINICAS

CEPAS BACTERIANAS

En los experimentos de clonado y manipulación de genes se usaron las siguientes cepas de *Escherichia coli*:

DH5aF' {endA1, hsdR17(rk⁻, mk⁺), supE44, thi-1, recA1, gyrA96, relA1, DlacU169 F'[f80 lacZDM15]}. Hanahan (1983); BRL (1986).

XL1-Blue {supE44, hsdR17, thi-1, recA1, endA1, gyrA96, relA1, lac⁻, F'[proAB, lacI^q, lacZΔM15, Tn10 (tet^r)]}. Bullock, *et al.*, 1987.

SURE {mcrA, mcrB, mrr, endA1, supE44, recA1, l, thi, gyrA96, umuC::Tn5 (kan^r), cbcC201, uvrC, lac⁻, relA1, recB, recJ, D(hsdRMS), F'[proAB, lacI^q, lacZDM15, Tn10 (tet^r)]}. StratageneTM (1989).

LÍNEAS CELULARES EUCARIOTICAS

Se utilizaron las siguientes líneas celulares, cuyo origen en todos los casos fue a partir de cepas de American Type Culture Collection (Rockville, EE.UU.):

BHK-21 (C-13): fibroblastos de riñón de hamster lactante (Stoker and McPherson, 1961).

VERO E6: fibroblastos de riñón de mono verde africano *Cercopithecus aegypticus* (Yasumura and Kawakita, 1963).

VIRUS

Las cepas del virus Junín utilizadas en esta tesis fueron la cepa MC2, aislada originariamente de un roedor de la zona de

Mones Cazón (Vilches *et al.*, 1965; De Mitri and Martínez Segovia, 1971), y la cepa Candid #1 (Barrera Oro and Eddy, 1982; McKee *et al.*, 1984; Maiztegui *et al.*, 1987).

Los inóculos virales de la cepa Junín-MC2, conservados en el Instituto de Bioquímica Biología Molecular (IBBM) de la Facultad de Ciencias Exactas de la UNLP, fueron obtenidos por pasaje en cerebro de ratón lactante de un inóculo originalmente provisto por el Instituto Nacional de Microbiología "Carlos G. Malbrán".

La cepa Junín-MC2 fue clonada por el método de placa aislada en dos pasajes sucesivos por células VERO y se amplificó en monocapas de células BHK-21. Estos virus clonados fueron utilizados como inóculos en los distintos experimentos.

La cepa Junín-Candid #1 fue amplificada en células FRhL-2 en el INEVH (Pergamino), siguiendo el protocolo de producción de la vacuna, y enviada al IBBM para su clonado y secuenciamiento.

MUESTRAS CLÍNICAS

Para los ensayos de ELISA se recibieron muestras de suero de diferentes individuos que, habitando dentro del área endémica, fueron hospitalizados en el INEVH. Los pacientes fueron caracterizados de acuerdo al diagnóstico clínico establecido en el hospital. De acuerdo a esa caracterización, los pacientes se dividieron en tres grupos. El primer grupo estaba compuesto por aquellos pacientes con diagnóstico clínico de FHA, el segundo, por pacientes con diagnóstico clínico de síndrome febril de etiología indeterminada (SFEI) y el último, por aquellos pacientes con otros diagnósticos, que fueron utilizados como controles del ensayo. Esta agrupación se realizó solamente después de concluidos los ensayos a ciego realizados independientemente en el INEVH y el IBBM. El día de comienzo de la infección fue calculado por anamnesis de cada uno de los pacientes.

MEDIOS DE CULTIVO DE BACTERIAS

MEDIO DE DESARROLLO

Medio LB

Extracto de levadura	5 g
Triptona	10 g
NaCl	10 g
H ₂ O csp	1 litro

Se ajusta el pH a 7,6 y se esteriliza durante 15 minutos a 120°C en un autoclave.

Para los medios sólidos: agregar 15 g de agar/litro.

MEDIOS PARA SELECCIÓN DE CLONES RECOMBINANTES

En general se utiliza un medio base de LB con el agregado de antibióticos según la característica fenotípica transmitida por el plásmido transformante. Los antibióticos más comunes son utilizados a las siguientes concentraciones:

Antibiótico	Solución stock	Concent. final
Ampicilina	25 mg/ml	50 µg/ml
Cloramfenicol	34 mg/ml (etanol)	30 µg/ml
Cloramfenicol (amplificación de plásmidos)	idem	170 µg/ml
Tetraciclina	12 mg/ml (etanol 50%)	12 µg/ml

Además, en el caso en que el sistema utilizado permita utilizar el ensayo de α -complementación es posible discriminar entre las bacterias transformadas con plásmidos sin inserto (α -complementación positiva) de las bacterias transformadas con plásmidos que contienen un inserto (α -complementación negativa). Para ello, se adiciona al medio de cultivo el inductor, IPTG (Isopropil-tio- β -D-galactósido) y el reactivo cromogénico, X-gal (5-Br-4-Cl-3-indolil- β -D-galactósido). Así, a las cajas de Petri ya preparadas conteniendo

agar/LB con el antibiótico adecuado, se les agregan 40 μ l de una solución de X-gal en dimetilformamida (20 mg/ml) y 4 μ l de una solución acuosa de IPTG (200 mg/ml). Los reactivos deben esparcirse rápidamente utilizando una espátula de vidrio estéril, para asegurar la homogeneidad.

MEDIOS PARA EL DESARROLLO DE COMPETENCIA

Medio SOB

Extracto de levadura	5 g
Triptona	20 g
NaCl	0,5 g
H ₂ O csp	980 ml

Se ajusta el pH a 7,5 con KOH y se esteriliza durante 15 minutos a 120°C en un autoclave.

Inmediatamente antes de utilizar, se agregan 20 ml de MgSO₄ 1M esterilizado en autoclave.

Tfb I

Acetato de potasio	30 mM
KCl	100 mM
CaCl ₂	10 mM
MnCl ₂	50 mM
glicerol	15 %.

Se ajusta el pH a 5.8 con ácido acético 0.2 M y se esteriliza por filtración.

Tfb II:

PIPES	10 mM
CaCl ₂	75 mM
KCl	10 mM
glicerol	15 %.

Se ajusta el pH a 6,5 con KOH y se esteriliza por filtración.

SOLUCIONES PARA LA PURIFICACION Y EL ANALISIS DE LOS ACIDOS NUCLEICOS

PURIFICACION DE RNA

TCG 2X

isotiocianato de guanidinio	8 M
Citrato de sodio	50 mM
Sarkosyl	1%
β -mercaptoetanol	180 mM

La solución de TCG 2X cristaliza a temperatura ambiente. Para disolverla se debe poner la botella en un baño María (100°C) durante 15 minutos, aflojando la tapa. Cuando se observa que la solución se torna translúcida, se debe agitar la botella por rotación hasta que la solución quede homogénea.

Fenol ácido

Fenol equilibrado con agua bidestilada estéril.

Se funde el fenol en un baño maría. Se agregan 0,3 volúmenes de agua. Se agita hasta formar una emulsión y se dejan decantar las fases. La fase superior corresponde a la solución acuosa y la fase inferior al fenol. El pH de la solución acuosa debe estar alrededor de 4,5.

NOTA: Las soluciones que contienen fenol se deben manejar con sumo cuidado porque el reactivo produce quemaduras profundas en piel y mucosas. Ante cualquier accidente lavar con abundante agua.

Cloroformo:alcohol isoamílico

24 partes de Cl₃CH + 1 parte de alcohol isoamílico. El isoamílico no es esencial y se utiliza para mejorar la separación de fases.

NaAc 3M

Se disuelven 408,1 g de acetato de sodio en 800 ml de H₂O. Se ajusta el pH a 5,2 con ácido acético glacial, se ajusta el volumen a 1 litro y se esteriliza por autoclave.

Alcoholes

Isopropanol y etanol 96% y 70 %.

PREPARACION DE PLASMIDOS SEGUN EL PROTOCOLO DE LISIS POR ALCALI

Alcoholes

etanol 96% y 70%

TE 1X

TRIS-HCl, pH 8	10 mM
EDTA	1 mM

1,0% (P/V) SDS.

10g de SDS en 100ml de agua destilada. Se mantiene a temperatura ambiente.

Solución I

Glucosa	50 mM
Tris ClH pH 8	25 mM
EDTA	10 mM

Se esteriliza por filtración. Se mantiene a 4°C

Solución II

NaOH	0,2 N
SDS	1 %

La solución debe prepararse fresca en el momento de su uso, a partir de soluciones de NaOH y SDS concentradas.

Solución III

AcK 5 M, pH 4,8.

La solución es 5 M con respecto al acetato y 3 M con respecto al potasio.

Fenol básico (pH 8)

Fenol equilibrado con Tris:HCl pH 8.

El fenol se funde en un baño maría. Se agrega 1 volumen de una solución de Tris base 1 M. Se agita hasta formar una emulsión y se deja decantar las fases. La fase superior corresponde a la solución acuosa y la fase inferior al fenol. Se elimina la solución acuosa. Se repite el procedimiento hasta que el pH de la solución acuosa alcance el valor de 8.

Fenol:cloroformo:alcohol isoamílico

24 partes de fenol, pH 8

24 partes de Cl₃CH

1 parte de alcohol isoamílico.

El alcohol isoamílico no es esencial. y se utiliza para definir mejor la interfase.

ANÁLISIS DE ÁCIDOS NUCLEICOS MEDIANTE ELECTROFORESIS EN GELES DE AGAROSA**Geles de Agarosa:**

Agarosa 0,7%, Agarosa 1%, Agarosa 2,5%, Agarosa NuSieve 4%. Estas soluciones contienen TAE 1X y Bromuro de etidio 0,16 mg/ml

Buffer de corrida: TAE 1X

Tris-acetato	0,04 M;
EDTA	0,001 M.

El TAE se puede preparar como solución concentrada 50X.

TAE 50X (para 1 litro)

Tris base	242 g
0,5 M EDTA, pH 8	100 ml
ácido acético glacial	57,1 ml

Buffer de corrida: RNA/formaldehído 5X

MOPS pH 7	0,1 M
NaAc	40 mM
EDTA pH 8	5 mM

Se prepara disolviendo 20,6 g de ácido 3-(N-morfolin)propansulfónico (MOPS) en 800 ml de 50 mM NaAc. Se ajusta el pH a 7,0 con NaOH, se agregan 10 ml de 0,5 M EDTA pH 8, y se ajusta el volumen a 1 litro con H₂O. Se esteriliza por filtración y se guarda a temperatura ambiente, protegido de la luz solar.

Bromuro de etidio

Se prepara una solución concentrada hasta 50 mg/ml de bromuro de etidio en H₂O y se mantiene a 4°C protegiéndola de la acción de la luz natural.

Soluciones para la siembra de muestra:

Se utilizaron 4 diferentes soluciones de acuerdo al tipo de gel empleado. Todas contienen un 0,2% (p/v) del (o los) colorante(s) indicado(s) más abajo.

A (para geles de agarosa 0,7% o 1%): TAE 1X, glicerol 50%, xilencianol y azul de bromofenol.

B (para geles de agarosa 2,5% o 4%): TAE 1X, glicerol 50% y naranja G.

C (para geles de RNA nativos): glicerol 50%, colorantes: xilencianol FF y azul de bromofenol.

D (para geles de RNA/formaldehído): glicerol 50%, 1 mM EDTA pH 8, xilencianol y azul de bromofenol.

SOUTHERN BLOT**SSC 20x:**

NaCl	3,0 M
Citrato de sodio pH 7,0	0,3 M

Se esteriliza por autoclave.

Denhardt 50X

Ficoll	1% p/v
Polivinilpirrolidona	1% p/v
Serolabúmina	1% p/v

Solución de prehibridación hibridación

Formamida	50%	45%
SSC	5x	5x
Denhardt	5x	1x
Pirofosfato de Na pH 6,5	25 mM	20 mM
DNA de esperma de salmón	0,5 mg/ml	0,2 mg/ml
Sonda	-	100 ng/ml

REACCIONES ENZIMATICAS

Buffer de ligación 10X

Tris·HCl pH 7,6	200 mM
MgCl ₂	50 mM
Ditiotreitol	50 mM
Seroalbúmina bovina	500 mg/ml

Buffer de DNAPol I y Klenow 10X

Tris-ClH pH 7.9	10 mM
NaCl	50 mM
MgCl ₂	10 mM
Ditiotreitol	1 mM

Buffer de PNK (polinucleotidoquinasa) 10X

Tris-ClH pH 7.5	40 mM
MgCl ₂	10 mM
Ditiotreitol	5 mM

Buffer de PCR 10X

Tris-ClH pH 9,0	100 mM
KCl	500 mM
Tritón X-100	1,0 %

Buffer de Transcripción 5X

Tris-ClH pH 7.5	200 mM
NaCl	50 mM
MgCl ₂	30 mM
Espermidina	10 mM

SECUENCIAMIENTO NUCLEOTIDICO DE DNA

Buffer de reacción 5X

Tris-ClH pH 7.5	200 mM
MgCl ₂	100 mM
NaCl	250 mM

Mezcla de marcación

Ditiotreitol 0,1M	0,5 ml
dGTP 7,5 mM, dCTP 7,5 mM, dTTP 7,5 mM	0,2 ml

Solución de finalización

formamida	95 %,
EDTA	20 mM,
azul de bromofenol	0,05 %
xilencianol	0,05 %

GELES DE POLIACRILAMIDA

Purificación de oligonucleótidos (minigeles)

40% acrilamida/bisacrilamida 38:2	5 ml
urea	5 gr
TBE 10X,	1 ml
H ₂ O csp	10 ml

Desgasificar durante 15 minutos o más a temperatura ambiente.

10% (p/v) persulfato de amonio	100 ml
TEMED	10 ml

Armar el gel y dejar 1 hora a temperatura ambiente para que polimerice.

Geles de secuenciamento (geles largos)

40% acrilamida/bisacrilamida 38:2	9 ml
urea	30 gr
TBE 10X,	6 ml
H ₂ O csp	60 ml

Desgasificar durante 15 minutos o más a temperatura ambiente.

10% (p/v) persulfato de amonio	500 ml
TEMED	50 ml

Armar el gel y dejar al menos 2 horas a temperatura ambiente para que polimerice.

Capítulo 16 Métodos

Preparación de arenavirus y células de mamífero infectadas

Cultivo de células

Las distintas líneas celulares se cultivaron en monocapa en recipientes de poliestireno (frascos o cajas de Petri). Los pasajes se efectuaron por tripsinización de la monocapa. El conteo de las células, cuando fue necesario, se realizó en una cámara de Neubauer y la viabilidad se determinó por exclusión de azul de tripan.

Los medios de cultivo usados fueron MEM (*Minimal Essential Medium*) de Gibco (Cat. 410-1500) para las células VERO y DMEM (*Dulbecco's Minimal Essential Medium*) de Gibco (Cat. 430-1600) para las células BHK-21. Ambos medios se suplementaron con aminoácidos no esenciales, glutamina hasta 2 mM final, gentamicina a 50 µg/ml y anfotericina B a 1,5 µg/ml. El DMEM se suplementó con agregado de glucosa hasta 4,5 g/l. Los medios se suplementaron, además, con 10% de suero fetal bovino, para el crecimiento de los cultivos, y 2% para el mantenimiento.

Todas las líneas se cultivaron a 37°C en frascos de plástico cerrados. En cajas de Petri, la incubación se realizó en atmósfera de aire saturada con agua y con 5% de CO₂, agregando HEPES pH 7,5, hasta 10 mM final a los medios de cultivo.

El material de plástico usado en el cultivo de células fue provisto por Nunc (Roskilde, Dinamarca) o Corning Glass Work (Corning, EE.UU). Para la reutilización de este material el lavado del mismo se efectuó con solución diluida de Tritón X-100, enjuagándose con abundante agua destilada y agua bidestilada. Una vez seco, se lo esterilizó en un

horno de microondas durante 3 ciclos de 5 minutos con la máxima potencia.

Propagación de virus

La producción de virus se llevó a cabo *in vitro* infectando células BHK-21, teniendo en cuenta los datos publicados por Martínez Segovia *et al.* (1967).

Brevemente, las células en crecimiento exponencial (con la monocapa cubriendo el 50% de la superficie de cultivo) se lavaron con PBS y se inocularon con el stock de virus diluido en medio MEM a una multiplicidad de infección (mdi) de 0,1 unidad formadora de placas (ufp) por célula. Luego de un período de adsorción de 1 h a 37°C las monocapas infectadas se lavaron con PBS y se incubaron a 37°C con medio de cultivo conteniendo 2% de suero fetal bovino.

A diferentes tiempos post-infección se renovó el medio de cultivo, tomando alícuotas para titular.

Condiciones de bioseguridad

Todos los procesos de manipulación de virus vivos o células infectadas se realizaron en condiciones de seguridad correspondientes a un laboratorio tipo BSL3, de acuerdo a las normas de bioseguridad para organismos patógenos establecidas por los Institutos Nacionales de Salud de los EE.UU. Este laboratorio cuenta con tres cuartos. Uno, para el manejo de material patógeno, donde están ubicadas las estufas para el cultivo de células. Este cuarto de 4,5 m² se conecta únicamente con el segundo cuarto, otra cámara de 12 m² donde están ubicadas las centrifugas (común y ultracentrífuga) y donde se encuentra, además, una cámara de guantes en la que se puede trabajar con animales infectados. El tercer

cuarto es una antecámara de 1 m² que se utiliza como entrada del laboratorio, y de donde se toman y se guardan los guardapolvos que se utilizan para entrar en el laboratorio de trabajo. El personal que trabaja con el material infectivo se coloca, además de los guardapolvos convencionales de la antecámara, la vestimenta de protección en el segundo cuarto. Esta consiste en una bata, cofia, barbijo, guantes de cirugía, máscara de protección facial y bolsas para recubrir el calzado y la zona inferior de las piernas. Los cuartos y la cámara de guantes están sometidos a una presión negativa durante el trabajo con material infectivo. De esta manera, el aire es extraído a presión del laboratorio, y es filtrado a través de unas membranas embebidas en alcohol y detergente, y es obligado a atravesar un laberinto con tubos fluorescentes que emiten luz ultravioleta.

Titulación de virus

Los títulos de los inóculos de virus se determinaron por el método de Dulbecco (1952) de ensayo de formación de placas de lisis celular sobre monocapas de células VERO de acuerdo a Rustici (1984). El tiempo de adsorción fue de 60 minutos para un volumen de inóculo de 100 ml por cada caja de Petri de 35 mm de diámetro.

Preparación de RNA

Control de la actividad de RNAsas

Para obtener buenas preparaciones de RNA es necesario minimizar la actividad de las RNAsas endógenas de la preparación y, además, evitar la introducción de trazas de las mismas desde otras fuentes potenciales, presentes en el laboratorio. Con respecto a lo último, existen pocas metodologías estandarizadas para evitarlo (como el empleo de guantes para evitar el contacto de las células de descamación de la piel con las muestras). Por ello, para lograr un manejo seguro de muestras conteniendo RNA es necesario llevar a cabo las manipulaciones con extremo cuidado.

En cambio, para solucionar el problema de la actividad de las RNAsas endógenas, es posible encontrar una gran variedad de métodos en la literatura. El más común es el que utiliza soluciones conteniendo dietilpircarbonato (DEPC) para inhibir la actividad enzimática. Sin embargo, este método presenta una serie de desventajas que es importante evitar. Las más importantes de las mismas son la toxicidad del DEPC, su actividad inespecífica sobre todas las actividades enzimáticas, y la dificultad de eliminarlo de la solución después del tratamiento.

Por ello, nosotros utilizamos este método únicamente como última alternativa. Además, encontramos que tanto el material de plástico o vidrio, así como las soluciones utilizadas, cuando son esterilizadas dos veces en autoclave, poseen una actividad de RNAsas lo suficientemente disminuidas, como para ser empleadas en este tipo de técnicas.

Extracción de RNA de células infectadas

El RNA total de células se obtiene mediante el siguiente procedimiento:

- (1) La monocapa de células se desprende del sustrato usando PBS sin Ca⁺² ni Mg⁺², a 4°C, y con 5 mM final de EDTA.
- (2) La suspensión de células se centrifuga a baja velocidad en una centrífuga clínica, descartándose el sobrenadante.
- (3) Se efectúa la lisis celular con solución TCG 1X, usando un volumen equivalente a 20 veces el volumen de las células obtenido en el paso anterior. Esta solución se extrae con fenol ácido (1 volumen) y cloroformo: alcohol isoamílico 24:1 v/v (0,2 volúmenes) y luego se precipita con 2 volúmenes de isopropanol en presencia de acetato de sodio 0,2 M pH 5,2; de acuerdo a la técnica de Chomczynski & Sacchi (1987), con modificaciones menores.

En nuestros experimentos el *pellet* de RNA obtenido fue resuspendido en 100 µl de

agua libre de RNasas, fue precipitado con 2,5 volúmenes de etanol en presencia de acetato de sodio 0,3 M, y lavado con etanol 75%. El RNA así aislado, fue secado, resuspendido en 10 µl de agua libre de RNasas y conservado congelado a -70°C. La calidad del RNA obtenido fue analizada por electroforesis en geles de agarosa.

Alternativamente, para obtener preparaciones de RNA con un mayor grado de pureza, se siguió el procedimiento descrito por Chirgwin *et al.* (1979).

La monocapa de células se recolecta de la misma manera que en el procedimiento anterior, y la lisis celular se realiza en una solución conteniendo: NaCl 150 mM; EDTA 1 mM; Tris-HCl 10 mM, pH 7,4 y NP40 0,2%. Esta solución se centrifuga a baja velocidad (4000 x g durante 4 minutos) para precipitar los núcleos y restos de membrana celulares. El sobrenadante se siembra en un gradiente preformado, entre 20% y 40% p/v, de CsCl, y se centrifuga a 45000 rpm durante 18 hs, en un rotor de ultracentrifuga, SW60. Este procedimiento puede ser utilizado tanto para obtener RNAs mensajeros celulares (mRNAs), que forman un precipitado en el fondo del gradiente, como para preparar nucleocápsides (RNA ligado a proteínas), que forman una banda en el centro del gradiente.

En nuestros experimentos, los mRNAs se resuspendieron en una solución conteniendo: 1 mM EDTA, 10 mM Tris-HCl pH 8 y 0,1% SDS, se precipitaron con 0,3M NaAc en etanol 75% y, finalmente, se resuspendieron en H₂O.

Preparación de baculovirus y células de insecto infectadas

Cultivo de células

Las células de *Spodoptera frugiperda* IPLB-SF21 (SF-21) (Vaughn *et al.*, 1977) fueron crecidas a 28°C en TC-100 (GIBCO BRL, Gaithersburg, MD, USA) conteniendo 10% de suero fetal bovino (Bioser, Buenos Aires, Argentina). Los pasajes se efectuaron

resuspendiendo la monocapa con un rastrillo de silicona (*cell scraper*). El conteo de las células, cuando fue necesario, se realizó en una cámara de Neubauer y la viabilidad se determinó por exclusión de azul de tripan.

Todos los cultivos se realizaron a 28°C en frascos de plástico cerrados y/o en *spinners* con agitación magnética.

El material de plástico usado en el cultivo de células fue provisto por Nunc (Roskilde, Dinamarca) o Corning Glass Work (Corning, EE.UU). Para la reutilización de este material el lavado del mismo se efectuó con solución diluida de Tritón X-100, enjuagándose con abundante agua destilada y agua bidestilada. Una vez seco, se lo esterilizó en un horno de microondas durante 3 ciclos de 5 minutos con la máxima potencia.

Propagación de virus

La producción de virus se llevó a cabo *in vitro* infectando células SF-21, de acuerdo a O'Reilly *et al.* (1992).

Brevemente, se retiró el medio cultivo de las monocapas en crecimiento exponencial (con las células cubriendo el 50% de la superficie de cultivo) y se inocularon con el stock de virus diluido en medio TC-100 sin suero a una multiplicidad de infección (mdi) de 0,1 unidad formadora de placas (ufp) por célula. Luego de un período de adsorción de 1 h a temperatura ambiente, se completó con medio TC-100 completo y se incubó a 28°C durante 5 a 7 días. Periódicamente, se controló la aparición de efectos citopáticos y/o poliedros.

La cuantificación de los stocks virales se realizó mediante ensayos de placas (O'Reilly *et al.*, 1992).

Purificación de DNA viral

El DNA viral fue purificado a partir de las formas brotantes obtenidas en los cultivos anteriores, de acuerdo a lo descrito en O'Reilly *et al.* (1992).

La integridad fue verificada mediante electroforesis en geles de agarosa 0,5 %.

Preparación de extractos proteicos

Para análisis por SDS-PAGE y Western blot

A diferentes tiempos post-infección se colectaron muestras (12, 24, 36, 48, 60 y 72 p.i.). Las monocapas de células infectadas fueron resuspendidas con un *cell scraper* en el medio de cultivo y recuperadas mediante centrifugación a baja velocidad. El *pellet* de células fue enjuagado con PBS 1 X, resuspendido en 400 µl de PBS 1X y centrifugado durante 1 min. a 14000 rpm. Estos *pellets* fueron resuspendidos en 40 µl de buffer de muestra 1X; 5 µl de cada muestra fueron resueltas mediante electroforesis en geles desnaturizantes de poliacrilamida 10% (SDS-PAGE) y visualizadas mediante tinción con *Coomassie Brilliant Blue R250*.

Para los ensayos de *Western blot*, las proteínas fueron electrotransferidas a membrana de nylon, utilizando un *buffer* Tris-glicina pH 8,3 con 0,1 % SDS. Posterior a la transferencia, las membranas fueron bloqueadas con 3 % de leche descremada Molico y reveladas con anticuerpos monoclonales específicos (Sanchez *et al.*, 1989).

Para ensayos de replegamiento

Las fracciones insolubles de los lisados proteicos fueron solubilizadas en un *buffer* desnaturizante: 6 M clorhidrato de guanidinio, 25 mM Tris-HCl pH 7,5 y 50 mM NaCl. En este *buffer*, la proteína N permanece soluble a temperatura ambiente durante bastante tiempo. Además, en estas condiciones la proteína N puede ser inmovilizada cuantitativamente en membranas de nylon o en policubetas de ELISA.

Las proteínas desnaturizadas pueden ser replegadas de dos maneras alternativas, en solución o inmovilizadas.

La proteína solubilizada en el *buffer* desnaturizante puede ser replegada en solución mediante la diálisis extensiva contra un *buffer* carente de clorhidrato de guanidinio.

Las proteínas inmovilizadas, en cambio, pueden ser renaturalizadas sobre la membrana mediante sucesivas incubaciones de 15 min., a 4°C, en *buffers* donde se va disminuyendo la concentración del clorhidrato de guanidinio a la mitad, desde 6 M hasta 0,375 M. Posteriormente, se incuba durante 1 hora a 4°C en *buffer* 25 mM Tris-HCl pH 7,5 y 50 mM NaCl.

Ensayos de interacción proteína-Zn y/o proteína-RNA

Interacción proteína-Zn

Las membranas con las proteínas inmovilizadas, fueron sometidas o no al tratamiento de renaturalización. A continuación, fueron incubadas con ⁶⁵ZnCl₂ en *buffer* Tris-HCl pH 7,5, durante 1 hora, a temperatura ambiente. Luego fueron lavadas extensivamente con agua y reveladas mediante autorradiografía.

Interacción proteína-RNA

Las membranas con las proteínas inmovilizadas, fueron sometidas o no al tratamiento de renaturalización. A continuación, fueron incubadas con sondas de RNA marcadas con ³²P, correspondientes a la región intergénica o la potencial estructura tipo *panhandle*, en *buffer* Tris-HCl pH 7,5, durante 1 ó 2 horas, a temperatura ambiente. Luego fueron lavadas extensivamente con *buffer* Tris-HCl pH 7,5 y reveladas mediante autorradiografía.

Ensayos de ELISA

Las fracciones insolubles de los lisados proteicos fueron inmovilizadas en policubetas de 96 *wells* en *buffer* carbonato, durante 1 hora a toda la noche, a 4°C.

Estas policubetas fueron utilizadas para detectar antígenos circulantes en sueros de ratón, cobayo y humanos; de acuerdo a los protocolos estándar de ELISA. El sistema utilizado para el revelado fue el de IgG conjugada a peroxidasa y *O*-fenilendiamina (OPD) como sustrato.

Preparación de DNA

Mini-preparaciones de plásmidos (lisis por álcali).

Para las preparaciones de DNA plasmídico en escala pequeña, se trabajó a partir de 5 ml de un cultivo saturado de cada clon de bacterias seleccionado.

Para realizar la técnica, se toma una ansada de una colonia de bacterias aislada conteniendo las características fenotípicas seleccionadas, y se la siembra sobre 5 ml de medio LB líquido conteniendo el antibiótico adecuado.

Se cultiva durante toda la noche con agitación, a 37°C.

Para la obtención del DNA plasmídico se sigue el siguiente protocolo:

- (1) Se recuperan 1,5 ml del cultivo (que debe estar saturado), en un tubo de microcentrifuga.
- (2) Se centrifuga a 12000 x g por 30 segundos en la microcentrifuga.
- (3) Se descarta el sobrenadante, eliminando al máximo posible las gotas de medio que queden en las paredes.
- (4) Se resuspende el *pellet* en 100 µl de Solución I.
- (5) Se repiten los pasos 2, 3 y 4.
- (6) Se agregan 200 µl de Solución II. Se mezclan las soluciones por inversión suave, 5 veces, cerrando bien el tubo. Se debe observar que el cultivo se aclara y pierde turbiedad.
- (7) Se incuba 2 minutos en agua-hielo.
- (8) Se agregan 150 µl de Solución III, preenfriada. Se agita suavemente, durante 10 segundos, asegurando que la solución se homogenice.
- (9) Se incuba por 5 minutos en agua-hielo.
- (10) Se centrifuga durante 5 minutos en microcentrifuga a máxima velocidad.
- (11) Se transfiere el sobrenadante a otro tubo eppendorf.

(12) Se agregan 2 volúmenes de etanol 96%, frío y se mezcla bien por inversión.

(13) Se incuba por 5 minutos a temperatura ambiente.

(14) Se centrifuga durante 5 minutos en microcentrifuga a máxima velocidad.

(15) Se vuelca el sobrenadante, eliminando al máximo posible las gotas de líquido que queden en las paredes.

(16) Se lava el *pellet* por agregado de 700 µl de etanol 70% y se repite el paso 15.

(17) Se resuspende el precipitado en 20 µl de H₂O.

(18) La preparación de DNA se analiza por electroforesis en gel de agarosa al 0,8-1%.

Midi-preparaciones de plásmidos.

Cuando debieron prepararse cantidades mayores de plásmidos se partió de cultivos líquidos de, al menos, 100 ml (Birnboim & Doly, 1979). La extracción se realizó por el mismo procedimiento que el usado en las minipreparaciones. El plásmido resultante, contaminado con oligonucleótidos de RNA o restos de DNA de alto peso molecular se purificó por ultrafiltración en geles de Sephacryl S-1000 (Pharmacia, Uppsala, Suecia)

Maxi-preparaciones de plásmidos.

Para preparar grandes cantidades de plásmidos con un alto grado de pureza, se siguió el protocolo de lisis alcalina y ultracentrifugación en gradientes de CsCl descrito por Gorman (1986).

A cultivos en gran volumen (al menos 500 ml) y en fase logarítmica ($DO_{600} = 0,5$) de los clones de *E. coli* que llevan el plásmido a purificar, se les agrega cloramfenicol (concentración final 170 µg/ml) con el fin de detener la replicación del DNA cromosomal sin interferir con la replicación del plásmido. Al cabo de 24 horas de crecimiento a 37°C con agitación (200 rpm en un agitador orbital) el cultivo se centrifuga 10 minutos a 3000 x g.

El *pellet* de bacterias obtenido se resuspende en 1/100 del volumen de cultivo de una solución que contiene 50 mM Tris-HCl pH 8, 50 mM glucosa y 5 mM EDTA, y se incuba durante 5 minutos a temperatura ambiente. Luego se agregan 2 volúmenes (respecto al buffer de resuspensión) de 0,2 M NaOH y 1 % SDS para producir la lisis celular. Al cabo de 5 minutos de incubación se agregan 1,5 volúmenes (respecto al buffer de resuspensión) de solución 3 M acetato de potasio pH 5,3 y se extiende la incubación otros 10 minutos. Al cabo de ese tiempo, se centrifuga la mezcla de lisis durante 15 minutos a 10000 x g. Se recoge el sobrenadante, se agregan 2,5 volúmenes (respecto del sobrenadante) de etanol y se incuba a -70°C durante 30 minutos. Luego se centrifuga durante 30 minutos a 10000 x g y el sedimento obtenido, después de secado por vacío, se resuspende en 5 ml de TE.

Esta solución incluye DNA plasmídico superenrollado, DNA relajado y linearizado, RNA y proteínas. Para separarlos, se utilizó un gradiente de densidad a través en CsCl.

Análisis de ácidos nucleicos mediante electroforesis

Electroforesis en geles de agarosa

La electroforesis en geles de agarosa permite la separación de las moléculas de ácidos nucleicos de acuerdo a su forma y su tamaño. Mediante esta técnica no es posible determinar, el peso molecular de un ácido nucleico cerrado circularmente, como es el caso de un DNA plasmídico, debido a que su migración en el gel es altamente dependiente de la forma de la molécula. Sin embargo, sí es posible determinar el peso molecular de las moléculas de ácidos nucleicos lineales. En particular, la electroforesis en geles de agarosa nativos se hace independiente de la forma de la macromolécula en el caso de las moléculas de ácidos nucleicos lineales de doble cadena. Justamente, este es el caso cuando se determina el peso molecular de fragmentos de DNA

obtenidos por digestión con enzimas de restricción o por amplificación enzimática por PCR. En cambio para realizar la determinación del peso molecular de un ácido nucleico de cadena simple, como por ejemplo un mRNA, es necesario utilizar una electroforesis en geles de agarosa en condiciones desnaturalizantes.

Análisis de DNA o RNA en geles de agarosa en condiciones nativas

Para los geles nativos, la solución de agarosa puede ser preparada en un medio TAE o TBE, sin agregados o conteniendo bromuro de etidio. Nosotros utilizamos TAE 1X con bromuro de etidio en el gel hasta una concentración de 0,16 µg/ml. La concentración de agarosa debe ser la necesaria para separar los fragmentos de DNA obtenidos. Existe una relación lineal entre el logaritmo de la movilidad electroforética del DNA (m) y la concentración del gel de agarosa (t), que se representa según la ecuación:

$$\log m = \log m_0 - K_r \cdot t$$

Donde m_0 es la movilidad electroforética libre del DNA y K_r es una constante, denominada coeficiente de retardo, que está relacionada con las propiedades del gel y el tamaño y la forma de las moléculas de DNA. En la **Tabla 16.a** se muestran las concentraciones de agarosa habitualmente utilizadas para separar fragmentos de DNA de diferentes tamaños.

Nosotros utilizamos, para resolver fragmentos de DNA con tamaños de más de 5 kb, agarosa 0,7%; para fragmentos con tamaños entre 1 kb y 5 kb, agarosa 1%; para fragmentos con tamaños entre 200 bases y 1,5 kb, agarosa 2,5%; y, para fragmentos con tamaños menores a 300 bases, agarosa 4%. En este último caso en particular, la agarosa utilizada debe ser de alto grado de pureza como es el caso de la agarosa NuSieve (FMC).

La solución de agarosa se funde en un horno a microondas o en un baño María, y, luego, se la deja tomar una temperatura de aproximadamente 60°C.

Tabla 16.a. Rango de separación en geles conteniendo diferentes cantidades de agarosa.

Concentración de agarosa en el gel %(p/v)	Rango de tamaños para la separación eficiente de moléculas de DNA lineal (kb)
0,3	5-60
0,6	1-20
0,7	0,8-10
1,0	0,5-7
1,5	0,4-6
2,0	0,2-3
2,5	0,1-2
4,0	0,05-1

Mientras tanto se arma el molde para el gel, utilizando el portageles y el peine para formar las fosas de siembra.

Se coloca la agarosa fundida en el molde y se deja a temperatura ambiente hasta que gelifique (aproximadamente media hora).

El gel se monta con el portageles, en la cuba de electroforesis, la que se llena con buffer de corrida, TAE 1X.

Las muestras conteniendo el DNA o RNA a ser analizado se mezclan con la solución de siembra correspondiente, y se siembran en cada una de las fosas. En una de las fosas de los extremos se siembra el DNA o el RNA que se utiliza como marcador de peso molecular.

Una vez conectados los bornes de la cuba a una fuente de poder, la electroforesis se corre a 70-100 V durante una hora.

Finalizada la corrida, se desconecta el aparato de electroforesis y el gel de agarosa se observa por fluorescencia de luz UV (El bromuro de etidio, intercalado entre las cadenas del DNA, fluoresce con luz naranja). Nosotros utilizamos un transiluminador Fotodyne de luz UV de 310 nm.

Las soluciones conteniendo bromuro de etidio son potenciales cancerígenos por lo que fueron manipuladas utilizando guantes de cirugía y protecciones especiales. Para la

protección contra la luz UV se utilizaron máscaras de policarbonato.

Alternativamente, si se resuelven en los geles moléculas marcadas radioactivamente, es posible detectar su ubicación después de la corrida electroforética por exposición a pantallas radiográficas y revelado fotográfico (Laskey, 1980).

Análisis de RNA en geles de agarosa en condiciones desnaturalizantes

El RNA se analizó por electroforesis en geles de agarosa en condiciones desnaturalizantes usando ya sea 10 mM de hidróxido de metilmercurio tanto en la muestra como en el gel (Bayley and Davidson, 1976), glioxal o formaldehído (Lehrach *et al.*, 1977; Wicks, 1986). Los mejores resultados se obtuvieron utilizando formaldehído o hidróxido de metilmercurio como agente desnaturalizante. Por ello, y considerando, además, la alta toxicidad del hidróxido de metilmercurio, se eligieron los geles de formaldehído para realizar el análisis de las muestras de RNA. La técnica se realizó según el siguiente protocolo:

- (1) Se funde la cantidad apropiada de agarosa en H₂O, se la deja enfriar a 60°C y se agrega el buffer de corrida para geles de formaldehído a una concentración final de

1X y el formaldehído a una concentración final de 2,2 M. El gel se prepara dentro de una campana y se lo deja enfriar durante 30 minutos.

(2) Las muestras se preparan mezclando:

RNA (hasta 10mg)	2 ml
Buffer de corrida 5X	1 ml
formaldehído	2 ml
formamida	5 ml

(3) Las muestras se incuban durante 15 minutos a 65°C.

(4) Las muestras preparadas se mezclan con 2 µl de buffer de siembra para geles de formaldehído.

(5) Se precorre el gel, sumergido en buffer de corrida para geles de formaldehído, durante 15 minutos a 70V.

(6) Luego se siembran las muestras y el gel se corre durante 2 horas a 50V.

Los geles se tiñen con bromuro de etidio (0,5 mg/ml en agua) durante 15 minutos y se lavan dos veces en 0,5 M acetato de amonio durante 15 minutos para disminuir la fluorescencia de fondo. El RNA se observa por fluorescencia usando un transiluminador Fotodyne de luz UV de 310 nm. Alternativamente, el RNA se puede teñir directamente en el buffer de muestra antes de la electroforesis por el procedimiento de Sheiness and Sullivan (1986).

Análisis de ácidos nucleicos por hibridación

Northern blot

La hibridación del RNA después de la electroforesis en geles, es utilizada para analizar la composición de los mismos. La técnica de *northern blot* se utilizó en esta tesis para identificar y caracterizar diferentes especies de RNAs, entre ellos los RNAs genómicos y los mRNAs virales (Denis *et al.*, 1988).

(1) Una vez realizada la electroforesis y fotografiado el gel, se eliminan las partes

innecesarias del gel y se transfiere el gel a una bandeja.

(2) El gel se lava 5 veces con 50 ml de H₂O durante 5 minutos cada vez y con agitación suave.

(3) Si los fragmentos de RNA son grandes (más de 2 kb), se realiza un tratamiento con una solución de 0,05 M NaOH. Se mantiene agitando durante 5 minutos.

(4) Se elimina la solución, se lava el gel con H₂O y se agrega una solución de SSC 20X. Se mantiene agitando durante 10 minutos.

(5) Se realiza la transferencia mediante flujo de *buffer* (el filtro debe estar mojado previamente en el *buffer* de transferencia):

(6) Se transfiere el RNA a la membrana de nitrocelulosa o nylon durante toda la noche. El buffer de transferencia es SSC 20X.

(7) Una vez realizada la transferencia, se lava la membrana en SSC 6X para eliminar restos de agarosa, y se la deja secar sobre papel de filtro 3MM durante 5-10 minutos a temperatura ambiente.

(8) Para fijar el RNA a la membrana se hornea durante 2 horas a 80°C. Para ello la membrana se cubre con dos papeles 3MM.

Alternativamente, se puede fijar exponiendo a la membrana a la luz UV (310 nm) durante 1 minuto (Kandijan, 1986).

Southern blot

La hibridación de fragmentos de DNA resueltos por su tamaño después de electroforesis en geles, inmovilizados sobre membranas, es utilizada para analizar la composición de los mismos (Denhardt, 1966). La técnica de *Southern blot* se utiliza para confirmar la secuencia del fragmento de DNA clonado en los plásmidos recombinantes después de los procedimientos de ligación y transformación, o para verificar la identidad de los fragmentos de DNA amplificados por PCR.

- (1) Una vez realizada la electroforesis y fotografiado el gel, se eliminan las partes innecesarias del gel y éste se transfiere a una bandeja.
- (2) Se trata con 0,25 M HCl. Se incuba a temperatura ambiente 10 minutos a partir del momento en que el azul de bromofenol vira al amarillo.
- (3) Después de descartar el HCl se lava con agua bidestilada.
- (4) Se descarta el agua y se agrega una solución de 0,5 M NaOH + 1,5 M NaCl. Se mantiene agitando durante 15 minutos.
- (5) Se elimina la solución y se agrega una solución de 0,5 M Tris·HCl pH 7,5 + 3 M NaCl. Se mantiene agitando durante 15 minutos.
- (6) Se prepara el dispositivo de transferencia esquematizado en la sección anterior (la transferencia del DNA a la membrana de nitrocelulosa o nylon toma toda la noche).
- (7) El buffer de transferencia es 6X SSC.
- (8) Una vez realizada la transferencia, se tiñe nuevamente el gel para comprobar si la transferencia fue eficiente.
- (9) Se deja secar la membrana sobre papel de filtro 3MM durante 5-10 minutos a temperatura ambiente.
- (10) Para fijar el DNA a la membrana se hornea durante 2 horas a 80°C o se expone a la luz UV de la misma manera que los filtros de RNA.

Southern blot de fragmentos pequeños

Para identificar por hibridación fragmentos de DNA menores de 300 pb se siguieron las recomendaciones de Sambrook *et al.* (1989) en la transferencia.

En particular:

- (1) El gel donde se resuelven las muestras se transfiere a una membrana de nitrocelulosa (NC) con poro de 0,2 μm (en lugar de los comunes de 0,45 μm).
- (2) Se utiliza 20X SSC como buffer de transferencia.

- (3) Una vez seco, el filtro se prehibrida a 42°C durante 2 horas la solución de prehibridación y se hibrida a 42°C en la solución de hibridación conteniendo sulfato de dextrano 5 % durante 16 hs.

Dot blot o slot blot

La técnica de *dot blot* es utilizada para detectar la presencia de secuencias complementarias a la sonda utilizada en una muestra sin dar ninguna información acerca del estado de esas secuencias.

La técnica se realiza sembrando directamente la muestra sobre la membrana de nitrocelulosa o nylon. La siembra puede ser en un punto (*dot blot*) o formando una línea sobre la membrana (*slot blot*). Las muestras pueden ser sembradas crudas o purificadas.

Una vez sembradas las muestras, la membrana se trata de la misma manera que en el caso del *northern blot* o *Southern blot* para fijar los ácidos nucleicos.

Revelado de secuencias nucleotídicas por hibridación.

El revelado consta de cuatro etapas. Estas etapas (prehibridación, hibridación, lavados y detección de los híbridos), son esencialmente las mismas tanto para *northern blot* como para *Southern* y *dot blot*.

PREHIBRIDACIÓN

Como bloqueante se utiliza DNA de esperma de salmón, que se desnaturaliza por calentamiento a 100°C durante 10 minutos. Se agrega el DNA desnaturalizado a la solución de prehibridación, se mezcla y se coloca en un recipiente plano junto con el filtro. Alternativamente, el filtro con la solución de prehibridación pueden sellarse en un sobre de polietileno. Se incuba 0,5-4 horas a 42°C. El volumen de solución utilizada debe ser de al menos 20-100 ml/cm² de filtro.

HIBRIDACIÓN

- (1) La sonda se desnaturaliza por hervor durante 10 minutos a 100°C y,

posteriormente se la enfría bruscamente en agua-hielo.

- (2) La sonda desnaturalizada se agrega a la solución de hibridación, a una concentración final de 100 a 200 ng/ml.
- (3) El filtro se embebe en la solución de hibridación. Se deben usar al menos 20-100 ml de solución por cada cm² de filtro.
- (4) Finalmente, se incuba durante toda la noche a 42°C si el *buffer* de hibridación contiene formamida, o a 65°C si no contiene agentes desnaturalizantes.

LAVADOS DE LA MEMBRANA

Una vez finalizada la hibridación, la membrana se lava según la siguiente secuencia:

- (1) SSC 2x/SDS 0,1% p/v, 3 minutos a temperatura ambiente.
- (2) SSC 0,2x/SDS 0,1% p/v, 3 minutos a temperatura ambiente.
- (3) SSC 0,16x/SDS 0,1% p/v, 15 minutos a 50°C.

Todos los lavados se realizan en un volumen de 50-100 ml.

DETECCIÓN DE LOS HÍBRIDOS SOBRE LA MEMBRANA

El sistema de detección depende de la sonda utilizada. Si la membrana es hibridada con sondas radioactivas, se expone durante toda una noche con una película radiográfica, a -70°C, con pantalla amplificadora. En los casos en que el nivel de detección es insuficiente, la membrana se expone durante, al menos, 3 días en las mismas condiciones.

En nuestros experimentos, las membranas hibridadas con sondas biotiniladas fueron reveladas siguiendo el protocolo de BluGene (BRL, USA). Brevemente, las membranas se bloquearon con una solución conteniendo albúmina, y, la membrana fue revelada con un reactivo que contenía estreptavidina (SA) unida covalentemente a una fosfatasa alcalina (PA).

La SA se une específicamente a la biotina presente en las sondas, por lo que la PA queda fija en esos puntos. Entonces, los sitios donde hibrida la sonda se ponen en evidencia por un sistema de desarrollo de color, que detecta los complejos muestra/sonda-biotina/SA-PA. En la reacción de desarrollo de color, interviene la PA, que cataliza la hidrólisis del grupo fosfato del reactivo *5-bromo-4-cloro-indolil-fosfato* (BCIP), soluble e incoloro, y permite que este reaccione con el reactivo *nitro blue tetrazolium* (NBT) convirtiéndose en un producto insoluble y azul.

Técnicas de clonado de DNA

Digestión de DNA con enzimas de restricción

La cantidad de unidades enzimáticas que se agrega al tubo de reacción debe ser calculada según la siguiente proporción: 1U enzimática/mg de DNA/hora de digestión.

Un protocolo típico es como el siguiente:

DNA plasmídico 1 µg/µl	20 µl
Buffer de restricción 10X	5 µl
Enzima de restricción 10 U/µl	1 µl
Agua	24 µl

Se incuba durante 2 horas a 37°C.

Cada enzima tiene un buffer determinado que se debe tener en cuenta en cada incubación.

La calidad de las reacciones de digestión se analiza por electroforesis en geles de agarosa nativos. La concentración de los fragmentos de DNA, se estima, en forma aproximada, por comparación con la intensidad de fluorescencia producida después de la electroforesis por los fragmentos presentes en patrones de PM de concentración conocida.

Ligación de los productos de la digestión

Para estimar la cantidad de cada DNA que se agrega al tubo de ligación se debe tener en cuenta la relación molar entre los fragmentos a ser ligados. Es probable que la relación molar

óptima sea 1:1. Sin embargo, dependiendo del método de selección de clones recombinantes utilizado, es posible que la autoligación de alguno de los fragmentos influya en el número de clones de colonias bacterianas obtenidas conteniendo el genotipo buscado, por lo que la relación óptima debe ser ajustada experimentalmente. Si uno de los fragmentos es un vector plasmídico, por regla general se utilizan 5 ng del mismo en la reacción frente a diferentes relaciones molares del otro fragmento.

Un protocolo típico es como el siguiente:

	Vol
Fragmento 1 (X ng)	2 µl
Fragmento 2 (Y ng)	2 µl
Buffer de ligación 10X	1 µl
ATP 10 mM (1 mM final)	1 µl
Agua csp 10 µl	3 µl
DNA ligasa del fago T4 (0,1U Weiss/µl)	1 µl

La temperatura óptima de ligación es de 16°C. La incubación se mantiene durante 1 a 16 horas. Para la ligación de extremos romos se utiliza ATP a una concentración final de 0,5 mM y 5 veces más de DNA ligasa.

Es importante tener en cuenta la posible controversia generada en cuanto al número de unidades de DNA ligasa agregadas al tubo de reacción, debido a que diferentes compañías comerciales utilizan distintas formas de dosar las mismas. La mayoría de las compañías utilizan las unidades Weiss, que es la cantidad de enzima que cataliza el intercambio de 1 nmol de ³²P del pirofosfato al γ-ATP en 20 minutos a 37°C. Una unidad Weiss corresponde a 67 unidades de extremos cohesivos, definida por la compañía New England Biolabs.

Preparación de células competentes (Método de cloruro de calcio)

Se utilizaron bacterias *Escherichia coli* DH5a, SURE o XL1-Blue. Las células bacterianas que se utilizan para ser transformadas con los plásmidos son llevadas a

un estado de competencia que facilita la incorporación de DNA. En nuestro laboratorio esto se realizó siguiendo el método de Hanahan (1983).

Para ello, las bacterias se sembraron en 2 ml de medio SOB y se incubaron a 37°C hasta que alcanzaron la fase estacionaria de crecimiento (durante toda una noche).

El cultivo fue diluido agregando 200 ml de medio SOB fresco y se continuó la incubación a 37°C hasta llegar a la fase logarítmica media del crecimiento. La densidad óptica (DO) inicial de este cultivo, medida a 500 nm, fue de 0,1. Cada 15 minutos se tomaron alícuotas de 1ml para medir DO_{500nm}. La DO_{500nm} del cultivo alcanza valores de 0,6 a 0,9 cuando el cultivo se encuentra en fase logarítmica de crecimiento. El crecimiento se detuvo, colocando el cultivo en baño de hielo, cuando la DO_{500nm} llegó a 0,8. La DO_{500nm} final puede variar de acuerdo a la cepa bacteriana utilizada.

Las bacterias fueron recolectadas por centrifugación en un rotor SS34 de una centrífuga Sorvall, a 10000 rpm durante 5 minutos a 4°C.

Las células se resuspendieron en Tfb I a 0°C (40 ml por cada 100 de cultivo), se dejaron 5 minutos en hielo y luego se centrifugaron 5 minutos a 10000 rpm. Por último, las bacterias se resuspendieron en 4 ml de Tfb II y se mantuvieron en un baño de hielo durante 15 minutos. Las bacterias competentes se mezclaron con glicerol estéril al 15% de concentración final y fueron mantenidas en una congeladora a -135°C hasta su utilización.

Transformación de células competentes

Para estimar la eficiencia de transformación de las bacterias competentes, éstas se transforman con cantidades conocidas de un DNA plasmídico circular cerrado y superenrollado, purificado mediante gradiente de CsCl. En particular, nosotros utilizamos 0,1, 1, 10 y 100 ng de DNA del plásmido pUC9

para transformar una muestra representativa del conjunto de bacterias competentes producidas. La eficiencia de transformación se expresa como el número de bacterias transformadas por cada mg de DNA plasmídico utilizado.

Existen diversos protocolos para la transformación de células (Sambrook *et al.*, 1989; Golub, 1988; Felgner *et al.*, 1987). En nuestro laboratorio se utiliza el siguiente:

- (1) Se colocan 50 μ l de bacterias competentes en tubos eppendorf previamente enfriados en hielo.
- (2) Se agrega el DNA a las bacterias competentes.
- (3) Se deja la mezcla durante 30 minutos en hielo.
- (4) Los tubos se mantienen durante 2 minutos a 42°C (ésta es la etapa crítica de la transformación).
- (5) Se pone la mezcla durante 2 minutos en hielo.
- (6) Se toma una alícuota para plaqueo de células viables.
- (7) El resto de la mezcla se incuba a 37°C durante 1 hora.
- (8) Se siembra la mezcla de transformación en cajas de Petri con medio rico conteniendo los componentes que permitan realizar la selección fenotípica de los recombinantes.

Caracterización de los clones recombinantes

Para caracterizar e identificar los clones recombinantes que llevan el inserto de cDNA buscado se han descrito una gran variedad de métodos. Entre ellos, se utilizan métodos basados en la velocidad de crecimiento (Karlovsky and Prell, 1987) o, más comúnmente, métodos basados en la detección de la secuencia de interés por hibridación (Grunstein and Hogness, 1975). En nuestro laboratorio se utilizan más frecuentemente sistemas de clonado que permiten detectar los clones recombinantes por medio de resistencia a antibióticos y mediante el ensayo de α -

complementación. Estos sistemas utilizan plásmidos que llevan, además de un gen que confiere resistencia a algún antibiótico, la región codificante de los 146 aminoácidos del extremo amino del gen de la β -galactosidasa de *E. coli* (*lacZ*). El sitio de clonado de estos plásmidos está insertado dentro de la región codificante mencionada. De esta manera, si se inserta alguna secuencia en el vector, el marco de lectura de la región codificante del *lacZ* se interrumpe mientras que la del gen que confiere resistencia a un antibiótico, permanece. Estos vectores se utilizan para transformar bacterias que lleven solo la región carboxiterminal del gen de la β -galactosidasa (por ejemplo, genotipo *lacZ* Δ M15). Solamente la combinación de los dos fragmentos de la β -galactosidasa (el del plásmido y el de la bacteria) posee actividad biológica, por lo que, por medio de un ensayo fenotípico, se puede discriminar entre las bacterias transformadas con plásmidos con inserto (resistentes al antibiótico, β -galactosidasa [-]) o sin inserto (resistentes al antibiótico, β -galactosidasa [+]). El medio para selección de recombinantes utilizado en este caso contiene un reactivo que produce un precipitado azul en el caso de un fenotipo β -galactosidasa [+].

Preparación de sondas de ácidos nucleicos

Síntesis de los nucleótidos radioactivos

Los nucleótidos radiactivos, utilizados en las reacciones de marcación de ácidos nucleicos, se sintetizaron a partir del 32 [P] H_3PO_4 , de acuerdo a la técnica de Walseth & Johnson (1979), modificada.

Esta técnica, puesta a punto en el laboratorio por P.D. Ghiringhelli, consta de cuatro etapas, que son descritas en detalles en el **Capítulo 17** (Ghiringhelli *et al.*, 1985).

Preparación de sondas de DNA

Para los experimentos de *Southern blot* se marcaron los fragmentos de DNA a utilizar como sonda con bio-11-dATP por la técnica de

nick translation (Rigby *et al.*, 1977) o con $\alpha^{32}\text{P}$ dATP por la misma técnica o por extensión de oligonucleótidos (Feinberg and Vogelstein, 1983). En otros experimentos se utilizó la técnica de marcación del extremo 5' por intercambio de fosfato utilizando $\gamma^{32}\text{P}$ ATP.

Los protocolos señalados en las 2 secciones siguientes indican la preparación de sondas radioactivas, pero también pueden ser utilizados para producir sondas no radioactivas, reemplazando el nucleótido radioactivo por biotina-11-dATP.

Preparación de sondas de DNA por *nick-translation*.

En esta reacción se utilizan las enzimas DNA polimerasa I (DNA pol I) y DNasa I de *E. coli*. La DNA pol I adiciona nucleótidos al extremo OH-3' del DNA molde merced a su actividad polimerasa 5'-3'. En la reacción, este extremo es generado cuando una de las cadenas de una molécula de DNA de cadena doble es hidrolizada por la DNasa I. Además, la DNA pol I puede eliminar nucleótidos desde el extremo 5' del corte merced a su actividad exonucleasa 5'-3'. Por ello, una vez generado un corte por la DNasa I, el resultado de la suma de las dos actividades enzimáticas de la DNA pol I, es el movimiento del corte a lo largo del DNA molde (*nick-translation*, Rigby *et al.*, 1977). La siguiente es una reacción típica:

(1) Para preparar una sonda radioactiva se mezclan en un tubo de reacción:

10 X <i>nick-translation buffer</i>	5 μl
DNA (1 $\mu\text{g}/\mu\text{l}$)	1 μl
3 dNTPs (-dATP) 1 mM	1 μl
$\alpha^{32}\text{P}$ dATP (400 Ci/mmol)	100 pmoles (~40 μCi)
H ₂ O csp	48 μl

(2) Se mantiene la mezcla a 0°C

(3) Se agrega 0,5 μl de una solución conteniendo DNasa I (0,1 $\mu\text{g}/\mu\text{l}$, en

glicerol 50%). Se mezcla utilizando un vortex.

(4) Se agregan 5 unidades Richardson (Richardson *et al.*, 1964) de DNA pol I. Se mezcla.

(5) Se incuba 60 minutos a 16°C.

(6) Se frena la reacción mediante el agregado de 2 μl de 0,5 M EDTA.

Preparación de sondas de DNA por extensión de oligonucleótidos.

Se utilizó la técnica de cebado al azar (*random priming*) (Feinberg and Vogelstein, 1983).

En esta técnica, la mezcla de reacción contiene *primers* que actúan como iniciadores de la síntesis de DNA en sitios del molde elegidos al azar. Estos *primers* (*random primers*) son oligonucleótidos de secuencia azarosa de 6 nucleótidos de longitud (pdN₆).

En nuestros experimentos, se utilizaron entre 100 ng y 1mg de DNA molde purificado. Previamente a la marcación, los fragmentos de DNA generados por restricción se purificaron mediante electroforesis en gel de agarosa de bajo punto de fusión.

(1) Para preparar una sonda radioactiva se mezclan en un tubo de reacción:

10 X buffer	2,5 μl
DNA (0,1-1 $\mu\text{g}/\text{ml}$)	1 μl
<i>primers</i> pdN ₆ (100 ng/ml)	1 μl
3 dNTPs (-dATP) 1mM	1 μl
$\alpha^{32}\text{P}$ dATP (400 Ci/mmol)	100 pmoles (40 μCi)
H ₂ O csp	25 μl

(2) Se mantiene la mezcla a 0°C.

(3) Se agregan 5 unidades del fragmento Klenow de la DNA polimerasa I de *E. coli* (Promega Corp. Madison Wisconsin, USA).

(4) Se incuba 90 minutos a 25°C.

(5) Se frena la reacción mediante el agregado de 1 μl de 0,5 M EDTA.

Preparación de sondas de RNA.

Las ribosondas se prepararon mediante una reacción de transcripción *in vitro*, utilizando una RNA polimerasa (Melton *et al.*, 1984). Los plásmidos utilizados como molde de la transcripción (pBS) poseen sitios promotores para las RNA polimerasas de los fagos T7 y T3. Así, el fragmento de DNA insertado adyacente a esos sitios puede ser copiado en RNA. Previo a la transcripción los plásmidos fueron linealizados con enzimas de restricción. Para producir una sonda de RNA uniformemente marcada, se mezclan en un tubo:

(1) (mantener el orden de agregado señalado)

H ₂ O csp	25 µl
buffer de transcripción 5X	5 µl
DNA (1 µg/µl)	2 µl
DTT 0,1 M	2,5 µl
3 NTPs (-ATP) 2,5 mM	5 µl
ATP 0,1 mM	2 µl
α ³² [P]ATP (400 Ci/mmol)	100 pmoles (40 µCi)

(2) Se mantiene la mezcla a 0°C

(3) Se agregan 25 unidades de RNasin y 10 unidades de RNA polimerasa (T7 oT3).

(4) Se incuba durante 60 minutos a 37°C.

(5) El RNA se purifica y la reacción se congela a -20°C.

En el caso de la producción de transcriptos no marcados fríos, se eliminan todas las líneas del protocolo donde se agregan nucleótidos y se las reemplaza por la siguiente:

4 NTPs (2,5 mM cada uno)	5 µl
--------------------------	------

Incorporación de marca por transferencia de fosfatos al extremo 5'.

Esta técnica se utiliza comúnmente para marcar oligonucleótidos. La incorporación del fosfato radioactivo al extremo 5' de los oligonucleótidos se realiza por fosforilación con polinucleótido quinasa (Sambrook *et al.*, 1989). Para tal fin, se utilizan 30 pmoles del oligonucleótido apropiado.

(1) Para preparar el oligonucleótido radioactivo se mezclan en un tubo de reacción:

10 X PNK buffer	2,5 µl
DNA (10 pmoles/ml)	3 µl
γ ³² [P]ATP (1000 Ci/mmol)	1 pmol (100 µCi)
H ₂ O csp	25 µl

(2) Se mantiene la mezcla a 0°C

(3) Se agregan 5 unidades de polinucleótido quinasa (PNK) del bacteriofago T4 (Promega Corp. Madison, Wisconsin, USA). Se mezcla.

(4) Se incuba durante 60 minutos a 37°C.

(5) La reacción se congela a -20°C

Purificación de los ácidos nucleicos.

En esta tesis, algunas de estas técnicas, se han utilizado mayoritariamente en la purificación de ácidos nucleicos marcados. Así, algunas de las descripciones presentadas más abajo hacen hincapié en este hecho. Sin embargo, es importante tener en cuenta que los métodos delineados pueden ser utilizados en general para purificar o concentrar DNA o RNA de cualquier origen.

Precipitación alcohólica en presencia de sales

Este método se utiliza para concentrar los ácidos nucleicos.

(1) Se ajusta la concentración salina de la solución acuosa con NaAcO hasta una concentración final de 0,3M.

(2) Se agregan 2,5-3 volúmenes de etanol, o 0,8 volúmenes de isopropanol.

(3) Se incuba durante 15-60 minutos a -70°C (o más tiempo a -20°C)

(4) Se centrifuga a 14000 rpm.

(5) Se elimina el sobrenadante y el sedimento se lava con etanol 70%.

(6) Se elimina el etanol residual y se resuspende en H₂O. Es importante no secar demasiado los sedimentos de ácidos nucleicos, pues estos son difíciles de redissolver.

En los casos en los que se quiere impedir que los oligonucleótidos coprecipiten junto con el DNA, se modifica el paso 1 de esta técnica de esta manera:

- (1) Se ajusta la concentración salina de la solución acuosa con NaAc hasta una concentración final de 0,1M y se agregan 5 µg de polímero lineal de acrilamida, que actúa como coprecipitante.

Purificación de ácidos nucleicos por gradiente de densidad

Para separar las diferentes formas topoisoméricas de los plásmidos, estos se centrifugan a través de un colchón de CsCl. Para ello se disuelve 1 g de CsCl por cada ml de solución conteniendo el DNA plasmídico. A esta solución se le agrega bromuro de etidio, a una concentración final de 0,5 mg/ml, y se la somete a ultracentrifugación a 40000 rpm durante 48 hs a 18°C en un rotor #50 de Beckmann. Al cabo de ese tiempo se observan, por transiluminación con luz UV, dos bandas separadas (la superior es DNA desenrollado y la inferior es DNA superenrollado) y un *pellet* (RNA). La fracción correspondiente al plásmido superenrollado se aísla por punción del tubo. Posteriormente, el bromuro de etidio se elimina mediante sucesivas extracciones con alcohol isoamílico saturado con 5 M NaCl. Finalmente, el CsCl se elimina por diálisis contra 2000 veces el volumen de la muestra de buffer 1 X TE, con varios cambios del buffer.

En nuestros experimentos, la muestra se concentró por precipitación alcohólica, y el sedimento de DNA obtenido se resuspendió en H₂O a una concentración final de 1 µg/µl de DNA (medido como absorbancia a 260 nm) y se conservó en alícuotas de 25 µl a -70°C. Finalmente, la pureza del DNA obtenido se controló por electroforesis en geles de agarosa, sin detectarse otra forma que no fuese plásmido superenrollado. Para obtener un buen rendimiento es importante trabajar en oscuridad, para evitar la fotólisis de la muestra

durante las etapas en las que el DNA se encuentra mezclado con bromuro de etidio.

Alternativamente, se utilizó el método de centrifugación en gradientes preformados descrito por Babykin and Zinchenko (1984)

Purificación mediante columnas de exclusión molecular

El DNA o el RNA marcado radioactivamente se purificó en columnas de Sephadex G-50 (Nick columnsTM, Pharmacia, Upsala, Suecia). La purificación se realizó haciendo pasar la mezcla de reacción a través de la columna, por gravedad. Posteriormente, se agregó buffer TE pH 8.0 y se recogieron fracciones de 400 µl cada una. De cada fracción se tomaron 2 µl y se agregaron a tubos de polipropileno conteniendo 1000 µl de agua destilada. Dichos tubos se pusieron en viales para su lectura en contador de centelleo. La radioactividad de cada fracción se contó por el método de Cerenkov, permitiendo verificar una incorporación total del orden de 10⁷ cpm. Usualmente, más del 90 % de la sonda está presente en las fracciones 2 y 3:

Las sondas no radioactivas se purificaron por analogía y se concentraron por precipitación etanólica en presencia de una baja concentración salina (0,1 M NaAcO) y de un agente coprecipitante (polímero lineal de acrilamida, 5 µg).

Purificación por electroforesis en agarosa de bajo punto de fusión.

Existen diversos métodos para purificar ácidos nucleicos después de una electroforesis en geles (Vogelstein and Gillespie, 1979; Girvitz *et al.*, 1980; Maxam and Gilbert, 1977). Nosotros utilizamos el procedimiento descrito por Wieslander (1979), en el que, los ácidos nucleicos se separan mediante electroforesis en geles de agarosa de bajo punto de fusión. En estos geles, una vez finalizada la electroforesis, se visualiza el sitio de migración del ácido nucleico que se desea purificar, y se corta el

bloque de gel de agarosa conteniéndolo.

Este bloque se funde a 50°C y la agarosa se elimina mediante extracciones con solventes orgánicos. En particular, se utiliza fenol, equilibrado a pH 8, en caliente (50°C) y fenol cloroformo-isoamílico. Finalmente, el DNA o RNA se recupera por precipitación etanólica en presencia de 0,3 M NaAc.

Tratamiento con DNasa RQ1

Este tratamiento se utiliza para eliminar el DNA molde de las reacciones de transcripción *in vitro*. Una vez finalizada la reacción de transcripción se agregan a la muestra 10 unidades de DNasa RQ1 (libre de RNasa) por cada µg de DNA molde sembrado originalmente en la reacción. Se incuba durante 15 minutos a 37°C y el RNA se recupera por precipitación etanólica en de 0,3 M NaAc.

Precipitación diferencial del RNA

Para separar mezclas de DNA y RNA, se precipita el RNA, mientras el DNA se mantiene en solución, con una solución de 3 M LiCl/ 6 M urea. A la mezcla RNA/DNA se le agrega un volumen igual de la solución de LiCl/urea. La mezcla se mantiene a -20°C durante 2 horas o a 4°C durante toda la noche y después se centrifuga durante 30 minutos a 14000 rpm.

El sedimento puede ser lavado con etanol 70% y el RNA se resuspende en H₂O. El RNA obtenido por esta metodología es muy puro, pero los rendimientos son escasos.

Purificación de los oligonucleótidos marcados.

A la muestra de oligonucleótido marcado con γ -³²[P]ATP, se le agrega un volumen igual de una solución de formamida conteniendo colorantes (0,05% azul de bromofenol y 0,05% xilencianol).

La mezcla se incuba durante 5 minutos a 95°C y se purifica mediante electroforesis en un gel de poliacrilamida.

Nosotros utilizamos un minigel (BioRad) de 20 % poliacrilamida /7 M urea. La

electroforesis se realizó a 150 volts durante 1 hora y luego de esto se envolvió el gel con una película de polietileno y se lo expuso durante 3 minutos con una película radiográfica, a temperatura ambiente y sin pantalla amplificadora. En los extremos del gel se colocaron fragmentos de cinta adhesiva conteniendo puntos marcados con tinta radiactiva para el alineamiento del gel con la película radiográfica luego de la exposición. Luego, se separó la región del gel de poliacrilamida conteniendo la banda mayoritaria del oligonucleótido marcado. Para eliminar la urea, el fragmento de poliacrilamida escindido se colocó dentro de la matriz de otro minigel de poliacrilamida 6 %, sin urea, previo a la polimerización del mismo. Una vez realizada una electroforesis a 150 volts durante 15 minutos, se determinó la localización del oligonucleótido marcado en el gel de la misma manera que en el caso anterior. El nuevo fragmento de poliacrilamida escindido se eluyó con 500 µl de agua bidestilada, manteniéndolo durante la toda noche a 55-60°C. La solución se centrifugó a 14000rpm durante 15 minutos, se tomó el sobrenadante, se pasó a otro tubo y se determinó la radiactividad por Cerenkov en contador de centelleo. La incorporación total de radiactividad fue del orden de 10⁶cpm.

Síntesis de cDNA

Para la reacción de síntesis de cDNA para RT-PCR se siguieron las recomendaciones publicadas por Gubler and Hoffman (1983) y Gubler (1988). Alternativamente, el cDNA de cadena simple se sintetizó utilizando entre 1 y 2 µl del RNA como molde. El RNA se desnaturalizó por calor en presencia de 2 µM de cada primer en un volumen de 2,5 µl. La reacción de extensión se llevó a cabo en presencia de 1 µM de desoxinucleósidos trifosfatos (dNTPs); 1,5 unidades/µl de RNAsin (Promega) y 4 unidades de transcriptasa reversa obtenida del virus de la mieloblastosis aviar (AMV) (Promega) o del virus de la leucemia

murina de Moloney (MoMuLV) (New England Biolabs), en un medio conteniendo 25 mM Tris-HCl pH 7,5; 60 mM KCl; 10 mM MgCl₂; 1 mM ditiotreitól; en un volumen total de 5 µl. La reacción se incubó durante 90 minutos a 42°C (AMV) o a 37°C (MoMuLV). Parar realizar la reacción en volúmenes finales mayores (hasta 25 µl), se aumentó en forma proporcional el volumen de cada etapa.

La síntesis de cDNA también se ensayó con *primers* de secuencia al azar (pdN₆), aunque, en este caso, los resultados finales fueron menos reproducibles.

El cDNA se purificó por precipitación con 2,5 volúmenes de etanol en presencia de 0,1 M acetato de sodio y 5 µg de poliacrilamida lineal, que actúa como coprecipitante. Luego se lavó con etanol 75 %, se secó y se resuspendió en 10 µl de H₂O bidestilada.

PCR

Una alícuota del cDNA, típicamente entre 1 y 4 ml, se sembró en la reacción de PCR en presencia de los mismos *primers* utilizados en su síntesis.

Las reacciones de PCR se llevaron a cabo en un volumen final de 10 µl, utilizando 0,125 unidades de Taq DNA polimerasa de Perkin Elmer-Cetus o de Promega, 0,2 a 1,0 µM de cada primer y los buffers provistos por los fabricantes.

Para controlar la existencia de contaminaciones cruzadas, se deben homogeneizar las condiciones de preparación de las distintas muestras. Para ello, todos los componentes de la reacción de PCR de las diferentes muestras (excepto los DNAs), se mezclan en una solución general de reactivos. Normalmente, se prepara un exceso (un tubo) de solución general, y se utiliza el volumen en exceso para preparar el control negativo de la amplificación. Por ejemplo, para amplificar 9 muestras, utilizando 1 µl de cDNA, se realiza el siguiente protocolo:

(1) Se prepara la siguiente Solución general:

Buffer de PCR 10X	10 µl
MgCl ₂ (25 mM)	6 µl
4 dNTPs (2 mM cada uno)	10 µl
<i>primers</i> 1 + 2 (10 µM cada uno)	10 µl
H ₂ O csp	90 µl

En algunos casos es necesario usar aditivos, como por ejemplo:

Perfect Match (10 unidades/ml)	1 µl
BSA (1 mg/ml)	10 µl

- (2) Se mantiene la mezcla a 0°C.
- (3) Se agregan 2,5 unidades de DNA polimerasa de *Thermus aquaticus*.
- (4) Se mezcla en un vortex durante 1 minuto y se centrifuga durante 15 segundos.
- (5) Se separa la mezcla general en alícuotas de 9 µl, cada una, en tubos eppendorf de 0,5 ml de capacidad.
- (6) Utilizando otra pipeta, se agrega a cada tubo 1 µl de la solución conteniendo el DNA o cDNA a ser analizado.
- (7) Se cubre la reacción con 25 µl de vaselina liviana. En el caso de trabajar con tubos capilares, esta etapa se evita.
- (8) Las muestras se colocan en la gradilla del reactor térmico y se corre la reacción.

En nuestro caso, los ciclados térmicos se llevaron a cabo mayoritariamente en un reactor térmico de Hybaid (Teddington, Gran Bretaña), ajustando las condiciones de ciclado de acuerdo al producto a amplificar.

Una alícuota de 4 µl de la PCR fue analizada por electroforesis en geles de agarosa. Los geles se tiñeron con bromuro de etidio (0,16 µg/ml en el gel). El DNA se observó por transiluminación con luz ultravioleta.

En algunos casos, para establecer la especificidad de las bandas de amplificación observadas, los geles fueron analizados por *Southern blot*, con sondas de secuencia conocida.

Análisis del DNA por secuenciamiento nucleotídico

Obtención de DNA de cadena simple

Los plásmidos que se utilizaron para el clonado de fragmentos de DNA (pBS, KS y SK) poseen secuencias que permiten rescatarlos en forma de DNA de cadena simple. Estas secuencias corresponden a las regiones de empaquetamiento del fago M13 (Yanish Perron *et al.*, 1985). Estos fagos, están presentes en las bacterias como formas replicativas de cadena doble (igual que los plásmidos) y el genoma empaquetado que brota de la célula es de cadena simple. Para que pueda obtenerse el DNA de cadena simple representando secuencias del pBS, el plásmido debe adquirir la capacidad de replicarse como un fago, empaquetarse y brotar de la célula. Para ello, se debe infectar a las bacterias transformadas con el pBS, con un fago defectivo que aporta todos los elementos que actúan en *trans* (proteínas) necesarios (Vieira & Messing, 1987; Messing, 1983, 1993). El fago es defectivo para que no compita con éxito con el plásmido por las proteínas necesarias.

Así, de los fagos presentes en el sobrenadante de un cultivo bacteriano infectado se purifica el DNA, de acuerdo a la técnica de Messing (Messing 1983; Vieira and Messing, 1987) por precipitación de las partículas con PEG-NaAc y posterior extracción fenólica para eliminar las proteínas.

Secuenciamiento de fragmentos de DNA por el método de Sanger.

La secuenciación de DNA se realizó mediante el método de terminación de cadenas mediada por dideoxynucleótidos (Sanger *et al.*, 1977). En el caso de la secuenciación de los fragmentos de DNA clonado en plásmidos pBS se utilizaron los oligonucleótidos universales, directo y reverso, del bacteriofago M13 (Vieira and Messing, 1982), complementarios al sitio de clonado múltiple del plásmido, y se

siguieron las recomendaciones publicadas por Zagursky *et al.* (1985).

Para la reacción de secuencia se utilizaron entre 1 y 2 µg del DNA plasmídico molde. Dichos DNAs se desnaturalizaron durante 10 minutos en presencia de 2N NaOH y posteriormente se purificaron en columna de Sephacryl S-400 (*Spun columns*, Pharmacia Upsala, Suecia).

Algunas veces, se utilizaron plásmidos sin purificar extraídos de bacterias mediante minipreparaciones, los que se desnaturalizaron por calor, previo a la reacción de secuencia.

La reacción de secuenciación se realizó utilizando el kit *Sequenase*TM versión 2.0 (United States Biochemical Corporation, Cleveland, USA).

- (1) Para la asociación de los *primers* con el DNA molde, se prepara la siguiente mezcla de apareamiento:

DNA (1 µg)	3 µl
buffer de reacción 5X	1 µl
<i>primer</i> (20ng/µl)	1 µl

- (2) La mezcla se incuba durante 2 minutos a 95°C, se deja enfriar a temperatura ambiente y se incuba durante 1 hora a la misma temperatura.
- (3) Para la reacción de incorporación de α^{32} [P]dATP se utilizan los 5 µl de la mezcla de apareamiento según el siguiente protocolo:

mezcla de apareamiento	5 µl
solución de marcación	2,7 µl
α^{32} [P]dATP (400 Ci/mmol)	4 µCi
T7DNApol (1,6 U/µl)	0,125 µl

- (4) La mezcla de marcación, así obtenida se incuba durante 2 minutos a 25°C.
- (5) Se separan cuatro alícuotas de 1,75 µl de la mezcla de marcación y se agregan a cada uno de los cuatro tubos conteniendo, separadamente, 1,25 µl de cada solución de terminación. Cada solución de

terminación contiene uno de los 4 dideoxinucleótidos.

- (6) Las reacciones de secuencia se incuban durante 5 minutos a 37°C.
- (7) La reacción se detiene por el agregado de 4 µl solución de finalización.
- (8) Se siembran 2 µl de cada reacción con cada dideoxinucleótido en calles separadas y se realiza la electroforesis en geles de 6% poliacrilamida/7 M urea (Sanger and Coulson, 1978; Sambrook *et al.*, 1989) a 80 watts y a tiempos que dependen del número de nucleótidos a leer.
- (9) Para su revelado, los geles se recubren con una película de polietileno y se ponen en contacto con una película radiográfica a -70°C durante toda la noche utilizando pantalla amplificadora.

Métodos computacionales

Gráficos cartesianos de homología

- (1) Se realizó un alineamiento múltiple utilizando el paquete de rutinas CLUSTAL V *Multiple Sequence Alignments* (Higgins and Sharp, 1988, 1989). El archivo de salida del alineamiento generado por el CLUSTAL V está formado por las secuencias y una línea con caracteres gráficos que indican la identidad (*), los cambios conservativos (·) y la ausencia de homología ().
- (2) Con las líneas de datos correspondientes a los caracteres gráficos se generó un archivo ASCII independiente, carente de cualquier formato.
- (3) Este archivo se procesó con la rutina HOMGRAPH (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), para generar un archivo ASCII numérico con valores de homología relativos.
- (4) El archivo ASCII con la tabla de datos de homología relativa fue importado en el

programa SigmaPlot para Windows v. 1.01 (Jandel Corporation), con el cual se confeccionó el gráfico definitivo.

Búsqueda de regiones codificantes

Existen diversos métodos que permiten predecir la capacidad de codificación de una determinada secuencia nucleotídica. La mayor parte de estos métodos corresponden a una de dos categorías: los métodos que buscan las señales que delimitan las regiones codificantes (búsqueda por señales) y los métodos que analizan el contenido informativo de las potenciales regiones (búsqueda por contenido) (Rice *et al.*, 1991).

Tabla 16.b. Entorno consenso de AUGs eucariotas

	Frecuencia de nucleótido										
	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4
A	81	49	34	57	170	62	43	217	0	0	59
C	58	29	86	117	4	102	119	0	0	0	31
G	23	89	35	24	41	21	37	0	0	217	87
T	54	48	62	18	2	32	18	0	217	0	38
con	A	G	t/c	C	a/g	C	C	A	T	G	G

En las secuencias eucariotas las señales que se buscan son, principalmente, codones de iniciación en un entorno adecuado, codones de terminación, sitios de *splicing*, señales de poliadenilación, etc.

Respecto de los codones de iniciación, Kozak (1978, 1983, 1984) ha demostrado que la eficiencia de traducción *in vitro* está en directa función del entorno nucleotídico del AUG. En función de esto, Kozak y otros autores han compilado y analizado el entorno de muchos AUGs eucariotas, construyendo una secuencia consenso (**Tabla 16.b**) y matrices de pesos logarítmicos que posibilitan la predicción de los codones de iniciación. Sin embargo, no debemos olvidarnos que estas predicciones deben concordar de alguna manera con los datos experimentales existentes.

En cuanto al contenido informativo, la periodicidad de tripletes impuesta por el código

genético, el uso preferencial de codones y las secuencias aminoacídicas particulares son más que suficientes para generar diferencias en las frecuencias nucleotídicas para cada una de las posiciones de los codones (Staden, 1984; McCaldon and Argos, 1988; Staden, 1990).

Ficket (1982) ha diseñado un método estadístico para determinar las regiones codificantes. Este análisis, llamado *TestCode analysis*, se basa en una desviación periódica de la composición nucleotídica en las regiones codificantes.

Sintéticamente, el método fue desarrollado utilizando 321 fragmentos de secuencias codificantes y 249 fragmentos de secuencias no codificantes, cada uno de ellos, como mínimo, de 200 nt. de longitud (el método pierde confiabilidad con segmentos mas cortos). Sobre cada segmento se aplica el siguiente algoritmo:

- (1) Para cada tipo de base (A, C, G, T) se calculan dos parámetros, la posición y el contenido. El parámetro posición se calcula teniendo en cuenta las tres posiciones en el codón. Así, para el residuo X:

$$X\text{-pos} = \text{Max}(X1, X2, X3) / \text{Min}(X1, X2, X3) + 1$$

Donde

X1 = n° de Xs en posición 1, 4, 7, 10, ...

X2 = n° de Xs en posición 2, 5, 8, 11, ...

X3 = n° de Xs en posición 3, 6, 9, 12, ...

Esto permite medir la asimetría en la distribución de la base.

- (2) El parámetro de contenido tiene en cuenta la contribución de la base X a la formación del posible marco de lectura. Para cada X se utilizan tablas de probabilidades (p1 a p8) y pesos (w1 a w8) asociados con la cantidad de veces que el residuo en cuestión predice regiones codificantes o no codificantes.
- (3) Se calcula el TestCode de acuerdo a:

$$\text{TestCode} = p1w1 + p2w2 + \dots + p8w8$$
- (4) En función del TestCode calculado se realiza la predicción (**Tabla 16.c**). Este método es muy útil y confiable para predecir la probabilidad de codificación, pero no para señalar los límites de una región codificante

En conclusión, la mejor manera de predecir marcos de lectura abiertos es utilizar una combinación de métodos de análisis de señales y de análisis de información. Obviamente, la última palabra la tienen siempre los datos experimentales.

Cálculo de LCC (complejidad local de composición)

- (1) Utilizando como archivos de entrada las secuencias nucleotídicas en texto ASCII sin formato, se calcula el perfil de LCC para cada una de ellas utilizando la rutina CALCLCC (Ghiringhelli, P.D.; descrita en el **Capítulo 18**). Esta rutina genera un archivo ASCII numérico que se puede importar con cualquier programa de graficación para su representación.
- (2) Para generar el gráfico consenso se debe trabajar con los archivos individuales de datos de LCC y con un alineamiento múltiple (generado con el Clustal X, o programa equivalente). En el archivo del alineamiento múltiple, los nombres de las secuencias individuales deben ser reemplazados por los nombres de los archivos individuales con la información de LCC. A continuación, este conjunto de

Tabla 16.c. Valores de TestCode y predicciones

TestCode	Prob.	Predic.	TestCode	Prob.	Pred.
0,32-0,43	0%	No codif.	0,84-0,95	77%	?
0,43-0,53	4%	No codif.	0,95-1,05	92%	?
0,53-0,64	7%	No codif.	1,05-1,16	98%	Codif.
0,64-0,74	29%	No codif.	1,16-1,26	100%	Codif.
0,74-0,84	40%	?	1,26-1,37	100%	Codif.

archivos es procesado con la rutina AJALNLCC (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), para generar nuevos archivos individuales, ajustados al alineamiento múltiple mediante la introducción de *gaps*. Estos archivos están escritos en texto ASCII, de manera tal que se puede importar con cualquier programa de graficación para su representación.

- (3) Por último, si se desea representar el perfil promedio y las dispersiones, se puede ejecutar la rutina STATLCC (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), utilizando los archivos del paso anterior como archivos de entrada. Nuevamente, el archivo de salida está escrito en texto ASCII, de manera tal que se pueden importar con cualquier programa de graficación para su representación.

Frecuencia relativa de nucleótidos (A+U vs C+G)

- (1) Utilizando como archivos de entrada las secuencias nucleotídicas en texto ASCII sin formato, se calcula el perfil de abundancias relativas A+U vs C+G para cada una de ellas utilizando la rutina FRENATCG (Ghiringhelli, P.D.; descrita en el **Capítulo 18**). Esta rutina genera un archivo ASCII numérico que se puede importar con cualquier programa de graficación para su representación.
- (2) Para generar el gráfico consenso se debe trabajar con los archivos individuales de datos de generados anteriormente y con un alineamiento múltiple (generado con el Clustal X, o programa equivalente). En el archivo del alineamiento múltiple, los nombres de las secuencias individuales deben ser reemplazados por los nombres de los archivos individuales con los datos de frecuencias. A continuación, este conjunto de archivos es procesado con la

rutina AJALNFRE (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), para generar nuevos archivos individuales, ajustados al alineamiento múltiple mediante la introducción de *gaps*. Estos archivos están escritos en texto ASCII, de manera tal que se pueden importar con cualquier programa de graficación para su representación.

- (3) Por último, si se desea representar el perfil promedio y las dispersiones, se puede ejecutar la rutina STATFRE (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), utilizando los archivos del paso anterior como archivos de entrada. Nuevamente, el archivo de salida está escrito en texto ASCII, de manera tal que se puede importar con cualquier programa de graficación para su representación.

Abundancia relativa de oligonucleótidos

- (1) Como archivos de entrada se utilizan las secuencias nucleotídicas en texto ASCII sin formato.
- (2) En el análisis global, se procesan las secuencias individuales completas con la rutina ABRLOLIG (Ghiringhelli, P.D.; descrita en el **Capítulo 18**). Los resultados se tabulan y/o grafican para el análisis comparativo.
- (3) En el análisis asociado con la posición, se trabaja con ventanas de longitud definida por el usuario, y se utiliza la misma rutina anterior para calcular las abundancias relativas para cada conjunto de residuos coincidente con la longitud de la ventana.
- (4) Por último, también se puede realizar el cálculo de abundancia relativa de dinucleótidos en regiones codificantes y no codificantes, por separado, utilizando la rutina ABRDINI (Ghiringhelli, P.D.; descrita en el **Capítulo 18**). Los resultados se tabulan y/o grafican para el análisis comparativo.

Manipulación de datos (Datmanip)

- (1) Esta rutina utiliza como archivos de entrada los archivos ASCII obtenidos con los programas PHDhtm, PHDsec y PHDacc (Rost *et al.*, 1994a,b, 1995; Rost and Sander, 1993a,b,c, 1994a,b,c,d), y permite tomar datos seleccionados y adecuarlos para su presentación gráfica.
- (2) Los datos se procesan con el programa DATMANIP (Ghiringhelli, P.D.; descrita en el **Capítulo 18**).
- (3) Selección de columnas y obtención de datos; cambios de signo en los valores; búsqueda de máximos y mínimos; normalización de valores a un rango entre 0 y 1 y cálculo de promedios por ventanas.
- (4) Los datos de salida, se escriben en formato ASCII, interpretables por la mayor parte de los programas de gráficos.

Análisis comparativo de estructuras secundarias de RNA (óptimas y subóptimas) y determinación de la estructura más frecuente

- (1) Esta rutina utiliza como archivos de entrada los archivos ASCII obtenidos con la rutina MFOLD (Zuker, 1981, 1989, 1994; Zuker & Sankoff, 1984; Zuker *et al.*, 1991), y exportados en formato GCG (Genetics Computer Group, Wisconsin).
- (2) Los archivos de las distintas formas (óptima y subóptimas) se procesan con la rutina COMP_CT (Ghiringhelli, P.D.; descrita en el **Capítulo 18**).
- (3) A continuación, se procesan los archivos de salida anteriores con la rutina ARM-TABL (Ghiringhelli, P.D.; descrita en el **Capítulo 18**).
- (4) Por último, utilizando una combinación de rutinas ANALIZCT, CALCPORC, MEXTCT-1, MEXTCT-2, MCONNECT, EXTR-SEQ y EXTSEQ-V (Ghiringhelli, P.D.; descrita en el **Capítulo 18**), se

rearman los distintos segmentos en un único archivo tipo GCG (*Genetics Computer Group*, Wisconsin, USA). Estos datos pueden ser representados gráficamente mediante el programa CONNECT del GCG.

Detección de puntos potenciales de recombinación en moléculas de RNA, construcción y selección de las moléculas recombinantes en base a criterios de presencia de marcos de lectura abiertos y variaciones en la longitud de las proteínas codificadas

- (1) Como primer paso se efectúa una comparación entre dos secuencias, en la misma polaridad o en polaridad complementaria, utilizando el programa Bl2seq (*Blast 2 sequences*, Karlin and Altschul, 1990, 1993; Altschul *et al.*, 1994, 1996, 1997; Zhang *et al.*, 1998), con la siguiente línea de comandos:

```
Bl2seq -i CD1-v.rnd -j Oli-vc.rnd -p
blastn -g T -o CD1vOlic.bl2 -W 7 -F F -q
-1 -e 100000 -S 1.
```

- (2) A continuación, se procesa el archivo de salida con las rutinas ANPARCOD, ARMAREC4 y REV TAB3 (Ghiringhelli, P.D.; descritas en el **Capítulo 18**), para analizar y construir las moléculas de RNA recombinantes, mediante el método de *copy choice*.
- (3) Si se desea obtener una representación gráfica global de las chances de recombinación, el archivo de salida de Bl2seq se debe procesar con las rutinas DIVIDE, ANA1 ó ANA2-1, PROCESA, PROCESA1, PROCESA2, JUNTA1, JUNTA2, JUNTA y HISTO (Ghiringhelli, P.D.; descritas en el **Capítulo 18**), para generar archivos importables en programas de graficación y representarlos como gráficos del tipo *dop plot* o histogramas.

11/5/86

Determinación de anfipaticidad en estructuras secundarias y perfiles de hidrofobicidad

- (1) Como archivo de entrada se utiliza un archivo en formato ASCII conteniendo la el datos de la estructura secundaria de la proteína (nombre, posiciones, etc.) y la secuencia aminoacídica de la misma.
- (2) Este archivo se procesa con la rutina ANFIPATI (Ghiringhelli, P.D.; descrita en el **Capítulo 18**) para generar los perfiles de amplitudes de Fourier y los datos de índice de anfipaticidad.
- (3) Si se desea obtener una información resumida, se puede procesar el archivo anterior con la rutina EXTRPEEK (Ghiringhelli, P.D.; descrita en el **Capítulo 18**).
- (4) Por último, si se desea obtener el perfil de hidrofobicidad de la proteína completa, se puede procesar un archivo ASCII con la secuencia aminoacídica total, con la rutina HIDROF (Ghiringhelli, P.D.; descrita en el **Capítulo 18**). La misma permite generar un archivo ASCII con datos de hidrofobicidad por ventanas, que puede ser importado con cualquier programa de gráficos.

Modelado de estructuras proteicas por homología con homólogos remotos

Para esta tarea se utilizó el servidor gratuito UCLA-Fisher que realiza el reconocimiento de plegamiento entre la secuencia proteica enviada y la base de datos de estructura PDB (*Protein Data Bank*).

La secuencia utilizada para el análisis corresponde a los residuos 453 a 564 de la proteína N de la cepa MC2 del virus Junín y comprende la región C-terminal de la misma, conteniendo el potencial *zinc finger*.

Con esta secuencia, el servidor predice la estructura secundaria utilizando el PHD y realiza una búsqueda en el banco PDB mediante

el programa PSI-BLAST (*position specific iterated-BLAST*).

Esta búsqueda la realiza utilizando cinco algoritmos diferentes en paralelo (gonp, gonpm, seqpprf, seqpmprf y prfseq). Para cada uno de ellos calcula el *Z-score* y arma un consenso de todos. Esencialmente, lo que se compara en esta etapa del análisis son las estructuras secundarias predichas para la proteína incógnita y las detectadas en las proteínas del PDB. Un *Z-score* igual o superior a 5 se considera significativo estadísticamente.

De esta manera se encontraron dos alineamientos significativos: 1LLD (*L-lactate dehydrogenase; Z-score 6,3*) y 1BJX (*human protein disulfide isomerase; Z-score 5,9*).

El esquema final de la posible estructura de la región C-terminal de la proteína N de la cepa MC2 del virus Junín se construyó utilizando las coordenadas de 1LLD, correspondientes a la región de homología estructural, para ubicar los residuos de la proteína N de MC2.

El modelo final tiene bastante certeza en cuanto a la disposición de los motivos de estructura secundaria tipo α -hélice y β -hoja plegada, pero es relativamente incierto en cuanto a la disposición de los *loops*.

PARTE VI

MISCELANEA

Capítulo 17

Puesta a punto, modificación y diseño de técnicas

SINTESIS ENZIMÁTICA DE $\gamma^{32}\text{P}$ Y $\alpha^{32}\text{P}$ NUCLEOSIDOS TRIFOSFATO

Los nucleósidos trifosfato marcados con isótopos radioactivos son herramientas esenciales para la investigación en bioquímica y biología molecular. En particular, los $\gamma^{32}\text{P}$ y $\alpha^{32}\text{P}$ nucleósidos trifosfato son reactivos casi imprescindibles para la mayor parte de los esquemas de clonado, secuenciamiento y caracterización de nuevos genomas.

Sin embargo, la compra de tales productos resulta bastante costosa y, algunas veces, el suministro regular de estos insumos, con isótopos de vida media muy corta, es problemático e ineficiente. Por estas razones, a fines de 1984, en el laboratorio de Química Biológica II, los Dres. Oscar Grau y Víctor Romanowski decidieron poner en marcha la síntesis enzimática de $\gamma^{32}\text{P}$ y $\alpha^{32}\text{P}$ nucleósidos trifosfato, basados en las publicaciones de Walseth & Johnson (1979) y Maia *et al.* (1982).

El proceso de síntesis enzimática se realizó siguiendo la ruta bioquímica delineada en la **Figura 17.1**.

Materiales y métodos

Los reactivos generales utilizados (para buffers, soluciones salinas, etc.) fueron comprados en SIGMA (St. Louis, EE.UU.), mientras que los reactivos específicos fueron comprados a los proveedores indicados en la **Tabla 17.a**.

Los protocolos de síntesis de $\gamma^{32}\text{P}$ ATP y distintos $\alpha^{32}\text{P}$ dNTPs ($\alpha^{32}\text{P}$ dATP, $\alpha^{32}\text{P}$ dCTP y $\alpha^{32}\text{P}$ dTTP) o $\alpha^{32}\text{P}$ ATP fueron optimizados en base a los reactivos disponibles (**Figuras 17.2, 17.3, 17.4 y 17.5**).

Es de destacar que para sintetizar $\alpha^{32}\text{P}$ CTP, $\alpha^{32}\text{P}$ dCTP y $\alpha^{32}\text{P}$ dTTP se deben utilizar los precursores adecuados (3' CMP, 3' dCMP o 3' dTMP) y, además, adicionar los fosfatos no radioactivos β y γ , en la última etapa, mediante

las enzimas nucleósido monofosfato quinasa y nucleósido difosfato quinasa, en lugar de las enzimas mioquinasa y piruvato quinasa.

El rendimiento de cada etapa de la síntesis se analizó mediante cromatografía ascendente en capa fina, en PEI-celulosa con un indicador fluorescente sobre una base de acetato.

Las placas de PEI-celulosa fueron recorridas con H_2O y secadas con una corriente de aire tibio. Se utilizaron volúmenes de muestra entre 0,3-0,5 μl . Como fase móvil se utilizó buffer KH_2PO_4 en concentraciones variables (entre 0,33 M y 0,75 M) y pH 3,5. Las posiciones relativas de los distintos compuestos radioactivos (NTP, NDP, NMP y Pi) varía en función de la molaridad del buffer. Para verificar la identidad de los distintos compuestos radioactivos, se sembraron en forma individual, en calles adicionales, muestras control no marcadas (1 μl de 10 mM ATP, 1 μl de 10 mM ADP y 1 μl de 10 mM AMP). La posición de cada compuesto fue visualizada por irradiación con luz UV.

Tabla 17.a Reactivos específicos

Reactivo	Proveedor	Nº de catálogo
<i>Reactivos químicos</i>		
^{32}P carrier free (en tricina)	NEN	NEX-053
Acido pirúvico (sal sódica). Tipo II	SIGMA	P-2256
L-cisteína (clorhidrato)	SIGMA	C-7880
2' deoxiadenosina 3' monofosfato (sal sódica)	SIGMA	D-3139
Esperrina (tetrahydroclorhidrato)	SIGMA	S-2876
Fosfoenolpiruvato (sal trisódica)	SIGMA	P-7002
L- α -glicerofosfato (sal de di(ciclohexilamonio))	SIGMA	G-7886
β -NAD (forma reducida) (sal sódica)	SIGMA	N-8129
<i>Enzimas</i>		
α -glicerofosfato dehidrogenasa. Tipo III	SIGMA	G-4381
3' fosfogliceroquinasa	SIGMA	P-7634
Gliceraldehido 3 fosfato dehidrogenasa	SIGMA	G-0763
L-láctico dehidrogenasa. Tipo II	SIGMA	L-2500
Mioquinasa. Grado III	SIGMA	M-3003
Nucleasa P1	SIGMA	N-8630
Nucleósido monofosfato quinasa	SIGMA	N-4379
Nucleósido 5' difosfato quinasa	SIGMA	N-0379
Piruvato quinasa. Tipo II	SIGMA	P-1506
Polinucleótido quinasa	BRL	800455
Triosa fosfato isomerasa. Tipo III S	SIGMA	T-2391

Las autorradiografías fueron realizadas con película X-OmatTM K de Kodak (Rochester, EE.UU.), y reveladas con revelador manual para placas radiográficas. La cuantificación de los productos fue realizada por medición con un

contador de centelleo líquido (efecto Cerenkov) sobre los trozos de PEI-celulosa con las muestras.

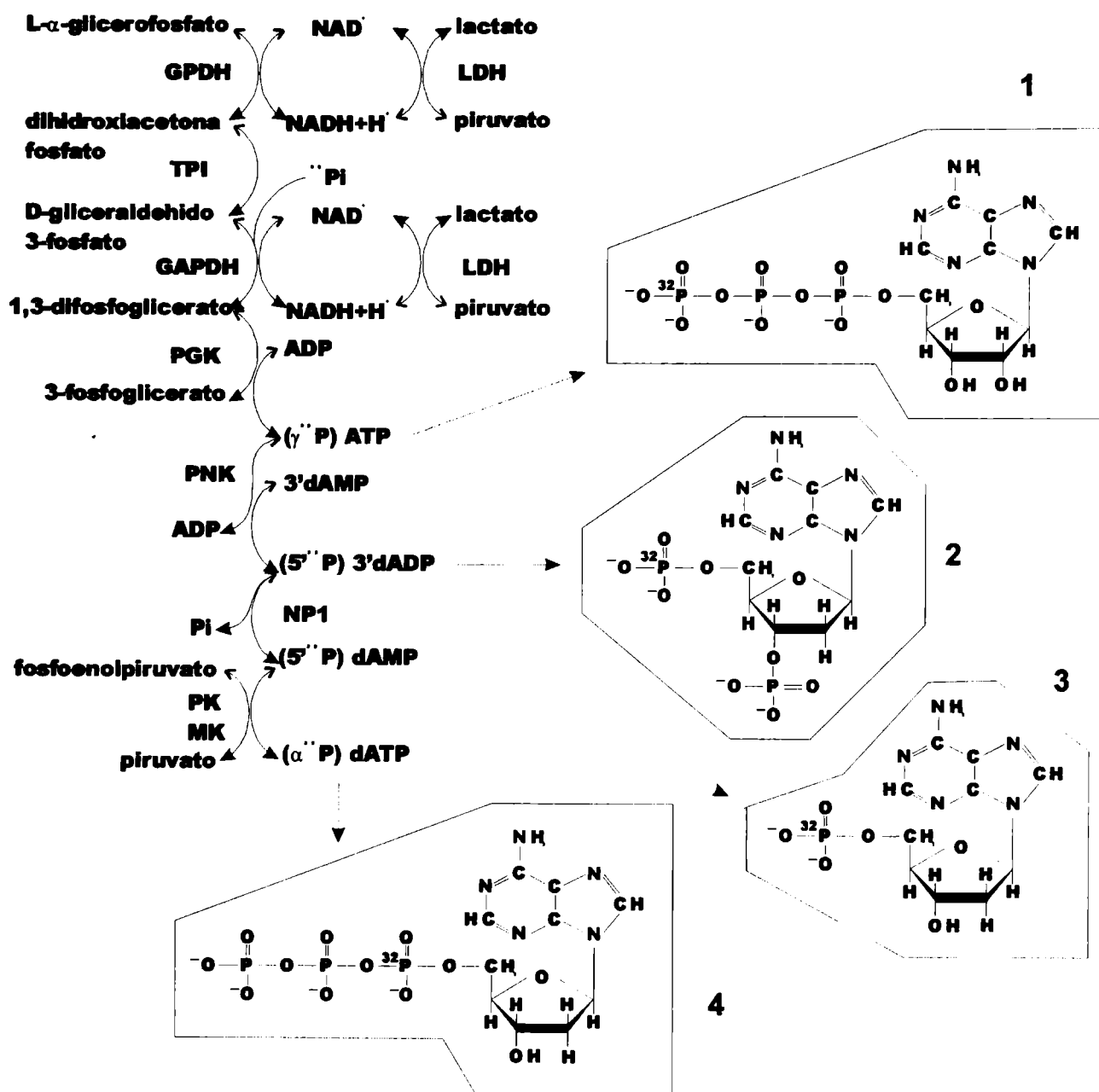


Figura 17.1. Ruta bioquímica de la síntesis enzimática de nucleósidos trifosfato marcados con ^{32}P . La producción de nucleósidos trifosfato puede dividirse en varias etapas: la síntesis de $(\gamma\text{-}^{32}\text{P})\text{ATP}$ (1), mediante un conjunto de reacciones típico de la glicólisis; posteriormente, se transfiere este ^{32}P a un precursor de tipo 3'dNMP o 3'NMP para generar un $(5'\alpha\text{-}^{32}\text{P})\text{3'dNDP}$ o $(5'\alpha\text{-}^{32}\text{P})\text{3'NDP}$ (2), se elimina el fosfato de posición 3' para generar un $(5'\alpha\text{-}^{32}\text{P})\text{dNMP}$ o $(5'\alpha\text{-}^{32}\text{P})\text{NMP}$ (3) y se adicionan fosfatos no marcados en las posiciones β y γ , para generar un $(5'\alpha\text{-}^{32}\text{P})\text{dNTP}$ o $(5'\alpha\text{-}^{32}\text{P})\text{NTP}$ (4). **GPDH:** glicerofosfato dehidrogenasa. **TPI:** Triosa fosfato isomerasa. **GAPDH:** gliceraldehido 3-fosfato dehidrogenasa. **PGK:** 3-fosfoglicerocquinasa. **PNK:** polinucleótido quinasa. **NP1:** Nucleasa P1. **MK:** mioquinasa. **PK:** Piruvato quinasa.

SINTESIS DE ($\gamma^{32}\text{P}$) ATP

Operaciones

1. Concentrar el ^{32}P i a sequedad en Speed-Vac (o equivalente).
2. Resuspender en agua, a una concentración final de 0,5-1 mCi/ μl .
3. Mezclar:

Reactivos I (ver apartado 1)	285 μl
Piruvato 4 mg/ml (en H_2O)	15 μl
^{32}P i (50 mCi) (en H_2O)	100 μl
H_2O	110 μl
<i>Tomar una alícuota de 0,3-0,5 μl</i>	
Mezcla enzimas/cisteína (ver apartado 2)	90 μl
TOTAL	600 μl

4. Incubar a temperatura ambiente 30 - 60 minutos. *Tomar una alícuota de 0,3-0,5 μl*
5. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas, usando KH_2PO_4 0,75 M pH 3,5. Secar la placa y visualizar por autorradiografía.
6. Si la incorporación es mayor del 90% inactivar las enzimas. Colocar 10 min en agua hirviendo.

Apartados

Apartado 1 – Reactivos I

Solución stock	Volumen en la mezcla	Concentración final en la reacción
500 mM Tris HCl pH 9,0	120 μl	50 mM
300 mM Mg Cl_2	48 μl	12 mM
20 mM espermina	108 μl	1,8 mM
100 mM dithiothreitol	72 μl	6 mM
2,4 mM L- α -glicerofosfato	60 μl	0,12 mM
1,54 mM ADP	39 μl	0,05 mM
10 mM NAD (o NADH)	60 μl	0,5 mM
H_2O	63 μl	
VOLUMEN TOTAL	570 μl	

Apartado 2 – Enzimas/cisteína

(a) Centrifugar 17 μl de la mezcla de enzimas (ver apartado 3), 5 min a 4°C, en microcentrífuga.	
(b) Retirar el sobrenadante y resuspender en 10 μl de Tris HCl 50 mM pH 9,0.	
Mezcla de enzimas (en Tris HCl 50 mM pH 9,0)	10 μl
H_2O	22 μl
Cisteína-HCl 57 mM (en Tris base 0,1 M)	94 μl
Tris-HCl 500 mM pH 9,0	14 μl
TOTAL	140 μl

Apartado 3 – Mezcla de enzimas

Enzimas	Volumen	Concentración final
α glicerofosfato dehidrogenasa (GPDH) (1,3 mg/ml – 250 U/mg)	70,8 μl	0,23 U/ μl
triosa fosfato isomerasa (TPI) (3,7 mg/ml – 5200 U/mg)	1,0 μl	0,19 U/ μl
gliceraldehido-3 P-dehidrogenasa (GAPDH) (26mg/ml – 88 U/mg)	13,4 μl	0,31 U/ μl
3-fosfogliceroquinasa (PGK) (2,2 mg/ml – 2050 U/mg)	3,8 μl	0,17 U/ μl
lactato dehidrogenasa (LDH) (12,3 mg/ml – 780 U/mg)	11,0 μl	1,06 U/ μl
TOTAL	100,0 μl	

Figura 17.2. Protocolo sintético para la síntesis enzimática de ($\gamma^{32}\text{P}$) ATP. La mezcla de reactivos I del *Apartado 1* es estable por largo tiempo, conservándola a -80°C , y soporta múltiples ciclos de congelación y descongelación. Uno de los stocks se ha usado durante 4 años sin perjuicios en el rendimiento final. Las distintas enzimas utilizadas en este protocolo, se comercializan como suspensiones en $(\text{NH}_4)_2\text{SO}_4$, en concentraciones específicas y en buffers *ad hoc*. Sin embargo, la mezcla de enzimas realizada mediante la simple adición de los volúmenes indicados en el *Apartado 3*, es estable por largo tiempo a 4°C . El mismo stock de mezcla se ha utilizado durante 2,5 años sin perjuicios en el rendimiento final. Por otra parte, la solución de piruvato, la solución de cisteína-HCl y la mezcla indicada en el *Apartado 2* deben ser preparadas en el momento de usar.

SINTESIS DE ($\alpha^{32}\text{P}$) ATP

Operaciones

1. Centrifugar el ($\gamma^{32}\text{P}$) ATP para decantar las enzimas inactivadas. *Tomar una alícuota de 0,3-0,5 μl .*
2. Mezclar:

($\gamma^{32}\text{P}$) ATP (25 mCi)	300,00 μl
3' AMP 18 mM (en H_2O)	30,00 μl
Tris-HCl 50 mM pH 9,0	28,50 μl
PNK (10 U/ μl)	1,50 μl
TOTAL	360,00 μl

3. Incubar a temperatura ambiente 12 - 16 horas (ó 4 hs. a 37°C). *Tomar una alícuota de 0,3-0,5 μl*
4. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 1 y 3, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
5. Si la transferencia es mayor del 95% continuar con la etapa siguiente. *Tomar una alícuota de 0,3-0,5 μl*
6. Mezclar:

(5' $\alpha^{32}\text{P}$) 3' ADP	360,00 μl
Tris-HCl 50 mM pH 8,0	75,00 μl
NPI (3,0 U/ μl) (en NaAcO 10 mM pH 5,0)	15,00 μl
TOTAL	450,00 μl

7. Incubar a temperatura ambiente 30 - 45 minutos. *Tomar una alícuota de 0,3-0,5 μl*
8. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 5 y 7, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
9. Si la hidrólisis es mayor del 95% continuar con la etapa siguiente. *Tomar una alícuota de 0,3-0,5 μl*
10. Mezclar:

Reactivos IV (ver apartado 1)	360,00 μl
Fosfoenolpiruvato 10 mg/ml (en H_2O)	45,00 μl
(5' $\alpha^{32}\text{P}$) AMP	450,00 μl
Mezcla enzimas (PK+MK) (ver apartado 2)	45,00 μl
TOTAL	900,00 μl

11. Incubar a temperatura ambiente 30 - 60 minutos. *Tomar una alícuota de 0,3-0,5 μl*
12. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 9 y 11, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
13. Si la incorporación es mayor del 95% inactivar las enzimas. Colocar 10 min en agua hirviendo.

Apartados

Apartado 1 – Reactivos IV

Solución stock	Volumen en la mezcla	Concentración final en la reacción
1 M Tris HCl pH 8,0	50 μl	55,6 mM
1 M Mg Cl ₂	10 μl	11,1 mM
1 M KCl	20 μl	22,2 mM
100 mM dithiothreitol	50 μl	5,6 mM
H_2O	770 μl	
VOLUMEN TOTAL	900 μl	

Apartado 2 – Enzimas (PK + MK)

- | |
|---|
| (a) Centrifugar 6,5 μl de PK y 7,0 μl de MK, 5 min. a 4°C , en microcentrífuga. |
| (b) Retirar el sobrenadante y resuspender en 45 μl de Reactivos IV. |

MK (mioquinasas): 2,7 mg/ml – 1900 U/mg
 PK (piruvato quinasa): 18,3 mg/ml – 515 U/mg

Figura 17.3. Protocolo sintético para la síntesis enzimática de ($\alpha^{32}\text{P}$) ATP. La mezcla de reactivos IV del *Apartado 1* es estable por largo tiempo, conservándola a -80°C , y soporta múltiples ciclos de congelación y descongelación. Uno de los stocks se ha usado durante 4 años sin perjuicios en el rendimiento final. La enzima PNK (polinucleótido quinasa), comercializada en solución con 50% glicerol, se almacena a -20°C . La enzima NPI (nucleasa P1) se comercializa liofilizada y debe solubilizarse en un buffer específico (NaAcO 10 mM pH 5,0, 50% glicerol) a una concentración final de 3,0 U/ μl (actividad sobre 3' AMP); almacenada a -20°C es estable durante largo tiempo (más de 4 años sin pérdidas apreciables de actividad). La NPI tiene un pH óptimo para RNA y ribonucleótidos entre 7 y 9. Las enzimas MK y PK, se comercializan como suspensiones en $(\text{NH}_4)_2\text{SO}_4$, en buffers *ad hoc*, y son estables por largo tiempo a 4°C . Los mismos stocks se han utilizado durante 2,5 años sin perjuicios en el rendimiento final. Por otra parte, la solución de fosfoenolpiruvato y la mezcla indicada en el *Apartado 2* deben ser preparadas en el momento de usar.

SINTESIS DE ($\alpha^{32}\text{P}$) dATP

Operaciones

1. Centrifugar el ($\gamma^{32}\text{P}$) ATP para decantar las enzimas inactivadas. *Tomar una alícuota de 0,3-0,5 μl .*
2. Mezclar:

($\gamma^{32}\text{P}$) ATP (25 mCi)	300,00 μl
3' dAMP 18 mM (en H_2O)	30,00 μl
Tris-HCl 50 mM pH 9,0	28,50 μl
PNK (10 U/ μl)	1,50 μl
TOTAL	360,00 μl

3. Incubar a temperatura ambiente 12 - 16 horas (ó 4 hs. a 37°C). *Tomar una alícuota de 0,3-0,5 μl*
4. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 1 y 3, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
5. Si la transferencia es mayor del 95% continuar con la etapa siguiente. *Tomar una alícuota de 0,3-0,5 μl*
6. Mezclar:

(5' $\alpha^{32}\text{P}$) 3' dADP	360,00 μl
AcOH 1M	30,00 μl
NaAcO 20 mM pH 5,0	50,00 μl
NP1 (3,0 U/ μl) (en NaAcO 10 mM pH 5,0)	10,00 μl
TOTAL	450,00 μl

7. Incubar a temperatura ambiente 30 - 45 minutos. *Tomar una alícuota de 0,3-0,5 μl*
8. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 5 y 7, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
9. Si la hidrólisis es mayor del 95%, agregar 8 μl de NaOH 1M (el pH debe ser $\sim 8,0$). *Tomar una alícuota de 0,3-0,5 μl*
10. Mezclar:

Reactivos IV (ver apartado 1)	368,00 μl
Fosfoenolpiruvato 10 mg/ml (en H_2O)	45,00 μl
(5' $\alpha^{32}\text{P}$) dAMP	458,00 μl
Mezcla enzimas (PK+MK) (ver apartado 2)	45,00 μl
TOTAL	916,00 μl

11. Incubar a temperatura ambiente 30 - 60 minutos. *Tomar una alícuota de 0,3-0,5 μl*
12. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 9 y 11, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.
13. Si la incorporación es mayor del 95% inactivar las enzimas. Colocar 10 min en agua hirviendo.

Apartados

Apartado 1 – Reactivos IV

Solución stock	Volumen en la mezcla	Concentración final en la reacción
1 M Tris HCl pH 8,0	50 μl	55,6 mM
1 M Mg Cl_2	10 μl	11,1 mM
1 M KCl	20 μl	22,2 mM
100 mM dithiothreitol	50 μl	5,6 mM
H_2O	770 μl	
VOLUMEN TOTAL	900 μl	

Apartado 2 – Enzimas (PK + MK)

- (a) Centrifugar 6,5 μl de PK y 7,0 μl de MK, 5 min. a 4°C , en microcentrífuga.
- (b) Retirar el sobrenadante y resuspender en 45 μl de Reactivos IV.

MK (mioquinasas): 2,7 mg/ml – 1900 U/mg
 PK (piruvato quinasa): 18,3 mg/ml – 515 U/mg

Figura 17.4. Protocolo sintético para la síntesis enzimática de ($\alpha^{32}\text{P}$) dATP. La mezcla de reactivos IV del *Apartado 1* es estable por largo tiempo, conservándola a -80°C , y soporta múltiples ciclos de congelación y descongelación. Uno de los stocks se ha usado durante 4 años sin perjuicios en el rendimiento final. La enzima PNK (polinucleótido quinasa), comercializada en solución con 50% glicerol, se almacena a -20°C . La enzima NP1 (nucleasa P1) se comercializa liofilizada y debe solubilizarse en un buffer específico (NaAcO 10 mM pH 5,0, 50% glicerol) a una concentración final de 3,0 U/ μl (actividad sobre 3' AMP); almacenada a -20°C es estable durante largo tiempo (más de 4 años sin pérdidas apreciables de actividad). La NP1 tiene un pH óptimo para DNA y deoxirribonucleótidos entre 4,5 y 6,0. Las enzimas MK y PK, se comercializan como suspensiones en $(\text{NH}_4)_2\text{SO}_4$, en buffers *ad hoc*, y son estables por largo tiempo a 4°C . Los mismos stocks se han utilizado durante 2,5 años sin perjuicios en el rendimiento final. Por otra parte, la solución de fosfoenolpiruvato y la mezcla indicada en el *Apartado 2* deben ser preparadas en el momento.

SINTESIS DE ($\alpha^{32}\text{P}$) dCTP

Operaciones

14. Centrifugar el ($\gamma^{32}\text{P}$) ATP para decantar las enzimas inactivadas. *Tomar una alícuota de 0,3-0,5 μl .*

15. Mezclar:

($\gamma^{32}\text{P}$) ATP (25 mCi)	300,00 μl
3' dCMP 18 mM (en H_2O)	30,00 μl
Tris-HCl 50 mM pH 9,0	28,50 μl
PNK (10 U/ μl)	1,50 μl
TOTAL	360,00 μl

16. Incubar a temperatura ambiente 12 - 16 horas (ó 4 hs. a 37°C). *Tomar una alícuota de 0,3-0,5 μl*

17. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 1 y 3, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.

18. Si la transferencia es mayor del 95% continuar con la etapa siguiente. *Tomar una alícuota de 0,3-0,5 μl*

19. Mezclar:

(5' $\alpha^{32}\text{P}$) 3' dCDP	360,00 μl
AcOH 1M	30,00 μl
NaAcO 20 mM pH 5,0	50,00 μl
NP1 (3,0 U/ μl) (en NaAcO 10 mM pH 5,0)	10,00 μl
TOTAL	450,00 μl

20. Incubar a temperatura ambiente 30 - 45 minutos. *Tomar una alícuota de 0,3-0,5 μl*

21. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 5 y 7, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.

22. Si la hidrólisis es mayor del 95%, agregar 8 μl de NaOH 1M (el pH debe ser $\sim 8,0$). *Tomar una alícuota de 0,3-0,5 μl*

23. Mezclar:

Reactivos IV (ver apartado 1)	368,00 μl
Fosfoenolpiruvato 10 mg/ml (en H_2O)	45,00 μl
(5' $\alpha^{32}\text{P}$) dCMP	458,00 μl
Mezcla enzimas (PK+NMK+NDK) (Apartado 2)	45,00 μl
TOTAL	916,00 μl

24. Incubar a temperatura ambiente 30 - 60 minutos. *Tomar una alícuota de 0,3-0,5 μl*

25. Analizar los productos mediante cromatografía ascendente en PEI-celulosa de las alícuotas de 9 y 11, usando KH_2PO_4 0,33 M pH 3,5. Secar la placa y visualizar por autorradiografía.

26. Si la incorporación es mayor del 95% inactivar las enzimas. Colocar 10 min en agua hirviendo.

Apartados

Apartado 1 – Reactivos IV

Solución stock	Volumen en la mezcla	Concentración final en la reacción
1 M Tris HCl pH 8,0	50 μl	55,6 mM
1 M Mg Cl_2	10 μl	11,1 mM
1 M KCl	20 μl	22,2 mM
100 mM dithiothreitol	50 μl	5,6 mM
H_2O	770 μl	
VOLUMEN TOTAL	900 μl	

Apartado 2 – Enzimas (PK + NMK + NDK)

- (a) Centrifugar 6,5 μl de PK, 5 min. a 4°C , en microcentrífuga.
 (b) Retirar el sobrenadante, resuspender en 45 μl de Reactivos IV y agregar 0,3 μl de NMK y 0,1 μl de NDK.

PK (piruvato quinasa): 18,3 mg/ml – 515 U/mg
 NMK (nucleósido monofosfato quinasa): 0,6 U/mg
 NDK (nucleósido difosfato quinasa): 1660 U/mg

Figura 17.5. Protocolo sintético para la síntesis enzimática de ($\alpha^{32}\text{P}$) dCTP. Las consideraciones acerca de la estabilidad de los reactivos o mezclas y las enzimas PNK, NP1 y PK son idénticas a las indicadas en la **Figura 17.4**. Por otra parte, la última etapa de esta síntesis requiere algunas consideraciones particulares. La adición de fosfatos no radioactivos en las posiciones β y γ es realizada con las enzimas NMK (nucleósido monofosfato quinasa) y NDK (nucleósido difosfato quinasa), ambas enzimas se comercializan liofilizadas y deben ser reconstituídas en un buffer con 50% glicerol para poder almacenarlas a 20°C . Las concentraciones finales en el buffer de reconstitución (50 mM Tris HCl pH 8,0; 20 mM KCl; 5 mM DTT; 50% glicerol), deben ser: NMK: 0,05 U/ μl y NDK: 0,25 U/ μl . El ATP necesario para la catálisis de estas enzimas deriva del ADP residual de la síntesis del ($\gamma^{32}\text{P}$) ATP, el fosfoenolpiruvato y la PK (piruvato quinasa) presente en esta etapa. La solución de fosfoenolpiruvato y la mezcla de enzimas PK+NMK+NDK deben ser preparadas en el momento.

Resultados

Usualmente, se realizó la síntesis de ($\gamma^{32}\text{P}$) ATP a partir de 25 o 50 mCi de (^{32}P) ortofosfato. A continuación, se sintetizaron los ($\alpha^{32}\text{P}$) dNTPs y/o ($\alpha^{32}\text{P}$) rNTPs, de acuerdo a las necesidades.

Los rendimientos típicos de cada etapa de síntesis fueron: 90-96% para la síntesis de ($\gamma^{32}\text{P}$) ATP, 100% para la síntesis de ($5' \alpha^{32}\text{P}$) 3' dADP o ($5' \alpha^{32}\text{P}$) 3' rADP, 100% para la síntesis de ($5' \alpha^{32}\text{P}$) dAMP o ($5' \alpha^{32}\text{P}$) rAMP y >95% para la síntesis de ($5' \alpha^{32}\text{P}$) dATP o ($5' \alpha^{32}\text{P}$) rATP (Figura 17.6). Por otra parte, las actividades específicas obtenidas fueron del orden de 8000-9000 Ci/mmol para el ($\gamma^{32}\text{P}$) ATP y el ($5' \alpha^{32}\text{P}$) dATP, del orden de 900-1200 Ci/mmol para el ($5' \alpha^{32}\text{P}$) rATP.

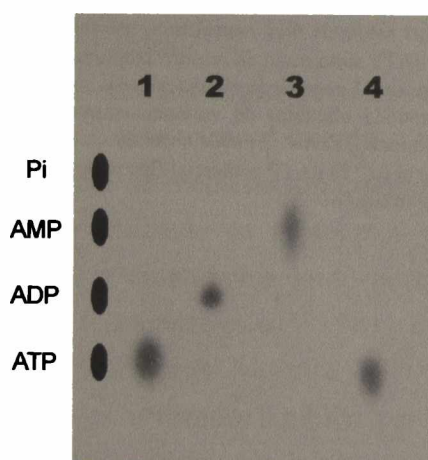


Figura 17.6. Autorradiografía analítica de las etapas de síntesis enzimática de ($\alpha^{32}\text{P}$) rATP. Sobre una placa de PEI-celulosa en capa fina se sembraron alicuotas de 0,3-0,5 μl correspondientes a cada una de las etapas de la síntesis. Los productos se resolvieron mediante cromatografía ascendente utilizando buffer KH_2PO_4 0,66 M pH 3,5. 1. Síntesis de ($\gamma^{32}\text{P}$) ATP. 2. Síntesis de ($5' \alpha^{32}\text{P}$) 3' rADP. 3. Síntesis de ($5' \alpha^{32}\text{P}$) rAMP. 4. Síntesis de ($5' \alpha^{32}\text{P}$) rATP. Con óvalos negros se indican las posiciones relativas correspondientes a ATP, ADP y AMP no radioactivos (visualizados por absorción de luz UV), y la correspondiente al ortofosfato (Pi) no incorporado.

Para evaluar la posible interferencia de los restos de reactivos, derivados de la síntesis, en la incorporación en moléculas de DNA y RNA, se realizaron distintos ensayos de utilización en técnicas comunes de biología molecular.

Utilizando el ($5' \alpha^{32}\text{P}$) dATP sintetizado en nuestro laboratorio y el ($5' \alpha^{32}\text{P}$) dATP comprado en New England Nuclear (USA) (Figura 17.7), se realizaron ensayos comparativos de síntesis de cDNA (datos no mostrados) y marcación de DNA por *nick translation* (Figura 17.8).

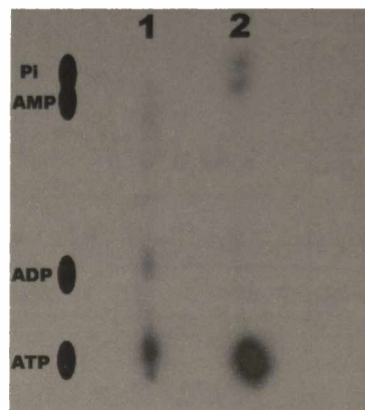


Figura 17.7. Autorradiografía analítica de diferentes síntesis de ($\alpha^{32}\text{P}$) dATP. Sobre una placa de PEI-celulosa en capa fina se sembraron alicuotas equivalentes (0,3-0,5 μl) de dos ($\alpha^{32}\text{P}$) dATPs de distinto origen. Los productos se resolvieron mediante cromatografía ascendente utilizando buffer KH_2PO_4 0,75 M pH 3,5. 1. ($\alpha^{32}\text{P}$) dATP comprado en New England Nuclear. 2. ($\alpha^{32}\text{P}$) dATP sintetizado en nuestro laboratorio. Con óvalos negros se indican las posiciones relativas correspondientes a ATP, ADP y AMP no radioactivos (visualizados por absorción de luz UV), y la correspondiente al ortofosfato (Pi) no incorporado.

En base a los resultados obtenidos, se decidió obviar la etapa de purificación final por columna (DEAE-Sephadex A-25, Pharmacia, Suecia).

Hasta el presente, se continúan utilizando en nuestro laboratorio nucleósidos trifosfato marcados con ^{32}P sintetizados de la manera descrita.

Nota: Este trabajo fue presentado en la SAIB

Síntesis enzimática de $\gamma^{32}\text{P}$ y $\alpha^{32}\text{P}$ nucleósidos trifosfato

V. Romanowski, P.D. Ghiringhelli y O. Grau
XXI Reunión Anual de la Sociedad Argentina de Investigación Bioquímica (SAIB).

Huerta Grande, Córdoba, 28-30 de octubre de 1985.

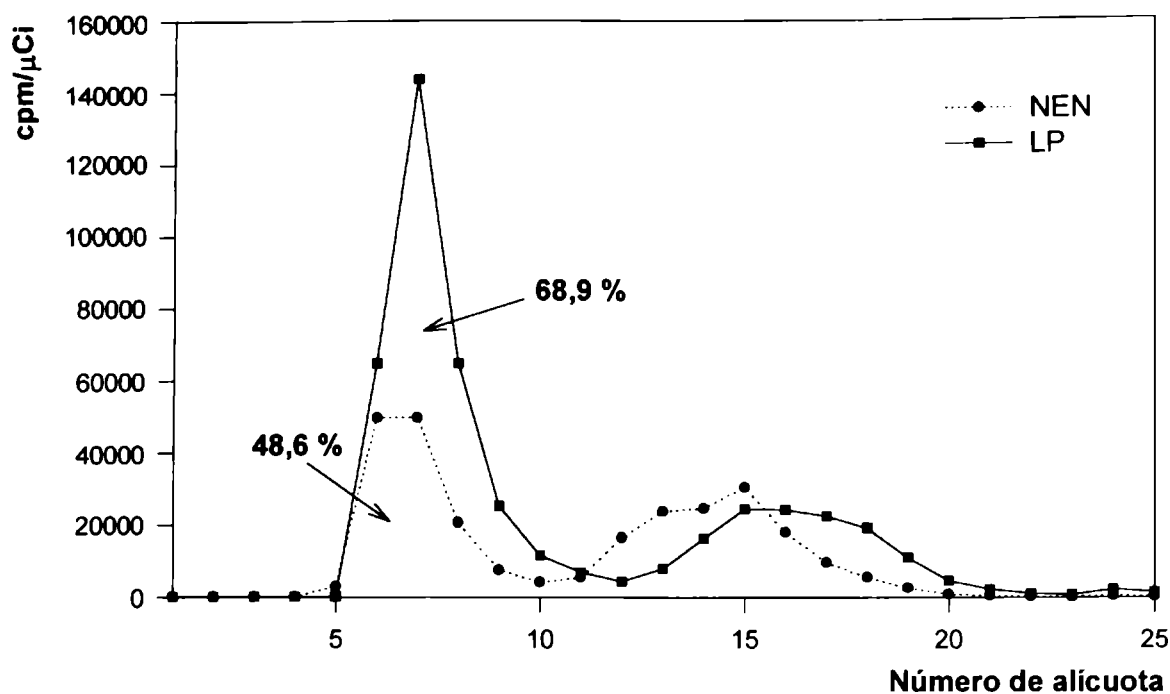


Figura 17.8. Análisis comparativo de la eficiencia de incorporación de dos ($\alpha^{32}\text{P}$) dATP de distinto origen. Sobre alícuotas equivalentes de un mismo molde de DNA se realizó una síntesis de sondas radioactivas mediante *nick translation*, utilizando $9,5 \mu\text{Ci}$ totales de ($\alpha^{32}\text{P}$) dATP de New England Nuclear (NEN) y $34,8 \mu\text{Ci}$ totales de ($\alpha^{32}\text{P}$) dATP sintetizado en nuestro laboratorio (L.P), las concentraciones finales de los demás reactivos y las condiciones de reacción fueron iguales. Luego de la incubación, las sondas fueron purificadas utilizando una columna de Sephadex G-50 (Pharmacia). Se recolectaron 25 alícuotas de volumen equivalente, cuya radioactividad fue medida en un contador de centelleo líquido (Rackbeta, Pharmacia, Suecia). Como se puede observar en el gráfico, el ($\alpha^{32}\text{P}$) dATP sintetizado en nuestro laboratorio fue incorporado más eficientemente que el ($\alpha^{32}\text{P}$) dATP comercial (los valores indican el porcentaje del área total representado por el pico correspondiente al DNA radioactivo sintetizado).

UN PROTOCOLO RAPIDO PARA VISUALIZAR PLACAS VIRALES EN EXPERIMENTOS DE TITULACION

Muchas situaciones experimentales requieren de una cuantificación precisa y confiable de los stocks de virus realizando titulaciones por plaqueo (por ej.: stocks de vacunas, estabilidad de preparaciones virales, inóculos virales para una variedad de propósitos incluyendo vectores de expresión recombinantes, etc.). Sin embargo, muchos virus eucariotas producen placas que no son detectables sin algún tipo de tinción, y algunos son de difícil visualización aún después de teñir las monocapas de células infectadas con rojo neutro o azul tripan, debido al bajo contraste obtenido con estos colorantes vitales. En cualquiera de estos casos, el conteo de placas se transforma en una tarea muy laboriosa y los datos obtenidos son poco confiables.

Para evitar estos problemas, hemos desarrollado un procedimiento alternativo utilizando azul de metileno, un colorante que es ampliamente utilizado en protocolos de tinción histoquímica y citoquímica tanto para sistemas procariotas (Gerhardt *et al.*, 1981) como para eucariotas (Martoja & Martoja, 1970). El azul de metileno es utilizado también para detectar específicamente ácidos nucleicos *in vivo* e *in vitro* (Clark, 1980; Jeppesen, 1974; Martoja & Martoja, 1970).

En esta publicación describimos un protocolo general que permite diferenciar claramente las células en una monocapa de las placas formadas por virus animales o de insectos. Como ejemplos de ambos tipos de sistemas hemos utilizado células Vero 76 infectadas con arenavirus y células Sf 21 (*Spodoptera frugiperda*) infectadas con AcNPV (*Autographa californica* nuclear polyhedrosis virus) tipo salvaje y recombinantes expresando diferentes genes foráneos.

En ambos plaqueos, las células fueron sembradas a una densidad de aproximadamente $1,5 \times 10^6$ células por placa de 35 mm. Una vez formada la monocapa subconfluente de células, y después de infectar con una dilución apropiada de virus, realizada en medio de cultivo, se retiró cuidadosamente el inóculo y las células fueron cubiertas con medio semisólido, de acuerdo a los protocolos disponibles (Dulbecco, R., 1952; Piwnica-Worms, H., 1990; Summers and Smith, 1987).

Las placas fueron incubadas el período de tiempo establecido (en nuestro caso 4-5 días) y teñidas de acuerdo al siguiente procedimiento:

- ⇒ Agregar 2 ml de 4% formaldehído en 1 X PBS (si se trabaja con células Sf 21 en medio TC100 es necesario eliminar la capa de medio líquido) e incubar 30 min. A temperatura ambiente.
- ⇒ Después de la fijación, descartar la solución de formaldehído y eliminar la capa de agarosa utilizando una espátula.
- ⇒ Agregar 0,5 ml de solución stock de azul de metileno por placa, incubar 15 seg. La solución stock contiene 0,8% de colorante, 48% etanol, 0,005% KOH en agua, y es estable a temperatura ambiente en condiciones estándar.
- ⇒ Enjuagar 4 veces con agua destilada y dejar secar las placas invertidas sobre toallas de papel.
- ⇒ Las placas pueden ser contadas inmediatamente.

Cuando la titulación se realiza con un sistema en el cual el virus produce lisis celular, tal como un arenavirus plaqueado en células Vero 76, las placas aparecen como áreas completamente transparentes sobre un fondo azul oscuro. En sistemas no líticos, tal como baculovirus-células Sf 21, las placas son áreas con una densidad celular muy inferior a las regiones de la monocapa con células no infectadas, y pueden ser fácilmente visualizadas como manchas translúcidas azul claro sobre un fondo azul oscuro (**Figura 17.9**)

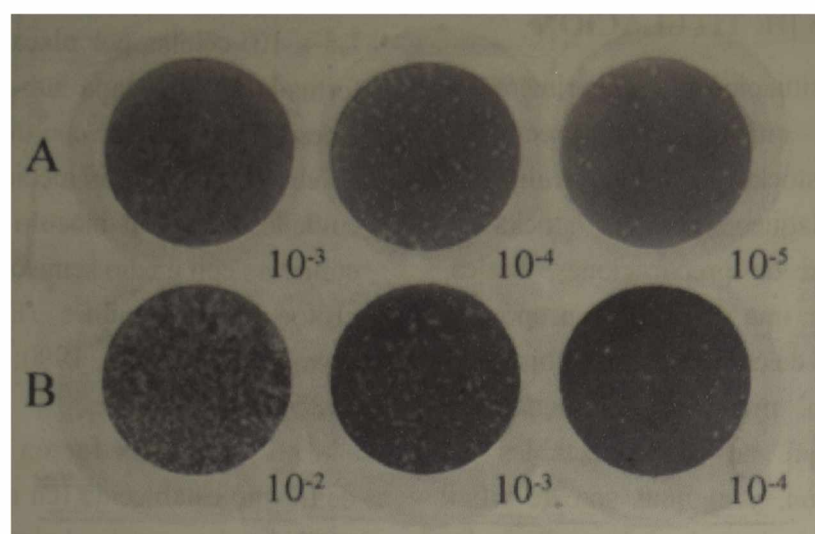


Figura 17.9. Titulación de virus eucarióticos utilizando el protocolo de tinción con azul de metileno.
A: Alícuotas de 0,1 ml de diluciones seriadas de un recombinante del virus de la poliedrosis nuclear de *Autographa californica* (AcNPV) fueron adsorbidas a monocapas de células de *Spodoptera frugiperda* (Sf 21), una vez retirados los inóculos las monocapas fueron recubiertas con 2 ml de 1,5% de agarosa Seaplaque™ conteniendo 0,5X de medio TC100. Una vez gelificada, la agarosa semisólida fue cubierta con 1 ml de 1X TC100 (conteniendo 10% de suero fetal bovino). Las placas fueron incubadas durante 4 días a 28°C y teñidas.
B: Alícuotas de 0,1 ml de diluciones seriadas del arenavirus Junín fueron adsorbidas a monocapas de células Vero 76, una vez retirados los inóculos las monocapas fueron recubiertas con 2 ml de 1,0% de agarosa Seaplaque™ conteniendo 1X de medio MEM y 2% de suero fetal bovino. Las placas fueron incubadas durante 5 días a 37°C y teñidas.

El procedimiento aquí descrito es simple y rápido (40 min.), y las placas pueden ser visualizadas inmediatamente, a diferencia de muchos protocolos de tinción previamente descritos (Tabla 17.b).

Este protocolo no puede ser utilizado cuando se necesita recuperar virus a partir de las placas, debido a la inclusión del paso de fijación con formaldehído. Sin embargo, esta aparente desventaja puede ser una ventaja en algunos casos, por ej.: los plaques pueden ser contados en un laboratorio común ya que los

virus son inactivados durante la fijación con formaldehído. Además, las monocapas de células fijadas no “desaparecen” después de la tinción (utilizando procedimientos de tinción con colorantes vitales, algunas células mueren y se despegan, produciendo un borronero de las placas luego de unas pocas horas).

Nota: Este trabajo fue publicado
A quick protocol for visualizing virus plaques in titration experiments
 P. D. Ghiringhelli and V. Romanowski
Biotechniques 17, 464-465 (1994)

Tabla 17.b. Comparación de tres protocolos de tinción

	MTT (Shafanet, A.B., 1991)	Rojo neutro (Dulbecco, R., 1952; Summers & Smith, 1987)	Azul de metileno
Fijación	NN	NN	30 min.
Tiempo de tinción	2 horas	2-4 horas	15 seg.
Recuperación de virus a partir de la placa	Sí	Sí	No
Detección de las placas	2 horas	2-12 horas (depende de las células)	inmediatamente

NN: no necesario

LIGACION DE OLIGONUCLEOTIDOS A HIBRIDOS cDNA/RNA, CATALIZADA POR LA DNA LIGASA DEL FAGO T4 Y AMPLIFICACION POR PCR: BASES PARA UN METODO RAPIDO DE CLONADO DE EXTREMOS 5' DE mRNAs Y RNAs VIRALES.

Introducción

La combinación de la transcripción reversa de RNAs con la PCR (RT-PCR, Doherty *et al.*, 1989) ha aumentado sensiblemente la capacidad de síntesis de cantidades clonables de cDNA a partir de un número muy escaso de células o de especies de mRNAs o RNAs varios que se encuentran en muy bajas proporciones.

Dependiendo del tipo de RNA buscado (mRNA o RNA no poliadenilado, celular o viral) y del grado de conocimiento previo de las secuencias buscadas o de otras homólogas, el diseño de un *primer* adecuado para iniciar la síntesis del cDNA no presenta grandes problemas. Este *primer* puede ser utilizado, también, como uno de los miembros del par a utilizar en la PCR. El *primer* que suele presentar mayor complejidad para el diseño es el otro miembro del par de la PCR. En particular, cuando se desean determinar con precisión los primeros nucleótidos del extremo 5' de cualquier molécula de RNA se enfrenta una situación de compromiso en la cual, en la mayor parte de las técnicas ya descritas, el grado de certeza acerca de la secuencia obtenida aumenta en la medida en que se repita iterativamente el procedimiento sobre el mismo molde.

Los métodos basados en el uso de transferasa terminal (TdT, *terminal deoxyribonucleotide transferase*) ya descritos, como: *anchor-PCR* (Delort *et al.*, 1989; Loh *et al.*, 1989; Belyavsky *et al.*, 1989), 5' RACE (*rapid amplification of cDNA ends*; Frohman *et al.*, 1988), *one-sided polymerase chain reaction* (Ohara *et al.*, 1989), y las distintas variantes de los mismos, tienen importantes limitaciones. Primero, la reacción de TdT es difícil de controlar y de baja eficiencia. Segundo,

el *primer* para PCR correspondiente a la zona de extensión con TdT tiene una cola homopolimérica, lo cual trae aparejado una seria disminución en la eficiencia y especificidad de la amplificación. Tercero, esta reacción se realiza utilizando como sustrato el cDNA de simple cadena, por lo tanto, para evitar la adición de deoxirribonucleótidos a productos parciales de extensión de *primer*, se debe purificar el/los productos de extensión de *primer* de mayor longitud, para lo cual es necesario disponer de mayores masas de material de partida. Cuarto, la adición de un deoxirribonucleótido determinado al extremo 3' del cDNA puede derivar en la pérdida de información acerca de uno o más nucleótidos idénticos al agregado, pero que sean parte de la secuencia del RNA; para solucionar esto, se debería repetir el mismo procedimiento con la adición de, al menos, tres nucleótidos distintos.

Por otra parte, la diversidad de métodos que tienen en común el uso de RNA ligasa del fago T4, aún cuando tienen mayor sensibilidad que los anteriores (se puede partir de masas de RNA inferiores) y dan resultados más claros, también presentan algunos inconvenientes. La RNA ligasa del fago T4 se caracteriza por tener mayor actividad sobre sustratos de simple cadena que sobre sustratos de doble cadena, acepta como moléculas donantes de 5'-P tanto RNA como DNA, mientras que tiene mayor preferencia por moléculas 3'-OHceptoras de RNA (Kornberg and Baker, 1992).

En la mayor parte de los métodos basados en el uso de RNA ligasa del fago T4, si se quiere disminuir la intensidad de trabajo posterior de selección de clones, es aconsejable partir de RNAs altamente purificados. En el método de Dumas *et al.* (1991), ligación entre un oligodeoxirribonucleótido de secuencia definida y el extremo 3'-OH del cDNA, para aumentar la eficiencia de esta reacción se debe eliminar el RNA molde, con lo cual se incorpora la posibilidad de ligaciones a productos parciales de extensión de *primer*. Una hidrólisis

alcalina extensiva es necesaria para minimizar la inserción de pequeños oligorribonucleótidos derivados del RNA molde inicial entre el *primer* a ligar y el extremo 3' del cDNA. Además, también se debe eliminar el *primer* utilizado en la síntesis del cDNA. La realización poco eficiente de cualquiera de estas etapas se traduce en una disminución en la cantidad de moldes efectivos para la PCR y, por tanto, en una pérdida de sensibilidad.

El método descrito por Meyer & Southern (1993) está limitado a RNAs carentes de modificaciones en los extremos y, además, no está exento de la posible inserción de pequeños oligorribonucleótidos entre ambos extremos.

Mientras que en el método de Lambden *et al.* (1992), se realiza la ligación de un único *primer* a ambos extremos 3' de las hebras de una molécula de RNA de doble cadena y la amplificación mediante RT-PCR con un oligonucleótido complementario al ligado. Las limitaciones de este método son obvias: en primera instancia, está restringido a un sustrato inicial de doble cadena y, en segunda instancia, la posibilidad de obtener el producto final depende de la longitud de la molécula sustrato y de la eficiencia de la reacción de PCR.

En 1993, Bertling *et al.* publicaron un método basado en el uso de la DNA ligasa del fago T4. Esta enzima, típicamente une dos moléculas de DNA de doble cadena, con extremos cohesivos o romos. Además, puede catalizar la ligación de homopolímeros de RNA alineados sobre un DNA complementario, la ligación de cadenas de DNA alineadas sobre un RNA complementario, la ligación del extremo 3' de una molécula de RNA (que forma parte del dúplex híbrido RNA/DNA) a DNA y es muy sensible a desapareamientos en la zona de unión o en sus cercanías (Kornberg and Baker, 1992). Sin embargo, Harada & Orgel (1993) han reportado que oligonucleótidos con desapareamientos en la zona de unión o en las cercanías son ligados con la pérdida de apenas el 50% de la eficiencia.

En el método de Bertling se utiliza un dúplex de oligodeoxirribonucleótidos, uno de los cuales tiene un extremo 3' extendido de 4 nucleótidos de longitud y secuencia aleatoria. Este dúplex es ligado al extremo 3' del cDNA de simple cadena (previa hidrólisis del RNA molde). En esta técnica es recomendable purificar los productos de longitud máxima de la extensión de *primer*, ya que, dada la naturaleza aleatoria del extremo extendido, que hibrida con los últimos nucleótidos del cDNA, una vez hidrolizado el RNA molde, el dúplex puede ser ligado a cualquier molécula de cDNA, independientemente de su longitud.

Para evitar estos inconvenientes, se desarrolló un método que se basa en la unión de oligodeoxirribonucleótidos sintéticos a los híbridos cDNA/RNA derivados de la extensión de *primer* y la posterior amplificación por PCR sin etapas de purificación intermedias.

Materiales y métodos

Virus y líneas celulares

La cepa MC2 del virus Junín fue crecida en células BHK-21 (ATTC). Las monocapas con un 50% de confluencia fueron infectadas con 0,1 pfu/cél. e incubadas a 37°C en medio Eagle's MEM conteniendo aminoácidos no esenciales y 10% de suero fetal bovino. El porcentaje de suero fue reducido a 2% un día después de la infección. Los virus fueron recuperados y purificados a partir del medio sobrenadante, recolectado cada 8 hs., a 3, 4 y 5 días post-infección (p.i.) (Grau *et al.*, 1981; Rosas *et al.*, 1988).

La cepa Candid #1 del virus Junín fue crecida en células diploides FRhL-2 (*fetal rhesus lung diploid*). Las monocapas con un 50% de confluencia fueron infectadas con 0,1 pfu/cél. e incubadas a 37°C en medio Eagle's MEM conteniendo aminoácidos no esenciales y 10% de suero fetal bovino. Los virus fueron recuperados y purificados, de la manera ya citada, a partir del medio sobrenadante al 4 día p.i..

Preparación de RNAs virales

Los viriones sedimentados por ultracentrifugación a partir del medio sobrenadante (rotor 70 Ti, 90 min., 54.000 rpm) fueron desarmados con isotiocianato de guanidinio y el RNA fue purificado de acuerdo al método de Chomczynski & Sacchi (1987). El precipitado de RNA fue resuspendido en sarkosyl 0,5%, extraído con fenol ácido y reprecipitado con 0,3 M de acetato de sodio y 2,5 vol. de etanol. El precipitado final fue resuspendido en agua y almacenado en alícuotas a -70°C . El RNA fue analizado por electroforesis en agarosa en la presencia de hidróxido de metilmercurio (II) como desnaturizante (Bayley & Davidson, 1976).

Extensión de primer, adición de nucleótidos con TdT y PCR (Figura 17.10)

Para iniciar la síntesis de cDNA se utilizaron, alternativamente, dos *primers*, que permiten determinar los extremos 5' de los dos mRNAs derivados del RNA S y/o ambos extremos del RNA viral. Para extender el extremo 5' del RNA viral complementario (VC) y el extremo 5' del mRNA, codificante para la proteína de la nucleocápside (N), se utilizó un oligonucleótido sintético de polaridad viral (V), correspondiente a las posiciones 3215-3231 (N 5'-GGCATCCTTCAGAACAT-3'). Por otra parte, para extender el extremo 5' del RNA viral (V) y el extremo 5' del mRNA, codificante para el precursor de las glicoproteínas (GPC), se utilizó un oligonucleótido sintético de polaridad VC, correspondiente a las posiciones 242-196 V (correspondiente a 3157-3204 VC; Gf 5'-GATCTTCCTGCTAGATCAAGAATTGAACAACCACTTTTGTACAGGTT-3').

En ambos casos, el RNA viral fue mezclado con el primer correspondiente y desnaturizado con hidróxido de metilmercurio (II) (10 mM) durante 5 min. Después de la inactivación del desnaturizante con 140 mM 2-mercaptoetanol, la síntesis de la primer cadena del cDNA fue catalizada por la transcriptasa reversa de AMV (*avian myeloblastosis virus*,

Promega, Madison, WI). A continuación, se degradó el RNA con 50 mM NaOH y se purificó el cDNA de simple cadena mediante precipitación con 2,5 vol. de etanol y en presencia de 0,3 M de acetato de sodio. El precipitado final fue resuspendido en agua. Alternativamente, los productos de extensión completa o mayores (de acuerdo a la secuencia descrita, Ghiringhelli *et al.*, 1991) fueron purificados mediante electroforesis en geles de 6% acrilamida-N,N' bisacrilamida (38:2), utilizando 8 M urea como desnaturizante. Los fragmentos fueron recuperados mediante elución en agua (toda la noche a 60°C) y posterior concentración mediante precipitación con etanol y acetato de sodio.

Utilizando dCTP como precursor y TdT (Promega, Madison, WI), se agregó una sucesión de Cs (8 a 12 nt. en promedio) al extremo 3' del cDNA, en las condiciones descritas por el proveedor.

El cDNA con la cola de Cs fue utilizado como molde para amplificación mediante PCR en una reacción de 35 ciclos, en las condiciones estándar, utilizando *Taq* DNA polimerasa (Promega, Madison, WI) y 1 μM final de los *primers* N y 5'EX (5'-GATCGATGAA-TTCGGGGGGGG-3'). Los productos de amplificación fueron resueltos en geles de agarosa 1,5% y recuperados utilizando NaI y adsorción a matriz de silica (Gene CleanTM, Bio 101, La Jolla, CA). Estos productos fueron ligados al plásmido pUC9 (New England Biolabs, Beverly, MA) digerido con *Sma* I. Utilizando una sonda nucleotídica homóloga, marcada radioactivamente, se identificaron clones recombinantes mediante hibridación de colonias (Grunstein & Hogness, 1975).

Preparación de oligonucleótidos sintéticos para la adición

Para la adición se utilizaron alternativamente un solo oligonucleótido, el *primer* NA (5'-GGATCCATGGCACACTCC-AAAGAGGTTCCA-3'), o un dúplex de oligonucleótidos, formado por el *primer* NA y el

primer NB (5'-AGCTTGGAACCTCTTTGGA-GTGTGCCATG-3'). El *primer* NA fue fosforilado con trazas de ($\gamma^{32}\text{P}$)ATP y ATP no radioactivo. El *primer* NB, por otra parte, no fue fosforilado. Las formas máximas respectivas de cada *primer* (NA 30 nt. y NB 29 nt.) fueron purificadas mediante electroforesis en geles de 20% acrilamida-N,N' bisacrilamida (38:2) y 8 M urea. Los *primers* fueron recuperados mediante elución en agua (8 hs. a 50°C) y concentrados posteriormente por evaporación. Para evitar la concatemerización, el extremo 3' del *primer* NA fue bloqueado mediante el agregado de un dideoxynucleótido, utilizando la deoxinucleotidil transferasa y ddCTP como precursor. El *primer* bloqueado fue purificado con MermAid™ (Bio 101, La Jolla, CA).

El dúplex fue formado mezclando cantidades equimolares de los *primers* NA y

NB, en presencia de buffer de hibridación 1 X (10 mM Tris-HCl pH 7,5, 0,1 mM EDTA pH 8,0, 50 mM NaCl). La mezcla fue calentada durante 3 min. a 100°C y luego se permitió un descenso lento de la temperatura hasta 37°C.

Extensión de primer, adición de oligonucleótidos sintéticos y PCR (Figura 17.11)

Para iniciar la síntesis de cDNA se utilizó un oligonucleótido de polaridad VC, complementario a las posiciones 215-196 V (correspondiente a 3185-3204 VC; G 5'-CAACCACTTTTGTACAGGTT-3'). La síntesis de cDNA se realizó de la manera descrita anteriormente.

Las reacciones de ligación se realizaron utilizando, alternativamente, el *primer* NA como simple cadena o el dúplex de oligonucleótidos NA/NB. A su vez, también se ensayaron, en forma independiente, diferentes ligasas (DNA

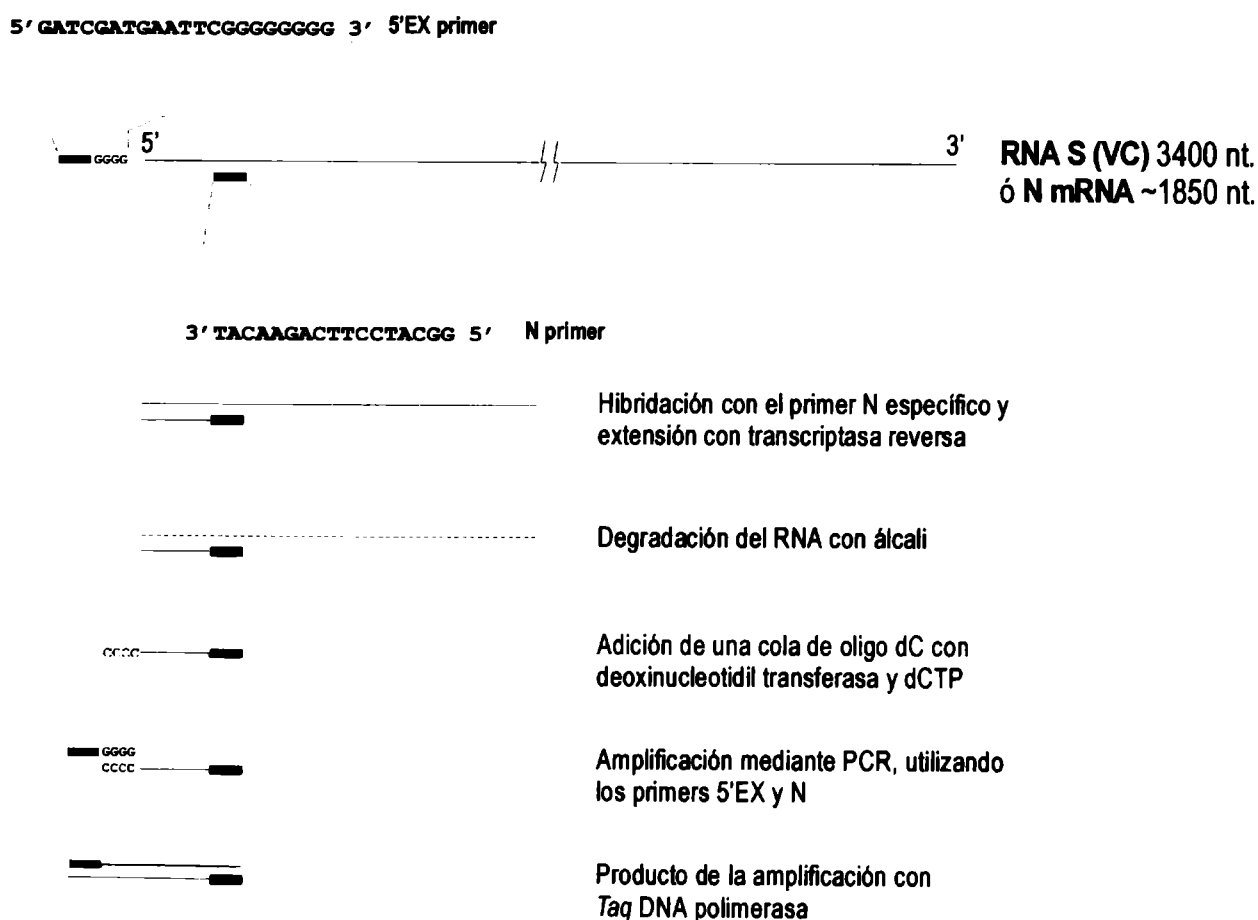


Figura 17.10. Extensión de primer, adición de nucleótidos con TdT y PCR. La reacción de extensión de primer se realizó utilizando un oligonucleótido sintético de polaridad viral (*primer* N), que permite extender los extremos 5' del RNA viral complementario (VC) y 5' del mRNA de N. La síntesis de la primer cadena del cDNA fue realizada con la transcriptasa reversa de AMV. A continuación, se degradó el RNA con 50 mM NaOH y se purificó el cDNA de simple cadena mediante precipitación con etanol y acetato de sodio. El precipitado final fue resuspendido en agua. Utilizando dCTP como precursor y TdT (deoxinucleotidil transferasa), se agregó una sucesión de Cs (8 a 12 nt. en promedio) al extremo 3' del cDNA. Este cDNA fue utilizado como molde para amplificación mediante PCR utilizando Taq DNA polimerasa y los *primers* N y 5'EX.

ligasa y RNA ligasa del fago T4), diferentes buffers: Buffer Mn (1 X: 50 mM Hepes-NaOH pH 7,9, 10 mM MgCl₂, 10 mM DTT, 100 μM ATP, 50 μg/ml BSA) y Buffer Mg (1 X: 30 mM Tris-HCl pH 7,8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP), en presencia o ausencia de 4% PEG 8000 (King *et al.*, 1986), y la incubación conjunta con DNA ligasa de T4 y fragmento de

Klenow de la DNA polimerasa I de *E. coli*, con la adición de 0,5 mM de cada uno de los 4 dNTPs. En todos los casos, las incubaciones se realizaron a 16°C durante toda la noche.

Los cDNAs ligados fueron utilizados como moldes para amplificación mediante PCR en una reacción de 35 ciclos, en las condiciones estándar, utilizando *Taq* DNA polimerasa

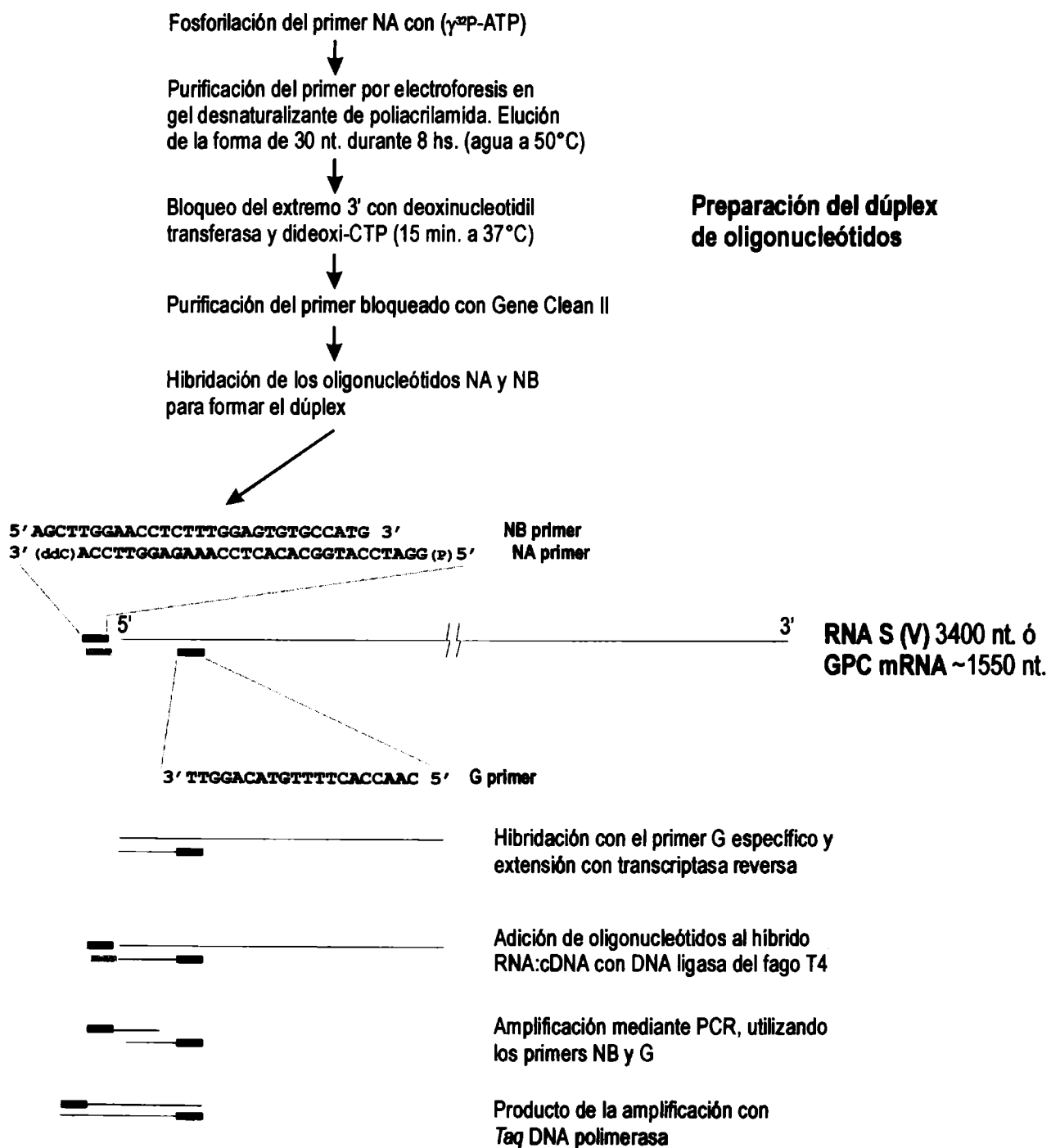


Figura 17.11. Extensión de primer, adición de oligonucleótidos sintéticos y PCR. La reacción de extensión de primer se realizó utilizando un oligonucleótido sintético de polaridad viral complementaria (*primer* G), que permite extender los extremos 5' del RNA viral (V) y 5' del mRNA de GPC. La síntesis de la primer cadena del cDNA fue realizada con la transcriptasa reversa de AMV. A continuación, se realizó la adición del dúplex de oligonucleótidos NA/NB con la DNA ligasa del fago T4 en buffer Mg 1 X. Los cDNAs ligados fueron utilizados como moldes para amplificación mediante PCR, utilizando *Taq* DNA polimerasa y los *primers* G y NB. Los productos de amplificación fueron resueltos en geles de agarosa 1,5%, visualizados mediante tinción con bromuro de etidio.

(Promega, Madison, WI) y 1 μ M final de los *primers* G y NB. Los productos de amplificación fueron resueltos en geles de agarosa 1,5%, visualizados mediante tinción con bromuro de etidio y recuperados utilizando Nal y adsorción a matriz de silica (Gene Clean™, Bio 101, La Jolla, CA), para su posterior clonado y caracterización.

Resultados

Extensión de primer, adición de nucleótidos con TdT y PCR

Para realizar las extensiones de *primer* se utilizaron RNAs virales obtenidos a partir de viriones purificados de las cepas MC2 y Candid #1 del virus Junín, RNAs totales obtenidos a partir de células infectadas, BHK o FRhL-2, con las respectivas cepas y RNA obtenido a partir de una muestra de sangre de un paciente de Fiebre Hemorrágica Argentina. La síntesis de cDNA y la adición de nucleótidos (Cs) se realizó de acuerdo al método descripto.

Las reacciones de PCR con el par de *primers* 5'EX y Gf (extremos 5' viral y 5' del mRNA de GPC) y con el par de *primers* 5'EX y N (extremos 3' viral y 5' del mRNA de N) dieron productos de longitudes acordes con lo esperado (Figura 17.12). En particular, en el caso de los extremos 3' virales y/o 5' de los mRNAs de N, las bandas obtenidas siempre se observaron difusas, lo cual puede deberse a una mayor heterogeneidad de tamaños respecto del otro extremo. Algunos de los fragmentos generados sobre RNA de la cepa MC2 con el par de *primers* 5'EX y Gf fueron recuperados, clonados y secuenciados, (datos no mostrados). En estas secuencias se observaron dos hechos principales, en primera instancia, se obtuvieron varios clones correspondientes a productos de extensión incompleta (problema inherente al método en sí mismo) y, por otra parte, aproximadamente el 50% de los clones obtenidos termina en un conjunto de 4 Cs distante 8 nt. del extremo 5' del RNA viral. Esto último puede deberse tanto a un producto de extensión incompleta como a la hibridación

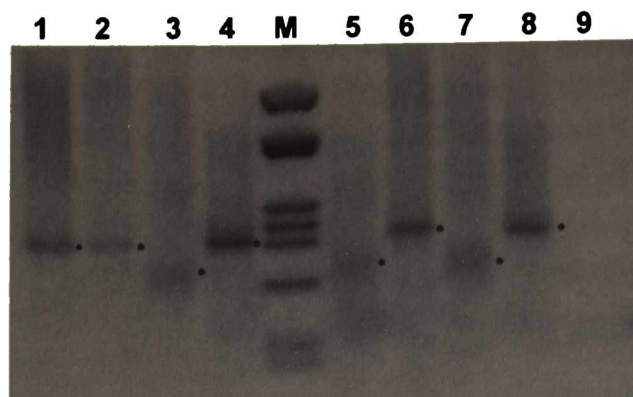


Figura 17.12. Amplificación de cDNAs con nucleótidos adicionados. En todos los casos los moldes corresponden al virus Junín. **Calle 1:** producto de PCR de la extensión de primer del extremo 5' a partir de RNA viral de la cepa MC2. **Calle 2:** producto de PCR de la extensión de primer del extremo 5' a partir de RNA total de células infectadas con la cepa MC2. **Calle 3:** producto de PCR de la extensión de primer del extremo 3' a partir de RNA viral de la cepa CD1. **Calle 4:** producto de PCR de la extensión de primer del extremo 5' a partir de RNA viral de la cepa CD1. **Calle 5:** producto de PCR de la extensión de primer del extremo 3' a partir de RNA total de células infectadas con la cepa CD1. **Calle 6:** producto de PCR de la extensión de primer del extremo 5' a partir de RNA total de células infectadas con la cepa CD1. **Calle 7:** producto de PCR de la extensión de primer del extremo 3' a partir de RNA de sangre de un paciente infectado. **Calle 8:** producto de PCR de la extensión de primer del extremo 5' a partir de RNA de sangre de un paciente infectado. **Calle 9:** control negativo de PCR (sin DNA). **M:** marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp, 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp). Con puntos negros se indican los fragmentos de amplificación esperados: ~240 bp para el extremo 5' y ~186 bp para el extremo 3'. Se muestra la imagen en negativo del gel de agarosa.

errónea del *primer* 5'EX durante la PCR (hibridación en el conjunto de Cs interno durante los primeros ciclos de PCR, en lugar de la hibridación en el sitio correcto)

Extensión de primer, adición de oligonucleótidos sintéticos y PCR

Las extensiones de *primer* se realizaron sobre RNAs totales obtenidos a partir de células BHK infectadas con distintos clones de la cepa MC2 del virus Junín. La síntesis de cDNA se realizó de acuerdo al método descripto.

En primera instancia, se realizaron experimentos tendientes a determinar las condiciones de reacción. Para ello, se realizaron las reacciones de ligación utilizando, alternativamente, el *primer* NA como simple

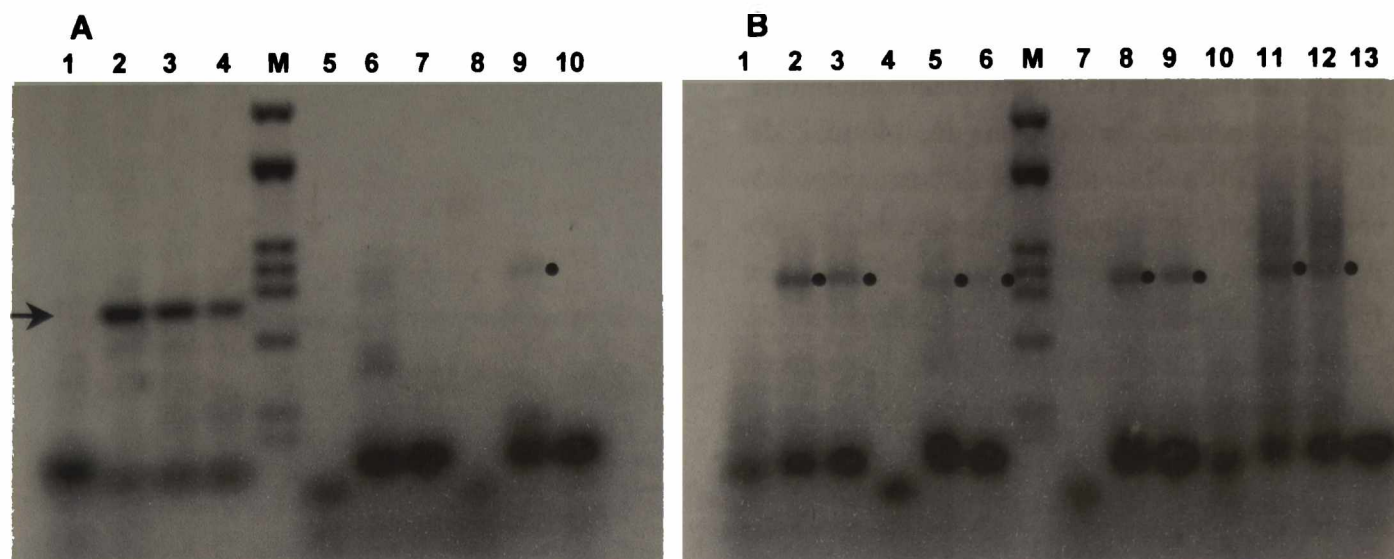


Figura 17.13. Amplificación de cDNAs ligados a oligonucleótidos. En todos los casos, las cantidades de molde en los distintos puntos de las diluciones seriadas corresponden a 200, 20 y 2 femtomoles totales de cDNA no ligado, respectivamente.

A. Calle 1: control negativo de PCR con *primers* NB y G (molde: cDNA sin ligar). **Calles 2 a 4:** controles de existencia del cDNA, amplificaciones con *primers* internos utilizando como moldes diluciones seriadas del cDNA. **Calles 5 a 7:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena. **Calles 8 a 10:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB. Con una flecha negra se indica el fragmento interno de amplificación: ~220 bp y con puntos negros se indica el fragmento de amplificación esperado: ~256 bp.

B. Calles 1 a 3: PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzimas: T4 DNA ligasa + Klenow). **Calles 4 a 6:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena, en presencia de 4% PEG 8000. **Calles 7 a 9:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB, en presencia de 4% PEG 8000. **Calles 10 a 12:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzimas: T4 DNA ligasa + Klenow), en presencia de 4% PEG 8000. **Calle 13:** control negativo de PCR (sin DNA). **M:** marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp, 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp. Con puntos negros se indica el fragmento de amplificación esperado: ~256 bp. Se muestran las imágenes en negativo de los geles de agarosa.

cadena o el dúplex de oligonucleótidos NA/NB. A su vez, también se ensayaron, en forma independiente, las actividades de la DNA ligasa y de la RNA ligasa del fago T4, tanto en el buffer tradicional (Buffer Mg) como en un buffer con manganeso (Buffer Mn). Todas las ligaciones se realizaron en presencia o ausencia de 4% PEG 8000. Por otra parte, como el dúplex NA/NB presenta extremos 5' extendidos, se ensayó también la ligación con DNA ligasa de T4 en presencia del fragmento de Klenow de la DNA polimerasa I de *E. coli* y 0,5 mM de cada uno de los 4 dNTPs. En todos los casos, las incubaciones se realizaron a 16°C durante toda la noche. El análisis mediante PCR de estos experimentos permitió revelar la existencia de productos de ligación del *primer* NA simple cadena o del dúplex NA/NB por acción de la

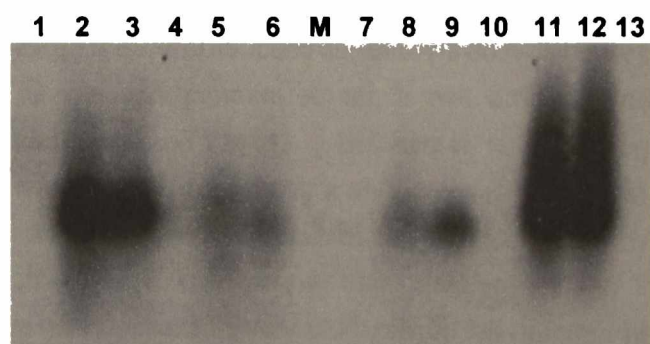


Figura 17.14. Southern blot de los productos de amplificación de cDNAs ligados a oligonucleótidos. La autorradiografía corresponde a la hibridación con una sonda homóloga marcada radioactivamente de una membrana de nylon obtenida mediante la transferencia del gel de la **figura 17.3. B**. Los números de las calles son coincidentes con los de la citada figura. Con una flecha negra se indica la posición del fragmento de amplificación de ~256 bp.

DNA ligasa del fago T4 en Buffer Mg, con o sin la adición del fragmento de Klenow y dNTPs, y en presencia o ausencia de 4% PEG 8000 (datos no mostrados, **Figura 17.13 y Figura 17.14**).

Para observar la existencia de una o más especies en longitud del producto de ligación, el cDNA fue marcado radioactivamente durante la síntesis mediante la adición de 14 μCi de $(\alpha^{32}\text{P})\text{dATP}$ a la mezcla de reacción. A continuación, se realizó la ligación del híbrido RNA-cDNA con el dúplex NA/NB, utilizando la DNA ligasa del fago T4 y el fragmento de Klenow de la DNA polimerasa I de *E. coli*, en presencia de 4% PEG 8000. Las tres cuartas partes de la reacción de ligación fueron resueltas mediante electroforesis en un gel 6% de acrilamida-N N' bisacrilamida y 8 M urea (Figura 17.15). De esta forma se pudieron observar varios productos principales y subproductos. El producto mayoritario de 186 bp no es amplificable por PCR con los *primers* NB y G, ni con los *primers* Arena y G, por lo tanto corresponde a una forma de extensión incompleta debida, probablemente, a la existencia de alguna estructura secundaria que actúa como señal de terminación de la transcripción reversa. El fragmento de 224 bp es 9 nt. más largo que el producto de extensión completa esperado (215 bp), esto puede deberse a la existencia de nucleótidos adicionales en el extremo 5' del mRNA de GPC (Raju *et al.*, 1990). Por otra parte, el producto de 224 bp es amplificable con el par de *primers* Arena y G, pero no con el par NB y G; por lo tanto, este fragmento corresponde a productos de extensión completa no ligados. El fragmento de 256 bp es 2 nt. más pequeño que el producto esperado de la ligación del dúplex NA/NB (224 bp + 34 bp = 258 bp), lo cual puede deberse a una actividad exonucleolítica durante la ligación, sin embargo este es el único producto amplificable por PCR con el par de *primers* NB y G.

Con el objeto de verificar la repetibilidad del método, se realizó la síntesis de cDNA con el *primer* G sobre RNAs de células BHK infectadas con diferentes clones (5₂ y 6₂) y una mezcla de varios clones de la cepa MC2 del virus Junín, RNA viral de la cepa Candid #1 del virus Junín y RNAs de células FRhL-2

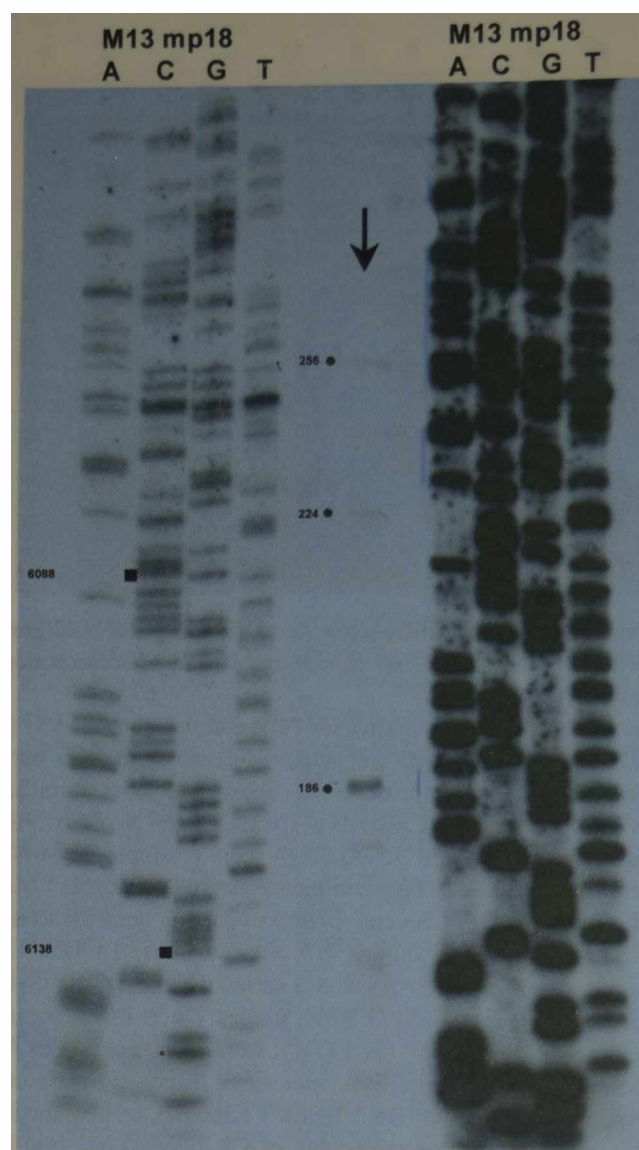


Figura 17.15. Extensión de *primer* y adición de oligonucleótidos con T4 DNA ligasa. Autorradiografía de un gel de acrilamida desnaturalizante con una mezcla de ligación del híbrido RNA-cDNA al dúplex de oligonucleótidos NA/NB. Para la determinación precisa de los tamaños, se sembraron a ambos lados diferentes cantidades de una reacción de secuenciación mediante el método de Sanger (1977), utilizando como molde DNA simple cadena de M13 mp18. Con rectángulos negros, y la numeración respectiva, se indican puntos de referencia en la secuencia de M13. Con puntos negros se indican los productos principales obtenidos.

infectadas con la cepa Candid #1 (#12 y #17). A continuación, se realizó la ligación del dúplex NA/NB con DNA ligasa del fago T4 en presencia de 4 % PEG 8000 y se realizó la amplificación por PCR con el par de *primers* NB y G (Figura 17.16).

En todos los casos se obtuvo el producto mayoritario esperado de ~256 bp y solo en un caso (CD1 #12) se obtuvieron dos fragmentos de amplificación de peso molecular inferior.

Esto puede deberse a una degradación parcial de los extremos 5' viral y/o 5' del mRNA de GPC, ocurrida durante la purificación del RNA, o a la existencia de mayor heterogeneidad en los extremos correspondientes de los RNAs de este lote viral.

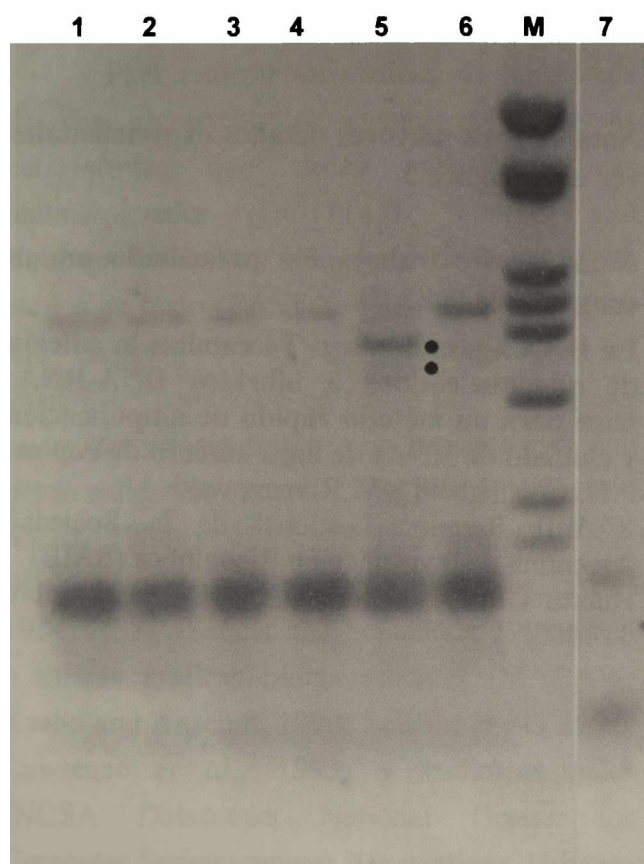


Figura 17.16. Amplificación de cDNAs ligados a oligonucleótidos. En todos los casos, los moldes utilizados corresponden al virus Junín y son diluciones 1/10 de las mezclas de ligación respectivas. **Calle 1:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 mezcla con el dúplex NA/NB. **Calle 2:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 clon 5₂ con el dúplex NA/NB. **Calle 3:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 clon 6₂ con el dúplex NA/NB. **Calle 4:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CD1 viral con el dúplex NA/NB. **Calle 5:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CD1 # 12 con el dúplex NA/NB. **Calle 6:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CD1 # 17 con el dúplex NA/NB. **Calle 7:** control negativo de PCR (sin DNA). Con una flecha negra se indica el fragmento de amplificación esperado: ~256 bp y con puntos negros en la calle 5 se muestran fragmentos adicionales obtenidos. **M:** marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp, 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp. Se muestra la imagen en negativo del gel de agarosa.

Discusión

Se han descrito numerosos métodos para determinar la secuencia nucleotídica 5' de RNAs virales y/o RNAs mensajeros. Sin embargo, algunas de ellas son de baja precisión, mientras que otras requieren mucho tiempo de procesamiento, con varias etapas de purificación de intermedias, para evitar trabajar con productos de longitud incompleta.

En este trabajo se ha descrito un método que no requiere de ninguna etapa de purificación intermedia y que, a diferencia de otras estrategias, elimina la incertidumbre en la secuencia nucleotídica del extremo 5' del RNA, ya que se adicionan oligonucleótidos de secuencia definida.

También se ha demostrado que la DNA ligasa del fago T4 es capaz de catalizar la adición de oligonucleótidos a híbridos RNA-cDNA, con baja eficiencia a oligonucleótidos de simple cadena y con mayor eficiencia a oligonucleótidos de doble cadena. En cualquiera de los casos la adición es más eficiente en la presencia de 4% PEG 8000.

Por otra parte, es interesante notar que la DNA ligasa del fago T4 no requiere imprescindiblemente extremos romos en el dúplex a ligar. En los experimentos presentados en este trabajo se utilizó un dúplex con extremos protruyentes (5 nt. en el extremo de ligación).

Por último, la adición de oligonucleótidos a híbridos RNA-cDNA y la posterior amplificación de estos fragmentos permite obtener cantidades fácilmente clonables de cDNA doble cadena a partir de masas ínfimas de RNA.

Bibliografía

- Bayley, J.M. and Davidson, N. (1976). Methylmercury as a reverside denaturing agent for agarose gel electrophoresis. *Anal. Biochem.* **70**, 75-85.
- Belyavsky, A., Vinogradova, T. and Rajewsky, K. (1989). *Nucleic Acids Res.* **17**, 2919-2932.
- Bertling, W.M., Beier, F. and Reichenberger, E. (1993). Determination of 5' ends of specific messenger-RNAs by DNA ligase-dependent

- amplification. PCR-METHODS AND APPLICATIONS 3, 95-99.
- Chomczynski, P. and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chlorophorm extraction. *Anal. Biochem.* **162**, 156-159.
- Delort, J., Dumas, J.B., Darmon, M.C. and Mallet, J. (1989). *Nucleic Acids Res.* **17**, 6439-6448.
- Doherty, P., Huesca Contreras, M., Dosch, M. and Pan, S. (1989). Rapid amplification of complementary DNA from small amounts of unfractionated RNA. *Anal. Biochem.* **177**, 7-10.
- Dumas, J.B., Milne, E., Delort, J. and Mallet, J. (1991). Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5'ends of mRNAs and for constructing cDNA libraries by *in vitro* amplification. *Nucleic Acids Res.* **19**, 5227-5232.
- Frohman, M.A., Dush, M.K. and Martin, G.R. (1988). *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.
- Ghiringhelli, P.D., Rivera Pomar, R.V., Lozano, M.E., Grau, O. and Romanowski, V. (1991). Molecular organization of Junín virus S RNA: complete nucleotide sequence, relationship with the other members of Arenaviridae and unusual secondary structures. *J. Gen. Virol.* **72**, 2129-2141.
- Grau, O., Franze Fernández, M.T., Romanowski, V., Rustici, S.M. and Rosas, M.F. (1981). Junín virus structure. En: *The replication of negative strand viruses.* (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 11-14.
- Grunstein, M. and Hogness, D.S. (1975). Colony hybridization: a method for the isolation of cloned DNA that contains a specific gene. *Proc. Natl. Acad. Sci. USA* **72**, 3961-3965.
- Harada, K. And Orgel, L.E. (1993). Unexpected substrate-specificity of T4 DNA-ligase revealed by *in vitro* selection. *Nucleic Acids Res.* **21**, 2287-2291.
- King, P.V. and Blakesley, R.W. (1986). Optimizing DNA ligations for transformation. *Focus* **8**(1), 1-3.
- Kornberg, A. and Baker, T.A. (1992). *DNA Replication* 2nd Edition. W.H. Freeman and company, pp. 194-196. pp. 307-322.
- Lambden, P.R., Cooke, S.J., Caul, E.O. and Clarke, I.N. (1992). Cloning of noncultivable human rotavirus by single primer amplification. *J. Virol.* **66**, 1817-1822.
- Loh, Y., Elliott, J.F., Cwirla, S., Lanier, L.L. and Davis, M.M. (1989). *Science* **243**, 217-220.
- Meyer, B.J. and Southern, P.J. (1993). Concurrent sequence analysis of 5' and 3' RNA termini by intramolecular circularization reveals 5' nontemplated bases and 3' terminal heterogeneity for Lymphocytic Choriomeningitis virus mRNAs. *J. Virol.* **67**, 2621-2627.
- Ohara, O., Dorit, R.L. and Gilbert, W. (1989). *Proc. Natl. Acad. Sci. USA* **86**, 5673-5677.
- Rosas, M.F., Romanowski, V. and Grau, O. (1988). The phospholipid composition of arenaviruses. *Anal. Asoc. Quím. Argent.* **76**, 269-283.

Nota 1: para mayores detalles experimentales ver el Apéndice 3

Nota 2: este trabajo fue presentado en un congreso de la SAIB

La DNA ligasa del fago T4 cataliza la adición de oligonucleótidos a híbridos DNA-RNA: base para un método rápido de amplificación y clonado de RNAs de bajo número de copias
P.D. Ghiringhelli y V. Romanowski
XXVIII Reunión Nacional de la Sociedad Argentina de Investigación Bioquímica (SAIB)
Huerta Grande, Córdoba, 18-21 de noviembre de 1992

Capítulo 18

Diseño y programación de rutinas de computación

Rutina HOMGRAPH

(Diseñada y escrita por P.D. Ghiringhelli)

Para realizar un análisis de homología de dos secuencias se cuenta con diversas herramientas que tienen salidas gráficas bidimensionales (DOTPLOT, Collins and Coulson, 1987; von Heijne, 1987; Lesk, 1988; etc.) o de texto (FASTA, BLAST, etc.) (States and Boguski, 1991), mientras que para el análisis de homología de tres o más secuencias se cuenta principalmente con herramientas que tienen salidas de texto (CLUSTAL, Higgins and Sharp, 1988 y 1989; Higgins *et al.*, 1996; PILEUP, UW-GCG package, Deveraux *et al.*, 1984; HSSP, Sander and Schneider, 1991; etc.) o salidas gráficas bidimensionales (MACAW, Karlin and Altschul, 1990; Schuler *et al.*, 1991; Lawrence *et al.*, 1993) y tridimensionales (NCSA Datascope, National Center for Computer Supercomputer Applications, Urbana-Champaign, IL and Spyglass, Inc., Champaign, IL; etc.).

El programa MACAW merece un comentario aparte. Este programa ha sido diseñado para realizar alineamientos locales múltiples (alineamiento que incluye solo las regiones con mayor similitud) y buscar en ellos los dominios conservados, aún cuando sean pequeños. Presenta el problema que solo optimiza los dominios conservados, y no las regiones variables, ni la disposición de los *gaps*.

Los otros programas de alineamiento múltiple mencionados realizan alineamientos globales (incluyen todos los residuos de las secuencias).

Objetivos

⇒ Visualizar gráficamente, en forma panorámica, las regiones constantes y las regiones variables que se encuentran en un

alineamiento global múltiple de secuencias, especialmente cuando estas son grandes.

Fundamentos

La determinación de regiones constantes y variables en proteínas, basada en alineamientos múltiples ya confeccionados, debe considerar la homología de residuos simples y su relación con los vecinos más próximos. Además, de alguna forma se deben maximizar las diferencias entre regiones con diferentes grados de homología (identidad perfecta, cambios conservativos, ausencia de homología) y distribución local de la misma. Esto último se puede lograr trabajando con las frecuencias acumuladas locales. Asimismo, sería útil distinguir, dentro de las regiones constantes, aquellas que presentan mayor grado de conservación (identidad perfecta) de las que presentan cierto grado de variabilidad (cambios conservativos).

Para lograr esto, se puede trabajar:

- directamente con un alineamiento múltiple: complica el manejo de los datos en la computación

- con *profiles* derivados de un alineamiento múltiple: implica una etapa más de computación. Es, tal vez, más sensible que otros métodos, pero algunas veces otorga resultados difíciles de interpretar.

- con archivos simples que representen mediante símbolos, números o cualquier tipo de carácter de texto la información derivada del alineamiento múltiple exclusivamente en tres categorías fijas de datos (identidad perfecta, cambios conservativos y ausencia de homología): aprovecha los datos de salida de la mayor parte de los programas de alineamiento múltiple. Es, tal vez, menos sensible, pero da resultados sencillos y fáciles de interpretar.

Algoritmo

Se trabaja con ventanas de longitud definida por el usuario y el cálculo de la frecuencia acumulada para cada conjunto de residuos coincidente con la longitud de la ventana, desde $1 + \text{Long}_{\text{ventana}}/2$, desplazando la ventana un residuo por vez, hasta $\text{Long}_{\text{secuencia}} - \text{Long}_{\text{ventana}}/2$.

Los valores que se usan son arbitrarios: identidad perfecta = 1,0, cambio conservativo = 0,5, ausencia de homología = -1,0. La frecuencia acumulada se calcula como la sumatoria de los valores individuales de los residuos que conforman la ventana. Esta se normaliza a un rango que va desde -1,0 hasta +1,0 ($\text{Frec. acum.} / \text{Long}_{\text{ventana}}$) y se asigna como valor particular del residuo central de la ventana.

Criterios

Selección de ventanas a utilizar

En principio, el usuario puede definir ventanas de la longitud que desee (> 1 y $< \text{Long}_{\text{secuencia}}$). La selección de la misma debe basarse en un equilibrio entre sensibilidad y sencillez. Las ventanas pequeñas aumentan mucho la sensibilidad, pero también aumentan mucho el “ruido”, complicando el gráfico y su interpretación. Por otra parte, las ventanas muy grandes tienen la ventaja de simplificar el gráfico, pero también reducen la sensibilidad de la búsqueda. Por último, utilizando una ventana

de 35 residuos, en secuencias mayores de 500, se obtienen resultados que conjugan ambos hechos, sensibilidad y simplicidad del gráfico (Figura 18.1)

Selección de regiones variables y conservadas

Una vez confeccionado el gráfico, las preguntas a contestar son:

- todos los picos inferiores a 0 son regiones variables?
- todos los picos superiores a 0 son regiones conservadas?.

Aquí también hay que determinar uno o más criterios para realizar la selección.

Los criterios utilizados fueron:

- (1) En todos los casos, que la base del pico sea $> \text{Long}_{\text{ventana}}$.
- (2) En las regiones variables, además, que el porcentaje de identidad sea $\leq 25\%$ y el porcentaje de variabilidad $\geq 60\%$.
- (3) En las regiones constantes, además, que el porcentaje de identidad sea $\geq 45\%$ y el porcentaje de variabilidad $\leq 20\%$ (ver Tabla 18.a y Figura 18.1).

Para controlar estos resultados se realizaron gráficos de matriz de puntos (Windows DotPlot v.1.0, Nakisa, 1993) con todos los pares de secuencias correspondientes al alineamiento múltiple sobre el cual se trabajó. Este programa, trabaja sólo sobre pares de secuencias y se basa en un algoritmo diferente,

Tabla 18.a. Características de las regiones variables y las regiones conservadas

	Posición	Identidad	Cambios Conserv.	Variabilidad	Total de residuos	% Identidad	% Camb. Conserv.	% Homolog	% Variab.
V 1	156-340	17	36	132	185	9,1	19,5	28,6	71,4
V 2	367-534	10	23	135	168	6,0	13,7	19,7	80,3
V 3	891-1128	38	49	151	238	16,0	20,6	36,6	63,4
V 4	1553-1621	6	17	46	69	8,7	24,6	33,3	66,7
V 5	1697-1804	17	16	75	108	15,7	14,8	30,5	69,5
V 6	1883-2314	42	83	307	432	9,7	19,2	28,9	71,1
C 1	744-789	36	5	5	46	78,4	10,8	89,2	10,8
C 2	1129-1210	46	21	15	82	56,1	25,6	81,7	18,3
C 3	1339-1394	29	17	9	55	52,7	30,9	83,6	19,7
C 4	1425-1500	46	15	15	76	60,6	19,7	80,3	19,7
C 5	1622-1696	35	25	15	75	46,7	33,3	80,0	20,0

pero debería dar resultados similares. Se utilizaron ventanas desde 21 a 35 y umbrales (*threshold*) desde 9 a 17. Los resultados

obtenidos fueron sustancialmente equivalentes a los presentados en esta tesis (datos no mostrados).

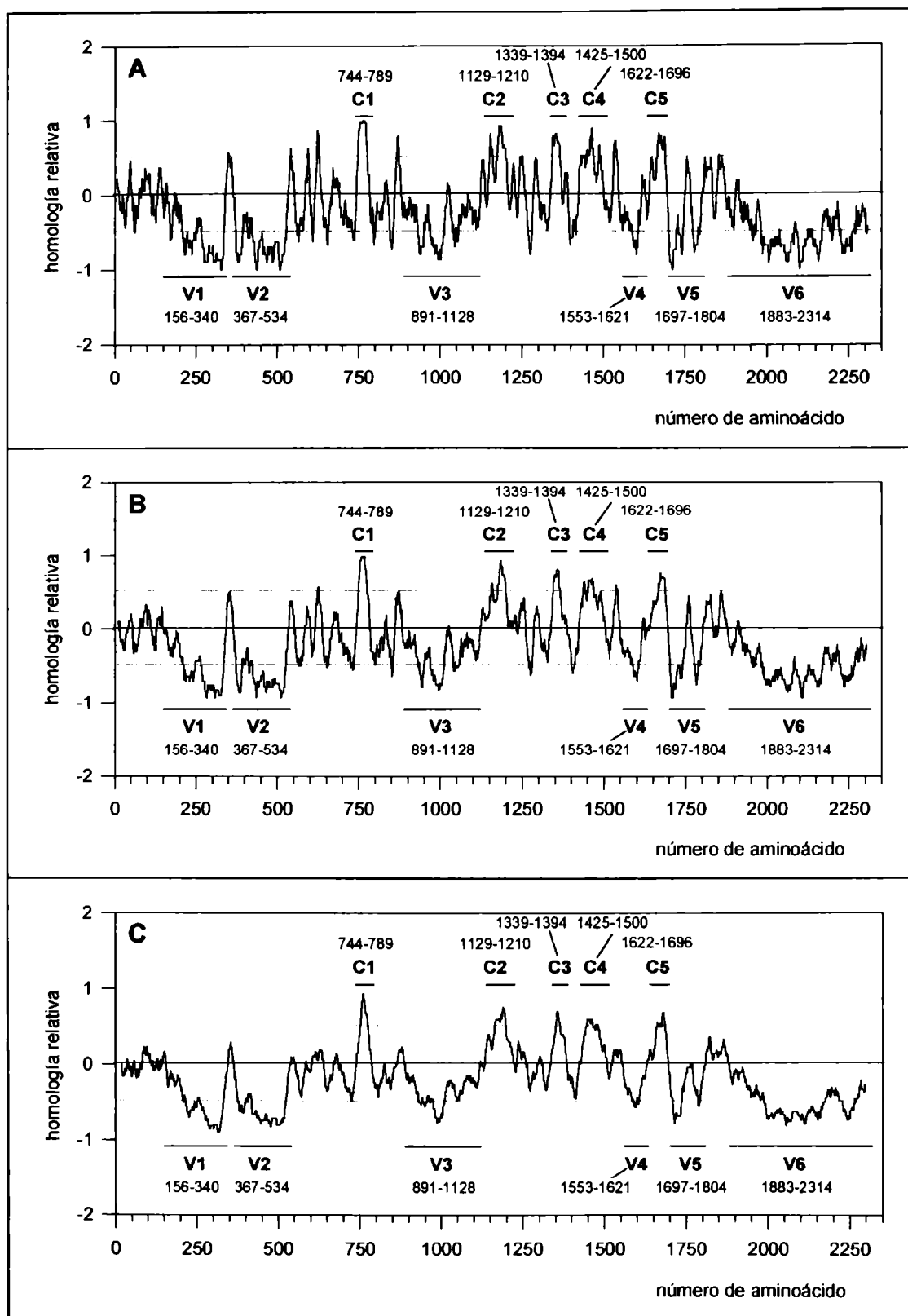


Figura 18.1. Regiones variables y regiones conservadas en las proteínas L de LCM, Pichindé y Tacaribe. Los gráficos, derivados de la rutina HOMGRAPH (Ghiringhelli, P.D., 2002, esta tesis), muestran 6 regiones variables y 5 regiones conservadas. Cada gráfico corresponde a una ventana de longitud diferente: A. ventana de 15 residuos. B. ventana de 21 residuos. C. ventana de 35 residuos. Como se puede observar, la complejidad del gráfico disminuye con la amplitud de la ventana utilizada, mientras que los picos principales se mantienen constantes. Las líneas de guiones (ubicadas en -0,5 y +0,5) sirven de guía rápida para la selección de los picos de importancia, ya que en la práctica todos los picos que las superan (cuando se usa una ventana de 35 residuos) cumplen las condiciones indicadas en el texto (ver **Tabla 18.a**).

Complejidad local de composición (LCC)

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

La complejidad local de composición de un mensaje cualquiera representa la función local de entropía o contenido informativo local.

Los ácidos nucleicos, al igual que las proteínas, son macromoléculas informativas por excelencia. Conocer una secuencia nucleotídica novedosa implica disponer de información particular (en la secuencia *per se*) y general (en la familia de secuencias a la cual pertenece).

Esta información puede ser analizada en forma individual (secuencia por secuencia) o grupal (conjuntos de secuencias relacionadas). En los programas o paquetes de programas de uso frecuente, el único que realiza un análisis de este tipo es el programa COMPLEX (PC/GENE, Intelligenetics; basado en Konopka and Owens, 1990). Dicho programa analiza la complejidad local de la información sobre secuencias individuales y en función de ventanas definidas por el usuario. Además, posee sólo salida gráfica, por lo cual se dificulta la comparación de varias secuencias.

Objetivos

⇒ Contar con herramientas que permitan un análisis grupal de la complejidad de la información sobre secuencias relacionadas.

Fundamentos

La teoría de la información, que estudia las propiedades de los mensajes, permite describir globalmente la información de una macromolécula mediante el cálculo del contenido informativo de la misma (Ec. 1), donde P_i representa la probabilidad de hallar el símbolo i en una determinada posición (States and Boguski, 1991).

$$I = \sum_{i=1}^n P_i \cdot \log_2 P_i \quad \text{Ec. 1}$$

Konopka and Owens (1990) describen una ecuación (Ec. 2), relacionada con la Ec. 1,

con la cual calculan la complejidad de la información en forma local, trabajando sobre ventanas definidas por el usuario. En esta ecuación: L =longitud de la ventana, N_1 =cantidad de A, N_2 =cantidad de C, N_3 =cantidad de G y N_4 =cantidad de T. Los autores utilizan el logaritmo en base 4 para que los resultados estén delimitados entre 0 y 1.

Ec. 2

$$LCC = -\left\{ \left(\frac{1}{L} \right) \left[N_1 \log_4 \left(\frac{N_1}{L} \right) + N_2 \log_4 \left(\frac{N_2}{L} \right) + N_3 \log_4 \left(\frac{N_3}{L} \right) + N_4 \log_4 \left(\frac{N_4}{L} \right) \right] \right\}$$

El patrón de LCC es una propiedad intrínseca de cada secuencia, sin embargo, se pueden formar grupos de diferente complejidad, como los descriptos por los autores:

genes bacterianos > genes de fagos > genes de RNA 5S > exones eucariotas > genes de RNA 16S > exones de levaduras > exones virales > intrones virales > repeticiones tipo Alu > genes de tRNA > regiones no codificantes 5' > intrones eucariotas > intrones de levaduras > repeticiones no Alu > regiones no codificantes 3' > exones de organelas > regiones espaciadoras > intrones de organelas

Por otra parte, el mensaje particular de cada macromolécula es unívoco y la probabilidad de encontrar dos mensajes iguales derivados de una evolución convergente es muy baja. El concepto expresado anteriormente es cierto cuando se consideran los mensajes en referencia a la información particular que ellos portan.

Ahora bien, cuando se considera la organización de la información, en forma independiente del contenido de la misma, es de esperar que la comparación de los patrones de complejidad local de composición (LCC) de macromoléculas relacionadas evolutivamente muestre un cierto grado de similitud estructural. Esta similitud debería guardar correlación con el grado de relación evolutiva.

La comparación y búsqueda de correlación de los patrones de LCC puede realizarse de varias maneras:

- (1) por simple comparación visual de los gráficos individuales
- (2) por análisis comparativo directo de los datos de LCC de cada macromolécula
- (3) por integración y ajuste de los datos de LCC individuales en función de un alineamiento múltiple y generación de un gráfico consenso para el conjunto.

Hipótesis

- ⇒ El patrón de LCC es una propiedad intrínseca de cada secuencia.
- ⇒ Un gráfico consenso generado a partir de un conjunto de secuencias relacionadas evolutivamente, debería presentar un alto grado de similitud con cada uno de los gráficos individuales.
- ⇒ Un gráfico consenso generado a partir de un conjunto de secuencias con escasa o nula relación evolutiva, debería presentar grandes diferencias con cada uno de los gráficos individuales.
- ⇒ El patrón consenso generado es altamente dependiente del alineamiento múltiple.

Algoritmo

Se trabaja con ventanas de longitud definida por el usuario y se utiliza una variante de la ecuación de Konopka and Owens (1990) para calcular el índice de LCC (Ec. 3) para cada conjunto de residuos coincidente con la longitud de la ventana, desde $1 + \text{Long}_{\text{ventana}}/2$, desplazando la ventana un residuo por vez, hasta $\text{Long}_{\text{secuencia}} - \text{Long}_{\text{ventana}}/2$. Este cálculo se realiza para cada una de las secuencias del grupo, en forma individual.

Ec. 3

$$\text{LCC} = \left\{ \left\{ \frac{1}{L} \cdot [N1 \cdot \log_2(N1/L) + N2 \cdot \log_2(N2/L) + N3 \cdot \log_2(N3/L) + N4 \cdot \log_2(N4/L)] \right\}^{-1} \right\} / 2$$

Donde: N1=cantidad de A; N2=cantidad de C;
N3=cantidad de G; N4=cantidad de T

Para el cálculo se usa el mismo valor (1,0) para cualquiera de los cuatro nucleótidos (N1 a N4). El índice de LCC calculado, se transforma en valor positivo, se normaliza a un rango que va desde 0,0 hasta 1,0 y se asigna como valor particular del residuo central de la ventana. Un ejemplo de archivo de salida de esta rutina se muestra en la **Figura 18.2**.

```

Posic.  LCC
      26 ;0,9729
      27 ;0,9729
      28 ;0,9729
      29 ;0,9721
      30 ;0,9601
.....
.....
    3383 ;0,9759
    3384 ;0,9771
    3385 ;0,9759
    3386 ;0,9819
    3387 ;0,9846
    
```

Figura 18.2. Ejemplo de archivo de salida de la rutina de cálculo individual de LCC.

A continuación, se trabaja sobre el conjunto de archivos individuales con datos de LCC de las secuencias del grupo y un alineamiento múltiple de las mismas secuencias (**Figura 18.3**), para ajustar los datos al alineamiento, generando un archivo de salida como el de la **Figura 18.4**.

En este punto es interesante considerar que, desde el punto de vista evolutivo, los *gaps* deben ser considerados como un carácter informativo adicional, razón por la cual se incorpora en la ecuación el término N5 (cantidad de GAPS) (Ec. 4). Dado que su importancia no es igual a la de cualquier nucleótido, se introduce una penalización a la existencia de *gaps* otorgándoles 1/5 del valor dado a un nucleótido (0,2).

Ec. 4

$$\text{LCC} = \left\{ \left\{ \frac{1}{L} \cdot [N1 \cdot \log_2(N1/L) + N2 \cdot \log_2(N2/L) + N3 \cdot \log_2(N3/L) + N4 \cdot \log_2(N4/L) + N5 \cdot \log_2(N5/L)] \right\}^{-1} \right\} / 2$$

Donde: N1=cantidad de A; N2=cantidad de C;
N3=cantidad de G; N4=cantidad de T;
N5=cantidad de GAPS

```

FILE-1.LCC -TGCAGTAAGGGGATCCTAGGC-GATTTT-GG--TAACGC-----
FILE-2.LCC CTGCAGTAAGGGGATCCTAGGC-GATTTT-GG--TAACGC-----
FILE-3.LCC -CGCA--CCGGGGATCCTAGGC---TTTTAA--CTAACGCATTACAGAGGGATTCAACCC
FILE-4.LCC -CGCA--CCGGGGATCCTAGGCATACCTT-GG---ACGC-----
FILE-5.LCC GCGCA--CCGGGGATCCTAGGCGTTTTTTAG--TC-ACGC-TTA-----
FILE-6.LCC C-GCA--CCGGGGATCCTAGGCATTCTT-G--TC-----
FILE-7.LCC GCGCA--CCGGGGATCCTAGGC-ATTTTT-GG-TT-GCGC-----
FILE-8.LCC GC--A--CCGGGGATCCTAGGC-A--TTTAGGATT-GCGC-----
FILE-9.LCC -CGCA--CCGGGGATCCTAGGC--TTTTT-GGATT-GCGC-----
FILE10.LCC -CGCA--CCGGGGATCCTAGGC--TTTTT-GGATT-GCGC-----
FILE11.LCC --GT-----GGATCCTAGGC--TTTTT-GG-TT-GCGC-----
.....*****.....

FILE-1.LCC TATA-AGTT----GTTACTGCTTTCTATTTGGACAACATCAAACCATCTATT-----GT
FILE-2.LCC TGTA-AGTT----GTTACTGCTTTCTATTTGGACAACATCAAACCATCTATT-----GT
FILE-3.LCC TGTGCAATTCAAAGTCACAGTTTTTGTGACAGTTAGTTGTCAACCCCTTTTTGACAACC
FILE-4.LCC -GCATATTTCTTGATCAAA-----
FILE-5.LCC -----AAT-----CTTT-----GATTGCGTCAA-----
FILE-6.LCC --CAT-ATT-----TGCCT---AACTGAACCAGGT-GAATCAC---TC-----CC
FILE-7.LCC -----AATTC-----AA-----GTGTCCTATTTAA--
FILE-8.LCC -----ATTTT-----AAAACCTCCTTTTTGG-----AAA--GTGTCGCAATCAG--
FILE-9.LCC -----TTTCC---TCTAGATC---AACTGG---GTGTCAGGCCCTATCCTACAGAAG-
FILE10.LCC -----TTTCC---TTTAGGAC---AACTGG---GTGCTGGATTCTATCCAGTAAAG-
FILE11.LCC -----ATTC-----TAGA--GCATCTCG---GA--
.....***

```

Figura 18.3. Ejemplo de alineamiento múltiple a emplear para ajustar los datos individuales de LCC. La cantidad de secuencias en el alineamiento, el número de nucleótidos por línea y el número de caracteres de identificación son variables dependientes del usuario. Desde el punto de vista del correcto funcionamiento de la rutina, es crucial que para identificar cada secuencia se utilice el nombre completo del archivo individual con los datos de LCC (incluyendo la extensión).

```

Posic. Sec.1; Sec.2;...; Sec.N
26 ;0.8576;0.8726;...;0.7071
27 ;0.8576;0.9729;...;0.7071
28 ;0.9772;0.9729;...;0.7071
29 ;0.9772;0.9729;...;0.6959
30 ;0.8221;0.8221;...;0.6863
31 ;0.9753;0.9721;...;0.6863
32 ;0.9633;0.9601;...;0.6863
33 ;0.7879;0.7879;...;0.6863
34 ;0.7729;0.7729;...;0.6863
35 ;0.9694;0.9648;...;0.6863
.....
3616 ;0.8446;0.8446;...;0.8033
3617 ;0.8473;0.8473;...;0.8086
3618 ;0.9759;0.9759;...;0.8009
3619 ;0.9759;0.9759;...;0.8053
3620 ;0.9798;0.9798;...;0.8053
3621 ;0.9759;0.9759;...;0.7985
3622 ;0.9771;0.9771;...;0.8135
3623 ;0.9759;0.9759;...;0.9996
3624 ;0.9819;0.9819;...;0.9996
3625 ;0.9846;0.9846;...;0.9985

```

Figura 18.4. Ejemplo de archivo de salida de la rutina de ajuste. En este archivo, cada columna corresponde a una secuencia particular, y las distintas filas a las sucesivas columnas de datos del alineamiento múltiple.

Por último, utilizando el archivo de salida de la etapa anterior (Figura 18.4), se calcula la media aritmética y la desviación estándar para cada línea de datos, generando un archivo ASCII (Figura 18.5) que puede ser importado con diversos programas de graficación (por ej. SIGMAPLOT) para construir el gráfico consenso final.

Posic	Total	X	$\sigma(n-1)$	$X-1\sigma$	$X+1\sigma$
26;	9,176;	0,834;	0,060;	0,774;	0,894
27;	9,248;	0,841;	0,075;	0,766;	0,916
28;	9,415;	0,856;	0,096;	0,759;	0,952
29;	9,972;	0,907;	0,107;	0,800;	1,014
30;	8,975;	0,816;	0,083;	0,733;	0,899
.....
3621;	10,064;	0,915;	0,084;	0,831;	0,998
3622;	10,131;	0,921;	0,077;	0,844;	0,998
3623;	10,495;	0,954;	0,059;	0,895;	1,013
3624;	10,759;	0,978;	0,019;	0,959;	0,998
3625;	10,765;	0,979;	0,020;	0,958;	0,999

Figura 18.5. Ejemplo de archivo de salida de la rutina de estadística.

Criterios

Selección de ventanas a utilizar

En principio, el usuario puede definir ventanas de la longitud que desee (> 1 y $< \text{Long}_{\text{secuencia}}$). La selección de la misma debe basarse en un equilibrio entre sensibilidad y sencillez. Utilizando una ventana de 51 residuos se obtienen resultados que conjugan ambos hechos, sensibilidad y simplicidad del gráfico (Figura 18.6)

Confiabilidad del método

Para comprobar la equivalencia entre las ecuaciones utilizadas aquí y la original de Konopka & Owens (1990), se efectuó el análisis con el programa COMPLEX (PC/GENE, Intelligenetics) y con las rutinas postuladas en esta tesis, utilizando el RNA S del virus Junín Candid #1 como modelo. Como se puede observar en la Figura 18.7 los resultados obtenidos con ambas rutinas son idénticos. También se comprobó la coincidencia con los postulados de Konopka & Owens (1990) en cuanto al orden de los niveles de complejidad (Figura 18.8).

Gráficos consenso y diferenciación de grupos

Con la finalidad de comprobar la sensibilidad del método para diferenciar grupos de secuencias con distinto grado de relación evolutiva, se analizó un conjunto de secuencias virales de RNAs S de la familia *Bunyaviridae* (Tabla 18.b).

Tabla 18.b. RNAs S de la familia Bunyaniviridae

Abrev.	Nombre	Longitud en nt.	Referencias
AIN	Aino	850	Akashi <i>et al.</i> , 1984
BAT	Batai	946	Dunn <i>et al.</i> , 1994
GER	Germiston	980	Gerbaud <i>et al.</i> , 1987
INK	Inkoo	986	Vapalahti <i>et al.</i> , 1996
JAM	Jamestown Canyon	991	Bowen <i>et al.</i> , 1995
JSU	Jerry Slough	935	Bowen <i>et al.</i> , 1995
KEY	Keystone	937	Bowen <i>et al.</i> , 1995
LAC	La Crosse	984	Patterson <i>et al.</i> , 1983
MAG	Maguari	945	Elliott <i>et al.</i> , 1989
MEL	Melao	941	Bowen <i>et al.</i> , 1995
MOR	Morro Bay	976	Fulhorst <i>et al.</i> , 1996
TAH	Tahyna	977	Vapalahti <i>et al.</i> , 1996
TRI	Trivittatus	973	Bowen <i>et al.</i> , 1995

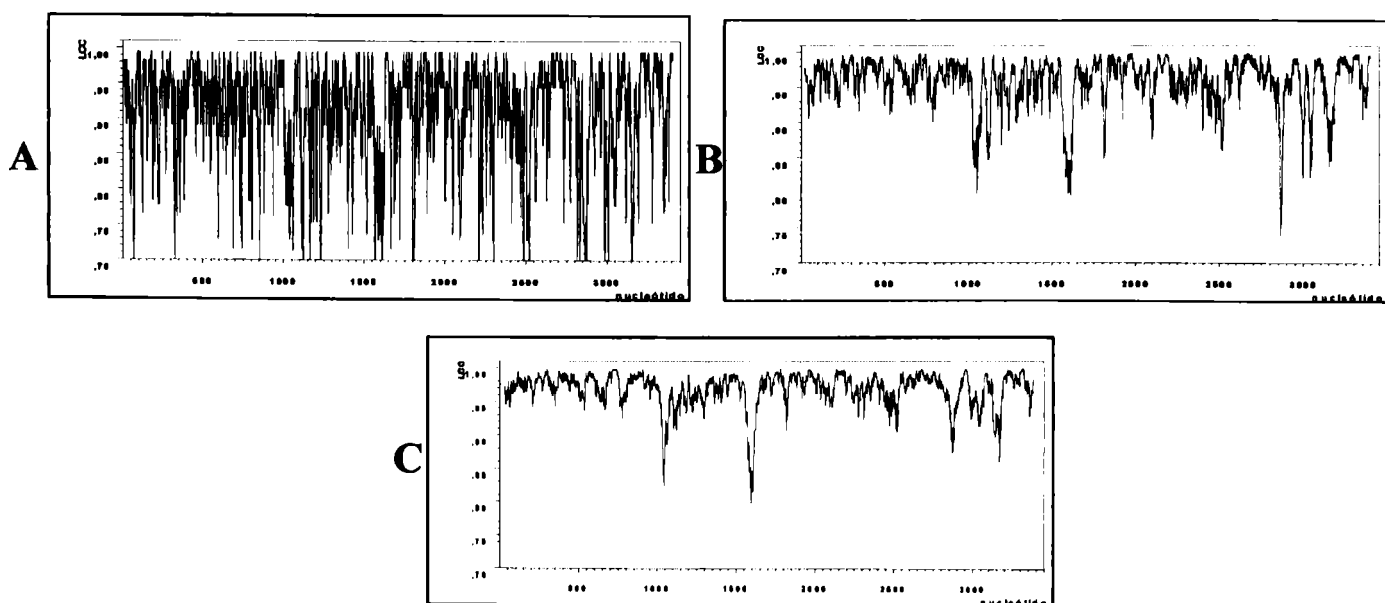


Figura 18.6. Gráficos de complejidad local de composición (LCC), derivados del paquete de rutinas de cálculo y análisis múltiple de LCC (Ghiringhelli, P.D., 2002, esta tesis). Los tres gráficos corresponden a la secuencia nucleotídica del virus Junin-Candid #1, en cada uno de ellos se utilizó una ventana de cálculo diferente. A. Ventana de 15 residuos. B. Ventana de 35 residuos. C. Ventana de 51 residuos. Como se puede observar, la complejidad del gráfico disminuye a medida que aumenta la amplitud de la ventana utilizada. Además, con las ventanas de 35 o 51 residuos, se logran discriminar mejor las principales regiones de escasa complejidad local.

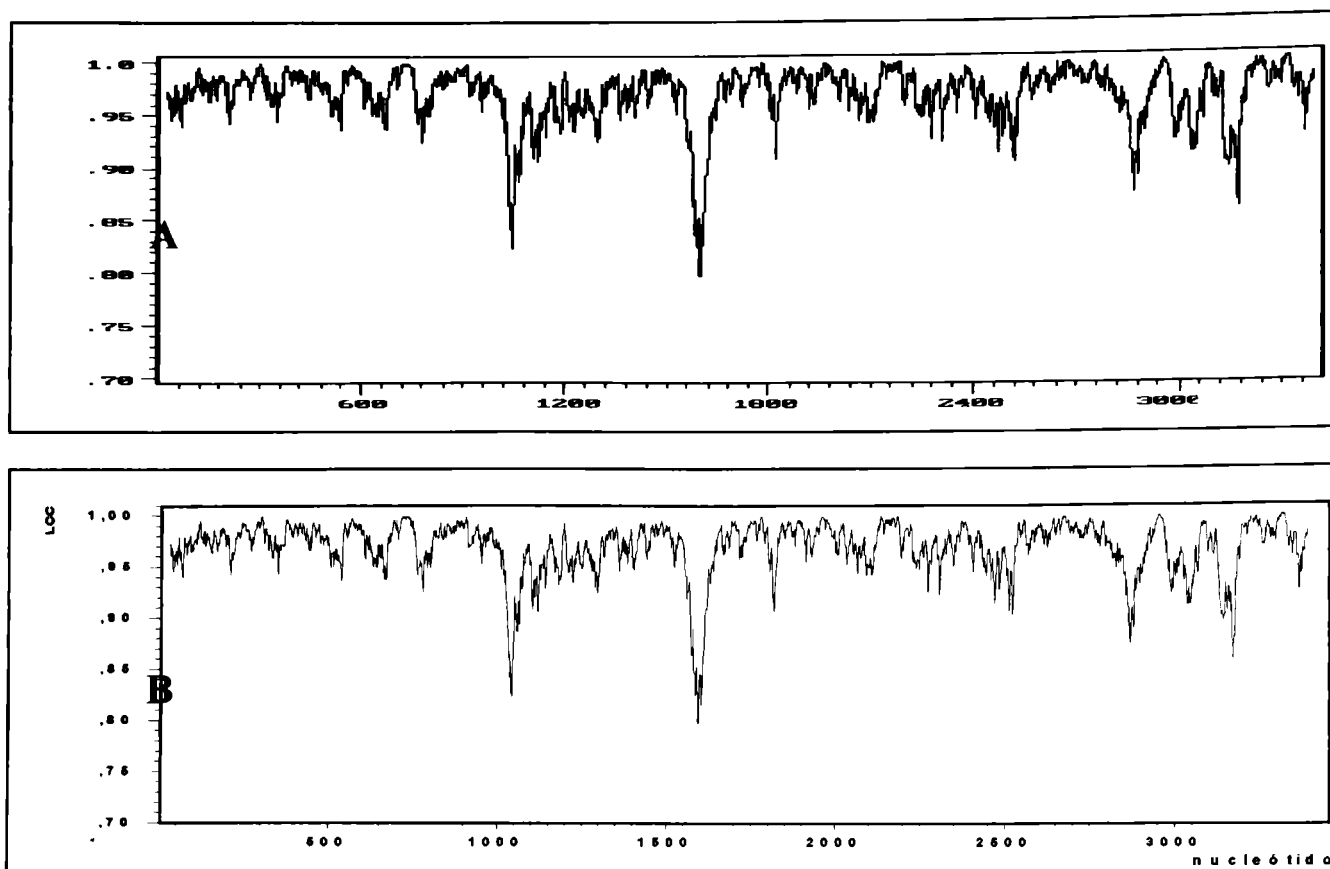


Figura 18.7. Gráficos de complejidad local de composición (LCC) del RNA S del virus Junin-Candid #1. **A.** Gráfico derivado del programa COMPLEX (Konopka and Owens, 1990; PC/GENE, Intelligenetics). **B.** Gráfico derivado del paquete de rutinas de cálculo y análisis múltiple de LCC (Ghiringhelli, P.D., 2002, esta tesis). En los dos gráficos se utilizó una ventana de 51 residuos. Como se puede observar, ambos gráficos son idénticos.

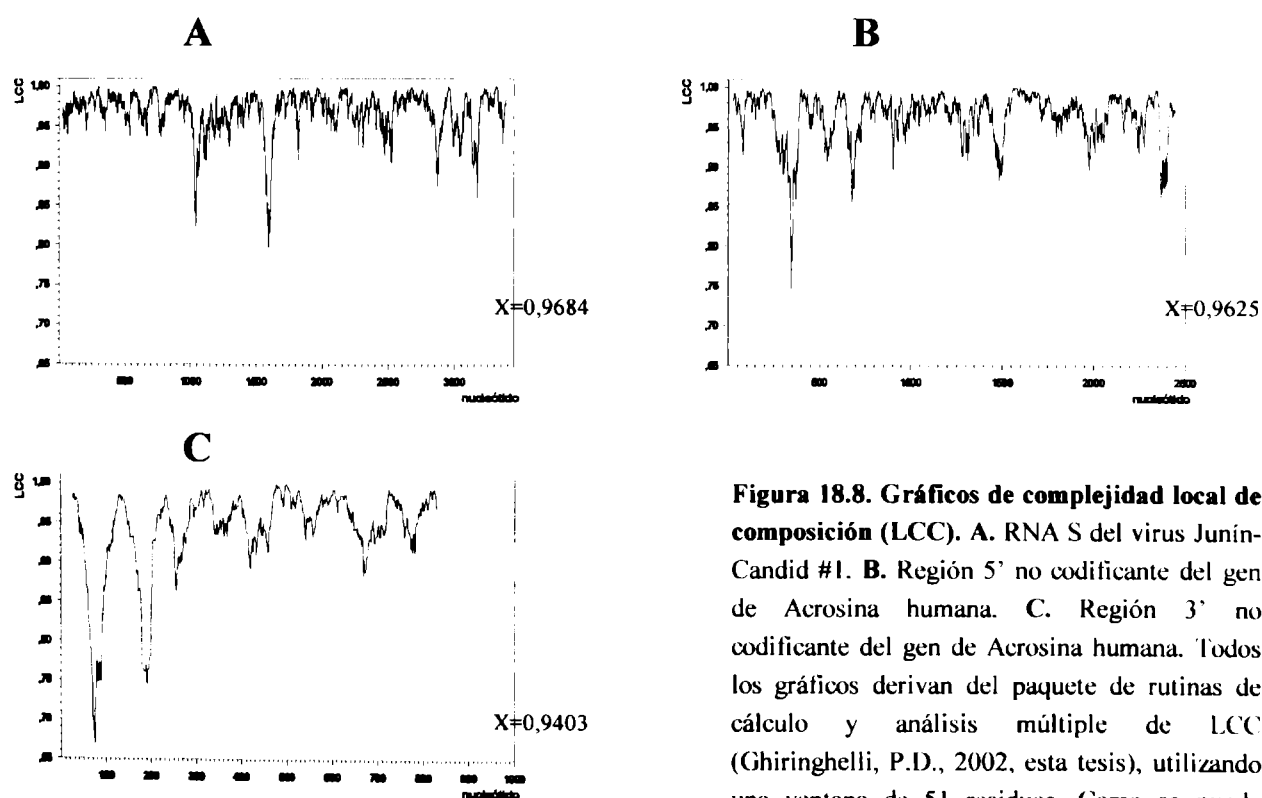


Figura 18.8. Gráficos de complejidad local de composición (LCC). **A.** RNA S del virus Junin-Candid #1. **B.** Región 5' no codificante del gen de Acrosina humana. **C.** Región 3' no codificante del gen de Acrosina humana. Todos los gráficos derivan del paquete de rutinas de cálculo y análisis múltiple de LCC (Ghiringhelli, P.D., 2002, esta tesis), utilizando una ventana de 51 residuos. Como se puede observar, el orden de niveles de complejidad es:

$$A > B > C$$

Con el conjunto de secuencias de bunyavirus se realizó el análisis individual y se confeccionó el gráfico consenso de LCC mediante el ajuste a un alineamiento múltiple de las 13 secuencias (**Figura 18.9.A**). A continuación se efectuó la comparación entre el

gráfico consenso y los gráficos individuales. El resultado de esta comparación (**Figura 18.9**) permite concluir que el gráfico consenso no se corresponde en su totalidad con ninguno de los gráficos individuales.

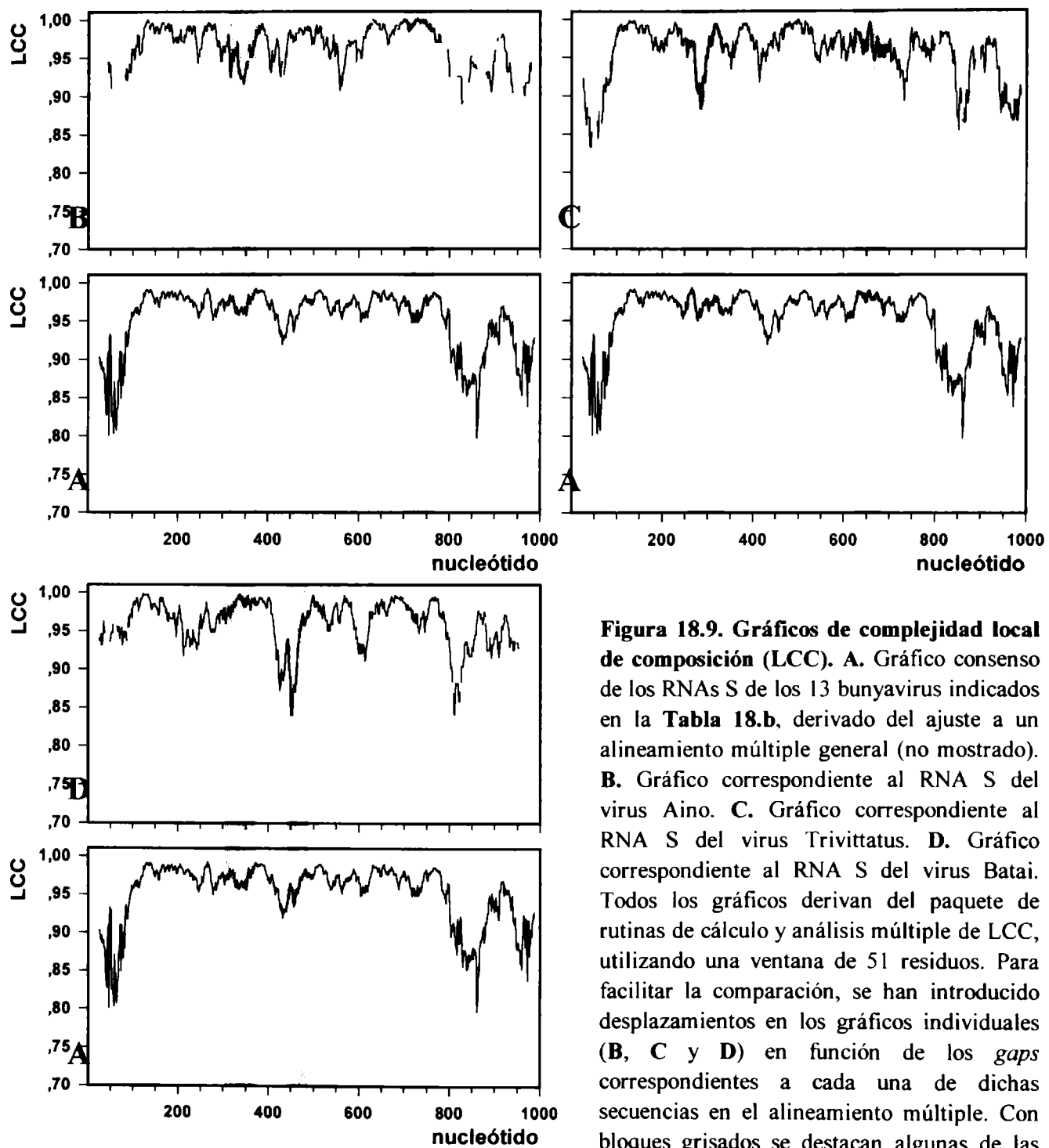


Figura 18.9. Gráficos de complejidad local de composición (LCC). A. Gráfico consenso de los RNAs S de los 13 bunyavirus indicados en la **Tabla 18.b**, derivado del ajuste a un alineamiento múltiple general (no mostrado). B. Gráfico correspondiente al RNA S del virus Aino. C. Gráfico correspondiente al RNA S del virus Trivittatus. D. Gráfico correspondiente al RNA S del virus Batai. Todos los gráficos derivan del paquete de rutinas de cálculo y análisis múltiple de LCC, utilizando una ventana de 51 residuos. Para facilitar la comparación, se han introducido desplazamientos en los gráficos individuales (B, C y D) en función de los *gaps* correspondientes a cada una de dichas secuencias en el alineamiento múltiple. Con bloques grisados se destacan algunas de las principales regiones en donde se encuentran diferencias entre los patrones individuales respectivos y el patrón consenso.

Tabla 18.c. Homología nucleotídica de 13 RNAs S de bunyavirus

	BAT	GER	INK	JAM	JSU	KEY	LAC	MAG	MEL	MOR	TAH	TRI
AIN	40,4	40,3	50,2	49,7	46,8	47,4	48,4	40,4	46,7	47,9	47,0	46,2
BAT		71,2	48,7	48,1	44,1	45,6	49,2	79,3	46,8	47,0	46,7	46,9
GER			50,8	49,6	46,4	47,6	50,1	69,8	46,8	50,2	48,9	49,5
INK				89,7	84,7	73,5	73,9	48,5	74,6	74,6	74,2	69,3
JAM					88,1	73,8	73,0	48,0	73,9	73,8	73,2	69,2
JSU						70,8	67,9	44,3	71,2	69,4	69,0	64,5
KEY							67,9	46,1	75,5	68,7	68,8	67,1
LAC								47,7	69,3	77,9	79,1	69,0
MAG									45,2	47,1	47,6	47,1
MEL										69,4	70,2	66,2
MOR											79,5	68,1
TAH												68,1

AIN: Aino; BAT: Batai; GER: Germiston; INK: Inkoo; JAM: Jamestown Canyon; JSU: Jerry Slough; KEY: Keystone; LAC: La Crosse; MAG: Maguari; MEL: Melao; MOR: Morro Bay; TAH: Tahyna; TRI: Trivittatus.
 En grisado oscuro: grupo mayoritario (INK, JAM, JSU, KEY, LAC, MEL, MOR, TAH, TRI).
 En grisado claro: grupo minoritario (BAT, GER, MAG).

En función de estos resultados, se analizó en detalle el grado de homología del conjunto de RNAs S de bunyavirus utilizado. Este análisis permite observar que la homología entre los distintos pares de secuencias varía en un rango muy amplio (40,4% - 89,7%). Con la finalidad de reducir la dispersión, se definieron tres subconjuntos de secuencias (Tabla 18.c): Aino, cuya homología con el resto varía entre 40,3% y 50,2%; grupo mayoritario, cuya homología dentro del subconjunto varía entre 64,5% y 89,7%; y grupo minoritario, cuya homología dentro del subconjunto varía entre 69,8% y 79,3%.

A continuación, se repitió el análisis individual y la confección de los gráficos consenso para los grupos mayoritario y minoritario, encontrándose que el gráfico consenso de cada grupo coincide en gran medida con los gráficos individuales de los cuales deriva (Figuras 18.10 y 18.11).

Los resultados de este estudio permiten concluir que:

El gráfico consenso generado a partir de un conjunto de secuencias muy relacionadas evolutivamente, presenta un alto grado de similitud con cada uno de los gráficos individuales.

El gráfico consenso generado a partir de un conjunto de secuencias relacionadas, biológicamente y evolutivamente, pero con gran

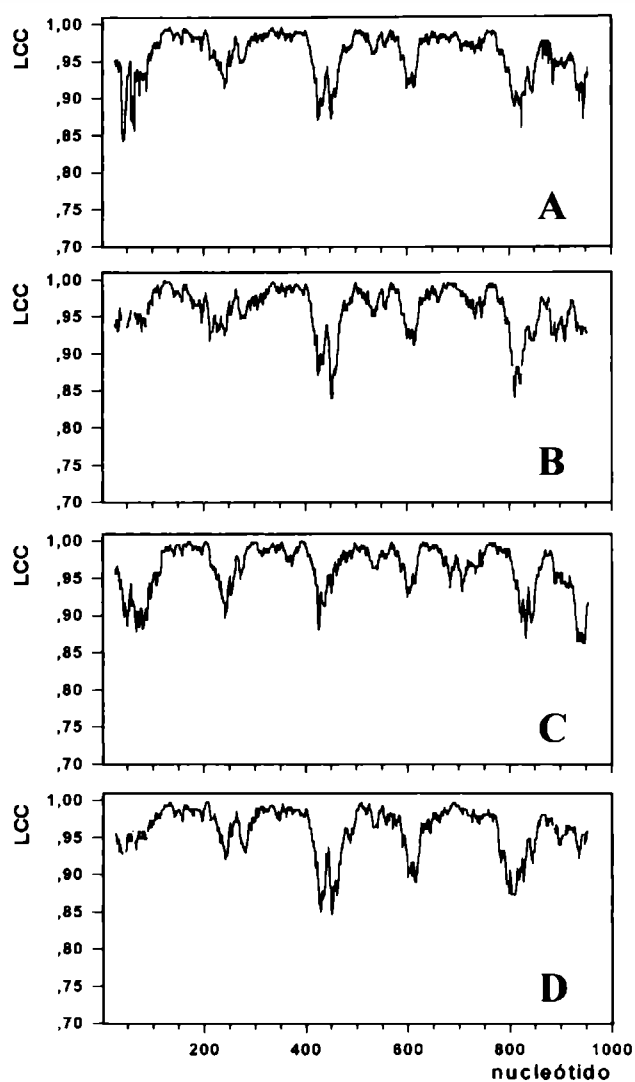


Figura 18.10. Gráficos de LCC del grupo minoritario. A. Gráfico consenso del grupo. B. Gráfico correspondiente al RNA S del virus Batai. C. Gráfico correspondiente al RNA S del virus Germiston. D. Gráfico correspondiente al RNA S del virus Maguari. Para facilitar la comparación, se han introducido desplazamientos en los gráficos individuales (B, C y D) en función de los *gaps* correspondientes a cada una de dichas secuencias en el alineamiento múltiple (ver Apéndice 2).

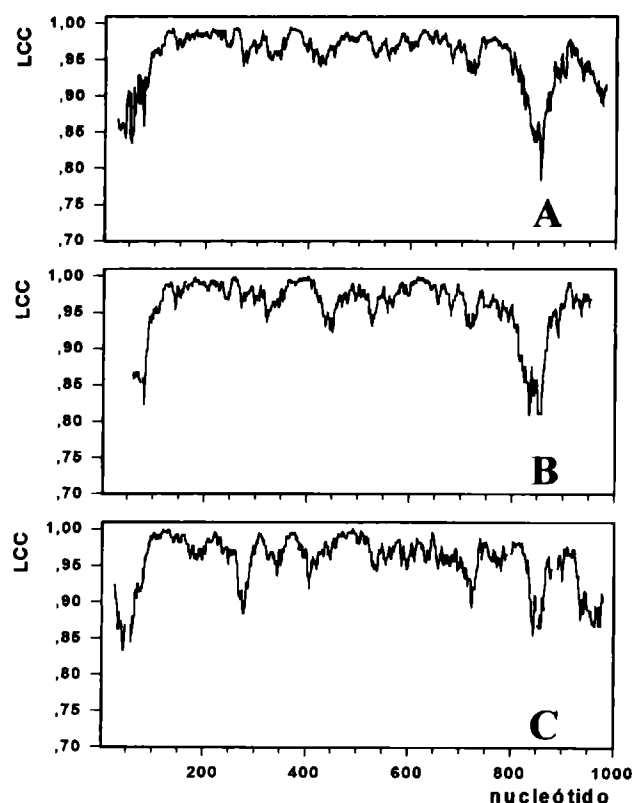


Figura 18.11. Gráficos de LCC del grupo mayoritario. A. Gráfico consenso del grupo. B. Gráfico correspondiente al RNA S del virus Jerry Slough. C. Gráfico correspondiente al RNA S del virus Trivittatus. Para facilitar la comparación, se han introducido desplazamientos en los gráficos individuales (B y C) en función de los *gaps* correspondientes a cada una de dichas secuencias en el alineamiento múltiple (ver Apéndice 2).

dispersión en las homologías individuales, presenta grandes diferencias con cada uno de los gráficos individuales.

Comparación cuantitativa de LCCs

Otra manera, menos intuitiva, de establecer relaciones entre patrones de LCC es cuantificar las comparaciones de LCCs, con independencia de los alineamientos múltiples.

Como los gráficos de LCC se caracterizan por ser de tipo sinusoidal, con variación en las intensidades, una de las alternativas para realizar la comparación cuantitativa es aplicar la transformada de Fourier sobre ventanas de longitud equivalente a las utilizadas para el cálculo de LCC. Esta estrategia se puede aplicar, tanto en ventanas solapadas como en ventanas adyacentes.

En principio, se calcula la transformada de Fourier utilizando como punto de partida un archivo ASCII sin formato, derivado del que

tiene los datos individuales de LCC (Figura 18.2). Para cada ventana, la transformada se calcula de acuerdo a la Ec. 5, entre 0° y 180° (Figura 18.12), con una frecuencia determinada por el usuario.

Ec. 5 (ver Apéndice 5, bases matemáticas)

$$F = \left\{ \left[\sum_{k=0}^K (L_k - L) \cos k\omega \right]^2 + \left[\sum_{k=0}^K (L_k - L) \text{sen } k\omega \right]^2 \right\}^{1/2}$$

Donde: K = posición en la ventana
 L = valor de LCC; L = media de los valores de LCC de la ventana
 ω = frecuencia de análisis

Además, en cada punto se filtra el ruido estándar, restando las amplitudes previamente calculadas sobre un grupo de secuencias de igual composición nucleotídica, pero de organización aleatoria. De esta manera, se generan nuevos archivos ASCII, individuales para cada secuencia, con los datos correspondientes a todas las ventanas (Figura 18.13).

A continuación, se realiza la comparación entre las secuencias, ventana a ventana y punto por punto, en base a los valores obtenidos con el cálculo anterior. Para ello, se utilizan como archivos de entrada los generados en la etapa anterior (Figura 18.13) y, para cada ventana, se calcula el grado de similitud aplicando las Ecs. 6 y 7.

Ec. 6

$$C_n = \frac{\sum_{f=1}^N A_f^{(n)} - B_f^{(n)}}{\sum_{f=1}^N A_f^{(n)} + B_f^{(n)}}$$

Donde: f = frecuencia
 A = amplitud de la frecuencia en la sec. A
 B = amplitud de la frecuencia en la sec. B
 N = total de frecuencias utilizadas
 C = valor adimensional de la comparación

Ec. 7

$$S_v = \left[1 - \frac{C_n}{N} \right] * 100$$

Donde: C = valor adimensional de la comparación de la ventana n
 S_v = similitud local de la ventana

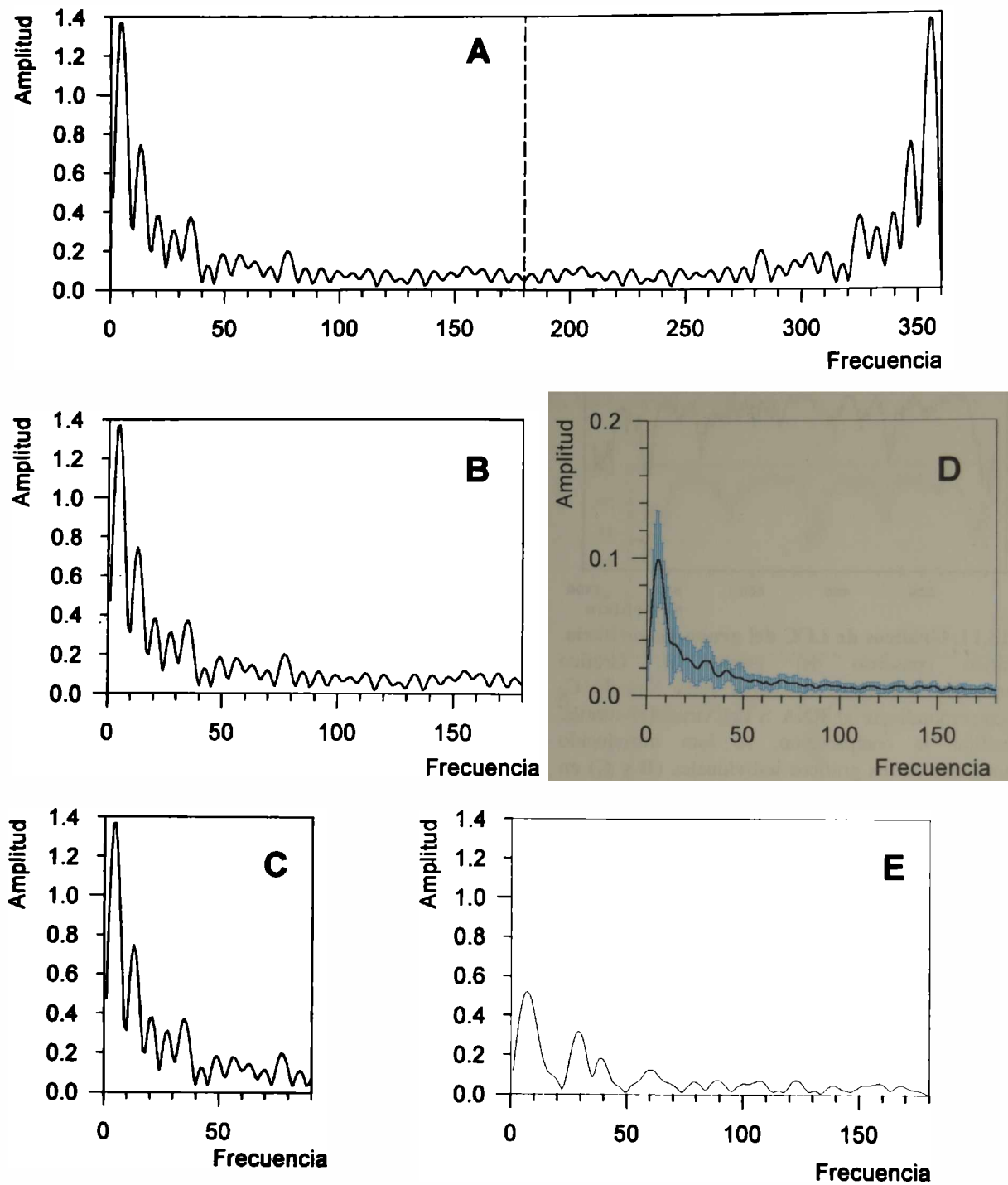


Figura 18.12. Estudio del rango de frecuencias para análisis de Fourier. En las ventanas A, B y C se grafican los perfiles de amplitudes características de una ventana de 51 nucleótidos (posiciones 1582 a 1632 V), correspondiente al RNA S de la cepa Candid #1 del virus Junín. En la ventana A se muestra el perfil correspondiente al rango completo de frecuencias (0° a 360°), observando que el perfil es especular a partir de los 180° (línea vertical de guiones). En la ventana B se muestra el perfil correspondiente al rango entre 0° y 180° . En la ventana C se muestra el perfil correspondiente al rango entre 0° y 90° . En la ventana D se muestra el perfil correspondiente al análisis de un conjunto de secuencias de composición nucleotídica y longitud equivalente a las otras, pero de organización aleatoria. La línea llena negra representa la media estadística de todas las secuencias y sus correspondientes ventanas de 51 nt., las líneas grisadas verticales representan la dispersión ($+1\sigma$ y -1σ) de los datos. Como puede observarse, en la mayor parte del gráfico la dispersión es pequeña. Por otra parte, a fines comparativos, debe notarse que la escala del eje Y está expandida 7 veces, respecto a los otros gráficos. En la ventana E se grafica el perfil de amplitudes características de una ventana de 51 nucleótidos (posiciones 358 a 408 V), correspondiente al RNA S de la cepa Nigeria del virus Lassa. En todos los casos se observa que la amplitud más importante corresponde al rango entre 0° y 15° . Sin embargo, en otras regiones de los genomas (datos no mostrados) existen amplitudes características, sensiblemente menores a la anterior, pero que deberían considerarse en el análisis.

FRECUENC.	POWER	SQRPOWER	FSQRPOWER
Ventana: 1			
0.00	0.0000	0.0000	0.0000
1.00	0.0534	0.2310	0.2050
2.00	0.1911	0.4372	0.3872
3.00	0.3564	0.5970	0.5260
4.00	0.4832	0.6951	0.6091
.....			
176.00	0.0014	0.0377	0.0317
177.00	0.0016	0.0399	0.0349
178.00	0.0014	0.0377	0.0327
179.00	0.0011	0.0336	0.0296
180.00	0.0010	0.0314	0.0274
.....			
Ventana: N			
0.00	0.0000	0.0000	0.0000
1.00	0.0035	0.0594	0.0334
2.00	0.0123	0.1110	0.0610
3.00	0.0220	0.1484	0.0774
4.00	0.0281	0.1677	0.0817
.....			
176.00	0.0000	0.0056	0.0000
177.00	0.0000	0.0062	0.0012
178.00	0.0000	0.0059	0.0009
179.00	0.0000	0.0050	0.0010
180.00	0.0000	0.0045	0.0005

Figura 18.13. Ejemplo de archivo de salida de la rutina de cálculo de la transformada de Fourier. FRECUENC.: frecuencia de análisis. POWER: transformada de Fourier. SQRPOWER: raíz cuadrada de la transformada. FSQRPOWER: SQRPOWER filtrado.

Archivos procesados:	
AIN-51.FOU	BAT-51.FOU
Ventana: 1	
Semejanza:	72.5 %
Ventana: 2	
Semejanza:	71.7 %
Ventana: 3	
Semejanza:	63.6 %
Ventana: 4	
Semejanza:	66.1 %
Ventana: 5	
Semejanza:	68.4 %
Ventana: 6	
Semejanza:	62.3 %
Ventana: 7	
Semejanza:	63.0 %
Ventana: 8	
Semejanza:	51.8 %
Ventana: 9	
Semejanza:	62.7 %
Ventana: 10	
Semejanza:	66.7 %
Ventana: 11	
Semejanza:	57.6 %
Ventana: 12	
Semejanza:	68.6 %
Suma total:	775.030
Similitud:	64.6 %

Figura 18.14. Ejemplo de archivo de salida de la rutina de comparación de amplitudes de Fourier. Las ventanas están numeradas en forma correlativa. Semejanza: es S_v . Suma total es: $\sum S_v$. Similitud: es la similitud global S_G .

Por último, se calcula un porcentaje de similitud global en función de la Ec. 8.

Ec. 8

$$S_G = \sum_{v=1}^V S_v$$

Donde: S_v = similitud local de cada ventana
 S_G = similitud global
 V = cantidad total de ventanas

De esta manera, se genera un nuevo archivo de salida, para cada par, a partir del cual se puede extraer la información de similitud de patrones de LCC (Figura 18.14).

Para comprobar la consistencia de esta metodología, se aplicó el análisis cuantitativo sobre las 13 secuencias de RNAs S de bunyavirus ya utilizadas (Tabla 18.b). En todos los casos, se aplicó el análisis en un esquema de ventanas adyacentes de 51 nt. de longitud.

En la Tabla 18.d se puede observar que las relaciones de similitud obtenidas son prácticamente coincidentes con las obtenidas mediante análisis de homología nucleotídica (Tabla 18.c). En la Figura 18.15 se muestra el análisis comparativo de algunas ventanas como ejemplo.

En síntesis, esta metodología de análisis resulta promisorio desde el punto de vista del análisis comparativo de los patrones de LCC individuales de secuencias relacionadas.

Por otra parte, también permite complementar, desde un punto de vista diferente, los datos derivados de homologías y alineamientos múltiples, posibilitando reagrupar las secuencias en nuevos subconjuntos con mayor grado de relación evolutiva interna.

Tabla 18.d. Homología de LCC de 13 RNAs S de bunyavirus

	BAT	GER	INK	JAM	JSU	KEY	LAC	MAG	MEL	MOR	TAH	TRI
AIN	64.6	63.5	67.1	66.4	67.3	67.6	69.1	65.0	65.6	62.5	66.8	66.5
BAT		72.8	66.6	64.2	65.2	67.1	66.8	78.0	67.9	67.3	69.2	67.6
GER			67.2	65.3	66.9	68.1	67.0	70.9	68.6	68.7	69.7	69.2
INK				78.6	75.3	74.2	73.3	66.7	73.8	72.5	71.4	70.0
JAM					75.9	73.5	71.6	65.2	72.7	69.6	69.3	67.4
JSU						71.5	70.8	65.3	72.0	69.8	68.8	68.8
KEY							72.8	69.5	73.8	72.1	74.4	72.6
LAC								66.6	70.6	71.3	74.2	71.6
MAG									68.0	68.0	70.1	67.4
MEL										73.6	72.8	69.1
MOR											73.3	69.8
TAH												71.2

AIN: Aino; BAT: Batai; GER: Germiston; INK: Inkoo; JAM: Jamestown Canyon; JSU: Jerry Slough; KEY: Keystone; LAC: La Crosse; MAG: Maguari; MEL: Melao; MOR: Morro Bay; TAH: Tahyna; TRI: Trivittatus.
 En grisado oscuro: grupo mayoritario (INK, JAM, JSU, KEY, LAC, MEL, MOR, TAH, TRI).
 En grisado claro: grupo minoritario (BAT, GER, MAG).

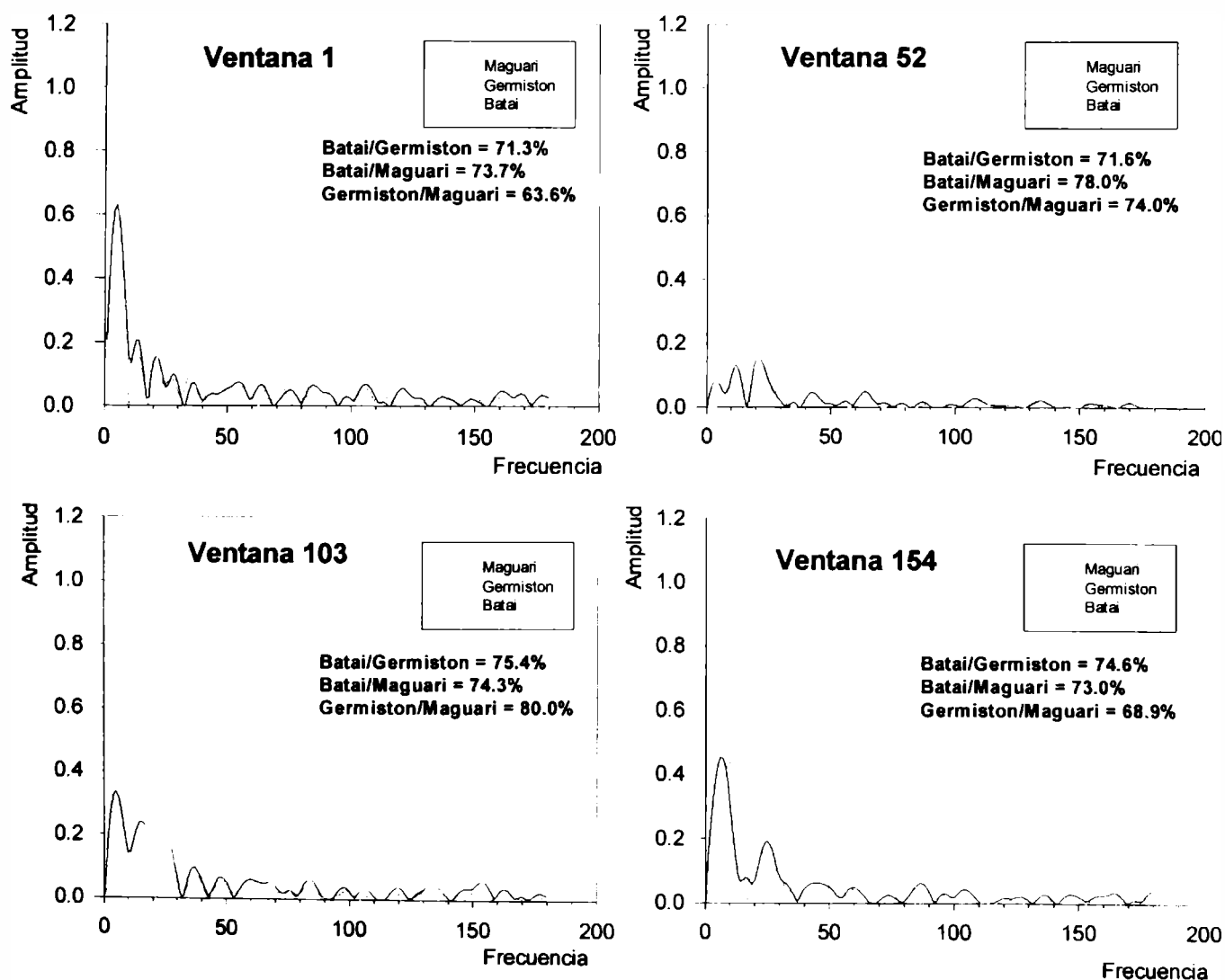


Figura 18.15. Diferentes ventanas de análisis de Fourier. Las cuatro ventanas corresponden al análisis comparativo de los patrones de LCC de los RNAs S de los bunyavirus Batai, Germiston y Maguari (grupo minoritario). Todas las ventanas corresponden a segmentos de 51 nucleótidos de longitud. La posición inicial es utilizada para identificar la ventana. En cada ventana se muestran, además, las relaciones de similitud local de LCC calculadas para cada par de secuencias.

Frecuencia relativa de nucleótidos

(A+U vs C+G)

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

La frecuencia relativa de nucleótidos y su distribución a lo largo de una secuencia constituyen una peculiaridad de la misma. En el caso de moléculas de ácidos nucleicos de simple cadena, esta distribución está asociada, además, con la potencial formación de estructuras secundarias estables.

Por otra parte, cuando se analiza una familia de secuencias nucleotídicas, el hallazgo de una buena correlación en la distribución de estas frecuencias, puede estar asociado con la importancia evolutiva del mensaje particular.

El análisis de las frecuencias de A+U vs C+G puede ser realizado en forma individual (secuencia por secuencia) o grupal (conjuntos de secuencias relacionadas). En los programas o paquetes de programas de uso frecuente, el único que realiza un análisis de este tipo es el programa PBASE (PC/GENE, Intelligentics, 1990). Dicho programa analiza la distribución de frecuencias en secuencias individuales, en función de ventanas definidas por el usuario, y la representa gráficamente. Esto último dificulta el análisis comparativo de varias secuencias.

Objetivos

⇒ El objetivo principal de este paquete de rutinas es contar con herramientas que permitan un análisis grupal de la distribución de frecuencias A+U vs C+G sobre secuencias relacionadas.

Fundamentos

Cuando se considera la distribución de frecuencias A+U vs C+G, es de esperar que la comparación de los patrones de distribución de frecuencias en macromoléculas relacionadas evolutivamente muestren un cierto grado de similitud. Esta similitud podría estar asociada con el grado de relación evolutiva.

La comparación y búsqueda de correlación de los patrones de distribución de frecuencias A+U vs C+G puede realizarse de varias maneras:

- (1) por simple comparación visual de los gráficos individuales
- (2) por análisis comparativo directo de los datos de cada macromolécula
- (3) por integración y ajuste de los datos individuales en función de un alineamiento múltiple y generación de un gráfico consenso para el conjunto de moléculas.

Hipótesis

El patrón de distribución de frecuencias A+U vs C+G es una propiedad intrínseca de cada secuencia.

Un gráfico consenso generado a partir de un conjunto de secuencias relacionadas evolutivamente, debería presentar un alto grado de similitud con cada gráfico individual.

Toda región cuya distribución de frecuencias real sea mayor (menor) que la esperable para un conjunto de secuencias de organización aleatoria (longitud y composición equivalente) debería considerarse un rasgo peculiar del grupo de secuencias en análisis.

El patrón consenso generado es altamente dependiente del alineamiento múltiple.

Algoritmo

Se trabaja con ventanas de longitud definida por el usuario y se utilizan las **Ecuaciones 8 y 9** para calcular las frecuencias de A+U y C+G, respectivamente para cada conjunto de residuos coincidente con la longitud de la ventana, desde $1 + \text{Long}_{\text{ventana}/2}$, desplazando la ventana un residuo por vez, hasta $\text{Long}_{\text{secuencia}} - \text{Long}_{\text{ventana}/2}$. A continuación, se aplica la regla del máximo (**Ec. 10**) y se asigna el valor a la posición central de la ventana. Este cálculo se realiza para cada una de las secuencias del grupo, en forma individual. Para

el cálculo se utilizan valores positivos para A+U y negativos para C+G.

Ec. 8

$$N_{AU} = \sum_{k=1}^K \frac{A+U}{K}$$

Ec. 9

$$N_{CG} = -\sum_{k=1}^K \frac{C+G}{K}$$

Ec. 10

$$F = \begin{cases} 0 & \text{si } |N_{AU}| = |N_{CG}| \\ N_{AU} & \text{si } |N_{AU}| > |N_{CG}| \\ N_{CG} & \text{si } |N_{AU}| < |N_{CG}| \end{cases}$$

A continuación, se trabaja sobre el conjunto de archivos individuales con datos de distribución de frecuencias de las secuencias del grupo y un alineamiento múltiple de las mismas,

para ajustar los datos al alineamiento. En este punto es necesario mencionar que a los *gaps* se les asigna el valor 0.

Por último, utilizando el archivo de salida de la etapa anterior, se calcula la media aritmética y la desviación estándar para cada línea de datos, generando un archivo ASCII que puede ser importado con diversos programas de graficación para construir el gráfico consenso final.

Criterios

Selección de ventanas a utilizar

En principio, el usuario puede definir ventanas de la longitud que desee (> 1 y $< \text{Long}_{\text{secuencia}}$). La selección de la misma debe basarse en un equilibrio entre sensibilidad y sencillez. Utilizando una ventana de 35 o 51 residuos se obtienen resultados que conjugan ambos hechos, sensibilidad y simplicidad del gráfico (Figura 18.16)

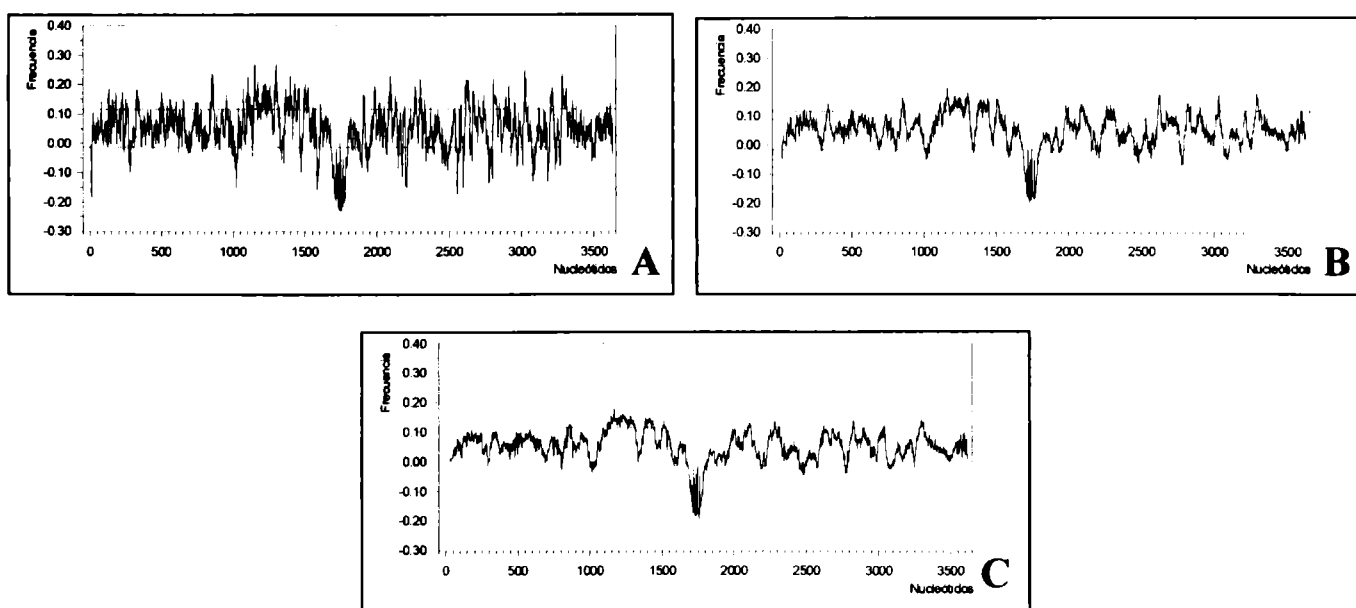


Figura 18.16. Gráficos de frecuencia nucleotídica relativa (A+U vs C+G). Los tres gráficos corresponden al análisis conjunto de los RNAs S de los arenavirus Junin-Candid #1, Junin-MC2, Oliveros, Pichindé, Sabiá, Tacaribe, Lassa-Nigeria, Lassa-Josiah, LCM-Armstrong, LCM-WE y Mopeia, en cada uno de ellos se utilizó una ventana de cálculo diferente. A. Ventana de 15 residuos. B. Ventana de 35 residuos. C. Ventana de 51 residuos. Como se puede observar, la complejidad del gráfico disminuye a medida que aumenta la amplitud de la ventana utilizada. Las líneas punteadas horizontales señalan el rango de frecuencias esperables para secuencias nucleotídicas de organización aleatoria.

Análisis estadístico y detección de regiones características

Con la finalidad de determinar regiones características que se destaquen sobre lo esperable por simple azar, generalmente, se utiliza un análisis de Monte Carlo.

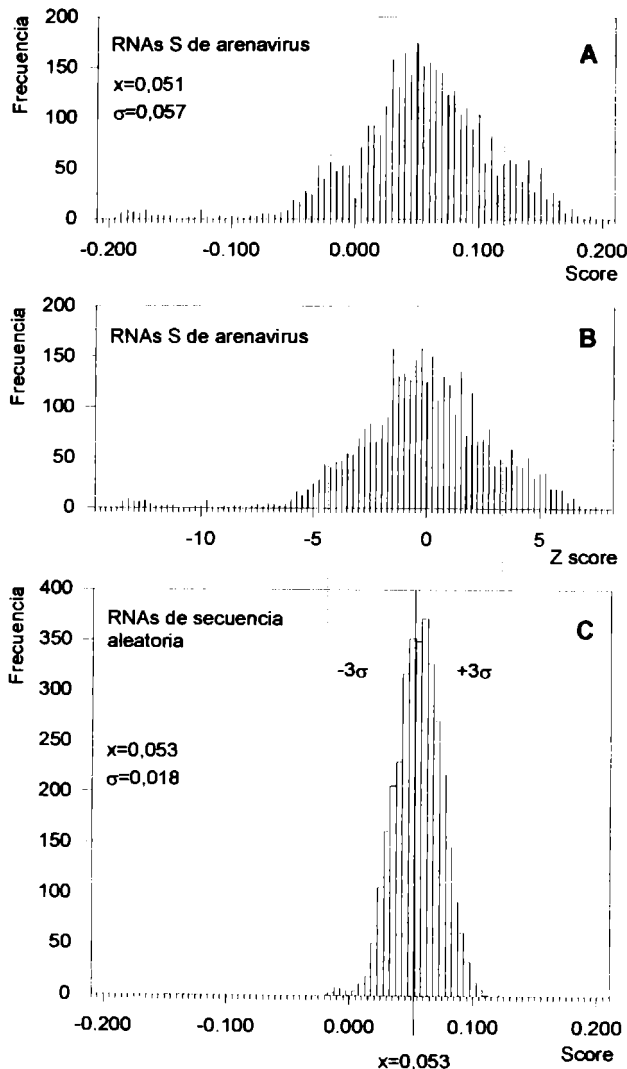


Figura 18.17. Análisis de Montecarlo de la distribución de frecuencias nucleotídicas relativas (A+U vs C+G). A. Histograma correspondiente a las secuencias de los RNAs S de los arenavirus Junín-Candid #1, Junín-MC2, Oliveros, Pichindé, Sabiá, Tacaribe, Lassa-Nigeria, Lassa-Josiah, LCM-Armstrong, LCM-WE y Mopeia. La distribución de frecuencias se representa asumiendo el valor de frecuencia nucleotídica relativa de cada ventana como *score*. B. Histograma correspondiente a las secuencias de los RNAs S de los mismos arenavirus. La distribución de frecuencias se representa en función del *Z score*. C. Histograma correspondiente a secuencias de RNAs de composición nucleotídica equivalente a la de los arenavirus, pero de organización aleatoria. La distribución de frecuencias se representa asumiendo el valor de frecuencia nucleotídica relativa de cada ventana como *score*.

Para ello, se construye un histograma con los *scores* individuales correspondientes a dos o más secuencias reales. A continuación, se generan una o más secuencias de organización aleatoria, manteniendo el tamaño y la composición, y se construye el histograma correspondiente con los *scores* individuales.

La simple comparación de estos histogramas permite visualizar los límites entre los conjuntos de *scores* significativos y los no significativos. En la práctica, y cuando se trabaja con valores positivos y negativos, los valores de *scores* de secuencias reales que cumplan cualquiera de las condiciones de la Ec. 11 son significativos.

Ec. 11

$$S_r > M_{S_a} + 3\sigma$$

ó

$$S_r < M_{S_a} - 3\sigma$$

Donde:

- S_r = *scores* de las secuencias reales
- M = media aritmética de los *scores* de las secuencias aleatorias (S_a)
- σ = desviación estándar

Una manera alternativa, que también establece relaciones entre *scores* reales y *scores* esperables es la utilización del *Z score*, calculado de acuerdo a la Ec. 12.

Ec. 12

$$Z_{score} = \frac{(S_r - M_{S_a})}{\sigma}$$

Donde:

- S_r = *scores* de las secuencias reales
- M = media aritmética de los *scores* de las secuencias aleatorias (S_a)
- σ = desviación estándar

Como se observa en la **Figura 18.17**, en la distribución de frecuencias A+U vs C+G correspondientes a los RNAs S de arenavirus, cualquier valor de *score* mayor de +0,12 o menor de -0,02, cuando se utiliza la frecuencia como *score*, es estadísticamente significativo. Cuando se utiliza el *Z score* los límites superior e inferior son +3,25 y -4,5, respectivamente.

Abundancia relativa de oligonucleótidos

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

En distintos organismos, con genomas basados en DNA o RNA, la organización de la información nucleotídica es sumamente variada. En una secuencia particular, esta heterogeneidad puede verse reflejada en la existencia de diferentes peculiaridades, tales como:

- (1) Una excesiva dispersión en la abundancia relativa de di-, tri- y tetranucleótidos, globalmente o con distribución diferencial

- en regiones codificantes y no codificantes.
- (2) La distribución de marcadores definidos: conjuntos de nucleótidos, motivos regulatorios, sitios de restricción, etc.
- (3) Presencia de oligonucleótidos o péptidos en cantidades excepcionales, muy raros o muy frecuentes.

En 1994, Karlin and Cardon describieron un conjunto de métodos, basados en los conceptos anteriores, que permiten determinar heterogeneidades de organización, nucleotídica o peptídica.

Ec. 13a. Análisis no simétrico de abundancia relativa de dinucleótidos

$$\rho_{XY} = \frac{f_{XY}}{f_X \times f_Y}$$

Ejemplo:

$$\rho_{GT} = \frac{f_{GT}}{f_G \times f_T}$$

Donde:

- f_{XY} = frecuencia del dinucleótido XY en la hebra en análisis
- f_X = frecuencia del nucleótido X
- f_Y = frecuencia del nucleótido Y

Ec. 13b. Análisis no simétrico de abundancia relativa de trinucleótidos

$$\gamma_{XYZ} = \frac{f_{XYZ} \times f_X \times f_Y \times f_Z}{f_{XY} \times f_{YZ} \times f_{XNZ}}$$

Ejemplo:

$$\gamma_{GTA} = \frac{f_{GTA} \times f_G \times f_T \times f_A}{f_{GT} \times f_{TA} \times f_{GNA}}$$

En las Ec. 13b y 13c:

- f_{XYZ} = frecuencia del trinucleótido XYZ
- f_{XYZW} = frecuencia del tetranucleótido XYZW

El resto de las frecuencias corresponden a las frecuencias de los componentes del tri- o tetranucleótido en análisis.

Ec. 13c. Análisis no simétrico de abundancia relativa de tetranucleótidos

$$\tau_{XYZW} = \frac{f_{XYZW} \times f_{XY} \times f_{XNZ} \times f_{XNNW} \times f_{YZ} \times f_{YNW} \times f_{ZW}}{f_{XYZ} \times f_{XYNW} \times f_{XNZW} \times f_{YZW} \times f_X \times f_Y \times f_Z \times f_W}$$

Ejemplo:

$$\tau_{GTAC} = \frac{f_{GTAC} \times f_{GT} \times f_{GNA} \times f_{GNNC} \times f_{TA} \times f_{TNC} \times f_{AC}}{f_{GTA} \times f_{GTNC} \times f_{GNAC} \times f_{TAC} \times f_G \times f_T \times f_A \times f_C}$$

En esencia, estos métodos se basan en el análisis de las frecuencias propias de los di-, tri- y tetranucleótidos y la relación existente entre ellas y las frecuencias de sus componentes mono-, di- y trinucleotídicos (abundancia relativa). Para ello, se pueden utilizar, alternativamente, dos tipos de análisis, de acuerdo al ácido nucleico en estudio (simple o doble cadena). En el caso de ácidos nucleicos de simple cadena se puede realizar un análisis de tipo no simétrico (Ec. 13a, 13b y 13c) y en el caso de ácidos nucleicos de doble cadena se

puede realizar un análisis de tipo simétrico (Ec. 14a, 14b y 14c).

Por otra parte, basándose en datos experimentales y en análisis estadísticos estrictos, los mismos autores destacan que la existencia de valores de abundancia relativa $\leq 0,78$ o $\geq 1,23$ tienen una probabilidad $\leq 0,001$, independientemente de la composición de bases del genoma. Por lo tanto, cualquier valor inferior a 0,78 se considera sub-representado y cualquier valor mayor a 1,23 se lo considera sobre-representado.

Ec. 14a. Análisis simétrico de abundancia relativa de dinucleótidos

$$\rho_{XY}^* = \frac{2(f_{XY} + f_{XY}^c)}{(f_X + f_X^c)(f_Y + f_Y^c)}$$

Ejemplo:

$$\rho_{GT}^* = \frac{2(f_{GT} + f_{GT}^c)}{(f_G + f_G^c)(f_T + f_T^c)}$$

Donde:

f_{XY} = frecuencia del dinucleótido XY en la hebra en análisis

f_{XY}^c = frecuencia del dinucleótido complementario al dinucleótido XY en la hebra complementaria a la analizada

$f_X f_X^c$ = frecuencia del nucleótido X y de su complementario, respectivamente

$f_Y f_Y^c$ = frecuencia del nucleótido Y y de su complementario, respectivamente

Ec. 14b. Análisis simétrico de abundancia relativa de trinucleótidos

$$\gamma_{XYZ}^* = \frac{(f_{XYZ} + f_{XYZ}^c)(f_X + f_X^c)(f_Y + f_Y^c)(f_Z + f_Z^c)}{2[(f_{XY} + f_{XY}^c)(f_{YZ} + f_{YZ}^c)(f_{XNZ} + f_{XNZ}^c)]}$$

Ejemplo:

$$\gamma_{GTA}^* = \frac{(f_{GTA} + f_{GTA}^c)(f_G + f_G^c)(f_T + f_T^c)(f_A + f_A^c)}{2[(f_{GT} + f_{GT}^c)(f_{TA} + f_{TA}^c)(f_{GNA} + f_{GNA}^c)]}$$

En las Ec. 14b y 14c, siempre:

f = frecuencia del oligonucleótido en la hebra en análisis

f^c = frecuencia del oligonucleótido complementario al de f en la hebra complementaria a la analizada

Ec. 14c. Análisis simétrico de abundancia relativa de tetranucleótidos

$$\tau_{XYZW}^* = \frac{2[(f_{XYZW} + f_{XYZW}^c)(f_{XY} + f_{XY}^c)(f_{XNZ} + f_{XNZ}^c)(f_{XNNW} + f_{XNNW}^c)(f_{YZ} + f_{YZ}^c)(f_{YNW} + f_{YNW}^c)(f_{YNW} + f_{YNW}^c)(f_{ZW} + f_{ZW}^c)]}{(f_{XYZ} + f_{XYZ}^c)(f_{XYNW} + f_{XYNW}^c)(f_{XNZW} + f_{XNZW}^c)(f_{XNZW} + f_{XNZW}^c)(f_{YZW} + f_{YZW}^c)(f_X + f_X^c)(f_Y + f_Y^c)(f_Z + f_Z^c)(f_W + f_W^c)}$$

Ejemplo:

$$\tau_{GTAC}^* = \frac{2[(f_{GTAC} + f_{GTAC}^c)(f_{GT} + f_{GT}^c)(f_{GNA} + f_{GNA}^c)(f_{GNNC} + f_{GNNC}^c)(f_{TA} + f_{TA}^c)(f_{TNC} + f_{TNC}^c)(f_{TNC} + f_{TNC}^c)(f_{AC} + f_{AC}^c)]}{(f_{GTA} + f_{GTA}^c)(f_{GTNC} + f_{GTNC}^c)(f_{GNAC} + f_{GNAC}^c)(f_{GNAC} + f_{GNAC}^c)(f_{TAC} + f_{TAC}^c)(f_G + f_G^c)(f_T + f_T^c)(f_A + f_A^c)(f_C + f_C^c)}$$

Objetivos

- ⇒ Contar con herramientas que permitan el análisis individual, sobre la secuencia completa y/o relacionada con la posición, de las abundancias relativas de los diferentes di-, tri- y tetranucleótidos.

Fundamentos

Cuando se considera la abundancia relativa de di-, tri- o tetranucleótidos de una secuencia determinada, es esperable que la mayor parte de los valores estén comprendidos entre 0,78 y 1,23 (probabilidad $\geq 0,999$; Karlin y Cardon, 1994). Por lo tanto, cualquier valor de abundancia relativa $\leq 0,78$ o $\geq 1,23$ debería ser considerado como una característica distintiva de la información nucleotídica analizada.

Si se extiende el estudio a un conjunto de secuencias relacionadas, el hallazgo de sub-representaciones y/o sobre-representaciones de oligonucleótidos cortos podría conducir a la búsqueda de su significado evolutivo (funcional o estructural).

Por otra parte, el cálculo de las abundancias relativas puede realizarse de varias maneras:

- (1) por análisis de las secuencias individuales.
Se trabaja con las secuencias completas.
- (2) por análisis de las secuencias individuales en un esquema asociado con la posición.
Se trabaja con ventanas adyacentes.

Hipótesis

- ⇒ El patrón de sub- y sobre-representaciones de oligonucleótidos cortos es una propiedad intrínseca de cada secuencia.
- ⇒ El análisis comparativo de un conjunto de secuencias relacionadas evolutivamente, debería permitir la detección de sub- y sobre-representaciones constantes. Esto último, debería considerarse un rasgo peculiar del grupo de secuencias en análisis.

Algoritmo

En el análisis global, se trabaja sobre las secuencias individuales completas utilizando las **Ecuaciones 13a, 13b y 13c** para el análisis no simétrico (o las **Ecuaciones 14a, 14b y 14c** para el análisis simétrico). A continuación, los resultados se tabulan y/o grafican para el análisis comparativo.

En el análisis asociado con la posición, se trabaja con ventanas de longitud definida por el usuario, y se utilizan las mismas **Ecuaciones** anteriores para calcular las abundancias relativas para cada conjunto de residuos coincidente con la longitud de la ventana. Como se utiliza un esquema de ventanas adyacentes, este análisis permite obtener un total de datos equivalente a $\text{Long}_{\text{secuencia}}/\text{Long}_{\text{ventana}}$.

Para minimizar los sesgos derivados del análisis de un número pequeño de datos, las ventanas deberían ser grandes y guardar alguna relación con la longitud total de la secuencia; por ejemplo: para las secuencias de los RNAs S de arenavirus (longitud promedio ~ 3500 nt.) las ventanas a utilizar deberían ser ≥ 200 nucleótidos, mientras que para las secuencias de los DNAs de baculovirus (longitud promedio ~ 130000 nt.) las ventanas a utilizar deberían ser ≥ 2000 nucleótidos.

En ambos tipos de análisis, el punto de partida es una colección de archivos de secuencia individuales sin formato.

Los datos de abundancia relativa de di-, tri- y tetranucleótidos de cada secuencia se guardan en un archivo ASCII que puede ser importado con diversos programas de graficación. Todos los valores generados son positivos. Sin embargo, para facilitar la discriminación rápida entre sub- y sobre-representaciones y la visualización gráfica, el archivo original se modifica, llevando a 0 todos los valores comprendidos entre 0,78 y 1,23, asignando valor positivo a la diferencia: valor $>1,23 - 1,23$ y valor negativo a la diferencia: $0,78 - \text{valor} < 0,78$.

En las figuras 18.18 y 18.19 se muestran algunos ejemplos de abundancia relativa de dinucleótidos.

La Figura 18.18 corresponde al análisis realizado sobre las secuencias completas, en forma global.

La Figura 18.19 corresponde al análisis realizado sobre el RNA S del virus Junín – Candid 1, utilizando un esquema de ventanas adyacentes.

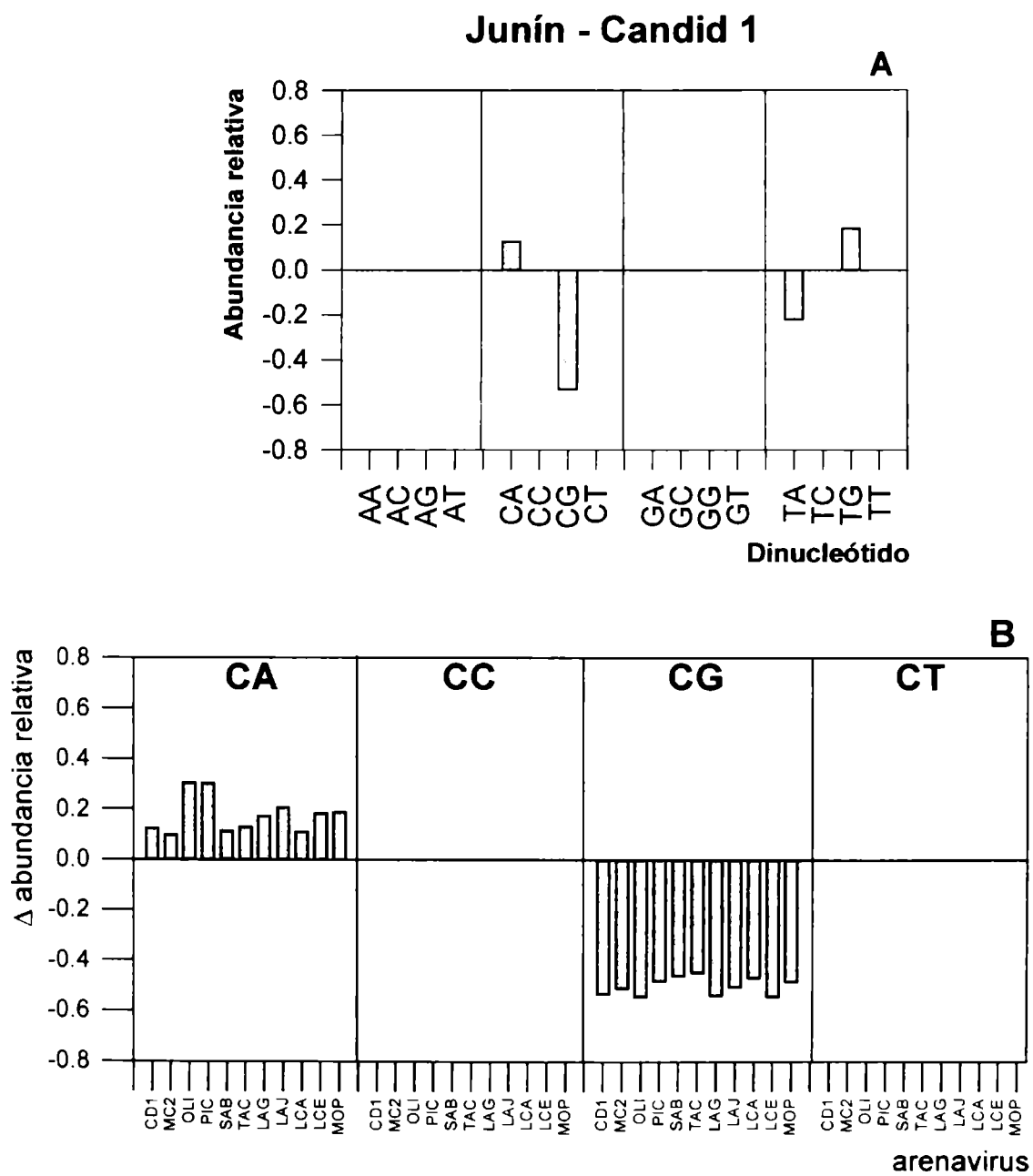


Figura 18.18. Análisis de abundancia relativa de los 16 dinucleótidos posibles. A. Histograma correspondiente a la secuencia del RNA S del virus Junín-Candid #1.

B. Histograma correspondiente a las secuencias de los RNAs S de los distintos arenavirus. CD1: Junín-Candid #1, MC2: Junín-MC2, OLI: Oliveros, PIC: Pichindé, SAB: Sabiá, TAC: Tacaribe, LAG: Lassa-Nigeria, LAJ: Lassa-Josiah, LCA: LCM-Armstrong, LCE: LCM-WE y MOP: Mopeia.

En ambos casos se representan sólo las sub- y sobre-representaciones de dinucleótidos.

Junín - Candid 1 Abundancia relativa de dinucleótidos

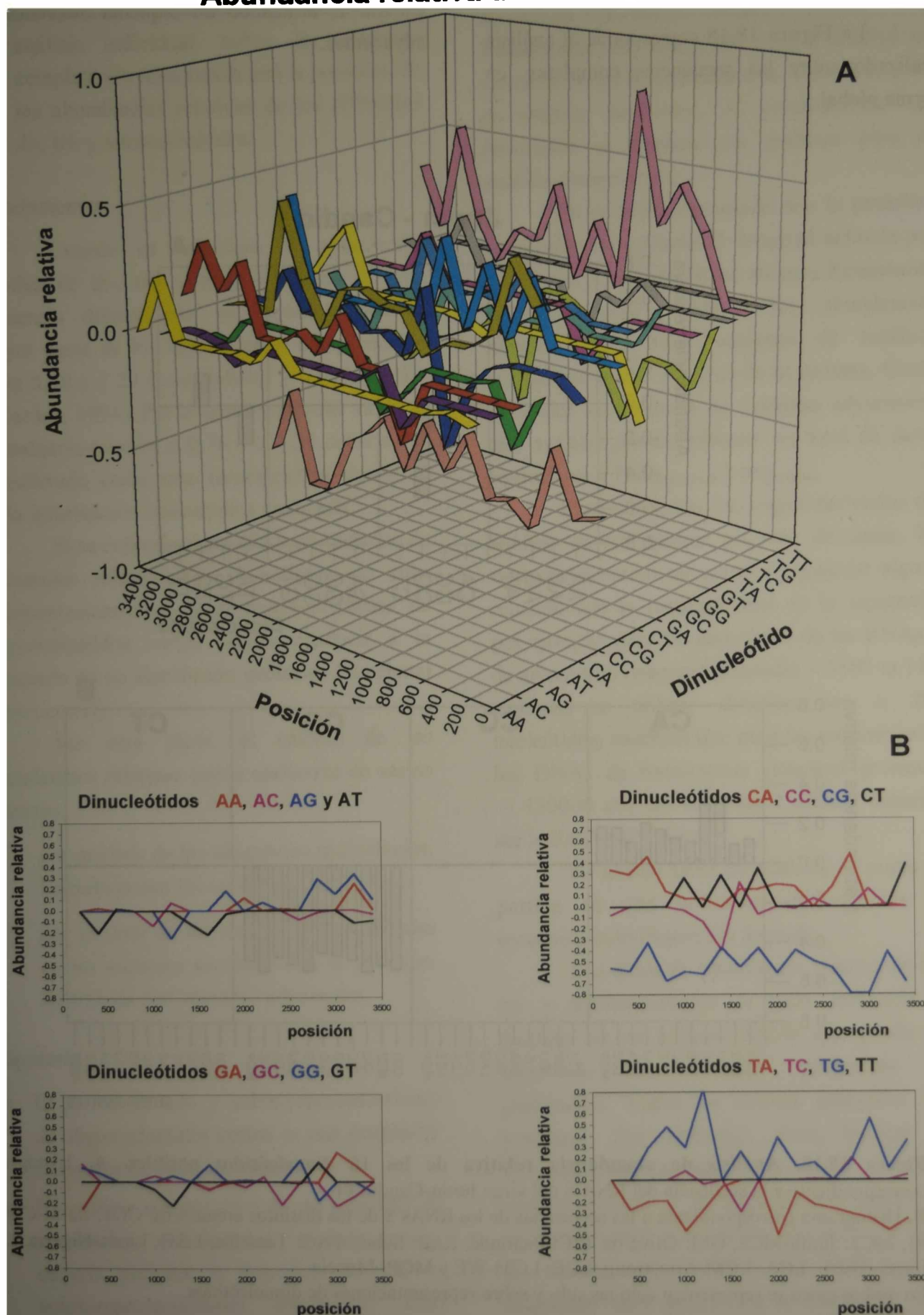


Figura 18.19. Abundancia relativa de los dieciséis dinucleótidos posibles en el RNA S del virus Junín - Candid 1. En los diferentes gráficos, se representan las sub- y sobre-representaciones correspondientes a ventanas adyacentes de 200 nucleótidos de longitud. **A.** Representación tridimensional de las abundancias relativas. **B.** Representaciones bidimensionales de las abundancias de cada uno de los cuatro subconjuntos de dinucleótidos existentes.

Datmanip (manipulación de datos)

(Rutina diseñada y escrita por P.D. Ghiringhelli)

En esencia esta rutina es un utilitario que permite tomar datos seleccionados, a partir de los archivos ASCII obtenidos con los programas PHDhtm, PHDsec y PHDacc (Rost *et al.*, 1994a,b, 1995; Rost and Sander, 1993a,b,c, 1994a,b,c,d), y adecuarlos para su presentación gráfica.

Las posibilidades que ofrece son:

- ⇒ Selección de columnas y obtención de datos.
- ⇒ Cambios de signo en los valores.
- ⇒ Búsqueda de máximos y mínimos.
- ⇒ Normalización de valores (entre 0 y 1).
- ⇒ Cálculo de promedios por ventanas.

Análisis comparativo de estructuras secundarias de RNA (óptimas y subóptimas) y determinación de la estructura más frecuente

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

La predicción de estructuras secundarias de RNA consiste en el cálculo de todos los posibles apareamientos de bases (G-C, A-U y G-U), sujetos a algunas reglas simples.

Actualmente, existe una diversidad de programas que permiten realizar estas predicciones. Entre ellos, se pueden mencionar el RNAFOLD (GCG, Wisconsin, USA), el RNADRAW (Ole Matzura, <http://broccoli.mfn.ki.se/rnadraw/rnadraw.html>) o el MULFOLD (Zuker, M., <http://ibc.wustl.edu/~zuker/mfold>) (Zuker and Stiegler, 1981; Zuker and Sankoff, 1984). Para predecir la estructura óptima, todos estos programas emplean una estrategia basada en minimizar la energía libre, utilizando las reglas desarrolladas por Turner *et al.* (1987, 1988). El MULFOLD permite, además, obtener una serie de estructuras subóptimas (Zuker, 1989; Zuker *et al.*, 1991).

Por otra parte, dados los requerimientos de espacio computacional (RAM y HD), todos estos programas tienen limitaciones en la longitud de los RNAs que pueden manejar para

realizar la predicción (por ej. el servidor con el programa MULFOLD puede manejar hasta 3000 nucleótidos).

Objetivos

Los objetivos principales de este paquete de rutinas son:

- ⇒ Contar con herramientas que permitan predecir la estructura secundaria completa de los RNAs S de arenavirus.
- ⇒ Contar con herramientas que permitan analizar comparativamente las estructuras óptimas y subóptimas.

Fundamentos

Cuando se analizan las predicciones de estructura secundaria de RNAs relacionados filogenéticamente, es esperable encontrar una conservación de algunos motivos estructurales. En el caso de los arenavirus, estos motivos podrían estar relacionados con funciones regulatorias: señales de encapsidación, controles de replicación/transcripción, controles de traducción, etc.

Hipótesis

El patrón de plegamiento de cada uno de los diferentes segmentos (hasta 3000 nucleótidos) en que se divide un RNA largo (mayor de 3000 nucleótidos), está dominado por las regiones de mínima energía. Por lo tanto, es factible construir una estructura secundaria consenso para la molécula completa de RNA a partir del análisis del grado de solapamiento de plegamientos parciales.

El análisis comparativo de un conjunto de estructuras secundarias, la óptima (de mínima energía libre) y las subóptimas (hasta un 5% de disminución en la energía libre), debería permitir la detección de la estructura más frecuente. Esto último, junto con información adicional (datos experimentales, similitud con otros RNAs de la misma familia, etc.) avalaría la decisión de considerar una estructura subóptima como la potencialmente existente.

Algoritmo

Se realiza la predicción de estructura secundaria sobre segmentos de RNA no mayores de 3000 nucleótidos de longitud; el solapamiento de los diferentes segmentos debe ser, al menos, de 500-1000 nucleótidos. Los datos se exportan en formato GCG (*Genetics Computer Group*, Wisconsin, USA), en los cuales la numeración es siempre correlativa y comenzando en 1.

Luego se hacen las comparaciones posición por posición, de a pares. Es importante que todos los archivos tengan la misma cantidad de líneas. Los archivos de los pares muestran con un 0 si la predicción de plegamiento para una posición dada es diferente en ambos archivos y con un 1 si la predicción de plegamiento para una posición dada es la misma en ambos archivos. En cualquiera de los casos no se discrimina si la predicción corresponde a interacción con otro nucleótido o no.

Por otra parte, también se calcula un *score* global de la comparación. Este *score* se construye mediante la sumatoria de los valores correspondientes a cada posición, utilizando un 2 para representar idéntica interacción con otro nucleótido, un 1 para representar idéntica ausencia de interacción y un 0 para representar las diferencias. El *score* puede tener un máximo teórico, que es el número de nucleótidos X 2, en el caso en que todos los nucleótidos estén apareados y las predicciones sean exactamente idénticas (muy poco probable); y un mínimo teórico, que sería la ausencia total de semejanza (muy poco probable). El valor del *score* es una medida de la semejanza entre dos plegamientos predichos.

Una vez construidos todos los pares deseados, se genera una tabla que tiene por primer columna la numeración de posición y las siguientes son tantas columnas como pares que existan (se transfieren los valores de 0 y 1. A continuación, se cuenta, línea por línea cuantos números 1 hay y se vuelca en un nuevo archivo que posee tres columnas: una con las posiciones,

otra con los totales de números 1 y la última con los porcentajes correspondientes a cada línea.

Predicción de la estructura (sin los extremos)

Se realiza el análisis comparativo de las estructuras calculadas para los diferentes segmentos en que se dividió el RNA (3 o más segmentos).

En el análisis de las regiones solapadas, se considera de mayor importancia la consistencia en la determinación de interacciones próximas (de mínimo ΔG_0 local) respecto de las interacciones entre los extremos libres de ambos segmentos en comparación.

Predicción de la estructura de extremos distales (5' y 3')

Se construyen moléculas quiméricas, formadas por 200-600 nucleótidos reales correspondientes a cada extremo, separados por una secuencia homopolimérica (≥ 200 Cs).

Se realizan dos tipos de análisis comparativo: se busca la estructura predominante mediante el estudio de un conjunto de predicciones que difieren, como máximo, un 5% en su ΔG_0 ; y se busca el solapamiento de interacciones entre la estructura predominante anterior y los segmentos más internos superpuestos.

En las Figuras 18.20, 18.21, 18.22 y 18.23 se pueden observar algunos ejemplos de estructuras calculadas de esta manera.

Predicción de la estructura completa

En base a los análisis anteriores, se rearma la estructura completa del RNA S, asignándole mayor importancia a la existencia de interacciones próximas (de mínimo ΔG_0 local), respecto de las interacciones entre los extremos libres de ambos segmentos en comparación, en el análisis de las regiones solapadas.

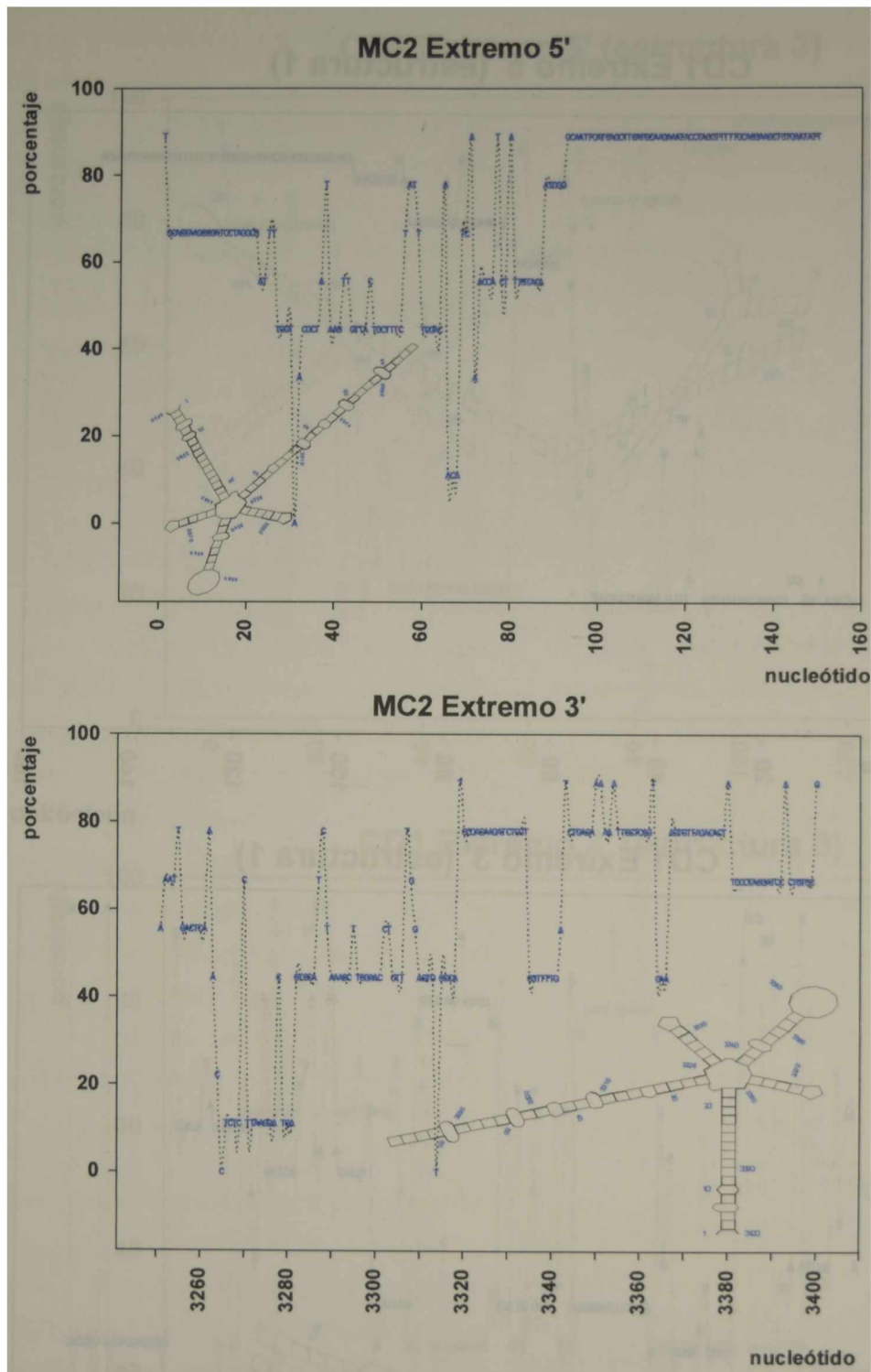


Figura 18.20. Análisis comparativo para la determinación de la estructura secundaria más probable en los extremos del RNA S viral del virus Junín-MC2. Los gráficos derivan del análisis comparativo de 9 estructuras secundarias predichas (incluyendo la óptima y las subóptimas), que difieren sólo un 5% en su ΔG_0 . La estructura de mayor probabilidad corresponde a la de máxima estabilidad. La estructura tipo *panhandle* predicha se ha superpuesto para facilitar la observación.

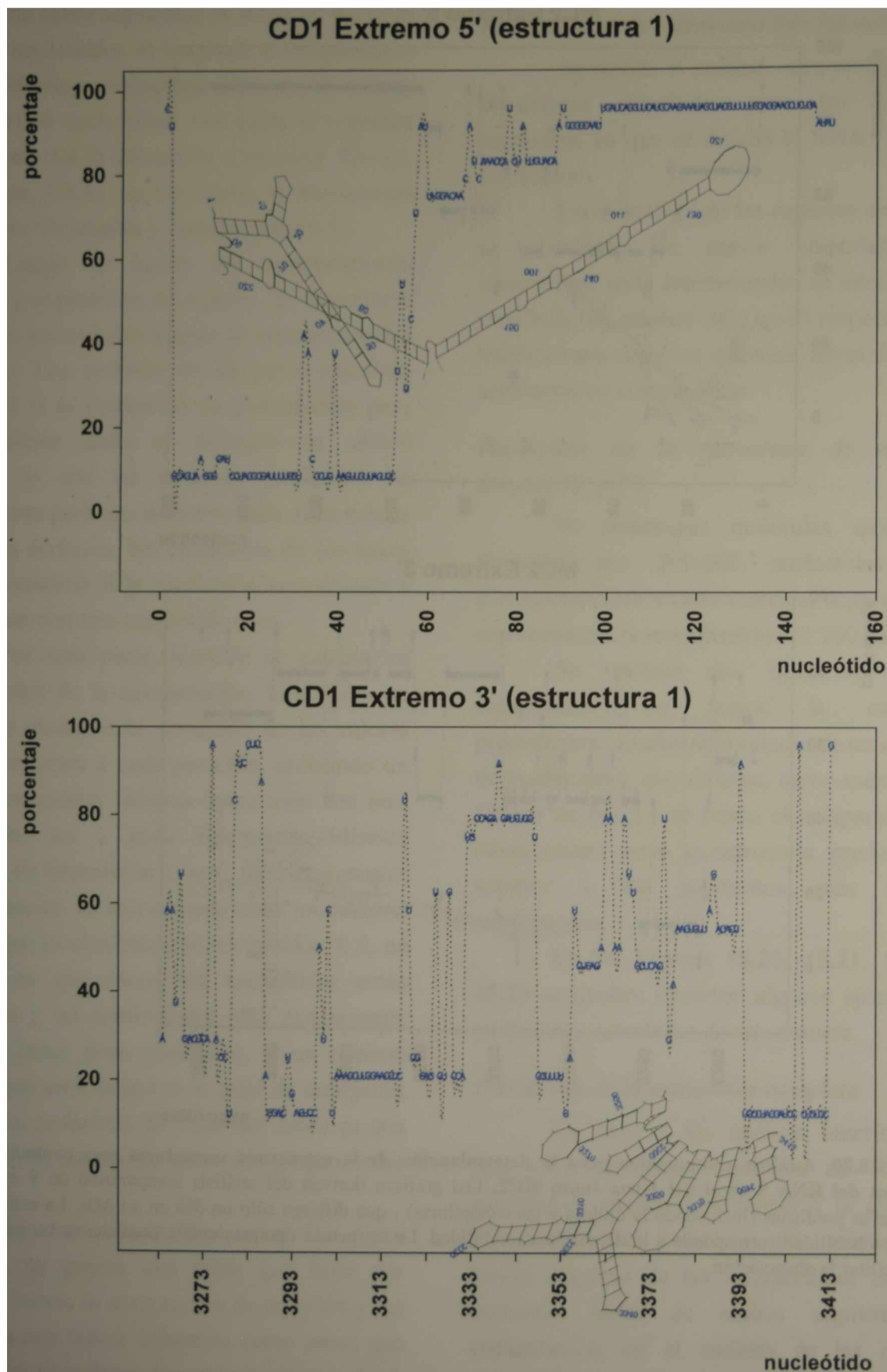


Figura 18.21. Análisis comparativo para la determinación de la estructura secundaria más probable en los extremos del RNA S viral del virus Junín-Candid #1. Los gráficos derivan del análisis comparativo de 24 estructuras secundarias predichas (incluyendo la óptima y las subóptimas), que difieren sólo un 5% en su ΔG_0 . Las estructuras superpuestas en cada gráfico corresponden a las de máxima estabilidad y baja probabilidad. En este caso no se constituye una estructura tipo *panhandle*.

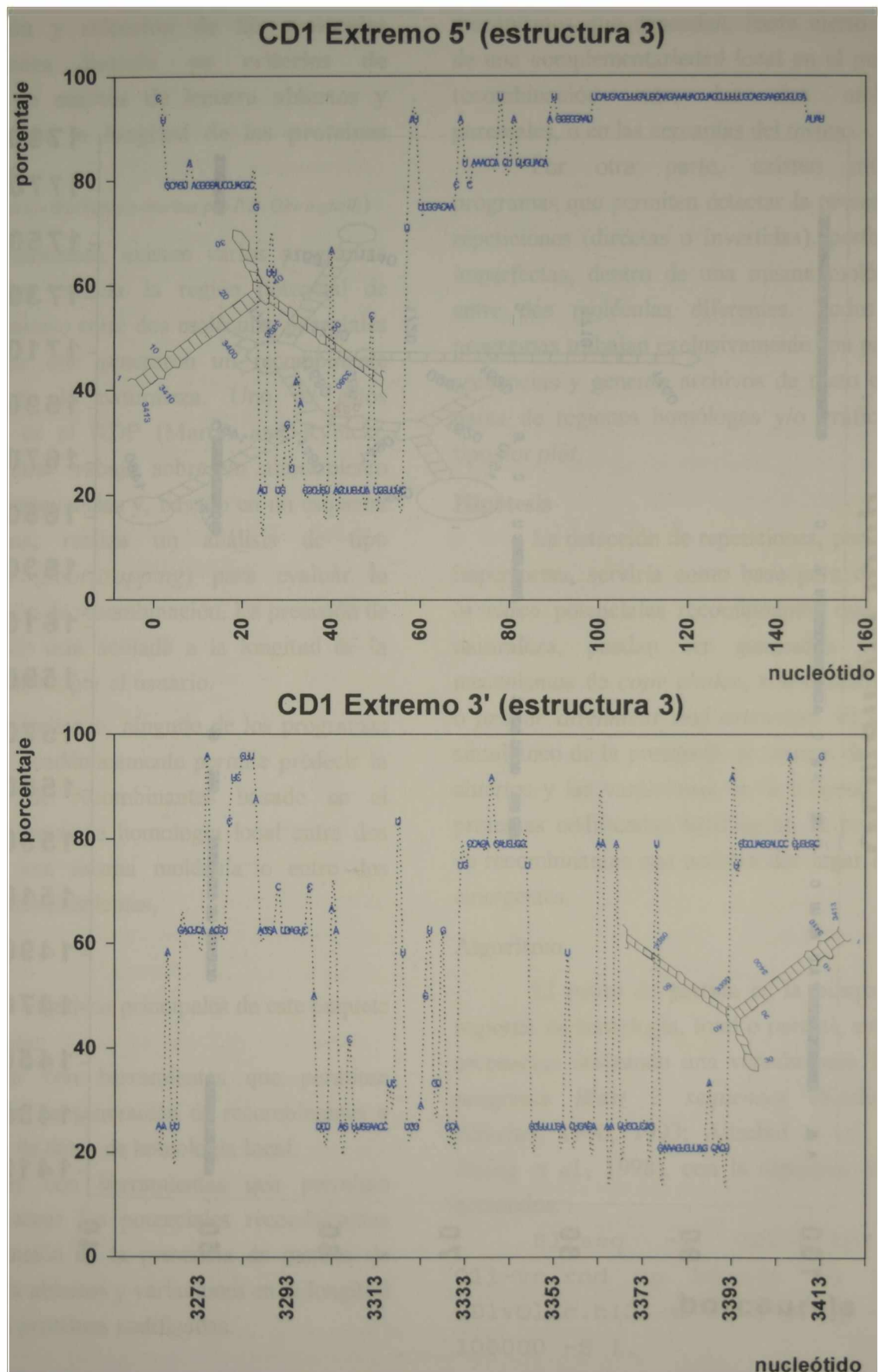


Figura 18.22. Análisis comparativo para la determinación de la estructura secundaria más probable en los extremos del RNA S viral del virus Junin-Candid #1. Los gráficos derivan del análisis comparativo de 24 estructuras secundarias predichas (incluyendo la óptima y las subóptimas), que difieren sólo un 5% en su ΔG_0 . La estructura de mayor probabilidad corresponde a la segunda subóptima. La estructura tipo *panhandle* predicha se ha superpuesto para facilitar la observación.

Detección de puntos potenciales de recombinación en moléculas de RNA, construcción y selección de las moléculas recombinantes basado en criterios de presencia de marcos de lectura abiertos y variaciones en la longitud de las proteínas codificadas

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

Actualmente, existen varios programas que permiten evaluar la región potencial de entrecruzamiento entre dos moléculas parentales desconocidas que generaron un recombinante detectado en la naturaleza. Uno de estos programas es el RDP (Martin and Rybicki, 2000), el cual trabaja sobre un alineamiento múltiple de secuencias y, basado en un esquema de ventanas, realiza un análisis de tipo filogenético (*bootstrapping*) para evaluar la posible región de recombinación. La precisión de la predicción está acotada a la longitud de la ventana definida por el usuario.

Sin embargo, ninguno de los programas disponibles académicamente permite predecir la generación de recombinantes basado en el análisis de datos de homología local entre dos copias de una misma molécula o entre dos moléculas independientes.

Objetivos

Los objetivos principales de este paquete de rutinas son:

- ⇒ Contar con herramientas que permitan predecir la generación de recombinantes a partir de datos de homología local.
- ⇒ Contar con herramientas que permitan seleccionar los potenciales recombinantes en función de la presencia de marcos de lectura abiertos y variaciones en la longitud de las proteínas codificadas.

Fundamentos

Como se ha mencionado en el **Capítulo 12** (páginas 146 y posteriores), existen varios mecanismos de recombinación posibles en los

genomas basados en moléculas de RNA. Entre ellos, los únicos predecibles son aquellos mecanismos que dependen, hasta cierto punto, de una complementariedad local en el punto de recombinación entre las dos moléculas parentales, o en las cercanías del mismo.

Por otra parte, existen múltiples programas que permiten detectar la presencia de repeticiones (directas o invertidas), perfectas o imperfectas, dentro de una misma molécula o entre dos moléculas diferentes. Todos estos programas trabajan exclusivamente con pares de secuencias y generan archivos de texto con los pares de regiones homólogas y/o gráficos del tipo *dot plot*.

Hipótesis

La detección de repeticiones, perfectas o imperfectas, serviría como base para construir *in silico* potenciales recombinantes que, en la naturaleza, puedan ser generados por los mecanismos de *copy choice*, transesterificación o *primer alignment and extension*. El análisis simultáneo de la presencia de marcos de lectura abiertos y las variaciones en la longitud de las proteínas codificadas habilitarían la predicción de recombinantes que podrían dar lugar a virus emergentes.

Algoritmo

El punto de partida es la búsqueda de regiones de homología, total o parcial, entre dos secuencias utilizando una versión para PC del programa *Blast 2 sequences* (Karlin and Altschul, 1990, 1993; Altschul *et al.*, 1997; Zhang *et al.*, 1998), con la siguiente línea de comandos:

```
B12seq -i CD1-v.rnd -j
Oli-vc.rnd -p blastn -g T -o
CD1vOlic.bl2 -W 7 -F F -q -l -e
100000 -S 1.
```

Donde: *-i* y *-j* son los nombres de los archivos de entrada; *-p* selecciona la subrutina a utilizar (*blastn* o *blastp*); *-g* permite operar con o sin introducción de *gaps* (*T: true; F: false*); *-o* es el

nombre del archivo de salida; *-W* indica la longitud mínima de palabra que se usa en la búsqueda; *-F* filtra las secuencias en análisis respecto de regiones de baja complejidad informativa (*T: true; F: false*); *-e* valor estadístico relacionado con las chances de similitud por azar (*expected value*); *-S* permite seleccionar si la búsqueda de regiones de homología en la segunda secuencia se hace utilizando la hebra ingresada (1), la complementaria (2), o las dos polaridades (3).

De esta manera, se detectan todas las regiones de homología existentes, con una longitud mínima de 7 residuos y 100 % de homología hasta la longitud máxima posible, acotada por la condición de que la homología debe ser igual o mayor a 60%.

El archivo de salida obtenido con el programa Bl2seq (Figura 18.24) se puede utilizar como archivo de entrada para diversas rutinas. Algunas de ellas tienen por finalidad analizar el archivo para seleccionar los recombinantes potencialmente viables y construir las moléculas de RNA S recombinantes, mientras que otras tienen por finalidad la visualización gráfica de todos los posibles puntos de recombinación.

En el caso de la selección y construcción de recombinantes, los datos deben ser procesados secuencialmente con las rutinas ANPARCOD, ARMAREC4 y REVTAB3, en las Figuras 18.25, 18.26 y 18.27 se muestran ejemplos de los archivos de salida obtenidos. El archivo de salida obtenido con la rutina ANPARCOD sirve de base para representar el mecanismo de generación del recombinante. Los archivos generados con la rutina ARMAREC4 sirven de base para evaluar la continuidad de marcos de lectura y obtener los productos de

traducción recombinantes, en el caso en que los puntos de entrecruzamiento correspondan a regiones codificantes; mientras que permite obtener las moléculas híbridas para efectuar predicciones de estructura, etc., en el caso en que los puntos de entrecruzamiento correspondan a regiones no codificantes. El archivo generado con la rutina REVTAB3 presenta sintéticamente todos los recombinantes encontrados, potencialmente viables. La Figura 18.28 muestra el procesamiento posterior de los archivos obtenidos con las rutinas ANPARCOD y ARMAREC4 que verifican la factibilidad de generación de un recombinante potencialmente viable mediante el mecanismo de *copy choice*.

En el caso de la visualización gráfica, los datos deben ser procesados secuencialmente con las rutinas DIVIDE, ANA2-1, PROCESA, PROCESA1, PROCESA2, JUNTA1, JUNTA2, JUNTA y HISTO. La Figura 18.29 muestra dos ejemplos de gráficos que presentan todos los posibles puntos de recombinación entre dos moléculas de RNA viral.

Dado que los datos mostrados corresponden a dos moléculas de RNA S en polaridad viral, todos los posibles puntos de entrecruzamiento detectados pueden ser utilizados para la generación de recombinantes mediante los mecanismos de *copy choice* o transesterificación. Para la evaluación de generación de recombinantes mediante el mecanismo de *primer alignment and extension*, el análisis debería realizarse entre una molécula en polaridad viral y la otra en polaridad viral complementaria.

```

>
      Length = 3534

      Score = 25.4 bits (15), Expect = 0.27
      Identities = 22/29 (75%)
      Strand = Plus / Plus

Query: 9  aggggatcctaggcgattttggtaacgct 37
      || ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 6  agtggatcctaggcatataggataacgct 34

      Score = 25.4 bits (15), Expect = 0.27
      Identities = 17/19 (89%)
      Strand = Plus / Plus

Query: 3395 gcctaggatccactgtgcg 3413
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 3516 gcctaggatccccggtgcg 3534

.....
.....
.....
.....

      Score = 12.7 bits (7), Expect = 1806
      Identities = 7/7 (100%)
      Strand = Plus / Plus

Query: 665  aagactg 671
      ||| ||| ||| ||| |||
Sbjct: 2591 aagactg 2597

Lambda      K      H
      1.10      0.333      0.550

Gapped
Lambda      K      H
      1.10      0.333      0.550

Matrix: blastn matrix:1 -1
Gap Penalties: Existence: 5, Extension: 2
Number of Hits to DB: 907
Number of Sequences: 0
Number of extensions: 907
Number of successful extensions: 907
Number of sequences better than 100000.0: 1
length of query: 3413
length of database: 3534
effective HSP length: 27
effective length of query: 3386
effective length of database: 3507
effective search space: 11874702
effective search space used: 11874702
T: 0
A: 0
X1: 4 ( 6.3 bits)
X2: 12 (19.0 bits)
S1: 14 (23.8 bits)
S2: 4 ( 7.9 bits)

```

Figura 18.24. Ejemplo de archivo derivado de la rutina BL2seq (*Blast 2 sequences*), derivado de la comparación entre los RNAs S de los virus Junín-Candid #1 y Oliveros. Este archivo es utilizado como archivo de entrada para las diversas rutinas descritas en este capítulo.

```

Región codificante para GPC
  K I S F D D I A V L L P E Y D V I I Q H P
  AAGAUCAGCUUUGAUGACAUCGCAGUGUUGUACCAGAAUUGACGUUAUAAUUCAGCAUCCG
Query: 416 AUCAGCUUUGAUGACAUCGCAGUGUUGUACCAGAAUUGACGUUAUAAUUCAGCAUCCGG 476
          ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct: 411 CCUGGAGAUUUCACACUACUCACAACAUAUGUCUUGUCCUCGACCCACGCAUCGGGUUUU 351
          UGGAGAUUUCACACUACUCACAACAUAUGUCUUGUCCUCGACCCACGCAUCGGGUUUUUGAC
          G R F Y H H T N N V S C S S P A Y G L F Q

Región codificante para GPC
  K M K C F G N T A V A K C N L N H D S E F
  AAAAUGAAGUGUUUUGGCAAUACUGCUGUAGCAAAUGCAAUCUGAAUCAUGACUCUGAAUUC
Query: 932 AUGAAGUGUUUUGGCAAUACUGCUGUAGCAAAUGCAAUCUGAAUCAUGACUCUGAAUUCUG 993
          ||| | ||| ||| ||| ||| ||| |||
Sbjct: 815 ACGUGAAGUUCACACUCGACACGAGAACGUAGUACGUUAGACGAGAGACUGUCAUUAGUUUG 754
          CGUGAAGUUCACACUCGACACGAGAACGUAGUACGUUAGACGAGAGACUGUCAUUAGUUUGGU
          C K L H S S H E Q M M C D A R Q C Y D F W
    
```

Figura 18.25. Ejemplo de archivo derivado de la rutina ANPARCOD. Este archivo es utilizado como archivo de entrada para la rutina ARMAREC4.

```

.....
GUAUAGAAGCCCAAGAGAGGUUAAAGCCUGUAUGACACUGUUCAUUGUCUCACCUCUUGAACAGUCAUGCAUGC
AAUUGUCAAUUGCAGGAACAGAACCAAACUGAUUGUUAAGUUUUGAAGGAUCUUUAACAUCCCAUACCCUCACCAC
ACCAUUCUUUUUAGUUCUUGCUGUUGAAAUCUUUCUCUCUUGUGAUGAAUUGGAGACAAGUGUGUUCAGCAG
ACAGCUUUUCAAGAUAUCAAGGUUGGGUAUUUGACUGUGUAAAGGAGCCCAAGUGUGGUCAAACCUUGUACCA
CAUCAUUCACUGUUCAGCUCUUGCACUGACAUGCAAGCAAUUGUCAAGCAGGCAUUGAUCCAAACUGAUUGA
UCAAUUGCUUCUGAUCCUUAUUGUCCAGAUCCUUAACACACCGUUGUUUUGACCUUGUCUGGACUGAAACCAA
.....
    
```

Figura 18.26. Ejemplo de archivo de salida de la rutina ARMAREC4. La zona mostrada corresponde a la región de unión entre los RNAs S virales de los virus Junín-Candid #1 y Oliveros. En rojo se indica la secuencia correspondiente a Junín-Candid #1 y en negro la correspondiente a Oliveros.

C1OI11I.RE1;	2	marco de lectura >	de 400 aa;	89	;	1541	;	484
C1OI23I.RE1;	2	marco de lectura >	de 400 aa;	89	;	1472	;	461
C1OI32D.RE1;	2	marco de lectura >	de 400 aa;	116	;	1496	;	460
C1OI33D.RE2;	1	marco de lectura >	de 400 aa;	82	;	1705	;	541
C1OI43I.RE2;	3	marco de lectura >	de 400 aa;	981	;	2577	;	532
C1OI44I.RE2;	2	marco de lectura >	de 400 aa;	89	;	1862	;	591
C1OI45D.RE1;	2	marco de lectura >	de 400 aa;	116	;	1499	;	461
C1OI46I.RE2;	2	marco de lectura >	de 400 aa;	89	;	1673	;	528

Figura 18.27. Ejemplo de tabla generada con la rutina ARMAREC4 y corregida con la rutina REVTAB3. Los datos mostrados corresponden a potenciales recombinantes entre diferentes regiones de los RNAs S de los virus Junín-Candid #1 y Oliveros.

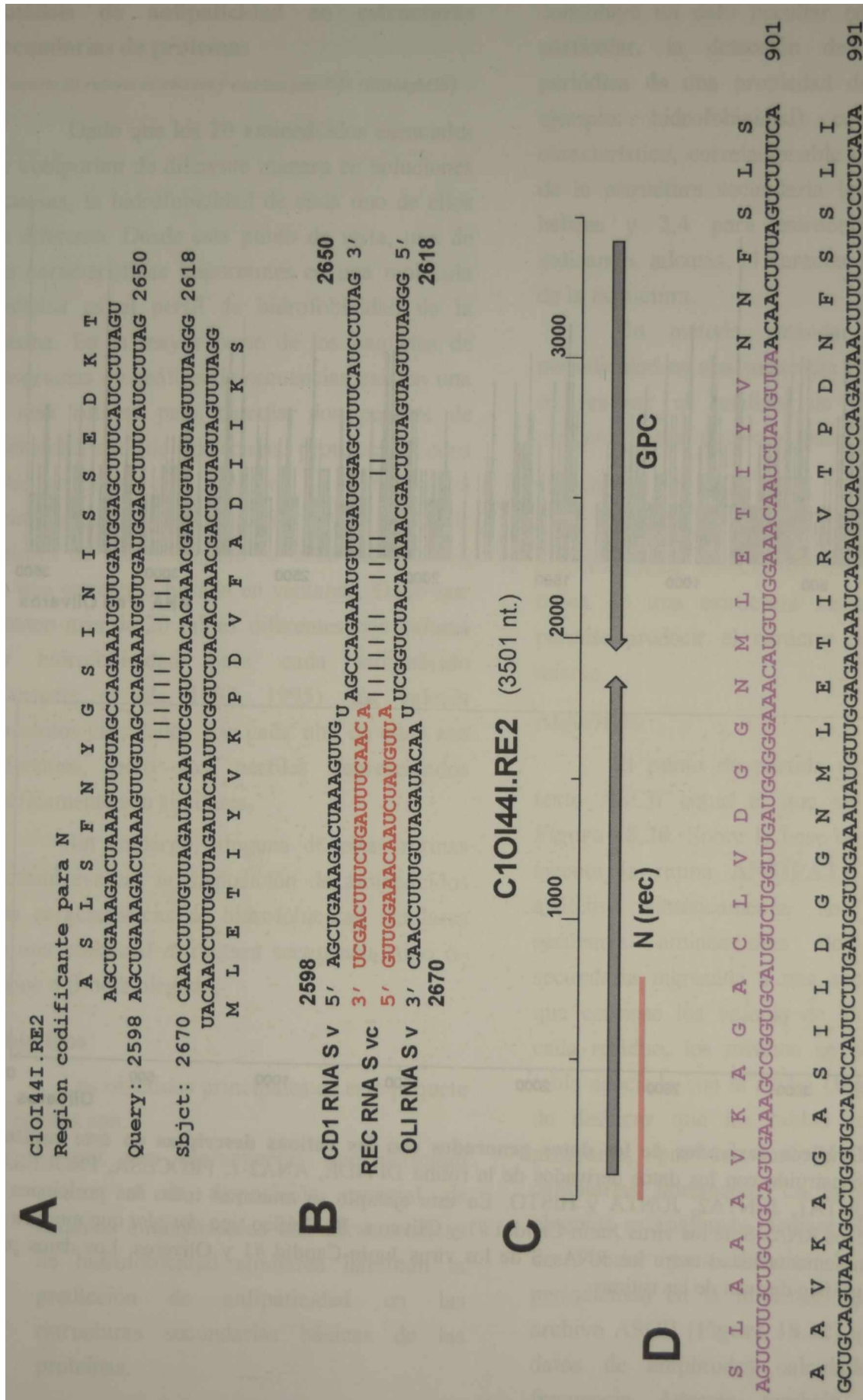


Figura 18.28. Ejemplo de procesamiento ulterior para mostrar la generación de un recombinante entre los RNAs S de Junin-Candid #1 y Oliveros, en la región codificante para la proteína de la nucleocápside (N), mediante el mecanismo de *copy choice*. A. Archivo de salida original de la rutina ANPARCOD. B. Modificaciones menores del panel A para mostrar el punto de salto de la RNA polimerasa. C. Mapa físico del recombinante. La línea rosa indica la porción de secuencia derivada del virus Oliveros. D. Segmento traducido de la secuencia nucleotídica del recombinante generado con la rutina ARMAREC4.BAS. En fucsia se indica la secuencia derivada del virus Oliveros.

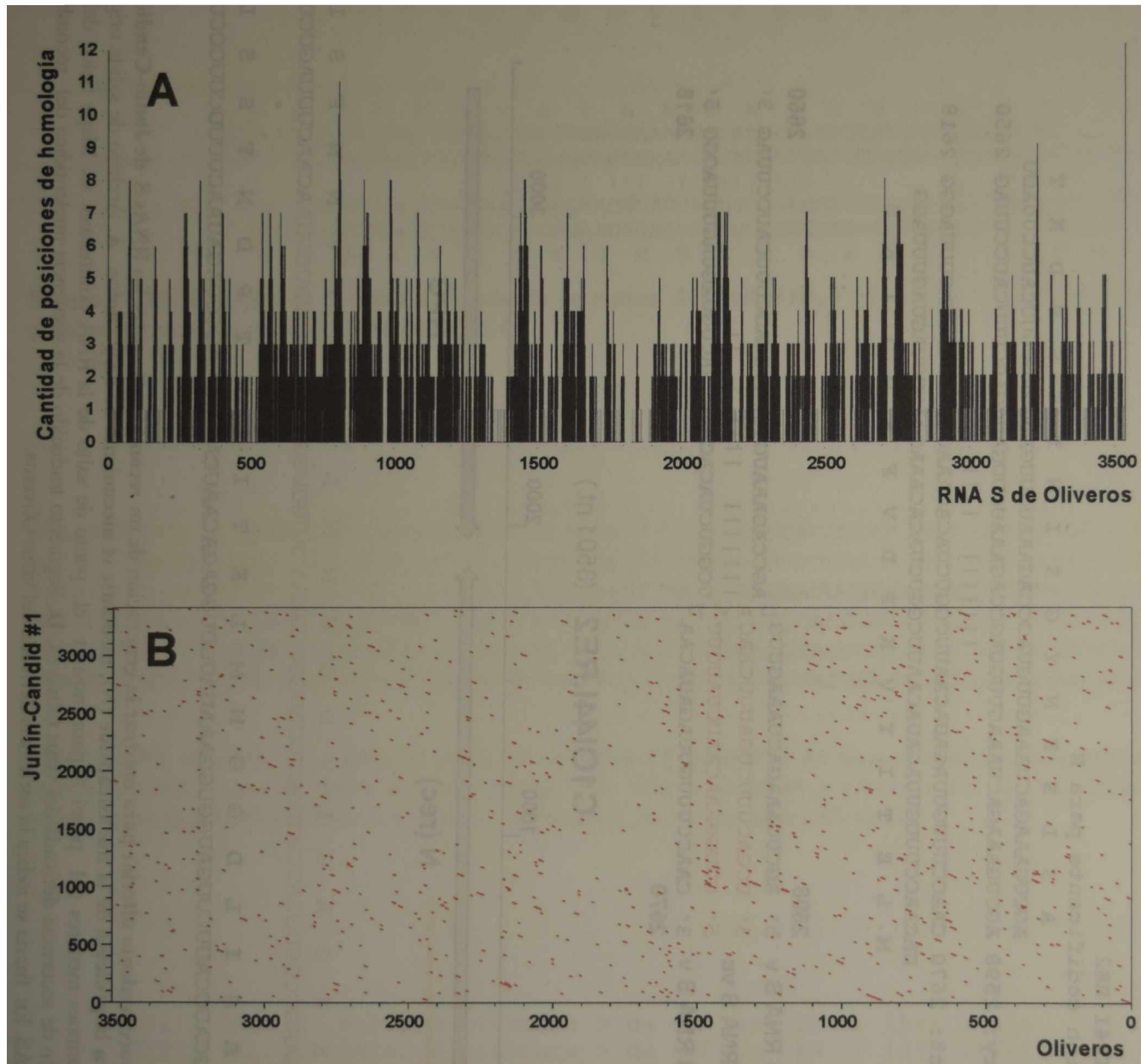


Figura 18.29. Gráficos derivados de los datos generados con las rutinas descritas en este capítulo. A. Histograma construido con los datos derivados de la rutina DIVIDE, ANA2-1, PROCESA, PROCESA1, PROCESA2, JUNTA1, JUNTA2, JUNTA y HISTO. En este ejemplo se muestran todas las posiciones de homología entre los RNAs S de los virus Junín-Candid #1 y Oliveros. **B.** Gráfico tipo *dot plot* que muestra las regiones de complementariedad entre los RNAs S de los virus Junín-Candid #1 y Oliveros. Los datos para confeccionar el gráfico derivan de las rutinas.

Análisis de anfipaticidad en estructuras secundarias de proteínas

(Paquete de rutinas diseñadas y escritas por P.D. Ghiringhelli)

Dado que los 20 aminoácidos esenciales se comportan de diferente manera en soluciones acuosas, la hidrofobicidad de cada uno de ellos es diferente. Desde este punto de vista, una de las características importantes en una molécula proteica es el perfil de hidrofobicidad de la misma. En la mayor parte de los paquetes de programas de análisis de secuencias existen una o más rutinas para calcular los perfiles de hidrofobicidad de moléculas proteicas. Todos ellos se basan en el análisis de la estructura primaria de la proteína en función de tablas de valores de hidrofobicidad de cada aminoácido, en una estrategia basada en ventanas. Dado que existen más de 20 tablas diferentes con valores de hidrofobicidad para cada aminoácido (Cornette *et al.*, 1987, 1995), los valores absolutos calculados con cada una de ellas son diferentes, pero los perfiles representados gráficamente son similares.

Sin embargo, ninguna de estas rutinas permite evaluar la disposición de aminoácidos con características de hidrofobicidad similares en una potencial estructura secundaria, tipo α -hélice o β -hoja plegada.

Objetivos

Los objetivos principales de este paquete de rutinas son:

- ⇒ Contar con herramientas que permitan evaluar si la disposición espacial de residuos aminoacídicos con características de hidrofobicidad similares habilitan la predicción de anfipaticidad en las estructuras secundarias básicas de las proteínas.

Fundamentos

La búsqueda y hallazgo de periodicidad en una secuencia numérica que representa alguna propiedad de una secuencia aminoacídica

constituye un dato peculiar en sí mismo. En particular, la detección de una repetición periódica de una propiedad determinada (por ejemplo: hidrofobicidad) con un período característico, correlacionable con la geometría de la estructura secundaria local (3,6 para α hélices y 2,4 para estructuras β), estaría indicando, además, el carácter de anfipaticidad de la estructura.

Un método estándar para detectar periodicidad en una secuencia finita de números es realizar el análisis de dicha secuencia mediante la transformada discreta de Fourier.

Hipótesis

El hallazgo de periodicidad en la hidrofobicidad, coincidente con la periodicidad típica de una estructura secundaria proteica permite predecir el carácter anfipático de la misma.

Algoritmo

El punto de partida es un archivo de texto ASCII como el que se muestra en la **Figura 18.30**. Sobre la base de este archivo se ejecuta la rutina ANFIPATI que realiza el análisis. Sintéticamente, se transforma la secuencia aminoacídica de la estructura secundaria ingresada a una sucesión numérica que contiene los valores de hidrofobicidad de cada residuo, los mismos se obtienen de una tabla asociada con la rutina (**Figura 18.31**). Es de destacar que las tablas son externas al programa, por lo cual el usuario puede cambiarlas libremente. La sucesión numérica obtenida es analizada mediante la transformada de Fourier para detectar la existencia de periodicidad en la información, generando un archivo ASCII (**Figura 18.32**) que contiene los datos de amplitudes calculados para cada frecuencia. Además, también se calcula un índice de anfipaticidad, que resulta del cociente entre el área bajo la curva entre dos frecuencias determinadas (85° y 110° si la estructura analizada es una α -hélice, y 160° y 180° si la

estructura analizada es una β -hoja plegada) y el área total bajo la curva (Cornette *et al.*, 1987).

Por otra parte, este archivo se puede importar con muchos de los programas de graficación disponibles, para visualizar el perfil de amplitudes obtenido en el análisis anterior (Figura 18.33).

A partir del archivo anterior, se puede generar una versión más sintética utilizando la rutina **EXTRPEEK**. El archivo ASCII generado con esta rutina contiene las estructuras

numeradas correlativamente, el valor del índice de anfipaticidad y los datos de todos los picos destacables (Figura 18.34).

Por último, una rutina relacionada con las anteriores es **HIDROF**, la cual permite calcular el perfil de hidrofobicidad, basado en un análisis por ventanas y una tabla de hidrofobicidad. El archivo ASCII generado tiene las posiciones de los residuos y el valor de hidrofobicidad asignado (Figura 18.35).

```

a1 (22-35)      # CLTRWMLIEAELKC
a2 (50-101)    # EEFC DMLTLDFDNKQAIQRLKAEAQMSIQLINKAVNALINDQLIMKNHLRDI
a3 (145-165)   # DIEQQADNMITEMLQKEYMER
a4 (182-192)   # SFYLISIFLHL
a1 (22-35)      # CLTRWMLIEAELKC
a2 (50-101)    # EEFC DMLRRLDFDNKQAIERLKAEAQMSIQLINKAVNALINDQLIMKNHLRDI
a3 (145-165)   # DIEQQADNMITEMLQKEYMER
a4 (182-192)   # SFYLISIFLHL
a1 (22-35)      # CLTKWMILAAELKC
a2 (50-101)    # AEFCDMLRLIDYNKAALSFKEDVESALHLFKTTVNSLISDQLLMRNHLRDL
a3 (144-165)   # DQIEQEADNMITEMLRKDYIKR
a4 (182-192)   # AYLVSIFLHL
.....
    
```

Figura 18.30. Ejemplo de archivo de ingreso para la rutina ANFIPMED. El archivo puede contener cualquier cantidad de secuencias a analizar, la única limitación es que cada línea no puede tener más de 256 caracteres. Cada una de las líneas corresponde a la secuencia aminoacídica de una estructura secundaria definida, y está compuesta por dos secciones, el símbolo # separa: a la izquierda el nombre y los datos de posición de la estructura secundaria en la proteína completa, y a la derecha la secuencia aminoacídica de la misma.

Kyte & Doolittle

A 1.00	C 0.17	D-3.00	E-2.60	F 2.50	G 0.67	H-1.70	I 3.10
K-4.60	L 2.20	M 1.10	N-2.70	P-0.29	Q-2.90	R-7.50	S-1.10
T-0.75	V 2.30	W 1.50	Y 0.08				

PRIFT

A 0.22	C 4.07	D-3.08	E-1.81	F 4.44	G 0.00	H 0.46	I 4.77
K-3.04	L 5.66	M 4.23	N-0.46	P-2.23	Q-2.81	R 1.42	S-0.45
T-1.90	V 4.67	W 1.04	Y 3.23				

Figura 18.31. Ejemplos de tablas de hidrofobicidad utilizadas. Dos de las tablas de hidrofobicidad más utilizadas. En ambas, se representan con valores positivos los aminoácidos hidrofóbicos o menos polares y con valores negativos los aminoácidos menos polares o hidrofílicos.

```

a3 (145-165) DIEQQADNMITEMLQKEYMER          Length: 21 aa.
Amphipathic Index:  2.32779          Mean hydrophobicity:  0.2633
Power Spectrum Data

```

Frequency	Power	Norm. Power (1000=1000)	Norm. Power (Max.Power=1000)
0.00	0.0000	0.0000	0.0000
1.00	1.8874	5.7216	2.1696
2.00	7.4203	22.4943	8.5299
3.00	16.2219	49.1760	18.6478
4.00	27.6984	83.9666	31.8405
5.00	41.0869	124.5532	47.2311
6.00	55.5170	168.2975	63.8191
7.00	70.0799	212.4443	80.5598
8.00	83.9000	254.3394	96.4465
9.00	96.2017	291.6315	110.5879
.....			
86.00	764.0360	2316.1435	878.2912
87.00	807.1655	2446.8889	927.8704
88.00	840.7764	2548.7789	966.5075
89.00	862.3281	2614.1120	991.2821
90.00	869.9120	2637.1022	1000.0000
91.00	862.3991	2614.3272	991.3637
92.00	839.5372	2545.0222	965.0829
93.00	801.9836	2431.1800	921.9135
94.00	751.2697	2277.4430	863.6158
95.00	689.6974	2090.7891	792.8358
96.00	620.1813	1880.0539	712.9242
97.00	546.0360	1655.2856	627.6911
98.00	470.7369	1427.0195	541.1317
99.00	397.6676	1205.5129	457.1355
100.00	329.8742	1000.0000	379.2041
101.00	269.8526	818.0470	310.2068
102.00	219.3798	665.0408	252.1862
.....			
177.00	169.5485	513.9792	194.9030
178.00	158.2782	479.8138	181.9474
179.00	150.8041	457.1564	173.3556
180.00	148.1901	449.2321	170.3507

Figura 18.32. Ejemplo de archivo de salida de la rutina ANFIPATI. El archivo está en formato texto ASCII. Los primeros renglones presentan: los datos de ingreso, tales como el nombre de la estructura, la secuencia y la longitud en aminoácidos; el índice de anfipaticidad calculado y la hidrofobicidad media de la estructura. A continuación se encuentra la información derivada del análisis de Fourier realizado. La primer columna (*Frequency*) corresponde a las frecuencia de Fourier utilizadas. La segunda columna (*Power*) corresponde a las amplitudes obtenidas para cada frecuencia durante el cálculo. La tercera y cuarta columnas (*Norm. Power*) corresponden a dos normalizaciones diferentes de los datos de *Power*. En la tercera, se asigna el valor de 1000 a la amplitud correspondiente a la frecuencia 100 y se normaliza el resto en función de ello. En la cuarta, se asigna el valor de 1000 a la máxima amplitud obtenida y se normaliza el resto en función de ello.

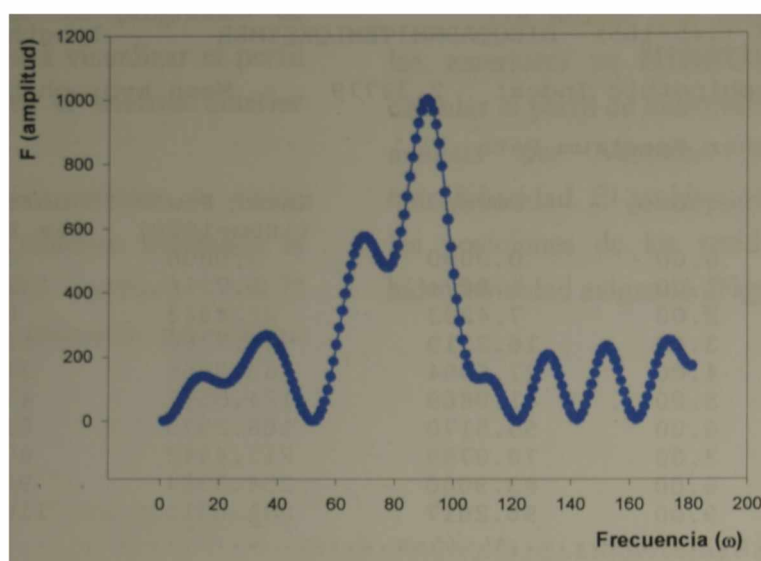


Figura 18.33. Ejemplo de perfil de Fourier derivado del análisis con la rutina ANFIPATI. Se representan las amplitudes normalizadas a un máximo de 1000 en función de las frecuencias utilizadas. Dado que el pico máximo está muy cerca de la frecuencia 100, y que la secuencia aminoacídica analizada corresponde a una predicción de α -hélice, se puede inferir que la estructura secundaria es anfipática.

Helix:	1	AI=	1.83171		
Peek	1	22.00	202.6712	5 ;	-0.35
Peek	2	36.00	260.0942	6 ;	-0.40
Peek	3	52.00	507.5008	7 ;	0.03
Peek	4	68.00	1000.0000	8 ;	-0.52
Peek	5	98.00	993.2395	9 ;	-0.30
Peek	6	123.00	405.2811	10 ;	-0.51
Peek	7	155.00	280.1031	11 ;	-0.28
Peek	8	172.00	330.8132	12 ;	-0.53
Peek	9	180.00	167.3970	13 ;	-0.42
				14 ;	-0.58
Helix:	2	AI=	1.70583	15 ;	-1.55
Peek	1	63.00	300.9302	16 ;	-1.25
Peek	2	104.00	733.1086	17 ;	-1.55
Peek	3	148.00	999.9999	18 ;	-1.03
				19 ;	-1.36
Helix:	3	AI=	2.39284	20 ;	-1.70
Peek	1	36.00	201.8276	21 ;	-1.70
Peek	2	110.00	1000.0000	22 ;	-0.62
Peek	3	180.00	204.1940	23 ;	-0.92

Figura 18.34. Ejemplo de archivo de salida de la rutina EXTRPEEK. Simplemente se indican el número de la estructura, el índice de anfipaticidad (AI) y todos los picos detectados tomando sólo el valor máximo de cada uno. Los picos se numeran en forma correlativa, y se indica la frecuencia de Fourier y el valor de amplitud correspondiente.

Figura 18.35. Ejemplo de archivo de salida de la rutina HIDROF. Simplemente se indican el número del residuo y el valor de hidrofobicidad calculado y asignado a la posición central de la ventana.

NOTA: Todas las rutinas descriptas y analizadas en este capítulo, no han sido incluidas en la Tesis debido al incremento excesivo en cantidad de páginas

PARTE VII

**CONCLUSIONES
GENERALES**

Conclusiones generales

En el trabajo de Tesis presentado se describieron diversos aspectos moleculares del RNA S de las cepas MC2 y Candid #1 del virus Junín y del resto de los arenavirus.

- ⇒ Se clonó el cDNA correspondiente al RNA S de las cepas MC2 y Candid #1 del virus Junín.
- ⇒ Se determinó la secuencia nucleotídica completa de ambas cepas. El RNA S de la cepa MC2 tiene 3400 nt. y el de la cepa Candid #1 tiene 3413 nt.
- ⇒ El marco de lectura identificado en polaridad viral en las posiciones 88 a 1533 (481 aminoácidos) en MC2 y 89 a 1546 (485 aminoácidos) en Candid #1 codifica para el precursor de las glicoproteínas virales.
- ⇒ El marco de lectura identificado en polaridad viral complementaria en las posiciones 82 a 1776 (564 aminoácidos) en MC2 y en Candid #1 codifica para la proteína de la nucleocápside.
- ⇒ La proteína de la nucleocápside de la cepa MC2 se expresó en el sistema baculovirus-células de insecto y se utilizó en ensayos de ELISA, permitiendo la detección de anticuerpos circulantes en sueros de ratón, cobayo y humanos. Además, con esta proteína recombinante se realizaron experimentos iniciales de interacción entre la proteína N y regiones seleccionadas del RNA viral (*panhandle* y *hairpin loops* intergénicos) y entre la proteína N y el catión Zn^{+2} .
- ⇒ Se realizó un estudio bioinformático exhaustivo de los ácidos nucleicos, en el cual se incluyeron algunos aspectos no estándar del análisis de secuencias, que llevaron al desarrollo de algoritmos particulares para poder efectuar los análisis correspondientes.
 - ◆ El análisis cuantitativo comparado de la complejidad informativa local (LCC: *local composition complexity*) permite reagrupar filogenéticamente los arenavirus, separando los patógenos netos de humanos (con la excepción de Mopeia) de los no patógenos. Esta correlación debería ser analizada nuevamente con un número mayor de secuencias de RNAs S arenavirales (no disponibles hoy día) para certificar estadísticamente su validez.
 - ◆ El análisis basado en las frecuencias de A+U vs C+G permitió detectar 3 regiones significativas, certificadas estadísticamente. Las dos primeras ricas en A+U, probablemente relacionadas con la frecuencia de uso de codones en la zona codificante para la glicoproteína G2, y la tercera rica en C+G, probablemente asociada con restricciones regulatorias que hacen necesaria la presencia de una o más estructuras tipo *hairpin loop* intergénicas que aseguren la terminación de la síntesis en la generación de los mRNAs durante el ciclo viral.
 - ◆ El análisis basado en las abundancias relativas de oligonucleótidos cortos (di-, tri- y tetranucleótidos), permitió determinar el perfil de abundancias relativas de los RNAs S de los arenavirus que constituyen la firma genómica (*genome signature*) de los mismos y la correlación entre la sub-abundancia del dinucleótido CG vs la sobreabundancia de los dinucleótidos UG y CA, sugiriendo la existencia de un mecanismo de deaminación preferencial que condujo a estas alteraciones durante la evolución.

- ◆ Se determinó la potencial estructura secundaria de los RNAs S completos de las cepas MC2 y Candid #1 del virus Junín y, en particular, de las estructuras tipo *hairpin loop* intergénicas y tipo *panhandle*, estabilizadas por interacciones entre los extremos 5' y 3', de todos los arenavirus analizados.
 - ◆ En función de las estructuras y de un análisis detallado de la presencia de repeticiones perfectas e imperfectas (directas o invertidas) de las secuencias de los RNAs S de todos los arenavirus analizados, se presentaron hipótesis y ejemplos que soportan la existencia de, al menos, tres posibles mecanismos de recombinación que permitirían generar nuevos genotipos y fenotipos arenavirales, mediante recombinación en las regiones intergénicas y/o en las regiones codificantes. En muchos de los casos es imposible distinguir si el recombinante final deriva de un mecanismo de *copy choice*, transesterificación o *primer alignment and extension*.
 - ◆ Se presentaron evidencias teóricas que avalan la hipótesis de que el RNA S del clon viral 6₂ de la cepa MC2 caracterizado en esta Tesis puede ser un recombinante con fragmentos de mRNAs celulares.
 - ◆ Se analizaron exhaustivamente las secuencias diferenciales de los extremos 5' de los RNAs S de las cepas MC2 y Candid #1 del virus Junín, comparativamente con los otros arenavirus y con otros sistemas virales, planteando hipótesis sobre los posibles mecanismos que permitirían la existencia de uno u otro tipo de extremos en los RNAs S virales y en los mRNAs.
 - ◆ Utilizando como secuencia de búsqueda la región intergénica del RNA S de los distintos arenavirus se evaluó la existencia de repeticiones imperfectas en los RNAs ribosomales humanos 18S y 28S. Este análisis permitió encontrar evidencias teóricas que sustentan nuevas hipótesis. La proteína N arenaviral podría interactuar con el rRNA 28S y/o complejos ribonucleoproteicos que lo contengan (ribosomas o subunidades ribosomales) posibilitando su incorporación en los viriones durante las etapas de encapsidación y brotación. Además, esta posible interacción también puede tener relación con el desnudamiento (*uncoating*) de los RNAs virales durante las etapas tempranas de la infección, un aspecto no estudiado hasta el presente en los arenavirus.
- ⇒ Respecto del estudio bioinformático de las proteínas GPC y N se realizó un análisis teórico detallado de las estructuras primarias y secundarias predichas, y de posibles modificaciones postraduccionales.
- ◆ Se identificaron diferentes sitios consenso de modificaciones postraduccionales conservados en todos los arenavirus. En particular, en el péptido señal se identificó una señal de miristoilación (₂GQIVTF₇) totalmente conservada que podría ser utilizada *in vivo*, permitiendo la interacción ribosoma-mRNA-proteína naciente con la SRP (*signal recognition particle*) celular para asegurar el correcto direccionamiento hacia el retículo endoplásmico. Por otra parte, se identificó una cantidad variable de sitios potenciales de N-glicosilación (4 en las glicoproteínas G1 de Junín y Tacaribe hasta 11 en la glicoproteína G1 de Pichindé, y 3 en la glicoproteína G2 de LCM hasta 5 en la glicoproteína G2 de Pichindé). De todos estos sitios, sólo uno está totalmente conservado

- en tipo y posición en las glicoproteínas G1, mientras que en las glicoproteínas G2 se conservan tres de los sitios hallados.
- ◆ Se identificó una región transmembrana totalmente conservada (aminoácidos 170 a 191), una señal de dimerización (residuos 73 a 81) y un patrón estructural común que explicarían la asociación de las glicoproteínas G2 en forma de tetrámeros que atraviesan la envoltura viral.
 - ◆ Se analizaron los posibles péptidos de fusión de las glicoproteínas G2 y se identificó al péptido de fusión II (residuos 17 a 21) como el de mayor probabilidad de cumplir esta función *in vivo*.
 - ◆ Se identificó un aminoácido diferencial en la región transmembrana de las glicoproteínas G2 de las cepas MC2 y Candid #1. La información analizada permitió sugerir una asociación entre el cambio aminoacídico y la atenuación de la virulencia.
 - ◆ Se analizó la presencia de sitios de fosforilación y se identificaron tres sitios conservados en tipo y posición en todas las proteínas N, uno de caseinquinasa II, uno de proteinquinasa C y uno de tirosinquinasa. La fosforilación/defosforilación de estos sitios podría estar relacionada con la transición transcripción/replicación en los arenavirus.
 - ◆ Se identificaron tres señales de dimerización que podrían asegurar la disposición compacta de las proteínas N en las nucleocápsides virales. alguna de estas señales también podría participar de la interacción con otra proteína viral (Z o L) o alguna proteína celular.
 - ◆ Se analizó el perfil de distribución de histidinas y cisteínas en las proteínas N. En base a lo anterior se predijo la existencia de un motivo tipo *zinc finger* en la porción C-terminal de las proteínas N. La interacción con el catión Zn^{+2} se confirmó en experimentos preliminares presentados en esta Tesis y en trabajos posteriores no presentados aquí.
 - ◆ Se estudió en las proteínas N la existencia de potenciales dominios de interacción con el RNA viral, y se detectaron tres dominios posibles. El dominio I ubicado en la región N-terminal, el dominio II ubicado en el primer tercio de la proteína y el dominio III ubicado en la región C-terminal, en el *loop* central del potencial *zinc finger*.
 - ◆ En base a las predicciones de estructura secundaria, la homogeneidad de patrones estructurales encontrada y utilizando metodologías de reconocimiento de plegamiento (*folding recognition*) se esquematizó en forma planar la posible disposición de los motivos estructurales que conforman la región tipo *zinc finger* de la proteína N de los virus Junín-MC2 y Junín-Candid #1.
- ⇒ En síntesis, en esta Tesis se resumieron datos experimentales y aspectos informativos que permitieron postular una serie de hipótesis explicativas de algunas de las peculiaridades de las cepas del virus Junín analizadas. En las mismas se intentó conjugar muchos datos experimentales dispersos a lo largo de los años con aspectos informativos de las macromoléculas y/o especulaciones sobre características no demostradas experimentalmente aún.

PARTE VIII

BIBLIOGRAFIA

1. Aaziz, R. and Tepfer, M. (1999). Recombination between genomic RNAs of two cucumoviruses under conditions of minimal selection pressure. *Virology*. **263**, 282-289.
2. Abelson, H.T., Smith, G.H., Hoffman, H.A. and Rowe, P.W. (1969). Use of enzyme-labelled antibody for electron microscope localization of LCM virus antigens in infected cell cultures. *Journal of National Cancer Institute*. **42**, 497-515.
3. Agrest, A., Avalos, J.C.S. and Slepoy, M.A. (1969). Fiebre hemorrágica argentina y coagulopatía por consumo. *Medicina (Buenos Aires)*. **29**, 194-201.
4. Ahmed, R. and Stevens, J.G. (1990). Viral Persistence. In: *Virology, second edition*. (Fields, B.N., Knipe, D.M. et al. Eds). Raven Press, New York. pp. 243-265
5. Ahmed, R., Hahn, C.H., Somasundaram, T., Villarete, L., Matloubian, M. and Strauss, J.H. (1991). Molecular basis of organ-specific selection of viral variants during chronic infection. *Journal of Virology*. **65**, 4242-4247.
6. Aissani, B., D'Onofrio, G., Mouchiroud, D., Gardiner, D., Gautier, C. and Bernardi, G. (1991). The compositional properties of human genes. *Journal of Molecular Evolution*. **32**, 493-503.
7. Akashi, H., Gay, M., Ihara, T. and Bishop, D.H.L. (1984). Localized conserved regions of the S RNA gene products of bunyaviruses are revealed by sequence analyses of the simbu serogroup Aino virus. *Virus Research*. **1**, 51-63
8. Albariño, C.G., Ghiringhelli, P.D., Grau, O. y Romanowski, V. (1992). Análisis de los genes estructurales de variantes atenuadas del virus Junín genéticamente relacionadas. XXVIII Reunión Nacional de la Sociedad Argentina de Investigación Bioquímica (Huerta Grande, Córdoba).
9. Albariño, C.G.; Ghiringhelli, P.D.; Posik, D.M.; Lozano, M.E.; Ambrosio, A.M.; Sanchez, A. and Romanowski, V. (1997). Molecular characterization of attenuated Junín virus strains. *Journal of General Virology*. **78** (7), 1605-1610.
10. Albariño, C. G. (1997). Tesis Doctoral: Caracterización molecular de variantes del virus Junín con diferentes grados de virulencia. Universidad Nacional de La Plata.
11. Albariño, C. G.; Posik, D.M.; Ghiringhelli, P.D.; Lozano, M.E. and Romanowski, V. (1998). Arenavirus Phylogeny: A New Insight. *Virus Genes*. **16**, 39-46.
12. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*. **215**, 403-410.
13. Altschul, S.F. (1991). Amino acid substitution matrices from an information theoretic perspective. *Journal of Molecular Biology*. **219**, 555-565.
14. Altschul, S.F. (1993). A protein alignment scoring system sensitive at all evolutionary distances. *Journal of Molecular Evolution*. **36**, 290-300.
15. Altschul, S.F.; Boguski, M. S.; Gish, W. and Wootton, J.C. (1994). Issues in searching molecular sequence databases. *Nature Genetics*. **6**, 119-129.
16. Altschul, S.F. and Gish, W. (1996). Local alignment statistics. *Methods in Enzymology*. **266**, 460-480.
17. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. **25**, 3389-3402.
18. Ambrosio, A., Enría, D. and Maiztegui, J.I. (1986). Junín virus isolation from lympho-mononuclear cells of patients with Argentine hemorrhagic fever. *Intervirology*. **25**, 97-102.
19. Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science*. **181**, 223-230.
20. Anzai, J., Takamatsu, F., Takeuchi, K., Kohno, T., Morimoto, K., Goto, H., Minamoto, N. and Kawai, A. (1997). Identification of a phosphatase-sensitive epitope of rabies virus nucleoprotein which is recognized by a monoclonal antibody. *Microbiology and Immunology*. **41**, 229-240.
21. Añón, M.C., Grau, O., Martínez Segovia, Z.M. and Franze-Fernández, M.T. (1976). RNA Composition of Junín virus. *Journal of Virology*. **18**, 833-838.
22. Arata, A.A. and Gratz (1975). The structure of rodents faunas associated with arenaviral infections. *Bulletin WHO*. **52**, 621-627
23. Aribalzaga, R.A. (1955). Una nueva enfermedad epidémica a germen desconocido: hipertermia nefrotóxica, leucopénica y enantemática. *Día Médico*. **27**, 1204-1210.
24. Armstrong, C. and Lillie, R.D. (1934). Experimental lymphocytic choriomeningitis of monkeys and mice produced by a virus encountered in studies of the 1933 St. Louis encephalitis epidemic. *Public Health Report*. **49**, 1019-1027.
25. Armstrong, C. and Sweet, L.K. (1939). Lymphocytic choriomeningitis. Report of two cases, with recovery of the virus from gray mice (*Mus musculus*) trapped in the two infected households. *Public Health Report*. **54**, 673-684.
26. Atkins, J.F. (1993). Contemporary RNA Genomes. In: *The RNA world*. (R.F. Gesteland and J.F. Atkins, eds.). Cold Spring Harbor Laboratory Press, New York. pp.: 535-556.
27. Auperin, D., Dimock, K., Cash, P., Rawls, W.E. and Bishop, D.H.L. (1982a). Analysis of the genome of prototype Pichinde arenavirus and a virulent derivative of Pichinde Munchique. Evidence for sequence conservation at the 3' termini of their viral RNA species. *Virology*. **116**, 363-367.
28. Auperin, D.D., Compans, R.W. and Bishop, D.H.L. (1982b). Nucleotide sequence conservation at the 3' termini of the virion RNA species of New World and Old World arenaviruses. *Virology*. **121**, 200-203.
29. Auperin, D.D., Galinski, M.S. and Bishop, D.H.L. (1984a). The sequences of the N protein gene and intergenic region of the S RNA of Pichinde arenavirus. *Virology*. **134**, 208-219.
30. Auperin, D.D., Romanowski, V., Galinski, M.S. and Bishop, D.H.L. (1984b). Sequencing studies of Pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA. *Journal of Virology*. **52**, 987-904.
31. Auperin, D.D., Sasso, D.R. and McCormick, J.B. (1986). Nucleotide sequence of the glycoprotein gene and intergenic region of the Lassa virus S genome RNA. *Virology*. **154**, 155-167.
32. Auperin, D.D. and McCormick, J.B. (1989). Nucleotide sequence of the Lassa virus (Josiah strain) S genome RNA and aminoacid sequence comparison of the N and GPC proteins to other arenaviruses. *Virology*. **168**, 421-425.
33. Babykin, M.M. and Zinchenko, V.V. (1984). Rapid separation of DNAs by buoyant density in three layer gradients. *Analytical Biochemistry*. **137**, 175-181.
34. Bairoch, A. (1993). The PROSITE dictionary of sites and patterns in proteins, its current status. *Nucleic Acids Research*. **21**, 3097-3103.
35. Bairoch, A. and Apweiler, R. (1997). The SWISS-PROT protein sequence data bank and its new supplement TrEMBL. *Nucleic Acids Research*. **25**, 31-36.

Referencias bibliográficas

36. Banerjee, S., An, S., Zhou, A., Silverman, R.H. and Makino, S. (2000). RNase L-independent specific 28S rRNA cleavage in murine coronavirus-infected cells. *Journal of Virology*. **74**, 8793-8802.
37. Balachandran, S., Roberts, P.C., Kipperman, T., Bhalla, K.N., Compans, R.W., Archer, D.R. and Barber, G.N. (2000). Alpha/Beta interferons potentiate virus-induced apoptosis through activation of the FADD/Caspase-8 death signaling pathway. *Journal of Virology*. **74**, 1513-1523.
38. Ball, L.A. (1997). Nodavirus RNA recombination. *Seminars in Virology*. **8**, 95-100.
39. Barber, G.N., Clegg, C.S. and Lloyd, G. (1990). Expression of the Lassa virus nucleocapsid protein in insect cells infected with a recombinant baculovirus: application to diagnostic assays for Lassa virus infection. *Journal of General Virology*. **71**, 19-28.
40. Barik, S. and Banerjee, A.K. (1991). Cloning and expression of the Vesicular Stomatitis Virus phosphoprotein gene in *Escherichia coli*: analysis of phosphorylation status versus transcriptional activity. *Journal of Virology*. **65**, 1719-1726.
41. Barik, S. and Banerjee, A.K. (1992). Sequential phosphorylation of the phosphoprotein of vesicular stomatitis virus by cellular and viral protein kinases is essential for transcription activation. *Journal of Virology*. **66**, 1109-1118.
42. Bar-Joseph, M., Yang, G., Gafny, R. and Mawasi, M. (1997). Subgenomic RNAs: the possible building blocks for modular recombination of *Closteroviridae* genomes. *Seminars in Virology*. **8**, 113-119.
43. Barrera Oro, J.G. and Eddy, G.A. (1982). Characteristics of candidate live attenuated Junin virus vaccine. Fourth International Conference on Comparative Virology, Banff, Alberta, Canada. 17-22 de octubre.
44. Barrera Oro, J.G., McKee Jr., K.T., Spisso, J., Mahlandt, B.G. and Maiztegui, J.I. (1990). A refined complement-enhanced neutralization test for detecting antibodies to Junin virus. *Journal of Virological Methods*. **29**, 71-80.
45. Barret, A.T.D. and Dimmock, N.J. (1986). Defective interfering viruses and infections of animals. *Current Topics in Microbiology and Immunology*. **128**, 55.
46. Barton, D.J., Morasco, B.J. and Flanagan, J.B. (1999). Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *Journal of Virology*. **73**, 10104-10112.
47. Barton, G.J. (1990). Protein multiple sequence alignment and flexible pattern matching. *Methods in Enzymology*. **183** 403-428.
48. Barton, G.J. (1993). An efficient algorithm to locate all locally optimal alignments between two sequences allowing for gaps. *CABIOS*. **9**, 729-734.
49. Barton, G.J. (1995). Protein secondary structure prediction. *Current Opininion in Structural Biology*. **5**, 372-376.
50. Bayley, J.M. and Davidson, N. (1976). Methylmercury as a reverside denaturing agent for agarose gel electrophoresis. *Analytical Biochemistry* **70**, 75-85.
51. Becher, P., Orlich, M. and Thiel, H-J. (1998). Ribosomal S27a coding sequences upstream of ubiquitin coding sequences in the genome of a pestivirus. *Journal of Virology*. **72**, 8697-8704.
52. Becher, P., Orlich, M., Köning, M. and Thiel, H-J. (1999). Nonhomologous RNA recombination in Bovine Viral Diarrhea Virus: molecular characterization of a variety of subgenomic RNAs isolated during an outbreak of fatal mucosal disease. *Journal of Virology*. **73**, 5646-5653.
53. Becher, P., Orlich, M. and Thiel, H-J. (2001). RNA recombination between persisting pestivirus and a vaccine strain: Generation of cytopathogenic virus and induction of lethal disease. *Journal of Virology*. **75**, 6256-6264.
54. Beck, B.N. and Ho, S.N. (1989). Increased specificity of PCR-amplified products by size fractionation of restriction enzyme-digested template genomic DNA. *Nucleic Acid Research*. **16**, 9051
55. Becker, Y. (2000). Evolution of viruses by acquisition of cellular RNA or DNA nucleotide sequences and genes: an introduction. *Virus Genes*. **21**, 7-12.
56. Belyavsky, A., Vinogradova, T. and Rajewsky, K. (1989). PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. *Nucleic Acids Research*. **17**, 2919-2932.
57. Benner, S.A. (1989). Patterns of divergence in homologous proteins as indicators of tertiary and quaternary structure. *Advances in Enzyme Regulation*. **28**, 219-236.
58. Benner, S.A. and Gerloff, D. (1990). Patterns of divergence in homologous proteins as indicators of secondary and tertiary structure of the catalytic domain of protein kinases. *Advances in Enzyme Regulation*. **31**, 121-181.
59. Benner, S.A., Badcoe, I., Cohen, M.A. and Gerloff, D.L. (1994). Bona Fide Prediction of Aspects of Protein Conformation. *Journal of Molecular Biology*. **235**, 926-958.
60. Benson, D.A., Boguski, M.S., Lipman, D.J. and Ostell, J. (1997). GenBank. *Nucleic Acids Research*. **25**, 1-6.
61. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977). The Protein Data Bank: a computer based archival file for macromolecular structures. *Journal of Molecular Biology*. **112**, 535-542.
62. Berria, M.I., Gutman Frugone, L.F., Girda, R. and Barrera Oro, J.G. (1967). Estudios inmunológicos con virus Junin. I. Formación de anticuerpos en cobayos inoculados con virus vivos. *Medicina (Buenos Aires)*. **27**, 93-98.
63. Bertling, W.M., Beier, F. and Reichenberger, E. (1993). Determination of 5' ends of specific messenger-RNAs by DNA ligase-dependent amplification. *PCR-METHODS AND APPLICATIONS*. **3**, 95-99.
64. Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*. **7**, 1513-1523.
65. Bishop, D.H.L., Gay, M.E. and Matsumoto, Y. (1983). Nonviral heterogeneous sequences are present at the 5' ends of one species os snowshoe hare bunyavirus complementary RNA. *Nucleic Acid Research*. **11**, 6409-6418.
66. Bishop, D.H.L. and Auperin, D.D. (1987). Arenavirus gene structure and organization. *Current Topics in Microbiology and Immunology*. **133**, 41.
67. Bishop, D.H.L. (1990). Arenaviruses and their replication. In: *Virology, second edition*. (Fields, B.N., Knipe, D.M. et al. Eds). Raven Press, Nueva York. pp: 1231-1243.
68. Blaschke, R., Andres, A.C., Reid, H.H., Zurcher, G., Friis, R.R. and Ziemiecki, A. (1991). Tyrosine kinases: from viral oncogenes to developmental regulators. *Behring Inst. Mitt.* **89**, 81-92.
69. Blinov, V., Ignatev, G., Sandkhchiev, L., Ghiringhelli, P.D. y Romanowski, V. (1994). Mecanismos de delección y duplicación de genes en las regiones intercistónicas de los RNAs de Arena y Bunyavirus. XXX Reunión Nacional de la Sociedad Argentina de Investigación Bioquímica (SAIB). Iguazú, Misiones, Argentina.

Referencias bibliográficas

70. Blumberg, B.M., Lepert, M. and Kolakofsky D. (1981). Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell*. **23**, 837-845.
71. Blumberg, B.M., Crowley, J.C., Silverman, J.L., Menoinna, J., Cook, S.D. and Dowling, P.C. (1988). Measles virus L protein evidences elements of ancestral RNA polymerase. *Virology*. **164**, 487-497.
72. Boersma, D.P., Saleh, F., Nakamura, K. and Compans, R.W. (1982). Structure and glycosylation of Tacaribe viral glycoproteins. *Virology*. **123**, 452-456.
73. Boersma, D.P. and Compans, R.W. (1985). Synthesis of Tacaribe virus polypeptides in an *in vitro* coupled transcription and translation system. *Virus Research*. **2**, 261-271.
74. Bolshoy, A., Shapiro, K., Trifonov, E.N. and Ioshikhes, I. (1997). Enhancement of the nucleosomal pattern in sequences of lower complexity. *Nucleic Acids Research*. **25**, 3248-3254.
75. Borden, K.L.B., Campbell Dwyer, E.J. and Salvato, M.S. (1998a). An arenavirus RING (zinc-binding) protein binds the oncoprotein promyelocyte leukemia protein (PML) and relocates PML nuclear bodies to the cytoplasm. *Journal of Virology*. **72**, 758-766.
76. Borden, K.L.B., Campbell Dwyer, E.J., Carlile, G.W., Djavani, M. and Salvato, M.S. (1998b). Two RING finger proteins, the oncoprotein PML and the arenavirus Z protein, colocalize with the nuclear fraction of the ribosomal P proteins. *Journal of Virology*. **72**, 3819-3826.
77. Bordo, D., Djinovic, K. and Bolognesi, M. (1994). Conserved Patterns in the Cu, Zn Superoxide Dismutase Family. *Journal of Molecular Biology*. **238**, 366-386.
78. Bork, P. and Gibson, T.J. (1996). Applying motif and profile searches. *Methods in Enzymology*. **266**, 162-184.
79. Borrow, P. and Oldstone, M.B.A. (1992). Characterization of Lymphocytic Choriomeningitis Virus-Binding Protein(s): a Candidate Cellular Receptor for the Virus. *Journal of Virology*. **66**, 7270-7281.
80. Borrow, P. and Oldstone, M.B.A. (1994). Mechanism of Lymphocytic Choriomeningitis Virus Entry into Cells. *Virology*. **198**, 1-9.
81. Bos, E.C.W., Luytjes, W., van Der Meulen, H., Koerten, H.K. and Spaan, W.J.M. (1996). The production of recombinant infectious DI-particles of a murine Coronavirus in the absence of helper virus. *Virology*. **218**, 52-60.
82. Bouloy, M., Plotch, S.J. and Krug, R.M. (1978). Globin mRNAs are primers for the transcription of influenza viral RNA *in vitro*. *Proceedings of National Academy of Sciences USA*. **75**, 4886-4890.
83. Bouloy, M., Morgan, M.A., Shatkin, A.J. and Krug, R.M. (1979). Cap and internal nucleotides of reovirus mRNAs primers are incorporated into influenza viral complementary RNA during transcription *in vitro*. *Journal of Virology*. **32**, 895-904.
84. Boutin, J.A. (1997). Myristoylation. *Cell Signal*. **9**, 15-35.
85. Bowen, M.D., Jackson, A.O., Bruns, T.D., Hacker, D.L. and Hardy, J.L. (1995). Determination and comparative analysis of the small RNA genomic sequences of California encephalitis, Jamestown Canyon, Jerry Slough, Melao, Keystone and Trivittatus viruses (Bunyaviridae, genus Bunyavirus, California serogroup). *Journal of General Virology*. **76**, 559-572.
86. Bowen, M.D., Peters, C.J., Mills, J.N. and Nichol, S.T. (1996). Oliveros virus: a novel arenavirus from Argentina. *Virology*. **217**, 362-366.
87. Bowen, M.D., Peters, C.J. and Nichol, S.T. (1997). Phylogenetic analysis of the *Arenaviridae*: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Molecular Phylogenetics and Evolution*. **8**, 301-316.
88. Bowie, J.U., Clarke, N.D., Pabo, C.O. and Sauer, R.T. (1990a). Identification of protein folds: matching hydrophobicity patterns of sequence sets with solvent accessibility patterns of known structures. *Proteins*. **7**, 257-264.
89. Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A. and Sauer, R.T. (1990b). Deciphering the message in protein sequences: tolerance to amino acid substitutions. *Science*. **247**, 1306-1310.
90. Bowie, J.U., Lüthy, R. and Eisenberg, D. (1991). A Method to Identify Protein Sequences That Fold into a Known Three-Dimensional Structure. *Science*. **253**, 164-169.
91. Bowie, J.U., Zhang, K., Wilmanns, M. and Eisenberg, D. (1996). Three-dimensional profiles for measuring compatibility of amino acid sequence with three-dimensional structure. *Methods in Enzymology*. **266**, 598-616.
92. Boxaca, M., Guerrero L.B. de, Parodi, A.S., Rugiero, H.R. and Gonzalez Capa S. (1965). Viremia en enfermos de fiebre hemorrágica Argentina. *Revista de la Asociación Médica Argentina*. **79**, 230-238.
93. Brandt, M., Yao, K., Liu, M., Heckert, R.A. and Vakharia, V.N. (2001). Molecular determinants of virulence, cell tropism, and pathogenic phenotype of infectious bursal disease virus. *Journal of Virology*. **75**, 11974-11982.
94. Brian, D.A. and Spaan, J.M. (1997). Recombination and Coronavirus defective interfering RNAs. *Seminars in Virology*. **8**, 101-111.
95. Bruno-Lobo, G.G., Bruno-Lobo, M., Johnson, K.M., Webbs, P.A. and De Paola, V. (1968). Pathogenesis of Junín virus infection in the infant hamster. *Annals in Microbiology (Rio de Janeiro)*. **15**, 11-33.
96. Bruns, M., Cihak, J., Müller, G. and Lehmann-Grube, F. (1983a). Lymphocytic choriomeningitis virus. VI. Isolation of a glycoprotein mediating neutralization. *Virology*. **130**, 247-251.
97. Bruns, M., Martinez Peralta, L. and Lehmann-Grube, F. (1983b). Lymphocytic choriomeningitis virus. III. Structural proteins of the virion. *Journal of General Virology*. **64**, 599-611.
98. Bruns, M., Zeller, W., Rohdewohld, H. and Lehmann-Grube, F. (1986). Lymphocytic choriomeningitis virus. IX. Properties of the nucleocapsid. *Virology*. **151**, 77-85.
99. Bruns, M., Kratzberg, T., Zeller, W. and Lehmann-Grube, F. (1990). Mode of replication of lymphocytic choriomeningitis virus in persistently infected cultivated mouse L cells. *Virology*. **177**, 615-624.
100. Brutlag, D.L. (1998). Genomics and computational biology. *Current Opinion in Microbiology*. **1**, 340-345.
101. Bruyere, A., Wantroba, M.m, Flasiński, S., Dzianott, A. and Bujarski, J.J. (2000). Frequent homologous recombination events between molecules of one RNA component in a multipartite RNA virus. *Journal of Virology*. **74**, 4214-4219.
102. Bryant, S.H. and Altschul, S.F. (1995). Statistics of sequence-structure threading. *Current Opinion in Structural Biology*. **5**, 236-244.
103. Buchmeier, M.J., Gee, S.R. and Rawls, W.E. (1977). Antigens of Pichinde virus. I. Relationship of soluble antigens derived from infected BHK-21 cells to the structural components of the virion. *Journal of Virology*. **22**, 175-186.
104. Buchmeier, M.J. and Oldstone, M.B.A. (1978). Identity of the viral protein responsible for serological cross-reactivity among the Tacaribe complex arenaviruses. In: *Negative Strand Viruses and the Host Cell* (Mahy, B.J.W. and Barry, R.D. Eds.). Academic Press, Nueva York. pp: 91-97.
105. Buchmeier, M.J., Elder, J.H. and Oldstone, M.B.A. (1978). Protein structure of lymphocytic choriomeningitis virus: identification of virus structural and cell-associated polypeptides. *Virology*. **89**, 133-145.
106. Buchmeier, M.J. and Oldstone, M.B.A. (1979). Protein structure of lymphocytic choriomeningitis virus: evidence for a cell-associated precursor of the virion glycopeptides. *Virology*. **99**, 111-120.

107. Buchmeier, M.J. and Oldstone, M.B.A. (1981). Molecular studies of LCM virus-induced immunopathology: development and characterization of monoclonal antibodies to LCM virus. In: The replication of negative strand viruses. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 71-78.
108. Buchmeier, M.J., Lewicki, H.A., Tomori, O. and Oldstone, M.B.A. (1981). Monoclonal antibodies to lymphocytic choriomeningitis and Pichinde viruses: generation, characterization and cross-reactivity with other arenaviruses. *Virology*. **113**, 73-85.
109. Buchmeier, M.J. and Parekh, B.S. (1987). Protein structure and expression among arenaviruses. *Current Topics in Microbiology and Immunology*. **133**, 41-57.
110. Buchmeier, M.J., Southern, P.J., Parekh, B.S., Wooddell, M.K. and Oldstone, M.B.A. (1987). Site-specific antibodies define a cleavage site conserved among arenaviruses GP-C glycoproteins. *Journal of Virology*. **61**, 982-985.
111. Buchholz, C.J., Spehner, D., Drillien, R., Neubert, W.J. and Homann, H.E. (1993). The conserved N-terminal region of Sendai virus nucleocapsid protein NP es required for nucleocapsid assembly. *Journal of Virology*. **67**, 5803-5812.
112. Bujarski, J.J. and Nagy, P.D. (1996). Different mechanism of homologous and nonhomologous recombination in brome mosaic virus: role of RNA sequences and replicase proteins. *Seminars in Virology*. **7**, 363-372.
113. Bullock, W.O., Fernández, J.M. and Short, J.M. (1987). XL1-Blue: a high efficiency plasmid transforming *recA Escherichia coli* strain with β -galactosidase selection. *BioTechniques*. **5**, 376
114. Burge, C., Campbell, A.M. and Karlin, S. (1992). Over- and Under- representation of short oligonucleotides in DNA sequences. *Proceedings of the National Academy of Sciences USA*. **89**, 1358-1362.
115. Burns, J.W., Salvato, M. and Buchmeier, M.J. (1990). Molecular architecture of lymphocytic choriomeningitis virus. VIII International Congress of Virology, Berlín (RFA). Resúmen W3-001.
116. Burns, J.W. and Buchmeier, M.J. (1991). Protein-protein interaction in lymphocytic choriomeningitis virus. *Virology*. **183**, 620.
117. Burns, J.W. and Buchmeier, M.J. (1993). Glycoproteins of the arenaviruses. In: The Arenaviridae. (Salvato, M.S. Ed.), Serie: The Viruses (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 17-35.
118. Busconi, L. and Michel, T. (1993). Endothelial nitric oxide synthase. N-terminal myristoylation determines subcellular localization. *Journal of Biological Chemistry*. **268**, 8410-8413.
119. Campbell, A., Mrázek, J. and Karlin, S. (1999). Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. *Proceedings of National Academy of Sciences USA*. **96**, 9184-9189.
120. Campbell Dwyer, E.J., Lai, H., MacDonald, R.C., Salvato, M.S. and Borden, K.L.B. (2000). The lymphocytic choriomeningitis virus RING protein Z associates with eukaryotic initiation factor 4E and selectively represses translation in a RING-dependent manner. *Journal of Virology*. **74**, 3293-3300.
121. Candurra, N.A., Damonte, E.B. and Coto, C.E. (1989). Antigenic relationships among attenuated and pathogenic strains of virus Junín. *Journal of Medical Virology*. **27**, 145.
122. Cao, W., Henry, M.D., Borrow, P., Yamada, H., Elder, J.H., Ravkov, E.V., Nichol, S.T., Compans, R.W., Campbel, K.P. and Oldstone, M.B. (1998). *Science*. **282**, 2079-2081.
123. Cardon, L.R., Burge, C., Clayton, D.A. and Karlin, S. (1994). Pervasive CpG suppression in animal mitochondrial genomes. *Proceedings of National Academy of Sciences USA*. **91**, 3799-3803.
124. Cardozo, T., Totrov, M. and Abagyan, R. (1995). Homology modeling by the ICM method. *Proteins*. **23**, 403-414.
125. Carpenter, C.D., Oh, J.-W., Zhang, C. and Simon, A.E. (1995). Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. *Journal of Molecular Biology*. **245**, 608-622.
126. Carpenter, C.D. and Simon, A.E. (1996). In vivo restoration of biologically active 3' ends of virus-associated RNAs by nonhomologous RNA recombination and replacement of a terminal motif. *Journal of Virology*. **70**, 478-486.
127. Carter, M.F., Biswal, N. and Rawls, W.E. (1973). Characterization of nucleic acid of Pichinde virus. *Journal of Virology*. **11**, 61-68.
128. Casals, J., Buckley, S.M. and Cedeno, R. (1975). Antigenic properties of the arenaviruses. *Bulletin WHO*. **52**, 421-427.
129. Casals, J. (1977). Serological reactions with arenaviruses. *Medicina (Buenos Aires)*. **37**, 59-68.
130. Casari, G., De Daruvar, A., Sander, C. and Schneider, R. (1996). Bioinformatics and the discovery of gene function. *Trends in Genetics*. **12**, 244-245.
131. Cascone, P.J., Haydar, T.F. and Simon, A.E. (1993). Sequences and structures required for recombination between virus-associated RNAs. *Science*. **260**, 801-805.
132. Caton, A.J. and Robertson, J.S. (1980). Structure of the host-derived sequences present at the 5' ends of influenza virus mRNA. *Nucleic Acids Research*. **8**, 2591-2603.
133. Chambers, P., Pringle, C.R. and Easton, A.J. (1990). Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *Journal of General Virology*. **71**, 3075-3081.
134. Charini, W.A., Todd, S., Gutman, G.A. and Semler, B.L. (1994). Transduction of a human RNA sequence by poliovirus. *Journal of Virology*. **68**, 6547-6552.
135. Charrel, R.N., de Lamballerie, X., De Micco, P. and Fulhorst, C.F. (2001a). Nucleotide sequence of the Pirital (Family Arenaviridae) small genomic segment. *Biochemical and Biophysical Research Communications*. **280**, 1402-1407.
136. Charrel, R.N., de Lamballerie, X. and Fulhorst, C.F. (2001b). The Whitewater Arroyo virus: natural evidence for genetic recombination among Tacaribe serocomplex viruses (family Arenaviridae). *Virology*. **283**, 161-166.
137. Chetverin, A.B., Chetverina, H.V., Demidenko, A.A. and Ugarov, V.I. (1997). Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. *Cell*. **88**, 503-513.
138. Chetverin, A.B. (1999). The puzzle of RNA recombination. *FEBS letter*. **460**, 1-5.
139. Chetverina, H.V., Demidenko, A.A., Ugarov, V.I. and Chetverin, A.B. (1999). Spontaneous rearrangements in RNA sequences. *FEBS letter*. **450**, 89-94.
140. Child, P.L., Mackenzie, R.B., Valverde, L.R. and Johnson, K.M. (1967). Bolivian hemorrhagic fever. *Virchows Archives [A]*. **368**, 1-9.
141. Childs, J.E. and Peters, C.J. (1993). Ecology and epidemiology of arenaviruses. In: The Arenaviridae. (Salvato, M.S. Ed.), Serie: The Viruses (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 325-384.
142. Chinault, D.N., Thompson, H.A. and Gangemi, J.D. (1981). Polypeptide synthesis catalyzed by components of Pichinde virus disrupted by detergent. *Journal of General Virology*. **55**, 213-217.
143. Chinae, G., Padron, G., Hooft, R.W.W., Sander, C. and Vriend, G. (1995). The use of position-specific rotamers in model building by homology. *Proteins*. **23**, 415-421.

Referencias bibliográficas

144. Chirgwin, J.M., Przybyla, A.E., McDonald, R.J. and Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*. **18**, 5294-5299.
145. Chomczynski, P. and Sacchi, N. (1987). Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*. **162**, 156-159.
146. Choppin, P.W. and Scheid, A. (1980). The role of viral glycoproteins in adsorption, penetration and pathogenicity of viruses. *Rev. Infect. Dis.* **2**, 40-61.
147. Chu, T.W. (1994). Cloning of gene encoding metal-dependent DNA binding domains from the Class I region of the MHC. Unpublished. Direct submission to GenBank (ACCESSION U09825).
148. Cianci, C., Tiley, L. and Krystal, M. (1995). Differential activation of the influenza virus polymerase via template RNA binding. *Journal of Virology*. **69**, 3995-3999.
149. Clark, G. (1980). Staining procedures. Ed. Williams and Wilkins, Baltimore, MD.
150. Clark, T., Gresser, I., Pfau, C., Moss, J. and Woodrow D. (1986). Antibody to mouse alpha/beta interferon abrogates Pichinde virus-induced liver lesions in suckling mice. *Journal of Virology*. **59**, 728-730.
151. Clegg, J.C.S. and Lloyd, G. (1983). Structural and cell-associated proteins of Lassa virus. *Journal of General Virology*. **64**, 1127-1136.
152. Clegg, J.C.S. and Oram, J.D. (1985). Molecular cloning of Lassa-virus RNA: nucleotide sequence and expression of the nucleocapsid protein gene. *Virology*. **144**, 363-372.
153. Clegg, J.C.S., Wilson, S.M. and Oram, J.D. (1990). Nucleotide sequence of the S RNA of Lassa virus (Nigerian strain) and comparative analysis of arenavirus gene products. *Virus Research*. **18**, 151-164.
154. Clegg, J.C.S. (1993). Molecular phylogeny of the arenaviruses and guide to published sequence data. In: *The Arenaviridae*. (Salvato, M.S. Ed.), Serie: *The Viruses* (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 175-187.
155. Cohen, F.E. and Presnell, S.R. (1996). In *Protein structure prediction* (Sternberg, M. J. E., eds.), *Protein structure prediction*. Oxford Univ. Press, Oxford, pp. 207-228.
156. Coimbra, T.L.M., Nassar, E.S. Burattini, M.N., de Souza, L.T.M., Ferreira, I.B., Rocco, I.M., Travassos da Rosa, A.P.A., Vasconcelos, P.F.C., Pinheiro, F.P., LeDuc, J.W., Rico Hesse, R., Gonzalez, J.P., Jahrling, P.B. and Tesh, R.B. (1994). New arenavirus isolated in Brazil. *Lancet*. **343**, 391-392.
157. Collett, M.S. (1986). Messenger RNA of the M segment RNA of Rift Valley fever virus. *Virology*. **151**, 151-156.
158. Collins, J.F. and Coulson, A.F.W. (1987). Molecular sequence comparison and alignment. In *Nucleic Acid and Protein Sequence Analysis: A Practical Approach*. Bishop, M.J. and Rawlings, C.J., eds. IRL Press, Oxford.
159. Compans, R.W., Boersma, D.P., Cash, P. et al. (1981). Molecular and genetic studies of Tacaribe, Pichinde, and lymphocytic choriomeningitis virus. In: *The replication of negative strand viruses*. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 31-42. Elsevier, Nueva York.
160. Compans, R.W. and Bishop, D.H.L. (1985). Biochemistry of arenaviruses. *Current Topics in Microbiology and Immunology*. **114**, 153-177.
161. Contigiani, M.S. and Sabattini, M.S. (1977). Virulencia diferencial de cepas del virus Junin por marcadores biológicos en ratones y en cobayos. *Medicina (Buenos Aires)*. **37**, 244.
162. Cordo, S.M., Candurra, N.A. and Damonte, E.B. (1999). Myristic acid analogs are inhibitors of Junin virus replication. *Microbes and Infections*. **1**, 609-614.
163. Cornette, J.L., Kemp, B.C., Margalit, H., Spouge, J.L., Berzofsky, J.A. and DeLisi, C. (1987). Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *Journal of Molecular Biology*. **195**, 659-685.
164. Cornette, J.L., Margalit, H., Berzofsky, J.A. and DeLisi, C. (1995). Periodic variation in side-chain polarities of T-cell antigenic peptides correlates with their structure and activity. *Proceedings of National Academy of Sciences USA*. **92**, 8368-8372.
165. Corrales, F.J. and Fersht, A.R. (1996). Kinetic significance of GroEL14. (GroES7)2 complexes in molecular chaperone activity. *Folding & Design*. **1**, 265-273.
166. Cossio, P.M., Laguens, R.P., Arana, R.M., Segal, A. and Maiztegui, J.I. (1975). Ultrastructural and immunohistochemical study of the human kidney in Argentine hemorrhagic fever. *Virchows Archives*. **368**, 1.
167. Coto, C.E., Help, G.I. and Tkaczewski, L.Z. (1972). Biological properties of Junin virus purified from infected mouse brain. *Medicina (Buenos Aires)*. **32**, 281-286.
168. Coto, C.E., León, M.E., Martínez Peralta, L., Help, G.I. and Laguens, P. (1979). Production of infectious virus and viral surface antigen in Vero cells persistently infected with Junin virus. In: *Humoral immunity in neurological diseases* (Karcher, D., Lowenthal, A. and Strosberg, A.D. Eds.). Plenum Press, Nueva York. pp: 405-415.
169. Coto, C.E., Damonte, E.B., Alche, L.E. and Sclaro, L. (1993). Genetic variation in Junin virus. In: *The Arenaviridae*. (Salvato, M.S. Ed.), Serie: *The Viruses* (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 85-101.
170. Cresta, B., Padula, P. and Martínez Segovia, Z.M. (1980). Biological properties of Junin virus proteins. I. identification of the immunogenic glycoprotein. *Intervirology*. **13**, 284-288.
171. Crothers, D.M. and Cole, P.E. (1978). In: *Transfer RNA* (S. Altman, ed.). MIT Press. Cambridge, Massachusetts, p. 196.
172. Cuervo, N.S., Guillot, S., Romanenkova, N., Combiescu, M., Aubert-Combiescu, A., Seghier, M., Caro, V., Crainic, R. and Delpyroux, F. (2001). Genomic features of intertypic recombinant Sabin poliovirus strains excreted by primary vaccinees. *Journal of Virology*. **75**, 5740-5751.
173. Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J. (1998). Jpred: a consensus secondary structure prediction server. *Bioinformatics*. **14**, 892-893.
174. Cuff, J.A. and Barton, G.J. (1999). Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins: Structure Function and Genetics*. **34**, 508-519.
175. Cummins, D., Molinas, F.C., Lerer, G., Maiztegui, J.I., Faint, R. and Machin, S.J. (1990). A plasma inhibitor of platelet aggregation in patients with Argentine hemorrhagic fever. *American Journal of Tropical Medicine and Hygiene*. **42**, 470-475.
176. D'Aiutolo, A.C. and Coto, C.E. (1986). Vero cells persistently infected with Tacaribe virus: role of interfering particles in the establishment of the infection. *Virus Research*. **6**, 235-244.
177. Dalton, A.J., Rowe, W.P. Smith, G.H. Wilsnack, R.E. and Pugh, W.E. (1968). Morphological and cytochemical studies on lymphocytic choriomeningitis virus. *Journal of Virology*. **2**, 1465-1478.

Referencias bibliográficas

178. Davis, N.L., Brown, K.W., Greenwald, G.F., Zajac, A.J., Zacny, V.L., Smith, J.F. and Johnston, R.E. (1995). Attenuated mutants of Venezuelan equine encephalitis virus containing lethal mutations in the PE2 cleavage signal combined with a second-site suppressor mutation in E1. *Virology*. **212**, 102-110.
179. de Bracco, M.M., Rimoldi, M.T., Cossio, P.M., Rabinovich, A., Maiztegui, J.I., Carballal, G. and Arana, R.M. (1978). Argentine hemorrhagic fever: Alterations of the complement system and anti-Junin virus humoral response. *New England Journal of Medicine*. **299**, 216-221.
180. De Filippis, V., Sander, C. and Vriend, G. (1994). Predicting local structural changes that result from point mutations. *Protein Engineering*. **7**, 1203-1208.
181. De Mitri, M.I. and Martínez Segovia, Z. (1971). Phytohemagglutinin unresponsiveness in leukopenia induced by Junin virus in Guinea pigs. *Acta Virologica*. **16**, 234-238.
182. De Mitri, M.I. and Martínez Segovia, Z. (1980). Biological activities of Junin virus proteins, II. Complement fixing polypeptides associated with the soluble antigen and purified virus particles. *Intervirology*. **14**, 84-90.
183. De Mitri, M.I. and Martínez Segovia, Z.M. (1985). Polypeptide synthesis in Junin virus-infected BHK-21 cells. *Acta Virologica*. **29**, 97-103.
184. Dejean, C.B., Ayerra, B.L. and Teyssie, A.R. (1987). Interferon response in the guinea pig infected with Junin virus. *Journal Medical Virology*. **23**, 83-91.
185. Dejean, C.B., Oubiña, J.R., Carballal, G. and Teyssie, A.R. (1988). Circulating interferon in the guinea pig infected with the XJ, prototype Junin virus strain. *Journal of Medical Virology*. **24**, 97-99.
186. Delort, J., Dumas, J.B., Darmon, M.C. and Mallet, J. (1989). An efficient strategy for cloning 5' extremities of rare transcripts permits isolation of multiple 5'-untranslated regions of rat tryptophan hydroxylase mRNA. *Nucleic Acids Research*. **17**, 6439-6448.
187. Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. *Biochemical and Biophysical Research Communication*. **23**, 641-653.
188. Denis, N., Corcos, D., Kruh, J. and Kitzis, A. (1988). A rapid and accurate method for quantitating total RNA transferred during Northern blot analysis. *Nucleic Acids Research*. **16**, 2354.
189. Deperieux, E. and Feytmans, E. (1992). MATCH-BOX: a fundamentally new algorithm for the simultaneous alignment of several protein sequences. *CABIOS*. **8**, 501-509.
190. Deshpande, K.L., Fried, V.A., Ando, M. and Webster, R.G. (1987). Glycosylation affects cleavage of an H5N2 influenza virus hemagglutinin and regulates virulence. *Proceedings of the National Academy of Sciences, USA*. **84**, 36-40.
191. Desport, M., Collins, M.E. and Brownlie, J. (1998). Genome instability in BVDV: an examination of the sequence and structural influences on RNA recombination. *Virology*. **246**, 352-361.
192. Deveraux, J., Haerberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research*. **12**: 387-395.
193. Di Francesco, V., Garnier, J. and Munson, P.J. (1996). Improving protein secondary structure prediction with aligned homologous sequences. *Protein Science*. **5**, 106-113.
194. Ding, A-W., Shi, B-J., Li, W-X. and Symons, R.H. (1996). An interspecies hybrid RNA virus is significantly more virulent than either parental virus. *Proceedings of National Academy of Sciences USA*. **93**, 7470-7474.
195. Di Simone, C., Zandonatti, M.A. and Buchmeier, M.J. (1994). Acidic pH triggers LCMV membrane fusion activity and conformational change in the glycoprotein spike. *Virology*. **198**, 455-465.
196. Di Simone, C. and Buchmeier, M.J. (1995). Kinetics of pH dependence of acid-induced structural changes in the Lymphocytic Choriomeningitis Virus glycoprotein complex. *Virology*. **209**, 3-9.
197. Doherty, P., Huesca Contreras, M., Dosch, M. and Pan, S. (1989). Rapid amplification of complementary DNA from small amounts of unfractionated RNA. *Analytical Biochemistry*. **177**, 7-10.
198. Dolja, V.V. and Carrington, J.C. (1992). Evolution of positive-strand RNA viruses. *Seminars in Virology*. **3**, 315-326.
199. Domingo, E. (1992). Genetic variations and quasispecies. *Current Opinion in Genetics & Development*. **2**, 61-63.
200. Domingo, E., Escarnis, C., Sevilla, N., Moya, A., Elena, S.F., Quer, J., Novella, I.S. and Holland, J.J. (1996). Basic concepts in RNA virus evolution. *FASEB Journal*. **10**, 859-864.
201. Domingo, E., Baranowski, E., Ruiz-Jarabo, C.M., Martín-Hernández, A.M., Sáiz, J.C. and Escarnis, C. (1998). Quasispecies structure and persistence of RNA viruses. *Emerging Infectious Diseases*. **4**, 521-527.
202. Domingo, E. (2002). Quasispecies theory in virology. *Journal of Virology*. **76**, 463-465.
203. Dong, B. and Silverman, R.H. (1997). A bipartite model of 2-5A-dependent RNase L. *Journal of Biological Chemistry*. **272**, 22236-22242.
204. Doolittle, R.F. (1986). Of URFs and ORFs: a primer on how to analyze derived amino acid sequences. University Science Books, Mill Valley California.
205. Doolittle, R.F. (1996). Computer methods for macromolecular sequence analysis. Academic Press, San Diego.
206. D'Souza-Schorey, C. and Stahl, P.D. (1995). Myristoylation is required for the intracellular localization and endocytic function of ARF6. *Experimental Cellular Research*. **221**, 153-159.
207. Duggal, R., Cuconati, A., Gromeier, M. and Wimmer, E. (1997). Genetic recombination of poliovirus in a cell-free system. *Proceedings of National Academy of Sciences USA*. **94**, 13786-13791.
208. Dulbecco, R. (1952). Production of plaques in monolayers tissue culture by single particles of an animal virus. *Proceedings of National Academy of Sciences USA*, **38**, 747-756.
209. Dulbecco, R. (1980). The nature of viruses. In: *Microbiology*, third edition (Davis, B.D., Dulbecco, R., Eisen, H.N. and Ginsberg, H.S., Eds.). Harper and Row Publishers, Philadelphia. pp: 853-884.
210. Dumas, J.B., Milne, E., Delort, J. and Mallet, J. (1991). Oligodeoxyribonucleotide ligation to single-stranded cDNAs: a new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by *in vitro* amplification. *Nucleic Acids Research*. **19**, 5227-5232.
211. Dunn, E.F., Pritlove, D.C. and Elliott, R.M. (1994). The S RNA genome segments of Batai, Cache Valley, Guaroa, Kairi, Lumbo, Main Drain and Northway bunyaviruses: sequence determination and analysis. *Journal of General Virology*. **75**, 597-608.
212. Dunnill, P. (1968). The use of helical net-diagrams to represent protein structures *Biophysical Journal* **8**, 865-875.
213. Dutko, F.J., Wright, E.A. and Pfau, C.J. (1976). The RNAs of defective interfering Pichinde virus. *Journal of General Virology*. **31**, 417-427.

Referencias bibliográficas

214. Dutko, F.J. and Pfau, C.J. (1978). Arenaviruses defective interfering particles mask the cell killing potential of standard virus. *Journal of General Virology*. **38**, 195-204.
215. Dutko, F.J., Kennedy, S.I. and Oldstone, M.B.A. (1981). Genome structure of lymphocytic choriomeningitis virus: complementary termini. In: *The replication of negative strand viruses*. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 43-48.
216. Dziafott, A., Flasiński, S. and Bujarski, J.J. (1995). Foreign complementary sequences facilitate genetic RNA recombination in brome mosaic virus. *Virology*. **208**, 370-375.
217. Dziafott, A., Rauffer-Bruyere, N. And Bujarski, J.J. (2001). Studies on functional interaction between Brome Mosaic virus replicase proteins during RNA recombination, using combined mutants *in vivo* and *in vitro*. *Virology*. **289**, 137-149.
218. Eddy, G.A., Wagner, F.S., Scott, S.K. and Mahlauth, B.J. (1975). Protection of monkeys against Machupo virus by the passive administration of Bolivian hemorrhagic fever immunoglobulin (human origin). *Bulletin WHO*. **52**, 723-727.
219. Efremov, R.G. and Vergoten, G. (1996). Recognition of transmembrane α -helical segments with environmental profiles. *Protein Engineering*. **9**, 253-263.
220. Eisenberg, D., Schwarz, E., Komaromy, M. and Wall, R. (1984a). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of Molecular Biology*. **179**, 125-142.
221. Eisenberg, D., Weiss, R.M. and Terwilliger, T.C. (1984b). The hydrophobic moment detects periodicity in protein hydrophobicity. *Proceedings of National Academy of Sciences USA*. **81**, 140-144.
222. Eisenberg, D., Lüthy, R. and McLachland, A.D. (1991). Secondary structure-based profiles: Use of structure-conserving scoring tables in searching protein sequence databases for structural similarities. *Proteins*. **10**, 229-239.
223. Elliot, R.M. and McGregor, A. (1989). Nucleotide sequence and expression of the small (S) RNA segment of Maguari bunyavirus. *Virology*. **171**, 516-524.
224. Elliot, R.M. (1990). Molecular biology of Bunyavirus. *Journal of General Virology*. **71**, 501-522.
225. Ellis, H.M.; Spann, D.R. and Posakony, J.W. (1990). Extramacrochaetae, a Negative Regulator of Sensory Organ Development in *Drosophila*, Defines a New Class of Helix-Turn-Helix Proteins. *Cell*. **61**, 27-38.
226. Elofsson, A., Le Grand, S.M. and Eisenberg, D. (1995). Local moves: an efficient algorithm for simulation of protein folding. *Proteins*. **23**, 73-82.
227. Elsner, B., Schwarz, E., Mando, O.G., Maiztegui, J.I. and Vilches, A. (1973). Pathology of 12 fatal cases of Argentinian hemorrhagic fever. *American Journal of Tropical Medicine and Hygiene*. **22**, 229-236.
228. Engelman, D.M., Stitz, T.A. and Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annual Review of Biophysical Chemistry*. **15**, 321-354.
229. Enría, D., Briggiler, A.M., Fernández, J.H., Levis, S. and Maiztegui, J.I. (1984). Importance of dose of neutralizing antibodies in treatment of Argentine hemorrhagic fever with immune plasma. *Lancet*. **2**, 255.
230. Enría, D., García-Franco, S., Ambrosio, A., Vallejos, D., Levis, S. and Maiztegui, J.I. (1986). Current status of the treatment of Argentine Hemorrhagic Fever. *Medical Microbiology and Immunology*. **175**, 173-176.
231. Esposito, G., Lesk, A.M., Molinari, H., Motta, A., Niccolai, N. and Pastore, A. (1994). In *Protein Structure by Distance Analysis* (Bohr, H. and Brunak, S., eds.), *Protein Structure by Distance Analysis*. IOS Press, Amsterdam, Oxford, Washington DC, pp. 51-63.
232. Estabrook, E.M., Tsai, J. And Falk, B.W. (1998). *In vivo* transfer of barley stripe mosaic hordeivirus ribonucleotides to the 5' terminus of maize stripe tenuivirus RNAs. *Proceedings of National Academy of Sciences USA*. **95**, 8304-8309.
233. Farber, F.E. and Rawls, W.E. (1975). Isolation of ribosome-like structures from Pichinde virus. *Journal of General Virology*. **26**, 21-31.
234. Fariselli, P. and Casadio, R. (1996). HTP: a neural network-based method for predicting the topology of helical transmembrane domains in proteins. *CABIOS*. **12**, 41-48.
235. Feinberg, A.P. and Volgelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry*. **132**, 6-14.
236. Feldmann, H., Klenk, H.D. and Sanchez, A. (1993). Molecular biology and evolution of filoviruses. *Archives in Virology Suppl*. **7**, 81-100
237. Felgner, P.L., Gadek, T., Holm, M., Roman, R., Chan, H.W., Wenz, M., Northrop, J.P., Ringold, G.M. and Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of National Academy of Sciences USA*. **84**, 7413-7417.
238. Felsenstein, J. (1995). Phylogeny Inferencing Package. Version 3.5. University of Washington. USA.
239. Feng, D.-F. and Doolittle, R.F. (1996). Progressive alignment of amino acid sequences and construction of phylogenetic trees from them. *Methods in Enzymology*. **266**, 368-382.
240. Fenner, F., Mc Auslan, B.R., Mims, C.A., Sambrook, J and White, D.D. (1974). *The biology of animal viruses*. Academic Press, Nueva York.
241. Fickett, J.W. (1982). Recognition of protein coding regions in DNA sequences. *Nucleic Acids Research*. **10**, 5303-5318.
242. Figlerowicz, M., Nagy, P.D. and Bujarski, J.J. (1997). A mutation in the putative RNA polymerase gene inhibits nonhomologous, but not homologous, genetic recombination in an RNA virus. *Proceedings of National Academy of Sciences USA*. **94**, 2073-2078.
243. Figlerowicz, M. (2000). Role of RNA structure in non-homologous recombination between genomic molecules of brome mosaic virus. *Nucleic Acids Research*. **28**, 1714-1723.
244. Fischer, D. and Eisenberg, D. (1996). Fold recognition using sequence-derived properties. *Protein Science*. **5**, 947-955.
245. Fischer, D., Rice, D.W., Bowie, J.U. and Eisenberg, D. (1996b). Assigning amino acid sequences to 3D protein folds. *FASEB Journal*. **10**, 126-136.
246. Flöckner, H., Braxenthaler, M., Lackner, P., Jaritz, M., Ortner, M. and Sippl, M.J. (1995). Progress in fold recognition. *Proteins*. **23**, 376-386.
247. Forss, S. and Schaller, H. (1982). A tandem repeat gene in a picornavirus. *Nucleic Acids Research*. **10**, 6441-6450.
248. Francis, S.J., Southern, P.J., Valsamakis, A. and Oldstone, M.B.A. (1987). State of viral genome and protein during persistently lymphocytic choriomeningitis virus infections. *Current Topics in Microbiology and Immunology*. **133**, 67-88.
249. Francis, S.J. and Southern, P.J. (1988). Molecular analysis of viral RNAs in mice persistently infected with lymphocytic choriomeningitis virus. *Journal of Virology*. **62**, 1251-1257.
250. Franze Fernández, M.T., Zetina, C., Iapalucci, S., Lucero, M.A., Boissou, C., López, R., Rey, O., Daheli, M., Cohen, G. and Zalein, M. (1987). Molecular structure and early events in the replication of Tacaribe Arenavirus S RNA. *Virus Research*. **7**, 309-324.

Referencias bibliográficas

251. Freier, S.M.; Kierzek, R.; Jaeger, J.A.; Sugimoto, N.; Caruthers, M.H.; Neilson, T. and Turner, D.H. (1986). Improved free-energy parameters for predictions of rna duplex stability. *Proceedings of National Academy of Sciences USA*. **83**, 9373-9377.
252. Frishman, D. and Argos, P. (1997). Seventy-five percent accuracy in protein secondary structure prediction. *Proteins*. **27**, 329-335.
253. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences USA*. **85**, 8998-9002.
254. Fuchs, R. (1991). MacPattern: protein pattern searching on the Apple Macintosh. *CABIOS*. **7**, 105-106.
255. Fuchs, R. (1994). Predicting protein function: a versatile tool for the Apple Macintosh. *CABIOS*. **10**, 171-178.
256. Fulhorst, C.F., Bowen, M.D., Hardy, J.L., Eldridge, B.F., Chiles, R.E., Jackson, A.O. and Reeves, W.C. (1996). Geographic distribution and serologic and genomic characterization of Morro Bay virus, a newly recognized bunyavirus. *American Journal of Tropical Medicine and Hygiene*. **54**, 563-569.
257. Fulhorst, C.F., Bowen, M.D., Ksiazek, Rollin, P.E., Nichol, S.T., Kosoy, M.Y. and Peters, C.J. (1996). Isolation and characterization of Whitewater Arroyo virus, a novel North American arenavirus. *Virology*. **224**, 114-120.
258. Fuller-Pace, F.V. and Southern, P.J. (1988). Temporal analysis of transcription and replication during acute infection with lymphocytic choriomeningitis virus. *Virology*. **162**, 260-263.
259. Fuller-Pace, F.V. and Southern, P.J. (1989). Detection of virus-specific RNA-dependent RNA polymerase activity in extracts from cells infected with lymphocytic choriomeningitis virus: *in vitro* synthesis in full-length viral RNA species. *Journal of Virology*. **63**, 1938-1944.
260. Galaktionov, S.G. and Marshall, G.R. (1996). Prediction of protein structure in terms of intraglobular contacts: 1D to 2D to 3D. Preprint, Inst for Biomedical Computing; Washington Univ., St. Louis, Missouri. (Nunca publicado, 2001)
261. Galarza, J.M., Peng, Q., Shi, L. and Summers, D.F. (1996). Influenza A virus RNA-dependent RNA polymerase: analysis of RNA synthesis *in vitro*. *Journal of Virology*. **70**, 2360-2368.
262. Galinski, M., Mink, M. and Pons, M. (1988). Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene encoding the L protein. *Virology*. **165**, 499-510.
263. Gangemi, J.D., Rosato, R.R., Connell, E.V., Johnson, E.M. and Eddy, G.A. (1978). Structural polypeptides of Machupo virus. *Journal of General Virology*. **41**, 183-188.
264. Garay, R.P. (1975). Patogenia de la encefalitis por virus Junin: estudios en ratón lactante. *Medicina (Buenos Aires)*. **31**, 383-392.
265. Garcia, J.B., Morzunov, S.P., Levis, S., Rowe, J., Caldrón, G., Enría, D., Sabattini, M., Buchmeier, M.-J., Bowen, M.D. and St. Jeor, S.C. (2000). Genetic diversity of the Junin virus in Argentina: geographic and temporal patterns. *Virology*. **272**, 127-136.
266. García Franco, S., Ambrosio, A.M., Feuillade, M.R. and Maiztegui, J.I. (1988). Evaluation of an enzyme-linked immunosorbent assay for quantitation of antibodies to Junin virus in human sera. *Journal of Virological Methods*. **19**, 299-306
267. Garcin, D. and Kolakofsky, D. (1990). A novel mechanism for the initiation of Tacaribe arenavirus genome replication. *Journal of Virology*. **64**, 6196-6203.
268. Garcin, D. and Kolakofsky, D. (1992). Tacaribe arenavirus RNA synthesis *in vitro* is primer dependent, and suggests an unusual model for the initiation of genome replication. *Journal of Virology*. **66**, 1370-1376.
269. Garcin, D., Rochat, S. and Kolakofsky, D. J. (1993). The Tacaribe arenavirus small zinc finger protein is required for both mRNA synthesis and genome replication. *Virology*. **67**, 807-812.
270. Garcin, D., Lezzi, M., Dobbs, M. Elliott, R.M., Schmaljohn, C., Kang, C.Y. and Kolakofsky, D. (1995). The 5' ends of Hantaan virus (*Bunyaviridae*) RNAs suggest a Prime-and-Realign mechanism for the initiation of RNA synthesis. *Journal of Virology*. **69**, 5754-5762.
271. Gard, G.P., Vezza, A.C., Bishop, D.H.L. and Compans, R.W. (1977). Structural proteins of Tacaribe and Tamiami virions. *Virology*. **83**, 85-95.
272. Garnier, J., Gibrat, J.-F. and Robson, B. (1996). GOR method for predicting protein secondary structure from amino acid sequence. *Methods in Enzymology*. **266**, 540-553.
273. Geourjon, C. and Deléage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *CABIOS*. **11**, 681-684.
274. Gerbaud, S., Vialat, P., Pardigon, N., Wychowski, C., Girard, M. and Bouloy, M. (1987). The S segment of the Germiston virus RNA genome can code for three proteins. *Virus Research*. **8**, 1-13.
275. Gerhardt, P. *et al.* (1981). Manual of methods for general bacteriology. American Society for Microbiology, Washington, DC.
276. Ghiringhelli, P.D., Romanowski, V. and Grau, O. (1985). Síntesis enzimática de $\gamma^{32}\text{P}$ y $\alpha^{32}\text{P}$ nucleósidos trifosfatos. XXI Reunión anual de la Sociedad Argentina de Investigación Bioquímica. Huerta Grande, Córdoba.
277. Ghiringhelli, P.D., Rivera Pomar, R.V., Baro, N., Grau, O. and Romanowski, V. (1987). Molecular cloning of Junin virus RNA. VII International Congress of Virology (Edmonton, Canada), abstract OP3.10, pp 253.
278. Ghiringhelli, P.D., Rivera Pomar, R.V., Baro, N.I., Rosas, M.F., Grau, O. and Romanowski, V. (1989). Nucleocapsid protein gene of Junin arenavirus (cDNA sequence). *Nucleic Acids Research*. **17**, 8001.
279. Ghiringhelli, P.D., Rivera Pomar, R.V., Lozano, M.E., Grau, O. and Romanowski, V. (1991a). Molecular organization of Junin virus S RNA: complete nucleotide sequence, relationship with the other members of Arenaviridae and unusual secondary structures. *Journal of General Virology*. **72**, 2129-2141.
280. Ghiringhelli, P.D., Albariño, C.G., Piboul, M.P. and Romanowski, V. (1991b). The glycoprotein precursor gene of the Junin vaccine strain. 10th Annual Meeting of the American Society for Virology (Fort Collins, USA).
281. Ghiringhelli, P.D., Rivera-Pomar, R.V., Albariño, C.G., Lozano, M.E. Tasso, S.M. y Romanowski, V. (1993). Interacción de la proteína N de arenavirus con Zn^{++} y RNA: bases moleculares para su función regulatoria. XXIX Reunión Nacional de la Sociedad Argentina de Investigación Bioquímica (SAIB). Villa Carlos Paz, Córdoba, 17-20 de noviembre.
282. Ghiringhelli, P.D. and Romanowski, V. (1994). A quick protocol for visualizing virus plaques in titration experiments. *Biotechniques*. **17**, 464-465.
283. Ghiringhelli, P.D., Albariño, C.G., Piboul, M.P. and Romanowski, V. (1997). The glycoprotein precursor gene of the attenuated Junin virus vaccine strain (Candid # 1). *American Journal of Tropical Medicine and Hygiene*. **56**, 216-225.
284. Ghiringhelli, P.D., Gofí, S., Lozano, M.E., Posik, D.M. and Romanowski, V. (2002). Sources of genetic variation in Arenaviridae: theoretical models and evidence in nature. XII International Congress of Virology. Paris, Francia. (aceptado para su presentación).
285. Gibbs, M.J. and Weiller, G.F. (1999). Evidence that a plant virus switched hosts to infect a vertebrate and then recombined with a vertebrate-infecting virus. *Proceedings of National Academy of Sciences USA*. **96**, 8022-8027.
286. Giménez, H.B. and Compans, R.W. (1980). Defective interfering Tacaribe virus and persistently infected cells. *Virology*. **107**, 229-239.

287. Giranda, V.L., Heinz, B.A., Oliveira, M.A., Minor, I., Kim, K.H., Kolatkar, P.R., Rossman, M.G. and Rueckert, R.R. (1992). Acid-induced structural changes in human rhinovirus 14: possible role in uncoating. *Proceedings of National Academy of Sciences USA*. **89**, 10213-10217.
288. Girvitz, S.C., Bacchetti, S., Rainbow, A.J. and Graham, F.I. (1980). A rapid and efficient procedure for the purification of DNA from agarose gels. *Analytical Biochemistry*. **106**, 492-498.
289. Gish, W. and States, D.J. (1993). Identification of protein coding regions by database similarity search. *Nature Genetics*. **3**, 266-272.
290. Glushakova, S.E. and Lukashovich, I.S. (1989). Early events in arenavirus replication are sensitive to lysosomotropic compounds. *Archives in Virology*. **104**, 157.
291. Glushakova, S.E., Lukashovich, I.S. and Baratova, L.A. (1990). Prediction of arenaviruses fusion peptides on the basis of computer analysis of envelope protein sequences. *FEBS Letter*. **269**, 145-147.
292. Gmyl, A.P., Belousov, E.V., Maslova, S.V., Khitrina, E.V. Chetverin, A.B. and Agol, V.I. (1999). Nonreplicative RNA recombination in poliovirus. *Journal of Virology*. **73**, 8958-8965.
293. Golub, E. (1988). "One minute" transformation of competent *E. coli* by plasmid DNA. *Nucleic Acids Research*. **16**, 1641.
294. Gombart, A.F., Hirano, A. and Wong, T.C. (1995). Nucleoprotein phosphorylated at both serine and threonine is preferentially assembled into the nucleocapsids of measles virus. *Virus Research*. **37**, 63-73.
295. González, J.P., McCormick, J.B., Saluzzo, J.F. and Herve, J.P. (1983). An arenavirus isolated from wild-caught rodents (*Pramys* species) in the Central African Republic. *Intervirology*. **19**, 105-112.
296. Gonzalez, J.P., Bowen, M.D., Nichol, S.T. and Rico-Hesse, R. (1996). Genetic characterization and phylogeny of Sabiá virus, an emergent pathogen in Brazil. *Virology*. **221**, 318-324.
297. Gorman, C. (1986). High efficiency gene transfer into mammalian cells. In: DNA cloning, Vol. II (Glover, D.M. Ed.). IRL Press, Oxford, Inglaterra. pp: 143-190.
298. Grand, R.J. (1989). Acylation of viral and eukaryotic proteins. *Biochemical Journal*. **258**, 625-638.
299. Grau, O., Franze Fernández, M.T., Romanowski, V., Rustici, S.M. and Rosas, M.F. (1981). Junin virus structure. In: The replication of negative strand viruses. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 11-14.
300. Grau, O., Romanowski, V., Baro, N.I., Remorini, P. G. and Ghiringhelli, P.D. (1987). Reasociación entre los arenavirus Junin y LCM. *Taller Latinoamericano de Virología Molecular (Santiago de Chile)*, pp. 21-26.
301. Gray, P.M.D., Kemp, G.J.L., Rawlings, C.J., Brown, N.P., Sander, C., Thornton, J.M., Orengo, C.M., Wodak, S.J. and Richelle, J. (1996). Macromolecular structure information and databases. *TIBS*. **21**, 251-256.
302. Gribskov, M. and Veretnik, S. (1996). Identification of sequence patterns with profile analysis. *Methods in Enzymology*. **266**, 198-227.
303. Griffiths, C.M., Wilson, S.M. and Clegg, J.C.S. (1992). Sequence of the nucleocapsid protein gene of Machupo virus: close relationship with another South American pathogenic arenavirus, Junin. *Archives in Virology*. **124**, 371-374.
304. Grunstein, M. and Hogness, D.S. (1975). Colony hybridization: a method for the isolation of cloned DNA that contains a specific gene. *Proceedings of National Academy of Sciences USA*. **72**, 3961-3965.
305. Gschwender, H.H. and Lehmann-Grube, F. (1973). Antigenic properties of the LCM virus: virion and complement-fixing antigen. In: Lymphocytic Choriomeningitis Virus and Other Arenaviruses. (F. Lehmann-Grube, ed.). Springer-Verlag, Berlin. pp: 25-35.
306. Gschwender, H.H., Brummund, M., Lehmann-Grube, F. (1975). Lymphocytic choriomeningitis virus. I. Concentration and purification of the infectious virus. *Journal of Virology*. **15**, 1317-1322.
307. Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene*. **25**, 263-269.
308. Gubler, U. (1988). A one tube reaction for the synthesis of blunt-ended double-stranded cDNA. *Nucleic Acids Research*. **16**, 2726.
309. Guerrero, L.B., Weissenbacher, M.C. and Parodi, A.S. (1969). Inmunización contra FHA con una cepa atenuada del virus Junin. I. Estudio de una cepa modificada del virus Junin. Inmunización de cobayos. *Medicina (Buenos Aires)*. **29**, 1.
310. Hahn, C.S., Dalrymple, J.M., Strauss, J.H. and Rice, C.M. (1987). Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. *Proceedings of National Academy of Sciences USA*. **84**, 2019-2023.
311. Hajjar, A.M. and Linial, M.L. (1993). A model system for nonhomologous recombination between retroviral and cellular RNA. *Journal of Virology*. **67**, 3845-3853.
312. Hajjou, M., Hill, K.R., Subramaniam, S.V., Hu, J.Y. and Raju, R. (1996). Nonhomologous RNA-RNA recombination events at the 3' nontranslated region of the Sindbis virus genome: hot spots and utilization of nonviral sequences. *Journal of Virology*. **70**, 5153-5164.
313. Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology*. **166**, 557-580.
314. Hansen, J.E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J.O., Hansen, J.-E.S. and Brunak, S. (1995). Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase. *Biochemistry Journal*. **308**, 801-813.
315. Hansen, J.E., Lund, O., Nielsen, J.O., Brunak, S. and hansen, J.-E.S. (1996). Prediction of the secondary structure of HIV-1 gp120. *Proteins*. **25**, 1-11.
316. Harada, K. And Orgel, L.E. (1993). Unexpected substrate-specificity of T4 DNA-ligase revealed by *in vitro* selection. *Nucleic Acids Research*. **21**, 2287-2291.
317. Harnish, D., Dimock, K., Leung, W.C., Rawls, W. (1981a). Immunoprecipitable polypeptides in Pichinde virus infected BHK-21 cells. In: The replication of negative strand viruses. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 23-29.
318. Harnish, D.G., Leung, W.C. and Rawls, W.E. (1981b). Characterization of polypeptides immunoprecipitable from Pichinde virus infected BHK-21 cells. *Journal of Virology*. **38**, 840-848.
319. Harnish, D.G., Dimock, K. and Bishop, D.H.L. (1983). Gene mapping in Pichinde virus: assignment of viral polypeptides to genomic L and S RNAs. *Journal of Virology*. **46**, 638-641.
320. Harnish, D.G., Shivaprakash, M., Sisson, G. and Rawls, W.E. (1989). Analysis of RNA during the replication of wild type and temperature-sensitive mutants of Pichinde virus. In: Genetics and Pathogenicity of Negative Stranded Viruses. (Kolakofsky, D. and Mahy, B. Eds.). Elsevier Science Publishers BV. pp: 182-188.
321. Harnish, D.G., Polyak, S.J. and Rawls, W.E. (1993). Arenavirus Replication. Molecular dissection of the role of viral protein and RNA. In: The Arenaviridae. (Salvato, M.S. Ed.), Serie: The Viruses (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 157-174.
322. Hartl, F.-U., Hlodan, R. and Langer, T. (1994). Molecular chaperones in protein folding: the art of avoiding sticky situations. *TIBS*. **19**, 20-25.

323. Hartly, R.N. and Palese, P. (1995). Measles virus phosphoprotein (P) requires the NH₂- and COOH-terminal domains for interactions with the nucleoprotein (N) but only the COOH terminus for interactions with itself. *Journal of General Virology*. **76**, 2863-2867.
324. Hein, J. (1990). Reconstructing evolution of sequences subject to recombination using parsimony. *Mathematical Biosciences*. **98**, 185-200.
325. Helenius, A., Kielian, M., Mellman, I. and Schmid, S. (1989). Entry of the enveloped animal viruses into their host cells. *UCLA Symp. Mol. Cell. Biol.* **90**, 145-161.
326. Help, G.I., Coto, C.E. and Tkaczewski, L.Z. (1970). Evaluación de un método de purificación del virus Junín. *Medicina (Buenos Aires)*. **30**, 15-21.
327. Help, G.I. and Coto, C.E. (1980). Génesis de partículas interferentes durante la multiplicación del virus Junín *in vivo*. *Medicina (Buenos Aires)*. **40**, 531-536.
328. Henikoff, S. and Henikoff, J.G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of National Academy of Sciences USA*. **89**, 10915-10919.
329. Henikoff, S. and Henikoff, J.G. (1993). Performance evaluation of amino acid substitution matrices. *Proteins*. **17**, 49-61.
330. Henikoff, J.G. and Henikoff, S. (1996a). Blocks database and its applications. *Methods in Enzymology*. **266**, 88-104.
331. Henikoff, J.G. and Henikoff, S. (1996b). Using substitution probabilities to improve position-specific scoring matrices. *CABIOS*. **12**, 135-143.
332. Higgins, D.G. and Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. *Gene*. **73**, 237-244.
333. Higgins, D.G. and Sharp, P.M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS*. **5**, 151-153.
334. Higgins, D.G., Thompson, J.D. and Gibson, T.J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods in Enzymology*. **266**, 383-402.
335. Hill, K.R., Hajjou, M., Hu, J.Y. and Raju, R. (1997). RNA-RNA recombination in Sindbis virus: roles of the 3' conserved motif, poly(A) tail, and nonviral sequences of templates RNAs in polymerase recognition and template switching. *Journal of Virology*. **71**, 2693-2704.
336. Hirschhorn, R., Tzall, S. and Ellenbogen, A. (1990). Hot spot mutations in adenosine deaminase deficiency. *Proceedings of National Academy of Sciences USA*. **87**, 6171-6175.
337. Holland, J.J., Semler, B.L., Jones, C., Perrault, J., Reid, L. and Roux, L. (1978). Role of DI₁ virus mutation, and host response in persistent infections by envelope RNA viruses. In: *Persistent viruses*. Academic Press, Inc., New York. pp. 57-72.
338. Holland, J.J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. and Vanderpol, S. (1982). Rapid evolution of RNA genomes. *Science*. **215**, 1577-1585.
339. Holm, L. and Sander, C. (1994). The fssp database of structurally aligned protein fold families. *Nucleic Acids Research*. **24**, 206-210.
340. Homann, H.E., Willenbrink, W., Buchholz, C.J. and Neubert, W.J. (1991). Sendai virus protein-protein interactions studied by a protein-blotting overlay technique: mapping of domains on NP protein required for binding to P protein. *Journal of Virology*. **65**, 1304-1309.
341. Hooft, R.W., Vriend, G., Sander, C. and Abola, E.E. (1996). Errors in protein structures. *Nature*. **381**, 272.
342. Hotchin, J. (1974). The role of transient infection in Arenavirus persistence. *Progress in Medical Virology*. **18**, 81-93.
343. Houge, G., Robaye, B., Eikhom, T.S., Golstein, J., Mellgren, G., Gjertsen, B.T., Lanotte, M. and Døskeland, S.O. (1995). Fine mapping of 28S rRNA sites specifically cleaved in cells undergoing apoptosis. *Molecular and Cellular Biology*. **15**, 2051-2062.
344. Howard, C. and Buchmeier, M.J. (1983). A protein kinase activity in lymphocytic choriomeningitis virus and identification of the phosphorylated product using monoclonal antibody. *Virology*. **126**, 538-547.
345. Hsu, M., Parvin, J.D., Gupta, S., Krystal, M. and Palese, P. (1987). Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proceedings of National Academy of Sciences USA*. **84**, 8140-8144.
346. Hu, H.M., O'Rourke, K., Boguski, M.S. and Dixit, V.M. (1994). A novel RING finger protein interacts with the cytoplasmic domain of CD40. *Journal of Biological Chemistry*. **269**, 30069-30072. (ACCESSION U15637)
347. Huang, A.S. and Baltimore, D. (1970). Defective viral particles and viral disease processes. *Nature (Londres)*. **226**, 325-327
348. Huang, A.S. and Baltimore, D. (1977). Defective interfering animal viruses. In: *Comprehensive Virology* (Fraenkel Conrat, H. and Wagner, R.R. Eds.). Plenum Press, Nueva York. Vol. **10**, pp. 73-116.
349. Hubbard, T.J.P. and Park, J. (1995). Fold recognition and ab initio structure predictions using Hidden Markov models and β -strand pair potentials. *Proteins*. **23**, 398-402.
350. Hubbard, T., Tramontano, A., Barton, G., Jones, D., Sippl, M., Valencia, A., Lesk, A., Moulton, J., Rost, B., Sander, C., Schneider, R., Lahm, A., Leplae, R., Buta, C., Eisenstein, M., Fjellström, O., Floeckner, H., Grossmann, J.G., Hansen, J., Helmer-Citterich, M., Joergensen, F.S., Marchler-Bauer, A., Osuna, J., Park, J., Reinhardt, A., Ribas de Pouplana, L., Rojo-Dominguez, A., Saudek, V., Sinclair, J., Sturrock, S., Vinclovas, C. and Vinals, C. (1996). Update on protein structure prediction: Results of the 1995 IRBM workshop. *Folding & Design*. **1**, R55-R63.
351. Huber, T. and Torda, A.E. (1998). Protein fold recognition without Boltzmann statistics or explicit physical basis. *Protein Science*. **7**, 142-149
352. Hughey, R. and Krogh, A. (1996). Hidden Markov models for sequence analysis: extension and analysis of the basic method. *CABIOS*. **12**, 95-107.
353. Huntley, C.C., De, B.P. and Banerjee, A.K. (1997). Phosphorylation of Sendai virus phosphoprotein by cellular protein kinases ζ . *Journal of Biological Chemistry*. **272**, 16578-16584.
354. Hwang, C.K., Svarovskaia, E.S. and Pathak, V.K. (2001). Dynamic copy choice: steady state between murine leukemia virus polymerase and polymerase-dependent RNase H activity determines frequency of *in vivo* template switching. *Proceedings of National Academy of Sciences USA*. **98**, 12209-12214.
355. Iapalucci, S., López, R., Rey, O., López, N., Franze-Fernández, M.T., Cohen, G.N., Lucero, M., Ochoa, A. and Zakin, M.M. (1989a). Tacaribe virus L gene encodes a protein of 2210 amino acid residues. *Virology*. **170**, 40-47.
356. Iapalucci, S., López, N., Rey, O., Zakin, M.M., Cohen, G.N. and Franze-Fernández, M.T. (1989b). The 5' region of Tacaribe virus L RNA encodes a protein with a potential metal binding domain. *Virology*. **173**, 357-361.
357. Iapalucci, S., López, N. and Franze Fernández, M.T. (1991). The 3' end of the Tacaribe arenavirus subgenomic RNAs. *Virology*. **182**, 269-278.
358. Innis, M.A., Myambo, K.B., Gelfand, D.H. and Brow, M.A. (1988). DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proceedings of National Academy of Sciences USA*. **85**, 9436-9440.

359. Inokuchi, Y. and Hirashima, A. (1987). Interference with viral infection by defective RNA replicase. *Journal of Virology*. **61**, 3946-3949
360. Inoue, S., Orimo, A., Hosoi, T., Kondo, S., Toyoshima, H., Kondo, T., Ikegami, A., Ouchi, Y., Orimo, H. and Muramatsu, M. (1993). Genomic binding-site cloning reveals an estrogen-responsive gene that encodes a RING finger protein. *Proceedings of National Academy of Sciences USA*. **90**, 11117-11121. (ACCESSION D21205).
361. Ishida, A., Asano, H., Hasegawa, M., Koseki, H., Ono, T., Yoshida, M., Taniguchi, M. and Kanno, M. (1993). Cloning and chromosome mapping of the human Mel-18 gene which encodes a DNA-binding protein with a new 'RING-finger' motif. *Gene*. **129**, 249-255. (ACCESSION D13969).
362. Israelachvili, J. and Wennerström, H. (1996). Role of hydration and water structure in biological and colloidal interactions. *Nature*. **379**, 219-225.
363. Jaeger, J.A., Turner, D.H. and Zuker, M. (1990). Predicting optimal and suboptimal secondary structure for RNA. In: *Molecular Evolution: Computer analysis of protein and nucleic acids sequences* (Methods in Enzymology, vol. 183). (Doolittle, R.F. ed.). Academic Press, San Diego, pp 281-306.
364. Jarvis, T.C. and Kirkegaard, K. (1992). Poliovirus RNA recombination: Mechanistic studies in the absence of selection. *EMBO Journal* **11**, 3135-3145.
365. Jensen, M.J. and Moore, D.M. (1993). Phenotypic and functional characterization of mouse attenuated and virulent variants of foot-and-mouth disease virus type O1 Campos. *Virology*. **193**, 604-613.
366. Jeppesen, P.G.N. (1974). A method for separating DNA fragments by electrophoresis in polyacrylamide concentration gradient slab gels. *Analytical Biochemistry*. **58**, 195-207.
367. Johnson, D.R., Cox, A.D., Solski, P.A., Devadas, B., Adams, S.P., Leimgruber, R.M., Heuckeroth, R.O., Buss, J.E. and Gordon, J.I. (1990). Functional analysis of protein N-myristoylation: metabolic labeling studies using three oxygen-substituted analogs of myristic acid and cultured mammalian cells provide evidence for protein-sequence-specific incorporation and analog-specific redistribution. *Proceedings of National Academy of Sciences USA*. **87**, 8511-8515.
368. Johnson, K.M., Wiebenga, N.H., Mackenzie, R.B., Juns, M.L., Tauraso, N.M., Shelokov, A., Webb, P.A., Justines, G. and Beye, H.K. (1965). Virus isolation from human cases of hemorrhagic fever in Bolivia. *Proceedings of Society in Experimental Biology and Medicine*. **118**, 113.
369. Johnson, K.M., Webb, P.A. and Justines, G. (1973). Biology of Tacaribe-complex viruses, In: *Lymphocytic choriomeningitis virus and other arenaviruses* (Lehmann Grube, F. Ed.). Springer Verlag, Berlin. pp: 241-258.
370. Johnson, M.s., May, A.C.W., Rodionov, M. and Overington, J.P. (1996). Discrimination of common protein folds: application of protein structure to sequence/structure comparisons. *Methods in Enzymology*. **266**, 575-598.
371. Jones, D.T., Taylor, W.R. and Thornton, J.M. (1992). A new approach to protein fold recognition. *Nature*. **358**, 86-89.
372. Jones, D.T., Taylor, W.R. and Thornton, J.M. (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry*. **33**, 3038-3049.
373. Jones, D.T., Miller, R.T. and Thornton, J.M. (1995). Successful protein fold recognition by optimal sequence threading validated by rigorous blind testing. *Proteins*. **23**, 387-397.
374. Justines, G. and Johnson, K.M. (1969). Immune tolerance in *Calomys callosus* infected with Machupo virus. *Nature*. **222**, 1090-1091.
375. Kabsch, N. and Sander, C. (1983). Dictionary of Secondary Structure of Proteins. *Biopolymers*. **22**, 2577-2637
376. Kamer, G. and Argos, G. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Research*. **12**, 7269.
377. Kandjian, E.W. (1986). UV crosslinking of RNA to nylon membranes enhances hybridization signals. *Molecular Biology Reports*. **11**, 107-115.
378. Karlin, S. and Altschul, S.F. (1990). Methods for Assessing the Statistical Significance of Molecular Sequence Features by Using General Scoring Schemes. *Proceedings of National Academy of Sciences USA*. **87**, 2264-2268.
379. Karlin, S., Blaisdell, B.E. and Brendel, V. (1990). Identification of significant sequence patterns in proteins. In: *Methods in enzymology*, vol. 183. (Doolittle, R.F. Ed). Academic Press, New York. pp. 388-402.
380. Karlin, S. and Altschul, S.F. (1993). Applications and statistics for multiple high-scoring segments in molecular sequences. *Proceedings of National Academy of Sciences USA*. **90**, 5873-5877.
381. Karlin, S. and Cardon, L.R. (1994). Computational DNA sequence analysis. *Annual Review in Microbiology*. **48**, 619-654.
382. Karlin, S., Mocarski, E.S. and Schachtel, G.A. (1994a). Molecular evolution of herpesvirus: genomic and protein sequence comparisons. *Journal of Virology*. **68**, 1886-1902.
383. Karlin, S., Doerfler, W. and Cardon, L.R. (1994b). Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses. *Journal of Virology*. **68**, 2889-2897.
384. Karlin, S., Ladunga, I. and Blaisdell, B.E. (1994c). Heterogeneity of genomes: measures and values. *Proceedings of National Academy of Sciences USA*. **91**, 12837-12841.
385. Karlin, S. and Ladunga, I. (1994). Comparisons of eukaryotic genomic sequences. *Proceedings of National Academy of Sciences USA*. **91**, 12832-12836.
386. Karlin, S. and Burge, C. (1996). Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proceedings of National Academy of Sciences USA*. **93**, 1560-1565.
387. Karlin, S., Mrázek, J. and Campbell, A.M. (1997). Compositional biases of bacterial genomes and evolutionary implications. *Journal of Bacteriology*. **179**, 3899-3913.
388. Karlin, S. and Mrázek, J. (1997). Compositional differences within and between eukaryotic genomes. *Proceedings of National Academy of Sciences USA*. **94**, 10227-10232.
389. Karlin, S., Brocchieri, L., Mrázek, J., Campbell, A.M. and Spormann, A.M. (1999). A chimeric prokaryotic ancestry of mitochondria and primitive eukaryotes. *Proceedings of National Academy of Sciences USA*. **96**, 9190-9195.
390. Karlovsky, P. and Prell, H.H. (1987). Screening of pUC plasmid clones for inserts based on growth rate (without X-gal). *Nucleic Acids Research*. **15**, 6753.
391. Kentsis, A., Campbell Dwyer, E.J., Perez, J.M., Sharma, M., Chen, A., Pan, Z.Q. and Borden, K.L.B. (2001). The ring domain of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E. *Journal of Molecular Biology*. **312**, 609-623.

Referencias bibliográficas

392. Kibler, K.V., Shors, T., Perkins, K.B., Zeman, C.C., Banaszak, M.P., Biesterfeldt, J., Langland, J.O. and Jacobs, B.L. (1997). Double-stranded RNA is a trigger for apoptosis in vaccinia virus-infected cells. *Journal of Virology*. **71**, 1992-2003.
393. Kiley, M.P., Tomori, O., Regnery, R.L. and Johnson, K.M. (1981). Characterization of the Arenaviruses Lassa and Mozambique. In: The replication of negative strand viruses. (Bishop, D.H.L. and Compans, R.W. Eds.) Elsevier Press, Nueva York. pp: 11-14.
394. Kiley, M.P., Wilusz, J., McCormick, J.B. and Keene, J.D. (1986). Conservation of the 3' terminal nucleotide sequences of Ebola and Marburg virus. *Virology*. **149**, 251-254.
395. King, P.V. and Blakesley, R.W. (1986). Optimizing DNA ligations for transformation. *Focus*. **8**, 1-3.
396. King, A.M. (1988). Preferred sites of recombination in poliovirus RNA: an analysis of 40 intertypic cross-over sequences. *Nucleic Acids Research*. **16**, 11705-11723.
397. King, R.D. and Sternberg, M.J. (1996). Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. *Protein Science*. **5**, 2298-2310.
398. King, R.D., Saqi, M., Sayle, R. and Sternberg, M.J. (1997). DSC: public domain protein secondary structure prediction. *CABIOS*. **13**, 473-474.
399. Kirkegaard, K. and Baltimore, D. (1986). The mechanism of RNA recombination in polio virus. *Cell*. **47**, 433-443.
400. Klaver, B. and Berkhout, B. (1994). Evolution of a disrupted TAR RNA hairpin structure in the HIV-1 virus. *EMBO Journal*. **13**, 2650-2659.
401. Klavinskis, L.S. and Oldstone, M.B.A. (1989). Lymphocytic choriomeningitis virus selectively alters differentiated but no housekeeping functions: block in the expression of growth hormone gene is at the level of transcription initiation. *Virology*. **168**, 232-235.
402. Klein, P., Kanehisa, M. and De Lisi, C. (1985). The detection and classification of membrane-spanning proteins. *Biochimie et Biophysic Acta*. **815**, 468-476.
403. Klovinis, J. Tsareva, N.A., de Smit, M.H., Berzins, V. and van Duin, J. (1997). Rapid evolution of translational control mechanisms in RNA genomes. *Journal of Molecular Biology*. **265**, 372-384.
404. Kocher, J.-P., Rooman, M.J. and Wodak, S.J. (1994). Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches. *Journal of Molecular Biology*. **235**, 1598-1613.
405. Kolakofsky, D. and Hacker, D. (1991). Bunyavirus RNA synthesis: genome transcription and replication. *Current Topics in Microbiology and Immunology*. **169**, 143-159.
406. Konopka, A.K. and Owens, J. (1990). Complexity charts can be used to map functional domains in DNA. *General Analytical Techniques and Applications*. **7**, 35-38.
407. Koonin, E.v., Tatusov, R.L. and Rudd, K.E. (1996). Protein sequence comparison at genome scale. *Methods in Enzymology*. **266**, 295-322.
408. Kornberg, A. and Baker, T.A. (1992). DNA Replication 2nd Edition. W.H. Freeman and company, pp. 194-196 y pp. 307-322.
409. Kozak, M. (1978). How do eucaryotic ribosomes select initiation regions in messenger RNA?. *Cell*. **15**, 1109-1123.
410. Kozak, M. (1983). Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbiological Reviews*. **47**, 1-45.
411. Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNAs. *Nucleic Acids Research*. **12**, 857-872.
412. Krickler, M.C., Drake, J.W. and Radman, M. (1992). Duplication-targeted DNA methylation and mutagenesis in the evolution of eukaryotic chromosomes. *Proceedings of National Academy of Sciences USA*. **89**, 1075-1079.
413. Krogh, A. and Mitchison, G. (1995). In Third International conference on Intelligent Systems for Molecular Biology (ISMB). (Rawlings, C., Clark, D., Altman, R., Hunter, L., Lengauer, T. and Wodak, S., eds.), pp. 215-221, AAAI Press, Cambridge, England
414. Krug, R.M. (1981). Priming of Influenza viral RNA transcription by capped heterologous RNAs. *Current Topics in Microbiology and Immunology*. **93**, 125-149.
415. Kuge, S., Saito, I. and Nomoto, A. (1986). Primary structure of poliovirus defective-interfering particle genomes and possible generation mechanism of the particles. *Journal of Molecular Biology*. **192**, 473-487.
416. Kunz, S., Sevilla, N., McGavern, D.B., Campbell, K.P. and Oldstone, M.B. (2001). Molecular analysis of the interaction of LCMV with is cellular receptor [alpha]-dystroglycan. *Journal of Cell Biology*. **155**, 301-310.
417. Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology*. **157**, 105-132.
418. Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**, 680-685.
419. Lai, M.M.C. (1992). RNA recombination in animal and plant viruses. *Microbiological Reviews*. **56**, 61-79.
420. Lai, M.M.C. (1996). Recombination in large RNA viruses: Coronaviruses. *Seminars in Virology*. **7**, 381-388.
421. Lake, J.A. (1993). Reconstructing evolutionary trees from DNA and protein sequences: paralinear distances. *Proceedings of the National Academy of Sciences USA*. **91**, 1455-1459.
422. Lamb, R.A. and Kolakofsky, D. (1995). Paramyxoviridae: the viruses and their replication. In: Fields Virology (3rd ed.). (Fields, B.N., Knipe, D.M. and Howley, P.M. eds.), chapter 40, Lippincott-Raven, New York.
423. Lamb, R.A. and Krug, R.M. (1995). Orthomyxoviridae: the viruses and their replication. In: Fields Virology (3rd ed.). (Fields, B.N., Knipe, D.M. and Howley, P.M. eds.), chapter 45, Lippincott-Raven, New York.
424. Lambden, P.R., Cooke, S.J., Caul, E.O. and Clarke, I.N. (1992). Cloning of noncultivable human rotavirus by single primer amplification. *Journal of Virology*. **66**, 1817-1822.
425. Lascano, E.F. and Berria, M.I. (1970). Microscopía electrónica de cultivos primarios de fibroblastos de ratón inoculados con virus Junin. *Medicina (Buenos Aires)*. **29**, 487-493.
426. Lascano, E.F. and Berria, M.I. (1974). Ultrastructure of Junin virus in mouse whole brain and mouse brain tissue culture. *Journal of Virology*. **14**, 965-974.
427. Lascano, E.F., Berria, M.I. and Candurra, N.A. (1981). Diagnosis of Junin virus in cell cultures by immunoperoxidase staining. *Archives in Virology*. **70**, 79-82.
428. Laskey, R.A. (1980). The use of intensifying screens or organic scintillators for visualizing radioactive molecules resolved by gel electrophoresis. *Methods in Enzymology*. **65**, 363-365.
429. Lawrence, C.E., Altschul, S.F., Boguski, M.S., Liu, J.S., Neuwald, A.F. and Wootton, J.C. (1993). Detecting Subtle Sequence Signals: A Gibbs Sampling Strategy for Multiple Alignment. *Science*. **262**, 208-214.

430. Lazzarini, R.A., Keene, J.D. and Schubert, M. (1981). The origins of defective interfering particles of negative-strand RNA viruses. *Cell*. **26**, 145-154.
431. Leahy, M.B., Dessens, J.T. and Nuttall, P.A. (1997). *In vitro* polymerase activity of Thogoto virus: evidence for a unique cap-snatching mechanism in a tick-borne Orthomyxovirus. *Journal of Virology*. **71**, 8347-8351.
432. Lee, K.J., Novella, I.S., Teng, M.N., Oldstone, M.B.A. and de la Torre, J.C. (2000). NP and L proteins of Lymphocytic Choriomeningitis virus (LCMV) are sufficient for efficient transcription and replication of LCMV genomic RNA analogs. *Journal of Virology*. **74**, 3470-3477.
433. Lehmann Grube, F., Martínez Peralta, L., Bruns, M. and Löhler, J. (1983). Persistent infection of mice with the lymphocytic choriomeningitis virus. In: *Comprehensive Virology* (Fraenkel Conrat, H. and Wagner, R.R. Eds.). Plenum Press, Nueva York. Vol. **18**, pp: 43-103.
434. Lehrach, H., Diamond, D., Wozney, J.M. and Boedtker, H. (1977). RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry*. **16**, 4743-4751.
435. Lemer, C.M.-R., Rooman, M.J. and Wodak, S.J. (1995). Protein structure prediction by threading methods: evaluation of current techniques. *Proteins*. **23**, 337-355.
436. Lentz, T.L. (1990). The recognition event between virus and host cell receptor: a target for antiviral agents. *Journal of General Virology*. **71**, 751-766.
437. Lenz, O., ter Meulen, J., Feldmann, H., Klenk, H-D. and Garten, W. (2000). Identification of a novel consensus sequence at the cleavage site of the Lassa virus glycoprotein. *Journal of Virology*. **74**, 11418-11421.
438. Lenz, O., ter Meulen, J., Klenk, H-D., Seidah, N.G. and Garten, W. (2001). The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proceedings of National Academy of Sciences USA*. **98**, 12701-12705.
439. Lepine, P., Mollaret, P. and Kreis, B. (1937). Receptivité de l'homme au virus murin de la choriomeningite lymphocytaire. Reproduction expérimentale de la meningite lymphocytaire benigne. *C.R. Academy of Science Paris*. **204**, 1846-1848.
440. Lesk, A.M. (1988). *Computational Molecular Biology: Sources and Methods for Sequence Analysis*. Oxford University Press, Oxford.
441. Leung, W.C. and Rawls, W.E. (1977). Virion-associated ribosomes are not required for the replication of Pichinde virus. *Virology*. **81**, 174-176.
442. Leung, W.C., Leung, M.F.K.L. and Rawls, W.E. (1979). Distinctive RNA transcriptase, polyadenylic acid polymerase and polyuridylic acid polymerase activities associated with Pichinde virus. *Journal of Virology*. **30**, 98-107.
443. Leung, W.C., Harnish, D., Ramsingh, A., Dimock, K. and Rawls, W.E. (1981). Gene mapping in Pichinde virus. In: *The replication of negative strand viruses*. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, Nueva York. pp: 51-57.
444. Levis, R., Weiss, B.G., Tsiang, M., Huang, H. and Schlesinger, S. (1986). Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential replication and packaging. *Cell*. **44**, 137-145.
445. Levis, S.C., Saavedra, M.C., Ceccoli, C., Falcoff E., Feuillade M.R., Enria D.A., Maiztegui J.I. and Falcoff R. (1984). Endogenous interferon in Argentine hemorrhagic fever. *Journal of Infectious Diseases*. **149**, 428-433.
446. Levis, S.C., Saavedra, M.C., Ceccoli, C., Feuillade, M.R., Enria, D.A., Maiztegui, J.I. and Falcoff, R. (1985). Correlation between endogenous interferon and the clinical evolution of patients with Argentine hemorrhagic fever. *Journal of Interferon Research*. **5**, 383-389.
447. Levitt, M. and Warshel, A. (1975). Computer simulation of protein folding. *Nature*. **253**, 694-698.
448. Levy-Mintz, P. and Kielian, M. (1991). Mutagenesis of the putative fusion domain of the Semliki Forest Virus spike protein. *Journal of Virology*. **65**, 4292-4300.
449. Lewin, B. (1997). Chapter 31: Catalytic RNA. In *Genes VI* (Benjamin Lewin Ed.). Oxford University Press, New York.
450. Li, Y.L. and Ball, L.A. (1993). Nonhomologous RNA recombination during negative-strand synthesis of Flock House virus RNA. *Journal of Virology*. **67**, 3854-3860.
451. Li, M.-L., Ramirez, B.C. and Krug, R.M. (1998). RNA-dependent activation of primer RNA production by influenza virus polymerase: different regions of the same protein subunit constitute the two required RNA-binding sites. *The EMBO Journal*. **17**, 5844-5852.
452. Li, X.-L., Blackford, J.A. and Hassel, B.A. (1998). RNase L mediates the antiviral effect of interferon through a selective reduction in viral RNA during encephalomyocarditis virus infection. *Journal of Virology*. **72**, 2752-2759.
453. Livingstone, C.D. and Barton, G.J. (1993). Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *CABIOS*. **9**, 745-756.
454. Livingstone, C.D. and Barton, G.J. (1994). Secondary Structure Prediction from Multiple Sequence Data: Blood Clotting Factor XIII and *Yersinia* Protein-Tyrosine Phosphatase. *International Journal of Peptide and Protein Research*. **44**, 239-244.
455. Löflering, B., Mühlberger, E., Tamura, T., Klenk, H-D. and Becker, S. (1999). The nucleoprotein of Marburg virus is target for multiple cellular kinases. *Virology*. **255**, 50-62.
456. Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L. and Davis, M.M. (1989). Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science*. **243**, 217-220.
457. López, R.A. and Franze Fernández, M. T. (1985). Effect of Tacaribe virus infection on host cell protein and nucleic acid synthesis. *Journal of General Virology*. **66**, 1753-1761.
458. López, R.A. (1986). *Arenavirus: su interacción con la célula hospedadora*. Tesis Doctoral. Facultad de Ciencias Exactas. Universidad Nacional de La Plata.
459. López, R.A., Grau, O. And Franze Fernández, M.T. (1986). Effect of actinomycin D on arenavirus growth and estimation of the generation time for a virus particle. *Virus Research*. **5**, 213-220.
460. López, N., Jácamo, R. and Franze-Fernández, M.T. (2001). Transcription and RNA replication of Tacaribe virus genome and antigenome analogs require N and L proteins: Z protein is an inhibitor of these processes. *Journal of Virology*. **75**, 12241-12251.
461. Lovering, R., Hanson, I.M., Borden, K.L., Martin, S., O'Reilly, N.J., Evan, G.I., Rahman, D., Pappin, D.J., Trowsdale, J., Freemont, P.S. (1993). Identification and preliminary characterization of a protein motif related to the zinc finger. *Proceedings of National Academy of Sciences USA*. **90**, 2112-2116. (ACCESSION A47380).
462. Lozano, M.E., Ghiringhelli, P.D., Romanowski, V. and Grau, O. (1993). A simple nucleic acids amplification assay for the rapid detection of Junin virus in whole blood samples. *Virus Research*. **27**, 37-53.
463. Lozano, M.E. (1995). *Desarrollo de un método de diagnóstico rápido para la fiebre hemorrágica argentina*. Tesis Doctoral. Facultad de Ciencias Exactas. Univ. Nacional de La Plata.

Referencias bibliográficas

464. Lucia, H.L., Coppenhaver, D.H. and Baron, S. (1989). Arenavirus infection in the guinea pig model: antiviral therapy with recombinant interferon-alpha, the immunomodulator CL246,738 and ribavirin. *Antiviral Research*. **12**, 279-292.
465. Lukashevich, I.S., Stelmakh, T.A., Golubev, V.P., Stchesljenok, E.P. and Lemeshko, N.N. (1984). Ribonucleic acids of Machupo and Lassa viruses. *Archives in Virology*. **79**, 189-203.
466. Lupas, A. (1996). Coiled coils: new structures and new functions. *TIBS*. **21**, 375-382.
467. Lüthy, R. and Eisenberg, D. (1991). Proteins. In: Sequence analysis primer. (Gribbskov, M. and Devereux, J. Eds). Stockton Press, New York. pp. 61-87.
468. Luytjes, W., Bredendbeek, P.J., Noten, A.F.H., Horzinek, M.C. and Spaan, W.J.M. (1988). Sequence of Mouse Hepatitis Virus A59 mRNA 2: indications for RNA recombination between Coronaviruses and Influenza C virus. *Virology*. **166**, 415-422.
469. Mackenzie, R.B., Beye, H.K., Valverde, C.L. and Garrón, H. (1964). Epidemic hemorrhagic fever in Bolivia. I. A preliminary report of the epidemiologic and clinical findings in a new epidemic area in South America. *American Journal of Tropical Medicine and Hygiene* **13**, 620.
470. Magee, A.I., Gutierrez, L., Marshall, C.J. and Hancock, J.F. (1989). Targeting of oncoproteins to membranes by fatty acylation. *Journal of Cell Science*. Suppl. **11**, 149-160.
471. Maiztegui, J.I., Sabattini, M.S. and Barrera Oro, J.G. (1972). Actividad del virus de la coriomeningitis linfocitaria (LCM) en el área endémica de Fiebre Hemorrágica Argentina (FHA). I. Estudios serológicos en roedores capturados en la ciudad de Pergamino. *Medicina (Buenos Aires)*. **32**, 131-137.
472. Maiztegui, J.I., Voëffrey, J.R., Fernández, N.J. and Barrera Oro, J.G. (1973). Aislamiento de virus Junín a partir de leche materna. *Medicina (Buenos Aires)*. **33**, 659.
473. Maiztegui, J.I. (1975). Clinical and epidemiological patterns of Argentine hemorrhagic fever. *Bulletin WHO*. **52**, 3385-575.
474. Maiztegui, J.I., Laguens, R.P., Cossio, P.M., Casanova, M.B., de la Vega, M.T., Ritacco, V., Segal, A., Fernández, N.J. and Arana, R.M. (1975). Ultrastructural and immuno-histochemical studies in five cases of Argentine hemorrhagic fever. *Journal of Infectious Disease*. **132**, 35-43.
475. Maiztegui, J.I. and Sabattini, M.S. (1977). Extensión progresiva del área endémica de fiebre hemorrágica argentina. *Medicina (Buenos Aires)*. **37** Supl. 3, 162-166.
476. Maiztegui, J.I., Fernández, N.J. and de Damilano, A.J. (1979). Efficacy of immune plasma in treatment of Argentine hemorrhagic fever and association between treatment and a late neurological syndrome. *Lancet*. **2**, 1216-1217.
477. Maiztegui, J.I., Feuillade, M. and Briggiler, A. (1986). Progressive extension of the endemic area and changing incidence of AHF. *Medical Microbiology and Immunology*. **175**, 149-152.
478. Maiztegui, J.I., Feinsod, F., Briggiler, A.M., Peters, C.J., Enría, D.A., Lupton, H.W., Ambrosio, A.M., Tiano, E., Feuillade, M.R., Gamboa, G., Conti, O., Vallejos, D., Mac Donald, C., and Barrera Oro, J.G. (1987). Inoculation of Argentine volunteers with a live-attenuated Junín virus vaccine. VII International Congress of Virology, Edmonton, Canada (Abstract book, p. 69, R.3.49).
479. Maiztegui, J.I., McKee, K.T. Jr., Barrera Oro, J.G., Harrison, L.H., Gibbs, P.H., Feuillade, M.R., Enría, D.A., Briggiler, A.M., Levis, S.C., Ambrosio, A.M., Halsey, N.A. and Peters, C.J. (1998). Protective of a live attenuated vaccine against Argentine hemorrhagic fever. AHF study group. *Journal of Infectious Diseases*. **177**, 277-283.
480. Makino, S., Shieh, C.-K., Soe, L.H., Baker, S.C. and Lai, M.M.C. (1988). Primary structure and translation of defective interfering RNA of murine coronavirus. *Virology*. **166**, 550-560.
481. Mando, O.G. (1977). Hepatic and enzymatic alterations in Argentine hemorrhagic fever. *Medicina (Buenos Aires)*. **37**, 190-192.
482. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual, 1st Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
483. Mannweiler, K. and Lehmann Grube, F. (1973). Electron microscopy of LCM virus-infected L cells. In: Lymphocytic choriomeningitis virus and other arenaviruses (Lehmann Grube, F. Ed.). Springer Verlag, Berlin. pp: 37-48.
484. Marión, R.M., Zürcher, T., de la Luna, S. and Ortín, J. (1997). Influenza virus NS1 protein interacts with viral transcription-replication complexes *in vivo*. *Journal of General Virology*. **78**, 2447-2451.
485. Marsh, M. and Helenius, A. (1989). Virus entry into animal cells. *Advances in Virus Research*. **36**, 107-151.
486. Martin, D. and Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics*. **16**, 562-563.
487. Martinand, C., Montavon, C., Salehzada, T., Silhol, M., Lebleu, B. and Bisbal, C. (1999). RNase L inhibitor is induced during human immunodeficiency virus type I infection and down regulates the 2-5A/RNase L pathway in human T cells. *Journal of Virology*. **73**, 290-296.
488. Martínez Peralta, L., Bruns, M. and Lehmann Grube, F. (1981). Biochemical composition of lymphocytic choriomeningitis virus interfering particles. *Journal of General Virology*. **55**, 475.
489. Martínez Segovia, Z.M., Holstein, B.A. and Grazioli, F. (1967). Multiplicación del virus Junín en cultivo de tejidos. *Ciencia e Investigación* **23**, 35-39.
490. Martínez Segovia, Z.M. and Díaz, A. (1968). Purification of Junín virus by an aqueous biphasic polymer system. *Applied Microbiology*. **15**, 1602-1604.
491. Martínez Segovia, Z.M. and Grazioli, F. (1969). The nucleic acid of Junín virus. *Acta Virologica*. **13**, 264-268.
492. Martínez Segovia, Z.M. and De Mitri, M.I. (1977). Junín virus structural proteins. *Journal of Virology*. **21**, 579-583.
493. Martoja, R. and Martoja, M. (1970). Técnicas de histología animal. Ed. Toray-Mason S.A., Barcelona, España.
494. Matloubian, M., Somasundaram, T., Kolhekar, S.P., Selvakumar, R. and Ahmed, R. (1990). Genetic basis of viral persistence: single amino acid change in the viral glycoprotein affects ability of Lymphocytic Choriomeningitis virus to persist in adult mice. *Journal of Experimental Medicine*. **172**, 1043-1048.
495. Matloubian, M., Kolhekar, S.P., Somasundaram, T. and Ahmed, R. (1993). Molecular determinants of macrophage tropism and viral persistence: importance of single amino acid changes in the polymerase and glycoprotein of Lymphocytic Choriomeningitis virus. *Journal of Virology*. **67**, 7340-7349.
496. Maxam, A.M. and Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of National Academy of Sciences USA* **74**, 560-582.
497. Maxfield, F.R. and Scheraga, H.A. (1979). Improvements in the Prediction of Protein Topography by Reduction of Statistical Errors. *Biochemistry*. **18**, 697-704.

498. May, A.C.W. and Blundell, T.L. (1994). Automated comparative modelling of protein structures. *Current Opinion in Biotechnology*. **5**, 355-360.
499. McCabe, J.B. and Berthiaume, L.G. (1999). Functional roles for fatty acylated amino-terminal domains in subcellular localization. *Molecular Biology of the Cell*. **10**, 3771-3786.
500. McCaldon, P. and Argos, O. (1988). Oligopeptide biases in protein sequences. *Proteins*. **4**, 99-122.
501. McClure, M.A. and Perrault, J. (1986). RNA virus genomes hybridize to cellular rRNAs and to each other. *Journal of Virology*. **57**, 917-921.
502. McCormick, J.B. (1990). Arenaviruses. In: *Virology*, second edition. (Fields, B.N., Knipe, D.M. *et al.* Eds). Raven Press, Nueva York. pp: 1245-1267.
503. McKee K.T. Jr, Barrera Oro, J.G., Kuehne, A.I., Spisso, J., and Mahlandt, B.G. (1984). Immunogenicity and protective efficacy of a live, attenuated Argentine hemorrhagic fever (AHF) vaccine in primates. XI International Congress for Tropical Medicine and Malaria, Calgary, Canada. Abstract book, p. 29.
504. McKee, K.T., Jr., Mahlandt, B.G., Maiztegui, J.I., Green, D.E., Eddy, G.A. and Peters, C.J. (1985). Experimental Argentine hemorrhagic fever in rhesus macaques: virus strain-dependent clinical response. *Journal of Infectious Disease*. **152**, 218-221.
505. Meegan, J., LeDue, J., García Franco, S. and Maiztegui, J.I. (1986). An ELISA test for IgG and IgM antibodies to Junin virus. II Congreso Argentino de Virología. Córdoba. Libro de resúmenes, comunicación 161.
506. Mehta, P.K., Heringa, J. and Argos, P. (1995). A simple and fast approach to prediction of protein secondary structure from multiply aligned sequences with accuracy above 70%. *Protein Science*. **4**, 2517-2525.
507. Meier, E., Harmison, G.G., Keene, J.D. and Schubert, M. (1984). Sites of copy choice replication involved in generation of vesicular stomatitis virus defective-interfering particle RNAs. *Journal of Virology*. **51**, 515-521.
508. Melcon, M.O. and Herskovits, E. (1981). Late neurological complications of Argentine hemorrhagic fever. *Medicina (Buenos Aires)*. **41**, 137.
509. Messing J. (1983). New M13 vectors for cloning. *Methods in Enzymology*. **101**, 20-78.
510. Messing J. (1993). M13 cloning vehicles. Their contribution to DNA sequencing. *Methods in Molecular Biology*. **23**, 9-22.
511. Mettler, N.E., Casals, J. and Shope, R.E. (1963). Study of antigenic relationship between Junin virus, the etiological agent of Argentine hemorrhagic fever, and other arthropode borne viruses. *American Journal of Tropical Medicine and Hygiene*. **12**, 647-652.
512. Meyer, B.J. and Southern, P.J. (1993). Concurrent sequence analysis of 5' and 3' RNA termini by intramolecular circularization reveals 5' nontemplated bases and 3' terminal heterogeneity for Lymphocytic Choriomeningitis virus mRNAs. *Journal of Virology*. **67**, 2621-2627.
513. Meyer, B.J. and Southern, P.J. (1994). Sequence heterogeneity in the termini of Lymphocytic Choriomeningitis virus genomic and antigenomic RNAs. *Journal of Virology*. **68**, 7659-7664.
514. Meyer, B.J. and Southern, P.J. (1997). A novel type of defective viral genome suggests a unique strategy to establish and maintain persistent Lymphocytic Choriomeningitis virus infections. *Journal of Virology*. **71**, 6757-6764.
515. Meyer, B.J. and Schmaljohn, C. (2000). Accumulation of terminally deleted RNAs may play a role in Seoul virus persistence. *Journal of Virology*. **74**, 1321-1331.
516. Miller, W.A. and Koev, G. (1998). Getting a handle on RNA virus recombination. *Trends in Microbiology*. **6**, 521-523.
517. Mills, J.N., Ellis, B.A., McKee Jr., K.T., Ksiazek, T.G., Barrera Oro, J.G., Maiztegui, J.I., Calderón, G.E. Peters, C.J. and Childs, J.E. (1991). Junin virus activity in rodents from endemic and nonendemic loci in central Argentina. *American Journal of Tropical Medicine and Hygiene*. **44**, 589-597.
518. Miralles, R., Moya, A. and Elena, S.F. (2000). Diminishing returns of population size in the rate of RNA virus adaptation. *Journal of Virology*. **74**, 3566-3571.
519. Mitra, A., Higgins, D.W., Langenberg, W.G., Nie, H., Sengupta, D.N. and Silverman, R.H. (1996). A mammalian 2-5A system functions as an antiviral pathway in transgenic plants. *Proceedings of National Academy of Sciences USA*. **93**, 6780-6785.
520. Molenkamp, R., Greve, S., Spaan, W.J.M. and Snijder, E.J. (2000). Efficient homologous RNA recombination and requirement for an open reading frame during replication of Equine Arteritis virus defective interfering RNAs. *Journal of Virology*. **74**, 9062-9070.
521. Molinas, F.C. and Maiztegui, J.I. (1981). Factor VIII:C and factor VIII R:Ag in Argentine hemorrhagic fever. *Thrombosis and Haemostasis*. **46**, 525.
522. Monath, T.P., Arroyo, J., Levenbook, I., Zhang, Z-X., Catalan, J., Draper, K. And Guirakhoo, F. (2002). Single mutation in the flavivirus envelope protein hinge region increases neurovirulence for mice and monkeys but decreases viscerotropism for monkeys: relevance to development and safety testing of live, attenuate vaccines. *Journal of Virology*. **76**, 1932-1943.
523. Moncayo, A.C., Hice, C.L., Watts, D.M., Travassos da Rosa, A.P.A., Guzman, H., Russell, K.L., Calampa, C., Gozalo, A., Popov, V.L., Weaver, S.C. and Tesh, R.B. (2001). Allpahuayo virus: a newly recognized arenavirus (Arenaviridae) from arboreal rice rats (*Oecomys bicolor* and *Oecomys paricola*) in Northeastern Peru. *Virology*. **284**, 277-286.
524. Monge, A., Friesner, R.A. and Honig, B. (1994). An algorithm to generate low-resolution protein tertiary structures from knowledge of secondary structure. *Proceedings of National Academy of Sciences USA*. **91**, 5027-5029.
525. Moody, M.D., Burg, J.L., DiFrancesco, R., Lovern, D., Stanik, W., Lin-Goerke, J., Mahdavi, K., Wu, Y. and Farrel, M.P. (1994). Evolution of host cellRNA into efficient template RNA by Q β replicase: the origin of RNA in untemplated reactions. *Biochemistry*. **33**, 13836-13847.
526. Morse, S.S. (1993). Examining the origins of emerging viruses. In: *Emerging viruses*. (S. S. Morse, ed.). Oxford University Press, New York. pp:10-28.
527. Mosimann, S., Meleshko, r. and James, M.N.G. (1995). A critical assessment of comparative molecular modeling of tertiary structures of proteins. *Proteins*. **23**, 301-317.
528. Moulton, J., Pedersen, J.T., Judson, R. and Fidelis, K. (1995). A large-scale experiment to assess protein structure prediction methods. *Proteins*. **23**, ii-iv.
529. Mrázek, J. and Karlin, S. (1998). Strand compositional asymmetry in bacterial and large viral genomes. *Proceedings of National Academy of Sciences USA*. **95**, 3720-3725.
530. Mumenthaler, Ch. and Braun, W. (1995). Predicting the helix packing of globular proteins by self-correcting distance geometry. *Protein Science*. **4**, 863-871.

531. Mundry, K.W., Watkins, P.A.C., Ashfield, T., Plaskitt, K.A., Eisele-Walter, S. and Wilson, M.A. (1991). Complete uncoating of the 5' leader sequence of tobacco mosaic virus RNA occurs rapidly and is required to initiate cotranslational virus disassembly *in vitro*. *Journal of General Virology*. **72**, 769-777.
532. Munishkin, I.Y., Voronin, L.A., and Chetverin, A.B. (1988). An *in vivo* recombinant RNA capable of autocatalytic synthesis by Q β replicase. *Nature*. **333**, 473-475.
533. Munishkin, I.Y., Voronin, L.A., Ugarov, V.I., Bondareva, L.A., Chetverina, H.V. and Chetverin, A.B. (1991). Efficient templates for Q β replicase are formed by recombination from heterologous sequences. *Journal of Molecular Biology*. **221**, 463-472.
534. Murphy, F.A., Webb, P.A., Johnson, K.M. and Whitfield, S.G. (1969). Morphological comparison of Machupo with lymphocytic choriomeningitis virus: basis for a new taxonomic group. *Journal of Virology*. **4**, 535-541.
535. Murphy, F.A., Webb, P.A., Johnson, K.M., Whitfield, S.G. and Chappel, W.A. (1970). Arenaviruses in Vero cells: ultrastructural studies. *Journal of Virology*. **6**, 507-518.
536. Murphy, F.A., Whitfield, S.G., Webb, P.A. and Johnson, K.M. (1973). Ultrastructural studies of arenaviruses. In: Lymphocytic choriomeningitis virus and other arenaviruses (Lehmann Grube, F. Ed.). Springer Verlag, Berlin. pp: 273-285.
537. Murphy, F.A. and Walker, D.H. (1975). Arenaviruses: Persistent infection and viral survival in reservoir host. In: Viruses and Environment. (E. Kurstat and K. Maramorosch, eds.). Academic Press, New York. pp.: 155-180.
538. Murphy, F.A. and Whitfield, S.G. (1975). Morphology and morphogenesis of arena-viruses. *Bulletin WHO*. **52**, 409-419.
539. Myers, G., R umenapf, T. and Thiel, H.-J. (1990). Insertion of ubiquitin-coding sequence identified in the RNA genome of a togavirus. In: New aspects of positive-stranded RNA viruses. (M.A. Brinton and F.X. Heinz, eds.). American Society for Microbiology, Washington. pp.:25-30.
540. Nagy, P.D. and Bujarsky, J.J. (1993). Targeting the site of RNA-RNA recombination in brome mosaic virus with antisense sequences. *Proceedings of National Academy of Sciences USA*. **90**, 6390-6394.
541. Nagy, P.D. and Bujarsky, J.J. (1995). Efficient system of homologous RNA recombination in Brome Mosaic virus: sequence and structure requirements and accuracy of crossovers. *Journal of Virology*. **69**, 131-140.
542. Nagy, P.D. and Bujarsky, J.J. (1996). Homologous RNA recombination in Brome Mosaic Virus: AU-rich sequences decrease the accuracy of crossovers. *Journal of Virology*. **70**, 415-426.
543. Nagy, P.D. and Simon, A.E. (1997). New insights into the mechanism of RNA recombination. *Virology*. **235**, 1-9.
544. Nagy, P.D. and Bujarsky, J.J. (1997). Engineering of homologous recombination hot spots with AU-rich sequences in Brome Mosaic Virus. *Journal of Virology*. **71**, 3799-3810.
545. Nagy, P.D. and Bujarsky, J.J. (1998). Silencing homologous RNA recombination hot spot with GC-rich sequences in Brome Mosaic Virus. *Journal of Virology*. **72**, 1122-1130.
546. Nagy, P.D., Zhang, C. and Simon, A.E. (1998). Dissecting RNA recombination *in vitro*: role of RNA sequences and the viral replicase. *EMBO Journal*. **17**, 2392-2403.
547. Nagy, P.D., Pogany, J. and Simon, A.E. (1999a). RNA elements required for RNA recombination function as replication enhancers *in vitro* and *in vivo* in a plus-strand RNA virus. *EMBO Journal*. **18**, 5653-5665.
548. Nagy, P.D., Ogiela, C. and Bujarsky, J.J. (1999b). Mapping sequences active in homologous RNA recombination in Brome Mosaic virus: prediction of recombination hot spots. *Virology*. **254**, 92-104.
549. Nakisa, R.C. (1993). DotPlot, a program for graphical comparison of nucleic acid and protein sequences for IBM PC. Published electronically on the Internet and available by anonymous ftp from ftp.bio.indiana.edu.
550. Negroni, M. and Bue, H. (1999). Recombination during reverse transcription: an evaluation of the role of the nucleocapsid protein. *Journal of Molecular Biology*. **286**, 15-31.
551. Neuwald, A.F., Liu, J.S. and Lawrence, C.E. (1995). Gibbs motif sampling: detection of bacterial outer membrane protein repeats. *Protein Science*. **4**, 1618-1631.
552. Nguyen, M., Ramirez, B-C., Goldbach, R. And Haenni, A-L. (1997). Characterization of the *in vitro* activity of the RNA-dependent RNA polymerase associated with the ribonucleoproteins of Rice Hoja Blanca Tenuivirus. *Journal of Virology*. **71**, 2621-2627.
553. Nielsen, H., Engelbrecht, J., von Heijne, G. and Brunak, S. (1996). Defining a similarity threshold for a functional protein sequence pattern: the signal peptide cleavage site. *Proteins*. **24**, 165-177.
554. Nilges, M. and Br unger, A. T. (1991). Automated modelling of coiled coils. Application to the GCN4 dimerization region. *Protein Engineering*. **4**, 649-659.
555. Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *Journal of Molecular Biology*. **245**, 645-660.
556. Nilges, M. (1996). Structure calculation from NMR data. *Current Opinion in Structural Biology*. **6**, 617-623.
557. Notredame, C. and Higgins, D.G. (1996). SAGA: sequence alignment by genetic algorithm. *Nucleic Acids Research*. **24**, 1515-1524.
558. Novotny, J., Rashin, A.A. and Bruccoleri, R.E. (1988). Criteria that discriminate between native proteins and incorrectly folded models. *Proteins*. **4**, 19-30.
559. Odgren, P.R., Harvie, L.W.J. and Fey, E.G. (1996). Phylogenetic occurrence of coiled coil proteins: implications for tissue structure in metazoa via a coiled coil tissue matrix. *Proteins*. **24**, 467-484.
560. O'Donoghue, S. I. and Nilges, M. (1997). Tertiary structure prediction using mean-force potentials and internal energy functions: successful prediction for coiled-coil geometries. *Folding & Design*. **2**, S47-52.
561. Ofodile, A., Padnos, M., Molomut, N. and Duffy, J.L. (1973). Morphological and biological characteristics of the M-P strain of lymphocytic choriomeningitis virus. *Infect. Immun*. **7**, 309-315.
562. Ohara, O., Dorit, R.I. and Gilbert, W. (1989). One-sided polymerase chain reaction: the amplification of cDNA. *Proceedings of the National Academy of Sciences USA*. **86**, 5673-5677.
563. Ohno, S. (1988). Universal rule for coding sequence construction: TA/CG deficiency-TG/CA excess. *Proceedings of the National Academy of Sciences USA*. **85**, 9630-9634.
564. Oldstone, M.B.A., Holmstoen, J. and Welsh, R.M., Jr (1977). Alteration in acetylcholine enzymes in neuroblastoma cells persistently infected with lymphocytic choriomeningitis virus. *Journal of Cellular Physiology*. **91**, 459-472.
565. Oldstone, M.B.A., Buchmeier, M.J. and Doyle, M.V. (1978). Immunobiology of persistent virus infection. In: Persistent viruses. Academic Press, Inc., New York. pp. 43-56.

Referencias bibliográficas

566. Oldstone, M.B.A. and Buchmeier, M.J. (1982). Restricted expression of viral glycoproteins in cells of persistently infected mice. *Nature (Londres)*. **300**, 360-362.
567. Oldstone, M.B.A. (1987a). Arenaviruses: biology and immunotherapy (edit). *Current Topics in Microbiology and Immunology*. **134**, 1.
568. Oldstone, M.B.A. (1987b). Immunotherapy for virus infection. *Current Topics in Microbiology and Immunology*. **134**, 211-229.
569. Oldstone, M.B.A., Salvato, M., Tishon, A. and Lewicki, H. (1988). Virus-Lymphocytic interactions. III. Biologic parameters of a virus variant that fails to generate CTL and establishes persistent infection in immunocompetent hosts. *Virology*. **164**, 507-516.
570. Oliver, J.L., Román-Roldán, R., Pérez, J. and Bernal-Galván, P. (1999). SEGMENT: identifying compositional domains in DNA sequences. *Bioinformatics*. **15**, 974-979.
571. Olsthoorn, R.C.L. and van Duin, J. (1996). Evolutionary reconstruction of a hairpin deleted from the genome of an RNA virus. *Proceedings of National Academy of Sciences USA*. **93**, 12256-12261.
572. Orlova, M., Newlands, J., Das, A., Goldfarb, A. and Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. *Proceedings of National Academy of Sciences USA*. **92**, 4596-4600.
573. O'Reilly, D. R.; Miller, L. K. & Luckow, V. A. (1992). *Baculovirus expression vectors: A laboratory manual*. New York, W.H. Freeman and Company.
574. Ou, J.H., Strauss, E.G. and Strauss, J.H. (1981). Comparative studies of the 3'-terminal sequences of several alpha virus RNAs. *Virology*. **109**, 281-289.
575. Oubiña, J.R., Alison, L. and Howard, C.R. (1990). ARN polimerasa ARN dependiente del virus Pichinde: detección de la actividad enzimática y síntesis *in vitro* de ARN viral. IIIer Congreso Argentino de Virología. Resumen 7.
576. Ouzounis, C., Sander, C., Scharf, M. and Schneider, R. (1993). Prediction of protein structure by evaluation of sequence-structure fitness: Aligning sequences to contact profiles derived from 3D structures. *Journal of Molecular Biology*. **232**, 805-825.
577. Pabo, C. and Sauer, R. (1992). Transcription factors: structural families and principles of DNA recognition. *Annual Review in Biochemistry*. **61**, 1053-1095.
578. Padula, P.J., Miguel, S. and De Mitri, M.I. (1996). Relación de las glicoproteínas estructurales del virus Junín con el gen de su precursor. V Congreso Argentino de Virología – II Encuentro de Virólogos Latinoamericanos. Tandil, 24-27 de abril de 1996.
579. Palmer, E.L., Obijeski, J.F., Webb, P.A. and Johnson, K.M. (1977). The circular segmented nucleocapsid of an arenavirus-Tacaribe virus. *Journal of General Virology*. **36**, 541-545.
580. Parisi, G., Echave, J., Ghiringhelli, P.D. and Romanowski, V. (1996). Computational characterization of potential RNA-binding sites in arenavirus nucleocapsid proteins. *Virus Genes*. **13**, 249-256.
581. Parodi, A.S., Greenway, D.J., Rugiero, H.R., Rivero, E., Frigerio, M.J., Mettler, N.E., Garzon, F., Boxaca, M., Guerrero L.B. de. and Nota, N.R. (1958). Sobre la etiología del brote epidémico en Junín. *Día Médico*. **30**, 2300-2302.
582. Parodi, A.S., de la Barrera, J.M., Rugiero, H.R., Greenway, D.J., Yerga, M., Mettler, N.E., Boxaca, M. and Frigerio, M.J. (1959a). Los reservorios del virus de la fiebre hemorrágica epidémica de la provincia de Buenos Aires. *Prensa Médica Argentina*. **46**, 554-556.
583. Parodi, A.S., Rugiero, H.R., Greenway, D.J., Mettler, N.E., Martínez, A., Boxaca, M. and de la Barrera, J.M. (1959b). Aislamiento del virus Junín (FHE) de los ácaros de las zonas epidémicas (*Echinolaeps echidninus* Berlese). *Prensa Médica Argentina*. **46**, 2242-2244.
584. Parodi, A.S., Rugiero, H.R., Greenway, D.J., Mettler, N.E. and Boxaca, M. (1961). El aislamiento del virus Junín en roedores de zonas no epidémicas. *Prensa Médica Argentina*. **48**, 2321.
585. Patterson, J.L., Cabradilla, C., Holloway, B.P., Obijeski, J.F. and Kolakofsky, D. (1983). Multiple leader RNAs and messenger RNAs are transcribed from the La Crosse virus small genome segment. *Cell*. **33**, 791-799.
586. Patterson, J.L., Holloway, B.P. and Kolakofsky, D. (1984). La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease. *Journal of Virology*. **52**, 215-222.
587. Pearson, W.R. (1996). Effective protein sequence comparison. *Methods in Enzymology*. **266**, 227-258.
588. Pedersen, I.R. (1970). Density gradient centrifugation studies on Lymphocytic Choriomeningitis virus and on viral ribonucleic acid. *Journal of Virology*. **6**, 414-420.
589. Pedersen, I.R. (1971). Lymphocytic Choriomeningitis virus RNAs. *Nature (Londres)*. **234**, 112-114.
590. Pedersen, I.R. (1973). LCM virus: its purification and its chemical and physical properties. In: Lymphocytic choriomeningitis virus and other arenaviruses. (Lehmann Grube, F. Ed.). Springer Verlag, Berlin. pp: 13-23.
591. Pedersen, I.R. and Konigshofer, E.P. (1976). Characterization of ribonucleoproteins and ribosomes isolated from Lymphocytic Choriomeningitis virus. *Journal of Virology*. **20**, 14-21.
592. Pedersen, I.R. (1979). Structural components and replication of arenaviruses. *Advances in Virus Research*. **24**, 277-330.
593. Pedersen, J.T. and Moulton, J. (1996). Genetic algorithms for protein structure prediction. *Current Opinion in Structural Biology*. **6**, 227-31.
594. Pedersen, J.T. and Moulton, J. (1997). Protein folding simulations with genetic algorithms and a detailed molecular description. *Journal of Molecular Biology*. **269**, 240-259.
595. Persson, B. and Argos, P. (1994). Prediction of transmembrane segments in proteins utilising multiple sequence alignments. *Journal of Molecular Biology*. **237**, 182-192.
596. Persson, B. and Argos, P. (1996). Topology prediction of membrane proteins. *Protein Science*. **5**, 363-371.
597. Peters, C.J., Webb, P.A. and Johnson, K.M. (1973). Measurement of antibodies to Machupo virus by the indirect fluorescent technique. *Proceedings of the Society in Experimental Biology and Medicine*. **142**, 526-531.
598. Peters, C.J., Kuehne, R.W., Mercado, R.R., Le Bow, R.H., Spertzel, R.O. and Webb, P.A. (1974). Hemorrhagic fever in Cochabamba, Bolivia, 1971. *American Journal of Epidemiology*. **99**, 425.
599. Peters, C.J., Jahrling, P.B., Liu, C.T., Kenyon, R.H., McKee Jr., K.T. and Barrera Oro, J.G. (1987). Experimental studies of arenaviral hemorrhagic fevers. *Current Topics in Microbiology and Immunology*. **134**, 5-68.
600. Peters, C.J. (1998). Hemorrhagic fevers: How they wax and wane. In: Emerging Infections I. (W.M. Scheld, D. Armstrong and J.M. Hughes, eds). American Society for Microbiology, Massachusetts. p. 20.
601. Pevear, D.C. and Pfau, C.J. (1989). Lymphocytic Choriomeningitis virus. In: Clinical and molecular aspects of neurotropic virus infections. (D.H. Gilden and H.L. Lipton, Eds.). Kluwer Academic Publishers, New York. pp. 141-172.
602. Pfau, C.J. (1974). Biochemical and biophysical properties of the arenaviruses. *Progress in Medical Virology*. **18**, 64-80.
603. Pfau, C.J., Bergold, G.H., Casals, J., Johnson, K.M., Murphy, F.A., Pedersen, I.R., Rawls, W.E., Rowe, W.P., Webb, P.A. and Weissenbacher, M.C. (1974). Arenaviruses. *Intervirology*. **4**, 207-210.

Referencias bibliográficas

604. Pierangeli, A., Bucci, M., Forzan, M., Pagnotti, P., Equestre, M. and Pérez Bercoff, R. (1999). Primer alignment-and-extension: a novel mechanism of viral RNA recombination responsible for the rescue of inactivated poliovirus cDNA clones. *Journal of General Virology*. **80**, 1889-1897.
605. Pilipenko, E.V., Blinov, V.M., Chernov, B.K., Dmitrieva, T.M. and Agol, V.I. (1989a). Conservation of the secondary structure elements of the 5'-untranslated region RNA of cardio- and aphthovirus RNAs. *Nucleic Acids Research*. **17** (14), 5701-5711.
606. Pilipenko, E.V., Blinov, V.M., Romanova, L.I., Sinyakov, A.N., Maslova, S.V. and Agol, V.I. (1989b). Conserved structural domains in the 5'-untranslated region of picornaviral genomes: an analysis of the segment controlling translation and neurovirulence. *Virology*. **168** (1), 201-209.
607. Pilipenko, E.V., Blinov, V.M. and Agol, V.I. (1990). Gross rearrangement within the 5'-untranslated region of the picornaviral genomes. *Nucleic Acids Research*. **18** (11), 13371-13375.
608. Pircher, H., Moskophidis, D., Rohrer, U., Bürki, K., Hengartner, H. and Zinkernagel, R.M. (1990). Viral escape by selection of cytotoxic T cell-resistant virus variants *in vivo*. *Nature*. **346**, 629-633.
609. Piroosky, I., Zuccarini, J., Molinelli, E.A., Di Pietro, A., Martini, P., Ferreyra, B., Gutman Frugone, I.F. and Vazquez, T. (1959). Virosis hemorrágica del noroeste bonaerense. Endemoepidémica, febril, enantemática y leucopénica. I. La primera inoculación experimental al hombre. *Orientación Médica*. **8**, 144-148.
610. Piwnica-Worms, H. (1990). Preparation of insect cell cultures and baculovirus stocks. In F.M. Ausubel (Ed.), *Current Protocols in Molecular Biology*, Vol. 2. John Wiley & Sons, NY.
611. Plotch, S.J., Bouloy, M., Ulmanen, I. and Krug, R.M. (1981). A unique cap (m⁷GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*. **23**, 847-858.
612. Poch, O., Blumberg, B.M., Bouguelert, I. and Tordo, N. (1990). Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *Journal of General Virology*. **71**: 1153.
613. Polo, J.M. and Johnston, R.E. (1991). Mutational analysis of a virulence locus in the E2 glycoprotein gene of Sindbis virus. *Journal of Virology*. **65**, 6358-6361.
614. Popeacu, M., Schaeffer, H. and Lehmann-Grube, F. (1976). Homologous interference of lymphocytic choriomeningitis virus detection and measurement of interference focus unit. *Journal of Virology*. **20**, 1-7.
615. Popeacu, M. and Lehmann-Grube, F. (1977). Defective interfering particles in mice infected with lymphocytic choriomeningitis virus. *Virology*. **77**, 78-83.
616. Prehaud, C., Coulon, P., Lafay, F., Thiers, C. and Flamand, A. (1988). Antigenic site II of the rabies virus glycoprotein: structure and role in viral virulence. *Journal of Virology*. **62**, 1-7.
617. Raju, R. and Kolakofsky, D. (1989). The ends of La Crosse virus genome and antigenome RNAs within nucleocapsids are base paired. *Journal of Virology*. **63**, 122-128.
618. Raju, R., Raju, L., Hacker, D., Garcin, D., Compans, R.W. and Kolakofsky, D. (1990). Non templated bases at the 5' ends of Tacaribe virus mRNAs. *Virology*. **174**, 53-59.
619. Raju, R., Subramaniam, S.V. and Hajjou, M. (1995). Genesis of Sindbis virus by *in vivo* recombination of nonreplicative RNA precursors. *Journal of Virology*. **69**, 7391-7401.
620. Raju, R.V., Kakkar, R., Radhi, J.M. and Sharma, R.K. (1997). Biological significance of phosphorylation and myristoylation in the regulation of cardiac muscle proteins. *Molecular and Cellular Biochemistry*. **176**, 135-143.
621. Ramirez, B.C., Garcin, D., Calvert, L.A., Kolakofsky, D. and Haenni, A-L. (1995). Capped nonviral sequences at the 5' end of the mRNAs of Rice Hoja Blanca Virus RNA4. *Journal of Virology*. **69**, 1951-1954.
622. Ramos, B.A., Courtney, R.J. and Rawls, W.E. (1972). The structural proteins of Pichinde virus. *Journal of Virology*. **10**, 661-667.
623. Ramsingh, A.I., Dimock, K., Rawls, W.E. and Leung, W.C. (1980). Size estimation of Pichinde virus RNA by gel electrophoresis under denaturing conditions. *Intervirology*. **14**, 31-36.
624. Rao, M.J.K. and Argos, P. (1986). A conformational preference parameter to predict helices in integral membrane proteins. *Biochimie et Biophysic Acta*. **869**, 197-214.
625. Rawls, W.E. and Buchmeier, M.J. (1976). Arenaviruses: purification and physicochemical nature. *Bulletin WHO*. **52**, 393-401.
626. Rawls, W.E., Chan, M.A. and Gee, S.R. (1981). Mechanism of persistence in arenavirus infections: a brief review. *Canadian Journal of Microbiology*. **27**, 568-574.
627. Rey, O., Rossi, J.P.F.C., López, R., Iapalucci-Espinoza, S. and Franze Fernández, M.T. (1988). Tacaribe virus infections may induce inhibition of the activity host cell Ca²⁺ and Na⁺/K⁺ pumps. *Journal of General Virology*. **69**, 951-954.
628. Rey, O. (1990). Efecto temprano en la interacción del virus Tacaribe con la célula hospedadora y mecanismos implicados. Tesis Doctoral. Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires.
629. Rice, P.M., Elliston, K. and Gribskov, M. (1991). Similarity and homology. In: *Sequence Analysis Primer*. (M. Gribskov and J. Deveraux, eds.) Serie: UWBC Biotechnical Resource Series (R.R. Burgess, ed.). Stockton Press, New York, pp.: 1-59.
630. Richardson, W.H. (1964). Frequencies of genotypes of relatives, as determined by stochastic matrices. *Genetica*. **35**, 323-354.
631. Riesner, D. and Romer, R. (1973). In: *Physicochemical properties of nucleic acids*. (J. Duchesne, ed.). Academic Press, New York, vol. 2, p. 237.
632. Rigby, P.W.J., Dickermann, M., Rhodes, C. and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *Journal of Molecular Biology*. **113**, 237-345.
633. Riis, S.K. and Krogh, A. (1996). Improving prediction of protein secondary structure using structured neural networks and multiple sequence alignments. *Journal of Computational Biology*. **3**, 163-183.
634. Rivera Pomar, R.V., Manzella, E.M., Ghiringhelli, P.D., Grau, O. and Romanowski, V. (1991). Patterns of transient expression of the arenavirus nucleocapsid protein gene in transfected cells. *Microscopía Electrónica y Biología Celular*. **15**, 41-56.
635. Rivera Pomar, R.V., Lozano, M.E., Albariño, C.G., Ghiringhelli, P.D., Grau, O. and Romanowski, V. (1993). *In vivo* antiterminator activity of Junin arenavirus N protein. IXth International Congress of Virology. Glasgow, Escocia.
636. Rivers, T.M. and Scott, T.F.M. (1935). Meningitis in man caused by a filterable virus. *Science*. **81**, 439-440.
637. Riviere, Y., Ahmed, R., Southern, P.J., Buchmeier, M.J., Dutko, F.J. and Oldstone, M.B.A. (1985a). The S RNA segment of lymphocytic choriomeningitis virus codes for the nucleoprotein and the glycoproteins 1 and 2. *Journal of Virology*. **53**, 966-968.
638. Riviere, Y., Ahmed, R., Southern, P.J., Buchmeier, M.J. and Oldstone, M.B.A. (1985b). Genetic mapping of lymphocytic choriomeningitis virus pathogenicity: virulence in guinea pigs is associated with the L RNA segment. *Journal of Virology*. **55**, 704-709.

639. Robinson, N.J., Robinson, P.J., Gupta, A., Bleasby, A.J., Whitton, B.A. and Morby, A.P. (1995). Singular over-representation of an octameric palindrome, HIP1, in DNA from many cyanobacteria. *Nucleic Acids Research*. **23**, 729-735.
640. Romanova, L.I., Blinov, V.M., Tolskaya, E.A., Viktorova, E.G., Kolesnikova, M.S., Guseva, E.A. and Agol, V.I. (1986). The primary structure of the crossover regions of intertypic poliovirus recombinants: A model of recombination between RNA genomes. *Virology*. **155**, 202-213.
641. Romanowski, V. (1981). Estructura bioquímica del virus Junin. Tesis doctoral. Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina.
642. Romanowski, V. and Bishop, D.H.L. (1983). The formation of Arenaviruses that are genetically diploid. *Virology*. **126**, 87-95.
643. Romanowski, V. and Bishop, D.H.L. (1985). Conserved sequences and coding of two strains of lymphocytic choriomeningitis virus (WE and ARM) and Pichinde arenavirus. *Virus Research*. **2**, 35-51.
644. Romanowski, V., Matsuura, Y. and Bishop, D.H.L. (1985). Complete sequence of the S RNA of lymphocytic choriomeningitis virus (WE strain) compared to that of Pichinde arenavirus. *Virus Research*. **3**, 101-114.
645. Romanowski, V. (1993). Genetic organization of Junin virus, the etiological agent of Argentine hemorrhagic fever. In: *The Arenaviridae*. (Salvato, M.S. Ed.), Serie: *The Viruses* (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 51-83.
646. Romanowski, V., Tortorici, M.A., Albariño, C.G. and Ghiringhelli, P.D. (1999a). Arenavirus nucleocapsid protein has a zinc finger region critical for RNA binding. *1999 Meeting of the American Society for Biochemistry and Molecular Biology (ASBMB)*, San Francisco (Estados Unidos).
647. Romanowski, V., Tortorici, M.A., Albariño, C.G. and Ghiringhelli, P.D. (1999b). Arenavirus nucleocapsid protein has a zinc finger region critical for RNA binding. *FASEB Journal* **13**, A1503
648. Rosas, M.F. (1984). Bioquímica de virus: estructura de la envoltura del virus Junin. Tesis doctoral. Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina.
649. Rosas, M.F., Romanowski, V. and Grau, O. (1988). The phospholipid composition of arenaviruses. *Analales de la Asociación Química Argentina*. **76**, 269-283.
650. Rost, B. and Sander, C. (1993a). Secondary structure prediction of all-helical proteins in two states. *Protein Engineering*. **6**, 831-836.
651. Rost, B. and Sander, C. (1993b). Improved prediction of protein secondary structure by use of sequence profiles and neural networks. *Proceedings of National Academy of Sciences USA*. **90**, 7558-7562.
652. Rost, B. and Sander, C. (1993c). Prediction of protein structure at better than 70% accuracy. *Journal of Molecular Biology*. **232**, 584-599.
653. Rost, B. and Sander, C. (1994a). Combining evolutionary information and neural networks to predict protein secondary structure. *Proteins*. **19**, 55-72.
654. Rost, B. and Sander, C. (1994b). Structure prediction of proteins - where are we now? *Current Opinion in Biotechnology*. **5**, 372-380.
655. Rost, B. and Sander, C. (1994c). Conservation and prediction of solvent accessibility in protein families. *Proteins*. **20**, 216-226.
656. Rost, B. and Sander, C. (1994d) 1D secondary structure prediction through evolutionary profiles. In: Bohr, H. and Brunak, S. (eds.) "Protein Structure by Distance Analysis", Amsterdam: IOS Press. pp 257-276.
657. Rost, B., Sander, C. and Schneider, R. (1994a). PHD - an automatic mail server for protein secondary structure prediction. *CABIOS*. **10**, 53-60.
658. Rost, B., Sander, C. and Schneider, R. (1994b). Redefining the goals of protein secondary structure prediction. *Journal of Molecular Biology*. **235**, 13-26.
659. Rost, B. (1995a). In *Protein folds: a distance based approach* (Bohr, H. and Brunak, S., eds.), *Protein folds: a distance based approach*. CRC Press, Boca Raton, Florida, pp. 132-151.
660. Rost, B. (1995b). In *Third International Conference on Intelligent Systems for Molecular Biology* (Rawlings, C., Clark, D., Altman, R., Hunter, L., Lengauer, T. and Wodak, S., eds.), pp. 314-321, Menlo Park, CA: AAAI Press, Cambridge, England
661. Rost, B. (1995c). TOPITS: Threading One-dimensional Predictions Into Three-dimensional Structures. In: C Rawlings, D Clark, R Altman, L Hunter, T Lengauer, & S Wodak (eds.). "The third international conference on Intelligent Systems for Molecular Biology (ISMB)", Cambridge, U.K. Menlo Park, CA: AAAI Press pp. 314-321.
662. Rost, B., Casadio, R., Fariselli, P. and Sander, C. (1995). Prediction of helical transmembrane segments at 95% accuracy. *Protein Science*. **4**, 521-533.
663. Rost, B. (1996a). Expected prediction accuracy of PHD. WWW document (<http://dodo.cpmc.columbia.edu/~rost/Res/96D-ExpAccuracyPHD.html>), EMBL Heidelberg, Germany.
664. Rost, B. (1996b). Protein fold recognition by merging 1D structure prediction and sequence alignments. WWW document (<http://dodo.cpmc.columbia.edu/~rost/Papers/96PreTopits.html>), EMBL Heidelberg, Germany.
665. Rost, B. (1996c). PHD: predicting one-dimensional protein structure by profile based neural networks. *Methods in Enzymology*. **266**, 525-539.
666. Rost, B. (1996d). Sneaking in genomes for helical transmembrane proteins. Talk presented at Distance based approaches to protein structure prediction III', Copenhagen, Denmark (manuscript in preparation), EMBL, Heidelberg, Germany.
667. Rost, B. and Sander, C. (1996). Bridging the protein sequence-structure gap by structure predictions. *Annual Review in Biophysic and Biomolecular Structures*. **25**, 113-136.
668. Rost, B. and Valencia, A. (1996). Pitfalls of protein sequence analysis. *Current Opinion in Biotechnology*. **7**, 457-461.
669. Rost, B., Casadio, R. and Fariselli, P. (1996a). Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Science*. **5**, 1704-1718.
670. Rost, B., O'Donoghue, S. and Sander, C. (1996b). Protein structures evolve at random - almost. (Preprint, nunca publicado)
671. Rost, B., Schneider, R. and Sander, C. (1997). Protein fold recognition by prediction-based threading. *Journal of Molecular Biology*. **270**, 471-480.
672. Rost, B. & O'Donoghue, S. (1997). Sisyphus and protein structure prediction. *Bioinformatics*. **13**, 345-356.
673. Rothe, M., Wong, S.C., Henzel, W.J. and Goeddel, D.V. (1994). A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. *Cell*. **78**, 681-682. (ACCESSION L35303).
674. Rowe, W.P., Pugh, W.F., Webb, P.A. and Peters, C.J. (1970a). Serological relationships of the Tacaribe complex of viruses to lymphocytic choriomeningitis virus. *Journal of Virology*. **5**, 289-292.

675. Rowe, W.P., Murphy, F.A., Bergold, G.H., Casals, J., Hotchin, J., Johnson, K.M., Lehmann-Grube, F., Mims, C.A., Traub, E. and Webb, P.A. (1970b). Arenaviruses: Proposed name for a newly defined virus group. *Journal of Virology*. **5**, 651-652.
676. Rowe, C.L., Fleming, J.O., Nathan, M.J., Sgro, J-I., Palmenberg, A.C. and Baker, S.C. (1997). Generation of Coronavirus spike deletion variants by high-frequency recombination at regions of predicted RNA secondary structure. *Journal of Virology*. **71**, 6183-6190.
677. Rozovskaya, T.A., Chenchik, A.A. and Beabealashvili, R.Sh. (1982). Processive pyrophosphorolysis of RNA by *Escherichia coli* RNA polymerase. *FEBS Letter*. **137**, 100-104.
678. Rozovskaya, T.A., Rechinsky, V.O., Bibilashvili, R.S., Karpeisky, M.Ya., Tarusova, N.B., Khomutov, R.M. and Dixon, H.B. (1984). The mechanism of pyrophosphorolysis of RNA by RNA polymerase. Endowment of RNA polymerase with artificial exonuclease activity. *Biochemistry Journal*. **224**, 645-650.
679. Rudd, M.D., Izbán, M.G. and Luse, D.S. (1994). The active site of RNA polymerase II participate in transcript cleavage within arrested ternary complexes. *Proceedings of National Academy of Sciences USA*. **91**, 8057-8061.
680. Rugiero, H.R., Cintora, A., Libonatti, E., Magnoni, C., Castiglione, E. and Locisera, R. (1960). Formas nerviosas de la fiebre hemorrágica epidémica. *Prensa Med. Argentina*. **47**, 1845.
681. Rugiero, H.R., Parodi, A.S., Ruggiero, H.G., Mettler, N., Boxaca, M., de Guerrero, A.L., Cintora, A., Magnoni, C., Milani, H., Maglio, F., Gonzalez Cambaceres, C., Astarloa, L., Squassi, G., Fernández, D. and Giacosa, A. (1964a). Fiebre hemorrágica argentina. I. Periodo de incubación e invasión. *Revista de la Asociación Médica Argentina*. **78**, 221.
682. Rugiero, H.R., Cintora, A., Magnoni, C., Ruggiero, H.G., Gonzalez Cambaceres, C., Maglio, F., Astarloa, L., Squassi, G. and Fernández, D. (1964b). Fiebre hemorrágica argentina. III. Aparato cardiovascular. *Revista de la Asociación Médica Argentina*. **78**, 360-371.
683. Rugiero, H.R., Cintora, A., Ruggiero, H.G., Mettler, N., Boxaca, M., de Guerrero, A.L., Cintora, A., Magnoni, C., Milani, H., Maglio, F., Gonzalez Cambaceres, C., Astarloa, L., Squassi, G., Fernández, D. (1964c). Fiebre hemorrágica argentina. IV. Formas clínicas. *Revista de la Asociación Médica Argentina*. **78**, 500.
684. Ruiz-Jarabo, C.M., Arias, A., Baranowski, E., Escarmis, C. and Domingo, E. (2000). Memory in viral quasispecies. *Journal of Virology*. **74**, 3543-3547.
685. Ruo, S.L., Mitchell, S.W., Kiley, M.P., Roumillat, L.F., Fisherhoch, S.P. and McCormick, J.B. (1991). Antigenic relatedness between arenaviruses defined at the epitope level by monoclonal antibodies. *Journal of General Virology*. **72**, 549-555
686. Russell, R.B. and Barton, G.J. (1994). Structural features can be unconserved in proteins with similar folds. *Journal of Molecular Biology*. **244**, 332-350.
687. Russell, R.B., Copley, R.R. and Barton, G.J. (1996). Protein fold recognition by mapping predicted secondary structures. *Journal of Molecular Biology*. **259**, 349-365.
688. Rustici, S.M. (1984). Desarrollo *in vitro* del virus Junín, proteínas intracelulares. Tesis doctoral. Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina
689. Sabattini, M.S., Gonzalez de Ríos, L.E., Díaz, G. and Vega, V.R. (1977). Infección natural y experimental de roedores con virus Junín. *Medicina (Buenos Aires)*. **37** Supl. 3, 149-161.
690. Salamov, A.A. and Solovyev, V.V. (1995). Prediction of protein secondary structure by combining nearest-neighbor algorithms and multiple sequence alignment. *Journal of Molecular Biology*. **247**, 11-15.
691. Salas, R., Manzione, N., Tesh, R.B., Rico-Hesse, R., Shope, R., Betancourt, A., Godoy, O., Bruzual, R., Pacheco, M., Ramos, B., Taibo, M.E., Garcia Tamayo, J., Jaimes, E., Vasquez, C., Araoz, F. and Quarales, J. (1991). Venezuelan hemorrhagic fever. *Lancet*. **338**, 1033-1036.
692. Salch, F., Gard, G.P. and Compans, R.W. (1979). Synthesis of Tacaribe viral proteins. *Virology*. **93**, 369-376.
693. Sali, A. and Blundell, T. (1994). In Protein Structure by Distance Analysis (Bohr, H. and Brunak, S., eds.), *Protein Structure by Distance Analysis*. IOS Press, Amsterdam, Oxford, Washington, pp. 64-87.
694. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins*. **23**, 318-326.
695. Salvato M., Shimomaye, E., Southern, P. and Oldstone, M.B.A. (1988). Virus-Lymphocyte interactions. IV. Molecular characterization of LCMV Armstrong (CTL⁺) small genomic segment and that of its variant, Clone 13 (CTL⁻). *Virology*. **164**, 517-522.
696. Salvato, M.S. and Shimomaye, E.M. (1989). The completed sequence of lymphocytic choriomeningitis virus reveals a unique RNA structure and a gene for a zinc finger protein. *Virology*. **173**, 1-10.
697. Salvato, M.S., Shimomaye, E.M. and Oldstone, M.B.A. (1989). The primary structure of the lymphocytic choriomeningitis virus L gene encodes a putative RNA polymerase. *Virology*. **169**, 377-384.
698. Salvato, M., Borrow, P., Shimomaye, E. and Oldstone, M.B.A. (1991). Molecular basis of viral persistence: a single amino acid change in the glycoprotein of Lymphocytic Choriomeningitis virus es associated with supression of the antiviral cytotoxic T Lymphocyte response and establishment of persistence. *Journal of Virology*. **65**, 1863-1869.
699. Salvato, M.S., Schweighofer, K.J., Burns, J.W. and Shimomaye, E.S. (1992). Biochemical and immunological evidence that the 11 kDa zinc-binding protein of lymphocytic choriomeningitis virus is a structural component of the virus. *Virus Research*. **22**, 185.
700. Salvato, M.S. (1993). Molecular biology of the prototype arenavirus, lymphocytic chorio-meningitis virus. In: The Arenaviridae. (Salvato, M.S. Ed.), Serie: The Viruses (Wagner, R. and Frankel Conrat, H. Eds.). Plenum Press, Nueva York. pp: 133-156.
701. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning. A laboratory manual. 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York.
702. Samudrala, R., Pedersen, J.T., Zhou, H.-b., Luo, R., Fidelis, K. and Moul, J. (1995). Confronting the problem of interconnected structural changes in the comparative modeling of proteins. *Proteins*. **23**, 327-336.
703. Sanchez, A., Pifat, D.Y., Kenyon, R.H., Peters, C.J., McCormick, J.B. and Kiley, M.P. (1989). Junin virus monoclonal antibodies: characterization and crossreactivity with other arenaviruses. *Journal of General Virology*. **70**, 1125-1132.
704. Sander, C. and Schneider, R. (1991). Database of homology-derived structures and the structurally meaning of sequence alignment. *Proteins*. **9**, 56-68.
705. Sander, C. and Schneider, R. (1994). The HSSP data base of protein structure-sequence alignments. *Nucleic Acids Research*. **22**, 3597-3599.
706. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of National Academy of Sciences USA* **74**, 5463-5466.
707. Sanger, F. and Coulson, A.R. (1978). The use of thin acrylamide gel for DNA sequencing. *FEBS Letter*. **87**, 107-110.

708. Sastry, S.S. and Ross, B.M. (1997). Nuclease activity of T7 RNA polymerase and the heterogeneity of transcription elongation complexes. *The Journal of Biological Chemistry*. **272**, 8644-8652.
709. Sato, T., Irie, S. and Reed, J.C. (1995). A novel member of the TRAF family of putative signal transducing proteins binds to cytosolic domain of CD40. *FEBS Letter*. **358**, 113-118. (ACCESSION L38509).
710. Schauser, L., Borg, S. and Poulsen, C. (1994). Pzf, a gene isolated from a Lotus japonicus nodule cDNA library, encodes a new plant member of the RING-finger family of zinc-binding proteins. Unpublished. Direct submission to GenBank (ACCESSION Z36749).
711. Schiffer, M. and Edmundson, A.B. (1967). Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophysical Journal*. **7**, 121-135.
712. Schneider, R., de Daruvar, A. and Sander, C. (1997). The HSSP database of protein structure-sequence alignments. *Nucleic Acids Research*. **25**, 226-230.
713. Schubert, M., Harmison, G.G. and Meier, E. (1984). Primary structure of the vesicular stomatitis virus polymerase (L) gene: evidence for a high frequency of mutations. *Journal of Virology*. **51**, 505-514.
714. Schuler, G.D., Altschul, S.F. and Lipman, D.J. (1991). A Workbench for Multiple Alignment Construction and Analysis. *Proteins Structure Function and Genetics*. **9**, 180-190.
715. Schwartz, R.M., and Dayhoff, M.O. 1979. Dayhoff, M.O., ed. *Atlas of protein sequence and structure* Washington D.C. National Biomedical Research Foundation, pp 353-358.
716. Scolaro, L., Castilla, V., Candurra, N., Mersich, S. and Damonte, E. (1990). Virus Junin: funcionalidad de la glicoproteína de la envoltura en la determinación de su comportamiento biológico. IIIer Congreso Argentino de Virología. Resumen 8.
717. Segev, Y., Ofir, R., Salzberg, S., Heller, A., Weinstein, Y., Isakov, N., Udem, S., Wolfson, M. and Rager-Zisman, B. (1995). Tyrosine phosphorylation of measles virus nucleocapsid protein in persistently infected neuroblastoma cells. *Journal of Virology*. **69**, 2480-2485.
718. Senkevich, T.G., Koonin, E.V. and Buller, R.M. (1993). A poxvirus protein with a RING zinc finger motif is of crucial importance for virulence. *Virology*. **198**, 118-128. (ACCESSION U01161).
719. Shanafelt, A.B. (1991). An improved method for visualizing baculovirus plaques in cell agarose overlays. *Biotechniques*. **11**, 330.
720. Shapiro, B.A., Bengali, D., Kasprzak, W. and Wu, J.C. (2001). RNA folding pathway functional intermediates: their prediction and analysis. *Journal of Molecular Biology*. **312**, 27-44.
721. Sheiness, D. and Sullivan, D. (1986). Staining RNA in formamide gels. *Focus*. **8**, 11.
722. Sheridan, R.P., Dixon, J.S. and Venkataraghavan, R. (1985). Generating plausible protein folds by secondary structure similarity. *International Journal of Peptide and Protein Research*. **25**, 132-143.
723. Shindyalov, I.N., Kolchanov, N.A. and Sander, C. (1994). Can three-dimensional contacts in protein structures be predicted by analysis of correlated mutations? *Protein Engineering*. **7**, 349-358.
724. Shioda, T., Iwasaki, K. and Shibuta, H. (1986). Determination of the complete nucleotide sequence of the Sendai virus genome RNA and the predicted amino acid sequences of the F, HN and L proteins. *Nucleic Acids Research*. **14**, 1545-1563.
725. Shiraishi, T., Misumi, S., Takama, M., Takahashi, I. and Shoji, S. (2001). Myristoylation of human immunodeficiency virus type 1 gag protein is required for efficient env protein transportation to the surface of cells. *Biochemical and Biophysical Research Communications*. **282**, 1201-1205.
726. Simon, A.E. and Bujarski, J.J. (1994). RNA-RNA recombination and evolution in virus infected plants. *Annual Reviews in Phytopathology*. **32**, 337-362.
727. Simmonds, P. and Smith, D.B. (1999). Structural constraints on RNA virus evolution. *Journal of Virology*. **73**, 5787-5794.
728. Singh, I.R. and Helenius, A. (1992). Role of ribosomes in Semliki Forest Virus nucleocapsid uncoating. *Journal of Virology*. **66**, 7049-7058.
729. Singh, I.R., Suomalainen, M., Varadarajan, S., Garoff, H. and Helenius, A. (1997). Multiple mechanisms for the inhibition of entry and uncoating of superinfecting Semliki Forest virus. *Virology*. **231**, 59-71.
730. Singh, M.K., Fuller-Pace, F.V., Buchmeier, M.J. and Southern, P.J. (1987). Analysis of the genomic L RNA segment of the lymphocytic choriomeningitis virus. *Virology*. **161**, 448-456.
731. Sippl, M.J. (1995). Knowledge-based potentials for proteins. *Current Opinion in Structural Biology*. **5**, 229-235.
732. Sippl, M.J. (1996). Helmholtz free energy of peptide hydrogen bonds in proteins. *Journal of Molecular Biology*. **260**, 644-648.
733. Southern, P.J. and Bishop, D.H.L. (1987). Sequence comparison among arenaviruses. *Current Topics in Microbiology and Immunology*. **133**, 19-39.
734. Smelt, S.C., Borrow, P., Kunz, S., Cao, W., Tishon, A., Lewicki, H., Campbell, K.P. and Oldstone, M.B. (2001). Differences in affinity of binding of Lymphocytic Choriomeningitis virus strains to the cellular receptor α -dystroglycan correlate with viral tropism and disease kinetics. *Journal of Virology*. **75**, 448-457.
735. Speir, R.W., Wood, O., Liebhafner, H. and Buckley, S.M. (1970). Lassa fever, a new virus disease of man from west Africa. IV. Electron microscopy of Vero cell cultures infected with Lassa virus. *American Journal of Tropical Medicine and Hygiene*. **19**, 692-694.
736. Spriggs, D.R., Bronson, R.T. and Fields, B.N. (1983). Hemagglutinin variants of reovirus type 3 have altered central nervous system tropism. *Science*. **220**, 505-507.
737. Staden, R. (1984). Computer methods to locate signals in nucleic acids sequences. *Nucleic Acids Research*. **12**, 505-519.
738. Staden, R. (1990). Searching for patterns in protein and nucleic acid sequences. *Methods in Enzymology*. **183**, 193-211.
739. Staneck, L.D., Trowbridge, R.S., Welsh, R.M., Wright, E.A. and Pfau, C.J. (1972). Arenaviruses: cellular response to long-term *in vivo* infection with Parana and lymphocytic choriomeningitis viruses. *Infect. Immun.* **6**, 444-450.
740. Staneck, L.D. and Pfau, C.J. (1974). Interfering particles from a culture persistently infected with Parana virus. *Journal of General Virology*. **22**, 437-440.
741. States, D.J. and Boguski, M.S. (1991). Similarity and homology. In: *Sequence Analysis Primer*. (M. Gribskov and J. Devereaux, eds.). Serie: UWBC Biotechnical Resource Series (R.R. Burgess, ed.). Stockton Press, New York, pp. 89-157.
742. States, D.J., Gish, W. and Altschul, S.F. (1991). Improved sensitivity of nucleic acid database similarity searches using application specific scoring matrices. *Methods, A companion to Methods in Enzymology*. **3**, 66-70.
743. States, D.J. and Gish, W. (1994). Combined use of sequence similarity and codon bias for coding region identification. *Journal of Computational Biology*. **1**, 39-50.
744. Stephens, E.B. and Compans, R.W. (1988). Assembly of animal viruses at cellular membranes. *Annual Review in Microbiology*. **42**, 489.
745. Sternberg, M.J.E. (1996). Protein structure prediction. Oxford Univ. Press, Oxford.

746. Stocker, C., Martinez Peralta, L., Kratzberg, T., Lohmann, F. and Bruns, M. (1994). Characterization of a virus variant produced by L cells persistently infected with lymphocytic choriomeningitis virus. *Journal of General Virology*. **75**, 3431-3439.
747. Stoker, M.G.P. and Mc Pherson, I.A. (1961). Studies on transformation of hamster cells by polyoma virus *in vitro*. *Virology*. **14**, 631-641.
748. Strauss, J.H. and Strauss, E.G. (1988). Evolution of RNA viruses. *Annual Reviews in Microbiology*. **42**, 657-683.
749. Summers, M.D. and Smith, G.E. (1987). A manual of methods for baculovirus vectors and insect cell culture procedures. Texas Agricultural Experiment Station. College Station, Tx.
750. Sun S, Thomas PD, Dill KA. (1995). A simple protein folding algorithm using a binary code and secondary structure constraints. *Protein Engineering*. **8**, 769-778.
751. Svarovskaia, E.S., Delviks, K.A., Hwang, C.K. and Pathak, V.K. (2000). Structural determinants of murine leukemia virus reverse transcriptase that affect the frequency of template switching. *Journal of Virology*. **74**, 7171-7178.
752. Sved, J. and Bird, A. (1990). The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proceedings of National Academy of Sciences USA*. **87**, 4692-4696.
753. Tatem, J.M., Weeks-Levy, C., Georgiu, A., DiMichele, S.J., Gorgacz, E.J., Racaniello, V.R., Cano, F.R. and Mento, S.J. (1992). A mutation in the amino terminus of Sabin 3 poliovirus VP1 protein is attenuating. *Journal of Virology*. **66**, 3194-3197.
754. Takehara, K., Min, M.K., Battles, J.K., Sugiyama, K., Emery, V.C., Dalrymple, J.M. and Bishop, D.H. (1989). Identification of mutations in the M RNA of a candidate vaccine strain of Rift Valley fever virus. *Virology*. **169**, 452-457.
755. Tashiro, M., Yamakawa, M., Tobita, K., Klenk, H.D., Rott, R. and Seto, J.T. (1990). Organ tropism of Sendai virus in mice: proteolytic activation of the fusion glycoprotein in mouse organs and budding site at the bronchial epithelium. *Journal of Virology*. **64**, 3627-3634.
756. Taylor, W.R. and Hatrick, K. (1994). Compensating changes in protein multiple sequence alignment. *Protein Engineering*. **7**, 341-348.
757. Taylor, W.R. (1996). Multiple protein sequence alignment: algorithms and gap insertion. *Methods in Enzymology*. **266**, 343-367.
758. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**, 4673-4680.
759. Thompson, M.J. and Goldstein, R.A. (1996a). Constructing amino acid residue substitution classes maximally indicative of local protein structure. *Proteins*. **25**, 28-37.
760. Thompson, M.J. and Goldstein, R.A. (1996b). Predicting solvent accessibility: higher accuracy using Bayesian statistics and optimized residue substitution classes. *Proteins*. **25**, 38-47.
761. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. **25**, 4876-4882.
762. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M., and Gralla, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature (New Biology)*. **246**, 40-41.
763. Tkaczewski, L.Z. and Boxaca, M.C. (1970). Cambios citoquímicos en células Vero infectadas con algunos virus del grupo Tacaribe. *Medicina (Buenos Aires)*. **30**, 28-37.
764. Tkaczewski, L.Z. and Correa, J. (1972). Cytopathogenic process in Junín virus infected Vero cells. *Medicina (Buenos Aires)*. **32**, 293-301.
765. Tolskaya, E.A., Romanova, L.I., Blinov, V.M., Viktorova, E.G., Sinyakov, A.N., Kolesnikova, M.S. and Agol, V.I. (1987). Studies of the recombination between RNA genomes of poliovirus: The primary structure and non-random distribution of crossover regions in the genomes of intertypic poliovirus recombinants. *Virology*. **161**, 54-61.
766. Torda, A.E. (1997). Perspectives on protein fold recognition. *Current Opinion in Structural Biology*. **7**, 200-205.
767. Tordo, M., Poch, O., Ermine, A., Keith, G. and Rougeon, F. (1988). Completion of the Rabies virus genome sequence determination: highly conserved domain among the L (polymerase) proteins of unsegmented RNA viruses. *Virology*. **165**, 565-576.
768. Tortorici, M.A., Ghiringhelli, P.D. and Romanowski, V. (1998). Virus Junín: interacción proteína de nucleocápside - RNA S - zinc. *XXXIV Reunión Anual de la Sociedad Argentina de Investigación en Bioquímica y Biología Molecular*. Mendoza (Mendoza).
769. Tortorici, M.A., Albariño, C.G., Ghiringhelli, P.D. and Romanowski, V. (1999). Identificación de las zonas de unión a zinc y RNA de la proteína N del virus Junín. *XXXV Reunión Anual de la Sociedad Argentina de Investigación en Bioquímica y Biología Molecular*. Mendoza (Mendoza).
770. Tortorici, M.A. (2000). Regulación de la transcripción del RNA S de arnavirus. análisis de la interacción de la proteína N del virus Junín con el RNA y el ion zinc. Tesis doctoral. Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Argentina
771. Tortorici, M.A., Ghiringhelli, P.D., Lozano, M.E., Albariño, C.G. and Romanowski, V. (2001a). Zinc-binding properties of Junín virus nucleocapsid protein. *Journal of General Virology*. **82**, 121-128.
772. Tortorici, M.A., Albariño, C.G., Posik, D.M., Ghiringhelli, P.D., Lozano, M.E., Rivera Pomar, R. and Romanowski, V. (2001b). Arnavirus nucleocapsid protein displays a transcriptional antitermination activity *in vivo*. *Virus Research*. **73**, 41-55.
773. Traub, E. (1935). A filterable virus recovered from white mice. *Science*. **81**, 198-299.
774. Tuckwell, D.S., Humphries, M.J. and Brass, A. (1995). Protein secondary structure prediction by the analysis of variation and conservation in multiple alignments. *CABIOS*. **11**, 627-632.
775. Turner, D.H.; Sugimoto, N.; Jaeger, J.A.; Longfellow, C.E.; Freier, S.M. and Kierzek, R. (1987). Improved parameters for prediction of rna structure. *CSH Symposia in Quantitative Biology*. **52**, 123-133.
776. Turner, D.H. and Sugimoto, N. (1988). Rna structure prediction. *Annual Review in Biophysical Chemistry*. **17**, 167-192.
777. Urban, S., Urban, S., Fischer, K.P. and Tyrrell, D.L. (2001). Efficient pyrophosphorolysis by a hepatitis B virus polymerase may be a primer-unblocking mechanism. *Proceedings of National Academy of Sciences USA*. **98**, 4984-4989.
778. Utsumi, T., Tokunaga, T., Horii, J., Edashige, K., Utsumi, K., Koga, D. and Ide, A. (1994). Myristoylation of protein at a distinct position allows its phosphorylation by protein kinase C. *Archives in Biochemistry and Biophysics*. **313**, 337-345.
779. Valencia, A., Hubbard, T.J., Muga, A., Bañuelos, S., Llorca, O., Carrascosa, J. and Valpuesta, J.M. (1995). Prediction of the structure of GroES and its interaction with GroEL. *Proteins*. **22**, 199-209.
780. Van der Zeijst, B.A.M., Bleumink, N., Crawford, L.V., Swyryd, A. and Stark, G.R. (1983). Viral proteins and RNAs in BHK cells persistently infected with lymphocytic choriomeningitis virus. *Journal of Virology*. **48**, 262-270.
781. Vapalahti, O., Plyusnin, A., Cheng, Y., Manni, T., Brummer-Korvenkontio, M. and Vaheri, A. (1996). Inkoo and Tahyna, the European California serogroup bunyaviruses: sequence and phylogeny of the S RNA segment. *Journal of General Virology*. **77**, 1769-1774.
782. Vaughn, J. L., Goodwin, R.H., Tompkins, G. J & McCawley, P. (1977). The establishment of two cell lines from the insect (Lepidoptera: Noctuidae). *In vitro*. **13**, 213-217.

783. Vezza, A.C., Gard, G.P., Compans, R.W. and Bishop, D.H.L. (1977). Structural components of the arenavirus Pichinde. *Journal of Virology*. **23**, 776-786.
784. Vezza, A.C., Clewley, J.P., Gard, G.P., Abraham, N.Z., Compans, R.W. and Bishop, D.H.L. (1978). Virion RNA species of the arenaviruses Pichinde, Tacaribe and Tamiami. *Journal of Virology*. **26**, 485-497.
785. Vezza, A.C., Cash, P., Jahrling, P., Eddy, G. and Bishop, D.H.L. (1980). Arenavirus recombination: The formation of recombinants between prototype Pichinde and Pichinde-Munchique viruses and evidence that arenavirus S RNA codes for N polypeptide. *Virology*. **106**, 250-260.
786. Vialat, P., Muller, R., Vu, T.H., Prehaud, C. and Bouloy, M. (1997). Mapping of the mutations present in the genome of the Rift Valley fever virus attenuated MP12 strain and their putative role in attenuation. *Virus Research*. **52**, 43-50.
787. Vieira, J. and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*. **19**, 259-268.
788. Vieira, J. and Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods in Enzymology*. **153**, 3-24.
789. Vidal, M.C. and Coto, C.E. (1980). Persistent infection of mouse cells with Junin virus: participation of interferon. *Medicina (B. Aires)*. **40**, 537-542.
790. Vilches, A.M., Barrera Oro, J.G. and Gutman Frugone, L.F. (1965). Una nueva área epidémica de fiebre hemorrágica argentina iniciada a continuación de una epizootia en población aumentada de roedores silvestres. *II Jornadas Epidemiológicas Argentinas*. **3**, 359-360.
791. Villareal, L.P., Defilippis, V.R. and Gottlieb, K.A. (2000). Acute and persistent viral life strategies and their relationship to emerging diseases. *Virology*. **272**, 1-6.
792. Vinals, C., De Bolle, X., Depiereux, E. and Feytmans, E. (1995). Knowledge-based modeling of the D-lactate dehydrogenase three-dimensional structure. *Proteins*. **21**, 307-318.
793. Vingron, M. and Waterman, M.S. (1994). Sequence alignment and penalty choice. *Journal of Molecular Biology*. **235**, 1-12.
794. Vogelstein, B. and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proceedings of National Academy of Sciences USA*. **76**, 615-619.
795. von Heijne, G. (1987). Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit. Academic Press, San Diego.
796. von Heijne, G. (1994). Membrane proteins: from sequence to structure. *Annual Review in Biophysics and Biomolecular Structure*. **23**, 167-192.
797. von Heijne, G. (1996). In Protein structure prediction (Sternberg, M. J. E., eds.), Protein structure prediction. Oxford Univ. Press, Oxford, pp. 101-110.
798. Vrati, S., Faragher, S.G., Weir, R.C. and Dalgarno, L. (1986). Ross River virus mutant with a deletion in the E2 gene: properties of the virion, virus-specific macromolecule synthesis, and attenuation of virulence for mice. *Virology*. **151**, 222-232.
799. Vriend, G. (1990). WHAT IF: A molecular modelling and drug design program. *Journal of Molecular Graphics*. **8**, 52-56.
800. Wako, H. and Blundell, T.L. (1994). Use of amino acid environment-dependent substitution tables and conformational propensities in structure prediction from aligned sequences of homologous proteins I. Solvent accessibility classes. *Journal of Molecular Biology*. **238**, 682-692.
801. Wallgren, A. (1925). Une nouvelle maladie infectieuse du system nerveux central. *Acta Paediatrica*. **4**, 158.
802. Walseth, T.F. and Johnson, R.A. (1979). The enzymatic preparation of [α -(32 P)]nucleoside triphosphates, cyclic [32 P] AMP, and cyclic [32 P] GMP. *Biochimie and Biophysic Acta*. **562**, 11-31.
803. Wang, D. and Hawley, D.K. (1993). Identification of a 3' \rightarrow 5' exonuclease activity associated with human RNA polymerase II. *Proceedings of National Academy of Sciences USA*. **90**, 843-847.
804. Weaver, S.C., Salas, R.A., de Manzione, N., Fulhorst, C.F., Travasos da Rosa, A.P.A. Duno, G., Utrera, A., Ksiazek, T.G., Tovar, D., Guzman, H., Kang, W. and Tesh, R.B. (2001). Extreme genetic diversity among Pirital virus (Arenaviridae) isolates from western Venezuela. *Virology*. **285**, 110-118.
805. Webb, P.A., Johnson, K.M. and MacKensie, R.B. (1969). The measurement of specific antibodies in Bolivian hemorrhagic fever by neutralization of virus plaques. *Proc. Soc. Exp. Biol. Med.* **130**, 1013-1016.
806. Webb, P.A., Justines, G. and Johnson, K.M. (1975). Infection of wild and laboratory animals with Machupo and Latino viruses. *Bulletin WHO*. **52**, 493-499.
807. Weiss, B.G. and Schlesinger, S. (1991). Recombination between Sindbis virus RNAs. *Journal of Virology*. **65**, 4017-4025.
808. Weissenbacher, M.C., de Guerrero, L.B. and Boxaca, M.C. (1975). Experimental biology and pathogenesis of Junin virus infection in animals and man. *Bulletin WHO*. **52**, 507-515.
809. Weissenbacher, M.C. and Damonte, E.B. (1983). Fiebre hemorrágica argentina. In: Adel. Microbiol. Enf. Infec. (Coto, C., Esparza, J. and de Torres, R.A. Eds.). Buenos Aires-Caracas. Vol. **2**, pp: 119-171.
810. Weissenbacher, M.C., Caiello, M.A., Carballal, G., Planes, N., de la Vega, M.T. and Kravetz, F. (1985). Actividad del virus Junin en humanos y roedores de áreas no endémicas de la Provincia de Buenos Aires. *Medicina (Buenos Aires)*. **45**, 263-268.
811. Weissenbacher, M.C., Laguens, R.P. and Coto, C.E. (1987). Argentine hemorrhagic fever. In: *Current Topics in Microbiology and Immunology*. Vol. 134. (M. Oldstone, ed.). Springer-Verlag, Berlin. pp.: 80-116.
812. Welsh, R.M. and Pfau, C.J. (1972). Determinants of lymphocytic choriomeningitis virus interference. *Journal of General Virology*. **14**, 177-183.
813. Welsh, R.M., O'Connell, C.M. and Pfau, C.J. (1972). Properties of defective Lymphocytic Choriomeningitis virus. *Journal of General Virology*. **17**, 355.
814. Welsh, R.M., Burner, P.A., Holland, J.J., Oldstone, M.B.A., Thompson, H.A. and Villareal, L.P. (1975). A comparison of biochemical and biological properties of standard and defective lymphocytic choriomeningitis virus. *Bulletin WHO*. **52**, 403-408.
815. Welsh, R.M. and Oldstone, M.B.A. (1977). Inhibition of immunologic injury of cultured cells infected with lymphocytic choriomeningitis virus: role of defective interfering virus in regulating viral antigen expression. *Journal of Experimental Medicine*. **145**, 1449-1468.
816. Welsh, R.M. and Buchmeier, M.J. (1979). Protein analysis of defective interfering lymphocytic choriomeningitis virus and persistently infected cells. *Virology*. **96**, 503-515.
817. Wengler, G. and Wengler, G. (1984). Identification of transfer of viral core protein to cellular ribosomes during the early stages of Alphavirus infection. *Virology*. **134**, 432-445.
818. Wengler, G., Gros, C. and Wengler, G. (1996). Analyses of the role of structural changes in the regulation of uncoating and assembly of Alphavirus cores. *Virology*. **222**, 123-132.

819. Whelan, A.P.J. and Wertz, G.W. (1999). Regulation of RNA synthesis by the genomic termini of Vesicular Stomatitis Virus: identification of distinct sequences essential for transcription but not replication. *Journal of Virology*. **73**, 297-306.
820. Whitaker-Dowling, P. and Youngner, J.S. (1987). Viral interference-dominance of mutant viruses over wild-type virus in mixed infections. *Microbiol. Rev.* **51**, 179-191.
821. White, J. (1992). Membrane fusion. *Science*. **258**, 917-924.
822. White, K.A. and Morris, T.J. (1994). Recombination between defective tombusvirus RNAs generates functional hybrid genomes. *Proceedings of National Academy of Sciences USA*. **91**, 3642-3646.
823. Whitton, J.L., Gebhardt, J.R., Lewicki, H., Tishon, A. and Oldstone, M.B.A. (1988a). Molecular definition of a major Cytotoxic T-lymphocyte epitope in the glycoprotein of lymphocytic choriomeningitis virus. *Journal of Virology*. **62**, 687-695.
824. Whitton, J.L., Southern, P.J. and Oldstone, M.B.A. (1988b). Analyses of the Cytotoxic T-lymphocyte responses to glycoprotein and nucleoprotein components of lymphocytic choriomeningitis virus. *Virology*. **162**, 321-327.
825. Wicks, R.J. (1986). RNA molecular weight determination by agarose gel electrophoresis using formaldehyde as denaturant: comparison of RNA and DNA molecular weight markers. *International Journal of Biochemistry*. **18**, 277-278.
826. Wieslander, L. (1979). A simple method to recover intact high molecular weight RNA and DNA after electrophoretic separation in low gelling temperature agarose gels. *Analytical Biochemistry*. **98**, 305-309.
827. Wilson, S.M. and Clegg, J.C.S. (1991). Sequence analysis of the S RNA of the African arenavirus Mopeia: an unusual secondary structure feature in the intergenic region. *Virology*. **180**, 543-552.
828. Wilmanns, M. and Eisenberg, D. (1995). Inverse protein folding by the residue pair preference profile method: estimating the correctness of alignments of structurally compatible sequences. *Protein Engineering*. **8**, 627-639.
829. Wodak, S.J. and Rooman, M.J. (1993). Generating and testing protein folds. *Current Opinion in Structural Biology*. **3**, 247-259.
830. Wootton, J.C. and Federhen, S. (1993). Statistics of local complexity in amino acid sequences and sequence databases. *Computers in Chemistry*. **17**, 149-163.
831. Wreschner, D., Melloul, D. and Herzberg, M. (1978). Interaction between membrane functions and protein synthesis in reticulocytes. Isolation of RNase M, a membrane component inhibiting protein synthesis through specific endonucleolytic activity. *European Journal of Biochemistry*. **89**, 341-352.
832. Wright, K.E., Salvato, M.S. and Buchmeier, M.J. (1989). Neutralizing epitopes of lymphocytic choriomeningitis virus are conformational and requires both glycosylation and disulphide bonds for expression. *Virology*. **171**, 417-426.
833. Wright, K.E., Spiro, R.C., Burns, J.W. and Buchmeier, J.M. (1990). Post-translational processing of the glycoproteins of lymphocytic choriomeningitis virus. *Virology*. **177**, 175-183.
834. Wu, X. and Shaw, J. (1996). Bidirectional uncoating of the genome RNA of a helical virus. *Proceedings of National Academy of Sciences USA*. **93**, 2981-2984.
835. Wu, J.-C., Chiang, T.-Y., Shiue, W.-K., Wang, S.-Y., Sheen, I.-J., Huang, Y.-H. and Syu, W.-Jr. (1999). Recombination of hepatitis D virus RNA sequences and its implications. *Molecular Biology Evolution*. **16**, 1622-1632.
836. Yamashita, M., Krystal, M. and Palese, P. (1989). Comparison of the three large polymerase proteins of Influenza A, B, and C viruses. *Virology*. **171**, 458-466.
837. Yanish Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strain: nucleotide sequence of M13 mp18 and pUC19 vectors. *Gene*. **33**, 103-119.
838. Yasumura, Y. and Kawatika, Y. (1963). Studies on SV40 virus in tissue culture cells. *Nippon Rinsho* **21**, 1201-1215.
839. Yeh, T.S., Lo, S.J., Chen, P.J. and Lee, Y.H. (1996). Casein kinase II and protein kinase C modulate hepatitis delta virus RNA replication but not empty viral particle assembly. *Journal of Virology*. **70**, 6190-6198.
840. Yeagle, P.L. (1993). Membrane fusion. In: *The membranes of the cells*. (Yeagle, P.L. Ed.). Academic Press, New York. pp. 270-274.
841. Young, P.R., Chanas, A.C. and Howard, C.R. (1981). Analysis of the structure and function of Pichinde virus polypeptides. In: *The replication of negative strand viruses*. (Bishop, D.H.L. and Compans, R.W. Eds.). Elsevier Press, New York. pp. 11-14.
842. Young, P.R. and Howard, C. (1983). Fine structure analysis of Pichinde virus nucleocapsids. *Journal of General Virology*. **64**, 833-842.
843. Young, P.R., Chanas, A.C., Lee, S.R. and Howard, C.R. (1987). Localization of an arenavirus protein in the nuclei of infected cells. *Journal of General Virology*. **68**, 2465-2470.
844. Yura, Y., Kusaka, J., Tsujimoto, H., Yoshioka, Y., Yoshida, H. and Sato, M. (1997). Effects of protein tyrosine kinase inhibitors on the replication of herpes simplex virus and the phosphorylation of viral proteins. *Intervirology*. **40**, 7-14.
845. Yusoff, K., Millar, N.S., Chambers, P. and Emmerson, P.T. (1987). Nucleotide sequence analysis of the L gene of Newcastle disease virus: Homologies with Sendai and vesicular stomatitis viruses. *Nucleic Acids Research*. **15**, 3961-3976.
846. Zagursky, R.J., Baumeister, K., Lomax, N. and Berman, M.L. (1985). Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. *Gene Analytical Techniques*. **2**, 89-94.
847. Zeller, W., Bruns, M. and Lehmann-Grube, F. (1986). Viral nucleoprotein can be demonstrated on the surface of lymphocytic choriomeningitis virus-infected cells. *Medical Microbiology and Immunology*. **175**, 89-92.
848. Zeller, W., Bruns, M. and Lehmann-Grube, F. (1988). Lymphocytic choriomeningitis virus. X. Demonstration of the nucleoprotein on the surface of infected cells. *Virology*. **162**, 90-97.
849. Zhang, Z., Schäffer, A.A., Miller, W., Madden, T.L., Lipman, D.J., Koonin, E.V. and Altschul, S.F. (1998). Protein sequence similarity searches using patterns as seeds. *Nucleic Acids Research*. **26**, 3986-3990.
850. Zhang, L., Marriott, K. and Aronson, J.F. (1999). Sequence analysis of the small RNA segment of guinea pig-passaged Pichinde virus variants. *American Journal of Tropical Medicine and Hygiene*. **61**, 220-225.
851. Zhang, J. and Sapp, C.M. (1999). Recombination between two identical sequences within the same retroviral RNA molecule. *Journal of Virology*. **73**, 5912-5917.
852. Zheng, D.X., Dickens, L., Liu, T.Y. and Nakhasi, H.L. (1989). Nucleotide sequence of the 24S subgenomic messenger RNA of a vaccine strain (HPV77) of rubella virus: comparison with a wild-type strain (M33). *Gene*. **82**, 343-249.
853. Zimmermann, K. (1994). When awaiting 'bio' champollion: dynamic programming regularization of the protein secondary structure predictions. *Protein Engineering*. **7**, 1197-1202.
854. Zinkernagel, R.M. and Doherty, P.C. (1974). Restriction of *in vitro* T cell-mediated cytotoxicity in lymphocytic choriomeningitis virus within a synergic or semiallogenic system. *Nature (Londres)*. **248**, 701-702.

Referencias bibliográficas

855. Zinkernagel, R.M. and Doherty, P.C. (1977). Major transplantation antigens, viruses and specificity of surveillance T cells. *Cont. Topics in Immunobiology*. 7, 179-220.
856. Zinkernagel, R.M., First, T., Hengartner, H. and Althage, A. (1982). Susceptibility to lymphocytic choriomeningitis virus isolates correlates directly with early and high cytotoxic T-cell activity, as well as footpad swelling reaction, and all these are regulated by H-2D. *Journal of Experimental Medicine*. 162, 2125-2141.
857. Zlatkine, P., Mehul, B. and Magee, A.I. (1997). Retargeting of cytosolic proteins to the membrane by the Lck protein tyrosine kinase dual acylation motif. *Journal of Cellular Sciences*. 110, 673-679.
858. Zuckerkandl, E. and Pauling, L. (1965). In *Evolving Genes and Proteins* (Bryson, V. and Vogel, H. J., eds.), *Evolving Genes and Proteins*. Academic Press, New York and London, pp. 97-166.
859. Zuker, M. and Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*. 9, 133-148.
860. Zuker, M. and Sankoff, D. (1984). RNA secondary structures and their prediction. *Bulletin of Mathematical Biology*. 46, 591-621.
861. Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science*. 244, 48-52.
862. Zuker, M.; Jaeger, J.A. and Turner, D.H. (1991). A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. *Nucleic Acids Research*. 19, 2707-2714.
863. Zuker, M. (1994). Prediction of RNA secondary structure by energy minimization. In: *Computer analysis of sequence data*, vol. 25, part II. (A.M. Griffin and H.G. Griffin, eds.). CRC Press, New Jersey, chapter 23, pp. 267-294.
864. Zvelebil, M.J., Barton, G.J., Taylor, W.R. and Sternberg, M.J. (1987). Prediction of protein secondary structure and active sites using the alignment of homologous sequences. *Journal of Molecular Biology*. 195, 957-961.

1000

PARTE IX

APENDICES

Alineamiento múltiple de los RNAs S de arenavirus

							60
Junín MC2	-UGCAGUAAG	GGGAUCCUAG	GC-GAUUUU-	GG--UAACGC	-----	-----	
Junín CD1	CUGCAGUAAG	GGGAUCCUAG	GC-GAUUUU-	GG--UAACGC	-----	-----	
Oliveros	-CGCA--CCG	GGGAUCCUAG	GC---UUUUA	A--CUAACGC	AUUACAGAGG	GAUUCAACCC	
Pichindé	-CGCA--CCG	GGGAUCCUAG	GCAUACCUU-	GG----ACGC	-----	-----	
Sabiá	GCGCA--CCG	GGGAUCCUAG	GCGUUUUUUA	G--UC-ACGC	-UUA-----	-----	
Tacaribe	C-GCA--CCG	GGGAUCCUAG	GCAUUUCUU-	G--UC-----	-----	-----	
Lassa Jos.	GCGCA--CCG	GGGAUCCUAG	GC-AUUUUU-	GG-UU-GCGC	-----	-----	
Lassa Nig.	--GCA--CCG	GGGAUCCUAG	GC-A--UUUA	GGAUU-GCGC	-----	-----	
LCM Arm.	-CGCA--CCG	GGGAUCCUAG	GC--UUUUU-	GGAUU-GCGC	-----	-----	
LCM WE	-CGCA--CCG	GGGAUCCUAG	GC--UUUUU-	GGAUU-GCGC	-----	-----	
Mopeia	-----GU---	-GGAUCCUAG	GC--UUUUU-	GG-UU-GCGC	-----	-----	
Cons. Est.GGAUCCUAG	GC.....UU.	
ConsensoGGAUCCUAG	GC....yUU.	r.....	

							120
Junín MC2	UAUA-AGUU-	---GUUACUG	CUUUCUAUUU	GGACAACAUC	AAACCAUCUA	UU-----GU	
Junín CD1	UGUA-AGUU-	---GUUACUG	CUUUCUAUUU	GGACAACAUC	AAACCAUCUA	UU-----GU	
Oliveros	UGUGCAAUUC	AAAGUCACAG	UUUUUGUGAC	AGUUAGUUGU	CAACCCCUU	UUUGACAACC	
Pichindé	-GCAUAUUUC	UUGAUCAAA-	-----	-----	-----	-----	
Sabiá	-----AAU--	-----	CUUU-----	-GAUUGCGUC	AA-----	-----	
Tacaribe	--CAU-AUU-	-----UG	CCU---AACU	GAACCAGGU-	GAAUCAC---	UC-----CC	
Lassa Jos.	-----AAUUC	-----AA--	-----	-----	-----GUGUC	CUAUUUAA--	
Lassa Nig.	-----AUUUU	-----AAAA	CCUCCUUUUU	GG-----	AAA--GUGUC	GCAAUCAG--	
LCM Arm.	-----UUUC	C---UCUAGA	UC---AACU	GG---GUGUC	AGGCCCUAUC	CUACAGAAG-	
LCM WE	-----UUUC	C---UUUAGG	AC---AACU	GG---GUGCU	GGAUUCUAUC	CAGUAAAAG-	
Mopeia	-----AUUC	-----	-----	-----U	AGA--GCAUC	UCG---GA--	
Cons. Est.U..	
ConsensoU..	

							180
Junín MC2	ACA	GGGC	AAUUCAUCAG	CUUCAUGCAA	GAAAUACCUA	CCUUUUUGCA	GGAAGCUCUG
Junín CD1	ACA	GGGC	AAUUCAUCAG	CUUCAUGCAA	GAAAUACCUA	CCUUUUUGCA	GGAAGCUCUG
Oliveros	AUU	GGUC	AAGUCAUUGG	GUUCUCCAA	UCCUUAACCA	ACAUCAUUA	UGAGGCACUA
Pichindé	--G	GGAC	AAAUUGUGAC	UUUGAUCCAG	UCCAUUCCUG	AAGUCCUGCA	AGAGGUCUUC
Sabiá	-UC	GGUC	AAUUGUUCAG	CUUUUUUGAA	GAAGUCCGA	AUAUCAUCCA	UGAGGCUAUC
Tacaribe	ACC	GGGC	AAUUCAUAG	UUUCAUGCAG	GAGAUACCUA	UUUUCUUGCA	AGAAGCACUC
Lassa Jos.	--A	GGAC	AAUAGUGAC	AUUCUCCAG	GAAGUGCCUC	AUGUAAUAGA	AGAGGUGAUG
Lassa Nig.	--G	GGAC	AGAUUGUGAC	AUUCUCCAA	GAAGUCCUC	AUGUUAUUGA	GGAAGUGAUG
LCM Arm.	--G	GGUC	AGAUUGUGAC	AAUGUUUGAG	GCUCUGCCUC	ACAUCAUCGA	UGAGGUGAUC
LCM WE	--G	GGUC	AGAUUGUGAC	AAUGUUUGAG	GCUUUGCCUC	ACAUCAUUGA	UGAGGUGAUC
Mopeia	--G	GGGC	AGAUAGUCAC	CUUCUUUCA	GAGGUGCCAC	ACAUCUUGA	AGAAGUGAUG
Cons. Est.	...AUGGG.C	A..U..U...	..U..U..A	G...U.CC..	...U..U..A	.GA.G...U.	
Consenso	...AUGGG.C	Ar.U..U.r.	..U..U..Ar	G...U.CC..	...U..U..A	.GA.rGy..U.	

							240
Junín MC2	AAUAUUGCUC	UUGUUGCAGU	CAGUCUCAUU	GCCAUCAUUA	AGGGUGUAGU	AAACCGUAC	
Junín CD1	AAUAUUGCUC	UUGUUGCAGU	CAGUCUCAUU	GCCAUCAUUA	AGGGUGUAGU	AAACCGUAC	
Oliveros	AACAUUGCCC	UCAUCUGUGU	GGCAUUAAUU	GCCAUUAUGA	AAGGUUAUGU	UAACAUUUGG	
Pichindé	AAUGUUGCUC	UGAUCAUUGU	CUCAGUUCUA	UGCAUUGUCA	AAGGAUUUGU	CAAUCUGAUG	
Sabiá	AACAUAGCUC	UGAUAGCAGU	GAGCUAAUU	GCUGCCUUGA	AAGGAUGAU	UAACUUGUGG	
Tacaribe	AACAUUGCAU	UGGUGGCAGU	CAGCCUCAUC	UGUAUAGUGA	AAGGCUUGGU	GAACCUAUAC	
Lassa Jos.	AACAUUGUUC	UCAUUGCACU	GUCUGUACUA	GCAGUGCUGA	AAGGUCUGUA	CAAUUUUGCA	
Lassa Nig.	AAUAUUGUCC	UUAUUGCACU	AUCCAUCCUA	GCAAUUCUGA	AGGGACUAUA	CAAUGUUGCC	
LCM Arm.	AACAUUGUCA	UUAUUGUGCU	UAUCGUGAUC	ACGGGUAUCA	AGGCUGUCUA	CAAUUUUGCC	
LCM WE	AACAUUGUCA	UUAUUGUGCU	CAUUAUAUC	ACGAGCAUCA	AAGCUGUGUA	CAAUUUCGCC	
Mopeia	AACAUUGUGC	UGAUGACCCU	CUCAAUCUUG	GCCAUCCUAA	AGGGCAUCUA	CAAUGUGAUG	
Cons. Est.	AA..U.G...	U..U.....UU....A	A.G...U...	.AA.....	
Consenso	AAyrU.Gy..	U.rU.....UU..U.	...r...U.A	ArG...U...	.AAy.U....	

Apéndice 1

300

Junín MC2	AAAAGUGGUU	-GUUC-AAU	CUUGAU----	CUAGCA-----	--GGAAGA--	UCCUGCCCG-
Junín CD1	AAAAGUGGUU	UGUUCCAAU	CUUUGUAUUC	CUAGCACUCG	CAGGAAGA--	UCCUGCACA-
Oliveros	AAAUCAGGUC	UGAUUCAGCU	GUUCAUAUUC	CUCAUCCUAG	CGGGGAGGAG	CUGUUCUCAC
Pichindé	AGAUGUGGCC	UUUUCAGCU	CGUCACCUUC	CUCAUUUUGU	CUGGCAGAAG	UUGUG---AC
Sabiá	AAGAGUGGCC	UUUUCCAAU	GAUAUUCUU	UUGACUCUAG	CAGGAAGA--	UCGUGUUC--
Tacaribe	AGGUGCGGAC	UGUUUCAAU	GAUGGUCUUC	CUUGUGUUGG	CUGGCAGA--	UCAUGUUCA-
Lassa Jos.	ACGUGUGGCC	UUGUUGGUU	GGUCACUUC	CUCCUGUUGU	GUGGUAG--G	UCUUG-----
Lassa Nig.	ACGUGUGGCU	UGAUAGGGCU	UGUCACAUUC	CUUCUACUUU	CAGGAAG--G	UCAUG-----
LCM Arm.	ACCUGUGGGA	UAUUCGCAU	GAUCAGUUC	CUACUUCUGG	CUGGCAG--G	UCCUG-----
LCM WE	ACCUGUGGGA	UAUUCGCAU	GGUCAGCUUC	CUUUUUCUGG	CUGGUAG--G	UCCUG-----
Mopeia	ACCUGUGGAA	UCAUCGGUU	GAUAACAUU	UUGUUUUUGU	GUGGGAG--A	UCAUG-----
Cons. Est.	A...G.GG..	...U.....UU	.U.....	..GG.AG...
Consenso	A...G.GG..	...U....yU	..U.....	yU..y.....	..GG.AG...	yy.U.....

360

Junín MC2	-----	-----AGAG	-----CU-	-----UUUAA	AAUCGGACUG	CACACAGAGG
Junín CD1	-----	-----GAAG	-----AAGCU-	-----UUUAA	AAUCGGACUG	CACACAGAG-
Oliveros	ACCUUUCAGA	UUGGCCGAAA	CCAUGAGUUC	CAAAGUAUCA	CUCUAAACUU	CACACAGUUU
Pichindé	AGCAUGAUGA	UUGAUAGAAG	GCACAAUCUC	ACCCAUGUUG	AAUUCAAUCU	CACACGAAUG
Sabiá	-----	-----	-----U-	-----UUUAG	AAUUGGAAGG	AGCACAGAA-
Tacaribe	-----	-----GAGG	-----AAACA-	-----UUUAA	GAUAGGUAUG	CACACAAAA-
Lassa Jos.	-----	-----	-----CAC	AA-----	--CCAGUCUU	UAUAAAGGGG
Lassa Nig.	-----	-----	-----CUC	A-----	--CUGAUC--	UACAAAGGGA
LCM Arm.	-----	-UGGC--AU	GUACGGUCUU	AA-----GGG	ACCCGACAUU	UACAAAGGAG
LCM WE	-----	-UGGC--AU	GUACGGCCUU	AA-----UGG	UCCCGAUUUC	UAUAAAGGGG
Mopeia	-----	-----	-----CUC	AA-----	--GCA--UC	UAUAAGGACA
Cons. Est.A.....
ConsensoryA.r.r...

420

Junín MC2	UUCCAGACUG	UGUCCUUCUU	CAAUGGUG--	GGUCUCUUUU	UCCAACAAUC	CACAUGACCU
Junín CD1	UUCCAGACUG	UGUCCUUCU-	CAAUGGUG--	GGUCUCUUUU	-CCAACAAUC	CACAUGACCU
Oliveros	UUGGGCUACG	CACCCAGCUC	CUGUUCUGUU	AACAACACUC	AUCACUACUU	UAGAGGUCCU
Pichindé	UUUGACAAUC	UGCCUCAUUC	AUGCAGCAAG	AACAACACAC	ACCAUUACUA	UAAAGGACCA
Sabiá	UUGCAAAACA	UAACGUU-UG	AUAUGUUG--	AAGGUUAUCG	AGGACCACC	CACAU--CUU
Tacaribe	UUCCAAGAGG	UGUCCUCUC	UUUAAGUG--	CACUU--UUG	ACCAUUCAGU	CUCAUGAACU
Lassa Jos.	UUUAUGA--G	CUUCAGACUC	UGGAACUAAA	CAUGGAGACA	CUCAUAUGA	C-CAUGCCUC
Lassa Nig.	CUUACGA--G	CUGCAAACCC	UUGAGUUGAA	UAUGGAGACU	CUUAAUAUGA	C-CAUGCCGC
LCM Arm.	UUUACCA--A	UUUAAGUCAG	UGGAGUUUGA	UAUGUCACAU	CUGAACCUGA	C-CAUGCCCA
LCM WE	UUUACCA--G	UUCAAAUCAG	UGGAGUUUGA	UAUGUCUCAC	UUAAAUCUGA	C-GAUGCCCA
Mopeia	ACUAUGA--G	UUCUUCUCUC	UCGACCUCGA	CAUGUCUUCA	CUGAAUGCAA	C-GAUGCCUC
Cons. Est.A.....	...A.....
Consenso	.y.....	y.....y...	.r.....	...A.....	y..A.....

480

Junín MC2	GCCUCUGU--	--UGUG----	-----UACCU	UAAACAAGAG	CCAUCUUUAC	-AUUAAGGGG
Junín CD1	GCCUCUGU--	--UGUG----	-----UACCU	UAAACAAGAG	CCAUCUUUAC	-AUUAAGGGG
Oliveros	GGCAAUGUCU	CUUGGGGCAU	CGAGCUCACA	UUGACCAACA	ACUCCGUGAU	CAAUGCAA--
Pichindé	UCCAACACGA	CAUGGGGAAU	UGAACUGACC	UUGACAAACA	CAUCUAUUGC	UAAUGAAACU
Sabiá	GC--AUGG--	--UGAA----	-----UCAUU	CCACCUACUA	UGUCCAUGAA	-AACAAAAAU
Tacaribe	UCCUAUGU--	--UAUG----	-----UUUAG	CCAACAAAAC	ACAUCUCUAU	-CUAAAGAGU
Lassa Jos.	UCUCCUG---	CACAAAGAAC	--AACAGUCA	UCAUUA----	-----UAU	AAUGGUGGGC
Lassa Nig.	UAUCCUG---	CACAAAGAAC	--AACAGUCA	UCAUUA----	-----UAU	AAGGGUGGGG
LCM Arm.	ACGCAUG---	UUCAGCCAAC	--AACUCCA	CCAUUA----	-----CAU	CAGUAUGGGG
LCM WE	AUGCGUG---	CUCAGUCAAC	--AACUCUCA	UCACUA----	-----CAU	CAGUAUGGGG
Mopeia	UCUCCUG---	CUCAAAGAAC	--AACUCCA	UCACUA----	-----CAU	CCAAGUUGGG
Cons. Est.
Consensoyr...	..yr.....	y.r.y.....r.r..

Apéndice 1

						540
Junín MC2	GGCAAUGCUU	CAUUCAAGAU	CAGCUUUGAU	G-ACAUCGCA	GUGUUGUUAU	CAGAAUUAU--
Junín CD1	GGCAAUGCUU	CAUUCAAGAU	CAGCUUUGAU	G-ACAUCGCA	GUGUUGUUAU	CAGAAUUAU--
Oliveros	----GCAACU	CACUGAAGGU	GUUCACCAAC	AUUCACCACA	ACAUCACCAA	CUGCGUUCAG
Pichindé	AGUGGGAACU	UCUCCAACAU	CGGGAGCCUU	GGAUUAUGGCA	AUAUUAGCAA	UUGUGA-CAG
Sabiá	GCCACUUGGU	GUCUUGAGGU	GUCUGUGACU	G-AUGUUACC	CUGCUCAUGG	CUGAACAU--
Tacaribe	GGUAGAUGCA	GUUUCAAAAU	CAACAUUGAU	A-GUGUCACU	GUGUUAACAA	GAUCCGCC--
Lassa Jos.	AAUGAGACAG	GACUAGA-AC	UGACCUUGAC	CAACAC--GA	GCAUUUAUAA	UCACAAAUUU
Lassa Nig.	AAUGAGACAG	GACUUGA-GC	UCACCUUGAC	CAACAC--CA	GCAUUCUAAA	UCACAAAUUC
LCM Arm.	ACU---UCUG	GACUAGA-AU	UGACCUUCAC	CAAUGA--UU	CCAUCAUCAG	UCACAACUUU
LCM WE	AGC---UCUG	GACUGGA-GC	CAACUUUCAC	CAACGA--CU	CCAUCCUAAA	UCACAACUUC
Mopeia	AAUGAGACAG	GCCUAGA-GC	UGACAUUGAC	AAACAC--UA	GCAUAAUAAA	UCAUAAGUUU
Cons. Est.
Consensoyy.r.r.ryy	...yr.....	.yry...yr.y..

						600
Junín MC2	GACGUUAUAA	U--UCAGCA-	UCCGGCAGA-	UA-----UGA	GC-UGGUGUU	-CUAAAAGUG
Junín CD1	GACGUUAUAA	U--UCAGCA-	UCCGGCAGA-	UA-----UGA	GC-UGGUGUU	-CUAAAAGUG
Oliveros	AAUAUCGAUG	AGCAAGACCA	CCUAAUGAAA	UGGCUCAUAG	AAACAAUGCA	UCUACAGAUC
Pichindé	GACAA-GGGA	AGCAGGAC-A	UACAUUGAAA	UGGUUACUCA	AUGAACUACA	CUU-CAA--U
Sabiá	GAUCGUCAAG	UCCUCAACAA	UCUGUCAAAAC	UG-----UGU	GCACCCUGCA	GUUGAGCACA
Tacaribe	GAUGUCUUUG	U--UCACUC-	UCC--UAAA-	CU-----UG-	GU-UCUUGUU	-UUGAAAGUG
Lassa Jos.	UGCAAUCUGU	CUGAUGCC--	CACAAAAGA	ACCUCUAUG-	--ACCACGCU	CUUAUGAGC-
Lassa Nig.	UGUAACCUCU	CUGAUGCC--	CACAAAAGGA	AUCUUUAUG-	--AUCACAGC	CUCAUGAGU-
LCM Arm.	UGCAAUCUGA	CCUCUGCC--	UUCAACAAAA	AGACCUUUG-	--ACCACACA	CUCAUGAGU-
LCM WE	UGCAACUAAA	CCUCCGCU--	CUCAACAAAA	AGUCUUUUG-	--ACCAUACA	CUCAUGAGU-
Mopeia	UGCAACCUGU	CUGAUGCU--	CACAGGAGGA	AUCUUUAUG-	--ACAAAGCA	CUUAUGUCA-
Cons. Est.U..
Consenso	.ry.....r.y..	y.y...rrr.U..r..	.yy.r....

						660
Junín MC2	AUG--AUCAA	AU---UAGGC	-UGUCUCAGU	--GGUUCAUG	AAUGCUGUGG	GGCA--UGAU
Junín CD1	AUG--AUCAA	AU---UUGGC	-UGUCUCAGU	--GGUUCAUG	AAUGCUGUGG	GGCA--UGAU
Oliveros	AUGAAACCAG	GAAAGAGACU	AC-CACCAAU	CCUGUGUGAG	AAAGACAAAG	GGUUACUAAU
Pichindé	GUGCU-CCAU	GUGACUCGUC	AUAUAGGUGC	CAGAUGUAAG	ACAGUUGAAG	G--UGCUGG-
Sabiá	GAA--GCAGG	AUGGUUGGCU	-UACUUGAGU	--GGAUUUUU	AGAGCCCUAA	AGUA--UGAC
Tacaribe	AUG--A--GG	AG---UGGGU	-UGUUGCU-U	--GGUGGAUU	GAAGCCAUCG	GUCA--CAGA
Lassa Jos.	AUAAUCUCAA	CUUUCACUU	GUCCAU---C	CCCAACUUCA	AUCAGUAUGA	GGCAAUGAGC
Lassa Nig.	AUCAUCUCUA	CCUUUCAUCU	GUCCAU---U	CCCAACUUCA	AUCAUAUGA	GGCAAUGAGC
LCM Arm.	AUAGUUUCGA	GCCUACACCU	CAGUAU---C	AGAGGGAAACU	CCAACUAUAA	GGCAGUAUCC
LCM WE	AUAGUCUCGA	GUCUACACCU	CAGUAU---C	AGAGGGAAAUU	CCAACUACAA	AGCAGUGUCU
Mopeia	AUCUUGACAA	CAUUCACCU	AAGCAU---U	CCAGAUUUUA	ACCAUAUGA	GGCAAUGUCA
Cons. Est.
Consenso	r.....y	...y.....y	...r.....	...r.y...r	r.....

						720
Junín MC2	UG-GUAUCUA	---GACCCAC	CAUU----UC	UGUGUAGGAA	CCGUACAAAG	ACAG--AAGG
Junín CD1	UG-GUAUCUA	---GACCCAC	CAUU----UC	UGUGUAGGAA	CCGUACAAAG	ACAG--AAGG
Oliveros	UG-AAUACAA	CCUAAC-CAA	CAUUGCAUCU	CGCGAAGAAA	AGCACUCUGA	GUACUGGAGC
Pichindé	-G-GUGUAAA	UCCAAUACAA	CUUGACAGUU	GGGGACAGAG	GAGG--UGA	GGUUGGCAGA
Sabiá	UU-CAAUCAU	---GAUCCAA	CACC----GU	UGUGUCAAAA	GCAAACUUCA	ACAGUGAAUG
Tacaribe	UG-GGACCAG	---GACCCAG	GUCU----UC	UAUGCAGGAA	UAAAACCAA	ACAG--AGGG
Lassa Jos.	UGCGAUUUUA	AUGGGGGGAA	GAUUAGUGUG	CA-GUACAAC	CUGAGUCACA	GCUAUGCUGG
Lassa Nig.	UGCGAUUUUA	AUGGGGGGAA	AAUCACUGUG	CA-AUACAAC	CUGAGUCAUA	GCUUCGCAGU
LCM Arm.	UGCGACUUA	ACAAUGGCA-	--UAACCAUC	CA-AUACAAC	UUGACAUUCU	CAGAUCGACA
LCM WE	UGUGAUUUUA	ACAAUGGCA-	--UCACCAU	CA-AUACAAC	UUGUCAUCUU	CGGACCCACA
Mopeia	UGUGAUUUUA	ACGGAGGGAA	GAUCUCAGUU	CA-AUACAAC	CUGUCCACU	CAAUUUAUGU
Cons. Est.A.A.
Consensoy..	...r....A.	..y.....	.r.r...rA.

12/11/13
56986

Apéndice 1

780

Junín MC2	CUUCAUCU--	UUCAA---GU	CAAUACCUCC	AAGA--CU--	GG--UAUCA	-----UGAAA
Junín CD1	CUUCAUCU--	UUCAA---GU	CAAUACCUCC	AAGA--CU--	GG--UAUCA	-----UGAAA
Oliveros	CAACUUCUC-	UAUGGUCUAU	CAAAACUAUU	AGGCUCCAGU	AAGUCAUUGU	GGUUUGAUUA
Pichindé	CAUCUUAUUG	CAUCACUCGC	UCAAAUCAUC	GGGGAUCCUA	AAAUUCGC-AU	GGGUUGGGAA
Sabiá	AAACACGU-G	UGCAG---AU	AAACAUCACU	GAGGGGUUUG	GG--UCUCAC	GGGUUUGAAG
Tacaribe	UAAACUGA--	UCCAG---AU	UAACAUUUCA	AGAG--CU--	GA--UGGCAA	-----UGUUC
Lassa Jos.	GGAUGCAGCC	AACCAUU-GU	-GGUACUGUU	GCAAA--UGG	UGUGUUACAG	ACUUUUUAUGA
Lassa Nig.	GGAUGCAGCA	GGUCACU-GC	-GGCACACUU	GCAAA--UGG	UGUCUUACAA	ACAUUUUAUGA
LCM Arm.	AAGUGCUCAG	AGCCAGU-GU	-AGAACCUUC	AGAGG--UAG	AGUCCUAGAU	AUGUUUAGAA
LCM WE	GAGCGCCAUG	AGCCAGU-GU	-AGGACUUUC	AGAGG--UAG	AGUCUUGGAC	AUGUUUAGAA
Mopeia	GGAUGCCGGC	AACCACU-GU	-GGCACUAUU	GCAAA--UGG	CAUUAUGGAU	GUUUUUAGAA
Cons. Est. A
Consenso y y.r ry Ayy.y r.r r r

840

Junín MC2	A--CUAUGCC	AAGAA-GUUU	AAGACUGGUA	UGCACCAUUU	AUAUAGAGAA	UACCCCGACU
Junín CD1	A--CUAUGCC	AAGAA-GUUU	AAGACUGGUA	UGCACCAUUU	AUAUAGAGAA	UACCCCGACU
Oliveros	CUGUCAGAGA	GCAGA-UUGC	AUGAUGCAAG	AGCACAGCUC	ACACUUGAAG	UGCAACUACU
Pichindé	GUGCUUCA--	-CA-A-CUGC	A----GUGGA	GAUACU-UGC	AGACU-GACA	---AACUG-U
Sabiá	AUACCAUCCU	UCAA--GACU	AGGGGUUCUA	UUCGGUUCAA	GAAUUGCAUU	UUCAAAUAUC
Tacaribe	A--CUAUGGA	UGGAG-AUUA	AAAAAUGGAC	UUGAUCAUUA	UUACAGAGGA	CGUGAAGAAC
Lassa Jos.	-----	--GGAUGGCU	UGGGGUGGGA	GCUACAUUGC	UCUUGACUCA	GGCCGUGGCA
Lassa Nig.	-----	--GAAUGGCU	UGGGGAGGGA	GUUAUUAUUG	UCUUGAUUCU	GGACGCGGUA
LCM Arm.	CUGCCUUCGG	GGGGAUUAC	AUGAGGAGUG	GCUGGGGCUG	GACAGGCUCA	GAUGGCAAGA
LCM WE	CUGCCUUUGG	AGGAAAGUAC	AUGAGAAGUG	GCUGGGGCUG	GACAGGUUCA	GAUGGCAAGA
Mopeia	-----	--GAAUGUAU	UGGAGC---A	CCUCCCUUUC	GGUUGCCUCU	GACAUAAUGU
Cons. Est.
Consenso r.r y r

900

Junín MC2	CUUGCU-UGG	AUGGCAAACU	GUGUUUGA--	---UGAAGGC	ACAACCCACC	AGUUGGCCUC
Junín CD1	CUUGCU-UGG	AUGGCAAACU	GUGUUUGA--	---UGAAGGC	ACAACCCACC	AGUUGGCCUC
Oliveros	CAGAGUGUUC	UGGCCAUACC	ACCUUCAAAU	ACCUCUUAUCU	ACAAAACACC	ACUUGGGAGA
Pichindé	GAGGGUG---	-GAACACAUU	AC----AAUU	UCUUAUAUCAU	ACAGAACACC	ACAUGGGAAA
Sabiá	CAGGAU-UUA	GGUAAAAAAA	GGUUUUUA--	---UUGAUUA	GAAAUUCAAC	--UUGGAAAA
Tacaribe	CCUGUU-UUG	AGGGUGAGCA	GUGUCUUA--	---UCAAAAU	UCAACCUGAA	GACUGGCCGA
Lassa Jos.	ACUGGGACUG	UAUUUAUGACU	AGUUAUCAAU	AUCUGAUAAU	CCAAAUAACA	ACCUGGGGAA
Lassa Nig.	ACUGGGACUG	UAUUAUGACC	AGUUACCAAU	AUCUUAUCAU	UCAGAAUACA	ACCUGGGGAA
LCM Arm.	CCACCUGGUG	UAGCCAGACG	AGUUACCAAU	ACCUGAUUAU	ACAAAUAAGA	ACCUGGGGAA
LCM WE	CCACUUGGUG	CAGCCAAACA	AGCUAUCAGU	ACCUAUAUCAU	ACAAAACAGG	ACUUGGGGAA
Mopeia	GGACUCAUUG	CAUACAGACC	GAUUAUAAGU	AUUUGAUUAU	UCAGAACACA	UCAUGGGGAG
Cons. Est. A U A UGG
Consenso r r A U.r Ar.yr UGG

960

Junín MC2	UCCAAUGUCC	ACUUG-A---	CCA-UGUCA	CACAUUACAU	UU-CCUCACA	AGAGGCAA-G
Junín CD1	UCCAAUGUCC	ACUUG-A---	CCA-UGUCA	CACAUUACAU	UU-CCUCACA	AGAGGCAA-G
Oliveros	ACCAUUGUGA	GUUCA-A---	CCAUCUA---	-AACACAAUU	CA-CCUGUUG	AUGAGCUCUA
Pichindé	AUCAUUGCAC	AUACACA---	CCAAUGGCAA	CAAUAAGAAU	GG-CUCUUCA	AAGAACU--G
Sabiá	AUCAUUGCGA	AAUGA-A---	UCA-UGUAAA	CUCCAUGCAU	UU-AAUGUUG	GCGAAUGCUG
Tacaribe	CUGACUGUAA	GGCUG-A---	CCA-CACAAA	CACAUUCCGC	UU-UUUUAUCA	AGGAGCCA-A
Lassa Jos.	AUCACUGCCA	AUUCUCGAGA	GCAUCUCCCA	UGGGUUAUCU	CGGGCUCGUC	UCACAAAGGA
Lassa Nig.	ACCACUGCCA	AUUUUCCAGA	CCAUCACCUA	UUGGCUACCU	UGGGCUUCUC	UCACAAAGAA
LCM Arm.	ACCACUGCAC	AUAUGCAGGU	CCUUU-----	UGGGAUGUCC	AGGAUUCUCC	UUUCCCAAGA
LCM WE	ACCACUGUAG	AUAUGCAGGC	CCUUU-----	UGGGAUGUCU	AGAAUCCUCU	UUGCUCAGGA
Mopeia	AUCAUUGCAU	GUUCUCAAGA	CCCUCACCCA	UGGGAUUUUU	AAGCCUUCUG	UCACAGAGAA
Cons. Est. A.UG C
Consenso y.A.UGy r C.y y y.y r

Apéndice 1

						1020
Junín MC2	AACAUCAGC	U--UCCAAGG	AGGUCUUUAA	AAGCAUUCUU	UUCCUGGUCU	CUGACAGACU
Junín CD1	AACAUCAGC	U--UCCAAGG	AGGUCUUUAA	AAGCAUUCUU	UUCCUGGUCU	CUGACAGACU
Oliveros	CUGGUCAGUC	UUUCAUAACU	AGAAGGUUGC	AGGCUUUUCU	AACUUGGACU	UUGUCAGAUG
Pichindé	CUUAU-AGUU	CUGU-UAGCA	GGAAACUUCU	UGGCUUCUUC	ACUUGGGACU	UGAGUGAUUC
Sabiá	GUCGCUCUUC	UGGUUCUAGA	AGACCACUC-	-GGCAUU-UU	CUCCUGGACA	AUAACUGAUG
Tacaribe	AAAAGCAUAG	C--UGUAGGA	AGGACUCUCA	AGGCAUUUUU	CUCCUGGUCG	CUUACUGACC
Lassa Jos.	CGAGAGAUAU	UUUAUUUAGU	AGAAGAUUGC	UAGGCACAUU	CACAUGGACA	CUGUCAGAUU
Lassa Nig.	CUAGAGACAU	AUACAUUAGC	AGAAGGUUGU	UGGGGACAUU	CACCUGGACA	CUAUCAGACU
LCM Arm.	-GAAGACUAA	GUUCUUCACU	AGGAGACUAG	CGGGCACAUU	CACCUGGACU	UUGUCAGACU
LCM WE	-AAAGACAAA	GUUCUUCACU	AGGAGACUUU	CAGGCACAUU	CACCUGGACC	CUGUCAGACU
Mopeia	CCAGGAAUUU	CUACAUCUCA	AGAAGACUUU	UGGGUCUUUU	UACAUGGACU	UUGAGUGACU
Cons. Est.G.....U..	..G.....GG.C.
Consenso	...r.....	...y.y....	rGr...yU..	.rG...y.yy	.y..GG.C.r.y.

						1080
Junín MC2	CAUCCGGCAA	GGACACCCCU	GGAGGCUAUU	GUCUAGAAGA	GUGGAUGCUC	GUUGCAGCCA
Junín CD1	CAUCCGGCAA	GGACACCCCU	GGAGGCUAUU	GUCUAGAAGA	GUGGAUGCUC	GUUGCAGCCA
Oliveros	CUACAGGCAA	UGACCUCCCU	GGGGGUUACU	GCCUAGAGCA	GUGGGCCAUA	GUUUGGGCAG
Pichindé	CUCUGGGCAA	C-AUGUCCCA	GGUGGCUACU	GUCUAGAGCA	GUGGGCCAUA	AUCUGGGCUG
Sabiá	CAGUGGGCAA	UGACAUGCCU	GGUGGUUAUU	GUCUUGAAAG	AUGGAUGCUA	GUGACGUCAG
Tacaribe	CUUUAGGGAA	CGAAGCCCCU	GGGGGAUACU	GUCUUGAAAA	GUGGAUGCUC	GUUGCAGCG
Lassa Jos.	CUGAAGGUAA	AGACACACCA	GGGGGAUUAU	GUCUGACCAG	GUGGAUGCUA	AUUGAGGCUG
Lassa Nig.	CAGAGGGAAA	UGAAACACCA	GGGGGAUACU	GCCUUACUAG	AUGGAUGUUG	AUUGAGGCCG
LCM Arm.	CUUCAGGGGU	GGAGAAUCCA	GGUGGUUAUU	GCCUGACCAA	AUGGAUGAUU	CUUGCUCGAG
LCM WE	CCUCAGGAGU	AGAAAUCCA	GGUGGUUAUU	GCCUGACCAA	AUGGAUGAUC	CUUGCUCGAG
Mopeia	CGGAAGGGAA	CGACAUGCCA	GGUGGUUACU	GUCUCACAAG	GUCUAUGCUA	AUAGGGCUUG
Cons. Est.	C....GG...	..A....CC.	GG.GG.UA.U	G.CU.....	.U.....U.	.U.....
Consenso	C...rGG.r.	..A....CC.	GG.GG.UAyU	GyCU.r...r	rU..ry..U.	.U.....r

						1140
Junín MC2	AAAUGAAGUG	UUUUGGCAAU	ACUGCUGUAG	CAAAAUGCAA	UCUGAAUCAU	GACUCUGAAU
Junín CD1	AAAUGAAGUG	UUUUGGCAAU	ACUGCUGUAG	CAAAAUGCAA	UCUGAAUCAU	GACUCUGAAU
Oliveros	GCAUCAAGUG	CUUUGGGAAC	ACAGCAGUUG	CAAAGUGUAA	UCAAAAUCAU	GACUCUGAAU
Pichindé	GAAUAAAAGUG	UUUCGACAAC	ACUGUAAUGG	CAAAAUGCAA	CAAAGAUCAC	AACGAGGAAU
Sabiá	AUCUUAAGUG	CUUUGGAAAC	ACAGCACUAG	CAAAAUGUAA	CCUUGACCAC	GAUUCGGAU
Tacaribe	AAUUAAGUG	UUUUGGGAAC	ACUGCAAUUG	CCAAGUGCAA	CCAAAUCAC	GACUCAGAGU
Lassa Jos.	AACUAAAAGUG	CUUCGGGAAC	ACAGCUGUGG	CAAAAUGUAA	UGAGAAGCAU	GAUGAGGAAU
Lassa Nig.	AAUUAAGUG	CUUUGGAAAC	ACUGCAGUAG	CCAAGUGCAA	UGAGAAACAU	GAUGAAGAAU
LCM Arm.	AGCUUAAGUG	UUUCGGGAAC	ACAGCAGUUG	CGAAAUGCAA	UGUAAAUCAU	GAUGCCGAAU
LCM WE	AGCUAAAAGUG	UUUUGGGAU	ACAGCUGUUG	CAAAAUGUAA	UGUCAUCAU	GAUGAAGAGU
Mopeia	AUCUGAAAAGUG	CUUUGGGAAC	ACUGCCAUUG	CGAAGUGUAA	UCAGGCACAU	GAUGAAGAAU
Cons. Est.	...U.AA.UG	.UU.G..AA.	AC.G...U.G	C.AA.UG.AACA.	.A....GA.U
Consenso	r..U.AArUG	yUyGr.AAy	AC.Gy..U.G	C.AArUGyAA	y...r..CAy	rAy...GARU

						1200
Junín MC2	UCUGUGACAU	GCUGAGGCUU	UUUGAUUACA	ACAAAAAUGC	UAUCAAAACC	UUAAAUGAU-
Junín CD1	UCUGUGACAU	GCUGAGGCUU	UUUGAUUACA	ACAAAAAUGC	UAUCAAAACC	UUAAAUGAU-
Oliveros	UCUGUGACAU	GUUGAGGCUC	UUUGAUUACA	AUAGGAAUGC	UAUUAAGAGU	CUGAAUGAUC
Pichindé	UUUGCGACAC	AAUGAGGUUA	UUUGAUUUCA	AUCAAAAUGC	CAUCAAAACC	UUGCA-ACUC
Sabiá	UCUGUGACAU	GUUGAAAUUG	UUUGAGUUCA	ACAAAAAAGC	GAUAGAGACA	UUGAAUGAC-
Tacaribe	UUUGUGACAU	GCUAAGGCUG	UUUGAUUACA	ACAAAAAUGC	UAUAAAACA	CUGAAUGAG-
Lassa Jos.	UUUGUGACAU	GCUGAGGCUG	UUUGACUUCA	ACAACAAGC	CAUUGAAAGG	UUGAA-AGCU
Lassa Nig.	UUUGUGACAU	GCUAAGGUUG	UUCGAUUUCA	ACAACAGGC	CAUAAGGAGG	CUCAA-AACA
LCM Arm.	UCUGUGACAU	GCUGCGACUA	AUUGACUACA	ACAAGGCUGC	UUUGAGUAAG	UUCAA-AGAG
LCM WE	UCUGUGACAU	GCUACGACUA	AUUGAUUACA	ACAAGGCUGC	CCUGAGUAAG	UUCAA-GCAA
Mopeia	UCUGUGACAU	GCUCCGCCUU	UUUGAUUUUA	AUAACAAGC	AAUUGCAAA	CUGAG-GUCU
Cons. Est.	..UG.GACA.	..U.....U.	.U.GA.U..A	A.....GC	.AU....A..	.U.....
Consenso	UyUGyGACAy	r.U..r.yU.	.UyGA.U.yA	Ay.rr...GC	.AU.rr.A..	yU..r.r...

Apéndice 1

1260

Junín MC2	GAAACUAAGA	AACAAGUAAA	UCUGAUGGGA	CAGACAAUCA	AUGCGCUGAU	AUCUGACAAU
Junín CD1	GAAACUAAGA	AACAAGUAAA	UCUGAUGGGA	CAGACAAUCA	AUGCGCUGAU	AUCUGACAAU
Oliveros	AGUCUCAA-A	GUCGUCUAAA	UCUUCUAACA	AAUACAAUAA	AUUCACUGAU	UUCGGAUAAU
Pichindé	AAUGUUGAGA	AUUCACUGAA	UCUUUUCAAA	AAGACCAUCA	ACGGACUCAU	UUCUGACUCA
Sabiá	AAUACAAAAA	ACAAGGUAAA	CUUGCUGACC	CACUCAAUUA	AUGCAUUAAU	AUCUGACAAC
Tacaribe	GAGACUAAGA	CGAGGGUGAA	UGUCCUGAGC	CACACCAUCA	AUGCUCUGAU	CUCAGACAAC
Lassa Jos.	GAAGCACAAA	UGAGCAUUCA	GUUGAUCAAC	AAAGCAGUAA	AUGCUUUGAU	AAAUGACCAA
Lassa Nig.	GAGGCCAAA	UGAGCAUACA	GCUGAUCAAC	AAGGCUGUCA	AUGCUUUAU	AAAUGAUCAG
LCM Arm.	GACGUAGAAU	CUGCCUUGCA	CUUAUUCAAA	ACAACAGUGA	AUUCUUUGAU	UUCAGAUCAA
LCM WE	GAUGUAGAGU	CUGCCUUGCA	UGUAUUCAAA	ACAACAUUAA	AUUCUCUGAU	UUCGGAUCAG
Mopeia	GAAGUUCAGC	AGAGCAUAAA	UUUGAUAAA	AAAGCUGUCA	ACGCCUCAU	CAAUGAUCAA
Cons. Est.A..U..A	..U..U....C..U.A	A..C..U.AUGA....
Consenso	rr..y..A..U..A	..U..U.r..C..U.A	Ay.C.yU.AUGay....

1320

Junín MC2	UUUUGAUGA	AAAACAAAAU	UAGGGAAUUG	AUGAGUGUCC	CUUACUGCAA	UUACACAAAA
Junín CD1	UUUUGAUGA	AAAACAAAAU	UAGGGAAUUG	AUGAGUGUCC	CUUACUGCAA	UUACACAAAA
Oliveros	CUGUUGAUGA	AAAACAGUU	GGCUGAGUA	AUGAACAUUC	CGUAUUGCAA	UUACACAAAA
Pichindé	CUUGUGAUUA	GAAACAGUCU	CAAACAACUU	GCUAAAAUUC	CUUACUGCAA	CUACACAAAA
Sabiá	UUACUGAUGA	AGAAUCGACU	UAAAGAAUUG	UUGAACACUC	CUUAUUGUAA	UUACACAAAA
Tacaribe	CUUCUGAUGA	AGAAUAAGAU	UCGAGAGCUG	AUGAGUGUAC	CUUAUUGCAA	CUACACACGA
Lassa Jos.	CUUAUAAUGA	AGAACCACUCU	ACGGGACAUC	AUGGGCCUUC	CAUACUGUAA	UUACAGCAAG
Lassa Nig.	CUCAUCAUGA	AGAACCACUU	GAGAGACAUC	AUGGGCAUAC	CAUAUUGUAA	CUACAGCAGA
LCM Arm.	CUACUGAUGA	GGAACCACUU	GAGAGAUCUC	AUGGGGGUGC	CAUAUUGCAA	UUACUCAAG
LCM WE	CUGUUGAUAG	AGAAUCAUCU	AAGAGAUCUA	AUGGGGGUAC	CAUACUGUAA	UUACUCAAG
Mopeia	CUGGUCAUGA	GGAACCACUU	GAGAGAUCUA	AUGGGUAUUC	CUUAUUGCAA	CUACUCAAG
Cons. Est.	.U..U.AU..	..AA.....UGA..U.U.C	C.UA.UG.AA	.UAC.....
Consenso	yU..U.AU.r	rrAAy.r..UGA..U.	.y.r.r..U.C	C.UAyUGyAA	.UAC....rr

1380

Junín MC2	UUUUGGUAUG	UCAACCACAC	ACUUUCAGGA	CAACACUCAU	UACCAAGGUG	CUGGUUAAUA
Junín CD1	UUUUGGUAUG	UCAACCACAC	ACUUUCAGGA	CAACACUCAU	UACCAAGGUG	CUGGUUAAUA
Oliveros	UUCUGGUACA	UCAACGACAC	UAGGACUGGA	AGGCACACCC	UCCCCAAUG	CUGGUUAAUC
Pichindé	UUUUGGUAUG	UCAAUGACAC	UAUCACAGGA	AGGCAUUCUU	UACCGCAGUG	UUGGCUGGUU
Sabiá	UUUUGGUAUG	UCAAUCACAC	AGCAUCAGGG	GAACACUCAU	UGCCACGGUG	CUGGCUUGUU
Tacaribe	UUUUGGUAUG	UCAAUCACAC	ACUUUCAGGA	CAGCACUCUC	UCCAAGAUG	CUGGAUGAUA
Lassa Jos.	UAUUGGUACC	UCAACCACAC	AACUACUGGG	AGAACAUCAC	UGCCCAAUG	UUGGCUUGUA
Lassa Nig.	UAUUGGUACC	UUAACCACAC	AUCAACAGGA	AAGACCUCAC	UACCAAGGUG	UUGGCUUUAUC
LCM Arm.	UUUUGGUACC	UAGAACAUGC	AAAGACCGGC	GAAACUAGUG	UCCCCAAGUG	CUGGCUUGUC
LCM WE	UUCUGGUUAC	UGGAACAUGC	UAAGACUGGU	GAGACUAGUG	UACCAAGGUG	UUGGCUUGUC
Mopeia	UUUUGGUACU	UGAAUGAUAC	UAGGACAGGG	AGAACAUCCC	UCCCCAAGUG	UUGGCUUGUG
Cons. Est.	U..UGGUA..	U..A..A..CC.GG.C..	U.CC....UG	.UGG.U..U.
Consenso	U.yUGGUAY.	U.rA..AyrCC.GG.	.rr....C..	U.CC..rrUG	yUGG.U.rU.

1440

Junín MC2	AAAAACAACA	GCUAUUUGAA	CAUUUCUGAC	UUCCGUAUUG	ACUGGAUACU	AGAAAGUGAC
Junín CD1	AAAAACAACA	GCUAUUUGAA	CAUCUCUGAC	UUCCGUAUUG	ACUGGAUUAU	AGAAAGUGAC
Oliveros	UCAAAUGGGU	CUUACUUGAA	UGAGACUAAA	UUCAGAACUC	AGUGGCUGUC	AGAGAGCAAU
Pichindé	CACAAUGGCU	CGUACCUCAA	UGAAACGCAU	UUCAAGAAUG	AUUGGUUGUG	GGAGAGCCAA
Sabiá	AGAAAUAAUA	GCUACUUGAA	UGAAAGUGAA	UUUAGAAAUG	AUUGGAUUUAU	UGAGAGUGAU
Tacaribe	AGAAACAAU	CUUACCUCAA	CUCAUCUGAA	UUUAGAAAUG	AGUGGAUUCU	UGAAAGUGAU
Lassa Jos.	UCAAAUGGUU	CAUACUUGAA	CGAGACCCAC	UUUUCUGAUG	AUAUUGAACA	ACAAGCUGAC
Lassa Nig.	UCAAAUGGAU	CAUAUCUCAA	UGAGACCAAG	UUUUCAGAUG	ACAUCGAACA	ACAAGCUGAC
LCM Arm.	ACCAAUGGUU	CUUACUUAU	UGAGACCCAC	UUCAGUGAUC	AAAUCGAACA	GGAAGCCGAU
LCM WE	ACUAAUGGCU	CCUACUUGAA	UGAGACCCAU	UUUAGUGAUC	AAAUCGAACA	AGAAGCAGAU
Mopeia	ACCAAUGGGU	CAUACCUCAA	UGAAACCCAG	UUCUCAACAG	AAAUUGAGCA	AGAAGCCAAU
Cons. Est.	...AA.....	..UA..U.AAA.	UU.....U.	A.....	..A.....A.
Consenso	...AAyrr..	..UAyyU.AAA.	UUy...r.U.	A.....	..Arr...A.

Apéndice 1

1500

Junín MC2	UUCUUAUUU	CUGAAAUGCU	AAGCAAAGAG	UAUUCGGACA	GGCAGGGCAA	AACUCCCUUG
Junín CD1	UUCUUAUUU	CUGAAAUGCU	AAGCAAAGAG	UAUUCGGACA	GGCAGGGUAA	AACUCCUUUG
Oliveros	GCACUCUUA	CAGAAAUGCU	AACAGAAGAU	UAUGACAAA	GGCAGGGUAG	CACACCACUC
Pichindé	AAUCUAUACA	AUGAGAUGCU	GAUGAAAGAG	UAUGAGGAAA	GACAAGGAAA	AACCCCUUUG
Sabiá	CAUUUAUUGU	CUGAAAUGCU	CAAUAAAGAA	UACAUAGAU	GACAGGGAAA	GACACCGUUG
Tacaribe	UCCUCAUUU	CAGUGAUGCU	GGGUAAGGAG	UAUUCAGAGA	GACAGGGGAG	AACUCCAAUC
Lassa Jos.	AAUAUGAUCA	CUGAGAUGUU	ACAGAAGGAG	UAUAUGGAGA	GGCAGGGGAA	GACACCAUUG
Lassa Nig.	AACAUGAUAA	CAGAGAUGCU	ACAAAAGGAA	UACAUAGAU	GACAGGGCAA	AACUCCACUG
LCM Arm.	AACAUGAUUA	CAGAGAUGUU	GAGGAAGGAU	UACAUAAAGA	GGCAGGGGAG	UACCCCCCUA
LCM WE	AACAUGAUCA	CAGAGAUGUU	GAGGAAGGAC	UACAUAAAA	GACAAGGGAG	UACUCCUUUA
Mopeia	AACAUGUUCA	CUGACAUGUU	GAGGAAGGAG	UAUGAGAAA	GGCAGAGCAC	AACACCUCUG
Cons. Est.U.....	..GA.AUG.UAA.GA.	UA.....A.A	G.CA..G.A.	.AC.CC..U.
ConsensoU.....	..GARAUgyUAARGA.	Uay...rA.A	GrCarrG.A.	.AC.CC..U.

1560

Junín MC2	ACUUUAGUUG	ACAUCUGUUU	UUGGAGCACA	GUAUUCUUCA	-CAGCGUCCC	UCUCCCUUCA
Junín CD1	ACUUUAGUUG	ACAUCUGUAU	UUGGAGCACA	GUAUUCUUCA	-CAGCGUCAC	UCUCCCUUCA
Oliveros	UCCUUGGUUG	AUCUAUGUUU	UUGGUCAACC	UUAUUCUAUG	UGACCA-CAU	UGUUUGCUCA
Pichindé	GCACUGACAG	AUAUAUGCUU	CUGGUCUUUG	GUUUUUUACA	CCAUAU-CAG	UAUUUCUCCA
Sabiá	ACUUUGGUGG	AUAUCUGUUU	CUGGAGCACU	UUGUUUUUCA	-CAACAACAC	UGUUUCUUCA
Tacaribe	ACUUUAGUGG	ACAUUUGUUU	UUGGAGUACA	GUGUUCUUCA	UCAGCA-CAC	UAUUCUUGCA
Lassa Jos.	GGUCUAGUUG	ACCUCUUUGU	GUUCAGUACA	AGUUUCUAUC	UUAUUAGCA-	UCUCCCUUCA
Lassa Nig.	GGUUUAGUUG	ACCUAUUUGU	UUUUAGCACA	AGUUUCUAUC	UGAUAAGCA-	UCUUUCUCCA
LCM Arm.	GCAUUGAUGG	ACCUUCUGAU	GUUUUCCACA	UCUGCAUAUC	UAGUCAGCA-	UCUCCUGCA
LCM WE	GCCUUAUUGG	AUCUUUUGAU	GUUUUCAACA	UCAGCAUACU	UGAUCAGCA-	UCUUUCUGCA
Mopeia	GGUUUAGUAG	AUCUUUUGU	UUUCUCCACU	AGCUUUUAU	UGAUCUCCG-	UGUUCCUCCA
Cons. Est.U....G	A..U.U...U	.U.....U...	..A....C..	U.UU....CA
Consenso	...yUrry.G	Ay.U.U...U	.U.....y.y.U.y.	..A....C..	U.UUy.y.CA

1620

Junín MC2	CUUGGUGGGC	AUACCCACCC	AUAGGCACAU	CAGAGGCGAG	GCAUGCCCUC	UGCCCCACAG
Junín CD1	CUUGGUGGGU	AUACCCUCCC	ACAGACACAU	CAGGGGCGAA	GCAUGCCCUU	UGCCACACAG
Oliveros	UCUUGUUGGU	UUCCCAACAC	ACAGACACAU	CUUAGAUGGU	CCUUGCCCUA	AACCCACACAG
Pichindé	CAUAGUUGGA	AUACCUACUC	AUAGGCACAU	CAUUGGUGAU	GGUUGCCCGA	AACCACACAG
Sabiá	CCUGGUAGGC	UUUCCAACUC	AUAGACACAU	ACGUGGUGAA	CCCUGCCCAC	UACCCCAUAG
Tacaribe	CCUCAUUGGU	UUCCCAACUC	AUGAACACAU	UCGAGGGGAA	GGUUGCCCUC	UCCCCAUAG
Lassa Jos.	CCUAGUCAAA	AUACCAACUC	AUAGGCAUUA	UGUAGGCAAG	UCGUGUGGGA	AACCUCACAG
Lassa Nig.	CCUGGUCAAA	AUACCAACCC	AUAGACACAU	UGUAGGUAAA	CCUUGCCCAA	AACCCACACAG
LCM Arm.	CCUUGUCAAA	AUACCAACAC	ACAGGCACAU	AAAAGGUGGC	UCAUGUCCAA	AGCCACACCG
LCM WE	UUUUGUGAGG	AUACCAACAC	AUAGACACAU	AAAGGGCGGU	UCAUGUCCAA	AGCCACAUCG
Mopeia	CCUAAUCAAA	AUUCCAACAC	AUAGACACAU	AAAAGGCAAG	CCUUGCCCGA	AACCACACAG
Cons. Est.	..U..U....	.U.CC..C.C	A....CACAUG.....	...UG.....	..CC.CA..G
Consenso	y.U.rU.rr.	.U.CC..C.C	AyrrrCACAUGr.rr.	...UGy....	..CC.CA..G

1680

Junín MC2	GCUAAAUAGC	UUGGGUGGUU	GCAGAUGUGG	UAAGUACCCC	AAUCUAAAGA	AACCAACAGU
Junín CD1	GUUGAACAGC	UUGGGUGGUU	GCAGAUGUGG	UAAGUACCCC	AAUCUAAAGA	AACCAACAGU
Oliveros	AUUAACAAAG	AAAGGAAUUU	GUUCCUGUGG	ACAUUUUGGA	AUACCGGGCA	AACC---UGU
Pichindé	GAUUACAAGA	AACUCCUGU	GUAGCUGCGG	GUAUUUAU-A	AUACCAAAGA	AACC-----
Sabiá	GCUCAACAGU	AGAGGAGGAU	GUAGAUGUGG	GAAAUACCCU	GAACUAAAA	AGCCAUCAC
Tacaribe	AUUGAACAGC	AUGGGAGGAU	GUAGGUGUGG	GAAGUAUCUG	CCUCUGAAGA	AGCCAACUAU
Lassa Jos.	AUUGAAUCAU	AUGGGCAUUU	GUUCCUGUGG	ACUCUACAAA	CAGCCUGGUG	UGCCUGUGAA
Lassa Nig.	GCUGAACCCAC	AUGGGCAUCU	GCUCCUGUGG	UCUAUACAAA	CAGCCAGGUG	UGCCUGUCAG
LCM Arm.	AUUAACCAAC	AAAGGAAUUU	GUAGUUGUGG	UGCAUUUAAG	GUGCCUGGUG	UAAAAACCGU
LCM WE	CUUGACCAAC	AAGGGGAUCU	GUAGUUGUGG	UGCAUUAAG	GUGCCUGGUG	UAAAAACUAU
Mopeia	ACUCAAUAC	AUGGCAAUCU	GUUCCUGUGG	CUUCUAUAAG	CAACCAGGUC	UCCCCACACA
Cons. Est.	..U.A.....U	G....UGUGGU.....	...C.....
Consenso	..U.A...r.U	Gy...UGUGGU.....	...Cy.rr..

Apéndice 1

								1740
Junín MC2	UUGGCGCAGA	GGA----CAC	--UAA---GA	-CCUCCC-G-	-AA--GG---	---U-C-CC	<u>C</u>	
Junín CD1	UUGGCGUAGA	GGA----CAC	--UAA---GA	-CCUCCU-G-	-AG--GG---	---U-C-CC	<u>C</u>	
Oliveros	UCGAUGGGUG	AAACGAUCAC	GGUGA---G-	---UCUU---	CA---GG---	-----CU--	<u>C</u>	
Pichindé	UUAC-----	AAAUGG----	-GUAA---G-	-----	-----	-----		
Sabiá	CUGGC--ACA	AGAAC--CAC	--UAGAAGGA	-----	CAAUCA----	UUGU-CUCC	<u>C</u>	
Tacaribe	UUGGCAUCGU	AGA----CAU	--UAG---GA	-----CAG-	CA--CA---U	UUGCGCUCC	<u>C</u>	
Lassa Jos.	AUGG-----	AAGAGA----	--UGA---GA	-CCC-UU-GU	CA---GG---	-----		
Lassa Nig.	AUGG-----	AAGAGG----	--UGAAA---	---UCCCA--	CA---GG---	-----		
LCM Arm.	CUGG-----	AAAAGA-CGC	--UGA-A-GA	A-----	CA---G---C	-----		
LCM WE	CUGG-----	AAAAGA-CGC	--UGA-----	---UC--AG-	CA---G---C	-----		
Mopeia	AUGG-----	AAAAGG----	--UGA-A-GA	A-----	CA---GGGGC	--A-----	<u>C</u>	
Cons. Est.	.U.....U.....		
Consenso	.Ur.....	rrr.....	..Urr.....		

								1800
Junín MC2	<u>ACC-----AG</u>	<u>CCC---GGGC</u>	<u>--AUU--GCC</u>	<u>C-GGGCU---</u>	<u>--GGUG----</u>	<u>-----</u>		
Junín CD1	<u>ACC-----AG</u>	<u>CCC---GGGC</u>	<u>--ACU--GCC</u>	<u>C-GGGCU---</u>	<u>--GGUG----</u>	<u>-----</u>		
Oliveros	<u>CCC--AUCA-</u>	<u>CUC--UGGGC</u>	<u>CCAUUGGGCC</u>	<u>CAGGGU-GAU</u>	<u>--GG-GG----</u>	<u>-----</u>		
Pichindé	<u>----G-U---</u>	<u>-----UGGG-</u>	<u>AAAAUAAGCU</u>	<u>CCG--CU--U</u>	<u>CCGAC----</u>	<u>-----</u>		
Sabiá	<u>ACC-GA----</u>	<u>CCC---GGGC</u>	<u>--AAU--GCC</u>	<u>C-GGGUC---</u>	<u>--GGUG----</u>	<u>-----</u>		
Tacaribe	<u>ACC-GG----</u>	<u>CCC---GGGC</u>	<u>--UUU--GCC</u>	<u>C-GGGCC---</u>	<u>--GGUG----</u>	<u>-----</u>		
Lassa Jos.	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>		
Lassa Nig.	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>		
LCM Arm.	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>		
LCM WE	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>	<u>-----</u>		
Mopeia	<u>CCCCGA--GA</u>	<u>CCCACCG--C</u>	<u>C--UUCGGCG</u>	<u>GUGGGUC--U</u>	<u>C-GGGGGCCU</u>	<u>GGAGUGACUG</u>		
Cons. Est.		
Consenso		

								1860
Junín MC2	-----UG	<u>G--CCCC-C</u>	<u>AG-----U</u>	<u>CCGCGGCC--</u>	<u>----U-GGCC</u>	<u>GCG-GAC--U</u>		
Junín CD1	-----UG	<u>G--CCCC-C</u>	<u>AG-----U</u>	<u>CCGCGGCC--</u>	<u>----U-GGCC</u>	<u>GCG-GAC--U</u>		
Oliveros	-----CG	<u>G--CCCCUC</u>	<u>-GA-----C</u>	<u>CCACGGC---</u>	<u>AAA----ACC</u>	<u>GUG-GGC---</u>		
Pichindé	-----AU-	<u>GGGCCUC---</u>	<u>-GACGUCACU</u>	<u>CC-C--C---</u>	<u>-AAU-AGG-G</u>	<u>GAGUGAC-GU</u>		
Sabiá	-----UG	<u>G--CCCC-C</u>	<u>AG-----U</u>	<u>CCGCGGC---</u>	<u>AAAU---GCC</u>	<u>GCG-GAC--U</u>		
Tacaribe	-----UG	<u>G--CCCC-C</u>	<u>-GA-----U</u>	<u>CCGCGGC--G</u>	<u>---UU--GCC</u>	<u>GCG-AA---U</u>		
Lassa Jos.	-----	<u>G--CCCCGU</u>	<u>-GAC-----</u>	<u>CCACCGCCU-</u>	<u>--AUU-GGCG</u>	<u>GUG-G---GU</u>		
Lassa Nig.	-----	<u>G--CCCCGU</u>	<u>-GAC-----</u>	<u>CCACCGCC--</u>	<u>-AAUU-GGCG</u>	<u>GUG-G---GU</u>		
LCM Arm.	-----	<u>G--CCUCC-C</u>	<u>UGAC-UC--U</u>	<u>CCACCUC--G</u>	<u>AAA----GAG</u>	<u>GUG-GAGAGU</u>		
LCM WE	-----	<u>G--CCUCC-C</u>	<u>UGAC-UC--U</u>	<u>CCACCUC--G</u>	<u>AAA----GAG</u>	<u>GUG-GAGAGU</u>		
Mopeia	UUCCUUGAUG	<u>G--CCUCCAC</u>	<u>-GAC-----</u>	<u>CCACCGCC--</u>	<u>AGA---GGCG</u>	<u>GUG-G---GU</u>		
Cons. Est.	G..CC.C...	.G.....	CC.C..C... G.G.....		
Consenso	G..CCyc...	.G.....	CC.C..C...r.. G.G.r.....		

								1920
Junín MC2	<u>-GGGGAGGC-</u>	<u>-AC-----U</u>	<u>GC-----</u>	<u>-----UUA</u>	<u>CAGUGCAUA-</u>	<u>--GGCUGCCU</u>		
Junín CD1	<u>-GGGGAGGC-</u>	<u>-AC-----U</u>	<u>GC-----</u>	<u>-----UUA</u>	<u>CAGUGCAUA-</u>	<u>--GGCUGCCU</u>		
Oliveros	<u>CGAGGGGG--</u>	<u>-----GGU</u>	<u>GCA-----</u>	<u>-U-----UUA</u>	<u>CAUUGCAAC-</u>	<u>--CACUGA-U</u>		
Pichindé	<u>C---GAGGCC</u>	<u>U-C-----</u>	<u>G-AGGA----</u>	<u>UUA-AGC</u>	<u>UCA</u>	<u>GAGGUUGA-U</u>	<u>CAAGUCUGUG</u>	
Sabiá	<u>-GGGGAG-C-</u>	<u>-ACC-----</u>	<u>-----AA--</u>	<u>-U-----CUA</u>	<u>GAUGGUGAA-</u>	<u>--UGCUGCCU</u>		
Tacaribe	<u>CGGGGAGGC-</u>	<u>-ACCUGUGGU</u>	<u>GC-GGAAGUC</u>	<u>UUACAGC</u>	<u>UCA</u>	<u>CAGUGCAAA-</u>	<u>--AGCUGUUU</u>	
Lassa Jos.	<u>CACGGGGGC-</u>	<u>-----</u>	<u>-----GUC</u>	<u>C-AU---UUA</u>	<u>CAGAACGACU</u>	<u>CUAGGUGUCG</u>		
Lassa Nig.	<u>CACGGGGGC-</u>	<u>-----</u>	<u>-----GUC</u>	<u>C-AU---CUA</u>	<u>CAGGACGACU</u>	<u>UUAGGUGUUG</u>		
LCM Arm.	<u>CAGGGAGGC-</u>	<u>--CC---AG-</u>	<u>--AGG--GUC</u>	<u>-----UUA</u>	<u>GAGUGUCACA</u>	<u>ACAUUUGGGC</u>		
LCM WE	<u>CAGGGAGGC-</u>	<u>--CC---AG-</u>	<u>-C-GG--GUC</u>	<u>-----UUA</u>	<u>GAGUGUCACA</u>	<u>ACAUUGGGUC</u>		
Mopeia	<u>CGUGGAGGU-</u>	<u>-----</u>	<u>-----GUC</u>	<u>CC-U-UC</u>	<u>UUA</u>	<u>CAGGACAACU</u>	<u>CUGGGAGGAC</u>	
Cons. Est.G.G...A	.A.....		
ConsensoGrG...yyA	.A...y...		

Apéndice 1

						1980
Junín MC2	UCGGGAGGAA	CAGCAAGCUC	GGUGG-UAAU	A-GAGGUGUA	AGUUCUUCU-	-UCAUAGAGC
Junín CD1	UCGGGAGGAA	CAGCAAGCUC	GGUGG-UAAU	A-GAGGUGUA	AGUUCUUCU-	-UCAUAGAGC
Oliveros	UCAGACUUA	ACACUAGAGC	ACUUGGAAGU	A-GAGGCUUG	GGCUCUAUG-	-CUUGGCACA
Pichindé	UUGUUCUGU	ACAGUGUGUC	AAUAGGUAA-	A-CAUCUCAU	GGGUUUCUG-	-AUCUCUGAC
Sabiá	CCAGUGUGAA	GAGCAAU-GC	AUCAGGCAAU	A-AACCUUCC	AGUUUGGUU-	-UGAGGUGGU
Tacaribe	UGGGAAAGAA	AAGGAA-CUC	UUUGGGUAGU	A-GUGGUUUU	GGUUCUCC-	-UUUGGGUAGU
Lassa Jos.	AUGUUC-UGA	ACUCCAUAUC	UCUGGGCAGC	A-CUGCUCUC	AAAACCGAU-	-GUGUUCAGU
Lassa Nig.	AAGUUC-UGA	ACACCAUGUC	UCUAGGGAGC	A-CAACUCUC	AGGACUUUU-	-GCAUCAAGU
LCM Arm.	CU----CUAA	AAAUUAGGUC	AUGUGGCAGA	AUGUUGUGAA	CAGUUUUCAG	AUCUGGGAGC
LCM WE	CU----CUGA	AGAUCAAAUC	AUGUGGCAGG	AUGUUGUGAA	CGGUCGGGAG	AUCAGGGAGU
Mopeia	CUGUUC-UAA	ACACCAGGUC	UCUACUCAAG	A-CUGGUCUG	GGGGUUGGG-	-AUUUGUGGA
Cons. Est.CA..	A.....
Consensor.r...CAr.	A.....y...	.r.y.....r..

						2040
Junín MC2	-UUCCCAUCC	AACACUGACU	GAAACAUUUAU	GCAGUCUAGC	AGAGCACAGU	GUGGCUCACU
Junín CD1	-UUCCCAUCC	AACACUGACU	GAAACAUUUAU	GCAGUCUAGC	AGAGCACAGU	GUGGCUCACU
Oliveros	-UUGCCCUCA	AAUGCUGCUU	GAAACAUGAG	ACAGUCAAGC	AGUGCACAAU	GAGGUUCACC
Pichindé	CCAGCCUGUC	ACUGUUGCAU	CAAACAUGAU	GGUGUCAAGC	AAUGCACAUG	GGGGU--UU
Sabiá	-UGGCCUGCU	AUGGCUGACU	GAAACAUUAG	ACAAUCAAGU	AGGGCACAGU	GUGGCUC--U
Tacaribe	-UCUCCACUC	ACGGCUGAGU	GGAACAUUUAU	GCAGUCCAGC	AGUGCACAGU	GAGGUUCAUC
Lassa Jos.	CCU-CCUGAC	ACUGCUGCAU	CAAACAUGAU	GCAGUCCAUC	AGUGCACAGU	GUGG-----
Lassa Nig.	CCU-CCUGAA	ACUGCUGCAU	CAAACAUAU	GCAAUCCAUC	AGUGCACAAU	GAGG-----
LCM Arm.	CUUGCUCUUGG	AG-GCGCUUU	CAAAAAUGAU	GCAGUCCAUG	AGUGCACAGU	GCGG-----
LCM WE	CUUGCUCUUGG	AA-GCACUCU	CAAAGAUGAU	GCAGUCCAUG	AGUGCACAGU	GUGG-----
Mopeia	CUC-CCUAUG	ACUGCUGCUU	CAAACAUGAG	ACAAUCCAGC	AAUGCACAGU	GUGG-----
Cons. Est.CC....	A.....U	..AA.AU.A.UC.A..	A..GCACA..	G.GG.....
Consenso	.y..CC....	A..ry....U	.rAA.AU.A.	r..rUC.A..	Ar.GCACA..	G.GG.....

						2100
Junín MC2	GGAGGCCAAC	UUAAGGGGAG	CAUCCUUAUC	UCUCUUUUUC	UUGCUGACAA	CCACUCCAUC
Junín CD1	GGAGGCCAAC	UUAAGGGGAG	CAUCCUUAUC	UCUCUUUUUC	UUGCUGACAA	CCACUCCAUC
Oliveros	UGA-----	-----	-AUUGCCACC	UUUCUUCUUU	CUUGAAAGCA	CAACACCAUC
Pichindé	GCAG--UGGU	UU-----GUA	CAGCU---CC	UUUCUUCUUC	UUCUUUAUGA	CUAAGCCUUC
Sabiá	GAGGACU---	----GUGGGA	UGUCUUUGUC	UUUCUUUUUC	UUUGGCACAA	UAACUCCAUC
Tacaribe	AGGACUGGGG	CUUGGUGGAG	AAUCUUUGCU	CUUAUUCUUC	UUGCUAUAUA	UCACCCCAUC
Lassa Jos.	---GGUUUAU	UCCU-----	CUUUACCGCC	UCUUUUUCUUC	UUUUAACAA	CGACACCUGU
Lassa Nig.	---AGUUUAU	UCCU-----	CUUUACACC	UCUCUUCUUU	UUCUCCACAA	CUACCCCAUC
LCM Arm.	---GGUGAUC	UCUU-----	-----	--UCUUCUUU	UUGUCCCUUA	CUAUUCCAGU
LCM WE	---GGUGAUU	UCUU-----	-----	--UCUUCUUU	UUGUCUCUCA	CUACCCCAUC
Mopeia	---UGUCACU	UCUU-----	CUUUGCCACC	UCUCUUCUUC	UUUUGCGUUA	CAAUCCCGU
Cons. Est.U.UU.UU.	.U.....A	..A..CC..U
ConsensoU.UUyUUy	yU.....A	y.A..CC..U

						2160
Junín MC2	GUGAUGUUUG	CAUAGGUGGC	CAAUUUUCUC	CCAGACCUGU	UGGUCGAACU	GCCUGGCUUG
Junín CD1	GUGAUGUUUG	CAUAGGUGGC	CAAUUUUCUC	CCAGACCUGU	UGGUCGAACU	GCCUGGCUUG
Oliveros	GUGUUUAGUG	CAAAGUGAUC	CAAUUUUCUC	CCAAACAACA	CUCUCAUAGA	AUCUUGCUUG
Pichindé	GUGCUUGGUG	CAGAGUUGGU	UGUAUCUCUC	CCAGAUCUCA	UCCUCAAGG	UGCGUGCCUG
Sabiá	GUGUUUCUUA	CAGAGGUGAC	CAAAAUCUGA	CCAGAUUGGC	UCCUCAACU	UUCUCGCCUG
Tacaribe	AUCUUGUUUG	CAGAGGUGGC	CGAAUUUAUC	CCACACUUGG	UGCUCGAAUA	GUCUUGCUUG
Lassa Jos.	GUGCAUGUGG	CAUAAGUCUU	UAUACUGGUC	CCAGACUGCA	UUUUCAUACU	UCCUGGAAUC
Lassa Nig.	GUGCAUGUGA	CAUAGAUCU	UGAAUUGAUC	CCAAACAGCA	UUCUCAACU	UUCUUGAAUC
LCM Arm.	AUGCAUCUUA	CACAACCAGC	CAUAUUUGUC	CCACACUUUG	UCUUCAUACU	CCCUCGAAGC
LCM WE	GUGCAUUUUG	CAUAGCCAGC	CAUAUUUGUC	CCACACUUUA	UCUUCAUAU	CUCUUGAGGC
Mopeia	AUGCAUUCUA	CAAAGGUGUU	UGUAUUCAUC	CCAGAUCUUA	UCCUCAAAU	UUCUUGCAUC
Cons. Est.	.U.....	CA.A.....	...A.....	CCA.A.....	...UC..A..	..C..G....
Consenso	rU.....r	CA.Ar.....	yAr.A.y....	CCA.Ay....	y..UCr.A..	..C..G....

Apéndice 1

2220

Junín MC2	UUCUGAUGUA	AGCCUAACAU	CAACCAGCUU	AAGAUCUCUU	CUUCCAUGGA	GGUCAAAACAA
Junín CD1	UUCUGAUGUA	AGCCUAACAU	CAACCAGCUU	AAGAUCUCUU	CUUCCAUGGA	GGUCAAAACAA
Oliveros	GUCUGUUGGG	AAGCUGAGAU	CAAUUGUGGA	GAUGUCAGGU	CUUCCAUUCA	GUUCCAUUAG
Pichindé	CUCUGCACUG	AGCUUCACGU	CAAGUACUUU	UAGAUCUCUU	CUUCCAUGCA	UUUCGAACAG
Sabiá	UUCAGAUGCC	AAUUUAAUGU	CGACAAUCUU	UAAGUCUUUG	CGACCGUGUG	UAUCUACUAA
Tacaribe	AUCUGAAGUG	AACUUAACAU	CAAUUAAUUU	GAGAUCUUUC	CUUCCGUGAA	UAUCUAGAAG
Lassa Jos.	AGUUUUGCUG	AGGGCAAUAU	CAAUUAGUUU	AAUGUCUUUU	CUUCCUUGUG	AUUCAAGGAG
Lassa Nig.	UGCUUUACUA	AGAGUGAUGU	CAAUCAGUUU	UAUGUCUCUC	CUCCCUUGUG	ACUCAAGGAG
LCM Arm.	UUCCCUUGUC	AUUUCAACAU	CGAUAAGCUU	AAUGUCCUUC	CUAUUCUGUG	AGUCCAGAAG
LCM WE	CUCCUACUC	AUCUCAACAU	CAAUGAGUUU	UAUGUCCCUU	CUAUUCUGUG	AGUCCAGAAG
Mopeia	GUCUUUCUGC	AUGGACACAU	CAAUCAGUUU	UAUGUCCUC	CUGUUCUGGG	AGUCAAGAAG
Cons. Est.	A.....A..U	C.A.....	.A..UC....	CU....U...	..UC.A..A.
Consenso	..y.....	A.....A.rU	CrA..r....	.A.rUC....	CU.yy.U..r	..UC.A..Ar

2280

Junín MC2	CUUCCUGAUG	UCAUCGGACC	CUUGAGUGGU	CACAACCAUG	UCCGGAGGCA	GCAAGCCAAU
Junín CD1	CUUCCUGAUG	UCAUCGGACC	CUUGAGUGGU	CACAACCAUG	UCCGGAGGCA	GCAAGCCAAU
Oliveros	UUUUUGAUG	UCAUCAGAAC	CUUGACAUGU	CAAAACUAAU	UUUGAAGGCA	GCUUGAUUAG
Pichindé	UCUGAUUAUA	UCAUCUGAAC	CUUGAGCAGU	GAAAACCAUG	UUCUGGGGCA	AAUGUCUGAU
Sabiá	GCGCCUUAUA	UCAUCUGAAC	CUUGGGUGGU	GAGGACCAUG	UUUUGAGGAA	GGAGCCCUAU
Tacaribe	CUUUCUGAUG	UCAUCAGAUC	CUUGUGUUGU	AAUCACCAUG	UUGGGAGGUA	GUAACCCUUA
Lassa Jos.	UUUCCUUAUG	UCAUCGGACC	CCUGACAGGU	AAUGACCAUA	UUCCGGGGGA	GUGCAUCAAU
Lassa Nig.	UUUUCUGAUA	UCCUCUGAUC	CUUGGCAAGU	GAUGACCAUG	UUCCGAGGAA	GGGCUUCUUA
LCM Arm.	CUUUCUGAUG	UCAUCGGAGC	CUUGACAGCU	UAGAACCAUC	CCCUGCGGAA	GAGCACCUAU
LCM WE	CUUUCUGAUG	UCAUCAGAAC	CUUGACAGCU	CAAGACCAUC	CCUUGUGGGA	GAGCACCUAU
Mopeia	CUUUCUGAUG	UCAUCAGCAC	CUUGACACGA	CAAGACCAUG	UUCUUUGGCA	GGCUCUCAAU
Cons. Est.	U.AU. UCAUC.G..C	CUUG....G.	.A..ACCA..	GG.A.....A.
Consenso	.y...U.AUr	UCAUC.G..C	CUUG....G.	.A..ACCA..	yy....GG.A	r.....y.A.

2340

Junín MC2	CACGUAACUA	AGAACUCCUG	GCAUUGCAUC	UUCUAUGUCU	UUCAUUAAGA	UGCCGUGAGA
Junín CD1	CACGUAACUA	AGAACUCCUG	GCAUUGCAUC	UUCUAUGUCU	UUCAUUAAGA	UGCCGUGAGA
Oliveros	AAUUCUGCUC	AGCAGGCCUG	AUCUUGCGUU	CUCGAUGUCU	CUCAGUAAA	UUCCAUGAGA
Pichindé	GAUUGAGGAG	AUCAGACCCG	GUUGAGCAUC	AGCCAAGUCC	UUUAAAAGGA	GACCAUGUGA
Sabiá	AACAUAGCUG	AGCAAGCCUG	GCUGUGCAGA	UUCUAGGUCC	UUUAGCAACA	UCCAUGUGA
Tacaribe	GACAUAAACUG	AGCACACCUG	GCACAGCAUC	CUCAAUGUCC	UUCAUUAAGA	UGCCGUGAGA
Lassa Jos.	GACAGCACUG	GUCAAGCCCG	GUUGUGUAGC	GAAGAGGUCU	GUGACAUCAA	UCCAUGUGA
Lassa Nig.	CACUGCACUG	GUUAACCCAG	GUUGGGCAGC	AAACAAAUCA	GUCACAUCAA	UACCAUGUGA
LCM Arm.	AACUGACGAG	GUCAACCCGG	GUUGCGCAUU	GAAGAGGUCG	GCAAGAUCCA	UGCCGUGUGA
LCM WE	AACUGAUGAG	GUCAGCCCG	CCUGUGCAUU	GAAGAGGUCA	GCAAGAUCCA	UGCCGUGUGA
Mopeia	UACAGCACUU	GUCAACCCCG	GCUGAACUGA	GAAUAGGUCU	UGGAUGUCA	UGCCAUGAGA
Cons. Est.	.A.....	...A..CC.GA..UC.	...A.....A	..CC.UG.GA
Consenso	.Ay.....	r..A..CC.G	.y...ry...A.rUC.	...A.....A	..CCrUG.GA

2400

Junín MC2	GUGUCUGCUA	CCAUUUUUA	ACCCUUCUC	AUCAUGUGGU	UUUCUGAAGC	AGUGAAUUA
Junín CD1	GUGUCUGCUA	CCAUUUUUA	ACCCUUCUC	AUCAUGUGGU	UUUCUGAAGC	AGUGAAUUA
Oliveros	GUUUUUGCUU	CCAUUUUUGA	AGCCUUCUC	AUCAUGUGGG	AGUCUGAAGC	AAUGUAUGUA
Pichindé	AUACUUGUUC	UGCUCUUUGA	AGGAUUUUUC	AUCAUGGGGA	AACCUAUAAC	AGUGUAUGUA
Sabiá	AUGCCUGCUA	CCAUUUUUGA	ACCCUUCUC	AUCAUGUGGU	UUCCUAUAAC	AGAGGAUUA
Tacaribe	AUGUCUGCUC	CCAUUUUUA	AUCCUUCUC	AUCAUGUGGC	UUUCUGAAGC	AGUGGAUGUA
Lassa Jos.	GUACUUAGCA	UCCUGCUUGA	ACUGCUUUA	AUCAGUAGGU	UCACGGAAGA	AGUGUAUGUA
Lassa Nig.	AUACUUUGCA	UCUUGUUUGA	AUUGCUUCA	AUCUGUUGGC	UCCCUAAAGA	AAUGUACAUA
LCM Arm.	GUACUUGGAA	UCUUGCUUGA	AUUGUUUUUG	AUCAACGGGU	UCCUGUAAA	AGUGUAUGAA
LCM WE	AUACUUGGAG	UCCUGCUUGA	AUUGCUUCUG	GUCCGUAGGU	UCUCUGUAAA	AAUGUAUGAA
Mopeia	GUGUUUGGAG	UCUUGUUUGA	AUUGUUUUAU	GUCUGUUGGU	UCCUGUAAA	AAUGAAUUA
Cons. Est.	.U...U....	...U..UU.A	A....UU...	.UC....GG.	...C...A..	A.UG.A...A
Consenso	rUry.U....	y..U.yUUrA	A...yUU...	rUC....GG.	...C.r.Ar.	ArUG.Ayr.A

Apéndice 1

2460

Junín MC2	CUUGCUACCU	GCAGGCUGGA	ACAACGCCAU	CUCAACAGGG	UCA-GUAGCU	GGU-CCUUCA
Junín CD1	CUGCUUACCU	GCAGGCUGGA	ACAACGCCAU	CUCAACAGGG	UCA-GUAGCU	GGU-CCUUCA
Oliveros	GUAACCAGUU	GUGGGCUGAU	ACAAAGCAAG	CUCAACAGGG	UCU-GUUGGU	GG-GCCCUCA
Pichindé	GUUGCCUGUA	UCAGGUUGGU	AGAUGGCCAU	UCCACCAGGG	UC-AUUUGGU	G-UCCCCUCA
Sabiá	AUUACCUGUG	UCUGGUUGGU	ACAAUGCUAA	UUCAACCAGG	UCA--UUACU	GGUACCCUCA
Tacaribe	UUGUUUCUUU	GCAGGUUGAU	AUAAUGCAAG	CUCUACUGGG	UCAAGUUGGA	GGU-CC-UCG
Lassa Jos.	GCAGCCUGAA	CUUGGUUGAU	AGAGGGCAAU	UCCACUGGGA	UCU--UCAGG	UCUUCUUUCA
Lassa Nig.	GCAACCCGAG	CUCGGUUGAU	AAAGGGCUAU	CUCAACUGGG	UCU--UCUGG	UCUUCUUUCA
LCM Arm.	CUGCCCUGUC	UGUGGUUGGA	AAAUUGCUAU	UCCACUGGGA	UCA--UUAAA	UCUACCCUCA
LCM WE	UUGCCCAUUU	UGUGGUUGAA	AUAUUGCUAU	CUCACUGGGA	UCA--UUGAA	CCUGCCUUUCA
Mopeia	CUGACCAUUA	UUAGGUUGAU	AGAUAGCUAU	CUCACAGGGG	UCU--UCAGG	UCUCCCUUCU
Cons. Est.GG.UG..	A.A..GC.A.	.UC.AC.GG.	UC...U....CC.UC.
Consenso	.y..yy....	...GGyUGr.	A.A..GC.A.	yUC.AC.GGr	UC...U....CC.UC.

2520

Junín MC2	AUGUCGAGCC	AAAGGGUAUU	GGUGGGG-UC	GAGUUUCCCC	ACU--GCCUC	UCUGAUGACA
Junín CD1	AUGUCGAGCC	AAAGGGUAUU	GGUGGGG-UC	GAGUUUCCCC	ACU--GCCUC	UCUGAUGACA
Oliveros	AUAUCCACCC	AGAUGGUUUC	AUCUGGA-UU	CAGCUUUGAG	AUU--GCUUC	UUUGAUCACU
Pichindé	AUAUCAAUCC	ACGUAGUGGC	UCUUGGA-UC	AAGCAUCUGA	AUU--GAGGA	C---ACAACA
Sabiá	AUGUCAAUCC	AUGUUGUCAG	AGUUGGG-UC	AAUCUUACCC	AUU--GCCUC	UCUCACAACU
Tacaribe	AUGUCCAACC	AAAUAGUGUU	GGAGGGG-UC	AAGUUCACUA	AUA--GCAGC	UCUGACUAUU
Lassa Jos.	AUGUCCAACC	AGGUCUUAGC	AUUUGGG-UC	AACUUGCAGC	AUU--GCAUC	CUUGAGGGUC
Lassa Nig.	AUAUCCAACC	AGGUUUUUGC	GUUGGGAAUC	AAGUUGAAGC	ACUUGAAAUC	UUUGAGAGUC
LCM Arm.	AUGUCAAUCC	AUGUAGGAGC	GUUGGGG-UC	AAUUCUCCC	A--UGAGGUC	UUUUAAAAGC
LCM WE	AUGUCAAUCC	AUGUGGGAGC	AUUGGGG-UC	AAUCCUCCC	A--UCAAGUC	UUUCAACAGC
Mopeia	AUGUCAAUCC	AUGUCUUAGC	AUUUGGAUCC	AA-UUGCAAC	A--UGCUGUC	UUUGAGUUC
Cons. Est.	AU.UC.A.CC	A.....GG....	.A.....	A.....A.....
Consenso	AUrUC.A.CC	A.r.....GGr.yy	.A.y.....	A.....	y...A.....

2580

Junín MC2	GCUUCUUGUA	UCUCUGUCA	GUUAGCCAAU	CUCAAUUCU	GACCGUUCUU	UUCGG-----
Junín CD1	GCUUCUUGUA	UCUCUGUCA	GUUAGCCAAU	CUCAAUUCU	GACCGUUCUU	UUCGG-----
Oliveros	GCUUCCUGUG	UGUCUGUCA	AGGAUUAAGU	CUUAUGUUGC	CACCAUUUUU	GACAGGUG--
Pichindé	GUGUCUUCUU	UUUC--UUAG	G--GAUUUGU	UUGAGAU-CU	GGUGAUCC--	-ACCGUU--U
Sabiá	GACUCUUGCA	ACUCAGUCA	GUGGGAGAGU	CUAUUGUUCU	GACCGUUCUU	UUCUGGUGUU
Tacaribe	AGUUCUUGCC	CCUCUGUGAG	GCUUGCCAGU	CUUAAAUUUU	GCCCUUUUCU	UUCUGCUGGC
Lassa Jos.	AUCAGCUGAG	AAGAGGUAAG	CCCAGC-GGU	AAACCCUGCC	GACUGCAGGG	AUUUACUGGA
Lassa Nig.	AUCAACUGAG	AGUAAGUUAA	UCCUGC-AGC	AAAGCCUGCA	GACUGUAAUG	ACUUGUUGGA
LCM Arm.	AUUGUCUGGC	UGUAGCUUAA	GCC--C-ACC	UGAG---GUG	GACC-UGCUG	CUCCA--GGC
LCM WE	AUUGUUUGAC	UGUAACUCA	GCC--C-ACC	UGAG---GUG	GGCC-UGCUG	CUCCA--GGC
Mopeia	AUUGUCUGAG	AGUAGGUCAA	ACCAGC-AGG	CAUGCCUUGG	GCCUGAGCGA	GUCUGUUGGA
Cons. Est.U...U.A.
Consenso	r....yU...U.Ary.....	..y.r.....

2640

Junín MC2	-----GUUG	U--CUAGGUC	CAGCAA----	-CUGGUUUC-	CUUGUCAGAU	CAUACUUGU
Junín CD1	-----GUUG	U--CUAGGUC	CAGCAA----	-CUGGUUUC-	CUUGUCAGAU	CAUACUUGU
Oliveros	---GGGAGG	ACCAUA----	---CAUGACA	UCU-----	UUUGAGAAGU	CAUAGUGGU
Pichindé	CUUACUGGUG	GC--UGGAU-	-----AGCA	GUUGGCUUU-	GAGUCCAAAU	CCACAGUGGU
Sabiá	CUCG-----G	CCCUUG---U	UGG-----	---GGCUUU-	GUGCUUAAAU	CAACAGUGGU
Tacaribe	CUUGGUGCUG	CUCCUGGUUC	UGGCACUGCA	UCCGGCUUC-	UUGGUCAGAU	CUACGCUCGU
Lassa Jos.	AUUGUUGCUG	UC---AG-CU	UUCUGUGGCU	UCCCAUCUGA	UU--CCAGAU	CAACGACAGU
Lassa Nig.	CCCCCAUUU	CC---AA-UU	UUCUGGGGCU	UGUUGUCUUG	CU--CAAGGU	CCACCACAGU
LCM Arm.	GCUGGCCUGG	GU---GAAUU	GACUGCAGGU	U----UCUCG	CUUGUGAGAU	CAAGUGUUGU
LCM WE	ACUGGCCUAG	AU---GAGUU	GGCCACAAGU	U----UUUCA	UUUGUGAGAU	CAAUUGUCGU
Mopeia	GUUACCUUUG	UC---AGGCU	UUGCAUUCUG	UUGCACAUCA	CCACUCAGAU	CAACAGUUGU
Cons. Est.A..U	C.A.....GU
Consensor.....ArrU	C.A...y.GU

Apéndice 1

2700

Junín MC2	GUUGUCCCAU	GACCUGCCUA	UGAUUUGUGA	UCUGGAACCA	AUAUAAGGCC	AACCAUCGCC
Junín CD1	GUUGUCCCAU	GACCUGCCUA	UGAUUUGUGA	UCUGGAACCA	AUAUAAGGCC	AACCAUCGCC
Oliveros	GUUAUCCCAA	GCUCUGCCUG	UGAUCUGGGA	UCUUGAGCCA	AUGUAUGGCC	AACCUUCACC
Pichindé	GUUGUCCCAA	GCCCUCCCUU	GAACUUGGGA	UCUUGAGCCA	AUGUAGGGCC	AGCCAUCCCC
Sabiá	GUUUUCCCAU	GACCUACCCU	UGAUCUGGGA	UCUGGAGCCA	AUGUAAGGCC	AACCUUCUCC
Tacaribe	GUUGUCCCAU	GAUCGCCCCA	UAAUUUGCGA	UCGAGAGCCU	AUAUAGGGCC	AGCCAUCACC
Lassa Jos.	GUUUUCCCAU	GCCCUUCCUG	UUAUUGAGGU	UGUUGAUGCA	AUAUAUGGCC	AUCCAUCUCC
Lassa Nig.	AUUUUCCCAU	GCUCUCCCA	CAAUCGAGGU	CCUUGAUGCA	AUAUAGGGCC	AUCCGUCUCC
LCM Arm.	GUUUUCCCAU	GCUCUCCCA	CAAUCGAUGU	UCUACAAGCU	AUGUAUGGCC	AUCCUUCACC
LCM WE	GUUCUCCCAU	GCUCUCCCA	CAACUGACGU	UCUACAGGCU	AUGUAUGGCC	AUCCUUCACC
Mopeia	GUUAUCCCAA	GCCUGCCGA	CGAUCGAUGU	UCUUGAUGCU	AUGUAAGGCC	AACCCUCUCC
Cons. Est.	.UU.UCCCA.	G..C..CC..	..A.....G.	U....A..C.	AU.UA.GGCC	A.CC.UC.CC
Consenso	rUU.UCCCA.	G.yC..CC..	..Ayy.r.G.	U....A..C.	AUrUA.GGCC	A.CC.UC.CC

2760

Junín MC2	AGAAAGGCAA	AGUUUGUACA	GAAGGUUUUC	AUAAGGGUUU	CUAUUGCCUG	GUUUCUCAUC
Junín CD1	AGAAAGGCAA	AGUUUGUACA	GAAGGUUUUC	AUAAGGGUUU	CUAUUGCCUG	GUUUCUCAUC
Oliveros	AGAGAGGCAA	AUCUUGUAAA	GGAUGUUCUC	AUAGGGAUUC	CUCAUCCUG	GUUUUUCUGA
Pichindé	UGAGAGACAG	AUUUUUACA	GCAGGUUUUC	AUAAGGGUUC	CUUUGCCCGG	GUGUAGUACU
Sabiá	UGAAAGGCAG	AUUUUUGUACA	GAAGGUUUUC	AUAAGGGUUU	CUAUUCCAG	GUUUCUCAUC
Tacaribe	UGACAAACAC	AGUUUGUAGA	GUAUUUUUC	GUAGGGGUUC	CUACUCCAG	GUCUGUCGUC
Lassa Jos.	UGACAAACAA	AUCUUGUAGA	GUAUGUUUC	AUAAGGAUUC	CUUUCACCAG	GGGUGUCUGA
Lassa Nig.	UGAGAGACAG	AUCUUGUAVA	GGAUGUUCUC	AUAAGGGUUC	CUUUCACCCG	GUGUGUCUGA
LCM Arm.	UGAAAGGCAA	ACUUUAUAGA	GGAUGUUUC	AUAAGGGUUC	CUGUCCCAA	CUUGGUCUGA
LCM WE	UGAAAGACAG	ACUUUAUAAA	GGAUGUUUC	AUAAGGAUUU	CUAUCCCAA	CUUGAUCUGA
Mopeia	AGAAAGACAU	AGUUUGUAGA	GGAGAUUUUC	AUAAGGGUUC	CUUUCCCUG	GCGUGUCAGA
Cons. Est.	.GA.A..CA.	A..UU.UA.A	G.A..UU.UC	.UA.GG.UU.	CU....CC..
Consenso	.GA.ArrCA.	A.yUUrUA.A	G.A.rUUyUC	rUArGGrUUy	CU....CC.ry...

2820

Junín MC2	AAUAAACAUG	CCUUCUCUUC	GUUUAACCUG	AAUGGUUGAU	UUUAUGAGGG	AAGAAAAGUU
Junín CD1	AAUAAACAUG	CCUUCUCUUC	GUUUAACCUG	AAUGGUUGAU	UUUAUGAGGG	AAGAAAAGUU
Oliveros	GACAAACAUU	CUUUCUUGAU	GCUUCACCCU	AAGGAGGGAU	UUGAUGAUGU	CAGCAAACAC
Pichindé	CACAAACAUA	CCCUCACGGU	UCUUAUCUG	CAGUAGACUU	UUUAUGAGGG	UGCUAACAUC
Sabiá	UACAAACAUG	CCUUCUUUUU	UCUUAUGGA	UAGGGUUGUC	UUUAUGAUCU	GAGAAAAGUU
Tacaribe	AAUAAACAUA	CCCUCUUUC	UCUUGACUUG	UAAGGUGUUU	UUGAUCAGGC	UAGAGAAGUU
Lassa Jos.	AAUGAACAUU	CCAAGAGCCU	UCUUGACCUU	UAAAUGGAU	UUGAGGAUAC	CAUCCAUIUGU
Lassa Nig.	UACAAACAUU	CCCAGACUCU	UCUUAACUUU	CAAGAUUGAC	UUCAAGAUAC	CAUCCAUGGU
LCM Arm.	AACAAACAUG	UUGAGUUUUC	UCUUGGCCCC	GAGAACUGCC	UUCAAGAGGU	CCUCGCUGUU
LCM WE	GACAAACAUG	UUGAGUUUCU	UCUUGGCCCC	AAGGACUGCU	UUUAGGAGAU	CCUCACUAUU
Mopeia	GACAAACAUU	CCCAAAGAAC	GCUUGACUUU	CAAGGCAGCC	UUCAAAUIUC	CUUCCAAAUU
Cons. Est.	.A..AACAU.UU.....	.A.....	UU.A..A...
Consenso	.AyrAACAU.	yy.....y	.yUU.ry...	.A.r....y	UU.A..A...y

2880

Junín MC2	AUCUGGGGUG	ACUCUGAUUG	UCUCCAACAU	AUUUCCAUCA	UCAAGAAUGG	AUGCACCAGC
Junín CD1	AUCUGGGGUG	ACUCUGAUUG	UCUCCAACAU	AUUUCCAUCA	UCAAGAAUGG	AUGCACCAGC
Oliveros	AUCUGGGCUUA	ACAUAGAUUG	UUUCCAACAU	GUUUCUUCCA	UCAACCAGAC	AUGCACCAGC
Pichindé	AGAGGGCUUC	ACUUGGAUGG	UCUCAAGCAU	GUUGCCACCA	UCAAUCAUAC	AAGCAGCUGC
Sabiá	GUCAGGAGUG	AUCCUUAUGG	UUUCCAGCAU	GUUACCACCA	UCCAGAAGCG	UUGCACCAGC
Tacaribe	GUCGGGACUC	ACUCUGAUUG	UCUCCAACAU	GUUACCACCA	UCAAGGAUGG	AGGCUCCUGC
Lassa Jos.	CUGAGGUGAC	ACCUUGAUUG	UCUCCAACAU	AUUGCCACCA	UCCAGCAUGC	AAGCUCCUGC
Lassa Nig.	CUGAGGUGAA	ACCUUAAUAG	UCUCUAACAU	GUUACCACCA	UCAAGCAUGC	AGGCCUCCUGC
LCM Arm.	GCUUGGGCUUG	AUCAAAAUUG	ACUCUAACAU	GUUACUUCCA	UCCAACAGGG	CUGCCCCUGC
LCM WE	GCUUGGGUUG	AUCAAAAUAG	AUUCAGCAU	GUUCCUCCA	UGUAGCAGAG	CUGCCCCUGC
Mopeia	GGAUUUUGUU	ACUUUGAUGG	UCUCCAGCAU	GUUUCGCGG	UCUAUUAGAC	AAGCACCAGC
Cons. Est.U.	A.....AU.G	..UC.A.CAU	.UU.CC..C.	UC.A..A...	..GC.CC.GC
Consenso	r.....U.	Ay....AU.G	.yUC.ArCAU	rUU.CC.yCr	UC.A..A...	..GC.CC.GC

Apéndice 1

2940

Junín MC2	CUUUACUGCA	GCUGAAAGAC	UAAAGUUGUA	GCCAGAAAUG	UUGAUGGAGC	UUUCAUCCUU
Junín CD1	CUUUACUGCA	GCUGAAAGAC	UAAAGUUGUA	GCCAGAAAUG	UUGAUGGAGC	UUUCAUCCUU
Oliveros	UUUCACUGCA	GCAGCAAGAC	UCAGAUUGUA	CCCAGAGAUG	UUGUUAGCAC	UUUCCUCUCU
Pichindé	UUUGACUGCU	GCAGACAGAC	UAAGGUUAUA	CCCAGAGAU	UUCAAGGCUG	AUGGCUCAUG
Sabiá	UUUAACAGCU	GCAGAAAGAC	UAUAGUUUAUA	UCCUGAGAUG	UUCAAGCCGC	UCUCAUCUUU
Tacaribe	UUUUACUGCU	GCCAACAGAC	UGAGGUUGUA	ACCUGAAAUG	UUAAUUGAAC	UCUCUUCUUU
Lassa Jos.	CUUCACAGCU	GCACCCAAGC	UAAAAUUUAUA	ACCUGAGAU	UUCAAGAGC	UUUUCUUGGU
Lassa Nig.	UUUGACAGCA	GCACCCAAC	UGAAAUUGUA	ACCAGAGAUG	UUGAGUGAAC	UUUUCUUGU
LCM Arm.	CUUCACGGCA	GCACCAAGAC	UAAAGUUUAUA	GCCAGAAAUG	UUGAUGCUGG	ACUGCUGUUC
LCM WE	UUUCACAGCC	GCACCAAGAC	UGAAAUUAUA	ACCAGAGAU	UUGAUACUAG	AUUGCUGUUC
Mopeia	UUUCACCGCA	GCUGACAGGC	UAAAAUUGUA	CCCUGAUUUG	UUUAUGGCAC	UUUUCUUGGG
Cons. Est.	.UU.AC.GC.	GC....A..C	U....UU.UA	.CC.GA.AU.	UU.....U....
Consenso	yUU.AC.GC.	GC....ArrC	U..rrUUrUA	.CC.GA.AUr	UU.....	.y...U....

3000

Junín MC2	AGUCACAAUC	UGGAGGCAGU	CAUGUCCUG	AGUCAUUCUG	UCAAGGUCAC	UCAAGUUUGG
Junín CD1	AGUCACAAUC	UGGAGGCAGU	CAUGUCCUG	AGUCAUUCUG	UCAAGGUCAC	UCAAGUUUGG
Oliveros	UGUGAUGAAU	UGGAGACAAG	UGUGUUCAGC	AGACAGCUUU	UCAAGAUCAU	CAAGGUUGGG
Pichindé	GCUGAUGAUU	UUUAGGGCGC	UGUGUUGUUG	UGUGAGCUUU	UCCAGAUCUG	UCAUGUUCGG
Sabiá	AGUCACUAUU	UGAAGACAGU	CAUGUUCUGC	UGUAAGUUUU	UCUAGGUCAA	UCAGGUUGGG
Tacaribe	GGUUUAUUUU	UGCAGACACU	CAUGGUUUGG	GAUGAGCUUG	UCAAGAUCAG	AAAGAUUUGG
Lassa Jos.	GUCAAUCAUA	UUUAGGAUGG	GAUGACUUUG	AGUCAGCCUG	UCUAAGUCUG	AAGUGUUGGG
Lassa Nig.	GUCAAUCAUA	UUCAGAAUUG	GAUGACUCUG	AGACAAUCUG	UCGAGAUCAG	AUGAGUUGGG
LCM Arm.	AGUGAUGACC	CCCAGAACUG	GGUGCUUGUC	UUUCAGCCUU	UCAAGAUCAU	UAAGAUUUGG
LCM WE	AGUAAUGACC	CCCAGAACUG	GGUGUUUAUC	UUUUAGCCUU	UCUAGGUCAC	UGAGAUCGG
Mopeia	CUCAAUGAUC	CCCAAGAUUG	GGUGGGUCUG	AGAGAGUUUG	UCUAAGUCAG	ACAUGUUUGG
Cons. Est.A..A..	...A.....	..UG.....A...U.	UC.A..UC..UU.GG
Consenso	..y.Ay.A..	y..Arr....	.rUG.....AryyU.	UC.ArrUC..	..r.rUU.GG

3060

Junín MC2	AUACUUCACA	GUGUAUAGAA	GCCCAAGAGA	GGUAAAAGCC	UGUAUGACAC	UGUUCAUUGU
Junín CD1	AUACUUCACA	GUGUAUAGAA	GCCCAAGAGA	GGUAAAAGCC	UGUAUGACAC	UGUUCAUUGU
Oliveros	GUAUUUGACU	GUGUAAAGGA	GCCCAAGUGU	GGUCAAAACCU	UGUACCACAU	CAUUCAUCUG
Pichindé	GAACUUGACA	GUGUAAAGCA	AACCAAGUGC	ACUCAGUGCU	UGGACAACAU	CAUUAAGCUG
Sabiá	GAACUUAACU	GUAUAAAGAA	GUCCCAAAGA	UGUUAGUGCC	UGAACAAACAU	CCCCCAUGGU
Tacaribe	GUAUUUCACU	GUAUAAAGAA	GGCCUAAAAGA	GGUCAUUGCU	UGAACUACAU	UGUUCAUAGU
Lassa Jos.	AUACUUUGCU	GUGUAGAUA	AACCCAAAUC	UGUCAUUGCU	UGUACUGCAU	CAUUCAAGUC
Lassa Nig.	GUAUUUGGCU	GUGUAAAUCA	GCCCUAAAUC	UGUCAAAAGCU	UGAACGGCAU	CAUUCAGGUC
LCM Arm.	AUACUUGACU	GUGUAAAGCA	AGCCAAGGUC	UGUGAGCGCU	UGUACAACGU	CAUUGAGCGG
LCM WE	GUAUUUGACU	GUGUAAAGUA	AGCCAAGGUC	UGUGAGUGCC	UGCACAACAU	CAUUGAGUGG
Mopeia	AUACUUUGCA	GUGUACACAA	GCCCAAGUC	UGACAACGAU	UGGAUCACAU	CAUUCAGAUC
Cons. Est.	..A.UU..C.	GU.UA.A..A	..CC.A....A.....	UG.A...C..A....
Consenso	r.AyUU.rC.	GUrUA.A..A	r.CC.Ar...Ar...y	UG.Ay.rCry	y.yy.A....

3120

Junín MC2	CUCACCUCCU	UGAACAGUCA	UGCAUGCAAU	UGUCAUUGCA	GGAACAGAAC	CAAACUGAUU
Junín CD1	CUCACCUCCU	UGAACAGUCA	UGCAUGCAAU	UGUCAUUGCA	GGAACAGAAC	CAAACUGAUU
Oliveros	UUCAGCUCCU	UGCACUGACA	UGCAAGCAAU	UGUCAAAAGCA	GGCAUUGAUC	CAAACUGAUU
Pichindé	CUCACCCCCU	UGCUCUGUCA	UACAGGCGAU	GGUCAAGGCU	GGCAUUGAUC	CAAUUGAUU
Sabiá	AUUGUCACUU	UGUUUAGUCA	UACAAGCGAU	UGUCAGUGCA	GGGACAGAUC	CAAUUGAUU
Tacaribe	CUCCCCUCCU	UGGACAGUCA	UACAUGCAAU	UGUUAAUGCA	GGCAUGGAUC	CAAUUGAUU
Lassa Jos.	AACCUGCCCC	UGUUUUGUCA	GACAUGCCAG	UGUCAGACUU	GGCAUGGUCC	CGAACUGAUU
Lassa Nig.	CACUUGCCCC	UGUUUGGUCA	GGCAUGCUAA	AGUUAGGCUU	GGCAUUGUUC	CGAACUGAUU
LCM Arm.	AGUCUGUGAC	UGUUUGGCCA	UACAAGCCAU	AGUUAGACUU	GGCAUUGUGC	CAAUUGAUU
LCM WE	GGUCUGUGAC	UGUUUGGCCA	UGCAAGCCAU	UGUCAGGCUU	GGCAUUGUGC	CGAACUGAUU
Mopeia	UGCCUGCCCC	UGUUUGCACA	UGCAAGCAAU	CGUCAGGCUU	GGCAUUGUCC	CAAACUGAUU
Cons. Est.UG.....	CA	U.CA.GC.A.	.GU.A....	GG.A..G..C	C.AA.UGAUU
Consenso	..y.....y	UG..y...CA	UrCA.GC.A.	.GUyAr..y.	GG.Ay.G..C	CrAAyUGAUU

3180

Junín MC2	GUUAAGUUUU	GAAGGAUCUU	UAACAUCCCA	UACCCUCACC	ACACCAUUUC	CCCCAGUUC
Junín CD1	GUUAAGUUUU	GAAGGAUCUU	UAACAUCCCA	UACCCUCACC	ACACCAUUUC	CCCCAGUUC
Oliveros	GAUCAAUUGC	UUCUGAUCCU	UAAUGUCCCA	GAUCCUUACC	ACACCGUUGU	UUUGACCUU-
Pichindé	AAUCAAUAGU	GUGUUAUCCU	UGAUGUCCCA	GACCUUCACA	ACACCAUCUC	UGUUACCUGC
Sabiá	AUUCAGCCUG	UUCGGAUCAG	CUACGUCCCA	AAUCCUUACA	ACCCUGGAG	ACCUCACUUG
Tacaribe	GUUCAGUUUG	CUGGAAUCUU	UCACAUCCCA	AACUCUAACA	AUUCCAUUGG	GACGACCUCC
Lassa Jos.	AUUGAGCAAC	UCUGCAUUUU	UCACAUCCCA	AACUCUCACC	ACUCCAUCUC	UCCAGCCCG
Lassa Nig.	GUUGAGUAAC	UCUGCAUUUU	UGACAUCCCA	AACUCUCACA	AUGCCAUCAC	UGGCACCCUU
LCM Arm.	GUUCAAAAGU	GAUGAGUCUU	UCACAUCCCA	AACUCUUACC	ACACCA---C	UUGCACCCUG
LCM WE	GUUCAGAAGU	GAUGAGUCCU	UCACAUCCCA	AACCCUUACU	ACACCA---C	UUGCACCCUG
Mopeia	GUUUAAAAGC	UCUGGGUUUC	GAACAUCCCA	CACUCUAACA	AUGCCAUCUC	CGAGAGUGUU
Cons. Est.	..U.A.....U...	..A..UCCCA	.A...U.AC.	A..CC.....
Consenso	r.U.Ar....rUy..	..AyrUCCCA	.AyyyU.AC.	Ay.CC.....Y...

3240

Junín MC2	-----UUG-	---CUGUUG-	-AAAUCCAG	UGUUCUCAU	AUCUCUGAUC	UCUUGGCCAG
Junín CD1	-----UUG-	---CUGUUG-	-AAAUCCAG	UGUUCUCAU	AUCUCUGAUC	UCUUGGCCAG
Oliveros	----GUCUG-	---G-ACUG-	-AAACCCAAG	GGAUCUCAGU	AUAGCUGAGC	GCCUGUCCAG
Pichindé	GG--GUCUG-	---GCAUUAG	CAAACCCAU	GGAGCGGAGA	AUCUCUGCCC	UUUGCUCCAG
Sabiá	-----CUG-	---CUGUUG-	-AAAACCAAG	UUUCCUUAAA	AGAUCAGACC	UUUUGGUGAG
Tacaribe	-----CUG-	---CUGAGC-	-AAAUCCGAG	UGUUCUCAAA	AUCUCACUCC	UCCUGUUGAG
Lassa Jos.	AGCCCCUUGA	UUACCACCAC	UCAUUCUUAU	CAUAUUCAGG	AGAGCUCUUC	UUUGGUCAAG
Lassa Nig.	UCCCCUCCA	CUUACUCCAG	UCAUGCCAAU	CAUGUUCAAA	AGGGCUCUCC	UUUGAUCAAG
LCM Arm.	-----CUGA	GGCUUUC---	UCAUCCCAAC	UAUCUGUAGG	AUCUGAGAUC	UUUGGUUAG
LCM WE	-----CUGA	GGCUUUC---	UCAUCCCAAC	CAUUGCAGU	CUUUGGGAUC	UCUGAUCAAG
Mopeia	UCCUGCCCGG	AACCCGCCAC	CCAUUCCGAC	CAUCUGUAGC	AAAACCCUCC	UCUGUUCUAG
Cons. Est.G.A..CC.A.A..CAG
ConsensoyyG.A..CC.A.y..Ar.C	.yy...y.AG

3300

Junín MC2	UUGUGACUGA	GACAAGUUAC	CCAUGUAAAC	CCCUUGAGAG	CCUG---UCU	CUGCUCUUCU
Junín CD1	UUGUGAUUGA	GACAAGUUAC	CCAUGUAAAC	CCCUUGAGAG	CCUG---UCU	CUGCUCUUCU
Oliveros	UUGAGCAGUU	GUGAGAUUCC	CUUGGUAAAC	ACCAGG--GG	UGCUC-CUUU	CAUUCGUGUU
Pichindé	CUGAGUGUUU	GUAAGAUUGC	CUUCAUAGAC	ACCAGGCUGA	GACAGGCCUU	CGGUUCUGGU
Sabiá	CUGUGAUUGG	GAAAGGUUUC	CCAGGUACAC	ACCUCCUGAG	GCUGAUUUU	CUGUUUGUCC
Tacaribe	UUGAGAUUGU	GGCAGAUUUC	CCAUGUAAGC	AUUGGUACCA	UUGC---UCU	CAGUUCUUUU
Lassa Jos.	UUGCUGUGAG	CUUAGGUUGC	CCAUAUAGAC	ACCUGCACUU	AAUGGCCUUU	CUGUUCUGAU
Lassa Nig.	CUGUUGUGAA	CUCAAAUUCC	CCAUUAUAAAC	UCCUGAACUC	AAAGGCCUUU	CUGUUCUGGU
LCM Arm.	UUGCUGUGUU	GUUAAGUUCC	CCAUUAUUAAC	CCUGAAGCC	UGGGGCCUUU	CAGACCUCAU
LCM WE	UUGUUGUGCU	GUCAAAUUUC	CCAUGUAGAC	UCCAGAAGCU	UGAGGCCUCU	CAGUUCUCAU
Mopeia	CUGCUGAGCU	GUUAGGUUGC	CCAUUAUAGAC	UCCAGCAGCA	AGAGGCCUCU	CACCUCUGAU
Cons. Est.	.UG.....	...A..UU.C	C....UA..CU	C.....
Consenso	.UG.....	...ArrUU.C	Cy..rUA.rC	.yy.....yyU	C...y...y

3360

Junín MC2	AAACUUGGUU	UUUAAAUUC	CAAGGCCAGA	CGCCAACUCC	AUCCGCUCAA	CCCUCCCCAG
Junín CD1	AAACUUGGUU	UUUAAAUUCU	CAAGGCCAGA	CGCCAACUCC	AUCCGCUCAU	CCCUCCCCAG
Oliveros	GAUCUUCUC	UUCAGUUUUU	CCAGAUAGU	UGCUAACUCC	AUGAGUUCAU	CCUUGGAAAG
Pichindé	GACUUUCUUC	CUUAAUUCU	CCAAUUCAGU	UGCAAGCUCC	AUGAGCUCCU	CUUAGCUAA
Sabiá	AACUUUUCU	UUCAGUUUCU	CCAGGUCUGA	UGCAAGAUC	AUCAGUUCAG	AUUUGUUGAG
Tacaribe	UAUUUUGUCU	UUUAGCUUUU	CAAGAUCCGA	GGCCAGUUC	AUUAGUUCAU	CCCUUGCCAG
Lassa Jos.	CACCUUUGAC	UUUAAUUCU	CUAGAUAGC	GGCUAAGAUU	AAUAAGUCAU	CUGAGGUUAG
Lassa Nig.	GACUUUUGAU	UUCAGUUUUU	CUAAUUCAGC	GGCCAGGAUU	AGUAGAUCCU	CUGAACUAA
LCM Arm.	GAUCUUGGCC	UUCAGCUUCU	CAAGGUCAGC	CGCAAGAGAC	AUCAGUUCU	CUGCACUGAG
LCM WE	AAUUUUGGCC	UUCAGCUUCU	CAAGAUAGC	UGCAAGGGUC	AUCAAUUCU	CUGCACUAG
Mopeia	GACUUUUGCU	UUGAGUCUGU	CCAGGUCGGC	AGCAAGGACG	AGGAGGUCAU	CAGAGGUUAG
Cons. Est.	.A..UU....	.U.A...U..	C.A...C.G.	.GC.A.....	A.....UC..A.
Consenso	.A.yUU...y	yU.Ar.yU.y	C.ArryC.G.	.GC.Ar....	A...r.UC..Ar

Apéndice 1

3420

Junín MC2	AUCUCCCGCC	UGAAAACCG	UGUUUCGUUG	AACACUCCUC	AUGGACAUGA	GUCUGUCAAC
Junín CD1	AUCUCCCGCC	UGAAAACUG	UGUUUCGUUG	AACACUCCUC	AUGGACAUGA	GUCUGUCAAC
Oliveros	ACCUCCAACC	UUCAGUGAAA	CAUUGCUUUG	UUUGUUUCUC	AUCAUCAUCA	AACCAUCAAC
Pichindé	GCCUCCACU	UUAAGUACAU	UAUCUCUUUG	GACUGAUCUC	AUAUUCAUCA	GAGCAUCAAC
Sabiá	AUCACCAAGU	UUUAAGAUUG	UGUCUCUUUG	GGAACUCUUC	AUGCUCAUCA	GGUUGUCCAC
Tacaribe	GUCACCAACU	UUGAAGAUUG	UGUUCUUCUG	AACACUUUUC	AUAGACAUCA	GCCUGUCGAC
Lassa Jos.	AGUCCCAACU	CUCAGUAUAC	UCUUUUGUUG	AGUUGAUUUU	AAUUCAACAA	GAUUGUUGAC
Lassa Nig.	GGUUCCAACU	CUUAAGACAC	UUUUCUGCUG	UGUGGAUUUG	AGCUCAACAA	GAUUGUUGAC
LCM Arm.	CCUCCCCACU	UUCAAAACAU	UCUUCUUUGA	UGUUGACUUU	AAAUCCACAA	GAGAAUGUAC
LCM WE	UCUCCCCACU	UUCAUAACAU	UUUUCUUUGA	UGUAGACUUC	GAUCAACAA	GAGAAUGCAC
Mopeia	UGUCCCCACU	CUCAAAAACAU	UCUUCUGUUG	GACUGACUUU	AACUCAACUA	GAUUGUUCAC
Cons. Est.CC..C.	.U.A.....	..U.....U.A..AU..AC
Consenso	..y.CC.rCy	yU.A..r...	y.Uy.y.y.ryyU.	r.....Ay.A	r...rU..AC

3480

Junín MC2	CUCUUUAUUC	AGGUCCCUCA	ACUUAUUGAG	GUCUUCUUC	CCCCUUUUG	UCUUUCUGAG
Junín CD1	CUCUUUAUUC	AGGUCCCUCA	ACUUAUUGAG	GUCUUCUUC	CCCCUUUUG	UCUUUCUGAG
Oliveros	UGC UUUGUUA	AGAUCUCUGA	GCUUCAGGAG	AUCAUCAUCA	GACCUCCGUU	CUUUCUCUUC
Pichindé	CUCUUUGUUC	AUGUCCCUCA	GUCUGGUCAG	AUCAGAAUCA	GUCCUCUUUAU	CUUUGCGCAC
Sabiá	UUCUUUAUUU	AGGUCUCUCA	GUUUAUCAAG	AUCUCCAUCA	UUCCUUUUUAG	ACUUUCUAAG
Tacaribe	CUCUAUAUUG	AGAUCUCCUCA	AUUUGUUCAA	GUCAUCAUCG	GUCCUUUUUGG	UCUUUCUAAG
Lassa Jos.	CGCUUGAUUU	AGGUCCCUCA	ACCGUUUCA	AUCAUUGUCA	UCCCUUCUCU	CCUUGCGCAU
Lassa Nig.	UGC UUUGAUUG	AGAUCUCUCA	GUCGUUUUAG	GUCGCCAUCA	UCUCUUUUUCU	GCUUGCGCAU
LCM Arm.	AGUCUGGUUG	AGACUUCUGA	GUCUCUGUAG	GUCUUUGUCA	UCUCUCUUUU	CCUUCUCUUC
LCM WE	AGUCUGGUUA	AGACUCCUGA	GUCUCUGCAA	GUCUUUAUCC	UCCCUCCUUU	CCUUUCUCAU
Mopeia	UGC CUGGUUU	AGGUCCCUCA	AUCUUUUCAG	GUCAGAGUCA	UCCCUUCUCU	CCUUUCUCAU
Cons. Est.UU.	A.....CU.AA.	.UC....UC.	...CU.....	..UU.C..A.
Consenso	..yy..rUU.	A.ryyyCU.A	ryy.....Ar	rUC....UC.	..yCUyy...	.yUU.C..A.

3540

Junín MC2	UGCCCGCUGC	ACCUGUGCCA	CUUGGUUGAA	GUCAAUGCUG	UCAGCAAUUA	GCUUGGCAUC
Junín CD1	UGCCCGCUGC	ACCUGUGCCU	CUUGGUUGAA	GUCAAUGCUG	UCAGCAAUUA	GCUUGGCAUC
Oliveros	CACUCGCUGA	ACUUGCGAGA	CUUGACUGAA	GUCAAGUCCG	UCCACCAGAG	ACUUUGCAUC
Pichindé	CAUUCUUUGG	ACUUGAGCAA	CUUUGUGAAA	GUCAAGAGCA	GACAACAGUG	CUCUUGUGUC
Sabiá	GAUCCUUUGA	ACAAGAGAGA	CUUGAUUGAA	AUCAAGACCA	UCAAGUAUCA	UUUUGGCAUC
Tacaribe	CACUCGUUGA	ACUUGAGCAA	CCUGGUUGAA	GUCAAUGCUG	UCGGCAAUCA	GCUUAGCGUC
Lassa Jos.	CAACCGUUGA	ACAUUACUGA	CUUCGGAGAA	GUCAAGUCCA	UGUAAAAGAG	CCUGGGCAUC
Lassa Nig.	CAAUCUCUGA	ACAUUACUGA	CCUCGGAGAA	GUCAAGACCA	UGAAGGAGAG	CUUGAGCGUC
LCM Arm.	GAUCCUCUGA	ACAUUGCUGA	CCUCAGAGAA	GUCCAACCCA	UUCAGAAGGU	UGGUUGCAUC
LCM WE	GAUCCUCUGA	ACGUUGCUGA	CUUCAGAAAA	GUCCAACCCA	UUUAGAAGAC	UGGUUGCGUC
Mopeia	CAACCUUUGA	ACAUUGGCAA	CUUCAGAGAA	GUCCAGCCCA	UGAAGAAGUG	CUUGAGCAUC
Cons. Est.C..UG.	AC....G...	C.U.....AA	.UC.A.....A...G..UC
Consenso	.r.yC.yUG.	AC....G...	CyU.r...rAA	rUC.A...yrA...GyrUC

3600

Junín MC2	CUUCAGAACA	UCCGACUUGA	CAGUCUGAGU	AAAUUGACUC	AAACCUCUCC	UUAAGGACUG
Junín CD1	CUUCAGAACA	UCCGACUUGA	CAGUCUGAGU	AAAUUGACUC	AAACCUCUCC	UUAAGGACUG
Oliveros	ACGCAAGACA	UCUCCUUUGG	UUUGAGUGGU	GUAGUGACUC	AGGCCUCGCC	UCAGUGACUG
Pichindé	UGACAGCACA	UCAGCCUUCA	CAGGAUGGGU	CCAGUUGGAC	AAGCCCCUCC	UAAGGGAUUG
Sabiá	CCUCAGAACA	UCGGUCUUCA	CCGGUGUUGU	GAACUCACUG	AGCCCUCUUC	UCAGGGAUUG
Tacaribe	CUUUAAAUA	UCUGACUUGA	CUGUUUGGGU	GAAUUGACUC	AAUCCUUUC	UAAGAGACUG
Lassa Jos.	UUUCACCACC	UGUAGUUUGA	UGUUGGAGCA	GUAACCAGAU	AAUCCCUCC	UCAAGAUUG
Lassa Nig.	UUUAACUACC	UGCAACUUUA	UGUUGGAACA	GUAGCCAGAU	AGUCCCUUC	UUAGGGAUUG
LCM Arm.	CUUAAUGACA	GCAGCCUUCA	CAUCUGAUGU	GAAGCUCUGC	AAUUCUCUUC	UCAUGCUUG
LCM WE	CUUGAUGACG	GCAGCCUUUA	CAUCUGAUGU	AAAACCCUGC	AACUCCUCC	UCAACGCCUG
Mopeia	CUUGAUGACU	UGGAUCUUUA	UGUUGGAGCA	GUACCCUGAG	AGUUCUCC	UCAGGCUCUG
Cons. Est.A..A..UU..A.....	A...C....C	U.A.....UG
ConsensoA..Ay.yUU.r	y.....	..A.y.....	Ar.yCyy.yC	U.Ar...yUG

						3660
Junín MC2	AGUCCAUCUA	AAGCUUGGAA	CCUCUUUGGA	GUGUGC	CAUG	CCAGAAGAUC -UGGUGGUU-
Junín CD1	AGUCCAUCUA	AAGCUUGGAA	CCUCUUUGGA	GUGUGC	CAUG	CCAGAAGAUC -UGGUGGUU-
Oliveros	AGUCCAUCUG	AAGCUGGGUA	UGUCACUUGC	ACCAGA	CAUU	GUGCGAGACA AUGGUGAGAG
Pichindé	CACCCAGC-G	GAAC--GAUG	GGAUUU---	GUCGGA	CAUU	UUG-GAGU-- -UGUUCACAC
Sabiá	AGUCCAUCUG	AAGCUGGGGA	UUUCCUUUGA	GUUGC	CAUU	GUGGAAAA-- -----
Tacaribe	GGUCCAUCUG	AAGCUUGGCA	CUUCCUUGGA	UUGAGC	CAUU	AUGGC-GAUC -AAA--GCA-
Lassa Jos.	UGUCCACAAA	AAGGAUUUUA	UUUCCUUUGA	GGCACU	CAUC	GCCA-GAUUG UUGUGUUGUA
Lassa Nig.	AGUCCACAAG	AAUGACCUCA	CUUCCUUGGA	AGCACU	CAUU	GUCGUGAUGG UUGUCUGACC
LCM Arm.	CGUCCAUUGG	AAGCUCUUA	CUUCCUUUGA	CAAGGA	CAUC	-----UUG UUGCUCAAU-
LCM WE	UGUCCACUGA	AAGCUUUUGA	CUUCCUUGGA	CAAAGA	CAUU	-----UUG UCACACAAU-
Mopeia	UGUCCACAAG	AAGGACUUCA	CUUCCUUUGA	AUUGGA	CAUU	GU--UGGAGG CUUUUCGGU-
Cons. Est.	...CCA....	.A.....U...CAU.CAU.
Consenso	.ryCCA...r	ra.....r	...y.yU...CAU.CAU.

						3720
Junín MC2	UGAUCUGAG	--AAAA----	AAUUGCUCAG	UGAA-----		-AGUGUUAGA CACUAU----
Junín CD1	UGAUCUGAG	--AAAA----	AAUUGCUCAG	UGAA-----		-AGUGUUAGA CACUAU----
Oliveros	GUGUCGUGUG	ACAAAA----	GGUUUGUAAA	AGAGCA----		AAGCGUUAUC CUAUAU----
Pichindé	UUCCUCUGAG	UCAGUA----	AAGAAGUGAA	CG--UA----		CAGCGUGAUC UAGUGUC---
Sabiá	--GACUUG--	-----	--CUGACCAG	UAAA-----		-GG---UAGA CAAUUU----
Tacaribe	AGGAUUUCAA	--AAAA----	GA--GCUCAG	UGAA-----		-AGAGUUAGA CAAUUU----
Lassa Jos.	UGCACGCAAC	AAAGAACUGA	GACUA-UCUG	CCAAAUGAC	AAAAGCAAAG	CGCAAUCCAA
Lassa Nig.	CUUGAGUGGG	UCUUGAAUGU	GGUCACUCCA	AAGGUUUGAU	UAGUGCAAAG	CGCAAUCCAA
LCM Arm.	-----	----GGUU--	-----UCUCA	A-----GAC	AA-----AUG	CGCAAUCAAA
LCM WE	-----	----GAUU--	-----UUCCA	A-----AU	AA-----AAG	CGCAAUCAAA
Mopeia	-----GCAG-	----AGCUG-	----AUUCCA	-----	-----AAGG	CGCAAUCAAU
Cons. Est.A..U....
Consensoy..rA..	y....U....

			3741
Junín MC2	--GCCUAGGA	UCCACUGUGC	G
Junín CD1	--GCCUAGGA	UCCACUGUGC	G
Oliveros	--GCCUAGGA	UCCACUGUGC	G
Pichindé	--GCCUAGGA	UCCACUGUGC	G
Sabiá	--GCCUAGGA	UCCACUGUGC	G
Tacaribe	--GCCUAGGA	UCCACUGUGC	G
Lassa Jos.	UAGCCUAGGA	UCCACUGUGC	G
Lassa Nig.	UAGCCUAGGA	UCCACUGUGC	G
LCM Arm.	U-GCCUAGGA	UCCACUGUGC	-
LCM WE	U-GCCUAGGA	UCCACUGUGC	G
Mopeia	UAGCCUAGGA	UCCACUGUGC	-
Cons. Est.	..GCCUAGGA	UCCACUGUGC	.
Consenso	..GCCUAGGA	UCCACUGUGC	.

Figura Ap1.1. Alineamiento múltiple de los RNAs S de arenavirus. El alineamiento múltiple fue realizado con el Clustal V, de manera independiente para las regiones codificantes y las no codificantes. Las regiones codificantes fueron alineadas utilizando los siguientes parámetros: *fixed gap penalty*=10, *floating gap penalty*=10, *toggle transitions=weighted*. Las regiones no codificantes fueron alineadas en corridas independientes, utilizando los siguientes parámetros: *fixed gap penalty*=7, *floating gap penalty*=7, *toggle transitions=unweighted*. Además, para las regiones intergénicas se tuvieron en cuenta las predicciones de estructura secundaria, alineándolas en bloques independientes de acuerdo a equivalencias estructurales.

Al pie del alineamiento se indica el Cons. Est. (consenso estricto), que tiene en cuenta sólo las identidades, y el Consenso, que tiene en cuenta las identidades y las transiciones.

Con recuadros grisados se destacan los codones de iniciación y terminación de GPC y los tripletes complementarios a los codones de iniciación y terminación de N. En negrita y subrayado se indican las secuencias intergénicas que, potencialmente, forman estructuras tipo *hairpin-loop* (terminadores de la transcripción).

Alineamiento de los RNAs S de las cepas MC2 y Candid #1 del virus Junín

						60
Junín MC2	-UGCAGUAAG	GGGAUCCUAG	GCGAUUUUGG	UAACGCUAUA	AGUUGUUACU	GCUUUCUAUU
Junín CD1	CUGCAGUAAG	GGGAUCCUAG	GCGAUUUUGG	UAACGCUAUA	AGUUGUUACU	GCUUUCUAUU
	*****	*****	*****	***** *	*****	*****
						120
Junín MC2	UGGACAACAU	CAAACCAUCU	AUUGUACA	GGGGCAAUUC	AUCAGCUUCA	UGCAAGAAAU
Junín CD1	UGGACAACAU	CAAACCAUCU	AUUGUACA	GGGGCAAUUC	AUCAGCUUCA	UGCAAGAAAU
	*****	*****	*****	*****	*****	*****
						180
Junín MC2	ACCUACCUUU	UUGCAGGAAG	CUCUGAAUUA	UGCUCUUGUU	GCAGUCAGUC	UCAUUGCCAU
Junín CD1	ACCUACCUUU	UUGCAGGAAG	CUCUGAAUUA	UGCUCUUGUU	GCAGUCAGUC	UCAUUGCCAU
	*****	*****	*****	*****	*****	*****
						240
Junín MC2	CAUUAAGGGU	GUAGUAAACC	UGUACAAAAG	UGGUU-GUUC	-AAUUCUU-G	-AU--CUAGC
Junín CD1	CAUUAAGGGU	GUAGUAAACC	UGUACAAAAG	UGGUUUGUUC	CAAUUCUUUG	UAUUCUAGC
	*****	*****	*****	***** *	***** *	*****
						300
Junín MC2	-----AGGA	AGAUCUGCC	C-GA-G-AGC	UUUUAAAAUC	GGACUGCACA	CAGAGGUUCC
Junín CD1	ACUCGCAGGA	AGAUCUGCA	CAGAAGAAGC	UUUUAAAAUC	GGACUGCACA	CAGAG-UUCC
	****	*****	* * * * *	*****	*****	*****
						360
Junín MC2	AGACUGUGUC	CUUCUCCAUA	GGUGGGUCUC	UUUUCCAAC	AAUCCACAUG	ACCUGCCUCU
Junín CD1	AGACUGUGUC	CUUCU-CAUA	GGUGGGUCUC	UUUUCCAAC	AAUCCACAUG	ACCUGCCUCU
	*****	***** *	*****	*****	*****	*****
						420
Junín MC2	GUUGUGUACC	UAAACAAGA	GCCAUCUUUA	CAUUAAGGGG	GGCAAUGCUU	CAUUCAAGAU
Junín CD1	GUUGUGUACC	UAAACAAGA	GCCAUCUUUA	CAUUAAGGGG	GGCAAUGCUU	CAUUCAAGAU
	*****	*****	*****	*****	*****	*****
						480
Junín MC2	CAGCUUUGAU	GACAUCGCAG	UGUUGUUACC	AGAAUAUGAC	GUUAUAAUUC	AGCAUCCGGC
Junín CD1	CAGCUUUGAU	GACAUCGCAG	UGUUGUUACC	AGAAUAUGAC	GUUAUAAUUC	AGCAUCCGGC
	*****	*****	*****	*****	*****	*****
						540
Junín MC2	AGAUUAGAGC	UGGUGUUCUA	AAAGUGAUGA	UCAAUUUAGG	CUGUCUCAGU	GGUUCAUGAA
Junín CD1	AGAUUAGAGC	UGGUGUUCUA	AAAGUGAUGA	UCAAUUUUGG	CUGUCUCAGU	GGUUCAUGAA
	*****	*****	*****	***** *	*****	*****
						600
Junín MC2	UGCUGUGGGG	CAUGAUUGGU	AUCUAGACCC	ACCAUUUCUG	UGUAGGAACC	GUACAAAGAC
Junín CD1	UGCUGUGGGG	CAUGAUUGGU	AUCUAGACCC	ACCAUUUCUG	UGUAGGAACC	GUACAAAGAC
	*****	*****	*****	*****	*****	*****
						660
Junín MC2	AGAAGGCUUC	AUCUUUCAAG	UCAAUACCUC	CAAGACUGGU	AUCAAUGAAA	ACUAUGCCAA
Junín CD1	AGAAGGCUUC	AUCUUUCAAG	UCAAUACCUC	CAAGACUGGU	AUCAAUGAAA	ACUAUGCCAA
	*****	*****	*****	*****	*****	*****

Apéndice 1

						720
Junín MC2	GAAGUUUAG	ACUGGUAUGC	ACCAUUUUAU	UAGAGAAUAC	CCCGACUCUU	GCUUGGAUGG
Junín CD1	GAAGUUUAG	ACUGGUAUGC	ACCAUUUUAU	UAGAGAAUAC	CCCGACUCUU	GCUUGGAUGG
	*****	*****	*****	*****	*****	*****
						780
Junín MC2	CAAACUGUGU	UUGAUGAAGG	CACAACCCAC	CAGUUGGCCU	CUCCAAUGUC	CACUUGACCA
Junín CD1	CAAACUGUGU	UUGAUGAAGG	CACAACCCAC	CAGUUGGCCU	CUCCAAUGUC	CACUUGACCA
	*****	*****	*****	*****	*****	*****
						840
Junín MC2	UGUCAACACA	UUACAUUUC	UCACAAGAGG	CAAGAACAUI	CAGCUUCCAA	GGAGGUCUUU
Junín CD1	UGUCAACACA	UUACAUUUC	UCACAAGAGG	CAAGAACAUI	CAGCUUCCAA	GGAGGUCUUU
	*****	*****	*****	*****	*****	*****
						900
Junín MC2	AAAAGCAUUC	UUUUCUGGU	CUCUGACAGA	CUCAUCCGGC	AAGGACACCC	CUGGAGGCUA
Junín CD1	AAAAGCAUUC	UUUUCUGGU	CUCUGACAGA	CUCAUCCGGC	AAGGACACCC	CUGGAGGCUA
	*****	*****	*****	*****	*****	*****
						960
Junín MC2	UUGUCUAGAA	GAGUGGAUGC	UCGUUGCAGC	CAAAAUGAAG	UGUUUUGGCA	AUACUGCUGU
Junín CD1	UUGUCUAGAA	GAGUGGAUGC	UCGUUGCAGC	CAAAAUGAAG	UGUUUUGGCA	AUACUGCUGU
	*****	*****	*****	*****	*****	*****
						1020
Junín MC2	AGCAAAAUGC	AAUCUGAAUC	AUGACUCUGA	AUUCUGUGAC	AUGCUGAGGC	UUUUUGAUUA
Junín CD1	AGCAAAAUGC	AAUCUGAAUC	AUGACUCUGA	AUUCUGUGAC	AUGCUGAGGC	UUUUUGAUUA
	*****	*****	*****	*****	*****	*****
						1080
Junín MC2	CAACAAAAAU	GCUAUCAAAA	CCUAAAAUGA	UGAAACUAAG	AAACAAGUAA	AUCUGAUGGG
Junín CD1	CAACAAAAAU	GCUAUCAAAA	CCUAAAAUGA	UGAAACUAAG	AAACAAGUAA	AUCUGAUGGG
	*****	*****	*****	*****	*****	*****
						1140
Junín MC2	ACAGACAAUC	AAUGCGCUGA	UAUCUGACAA	UUUAUUGAUG	AAAAACAAAA	UUAGGGAAUU
Junín CD1	ACAGACAAUC	AAUGCGCUGA	UAUCUGACAA	UUUAUUGAUG	AAAAACAAAA	UUAGGGAAUC
	*****	*****	*****	*****	*****	*****
						1200
Junín MC2	GAUGAGUGUC	CCUUACUGCA	AUUACACAAA	AUUUUGGUAU	GUCAACCACA	CACUUUCAGG
Junín CD1	GAUGAGUGUC	CCUUACUGCA	AUUACACAAA	AUUUUGGUAU	GUCAACCACA	CACUUUCAGG
	*****	*****	*****	*****	*****	*****
						1260
Junín MC2	ACAACACUCA	UUACCAAGGU	GCUGGUUAAU	AAAAACAAC	AGCUAUUUGA	ACAUUUCUGA
Junín CD1	ACAACACUCA	UUACCAAGGU	GCUGGUUAAU	AAAAACAAC	AGCUAUUUGA	ACAUUCUGA
	*****	*****	*****	*****	*****	*****
						1320
Junín MC2	CUUCCGUAAU	GACUGGAUAC	UAGAAAGUGA	CUUCUUAUU	UCUGAAAUGC	UAAGCAAAGA
Junín CD1	CUUCCGUAAU	GACUGGAUAC	UAGAAAGUGA	CUUCUUAUU	UCUGAAAUGC	UAAGCAAAGA
	*****	*****	*****	*****	*****	*****
						1380
Junín MC2	GUAUUCGGAC	AGGCAGGGCA	AAACUCCCUU	GACUUUAGUU	GACAUCUGUU	UUUGGAGCAC
Junín CD1	GUAUUCGGAC	AGGCAGGGUA	AAACUCCUUU	GACUUUAGUU	GACAUCUGUA	UUUGGAGCAC
	*****	*****	*****	*****	*****	*****

Apéndice 1

						1440
Junín MC2	AGUAUUCUUC	ACAGCGUCCC	UCUUCUUUCA	CUUGGUGGGC	AUACCCACCC	AUAGGCACAU
Junín CD1	AGUAUUCUUC	ACAGCGUCAC	UCUUCUUUCA	CUUGGUGGGU	AUACCCUCCC	ACAGACACAU
	*****	***** *	*****	*****	***** **	* * *****
						1500
Junín MC2	CAGAGGCGAG	GCAUGCCCUC	UGCCCCACAG	GCUAAAUAGC	UUGGGUGGUU	GCAGAUGUGG
Junín CD1	CAGGGGCGAA	GCAUGCCCUU	UGCCACACAG	GUUGAACAGC	UUGGGUGGUU	GCAGAUGUGG
	*** *****	*****	**** *****	* * * * *	*****	*****
						1560
Junín MC2	UAAGUACCCC	AAUCUAAAGA	AACCAACAGU	UUGGCGCAGA	GGACACUAAAG	ACCUCCCGAA
Junín CD1	UAAGUACCCC	AAUCUAAAGA	AACCAACAGU	UUGGCGUAGA	GGACACUAAAG	ACCUCCUGAG
	*****	*****	*****	***** *	*****	***** **
						1620
Junín MC2	GGUCCCCACC	AGCCCGGGCA	UUGCCCGGGC	UGGUGUGGCC	CCCCAGUCCG	CGGCCUGGCC
Junín CD1	GGUCCCCACC	AGCCCGGGCA	CUGCCCGGGC	UGGUGUGGCC	CCCCAGUCCG	CGGCCUGGCC
	*****	*****	*****	*****	*****	*****
						1680
Junín MC2	GCGGACUGGG	GAGGCACUGC	UUAACAGUGCA	UAGGCUGCCU	UCGGGAGGAA	CAGCAAGCUC
Junín CD1	GCGGACUGGG	GAGGCACUGC	UUAACAGUGCA	UAGGCUGCCU	UCGGGAGGAA	CAGCAAGCUC
	*****	*****	*****	*****	*****	*****
						1740
Junín MC2	GGUGGUAUA	GAGGUGUAAG	UUCUUCUUA	UAGAGCUUCC	CAUCCAACAC	UGACUGAAAC
Junín CD1	GGUGGUAUA	GAGGUGUAAG	UUCUUCUUA	UAGAGCUUCC	CAUCCAACAC	UGACUGAAAC
	*****	*****	*****	*****	*****	*****
						1800
Junín MC2	AUUAUGCAGU	CUAGCAGAGC	ACAGUGUGGC	UCACUGGAGG	CCAACUAAA	GGGAGCAUCC
Junín CD1	AUUAUGCAGU	CUAGCAGAGC	ACAGUGUGGC	UCACUGGAGG	CCAACUAAA	GGGAGCAUCC
	*****	*****	*****	*****	*****	*****
						1860
Junín MC2	UUAUCUCUCU	UUUUCUUGCU	GACAACCACU	CCAUUGUGAU	GUUUGCAUAG	GUGGCCAAAU
Junín CD1	UUAUCUCUCU	UUUUCUUGCU	GACAACCACU	CCAUUGUGAU	GUUUGCAUAG	GUGGCCAAAU
	*****	*****	*****	*****	*****	*****
						1920
Junín MC2	UUCUCCCAGA	CCUGUUGGUC	GAACUGCCUG	GCUUGUUCUG	AUGUAAGCCU	AAAUCAACC
Junín CD1	UUCUCCCAGA	CCUGUUGGUC	AAACUGCCUG	GCUUGUUCUG	AUGUAAGCCU	AAAUCAACC
	*****	*****	*****	*****	*****	*****
						1980
Junín MC2	AGCUUAAGAU	CUCUUCUUC	AUGGAGGUCA	AACAACUUC	UGAUGUCAUC	GGACCCUUGA
Junín CD1	AGCUUAAGAU	CUCUUCUUC	AUGGAGGUCA	AACAACUUC	UGAUGUCAUC	GGACCCUUGA
	*****	*****	*****	*****	*****	*****
						2040
Junín MC2	GUGGUCACAA	CCAUGUCCGG	AGGCAGCAAG	CCAUACACGU	AACUAAGAAC	UCCUGGCAUU
Junín CD1	GUGGUCACAA	CCAUGUCCGG	AGGCAGCAAG	CCAUACACGU	AACUAAGAAC	UCCUGGCAUU
	*****	*****	*****	*****	*****	*****
						2100
Junín MC2	GCAUCUUCUA	UGUCUUUCAU	UAAGAUGCCG	UGAGAGUGUC	UGCUACCAUU	UUUAAACCCU
Junín CD1	GCAUCUUCUA	UGUCUUUCAU	UAAGAUGCCG	UGAGAGUGUC	UGCUACCAUU	UUUAAACCCU
	*****	*****	*****	*****	*****	*****

Apéndice 1

						2160
Junín MC2	UUCUCAUCAU	GUGGUUUUCU	GAAGCAGUGA	AUAUACUUGC	UACCUGCAGG	CUGGAACAAC
Junín CD1	UUCUCAUCAU	GUGGUUUUCU	GAAGCAGUGA	AUAUACUGCU	UACCUGCAGG	CUGGAACAAC
	*****	*****	*****	*****	*****	*****
						2220
Junín MC2	GCCAUCUCAAA	CAGGGUCAGU	AGCUGGUCCU	UCA AUGUCGA	GCCAAAGGGU	AUUGGUGGGG
Junín CD1	GCCAUCUCAAA	CAGGGUCAGU	AGCUGGUCCU	UCA AUGUCGA	GCCAAAGGGU	AUUGGUGGGG
	*****	*****	*****	*****	*****	*****
						2280
Junín MC2	UCGAGUUUCC	CCACUGCCUC	UCUGAUGACA	GCUUCUUGUA	UCUCUGUCA	GUUAGCCAAU
Junín CD1	UCGAGUUUCC	CCACUGCCUC	UCUGAUGACA	GCUUCUUGUA	UCUCUGUCA	GUUAGCCAAU
	*****	*****	*****	*****	*****	*****
						2340
Junín MC2	CUCAAAUUCU	GACCGUUCUU	UCCCGGUUGU	CUAGGUCCAG	CAACUGGUUU	CCUUGUCAGA
Junín CD1	CUCAAAUUCU	GACCGUUCUU	UCCCGGUUGU	CUAGGUCCAG	CAACUGGUUU	CCUUGUCAGA
	*****	*****	*****	*****	*****	*****
						2400
Junín MC2	UCAAUACUUG	UGUUGUCCCA	UGACCUGCCU	AUGAUUUGUG	AUCUGGAACC	AAUAUAAGGC
Junín CD1	UCAAUACUUG	UGUUGUCCCA	UGACCUGCCU	AUGAUUUGUG	AUCUGGAACC	AAUAUAAGGC
	*****	*****	*****	*****	*****	*****
						2460
Junín MC2	CAACCAUCGC	CAGAAAGGCA	AAGUUUGUAC	AGAAGGUUUU	CAUAAGGGUU	UCUAUUGCCU
Junín CD1	CAACCAUCGC	CAGAAAGGCA	AAGUUUGUAC	AGAAGGUUUU	CAUAAGGGUU	UCUAUUGCCU
	*****	*****	*****	*****	*****	*****
						2520
Junín MC2	GGUUUCUCAU	CAUUAACA	GCCUUCUCU	CGUUUAACCU	GAAUGGUUGA	UUUUUAGAGG
Junín CD1	GGUUUCUCAU	CAUUAACA	GCCUUCUCU	CGUUUAACCU	GAAUGGUUGA	UUUUUAGAGG
	*****	*****	*****	*****	*****	*****
						2580
Junín MC2	GAAGAAAAGU	UAUCUGGGGU	GACUCUGAUU	GUCUCCAACA	UAUUCCAUC	AUCAAGAAUG
Junín CD1	GAAGAAAAGU	UAUCUGGGGU	GACUCUGAUU	GUCUCCAACA	UAUUCCAUC	AUCAAGAAUG
	*****	*****	*****	*****	*****	*****
						2640
Junín MC2	GAUGCACCAG	CCUUUACUGC	AGCUGAAAAGA	CUAAAGUUGU	AGCCAGAAAU	GUUGAUGGAG
Junín CD1	GAUGCACCAG	CCUUUACUGC	AGCUGAAAAGA	CUAAAGUUGU	AGCCAGAAAU	GUUGAUGGAG
	*****	*****	*****	*****	*****	*****
						2700
Junín MC2	CUUUCAUCCU	UAGUCACAAU	CUGGAGGCAG	UCAUGUCCU	GAGUCAUCU	GUCAAGGUCA
Junín CD1	CUUUCAUCCU	UAGUCACAAU	CUGGAGGCAG	UCAUGUCCU	GAGUCAUCU	GUCAAGGUCA
	*****	*****	*****	*****	*****	*****
						2760
Junín MC2	CUCAAGUUUG	GAUACUUCAC	AGUGUAUAGA	AGCCCAAGAG	AGGUUAAAGC	CUGUAUGACA
Junín CD1	CUCAAGUUUG	GAUACUUCAC	AGUGUAUAGA	AGCCCAAGAG	AGGUUAAAGC	CUGUAUGACA
	*****	*****	*****	*****	*****	*****
						2820
Junín MC2	CUGUUCAUUG	UCUCACCUC	UGAACAGUC	AUGCAUGCAA	UUGUCAUUGC	AGGAACAGAA
Junín CD1	CUGUUCAUUG	UCUCACCUC	UGAACAGUC	AUGCAUGCAA	UUGUCAUUGC	AGGAACAGAA
	*****	*****	*****	*****	*****	*****

Apéndice 1

						2880
Junín MC2	CCAAACUGAU	UGUUAAGUUU	UGAAGGAUCU	UUAACAUCCC	AUACCCUCAC	CACACCAUUU
Junín CD1	CCAAACUGAU	UGUUAAGUUU	UGAAGGAUCU	UUAACAUCCC	AUACCCUCAC	CACACCAUUU
	*****	*****	*****	*****	*****	*****
						2940
Junín MC2	CCCCAGUUC	CUUGCUGUUG	AAAUCCAGU	GUUCUCAAUA	UCUCUGAUCU	CUUGGCCAGU
Junín CD1	CCCCAGUUC	CUUGCUGUUG	AAAUCCAGU	GUUCUCAAUA	UCUCUGAUCU	CUUGGCCAGU
	*****	*****	*****	*****	*****	*****
						3000
Junín MC2	UGUGACUGAG	ACAAGUUACC	CAUGUAAACC	CCUUGAGAGC	CUGUCUCUGC	UCUUCUAAAC
Junín CD1	UGUGAUUGAG	ACAAGUUACC	CAUGUAAACC	CCUUGAGAGC	CUGUCUCUGC	UCUUCUAAUC
	*****	*****	*****	*****	*****	*****
						3060
Junín MC2	UUGGUUUUUA	AAUUCCCAAG	GCCAGACGCC	AACUCCAUC	GCUCAACCCU	CCCAGAUUC
Junín CD1	UUGGUUUUUA	AUUUCUCAAG	GUCAGACGCC	AACUCCAUCA	GCUCAACCCU	CCCAGAUUC
	*****	* * * * *	* * * * *	*****	*****	*****
						3120
Junín MC2	CCCGCCUUGA	AAACCGUGUU	UCGUUGAACA	CUCCUCAUGG	ACAUGAGUCU	GUCAACCUCU
Junín CD1	CCCGCCUUGA	AAACUGUGUU	UCGUUGAACA	CUCCUCAUGG	ACAUGAGUCU	GUCAACCUCU
	*****	*****	*****	*****	*****	*****
						3180
Junín MC2	UUAUUCAGGU	CCCUCAACUU	AUUGAGGUCU	UCUUCUCCCC	UUUUAGUCUU	UCUGAGUGCC
Junín CD1	UUAUUCAGGU	CCCUCAACUU	AUUGAGGUCU	UCUUCUCCCC	UUUUAGUCUU	UCUGAGUGCC
	*****	*****	*****	*****	*****	*****
						3240
Junín MC2	CGCUGCACCU	GUGCCACUUG	GUUGAAGUCA	AUGCUGUCAG	CAAUUAGCUU	GGCAUCCUUC
Junín CD1	CGCUGCACCU	GUGCCUCUUG	GUUGAAGUCA	AUGCUGUCAG	CAAUUAGCUU	GGCAUCCUUC
	*****	*****	*****	*****	*****	*****
						3300
Junín MC2	AGAACAUCCG	ACUUGACAGU	CUGAGUAAAU	UGACUCAAAAC	CUCUCCUUA	GGACUGAGUC
Junín CD1	AGAACAUCCG	ACUUGACAGU	CUGAGUAAAU	UGACUCAAAAC	CUCUCCUUA	GGACUGAGUC
	*****	*****	*****	*****	*****	*****
						3360
Junín MC2	CAUCUAAAGC	UUGGAACCUC	UUUGGAGUGU	GC <u>CAUG</u> GCCAG	AAGAUCUGGU	GGUUUUGAUC
Junín CD1	CAUCUAAAGC	UUGGAACCUC	UUUGGAGUGU	GC <u>CAUG</u> GCCAG	AAGAUCUGGU	GGUUUUGAUC
	*****	*****	*****	*****	*****	*****
						3415
Junín MC2	UGAGAAAAAA	UUGCUCAGUG	AAAGUGUUAG	ACACUAUGCC	UAGGAUCCAC	UGUGCG
Junín CD1	UGAGAAAAAA	UUGCUCAGUG	AAAGUGUUAG	ACACUAUGCC	UAGGAUCCAC	UGUGCG
	*****	*****	*****	*****	*****	*****

Figura Ap1.2. Alineamiento múltiple de los RNAs S de las cepas MC2 y Candid #1 del virus Junín. El alineamiento fué realizado con el programa Clustal X, utilizando los siguientes parámetros: *gap opening*=10, *gap extension*=0.1, *transitionsweight*=0.5. Al pie del alineamiento se indica el consenso estricto con asteriscos. Con recuadros grisados se destacan los codones de iniciación y terminación de GPC y los tripletes complementarios a los codones de iniciación y terminación de N. En negrita y subrayado se indican las secuencias intergénicas que, potencialmente, forman estructuras tipo *hairpin-loop* (terminadores de la transcripción). En rojo se indica la zona de mayores diferencias, correspondiente a la región codificante para el final del péptido señal y el principio de la glicoproteína G1.

**Alineamiento múltiple de los RNAs S de bunyavirus
Grupo mayoritario**

60

INKOO.LCC AGTAGTGTGCTCCACTTGAATACTTTGAGAATATATTGAAATAAACTTCTA-----TAA
 JAMEST.LCC AGTAGTGTACTCCACTTGAATACTTTGAGAATATATTGAAATAAACCCCTA-----TAG
 JSU.LCC -----GAATATATTGAAATAAACTTTTA-----TAG
 KEY.LCC GA-----ATACTTTGATATAA-----AC---CTGACATAAAGTCC-A-----T--
 LACRO.LCC ATTAGTGTACTCCACTTGAATACTTTGAAAATAAAATTGTTGTTGACTGTTTTTTACCTAA
 MELAO.LCC GA-----ATTTGTTTTAAATAC-----T---TTGCTATAAGGCCA-A-----T--
 MORRO.LCC AGTAGTGTACTCCACTTGAATACTTTGAAAATATCGTGTTATTACTTGTGCTT--CTGG
 TAHYNA.LCC AGTAGTGTACTCCACTTGAATACTTTGAAAATAGTTTGTCTATTACTTGA-GATA--CCGA
 TRIVITA.LCC AGTAGTGTACXCCACGTGAATACTTTGAAAATACTTTGAAATAAATT-----CTAT
*.....*

120

INKOO.LCC AAGGCAT-ACTTGGT--TGATATGGGAGATTTGGTTTTCTATGATGAAGCATCCACAGGT
 JAMEST.LCC AAGGAAT-ACCTGGT--TGATATGGGAGATTTGGTTTTCTATGATGTCGCATCCACAGGT
 JSU.LCC AAGGAAT-ACTTGGT--TGATATGGGAGATTTAGTTTTTTATGATGTCGCATCCACAGGT
 KEY.LCC ----TAC-ACT--AT--TGCAATGGGTGATTTGGTTTTCTATGATGTCGCATCCACAGGT
 LACRO.LCC GGGGAAATTATCAAGAGTGTGATGTCGGATTTGGTGTTTTTATGATGTCGCATCAACAGGT
 MELAO.LCC ----TAC-A-----T--TGTGATGGGAGATTTGATCTTTTTATGATGTCGCATCAACAGGT
 MORRO.LCC AAAGAATTGACTAGA-GTGTTATGTCGGATTTGGTTTTTTATGATGTCGCATCCACAGGT
 TAHYNA.LCC GAGGAATAACTTGGAAAGTGTGATGTCGATTTGGTTTTTTATGATGTCGCATCCACCGGT
 TRIVITA.LCC AAAAAATTGTTGCATAGTGAGATGTCAGAGCTTGTCTTTTTATGATGCTCCATCAACAGGT
*.....*

180

INKOO.LCC GCAAATGGATTTGATCCAGATGCAGGGTTTGTGGCATTATG-GCTGACCATGGAGAATC
 JAMEST.LCC GCAAATGGATTTGATCCAGATGCAGGGTTTGTGGCATTATG-GCTGACCACGGAGAGTC
 JSU.LCC GCAAATGGATTTGATCCAGATGCAGGGTTTGTGGCATTATG-GCTGACCACGGAGAGTC
 KEY.LCC GCAAATGGATTTGATCCTGATGCAGGGTATGTGGCATTATG-GCTAACCATGGGGAGTC
 LACRO.LCC GCAAATGGATTTGATCCTGATGCAGGGTATATGG-ACTTCTGTGTTAAAAATGCAGAATC
 MELAO.LCC GCAAATGGATTTGATCCAGATGCAGGGTATCTGGCATTTACA-ATTGCGCATGGGGAAGC
 MORRO.LCC GCAAATGGATTTGATCCTGATGCAGGGTATGTGG-ACTTCTGTGCTAAACATGGGGAATC
 TAHYNA.LCC GCAAATGGATTTGATCCTGATGCAGGGTATGTGG-ACTTCTGTATTAAACATGGGGAAGC
 TRIVITA.LCC GCAAACGGATTTGATCCCGATGCAGGGTATGTGGCATTAT-TGCTGCACATGCCGGATC
 *****.*****.*****.*.....*.....*

240

INKOO.LCC TATCAATTTGGCAGCCGTTAGGATCTTCTTCCTTAATGCAGCAAAGCCAAAGCTGCTCT
 JAMEST.LCC TATCAATCTGTCAGCCGTTAGGATCTTCTTCCTTAATGCAGCAAAGCCAAAGCTGCTCT
 JSU.LCC TATCAATCTGTCAGCCGTTAGGATCTTCTTCCTTAATGCAGCAAAGCCAAAGCTGCTCT
 KEY.LCC GATCAGTCTGTCAACCGTTAGGATCTTCTTCCTTAATGCCGCAAAGCCAAAGCTGCTCT
 LACRO.LCC ACTCAACCTTGCTGCAGTTAGGATCTTCTTCCTCAATGCCGCAAAGGCCAAGGCTGCTCT
 MELAO.LCC TATCAATTTGTCAGCCGTTAGGATCTTCTTCCTTAATGCCGCAAAGCCAAAGCTGCTCT
 MORRO.LCC GATTAACCTTGCTGCAGTTAGGATCTTCTTCCTCAATGCCGCAAAGGCCAAGGCTGCTCT
 TAHYNA.LCC AATTAATCTCCATTCCGTTAGGATCTTCTTCCTTAATGCCGCAAAGCCAAAGCTGCTCT
 TRIVITA.LCC GTACGATCTTCTGCTGTTAGGATCTTCTTCCTCAATGCTGCCAAGGCCAAGAATGCTCT
*.....*.....*.....*.....*.....*.....*.....*.....*

300

INKOO.LCC CGCTCGTAAACCGGAGCGGAAAGCTACTCCTAAATTTGGAGAGTGGCAGGTGGAGATTGT
 JAMEST.LCC CGCTCGTAAACCGGAGCGGAAAGCTACTCCTAAGTTTGGAGAGTGGCAGGTGGAGATCGT
 JSU.LCC CGCTCGTAAACCGGAGCGGAAAGCTACTCCTAAGTTTGGAGAGTGGCAGGTGGAGATCAT
 KEY.LCC CACTCGTAAGCCGGAGCGGAAAGCTACACCTAAGTTTGGAGAGTGGCAGGTGGAGATCGT
 LACRO.LCC CTCGCGTAAGCCAGAGAGGAAAGCTAACCTAAATTTGGAGAGTGGCAGGTGGAGTTAT
 MELAO.LCC CTCTCGTAAACCGGAGAGGAAAGCTACTCCTAAATTTGGAGACTGGCAGGTGGAAATTGT
 MORRO.LCC CTCGCGTAAGCCAGAGAGGAAAGCTAATCCTAAATTTGGCGAGTGGCAGGTGGAGATTGT
 TAHYNA.LCC CGCTCGTAAAGCCGGAGAGGAAAGCTAGTCTAAATTTGGAGAGTGGCAGGTGGAGTTCGT
 TRIVITA.LCC CTCGAGAAAACAGAGGGGAAAGGTTAGTATTAAATTTGGCGAGTGGTCCGGTGGAGGTGGT
*.....*.....*.....*.....*.....*.....*.....*

660

INKOO.LCC GGATCCTCAGTACCTAAAGAAGGCACTTAGGCAACGTTATGGCACCTTGACTGCAGAGAA
 JAMEST.LCC GGATCCCCAGTACTTGAAGAAGGCTCTCAGGCAACGCTATGGCACCTTGACTGCAGAGAA
 JSU.LCC GGATCCCCAGTATCTGAAGAAGGCTCTCAGGCAACGCTATGGCACCTTGACTGCAGAGAA
 KEY.LCC GGATGCTCAGTACCTAAAAAAGCTCTGAGACAGCGCTATGGGACATTGACTGCAGATAA
 LACRO.LCC GGACCCTCAATACCTGAAGAAGGCCTTAAGGCAACGCTATGGCACTCTCACAGCAGATAA
 MELAO.LCC GGACCCTCAGTATCTAAAGAAGGCTCTACGACAGCGCTATGGCACCTTGACTGCAGACAA
 MORRO.LCC GGACCCCAATATCTAAAGAAAGCATTAAAGGCAGAGATATGGCTCCTTAACTGCTGAAAA
 TAHYNA.LCC GGACCCCAAGTATCTGAAGAAGCATTGAGGCAGCGCTATGGCTCTCTTACTGCAGACAA
 TRIVITA.LCC GGATGTCCAATACCTGAAAAAAGCTCTCCGTCAAAGATACGGGACTATGACAGCAGATAA

720

INKOO.LCC ATGGATGGCACAGAAGACAGTGCTCATTGCTAAAAGCCTAAAGGATGTTGAGCAGCTTAA
 JAMEST.LCC GTGGATGGCGCAGAAGACAGTGCTCATTGCTAAGAGCTTGAAGGATGTCGAGCAGCTCAA
 JSU.LCC ATGGATGGCACAGAAGACAGTGCTCATTGCTAAGAGCTTGAAGGATGTTGAGCAGCTCAA
 KEY.LCC ATGGATGGCTCAAAAGACATCTATGATTACAAAGAGCCTCAAAGATGTGGAGCAGCTAAA
 LACRO.LCC GTGGATGTCACAGAAGGTTGCAGCAATTGCTAAGAGCCTGAAGGATGTAGAGCAGCTTAA
 MELAO.LCC ATGGATGGCACAAAAACCCTATGATTGCTAAAAGCCTGAAGGATGTAGAGCAGCTGAA
 MORRO.LCC GTGGATGTCCAGAAGACAGGCATGATTGCTAAAAGCCTAAAGGAGGTTGAGCAACTCAA
 TAHYNA.LCC GTGGATGTCGCAGAAAACCTACTGCCATAGCTAAGAGCCTAAAGGATGTTGAGCAGCTTAA
 TRIVITA.LCC GTGGATGTCTACTAAGACAACAGTAATTGCAAAGACCCTGAAGAGAGTTGAGAGCTTCAA

780

INKOO.LCC ATGGGGAAAGAGGAGGACTTAGTGATGCTGCAAGAACCTTCTTGATCAAATTTGGTGTAAA
 JAMEST.LCC ATGGGGAAAGAGGAGGACTTAGTGATGCTGCAAGAACCTTCTTGATCAAATTTGGTGTAAA
 JSU.LCC ATGGGGAAAGAGGAGGCTTCTAGTGATGCTGCAAGAACCTTCTTGATCAAATTTGGTGTAAA
 KEY.LCC GTGGGGAAAGGGTGGTCTCAGCGATACTGCCAGAGCTTTCCTTGCTAAATTCGGCGTGAG
 LACRO.LCC ATGGGGAAAAGGAGGCCTGAGCGATACTGCTAAAACATTCCTGCAGAAATTTGGCATCAG
 MELAO.LCC GTGGGGTAAGGGGGTCTTAGTGATGCTGCTAGAACATTCCTTCAGAAAGTTCCGGTGTAG
 MORRO.LCC GTGGGGCAGAGGTGTTTGGAGCGATACTGCCAGGACTTTCCTTCAGAAATTTGGAATAAG
 TAHYNA.LCC ATGGGGTAGAGGCGCCTTAGCGACACTGCCAAGAACATTCCTGCAGAAATTTGGAATCAG
 TRIVITA.LCC ATGGGGTAAGGGAGGACTAAGTGAGGCAGCCAGAGCATTCCTCTCAAATTTAATGTAAA

840

INKOO.LCC GCTGCCTTAAACAAGAGCAGTGACGAAATTTGAGTAAATTTTGGATTAATAAATTGACAA
 JAMEST.LCC GTTGCCATAAACAATAAGCAATGGCATTTCAGTAA-TATTTGATATATAAATTGGTAA
 JSU.LCC ACTGCCATAAACAAGA-TAGCAATGGCATTTCAGTAA-TTTTCAGTATATAAATTGATAA
 KEY.LCC ACTTCCATGAGTC-----TGATGCTCTCCTGGTAGATTTCGGCATTT-TAA--T--TAG
 LACRO.LCC GCTTCCATAAATATG-GCATGAGG-----CATTCAAATT--AGGTTCTAAATTCATAA-A
 MELAO.LCC ACTTCCATAAAT-----TGATGGTCTCTAATGAATTCGGCATTT-TAAAT--TAA
 MORRO.LCC ACTGCCATAAGCTTG---AGCAGT-----CAGG-ATATT--TGGGACAAAATTTAAACA
 TAHYNA.LCC ACTCCCATAAATGGG---ATTAAG-----TCTG-TAATT--AGGCTCAAATTCAA--A
 TRIVITA.LCC AATACCATGAGTA-----TGATGCTCCCCTCGTGGAATTGGAATTCCAAATTCAA-AA
 ..*.*.*.*.....*

900

INKOO.LCC ATTTTACA-----ATTGGTTAAATTCATTTCTTATAATTGGCTTCCAAAAGGTC--T
 JAMEST.LCC ATTTTAAATTTTAAATTTGGTTCAAATTCATTTTT-ATAATTGGCTTCTAAAAGGTC--T
 JSU.LCC ATTTTAAATCTTGAAATTTGGTCCAAATTCATTTTTTATAATTGGCTCCAAAAGGTC--T
 KEY.LCC GTCT-AAATTTT---A--ATTTTAAATT-----GGC----TTCCAAAAGGTTCCAT
 LACRO.LCC TTTATATA-TGTCAATTTGAT--TAATTGGTTATCCAAAAGGGTTTTCTTAAGGG-----
 MELAO.LCC ATTTTAAATCAC---ATTATATTXAATT-----GGTCA--TTCCAAAAGGTTCTT
 MORRO.LCC ATTGGATATTTCAAATTAAT--TAATTGGCTAACTAAAAGGGTTCTT----AGG-----
 TAHYNA.LCC GTTTTATTGTAATAATTTGGTCATAATTGGCTATCCAAAAGGGTTCCCC---AAG-----
 TRIVITA.LCC TTTTCGTAT--AAATATTTGGTAAATTTG--ACCCAAA-----AAGAGTCCAT
 .*.....*

360
BATAI.LCC GTATTTAATACAAATTTTCCTGGGAACAGGAACAGTCCAGTTCAGACGATGGTCTTACC
GERMIST.LCC GTECTCAATACGAATTTCCCTGGAAACAGGAACAATGCAGTGCCTGACTACGGTCTTACC
MAGUARI.LCC GTATTTAATACAAATTTTCCTGGCAACCGGAACAGTCCAGTTCAGACGATGGTCTTACC
.*.**.*****.*****.***.*****.*.*****.*.*****.***.*.*****

420
BATAI.LCC CTCCACAGACTCAGTGGATTCCCTTGCCAGGTACTTACTTGAGAAGATTCTAAAAGTGAGT
GERMIST.LCC TTCCACCGTATCAGTGGATACCTTGCCCGGTACCTACTTGAAAGTACCTTGCCAGAGACA
MAGUARI.LCC CTCCACAGACTCAGTGGATTCCCTTGCCAGGTACTTACTTGAGAAAATTTAAAAGTGAGT
.*****.*.*****.*****.*****.*****.*****.*****.*****.*****.*****

480
BATAI.LCC GACCCGAAAAAATTGATTATAAAATCAAAAATTGTTAATCCACTTGCAGAAAAAACGGT
GERMIST.LCC GAACCAGAAAAGCTCATAATGAGAACA AAAAATTGTGAACCCTCTTGCTGAGAAGAATGGC
MAGUARI.LCC GACCCAGAAAACTTATCATCAAATCAAAAATAATCAACCCTTTGGCTGAAAAAATGGA
.*.*.***.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

540
BATAI.LCC ATAACCTGGGCAGATGGGGAGGAGTTTACCTCTCATTCTTTCCAGGGTCTGAAATGTTT
GERMIST.LCC ATCACATGGGAGAGCGGGCCAGAAGTCTACCTGTCTTCTTTCCAGGTGCTGAAATGTTT
MAGUARI.LCC ATAACATGGGCTGATGGTGAGGAGGTATACCTCTCATTCTTTCCAGGCTCAGAAATGTTT
.*.*.***.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

600
BATAI.LCC TTAGGAACCTTTCAGATTCTATCCCCTAGCAATTGGCATTTACAAGGTGCAGAGGAAAGAG
GERMIST.LCC CTTGGAACCTTTCAGATTCTACCCACTTGCCATTGGGATCTACAAAGTGCAGAGGAAAGAA
MAGUARI.LCC TTGGGAACATTCAAATTCTACCCATTGGCTATCGGAATTTACAAGTTTCAAGAAGGAA
.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

660
BATAI.LCC ATGGAGCCAAAGTATCTGGAGAAGACCATGCGCCAAAGATATATGGGCTTGGAAAGCATCA
GERMIST.LCC ATGGACCCAAATTTCTGGAGAAGACTATGCGCCAGAGGTACCTTGGCATTGATGCTCAG
MAGUARI.LCC ATGGAACCTAAATATCTGGAGAAAACAATGCGACAGAGGTACATGGTCTTGAAGCAGCT
*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

720
BATAI.LCC ACTTGGACAGTCAGTAAGCTGAACGAAGTTCAGTCTGCACTGACTGTAGTATCAGGACTT
GERMIST.LCC ACATGGACAACAACCTAAGCTTGGAGAAGTGAAGCAGCTTTAAAAGTAGTCTCAGGGCTC
MAGUARI.LCC ACATGGACAGTTAGCAAAGTCAATGAAGTCCAGGCTGCACTTACAGTTGTCTCTGGATTA
.*.***.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

780
BATAI.LCC GGGTGAAGAAGACAAATGTGAGCAGTGCCGCTAGAGAATTCCTAGCAAAGTTTGGCATC
GERMIST.LCC GGATGGAAGAAGACAAATGTCAGCAGTGCTGCAAGGGAGTTCCTATCCAAATTCGGCATC
MAGUARI.LCC GGTGGAAGAAAACCTAATGTGAGTCTGCGGCCAGAGATTTCTGGCTAAGTTTGAAT
.*.***.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

840
BATAI.LCC AGCATGTAAATGCATTGTCAGAATTGT-AATCTGAT-AAATTGT-ATTTTTAAAATTTGG
GERMIST.LCC AGGATGTGAACACCCAGAGAAGTGCATTAATCGGGCTAAATTTTTATTTTATTAATTTGA
MAGUARI.LCC AACATGTAAAACAT-GCAGTAATT-T-AATCTGTTCAAATTT-ATGCATTAA-TTTGG
..*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

900
BATAI.LCC CTAAA--AGAGTCCCTGGACTCAAAAAGTAAAACAGC-TTGTT--GGGTGGGTGGTTG
GERMIST.LCC CTAAACAAGGGGTTTTACCCCAAACAATCTAAGCAGCATTGTAAATGGGTGGGTGGTTG
MAGUARI.LCC CTAAA--AGAGTCTTCGGACTCAAAA--TAAA--CAGC-TTTC--GGGTGGGTGGTTG
*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****.*.*****

Bunyavirus (grupo mayoritario)

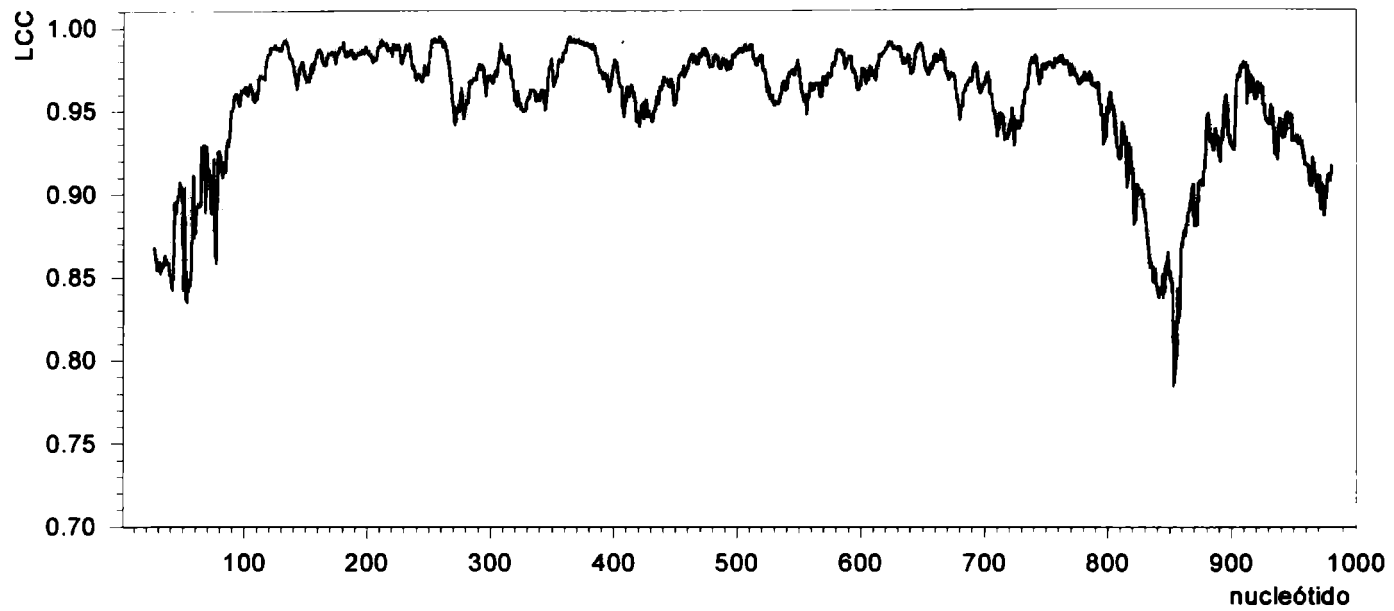


Figura Ap2.3. Gráfico consenso de LCC del grupo mayoritario. La línea negra representa las medias aritméticas de los valores de LCC de los 9 miembros del grupo. Las líneas grises verticales representan los errores estadísticos ($+1\sigma$ y -1σ) cada dos posiciones. El gráfico fue construido utilizando las rutinas descritas en el Capítulo 18.

Bunyavirus (grupo minoritario)

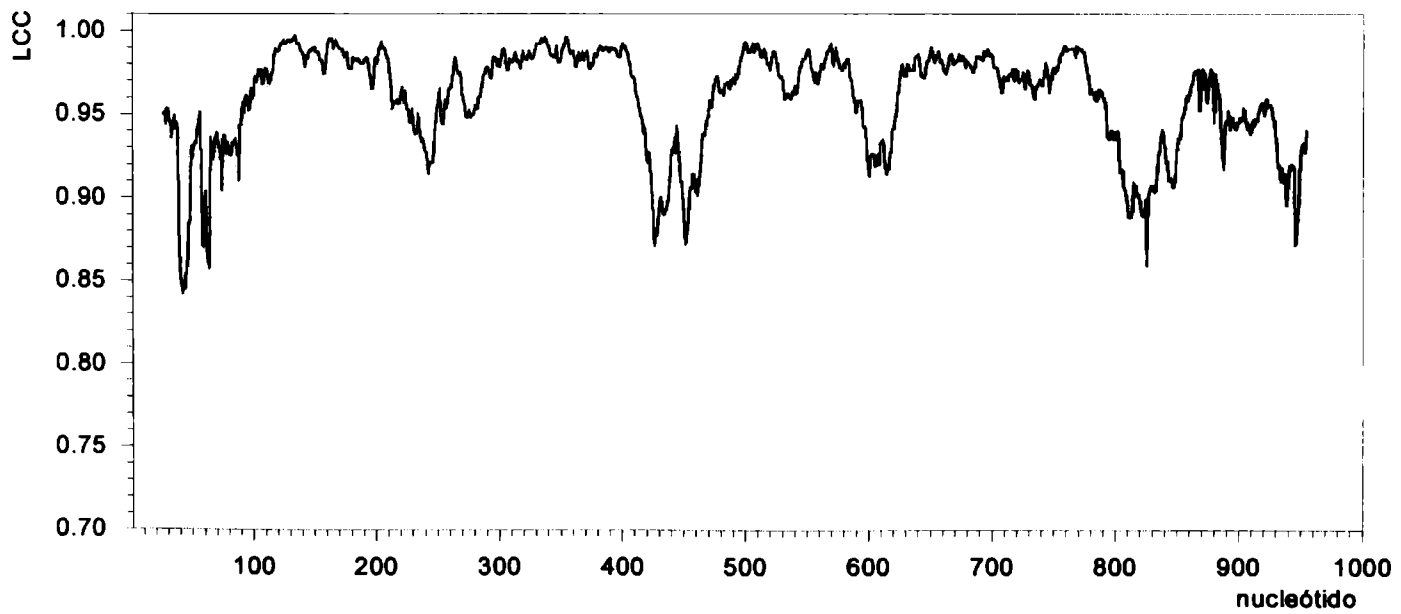


Figura Ap2.4. Gráfico consenso de LCC del grupo minoritario. La línea negra representa las medias aritméticas de los valores de LCC de los 3 miembros del grupo. Las líneas grises verticales representan los errores estadísticos ($+1\sigma$ y -1σ) cada dos posiciones. El gráfico fue construido utilizando las rutinas descritas en el Capítulo 18.

Ligación de oligonucleótidos a híbridos cDNA/RNA, catalizada por la DNA ligasa del fago T4 y amplificación por PCR: bases para un método rápido de clonado de extremos 5' de mRNAs y RNAs virales

Detalle experimental

Preparación del oligonucleótido (primer NA fosforilado y bloqueado en el extremo 3')

Reacción de fosforilación

Primer NA 3,2 µg/µl	6,3 µl
Premix I	10,0 µl
agua	3,7 µl
TOTAL	20,0 µl
15 min a 37°C	
Premix II	20,0 µl
TOTAL	40,0 µl
60 min. a 37°C	

Premix I

10 X buffer PNK	5 µl
(γ ³² P)ATP 30 µCi/ µl	3 µl
PNK I U/ µl	3 µl
agua	14 µl
TOTAL	25 µl

Premix II

10 X buffer PNK	8 µl
100 mM ATP	1 µl
PNK I U/ µl	3 µl
agua	28 µl
TOTAL	40 µl

Gel de poliacrilamida 10%-8 M urea
Autorradiografía 30 min. a temperatura ambiente, con pantalla (Fig. Ap3.1A)
Cortar el trozo de gel con la especie de 30 nt.
Colocar el trozo de gel entre vidrios de minigeles. Armar un gel de acrilamida 6% nativo alrededor. Correr 15 min.
Autorradiografía 30 min. a temperatura ambiente, con pantalla (Fig. Ap3.1B)
Cortar el trozo de gel con la forma de 30 nt.
Colocar el trozo de gel en un Eppendorf y eluir el primer en agua: Toda la noche a 50°C en 300 µl 2 x 1 h. a 50°C en 200 µl
Concentrar en Speed Vac

Reacción de bloqueo

Primer (³² P)NA en agua	19,0 µl
TdT buffer 5 X	6,0 µl
10 mM ddCTP	3,0 µl
agua	1,5 µl
TdT Promega 5 U/µl	0,5 µl
TOTAL	30,0 µl
5 min. a 37°C	
0,5 M EDTA pH 8,0	1,0 µl

Merm Aid High binding salt solution	300,0 µl
Merm Aid Glassfog	40,0 µl
20 min. con agitación fuerte	
3 x lavados con Merm Aid wash sol.	400,0 µl
2 x eluciones a 50°C con agua	50,0 µl
Concentración con Speed Vac	

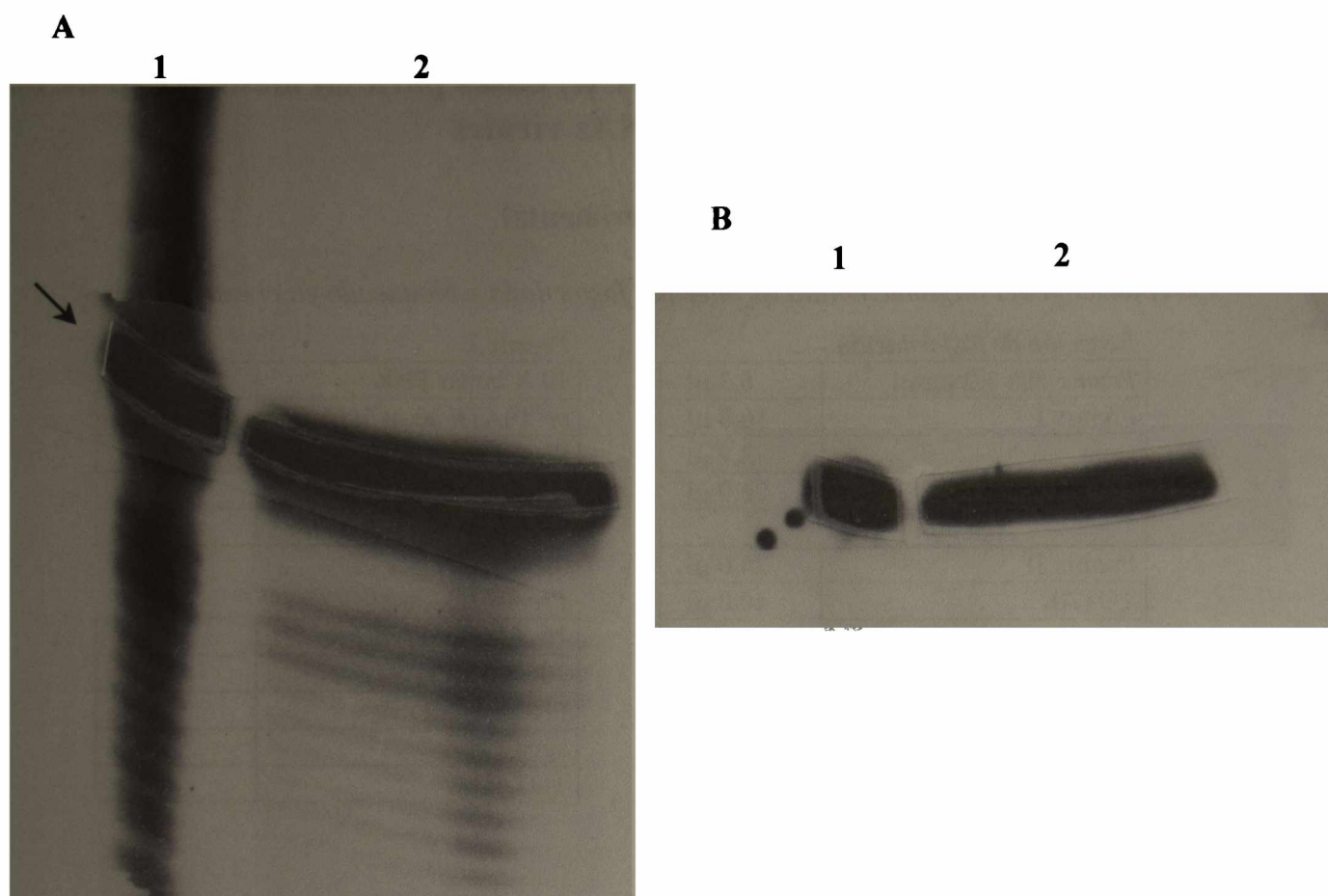


Figura Ap3.1. Purificación de la especie de 30 nucleótidos del primer NA. **A.** Autorradiografía del gel 10% de acrilamida-N,N' bisacrilamida (38:2) y 8 M urea como desnaturalizante. Calle 1: oligonucleótido sintético NA fosforilado solo con ($\gamma^{32}\text{P}$)ATP. Calle 2: oligonucleótido sintético NA fosforilado con ($\gamma^{32}\text{P}$)ATP (pulso de 15 min.) y luego con ATP no radioactivo (1 hora). Con una flecha se indica la especie de longitud máxima. También se observan las ventanas generadas en la autorradiografía para poder cortar con mayor precisión la especie deseada. **B.** Autorradiografía del gel 6% de acrilamida-N,N' bisacrilamida (38:2) nativo. Calle 1: especie de 30 nt. del primer NA fosforilado solo con ATP radioactivo. Calle 2: especie de 30 nt. del primer NA fosforilado con ATP radioactivo y ATP no radioactivo. También se observan las ventanas generadas en la autorradiografía para poder cortar con mayor precisión el fragmento de gel.

<i>Hibridación de oligonucleótidos</i>	
<i>Primer (^{32}P)NA en agua 11 ng/μl</i>	4,0 μl
<i>Primer NB en agua 10 ng/μl</i>	4,0 μl
<i>10 X buffer hibridación</i>	1,0 μl
<i>Agua</i>	1,0 μl
TOTAL	10,0 μl
3 min. a 100°C	
Dejar descender la temperatura hasta 37°C	

<i>10 X Buffer hibridación</i>
100 mM Tris-HCl pH 7,5
1 mM EDTA pH 8,0
500 mM NaCl

Experimento 1

Síntesis de cDNA y ligaciones a oligonucleótidos simple cadena y doble cadena

<i>Reacción de síntesis de cDNA</i>	
Transcripto <i>in vitro</i> 50 ng/μl	5,0 μl
Primer G en agua 20 ng/μl	10,0 μl
agua	20,0 μl
TOTAL	35,0 μl
3 min. a 100°C → agua-hielo	
0,1 M HOCH ₃ Hg	3,0 μl
5 min. a temperatura ambiente	
1,4 M βME	3,0 μl
5 min. a temperatura ambiente	
RNAsin 40 U/μl	2,5 μl
5 min. a temperatura ambiente	
5 X buffer transcripción reversa	19,3 μl
10 mM 4 dNTPs	9,6 μl
Agua	26,3 μl
AMV-RT 8 U/μl	5,0 μl
TOTAL	103,7 μl
80 min. a 42°C	
1 extracción con fenol y precipitación con etanol	
Resuspensión en 25 μl de agua	

Reacciones de ligación

Reactivos	1	2	3	4	5	6	7
Híbrido cDNA/RNA (10 ng/μl en RNA)	2,0 μl	2,0 μl	2,0 μl	2,0 μl	2,0 μl	2,0 μl	2,0 μl
Primer NA (11 ng/μl)	2,0 μl	2,0 μl	2,0 μl	2,0 μl	---	---	---
Dúplex NA/NB (4,4 ng/μl en NA)	---	---	---	---	3,0 μl	3,0 μl	3,0 μl
agua	13,0 μl	13,0 μl	13,0 μl	13,0 μl	12,0 μl	12,0 μl	10,5 μl
10 X buffer Mg	2,0 μl	2,0 μl	---	---	2,0 μl	---	2,0 μl
10 X buffer Mn	---	---	2,0 μl	2,0 μl	---	2,0 μl	---
T4 RNA ligasa 1,4 U/μl	---	1,0 μl	---	1,0 μl	---	---	---
T4 DNA ligasa 100 U/μl (NEB)	1,0 μl	---	1,0 μl	---	1,0 μl	1,0 μl	1,0 μl
Klenow 2,5 U/μl (BM)	---	---	---	---	---	---	0,5 μl
10 mM 4 dNTPs	---	---	---	---	---	---	1,0 μl
TOTAL	20,0 μl	20,0 μl	20,0 μl	20,0 μl	20,0 μl	20,0 μl	20,0 μl
Incubación a 16°C toda la noche							

<i>10 X Buffer Mg</i>
300 mM Tris-HCl pH 7,8
100 mM MgCl ₂
100 mM DTT
10 mM ATP

<i>10 X Buffer Mn</i>
500 mM Hepes-NaOH pH 7,9
100 mM MgCl ₂
100 mM DTT
1 mM ATP
500 μg/ml BSA

Reacciones de PCR para detectar productos de ligación de oligonucleótidos al híbrido cDNA/RNA (Figura Ap3.2)

<i>Premix</i>	
10 X buffer PCR	10,0 μl
Tritón X-100	0,1 μl
2 mM 4 dNTPs	10,0 μl
agua	58,4 μl
Taq polimerasa 2,5 U/μl	1,5 μl
TOTAL	80,0 μl

<i>Perfil de ciclado</i>	
92°C 2 min.	1 ciclo
92°C 30 seg.	40 ciclos
54°C 1,5 min.	
72°C 1 min.	
72°C 5 min.	1 ciclo

Apéndice 3

Reacciones de PCR

Reactivos	1	2	3	4	5	6	7	8	9	10
Clon de cDNA (GPC Junín MC2)	---	---	---	---	---	---	---	0,1	---	---
Híbrido cDNA/RNA (10 ng/μl en RNA)	---	---	---	---	---	---	---	---	0,5	0,5
Ligación 1 (experimento 1)	1,0	---	---	---	---	---	---	---	---	---
Ligación 2 (experimento 1)	---	1,0	---	---	---	---	---	---	---	---
Ligación 3 (experimento 1)	---	---	1,0	---	---	---	---	---	---	---
Ligación 4 (experimento 1)	---	---	---	1,0	---	---	---	---	---	---
Ligación 5 (experimento 1)	---	---	---	---	1,0	---	---	---	---	---
Ligación 6 (experimento 1)	---	---	---	---	---	1,0	---	---	---	---
Ligación 7 (experimento 1)	---	---	---	---	---	---	1,0	---	---	---
Primers Arena/G (10 μM de c/u)	---	---	---	---	---	---	---	1,0	1,0	---
Primers NB/G (10 μM de c/u)	1,0	1,0	1,0	1,0	1,0	1,0	1,0	---	---	1,0
Premix PCR	8,0	8,0	8,0	8,0	8,0	8,0	8,0	8,0	8,0	8,0
agua	---	---	---	---	---	---	---	0,9	0,5	0,5
TOTAL	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0	10,0

Nota: el *primer* Arena es el primer universal de los arnavirus y corresponde a los primeros nucleótidos de los extremos 5' viral y 5' viral complementario.

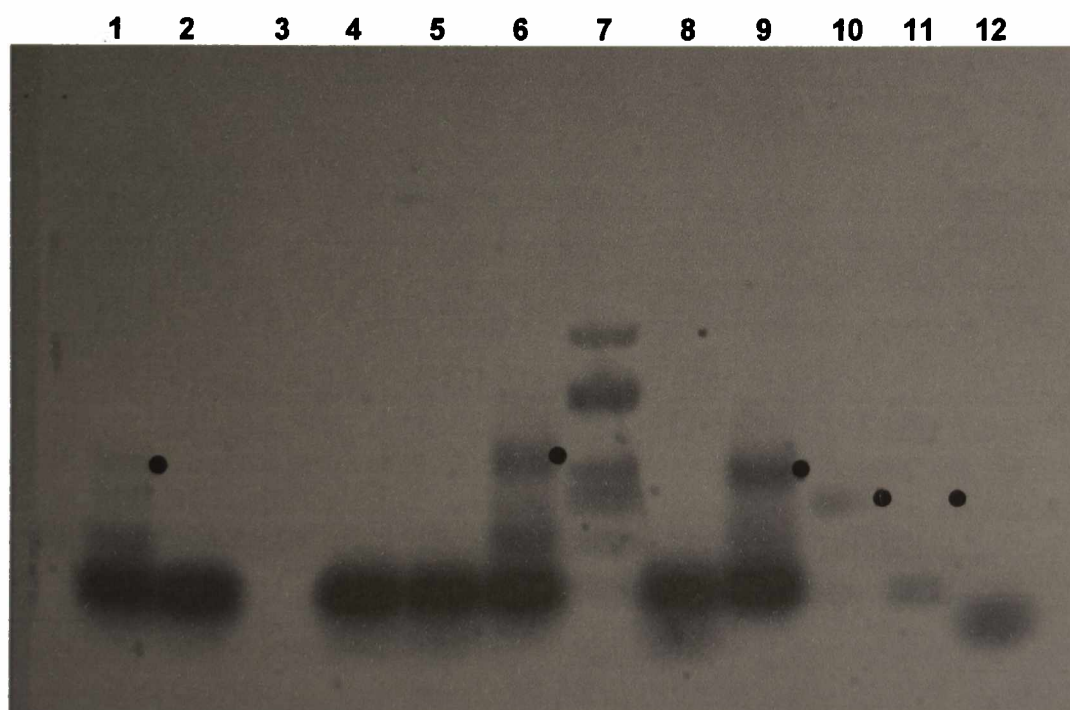


Figura Ap3.2. Amplificación de cDNAs ligados a oligonucleótidos. Calle 1: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena (enzima: T4 DNA ligasa en buffer Mg). Calle 2: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena (enzima: T4 RNA ligasa en buffer Mg). Calle 3: vacía. Calle 4: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena (enzima: T4 DNA ligasa en buffer Mn). Calle 5: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena (enzima: T4 RNA ligasa en buffer Mn). Calle 6: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzima: T4 DNA ligasa en buffer Mg). Calle 7: marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp + 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp. Calle 8: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzima: T4 DNA ligasa en buffer Mn). Calle 9: PCR con *primers* NB y G, el molde fue la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzimas: T4 DNA ligasa + Klenow, en buffer Mg + 4 dNTPs). Calle 10: PCR con *primers* internos, utilizando como molde un clon de cDNA. Calle 11: PCR con *primers* internos, utilizando como molde una alícuota del cDNA. Calle 12: PCR con *primers* NB y G, utilizando como molde una alícuota del cDNA. Con puntos negros se indican los fragmentos de amplificación esperados (~320 bp en el producto de ligación y ~220 bp con los *primers* internos)

Experimento 2

Síntesis de cDNA y ligaciones a oligonucleótidos simple cadena y doble cadena

<i>Reacción de síntesis de cDNA</i>	
RNA total de células infectadas (~5 µg)	5,0 µl
Primer G en agua 20 ng/µl	10,0 µl
agua	20,0 µl
TOTAL	35,0 µl
3 min. a 100°C → agua-hielo	
0,1 M HOCH ₃ Hg	3,0 µl
5 min. a temperatura ambiente	
1,4 M βME	3,0 µl
5 min. a temperatura ambiente	
RNAasin 40 U/µl	2,5 µl
5 min. a temperatura ambiente	
5 X buffer transcripción reversa	19,3 µl
10 mM 4 dNTPs	9,6 µl
Agua	26,3 µl
AMV-RT 8 U/µl	5,0 µl
TOTAL	103,7 µl
80 min. a 42°C	
1 extracción con fenol y precipitación con etanol	
resuspensión en 20 µl de agua	

Reacciones de ligación

Reactivos	1	2	3	4	5	6
Híbrido cDNA/RNA (~0,15 picomoles/µl)	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl
Primer NA (~12,2 picomoles/µl)	0,25 µl	---	---	0,25 µl	---	---
Dúplex NA/NB (~0,6 picomoles/µl)	---	2,0 µl	2,0 µl	---	2,0 µl	2,0 µl
agua	7,75 µl	6,0 µl	4,5 µl	6,75 µl	5,0 µl	3,5 µl
10 X buffer Mg	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl
T4 DNA ligasa 100 U/µl (NEB)	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl	0,5 µl
Klenow 2,5 U/µl (BM)	---	---	0,5 µl	---	---	0,5 µl
10 mM 4 dNTPs	---	---	1,0 µl	---	---	1,0 µl
40 % PEG 8000	---	---	---	1,0 µl	1,0 µl	1,0 µl
TOTAL	10,0 µl	10,0 µl	10,0 µl	10,0 µl	10,0 µl	10,0 µl
Incubación a temperatura ambiente toda la noche						

Reacciones de PCR para detectar productos de ligación de oligonucleótidos al híbrido cDNA/RNA (Figura Ap3.3)

<i>Premix</i>	
10 X buffer PCR	25,0 µl
2 mM 4 dNTPs	25,0 µl
agua	98,2 µl
Taq polimerasa 2,5 U/µl	1,8 µl
TOTAL	150,0 µl

<i>Perfil de ciclado</i>	
92°C 2 min.	1 ciclo
92°C 30 seg.	40 ciclos
54°C 1,5 min.	
72°C 1 min.	
72°C 5 min.	1 ciclo

Reacciones de PCR

Reactivos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
cDNA/RNA no lig. (~ 200 femtomoles)	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cDNA/RNA no lig. (~ 20 femtomoles)	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
cDNA/RNA no lig. (~ 2 femtomoles)	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 1 (exp. 2) (~ 200 femtomoles)	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 1 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 1 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 2 (exp. 2) (~ 200 femtomoles)	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 2 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 2 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 3 (exp. 2) (~ 200 femtomoles)	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-
Ligación 3 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-
Ligación 3 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-
Ligación 4 (exp. 2) (~ 200 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-	-
Ligación 4 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	-
Ligación 4 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-
Ligación 5 (exp. 2) (~ 200 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-
Ligación 5 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-
Ligación 5 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
Ligación 6 (exp. 2) (~ 200 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-	-
Ligación 6 (exp. 2) (~ 20 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-	-
Ligación 6 (exp. 2) (~ 2 femtomoles)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	-
Primers NB/G (10 µM de c/u)	1	-	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Primers Arena/G (10 µM de c/u)	-	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Premix PCR	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
agua	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3
TOTAL	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

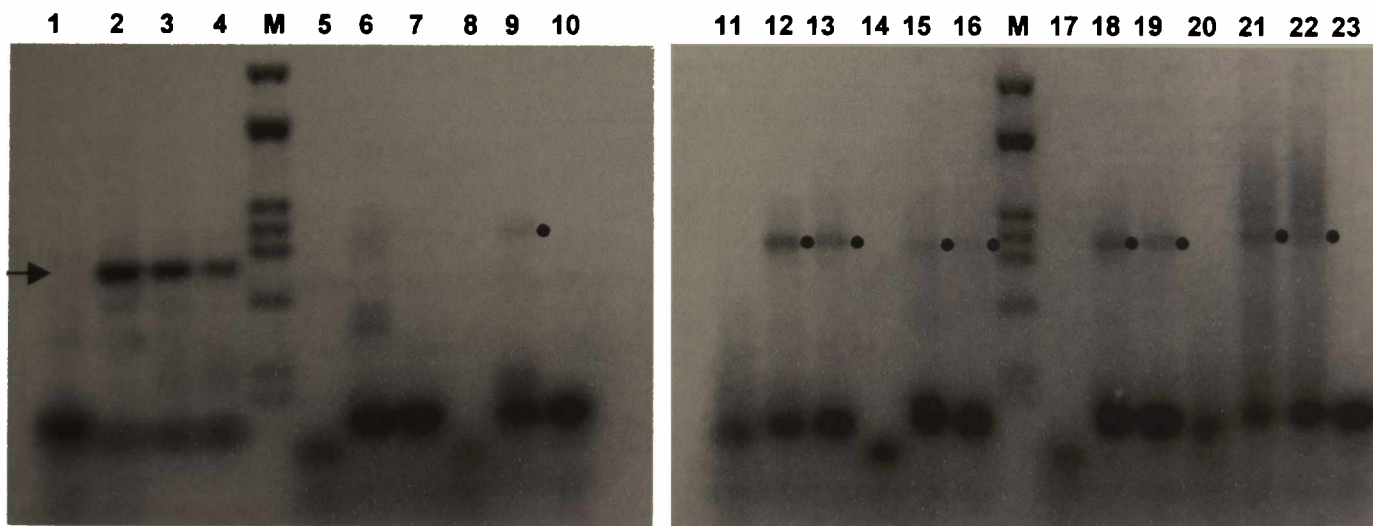


Figura Ap3.3. Amplificación de cDNAs ligados a oligonucleótidos. En todos los casos, las cantidades de molde en los distintos puntos de las diluciones seriadas corresponden a 200, 20 y 2 femtomoles totales de cDNA no ligado, respectivamente. **Calle 1:** control negativo de PCR con *primers* NB y G (molde: cDNA sin ligar). **Calles 2 a 4:** controles de existencia del cDNA, amplificaciones con *primers* Arena y G utilizando como moldes diluciones seriadas del cDNA (con una flecha negra se indica el fragmento de amplificación esperado: ~220 bp). **Calles 5 a 7:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena. **Calles 8 a 10:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB. **Calles 11 a 13:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzimas: T4 DNA ligasa + Klenow). **Calles 14 a 16:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el *primer* NA simple cadena, en presencia de 4% PEG 8000. **Calles 17 a 19:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB, en presencia de 4% PEG 8000. **Calles 20 a 22:** PCRs con *primers* NB y G, amplificaciones sobre diluciones seriadas de la mezcla de ligación del híbrido RNA-cDNA con el dúplex NA/NB (enzimas: T4 DNA ligasa + Klenow), en presencia de 4% PEG 8000. **Calle 23:** control negativo de PCR (sin DNA). Con puntos negros se indica el fragmento de amplificación esperado: ~256 bp. **M:** marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp, 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp.

Experimento 3

Síntesis de cDNA y ligaciones a oligonucleótidos simple cadena y doble cadena

Reacciones de síntesis de cDNA	MC2			Candid #1		
	clon 5 ₂	clon 6 ₂	mezcla	Viral	# 12	# 17
RNA total de células infectadas o viral	5,0 µl	5,0 µl	5,0 µl	2,5 µl	2,5 µl	2,5 µl
Primer G en agua 20 ng/µl	10,0 µl	10,0 µl	10,0 µl	5,0 µl	5,0 µl	5,0 µl
agua	20,0 µl	20,0 µl	20,0 µl	10,0 µl	10,0 µl	10,0 µl
TOTAL	35,0 µl	35,0 µl	35,0 µl	17,5 µl	17,5 µl	17,5 µl
3 min. a 100°C → agua-hielo						
0,1 M HOCH ₃ Hg	3,0 µl	3,0 µl	3,0 µl	1,5 µl	1,5 µl	1,5 µl
5 min. a temperatura ambiente						
1,4 M βME	3,0 µl	3,0 µl	3,0 µl	1,5 µl	1,5 µl	1,5 µl
5 min. a temperatura ambiente						
RNAsin 40 U/µl	2,5 µl	2,5 µl	2,5 µl	1,25 µl	1,25 µl	1,25 µl
5 min. a temperatura ambiente						
5 X buffer transcripción reversa	19,3 µl	19,3 µl	19,3 µl	9,6 µl	9,6 µl	9,6 µl
10 mM 4 dNTPs	9,6 µl	9,6 µl	9,6 µl	4,8 µl	4,8 µl	4,8 µl
Agua	26,3 µl	26,3 µl	26,3 µl	13,2 µl	13,2 µl	13,2 µl
AMV-RT 8 U/µl	5,0 µl	5,0 µl	5,0 µl	2,5 µl	2,5 µl	2,5 µl
TOTAL	103,7 µl	103,7 µl	103,7 µl	51,85 µl	51,85 µl	51,85 µl
90 min. a 42°C						
1 extracción con fenol y precipitación con etanol						
resuspensión en 20 µl de agua						

Reacciones de ligación

Reactivos	1	2	3	4	5	6
Híbrido cDNA/RNA MC2 mezcla (~0,15 picomoles/µl)	2,0 µl	---	---	---	---	---
Híbrido cDNA/RNA MC2 clon 5 ₂ (~0,15 picomoles/µl)	---	2,0 µl	---	---	---	---
Híbrido cDNA/RNA MC2 clon 6 ₂ (~0,15 picomoles/µl)	---	---	2,0 µl	---	---	---
Híbrido cDNA/RNA CD1 Viral (~0,15 picomoles/µl)	---	---	---	2,0 µl	---	---
Híbrido cDNA/RNA CD1 # 12 (~0,15 picomoles/µl)	---	---	---	---	2,0 µl	---
Híbrido cDNA/RNA CD1 # 17 (~0,15 picomoles/µl)	---	---	---	---	---	2,0 µl
Dúplex NA/NB (~0,6 picomoles/µl)	3,0 µl	3,0 µl	3,0 µl	3,0 µl	3,0 µl	3,0 µl
agua	11,0 µl	11,0 µl	11,0 µl	11,0 µl	11,0 µl	11,0 µl
10 X buffer Mg	2,0 µl	2,0 µl	2,0 µl	2,0 µl	2,0 µl	2,0 µl
T4 DNA ligasa 100 U/µl (NEB)	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl
40 % PEG 8000	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl	1,0 µl
TOTAL	20,0 µl	20,0 µl	20,0 µl	20,0 µl	20,0 µl	20,0 µl
Incubación a 22°C 4 horas						

Reacciones de PCR para detectar productos de ligación de oligonucleótidos al híbrido cDNA/RNA (Figura Ap3.4)

Premix	
10 X buffer PCR	7,0 µl
2 mM 4 dNTPs	7,0 µl
agua	27,5 µl
Taq polimerasa 2,5 U/µl	0,5 µl
Primers NB/G (10 µM de c/u)	7,0 µl
TOTAL	49,0 µl

Perfil de ciclado	
92°C 2 min.	1 ciclo
92°C 30 seg.	40 ciclos
54°C 1,5 min.	
72°C 1 min.	
72°C 5 min.	1 ciclo

Reacciones de PCR

Reactivos	1	2	3	4	5	6	7
Híbrido cDNA/RNA (10 ng/μl en RNA)	---	---	---	---	---	---	3,0
Ligación 1 (experimento 3) dilución 1/10	3,0	---	---	---	---	---	---
Ligación 2 (experimento 3) dilución 1/10	---	3,0	---	---	---	---	---
Ligación 3 (experimento 3) dilución 1/10	---	---	3,0	---	---	---	---
Ligación 4 (experimento 3) dilución 1/10	---	---	---	3,0	---	---	---
Ligación 5 (experimento 3) dilución 1/10	---	---	---	---	3,0	---	---
Ligación 6 (experimento 3) dilución 1/10	---	---	---	---	---	3,0	---
Premix PCR	7,0	7,0	7,0	7,0	7,0	7,0	7,0
TOTAL	10,0	10,0	10,0	10,0	10,0	10,0	10,0

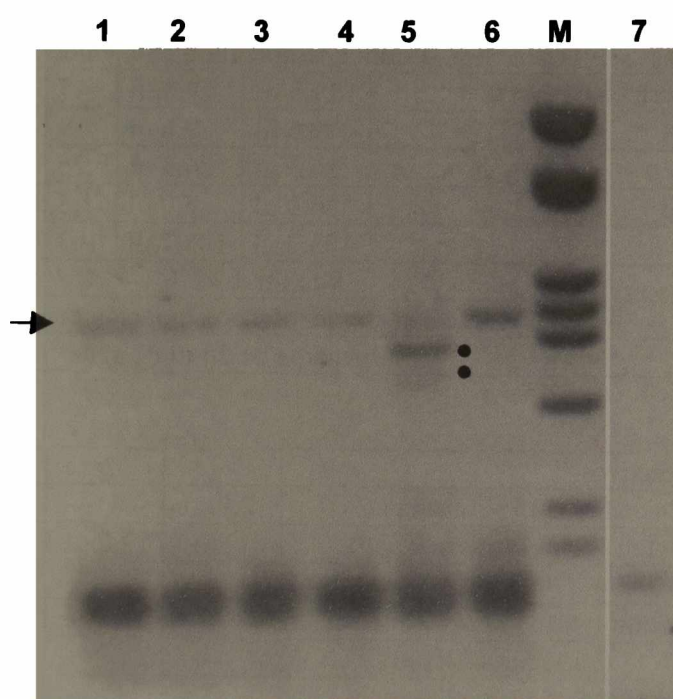


Figura Ap3.4. Amplificación de cDNAs ligados a oligonucleótidos. En todos los casos, los moldes utilizados corresponden a diluciones 1/10 de las mezclas de ligación respectivas. **Calle 1:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 mezcla con el dúplex NA/NB. **Calle 2:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 clon 5₂ con el dúplex NA/NB. **Calle 3:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA MC2 clon 6₂ con el dúplex NA/NB. **Calle 4:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CDI viral con el dúplex NA/NB. **Calle 5:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CDI # 12 con el dúplex NA/NB. **Calle 6:** amplificación con *primers* NB y G de la mezcla de ligación del híbrido RNA-cDNA CDI # 17 con el dúplex NA/NB. **Calle 7:** control negativo de PCR (sin DNA). Con una flecha negra se indica el fragmento de amplificación esperado: ~256 bp y con puntos negros en la calle 5 se muestran fragmentos adicionales obtenidos. **M:** marcador de peso molecular, pUC9 digerido con *Hae* III: 80 bp, 102 bp, 174 bp, 234 bp, 267 bp, 298 bp, 434 bp + 458 bp y 587 bp. **Nota:** La calle 7 corresponde al mismo gel.

Gráficos de LCC individuales de los RNAs S de arenavirus

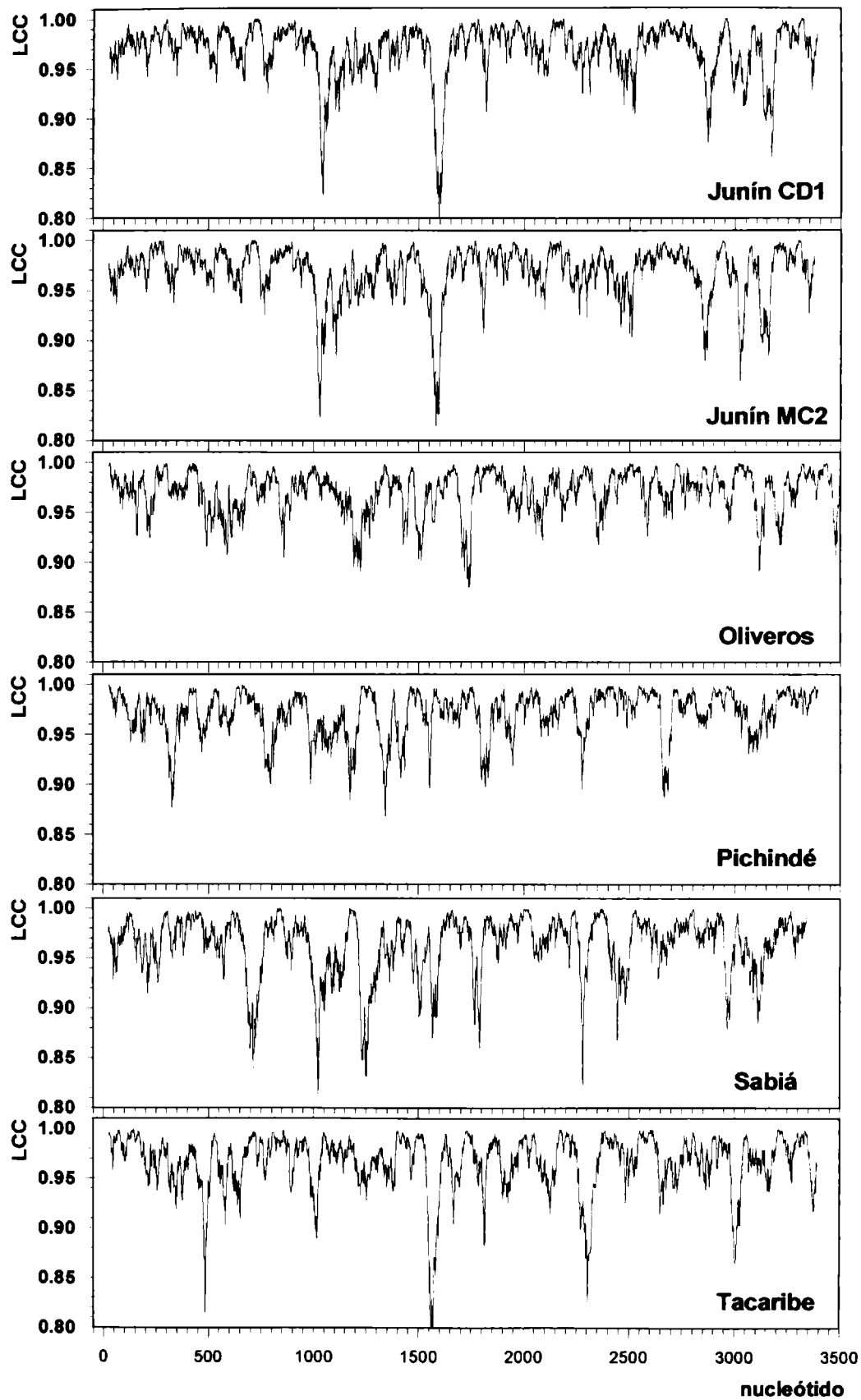


Figura Ap4.1. Gráficos de complejidad local de composición (LCC) de los RNAs S de los arenavirus del Nuevo Mundo Junín-Candid #1, Junín-MC2, Oliveros, Pichindé, Sabiá y Tacaribe. El gráfico deriva del paquete de rutinas de cálculo y análisis múltiple de LCC (Ghiringhelli, P.D., 2002, esta tesis), utilizando una ventana de 51 residuos.

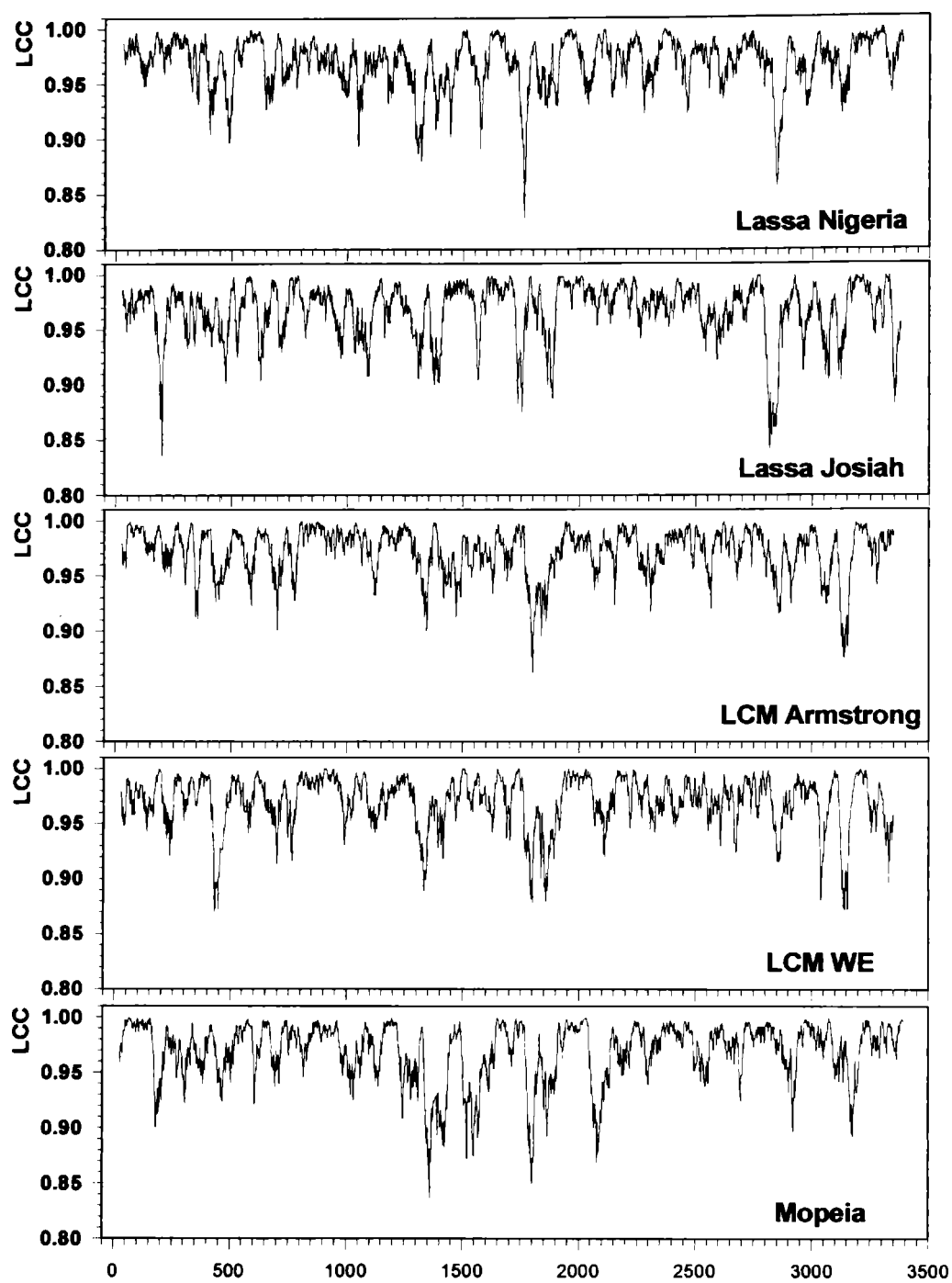


Figura Ap4.2. Gráficos de complejidad local de composición (LCC) de los RNAs S de los arenavirus del Viejo Mundo Lassa-Nigeria, Lassa-Josiah, LCM-Armstrong, LCM-WE y Mopeia. El gráfico deriva del paquete de rutinas de cálculo y análisis múltiple de LCC (Ghiringhelli, P.D., 2002, esta tesis), utilizando una ventana de 51 residuos.

Parte I - Transformada discreta de Fourier

(Cornette *et al.*, 1987)

La búsqueda y hallazgo de periodicidad en una secuencia numérica que representa alguna propiedad de una secuencia aminoacídica (o nucleotídica) constituye un dato peculiar en sí mismo.

En particular, en el caso de las secuencias aminoacídicas, la detección de una repetición periódica de una propiedad determinada (por ejemplo: hidrofobicidad) con un período característico, correlacionable con la geometría de la estructura secundaria local (3,6 para α hélices y 2,4 para estructuras β), estaría indicando, además, el carácter de anfipaticidad de la estructura.

Un método estándar para detectar periodicidad en una secuencia finita de números: $h_0, h_1, h_2, \dots, h_{i-1}$ $= \{h_k\}_{k=0}^{i-1}$ es comparar la secuencia numérica incógnita con una secuencia control de periodicidad conocida, típicamente una secuencia armónica de frecuencia ω y período $2\pi/\omega$:

$$\{A \cos(k\omega) + B \sin(k\omega)\}_{k=0}^{i-1}$$

Existen, al menos, dos métodos alternativos para comparar la secuencia incógnita con la secuencia control:

- 1) calcular la correlación entre ellas (asociada con la transformada discreta de Fourier)
- 2) calcular la suma de los cuadrados de las diferencias de entradas correspondientes (usada en el análisis de cuadrados mínimos) **(No desarrollado en este apéndice).**

Transformada discreta de Fourier

Para cualquier secuencia finita de números $\{h_k\}_{k=0}^{i-1}$ y frecuencia angular ω (medida en radianes), se definen $A(\omega)$, $B(\omega)$, $P(\omega)$ y $F(\omega)$ mediante:

$$A(\omega) = \sum_{k=0}^{i-1} h_k \cos(k\omega) \quad \text{Ec. 1} \quad B(\omega) = \sum_{k=0}^{i-1} h_k \sin(k\omega) \quad \text{Ec. 2}$$

$$P(\omega) = (A(\omega)^2) + (B(\omega)^2) = \left(\sum_{k=0}^{i-1} h_k \cos(k\omega) \right)^2 + \left(\sum_{k=0}^{i-1} h_k \sin(k\omega) \right)^2 \quad \text{Ec. 3}$$

$$F(\omega) = \left\{ (A(\omega)^2) + (B(\omega)^2) \right\}^{1/2} = \left\{ \left(\sum_{k=0}^{i-1} h_k \cos(k\omega) \right)^2 + \left(\sum_{k=0}^{i-1} h_k \sin(k\omega) \right)^2 \right\}^{1/2} \quad \text{Ec. 4}$$

Varios autores (McLachlan & Stewart, 1976; Eisenberg *et al.*, 1984) han examinado $P(\omega)$ (o su raíz cuadrada: $F(\omega)$, llamada intensidad) para el valor de $\omega = \hat{\omega}$ que maximiza $P(\omega)$, e interpretan a $\hat{\omega}$ como la frecuencia característica (dominante) de $\{h_k\}_{k=0}^{i-1}$ y $2\pi/\hat{\omega}$ como su período característico.

La razón principal es que $\sum_{k=0}^{i-1} A(\omega)\cos(k\omega) + B(\omega)\sin(k\omega) = h_k$ siempre es una buena aproximación

hacia la secuencia armónica que tiene establecida una correlación (no normalizada) con $\{h_k\}_{k=0}^{i-1}$. Por lo tanto, la secuencia original puede ser escrita como una suma senos y cosenos periódicos que tienen frecuencia ω y período $2\pi/\omega$.

Ejemplos

Como ejemplos prácticos, podemos considerar las secuencias numéricas correspondientes a los valores de hidrofobicidad de dos segmentos de la proteína de la nucleocápside del virus Junín-MC2, predichos como α -hélices utilizando el programa PHDsec (Rost and Sander, 1993).

Si consideramos a la hidropatía como una función discreta, entonces $f(x) = h_1, h_2, h_3, \dots, h_{18}$, en el caso de la **Figura Ap5.1.A** y $f(x) = h_1, h_2, h_3, \dots, h_{16}$, en el caso de la **Figura Ap5.2.A**.

Por otra parte, si tomamos intervalos $\Delta x=1$ y suponemos que los senos y cosenos son constantes en cada intervalo, tomando sólo el valor central del intervalo (Figuras Ap5.1.B y Ap5.2.B), los coeficientes $A(\omega)$ y $B(\omega)$ se pueden escribir como:

$$A\omega = \int_{-\infty}^{\infty} f(x) \cos(x\omega) dx = \sum_{k=0}^{n-1} h_k \cos(k\omega) \times 1 \quad \text{es la Ec. 1}$$

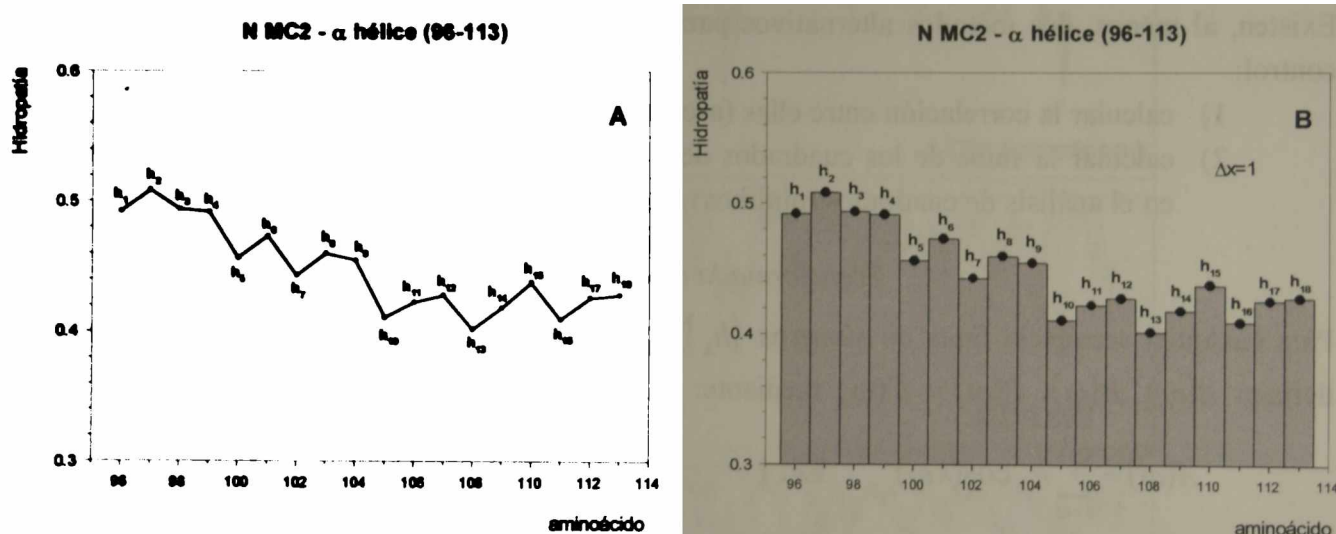


Figura Ap5.1. Perfil de hidropatía correspondiente a la secuencia $_{96}$ RVERMELASGLGNLKTKF $_{113}$ de la proteína N del virus junín-MC2, predicha como α hélice con el programa PHDsec. A. Representación tradicional del perfil de hidropatía. B. Representación en forma de histograma.

El perfil de hidropatía se calculó utilizando la tabla de hidrofobicidades PRIFT (Cornette *et al.*, 1987).

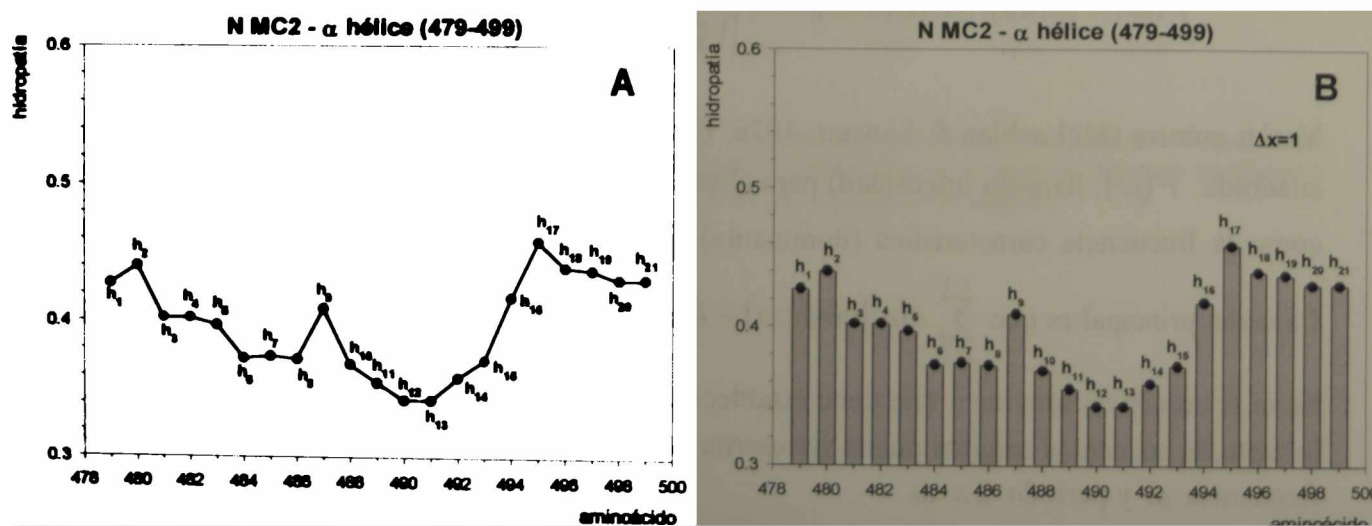


Figura Ap5.2. Perfil de hidropatía correspondiente a la secuencia $_{479}$ SEQARQFDQQVWEKFGHLCKH $_{499}$ de la proteína N del virus junín-MC2, predicha como α hélice con el programa PHDsec. A. Representación tradicional del perfil de hidropatía. B. Representación en forma de histograma.

El perfil de hidropatía se calculó utilizando la tabla PRIFT (Cornette *et al.*, 1987).

El análisis discreto de Fourier nos permite visualizar gráficamente el ω dominante (Figuras Ap5.3.A y Ap5.4.A) en cada caso y el período característico ($2\pi/\omega$) de las mismas (Figuras Ap5.3.B y Ap5.4.B).

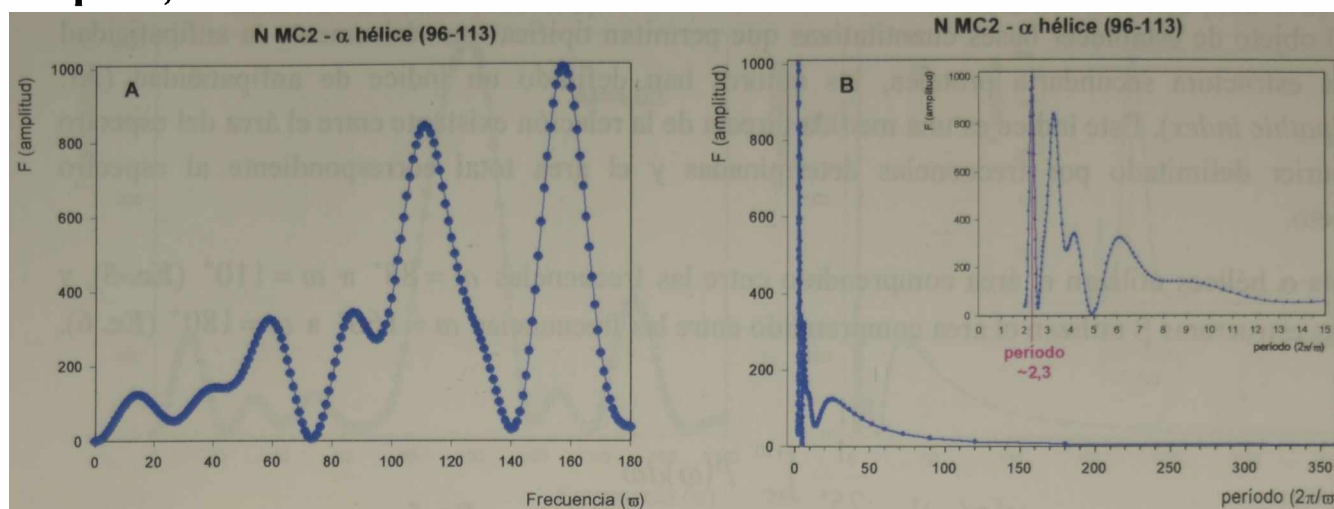


Figura Ap5.3. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la figura Ap5.1.

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a 158° . Además, existen otras amplitudes de cierta importancia a 112° , 88° y 60° .

B. Representación de F vs. $2\pi/\omega$. El gráfico completo corresponde al total de períodos, mientras que el gráfico insertado corresponde a una ampliación de períodos 0 a 15. En el mismo se observa que el período de mayor amplitud corresponde a un valor $\sim 2,3$. Si este valor se considera como cantidad de residuos, y teniendo en cuenta que la estructura secundaria predicha es una α hélice, se puede afirmar que la distribución de residuos aminoacídicos polares y no polares es homogénea en toda la estructura.

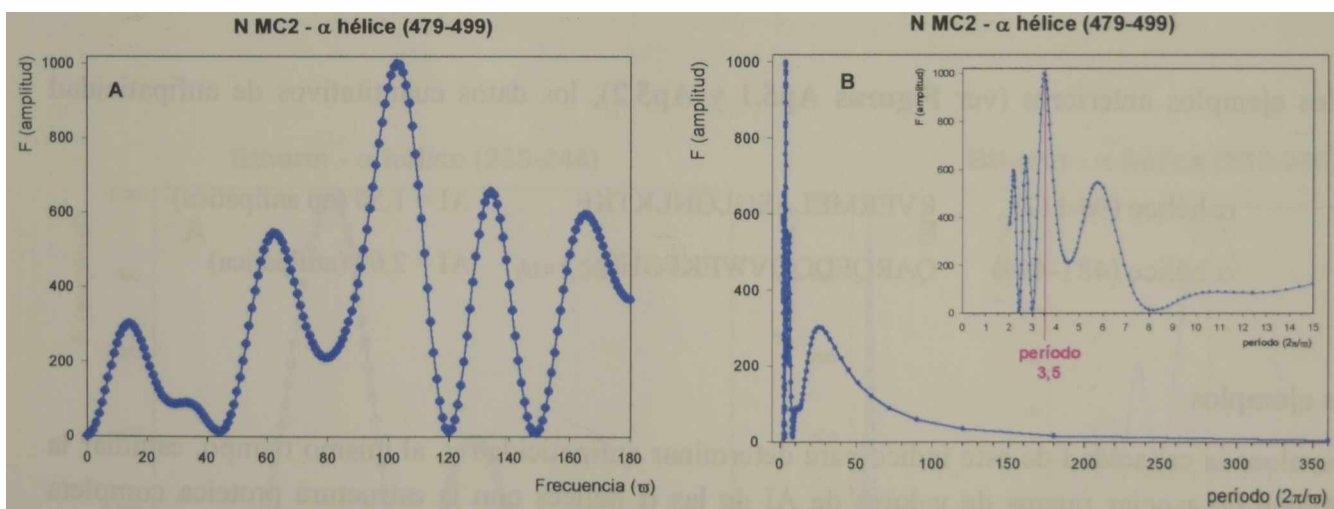


Figura Ap5.4. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la figura Ap5.2.

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a $\sim 100^\circ$.

B. Representación de F vs. $2\pi/\omega$. Se observa que el período de mayor amplitud corresponde a un valor $\sim 3,6$. Si este valor se considera como cantidad de residuos, y considerando que la estructura secundaria predicha es una α hélice, se puede afirmar que los residuos aminoacídicos polares se encuentran agrupados en una cara de la hélice y los no polares en la otra. Por lo tanto, se puede considerar que esta hélice es anfipática.

Los resultados obtenidos con este análisis obtienen con las representaciones gráficas Edmundson, 1967) o *helical net* (Dunnill, 1968). han sido desarrolladas sólo para las α hélices₁₀, el carácter de anfipaticidad en cualquiera de los tipo β .

cuantitativo son equivalentes a los que se cualitativas del tipo *helical wheel* (Schiffer and Por otra parte, las representaciones cualitativas mientras que el análisis de Fourier permite medir tipos de α hélice y, además, en estructuras de

Parte II – Índice de anfipaticidad

(Cornette *et al.*, 1987)

Con el objeto de establecer bases cuantitativas que permitan tipificar numéricamente la anfipaticidad de una estructura secundaria proteica, los autores han definido un índice de anfipaticidad (AI: *amphipathic index*). Este índice es una medida directa de la relación existente entre el área del espectro de Fourier delimitado por frecuencias determinadas y el área total correspondiente al espectro completo.

Para las α hélices utilizan el área comprendido entre las frecuencias $\omega = 85^\circ$ a $\omega = 110^\circ$ (Ec. 5), y para las estructuras β utilizan el área comprendido entre las frecuencias $\omega = 160^\circ$ a $\omega = 180^\circ$ (Ec. 6).

$$AI[P(\omega)] = \frac{\int_{85^\circ}^{110^\circ} P(\omega) d\omega}{\int_{0^\circ}^{180^\circ} P(\omega) d\omega} \quad \text{Ec. 5}$$

$$AI[P(\omega)] = \frac{\int_{160^\circ}^{180^\circ} P(\omega) d\omega}{\int_{0^\circ}^{180^\circ} P(\omega) d\omega} \quad \text{Ec. 6}$$

Para los ejemplos anteriores (ver Figuras Ap5.1 y Ap5.2), los datos cuantitativos de anfipaticidad son:

α hélice (96-113)	RVERMELASGLGNLTKKF	AI = 1,50 (no anfipática)
α hélice (481-496)	QARQFDQQVWEKFGHL	AI = 2,09 (anfipática)

Otros ejemplos

Para evaluar la capacidad de este índice para determinar anfipaticidad y, al mismo tiempo, estudiar la posibilidad de asociar rangos de valores de AI de las α hélices con la estructura proteica completa (proteína globular, de superficie o de membrana) y su ubicación en la misma, se analizó un conjunto de α hélices correspondientes a proteínas con estructura cristalográfica conocida.

En las Figuras Ap5.5, Ap5.6, Ap5.7 y Ap5.8 se pueden observar algunos ejemplos seleccionados.

En la Figura Ap5.5 se puede apreciar que el AI es excepcionalmente alto (4,02) y corresponde a una α hélice externa, netamente anfipática, de una proteína globular de *Bacillus thermophilus*.

En la Figura Ap5.6 se puede apreciar que el AI es menor que 2 (1,38) y corresponde a una α hélice no anfipática de la región del *core* de una proteína globular de *Bacillus thermophilus*.

En la Figura Ap5.7 se puede apreciar que el AI es excepcionalmente alto (3,93) y corresponde a una α hélice externa, netamente anfipática, de una proteína superficial de *Apis mellifera*.

En la Figura Ap5.8 se puede apreciar que el AI es excepcionalmente bajo (0,42) y corresponde a una α hélice no anfipática de la región transmembrana de una proteína integral humana.

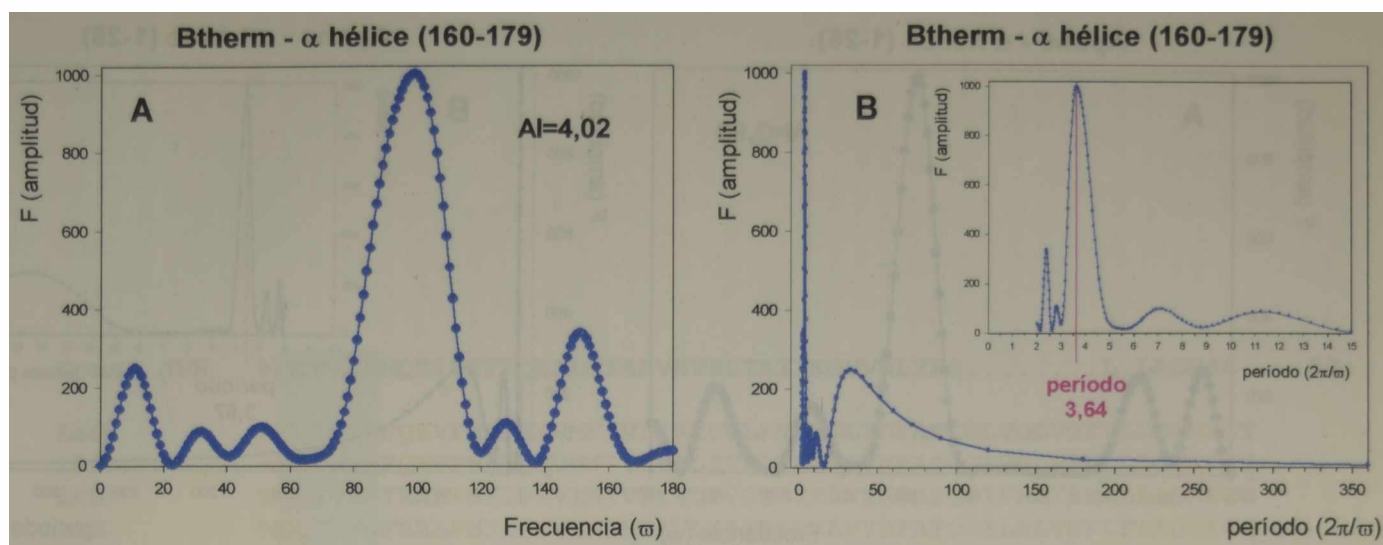


Figura Ap5.5. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la α hélice ${}_{160}\text{ESGAINGAISDIFGTLVEFY}_{179}$ de una proteína globular de *Bacillus thermophilus*.

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a $\sim 100^\circ$, y el índice de anfipaticidad correspondiente es de 4,02.

B. Representación de F vs. $2\pi/\omega$. Se observa que el período de mayor amplitud corresponde a un valor $\sim 3,6$. Si este valor se considera como cantidad de residuos, y teniendo en cuenta que la estructura secundaria predicha es una α hélice, se puede afirmar que los residuos aminoacídicos polares se encuentran agrupados en una cara de la hélice y los no polares en la otra. Por lo tanto, se puede considerar que esta hélice es anfipática.

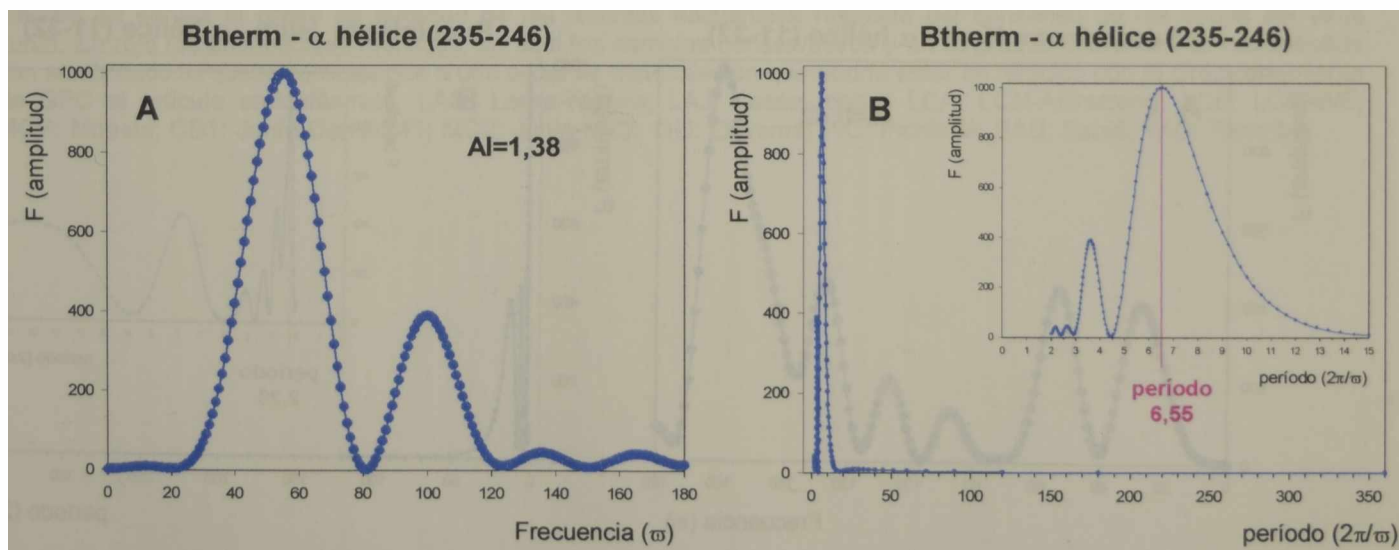


Figura Ap5.6. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la α hélice ${}_{235}\text{GIINKAAYLISQ}_{246}$ de una proteína globular de *Bacillus thermophilus*.

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a $\sim 55^\circ$, y el índice de anfipaticidad correspondiente es de 1,38.

B. Representación de F vs. $2\pi/\omega$. Se observa que el período de mayor amplitud corresponde a un valor $\sim 6,5$. Si este valor se considera como cantidad de residuos, y teniendo en cuenta que la estructura secundaria predicha es una α hélice, se puede afirmar que la distribución de residuos aminoacídicos polares y no polares es homogénea en toda la estructura.

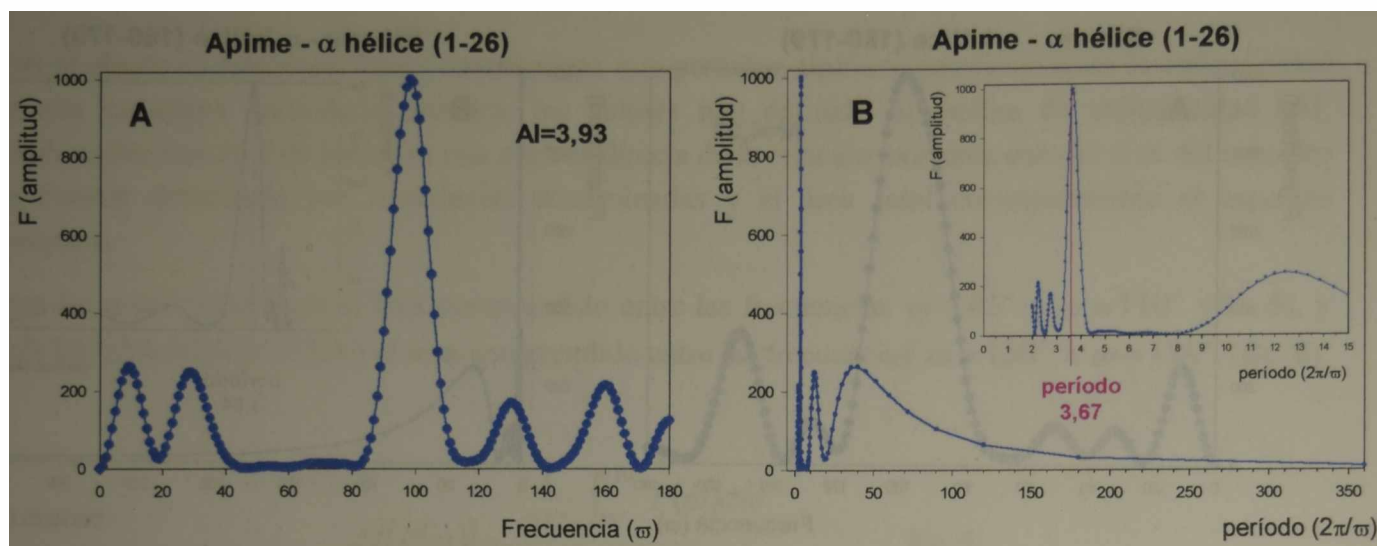


Figura Ap5.7. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la α hélice $_1\text{GIGAILKVLATGLPTLISWIKNRKQ}_{26}$ de una proteína superficial de *Apis mellifera*.

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a $\sim 100^\circ$, y el índice de anfipicidad correspondiente es de 3,93.

B. Representación de F vs. $2\pi/\omega$. Se observa que el período de mayor amplitud corresponde a un valor $\sim 3,7$. Si este valor se considera como cantidad de residuos, y teniendo en cuenta que la estructura secundaria predicha es una α hélice, se puede afirmar que los residuos aminoacídicos polares se encuentran agrupados en una cara de la hélice y los no polares en la otra. Por lo tanto, se puede considerar que esta hélice es anfipática.

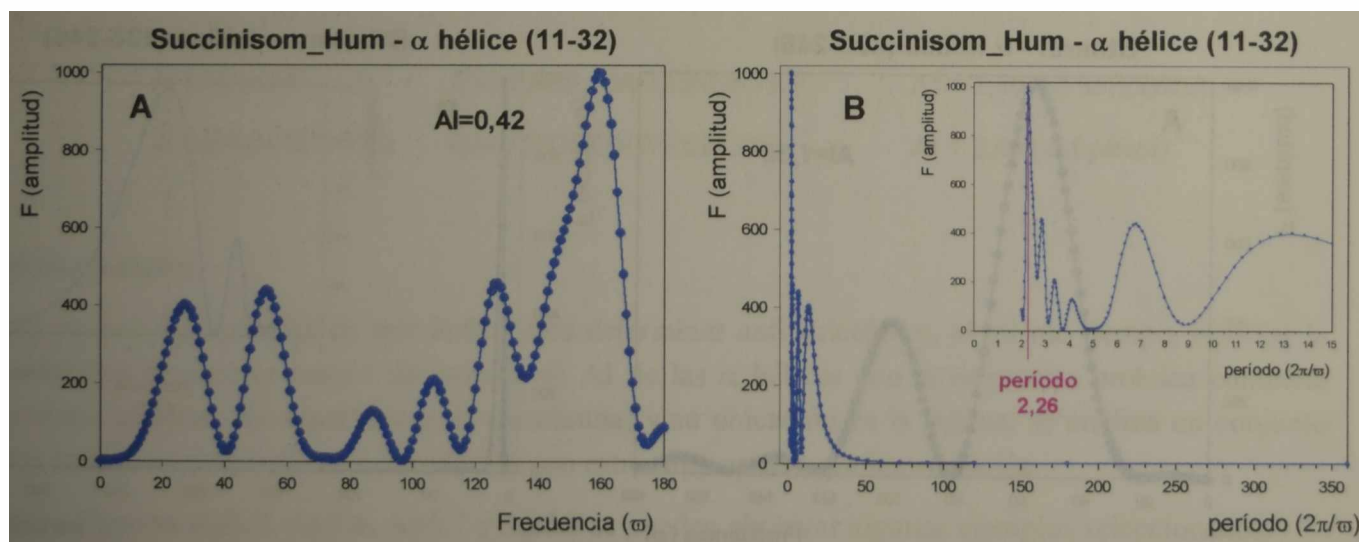


Figura Ap5.8. Análisis discreto de Fourier aplicado a la secuencia numérica de valores de hidrofobicidad de la α hélice $_{11}\text{ISLIVLFVTVTIHALIVVLA}_{32}$ de una proteína de membrana humana (Succinato isomerasa).

A. Representación de F vs. ω . Se observa que la frecuencia de mayor amplitud corresponde a $\sim 159^\circ$, y el índice de anfipicidad correspondiente es de 0,42.

B. Representación de F vs. $2\pi/\omega$. Se observa que el período de mayor amplitud corresponde a un valor $\sim 2,3$. Si este valor se considera como cantidad de residuos, y teniendo en cuenta que la estructura secundaria predicha es una α hélice, se puede afirmar que la distribución de residuos aminoacídicos polares y no polares es homogénea en toda la estructura.

Cons . JUN	MGQFISFMQEIP TFLQEALNIALVAVSLIAI IKGVVNL YKS L. LAGRSC .	58
LAG	MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLRGLYNFATCGLVGLVTFLLLCGRSCT	
LAJ	MGQIVTFFQEVPHVIEEVMNIVLIALSVLAVLRGLYNFATCGLVGLVTFLLLCGRSCT	
LCA	MGQIVTMFEALPHIIDEVINIVIIVLIVITGIKAVYNFATCGIFALISFLLAGRSCG	
LCE	MGQIVTMFEALPHIIDEVINIVIIVLIIITSIKAVYNFATCGILALVSFIFLAGRSCG	
MOP	MGQIVTFFQEVPHILEEVMNIVIMTLSILAILKGIYVMTCGIIGLITFLFLCGRSCS	
CD1	MGQFISFMQEIP TFLQEALNIALVAVSLIAI IKGVVNL YKSGLFQFFVFLALAGRSCS	
MC2	MGQFISFMQEIP TFLQEALNIALVAVSLIAI IKGVVNL YKSG---CSILDLAGRSCP	
OLI	MGQVIGFFQSLENIINEALNTALICVALIALKGIYNIWASGLIQLFIFLILAGRSCS	
PIC	MGQIVTLIQSIEEVLQEVFNVALIIVSVLCIVKGFVNI MRCGLFQLVTFLIISGRSCD	
SAB	MGQLFSFFEEVFNIIHEAINTALIAVSLIAALKGMINLWKSGLFQLIFFLTLAGRSCS	
TAC	MGQFISFMQEIP IFLQEALNIALVAVSLICIVKGLVNLVRCGLFQLMVFLVLAGRSCS	

Figura Ap6.1. Alineamiento múltiple correspondiente al péptido señal del precursor de las glicoproteínas (GPC) de los arenavirus. El alineamiento fué realizado con el programa Clustal X, utilizando los siguientes parámetros: *gap opening*=10, *gap extension*=0,1 en los pares, *gap opening*=10, *gap extension*=0,05 en el alineamiento múltiple y las tablas de cambios conservativos de la serie PAM en ambos casos. La línea individual en la parte superior es el consenso estricto entre las cepas MC2 y Candid #1 del virus Junín. Los puntos corresponden a las regiones con ausencia de consenso y/o gaps entre las cepas de Junín. Con letras de diferentes colores, se señala el grado de similitud de las distintas secuencias respecto del consenso de las cepas del virus Junín. En rojo los aminoácidos idénticos, en azul los cambios conservativos y en negro los diferentes. El hexapéptido con sombreado turquesa corresponde a una señal de miristoilación que podría estar en relación con el direccionamiento de GPC al retículo endoplásmico. LAG: Lassa-Nigeria; LAJ: Lassa-Josiah; LCA: LCM-Armstrong; LCE: LCM-WE; MOP: Mopeia; CD1: Junín-Candid #1; MC2: Junín-MC2; OLI: Oliveros; PIC: Pichindé; SAB: Sabiá; TAC: Tacaribe.

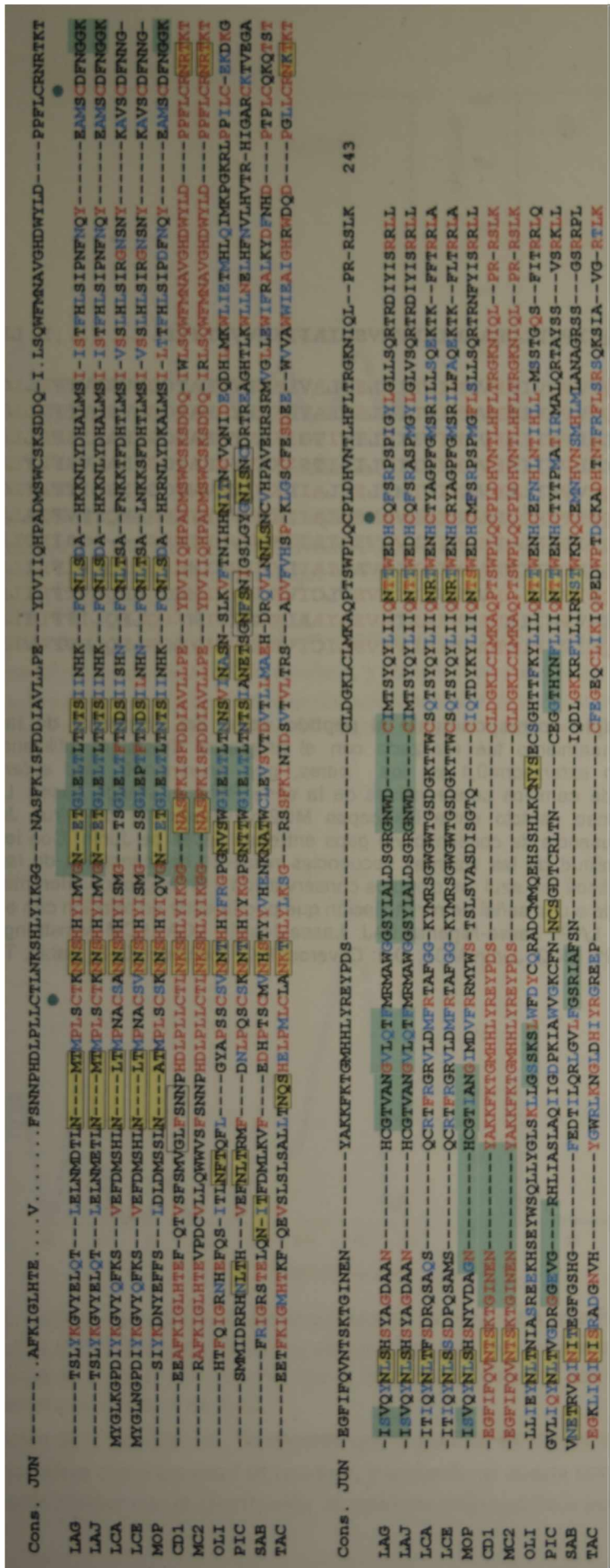
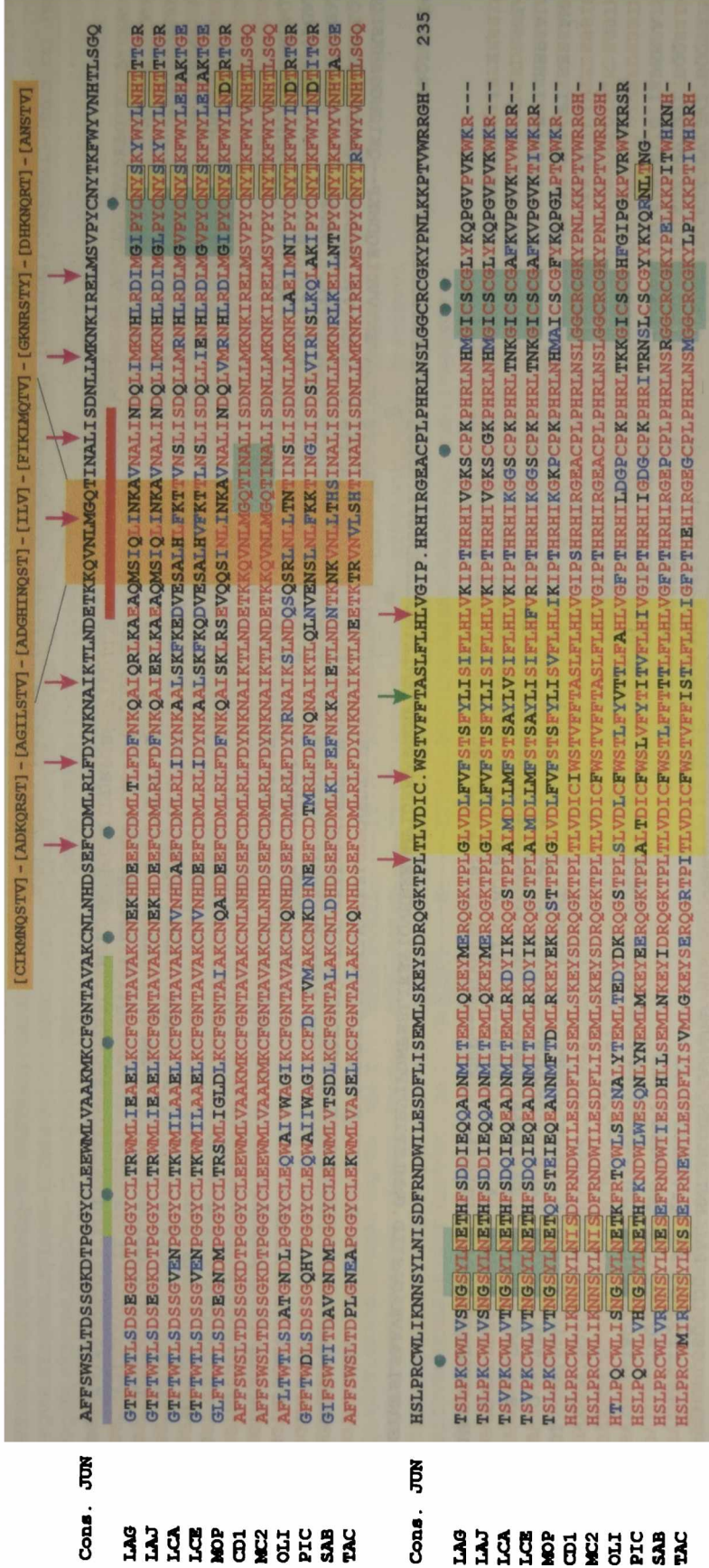


Figura Ap6.2. Alineamiento múltiple correspondiente a la glicoproteína G1 de los arenavirus. El alineamiento fue realizado con el programa Clustal X, utilizando los siguientes parámetros: gap opening=10, gap extension=0,1 en los pares, gap opening=0,05 en el alineamiento múltiple y las tablas de cambios conservativos de la serie PAM en ambos casos. La línea individual en la parte superior es el consenso estricto entre las cepas MC2 y Candid #1 del virus Junin. Los puntos corresponden a las regiones con ausencia de consenso entre las cepas de Junin y/o gaps entre ellas. Los guiones corresponden a gaps introducidos en las cepas de Junin para maximizar la homología con los otros arenavirus. Con letras de diferentes colores, se señala el grado de similitud de las distintas secuencias respecto del consenso de las cepas del virus Junin. En rojo los aminoácidos idénticos, en azul los cambios conservativos y en negro los diferentes. Los recuadros con fondo amarillo claro indican los sitios potenciales de N-glicosilación. La cantidad y posición de estos sitios es variable (LAG: 7; LAJ: 7; LCA: 6; LCE: 6; MOP: 7; CD1: 4; MC2: 4; OLI: 9; PIC: 11; SAB: 7; TAC: 4). Con un punto verde sobre la columna se indican las 3 cisteínas totalmente conservadas. La cantidad de estos residuos varía entre 5 y 8 (LAG: 6; LAJ: 6; LCA: 6; LCE: 6; MOP: 6; CD1: 6; MC2: 7; OLI: 8; PIC: 8; SAB: 5; TAC: 6). En turquesa claro se indican sitios potenciales de minioctilación. LAG: Lassa-Nigeria; LAJ: Lassa-Josiah; LCA: LCM-Armstrong; LCE: LCM-WE; MOP: Mopeia; CD1: Junin-Candid #1; MC2: Junin-MC2; OLI: Oliveros; PIC: Pichindé; SAB: Sabia; TAC: Tacaribe.

dominio de dimerización

120



hélice transmembrana

Figura Ap6.3. Alineamiento múltiple correspondiente a la glicoproteína G2 de los arenavirus. El alineamiento fué realizado con el programa Clustal X, utilizando los siguientes parámetros: *gap opening*=10, *gap extension*=0,1 en los pares, *gap opening*=10, *gap extension*=0,05 en el alineamiento múltiple y las tablas de cambios conservativos de la serie PAM en ambos casos. La línea individual en la parte superior es el consenso estricto entre las cepas MC2 y Candid #1 del virus Junin. Los puntos corresponden a las regiones con ausencia de consenso entre las cepas de Junin y/o *gaps* entre ellas. Los guiones corresponden a *gaps* introducidos en las cepas de Junin para maximizar la homología con los otros arenavirus. Con letras de diferentes colores, se señala el grado de similitud de las distintas secuencias respecto del consenso de las cepas del virus Junin. En rojo los aminoácidos idénticos, en azul los cambios conservativos y en negro los diferentes. Los recuadros con fondo amarillo claro indican los sitios potenciales de N-glicosilación. La cantidad y posición de estos sitios es casi constante (LAG: 4; LAJ: 4; LCA: 3; MOP: 4; CD1: 4; MC2: 4; OLI: 4; PIC: 5; SAB: 4; TAC: 4). Con un punto verde sobre la columna se indican las 9 cisteínas totalmente conservadas. La cantidad de estos residuos varía entre 9 y 10 (todos los arenavirus del Viejo Mundo: 9; todos los arenavirus del Nuevo Mundo: 10). Con barras de diferentes colores (lila, verde claro y rojo), ubicadas sobre el alineamiento, se indican tres regiones aminoacídicas potencialmente asociadas con la actividad fusogénica de G2. Con un sombreado naranja claro se señala el nonapéptido potencialmente responsable de estabilizar la asociación tetramérica de la porción extracelular de G2 (señal de dimerización) y, en el recuadro superior un consenso de la señal de dimerización, ampliado de acuerdo a los sitios hallados en la G2 y las proteínas N (ver Figura Ap.7.4). Las flechas de color fucsia en el primer bloque señalan la presencia de aminoácidos hidrofóbicos repetidos cada 7 residuos (*heptad repeats*); según lo descripto (Yu *et al.*, 1996), en entornos acuosos, los *heptad repeats* colaboran en la estabilización de interacciones proteína-proteína. Con un sombreado amarillo se indica la hélice transmembrana, responsable del anclaje de G2. Los aminoácidos indicados con flechas color fucsia en la zona transmembrana (en verde se señala una columna no conservada totalmente), podrían corresponder a *heptad repeats* que, en este caso, colaborarían en la estabilización del anclaje de G2 en la bicapa lipídica. En turquesa claro se indican sitios potenciales de miristilación. LAG: Lassa-Nigeria; LAJ: Lassa-Josiah; LCA: LCM-Armstrong; LCE: LCM-WE; MOP: Mopeia; CD1: Junin-Candid #1; MC2: Junin-MC2; OLI: Oliveros; PIC: Pichindé; SAB: Sabiá; TAC: Tacaribe.

Cons. JUN

LAG MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 LAJ MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 LCA MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 LCE MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 MOP MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 CD1 MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 MC2 MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 MAC MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 OLI MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 PIC MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 SAB MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG
 TAC MSASKEVPSFRWTQSLRRGLSQFTQTVKSDVLDKAKLIADSIDFNQ.AQVQRALRKRTRGEEEDLNKLRDLNKEVDRLMSRVSQRTVFKAGDLGR.E.MELAS.L.LKTK.RRAE-TG

Cons. JUN

LAG SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 LAJ SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 LCA SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 LCE SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 MOP SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 CD1 SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 MC2 SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 MAC SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 OLI SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 PIC SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 SAB SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN
 TAC SGGVYMGNLSSQQLAKRSEILRTLGFQ--QQFTGG--NGVVRVWDVKDP SKLNNQFGSVPALTIACMTVQGGETMNSVI QALWTLGLLYTVKYPNLSDLRLTQEHDCIQIVTKDESSIN

Cons. JUN

LAG ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 LAJ ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 LCA ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 LCE ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 MOP ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 CD1 ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 MC2 ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 MAC ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 OLI ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 PIC ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 SAB ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS
 TAC ISGYNE SILGAAVKAGACMLDGGNMLETIKVSFQTMGDIKLSILKVKKSLGMFVSDTPCE RNPYEN ILYKI CLSGDGPYI ASRFSIVGRAWENTVVDLEQDNKP--QKINGGSKSLQS

[CIKMNQSVV] - [ADKQRF] - [AGILSV] - [ADGHINQSF] - [ILV] - [FIKMQFV] - [GKRSY] - [DHRGQRT] - [ANSTV]

dominio de dimerización

Figura Ap6.4. Alineamiento múltiple correspondiente a la nucleoproteína de los arnavirus. (Ver leyenda en la página siguiente).

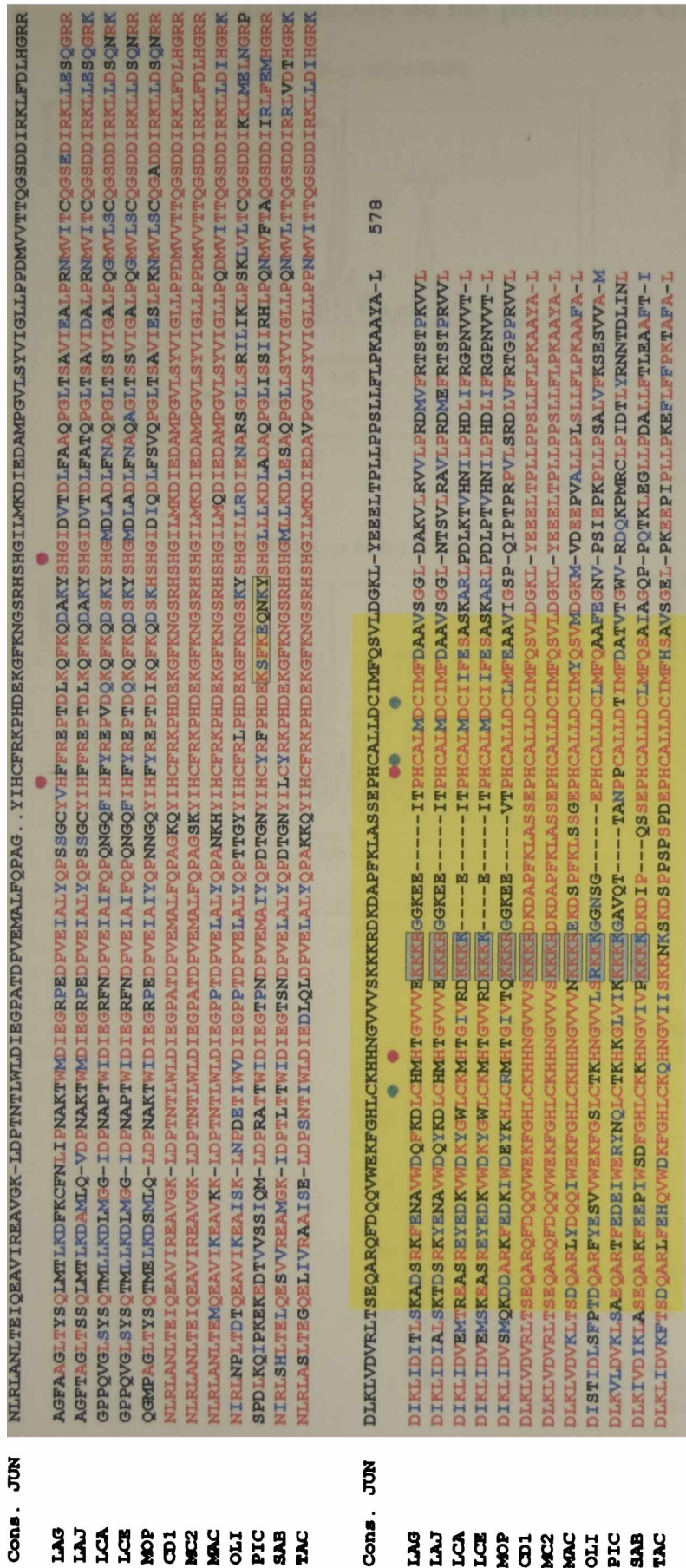


Figura Ap6.4. Alineamiento múltiple correspondiente a la nucleoproteína de los arenavirus. El alineamiento fué realizado con el programa Clustal X, utilizando los siguientes parámetros: *gap opening*=10, *gap extension*=0,1 en los pares, *gap opening*=10, *gap extension*=0,05 en el alineamiento múltiple y las tablas de cambios conservativos de la serie PAM en ambos casos. La línea individual en la parte superior es el consenso estricto entre las cepas MC2 y Candid #1 del virus Junín. Los puntos corresponden a las regiones con ausencia de consenso entre las cepas de Junín y/o *gaps* entre ellas. Los guiones corresponden a *gaps* introducidos en las cepas de Junín para maximizar la homología con los otros arenavirus. En rojo los aminoácidos idénticos, se señala el grado de similitud de las distintas secuencias respecto del consenso de las cepas del virus Junín. En azul los cambios conservativos y en negro los diferentes. Los recuadros con fondo amarillo claro indican los sitios potenciales de fosforilación por tirosininasas. En todos los arenavirus se encuentra 1 sitio conservado en posición (PIC es el único que presenta 2 sitios). Con un punto verde sobre la columna se indican las cisteínas conservadas en, al menos, 11 de los 12 arenavirus. La cantidad de estos residuos varía entre 7 y 10 (LAG: 10; LAJ: 9; LCA: 7; LCE: 7; MOP: 9; CD1: 7; MC2: 7; MAC: 7; PIC: 7; SAB: 7; TAC: 7). Con un punto fucsia sobre la columna se indican las histidinas conservadas en, al menos, 11 de los 12 arenavirus. La cantidad de estos residuos varía entre 7 y 13 (LAG: 7; LAJ: 7; LCA: 8; LCE: 9; MOP: 8; CD1: 11; MC2: 11; MAC: 13; OLI: 8; PIC: 10; SAB: 9; TAC: 10). Con un sombreado naranja claro se señalan tres regiones (numeradas en la parte superior) que presentan nona péptidos potencialmente responsables de estabilizar las asociaciones proteína-proteína (señales de dimerización) y, en el recuadro inferior del nonapéptido 3, un consenso general de la señal de dimerización, ampliado de acuerdo a los sitios hallados en las proteínas N y la glicoproteína G2 (ver Figura Ap.7.3). Los sitios que responden al consenso existente en el PROSITE están subrayados con una línea violeta gruesa. Con un sombreado amarillo se indica la región aminoacídica con capacidad de unir zinc (motivo tipo zinc *finger*) y, responsable de la interacción con el RNA, al menos, *in vitro*. Los recuadros con fondo verde claro señalan grupos de más de 3 residuos básicos consecutivos. LAG: Lassa-Nigeria; LAJ: Lassa-Josiah; LCA: LCM-Armstrong; LCE: LCM-WE; MOP: Mopeia; CD1: Junin-Candid #1; MC2: Junin-MC2; MAC: Machupo; OLI: Oliveros; PIC: Pichindé; SAB: Sabiá; TAC: Tacaribe.

Análisis de períodos dominantes en estructuras secundarias predichas de las proteínas GPC de arnavirus

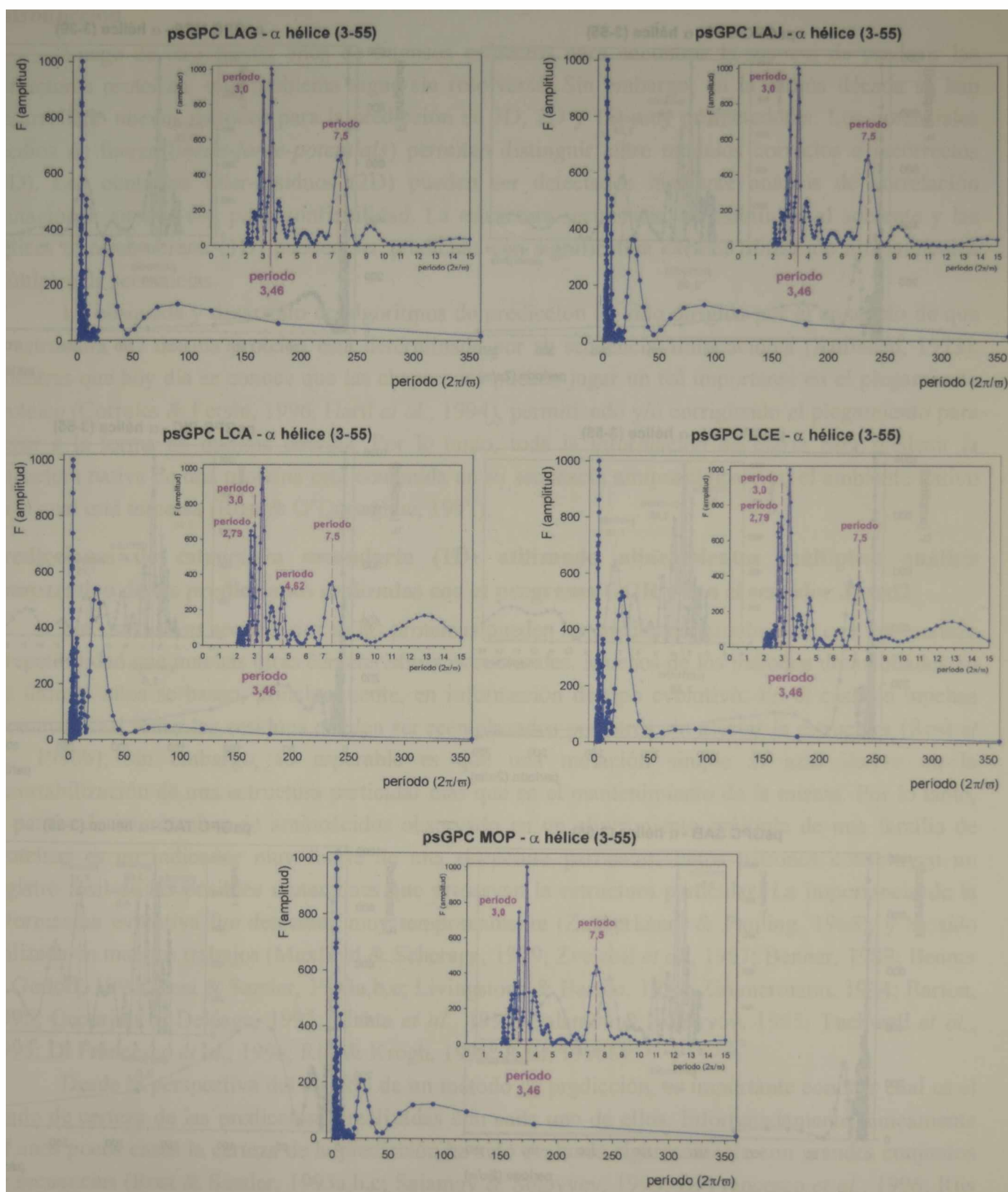


Figura Ap7.1. Análisis de períodos dominantes en las α hélices predichas en el péptido señal de los arnavirus del Viejo Mundo. Como se puede observar, en todos los gráficos aparece un período dominante de 3,46. Esto explica el valor de AI superior a 2, ya que el área comprendido entre 85 y 110 se hace significativa *versus* el área total (ver Figura 13.7 y Apéndice 5). Sin embargo, también se observa la presencia de otros períodos de importancia. Probablemente, esto último está en relación con la longitud de la estructura (40 o 50 residuos) y su carácter predominantemente apolar (~50% de aminoácidos hidrofóbicos y ~30% de aminoácidos ambivalentes), lo cual permite detectar repeticiones periódicas de diferente magnitud. En conclusión, para estructuras superiores a 15 o 20 residuos, el valor de AI que permite definir anfipaticidad debería ser muy superior a 2.

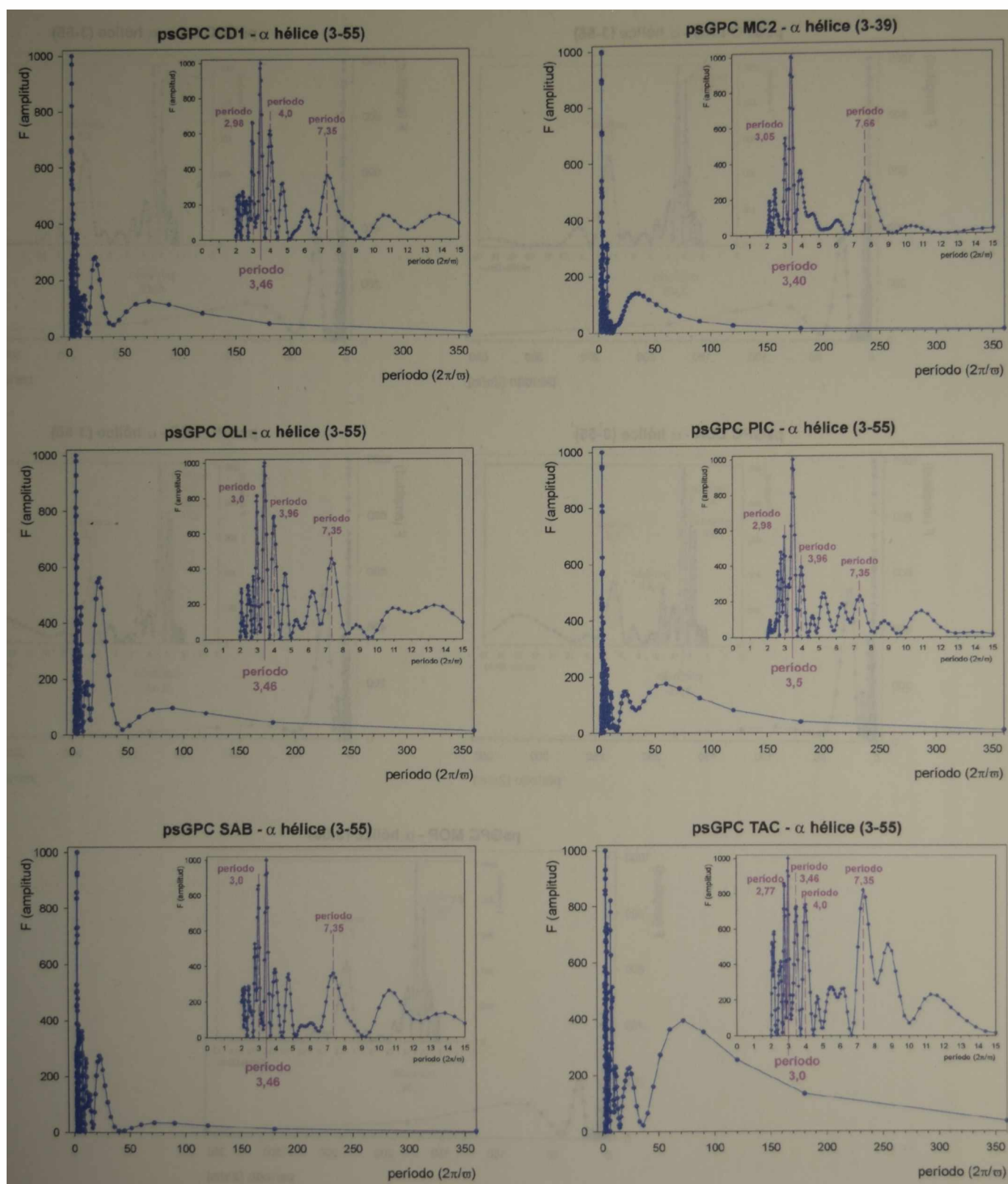


Figura Ap7.2. Análisis de periodos dominantes en las α hélices predichas en el péptido señal de los arenavirus del Nuevo Mundo. Como se puede observar, en todos los gráficos aparece un periodo dominante de $\sim 3,5$. Esto explica el valor de AI superior a 2, ya que el área comprendido entre 85 y 110 se hace significativa *versus* el área total (ver Figura 13.7 y Apéndice 5). Sin embargo, también se observa la presencia de otros periodos de importancia. Probablemente, esto último está en relación con la longitud de la estructura (40 o 50 residuos) y su carácter predominantemente apolar ($\sim 50\%$ de aminoácidos hidrofóbicos y $\sim 30\%$ de aminoácidos ambivalentes), lo cual permite detectar repeticiones periódicas de diferente magnitud. Como ejemplos particulares se pueden tomar en cuenta los valores de Sabiá t Tacaribe, que son inferiores a 2. En Sabiá (AI=1,83), el periodo predominante es $\sim 3,5$, pero existe un periodo de 3,0 de aproximadamente la misma magnitud. Mientras que en Tacaribe (AI=1,55) el periodo dominante es 3,0 y existen, al menos, 3 periodos adicionales de mayor o igual magnitud que el de 3,46. En conclusión, para estructuras superiores a 15 o 20 residuos, el valor de AI que permite definir anfipaticidad debería ser muy superior a 2.

Predicciones estructurales en proteínas

Introducción

Luego de casi treinta años de intensos esfuerzos para encontrar la manera de predecir las estructuras proteicas, este problema sigue sin resolverse. Sin embargo, en la última década se han desarrollado nuevos métodos para la predicción en 3D, 2D y 1D muy prometedores. Los potenciales medios de fuerza (*mean-force-potentials*) permiten distinguir entre modelos correctos e incorrectos (3D). Los contactos inter-residuos (2D) pueden ser detectados mediante análisis de correlación mutacional, aunque con poca confiabilidad. La estructura secundaria, accesibilidad al solvente y las hélices transmembrana (1D) pueden ser predichas con significativa certeza utilizando alineamientos múltiples de secuencias.

La búsqueda y desarrollo de algoritmos de predicción ha sido dirigida por el concepto de que la estructura 3D de una proteína está determinada por su secuencia aminoacídica (Anfinsen, 1973). Mientras que hoy día se conoce que las chaperonas pueden jugar un rol importante en el plegamiento proteico (Corrales & Fersht, 1996, Hartl *et al.*, 1994), permitiendo y/o corrigiendo el plegamiento para llegar a la forma de mínima energía. Por lo tanto, toda la información necesaria para predecir la estructura nativa de una proteína está contenida en su secuencia aminoacídica y en el ambiente nativo en el cual está inmersa (Rost & O'Donoghue, 1997).

Predicciones de estructura secundaria (1D) utilizando alineamientos múltiples: análisis comparativo de las predicciones realizadas con el programa GOR y con el servidor Jpred2

Las estructuras secundarias en las proteínas pueden ser predichas, usualmente, con más certeza y repetibilidad que muchas otras características estructurales. Muchos de los métodos desarrollados en los últimos años se basan, principalmente, en información de tipo evolutivo. En el caso de muchas proteínas, casi todos los residuos pueden ser reemplazados por otros sin afectar la estructura (Rost *et al.*, 1996b); sin embargo, lo esperable es que una mutación simple al azar derive en la desestabilización de una estructura particular más que en el mantenimiento de la misma. Por lo tanto, el patrón de intercambio de aminoácidos observado en un alineamiento múltiple de una familia de proteínas es un indicador muy fuerte de una estructura particular. Estos patrones constituyen un registro fósil de las posibles mutaciones que preservan la estructura particular. La importancia de la información evolutiva fue detectada muy tempranamente (Zuckerandl & Pauling, 1965), y ha sido utilizada en muchos trabajos (Maxfield & Scheraga, 1979; Zvelebil *et al.*, 1987; Benner, 1989; Benner & Gerloff, 1990; Rost & Sander, 1993a,b,c; Livingstone & Barton, 1994; Zimmermann, 1994; Barton, 1995; Geourjon & Deléage, 1995; Mehta *et al.*, 1995; Salamov & Solovyev, 1995; Tuckwell *et al.*, 1995; Di Francesco *et al.*, 1996; Riis & Krogh, 1996; Rost, 1996b).

Desde la perspectiva del usuario de un método de predicción, es importante conocer cual es el grado de certeza de las predicciones realizadas con cada uno de ellos. Infortunadamente, únicamente en unos pocos casos la certeza de la predicción ha sido evaluada rigurosamente con grandes conjuntos de secuencias (Rost & Sander, 1993a,b,c; Salamov & Solovyev, 1995; Di Francesco *et al.*, 1996; Riis & Krogh, 1996; Rost, 1996b); en estos casos el promedio de certeza de las predicciones en tres estados (α -hélice, β -hoja plegada y *loop*) es superior al 70 %, con una desviación estándar típica cercana al 10% (una desviación estándar sobre más de 500-700 cadenas proteicas únicas; Rost, 1996a,b). Para algunos de los métodos, los altos *scores* de la predicción han sido correlacionados con la certeza de la misma (Rost & Sander, 1993a,b,c; Di Francesco *et al.*, 1996; Rost, 1996b). En la práctica, esto permite que el usuario enfoque su atención sobre regiones en las cuales la predicción tiene mayor probabilidad

de ser correcta (alrededor del 45 % de todos los residuos son predichos con niveles de certeza comparables con el modelado por homología (*homology modelling*, Rost, 1996b).

Los métodos de predicción de estructura secundaria que no usan información evolutiva, tales como el GOR IV (Garnier *et al.*, 1996), tienen una certeza promedio del 60 %, y menos del 10 % de los residuos son predichos con la misma certeza que en la predicción derivada del modelado por homología. A manera de ejemplo, en la **Figura Ap8.1** se muestran las diferencias en la predicción obtenidas con un método de generación de consenso (que usa información evolutiva) y el GOR IV (ver también **Capítulo 13, Figura 13.2**).

Por lo tanto, los métodos actuales son mucho más precisos que los anteriores. Una de las bases de este incremento en la precisión radica en el hecho de utilizar como punto de partida alineamientos múltiples; sin embargo, hay que tener sumo cuidado en cómo se hacen los alineamientos, ya que errores en los mismos generan una sensible disminución en la certeza de la predicción (Di Francesco *et al.*, 1996; Rost, 1996b; Rost & Valencia, 1996).

De la misma manera, otras predicciones, tales como la accesibilidad al solvente y el análisis de hélices transmembrana, se han visto sensiblemente mejoradas por la incorporación de información evolutiva en los análisis (Benner, 1989; Benner & Gerloff, 1990; Jones *et al.*, 1992, 1994, 1995; Persson & Argos, 1994, 1996; von Heijne, 1994, 1996; Benner *et al.*, 1994; Esposito *et al.*, 1994; Monge *et al.*, 1994; Rost & Sander, 1994a,c; Russell & Barton, 1990, 1993, 1994; Wako & Blundell, 1994; Cornette *et al.*, 1995; Hansen *et al.*, 1995, 1996; Mumenthaler & Braun, 1995; Nilges, 1991, 1995; Neuwald *et al.*, 1995; Rost, 1995a,b,c, 1996a,b,c; Cohen & Presnell, 1996; Efremov & Vergoten, 1996; Fariselli & Casadio, 1996; Galaktionov & Marshall, 1996; Hansen *et al.*, 1996; Nielsen *et al.*, 1996; Rost *et al.*, 1996a,b; Thompson & Goldstein, 1996a,b).

En síntesis, uno de los servidores actualmente más utilizados para predecir estructuras secundarias es el JPred2 (Cuff & Barton, 1999). El primer paso es la generación de un alineamiento múltiple con baja redundancia, las secuencias redundantes en el alineamiento generan desviaciones importantes en los *profiles* que se utilizan para los métodos de predicción. A continuación, se efectúa la predicción con distintos algoritmos en forma individual y se construye un consenso general.

En principio, el JPred fue desarrollado como resultado de un estudio de certeza de diferentes métodos de predicción. Durante esta evaluación, se generaron dos subconjuntos independientes de proteínas no relacionadas, uno de ellos con 126 proteínas y el otro con 396 secuencias. En la **Tabla Ap8.a** se muestran los porcentajes de certeza obtenidos para cada subconjunto de proteínas.

Tabla Ap8.a. Certeza en la predicción de distintos métodos

Método	Conjunto de 126 proteínas	Conjunto de 396 proteínas
PHD	73,5 %	71,9 %
DSC	71,1 %	68,4 %
PREDATOR	70,3 %	68,6 %
NNSSP	72,7 %	71,4 %
Mulpred	67,2 %	66,1 %
Zpred	66,7 %	64,8 %
Consenso (Jpred)	74,8 %	72,9 %

En conclusión, hasta el presente, la mejor opción para evaluar las posibles estructuras secundarias en una proteína es utilizar alguno de los algoritmos que incorporan información evolutiva y que generan un consenso entre diferentes programas de predicción.

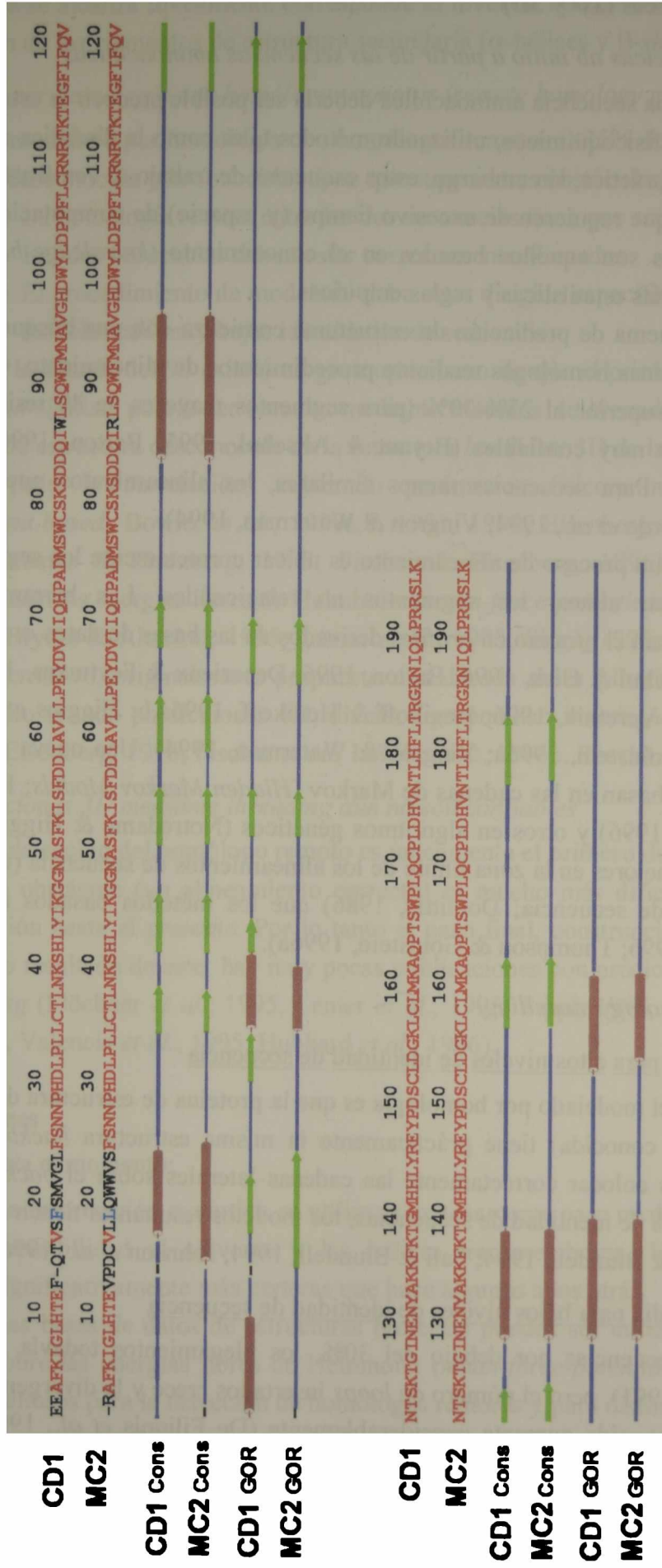


Figura Ap8.1. Comparación de las secuencias aminoacídicas de la glicoproteína G1 de las cepas MC2 y Candid #1 del virus Junin. En los dos primeros renglones se muestran las correspondientes secuencias aminoacídicas, en marrón las identidades, en verde los cambios conservativos y en negro las diferencias. Las líneas CD1 Cons y MC2 Cons muestran esquemáticamente las predicciones de estructura secundaria obtenidas con el servidor JPred2, con un cilindro rojo se indican estructuras tipo α hélice, con flechas verdes se indican estructuras tipo β hoja plegada y con líneas azules los loops. Las otras líneas gráficas derivan de la predicción con el programa GOR IV, los códigos son iguales que los anteriores. Como se puede observar, las predicciones obtenidas con el programa GOR IV son diferentes a las obtenidas con el JPred2, inclusive en algunas zonas de total identidad entre las dos proteínas, el GOR IV predice diferentes estructuras para Candid #1 (CD1) y para MC2 (zona resaltada en amarillo claro).

Modelado de estructuras proteicas (2D y 3D)

Predicciones de estructuras proteicas ab initio a partir de las secuencias aminoacídicas

En principio, conocida una secuencia aminoacídica debería ser posible predecir la estructura de la proteína a partir de principios fisicoquímicos, utilizando métodos tales como la dinámica molecular (Levitt & Warshel, 1975). En la práctica, sin embargo, estos esquemas de trabajo se ven frustrados por la complejidad de los cálculos, que requieren de excesivo tiempo (y espacio) de computación. Por lo tanto, los métodos más exitosos son aquellos basados en el conocimiento (*knowledge-based*) que utilizan una combinación de teorías estadísticas y reglas empíricas.

Por ende, cualquier esquema de predicción de estructuras comienza con una búsqueda en las bases de datos para buscar proteínas homólogas mediante procedimientos de alineamiento. Cuando la identidad de las secuencias es superior al 25%-30% (para segmentos mayores de 80 residuos), los alineamientos son normalmente muy confiables (Bryant & Altschul, 1995; Barton, 1996; Taylor, 1996; Schneider *et al.*, 1997). Para secuencias menos similares, los alineamientos pueden fallar (Henikoff & Henikoff, 1993; Bordo *et al.*, 1994; Vingron & Waterman, 1994).

El objetivo principal de un proceso de alineamiento es ubicar correctamente los segmentos de secuencia relacionados y evitar alinear los segmentos no relacionados. Las herramientas de alineamiento más avanzadas basan el proceso en *profiles* derivados de las bases de datos o de familias particulares de secuencias (Altschul & Gish, 1996; Barton, 1996; Deperieux & Feytmans, 1992; Feng & Doolittle, 1996; Gribskov & Veretnik, 1996; Henikoff & Henikoff, 1996a,b; Higgins *et al.*, 1996; Pearson, 1996; Thompson & Goldstein, 1996a; Vingron & Waterman, 1994). Una nueva generación de métodos de alineamiento se basan en las cadenas de Markov (*Hidden Markov Models*; Hubbard & Park, 1995; Hughey & Krogh, 1996) y otros en algoritmos genéticos (Notredame & Higgins, 1996). Esto últimos dos métodos son mejores en la zona límite de los alineamientos de secuencia (usualmente entre 20 y 30% de identidad de secuencia; Doolittle, 1986) que los métodos basados en *profiles* (Higgins *et al.*, 1996; Taylor, 1996; Thompson & Goldstein, 1996a).

Modelado por homología (Homology modelling)

Predicción con certeza atómica para altos niveles de identidad de secuencia

La presunción básica del modelado por homología es que la proteína de estructura desconocida y la homóloga (de estructura conocida) tiene prácticamente la misma estructura *backbone* en las regiones alineadas. La tarea es colocar correctamente las cadenas laterales sobre el *backbone*. Para niveles de alrededor del 70-90% de identidad de secuencias, los modelos resultantes tienen alta certeza (De Filippis *et al.*, 1994; May & Blundell, 1994; Sali & Blundell, 1994; Johnson *et al.*, 1996).

Predicción con certeza intermedia para bajos niveles de identidad de secuencia

Para identidades de secuencias por debajo del 30%, los plegamientos todavía pueden ser iguales (Sander & Schneider, 1991), pero el número de *loops* insertados crece y la divergencia entre la estructura desconocida y la conocida aumenta considerablemente (De Filippis *et al.*, 1994; May & Blundell, 1994; Chinea *et al.*, 1995; Mosimann *et al.*, 1995; Moulton *et al.*, 1995; Samudrala *et al.*, 1995; Vinals *et al.*, 1995). El modelado de las regiones de *loops* es un problema muy difícil (Cardozo *et al.*, 1995; Mosimann *et al.*, 1995; Sali *et al.*, 1995); aún los mejores métodos raramente alcanzan el nivel de certeza atómico y siempre han dado resultados completamente diferentes a la estructura correcta. Desde un punto de vista pesimista, en las predicciones 3D en estos niveles de identidad la

certeza que se alcanza típicamente corresponde al nivel de un *ribbon plot*; por ejemplo, la posición y orientación de los elementos de estructura secundaria (α -hélices y β -*sheets*) puede ser identificada.

Modelado por homología con homólogos remotos (remote homology modelling)

El modelado por homología con homólogos remotos (<25% de identidad de secuencia entre la estructura desconocida y la conocida) tiene que superar tres obstáculos principales:

- (1) La homología remota entre las dos secuencias tiene que ser detectada
- (2) Las dos secuencias deben estar correctamente alineadas
- (3) El procedimiento de modelado debe estar dirigido específicamente a la difícil tarea de modelar con muy bajas identidades de secuencia.

En la década del 90 existió un gran optimismo de que el primer obstáculo, la detección de un plegamiento similar, podría ser resuelta mediante métodos de *threading*. La idea básica es ubicar la secuencia de estructura desconocida en la estructura *backbone* 3D de la estructura conocida, evaluando en cada paso la correlación, utilizando parámetros fisicoquímicos basados en el entorno (*environment-based*; Bowie, *et al.*, 1990a,b, 1991, 1996; Eisenberg, *et al.*, 1991; Ouzounis, *et al.*, 1993; Wilmanns & Eisenberg, 1995; Fischer & Eisenberg, 1996). Muchos de los métodos de *threading* usan las energías libres de Helmholtz (*mean-force-potentials*) derivados del PDB (Kocher *et al.*, 1994; Bryant & Altschul, 1995; Lemer *et al.*, 1995; Sippl, 1995; Wodak & Rومان, 1993). Un método alternativo, originalmente propuesto una década antes (Sheridan *et al.*, 1985), es realizar el apareamiento usando predicciones 1D; diversos grupos han investigado este concepto (Rost, 1995a; Fischer & Eisenberg, 1996; Fischer *et al.*, 1996b; Rost, 1995b, Rost *et al.*, 1996c; Russell *et al.*, 1996).

Las predicciones 3D mediante threading aún no son confiables

La detección del homólogo remoto es únicamente el primero de los tres obstáculos. Parece que el segundo obstáculo (un alineamiento correcto) es mucho más difícil e, infortunadamente, no tiene solución hasta el presente. Por lo tanto el paso final, construcción del modelo 3D, usualmente falla. Como resultado de esto, hay muy pocas publicaciones con predicciones 3D certeras por métodos de *threading* (Flöckner *et al.*, 1995, Lemer *et al.*, 1995, Sippl, 1995, Rost *et al.*, 1996c, Hubbard & Park, 1995, Valencia *et al.*, 1995, Hubbard *et al.*, 1996).

Conclusiones

Hasta el momento:

- (1) La información evolutiva es utilizada exitosamente para predecir la estructura secundaria, la accesibilidad al solvente y las hélices transmembrana. Estas predicciones en 1D son significativamente más certeras que hace algunos años atrás.
- (2) Las bases de datos de estructuras proteicas pueden ser utilizadas para obtener información sobre las energías libres de Helmholtz (*mean-force-potentials*). Estos parámetros son muy valiosos para la detección de homólogos remotos y para distinguir entre modelos alternativos (generados teóricamente o experimentales).
- (3) El modelado por homología permite predecir la estructura 3D de alrededor del 10% de las proteínas expresadas.
- (4) Algunas de las mejoras en las técnicas de reconocimiento de plegamiento (*threading*), y en las técnicas de alineamiento facilitan el modelado 3D de homólogos remotos para una fracción adicional considerable de las proteínas expresadas.

Reglas de energía libre para el cálculo de plegamiento de RNA
(Freier *et al*, 1987; Turner *et al*, 1988)

Energías libres de *stacking*, a 37 °C

Organización de los datos:

	Y			
(X)	A	C	G	U
	5' ==> 3'			
	AX			
	AY			
	3' <== 5'			
(A)
(C)
(G)
(U)	-0.7	-0.1	-0.7	-0.1

	Y					Y					Y					Y				
(X)	A	C	G	U		A	C	G	U		A	C	G	U		A	C	G	U	
	5' --> 3'					5' --> 3'					5' --> 3'					5' --> 3'				
	AX					AX					AX					AX				
	AY					CY					GY					UY				
	3' <-- 5'					3' <-- 5'					3' <-- 5'					3' <-- 5'				
(A)	-0.9	
(C)	-2.1	.
(G)	-1.7	-0.5
(U)	-0.9	-0.7

	Y					Y					Y					Y					
(X)	A	C	G	U		A	C	G	U		A	C	G	U		A	C	G	U		
	5' --> 3'					5' --> 3'					5' --> 3'					5' --> 3'					
	CX					CX					CX					CX					
	AY					CY					GY					UY					
	3' <-- 5'					3' <-- 5'					3' <-- 5'					3' <-- 5'					
(A)	-1.8	
(C)	-2.9	
(G)	-2.0	-1.5
(U)	-1.7	-1.5

	Y					Y					Y					Y							
(X)	A	C	G	U		A	C	G	U		A	C	G	U		A	C	G	U				
	5' --> 3'					5' --> 3'					5' --> 3'					5' --> 3'							
	GX					GX					GX					GX							
	AY					CY					GY					UY							
	3' <-- 5'					3' <-- 5'					3' <-- 5'					3' <-- 5'							
(A)	-2.3		
(C)	-3.4		
(G)	-2.9	-1.3	
(U)	-2.1	-1.9

Apéndice 9

(X)	Y				Y				Y				Y			
	A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
	5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	UX AY				UX CY				UX GY				UX UY			
	3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
(A)	.	.	.	-1.1	-0.7
(C)	.	.	-2.3	-1.3
(G)	.	-1.8	.	-0.7	-1.5	.	-0.6
(U)	-0.9	.	-0.5	-0.5	.	-0.5

Energías libres de *mismatch* terminales y *stacking*, a 37 °C

(X)	Y				Y				Y				Y			
	A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
	5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	AX AY				AX CY				AX GY				AX UY			
	3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
(A)	.	.	.	-1.0	.	.	-0.7	.	.	.	-1.1	-0.8	-1.0	-1.0	-1.0	-1.0
(C)	.	.	-1.1	.	.	-1.1	.	.	.	-1.6	.	-0.7	-0.7	-0.7	-0.7	-0.7
(G)	.	-1.9	.	-1.5	-1.0	-0.9	.	.	-1.9	.	-1.5	-0.8	-1.0	-1.0	-1.0	-1.0
(U)	-0.8	.	-0.8	.	-0.7	-0.7	.	.	-0.8	-0.8	.	-0.8	-0.8	-0.8	-0.8	-0.8

(X)	Y				Y				Y				Y			
	A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
	5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	CX AY				CX CY				CX GY				CX UY			
	3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
(A)	.	.	.	-0.8	.	.	-0.6	-1.9	-2.0	-1.9	-1.9	-0.6
(C)	.	.	-1.3	.	.	-0.6	.	-1.0	-1.1	-1.0	-0.8	-0.8
(G)	.	-2.0	.	-1.4	-1.1	-0.9	-1.9	-1.9	-1.9	-1.9	-1.9	-1.5	.	.	.	-1.1
(U)	-1.0	.	-1.0	.	-0.7	-0.7	.	-1.4	-1.5	-1.4	-1.2	-0.8	.	.	.	-0.8

(X)	Y				Y				Y				Y			
	A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
	5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	GX AY				GX CY				GX GY				GX UY			
	3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
(A)	.	.	.	-1.1	-1.1	-1.3	-1.3	-1.3	.	.	-1.2	-0.8	-1.0	-1.0	-1.0	-1.0
(C)	.	.	-1.3	.	-1.1	-0.6	-0.6	-0.5	.	-1.4	.	-0.7	-0.7	-0.7	-0.7	-0.7
(G)	.	-1.9	.	-1.5	-1.6	-1.5	-1.4	-1.5	-1.9	.	-1.6	-0.8	-1.0	-1.0	-1.0	-1.0
(U)	-1.0	.	-1.0	.	-0.8	-0.8	-0.8	-0.7	-1.0	-1.0	.	-0.8	-0.8	-0.8	-0.8	-0.8

(X)	Y				Y				Y				Y			
	A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
	5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	UX AY				UX CY				UX GY				UX UY			
	3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
(A)	-1.0	-0.8	-1.1	-0.9	.	.	-0.5	-1.5	-1.4	-1.5	-1.4	-0.5
(C)	-0.7	-0.6	-0.3	-0.5	.	-0.5	.	-0.9	-0.9	-0.7	-0.7	-0.7
(G)	-1.1	-0.9	-1.2	-0.9	-0.8	-0.7	-1.5	-1.4	-1.6	-1.4	-1.4	-1.2	.	.	.	-0.9
(U)	-0.3	-0.6	-0.3	-0.5	-0.7	-0.7	-0.9	-1.1	-0.9	-0.9	-0.9	-0.8	.	.	.	-0.8

Energías libres de *single base stacking*, a 37 °C

X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	AX				AX				AX				AX		
	A				C				G				U		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
												-0.8	-0.5	-0.8	-0.6
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	CX				CX				CX				CX		
	A				C				G				U		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
								-1.7	-0.8	-1.7	-1.2				
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	GX				GX				GX				GX		
	A				C				G				U		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
				-1.1	-0.4	-1.3	-0.6					-0.8	-0.5	-0.8	-0.6
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	UX				UX				UX				UX		
	A				C				G				U		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
								-1.2	-0.5	-1.2	-0.7				
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	A				A				A				A		
	AX				CX				GX				UX		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
												-0.3	-0.1	-0.2	-0.2
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	C				C				C				C		
	AX				CX				GX				UX		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
								-0.2	-0.3	0.0	0.0				
X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
	G				G				G				G		
	AX				CX				GX				UX		
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
				-0.5	-0.2	-0.2	-0.1					-0.2	-0.2	-0.2	-0.2

Apéndice 9

X				X				X				X			
A	C	G	U	A	C	G	U	A	C	G	U	A	C	G	U
5' --> 3'				5' --> 3'				5' --> 3'				5' --> 3'			
U				U				U				U			
AX				CX				GX				UX			
3' <-- 5'				3' <-- 5'				3' <-- 5'				3' <-- 5'			
-0.3	-0.3	-0.4	-0.2	-0.2	-0.2	-0.2	-0.2

Energías libres de loops, a 37 °C (se interpola cuando se necesita)

TAMAÑO	INTERNAL	BULGE	HAIRPIN
1	.	3.9	.
2	4.1	3.1	.
3	4.5	3.5	4.5
4	4.9	4.2	5.5
5	5.3	4.8	4.9
6	5.7	5.0	5.1
7	5.9	5.2	5.2
8	6.0	5.3	5.5
9	6.1	5.4	5.8
10	6.3	5.5	5.9
11	6.4	5.7	6.0
12	6.4	5.7	6.1
13	6.5	5.8	6.2
14	6.6	5.9	6.3
15	6.7	6.0	6.4
16	6.8	6.1	6.4
17	6.8	6.1	6.5
18	6.9	6.2	6.5
19	6.9	6.2	6.6
20	7.0	6.3	6.7
21	7.1	6.3	6.7
22	7.1	6.4	6.8
23	7.1	6.4	6.8
24	7.2	6.5	6.9
25	7.2	6.5	6.9
26	7.3	6.5	6.9
27	7.3	6.6	7.0
28	7.4	6.7	7.0
29	7.4	6.7	7.1
30	7.4	6.7	7.1

Energías libres de tetra-loops ultraestables, a 37 °C

Sec.	Energía
GAAA	-2.0
GCAA	-2.0
GAGA	-2.0
GUGA	-2.0
GGAA	-2.0
UUCG	-2.0
UACG	-2.0
GCGA	-2.0

Miscelánea de energías libres, a 37 °C

Energías libres por tamaño del loop

misc. loop functions				
> internal, bulge or hairpin loops > 30: $dS(T)=dS(30)+param*\ln(n/30)$				1.079
asymmetric internal loops: the ninio equation				
> the maximum correction				3.0
> the $f(m)$ array	0.4	0.3	0.2	0.1
multibranch loops				
> offset, free base penalty, helix penalty	4.6	0.4	0.1	

Esperanza probabilística de oligonucleótidos en los RNA S de arenavirus y genomas de mamíferos

Los virus con genomas de RNA lineal y cuya replicación y/o transcripción es realizada por RNA polimerasas RNA dependientes codificadas por el genoma viral, se caracterizan por presentar alta variabilidad, especialmente en los extremos.

Teniendo en cuenta que las RNA polimerasas RNA dependientes virales, usualmente, requieren de un primer para iniciar la síntesis, la mayor parte de los autores se remiten a realizar experimentos *in vitro* con oligonucleótidos sintéticos y/o mRNAs (o RNAs) definidos para comprobar la capacidad de las RNA polimerasas virales de utilizarlos como *primers*.

Sin embargo, los resultados obtenidos con estos experimentos constituyen una ligera aproximación a la verdad. Es de valor muy relativo la extrapolación de la información obtenida mediante ensayos *in vitro* a los mecanismos *in vivo*. Esto es especialmente cierto cuando, en los ensayos *in vitro*, la oferta de *primers* es definida por el usuario; mientras que *in vivo* la oferta de *primers* es definida por el sistema celular en el cual se replica y/o transcribe el genoma viral. En particular, es de especial importancia la organización informativa del genoma de la célula huésped y la dimensión del mismo.

Dado que la disponibilidad de información acerca de genomas completos de mamíferos es inexistente, las únicas herramientas disponibles para intentar evaluar modelos alternativos (utilizados *in vivo*) son las del cálculo probabilístico.

Desde este punto de vista, se deben tener en cuenta tres premisas:

- ◆ La probabilidad de existencia de un oligonucleótido determinado depende de su longitud y de la frecuencia de cada uno de los residuos que lo componen.
- ◆ La esperanza probabilística (número de veces que es probable encontrarlo) de encontrar un oligonucleótido determinado en un genoma de longitud determinada, depende de la probabilidad de existencia del oligonucleótido particular y del tamaño del genoma en cuestión.
- ◆ La probabilidad de encontrar un oligonucleótido determinado un número finito de veces depende de las dos premisas anteriores.

En función de lo anterior, y tomando como ejemplo el RNA S genómico de los arenavirus y sus extremos 5', se puede calcular la esperanza probabilística de encontrar los extremos 5' de los RNAs S descritos para las cepas del virus Junín y para el resto de los arenavirus, tanto en los genomas de arenavirus y como en los genomas de mamíferos.

Para ello, debemos tener en cuenta que la longitud promedio de un RNA S es de 3413 nt. y las frecuencias promedio de los distintos nucleótidos en los RNA S de los arenavirus es:

A: 0,26	G: 0,22
C: 0,23	U: 0,29

Extremo 5' del RNA S genómico descrito para la cepa MC2 del virus Junín

Secuencia: 5' UGCAGUAAGGGG 3'

La probabilidad de encontrar un oligonucleótido de 12 residuos con esa composición en base a las frecuencias promedio de cada residuo en los arenavirus es:

$$(0,26)^3 * (0,23) * (0,22)^6 * (0,29)^2 = 3,9 \times 10^{-8}$$

por lo tanto 1 en 25,6 millones ($1/3,9 \times 10^{-8}$) de combinaciones va a tener esa secuencia

La frecuencia con que se puede esperar encontrar un oligonucleótido con esa composición en el genoma de los arenavirus es:

$$3,9 \times 10^{-8} * 3413 = 1,33 \times 10^{-4}, \text{ lo cual es equivalente a 1 vez cada 7513 genomas}$$

Asumiendo que las frecuencias de las distintas bases sean las mismas, la expectativa promedio de encontrar un oligonucleótido con esa composición en un genoma de $\sim 10^9$ bp es:

$$3,9 \times 10^{-8} * 10^9 = 39 \text{ veces y, si consideramos las 2 cadenas, es: } (3,9 \times 10^{-8} * 10^9) * 2 = 78 \text{ veces}$$

Y las probabilidades de encontrarlo un número finito de veces son:

Nº de veces	Probabilidad
2	$8,8 \times 10^{-15}$
28	$1,3 \times 10^{-2}$
36	$5,9 \times 10^{-2}$
39	$6,4 \times 10^{-2}$
100	$1,6 \times 10^{-16}$
370	$3,4 \times 10^{-220}$

Fundamentos matemáticos

- Siendo:
- N** = longitud del genoma
 - μ** = longitud del fragmento
 - p** = probabilidad de tener el fragmento en una posición dada
 - q = 1-p**
 - X (ó k)** = número de veces que aparece el fragmento de longitud μ en el genoma de longitud **N**
 - N- μ +1** = número de posiciones potenciales

Entonces **X** sigue una distribución binomial: **X ~ Bi(N- μ +1;p)**

$$p(X = k) = \binom{N - \mu + 1}{k} p^k \times q^{N - \mu + 1 - k}$$

$$p(X = 0) = q^{N - \mu + 1}$$

si

$$p(X \geq 1) = 1 - p(x = 0) = 1 - q^{N - \mu + 1} = 1 - (1 - p)^{N - \mu + 1} \approx N - \mu + 1 \times p \approx N \times p$$

Taylor de 1^{er} orden
 F(x) = término de orden 0 + término de orden 1 + resto
 El máximo error posible se determina acotando el resto

Máximo error cometido $\approx \frac{(N \times p)^2}{2}$

Sí $p = 3,9 \times 10^{-8}$

N	Límite máximo al error cometido por usar N x p
3400	$\sim 10^8$ (la aproximación sirve)
10^9	760 (la aproximación no sirve)

$N \gg \mu + 1$

Utilizando la fórmula de Stirling

$$p(X = k) \approx \frac{(N - \mu + 1)^{N - \mu + 1} \times e^{-k} \times p^k \times q^{N - \mu + 1 - k}}{N - \mu + 1 - k^{N - \mu + 1 - k} \times k!} =$$

$$= \left(1 + \frac{k}{N - \mu + 1 - k}\right)^{N - \mu + 1 - k} \times \frac{(N - \mu + 1)^k \times e^{-k} \times p^k \times q^{N - \mu + 1 - k}}{k!} \approx \frac{(N - \mu + 1)^k \times p^k \times q^{N - \mu + 1 - k}}{k!}$$

$$\left(1 + \frac{k}{N - \mu + 1 - k}\right)^{N - \mu + 1 - k} \approx e^k \quad \text{si} \quad N - \mu + 1 - k \gg k$$

En el caso de $N = 10^9$ $\mu = 12$ $p = 3,9 \times 10^{-8}$ y $k = 2$

$$p(X = 2) \approx \frac{10^{9 \times 2} \times (3,9 \times 10^{-8})^2 \times (1 - 3,9 \times 10^{-8})^{10^9}}{2} = 8,78 \times 10^{-15}$$

Una alternativa más sencilla para calcular la probabilidad de existencia del número **X** de sitios en genomas muy grandes es:

Siendo: **N** = longitud del genoma
 p = probabilidad de tener el fragmento en una posición dada
 N x p = cantidad promedio de posiciones que se espera tengan la secuencia del fragmento

Y, asumiendo que esos sitios se distribuyen aleatoriamente a lo largo de la secuencia

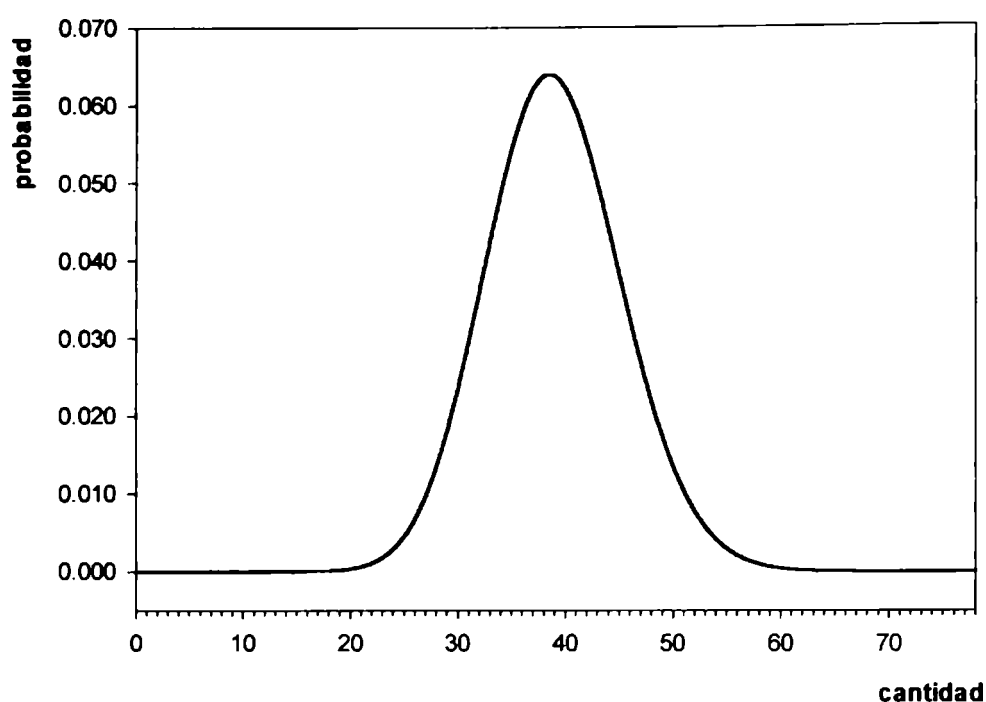
$$X \approx p(N \times p) \Rightarrow p(x = k) = \frac{(N \times p)^k \times e^{-N \times p}}{k!}$$

En el caso de:

$N = 10^9$
 $\mu = 12$
 $p = 3,9 \times 10^{-8}$
 $k = 2$

$$p(x = 2) = \frac{[10^9 \times (3,9 \times 10^{-8})]^2}{2} \times e^{-10^9 \times (3,9 \times 10^{-8})} = 8,78 \times 10^{-15}$$

Entonces, la distribución de probabilidades para distintos números finitos de veces, en un genoma de 10^9 residuos, es:



Extremo 5' del RNA S genómico descrito para el resto de los arenavirus

Secuencia: 5' CGCACCGGGG 3'

La probabilidad de encontrar un oligonucleótido de 12 residuos con esa composición en base a las frecuencias promedio de cada residuo en los arenavirus es:

$$(0,26) \cdot (0,23)^4 \cdot (0,22)^5 = 3,7 \times 10^{-7}$$

por lo tanto 1 en 2,7 millones ($1/3,7 \times 10^{-7}$) de combinaciones va a tener esa secuencia

La frecuencia con que se puede esperar encontrar un oligonucleótido con esa composición en el genoma de los arenavirus es:

$$3,7 \times 10^{-7} \cdot 3413 = 1,26 \times 10^{-3} \text{ lo cual es equivalente a 1 vez cada 792 genomas}$$

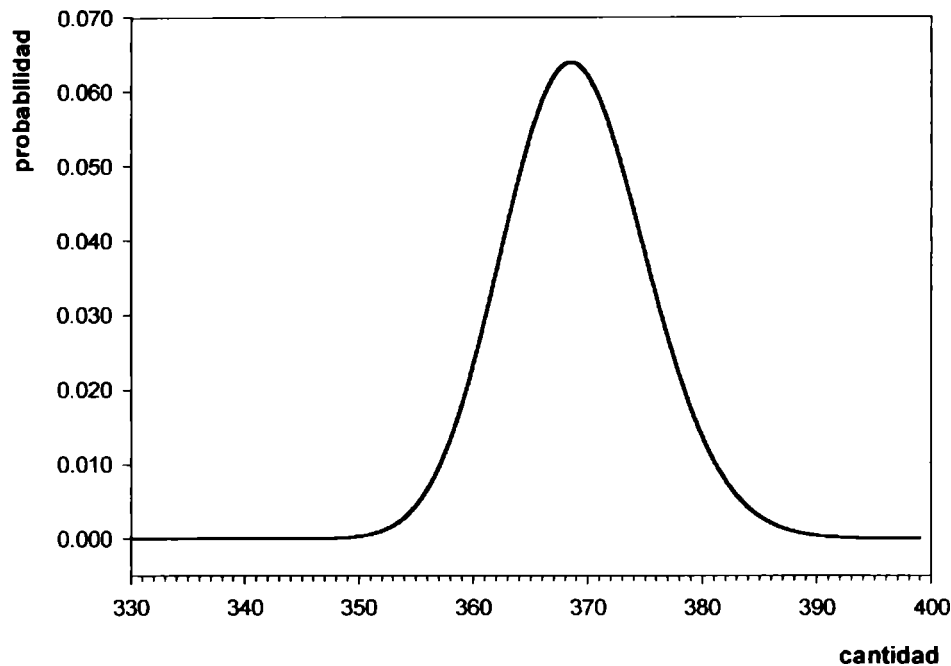
Asumiendo que las frecuencias de las distintas bases sean las mismas, la expectativa promedio de encontrar un oligonucleótido con esa composición en un genoma de $\sim 10^9$ bp es:

$$3,7 \times 10^{-7} \cdot 10^9 = 370 \text{ veces y, considerando las 2 cadenas, es: } (3,7 \times 10^{-7} \cdot 10^9) \cdot 2 = 740 \text{ veces}$$

Y las probabilidades de encontrarlo un número finito de veces son:

Nº de veces	Probabilidad
2	$1,4 \times 10^{-156}$
28	$5,4 \times 10^{-119}$
36	$1,6 \times 10^{-110}$
39	$1,2 \times 10^{-128}$
100	$1,4 \times 10^{-62}$
370	$2,1 \times 10^{-2}$

Entonces, la distribución de probabilidades para distintos números finitos de veces, en un genoma de 10^9 residuos, es:



Secuencia artificial (concatenación extremo 5' consenso de los arenavirus + extremo 5' de las cepas del virus Junín)

Secuencia: 5' CGCACCGGCUGCAGUAAG 3'

La probabilidad de encontrar un oligonucleótido de 18 residuos con esa composición en base a las frecuencias promedio de cada residuo en los arenavirus es:

$$(0,26)^4 * (0,23)^6 * (0,22)^6 * (0,29)^2 = 6,5 \times 10^{-12}$$

por lo tanto 1 en $1,6 \times 10^{11}$ ($1/6,5 \times 10^{12}$) de combinaciones va a tener esa secuencia

La frecuencia con que se puede esperar encontrar un oligonucleótido con esa composición en el genoma de los arenavirus es:

$$6,5 \times 10^{-12} * 3413 = 2,22 \times 10^{-8} \text{ lo cual es equivalente a 1 vez cada } 4,5 \times 10^7 \text{ genomas}$$

Asumiendo que las frecuencias de las distintas bases sean las mismas, la expectativa promedio de encontrar un oligonucleótido con esa composición en un genoma de $\sim 10^9$ bp es:

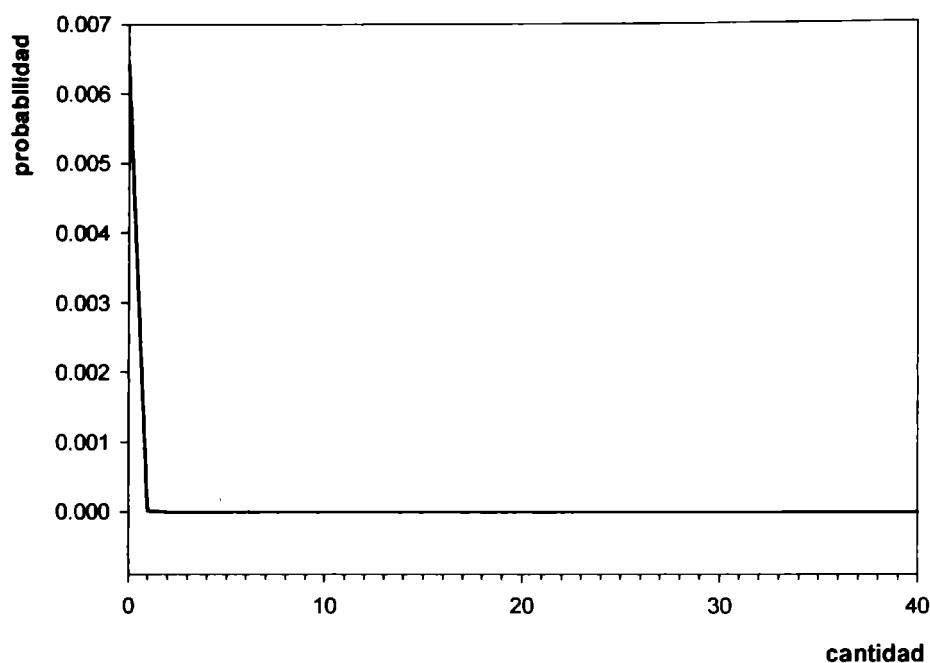
$$6,5 \times 10^{-12} * 10^9 = 6,5 \times 10^{-3} \text{ veces y, considerando las 2 cadenas, es:}$$

$$(6,5 \times 10^{-12} * 10^9) * 2 = 1,3 \times 10^{-2} \text{ veces}$$

Y las probabilidades de encontrarlo un número finito de veces son:

Nº de veces	Probabilidad
1	$6,5 \times 10^{-3}$
2	$2,1 \times 10^{-5}$
5	$9,2 \times 10^{-14}$
10	$3,4 \times 10^{-29}$
40	$2,9 \times 10^{-136}$

Entonces, la distribución de probabilidades para distintos números finitos de veces, en un genoma de 10^9 residuos, es:



Secuencia artificial (concatenación extremo 3' consenso de los arenavirus + extremo 3' de las cepas del virus Junín)

Secuencia: 5' CGCACAGUGGCGCACAGUGG 3'

La probabilidad de encontrar un oligonucleótido de 20 residuos con esa composición en base a las frecuencias promedio de cada residuo en los arenavirus es:

$$(0,26)^4 * (0,23)^6 * (0,22)^8 * (0,29)^2 = 3,1 \times 10^{-13}$$

por lo tanto 1 en $3,2 \times 10^{-12}$ ($1/3,1 \times 10^{-13}$) de combinaciones va a tener esa secuencia

La frecuencia con que se puede esperar encontrar un oligonucleótido con esa composición en el genoma de los arenavirus es:

$$3,1 \times 10^{-13} * 3413 = 1,05 \times 10^{-9} \text{ lo cual es equivalente a 1 vez cada } 9,5 \times 10^8 \text{ genomas}$$

Asumiendo que las frecuencias de las distintas bases sean las mismas, la expectativa promedio de encontrar un oligonucleótido con esa composición en un genoma de $\sim 10^9$ bp es:

$$3,1 \times 10^{-13} * 10^9 = 3,1 \times 10^{-4} \text{ veces y, considerando las 2 cadenas, es:}$$

$$(3,1 \times 10^{-13} * 10^9) * 2 = 6,2 \times 10^{-4} \text{ veces}$$

Y las probabilidades de encontrarlo un número finito de veces son:

Nº de veces	Probabilidad
1	$9,6 \times 10^{-8}$
2	$4,8 \times 10^{-8}$
5	$8,0 \times 10^{-10}$
10	$2,6 \times 10^{-14}$
40	$1,2 \times 10^{-55}$

Repeticiones de secuencias de la región intergénica del RNA S de la cepa Candid #1 del virus Junín y los rRNAs 18S y 28S humanos

Las repeticiones mostradas derivan de una búsqueda de homologías con el programa BL2seq, utilizando una longitud de palabra de 7 residuos y la máxima estrictez.

En todas las tablas se indican las posiciones nucleotídicas de las repeticiones, respectivas a cada secuencia, la identidad y la probabilidad de encontrar esa repetición por azar en el rRNA correspondiente, al menos 1 vez (P1) o 5 veces (P5).

Tabla Ap11.a. CD1 RNA S (v) vs 18S rRNA Humano

CD1 RNA S (v) (3413 nt.)	18 S rRNA Humano (1871 nt.)	Identidad	Probabilidad	
			P1	P5
1493 ugugguaa 1500	152 ugugguaa 159	100,0 %	2,3x10 ⁻²	4,9x10 ⁻¹¹
1516 gaaaccaa 1523	1142 gaaaccaa 1149	100,0 %	2,0x10 ⁻²	2,9x10 ⁻¹¹
1517 aaaccaac 1524	230 aaaccaac 237	100,0 %	1,9x10 ⁻²	2,0x10 ⁻¹¹
1498 uaaguac 1504	59 uaaguac 65	100,0 %	7,0x10 ⁻²	1,9x10 ⁻⁸
1534 uagagga 1540	1814 uagagga 1820	100,0 %	9,6x10 ⁻²	1,0x10 ⁻⁷
1563 ccaccag 1569	1215 ccaccag 1221	100,0 %	0,12	3,8x10 ⁻⁷
1571 ccgggca 1577	1127 ccgggca 1133	100,0 %	0,16	2,1x10 ⁻⁶
1595 gcccccc 1601	320 gcccccc 326	100,0 %	0,17	2,8x10 ⁻⁶
1615 gccgcgg 1621	620 gccgcgg 626	100,0 %	0,20	7,7x10 ⁻⁶
1625 ggggagg 1631	504 ggggagg 510	100,0 %	0,19	5,7x10 ⁻⁶
1639 uacagug 1645	76 uacagug 82	100,0 %	9,0x10 ⁻²	7,4x10 ⁻⁸

Tabla Ap11.b. CD1 RNA S (vc) vs 18S rRNA Humano

CD1 RNA S (vc) (3413 nt.)	18 S rRNA Humano (1871 nt.)	Identidad	Probabilidad	
			P1	P5
1755 gaaggcagc 1763	467 gaaggcagc 475	100,0 %	1,0x10 ⁻²	9,4x10 ⁻¹³
1889 uguugguuu 1897	890 uguugguuu 898	100,0 %	5,1x10 ⁻³	3,1x10 ⁻¹⁴
1915 uaccacauc 1923	454 uaccacauc 462	100,0 %	5,1x10 ⁻³	2,9x10 ⁻¹⁴
1829 ccgggcag 1836	1127 ccgggcag 1134	100,0 %	5,4x10 ⁻²	4,9x10 ⁻⁹
1803 gccgcgg 1809	620 gccgcgg 626	100,0 %	0,20	7,7x10 ⁻⁶
1813 ggggggc 1819	265 ggggggc 271	100,0 %	0,22	1,5x10 ⁻⁵
1866 ucuuagu 1872	1340 ucuuagu 1346	100,0 %	7,0x10 ⁻²	1,9x10 ⁻⁸
1870 agugucc 1876	776 agugucc 782	100,0 %	0,11	2,0x10 ⁻⁷
1878 cuacgcc 1884	1559 cuacgcc 1565	100,0 %	0,12	3,8x10 ⁻⁷

Tabla Ap11.c. CD1 RNA S (v) vs 28S rRNA Humano

CD1 RNA S (v) (3413 nt.)	28 S rRNA Humano (5035 nt.)	Identidad	Probabilidad	
			P1	P5
1603 guccgcgcc 1612	3998 guccgcgcc 4007	100,0 %	$4,7 \times 10^{-2}$	$2,4 \times 10^{-9}$
1600 ccaguccgc 1608	1428 ccaguccgc 1436	100,0 %	$5,7 \times 10^{-2}$	$6,8 \times 10^{-9}$
1513 aaagaaac 1520	60 aaagaaac 67	100,0 %	$9,9 \times 10^{-3}$	$8,4 \times 10^{-13}$
1560 uccccacc 1567	3020 uccccacc 3027	100,0 %	0,13	$6,6 \times 10^{-7}$
1566 ccagcccg 1573	213 ccagcccg 220	100,0 %	0,27	$6,3 \times 10^{-5}$
1583 cgggcugg 1590	2896 cgggcugg 2903	100,0 %	0,30	$1,5 \times 10^{-4}$
1595 gccccca 1602	3310 gccccca 3317	100,0 %	0,26	$1,2 \times 10^{-5}$
1601 caguccgc 1608	442 caguccgc 449	100,0 %	0,15	$1,5 \times 10^{-6}$
1614 ggccgcg 1621	3498 ggccgcg 3505	100,0 %	0,37	$4,4 \times 10^{-3}$
	3536 ggccgcg 3543	100,0 %	0,37	$4,4 \times 10^{-3}$
1615 gccgcgga 1622	4706 gccgcgga 4713	100,0 %	0,30	$1,4 \times 10^{-4}$
1616 ccgcggac 1623	2244 ccgcggac 2251	100,0 %	0,29	$9,4 \times 10^{-5}$
1617 cgcggaacu 1624	3213 cgcggaacu 3220	100,0 %	0,17	$2,3 \times 10^{-6}$
1618 gcggacug 1625	1158 gcggacug 1165	100,0 %	0,18	$3,5 \times 10^{-6}$
1621 gacugggg 1628	4161 gacugggg 4168	100,0 %	0,19	$5,3 \times 10^{-6}$
1656 cuucggga 1663	2859 cuucggga 2866	100,0 %	$8,4 \times 10^{-2}$	$5,0 \times 10^{-8}$
1658 ucgggagg 1665	814 ucgggagg 821	100,0 %	0,19	$5,3 \times 10^{-6}$
1494 gugguaa 1500	303 gugguaa 309	100,0 %	0,12	$3,3 \times 10^{-7}$
1495 ugguaag 1501	1850 ugguaag 1856	100,0 %	0,12	$3,3 \times 10^{-7}$
1499 aaguacc 1505	353 aaguacc 359	100,0 %	0,11	$1,9 \times 10^{-7}$
1515 agaaacc 1521	4223 agaaacc 4229	100,0 %	0,11	$2,7 \times 10^{-7}$
1516 gaaacca 1522	3838 gaaacca 3844	100,0 %	0,11	$2,7 \times 10^{-7}$
1529 uggcgua 1535	149 uggcgua 155	100,0 %	0,20	$7,1 \times 10^{-7}$
1536 gaggaca 1542	4217 gaggaca 4223	100,0 %	0,23	$2,0 \times 10^{-5}$
1551 uccugag 1557	2431 uccugag 2437	100,0 %	0,20	$7,1 \times 10^{-6}$
1558 ggucccc 1564	3168 ggucccc 3174	100,0 %	0,36	$5,5 \times 10^{-3}$
	3509 ggucccc 3515	100,0 %	0,36	$5,5 \times 10^{-3}$
	4737 ggucccc 4743	100,0 %	0,36	$5,5 \times 10^{-3}$
1559 gucccca 1565	1165 gucccca 1171	100,0 %	0,31	$1,8 \times 10^{-4}$
1561 ccccacc 1567	3313 ccccacc 3319	100,0 %	0,37	$3,6 \times 10^{-3}$
	3319 ccccacc 3325	100,0 %	0,37	$3,6 \times 10^{-3}$
	2993 ccccacc 2999	100,0 %	0,37	$3,6 \times 10^{-3}$
1568 agcccgg 1574	1792 agcccgg 1798	100,0 %	0,35	$9,7 \times 10^{-3}$
	3289 agcccgg 3295	100,0 %	0,35	$9,7 \times 10^{-3}$
1569 gcccggg 1575	3986 gcccggg 3992	100,0 %	0,15	0,10
	2309 gcccggg 2315	100,0 %	0,15	0,10
	254 gcccggg 260	100,0 %	0,15	0,10
	1200 gcccggg 1206	100,0 %	0,15	0,10
	2961 gcccggg 2967	100,0 %	0,15	0,10
1570 cccgggc 1576	1178 cccgggc 1184	100,0 %	0,17	$8,6 \times 10^{-2}$
1580 gcccggg 1586	1200 gcccggg 1206	100,0 %	0,15	0,10
	254 gcccggg 260	100,0 %	0,15	0,10
	2309 gcccggg 2315	100,0 %	0,15	0,10
	2961 gcccggg 2967	100,0 %	0,15	0,10
	3986 gcccggg 3992	100,0 %	0,15	0,10
1581 cccgggc 1587	1178 cccgggc 1184	100,0 %	0,17	$8,6 \times 10^{-2}$

Apéndice 11

(continuación)

1584 gggcugg 1590	2915 gggcugg 2921	100,0 %	0,33	$1,4 \times 10^{-2}$
	2883 gggcugg 2889	100,0 %	0,33	$1,4 \times 10^{-2}$
1588 uggugug 1594	2124 uggugug 2130	100,0 %	0,21	$1,2 \times 10^{-5}$
1589 ggugugg 1595	2163 ggugugg 2169	100,0 %	0,34	$6,3 \times 10^{-4}$
1594 ggcccc 1600	3952 ggcccc 3958	100,0 %	0,20	$7,0 \times 10^{-2}$
1595 gcccccc 1601	3440 gcccccc 3446	100,0 %	0,23	$5,6 \times 10^{-2}$
	2950 gcccccc 2956	100,0 %	0,23	$5,6 \times 10^{-2}$
	509 gcccccc 515	100,0 %	0,23	$5,6 \times 10^{-2}$
	3962 gcccccc 3968	100,0 %	0,23	$5,6 \times 10^{-2}$
	905 gcccccc 911	100,0 %	0,23	$5,6 \times 10^{-2}$
	4766 gcccccc 4772	100,0 %	0,23	$5,6 \times 10^{-2}$
	791 gcccccc 797	100,0 %	0,23	$5,6 \times 10^{-2}$
	2182 gcccccc 2188	100,0 %	0,23	$5,6 \times 10^{-2}$
1596 cccccca 1602	2953 cccccca 2959	100,0 %	0,37	$3,6 \times 10^{-3}$
	2188 cccccca 2194	100,0 %	0,37	$3,6 \times 10^{-3}$
1598 ccccagu 1604	1167 ccccagu 1173	100,0 %	0,31	$1,8 \times 10^{-4}$
1603 guccgcg 1609	435 guccgcg 441	100,0 %	0,36	$7,6 \times 10^{-3}$
1605 ccgcggc 1611	3646 ccgcggc 3652	100,0 %	0,17	$8,6 \times 10^{-2}$
	684 ccgcggc 690	100,0 %	0,17	$8,6 \times 10^{-2}$
	2929 ccgcggc 2935	100,0 %	0,17	$8,6 \times 10^{-2}$
1606 cgcggcc 1612	3272 cgcggcc 3278	100,0 %	0,17	$8,6 \times 10^{-2}$
	2978 cgcggcc 2984	100,0 %	0,17	$8,6 \times 10^{-2}$
	4865 cgcggcc 4871	100,0 %	0,17	$8,6 \times 10^{-2}$
1614 ggccgcg 1620	3131 ggccgcg 3137	100,0 %	0,15	0,10
1615 gccgcgg 1621	4787 gccgcgg 4793	100,0 %	0,15	0,10
	2928 gccgcgg 2934	100,0 %	0,15	0,10
1624 uggggag 1630	4151 uggggag 4157	100,0 %	0,35	$8,4 \times 10^{-4}$
1625 ggggagg 1631	1474 ggggagg 1480	100,0 %	0,30	$2,4 \times 10^{-2}$
	1205 ggggagg 1211	100,0 %	0,30	$2,4 \times 10^{-2}$
1626 gggaggc 1632	3948 gggaggc 3954	100,0 %	0,32	$1,8 \times 10^{-2}$
1640 acagugc 1646	4140 acagugc 4146	100,0 %	0,21	$9,6 \times 10^{-6}$

Tabla Ap11.d. CD1 RNA S (vc) vs 28S rRNA Humano

CD1 RNA S (vc) (3413 nt.)	28 S rRNA Humano (5035 nt.)	Identidad	Probabilidad	
			P1	P5
1782 gccuccccaguccgc 1796	1422 gccucuccaguccgc 1436	93,3 %	$1,8 \times 10^{-5}$	$1,5 \times 10^{-26}$
1791 guccgcggcc 1800	3998 guccgcggcc 4007	100,0 %	$4,7 \times 10^{-2}$	$2,4 \times 10^{-9}$
1812 uggggggcc 1820	184 uggggggcc 192	100,0 %	0,15	$1,3 \times 10^{-6}$
1855 ccucagga 1863	1661 ccucagga 1669	100,0 %	$2,9 \times 10^{-2}$	$1,9 \times 10^{-10}$
1773 guaagcag 1780	1852 guaagcag 1859	100,0 %	$4,5 \times 10^{-2}$	$1,9 \times 10^{-9}$
1783 ccucccca 1790	3018 ccucccca 3025	100,0 %	0,13	$6,6 \times 10^{-7}$
1785 ucccaggu 1792	1166 ucccaggu 1173	100,0 %	$7,2 \times 10^{-2}$	$2,1 \times 10^{-8}$
1789 caguccgc 1796	442 caguccgc 449	100,0 %	0,15	$1,5 \times 10^{-6}$
1802 ggccgcgg 1809	3498 ggccgcgg 3505	100,0 %	0,37	$4,4 \times 10^{-3}$
	3536 ggccgcgg 3543	100,0 %	0,37	$4,4 \times 10^{-3}$
1803 gccgcgga 1810	4706 gccgcgga 4713	100,0 %	0,30	$1,4 \times 10^{-4}$
1804 ccgcggac 1811	2244 ccgcggac 2251	100,0 %	0,29	$9,4 \times 10^{-5}$
1805 cgcggacu 1812	3213 cgcggacu 3220	100,0 %	0,17	$2,3 \times 10^{-6}$
1806 gcggacug 1813	1158 gcggacug 1165	100,0 %	0,18	$3,5 \times 10^{-6}$
1809 gacugggg 1816	4161 gacugggg 4168	100,0 %	0,19	$5,3 \times 10^{-6}$
1813 gggggggcc 1820	747 gggggggcc 754	100,0 %	0,35	$8,4 \times 10^{-4}$
	3197 gggggggcc 3204	100,0 %	0,35	$8,4 \times 10^{-4}$
	3179 gggggggcc 3186	100,0 %	0,35	$8,4 \times 10^{-4}$
	1052 gggggggcc 1059	100,0 %	0,35	$8,4 \times 10^{-4}$
1815 gggggccac 1822	1219 gggggccac 1226	100,0 %	0,30	$1,4 \times 10^{-4}$
1824 ccagcccg 1831	213 ccagcccg 220	100,0 %	0,27	$6,3 \times 10^{-5}$
1841 cgggcugg 1848	2896 cgggcugg 2903	100,0 %	0,30	$1,5 \times 10^{-4}$
1847 ggugggga 1854	4149 ggugggga 4156	100,0 %	0,20	$8,1 \times 10^{-6}$
1850 ggggaccc 1857	3189 ggggaccc 3196	100,0 %	0,30	$1,4 \times 10^{-4}$
1859 caggaggu 1866	4335 caggaggu 4342	100,0 %	$9,6 \times 10^{-2}$	$1,1 \times 10^{-7}$
1862 gaggucuu 1869	1737 gaggucuu 1744	100,0 %	$4,0 \times 10^{-2}$	$1,0 \times 10^{-9}$
1889 uguugguu 1896	1934 uguugguu 1941	100,0 %	$1,8 \times 10^{-2}$	$1,5 \times 10^{-11}$
1914 uuaccaca 1921	4353 uuaccaca 4360	100,0 %	$1,6 \times 10^{-2}$	$1,5 \times 10^{-11}$
1774 uaagcag 1780	4331 uaagcag 4337	100,0 %	0,12	$3,0 \times 10^{-7}$
1778 cagugcc 1784	4141 cagugcc 4147	100,0 %	0,32	$2,7 \times 10^{-4}$
1779 agugccu 1785	1830 agugccu 1836	100,0 %	0,20	$7,1 \times 10^{-6}$
1783 ccucccc 1789	1099 ccucccc 1105	100,0 %	0,37	$2,8 \times 10^{-3}$
	568 ccucccc 574	100,0 %	0,37	$2,8 \times 10^{-3}$
	3072 ccucccc 3078	100,0 %	0,37	$2,8 \times 10^{-3}$
	2200 ccucccc 2206	100,0 %	0,37	$2,8 \times 10^{-3}$
	2984 ccucccc 2990	100,0 %	0,37	$2,8 \times 10^{-3}$
	3065 ccucccc 3071	100,0 %	0,37	$2,8 \times 10^{-3}$
777 ccucccc 783	100,0 %	0,37	$2,8 \times 10^{-3}$	
1791 guccgcg 1797	435 guccgcg 441	100,0 %	0,36	$7,6 \times 10^{-3}$
1793 ccgcggc 1799	684 ccgcggc 690	100,0 %	0,17	$8,6 \times 10^{-2}$
	2929 ccgcggc 2935	100,0 %	0,17	$8,6 \times 10^{-2}$
	3646 ccgcggc 3652	100,0 %	0,17	$8,6 \times 10^{-2}$
1794 cgcggcc 1800	2978 cgcggcc 2984	100,0 %	0,17	$8,6 \times 10^{-2}$
	3272 cgcggcc 3278	100,0 %	0,17	$8,6 \times 10^{-2}$
	4865 cgcggcc 4871	100,0 %	0,17	$8,6 \times 10^{-2}$
1795 gcggcca 1801	4378 gcggcca 4384	100,0 %	0,35	$9,7 \times 10^{-3}$

Apéndice 11

(continuación)

1802 ggcgcg 1808	3131 ggcgcg 3137	100,0 %	0,15	0,10
1803 gccgcgg 1809	2928 gccgcgg 2934	100,0 %	0,15	0,10
	4787 gccgcgg 4793	100,0 %	0,15	0,10
1811 cuggggg 1817	3352 cuggggg 3358	100,0 %	0,33	1,4x10 ⁻²
1812 ugggggg 1818	1050 ugggggg 1056	100,0 %	0,32	1,9x10 ⁻²
1813 ggggggc 1819	4068 ggggggc 4074	100,0 %	9,7x10 ⁻²	0,14
	1148 ggggggc 1154	100,0 %	9,7x10 ⁻²	0,14
1816 gggccac 1822	1060 gggccac 1066	100,0 %	0,35	9,7x10 ⁻³
	1840 gggccac 1846	100,0 %	0,35	9,7x10 ⁻³
1826 agcccgg 1832	3289 agcccgg 3295	100,0 %	0,35	9,7x10 ⁻³
	1792 agcccgg 1798	100,0 %	0,35	9,7x10 ⁻³
1827 gcccggg 1833	1200 gcccggg 1206	100,0 %	0,15	0,10
	2309 gcccggg 2315	100,0 %	0,15	0,10
	3986 gcccggg 3992	100,0 %	0,15	0,10
	2961 gcccggg 2967	100,0 %	0,15	0,10
	254 gcccggg 260	100,0 %	0,15	0,10
1828 cccgggc 1834	1178 cccgggc 1184	100,0 %	0,17	8,6x10 ⁻²
1834 cagugcc 1840	4141 cagugcc 4147	100,0 %	0,32	2,7x10 ⁻⁴
1835 agugccc 1841	4839 agugccc 4845	100,0 %	0,32	2,7x10 ⁻⁴
1838 gcccggg 1844	2961 gcccggg 2967	100,0 %	0,15	0,10
	3986 gcccggg 3992	100,0 %	0,15	0,10
	1200 gcccggg 1206	100,0 %	0,15	0,10
	2309 gcccggg 2315	100,0 %	0,15	0,10
	254 gcccggg 260	100,0 %	0,15	0,10
1839 cccgggc 1845	1178 cccgggc 1184	100,0 %	0,17	8,6x10 ⁻²
1842 gggcugg 1848	2883 gggcugg 2889	100,0 %	0,33	1,4x10 ⁻²
	2915 gggcugg 2921	100,0 %	0,33	1,4x10 ⁻²
1845 cuggugg 1851	1555 cuggugg 1561	100,0 %	0,34	4,3x10 ⁻⁴
1847 ggugggg 1853	623 ggugggg 629	100,0 %	0,32	1,9x10 ⁻²
	429 ggugggg 435	100,0 %	0,32	1,9x10 ⁻²
1848 gugggga 1854	4914 gugggga 4920	100,0 %	0,35	8,4x10 ⁻⁴
1850 ggggacc 1856	636 ggggacc 642	100,0 %	0,34	1,3x10 ⁻²
1855 ccucag 1861	79 ccucag 85	100,0 %	0,31	1,8x10 ⁻⁴
1861 ggagguc 1867	1560 ggagguc 1566	100,0 %	0,34	5,7x10 ⁻⁴
1871 guguccu 1877	4196 guguccu 4202	100,0 %	0,19	5,2x10 ⁻⁶
1878 cuacgcc 1884	2253 cuacgcc 2259	100,0 %	0,31	1,8x10 ⁻⁴
1895 uuucuuu 1901	2618 uuucuuu 2624	100,0 %	1,9x10 ⁻²	2,0x10 ⁻¹¹

