

Tesis Doctoral

Secuencia nucleotídica completa del Adenovirus 7h: análisis de los polimorfismos del gen del hexón de los Adenovirus circulantes en Buenos Aires entre los años 1990-2002

Barrero, Paola Roxana

2004

Tesis presentada para obtener el grado de Doctor en Ciencias
Biológicas de la Universidad de Buenos Aires

Este documento forma parte de la colección de tesis doctorales y de maestría de la Biblioteca Central Dr. Luis Federico Leloir, disponible en digital.bl.fcen.uba.ar. Su utilización debe ser acompañada por la cita bibliográfica con reconocimiento de la fuente.

This document is part of the Master's and Doctoral Theses Collection of the Central Library Dr. Luis Federico Leloir, available in digital.bl.fcen.uba.ar. It should be used accompanied by the corresponding citation acknowledging the source.

Cita tipo APA:

Barrero, Paola Roxana. (2004). Secuencia nucleotídica completa del Adenovirus 7h: análisis de los polimorfismos del gen del hexón de los Adenovirus circulantes en Buenos Aires entre los años 1990-2002. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires.
http://hdl.handle.net/20.500.12110/tesis_n3717_Barrero

Cita tipo Chicago:

Barrero, Paola Roxana. "Secuencia nucleotídica completa del Adenovirus 7h: análisis de los polimorfismos del gen del hexón de los Adenovirus circulantes en Buenos Aires entre los años 1990-2002". Tesis de Doctor. Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires. 2004. http://hdl.handle.net/20.500.12110/tesis_n3717_Barrero

EXACTAS UBA

Facultad de Ciencias Exactas y Naturales



UBA

Universidad de Buenos Aires

FCE y N BIBLIOTECA

**Secuencia nucleotídica completa del Adenovirus 7h: análisis de los
polimorfismos del gen del hexón de s Adenovirus circulantes en
Buenos Aires entre os 1990-2002.**

**Tesista: Paola Roxa a Barrero
Licenciada en Ciencias Biológicas**

Director: Dra. Alicia Susana Mistchenko

Laboratorio de Virología, Hospital de Niños Dr. R. Gutiérrez

71420

VI

№3717

Indice		
Sección	Subsección	Páginas
Resumen		2
Abstract		3
Introducción		4-45
	Filogenia y taxonomía	4-9
	Estructura del virión	10-12
	Ciclo replicativo	12-16
	Expresión génica	17-36
	Interacción con el huésped	37
	Mecanismos de defensa	38-40
	Epidemiología	41-42
	Patogénesis	42-44
	Naturaleza del aporte	45
Materiales y Métodos		46-59
	Muestras y datos epidem.	47
	Datos meteorológicos	48
	Cultivo y purificación de DNA	48-49
	Obtención de biblioteca gen.	49-51
	Epidemiología molecular	51
	Genotipificación AV7	51-52
	Análisis de polimorfismos....	52-53
	Diseño de primers específicos	53-54
	Amplificación y secuenciación	55-57
	Análisis de secuencias....	58-59
Resultados		60-112
	Aspectos clínicos y epidemiol...	61-65
	Datos ambientales	66-67
	Datos moleculares: hexón AV7h	68-78
	Polimorfismos gen hexón	79-89
	Polimorfismos HVRs gen hexón	90-94
	Inferencia filogenética	95-96
	Secuenciación genoma AV7h	97-112
Conclusiones		113
Bibliografía		114-118
Addendum		119-135
	Secuencia genoma AV7h	119-126
	Polimorfismos y epidemiología...	127-133
	Primers diseñados	133-135

Resumen

Los Adenovirus circularon en el período 1999-2002 durante todo el año, registrándose un mayor número de casos en el período invernal de baja temperatura y primaveral con cambios de temperatura frecuentes.

El diagnóstico virológico se realizó tempranamente luego de la internación siendo la bronquiolitis el diagnóstico clínico al ingreso más frecuente. Se registró un mayor número de casos en menores de 12 meses y la evolución en general fue favorable registrándose 9 casos fatales en un total de 362 muestras analizadas.

Se encontró una correlación negativa de grado medio con respecto a la temperatura, indicando un aumento en el número de casos cuando el registro térmico desciende. En cuanto a la humedad relativa y el índice UVB, la relación no es tan clara aunque es siempre de grado menor. Este parece ser un parámetro no necesariamente asociado a la aparición de casos de Adenovirus.

El análisis filogenético del hexón completo permitió agrupar las muestras fatales y las separó de las graves y leves en el período 1990-1999. El análisis de las regiones hipervariables en el período 1990-2002 posicionó a las muestras analizadas en el *cluster* GTC2 junto al AV7h.

La secuencia completa del AV7h se llevó a cabo por una estrategia de ensamblado de 8 contigs que permitieron elaborar una secuencia consenso de 35259 pb. con un porcentaje G+C del 51%. Se conservaron intactos los genes para las regiones tempranas, intermedias y tardías. El análisis filogenético ubicó al genoma AV7h junto con el AVB con el que tiene una homología del 84%. Se hallaron ORFs relacionados con proteínas pulmonares que podrían inducir autoanticuerpos o daño pulmonar crónico postviral.

Abstract

Adenovirus 7h Sequencing: Hexon Gene Polymorphism Analysis In A Twelve-Year Period In Buenos Aires, Argentina.

In Argentina, Adenovirus 7h (AV7h) has been frequently associated with fatal cases and chronic lung disease. In to order to understand the characteristics of AV7h we accomplished the entire reference sequencing and analysed sample polymorphism in the hexon gene from a twelve-year period (1990-2002). Nasopharyngeal aspirates from children with acute lower respiratory infections were were characterized by multiplex-PCR. Primers were designed and PCRs were performed to obtain the full genome for the reference 7h strain 87-922 and the entire hexon gene or hipervariable regions for samples. Amplicons were sequenced, contigs were aligned, a consensus full-length sequence was obtained for the reference strain 87-922 and phylogenetic relationships for hexon gene were inferred by maximum-likelihood criteria. Overall differences between AV7h hexon gene and prototype for serotype 7 (Gomen) sequence were <4% at nucleotide level and <3.5% at amino acid level. Phylogenetic analysis clustered together fatal cases apart from severe and mild cases. The entire genome was completed in 8 contigs with 44, 16, 7, 15, 14, 26, 18 and 15 primer reactions respectively. AV7h was 35263 bp in length and had a GC content of 51% and genes from the early, intermediate, and late regions were present in the expected human adenovirus locations. The genome differs from prototype strain for AVB (subspecie B2) in 16%, and 4% from vaccinal strain AV7(subspecie B1). This work constitutes the first attempt to elucidate intragenotypic variations within AV7h and to reveal particular features of this strain.

Introducción

Filogenia y taxonomía

Los miembros de la familia *Adenoviridae* son virus icosaédricos, sin envoltura que se replican en el núcleo. Su DNA lineal de doble cadena contiene 26-45 kpb y se clasifica como de tamaño medio dentro de los DNA virus. El genoma se caracteriza por llevar terminales extremos repetitivos invertidos (ITRs) y los extremos 5' se unen a una proteína terminal (TP). Los Adenovirus infectan todas las clases de Vertebrados, peces, anfibios, reptiles, aves y mamíferos (Davison *et al.*, 2003). Las relaciones filogenéticas entre los distintos Adenovirus pueden resumirse en la Figura 1.

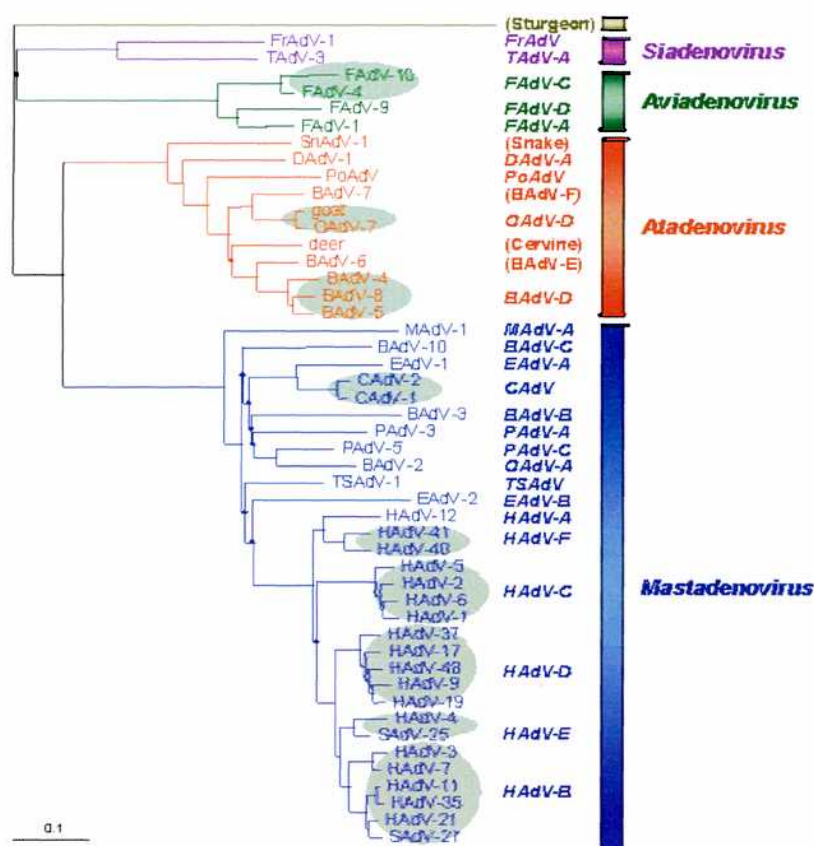


Figura 1: Filogenia de la proteína hexón de la familia *Adenoviridae*. Arbol de distancia con raíz en el esturión. Los géneros se marcan en diferentes colores y se destacan en verde los virus que pertenecen a la misma especie. B: bovino, C: canino, D: pato, E: equino, F: aviar, Fr: rana, H: humano, M: murino, O: ovino, P: porcino, Po: zarigüeyas, Sn: serpiente, T: pavo, TS: musaraña. Tomado de Davison *et al.*, 2003

Actualmente el ICTV reconoce 4 géneros de Adenovirus que coinciden con los cuatro clados principales, Siadenovirus (anfibios), Atadenovirus (reptiles), Aviadenovirus (aves) y Mastadenovirus (mamíferos) y se propone un nuevo quinto género Ichtadenovirus para Peces (Benkö y Harrach, 2003, Büchen-Osmond, 2003; Harrach y Benkö, 1998). Aviadenovirus y Mastadenovirus se originaron en Aves y Mamíferos respectivamente y los otros dos géneros tienen un rango más amplio de huéspedes. Los Atadenovirus se caracterizan por su alto contenido en A+T e infecta rumiantes, aves, reptiles y marsupiales. Se conocen dos miembros del género Siadenovirus: uno de aves y otro de rana (Davison *et al.*, 2000). Dentro de cada género, los virus se clasifican en especies (Tabla 1).

Familia	Género	Especie Tipo
<u>Adenoviridae</u>	<u>Mastadenovirus</u>	<u>Adenovirus humano C</u>
	<u>Aviadenovirus</u>	<u>Adenovirus aviar A</u>
	<u>Atadenovirus</u>	<u>Adenovirus ovino D</u>
	<u>Siadenovirus</u>	<u>Adenovirus de pavo B</u>

Tabla 1: Clasificación de la familia Adenoviridae

Se ha llegado a la conclusión que los Adenovirus coevolucionaron y coespeciaron con sus hospedadores por similitud entre relaciones filogenéticas inferidas en base a la proteasa viral y a la subunidad pequeña de RNAr mitocondrial de los hospedadores (Benkö y Harrach, 2003). A su vez, se postula que existió un ancestro común previo al origen de los eucariotas compartido entre Adenovirus y bacteriófagos del tipo de PRD1 por similitudes estructurales y de estrategias de replicación y organización genómica (San Martín y Burnett, 2003). Estos predecesores pueden haber adquirido características distintivas como el *splicing* y genes propios que se conservan en todos los géneros. A partir de allí, cada linaje de Adenovirus debe haber incorporado distintos grupos de genes que favorecieron su *fitness* en los diferentes nichos que ocupan.

De esta manera, se pueden diferenciar diferentes organizaciones genómicas que se corresponden con los cuatro géneros de Adenovirus reconocidos (Figura 2). Existen 16 genes comunes a todos los géneros ubicados en la región central del genoma relacionados con la replicación (pol, pTP, DBP), con la encapsidación del DNA (52K, IVa2) y con la formación y estructura del virión (pIIIa, III, pVII, pIX, pVI, hexón, proteasa, 100K, 33K, pVIII y fibra). Los genes específicos de género se ubican en las regiones terminales del genoma y se supone que se han adquirido por captura, recombinación y/o duplicación (Davison *et al.*, 2003). La estructura y función de los genes se detallará más adelante, así como los genes que son particulares a Mastadenovirus.

Se reconocen 51 serotipos de Adenovirus en humanos que se distinguen en base a su resistencia a neutralización por antisueros anti otros serotipos conocidos. La neutralización tipo específica resulta de la unión de un anticuerpo a epitopes de la proteína hexón y al nudo terminal de la proteína fibra. Regiones hipervariables en el hexón conforman *loops* serotipo específico en la superficie de la proteína (Shenk, 2001).

Los serotipos humanos se clasifican en seis especies (A-F) de acuerdo a diferentes propiedades como la oncogenicidad en roedores y la hemaglutinación. Aunque todos los serotipos humanos pueden transformar cultivos primarios de ratón, sólo los virus de las especies A y B son oncogénicos en ratón (Tabla 2).

Especie	Serotipos:	Potencial Oncogénico	Hemaglutinación	
			Rhesus	Rata
A	12, 18, 31	Alto	-	+ / -
B subespecie B1	3, 7, 16, 21, 50(*)	Débil	+	-
B subespecie B2				
C	1, 2, 5, 6	Ninguno	-	+ / -
D	8-10, 13, 15, 17, 19, 20, 22-30, 32, 33, 36-39, 42-49, 51(*)	Ninguno	+ / -	+
E	4	Ninguno	-	+ / -
F	40, 41	Desconocido	-	+ / -

Tabla 2: Clasificación de los Adenovirus humanos.
Modificado de Cann, 1997

Los Adenovirus pueden infectar y replicarse en varios sitios del tracto respiratorio así como en ojos y tracto gastrointestinal. De manera menos frecuente infectan el tracto urinario y el hígado. Ocasionalmente, pueden infectar y causar enfermedad en páncreas, miocardio, o sistema central nervioso, llevando a meningoencefalitis como complicación de una infección respiratoria. Sólo un tercio de los serotipos descritos están asociados con enfermedad, siendo muchas infecciones subclínicas que generan anticuerpos protectivos ante la reintroducción exógena del mismo serotipo. Los virus pueden hacer infecciones latentes en tejido linfoide y reactivarse en el tiempo (Tabla 3) (Lukashok y Horwitz, 1998).

Tabla 3: Patologías causadas por Adenovirus en humanos

Enfermedad	Individuos a riesgo	Serotipos
Faringitis aguda febril	Infantes, niños	1-3, 5-7
Fiebre faringoconjuntival	Escolares	3, 7, 14
Enfermedad respiratoria aguda	Reclutas militares	3, 4, 7, 14, 21
Neumonía	Reclutas militares,	4, 7
	Infantes, niños	1-3, 7
Queratoconjuntivitis epidémica	Todos	8, 11, 19, 37
Síndrome coqueluchoide	Infantes, niños	5
Cistitis aguda hemorrágica	Niños	11, 21
Gastroenteritis	Infantes, niños	40, 41
Meningoencefalitis	Niños e inmunocomprometidos	7, 12, 32
Hepatitis	Infantes y niños transplantados	1, 2, 5
Miocarditis	Niños	?
<i>Persistencia</i>	En inmunosuprimidos	
<i>En tracto urinario</i>		34, 35
<i>En colon</i>		42-49

Estructura del virión

Los Adenovirus tienen una morfología característica con cápside icosaédrica compuesta por tres proteínas principales: hexón (II), base pentón (III) y la fibra (IV), junto con otras proteínas VI, VIII, IX, IIIa y IVa2. El genoma viral es lineal, doble cadena, con uniones covalentes a la proteína terminal (TP) en sus extremos 5' que tienen terminales invertidos repetitivos (ITRs). El DNA viral está muy relacionado con la proteína básica VII y con un pequeño péptido *mu* (μ , X). Este complejo DNA-proteína se acopla a la proteína V, y provee un vínculo estructural con la cápside mediante la proteína VI. El virus lleva una proteasa (Pr) necesaria para procesar algunas proteínas estructurales (precursores) y producir viriones maduros infecciosos. Nota: no existe la proteína I. (Figura 3, Tabla 4) (San Martín y Burnett, 2003).

Tabla 3: Proteínas tardías de los Adenovirus. Modificado de Cann, 1997

Nombre:	Localización:	Funciones conocidas:
II	Hexón (monómero)	Estructural
III	Base Pentón	Penetración
IIIa	Asociada con base pentón	Penetración
IV	Fibra	Unión receptor; hemaglutinación
V	Core: asociada con DNA y base pentón	<i>Histone-like; packaging?</i>
VI	Polipéptido menor Hexón	Estabilización/ensamblado?
VII	Core	<i>Histone-like</i>
VIII	Polipéptido menor Hexón	Estabilización/ensamblado?
IX	Polipéptido menor Hexón	Estabilización/ensamblado?
TP	Genoma - <u>Pro</u> teína <u>T</u> erminal	Replicación Genoma

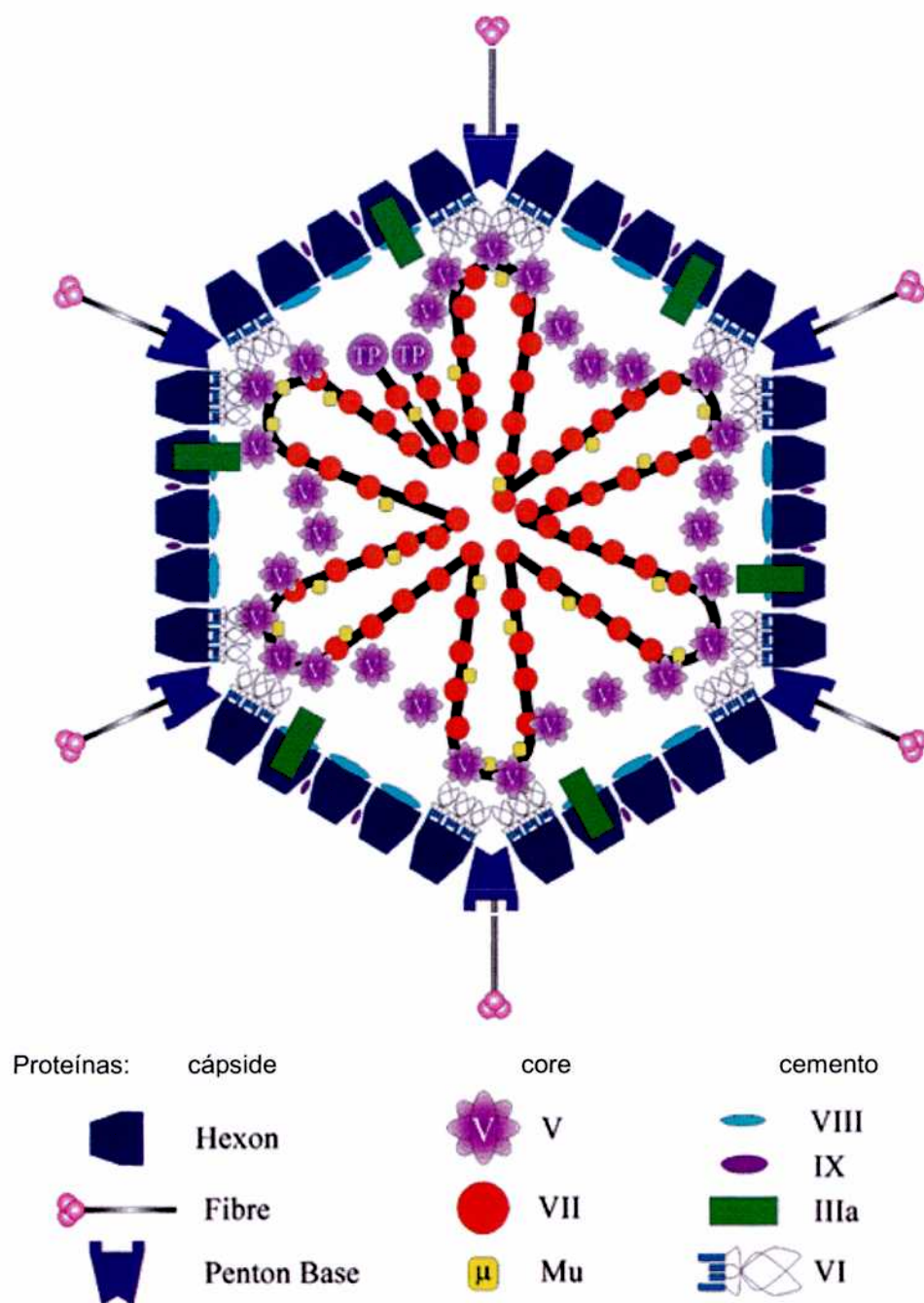


Figura 3: Estructura del virión de Adenovirus. Modificado de Russell, 2000.

La cápside está compuesta por 252 capsómeros, divididos en dos tipos principales: 240 hexones conformando las 20 caras triangulares del icosaedro y 12 pentones formando los 12 vértices. Desde la base de cada pentón protruye una fibra. La estructura puede verse en la Figura 4. Hexón, pentón y fibra constituyen los principales determinantes antigénicos del virus (Valentine y Pereira, 2003).

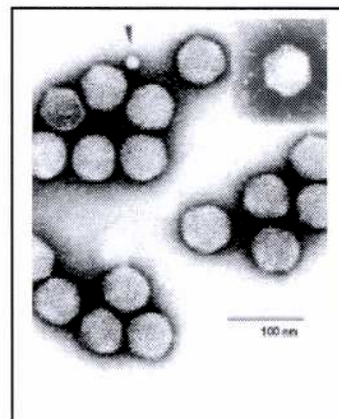
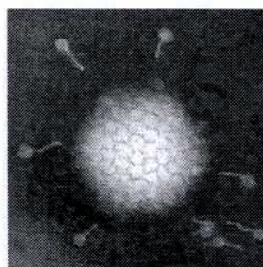


Figura 4: Estructura del virión de Adenovirus. Tomado de Valentine & Pereira, 2003.

Ciclo replicativo

El ciclo de replicación se divide por convención en dos fases separadas por el comienzo de la replicación del DNA. Los eventos tempranos comienzan con la interacción del virus con la célula e incluyen adsorción, penetración, movimiento del DNA viral al núcleo y expresión de un grupo de genes tempranos (E=early). Los productos de estos genes tempranos median la expresión viral posterior y la replicación del DNA, inducen la progresión del ciclo celular, bloquean la apoptosis y antagonizan las defensas antivirales del huésped. En células HeLa infectadas a una multiplicidad de 10 unidades

formadoras de placa por célula, la fase temprana dura 5-6 horas, luego se detecta la replicación del DNA viral. En ese momento, la fase tardía del ciclo comienza con la expresión de genes virales tardíos (L= *late*) y el ensamblado de los viriones progenie. El ciclo infeccioso se completa luego de 20-24 horas (Figura 4). Al final del ciclo se produjeron casi 10^4 partículas virales progenie por célula, junto con la síntesis excesiva de proteínas virales y DNA que no son ensamblados en viriones. Las células infectadas se dividen duplicando, al final del ciclo de replicación, el contenido de DNA y proteína (Shenk, 2001).

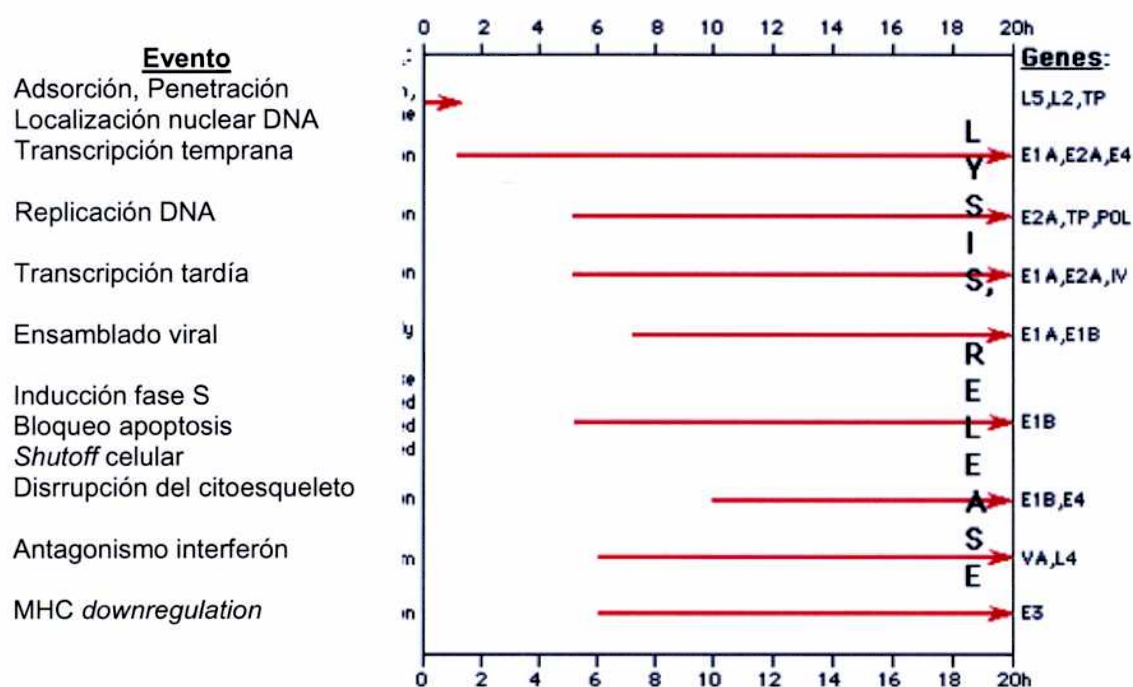
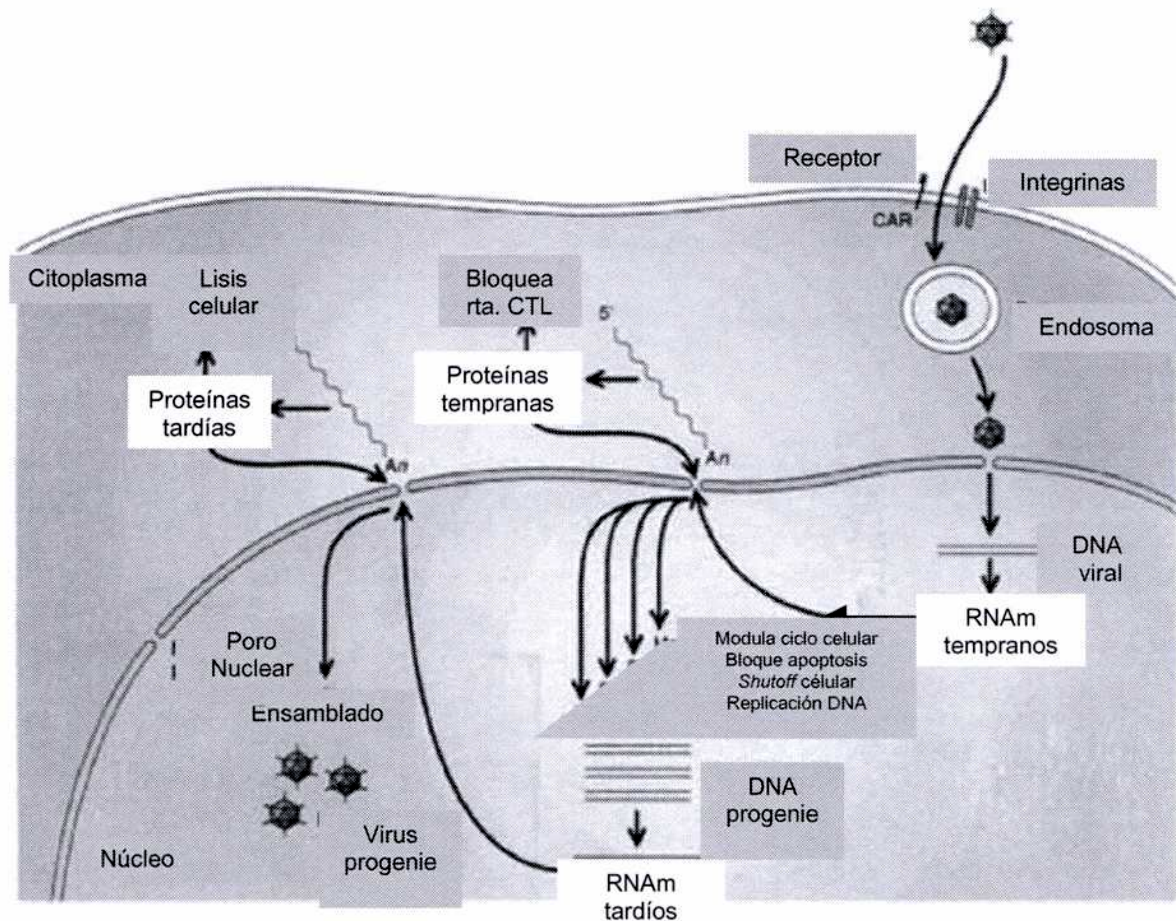


Figura 4: Ciclo replicativo de Adenovirus. Modificado de Cann, 1997.

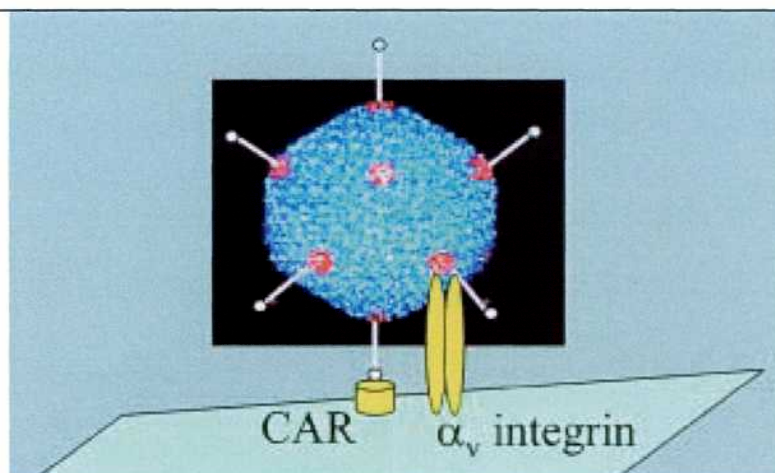
La replicación de todos los Adenovirus es similar y ocurre en el núcleo. En la Figura 5 se resume el ciclo replicativo.



La entrada de la partícula se produce en dos pasos: interacción de la fibra con el receptor y unión de la base del pentón con integrinas. Existen varios receptores celulares para Adenovirus incluyendo el receptor coxsackievirus-adenovirus CAR (especies A, C, D, E, y F, no B), la molécula MHC clase I alfa2 (especies A, B, C, D, E, y F), CD46 (especie B), ácido siálico (especie F serotipos Ad8, Ad19 y Ad37) VCAM (especie C serotipo Ad5p) y sBAR/ sB2AR (común a especie B y exclusivo para subespecie B2 respectivamente)(Boulanger, 1999; Gaggar *et al.*, 2003; Howitt *et al.*, 2003; Segerman *et al.*, 2003). La proteína pentón base se pliega entonces hacia la familia de heterodímeros de integrina de la superficie celular ($\alpha_v\beta_3$ y $\alpha_v\beta_5$) mediante una secuencia RGD (Arg-Gly-Asp) presente en cada molécula de polipéptido III, permitiendo la internalización via endocitosis mediada por receptor (Figura 6). Los complejos AV-receptor difunden en *coated pits*. El

proceso requiere dinamina, una proteína que regula específicamente la endocitosis mediada por clatrina; y requiere la activación de kinasa lipídica fosfatidilinositol-3-OH kinase quien induce la reorganización de los filamentos de actina (Nemerow y Stewart, 1999; Nemerow , 2000).

Figura 6: Entrada a la célula de los Adenovirus. Modificado de Nemerow & Stewart, 1999



Como dato interesante, los receptores para Adenovirus no se encuentran en la cara apical del epitelio respiratorio sino en las caras basolaterales (Figura 7). Este puede reflejar una presión evolutiva para bloquear la entrada por esta puerta. (Pickles *et al.*, 1998; Zhang *et al.*, 2002).

Cara apical

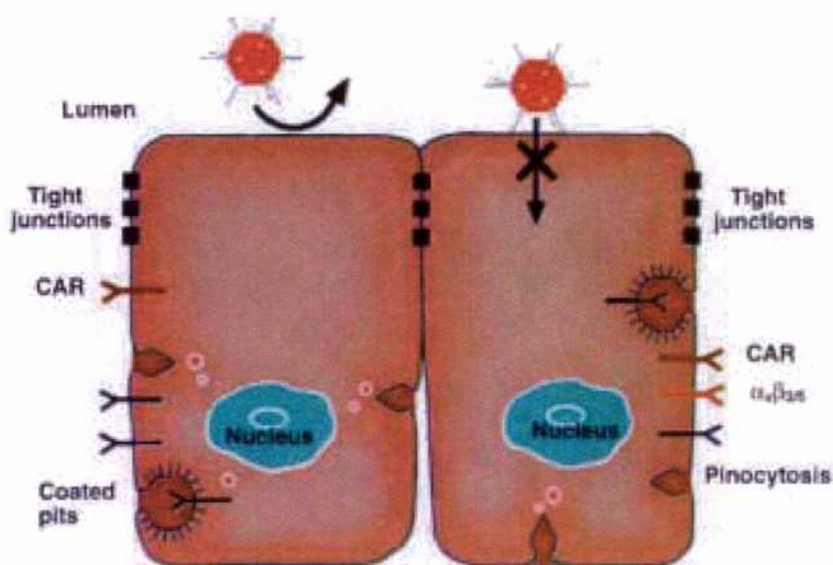
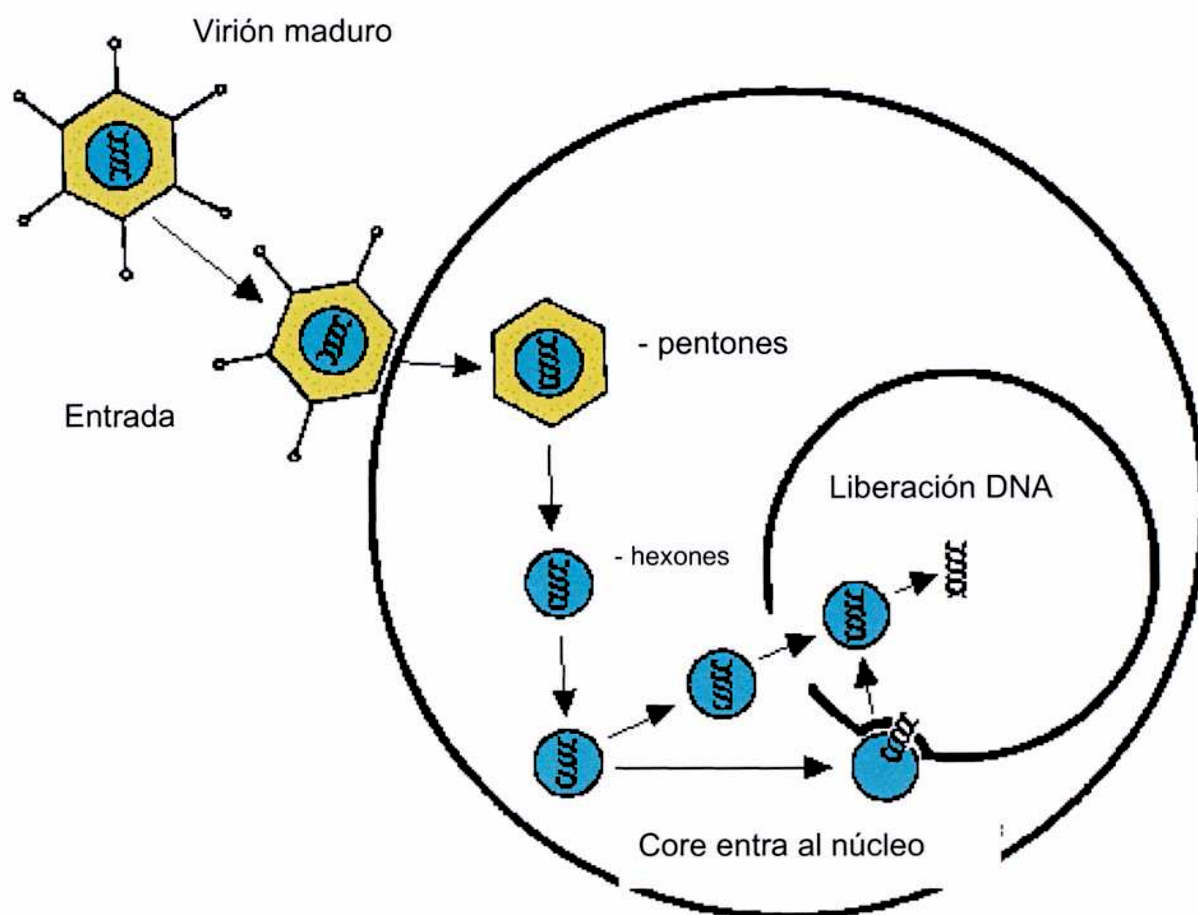


Figura 7: Ubicación de los receptores para Adenovirus. Modificado de Zhang *et al.*, 2002

La penetración involucra fagocitosis en vacuolas cuyas membranas luego son rotas por la actividad tóxica del pentón liberándose así la partícula en el citoplasma.

El *uncoating* sigue una secuencia ordenada: primero los pentones, dejando una partícula esférica parcialmente desnuda en el citoplasma. El core migra al núcleo y el DNA junto con la proteína VII entra por los poros nucleares, a partir de allí se convierte en un complejo DNA viral-histona asociándose con la matriz nuclear a través la proteína terminal unida covalentemente al DNA viral (PT) (Figura 8) (Shenk, 2001).

Figura 8: Penetración a la célula y *uncoating* de Adenovirus. Modificado de Cann, 1997.



Expresión Génica:

Se han descripto 16 genes comunes a todos los géneros que se ubican en la región central del genoma y están relacionados con la replicación (pol, pTP, DBP), con la encapsidación del DNA (52K, IVa2) y con la formación y estructura del virión (pIIIa, III, pVII, pIX, pVI, hexón, proteasa, 100K, 33K, pVIII y fibra). En particular los Mastadenovirus presentan un gen específico (V) entre pVII y pX. Los genes específicos de género se ubican en las regiones terminales del genoma E1 E3 y E4. En particular los Mastadenovirus presentan E1A, pIX, E3 (entre pVII y exón U) compartido con Siadenovirus, E4 orientado hacia la izquierda y los VARNA genes entre pTP y 52K compartidos con Aviadenovirus. Adicionalmente los ITRs aumentan de tamaño (93-369 nt).

El genoma de los Adenovirus género Mastadenovirus presenta un bloque central de genes tardíos orientados hacia la derecha desde 52K hasta la fibra interrumpidos en la misma hebra por un grupo de genes tempranos en la región E3 y en la hebra opuesta por E2, como DBP en la región E2A y pTP y DNA polimerasa en la región E2B (Figura 9). Las regiones terminales están ocupadas a la derecha por los genes E4 y a la izquierda por E1A y E1B más dos genes intermedios (IX y IVa2).

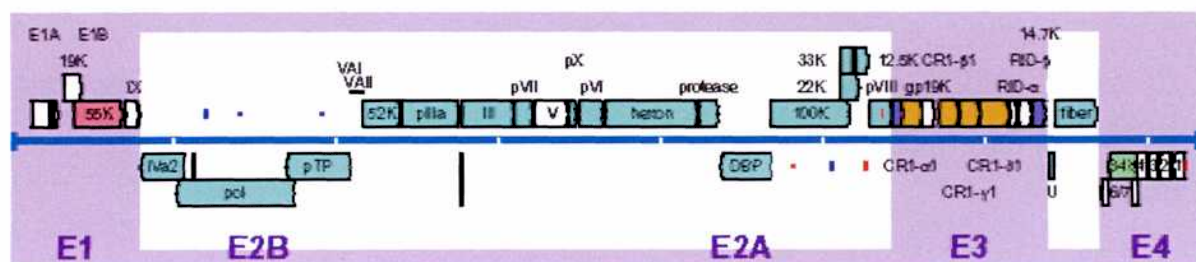


Figura 9: Genoma de Mastadenovirus. Los genes comunes se marcaron en celeste, los específicos en otro color. Modificado de Davison *et al.*, 2003.

El mecanismo de *splicing* fue descubierto en Adenovirus humanos como el proceso por el cual varios genes tardíos se expresan de un líder compartido no traducido tripartito (Berget *et al.*, 1977). Luego se vió también en genes tempranos e intermedios (E1A, DBP, pTP y pol en E2B, IVa2, E3 y E4). Este mecanismo no es universal en Adenovirus siendo en humanos E1B 19K, E1B55K y IX genes de un exón. El patrón "líder-exon" facilita la evolución de nuevos genes por captura o duplicación sólo requiriendo la presencia de un sitio aceptor de *splicing* a la izquierda de la nueva región codificante a insertar o duplicar. Un ejemplo de esto es la presencia del gen de la fibra duplicado en los Adenovirus humanos F, aunque sólo una copia sea funcional, y la presencia de una dUTPasa celular duplicada sin sitio activo en E4 ORF1 (Benkö y Harrach, 2003, Davison *et al.*, 2003).

El cromosoma viral lleva 5 unidades de transcripción temprana (E1A, E1B, E2, E3, y E4), dos intermedias (IX y IVa2), y una unidad tardía (*major late*) que es procesada para generar 5 familias de RNAm tardíos (L1 to L5), que serán transcritos por la RNA polimerasa II (Tabla 6). El cromosoma presenta también una o dos copias (dependiendo del serotipo) de genes asociados al virus (VA) que serán transcritos por la RNA polimerasa III. Por convención, el mapa se dibuja con el gen E1A hacia el extremo izquierdo. Ambas hebras del DNA viral se transcriben hacia la derecha (*rightward* o 5'-3') en el mapa convencional codificando las unidades E1A, E1B, IX, *major late*, VA RNA, y E3 y hacia la izquierda (*leftward* o 3'-5') las unidades E4, E2, y IVa2 (Shenk, 2001).

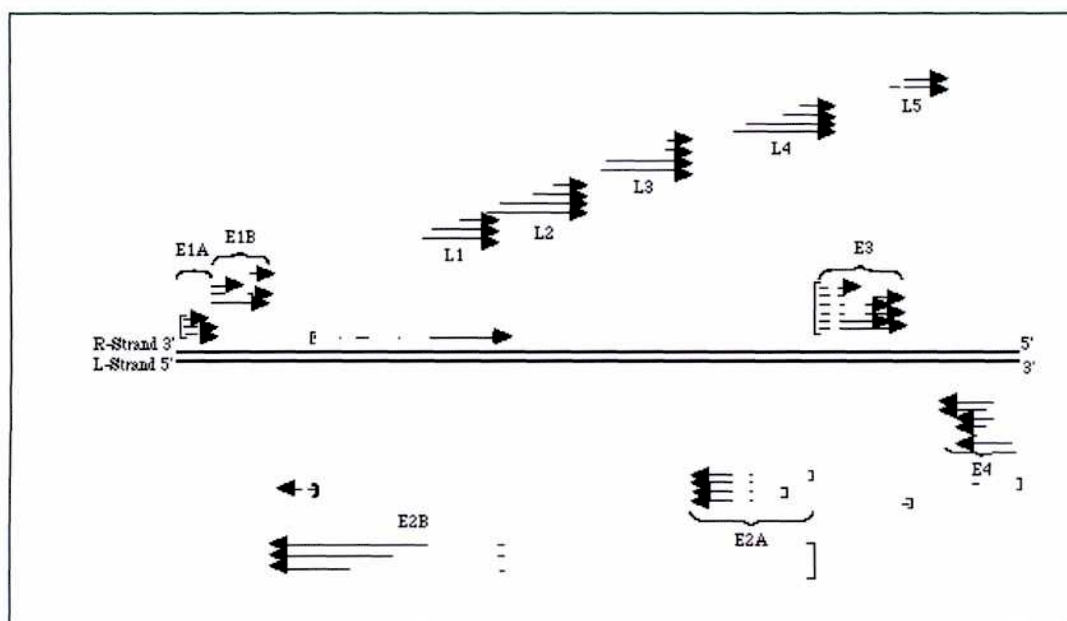
Fase:	Genes Transcritos:
Immediatamente temprana	E1A
Temprana	E1B, E2A, E2B, E3, E4
Intermedia	IX y IVa2
Tardía	Genes <i>Late</i> , proteínas del virion

Tabla 6: Expresión génica en Adenovirus. Modificado de Cann, 1997.

La organización de las unidades de transcripción en el cromosoma viral serviría como reguladores en el tiempo. Así, la transcripción del genoma está regulada por factores virales actuando en trans: los productos de los genes inmediatamente tempranos regulan la expresión de los genes tempranos y éstos a su vez de los intermedios y tardíos (Figura 5)(Cann, 1997).

Cada gen transcripto por la RNA polimerasa II da lugar a múltiples RNAm que se diferencian por *splicing* alternativo y en algunos casos por el uso de diferentes sitios poly(A) (Figura 10) (Berget, 1977).

Figura10: Mapa de transcripción de Adenovirus.
Modificado de Cann, 1997.



Cada transcripto da origen a un número diferente de proteínas, que puede o no coincidir con el número de RNAm que se generen. Algunas proteínas generadas del mismo transcripto tienen secuencias en común (E1A por ejemplo) mientras que otras no (E1B por ejemplo). No existe un consenso para la nomenclatura de las proteínas llamándose algunas por su masa molecular, otras por sus funciones, otras por su ORF y también por su unidad de transcripción (Figura 11).

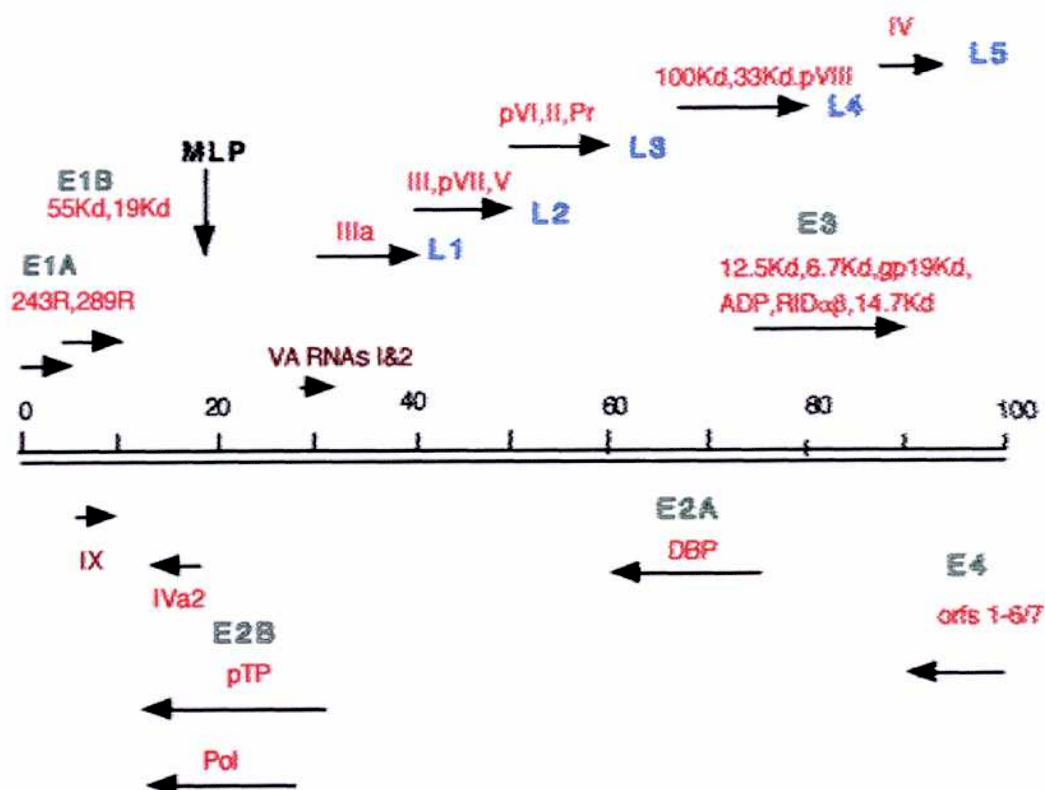


Figura 11: Expresión del genoma en Adenovirus. Los transcriptos tempranos se marcan en verde, los tardíos en azul, en marrón la ubicación de los genes VA RNA y en rojo las proteínas que codifica cada uno. Las flechas indican la dirección de la transcripción. MLP: Major late promoter, promotor tardío mayor. Modificado de Russel, 2000.

Expresión de genes tempranos

En general, la expresión génica temprana tiene tres objetivos:

1. Inducir la entrada de la célula en fase S del ciclo celular, dándole al virus un ambiente óptimo para su replicación.
2. Generar mecanismos de escape que protejan la célula infectada de la defensa antiviral del organismo.
3. Sintetizar proteínas virales necesarias para la replicación.

E1A codifica dos proteínas que activan la transcripción e inducen la entrada de la célula en fase S del ciclo celular; E1B codifica a su vez dos proteínas que previenen el arresto mediado por p53 de la fase G1 y bloquean la apoptosis; E2 codifica tres proteínas que modifican la replicación del DNA; E3 expresa productos que modulan la respuesta del hospedador a la infección y la familia de RNAm tardíos está involucrada en la producción y ensamblado de la cápside viral. Sólo E4 codifica proteínas para funciones diferentes: mediadores de la regulación de la transcripción y transporte del RNAm y moduladores de la replicación del DNA y apoptosis.

El producto del gen E1 se subdivide en E1A y E1B. El primer RNAm y proteína que se detecta ~1h luego de la infección es **E1A** (inmediatamente temprana). Esta proteína es un factor transcripcional que actúa en trans sobre los demás genes tempranos (transactivador) siendo capaz de activar la transcripción de promotores tanto virales como celulares de manera secuencia inespecífica, modificando el entorno celular (Cann, 1997).

E1A tiene dos componentes principales en 5' y en 3' llamadas 289R (o 13S) y 243R (o 12S), en base al número de residuos de aminoácidos o en base al coeficiente de sedimentación del RNAm que las codifica. Ambas se generan por splicing alternativo dando productos con 46 aa de diferencia que comparten motivos conservados CR1, CR2 y CR3 separados por secuencias no conservadas. Durante la fase tardía del ciclo infeccioso se acumulan 3 RNAm adicionales para E1A con función desconocida.

Los productos de E1A modulan el metabolismo celular tornando a la célula más susceptible a la replicación viral. El metabolismo de la célula se caracteriza básicamente por llevar a una división celular precisa reteniendo intactas las funciones celulares. Diversos mecanismos ayudan a que este proceso no sufra interferencia externa y a la remoción de las células defectivas. El primero invoca al sistema inmune innato y adaptativo, mediante la regulación del factor de transcripción NF- κ B, y el segundo opera llevando las células a apoptosis por numerosas rutas, siendo una de ellas el factor transcripción p53 (Russell, 2000; Shenk, 2001).

Brevemente, NF- κ B es un transactivador nuclear que es liberado en el citoplasma por clivaje proteolítico de su factor inhibitorio asociado, I κ B migrando hacia el núcleo y activando los genes de respuesta NF- κ B, entre otros el promotor del gen E3. La fosforilación de por el complejo kinasa, IKK, resulta crucial para la proteólisis de I κ B (Hiscott *et al.*, 2001).

La proteína p53 es un supresor de tumores que regula la transcripción de numerosos genes relacionados con el arresto del ciclo celular y la apoptosis. En células normales, p53 está presente en bajas cantidades, pero sus niveles

aumentan en respuesta a daño genético response y otras injurias. La regulación de los niveles de p53 se realiza a nivel proteico, utilizando la proteína celular mdm2, que se une a p53 actuando como ubiquitina ligasa, llevando a p53 al proteosoma para degradación. Otra proteína celular, p19arf, contribuye a este sistema uniéndose a mdm2, bloqueando su actividad ligasa y estabilizand a p53. Un factor adicional es que p53 puede ser activado por la proteína SUMO (*small ubiquitin-like modifier*) .

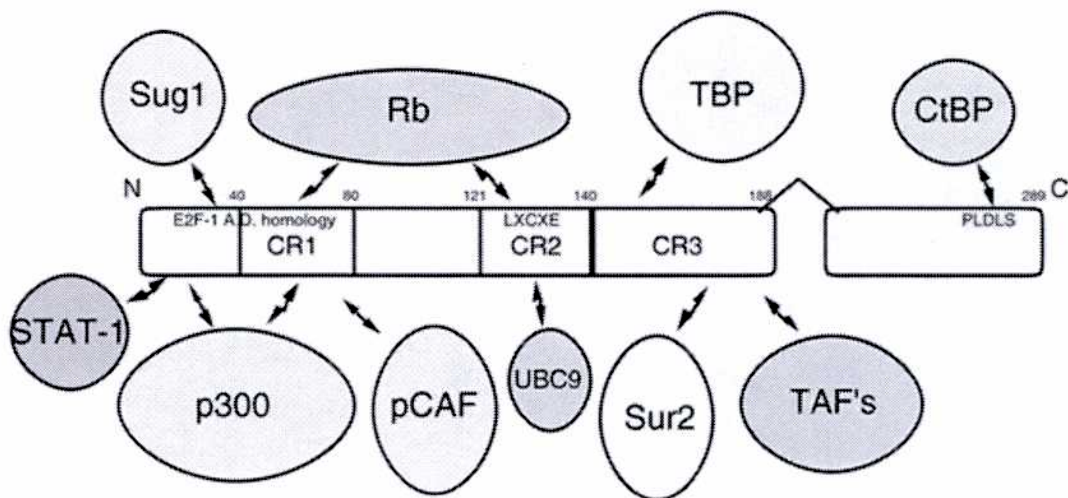


Figura 12: Unión de proteínas celulares a E1A. PLDLS y LXCXE son motivos de reconocimiento. E2F-1 A.D. *homology* se refiere a secuencias que desplazan E2F-1 de Rb y p300. Modificado de Russell, 2000.

La proteína E1A interacciona con: (Figura12)

- Se une a p21 e inhibidores ciclina kinasa dependientes (CDK), estimulando división celular y crecimiento. Se une a ciclinas A y E (complejos CDK), que regulan el pasaje a fase celular S (síntesis de DNA).

- Se une a la familia de transactivadores p300/CBP, que regulan la transcripción celular. Así modula la actividad asociada con acetiltransferasas p300/CBP (pCAF), alterando el nucleosoma. Inhibe la actividad del transactivador STAT-1, requerido para la inducción de la respuesta interferón y la activación de la apoptosis vía caspasas independiente de p53.
- Se une a la familia de proteínas de oncogenes Rb/p130, interfiriendo en la actividad acetiltransferasa de Rb y promoviendo la activación de transcriptos como E2F; crítico para la síntesis de componentes de la fase S. Luego de la unión a pRb, la liberación de E2F y ETF lleva a la transcripción de p53 y p19arf. Así promueve la síntesis de p53 y mantiene su integridad.
- E1A secuestra CBP, que en células normales evita la unión de mdm2, activando a p53.
- E1A se une a UBC9, proteína del SUMO activando p53
- Interacciona vía motivos PLDLS con la proteína CtBP que se une a su vez a Rb y funciona como co-represor transcripcional
- Interacciona con el complejo multiproteico Sur-2 estimulando la transcripción de genes virales.
- Se une a *TATA-box-binding protein* (TBP) y *TBP-associated protein* (TAF) regulando la transcripción de genes celulares.

E1A también bloquea la vía de NF- κ B interfiriendo en los procesos de regulación del ciclo celular (Figura 13).

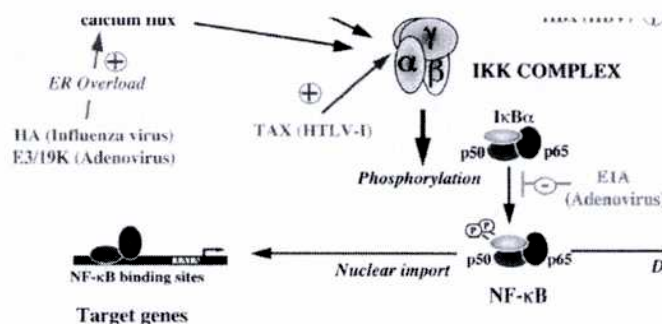


Figura 13: Interacción de E1A con NF- κ B. Modificado de Hiscott *et al.*, 2001.

Adicionalmente otras proteínas virales modulan y cooperan con la actividad de E1A como E4, y E1B.

E1B codifica 2 proteínas de 19 kDa y 55 kDa no relacionadas en secuencia de aminoácidos. Estas dos proteínas actúan de manera independiente y por mecanismos diferentes pero ambas tienen función anti-apoptótica. La proteína E1B 58 kDa directamente detiene la actividad supresora de tumores de p53 inhabilitándola para transactivar la transcripción. E1B 19-kDa inhibe apoptosis por un mecanismo similar a la proteína bcl-2 involucrada. En la supervivencia celular por interacción con miembro de la familia Bax (de transcripción promovida por p53) que induce apoptosis y necrosis.

En la Figura 14 se resume la acción sobre la célula infectada de E1A y E1B.

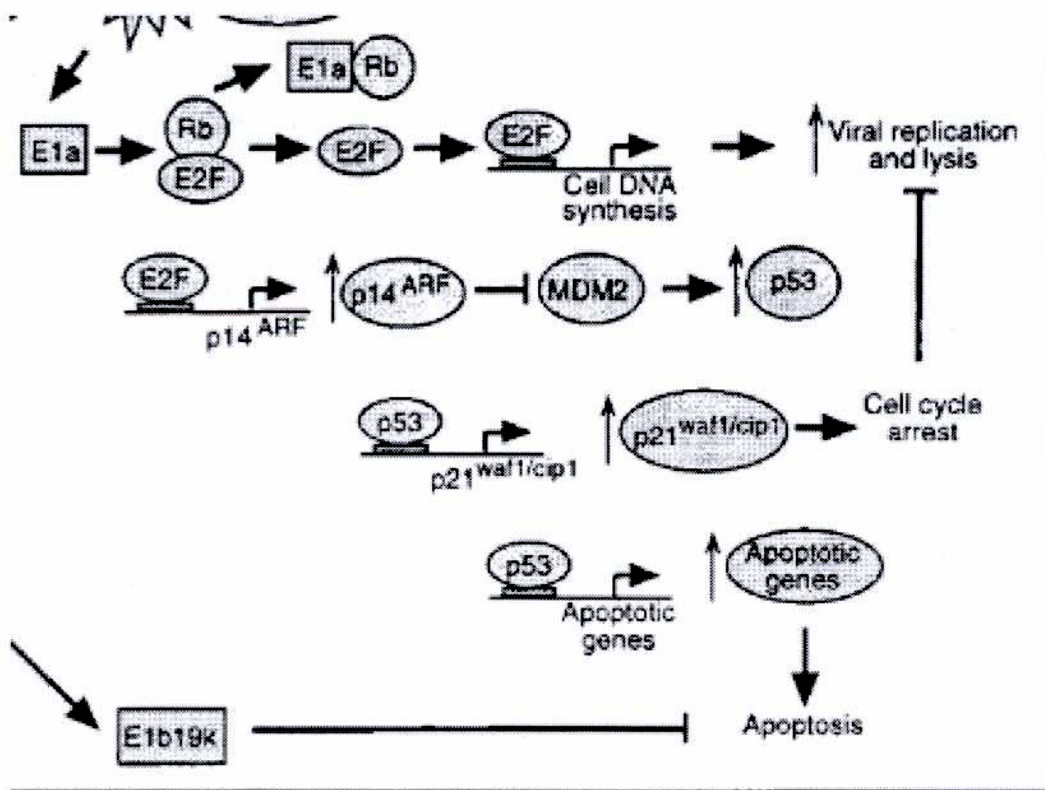
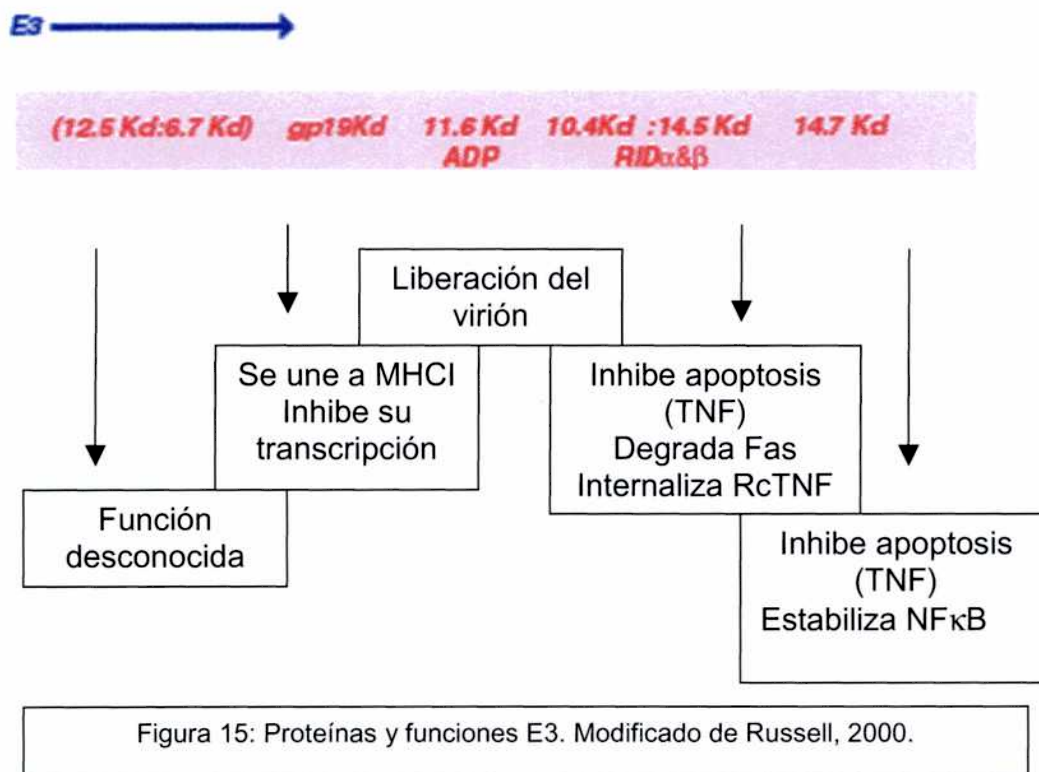


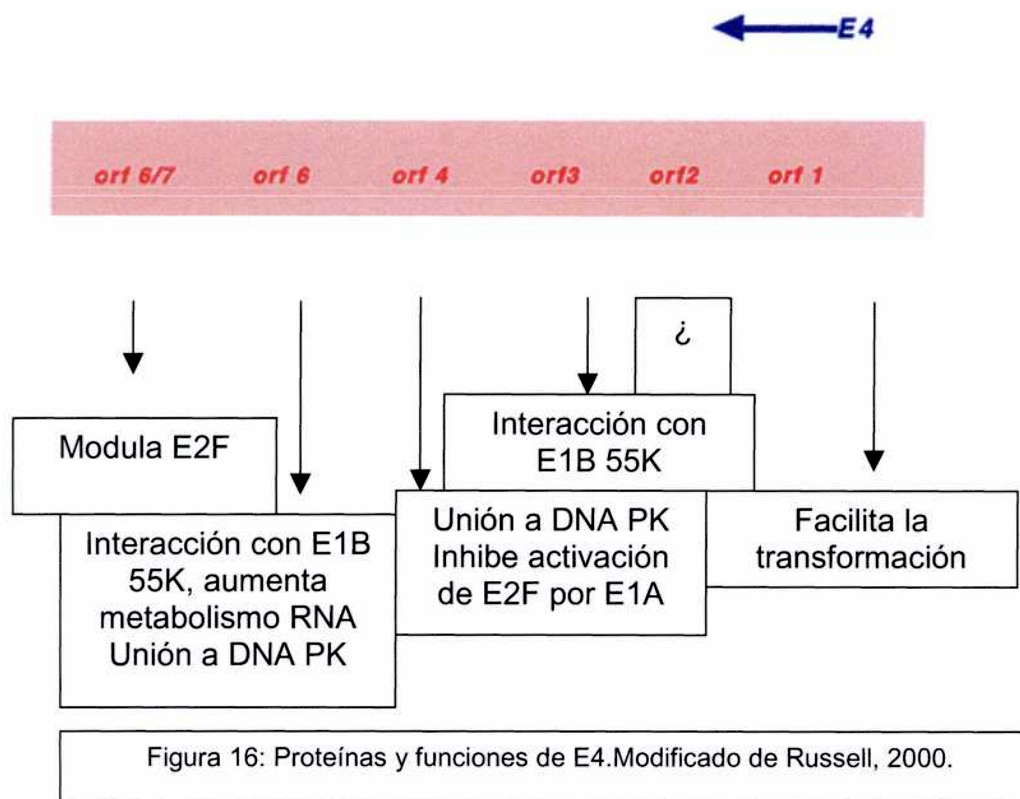
Figura 14: Interacción de E1A y E1B con célula hospedadora. Modificado de Dix *et al.*, 2001.

El gen E2 se divide en dos partes: E2A (DBP) y E2B (pTP y Pol). Provee la maquinaria para la replicación del DNA viral y la transcripción subsiguiente de los genes tardíos

Otras proteínas virales modulan mecanismos de apoptosis, mayormente los productos derivados de la unidad de transcripción E3. Existen 7 proteínas E3, ninguna de ellas es esencial para la replicación en cultivo de células, implicando que su función está relacionada con el bloqueo del sistema inmune. Una de estas proteínas (*Adenovirus death protein*=ADP) facilita la citólisis y liberación de los virus progenie. E3-gp19K, conservada en todos los Adenovirus respiratorios, es la primera línea de defensa contra CTL, se localiza en la membrana celular y se une a todos los MHC clase I evitando su transporte a superficie para ser reconocida por los CTLs. Los Adenovirus que infectan tracto gastrointestinal carecen de esta proteína, pero *downregula* el MHC clase I reprimiéndolo a nivel transcripcional, mediante E1A. Otras proteínas E3, RID y E3-14.7K inhiben apoptosis inducida por CTL en células infectadas. E3 permite a los Adenovirus evadir la respuesta inmune del hospedador (Figura 15) (Russell, 2000).



Los productos del *cassette* E4 se llaman orfs 1–6/7 y principalmente facilitan el metabolismo de RNAm viral en asociación con E1B promueven la replicación del DNA viral y el *shutoff* de la síntesis proteica celular. Se asocian también a la resistencia a lisis por CTLs. Se han asociado a proteínas kinasas activadas por DNA evitando la concatenación del DNA viral, como lo hacen bajo activación de p53 sugiriendoun bloqueo de la apoptosis. E1B y E4 contrarrestan las actividades de E1A. E4 orf4 inhibe la activación del promotor E2F. E4 orf6 tiene un efecto directo en la transactivación de p53 por interacción con E1B 55k cuando se une a p53, liberándolo para su degradación. E4 orf3 alivia la inactivación de p53 por E1B (Figura 16).



Todas las observaciones hechas indican que los Adenovirus, en el curso del ciclo infeccioso secuestran la maquinaria celular y alteran el ambiente celular para favorecer la replicación viral, teniendo profundos efectos en las funciones celulares normales. En la Figura 17 se resume lo explicado hasta el momento.

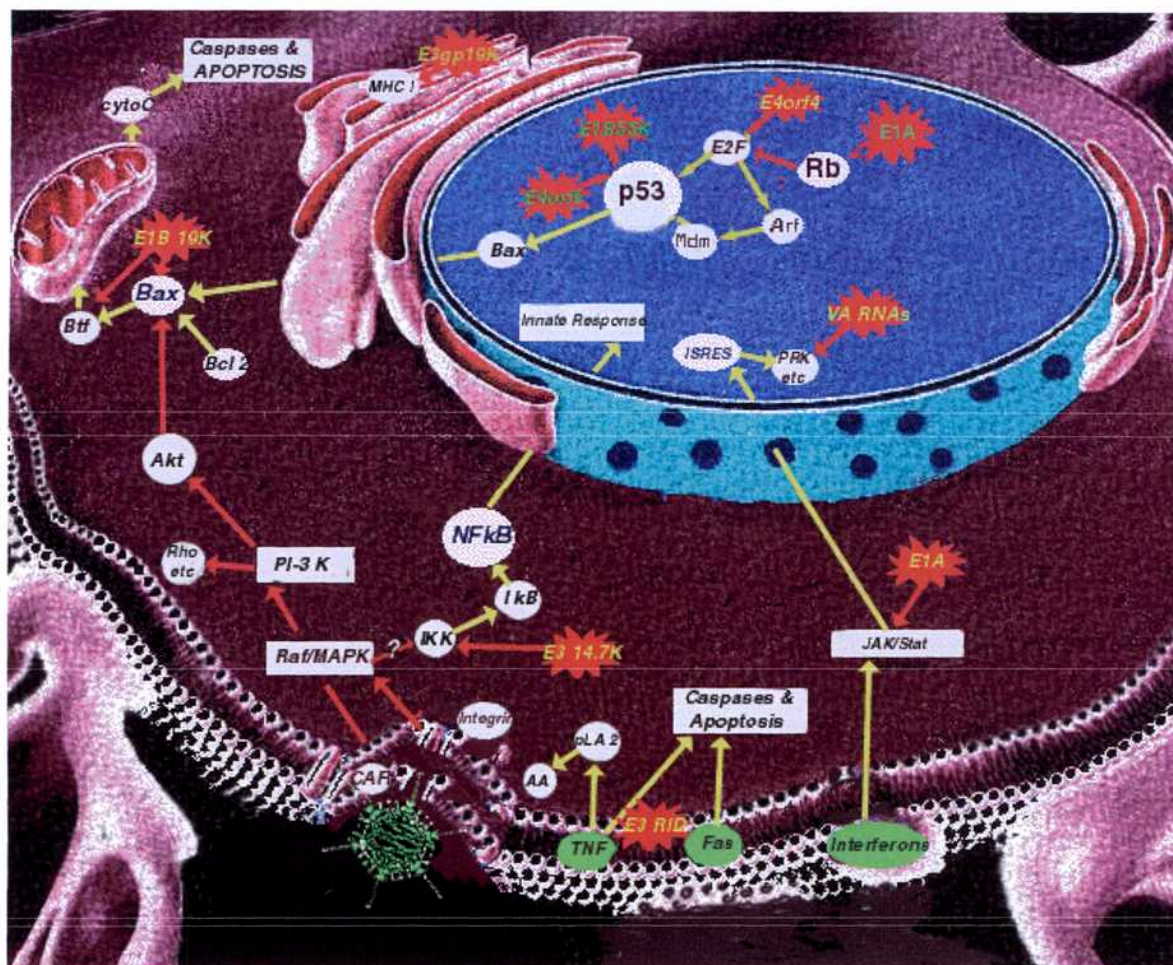


Figura 17: Mecanismos de interacción del Adenovirus con la célula.
Modificado de Russell, 2000.

Replicación del DNA Viral

Con la acumulación de las proteínas de E2 comienza la replicación del DNA 5-8 horas post-infección y continúa hasta la muerte de la célula.

Activación de expresión de genes tardíos y *shutoff* celular

Una vez que los RNAm tempranos se sintetizaron se traducen en polisomas junto con los RNAm celulares. Inicialmente, no tienen ninguna ventaja competitiva, pero luego de la replicación del DNA, al entrar en la fase tardía, los RNAm celulares son excluidos de los polisomas.

Los genes tardíos comienzan a ser expresados de manera eficiente al comienzo de la replicación del DNA viral. Las regiones tardías están organizadas en una sola unidad de transcripción grande cuyo transcripto primario tiene 29,000 nucleótidos. Este transcripto se procesa por *splicing* alternativo o uso de diferentes sitios poly(A) para generar múltiples RNAm, agrupados en 5 familias L1 a L5, según el uso de sitios de adición poly(A) común. La expresión de esta amplia familia de RNAm se halla bajo el control del promotor tardío principal o *major late promoter*. Este promotor tiene actividad baja pocas horas post-infección pero luego se vuelve varios cientos de veces más activo

Cuando comienza la replicación del DNA y se sintetizan todos los RNAm tardíos, la acumulación citoplasmática de RNAm celular se bloquea aunque la síntesis y procesamiento de los transcriptos celulares continúe, sugiriendo que el transporte está bloqueado. Este proceso está mediado por E1B 55-KD y por E4 34-KD que existen como complejo. Las mismas proteínas son esenciales

para una acumulación RNAm viral citoplasmática eficiente post infección. Sólo RNAm para β -tubulina continúa acumulándose en el citoplasma pero no es traducido. Varios componentes regulatorios cooperan para facilitar la traducción selectiva de los RNAm virales tardíos post infección. El primero involucra a proteína kinasa celular R (PKR), activada por RNA doble cadena. Los VA RNAs inhiben la activación de la PKR, protegiendo la síntesis de proteínas virales y proveyendo una compartimentalización funcional.

Finalmente, hay una activación selectiva de la síntesis proteínas virales tardías por la proteína L4 100-kd que puede unirse a RNAm, funcionando en el polisoma favoreciendo la traducción (Shenk, 2001).

El ensamblado ocurre en el núcleo, pero comienza en el citoplasma cuando los monómeros individuales forman los capsómeros hexón y pentón. Las cápsides vacías inmaduras a partir de estos protómeros son ensambladas en el núcleo, donde se forma el core de DNA genómico asociado a proteínas del core. La infección puede ser lítica facilitada por la proteína ADP (*adenovirus death protein*) (E3-11.6K), que también induce apoptosis. Las partículas virales tienden a acumularse en el núcleo y se visualizan en el microscopio como cristales eosinófilos llamados cuerpos de inclusión (Figura 18). Se cree que son la base de las infecciones latentes siendo la reactivación provocada por lisis accidental de las células infectadas, liberando partículas virales desde el núcleo. Este tipo de mecanismo de persistencia se conoce más como infección oculta.

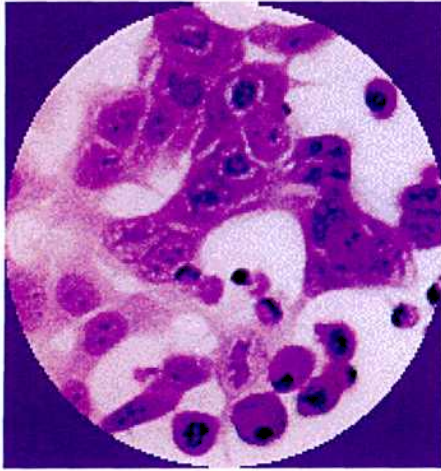


Figura 18: Infección lítica por Adenovirus. Las células infectadas con Adenovirus tienen afinidad por la Hematoxilina (colorante azul-purpúreo). Las células infectadas se redondean y la monocapa se desintegra. Las inclusiones basofílicas oscuras en el núcleo representan el cúmulo de proteínas virales en el sitio de ensamblado del virus.

Descripción del hexón:

El hexón es un homotrímero, consistiendo cada monómero en una proteína compleja de aproximadamente 900 aa. El estudio por cristalografía de rayos X y microscopía electrónica del trímero (hexón) de AV2 revela una base densa hexagonal que contiene tres copias de los dominios P1 y P2 (internos), desde la cual se proyectan tres *loops* superficiales: L1, L2 y L4, que contienen varios epitopes. De los tres *loops*, L1 es el más largo y complejo, proyectándose hacia el exterior, lo que le permite una mayor interacción con el medio (Figura 19) (Roberts *et al.*, 1986; Stewart *et al.*, 1991 y Athappilly *et al.*, 1994).

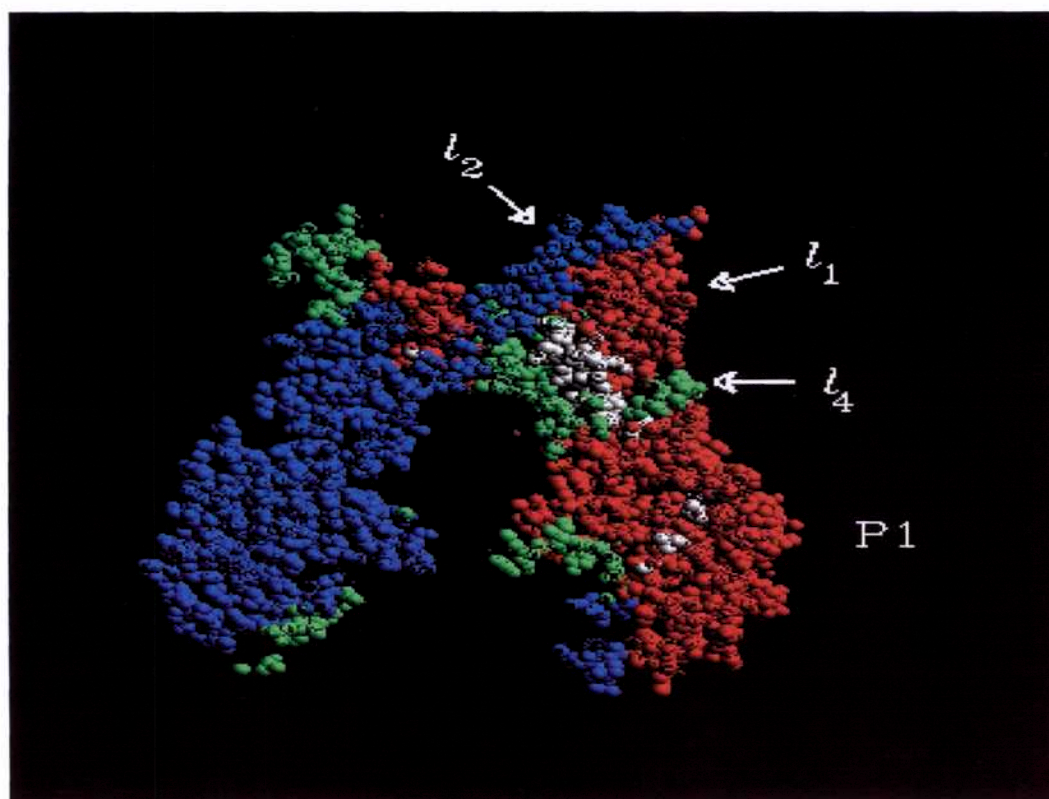


Figura 19 :Estructura del Hexón de Adenovirus. Tomado de San Martin & Burnett (2003).

El gen del hexón se ubica entre 0.51 y 0.6 unidades de mapa de la hebra positiva del genoma y se puede dividir en tres segmentos diferentes: una región variable central (nucleótidos 403-1356) y dos regiones altamente conservadas en los flancos. Las regiones más variables se hallan en L1 y L2 (entre los nucleótidos 403-912 y 1219-1356) (Li y Waddell, 1999). L1 y L2 son determinantes de serotipo y L4 es determinante de subgénero. El loop interno L3, en cambio, es compartido por serotipos de los subgéneros B, E y F (Pring-Akerblom *et al.*, 1995). Dentro de la región variable se definen siete regiones hipervariables (HVR) que difieren tanto en secuencia como en longitud. Las HVR 1 a 5 se ubican en L1 y HVR 7 en la punta de L2. HVR1 tiene la mayor variabilidad, relacionada con el tropismo (Crawford-Mikszá *et al.*, 1996).

HVR1 y HVR6 constituyen la base de L1; se ha sugerido que su variabilidad puede compensar el efecto topológico de otros cambios en la proteína. La variable longitud de HVR4 y HVR5 sugiere mecanismos de inserción en ellas (Figura 20) (Crawford-Mikszá *et al.*, 1996).

La neutralización es específica de serotipo y se dirige a los epitopes ϵ . Una sola molécula de anticuerpo anti-hexón por virión es necesaria para neutralizar la infectividad. L1 y L2 parecen buenos candidatos para epitopes de neutralización conformacionales, no lineales, compuestos por dos o más HVRs (Crawford-Mikszá *et al.*, 1996 y Wohlfart *et al.*, 1988).

Las cepas intermedias y variantes genotípicas aparecen por recombinación entre dos serotipos parentales; además, el gen del hexón evoluciona rápidamente. Las HVR son *hotspots* para recombinación ilegítima (inserciones, deleciones, duplicaciones y translocaciones en regiones de

homología limitada) y para mutaciones de una sola base, mayormente en sitios ricos en purinas. Sectores de ADN donde una hebra es de polipurinas y la otra de polipirimidinas distorsionan la estructura de la hélice, volviéndose sitios determinantes de mutaciones (Crawford-Mikszá *et al.*, 1996).

Tipos genómicos de AV7

Por análisis de variaciones génicas en el hexón de los diferentes genotipos que integran el serotipo 7 se ha determinado la presencia de dos *clusters* genómicos

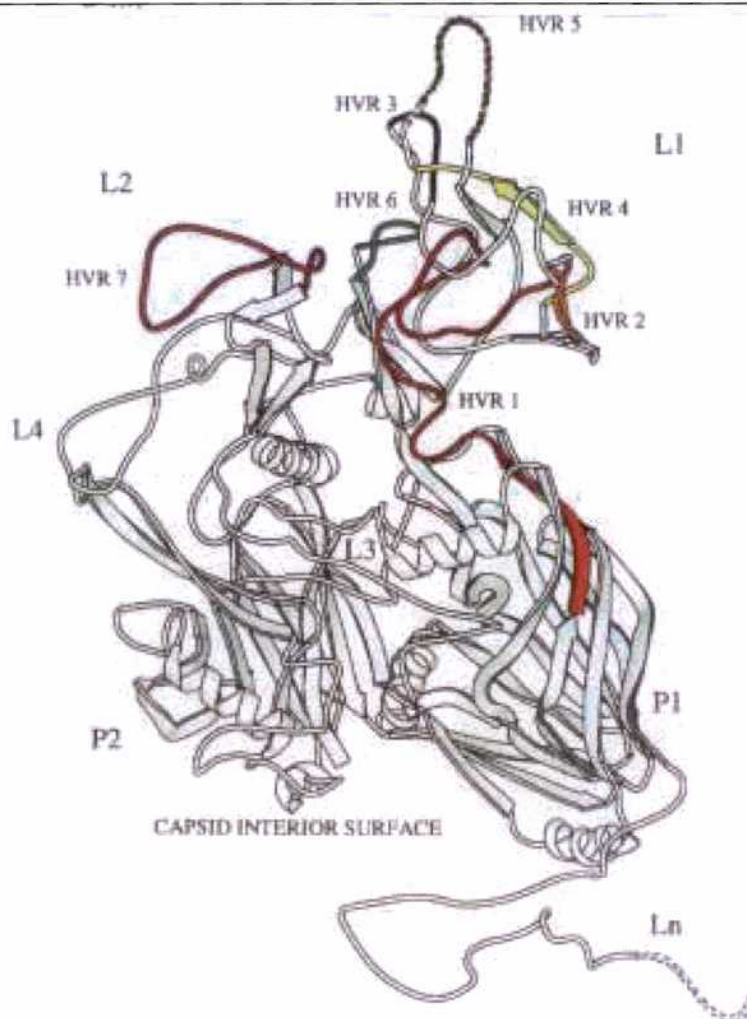
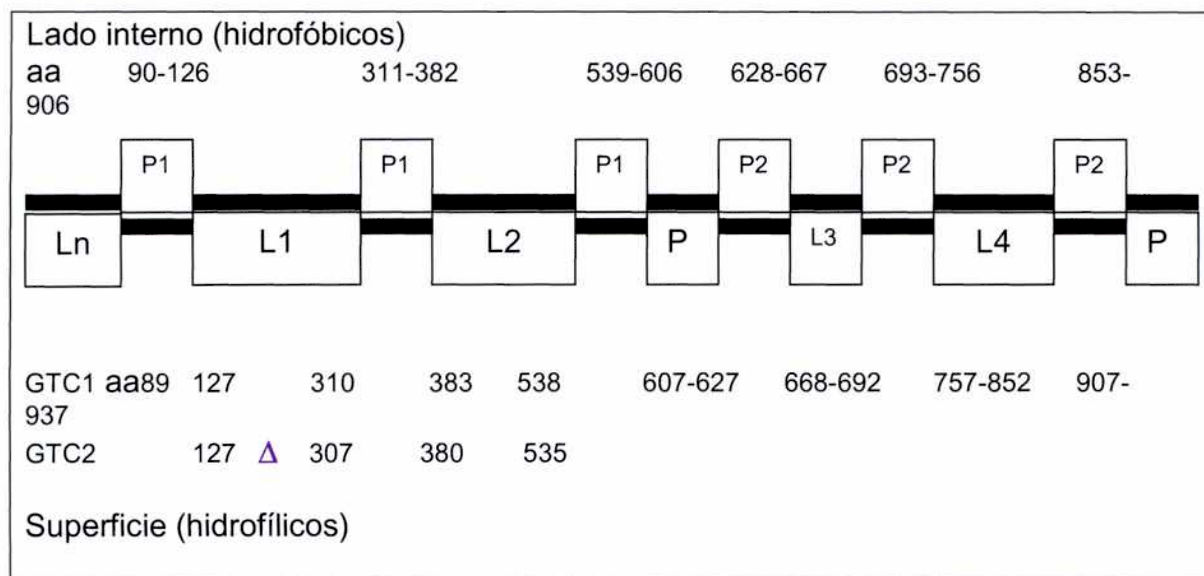
- *Cluster* de tipo genómico 1 (GTC1): incluye Ad7p y Ad7p1
- *Cluster* de tipo genómico 2 (GTC2): incluye Ad7b Ad7c Ad7d Ad7g y Ad7h

El grado de homología es mayor al 98% dentro de un *cluster* y entre 93 y 94.5% entre *clusters*. Dada la presencia de dos motivos repetidos cortos en el GTC2 se sugiere que éste podría derivar del GTC1 por recombinación ilegítima (13). Existe una delección de 9 nt en L1 del GTC2 flanqueada por estas repeticiones cortas.

El prototipo del serotipo es la cepa Gomen (AV7p) perteneciente al GTC1, su hexón ha sido secuenciado completamente y consta de 2814 nucleótidos y la proteína de 937 aminoácidos (aa). Comparado con otros miembros del subgénero B es más corto (15 y 9 nucleótidos que AV3 y AV16 respectivamente) (Li Q. *et al.*, 1997)

La ubicación de las 7 HVRs de acuerdo al perfil hidropático se puede ver en la Figura 20.

Figura 20: a- Esquema del perfil hidropático del Hexón de Adenovirus.
 b- Esquema del hexón de Adenovirus. Abreviaturas: L (loop), P (pedestal), Pc (conector del pedestal), Ln (loop amino terminal) Tomado de Crawford-Miksza *et al.*, 1996.



Interacción con el huésped

Latencia

Los Adenovirus pueden mantener una asociación con su hospedador humano persistiendo por años luego de la infección inicial, probablemente en linfocitos. Existen tres proteínas que podrían facilitar la persistencia por antagonizar la respuesta antiviral del hospedador: E1A y VA RNAs, que inhiben la respuesta celular a interferon α y β , y las proteínas E3 que protegen las células infectadas de la muerte por CTLs y *tumor necrosis factor* (TNF). La infección latente por Adenovirus se ha postulado como elemento importante en la patogénesis de la enfermedad pulmonar obstructiva crónica (*chronic obstructive pulmonary disease*= COPD) que ocurre en adultos fumadores. Se encontró DNA E1A en células epiteliales, sugiriendo un rol en COPD. En estudios en niños, la infección persistente se asoció con asma y con enfermedad pulmonar crónica postviral (Hogg, 2001). El significado de estas observaciones requiere mayor estudio.

Se demostró que el promotor de E1A y los VA RNAs transcriptos por la RNA polimerasa III se expresan constitutivamente en células infectadas persistentemente, y el promotor E3 contiene varios sitios de unión para el factor de transcripción linfoide NF κ B, facilitando su expresión constitutiva en células linfoides. Estos genes virales protectivos se inducen en respuesta a interleuquina-6 (IL-6), quimocina producida en el sitio de infección. La transcripción de genes E2 se activa por IL-6 produciendo DBP (*DNA-binding protein*), que a su vez activa los promotores de E1A y E3 (Shenk, 2001).

Mecanismos de defensa

La morbilidad asociada con Adenovirus es baja y resulta evidente que existe una defensa bien orquestada del hospedador tanto innata como adaptativa

La respuesta innata está constituida por péptidos de 3–4 kDa llamados defensinas, quemoquinas que reclutan neutrófilos y macrófagos, activación de complemento y células NK (*natural killer*). La transcripción del factor NF- κ B es también importante en estadios tempranos de la infección activando la transcripción de citoquinas proinflamatorias y moléculas de adhesión (Russell *et al.*, 2000).

Antagonistas virales de Interferon α y β

Los interferones funcionan por unión a receptores celulares, activan las cascadas celulares de Jak/STAT, llevando los complejos STAT al núcleo donde se unen a elementos de respuesta a interferón (*interferon-response elements* =ISREs) en el DNA celular. Los ISREs regulan la transcripción de genes como la protein kinasa inducida por RNA doble cadena (PRK) y la 2'-5' oligoadenilato sintetasa (OAS) entre otros inmunomoduladores, que funcionan como respuesta antiviral. Diferentes serotipos de Adenovirus llevan uno o dos genes VA RNA; la especie C lleva 2 (VA RNA_I y VA RNA_{II}). La síntesis de VA RNA comienza durante la fase temprana del ciclo infeccioso y se acelera durante la fase tardía. La síntesis de una forma inactiva de PKR se induce por interferón, y se activa por la presencia de RNA doble cadena. PKR activada activa el factor de elongación eIF-2 α que secuestra eIF-2B y se

produce un arresto de la traducción (Ma y Mathews, 1996; Goodbourn *et al.*, 2000).

Antagonistas virales de linfocitos T citotóxicos (CTL) y TNF- α

La histopatología de la enfermedad pulmonar presenta dos estados. El primero se caracteriza por daño leve a moderado en células del epitelio bronquiolar y difusa infiltración celular en regiones peribronquiolares y alveolares. En ese estadio aparecen en tejido pulmonar infectado TNF- α , IL-1, e IL-6. El segundo estadio de la enfermedad consiste meramente en infiltración por linfocitos. Estos dos estadios de patogénesis resultan de una respuesta inicial no específica del hospedador a la infección viral que incluye síntesis de citoquinas seguida por respuesta específica CTL.

El reconocimiento y posterior lisis por CTLs de una célula infectada requiere que los antígenos virales se presenten en un contexto de MHCI en la superficie de la célula infectada. Las células infectadas que expresan E3 gp19-KD son menos sensibles a lisis mediada por CTL. Como señalamos anteriormente, esta proteína reside en la membrana del retículo endoplásmico y su dominio luminal se une directamente al dominio de unión a péptido de MHC I, reteniéndolo. CTLs portan Fas ligando en superficie celular e induce apoptosis cuando el Fas ligando interacciona con el receptor Fas (CD95) en la célula target. Las proteínas E3 bloquean la apoptosis mediada por Fas. E3 14.7-KD o el complejo E3 14.5-KD–E3 10.4-kd proteína puede evitar citolisis por TNF- α inhibiendo la activación de fosfolipasa citosólica A2, que actúa *downstream* del receptor de TNF- α . E3 14.7KD interacciona con FLICE

(caspasa 8) que actúa *downstream* de los receptores FAS y TNF, y esta interacción es la responsable del parte del bloqueo de apoptosis por ambos receptores. Se sugiere que existe reactividad cruzada en CTLs humanos para reconocer diferentes subgrupos de Adenovirus (Smith *et al.*, 1998).

Las células CD4 *helper* son importantes para la respuesta proliferativa a la infección, por reconocimiento de antígenos virales presentados en un contexto MHC clase II. Estas células T *helper* pueden estimulan la proliferación de células B y proveer inmunoglobulinas para la respuesta humoral.

Respuesta humoral

La respuesta humoral es el componente principal de la defensa del hospedador y depende de la capacidad de reconocimiento de antígenos extraños por las células B. El reconocimiento mediado por células T *helper* inicia una masiva proliferación y libera inmunoglobulinas específicas para que interactúen directamente con los antígenos. Los anticuerpos neutralizantes se dirigen contra epitopes componentes de la cápside. Los 51 serotipos de Adenovirus humanos se clasifican en base a su capacidad neutralizante específica siendo la protección humoral muy restringida a cada serotipo (Russell, 2000; Valentine y Pereira, 2003).

Epidemiología

Entre los agentes causantes de infección respiratoria aguda en niños, los Adenovirus (AV) tienen un rol importante en la infección severa y/o fatal y en el desarrollo de enfermedad pulmonar crónica. Entre los serotipos de AV más frecuentemente aislados, los miembros del subgénero B, especialmente del serotipo 7 (AV7), han sido asociados con manifestaciones clínicas severas (Hogg, 2001).

Por análisis con enzimas de restricción los AV7 pueden subdividirse en varios genotipos diferentes abarcando desde 7a hasta 7i, con un patrón de distribución geográfica particular: AV7b en Australia, Europa y Estados Unidos; AV7c en Africa y AV7d en China (Kajon y Waddel, 1992).

Estudios epidemiológicos y del impacto clínico de enfermedades respiratorias por AV7 llevados a cabo en Argentina, Chile y Uruguay entre los años 1984 y 1990 revelaron que el 5,6% correspondía a AV7b, 9,9% a AV7c y 84,5% a AV7h. En ese estudio sólo AV7h se asoció a casos fatales de enfermedad respiratoria aguda, posiblemente dada su alta prevalencia y/o su alta patogenicidad (Kajon *et al.*, 1994).

La distribución temporal indica que ocurrió un desplazamiento de Ad7c, prevalente en 1984 hacia el genotipo 7h en 1986, el cual continuó siendo el predominante en los 10 años siguientes (Waddel *et al.*, 1985). Sin embargo, a partir de 1991 comenzó a aparecer otro tipo genómico (7i) que se asocia también a casos fatales (Mistchenko *et al.*, 1998.)

Diversos organismos poseen mecanismos endógenos de mutación, incluyendo recombinación ilegítima, los cuales les otorgan una ventaja para su supervivencia. Se ha sugerido que los AV han adoptado un mecanismo

evolutivo que les confiere capacidad de mutar en regiones inmunogénicas, lo que les permite escapar del sistema inmune del huésped y asegurar así su perpetuación (Crawford-Miksza *et al.*, 1996). Sin embargo, se desconocen los mecanismos evolutivos precisos que han actuado en la diversificación de los AV7 y en la prevalencia de un tipo genómico sobre otro, a través del tiempo.

Patogénesis

Durante la infección por Adenovirus se producen inmunocomplejos circulantes que se depositan en riñón y pulmón, causando una respuesta inflamatoria que dificulta el intercambio gaseoso en la membrana alvéolo-capilar (Mistchenko *et al.*, 1992). Los niveles séricos elevados de las citoquinas proinflamatorias IL-6, IL-8 y TNF- α se correlacionan con una evolución desfavorable, explicando la similitud entre un cuadro séptico bacteriano y la infección diseminada por Adenovirus (Mistchenko *et al.*, 1994). En las infecciones severas o fatales hay una severa disminución de linfocitos T CD4⁺ y CD8⁺ y de efectores *natural killer*, con valores elevados del antagonista del receptor de IL-1 (IL-1ra). A diferencia del resto de las citoquinas, el TGF- β 1 está significativamente elevado en pacientes con evolución favorable (Mistchenko *et al.*, 1998).

Desde el punto de vista histológico, durante el período agudo el virus se multiplica en bronquios pequeños y bronquiólos terminales, causando necrosis parcial del epitelio, con abundante cantidad de proteínas virales (fundamentalmente hexón) en los neumonocitos de tipo II y macrófagos alveolares (Ladenheim *et al.*, 1995; Hogg *et al.*, 1989). Estas alteraciones están a menudo asociadas con áreas de atelectasia, dando el patrón

característico de daño alveolar difuso. El análisis de la matriz extracelular muestra un discreto incremento de las fibras intersticiales reticulares, junto con colágeno y fibras elásticas, más pronunciado en las paredes de los bronquiólos respiratorios y ductos alveolares. Los septos intraalveolares muestran una red delicada de fibras colágenas de tipo I y III (Rosman *et al.*, 1996).

En las etapas finales de la infección, la cantidad de células infectadas que expresan antígenos virales es pequeña y el daño está conformado por fibrosis extensa y difusa del parénquima pulmonar, bronquios, bronquiólos terminales y respiratorios y septos alveolares, con paredes engrosadas, proliferación de células similares a fibroblastos, positivas para vimentina y α -actina de músculo liso, usualmente llamadas miofibroblastos (Rosman *et al.*, 1996). La diferenciación de fibroblasto a miofibroblasto se asocia con la expresión de α -actina de músculo liso (Thibeault *et al.*, 2000). Los miofibroblastos son células del intersticio pulmonar de mamíferos y proliferan en el tejido pulmonar con fibrosis. A menudo, los bronquios y bronquiólos muestran hiperplasia epitelial o metaplasia escamosa, algunas veces asociada con hipertrofia de las uniones parietales de las células del músculo liso. Los abundantes infiltrados linfomonocitarios en las paredes alveolares y lesiones residuales de fibrosis septal sugieren la participación de mecanismos de daño adicionales a la lesión que *per se* produce el virus. Estos mecanismos podrían intervenir en la patogénesis de la enfermedad. El análisis de la matriz extracelular muestra un incremento de distribución aleatoria de las fibras de colágeno tipo I y III en los septos intraalveolares, con aumento difuso de la red de reticulina. El TGF- β 1 estimula la

diferenciación de fibroblasto a miofibroblasto por un mecanismo antiapoptótico, lo que contribuiría al efecto fibrogénico de esta citoquina (Zhang y Phan, 1999).

El desarrollo del modelo animal para el estudio de la patogenia de la infección por Adenovirus, en la rata *Sigmodon hispidus*, mostró un patrón similar, y permitió dividir la infección dos fases (Prince *et al.*, 1993; Ginsberg *et al.*, 1991). Una temprana, que aparece en las primeras 24 horas, se caracteriza por la acumulación de neutrófilos, macrófagos y monocitos en los alvéolos pulmonares. Comienza con el influjo de polimorfonucleares en las primeras 6 horas, llegando a un máximo a las 24 horas. Es reemplazado por un infiltrado mononuclear compuesto de linfocitos T CD4⁺ y CD8⁺ y fagocitos mononucleares, que llega al máximo al 5° día, fase tardía. La eliminación del virus también sigue un curso bifásico, con una fase de eliminación rápida en la cual se clarifica el 60% del virus dentro de las primeras 24 horas, seguida de una fase lenta en la cual el DNA viral se detecta por semanas o meses. Mientras en la fase temprana la eliminación sería efectuada por polimorfonucleares neutrófilos, en la tardía intervendrían los linfocitos T citotóxicos (CTL), eliminando las células infectadas.

Naturaleza del aporte original proyectado

Los Adenovirus constituyen aproximadamente el 4-9% de los agentes etiológicos de la infección respiratoria aguda baja (Viegas *et al.*, 2003). Los serotipos 3, 7 y 21, pertenecientes al subgénero B, están frecuentemente asociados con infección severa y/o fatal y con daño pulmonar residual, en particular la entidad conocida como "enfermedad pulmonar crónica" (*chronic lung disease*) (Macek *et al.*, 1994; Wenman, 1982). Estos rasgos diferencian la infección por Adenovirus de las causadas por otros virus respiratorios, siendo especialmente característicos de ciertos serotipos. Desde el aislamiento de la cepa prototipo de Av7 (Av7p) en 1958, se ha descrito 20 tipos genómicos diferentes, distribuidos en todo el mundo (Wadell, 1984; Wadell *et al.*, 1985). Estudios de epidemiología molecular mostraron que desde el año 1984 hasta el presente el Adenovirus 7h circula en nuestro país y se encuentra frecuentemente asociado a infección severa y/o fatal (Niel *et al.*, 1991; Kajon y Wadell, 1992). Recientemente ha sido encontrado también en Brasil, Japón y EEUU (Erdman *et al.*, 2002; Hashido *et al.* 1999; Tanaka *et al.*, 2000).

Por otra parte se conocen las secuencias completas de sólo 4 de los 51 Adenovirus humanos AV2, AV5, AV 11, 12 y AV40 y de varios Adenovirus que afectan otras especies de Vertebrados como el Adenovirus Aviar CELO y Adenovirus canino 1 entre otros (Gingeras *et al.*, 1982; Roberts *et al.*, 1984; Chroboczek *et al.*, 1992; Chiocca *et al.*, 1996; Morrison *et al.*, 1997, Stone *et al.*, 2003). Mientras el número de secuencias disponible aumente, mejor será la predicción sobre el origen y la evolución de los Adenovirus.

Materiales y métodos.

Materiales y métodos

Muestras y datos epidemiológicos:

Para la determinación de la secuencia completa se empleó el stock original de la cepa patrón de Adenovirus serotipo 7 genotipo h (cepa 87-922) que fue aislada en 1987 en el Laboratorio de Virología del Hospital de Niños Dr. R. Gutiérrez a partir de un aspirado nasofaríngeo de un paciente con neumonía. La misma fue caracterizada por restricción enzimática (Niel *et al*, 1991; Kajon y Waddell, 1992).

Adicionalmente para el estudio de polimorfismos del gen hexón se estudiaron retrospectivamente muestras que fueron remitidas al Laboratorio de Virología del Hospital de Niños Dr. Ricardo Gutiérrez por distintas instituciones de la Ciudad de Buenos Aires y conurbano bonaerense entre los años 1999 y 2002 para el análisis de etiología en infecciones respiratorias agudas. Se recuperaron también muestras del año 1990 (1 muestra de evolución fatal), 1991 (1 muestra de evolución fatal), 1992 (2 muestras de casos fatales) y 1998 (1 muestra caso grave). Las secreciones nasofaríngeas (SNF) de pacientes hospitalizados fueron enviadas junto con una ficha epidemiológica y procesadas en el día por técnicas de inmunofluorescencia directa e indirecta (IFI) para detectar Adenovirus (AV), Virus Sincicial Respiratorio, Parainfluenza 1-2-3, e Influenza A y B. Las secreciones fueron conservadas en freezer a -20 °C para su posterior estudio. Los datos epidemiológicos y clínicos se organizaron en bases de datos (Epi6 y MS Excel).

Datos meteorológicos

Se obtuvieron datos de temperaturas medias mensuales (°C), humedad relativa media (%) y radiación solar UVB (dosis eritémica mínima en piel sensible blanca tipo II) del Servicio meteorológico Nacional (Fuerza Aérea Argentina). Los datos se graficaron y se realizó un test de correlación de Spearman para evaluar la relación existente entre los parámetros meteorológicos y la frecuencia de los Adenovirus. Para el cálculo de p se realizó un test a dos colas con un nivel de significación de 0.05. Se tomaron adicionalmente datos de 1998 para ampliar esta parte del estudio.

Cultivo y purificación de DNA

Para el estudio de los polimorfismos del gen hexón las muestras que resultaron positivas por IFI para Adenovirus fueron preservadas en *freezer* a -20 °C hasta el momento de inoculación así como también aquellas muestras que presentaron clínica compatible con Adenovirus e IFI negativa. Previo al pasaje en células se procedió a la clarificación del material por centrifugación a 10000 r.p.m. por 30 minutos, luego se inocularon 200 µL del sobrenadante en botellas de 3 mL de cultivo de células 293 o Hep-2 (ATCC) mantenidas en MEM Hanks al 2% SFB y se observó diariamente hasta la aparición de efecto citopático. Se realizaron hasta tres pasajes en ciego antes de considerar un cultivo como negativo. Cuando el efecto citopático fue visible, se procedió al congelamiento del cultivo para su posterior análisis molecular. El DNA de las muestras 1990-2002 se extrajo con una resina comercial que secuestra los inhibidores de PCR presentes tanto en el cultivo como en el inóculo.

Para la secuenciación del genoma, se inoculó en células 293 la cepa patrón de Adenovirus 7h (87-922), llegando a un volumen total de 300 mL. Al tener

100 % de efecto citopático, las células se congelaron y se procedió a la purificación del virus en gradiente de Cloruro de Cesio. Brevemente, el cultivo se resuspendió en PBS y se centrifugó a 35000 r.p.m. en una ultracentrífuga por 3 hs., se resuspendió en Tris-HCl 20 mM y se agregó buffer TES (Tris-HCl 20 mM, pH8, EDTA 10mM, 0.5% p/V SDS) conteniendo RnasaA (10 µg/mL). Luego se incubó 30 min a 37 °C y se digirió con Proteinasa K (100 µg/mL) por 3 hs más. Se agregó NaCl hasta concentración de 1M y se incubó a 65 °C para reducir la viscosidad. Luego se extrajo con fenol-cloroformo y se precipitó con etanol. La pureza del DNA se midió por espectrofotometría (Tollefson *et al.*, 1998, Sambrook y Russell, 2001).

Obtención de biblioteca genómica de DNA viral

Se digirió de 1µg del DNA extraído con las enzimas de restricción *BamH* I, *Bgl* II, *Hind* III y *Sma* I por 1 h en el buffer correspondiente y a la temperatura indicada por el protocolo (37 °C para *BamH* I, *Bgl* II y *Hind* III y 25 °C para *Sma* I). Los fragmentos se corrieron en geles de agarosa al 1% (p/v) en buffer TAE a 4 °C por 3 hs se tiñeron con bromuro de etidio, se purificaron y clonaron en el plásmido pGemT (Zhou M. *et al.*, 1995). Se aplicó la fórmula siguiente para calcular la cantidad de inserto necesario para cada fragmento usando como constantes una relación molar inserto vector de 3:1, tamaño del vector 3 kb y 50 ng de vector en todos los casos.

$\text{Inserto (ng)} = \frac{\text{ng de vector} \times \text{tamaño del inserto (kb)} \times \text{relación molar inserto vector}}{\text{Tamaño del vector (kb)}}$

Se cotejaron los cortes esperados de acuerdo a Niel *et al.*, 1991

*Bam*H I: 10 fragmentos

Bgl II: 12 fragmentos

Hind III: 9 fragmentos

Sma I: 7 fragmentos

Se transformaron y plaquearon bacterias *E. coli* JM109 (*endA*⁻) competentes en medio Luria-Bertani (LB) con ampicilina (sal de sodio 100ug/ mL), 0.5mM isopropil- β -D-tiogalactopiranosido (IPTG) y 80 ug/ml de 5-bromo-4-cloro-3-indolil- β -D-galactósido (X-gal). Las colonias blancas obtenidas se propagaron en medio líquido LB-ampicilina (2 mL). Dada la heterogeneidad del tamaño de los fragmentos se usó para obtener una cantidad homogénea de DNA plasmídico un sistema de amplificación de DNA circular llamado *rolling cycling*. Este sistema utiliza hexámeros al azar como *primers* y la enzima Phi29 DNA polimerasa con actividad *proofreading*. Brevemente, los hexámeros al azar se unen al DNA circular en sitios múltiples, y la enzima polimerasa extiende cada uno de esos fragmentos. Cuando la enzima encuentra otro fragmento extendido, se desplaza y sintetiza un fragmento monocatenario que será alcanzado por los hexámeros al azar y usado como templado para la siguiente ronda de amplificación. El proceso continúa, la amplificación resulta exponencial y se realiza a temperatura constante (amplificación isotérmica a 30 °C). Puede usar como templados la colonia bacteriana, cultivo líquido, o *glycerol stock*. Se obtienen 1- 1.5 μ g de producto final en 10 μ L (100-150 ng/ μ L) que se aplica directamente, sin necesidad de purificación posterior como templado de secuenciación rindiendo 5

reacciones de secuenciación por cada reacción de amplificación de DNA plasmídico.

Epidemiología molecular período 1990-2002:

Para el estudio de los polimorfismos del gen hexón, una vez obtenido el efecto citopático, se extrajo el ADN y se procedió a la caracterización molecular de la siguiente manera:

1. PCR genérica para confirmar el cultivo (Hierholzer *et al.*, 1993)
2. PCR especie específica (Xu *et al.*, 2000)
3. PCR-RFLP para determinar especie-*subespecie* y patrones de restricción enzimática (Kidd *et al.*, 1996)
4. PCR multiplex especie B:1 (Xu y Erdman, 2001)

Los productos se visualizaron en geles de agarosa al 2% (p/v) en buffer Tris-acetato-EDTA (TAE) teñidos con bromuro de etidio.

Se seleccionaron para este estudio sólo los Adenovirus serotipo 7, de diferentes años en el período 1990-2002.

Genotipificación de los AV7 del período 1990-2002

Se compararon las secuencias obtenidas entre sí y con las disponibles en bases de datos públicas (GenBank).

Se obtuvo una matriz de distancia con los componentes DNAdist para nucleótidos y Protdist para aminoácidos del *software* PHYLIP v3.5c (Felsestein, 1989), que permitió establecer homologías entre las cepas y se analizó la presencia de marcadores distintivos de acuerdo a la gravedad del paciente.

Se construyó un árbol por el método de *Maximun-Likelihood* que refleje las relaciones filogenéticas entre las muestras tanto en hexones completos como

en la región hipervariable. Se visualizó el mismo con el programa Treeview v1.5.2 (Page, 1996).

Análisis de los polimorfismos hallados en el gen hexón entre 1990-2002

Las secuencias obtenidas se alinearon de afuera hacia adentro en una estrategia de tipo *contig* y los polimorfismos se analizaron con el programa DNASP v3.5 (Rozas y Rozas, 1999). La diversidad se midió como el número total de sitios segregantes o polimórficos (S) dividido la longitud del fragmento (L) dividido el número de ciclos de PCR aplicados (S/L/ciclos). La diversidad media se calculó como el número total de sustituciones sobre el tamaño del amplicón (S/L) y se expresó como porcentaje (S/L*100). Se calculó κ como la relación entre transiciones y transversiones. El coeficiente ω se determinó como el cociente entre las mutaciones no sinónimas y las sinónimas (d_{ns}/d_s). Las secuencias de aminoácidos fueron inferidas usando el código universal y analizadas con el componente Protean del programa Lasergene. Se tomó en cuenta si los cambios de reemplazo de aminoácidos fueron conservativos o no conservativos con respecto a la carga. Se consideró una sustitución como fijada cuando estuvo presente en todo el período analizado (desde la caracterización del AV7h en 1987 hasta 2002) y fijada sólo en muestras cuando no estuvo presente en la cepa patrón de AV7h pero sí en todas las muestras analizadas desde 1990 hasta 1999 para el hexón completo y 1999 hasta 2002 para la región hipervariable del hexón.

Todas las comparaciones se realizaron con respecto a la cepa prototipo del serotipo 7 (7p) cepa Gomen. Se analizaron por separado muestras con secuenciación completa del hexón (1990-1999) y muestras con región hipervariable secuenciada (1999-2002), tomando como corte de regiones

hipervariables los nucleótidos 403 a 1356 y los aminoácidos 136 a 452 (numerados de acuerdo al hexón de la cepa AV7p Gomen).

Diseño de *primers* específicos:

Los *primers* específicos fueron diseñados con el programa Primer Select del paquete de *software* Lasergene (DNASTar Inc., Madison WI). Se tuvo en cuenta para el diseño la estabilidad del *primer* (que no tuviera estructura secundaria) y su composición, ya que influye en temperatura de *annealing* y auto-apareamiento en dímeros.

Hexón

A partir del esquema de los *primers* conocidos para secuenciación (Takeuchi *et al.*, 1999) se diseñaron *primers* específicos a partir de la secuencia completa disponible del Adenovirus 7 patrón cepa Gomen (Figura 21).

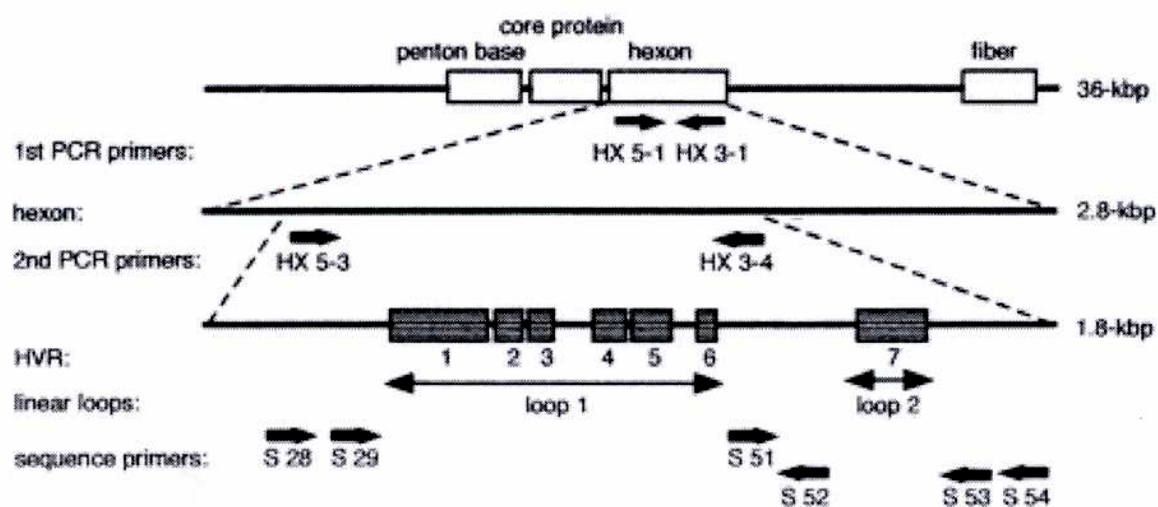


Figura 21: Esquema de la estrategia de secuenciación del gen hexón. Tomada de Takeuchi *et al.*, 1999.

Genoma AV7h

Las regiones del genoma que no pudieron ser secuenciadas por la estrategia de *shotgun* se secuenciaron usando productos de PCR generados con *primers* internos diseñados *ad hoc*.

Se tomaron como templados principales para el diseño de *primers* las secuencias del prototipo de AV7 cepa Gomen disponibles en las bases de datos públicas. Para completar el genoma se tomó como templado la secuencia completa del Adenovirus especie B, prototipo serotipo 11, cepa Slobitski (Ad11p) (Stone *et al.*, 2003). Se tomaron subregiones para completar lo diseñado sobre AV7 cepa Gomen.

Se muestran en la Tabla 7 las regiones abarcadas por secuencias disponibles de AV7 Gomen y las subregiones tomadas como molde de AV11p.

<i>Molde</i>	<i>GenBank</i>	<i>Región (pb)</i>
AV7 Gomen	X03000	0-10704
AV11p	AY163756	10600-13750
AV7 Gomen	AD001675	13684-15367
AV11p	AY163756	15100-18700
AV7 Gomen	AF065065	18115-21387
AV11p	AY163756	21000-27200
AV7 Gomen	AF104382	26578-31899
AV11p	AY163756	31300-34794

Tabla 7: Moldes tomados para el diseño de *primers* para amplificación y secuenciación del genoma del AV7h.

Amplificación por PCR y secuenciación

La amplificación por PCR fue realizada en termocicladores usando protocolo general (5 μ L de DNA, 1U *Taq* polimerasa, $MgCl_2$ 1.5 mM, KCl 50 mM, dNTPS 200 μ M, *primers* 0.2 μ M, 10 mM Tris-HCl pH8.3 concentración final) por 30 ciclos de 95 °C 20 seg, 50 °C 20 seg y 72 °C 1min. Los productos se visualizaron en geles de agarosa al 2% (p/v) en buffer TAE teñidos con bromuro de etidio.

Los productos de amplificación se purificaron con columnas comerciales y las reacciones de secuencia fueron realizadas de acuerdo al método de Sanger (Sanger *et al.*, 1977) con cada uno de los *primers* que correspondiere a una concentración de 10 μ M y los productos marcados con distintos fluorocromos se analizaron en secuenciador automático por electroforesis capilar. Se marcaron 300-500 ng de DNA para las regiones hipervariables y fragmentos genómicos, 800ng para los hexones completos y 1,2 ug para los insertos clonados por 30 ciclos de acuerdo al protocolo de secuenciación (95 °C 20 seg, 50 °C 20 seg y 60 °C 1min). Las muestras se inyectaron 112 seg a 2 kV y se corrieron por 5 hs a 6 kV.

Hexón

Los AV que resultaron tipificados como AV7 y la cepa patrón de AV7h (87-922) se analizaron por secuenciación para determinar su genotipo mediante una PCR de 2,8 kb (*primers* 5-1 y 3-1) con un sistema comercial que utiliza una mezcla de *Taq* DNA polimerasa y *Pfu* DNA polimerasa permitiendo tener un mayor rendimiento de la reacción e incorporando la actividad exonucleasa

3'5' (*proofreading*) (Figura 21). Para el análisis de los polimorfismos se llevó a cabo la amplificación de 953 pb. correspondiente a regiones hipervariables del gen hexón que permiten separar los AV7 en dos *clusters* genómicos: *Cluster* de tipo genómico 1 que incluye Ad7p y Ad7p1 y *Cluster* de tipo genómico 2 que incluye Ad7b Ad7c Ad7d Ad7g y Ad7h (Crawford-Mikszá y Schnurr, 1996 y Li y Waddell, 1999). Una vez obtenido el producto de PCR, se llevó un volumen de 200uL que luego se purificó con columnas de silica comerciales directamente o por recuperación de banda de gel agarosa de bajo punto de fusión al 1% (p/v) si se visualizaban productos no específicos adicionales. La imagen se digitalizó y se cuantificó con el *software* AlphaEase v5.04 (Alpha Innotech Corporation, San Leandro, CA, USA) por densitometría comparativa de la banda de gel de agarosa al 2% teñido con bromuro de etidio tomando como referencia marcadores comerciales de peso molecular y de masa conocidos.

Alternativamente se planificó una estrategia de clonado del producto de PCR obtenido en un vector pGemT (Zhou M. *et al.*, 1995). Para el tamaño del inserto de PCR (2,8 kb) fue necesario ligar 140 ng de inserto con 50 ng de vector linearizado. Esta estrategia se utilizó pensando en la conservación del material original de las muestras más viejas (1990-1999). Se transformaron y plaquearon bacterias *E. coli* JM109 (endA⁻) competentes en medio Luria-Bertani (LB) con ampicilina (sal de sodio 100ug/ mL), 0.5mM isopropil-β-D-tiogalactopiranosido (IPTG) y 80 ug/ml de 5-bromo-4-cloro-3-indolil-β-D-galactósido (X-gal). Las colonias blancas obtenidas se crecieron en medio líquido LB-ampicilina (25 mL) y el DNA plasmídico se obtuvo por una preparación de volumen medio en columnas comerciales. Se verificó la

presencia del inserto por corte con *EcoR* I 1h a 37°C y resolución en gel de agarosa al 1% teñido con bromuro de etidio en buffer TAE. Se aplicó el mismo protocolo de cuantificación que el usado para los insertos de PCR.

Genoma AV7h

Secuenciación de los clones obtenidos de la biblioteca genómica

El DNA plasmídico se secuenció en equipo automático capilar usando *primers* universales (*forward* y *reverse* de M13, T7 y Sp6). De acuerdo a los procedimientos estándar de la estrategia de *shotgun*, para tener certeza de los resultados obtenidos sería necesario secuenciar 375 clones con insertos de 800 pb cada uno (Sambrook y Russell, 2001).

PCR para genoma AV7h

Se estandarizó un protocolo general de PCR y se obtuvieron fragmentos desde 400 a 1200 pb. se llevó un volumen de 200uL que luego se purificó con columnas de silica comerciales directamente o por recuperación de banda de gel agarosa de bajo punto de fusión al 1% (p/v) si se visualizaban productos no específicos adicionales. La imagen se digitalizó y se cuantificó de la misma manera que lo explicado anteriormente para el hexón.

Análisis de secuencias del Genoma de AV7h

Los electroferogramas obtenidos se visualizaron con el programa Chromas (v 1.45, 32-bit) de la Universidad de Griffith (Queensland, Australia), se alinearon con el programa ClustalX v1.83 (Thompson *et al.*, 1997) y se analizaron con los distintos componentes del paquete Lasergene (DNASTar Inc., Madison WI) y Phylip v 3.5c (Felsenstein, 1989).

Datos de biblioteca genómica

Para el ensamblado de las secuencias obtenidas se adoptó una estrategia de tipo *contig* donde se alinean los fragmentos secuenciados teniendo en cuenta la calidad de la secuencia y la contaminación con secuencias provenientes del vector si las hubiera. Se determinaron las regiones faltantes y se solucionaron regiones de conflicto secuenciando con *primers* específicos.

Análisis de las secuencias nucleotídicas

Se determinó la longitud total del genoma, se calculó la frecuencia de las bases y el porcentaje de G+C, se ubicaron los marcos abiertos de lectura (ORF), regiones intergénicas, promotores, señales de *splicing* y poliadenilación por los programas ORF Finder (NCBI) y GeneMark v2.4. (Borodovsky y McIninch, 1993).

Deducción de la secuencia de aminoácidos de las proteínas virales

Se utilizó el código universal para inferir la secuencia de aminoácidos de acuerdo a los marcos abiertos de lectura predichos.

Análisis de recombinación

Se analizó la presencia de variantes intertípicas generadas por recombinación utilizando secuencias de referencia y los programas Simplot v2.4 y RDP v1.07 (Lole *et al.*, 1999; Martin y Rybicki, 2000).

Antigenicidad y homología con proteínas pulmonares

Se predecirá para las proteínas hexón, pentón, fibra y E1A los potenciales determinantes antigénicos (epitopes B y T) por métodos basados en la estructura tridimensional de la proteína y localización de regiones de mayor hidrofiliidad con la ayuda del programa Protean del *software* Lasergene.

Se buscarán en GenBank las proteínas de mayor homología con el péptido predicho, que puedan funcionar como mimotopes de proteínas virales generando autoanticuerpos.

Resultados

Aspectos clínicos y epidemiológicos de las infecciones por Adenovirus en niños

Muestras período 1999-2002

En el transcurso del período 1999-2002 se analizaron un total de 15103 muestras, de las cuales 5334 resultaron positivas para alguno de los virus analizados (AV, Virus Sincicial Respiratorio, Parainfluenza 1-2-3, e Influenza A y B). Se registraron 362 muestras positivas para AV. Las frecuencias de los AV fueron 2,42% en 1999, 1.95% en 2000, 1.56% en 2001 y 3.56% en 2002. La prevalencia resultó 9.24% en 1999, 4.8% en 2000, 4.1% en 2001 y 9.73% en 2002 (Tabla 8).

Tabla 8: Resumen de muestras. Modificado de Viegas *et al.*, 2003

Año	Muestras totales	Etiología viral	Adenovirus
1999	4008	1050 (26.2)	97 (9.2)
2000	3788	1542 (40.7)	74 (4.8)
2001	3458	1326 (38.3)	54 (4.1)
2002	3849	1416 (36.8)	137 (9.7)

En general la prevalencia es menor que otros agentes etiológicos virales (RSV 78.8%, Influenza A 8.3%), ocupando los Adenovirus el tercer puesto (7.8%) en un análisis generalizado realizado en nuestro laboratorio (Viegas *et al.*, 2003). Allí también pudimos observar que el aumento en el número de casos de AV en 1999 y en 2002 ($\chi^2 = 33,19$; $p < 0,00001$) se debió probablemente a la co-ocurrencia de una epidemia de sarampión en 1998-9 que disminuyó el número de casos de RSV en 1999 y una campaña masiva de vacunación anti-influenza en 2001 que probablemente provocó la desaparición de casos en 2002, alterando así el orden en los patrones de prevalencia que normalmente hemos observado.

De los pacientes analizados, 45.01% provenían del Hospital de Niños Ricardo Gutiérrez y el resto de otras Instituciones municipales y de la Provincia de Buenos Aires.

En el período analizado se detectaron más pacientes con AV de sexo masculino (56.6%) que femenino (43.4%).

El diagnóstico fue realizado en promedio a los 2 días post-internación aunque el rango de detección varió desde día 1 al día 90 (N=230).

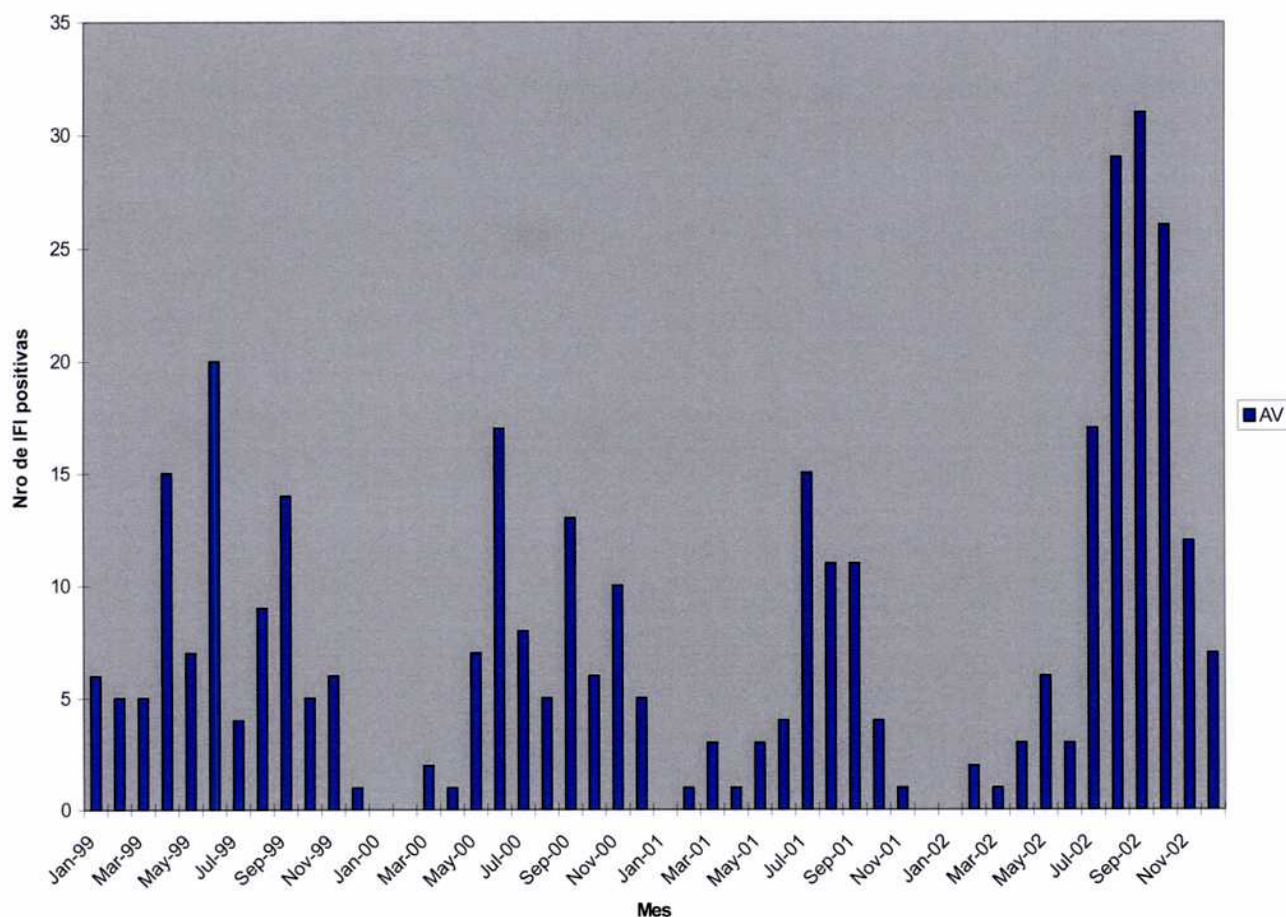
Se registró dificultad respiratoria moderada en el 65.5% de los casos, grave en el 22.1% y leve en 12.4 pacientes (N=290).

Si bien se detectó infección por AV a lo largo del año (Gráfico 1), la mayor ocurrencia se observó durante los meses de invierno seguido por los meses cálidos (fines de octubre y noviembre).

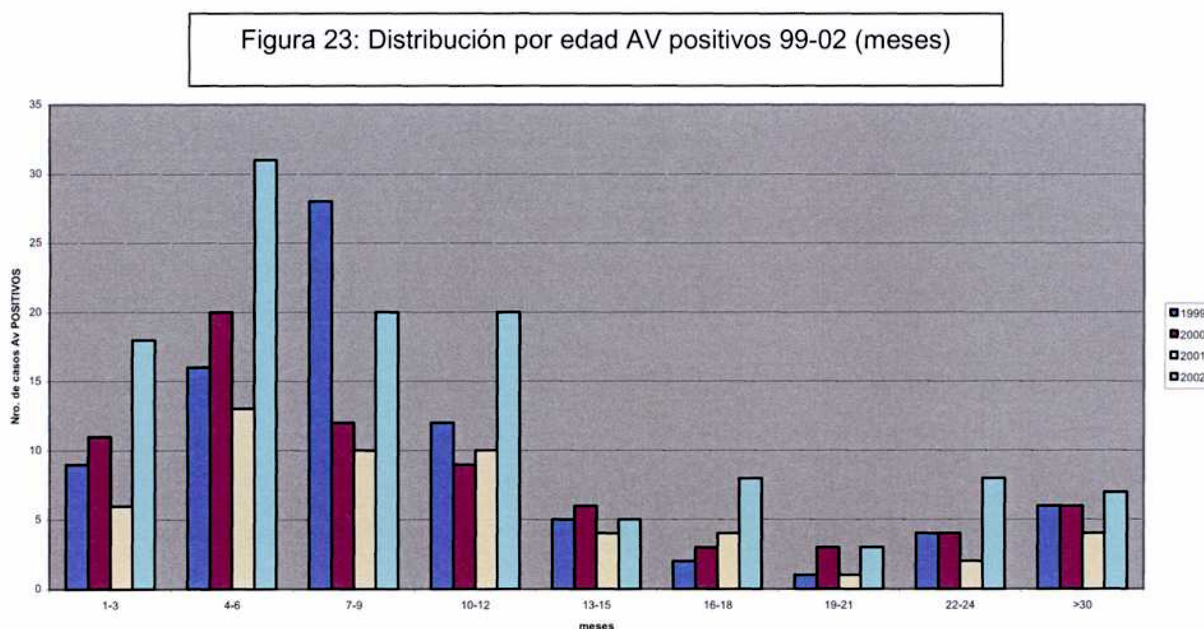
Los Adenovirus, como se puede observar en la Figura 22, no presentan una estacionalidad marcada a comparación de otros virus como el Sincicial Respiratorio siendo su patrón carácter endémico más que epidémico en

general. Se registró un mayor número de casos en el período invernal de baja temperatura. Se analizará con mayor detenimiento la estacionalidad en la sección de datos ambientales.

Figura 22: Distribución por mes AV positivos 99-02 (meses)



La distribución por edades de los pacientes infectados con Adenovirus se puede ver en el Figura 23. La mediana de las edades fue de 1.5 meses en 1999, 2 meses en 2000 y 2001 y 3 meses en 2002 (rango desde 1 mes hasta 6 años).



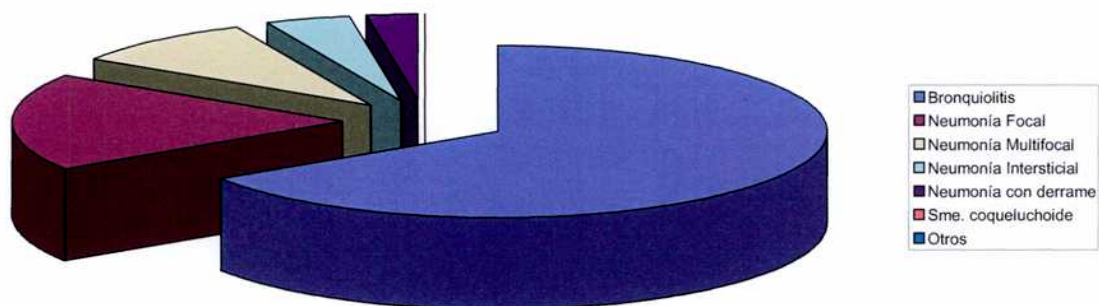
Las causas de internación pueden observarse en la Figura 24. La mayor causa de internación fue bronquiolitis (189 pacientes) seguida por neumonía focal (51 pacientes), neumonía multifocal (24 pacientes), neumonía intersticial (11 pacientes), neumonía con derrame (11 pacientes) y síndrome coqueluchoide (3 pacientes). Otros motivos de internación fueron catarro de vías aéreas superiores, displasia bronco-pulmonar, neumonitis, distress respiratorio, fístula pleuro-pulmonar, supuración pleuro-pulmonar epiglottitis, shock séptico, faringitis y otras. Se registraron enfermedades de base como leucemia linfoblástica aguda, linfangioma, síndrome de Down, Kawasaki, Fibrosis quística y Varicela.

Diez casos presentaron infección mixta, 7 con Virus Sincicial Respiratorio y 3 con Virus Parainfluenza 3

Nueve casos resultaron fatales, tres con nexo epidemiológico (a partir de una infección intrahospitalaria) y uno con enfermedad pulmonar secuelar.

Gráfico 24: Causa de Internación por AV muestras 99-02

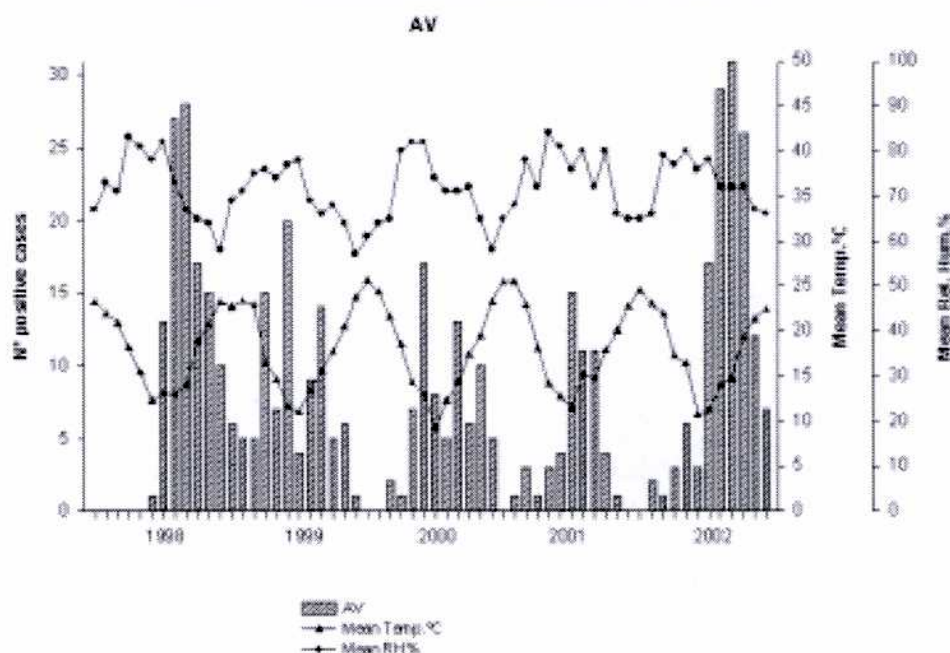
N=321



Datos ambientales

Se recopilaron datos disponibles sobre número de casos mensuales de Adenovirus desde 1998 hasta 2002. Se graficaron juntamente con los datos de temperatura y humedad relativa % promedio mensual (Figura 25)

Figura 25: Relación entre datos meteorológicos y frecuencias de AV 1998-2002
Modificado de Viegas *et al.*, 2003.



Para evaluar la correlación entre infección por Adenovirus y los distintos parámetros meteorológicos se realizó un test de correlación de Spearman divariado con respecto a la temperatura media, humedad relativa % y radiación solar UVB (Tabla 8).

Tabla 8: AV- Coeficientes de correlación de Spearman
Modificado de Viegas *et al.*, 2003.

Coeficientes

<i>Temperatura media</i>	-0.41 ($p = 0.0021$)
<i>Humedad relativa (%)</i>	-0.1 ($p = 0.4892$)
<i>UVB</i>	-0.32 ($p = 0.0172$)

Se encontró una correlación negativa de grado medio significativa con respecto a la temperatura media, coincidiendo con lo deducido gráficamente en la sección anterior. En cuanto a la humedad relativa % la correlación resultó ser levemente negativa pero no significativa. De la misma manera la correlación negativa fue no significativa para UVB. Todo tiende a evidenciar que la naturaleza del virus no se ve afectada por estos parámetros climáticos dado su patrón anual ininterrumpido de circulación y señala como principal causante de los brotes a la susceptibilidad de los huéspedes a condiciones ambientales de temperatura desfavorable como los registros bajos (invierno) y los cambios bruscos (primavera).

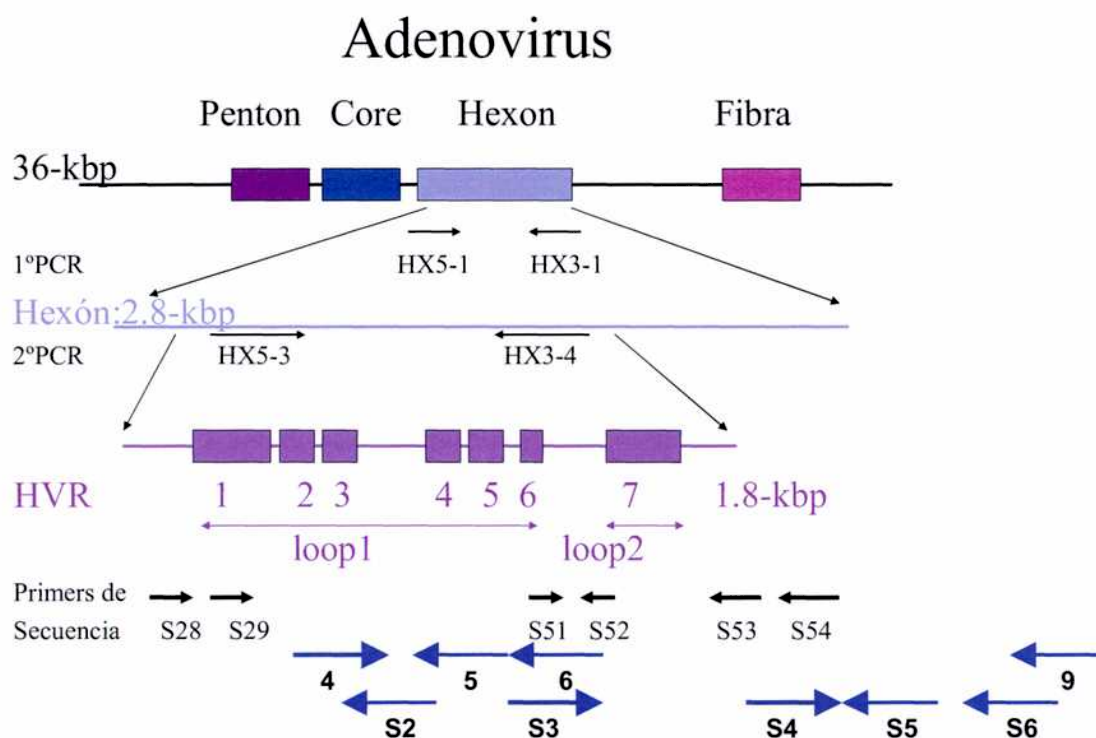
Datos moleculares

Secuenciación del hexón del AV7h

Diseño de primers específicos:

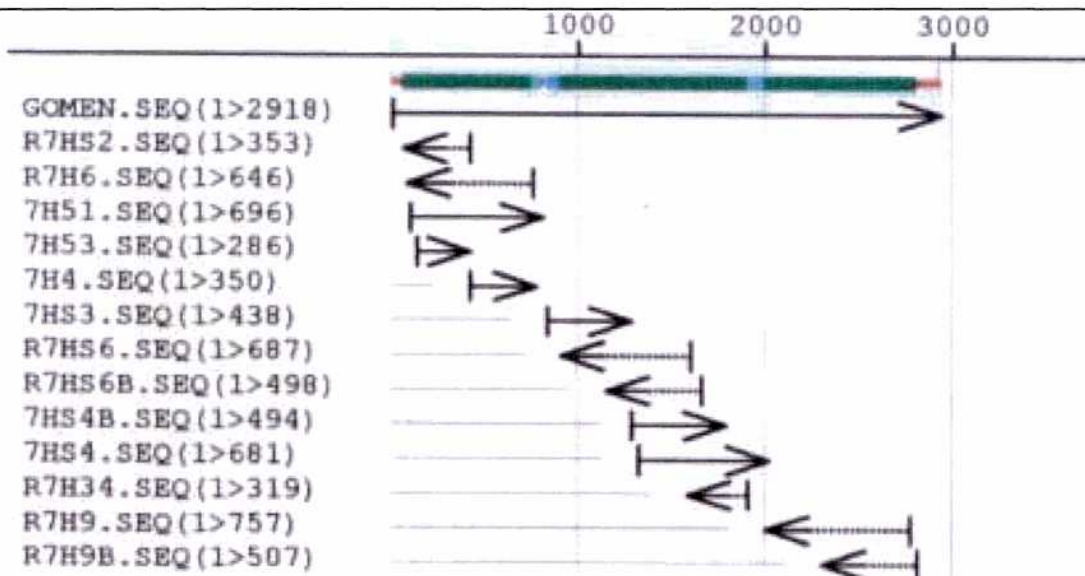
Se diseñaron *primers* específicos con el programa Primer Select del paquete de *software* Lasergene (DNASTar Inc., Madison WI) a partir de la secuencia completa disponible del Adenovirus 7 patrón cepa Gomen y la PCR se generalizó en treinta ciclos a 94 °C por 15 seg, 50 °C por 1 min y 72°C por 2 min con una extensión final a 72°C por 10 min. El esquema gráfico ampliado de los *primers* usados se ve en la Figura 26.

Figura 26: Primers usados para amplificación y secuenciación. Los *primers* señalados en negro fueron tomados de Takeuchi *et al.*, 1999; en azul se señalan los *primers* diseñados *de novo* para Av7h.



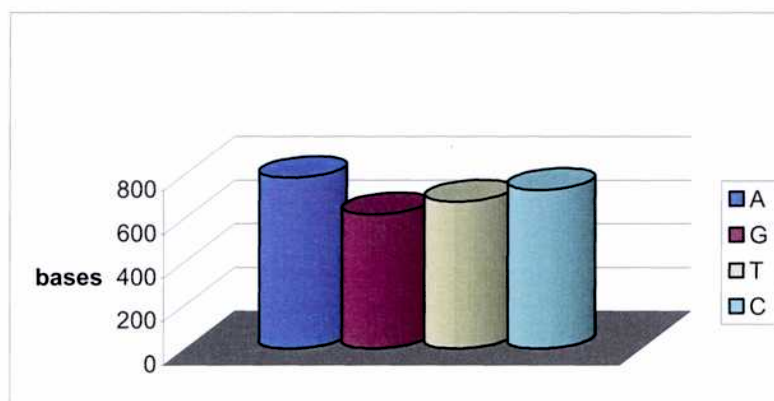
La cepa patrón de Adenovirus 7h (87-922) fue secuenciada íntegramente (2808 nt) siguiendo la estrategia de “afuera hacia adentro” desde las regiones más conservadas hacia las más variables en una “caminata” a lo largo de los diferentes *primers* utilizados (Figura 27).

Figura 27: Esquema del armado de la secuencia del gen hexón Av7h.
Las flechas indican la dirección de los *primers*



Se dedujo la secuencia de aminoácidos de 2808 bases cuya composición fue A+T=52.03% y G+C= 47.97% sin bases ambiguas (Figura 28).

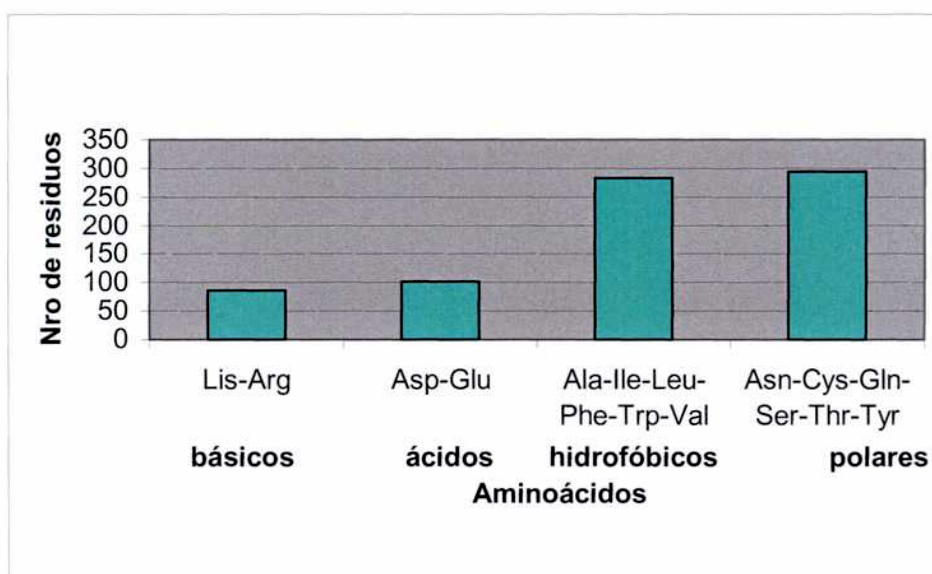
Figura 28: Composición nucleotídica hexón AV7h



Predicción del perfil proteico

De los 934 aa se predice una proteína estable de 105,28 kDa con un pI teórico de 5.43 de fórmula: $C_{4712}H_{7129}N_{1249}O_{1427}S_{36}$ cuya composición se ve en la Figura 29.

Figura 29: Composición aminoacídica hexón AV7h



En cuanto al uso del codon, se detectó un leve sesgo mayormente en sitios sinónimos siendo los codones NNC altamente preferidos salvo en el caso de los aminoácidos Ile (I), Leu (L) y Val(V). Se conoce que la presencia de

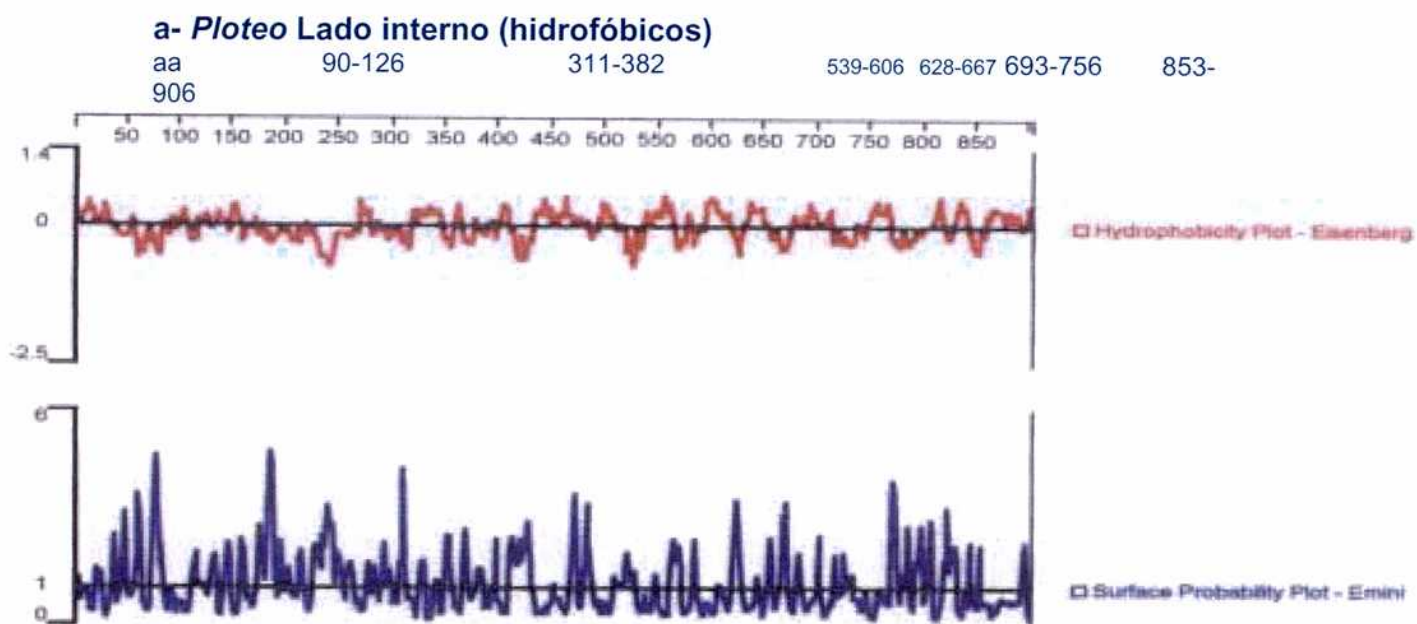
purinas (A-G) y pirimidinas (T-C) en la segunda posición se relacionan respectivamente con hidrofiliidad e hidrofobicidad. Se usaron, en distinta medida, todos los codones disponibles (61 tRNA disponibles) para la traducción (Tabla 9).

Tabla 9: Usos del codón hexón AV7h

UUU-F	18 (0,82)	UCU-S	16 (1,55)	UAU-Y	15 (0,55)	UGU-C	1 (0,50)
UUC-F	26 (1,18)	UCC-S	15 (1,45)	UAC-Y	40 (1,45)	UGC-C	3 (1,50)
UUA-L	2 (0,18)	UCA-S	9 (0,87)	UAA-*	0 (0,00)	UGA-*	0 (0,00)
UUG-L	7 (0,62)	UCG-S	4 (0,39)	UAG-*	0 (0,00)	UGG-W	13 (1,00)
CUU-L	24 (2,12)	CCU-P	18 (1,22)	CAU-H	4 (0,67)	CGU-R	5 (0,71)
CUC-L	7 (0,62)	CCC-P	15 (1,02)	CAC-H	8 (1,33)	CGC-R	7 (1,00)
CUA-L	3 (0,26)	CCA-P	22 (1,49)	CAA-Q	16 (1,07)	CGA-R	2 (0,29)
CUG-L	25 (2,21)	CCG-P	4 (0,27)	CAG-Q	14 (0,93)	CGG-R	3 (0,43)
AUU-I	19 (1,54)	ACU-T	24 (1,41)	AAU-N	32 (0,94)	AGU-S	7 (0,68)
AUC-I	11 (0,89)	ACC-T	24 (1,41)	AAC-N	36 (1,06)	AGC-S	11 (1,06)
AUA-I	7 (0,57)	ACA-T	15 (0,88)	AAA-K	22 (1,13)	AGA-R	18 (2,57)
AUG-M	32 (1,00)	ACG-T	5 (0,29)	AAG-K	17 (0,87)	AGG-R	7 (1,00)
GUU-V	15 (1,22)	GCU-A	22 (1,54)	GAU-D	31 (1,02)	GGU-G	9 (0,57)
GUC-V	8 (0,65)	GCC-A	21 (1,47)	GAC-D	30 (0,98)	GGC-G	23 (1,46)
GUA-V	4 (0,33)	GCA-A	10 (0,70)	GAA-E	22 (1,19)	GGA-G	21 (1,33)
GUG-V	22 (1,80)	GCG-A	4 (0,28)	GAG-E	15 (0,81)	GGG-G	10 (0,63)

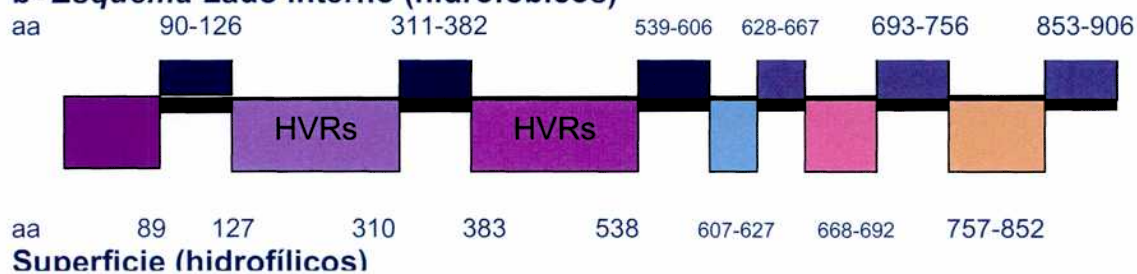
Se localizaron las regiones del gen descriptas y se analizaron los perfiles de hidrofobicidad e hidrofiliidad y se observó que, conforme a la literatura, las regiones hipervariables (HVRs) contienen residuos hidrofílicos (Figura 30).

Figura 30: Esquema del perfil hidropático del Hexón de AV7h.
a- Ploteo de hidrofobicidad. b- Esquema del ploteo



aa 89 127 310 383 538 607-627 668-692 757-852
Superficie (hidrofílicos) Comparación con secuencias reportadas

b- Esquema Lado interno (hidrofóbicos)



Referencias:



hidrofóbicos

Pedestal 1
Pedestal 2
Conector del pedestal

hidrofílicos

Loop amino terminal
Loop 1
Loop 2
Loop 3
Loop 4



Se comprobó la presencia de residuos hidrofóbicos en las regiones de aa90-126, 311-382, 539-606, 628-667, 693-756 y 853-906. A su vez se vió que las regiones hidrofílicas fueron aa1-89,127-310 (HVR), 383-538 (HVR), 607-627, 668-692 y 757-852. Las regiones hidrofílicas presentan mayor exposición en superficie y resultan portenciales antígenos.

Se comparó la secuencia del hexón del AV7h obtenida de la cepa 87-922 en nucleótidos y aminoácidos con las siguientes cepas

<i>Cepa</i>	<i>Acceso GenBank</i>	<i>Características</i>
Ad7p (Gomen) 2811nt	X76551	secuencia completa=
Ad7h	U76956	HVR=1419 nt

Observamos que a nivel nucleotídico la cepa Gomen (Ad7p X76551) y la secuencia obtenida de Ad7h (87-922) difieren en un 4% y a nivel aminoacídico en un 3.5% (Tabla 9). Esto es esperable ya que pertenecen a distintos GTC, el Ad7p al GTC1 y el Ad7h al GTC2. Vemos también que la secuencia obtenida del Ad7h completa difiere de la reportada en las HVR (Ad7h HVR X76551) en 1.57% a nivel nucleotídico (22 nt) y 2.22% a nivel aminoacídico (9 aa) permaneciendo en el mismo GTC. La cepa que se usó en este estudio era la original, que no había sido repicada desde su aislamiento en 1987.

	X76551	87-922	U76956
X76551	0.00000	0.0400	0.0474
87-922	0.03534	0.0000	0.0157
U76956	0.05301	0.02211	0.00000

Tabla 9 Matriz de distancia, se muestra en negro la matriz correspondiente a aminoácidos y en color a nucleótidos.

X76551	VAFKMATPSM	MPQWAYMHIA	GQDASEYLSP	GLVQFARATD	TYFSMGNKFR	NPTVAPTHDV	
87-922	---						
U76956	---						
X76551	TTDRSQRLML	RFVPVDREDN	TYSYKVRYTL	AVGDNRVLDM	ASTFFDIRGV	LDRGPSFKPY	
87-922							
U76956					---Y.....		
X76551	SGTAYNSLAP	KGAPNTSQWI	VTAGEERAVT	TTNTTFGIAS	MKGDNITKEG	LEIGKDITAD	HVR 1/A1
87-922			..T..D---N	A..Y.....	T.....		
U76956			..T..D---N	A..Y.....	T.....D		
X76551	NKPIYADKTY	QPEPQVGEE	WTDTDGTNEK	FGGRALKPAT	KMKPCYGSFA	RPTNIKGGQA	HVR 2/A2
87-922			..I.....			..T..T....	
U76956			..I.....				
X76551	KNRKVKPTG	DVETEEDID	MEFFDGREAA	DAFSPEIVLY	TENVNLETPD	SHVVYKPGTS	HVR 3/A3
87-922		N.....	..S.....				
U76956		..T.....	..A.....				
X76551	DDNSHANLGQ	QAMPNRPNYI	GFRDNFVGLM	YYNSTGNMGV	LAGQASQLNA	VVDLQDRNTE	HVR 4/A4
87-922	.G.....						
U76956	.G.....						
X76551	LSYQLLLDSL	GDRTRYFSMW	NQAVDSYDPD	VRIENHGIE	DELPNYCFPL	DGIGPAKTYQ	HVR 5/B1
87-922		..S.....		V.		G.K..	
U76956		..S.....		V.		GNK..	
X76551	GIKSKDNGWE	KDDNVKSNE	IAIGNNQAME	INIQANLWRS	FLYSNVALYL	PDVYKYTPTN	
87-922	...PR.TA..	..TK..TA..	L...				
U76956	...PR.TA..	..TK..TA..	L...				
X76551	ITLPANTNTY	EYMNGRVVSP	SLVDSYINIG	ARWSLDPMDN	VNPFNHHRNA	GLRYRSMLLG	HVR 6/B2
87-922					...D..D.		
U76956							
X76551	NGRYVPFHIQ	VPQKFFAVKN	LLLLPGSYTY	EWNFRKDVNM	VLQSSLGNDL	RTDGATISFT	
87-922							
U76956							
X76551	SINLYATFFP	MAHNTASTLE	AMLRNDTNDQ	SFNDYLSAAN	MLYPIPANAT	NIPISIPSRN	
87-922			..G.....	..K.....			
U76956							
X76551	WAAFRGWSFT	RLKTKETPSL	SGGFDPPYFVY	SGSIPYLDGT	FYLNHTFKKV	SIMFDSSVS	
87-922	.PP.....		R.				
U76956							
X76551	PGNDRLLSPN	EFEIKRTVDG	EGYNVAQCNM	TKDWFLVQML	ANYNIGYQGF	YIPEGYKDRM	
87-922							
U76956							
X76551	YSFFRNQPM	SRQVDEVNY	TDYKAVTLPY	QHNSGFGVGY	LAPTRMQGEP	YPANYPYPLI	HVR 7/D
87-922							
U76956							
X76551	GTTAVKSVTQ	KKFLCDRTMW	RIPFSSNFMS	MGALTDLGQN	MLYANSAHAL	DMTFEVDPM	
87-922					L.....		
U76956							
X76551	EPTLLYLLFE	VFDVVRVHQP	HRGVIEAVYL	RTPFSAGNAT	T		
87-922							
U76956							

Figura 31: Comparación de las secuencias inferidas de aminoácidos Cepas Gomen (Ad7p X76551) y 87-922 (Ad7h HVR X76551). Los puntos (.) significan igual residuo, los guiones (-) no hay secuencias disponible y los cambio de aminoácidos están

Se muestran en la Figura 31 las diferencias en aminoácidos obtenidas entre las cepas. Se hace especial hincapié en la localización de las regiones hipervariables del gen hexón. La nomenclatura y localización de las mismas varía según el autor. Se localizan HVR1-6 en el loop 1 y HVR 7 en el loop 2 para Crawford-Mikszsa (1996), Takeuchi y Li & Waddell (1999) pero cambia cuando Li *et al.* (1997) definen el perfil hidropático de los hexones adenovirales y se nombran A1-A4 en loop1, B1 y 2 en loop 2 y D en loop4.

Si graficamos la diversidad en nucleótidos P_i , tomada como el número promedio de diferencias por sitio con respecto a la cepa prototipo del serotipo 7 (Gomen) avanzando de a 10 nucleótidos en ventanas de 50 bases de longitud vemos que la mayor diversidad se encuentra localizada en zonas definidas como hidrofílicas, ubicadas en los loops del hexón hacia la superficie (Figura 32 y Tabla 10).

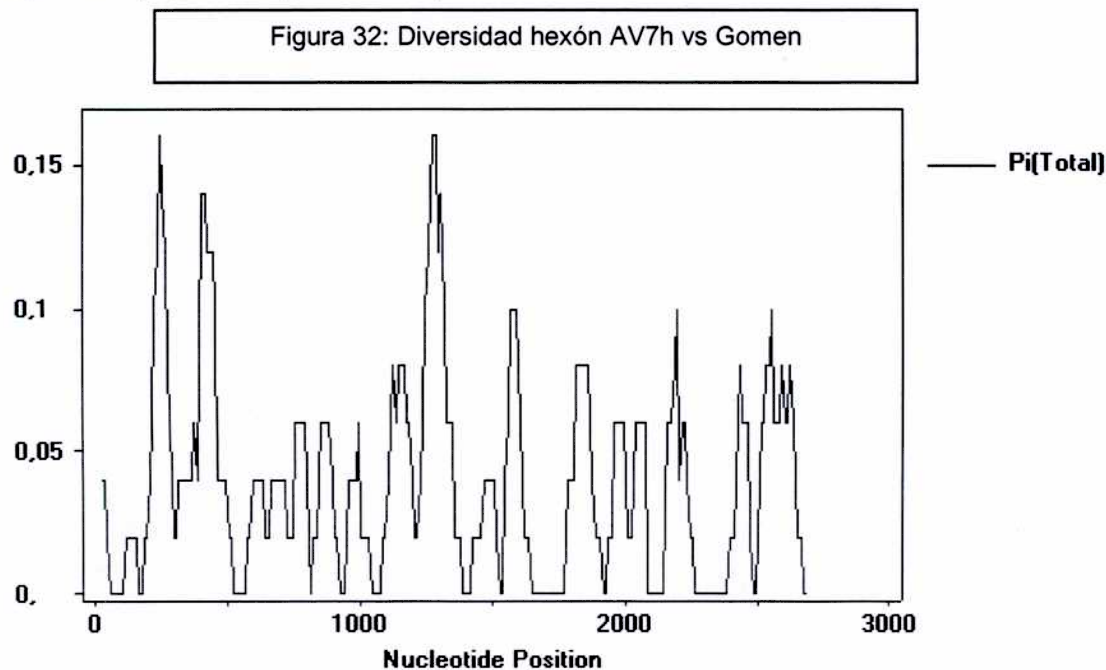


Tabla 10: Localización de picos de Diversidad hexón AV7h vs Gomen

Ventana	Punto medio	Pi	Región (aa)
201-250	225	0,10000	75-85
211-260	235	0,12000	Loop
221-270	245	0,16000	amino
231-280	255	0,14000	terminal
241-290	265	0,12000	
381-430	405	0,14000	135-151
391-449	415	0,14000	Loop 1
401-459	425	0,12000	
411-469	444	0,12000	
421-479	454	0,10000	
1220-1269	1244	0,10000	414-438
1230-1279	1254	0,12000	Loop 2
1240-1289	1264	0,14000	
1250-1299	1274	0,16000	
1260-1309	1284	0,16000	
1270-1319	1294	0,12000	
1280-1329	1304	0,14000	
1290-1339	1314	0,10000	
2160-2209	2184	0,08000	728-731
2170-2219	2194	0,10000	Loop 4

Regiones antigénicas

Los epitopes B y T fueron predichos y se pueden ver en la Figura 33. El índice de antigenicidad que los define integra cargas y polaridad para predecir los epitopes T.

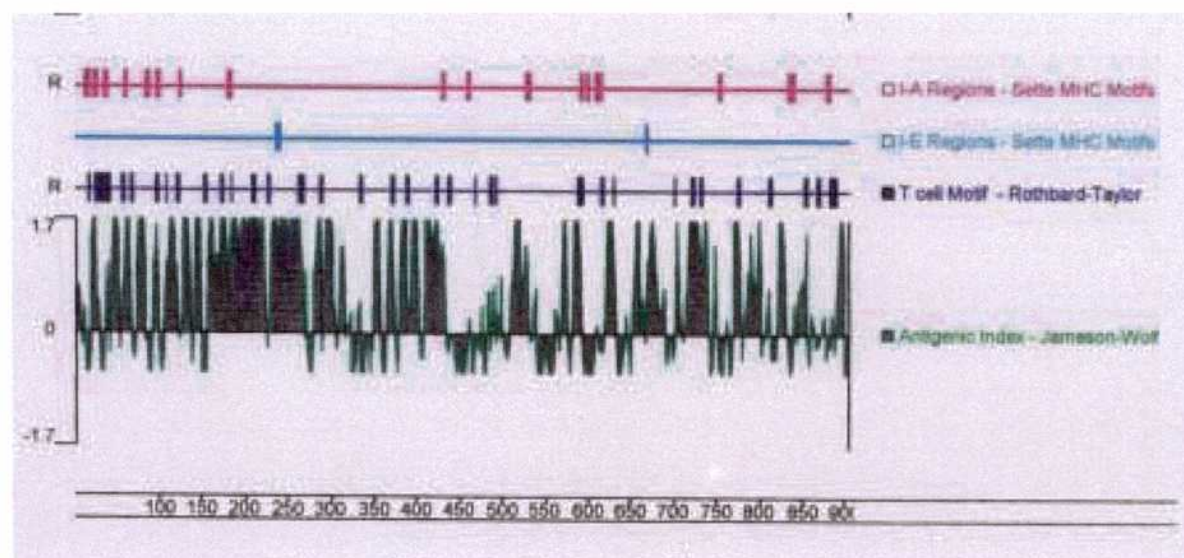


Figura 33: Localización gráfica de las regiones antigénicas Hexón AV7h

Epitopes B

Hexapéptidos: aa12-17, 21-26, 33-38, 57-62, 83-88, 95-100, 121-127, 178-183, 427-432, 456-461, 526-531, 589-594, 596-601, 750-755.

Octapéptido: aa876-883

Nonapéptido: aa833-841

Decapéptido: aa 606-615

Epitopes T

Tetrapéptidos: aa15-18, 23-26, 28-31, 106-109, 141-154, 182-185, 208-211, 225-228, 369-372, 386-389, 466-469, 485-488, 490-493, 628-631, 699-702, 719-722, 730-733, 773-776, 810-813, 851-855

Pentapéptidos: aa56-60, 64-68, 119-123, 170-174, 286-290, 333-337, 420-424, 434-438, 588-592, 614-618

Heptapéptidos: 93-99, 262-268

Octapéptidos: aa31-41, 882-889

En algunos casos los epitopes B y T predichos colocalizan o se solapan como se muestra en la tabla 11.

Tabla 11: Localización de los epitopes predictivos para Hexón AV7h. se muestra resaltado en verde la colocación de epitopes B y T. Se indica la localización en el hexón en aa y en estructura.

	<i>Epitopes B</i>	<i>Epitopes T</i>	<i>Epitopes B</i>	<i>Epitopes T</i>	
Loop amino terminal	12-17	15-18		434-438	Loop 2
	21-26	23-26	456-461		
		28-31		466-469	
	33-38	31-41		485-488	
	57-62	56-60		490-493	
Pedestal 1		64-68	526-531		Pedestal 1
	83-88			588-592	
	95-100	93-99	589-594		
		106-109	596-601		
Loop 1	121-127	119-123	606-615		conector Pedestal 2
		141-154		614-618	
		170-174		628-631	
	178-183			699-702	Pedestal 2
		182-185		719-722	
		208-211		730-733	
		225-228	750-755		Loop 4
Pedestal 1		262-268		773-776	
		286-290		810-813	
		333-337	833-841		Pedestal 2
Loop 2		369-372		851-855	
		386-389	876-883		
		420-424		882-889	
	427-432				

Polimorfismos del gen hexón en muestras clínicas: 1990-2002

Cultivo

Se aislaron 273/362 muestras positivas para AV en el período 1999-2002 en células 293 o Hep-2 (75.4%). Cabe destacar que 10 muestras con IFI negativa resultaron positivas por cultivo y PCR confirmatoria para AV. Se realizó un testeo adicional de los anticuerpos monoclonales comerciales usados para diagnóstico en cultivo con 38 muestras por duplicado positivos para PCR genérica.

Tabla 12: Evaluación de Mabs comerciales para AV

<i>Cultivo</i>	<i>Coincidencia (+)</i>	<i>Coincidencia (-)</i>	<i>No coincidencia</i>
Positivo	19	3	16
Negativo	1	0	4
Totales	20	3	20

Se vió que sólo el 50 % de los casos concuerdan en el diagnóstico, evidenciando que los epitopes usados comercialmente para detección de Adenovirus pueden no detectar algunos serotipos. Algunas de estas muestras no concordantes se caracterizaron como AV7 posteriormente por PCR.

Epidemiología molecular:

Se determinó semicuantitativamente mediante la PCR genérica, por comparación entre la banda obtenida a partir del DNA extraído del aspirado nasofaríngeo y la obtenida del DNA extraído del cultivo celular, si el título viral había aumentado, condición que resultó indispensable para la posterior caracterización molecular donde la sensibilidad del método predomina sobre

la especificidad. La PCR especie específica nos permitió distinguir entre las 6 especies de Adenovirus (A-F) y verificar los de mayor importancia en infecciones respiratorias agudas pediátricas. La determinación del especie-*subespecie* nos permitió separar los AV B:1 (AV 3,7,16 y 21) de los AV B:2 (AV 11,14,34 y 35). Finalmente la PCR especie B:1 diferenció los AV serotipo 3,7 y 21.

Se analizaron por PCR 312 muestras por uno o más de los métodos utilizados resultando 252 positivas (80.7%) (Figura). En algunos casos se procesaron aspirados nasofaríngeos sin pasaje por cultivo. Los resultados fueron

PCR especie específica: 41 Especie C y 16 Especie B (Figura 34 y 35)

PCR-RFLP para determinar especie-subespecie: 96 B:1 y 10 B:2

PCR multiplex especie B:1: 149 AV7 y 38 AV3.

Figura 34: PCR especie específica. Calle 1: marcador de 100 pb, calles 2 a 9, 11, 13 a 18 producto a cortar con Taq I, calles 10 y 12 especie B:2 (290 pb), calle 19 control negativo y calle 20 control positivo

1 * 2 * 3 * 4 * 5 * 6 * 7 * 8 * 9 * 10 * 11 * 12 * 13 * 14 * 15 * 16 * 17 * 18 * 19 * 20

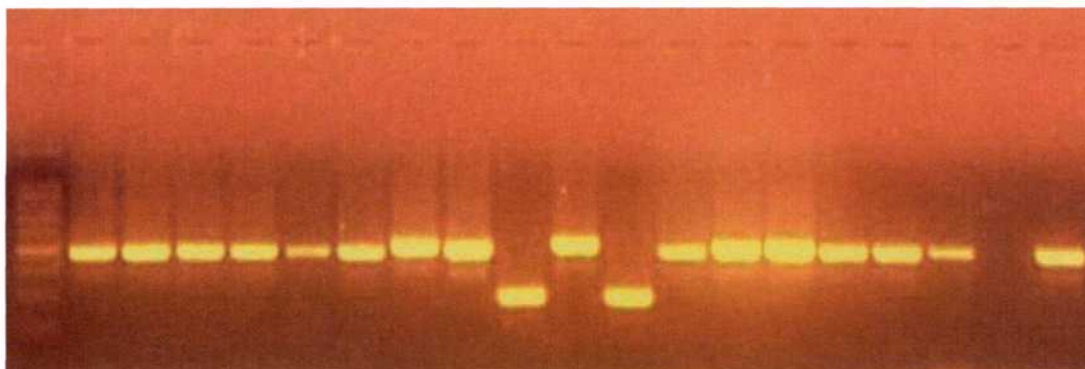
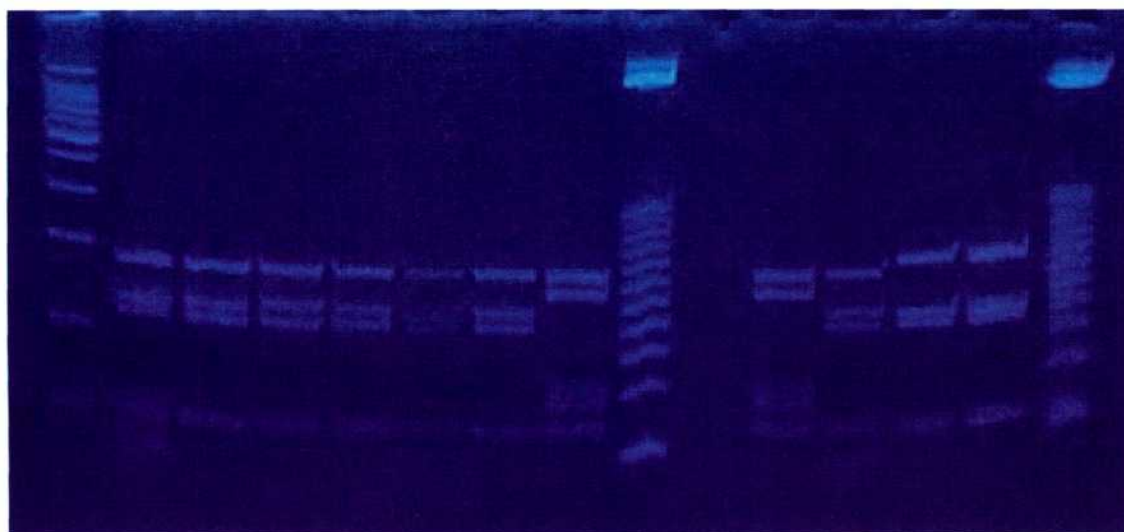


Figura 35: Restricción con Taq I de los productos de 420 a 510 bp. Calle 1 marcador de 100 pb, calle 9 marcador de 25 pb, calle 15 marcador de 10 pb; calle 2 AV7h calles 3 a 7 y 12 patrón similar a AV7h, calles 8 y 11 patrón simil AV7g y p, calle 10, control negativo, calle 13 patrón especie C, calle 14 AV2.



1 * 2 * 3 * 4 * 5 * 6 * 7 * 8 * 9 * 10 * 11 * 12 * 13 * 14 * 15

Las PCRs utilizadas demostraron tener diferente sensibilidad siendo ésta mayor en la PCR multiplex especie B:1 detectando tanto DNA viral extraído de aspirados nasofaríngeos como de sobrenadante de cultivo y menor en la PCR especie específica detectando sólo DNA viral extraído de cultivo celular. La tipificación por PCR mostró que el especie B:1 (AV7 y AV3) es el más importante, seguido por el C (AV5 y AV2) y el B:2 (AV 11).

Se continuaron estudiando 32 muestras de AV7 pertenecientes al período 1999-2002 por secuenciación. Adicionalmente se agregaron a este estudio de polimorfismos del hexón para ampliar el período abarcado muestras del año 1990 (1 muestra de evolución fatal), 1991 (1 muestra de evolución fatal), 1992 (2 muestras de casos fatales) y 1998 (1 muestra caso grave) como se detalla en la Tabla 13. Se analizaron hexones completos y regiones hipervariables por separado. El corte para las regiones hipervariables fueron los nucleótidos 403 a 1356 y los aminoácidos 136 a 452 (numerados de acuerdo a cepa Gomen). Los hexones completos se estudiaron desde 1990 hasta 1999, incluyendo la secuencia aportada por este trabajo de la cepa patrón de AV7h. Los polimorfismos en las regiones hipervariables fueron analizadas en el período 1990-2002, teniendo sólo información de las regiones hipervariables para el período 2000-2002

Tabla 13: Muestras seleccionadas para análisis de polimorfismos AV7
 La evolución clínica (evol) se evaluó con la dificultad respiratoria, mod corresponde a moderada en algunos casos la evolución resultó fatal. Se resalta la cepa patrón de Av7h usada para comparaciones.

Año	Nº Adeno	Evol	Código
1987	922	grave	87-922
1990	F55	fatal	BA55-90
1991	F10	grave	BA10-91
1992	F22	fatal	BA22-92
	F23	fatal	BA23-92
1998	3	fatal	BA3-98
1999	1	mod	BA1-99
	4	mod	BA4-99
	496	mod	BA496-99
	991	mod	BA991-99
	1325	mod	BA1325-99
	1390	mod	BA1390-99
	2972	fatal	BA2972-99
	3106	grave	BA3106-99
	3150	mod	BA3150-99
	3187	fatal	BA3187-99
	3434	mod	BA3434-99
2000	3017	mod	BA3017-00
2001	6800	grave	BA6800-01
2002	148	mod	BA148-02
	1924	mod	BA1924-02
	2026	mod	BA2026-02
	2169	mod	BA2169-02
	2268	mod	BA2268-02
	2687	mod	BA2687-02
	2688	mod	BA2688-02
	2966	mod	BA2966-02
	3029	mod	BA3029-02
	3037	mod	BA3037-02
	3270	mod	BA3270-02
	3282	mod	BA3282-02
	3387	mod	BA3387-02
	3478	mod	BA3478-02

Los hexones completos fueron amplificados con los *primers* hx5-1 y 3-1 dando un amplicón de 2,8Kb y secuenciados con los *primers* internos diseñados. Las regiones hipervariables fueron amplificadas y secuenciadas con los *primers* hx4 y S5 dando un producto de 953 pb.. Adicionalmente se analizaron muestras de AV3 para chequear el error del método. Diferentes productos finales fueron logrados y gráficamente podemos observar el comienzo y el alcance de cada secuencia obtenida (Figura 36).

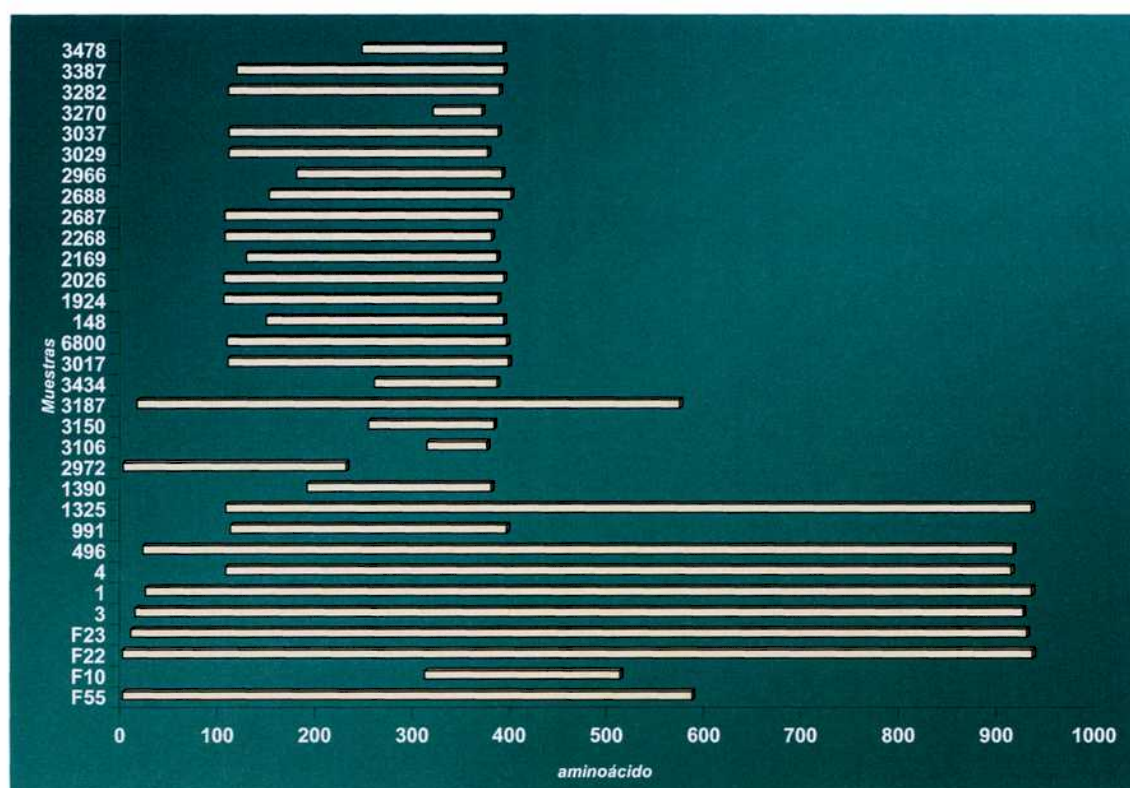


Figura 36: Alcance de las secuencias AV7 1990-2002 medido en amonoácidos

Se realizaron 2 PCRs diferentes para cada fragmento y 2 lecturas de cada *primer* para cada PCR, resultando cada base leída en dos sentidos dos veces (redundancia de 4). En los casos en que se clonó (muestras de 1990-1999) no se produjo ninguna diferencia entre las secuencias obtenidas de clones y las secuencias de PCR directa. Las secuencias de muestras AV3 coincidieron con lo reportado minimizando el error del método.

Se detallan a continuación las mutaciones encontradas en el período analizado para las 9 muestras que tuvieron secuencia completa del hexón. (Tabla 14). Se tomó para comparación la cepa prototipo del serotipo 7 (Gomen) y se incluyó la cepa patrón de AV7h. Se contabilizaron las transiciones (ts) y transversiones (tv). Las sustituciones se clasificaron en sinónimas (S) y no sinónimas (NS), estas últimas fueron analizadas de acuerdo a su carga y se consideró un cambio como no conservativo (NC) cuando resulta en un aminoácido de polaridad diferente y conservativo (C) cuando la carga permanece igual. Se consideraron como mutaciones fijadas aquellas que se presentaron en todo el período analizado.

Tabla 14: Mutaciones en hexones 1990-1999. Nd: no disponible, Ts: transición, Tv: transversión, S: sinónimo, NS: no sinónimo, C: conservativo, NC: no conservativo, P: positivo, N: negativo, PSC: polar sin carga, NP: no polar, Negrita: fijaciones: FS: fijación sinónima, NS: fijación no sinónima, Itálica: Fijaciones en muestras, FSm: fijación sinónima en muestras, FNSm: fijación no sinónima en muestras, In: inserción, Del: delección, (*) base sin cambio

Hexones		1987	1990	1992		1998	1999					Mut	Sust	Nat	Reemplazo			
nt	Gomen	av7H	55	22	23	3	1	4	496	1325	3187	Ts/Tv	S/NS	fijado	AA	cambio	C/NC	carga
9	C	*	T	*	nd	nd	nd	nd	nd	nd	nd	ts	S					
12	A	C	C	C	nd	nd	nd	nd	nd	nd	nd	tv	S					
24	C	G	*	G	nd	nd	nd	nd	nd	nd	nd	tv	S					
25	C	*	G	*	nd	nd	nd	nd	nd	nd	nd	tv	NS		9	Q-E	NC	PSC-N
70	C	*	*	T	*	*	nd	nd	*	nd	*	ts	S					
87	G	*	*	*	*	C	nd	*	nd	*	*	tv	S					
102	A	*	*	*	C	*	nd	*	nd	*	*	tv	S					
130	A	*	*	*	*	C	nd	*	nd	*	*	tv	NS		44	K-Q	NC	P-PSC
132	A	G	G	G	*	G	*	nd	*	nd	G	ts	S					
177	C	*	*	*	*	*	nd	T	nd	*	*	ts	S					
189	A	*	*	*	*	*	nd	C	nd	*	*	tv	S					
191	T	*	*	*	C	*	nd	*	nd	*	*	ts	NS		64	L-P	C	NP
207	A	G	G	G	G	G	nd	G	nd	G	G	ts	S	FS				
222	A	G	G	G	*	G	G	nd	*	nd	G	ts	S					
223	G	*	*	*	*	*	*	nd	A	nd	*	ts	NS		75	D-T	NC	N-PSC
224	A	*	*	*	*	*	nd	C	nd	*	*	tv	NS		75			
231	C	A	A	A	*	A	A	nd	*	nd	A	tv	S					
246	T	G	G	G	G	G	G	nd	*	nd	G	tv	S					
249	C	G	G	G	G	G	G	nd	*	nd	G	tv	S					
255	G	C	C	C	C	C	C	nd	*	nd	C	tv	S					
258	G	C	C	C	C	C	C	nd	*	nd	C	tv	S					
261	T	C	C	C	C	C	C	nd	*	nd	C	ts	S					
264	A	G	G	G	G	G	G	nd	*	nd	G	ts	S					
281	T	*	*	*	G	*	nd	*	nd	*	*	tv	NS		94	L-R	NC	NP-P
310	C	A	A	A	A	A	A	nd	*	nd	A	tv	S					
339	T	C	C	C	C	C	C	C	C	C	C	ts	S	FS				
351	T	C	C	C	C	C	C	C	C	C	C	ts	S	FS				
372	G	*	*	*	*	*	*	C	*	*	*	tv	S					
390	C	T	T	T	T	T	T	T	*	T	T	ts	S					
393	T	C	C	C	C	C	C	C	*	C	C	ts	S					
415	G	A	A	A	A	A	A	A	A	A	A	ts	NS	FNS	139	A-T	NC	NP-PSC
417	A	G	G	G	G	G	G	G	G	G	G	ts	S	FS				
418	G	*	*	*	*	*	*	*	A	*	*	ts	NS		140	G-N	C	PSC
419	G	*	*	*	*	*	*	*	A	*	*	ts	NS		140			
420	A	*	*	*	*	*	*	*	T	*	*	tv	S					
421	G	*	*	*	*	*	*	*	C	*	*	tv	NS		141	E-R	NC	N-P
422	A	*	*	*	*	*	*	*	G	*	*	ts	NS		141			
426	A	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	143	E-D	C	N
428	G	A	A	A	A	A	A	A	A	A	A	ts	NS	FNS	143			
429	A	T	T	T	T	T	T	T	T	T	T	tv	NS	FNS	144	R-N	NC	P-PSC
430-438	in	del	del	del	del	del	del	del	in	del	del		NS		147	T-A	NC	PSC-NP
448	A	T	T	T	T	T	T	T	*	T	T	tv	NS		150	N-Y	C	PSC
459	C	*	*	*	*	*	*	*	*	*	*	tv	S					
470	T	C	C	C	C	C	C	C	*	C	C	ts	NS		157	M-T	NC	NP-PSC
492	A	G	G	G	G	G	G	G	*	G	G	ts	S					
522	T	*	*	*	*	*	*	*	C	*	*	ts	S					
599	C	T	T	T	T	T	T	T	T	T	T	ts	NS	FNS	200	T-I	NC	PSC-NP
618	G	A	A	*	A	A	A	A	*	A	A	ts	S					
624	A	*	*	*	*	*	*	*	T	*	*	tv	S					
627	T	*	*	*	*	*	*	*	A	*	*	tv	S					
633	T	*	*	*	*	*	*	*	C	*	*	ts	S					
648	C	T	T	*	T	T	T	T	*	T	T	ts	S					
669	G	*	*	*	*	*	*	*	*	*	T	tv	S					
674	T	*	*	*	*	*	A	*	*	*	*	tv	NS		225	F-Y	NC	NP-PSC
682	C	A	*	*	*	*	*	*	*	*	*	tv	NS		228	P-T	NC	NP-PSC
687	A	*	*	*	G	*	*	*	*	*	*	ts	S					
690	T	*	C	*	*	C	C	C	C	C	C	ts	S					
692	T	C	*	*	*	*	*	*	*	*	*	ts	NS		231	I-T	NC	NP-PSC
725	A	*	C	C	C	C	C	C	C	C	C	tv	NS	FNSm	242	K-T	NC	P-PSC
739	G	A	*	*	*	*	*	*	*	*	*	ts	NS		247	D-N	NC	N-PSC
747	G	*	*	*	*	A	*	A	A	A	*	ts	S					
748	A	*	G	G	G	G	G	G	G	G	G	ts	NS	FNSm	250	T-A	NC	PSC-NP
756	G	*	*	*	*	*	*	*	A	*	*	ts	S					
777	T	C	*	*	*	*	*	*	*	*	*	ts	S					

Hexones		1987	1990	1992		1998	1999					Mut	Sust	Nat	Reemplazo			
nt	Gomen	av7H	55	22	23	3	1	4	496	1325	3187	Ts/Tv	S/NS	fijado	AA	cambio	C/NC	carga
779	T	C	*	*	*	*	*	*	*	*	*	ts	NS		260	F-S	NC	NP-PSC
780	T	C	*	*	*	C	*	C	C	C	C	ts	S					
828	C	*	*	*	*	*	*	*	T	*	*	ts	S					
846	A	G	*	G	*	G	*	G	*	G	G	ts	S					
867	A	G	*	G	*	G	*	G	*	G	G	ts	S					
870	T	A	*	A	A	A	*	A	*	A	A	tv	S					
893	A	G	G	G	G	G	G	G	G	G	G	ts	NS	FNS	298	D-G	NC	N-PSC
972	C	T	*	T	*	T	*	T	*	T	T	ts	S					
975	A	T	*	T	*	T	*	T	*	T	T	tv	S					
1014	T	C	*	C	C	C	*	C	*	C	C	ts	S					
1109	C	G	G	G	G	G	G	G	G	G	G	tv	NS	FNS	370	T-S	C	PSC
1122	T	C	*	C	C	C	C	C	*	C	C	ts	S					
1131	C	T	*	T	T	T	T	T	*	T	T	ts	S					
1146	T	C	*	C	C	C	C	C	*	C	C	ts	S					
1164	C	T	*	T	T	T	T	T	*	T	*	ts	S					
1166	T	*	*	*	*	*	*	*	*	A	*	tv	NS		389	I-N	NC	NP-PSC
1179	C	T	*	T	T	T	T	T	*	*	T	ts	S					
1183	A	G	G	G	G	G	G	G	G	G	G	ts	NS		395	I-V	C	NP
1235	C	G	G	G	G	G	G	G	G	G	G	tv	NS	FNS	412	A-G	NC	NP-PSC
1239	A	*	C	C	C	C	C	C	T	C	C	tv	NS	FNSm	413	K-N	NC	P-PSC
1241	C	A	A	A	A	A	A	A	A	A	A	tv	NS	FNS	414	T-K	NC	PSC-P
1258	T	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	420	S-P	NC	PSC-NP
1262	A	G	G	G	G	G	G	G	G	G	G	ts	NS	FNS	421	K-R	C	P
1268	A	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	423	N-T	C	PSC
1271	G	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	424	G-A	NC	PSC-NP
1285	G	A	A	A	A	A	A	A	A	A	A	ts	NS	FNS	429	D-T	NC	N-PSC
1286	A	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	429			
1290	T	A	A	A	A	A	A	A	A	A	A	tv	NS	FNS	430	N-K	NC	PSC-P
1298	A	C	C	C	C	C	C	C	C	C	C	tv	NS	FNS	433	K-T	NC	P-PSC
1300	T	G	G	G	G	G	G	G	G	G	G	tv	NS	FNS	434	S-A	NC	PSC-NP
1320	C	*	*	*	*	*	*	*	A	*	*	tv	S					
1326	C	T	T	T	T	T	T	T	T	T	T	ts	S	FS				
1328	A	T	T	T	T	T	T	T	T	T	T	tv	NS	FNS	443	Q-L	NC	PSC-NP
1356	C	T	T	T	T	T	T	T	T	T	T	ts	S					
1358	T	*	*	*	*	*	*	*	*	C	*	ts	NS		453	L-P	C	NP
1418	C	*	*	*	*	*	*	*	*	G	*	tv	NS		473	T-R	NC	PSC-P
1419	G	*	*	*	*	*	*	*	A	*	*	tv	S					
1422	A	*	*	*	*	*	*	*	C	*	*	tv	S					
1435	C	*	*	*	T	*	*	*	*	*	*	ts	S					
1436	T	*	*	*	G	*	*	*	*	*	*	tv	NS		479	L-W	C	NP
1443	C	T	T	T	T	*	T	T	*	T	T	ts	S					
1482	A	T	T	T	*	*	T	T	*	T	T	tv	S					
1497	T	C	*	C	C	*	C	C	*	C	C	ts	S					
1538	C	*	A	*	*	*	*	*	*	*	*	tv	NS		513	P-Q	NC	NP-PSC
1558	T	*	*	*	*	*	*	*	*	G	*	tv	NS		520	F-V	C	NP
1560	C	T	T	T	*	*	T	T	*	T	T	ts	S					
1561	A	G	*	*	*	*	*	*	*	*	*	ts	NS		521	N-D	NC	PSC-N
1569	C	*	*	A	*	*	*	*	*	*	*	tv	NS		523	H-Q	NC	P-PSC
1572	C	*	*	A	*	*	*	*	*	*	*	tv	S					
1573	A	G	*	*	*	*	*	*	*	*	*	ts	NS		525	N-D	NC	PSC-N
1574	A	*	*	T	*	*	*	*	*	*	*	tv	NS		525	N-I	NC	PSC-NP
1581	T	C	C	A	*	*	C	G	*	G	C	ts	S					
1584	G	A	A	A	*	*	A	A	*	A	A	ts	S					
1614	T	C	C	C	*	*	C	C	*	G	C	ts	S					
1649	A	*	C	*	*	*	*	*	*	*	*	tv	NS		550	K-T	NC	P-PSC
1681	C	*	*	*	*	*	*	*	*	A	*	tv	NS					
1686	C	*	*	*	*	*	*	*	*	G	*	tv	S					
1715	A	*	C	*	*	*	*	*	*	*	C	tv	NS		572	K-T	NC	P-PSC
1724	A	*	*	C	*	*	*	*	*	*	nd	tv	NS		575	N-T	C	PSC
1735	C	*	*	*	*	*	*	G	*	*	nd	tv	NS		579	Q-E	NC	PSC-N
1762	A	*	nd	C	*	*	*	*	*	*	nd	tv	NS		588	T-P	NC	PSC-NP
1772	C	*	nd	*	*	*	*	G	*	*	nd	tv	NS		591	A-G	NC	NP-PSC
1791	C	*	nd	G	G	G	G	G	G	G	nd	tv	NS	FNSm	597	S-R	NC	PSC-P
1797	T	C	nd	*	*	*	*	*	*	C	nd	ts	S					

Tabla 14 (cont): Mutaciones en hexones 1990-1999. Nd: no disponible, Ts: transición, Tv: transversión, S: sinónimo, NS: no sinónimo, C: conservativo, NC: no conservativo, P: positivo, N: negativo, PSC: polar sin carga, NP: no polar, Negrita: fijaciones, FS: fijación sinónima, NS: fijación no sinónima, Itálica: Fijaciones en muestras, FSm: fijación sinónima en muestras, FNSm: fijación no sinónima en muestras, In: inserción, Del: delección, (*) base sin cambio

Hexones		1987	1990	1992		1998	1999					Mut	Sust	Nat	Reemplazo			
nt	Gomen	av7H	55	22	23	3	1	4	496	1325	3187	Ts/Tv	S/NS	fijado	AA	cambio	C/NC	carga
1799	T	*	nd	C	C	C	C	C	C	C	nd	ts	NS	FNSm	600	L-P	C	NP
1800	G	C	nd	A	A	A	A	A	A	A	nd	ts	NS	FNSm				
1812	C	*	nd	A	*	*	*	*	*	*	nd	tv	NS		604	F-L	C	NP
1815	C	*	nd	*	*	*	*	*	*	T	nd	ts	S					
1823	C	*	nd	G	G	G	G	G	G	G	nd	tv	NS	FNSm	608	A-G	NC	NP-PSC
1833	A	C	nd	*	*	*	C	*	*	G	nd	tv	S					
1835	C	*	nd	G	G	G	G	G	G	G	nd	tv	NS	FNSm	612	A-G	NC	NP-PSC
1839	C	A	nd	A	*	*	A	A	*	A	nd	tv	S					
1842	C	T	nd	*	*	*	T	T	*	T	nd	ts	S					
1858	C	G	nd	*	*	*	*	*	*	*	nd	tv	NS		620	R-G	NC	P-PSC
1887	C	A	nd	*	*	*	*	*	*	*	nd	tv	NS		629	N-K	NC	PSC-P
1899	T	*	nd	G	*	*	*	*	*	*	nd	tv	S					
1912	C	*	nd	G	*	*	*	*	*	*	nd	tv	NS		638	L-V	C	NP
1950	T	C	nd	*	*	*	*	*	*	*	nd	ts	S					
1968	C	*	nd	*	*	*	*	*	*	T	nd	tv	S					
1972	G	C	nd	*	*	*	*	*	*	*	nd	tv	NS		658	A-P	C	NP
1975	G	C	nd	*	*	*	*	*	*	*	nd	tv	NS		659	A-P	C	NP
1986	C	*	nd	*	*	*	*	*	T	*	nd	ts	S					
1992	C	*	nd	*	*	T	*	*	*	*	nd	ts	S					
1998	T	*	nd	*	*	C	*	C	C	C	nd	ts	S					
2001	G	A	nd	*	*	A	*	A	*	A	nd	ts	S					
2004	C	*	nd	*	*	*	*	*	A	*	nd	tv	S					
2014	G	*	nd	A	A	A	A	A	A	A	nd	ts	NS	FNSm	672	E-K	NC	N-P
2040	C	*	nd	*	*	*	*	*	T	*	nd	ts	S					
2053	G	C	nd	*	*	*	*	*	*	*	nd	tv	NS		685	V-R	NC	NP-P
2054	T	G	nd	*	*	*	*	*	*	*	nd	tv	NS		685			
2055	A	T	nd	*	*	T	*	T	*	T	nd	tv	S					
2088	C	*	nd	*	*	T	*	T	*	T	nd	ts	S					
2093	A	*	nd	*	*	*	*	T	*	*	nd	tv	NS		698	Y-F	NC	PSC-NP
2121	C	*	nd	A	*	*	A	*	*	*	nd	tv	S					
2170	A	T	nd	T	*	T	T	T	*	T	nd	tv	S					
2171	G	C	nd	C	*	C	C	C	*	C	nd	tv	S					
2172	C	T	nd	T	*	T	T	T	*	T	nd	ts	S					
2201	T	*	nd	C	*	*	*	*	*	*	nd	ts	NS		734	V-A	C	NP
2205	T	C	nd	T	*	T	T	T	*	T	nd	ts	S					
2214	G	A	nd	A	A	A	A	A	A	A	nd	ts	S	FS				
2220	T	*	nd	*	*	*	*	*	C	*	nd	ts	S					
2226	C	*	nd	*	*	*	*	*	A	*	nd	tv	S					
2232	T	C	nd	C	C	C	C	C	C	C	nd	ts	S	FS				
2233	A	*	nd	C	*	*	*	*	*	*	nd	tv	NS		745	N-H	NC	PSC-P
2256	G	*	nd	*	*	*	*	*	A	*	nd	ts	S					
2308	G	*	nd	*	*	*	T	*	*	*	nd	tv	NS		770	E-W	NC	N-NP
2309	A	*	nd	*	*	*	G	*	*	*	nd	ts	NS		770			
2312	G	*	nd	*	*	*	A	*	*	*	nd	ts	NS		771	G-E	NC	PSC-N
2334	C	*	nd	*	*	*	*	*	T	*	nd	ts	S					
2360	G	*	nd	*	*	*	C	*	*	*	nd	tv	NS		787	S-T	C	PSC
2418	C	T	nd	T	T	T	T	T	*	T	nd	ts	S					
2444	G	*	nd	*	*	*	A	*	*	*	nd	tv	NS		815	G-E	NC	PSC-N
2445	G	A	nd	A	A	A	A	A	*	A	nd	tv	S					
2448	T	C	nd	C	C	C	C	C	C	C	nd	ts	S	FS				
2454	A	G	nd	G	G	G	G	G	*	G	nd	ts	S					
2463	G	*	nd	*	*	A	*	*	*	*	nd	ts	NS		821	M-I	C	NP
2526	G	A	nd	A	A	A	A	A	*	A	nd	ts	S					
2532	C	T	nd	T	T	T	T	T	T	T	nd	ts	S	FS				
2538	G	A	nd	A	A	A	A	A	*	A	nd	ts	S					
2553	T	C	nd	C	C	C	C	C	C	C	nd	ts	S	FS				
2559	G	*	nd	*	*	*	*	A	*	*	nd	ts	S					
2574	T	A	nd	A	A	A	A	A	*	A	nd	tv	S					
2577	C	G	nd	G	G	G	G	G	*	G	nd	tv	S					
2604	G	A	nd	A	A	A	A	A	*	A	nd	ts	S					
2613	C	G	nd	G	G	G	G	G	*	G	nd	tv	S					
2629	A	C	nd	C	C	C	C	C	C	C	nd	tv	NS	FNS	877	M-L	C	NP
2646	A	G	nd	G	G	G	G	G	*	G	nd	ts	S					
2728	G	*	nd	*	A	*	*	*	*	*	nd	ts	NS		910	V-K	NC	PSC-P
2729	T	*	nd	*	A	*	*	*	*	*	nd	tv	NS		910			
2760	C	*	nd	*	A	*	*	nd	nd	*	nd	tv	S					

Tabla 14 (cont): Mutaciones en hexones 1990-1999. Nd: no disponible, Ts: transición, Tv: transversión, S: sinónimo, NS: no sinónimo, C: conservativo, NC: no conservativo, P: positivo, N: negativo, PSC: polar sin carga, NP: no polar, Negrita: fijaciones: , FS: fijación sinónima, NS: fijación no sinónima, Itálica: Fijaciones en muestras, FSm: fijación sinónima en muestras, FNSm: fijación no sinónima en muestras, In: inserción, Del: deletión, (*) base sin cambio

Las muestras pertenecieron al GTC2 y presentaron la delección típica de este *cluster*. Las transiciones y transversiones resultaron equivalentes, con 3 sitios donde los polimorfismos fueron mayores a 2, es decir que se produjo más de una sustitución en la misma posición. Se registraron 193 sustituciones siendo 110 sinónimas y 83 no sinónimas. En los cambios de reemplazo aminoacídico no se conservó la carga en 53 mutaciones. En siete sitios el cambio de aminoácido es provocado por la sustitución de más de un nucleótido por codón. Se encontraron fijados, presentes desde 1987 hasta 1999 diez (10) cambios no sinónimos más uno encontrado sólo en muestras. El número de sitios no sinónimos fijados fue mayor (20) sumándole 8 sustituciones presentes sólo en muestras (ausentes en la cepa patrón AV7h). Cabe destacar que se encontraron cambios aminoacídicos fijados desde 1990 hasta 1999 en las cepas de AV7h de evolución fatal (Tabla 15). Estos cambios se encuentran fuera de las HVR y permiten separar casos leves y moderados de los fatales en el análisis filogenético (ver sección *ad hoc*).

Tabla 15: Reemplazos de aminoácidos en muestras fatales AV7 1990-1999.

Aa.	7h	Muestras fatales	Región antigénica afectada	Región Hexón
597	Ser (polar sin carga)	Arg (carga positiva)	Epitope B 596-601 Epitope T 588-592	P1
608	Ala (hidrofóbico)	Gly (polar sin carga)	Epitope B 606-615	PC
612	Ala (hidrofóbico)	Gly (polar si carga)	Epitope B 606-615	PC
672	Asp (carga negativa)	Lys (carga positiva)	Regiones hidrofílicas en superficie	L3
877	Met (hidrofóbico)	Leu (hidrofóbico)	Epitope B 876-883	P2

Polimorfismos en las regiones hipervariables del hexón 1990-2002

Adicionalmente se analizaron para 23 muestras AV7 las regiones hipervariables comprendidas entre los nucleótidos 403 a 1356 y los aminoácidos 136 a 452 (numerados de acuerdo a cepa Gomen). En este caso se agregaron 14 mutaciones, de las cuales 8 fueron no sinónimas y 5 de éstas no conservativas. Se adicionaron 5 sitios sinónimos fijados más 1 presente sólo en muestras y dos sitios fijados no sinónimos. Se analizaron todos los fragmentos disponibles independientemente de su longitud final.

Los cambios obtenidos se detallan en la Tabla 16.

Podemos resumir los polimorfismos hallados diciendo que se encontraron en total 207 sitios polimórficos, siendo 91 no sinónimos y 58 de éstos no conservativos. Las transiciones y transversiones fueron parejas dando un índice $\kappa = T_s/T_v = 1.02$, indicando que los sitios explorados están saturados (se probaron al azar todas las sustituciones posibles). Se fijaron en el período 1987-2002 sitios nuevos en 37 posiciones de las cuales 22 se expresaron como aminoácidos diferentes a la cepa prototipo del serotipo 7 (Gomen) La tabla 17 resume lo visto hasta el momento.

Hexon HVRs		1987	1991	1999								2000	2001	2002													Mut	Sust	Nat	Reemplazo				
nt	Gomen	av7H	10	991	1390	2972	3106	3150	3434	3017	6800	148	1924	2026	2169	2268	2687	2688	2966	3029	3037	3270	3282	3387	3478	Ts/Tv	S/NS	fijado	AA	cambio	C/NC	carga		
339	T	C	nd	C	nd	*	nd	nd	nd	C	C	nd	C	C	nd	C	C	nd	nd	C	C	nd	C	nd	nd	ts	S	no more						
351	T	C	nd	C	nd	*	nd	nd	nd	C	*	nd	C	C	nd	C	C	nd	nd	C	C	nd	C	nd	nd	ts	S	no more						
390	C	T	nd	*	nd	T	nd	nd	nd	T	*	nd	T	T	T	T	*	nd	nd	*	T	nd	T	*	nd	ts	S							
393	T	C	nd	C	nd	C	nd	nd	nd	C	C	nd	C	C	C	C	C	nd	nd	C	C	nd	C	*	nd	ts	S							
415	G	A	nd	A	nd	A	nd	nd	nd	A	A	nd	A	A	A	A	A	nd	nd	A	A	nd	A	A	nd	ts	NS	FNS	139	A-T	NC	NP-PSC		
417	A	G	nd	G	nd	G	nd	nd	nd	G	G	nd	G	G	G	G	G	nd	nd	G	G	nd	G	G	nd	ts	S	FS						
426	A	C	nd	C	nd	C	nd	nd	nd	C	C	nd	C	C	C	C	C	nd	nd	C	C	nd	C	C	nd	tv	NS	FNS	143	E-D	C	N		
428	G	A	nd	A	nd	A	nd	nd	nd	A	A	nd	A	A	A	A	A	nd	nd	A	A	nd	A	A	nd	ts	NS	FNS	143					
429	A	T	nd	G	nd	G	nd	nd	nd	G	G	nd	G	G	G	G	G	nd	nd	G	G	nd	G	G	nd	tv	NS	FNS	144	R-N	NC	P-PSC		
430-438	in	del	nd	del	nd	del	nd	nd	nd	del	del	nd	del	del	del	del	del	nd	nd	del	del	nd	del	del	nd	NS		147	T-A	NC	PSC-NP			
448	A	T	nd	T	nd	T	nd	nd	nd	T	T	T	T	T	T	T	T	nd	nd	T	T	nd	T	T	nd	tv	NS	FNSH	150	N-Y	C	PSC		
465	T	*	nd	*	nd	*	nd	nd	nd	*	*	*	*	*	*	*	*	C	nd	*	*	nd	*	*	nd	ts	S*							
470	T	C	nd	C	nd	C	nd	nd	nd	C	C	C	C	C	C	C	C	C	nd	C	C	nd	C	C	nd	ts	NS	FNSH	157	M-T	NC	NP-PSC		
492	A	G	nd	G	nd	G	nd	nd	nd	G	G	*	G	G	G	G	*	nd	G	G	nd	G	G	nd	ts	S								
522	T	*	nd	*	nd	*	nd	nd	nd	*	*	*	*	*	*	*	*	nd	*	*	nd	G	*	nd	tv	S	con 496							
525	A	*	nd	*	nd	*	nd	nd	nd	*	*	*	*	*	*	*	*	T	nd	*	*	nd	*	*	nd	tv	S*							
529	A	*	nd	T	nd	*	nd	nd	nd	*	*	*	*	*	*	*	*	T	nd	*	*	nd	*	*	nd	tv	NS*		177	N-Y	C	PSC		
537	C	*	nd	*	nd	*	nd	nd	nd	*	*	*	*	*	*	*	*	*	*	*	*	nd	*	A	nd	ts	S*							
599	C	T	nd	T	T	T	nd	nd	nd	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	nd	ts	NS	FNS	200	T-I	NC	PSC-NP		
614	A	*	nd	*	*	*	nd	nd	nd	*	G	*	G	*	*	*	*	*	*	*	*	nd	*	*	nd	ts	NS*		205	E-G	NC	N-PSC		
618	G	A	nd	A	A	A	nd	nd	nd	A	A	A	A	A	A	A	A	*	A	A	A	nd	A	A	nd	ts	S							
648	C	T	nd	*	T	T	nd	nd	nd	T	T	T	T	T	T	T	T	T	T	T	T	T	T	nd	ts	S								
649	A	*	nd	*	*	*	nd	nd	nd	*	G	*	*	*	*	*	*	*	*	*	*	nd	*	*	nd	ts	NS*		217	K-E	NC	P-N		
690	T	*	nd	C	*	*	nd	nd	nd	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	nd	ts	S							
725	A	*	nd	C	C	C	nd	nd	nd	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	nd	tv	NS	FNSm	242	K-T	NC	P-PSC		
747	G	*	nd	A	A	nd	nd	nd	nd	A	A	A	A	A	A	A	A	A	A	A	A	A	nd	A	A	A	ts	S	FSHm					
748	A	*	nd	G	C	G	nd	nd	nd	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	nd	ts	NS	FNSm	250	T-A	NC	PSC-NP		
780	T	C	nd	C	C	C	nd	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	ts	S							
822	G	*	nd	*	*	nd	nd	*	C	*	*	*	*	*	*	*	*	*	*	*	*	nd	*	*	*	tv	S*							
846	A	G	nd	G	G	nd	nd	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	nd	ts	S	FSH						
867	A	G	nd	G	G	nd	nd	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	nd	ts	S	FSH						
870	T	A	nd	A	A	nd	nd	A	*	A	A	A	A	A	A	A	A	A	A	A	A	nd	A	A	nd	ts	S							
893	A	G	nd	G	G	nd	nd	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	nd	ts	NS	FNS	298	D-G	NC	N-PSC		
972	C	T	*	T	T	nd	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	ts	S							
975	A	T	*	T	T	nd	T	T	T	T	T	*	T	T	T	T	T	T	T	T	T	T	T	T	T	ts	S							
986	A	*	*	C	*	nd	*	*	*	*	C	*	*	*	*	*	*	*	*	*	*	*	*	*	*	tv	NS*		329	N-T	C	PSC		
1014	T	C	C	C	*	nd	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	ts	S							
1028	A	*	*	T	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	tv	NS*		343	Q-L	NC	PSC-NP		
1029	A	*	G	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	ts	S*							
1033	A	*	*	*	T	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	tv	NS*		345	N-Y	C	PSC		
1049	T	*	*	*	C	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	ts	NS*		350	L-S	NS	NP-PSC		
1084	T	*	*	C	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	ts	S*							
1109	C	G	G	G	G	nd	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	nd	ts	NS	FNS	370	T-S	C	PSC		
1122	T	C	C	C	C	nd	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	ts	S	FSH						
1131	C	T	T	T	T	nd	nd	T	T	T	T	T	T	T	T	T	T	T	T	T	T	nd	T	T	T	ts	S	FSH						
1135	G	*	*	A	*	nd	nd	*	*	*	*	A	*	*	*	*	*	*	*	*	nd	A	nd	*	*	ts	NS*		379	A-T	NC	NP-PSC		
1146	T	C	C	*	C	nd	nd	*	C	C	C	*	C	C	C	C	C	C	nd	*	nd	C	C	C	C	ts	S							
1164	C	T	T	T	T	nd	nd	T	nd	T	T	*	nd	T	nd	nd	nd	T	T	nd	nd	nd	nd	*	T	ts	S							
1166	T	*	*	A	A	nd	nd	*	nd	*	*	*	nd	*	nd	nd	nd	*	nd	nd	nd	nd	nd	*	*	tv	NS	con 496	389	I-N	NC	NP-PSC		
1179	C	T	T	T	nd	nd	nd	nd	nd	T	T	nd	nd	nd	nd	nd	nd	T	T	nd	nd	nd	nd	nd	nd	ts	S	FSH						
1183	A	G	G	G	nd	nd	nd	nd	nd	G	G	nd	nd	nd	nd	nd	nd	G	G	nd	nd	nd	nd	nd	nd	ts	NS	FNS	395	I-V	C	NP		

Tabla 16: Mutaciones en HVRs hexones 1991-2002

Nd: no disponible, Ts: transición, Tv: transversión, S: sinónimo, NS: no sinónimo, C: conservativo, NC: no conservativo, P: positivo, N: negativo, PSC: polar sin carga, NP: no polar, Negrita: fijaciones, FS: fijación sinónima, FNS: fijación no sinónima, Itálica: Fijaciones en muestras, FSm: fijación sinónima en muestras, FNSm: fijación no sinónima en muestras, In: inserción, Del: delección. Se marcaron un (*) las sustituciones nuevas y las bases sin cambio según corresponda.

Tabla 17 Resumen polimorfismos hexones 1987-2002
--

a-Análisis de mutaciones

	HEXON	HVR	Totales
Sinónimas	110	6	116
No Sinónimas	83	8	91
Conservativas	21	3	
No Conservativas	53	5	
			207

b-Análisis de sustituciones

	HEXON	HVR	SubTotales	Totales
Transiciones	97	8	105	
polim>2	1	0	1	106
Transversiones	95	6	101	
polim>2	2	1	3	104
Totales	195	15		210
			Ts/Tv=	1,02

c-Análisis de sitios fijados

	HEXON	HVR	Totales
Sinónimos	10	5	
en muestras	1	1	17
No Sinónimos	20	2	
en muestras	8	0	30
Totales	39	8	47

Si tenemos en cuenta la longitud del fragmento (L) y los sitios polimórficos o segregantes (S) podemos calcular la diversidad encontrada en el período como S/L y expresarla como porcentaje. Esta relación se calculó para todas las muestras analizadas desde 1990 hasta 2002 y dio una mediana de 3.49% (rango 1.83-4.65), aunque si se calcula tomando en cuenta todos los polimorfismos hallados desde la descripción del AV7h en 1987 se eleva notoriamente a 7.36%. Para tener en cuenta los errores introducidos por el uso de DNA polimerasas en el proceso de PCR y secuenciación se normaliza este cociente dividiéndolo por el número total de ciclos aplicados. En las muestras este parámetro da mayor que lo esperado indicando que los cambios no son artefactos de la técnica aplicada, teniendo en cuenta que el error de un Taq DNA polimerasa ronda $4-9 \times 10^{-6}$ (Smith *et al.*, 1997).

Se calculó también para todas las muestras el índice ω que relaciona las sustituciones no sinónimas y sinónimas. Este cociente es indicador del tipo de selección que está operando sobre el gen analizado. En las muestras 1990-2002 dio una mediana de 0.46 (rango 0.2-1.16) y tomando en cuenta los polimorfismos desde el hallazgo de la cepa patrón del AV7h en 1987 da 0.78. En ambos casos es menor que 1 indicando que es un mecanismo de selección purificadora el que se está observando.

Todas las muestras analizadas conservan las mutaciones típicas de genotipo 7h entre los aa127-310, aa383-538 y aa757-852. Se detallan en la Tabla 18 los datos y cálculos obtenidos.

Tabla 18: Datos y cálculos sobre mutaciones hexones 1987-2002. L: longitud del fragmento, S: sustitución sinónima, NS: sustitución no sinónima, ω =NS/S, Diversidad media (%)= (sitios polimórficos/L)*100, diversidad por ciclo PCR= Diversidad media/60 (30 de PCR+ 30 secuenciación).

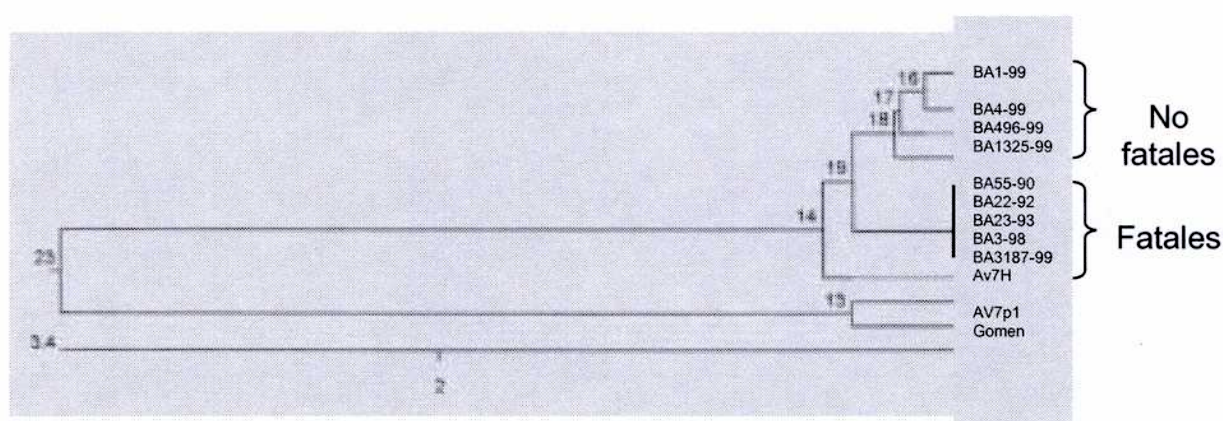
Hexones			Sitios		Mutaciones		ω	Diversidad	
Año	Muestra	L (pb)	monomórficos	polimórficos	S	NS		media (%)	por ciclo PCR
1987	av7h	2808	2703	105	70	35	0,50	3,74	$6,23 \times 10^{-4}$
1990	55	1748	1689	59	30	29	0,97	3,38	$5,63 \times 10^{-4}$
1991	10	436	428	8	6	2	0,33	1,83	$3,06 \times 10^{-2}$
1992	22	2802	2700	102	63	39	0,62	3,64	$6,07 \times 10^{-4}$
	23	2748	2664	84	48	36	0,75	3,06	$5,09 \times 10^{-4}$
1998	3	2727	2636	91	59	32	0,54	3,34	$5,56 \times 10^{-4}$
1999	1	2727	2628	99	59	40	0,68	3,63	$6,05 \times 10^{-4}$
	4	2414	2322	92	58	34	0,59	3,81	$6,35 \times 10^{-4}$
	496	2679	2610	69	36	33	0,92	2,58	$4,29 \times 10^{-4}$
	991	774	741	33	20	13	0,65	4,26	$7,11 \times 10^{-2}$
	1325	2484	2387	97	61	36	0,59	3,90	$6,51 \times 10^{-4}$
	1390	602	587	15	12	3	0,25	2,49	$4,15 \times 10^{-2}$
	2972	279	266	13	6	7	1,17	4,66	$7,77 \times 10^{-2}$
	3106	190	185	5	4	1	0,25	2,63	$4,39 \times 10^{-2}$
	3150	423	411	12	10	2	0,20	2,84	$4,73 \times 10^{-2}$
	3187	1665	1600	65	39	26	0,67	3,90	$6,51 \times 10^{-4}$
2000	3434	381	370	11	9	2	0,22	2,89	$4,81 \times 10^{-2}$
	3017	780	749	31	21	10	0,48	3,97	$6,62 \times 10^{-2}$
	6800	775	744	31	19	12	0,63	4,00	$6,67 \times 10^{-2}$
	148	739	721	18	12	6	0,5	2,44	$4,06 \times 10^{-2}$
	1924	746	718	28	18	10	0,56	3,75	$6,26 \times 10^{-2}$
	2026	766	737	29	20	9	0,45	3,79	$6,31 \times 10^{-2}$
	2169	744	718	26	16	10	0,63	3,49	$5,82 \times 10^{-2}$
	2268	727	701	26	17	9	0,53	3,58	$5,96 \times 10^{-2}$
	2687	749	724	25	16	9	0,56	3,34	$5,56 \times 10^{-2}$
	2688	747	726	21	15	6	0,4	2,81	$4,69 \times 10^{-2}$
	2966	668	649	19	15	4	0,27	2,84	$4,74 \times 10^{-2}$
	3029	715	690	25	16	9	0,56	3,50	$5,83 \times 10^{-2}$
	3037	747	719	28	18	10	0,56	3,75	$6,25 \times 10^{-2}$
	3270	156	153	3	3	0	0	1,92	$3,21 \times 10^{-2}$
	3282	751	722	29	20	9	0,45	3,86	$6,44 \times 10^{-2}$
	3387	766	741	25	16	9	0,56	3,26	$5,44 \times 10^{-2}$
	3478	438	423	15	12	3	0,25	3,42	$5,71 \times 10^{-2}$
Total		2808	2601	207	116	91	0,78	7,37	0,12

El total de la Tabla 18 se refiere a la suma de todos los polimorfismos (S) hallados sobre una longitud óptima de 2808 bases y trata de integrar, de manera grosera lo que pudo haber sucedido en el período desde la caracterización de la cepa AV7h hasta el año 2002.

Inferencia filogenética

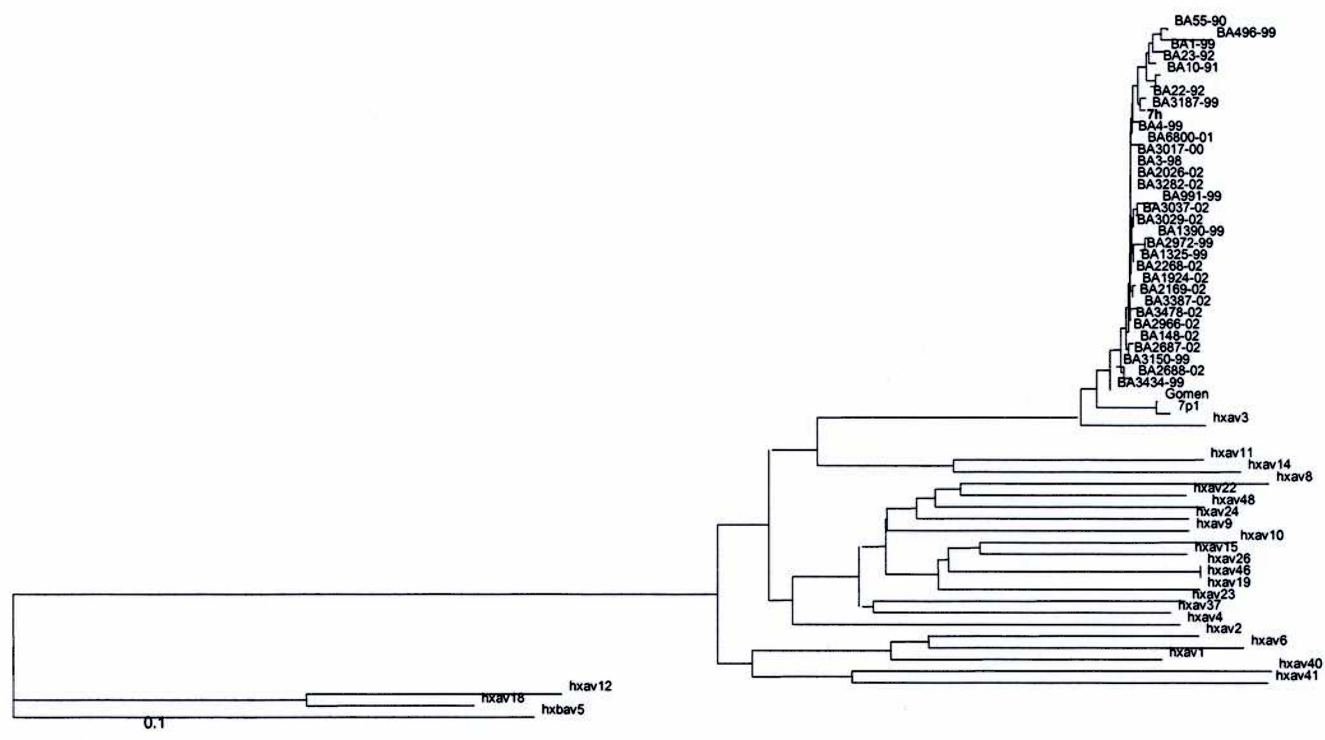
El análisis filogenético del hexón completo permitió agrupar las muestras fatales y las separó de las graves y leves como se puede ver en el siguiente árbol (Figura 37).

Figura 37: Inferencia filogenética hexones completos 1990-1999: La escala indica en el árbol los % de distancia y los números las sustituciones. Se utilizó la tabla de residuos PAM250.



Si analizamos las regiones hipervariables obtenemos la Figura 38. Allí se constata la asociación del genotipo y la relación con los serotipos de Adenovirus que tienen secuencia de hexón descriptas. No se pudieron analizar las muestras BA3106-99 y BA3270-99 por datos insuficientes para este análisis. Se separan claramente las especies de Adenovirus humanos A (AV12 y 18), C (AV1, 2 y 6), D (AV8, 9, 10, 15, 19, 22, 24, 26, 37, 46 y 48), E (AV4). Dentro de la especie B se separa la subespecie B2 (AV11 y 14) de la B1 (AV3 y 7). Dentro de la subespecie B1, podemos separar los *clusters* genómicos GTC1 (AV7p1 y Gomen) del GTC2 (AV7h y todas las muestras analizadas 1990-2002).

Figura 38: Inferencia filogenética de las regiones hipervariables de los hexones 1990-2002.



Resultados secuenciación genoma AV7h

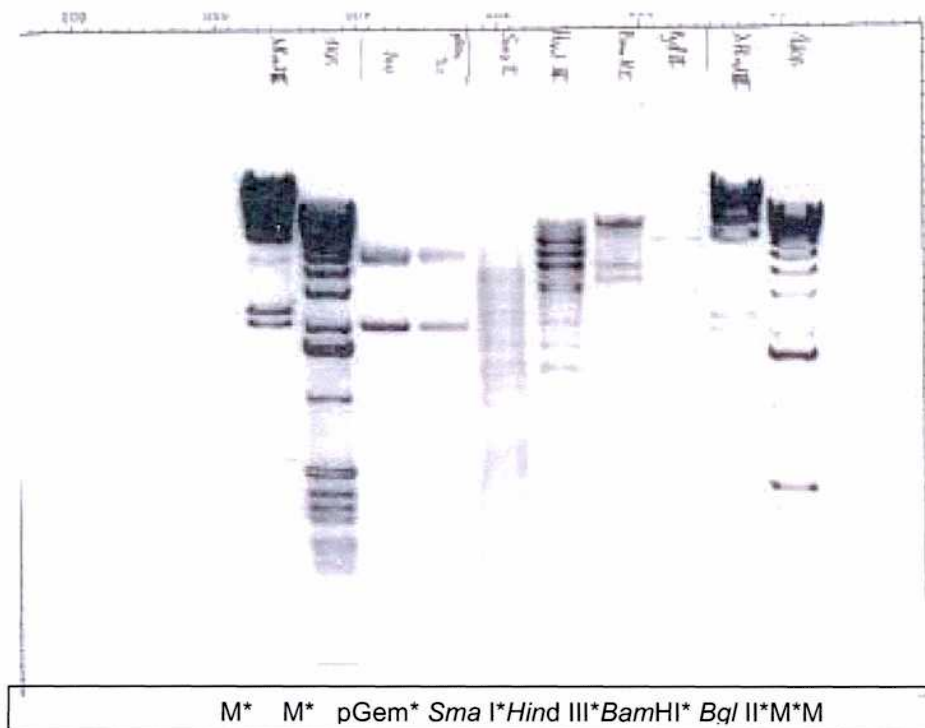
Clonado

Se digirió el DNA del AV7h con enzimas de restricción *Bam*H I, *Bg* II, *Hind* III y *Sma* I. Los fragmentos coincidieron con Niel *et al.*, 1991 (Figura 39).

Los insertos se purificaron, se cuantificaron y se hicieron 2 ligaciones L1 y L2, en cada caso ligando vector a los productos de digestión *Bam*H I, *Bg* II, *Hind* III. Se descartaron los fragmentos generados por corte con *Sma* I por no tener cantidad suficiente de inserto para ligar. Se hicieron 2 transformaciones T1 (con L1) y T2 (con L2). De la ligación L1-*Hind* III, obtuvimos diez colonias blancas (positivas).

Las colonias se analizaron por electroforesis en geles de agarosa al 1% y tinción con bromuro de etidio y resultaron positivas 6 colonias que tenían insertos cuyo peso molecular correspondería al fragmento grande de *Hind* III de 7000 pb.

Figura 39: Corte de restricción para DNA genómico AV7h



Secuenciación de clones

Se secuenciaron 2 µL (200ng) del producto de amplificación isotérmica del *rolling cycling*. Se realizó una búsqueda en la base de datos pública para comprobar los fragmentos obtenidos (Altschul *et al.*, 1997).

1. Clon H53 T7 homología con AV7 Gomen nucleótidos 22360-23050
2. Clon H22 T7 homología con AV7 Gomen nucleótidos 27550-28271.
3. Clon H12 Sp6 homología con AV7 Gomen nucleótidos 1893-2642.
4. Clon 254 Sp6 homología con AV7 Gomen nucleótidos 2506-2800.
5. Clon 254 T7 homología con AV7 Gomen nucleótidos 5351-6140.

Esta estrategia no resultó muy productiva, teniendo poco rendimiento tanto en las ligaciones como en la secuenciación. Estos datos fueron integrados en el esquema general de trabajo pero la estrategia se descartó y se continuó con la estrategia de PCR.

Diseño de *primers* específicos y amplificación

Se diseñaron 155 *primers* específicos con el programa Primer Select del paquete de software Lasergene (DNASTar Inc., Madison WI) para la secuenciación completa del AV7h a partir de los dos moldes disponibles en GenBank. (Tabla 19)

Tabla 19: Moldes para el diseño de *primers* para secuenciación genoma AV7h

<i>Molde</i>	<i>Región (pb)</i>	<i>Longitud</i>	<i>Primers</i>
AV7 Gomen	0-10704	10704	44
AV11p	10600-13750	3150	16
AV7 Gomen	13684-15367	1683	7
AV11p	15100-18700	3600	15
AV7 Gomen	18115-21387	3272	14(*)
AV11p	21000-27200	6200	26
AV7 Gomen	26578-31899	5321	18(*)
AV11p	31300-34794	3494	15

En las celdas marcadas con asteriscos adicionalmente se hizo uso de *primers* de PCR de tipificación disponibles para familia, especie, subespecie y cluster (Hierholzer *et al.*, 1993; Kidd *et al.*, 1996, Xu *et al.*, 2000y Xu & Erdman, 2001)

Se obtuvieron templados de PCR por duplicado de 400 a 1200 pb., se corrieron en geles de agarosa al 2% en buffer TAE teñidos con bromuro de etidio, se purificaron por columnas comerciales y se cuantificaron por densitometría comparativa.

Secuenciación de productos de PCR

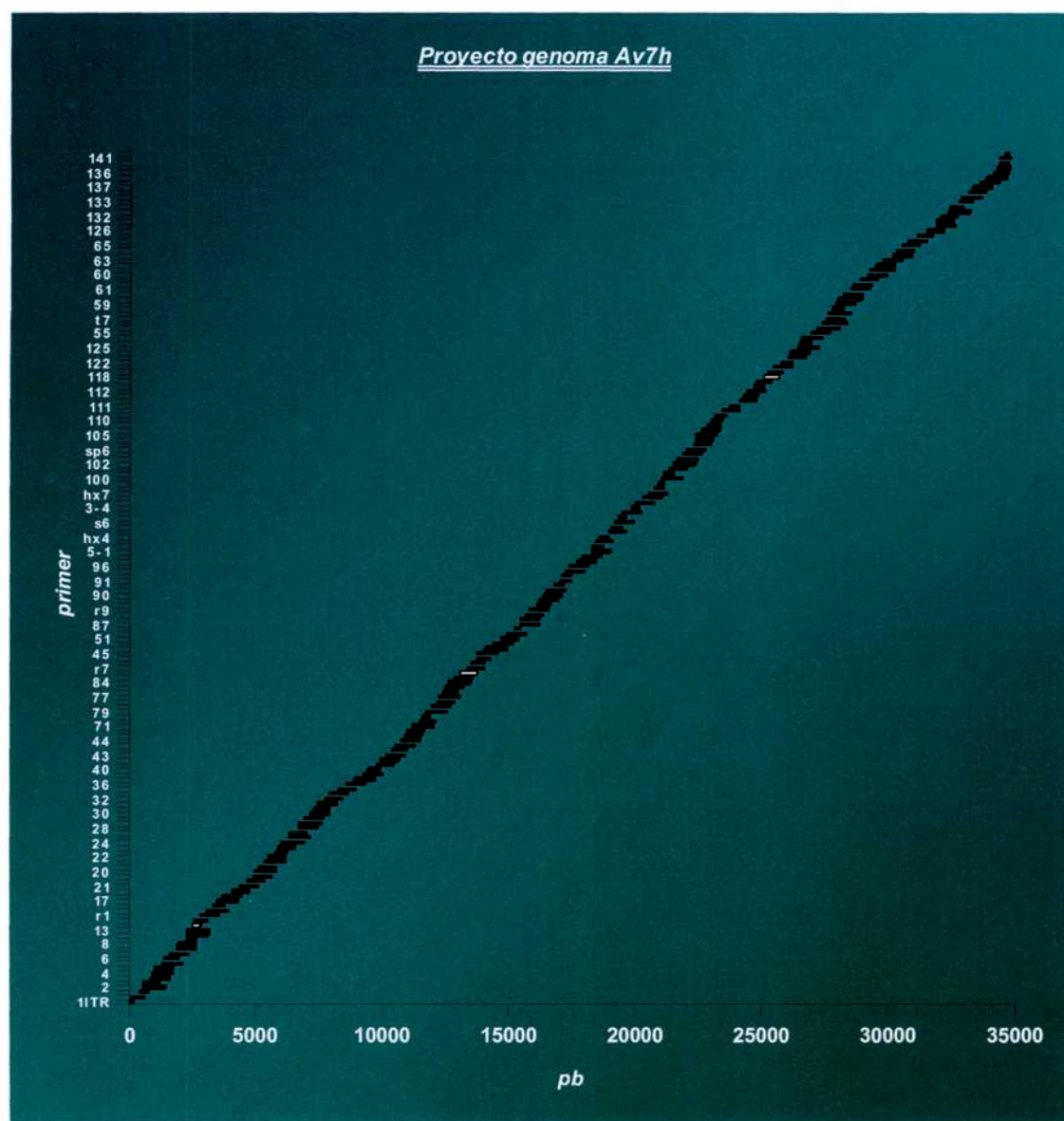
Se secuenciaron 80-150 ng de producto de PCR de acuerdo al peso molecular bajo las condiciones estándar en equipo automático capilar. Los electroferogramas obtenidos se visualizaron con el programa Chromas v 1.45 y se alinearon con el programa ClustalX v1.83.

Los resultados de secuenciación pueden verse en la Figura 40. Se tomó para cada *primer* (eje y) su punto de origen putativo (referido a la cepa AV11p) y la longitud del fragmento leído (eje x). Se puede apreciar que la redundancia es alta y la cobertura es total.

Obtención de la secuencia consenso para el genoma AV7h

Se alinearon primero las secuencias realizadas con el mismo *primer* para el mismo templado amplificado en dos reacciones de PCR independientes (duplicados). Se obtuvo una secuencia consenso para ese *primer*. Luego se integró esta secuencia en un alineamiento general ordenado en base al molde. Se realizó el mismo proceso para el *primer* complementario de ese par y así sucesivamente con los 155 *primers* diseñados. De esta manera se trabajó con una estrategia de *contigs* apilando pares de amplicones solapados hasta obtener una redundancia mínima de 4 (la misma base leída en dos amplificaciones independientes por dos *primers* 5'3' y dos *primers* 3'5'). A cada *contig* se lo designó de acuerdo a su punto de origen en nucleótidos basado en la secuencia completa del AV11p. Para cada *contig* se obtuvo una secuencia consenso.

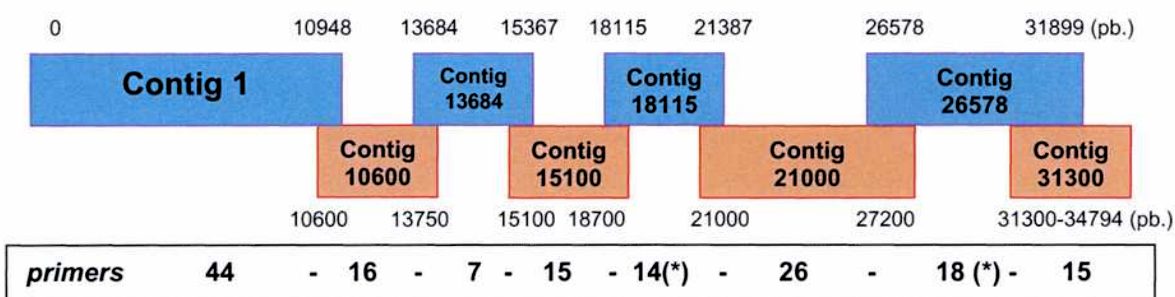
Figura 40: Cobertura de genoma AV7h



A su vez los *contigs* se solaparon entre sí en un alineamiento donde sus extremos aparecían en el *contig* continuo. Se obtuvo así una secuencia consenso final.

Se señalan en la Figura 41 en celeste los *contigs* cuyos *primers* fueron diseñados en base a secuencias de AV7 Gomen y en naranja los que fueron diseñados en base a AV11p. Se muestra el alcance de cada uno de ellos (en pb), cómo se solapan y se recuerda el número de *primers* utilizado para armar cada *contig*.

Figura 41: Esquema de los contigs del armado del genoma AV7h



Análisis de la secuencia completa del genoma de AV7h

La secuencia completa del AV7h consta de 35263 nucleótidos. La composición en bases se ve en la Tabla 20

Tabla 20: Composición de bases del genoma AV7h

Base	Cantidad	Porcentaje
G	8891	25
A	8961	25
T	8361	24
C	9046	26

El contenido de G+C es 51%. Se comparó la secuencia completa obtenida con los genomas de la especie B disponibles en bases de datos públicas (Tabla 21).

Tabla 21: Comparación con secuencias de especie B del genoma AV7h

Genoma	GenBank	Longitud	Homología
HAdV-B	NC_004001	35524	84%
AV11p	AY163756	34794	84%
AV7vacunal	AY495969	35514	96%

La diferencias en la homología están dadas porque la cepa Av11p y HAdV-B (Adenovirus de simio 21, o humano especie B) pertenecen a la subespecie B:2 mientras que la cepa AV7 vacunal pertenece a la misma subespecie que el AV7h (subespecie B:1).

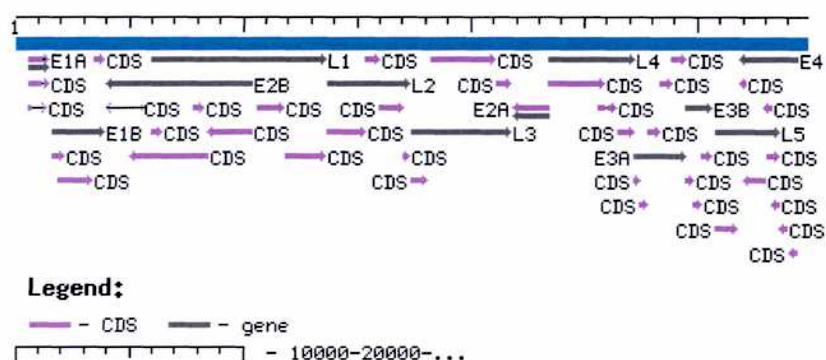


Figura 44: Mapa de transcripción de AV/h (genes en gris y proteínas en rosa).

Como hemos visto gráficamente, la organización genómica corresponde a la descrita para Adenovirus humanos. Se constata la presencia de todos los genes tempranos, intermedios y tardíos en las ubicaciones previstas.

Los ITRs se ubicaron en ambos extremos con una longitud de 136 pb., conservando el motivo ATAATATACC de unión a la proteína terminal (TP) y a la DNA polimerasa durante la replicación.

La señal de ensamblado (*packaging*) descrita para los nucleótidos 247-387 para AV11p se encuentra conservada, manteniendo una homología de 83.6% con éste y 100% con la cepa reportada de AV7 vacunal.

Se encontró también la secuencia de VA RNA_I y VA RNA_{II}. Se localizan desde los nucleótidos 10306 hasta 10890.

Expresión génica

Genes tempranos

Para el gen E1, se encontraron dos unidades de transcripción: E1A y E1B. E1A presentó dos proteínas de 24K y 28K y E1B dos proteínas de 21K y 55K. La proteína de 21K es el precursor de la de 19K.

E2 presenta dos regiones, E2A y E2B separadas físicamente en la cadena negativa. E2A codifica la DBP (*DNA binding protein*) proteína de unión al DNA viral y E2B codifica la DNA polimerasa y la proteína IVa2. Estas proteínas son esenciales para la replicación.

E3 codifica 8 proteínas (12.1K, 16.1K, 18.5K, 20.1K, 20.5K, 10.3K, 14.9K y 15.3K).

E4 presenta 7 marcos de lectura diferentes en la cadena negativa.

Genes intermedios

Los genes intermedios producen dos proteínas pIX y pIVa2. La proteína pIVa2 comparte el extremo N-terminal con la DNA polimerasa.

Genes tardíos

Se hallaron las 5 familias de genes tardíos (L1-L5) en la región central del genoma.

L1 codifica las proteínas de 52/55K y pIIIa como productos principales.

L2 codifica las proteínas III (base del pentón), V, VI, pVII y X (μ).

L3 codifica el hexón (II) y la proteasa viral de 23K.

L4 presenta 4 ORFs, con proteínas 100K, 33K y pVIII.

L5 codifica la proteína IV (fibra).

Se detallan en la Tabla 22 la localización de los genes, de sus ORFs y las proteínas que codifican.

Tabla 22: Localización física y propiedades de los ORFs de AV/h

Gen	Cadena	Principio	Fin	Longitud	ORF	Proteína	aa
E1A	+	576	1169	594	3	28K+24K	197
	+	1212	1454	243	3	28K+24K	80
E1B	+	1602	2138	537	3	21K(19K)	178/173
	+	1907	3184	1278	2	55k	425
INTERMEDIO	+	3405	3896	492	3	IX	163
E2B	(-)	3949	5160	1212	-1	IVa2	448
	(-)	5052	8420	3369	-2	DNA polim	1122
L1	+	6145	6465	321	1	11,5K	106
	+	7830	8585	756	3	15,3 agnoproteína	251
E2B	(-)	8423	10345	1923	-3	pTP	640
L1	+	10834	11412	579	1	52/55k	192
	+	11366	11752	387	2	55K	128
	+	11694	11993	300	3	52/55K	99
	+	12019	12378	360	1	pIIIa(43K)	119
	+	12372	13787	1416	3	pIIIa	471
L2	+	13867	15498	1632	1	Base penton(III)	543
	+	15510	16088	579	3	pVII	192
	+	16641	17090	450	1	V (39K)	149
	+	17264	17491	228	3	Px(μ)	75
L3	+	17565	18317	753	1	pVI	250
	+	18430	21234	2805	2	Hexón (II)	934
	+	21271	21882	612	3	23K Pr	203
E2A	(-)	21990	23543	1554	-1	DBP	143
L4	+	23704	25509	1806	1	100k	590
	+	25272	25505	234	3	7,3K(Asoc Hx)	77
	+	25771	26337	567	3	22K	188
	+	26700	27383	684	2	VIII (33K)	227
E3	+	27383	27703	321	1	12,1K	106
	+	27657	28097	441	2	16,1K	146
	+	28082	28600	519	1	gp19k(18,5K)	172
	+	28630	28974	345	3	20,1K	114
	+	29182	29751	570	3	gp20,5k	189
	+	29983	30258	276	3	10,3K	91
	+	30230	30667	438	1	14,9K(p)	146
	+	30660	31070	411	2	15,3K-p14K	136
U	(-)	31115	31276	162	-1	U exon	53
L5	+	31291	32250	960	3	Fibra(IV)	319
E4	(-)	32291	32542	252	-1	ORF6/7	83
	(-)	32539	33438	900	-2	ORF6	299
	(-)	33341	33709	369	-1	ORF4(13,6K)	122
L5	+	33564	34073	510	2	IV agnoproteína	169
E4	(-)	33718	34071	354	-2	ORF3(13K)	117
	(-)	34068	34502	435	-3	ORF2(14,3K)	144
	(-)	34499	34876	378	-1	ORF1(13,9K)	125

Análisis de recombinación

Se utilizaron dos métodos para medir si existía recombinación entre los genomas secuenciados de Adenovirus. Se analizaron las secuencias de un genoma de cada especie humana disponible en las bases de datos públicas (A-E) junto con los del AV11p y AV7 vacunal. Los resultados fueron negativos por ambos métodos. Esto no es sorprendente ya que un ancestro no muy lejano como el AV11p se postula como derivativo del AV35 (D) con la sustitución de la fibra por AV7 (Stone *et al.*, 2003). La probabilidad que cambios tan abruptos ocurran en un lapso de tiempo corto en virus DNA no debería ser alta, a diferencia de lo que sucede en virus RNA (Domingo & Holland, 1997).

Inferencia filogenética

Para reflejar las relaciones filogenéticas entre los genomas disponibles en las bases de datos públicas se testearon los 56 modelos de sustituciones posibles para develar cuál de ellos estaba operando en Adenovirus mediante el programa Modeltest v3.06 (Posada & Crandall, 1998). La composición estimada de bases fue

freqA = 0.4043

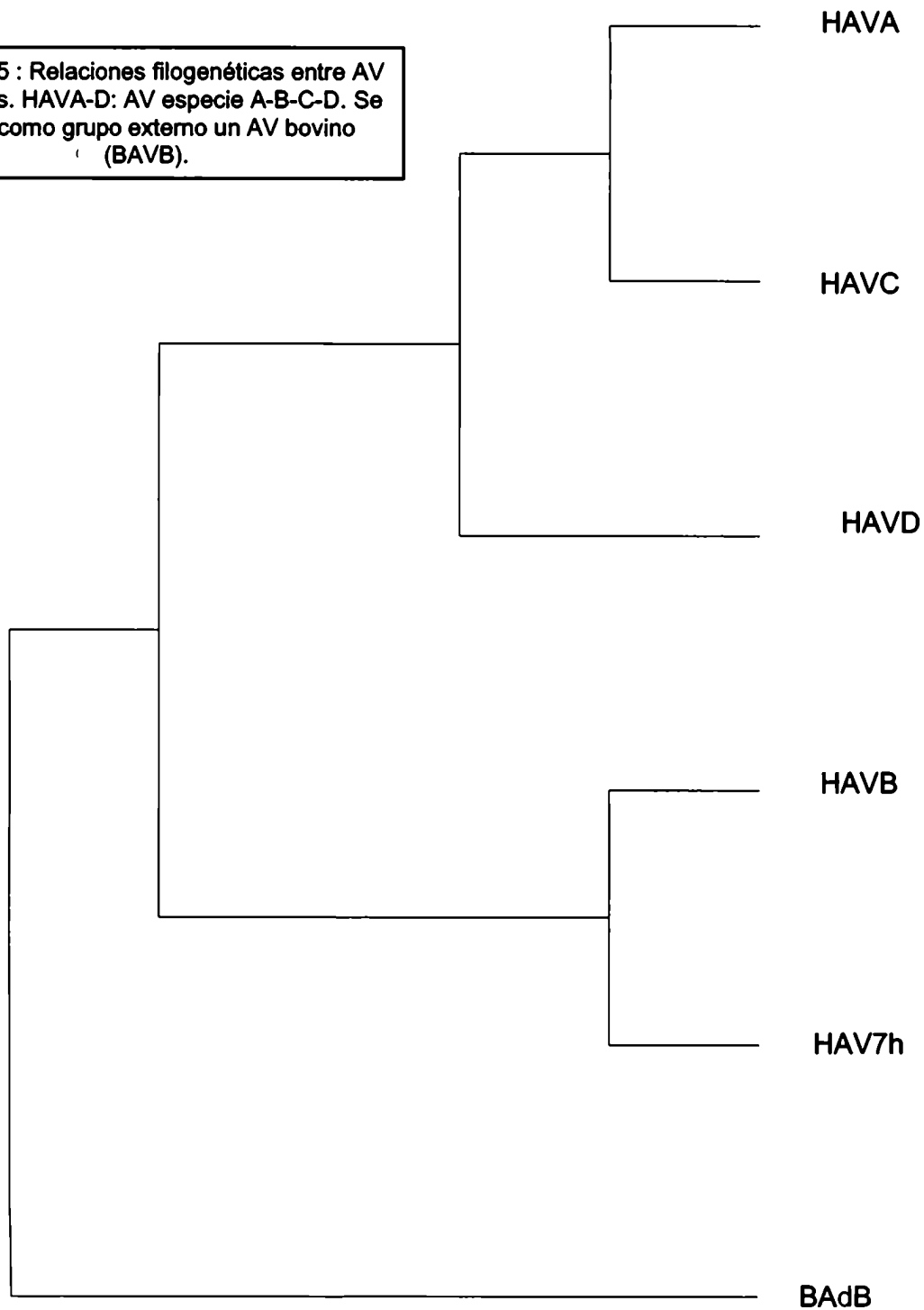
freqC = 0.1639

freqG = 0.1654

freqT = 0.2665

El modelo más probable resultó F81+G, que toma todas las tasas de sustituciones como equiprobables. Estos parámetros se utilizaron para construir un árbol por el método de *Maximum Likelihood* que genera la filogenia más probable de acuerdo al modelo evolutivo que esté operando (Figura 45).

Figura 45 : Relaciones filogenéticas entre AV humanos. HAVA-D: AV especie A-B-C-D. Se tomó como grupo externo un AV bovino (BAVB).



Regiones antigénicas

Se analizaron las regiones antigénicas de las proteínas fibra, base del pentón y E1A con el programa Protean. Este programa permite analizar la probabilidad de un péptido de ser presentado en un contexto T o B. Los resultados obtenidos se detallan en la Tabla 23.

Tabla 23: Regiones antigénicas en las proteínas E1A, Fibra, Penton y Hexón

E1A (197aa)		Fibra (319aa)		Pentón (543aa)		Hexón (934aa)	
B	T	B	T	B	T	B	T
	5-8, 10-13,19-26 47-50, 62-70, 87-90 93-97, 109-112 137-141, 145-152 190-194	5-10 54-59 69-74 84-89 156-161 167-172 207-212 216-221 267-272 303-311	31-34 37-40 44-47 68-75 117-121 165-168 178-181 194-205 228-231 295-302	144-149 238-243 312-317 421-426 458-463 513-518 524-534 537-542	3-7, 64-67 77-80 123-127 135-139 145-149 155-163 167-170 172-175 190-193 208-211 242-245 257-260 280-284 306-310 330-334 345-345 415-418 466-469 528-531	12-17 21-26 33-38 57-62 83-88 121-126 178-183 427-432 456-461 526-531 589-594 750-755 833-841 876-883 921-929	15-18 23-26 28-31 34-41 56-60 64-68 93-99 106-109 119-123 151-154, 170-174 182-185 208-211, 225-228, 262-268 286-290, 333-337, 369-372 386-389, 420-424 434-438 466-469, 485-488, 490-493 588-592 614-618 628-631, 699-702, 719-722 730-733 773-776, 810-813 851-854, 866-870 882-889 904-910 917-920

En algunos casos los epitopes B y T predichos colocalizan o se solapan como se muestra en la tabla. Es importante notar que la región E1A no presenta epitopes B, probablemente por ser un antígeno inmediatamente temprano. Las regiones antigénicas se sometieron a una búsqueda en la base de datos pública.

Mimetismo molecular

Los péptidos hallados se cotejaron contra la base de datos pública de proteínas para comprobar si alguno de ellos presentaba homología con proteínas pulmonares. Los resultados fueron negativos.

Cabe destacar que cuando se realizó el análisis de los ORFs probables para el genoma completo descrito en una sección anterior, 2 de ellos presentaban homología con proteínas pulmonares (Tabla 24)

Tabla 24: Péptidos de AV7h homólogos con proteínas de mamíferos

Gen	Cadena	Principio	Fin	Longitud	ORF	Proteína	aa
HUMANO	(-)	18095	18271	177	-2	ELASTINA	58
MURINO	(-)	22564	22740	177	-1	MIOSINA	58

Es importante destacar que estos péptidos no se encuentran dentro de ninguna proteína viral, ni forman parte del virión. Sólo se postula que podrían traducirse preferencialmente dentro del ciclo replicativo viral, aprovechando el *shutoff* de la traducción proteica celular. Estos péptidos al ser presentados dentro de un contexto infeccioso, ante un sistema inmune alerta podrían burlar la barrera de tolerancia y generar una primera respuesta de alerta contra ellos, que una vez desaparecida la infección productiva permanezca y se instale en el epitelio respiratorio. Así, los anticuerpos contra esos péptidos, por ejemplo, podrían actuar como autoanticuerpos y reconocer proteínas como la elastina y la miosina presentes en músculo y matriz del epitelio pulmonar llevando un deterioro irreversible del órgano, y una hiperreactividad generalizada. Este pulmón, si no lograra deshacerse

prontamente de la infección podría sufrir secuelas que permanezcan de por vida, disminuyendo su capacidad y generando daño pulmonar crónico postviral.

Conclusiones

Los Adenovirus circularon en el período 1999-2002 durante todo el año, registrándose un mayor número de casos en el período invernal de baja temperatura y primaveral con cambios de temperatura frecuentes.

El diagnóstico virológico se realizó tempranamente luego de la internación siendo la bronquiolitis el diagnóstico clínico al ingreso más frecuente. Se registró un mayor número de casos en menores de 12 meses y la evolución en general fue favorable registrándose 9 casos fatales en un total de 362 muestras analizadas.

Se encontró una correlación negativa de grado medio con respecto a la temperatura, indicando un aumento en el número de casos cuando el registro térmico desciende. En cuanto a la humedad relativa y el índice UVB, la relación no es tan clara aunque es siempre de grado menor. Este parece ser un parámetro no necesariamente asociado a la aparición de casos de Adenovirus.

El análisis filogenético del hexón completo permitió agrupar las muestras fatales y las separó de las graves y leves en el período 1990-1999. El análisis de las regiones hipervariables en el período 1990-2002 posicionó a las muestras analizadas en el *cluster* GTC2 junto al AV7h.

La secuencia completa del AV7h se llevó a cabo por una estrategia de ensamblado de 8 contigs que permitieron elaborar una secuencia consenso de 35259 pb. con un porcentaje G+C del 51%. Se conservaron intactos los genes para las regiones tempranas, intermedias y tardías. El análisis filogenético ubicó al genoma AV7h junto con el AVB con el que tiene una homología del 84%. Se hallaron ORFs relacionados con proteínas pulmonares que podrían inducir autoanticuerpos o daño pulmonar crónico postviral.

Bibliografía

- Athappilly FK, et al. (1994). J Mol Biol 242:430-455.
- Benkö M, Harrach B (2003). Curr Top Microbiol Immunol 272:3-35.
- Berget SM, Moore C, Sharp PA (1977). Proc Natl Acad Sci USA 74:3171-3175.
- Boulanger P. (1999). J Soc Biol 193:77-84.
- Borodovsky M y McIninch J (1993). Computers & Chemistry 17:123-133.
- Büchen-Osmond C. (2003). Computing in Science and Engeneering 5:16-25.
- Cann AJ (1997). Principles of molecular virology. Academic Press, 2nd Edition
- Crawford-Miksza L y Schnurr DP (1996). J Virol 70:1836-44.
- Chiocca S, Kurzbauer R, Schaffner G et al. (1996).J Virol 70:2939-2949
- Chroboczek J, Bieber F, y Jacrot B(1992). Virol.186:280-285.
- Davison AJ, Wright KM, Harrach B. (2000). J Gen Virol. 81:2431-9.
- Davison AJ, Benkö M, Harrach B. J Gen Virol. 2003 84:2895-908.
- Dix BR, Edwards SJ y Braithwaite AW (2001). J. Virol 75: 5443-5447.
- Domingo E y Holland JJ (1997). Ann Rev Microbio. 51, 151-178.
- Erdman DD, Xu W, Gerber SI, Gray GC, Schnurr D, Kajon AE y Anderson LJ (2002). Emerg Infect Dis 8: 269-277.
- Felsestein J (1989). Cladistics 5, 164-166.
- Gaggar A, Shayakhmetov DM, Lieber A (2003). Nat Med. 9:1408-12.
- Gingeras TR, Sciaky D, Gelinas RE, et al. J Biol Chem (1982) 257: 13475-13491
- Ginsberg HS, Moldawer LL, Sehgal PB et al. (1991). Proc Natl Acad Sci USA, 88:1651-1655
- Goodbourn S, Didcock L y Randall RE (2000). J Gen Virol 81:2341-2364.
- Harrach B y Benkö M (1998).Methods Mol Med 21:309-339.

- Howitt J, Anderson CW, Freimuth P (2003). *Curr Top Microbiol Immunol* 272:331-64.
- Hashido M, Mukouyama A, Sakae K, Tsuzuki H, Yamashita T, Inada T y Inouye S. (1999). *Epidemiol Infect.* 122:281-6.
- Hierholzer JC, et al.. *J Clin Microbiol.* (1993) 31:1886-91.
- Hogg JC, Irving WL, Porter H. et al (1989). *Am. Rev. Respir. Dis.* 139:1531-35.
- Hogg JC (2001). *Am J Respir Crit Care Med* 164:S71–S75.
- Kajon AE, Wadell G (1992). *Intervirology* 33:86-90.
- Kajon A, et al(1994). *J Clin Microb* 32 :2321-2323.
- Kidd AH, et al..*J Clin Microbiol* (1996) 34:622-7.
- Ladenheim HS, Mistchenko AS, Drut R (1995).*Pediatr. Pathol. y Lab. Med.* 15:291-298.
- Li QG, Lindman K, Wadell G (1997). *Arch Virol*;142:1307-22.
- Li QG y Waddell G (1999). *Arch Virol* 144:1739-1749.
- Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC (1999). *J Virol* 73:152–160.
- Lukashok SA, Horwitz MS (1998). *Curr Clin Top Infect Dis* 18:286–305.
- Ma Y y Mathews MB (1996). *J Virol* 70:5083–5099.
- Macek V, Sorli J, Kopriva S, Marin (1994). *J Am J Respir Crit Care Med* 150:7-10.
- Martin D y Rybicki E (2000). *Bioinformatics* 16, 562-563.
- Mistchenko AS, Lenzi HL, Mota EM, et al. (1992). *Acta Pediatr* 81:993-98.
- Mistchenko AS, Diez RA, Mariani AL, et al (1994). *J Pediatr* 124:714
- Mistchenko AS, Robaldo JF, Rosman FC, Koch ER, Kajon AE (1998). *J Med Virol* 54:233-6.
- Morrison MD, Onions DE, Nicolson L (1997)..*J Gen Virol* 78: 873-878.

- Nemerow GR y Stewart PL (1999). *Microbiol Mol Biol Rev* 63:725–734.
- Nemerow GR (2000). *Viol* 274:1–4.
- Niel C, Moraes MTB, Mistchenko AS, Leite JP, Gomes S (1991). *J Med Virol* 33:123-27.
- Page RDM (1996). *Comput. Appl. Biosci.* 12, 357–358. 589.
- Pickles RJ, McCarty D, Matsui H, Hart PJ, Randell SH y Boucher RC (1998). *J Virol* 72:6014-6023.
- Posada D y Crandall KA (1998). *Bioinformatics* 14: 817–818.
- Prince GA, Porter DD, Jenson AB et al. (1993). *Viol* 67:101-111.
- Pring-Akerblom P, et al. (1995). *Res Virol* 146:383-388.
- Roberts RJ, O'Neill KE y Yen CT (1984). *J Biol Chem.* 259:13968-13975.
- Roberts MM, et al (1986). *Science* 232:1148-1151.
- Rosman FC, Mistchenko AS y Ladenheim HS (1996). *Pediatr Pathol Lab Med* 16:521-541.
- Rozas, J., Rozas, R., (1999). *DNASP version 3: Bioinformatics* 15, 174–175.
- Russell WC (2000). *J Gen Virol* 81:2573-604.
- Sambrook J y Russell D (2001) *Molecular Cloning: A Laboratory Manual*. 3rd Edition Cold Spring Harbor Laboratory Press, Woodbury NY, USA.
- Sanger F et al. (1977) *Proc Nat Acad Sc* 74 :5463-5467.
- San Martin C y Burnett RM (2003). *Curr Top Microbiol Immunol* 272, 57–94.
- Segerman A, Arnberg N, Erikson A, Lindman K, Wadell G (2003). *J Virol* 77:1157-62.
- Shenk T (2001). *Adenoviridae: the viruses y their replication*. In *Fields Virology*, 4th edn, vol. 2: Cap67. Editado por D. M. Knipe y P. M. Howley. Philadelphia: Lippincott Williams y Wilkins.

Smith CA, Woodruff LS, Rooney C y Kitchingman GR (1998). *Hum Gene Ther* 9:1419–1427.

Stewart PL, et al. (1991). *Cell* 67:145-154.

Stone D, Furthmann A, Sandig V y Lieber A (2003). *Virology* 309:152-165.

Takeuchi S et al. (1999) *J Clin Microbiol* 37:1839–1845.

Tanaka K, Itoh N, Saitoh-Inagawa W et al. (2000) *J Med Virol* 61:143-149.

Thibeault DW, Ekekezie MSM y Truog WE (2000). *Pediatrics* 106:1452-59.

Thompson, JD, et al. (1997). *Nucleic Acids Res* 24:4876-4882.

Tollefson AE, Hermiston TW y Wold WSM (1998). Preparation and titration of CsCL-Banded Adenovirus Stock in *Methods in Molecular Medicine Series vol21: Adenovirus Methods and Protocols*, edited by WSM Wold, Humana Press, Totowa, NJ, USA.

Valentine RC, Pereira HG (2003). Reprinted from *J. Mol. Biol.* 196513: 13-20. *Rev Med Virol* 13:71-82.

Viegas M, Barrero PR, Maffey AF y Mistchenko AS. (2003) *Journal of Infection en prensa*.

Wadell G (1984). *Curr Top Microbiol Immunol* 110:191-220.

Wadell G, Cooney MK, da Costa Linhares A, et al. (1985). *J Clin Microbiol.* 21:403-408.

Wenman WM (1982). *Chest* 81:606-9

Wohlfart C, et al. (1988) *J Virol* 62:2321-2328.

Xu W, et al.. *J Clin Microbiol* (2000) 38:4114-20

Xu W y Erdman DD. *J Med Virol* (2001) 64:537-42.

Zhang H y Phan SH (1999). *Cell Mol Biol* 21:658-665.

Zhang L, Peebles M, Boucher RC, Collins PL y. Pickles RJ (2002). J Virol 76:5654-5666.

Zhou MY, Clark SE y Gomez-Sanchez CE (1995). BioTechniques 19, 34–35.

Addendum

>AV7h cepa 87-922, secuencia completa

```

1 CTCTCTATAT AATATACCTT ATAGATGGAA TGGTGCCAAC ATGTAAATGA GGTAATTTAA
61 AAAAGTGC GC GCTGTGTGGT GATTGGCTGT GGGGTGAACG GCTAAATGG GCGGGGCGGC
121 CGTGGGAAAA TGACGTGACT TATGTGGGAG GAGCTATGTT GCGAATTATT GCGGTAAATG
181 TGACGTAAAA CGAGGTGTGG TTTGAACACG GAAGTAGACA GTTTCCAC GCTTACTGAC
241 AGGATATGAG GTAGTTTTGG GCGGATGCAA GTAAAAATTC TCCATTTTCG CGCGAAAACT
301 GAATGAGGAA GTGAATTTCT GAGTCATTTT GCGGTATATGA CAGGGTGGAG TATTTGCCGA
361 GGGCCGAGTA GACTTTGACC GTTTACGTGG AGGTTTCGAT TACCGTGGTT TTTCCCTAAA
421 TTTCCGCGTA CGGTGTCAAA GTCCTGTGTT TTTACGTAGG TGTCAAATGA TCGCTAGGGT
481 ATTTAAACCT GACGAGTTCC ATCAAGAGGC CACTCTTGAG TGCCAGCGAG AAGAGTTTTT
541 TCCTCCGCGC CGCAAGTCAG TTCTGCGCTT TGAAATGAG ACACCTGCGT TTCCTGCCAC
601 AGGAGATTAT CTCCAGTGAA ACCGGGATTG AAATACTGGA GTTTGTGGTA AATACCCTAA
661 TGGGAGACGA CCCGGAACCG CCAGTGCAGC CTTTGTATCC ACCTACGCTG CACGATCTGT
721 ATGATTAGGA GGTAGACGGG CCGGAGGATC CCAATGAGGA AGCTGTGAAT GGGTTTTTTA
781 CTGATTCTAT GCTGCTAGCT GCCGATGAAG GATTGGACAT ATACCCTCCT CCTGAGACCC
841 TTGATACCCC AGGGGTGGTT GTGGAAAGCG GCAGAGGTGG GAAAAAATTG CCTGATCTGG
901 GAGCAGCTGA AATGGACTTG CGTTGTTATG AAGAGGGTTT TCCTCCGAGT GATGATGAAG
961 ATGGGGAAAT TGAGCAGTCC ATCCATACCG CAGTGAATAA GGGAGTAAAA GCTGCCAGCG
1021 ATGTTTTTAA GTTGGACTGT CCGGAGCTGC CTGGACATGG CTGTAAGTCT TGTGAATTTT
1081 ACAGGAATAA CACTGGAATG AAAGAACTAT TGTGCTCGCT TTGCTATATG AGAATGCACT
1141 GCCACTTTAT TTACAGTAAG TGTATTTAAG TGAATTTAA AGGAATAGTG TAGCTGTTTA
1201 ATAACTTTTG AATGGTAGAT TTATATTTTT ACTTGCATTT TTTGTAGGT CCTGTGCTG
1261 ATGATGAGTC ACCTTCTCCT GATTCAACTA CTTCACTTCC TGAAATTCAG GCGCCCGCAC
1321 CTGCAACCGT ATGCAAGCCC ATTCTGTGTA AGCCTAAGCC TGGGAAACGC CCTGTGTGG
1381 ATAAGCTTGA GGACTTGTG GAGGGTGGGG ATGGACCTTT GGACCTTAGT ACCCGGAAAC
1441 TCCCAAGGCA ATGAGTGCCC TGCACTGTG TTTATTTAAT GTGACGTCAT GTAATAAAAT
1501 TATGTCAGCT GCTGAGTGT TTATTACTTG TTGGGTGGGG ACTTGGATAT ATAAGTAGGA
1561 GCAGATCTGT GTGGTTAGCT CACAGCAACC TGCTGCCATC CATGGAGGTT TGGGTATCT
1621 TGAAGACCT CAGACAGACT AGGCTACTGC TAGAAAACGC CTCGGACGGA GTCTCTGGCC
1681 TTTGGAGATT CTGGTTCCGT GGTGATCTAG CTAGGCTAGT GTTTAGGATA AAACAGGACT
1741 ACAGGAAGA ATTTGAAAAG TTATTGGACG ACAGTCCAGG ACTTTTTGAA GCTCTTAACT
1801 TGGGTATCA GGCTCATTTT AAGGAGAAGG TTTATCAGT TTTAGATTTT TCTACTCCTG
1861 GTAGAAGTGC TGCTGCTGTA GCTTTTCTTA CTTTATATTT GGATAAATGG ATCCGCCAAA
1921 CTCATTTCAG CAAGGGATAC GTTTTGGATT TCATAGCAGC AGCTTTGTGG AGAACATGGA
1981 AGGCTCGCAG GATGAGGACA ATCTTAGATT ACTGGCCAGT GCAGCCTCTG GGAGTAGCAG
2041 GGATACTGAG ACACCCACCG ACCATGCCAG CGTTCTGCA GGAGGAGCAG CAGGAGGACA
2101 ATCCGAGAGC CGGCCCTGGAC CCTCCGGTGG AGGAGTAGCT GACCTGTTTC CTGAAGTGGC
2161 ACGGGTGCTT ACTAGGTCTA CGACCACTGG ACAGAACAGG GGCATTAGA GGGAGAGGAA
2221 TCCGTATGGG AACAAATCAA GAACCGAGTT GGCTTTAAGT TTAATGAGCC GCAGGCGTCC
2281 TGAAGCTGTT TGGTGGCATG AGGTTCAAGC CGAAGGCAGG GATGAAGTTT CAATATTGCA
2341 GGAGAAATAT TCACTAGAAC AACTTAAGAC CTGTTGGTTG GAACCTGAGG ATGATTGGGA
2401 GGTGGCCATT AGGAATTATG CTAAGATATC TCTGAGGCTT GATAAACAA ATAGAATTAC
2461 TAAGAAGATT AATATTAGAA ATGCATGCTA CATATCAGGG AATGGGCGAG AGGTTATAAT
2521 AGATACACAA GATAAAGCAG CTTTATAGAT TTGTATGATG GGTATGTGGC CAGGGGTGTG
2581 CGGCATGGAA CGAGTAACAC TTATGAATAT TAGGTTTAGA GGGGATGGGT ATAATGGCAT
2641 TGTATTTATG GCTAACACTA AGCTGATTCT ACATGGTTGT AGCTTTTTTG GGTTTAATAA
2701 TACGTGTGTA AAGCTTGGG GGCAAGTTAG TGTGAGGGGT TGTAGTTTTT ATGCATGCTG
2761 GATTGCAACA TCAGGTAGGG TGAAGAGTCA GTTGTCTGTA AAGAAATGCA TGTTTGAGAG
2821 ATGTAATCTT GGCATACTGA ATGAAGGTGA AGCAAGGGTC CGCCACTGCG CGGCTACACA
2881 AACTGGCTGC TTCATTCTAA TAAAGGGAAA TGCCAGTGTA AAGCATAATA TGATCTGTGG
2941 ACATTCCGAT GAGAGGCCTT ATCAGATGCT GACCTGCGCT GGTGGACATT GCAATATTCT
3001 TGCTACCGTG CATATCGTTT CACATGCACG CAAGAAATGG CCTGTATTTG AACATAATGT
3061 GATTACCAAG TGCACCATGC ACATAGGTGG TCGCAGGGGA ATGTTTATGC CTTACCAAGT
3121 TAACATGAAT CATGTGAAGG TAATGTTGGA ACCAGATGCC TTTTCCAGAG TGAGCGTACA
3181 ATGAATCTTT GATATGAATG TTCAACTATG GAAGATCCTG AGATATGATG AACTAAACC
3241 AAGGGTGC GC CATGCGAAT GCGGAGGCTA GCATGCTAGA TTCCAGCCGG TGTGCGTGGA
3301 TGTGACTGAA GACCTGAGAC CCGATCATT GGTGCTTGCC TGCACTGGAG CGGAGTTTGG
3361 TTCTAGTGGT GAAGAACTG ACTAAAGTGA GTAGTGGGGG CAAGATGTGG ATGGGACTT
3421 TCAGGTTGGT AAGGTGGACA AATTGGGTAA ATTTGTGTTA TTCTGTCTT GCAGCTGCCA
3481 TGAGTGAAG TGCTTCTTTT GAGGGGGGAG TATTTAGCCC TTATCTGACG GGCAGACTCC
3541 CACCATGGGC AGGAGTTTCGT CAGAATGTCA TGGGATCTAC TGTGGATGGG AGACCCGTCC
3601 AGCCCGCCAA TTCCTCAACG CTGACCTATG CCACTTTGAG TTCGTACCA TTGGATGCAG
3661 CTGCAAGCAG CGCCGCTACT CTGCGCGCCA ACACCATCCT TGGAATGGGC TATTACGGAA
3721 GCATCGTTGC CAATTCCAGT TCCTCTAATA ACCCTTCAAC CCTGGCTGAG GACAAGCTAC
3781 TTGTTCTCTT GGCTCAGCTC GAAGCCTTAA CCCAACGCTT AGGCGAACTG TCTAAGCAGG
3841 TGGCCAGTT GCGTGAGCAA ACTGAGTCTG CTGTTGCCAC AGCAAAGTCT AAATAAAGAT
3901 CTCAAATCAA TAAATAAAGA AATACTGATT ATGAAACCAA ATGAATGTTT ATTTGATTTT
3961 TCGCGCGCGG TATGCCCTGG ACCATCGGTC TCGATCATTG AGAACTCGGT GGATCTTTTC
4021 CAGTACCCTG TAAAGGTGGG ATTGAATGTT TAGATACATG GGCATTAGTC CGTCTCGGGG
4081 GTGGAGATAG CTCATTGAA GAGCCTCTTG CTCCGGGGTA GTGTTATAAA TCACCCAGTC
4141 ATAGCAAGGT CCGAGTGCAT GGTGTTGCAC AATATCTTTT AGGAGCAGAC TAATTGCAAC
4201 GGGGAGGCCC TTAGTGTAGG TGTTTACAAA TCTGTTAAGC TGGGACGGGT GCATTGCGGG

```

4261 TGAAATTATA TGCATTTTGG ACTGGATCTT AAGGTTGGCA ATGTTGCCGC CTAGATCCCC
 4321 TCTCGGGTTC ATATTGTGCA GGACCACCAA GACAGTGAT CTGTGCACT TGGGAAATTT
 4381 ATCATGCAGC TTAGAGGGAA AAGCATGAAA AAATTTGGAG ACGCCTTTGT GACCCCCCAG
 4441 ATTCTCCATG CACTCATCCA TAATGATAGC GATGGGGCCG TGGGCAGCGG CACGGGCGAA
 4501 CACGTTGCGG GGGTCTGAAA CATCATAGTT ATGCTCCTGA GTCAGGTCAT CATAAGCCAT
 4561 TTTAATAAAC TTTGGGCGGA GGGTGCCAGA TTGGGGAATG AAAGTTCCCT CTGGCCCGGG
 4621 AGCATAGTTT CCCTCACATA TTTGCATTTC CCAGGCTTTC AGTTCAGAGG GGGGGATCAT
 4681 GTCCACCTGC GGGGCTATAA AAAATACCGT TTCTGGAGCC GGGGTGATTA ACTGGGATGA
 4741 GAGCAAATTC CTAAGCAGCT GAGACTTGCC GCACCCAGTG GGACCGTAAA TGACCCCAAT
 4801 TACGGGTGTC AGATGGTAGT TTAGGGAGCG ACAGCTGCCG TCCTCCCGGA GCAGGGGGGC
 4861 CACTTCGTTT ATCATTTGCC TTACATGGAT ATTTTCCCGC ACCAAGTCCG TTAGGAGGCG
 4921 CTCTCCCCCA AGGGATAGAA GCTCCTGGAG CGAGGAGAAG TTTTTCAGCG GCTTTAGCCC
 4981 GTCAGCCATG GGCATTTTGG AAAGAGTCTG TTGCAAAGC TCGAGCCGGT CCCAGAGCTC
 5041 GGCAGTGTGC TCTATGGCAT CTCGATCCAG CAGACCTCCT CGTTTCGCGG GTTGGGACGG
 5101 CTCTCCGAGT AGGGAATCAG ACCGATGGCG TCCAGCGCTG CCAGGTCGCG ATCCTTCCAT
 5161 GGTGCGCAGC TCCGAGTTAG GGTGTGTTCC GTCACGGTGA AGGGGTGCGC GCCTGGTTGG
 5221 GCGCTTGCGA GGGTGCGCTT CAGACTCATC CTGCTGGTCG AGAACCCTG CCGATCGGCG
 5281 CCCTGCATGT CGGCCAGGTA GCAGTTTACC ATAAGTTCGT AGTTAAGCGC CTCGGCCGCG
 5341 TGGCCTTTGG TGGTGAAGT ACCTTTGGAA GTTTTATGGC AGGCGGGGCA GTAGATACAT
 5401 TTGAGGGCAT ACAGCTTGGG CGCGAGGAAA ATGGATTCGG GGGAGTATGC ATCCGCACCG
 5461 CAGGAGGCGC AGACGGTTTC GCACTCTACG AGCCAGGTCA GATCCGGCTC ATCCGGGTCA
 5521 AAAACAAGTT TTCCGCCATG TTTTGTGATG CGTTTCTTAC CTTTGGTTTC CATGAGTTCT
 5581 GTTCCCGGCT GAGTGACAAA GAGGCTGTCC GTGTCCCGT AGACCGACTT TATGGTCTG
 5641 TCCTCGAGCG GAGTGCCTCG GTCTCTTCG TAGAGGAACC CAGCCCACTC TGATACAAAA
 5701 GCGCGTGTCC AGGCCAGCAC AAAGGAGGCC ACGTGGGAGG GGTAGCGGTC GTTGTCAACC
 5761 AGGGGGTCCA CCTTCTCTAC GGTATGTAAA CACATGTCCC CCTCTCCAC ATCCAAGAAT
 5821 GTGATTGGCT TGTAAAGTGA GGCCACGTGA CCAGGGGTCC CCGCCGGGGG GGTATAAAAAG
 5881 GGGGCGGGCC TCTGTTCTGC CTCACTGTCT TCCGGATCGC TGTCCAGGAG CGCCAGCTGT
 5941 TGGGGTAGGT ATTCCCTCTC GAAGGCGGGC ATGACCTCTG CACTCAGGTT GTCAGTTTCT
 6001 AGGAACGAGG AGGATTGAT ATTGACAGTA CCAGCAGAGA TGCCTTTCAT AAGACTCTCG
 6061 TAGCCCTTGG CAGAAAACAC AATCTTCTTG TTGTCCAGCT TGGTAGCAAA TGATCCATAG
 6121 AGGGCATTGG ATAGAAGCTT GCGCATGGAG CGCATGGTTT GGTTCCTTTC CTTGTCCGCG
 6181 CGCTCCTTGG CGGCGATGTT AAGCTGGACG TACTCGCGCG CCACACATTT CCATTAGGG
 6241 AAGATGGTTG CAGTTCATC CGGAACTATT CTGACTCGCC ATCCCTTATT GTGCAGGGTT
 6301 ATCAGATCCA TACTGTGGC CACTCGCCT CGGAGGGGCT CATTGGTCCA GCAGAGTCGA
 6361 CCTCCTTTTC TTGAACAGAA AGGGGGGAGG GGGTCTAGCA TGAGCTCATC AGGGGGGTCC
 6421 GCATCTATGG TAAATATTCC CGGTAACAAA TCTTTGTCAA AATAGCTAAT GGTGGTGGGA
 6481 TCATCCAAGG TCATCTGCCA TTCTCGAACT GCCAGCGCGC GCTCATAGGG GTTAAGAGGG
 6541 GTGCCCGAGT GCATGGGTG GGTGAGCGCG GAGGCATACA TGCCACAGAT ATCGTATACA
 6601 TAGAGGGCTT CTTGAGGAT GCGCATGTAA GTGGGATAAC AGCGCCCCC TCTGATGCTT
 6661 GCTCGCATAT AGTCATAGAG TTCATGTGAG GGGGCGAGAA GACCCGGGCC CAGATTGGTG
 6721 CGGTTGGGTT TTTCCGCCCT GTAAACGATC TGGCGAAAGA TGGCATGGGA ATTTGAAGAG
 6781 ATAGTAGGTC TCTGGAATAT GTTAAATGG GCATGAGGTA GGCCTACAGA GTCCCTTATG
 6841 AAGTGGGCTT ATGACTCTTG CAGCTTGGCT ACCAGCTCGG CGGTGACGAG TACATCCAGG
 6901 GCACAGTAGT CGAGAGTTTC CTGGATGATG TCATAACCGG GTTGGCTTTT CTTTTCCAC
 6961 AGCTCGCGGT TGAGAAGGTA TTCTTCGCGA TCCTTCCAGT ACTCTTCGAG GGGAAACCCG
 7021 TCTTTTCTG CACGGTAAGA GCCCAACATG TAGAACTGAT TGAAGTCTT GTAGGGACAG
 7081 TAGGTGAGTA AGAGTATGCT TGGGCTGCAT TGCGCAGCGA GGTATGAGTG
 7141 AGGGCAAAAG TGTCCCTGAC CATGACTTTG AGGAATTGAT ACTTGAAGTC CATGTCATCA
 7201 CAGGCCCCCT GTTCCAGAG TTGGAAGTCC ACCCGCTTCT TGTAGGCGGG GTTGGGCAAA
 7261 GCGAAAGTAA CATCATGAA GAGGATCTTG CCGGCCCTGG GCATGAAAT TCGGGTGATT
 7321 CTGAAAGGCT GAGGACCTC TGCTCGGTTA TTGATAACCT GAGCGGCCAA GACGATCTCA
 7381 TCAAAGCCAT TGATGTTGTG CCCCCTATG TACAGTTCTA AGAATCGAGG GGTGCCCTTG
 7441 ACATGAGGCA GCTTCTTGAG TTCTTCAAAA GTGAGATCTG TAGGGTCAGT GAGAGCATAG
 7501 TGTTCGAGGG CCCATTCTGT CACGTGAGGG TTCGCTTTGA GGAAGGAGGA CCAGAGGTCC
 7561 ATGCGGAGTG CTGTTGTAA CTGTCCCGG TATTGACGAA AATGCTGCC GACTGCCATT
 7621 TTTTCTGGGG TGACGCAATA GAAGGTTTGG GGGTCTGCC GCCAGCGATC CCACTTAAGT
 7681 TTCATGGCGA GGTCTAGGCG GATGTTGACG AGCCGCTGGT CTCCAGAGAG TTTCATGACC
 7741 AGCATGAAGG GGATTAGCTG CTTGCCAAG GACCCCATCC AGGTGTAGGT TTCCACATCG
 7801 TAGGTGAGGA AGAGCTTTTC TGTGCGAGGA TGAGAGCCAA TCGGGAAGAA CTGGATCTCC
 7861 TGCCACCAGT TGGAGGAATG GCTGTTGATG TGATGGAAGT AGAACTCCCT GCGACGCGCC
 7921 GAGCATTCAT GCTTGTGCTT GTACAAACCG CCGCAGTACT CGCAGCGATT CACGGGATGC
 7981 ACCTCATGAA TGAGTTGTAC CTGACTTCCT TTGACGAGAA ATTTCAAGTG AAAATTGAGG
 8041 CCTGGCGTTT GTACCTGGCG CTCTACTATG TTGTCTGCAT CGGCATGACC ATCTTCTGTC
 8101 TCGATGGTGG TCATGCTGAC GAGCCCTCGC GGGAGGCAAG TCCAGACCTC GCGCGGCGAG
 8161 GGGCGGAGCT CGAGGACGAG AGCGCGCAGG CCGGAGCTGT CCAGGGTCCT GAGACGCTGC
 8221 GGAGTCAGGT TAGTAGGCAG TGTCAGGAGA TTGACTTGCA TGATCTTTTG GAGGGCGTGA
 8281 GAGAGGTTCA GATGTTACTT GATCTCCACG GGTCCGTTGG TGGAGATGTC AATGGCTTGC
 8341 AGGGTTCCGT GCGGCTTGGG CGCTACCACC GTGCCCTTGT TTTTCTTTT GGGCGGCGGT
 8401 GGCTCTGTTG CTTCTTGCA TTTTACGAGC GGTGTCGAGG GCGCGCACCG GCGGCTGGG
 8461 GCGGCTCGGG ACCCGCGGCG ATGGCTGGCA GTGGTACGTC GGCACCGCGC GCGGGTAGGT
 8521 TCTGGTACTG CGCCCTGAGA AGACTCGCAT GTGCGACGAC GCGGCGGTTG ACATCCTGGA
 8581 TCTGACGCCCT CTGGGTGAAA GCTACCGGCC CCGTGAGCTT GAACCTGAAA GAGAGTTCAA
 8641 CAGAATCAAT CTCGGTATCG TTGACGGCGG CTGCGCTAAG GATTTCTTGC ACGTCGCCAG
 8701 AGTTATCCTG TAGGGCGATC TCGACCATGA ACTGCTGGAT CTCTTCTCT TGAAGATCTC
 8761 CGCGGCGCGC TCTCTCGACG GTGCGGCGGA GGTGCTTGGG GATGCGCCCA ATGAGTTGAG
 8821 AGAATGCATT CATGCCCGCC TCGTTCAGGA CGCGGCTGTA GACCACAGCC CCCACGGGAT

8881 CTCTCGCGCG CATAACCACC TGGGCGAGGT TAAGCTCTAC GTGGCGGGTG AAGACCGCAT
 8941 AGTTGCGATAG GCGCTGGAAG AGGTAGTTGA GTGTGGTGGC GATGTGCTCG GTGACGAAGA
 9001 AATACATGAT CCATCGTCTC AGCGGCATCT CGTGACATC GCCAGCGCT TCCAAGCGCT
 9061 CCATGGCCTC GTAGAGTCC ACGGCAAAGT TGAAGAACTG GGAGTTACGC GCGGACACGG
 9121 TCAACTCCTC TTCCAGAAGA CGGATGAGTT CGGCAATGGT GGTGCGCACC TCGCGCTCGA
 9181 AATCCCGCGG GATTCTCTCC TCAATCTCTT CTCTCTCCAC TAACATCTCT TCCTCTTCAG
 9241 TTGGGGCTGC AGGAGGAGGG GGAACGCGGC GACGCGGCG GCGCACGGGC AGACGGTCTGA
 9301 TGAATCTTTC AATGACCTCT CCGCGCGCGC GCGCATGGT CTCGGTGACG GCACGACCGT
 9361 TCTCCCTGGG TCTCAGAGTG AAGACACCTC CCGCATCTC CCTGAAGTGG TGACTGGGAG
 9421 GCTCTCGGTT GGGCAGGGAC ACCGCGCTGA TTATGCATT TATCAATTGC CCCGTAGGTA
 9481 CTCGCGCGAA GGACCTGATC GTTCAAGAT CCACGGGATC TGAAAACCTT TCGACGAAAG
 9541 CGTCTAACCA GTCGCAATCG CAAGGTAGGC TGAGCACTGT ATCTTGCGGG CGGGGCGGGC
 9601 TAGACGCTCG GTGCGGGGTT TCTCTTTCTT TTCTTCTCTC CTCTTGGGAG GATGAGACGA
 9661 TGCTGCTGGT GATGAAATTA AAATAGGCAG TTTTGAGACG GCGGATGGTG GCGAGGAGCA
 9721 CCAAGTCTTT GGGTCCGGCT TGTGGATGC GCAGGCGATG GGCCATCCCC CAAGCATTAT
 9781 CCTGACATCT GGCCAGATCT TTATAGTAGT CTTGCATGAG TCGTTCCACG GGCACCTCTT
 9841 CTTGCGCCGC CCTGCCATGC ATGCGAGTGA TCCCGAACCC GCGCATGGGC TGGACAAGTG
 9901 CGAGGTCCGC TACAACCTT TCGGCGAGGA TGGCTTGCTG CACCTGGGTG AGGGTGGCTT
 9961 GGAAGTCTGC AAAGTCTACA AAGCGGTGGT AGGCCCCGGT GTTGATTGTG TAGGAGCAGT
 10021 TGGCCATGAC TGACCAGTTG ACTGTCTGGT GCCCAGGGCG CACGAGCTCG GTGTACTTGA
 10081 GCGCGCAGTA TGCGCGGGTG TCAAAGATGT AATCGTTGCA GGTGCGCACC AGGTACTGGT
 10141 AGCCGATGAG AAAGTGTGGC GGTGGCTGGC GGTACAGGGG CCATCGCTCT GTAGCCGGGG
 10201 CTCGCGGGCG AAGTTCTTCC AGCATGAGGC GGTGGTAACC GTAGATGTAC CTGGACATCC
 10261 AGGTGATACC GGAGGCGGTG GTGGATGCCC GCGGGAACCTC GCGTACGCGG TTCCAGATGT
 10321 TGCGCAGCGG CATGAAGTAG TTCAATGGTGC GCACGGTTTG GCCCGTGAGA CGTGACAGT
 10381 CGTTGATGCT CTAGACATAC GGGCAAAACG AAAGCGGTCA GCGGCTCGTC TCCGTGGCCT
 10441 GAGGCTTAAG CGAACGGGTT GGGTCCGCG TGTACCCCGG TTCGAATCTC GGATCAGGCT
 10501 GGAGCCCGAG CTAACGTGGT ACTGGCACTC CCGTCTCGAC CCAGGCCTGC ACAAACCTC
 10561 CAGGATACGG AGGCGGGTCG TTTTCTTGCT TTTTCTTGG ATGGGAGCCA GTGTGTCAAG
 10621 CTTTAGAACA CTCAGTTCTC GCGCCCGTAG TCTGGAGAA TAAATCGCAA GGTGCGTTG
 10681 CTTAGTCCCG CGTTGATATA AGCGCGCAT CGCGTCCCA CGGTTTCCGC GACAAGCGAG
 10741 GGTGTTGCGC CCCAGTCAAT TCCAAGACCC CGCCAGCCGA CTTCTCCAGT TTACGGGAGC
 10801 GAGCCCTTTT TTTTCTTGT TTTTGTGCGC CAGATGCATC CAGTGCTGCG ACAGATGCGC
 10861 CCCAGCAAC AGGCCCTTTC TCAGCAACAG CAGCAGCAAT CACAAAAGGC TGTCCCTGCA
 10921 ACTACTGCAA CTGCGCGCGT GAGCGGTGCG GGACAGCCCG CCTATGATCT GGACTTGGAA
 10981 GAGGGCGAAG GACTGGCAGC TCTAGGTGCG CTTCAACCCG AGCGGCATCC GCGAGTTCAA
 11041 CTGAAAAAAG ATTCTCGCGA GCGGTATGTG CCCCAACAGA ACCTATTTAG AGACAGAAGC
 11101 GCGGAGGAGC CGGAGGAGAT GCGAGCTTCC CGCTTAAACG CCGGTCTGTA GCTGCGTCAC
 11161 GGTGTTGACC GAAGACGAGT GTTGCGGGAC GAGGATTTCG AAGTTGATGA AATGACAGGG
 11221 ATCAGTCCCG CTAAGGACCA CGTGGCTGCA GCCAACCTTG TATCGGCTTA CGAGCAGACA
 11281 GTAAAGGAAG AGCGTAACCT CCAAAGTCT TTTAATAATC ATGTGCGAAC CCTGATTGCC
 11341 CGCGAAGAAG TTACCCCTGG GTCTGATGCA TCTGTGGGAC CTGATGGAGG CTATCAACCA
 11401 GAACCCCACT AGCAAAACCC TGACAGCTCA GCTGTTTCTG GTGGTTCAAC ATAGCAGGGA
 11461 CAACGAGGCA TACTGGGAGG CGTTGTTGAA CATCACCGAG CCTGATGGGA GATGGCTGTA
 11521 TGATCTGATC AACATCCTGC AAAGTATTAT AGTGCAGGAA CGTAGCCTGG TTTGGCTGAG
 11581 AAGTGGCAGC TATCAACTAC TCGGTCTTGA GCCTGGGCAA ATACTACGCT CGCAAGATCT
 11641 ACAAGACCCC CTACGTACCC ATAGACAAGG AGGTAAAGAT AGATGGGTTT TACATGCGCA
 11701 TGAAGTCTCA GGTGCTGACT TTGAGCGAGC ATCTGGGGGT GTATCGCAAT GACAGGATGC
 11761 ACCGTGCGGT GAGCGCCAGC AGGAGGCGCG AGCTGAGCGA CAGAGAACTT ATGCACAGCT
 11821 TGCAAGAGAG TCTAACGGGG GCTGGGACCG ATGGGGAGAA CTACTTTGAC ATGGGAGCGG
 11881 ACTTGCAATG GCAACCCAGT CGCAGGGCCA TGAGGGCTGC AGGGTGTGAG CTTCTTTACA
 11941 TGAAGAGAGT GATGAAGTC GAGGACGAGG AGGCGAGTA CTTGGAAGAC TGATGGCGCG
 12001 ACCCGTATT TTTGCTAGAT GGAACAGCAG CAGGCACCGG ACCCGCAAT GCGGGCGGGC
 12061 CTGCAGAGCC AGCCGTCCCG CATTAACCTC TCGGACGATT GGACCCAGGC CATGCAACGC
 12121 ATAATGGCGC TGACGACCCG CAACCCCGAA GCCTTTAGAC AGCAACCCCA GGCCAACCCG
 12181 CTTTCCGCTA TACTGGAGGC CGTAGTGCCC TCCCGCTCCA ACCCAACCCA CGAGAAGGTC
 12241 CTGGCTATCG TGAACGCTGC TGGTTGGAGA ACAAGGCCAT CCGTCCCGAT GAGGCGGGG
 12301 TGGTATACAA TGCTCTCTTG GAGCGCGTGG CCCGTTACAA CAGCAGCAAC GTGCAGACCA
 12361 ACCTGGACCG GATGGTGACC GATGTGCGCG AGGCCGTGTC TCAGCGCGAG CGGTTCCAGC
 12421 GCGAGGCCAA CTTGGGATCG CTGGTAGCGC TAAACGCTTT CCTCAGCACC CAGCCCGCTA
 12481 ACGTGCCCCG TGGTCAGCAA GACTATACAA ACTTTTGTAG TGCATTGAGA CTGATGGTAG
 12541 CTGAGGTGCC TCAGAGCGAG GTGTACCACT CCGGGCCAGA TTAATTCTTC CAGACCAGCA
 12601 GACAGGGCTT GCAGACAGTG AACCTGACTC AGGCTTTCAA GAACCTGAAG GGTCTCTGGG
 12661 GAGTGCAGCG CGCAGTAGGG GATCGCGCGA CCGTGTCTAG CTTGCTGACT CCCAACTCCC
 12721 GCCTGCTGCT GCTGCTGGTA TCCCCCTTTA CTGACAGCGG TAGCATCGAC CGCAACTCGT
 12781 ACTTGGGCTA CCGTCTTAA CCGTATCGCG AGGCCATAGG GCAGAGCCAG GTGGACGAGC
 12841 AGACCTATCA AGAAATCACC CAAGTGAGCC GCGCCCTGGG TCAGGAAGAC ACGGGCAGTT
 12901 TGGAGGCCAC CTGAACTTC TTGCTAACCA ACCGGTCGCA GAAGATCCCT CCTCAGTAGT
 12961 CGCTTACCGC TGAGGAGGAG CGGATCCTCA GATACGTGCA ACAGAGCGTT GGAAGTTTC
 13021 TGATGCAGGA GGGGGCGACA CCTACCGCGC CGCTGGACAT GACAGCTCGA AACATGGAGC
 13081 CCAGCATGTA TGCTAGTAAC AGGCCCTTCA TTAACAACT GCTGGACTAC CTGCACAGGG
 13141 GGGCCGCGAT GAACCTGAT TATTTCACCA ATGCTATCCT GAACCCACAC TGGCTGCCCC
 13201 CACCTGGTTT CTACACTGGC GAGTACGACA TGCCCGACCC CAATGACGGG TTCCTGTGGG
 13261 ACGATGTGGA CAGCAGCATA TTCTCCCCGC CTCCCGGTTA TACAGTTTGG AAGAAGGAAG
 13321 GGGGCGATAG AAGACACTCT TCCGTGTGCG TGTCCAGAAC GGCTGGTGCT GCCGCCACCG
 13381 TCCCGAAGC TGCAAGTCTT TCCCTAGCT TGCCCTTTTC ACTAAACAGC GTTCGAGCA
 13441 GTGAAGTGGG GAGAATAACC CGCCCGCGCT TGATGGGCGA GGATGAGTAC TTGAATGACT

13501 CTTTGTGTGAG GCCAGAGAGG GAAAAGAAGT TCCCCAACAA TGGAATAGAG AGCCTGGTGG
 13561 ATAAAAAGAG TAGGTGGAAG ACTTATGCTC AGGATCAAG AGACGAGCCT GGGATCATGG
 13621 GGATTACAAG TAGAGCGAGC CGTAGACGCC AGCGCCATGA CAGACAGAGG GGTCTTGTGT
 13681 GGGACCATGA GGATTGCGCC GATGATAGCA GCGTGCTGGA CTGGGGTGGG AGAGGAAGGG
 13741 GCAACCCGTT TGCTCATTTG CGCCCTCGCT TGGGTGGTAT GTTGTAAAAA AAAATAAAAA
 13801 AAAAACTCAG CAAGGCCATG GCGACGAGCG TACGTTCTGT CTCTTTTATT ATCTGTGTCT
 13861 AGTATAATGA GCGGAGTCGT GCTAGGCGGA GCGGTGGTGT ATCCGGAGGG TCCTCCTCCT
 13921 TCTTACGAGA GCGTGATGCA GCAACAGGCG GCGATGCTAC AGCCCCCACT GGAGGCTCCC
 13981 TTCGTACCCC CGCGGTACCT GGCGCTACG GAAGGGAGAA ACAGCATTCG TTAACGGAG
 14041 CTGTGCGCTC TGTACGATAC CACCAAGTTG TATCTGGTGG ACAACAAGTC GGCGGACATT
 14101 GGCTCCCTGA ACTATCAGAA CGACCACAGC AACTTCCTGA CCACGGTGGT GCAGAACAA
 14161 GACTTTACCC CCACGGAGGC TAGCACCCAG ACCATCAACT TTGACGAGCG GTCGCGATGG
 14221 GGCGGTGAGC TGAAGACCAT CATGCACACC AACATGCCCA ACGTGAACGA GTACATGTTT
 14281 AGCAACAAGT TCAAGGCGAG GGTGATGGTG TCCAGAAAAG CTCCTGAAGG TGTATAGTA
 14341 AATGACACCT ATGATCATAA AGAGGATATC TTAAAGTATG AGTGGTTTGA GTTCACTTTA
 14401 CCAGAAGGCA ACTTCTCAGC CACCATGACC ATTGACCTGA TGAACAATGC CATCATTTGAC
 14461 AACTACTTGG AAATTTGGCAG ACAAATGGA GTGCTGGAAA GTGACATTGG TGTAAAGTTT
 14521 GACACTAGAA ACTTTAGGCT CGGGTGGGAC CCCGAACTA AGTTGATTAT GCCAGGAGTC
 14581 TACACTTAGT AGGCATTCCA TCCTGACATT GTATTGCTGC CTGGTTGCGG GGTAGACTTT
 14641 ACTGAAAGCC GACTTAGCAA CTTGCTTGGC ATCAGGAAAA GACATCCATT CCAGGAGGGT
 14701 TTCAAATCA TGTATGAAGA TCTTGAAGGG GGTAAATATC CTGCCCTTTT GGATGTCACT
 14761 GCCTATGAGG AAAGCAAAAA GGATACCACT ACTGAAACAA CCACACTGGC TGTTCGAGAG
 14821 TACACTTAGT AAGATGATAA TATAACTAGA GGAGATACCT ATATAAGTGA AAAACAACAA
 14881 CGTGAAGCTG CAGCTGCTGA AGTTACAAAA GAGTTAAAGA TCCAACCTCT AGAAAAAGAC
 14941 AGCAAGAGTA GAAGCTACAA TGTCTTGGAA GACAAATCA ACACGGCCTA CCGCAGCTGG
 15001 TACCTGTCCT ACAATTACGG TAACCCCGAG AAAGGACTAA GGTCTTGGAC ACTGCTTACC
 15061 ACTTACAGTG TCACCTGTGG GCGACGAGCAG GTCTACTGGT CGCTCCCTGA CATGATGCAT
 15121 GACCCAGTCA CTTTCCGCTC CACAAGACAA GTCAACAACT ACCCAGTGGT GGGTGCGAGG
 15181 CTTATGCCCC TCTTCTCAA GAGTTTCTAC AATGAGCAAG CCGTGACTC TCAGCAGCTC
 15241 CGACAGGCCA CTTGCTCTAC GCACGTCTTC AACCGCTTCC CTGAGAAACA GATCCTCATC
 15301 GGGCCGCGG CGCCCACAAT TACCACCGTC AGTGAACCG TTCTGCTCT CACAGATCAC
 15361 CGGACCTGCG CGTTACGCG CAGTATCCGG GGAGTCCAGC GCGTGACCGT TACTGACGCC
 15421 AGACGCGCA CTTGCTCCTA CGTTTACAAG GCCCTTGGCA TTGTCGCGCC GTGCGTTCTT
 15481 TCAAGCCGCA CTTTCTAAAA AAAAAAATAA TGTCATTCT TATCTCGCCC AGTAATAATA
 15541 CGCGTTGGG TCTGCGCGCT CCAAGCAAGA TGTACGAGG CGCCCGCATA CGCTGTACCC
 15601 AGCATCTGTT GCGCGTTCCG GGTCAATTTC GCGCTCCATG GGGCGCACTA AAGGGTCGTA
 15661 CCCGCACTCG GACCACGGTC GATGATGTGA TCGACCAGGT GGTGCGCGAT GCTCGTAATT
 15721 ATACTCCTAC TGCGCCTACA TCTACTGTGG ATGCAGTTAT TGACAGTGTG GTGGCAGACG
 15781 CCCGCGCCTA TGCTGCGCGG AAGAGCCGAA GGAGGCGCAT TGCCAGGCGC CACAGGGCTA
 15841 CTACCCGCTG CCGAGCTGCA AAAGCTATTG TGCGGAGGGC CAAACGTGTG GGGCGAAGAG
 15901 CCATGCTTAG AGCGGCCAGA CGCGCGGCTT TAGGTGCCAG CAGCGGCAGG TCCCGCAGGC
 15961 GCGCGGCCAC GGCGGCAGCA GCGGCCATTG CCAACATGGC CCAACCCGCA AGAGGCAATG
 16021 TGTACTGGGT GCGTGATGCC ACTACCGGCC AGCGCGTGCC TGTGCGCACC CGCCCCCTC
 16081 GCACTTAGAA AGTACTGAGC AGTCTCCGAT GTTGTGTCCC AGCGGCAAGT ATGTCCAAGC
 16141 GCAAATACAA GGAAGAGATG CTCCAGGTCA TCGCGCCTGA AATCTACGGT CCACCGGTGA
 16201 AGGATGAAAA AAAGCCCCGC AAAATCAAGC GGGTCAAAAA GGACAAAAAA TAAGAGGACT
 16261 ATGACGATGA TGGGCTGGTG GAGTTTGTGC GCGAGTTCCG CCCAAGACGG CGCGTGCAGT
 16321 GGCGCGGGCG CCAAGTGCAT CAAGTGTCTA GACCCGGAAC ACGTGTGGTC TTTACACCCG
 16381 GCGAGCGTTC AAGCACTACT TTTAAACGGT CCTATGATGA GGTGTACGGG GATGATGATA
 16441 TTCTTGAGCA GGCGGCAGAC CGCCTTGATG AGTTTGCTTA TGGCAAGCGC ACTAGATCCA
 16501 GTCCCAAAGA GAGGCTGTG TCCATTCCCT TAAATCATGG AAATCCCACC CCCAGCCTCA
 16561 AACCACTATC CTTGACGCAA GTGCTGCGCG TGCTTCCGCG GAGAGGCGTA AAGCGGAGG
 16621 GTGAGGACCT ATATCTTACC ATGCAGCTAA TGGTGCCCAA GCGCCAGAGG CTAGAAGACG
 16681 TACTGAGGAA AATGAAAGTG GATCCCGATA TCCAGCCTGA GGTCAAAGTA AGACCTATCA
 16741 AGGAAGTGGC GCCAGGTTTG GGAGTACAAA CCTTCGACAT CAAGCCGAGT CCATGGAAGT
 16801 GCATACGAA CCGTCAAAACC CTACACTCAC TTGAGGTGCA AACGGAACCC TGGACGCCCG
 16861 CGCCCGTTGT TCGCCCCCAG CACCACTCGA AGATGCAAAC GGAACCCCTG ACGCCCGCGC
 16921 CCGTTGTTCTG CCCCAGCAC CACTCGAAGA TCACGACGAA AGTACGGCCC AGCAAGTCTT
 16981 CTAATGCCCA ATACTGCTCT GCACCCATCC ATCATTCCTA CTCCCGGGTA CAGAGGCACT
 17041 CGATACTATC GAAAACGAAA CAGTACCTGG TCCGATGGA TCCATCCTAG TCAGCAGCTT
 17101 CTCACCTAGT GACGTCGCGA ACATTGCCAG CAAACTGACT CCGAGCCAGC CCTGGTGCAG
 17161 GAAGAGTGAA CCGGATGTCG TCGGGAACCT TTGACACTGC CGCGCGCGCG CTACCATCCA
 17221 AGCATCACCA CTTAATGACT GTTGACGCTG CCTCCTTGCA GATATGGCCC TCACTTGCCG
 17281 CCTTCGCGTC CCCATTACTG GCTACCGAGG AAGAACTCG CGCCGTAGAA GGATGTTGGG
 17341 GCGAGGGATG CGCCGCCACA GACGAAGGCG CGCTATCAGC AGACGATTAG GGGGTGGCTT
 17401 TTGCGCAGCT CTTATACCCA TCATCGCCGC AGCGATCGGG GCGATACCAG GCATAGCTTC
 17461 CGTGGCGGTT CAGGCTCGC AGCGCCACTA ACATTGGAAA AAAACTTATA AATAACACTA
 17521 GAATGAGTCT TGAGCTCCTG GTCCTGTGAC TATGTTTTGT AGAGATGGAA GACATCAATT
 17581 TTTATCCTTT GGCTCCGCGA CACGGCACGA AGCCGTACAT GGGCACCTGG AGCGACATCG
 17641 GCACGAGCCA ACTGAACGGG GCGCCTTCA ATTGGAGCAG TATCTGGAGC GGGCTTAAAA
 17701 ATTTTGGCTC AACCATAAAA ACATACGGGA ACAAGCTTG GAACAGCAGT ACAGGACAGG
 17761 CGCTTAGAAA TAAACTTAAA GACCAGAACT TCCAACAAAA AGTAGTCGAT GGGATAGCTT
 17821 CCGGCATCAA TGGAGTGGTA GATTTGGCTA ACCAGGCTGT GCAGAAAAAG ATAAACAGCG
 17881 CTTTGAGACC GCGGCCAGCA ACCCTGGTG AAATGGAAGT GGAGGAAGAA CTTCTCCGCG
 17941 TGGAAAAGCG GCGCGACAAG CGTCCGCGAC CCGAGCTGGA GCAGACACTG GTGACGCGCG
 18001 CACAGGAGCC CCGTTCATAC GAGGAGGCAG TAAAGCTCGG AATGCCCACT ACCAGGCTCG
 18061 TAGCTCACAT GGCTACCGGG GTAATGAAAC CTCTCAGAC ACATCGACCC GCCACCTTGG

18121 ACTTGCCCTCC TCCCCCTGCT TCTGCGGCAC CTGTTCCCAA ACCTGTGCT ACCAGAAAGC
 18181 CCACCGCCGT ACAGCCCGTC GCCGTAGCCA GACCGCGTCC TGGGGGCACA CCGCGCCCGA
 18241 AAGCAAAGTGC GCAAAGTACT CTGAACAGCA TCGTGGGTCT GGGCGTGCAG AGTGTAAAGC
 18301 GCCGTGCTGT CTATTAAATTA AATATGGAGT AGCGCTTAAC TTGCTTGTCT GTGTGTATGT
 18361 ATCATCACCA CGCCGCGCGA GCAGAGGAGA AAGGAAGAGG TCGCGCGCCG AGGCTGAGTT
 18421 GCTTTCAAGA TGGCCACCCC ATCGATGATG CCCCAGTGGG CATACATGCA CATCGCCGGA
 18481 CAGGATGCTT CGGAGTACCT GAGTCCGGGT CTGGTGCAGT TCGCCCGTGC AACAGACACC
 18541 TACTTCAGTA TGGGGAACAA GTTTAGAAAC CCCACAGTGG CGCCCAACCA CGATGTGACC
 18601 ACCGACCGTA GCCAGCGACT GATGCTGCGC TTCGTGCCCC TTGACCGGGA GGACAATACA
 18661 TACTCTTACA AAGTGCAGTA CACCCTCGCC GTGGGCGACA ACAGAGTGCT TGACATGGCC
 18721 AGCACATTCT TTGACATTAG GGGGGTGCTT GATAGAGGTC CTAGCTTCAA GCCATATTCC
 18781 GGCACAGCTT ACAATTCACT GGCTCCTAAG GGCGCGCTTA ACACATCTCA GTGGATAGTT
 18841 ACAACGGGAG AAGACAATGC CACCACATAC ACATTTGGCA TTGCTTCCAC GAAGGGAGAC
 18901 AATATTACTA AGGAAGGTTT AGAAATTGGG AAAGACATTA CTGCAGACAA CAAGCCCAT
 18961 TATGCCGATA AAACATATCA GCCAGAGCCT CAAGTTGGAG AAGAATCATG GACTGATATT
 19021 GATGGAACAA ATGAAAAATT TGGAGGTAGA GCTCTTAAAC CAGCTACTAA AATGAAGCCA
 19081 TGCTACGGGT CTTTTCGAAG AACTACAAAT ACAAAGGGG GCCAAGCTAA AAACAGAAAA
 19141 GTAAACACCA CCGAAGGAAA TGTGAGACT GAGGAGCCAG ATATTGATAT GGAATTCTCC
 19201 GATGGTAGAG AAGCTGCTGA CGCTTTTCG CCTGAAATTG TGCTTTACAC GGAAAATGTC
 19261 AATTTGGAAG CTCACAGACG CCATGTGGTA TACAAGCCAG GAACCTCTGA TGGTAACTCT
 19321 CATGCAAAAT TGGGTCAACA AGCCATGCCT AACAGACCCA ATTACATTGG CTTCAAGGAT
 19381 AACCTTGTAG GTCCTATGTA CTACAACAGT ACTGGAATA TGGGAGTTT GGCAGGCCAA
 19441 GCATGCAAC TGAAATGCACT GGTGACTTG CAGGACAGAA AACTTGAACT GTCATATCAG
 19501 CTTTGTCTTG ATTCTCTGGG AGACAGAAGC AGATACTTCA GCATGTGGAA TCAGGCTGTG
 19561 GACAGCTATG ATCCCGATGT TCGTATTATT GAAAATCATG GCGTCGAGGA TGAATGCGCT
 19621 AATTACTGTT TTCCTCTGGA TGGCATAGGA CCAGGAAAAA AATATCAAGG CATTAAACCT
 19681 AGAGACATG ATGCGGAAAA AGATACTAAA GTTTCTACAG CTAATGAAAT AGCCATAGGC
 19741 AACAACTCTG CTATGGAAT TAATATCCAA GCTAATCTTT GGAGAAGTTT TCTGTACTCC
 19801 AATGTGGCTT TGTACCTTCC AGATGTTTAC AAGTACACGC CAACTAACAT TACTCTGCCC
 19861 GCTAACACCA ACACCTATGA GTACATGAAC GGGCGAGTGG TTTCCCATC TCTGGTCGAT
 19921 TATACATCAT CATTTGGCGC CAGGTGGTCT CTTGACCCAA TGGACAATGT GAATCCATT
 19981 GACCACCACC GCGATGCTGG CCTACGCTAC CGGTCCATGC TTCTGGGCAA TGGCCGTTAT
 20041 GTGCCTTTCC ACATACAAGT GCCTCAAAAA TTCTTTGCTG TCAAGAACCT ACTTCTTCTA
 20101 CCTGGCTCCT ACACCTATGA GTGGAACCTC AGAAAGGATG TGAACATGGT CCTGCAAGT
 20161 CCTCTTGAA ATGACCTCAG AACAGATGGT GCTACCATAA GTTTCACAG CATCAACCTC
 20221 TATGCCACCT TCTTCCCAT GGCTCACAAC ACCGCTTCAA CTCTTGAAGC CATGCTGGGC
 20281 AACGATACCA ATGATCAGTC ATTCAAAGAC TACCTCTCTG CAGCTAACAT GCTTTACCCC
 20341 ATCCCTGACA CTGCAACCAA CATTCCAATC TCCATCCCAT CTCGCAACTG GCCACCCCTC
 20401 AGGGGCTGGT CCTCACTAG ACTCAAAACC AAGGAGACTC CATCTCTTGG ATCAGGGTTC
 20461 GATCCCTACT TCCGTTATTC TGGATCTATT CCCTACCTGG ATGGCACCTT TTACCTAAC
 20521 CACACTTTCA AGAAGGTCTC CATCATGTTT GACTCCTCAG TCAGCTGGCC TGGCAATGAC
 20581 AGGCTGTTGT CTCCAAATGA GTTTGAAATC AAGCGCACTG TGGATGGGGA AGGATACAAT
 20641 GTGGCCCAAT GCAACATGAC CAAAGACTGG TTCCTGGTTC AGATGCTTGC CAACTACAA
 20701 ATTGGCTACC AGGGCTTTTA CATCCCTGAG GGATACAAGG ATCGCATGTA CTCCTTTTTC
 20761 AGAACTTCC AGCCTATGAG CAGGCAGGTG GTTGATGAGG TTAATTACAC TGAATACAA
 20821 GCCGTCACTT TACCATATCA ACACAACAAC TCTGGCTTG TAGGATACCT TGCGCTACT
 20881 ATGAGACAAG GGGAACTTA CCCAGCCAAT TATCCATACC CGCTCATCG AACTACTGCC
 20941 GTTAAAGTGT TTACCCAAAA AAAGTTCTCTG TCGGACAGGA CCATGTGGCG CATACCGTTC
 21001 TCCAGCAACT TCATGTCCAT GGGAGCCCTT ACGGACCTGG GACAGAACCT GCTCTATGCC
 21061 AACTCGGCCC ATGCGCTGGA CATGACTTTT GAGGTGGATC CCATGGATGA GCCCACCTG
 21121 CTTTATCTTC TTTTCGAAGT CTTGACGTG GTCAGAGTGC ACCAGCCACA CCGCGGCGTC
 21181 ATCGAGGCA GCTACCTGCG CACACCGTTC TCGGCCGGCA ACGCCACCAC ATAAGAAGCC
 21241 TCTTGCTTCT TGCAAGCAGC AGCTGCAGCC ATGTCATGCG GGTCCGGAAG CGGCTCCAGC
 21301 GAGCAAGAGC TCAAGGCCAT CGTCCGAGAC CTGGGTTGCG GACCCATTTT CCTGGGAACC
 21361 TTTGACAAGC GTTTCCCGGG GTTCATGGCC CCCGACAAGC TCGCCTGCGC CATAGTCAAC
 21421 ACTGCGGAG GCGAGACGGG GGGAGAGCAC TGCTGGCTT TTGGTTGGAA CCCGCGCTCC
 21481 AACACCTGCT ACCTTTTGA TCCTTTTGGG TTCTCGATG AGCGACTCAA ACAGATTAC
 21541 CAGTTTGAGT ACGAGGGGCT CCTGCGCGCG AGTGCCCTTG CTACCAAAAG ACCGCTGCAT
 21601 CACCTGGAAG AAGTCCACCC AGAGCGTGCA GGGCCCACGC TCAGCCGCT GTGGACTTTT
 21661 TTGCTGATG TTTCTTCATG CCTTTGTGCA CTGGCCCGAC CGCCCATGA ACGGAAACCC
 21721 CACCATGAAG TTGCTGACTG GGTGCCCCAA CAGCATGCTC CAATCTCCCC AAGTGCAGCC
 21781 CACCTGCGC CGCAACCAGG AGGCGCTATA TCGCTTCTTA AACACCACT CATCTTACTT
 21841 TCGTTCTCAC CGCGCACGCA TCGAAAGGGC CACCGCGTTT GACCGTATGG ATATGCAATA
 21901 AGTCATGTAA AACCCTGTTT AATAAAAAGC ACTTTATTTT TACATGCACT AAGGCTCTCG
 21961 TTTTACTACT ATTCTGTTT ATTATTCACT CAGAAATCAA ATGGGTCTG GCGGGAGTCA
 22021 AAGTGACCCG CGGCGAGGGA TACGTTGCGG AACTGTAAAC TGTTCTGCCA CTTGAACTCG
 22081 GGGATCACCA ACTTGGGAAC TGGAACTCTG GGAAGGTGT CTTGCCACAA CTTTCTGGTC
 22141 AGTGCAGGCG GCGCAAGTAG GTCAGGAGCA GAGATCTTGA AATCAGACTT GGGACCGGCA
 22201 TTCTGGACAC GGGAGTTGCG GTACACTGGG TTGCAACACT GGAACACCAT CAAGGCTGGG
 22261 TGTCTCACGC TTGCGAGCAC GGTGCGGTCA CTGATGAGTAC TCACATCCAA GTCTTCAGCA
 22321 TTGGCCATCC CAAAGGGGGT CATCTTACAG GTCTGCTGTC CCATCACGGG AGCGCAGCCT
 22381 GGCTTGTGGT TGAATCGCA ATGAATGGGG ATCAGCATCA TCCTGGCTTG GTGCGGGGTT
 22441 ATCCCTGGGT ACACGGCCTT CATGAAGGCT TCGTACTGCT TGAAAGCTTC CTGAGCCTTA
 22501 CTTCCCTCGG TATAGAACAT CCCACAGGAC TTGCTGGAAA ATTGATTAGT AGCACAGTTG
 22561 CGATCATTTA CACAGCAGCG GGCATCGTTG TTGGCCAAC GGACCACATT TCTGCCCCAG
 22621 GGTTCCTGGG TGATCTTGGC TCTGTCTGGG TTCTCTTCA TAGCGCGCTG TCCGTTCTCG
 22681 CTCGCCACAT CCATCTCGAT AATGTGGTCC TTCTGAATCA TGATAGTGCC ATGCAGGCAT

22741 TTCACCTTGC CTTTCATAATC GGTGCATCCA TGAGCCACACA GAGCGCACCC GGTGCACCTCC
 22801 CAACTATTGT GGGCGATCTC AGAATAAGAA TGTACCAATC CCTGCATGAA TCTTCCCCTC
 22861 ATCGCTGTCA GGGTCTTCAT GCTACTAAAT GTCAGCGGGA TGCCACGGTG CTCCTCGTTC
 22921 ACATACTGGT GGCAGATACG CTTGTACTGC TCGTGTCTGT CTGGCATCAG CTTGAAAGAG
 22981 GTTCTCAGGT CATTATCCAG CCTGTACCTC TCCATTAGCA CAGCCATCAC TTCCATGCCC
 23041 TTCTCCAGG CAGATACCAG GGGCAAGCTC AAAGGATTCC TAACAGCAAT AGAAGTAGCT
 23101 CCTTTAGCTA TAGGGTCATT CTTGTGATC TTCTCAACAC TTCTCTTGCC ATCCTTCTCA
 23161 ATGATGCGCA CCGGGGGTA GCTGAAGCCC ACGGCCACCA ACTGAGCCTG TTCTCTTTCT
 23221 TCTTCGCTGT CGTGGCCGAT GTCTTGACAGA GGGACATGCT TGGTCTTTCT GGGCTTCTTC
 23281 TTGGGAGGGA TCGGGGAGG ACTGTTGCTC CGTTCGGAG ACAGGGATGA CCGCGAAGTT
 23341 TCGCTTACCA GTACCACTG GCTCTCGATA GAAGAATCGG ACCCCACGCG ACGGTAGGTG
 23401 TTCCTCTTCG GGGGACAGG TGGAGGCGAC TGAGATGGGC TGGGGTCTGG CCTTGAAGC
 23461 GGATGGCTGA CAGAGCCCAT TCCGCGTTG GGGGTGTGCT CCCGTGGCG GTCCGTTGAA
 23521 TGATTTCCTC CGCGGGTGGC CATTGTGTTT TCCTAGGCAG AGAAACAACA GACATGGAAA
 23581 CTCAGCCATC ACTGCCAACA TCGCTGCAAG CGCCATCACA CCTCGGCCCC CAAGAAGCGA
 23641 CGAAGGAGGA GAGCTTAACC ACCCCACCA CCCAGTCCAG CTACCACCAA CCTCTACCCC
 23701 TCGATGATGA AGGAGGAGGA GGTGACGACA GCCCAGGAGA TGCAGGCGCA GGATAATGTG
 23761 AAAGCGGAAG AGATTGAGGC AGATGTCGAG CAAGACCCGG GCTATGTGAC ACCGGCGGAG
 23821 CACGAGGAGG AGCTGAAACG TTTCGCTTCA CATACCCCCC AAGCGGCAGC CAAACGCGAC CTGCGAGGCC
 23881 GCAGATGGCG ATCACCAGGA GGCTGGCATC GGGGATCAAG TTGCCGACTA CCTCACCGGG
 23941 CTTGGGGGGG AAGACGTGCT CCTCAAACAT CTAGCAAGGC AGTCGAACAT AGTTAAAGAC
 24001 GCACTAGTCG ACTCAACCGA AGTGCCCATC AGTGTGGAAG AGCTTAGCCG CGCCTACGAG
 24061 CTGAACCTCT TTTGCGCTCA CATACCCCCC AAGCGGCAGC CAAACGCGAC CTGCGAGGCC
 24121 AACCTCTGAC TGAACCTCTA TCCAGCTTTT ACTGTCCCCG AAGTGTGGA CACATACCAC
 24181 ATCTTTTTTA AGAAGCCAAA AATTCCAGTC TCCTGCCGCG CTAATCGCAC CCGCGCCGAT
 24241 CCCCTACTCA ATCTGGGACC TGGTTCACGC TTACCTGATA TAGCTTCCTT GGAAGAGGTT
 24301 CCAAAAATCT TTTGAGGTTCT GGGAAAGTAT GAGACTCGGG CCGCAAATGC TCTGCAACAG
 24361 GGAGAGAATG GCATGGATGA ACATCACAGC GCTTTAGTGG AACTGGAGGG TGACAATGCC
 24421 CGGCTTGCAG TGCTCAAGCG CAGTATCGTG GTCAACCATT TTGCCTACCC CGCTGTTAAC
 24481 CTGCCCCCCA AAGTTATGAG CGCTGTTATG GACCATCTGC TCATCAAAACG AGCAAGACCC
 24541 ATTTCCAGAA AGCAGAACAT GCAGGATCCA GACGCTCGG ACGAGGGCAA GCCGGTAGTC
 24601 AGTGACGAGC AGCTATCTCG CTGGCTGGGT ACCAACTCCC CCCGAGATTG GGAAGAGAGG
 24661 CGCAAGCTTA TGATGGCTGT AGTGCTAGTA ACTGTGGAGC TGGAGTGTCT GCGCCGCTTT
 24721 TTCACCGACC CTGAGACCCT GCGCAAGCTA GAGGAGAACC TGCCTACAC CTTTAGACAT
 24781 GGGTTCGTCG GGCAGGCATG CAAGATCTCC AACGTGGAGC TTACCAACCT GGTTCCTTAC
 24841 ATGGCGATT TGCATGAGAA CCGCTAGGG CAGAGCGTCC TGCACACCAC CTTTAAAGGG
 24901 GAGGCCCGCC GTGACTACAT CCGAGACTGT GTCTACCTTT ACCTCTGCCA TACCTGGCAG
 24961 ACTGGCATGG GTGTATGGCA ACAGTGTGTT GAAGAGCAGA ACCTTAAAGA GCTGGACAAG
 25021 CTCTTGCAAA GATCCCTCAA AGCCCTGTGG ACAGGTTTTG ATGAGCGCAC CGTCGCTCG
 25081 AACCTGGCAG ACATCATCTT CCCCAGCGGT CTCAGGGTTA CTCTGCGAAA CCGCCTGCCA
 25141 GACTTTATGA GCCAGAGCAT GCTTAACAAC TTTGCTCTT TCATCCTGGA ACGCTCCGGT
 25201 ATCCTGCCTG CCACCTGCTG TGGCTGCCCC TCCGACTTTG TGCCTCTCAC CTACCGCGAG
 25261 TGCCACCCGC CGCTATGGAG CCACTGCTAC CTGTTCCGCC TGGCCAACCTA CCTCTCCTAC
 25321 CATCGGCTAG TTATAGAGGA TGTGAGCGGA GACGGTCTGC TGGAAATGCCA CTGCGCTGCTG
 25381 AATCTTTGCA CACCCACCG CTCCCTTGCC TGCAACCCCC AGTTGCTGAG CGAGACCCAG
 25441 ATCATCGGCA CCTTCGAGTT GCAGGGTCCA GCAGTGAAGG CGAGGGGTCT TCTCCGGGGC
 25501 AGAGTCTAAA ACTGACACCG GGGCTGTGGA CCTCCGCTA CCTGCGCAAG TTTTATCCCG
 25561 AGGATTACCA CCCCTATGAG ATCAGGTTCT ATGAGGACCA GTCACATCCT CCCAAAGTCG
 25621 AGCTCTCAGC CTGCGTCATC ACCCAGGGAG CAATTCTGGC CCAATTGCAA GCCATCCAAA
 25681 AATCCGCGCA AGAATTTCTA CTGAAAAAGG GAAGCGGGT CTACCTTGAG CCCAGACCG
 25741 GTGAGGAGCT CAACACAAGG TTCCCCCAGG ATGTCCCATC GCCGAGGAAG CAAGAAGCTG
 25801 AAGTTCGAGC TGACGCCCCC AGAGGATATG GAGGAAGACT GGGACAGCAA GGCAGAGGAG
 25861 GAGATGGAAG ATTGGGACAG CCAGGCAGAG GAGGAAGACA GTTTGGAGGA GGAAGACGAG
 25921 GAGGCAGAGG AGGTGGAAGA AGCAACCGCC GCCAAACAGT TGTATCGGC GCGGAGACA
 25981 AGCAAGTCCC CAGACAGCAG CACGGCTACC ATCTCCGCTC CGGGTCCGGG GGTCCAGCGG
 26041 CGGCCCAACA GTAGATTGGA GCAGACCGGG CGATTCCCAA ACCCGACCAA CGGTTCCAAG
 26101 ACCGTTAAGA AGGAGCGACA GGGATACAAG TCCTGGCGTG GACACAAAAA CGCTATCATC
 26161 TCCTGCTTGC ATGAATGCGG GGGCAACATA TCCTTACCCC GGCATACCT GCTCTCCAC
 26221 CACGGTGTGA ACTTCCCCCG CAATATCTTG CATTACTACC GTCAACTCCA CAGCCCTTAC
 26281 TGCACTCAGC AAGTCCCGGC AACCCGACA GAAAAAGACA GCAGCGACAA CGGGTGACCA
 26341 GAAACCAGCA GTTAGAAAAT CTACTACAAG TGCAGCAGGA GGAGGACTGA GGATCACAGG
 26401 CGAACGAGCC AGCGCAGACC AGAGAGCTGA GGAACCGGAT CTTTCCAACC CTCTATGCCA
 26461 TCTTCCAGCA GAGTCGGGGG CAAGAGCAGG AACTGAAAGT AAAAAACCGA TCTCTGCGCT
 26521 CGCTCACCAG AAGTTGTTTG TATCACAAGA GCGAAGACCA ACTTCAGCGC ACTCTCGAGG
 26581 ACGCCGAGGC TCTCTTCAAC AAGTACTGCG CGCTGACTCT TAAAGAGTAG CCCTTGCCCG
 26641 CGCTTATTTC AAAACGGCGG GAATCACGTC ACCATTGGCA CCTGTCTTTT GCCCTAGTCA
 26701 TGAGTAAAGA GATTCCACG CCTTACATGT GGAGCTATCA GCCCCAAATG GGGTTGGCAG
 26761 CAGGCCCTCC CCAGGACTAC TCCACCCGCA TGAATTGGCT TAGCGCGGGG CCTCAATGA
 26821 TATCAGGGT TAATGATATA CGAGCTTATC GAAACAGTT ACTCTAGAA CAGTCAGCTC
 26881 TCACCAACCAC ACCCGGCCAA CACCTTAATC CCCGAAATTG GCCCGCGGCC CTGGTGTACC
 26941 AGGAAACTCC CGCTCCACC ACCGTACTAC TTCCTCGAGA CGCCAGGCC GAAGTTCAGA
 27001 TGACTAACGC AGGTGTACAG GTTCCGCGCG GTTCCGCCCT ATGTCGTAC CGACCTCAAC
 27061 AGAGTATAAA ACGCCTGGTG ATCAGAGGCC GAGGTATCCA GCTCAACGAC GAGTCGGTTA
 27121 GCTCTTCGCT TGGTCTGCGA CCAGACGAG TCTTCCAGAT CGCCGCTGT GGGAGATCTT
 27181 CCTTCAACC CCGTCAGGCT GTACTGACTT TGGAGAGTTC GTCTCTCGAG CCACGCTCGG
 27241 GCGGCATCGG AACTCTCCAG TTCGTGGAGG AGTTTACTCC CTCTGTCTAC TTCAACCCCT
 27301 TCTCCGGCTC TCCTGGCCAG TACCCAGACG AGTTCATACC GAACTTCGAC GCAATCAGCG

27361 AGTCAGTGGG TGGCTATGAT TGATGTCTAA TGGTGGCGCG GCTGAGCTAG CTCGACTGCG
 27421 ACACCTAGAC CACTGCCGCC GCTTTCGCTG CTTGCGCCGG GAACTCACCG AGTTCACTCTA
 27481 CTTGGAAGTC CCCGAGGAGC ACCCTCAGGG TCCGCCCCAC GGAGTGCGGA TTACCATCGA
 27541 AGGGGGGAATA GACTCTCGCC TGCATCGAAT CTTCTCCAG CGACCCGTGC TGATTGAGCG
 27601 CGACCAGGGA AATACAACCA TCTCCATTTA CTGCATCTGT AACCACCCCG GATTGTCATGA
 27661 AAGCCTTTGC TGTCTTGTGT GTGTGAGTT TAATAAAAC TGAGTTAAGA CCTCCTACG
 27721 GACTACCGCT TCTTCAATCA GGACTTTACA ACACCAACCA GATCTTCCAG AAGACCCAGA
 27781 CCCTTCCTCC TCTGATCCAG GACTCTAACT CTACCTTACC AGCACCATCC ACTACTAACC
 27841 TTCCCGAAAC TAACAAGCTT GGATCTCATC TGCAACACCG CCTTTCACGA AGCCTTCTTT
 27901 CTGCCAATAC TACCACTCCC AAAACCGGAG GTGAGCTCCG CGGTCTCCCT ACTGACGACC
 27961 CTTGGTGGT AGCGGTTTT GTAACTTAG GAGTAGTTGC GGGTGGGCTT GTGCTAATCC
 28021 TTTGCTACCT ATACATACCT TGCTGTGCAT ATTTAGTCAT ATTGCGCTGT TGGTTTAAAA
 28081 AATGGGGGCC ATATTAGTCG TGCTTGCTTT ACTTTCGCTT TGGGTCTGG GCTCTGCTAA
 28141 TCTCAATCCT CTTGATCAGC ATCCATGTCT AGACTTCGAC CCAGAAAAC GCACACTTAC
 28201 TTTTGACCCC GACACAAGCC GTCTCTGTGG AGTTCCTATT AAGTGCAGG GGGACTGCAG
 28261 GTCCGTTGAA ATTACACATA ATAACAAAC ATGGAACAAT ACCTTATCCA CCACATGGGA
 28321 GCCAGGAGTT CCCGAGTGGT ATACTGTCTC TGCCGAGGT CCTGACGGT CCATCCGCAT
 28381 TAGTAACAAC ACTTTCATTT TTTCTGAAAT GTGCGATCTG GCCATGTTCA TGAGCAGACA
 28441 GTATGACCTA TGGCTCCCCA TGCCCTCCCA CATTTGGGCA TTTTCCATTG TTTATTGCTT
 28501 GGTAACATGC ATCATCACTG CTATCATTTG TGTGTGCATA CACTTGCTTA TAGTTATTCTG
 28561 CCTAGACAA AGCAATGAGG AAAAAGAGAA AATGCCTTAA CCTTTTCCCT CATACCTTTT
 28621 CTTTACAGCA TGGCTTCTGT TACAGCTCTA ATTATTGCCA GCATTGTCAC TGTCGTTTAC
 28681 GGGCAACCAA TTGTCCATAT TACCTTAGGA CATATCACA CTCTGTAGG GCCCCCAATT
 28741 ACTTCAGAGG TTATTGGAC CAACTTGGGA AGTGTGATT ATTTTGATAT AATTTGCAAC
 28801 AAAACTAAAC CAATATTTGT AATCTGTAAC AGACAAATC TCACGTTAAT TAATGTTAGC
 28861 AAAATTATA ACGGTTACTA TTATGGTTAT GACAGATCCA GTAGTCAATA TAAAAATTAC
 28921 TTAGTCCGCA TAACTCAGCC CAAATTAACA GTGCCAATA TGACAATAAT TTAATGGCT
 28981 AATAAAGCAT TTGAAAATTT TACATCACCA ACAACGCCCA ATGAAAAAAA CATTCCAAAT
 29041 TCAATGATTG CAATTATTGC GGCCTGGGCA TTGGGAATGG CACTAATAAT AATATGCATG
 29101 TTCCTATATG CTTGTTGCTA TAAAAAGTT CAACATAAAC AGGATCCACT ACTAAATTTT
 29161 AACATTAACT TTTTATACA GATGATTTCC ACTACAATTT TTATCATTAT TAGCCTTGCA
 29221 CTTGTAACCT ATGGCCGTTT ACACCTAACT GTACCTGTTG GCTCAACATG TACACTACAA
 29281 GGACCCCAAG AAGGCTATGT CACTTGGTGG AGAATATATG ATAATGGAGG GTTCGCTAGA
 29341 CCATGTGATC AGCCTGGTAT AAAATTTTCA TGCAACGGAA GAGACTTGAC CATTATTAAC
 29401 ATAACATCAA ATGAGCAAGG CTTCTATTAT GGAACCACT ATAAAGATAG TTTAGATTAC
 29461 AACATTATTG TATTGTCAGC CACCACGCTC GCTCCCGCA AATCCACTTT CTCTAGCAGC
 29521 AGTGCCAAAG CAAGCACAAT TCCTAAAACA GCTTCTGCTA TTTTAAAGCT TCCAAAAATC
 29581 GCTTTAAGTA ATTCACAGC CGCTCCCAAT ACAATTCCTA AATCAACAAT TGGCATCATT
 29641 ACTGCCGTGG TAGTGGGATT AATGATTATA TTTTGTGCA TAATGTACTA CGCCTGCTGC
 29701 GTTGTGCGAA ATGAACAAA AGGTGATGCA TTACTAAATT TTGATGTTTA ATTTTATATA
 29761 GAATTATTAT ATTGTTTCAA TCAAATACCA CTAACACTAT CAATGTGCAG ACTACTTTAA
 29821 ATCATGACAT GGAAAAACCA CACTACCTCC TATGCATACC CAATGTTTAT AGAAATTTCTA
 29881 AAAGACGTCC CATCTATTCT TCTATGATTA GTCGTCCCCA TATGGCTCTG AATGAAATCT
 29941 AAGATCTTTT TTTTCTCTC TACAGTATGG TGAACATCAA TCATGATCCC TAGAAATTTT
 30001 TTCTTACCA TACTCATCTG TGCTTTCAAT GTCTGTGCTA CTTTACAGC AGTAGCCACT
 30061 GCAAGCCAG ACTGTATAGG ACCATTGCTC TCCTATGCAC TTTTGCCTT CGTTACTTGC
 30121 ATCTGCGTGT GTAGCATAGT CTGCTGGTT ATTAATTTT TCCAACGGT AGACTGGATC
 30181 TTTGTGCGAA TTGCTTACCT ACCTCACCAT CCCGAATACC GCAATCAAAA TGTGCGGCA
 30241 CTTCTTAGGC TTATTTAAAA CCATGCAGGC TATGCTACCA GTCATTTTAA TTCTGCTACT
 30301 ACCCTGCATT GCCCTAGCTT CCACGCCAC TCGCGCTACA CCTGAACAAC TTAGAAAATG
 30361 CAAATTTCAA CAACCATGGT CATTTCTTGA TTGCTACCAT GAAAAATCTG ATTTTCCCAC
 30421 ATAGTGGATA GTGATTGTTG GAATAATTAA CATACTTTCA TGTACCTTTT TCTCAATCAC
 30481 AATATACCCC ACATTTAATT TTGGGTGGAA TTCTCCCAAT GCACTGGGTT ACCCACAAGA
 30541 ACCAAATGAA CATATCCAC TACAACACGT ACAACAACCA CTAGCACTGG TAGAGTATGA
 30601 AAATGAGCCA CAACCTTAC TGCCCCCTGC CATTAGTTAC TTCAACCTAA CCGGCGGAGA
 30661 AACTGACCC ATTCGCCACA TCATCCACCG CTGCCAAGGA GCTGCTGGAC ATGGACGGAC
 30721 GTGCCTCAGA ACAGCGACTC ATCCAACTAC GCATTCTGTA GCAGCAGGAA CGAGCAGTAA
 30781 AAGAGCTAAG GGATGCCATT GGGATTACAC AGTGCAAAAA AGGCATATTC TGCTTAGTAA
 30841 AACATCCAA AATCTCTTAC GAGATCACCG CTAAGTACCA TCGTCTCTCA TACGAGCTCG
 30901 GTCCGCGAGC ACAAATAATC ACCTGCATGG TGGGAATCAA CCCCATAGTT ATCACCAGC
 30961 AGTCTGGAGA TACTAAGGGT TGCACTCCACT GTTCTGTGA TTCCACCGAG TGCATCTACA
 31021 CACTGTGAA GACCTCTGCG GGCCTTCGAG ACCTCCTACC CATGAACATA TCATTGCCCC
 31081 CTCCCTTACC CAATCCAAAT ATTAATAAAG ACACTTACTT GAAATCAGCA ATACAGTCTT
 31141 TGTCAAAACT TTCTACCAGC AGCACTTAC CTTCTTCCCA ACTCTGGTAC TCTAAACGTC
 31201 GGAGGGTGGC ATACTTTCTC CACACTTTGA AAGGGATGTC AAATTTTATT TCCTCTTCTT
 31261 TGCCACAAAT CTTCAATTTCT TTATCCCCAG ATGGCCAAGC GAGCTCGGCT AAGCACTTCC
 31321 TTCAACCCGG TGTACCCCTTA TGAAGATGAA AGCAACTTAC AACACCCATT TATAAATCCT
 31381 GTTTCATTTT CCGGAAATGG GTTCACACAA AGTCCAAACG GGGTTTTAAG TCTTAAATGT
 31441 GTTAATCCAC TTACCACTGC AAGCGGCTCC CTCCAACCTA AAGTGGGAAG TGGTCTTACA
 31501 GTAGACACTA CTGATGGATC CTTAGAAGAA AACATCAAAG TTAACACCCC CCTAACAAAG
 31561 TCAAACCAT CTATAAATTT ACCAATAGGA AACGGTTTGC AAATAGAACA AAACAAACTT
 31621 TGCAGTAAGC TCGGAAATGG TCTTACATTT GACTCTTCCA ATTCTATTGC ACTCAAAAT
 31681 AACACTTTAT GGACAGGTCC AAAACCGAGAA GCCAACTGCA TAATTGAATA CGGGAAGAGAA
 31741 AACCCAGATA GCAAACTAAC TTTAATCCTT GTAAAAATG GAGGAATTGT TAATGGATAT
 31801 GTAACGCTAA TGGGAGCCTC AGACTATGTT AACACCTTAT TAAAAACAA AAATGTCTCC
 31861 TTAATGTAG AATTATACCT TGATGCCACT GGTATATAT TACCAGACTT ATCTTCTCTT
 31921 AAAACAGATC TACAACATAA ATACAAGCAA ACCACTCACT TTAGTGCAAG AGGTTTTATG

31981 CCAAGTACTA CAGCGTATCC ATTGTGCTT CCTAATGCGG GAACAGATAA TGAATAATTAT
 32041 ATTTTGGTGC AATGCTACTA CAAAGCAAGC GATGGCGCCC TTTTCCGTT GGAAGTTACT
 32101 GTTACGCTTA ATAAACGCCT GCCAGATAGT CGCACATCCT ATGTTATGAC TTTTATGAG
 32161 TCCTTGAATG CTGCTCTAGC TCCAGAACT ACTCAGGCAA CCCTGATAAC CTCCCATTT
 32221 ACCTTTTCCCT ATATTACAGA GGATGACTGA CAACAAAAAT AAAGTTCAAC ATTTTATTAT
 32281 GAAATTCCCTT TTACAGTATT CGAGTAGTTA TTTGCTCTCC CCCTTCCCAT TTAACAGAAT
 32341 ACACCAATCT CTCCCCACGC ACAGCTTTAA ACATTGGAT ACCATTAGAG ATAGACATAG
 32401 TTTTAGATTG CACATTCCAA ACAGTTTCAG AGCGAGCCAA TCTGGGGTCA GTGATACATA
 32461 AAAATGCATC GGGATAGTCT TTTAAAGCGC TTTCACAGTC CAACTGTTGC GGATGCGACT
 32521 CCGGAGTCTG AATCACGGTC ATCTGGAAGA AGAACGATGG GAATCATAAT CCGAAAACGG
 32581 AATCGGGCGA TTGTGTCTCA TCAACCCAC AAGCAGCCGC TGTCTGCGTC GCTCCGTGCG
 32641 ACTGCTGTTT ATAGGATCGG GATCCACAGT GTCCTGAAGC ATGATTTTAA TAGCCCTTAA
 32701 CATTAACTTT CTGGTGCAGT GCGCGCAGCA ACGCATCTTT ATTTCACTTA GATTACTACA
 32761 GTAGGTACAG CACATTATCA CAATATTGTT TAATAACCA TAATTAAAAG CGCTCCAGCC
 32821 AAAACTCATA TCAGATATAA TCGCCCTGTC ATGACCATCA TACCAAATTT TAATATAAAT
 32881 TAAATGTCGT TCCCTCAAAA ACACACTACC CACATACATA ATCTCTTTTG GCATGTGCAT
 32941 ATTAACAATC TGTCTGTACC ATGGACAACG TTGGTTAATC ATGCAACCCA ATATAACCTT
 33001 CCGGAACCA AATGCCAACA CCGCTCCCC AGCCATGCAT TGAAGTGAAC CCTGCTGATT
 33061 ACAATGACAA TGAAGAACCC AATTCTCTCG ACCATGAATC ACTTGAGACT GAAAAATATC
 33121 TATAGTAGCA CAACAAAGAC ATAAATGCAT GCATCTTTTC ATAATTTTAA ACTCTCTGCG
 33181 ATTTAAAAAC ATATCCCAAG GAATGGGAAA CTCTTGAGCA ACAGTAAAGC TGGCAGAACCA
 33241 AGGAAGACCA CGAACACAAC TTACACTATG CATAGTCATA GTATCAAT CTGGCAACAG
 33301 CGGGTGGTCT TCAGTCATAG AAGCTCGGGT TTCATTTTCC TCACATCGTG GTAACGGGCG
 33361 TCTGGGTGTA GGGTGTATGTC TGGCGCATGA TGTGGAGCGT GCGCGCAACC TTGTCATAAT
 33421 GGAGTGTGCT CCTGACATTC TCGTATTTTG TATAGCAAAA CGCGGCCCTG GCACAACACA
 33481 CTCTTCTTCG TCTTCTATCC TGCCGCTTAG TGTGTTCCGT CTGATAATTC AAGTACAGCC
 33541 ACACCTTTAA GTTGGTCAAA AGAATGCTGG CTTCAGTTGT AATTAAAACT CCATCATATT
 33601 TAATTGTTCT AAGGAAATCA TCCACGGTAG CATATGCAAA TCCCAACCAA GCAATGCAAC
 33661 TGGATTGCGT TTCAAGCAGC AGAGGAGAGG GAAGAGACGG AAGAATCATG TTAATTTTAA
 33721 TTCCAAACGA TCTCGCAGTA CTTCAAATTG TAGATCGCGC AGATGGCATC TATCGCCCCC
 33781 ACTGTGTTGG TGAAGAAAGCA CAGCTAAATC AAAAGAAATG CGATTTTCAA GGTGCTCAAC
 33841 GGTGGCTTCC AACAAAGCCT CCACGCGCAC ATCCAAAAAC AAAAGAATAC CAAAAGAAGG
 33901 AGCATTTTCT AACTCTCTCA ACATCATATT ACATTCTGTC ACCATTCCCA GATAATTTTC
 33961 AGCTTTCCAG CCTTCAATTA TTCGTGTGTC TTCTTGTGGT AAATCCAAAC CACACATTAC
 34021 AAACAGGTCC CGGAGGGGCG CCTCCACCAC CATTCTTAAA CACACCCTCA TAATGACAAA
 34081 ATATCTTGCT CCTGTGTGTC CTGTAGCAAA TTAAGAATGG CATTTCATCA TTGACATGCC
 34141 CTGGCTCTAA GTTCTTCTCT AAGTTCTAGT TGTAGATACT CTCTCATATT ATCACCAAAC
 34201 TGCTTAGCCA GAAGCCCCC GGAACAATA GCAGGGGACG CTACAGTGCA GTACAAGCGC
 34261 AGACCTCCCC AATTGGCTCC AGCAAAAACA AGATTAGAAT AAGCATACTG GCAACCACCA
 34321 GTAATATCAT CAAAGTTGCT GGAATATATA TCAGGCAGAG TTTCTTGTA GAATTGAATA
 34381 AAAGAAAAAT TTTCAAAAAA AACATTCAAA ACCTCTGGGA TGCAAAATGCA ATAGGTTACC
 34441 GCGCTGCGCT CCAACATTGT TAGTTTGTAA TTAGTCTGCA AAATAAAAGA AACAGCGTC
 34501 ATATCATAGT AGCCTGTGCA ACAGGTGGAA AAATCAGTCT TTCCATCACA AGACAAGCCA
 34561 CAGGGTCTCC AGCTCTACCC TCGTAAACC TGTATCGTG ATTAACAAC AGCACCAGAA
 34621 GTTCTCGCG GTGCCAGCA TGAATAATTC TTGATGAAGC ATATAATCCA GACATGTTAG
 34681 CATCAGTTAA AGAGAAAAA CAGCCAACAT AGCCTCTGGG TATAATTATG CTTAATCTTA
 34741 AGTATAGCAA AGCCACCCCT CGCGGATACA AAGTAAAGG CACAGGAGAA TAAAAATAT
 34801 AATTATTTCT CTGCTGCTGT TCAGGCAACG TCGCCCCCG TCCCTCTAAA TACACATACA
 34861 AAGCCTCATT AGCCATGGCT TACCAGGCAA AGTACAGCCA ACGCGCAAAG CACAAGCTCT
 34921 AAAGAAGCTC TAAACACTC TCCAACCTCT CCACAATATA TACACAAGCC CTAACTGAC
 34981 GTAATGAGAG TAAAGTGTA AAAATCCCGC CAAACCAAC ACACACCCCG AAACGCGTC
 35041 AGCAGGAAA AGTACAGTT CACTTCTCA ATCCCAACAA GCGTAACTTC CTCTTCTCA
 35101 CGATACGTCA CATCCGATTA ACTTACAACG TCATTTTCCC ACGGCGGCC CGCCATTTT
 35161 AGCCGTTTAC CCCACAGCCA ATCACCACAC AGCGCGCACT TTTTAAATT ACCTCATTTA
 35221 CATATTGGCA CCATTCCATC TATAAGGTAT ATTATTGATG ATG

Polimorfismos y epidemiología molecular AV7

1990

>BA55-90

AAGATGGCCACTCCCTCGATGATGCCGAATGGGCATACATGCACATCGCCGGACAGGATGCTTCGGAGTACCTGAGTCCGGGCTCTGG
 TGCAGTTCGCCCGTGCAACAGACACCTACTTCACTATGGGGAACAAGTTTAGAAACCCACAGTGGCGCCACCCACGATGTGACCAC
 CGACCGTAGCCAGCGACTGATGCTGCGCTTCGTGCCGTTGACCGGGAGGACAATACATACTCTTACAAAGTGGGTACACCTCGCC
 GTGGCGACAACAGAGTGTGACATGGCCAGCACATTCTTTGACATTAGGGGGGTGCTTGATAGAGGTCTAGCTTCAAGCCATATT
 CCGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCAC
 ATACACATTGGCATTGCTTCCAGGAAGGGAGACAATATTACTAAGGAAGGTTAGAAATTGGGAAAGACATTACTGCAGACAACAAG
 CCCATTATCAAGGCATTAAACATATCAGCCAGAGCCTCAAGTTGGAGAGAAGAACTATGGACTGATATTGATGGAAACCAATGAAACCTTTG
 GAGGTAGAGCTCTTAAACCAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTGCAAGACCTACAACATAAAAGGGGGCCAAAGCTAA
 AAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAGACTGAGGAGCCAGATATTGATATGGAATTTTGTGATGGTAGAGAAGCTGCT
 GACGCTTTTTCGCTGAAATTTGCTTTACACGGAATGTCAATTTAGAACTCCAGACAGCCATGTAGTTTACAAGCCAGGAACTT
 CTGATGATAACTCTCATGCAAAATTTGGGTCAACACCCATGCCATAACAGACCAATTACATTGGCTTCAGGGATAAATTTGTAGGCCT
 AATGTACTACAACAGTACTGGAATATGGGAGTTTGGCTGGCCAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAAC
 ACTGAACTGTATATCAGCTTTTGTGATTCTCTGGGAGACAGAACAGATACTTCACTATGTGGAACAGGCTGTGGACAGTTATG
 ATCCCGATGTTGCGATTATTGAAATCAGGCATCGAGGATGAACCTGCTAATTTACTGTTTTCTCTGGATGGCATAGGACAGCGAA
 AACATATCAAGCATTAAACATAGAGACATGCTGGAAGAAAGATAAAGTTTCTACAGCTAATGGAATAGCAATGGAACCAAT
 CTGGCTATGGAATTAATATCAAGCTAATCTTGGAGAAGTTTCTGTACTCAATGTGGCTTTGTACCTTCCAGATGTTTACAAGT
 ACACGCCAATTAACATTACTCTGCCGCTAACACCAACACCTATGAGTACATGAACGGGCGAGTGGTTCCCATCTCTGGTTGATT
 ATACATCAACATTGGCGCCAGGTGGTCTCTTGACCAATGGACAATGTGAATCCATTAAACCACACCGCAATGCTGGCCTACGCTAC
 CGGTCCATGCTTCTGGGCAATGGCGTTTATGTGCTTTCCACATACAAGTGCTTCAACATTCTTTGCTGTCAAGAACCTACTTCTT
 TACCTGGCTCTTACACCTATGAGTGGAACTTCAGAACGGATGTGAACATGGTCTGCAAGTTCCCTTGGAAATGACCT

1991

>BA10-91

ACAGACCAATTACATTGGCTTCAGGGATACTGTTTAGTCTTAATGTACTACAACACTACTGGAATATGGGAGTTTGGCCGGCCA
 AGCATCAGCTGAATGCGAGTGGTGTGACTTGCAGGACAGAAACACTGAACCTGTATATCAGCTTTTCTGATTCTCTGGGAGACAGA
 AGCAGATACTTCACTATGTGAATCAGGCTGTGGACAGCTATGATCCCGATGTTCTGATTATTGAAATCATGGCGTCGAGGATGAAC
 TGCCTAATTACTGTTTTCTCTGGATGGCATAGGACAGGGAACAATATCAAGGCATTAAACCTAGAGACACTGCTATGGGAAAAAGA
 TACTAAAGTTTCTACAGCTAATGAATAGCCATAGGCAACAATCTGGCTATGGAATTAATATCAAGCTAATCTTTGGAGAAGTTT
 CTGTACTCCAATGTGCTTTGTACCTTCCAGATGTTTACAGTATGTTACAGTACGCCAATCAACATTACTCTGCCCGCTAACCAACCTATG
 AGTACATGAACGGGCGAGTGGTTTCCCATCTCTGGTCGATTACATCAACAATGGCGCCAGGTGGTCTCTTGACC

1992

>BA22-92

AGATGGCCACCCCTCGATGATGCCGAATGGGCATACATGCACATCGCCGGACAGGATGCTTCGGAGTACCTGAGTCCGGGTCTGGT
 GCAGTTCGCCCGTGCAACAGACACCTACTTCACTATGGGGAACAAGTTTAGAAACCCACAGTGGCGCCACCCACGATGTGACCACC
 GACCGTAGCCAGCGACTGATGCTGCGCTTCGTGCCGTTGACCGGGAGGACAATACATACTCTTACAAAGTGGGTACACCTCGCCG
 TGGGCGACAACAGAGTGTGACATGGCCAGCACATTCTTTGACATTAGGGGGGTGCTTGATAGAGGTCTAGCTTCAAGCCATATT
 CGGCACAGCTTCAAACTCACTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGGATAGTTACAACGGGAGACATGCTCCACACA
 TACACATTTGGCATTGCTTCCAGGAAGGGAGACAATATTACTAAGGAAGGTTAGAAATTGGGAAAGACATTACTGCAGACAACAAGC
 CCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATACTGATGGAACAAATGAAAAGTTTGG
 AGGTAGAGCTCTTAAACCAGCTACCAAAATGAAGCCATGTACGGGTCTTTTGCAAGACCTACAAATATAAAGGGGGCCAGCTAAA
 AACAGAAAAGTAAACCAACCGAAGGAGATGTTGAGACTGAGGAGCCAGATATTGATGGAATTTTGTGATGGTAGAGAAGCTGCTG
 ACGTTTCTTCCGCTGAAATTTGCTTTACACGGAATGTCTCAATTTGGAACTCCAGACAGCCATGTGGTATACAGCCAGAACTTC
 TGATGGTAACTCTCATGCAATTTGGGTCAACAAGCCATGCCTAACAGACCAATTACATTGGCTTCAGGGATACTTTGTAGGTCTT
 ATGTACTACAACAGTACTGGAATATGGGAGTTTGGCCGGCCAAGCATCACACTGAATGCAGTGGTTGACTTGCAGGACAGAAACA
 CTGAAGTGTATATCAGCTTTTGTGATTCTCTGGGAGACAGAAGCAGATACTTCAAGTATGGAATCAGGCTGTGGACAGCTATGA
 TCCCGATGTTCTGATTATTGAAAATCATGGCGTCGAGGATGAATGCTTAACTTACTGTTTCTCTGAGTGCATAGGACAGGGAAC
 AAATATCAAGGCATTAAACCTAGAGACACTGCTATGGGAAAAAGATACTAAAGTTTCTACAGCTAATGAAATAGCCATAGGCAACAATC
 TGGCTATGGAATTAATATCAAGCTAATCTTTGGAGAAGTTTCTGTACTCAATGTGGCTTTGTACCTTCCAGATGTTTACAAGTA
 CAGCCAACTAACATTACTCTGCCGCTAACCAACACCTATGAGTACATGAACGGGCGAGTGGTTTCCCATCTCTGGTCGATTCA
 TACATCAACATTGGCGCCAGGTGCTCTTGGACCAATGGCAATGTCAATTTAACCACCAACGAATTTGCTGGCAAGCACTACC
 GGTCCATGCTTCTGGCAATGGCGTTATGTGCTTTCCACATACAAGTGCTTCAAAAATCTTGTGTCAAGAACCTACTTCTTCT
 ACCTGGCTCCTACCTATGAGTGGAACTTCAAGAAAGGATGTGACCATGGTCTGCAAGTTCCCTTGGAAATGACCTCAGACAGAT
 GGTGCTACCAATAAGTTTCAACAGCATCAATCCATATGCCACCTTATCCCATGGCTCACAACACAGCTTCAACCTTGAAGCCATGG
 TGCGCAACATGGCGCCAGGTGCTCTTGGACCAATGGCAATGTCAATTTAACCACCAACGAATTTGCTGGCAAGCACTACC
 ACCAATTTCCATCCCATCTCGCACTGGGCAGCCTTCAGGGGCTGGTCTTCACTAGGCTCAAAACCAAGGAGACTCCATCTCTGGA
 TCAGGTTTGCATCCCTACTTCTGATATTCTGGATCTATCCCTACCTGATGGCACCTTTTACCTTAACCACACTTTCAAGAAGGTCT
 CAATCATGTTTGAATCTCTCAGTCACTGGCTGGCAATGACAGGCTGTGTCTCCAAATGAGTTTGAATCAAGCGCACTGGCGATGG
 GGAAGGATACAAATGTGGCCCAATGCCATGACCAAGACTGGTTCTGTTGATGCTTGCCTTCAACATGCTTGCCTACAGGCTAC
 TTTTACATCCCTGAGGATACAAGGATCGCATGTACTCTTTTTCAGAACTTCCAGCCTATGAGCAGGCGAGTGGTTGATGAGGTTA
 ATTACACTGACTACAAAGCCGTACCTTACCATATCAACACAACACTCTGGCTTTGTAGGATACCTTGGCGCTACTATGAGACAAGG
 GGAACCTTACCAGCCAATTTATCCATACCCGCTCATCGGAACCTATGCGGTTAAAGTGTATCCCAAAAAAGTTTCTGTGCGACAGG
 ACCATGTGGCGCATACCGTTTCTCAGCAACTTCTATGTCTGGGAGCCCTTACGGACCTGGGACAGAACTGCTCTATGCCAATCGG
 CCCATGCGCTGGACATGACTTTTGGGTGGATCCCATGGATGAGCCACCTGCTTTATCTTCTTTTGAAGTCTTGCAGTGGTCTCAG
 AGTGACACAGCCACACCGCGCGTCTACAGGCGCTTACCTGCGCACACCGTTCTCGGCCGGCAACGCCACCAACATAAG

>BA23-92

ATGGGCATACATGCACATCGCCGGACAGGATGCTTCGGAGTACTTGAAGTCCGGGTCTGGTGCAGTTCCGCCGTGCCACAGGACACTAC
 TTCAGTATGGGGAACAATTTAGAAACCCACAGTGGCGCCACCCACGATGTGACCACCGACCGTAGCCAGCGACCGATGTGCGCT
 TCGTGCCCGTTGACCGGGAAGACAATACCTACTCTTACAAAGTGGGTACACCTCGCCGTGGGCGACAACAGAGTGGGTGACATGGC
 CAGCAGATCTTTGACATTAGGGGGGTGCTTGATAGAGGTCTAGCTTCAAGCCATATTCCGGCACAGCTTACAATTCATCGGCTCT
 AAGGCGCGCTTCAACATCTCAGTGGATAGTTTACACGGGAGAAGACAATGCCACACATACACATTGGCTTGGCTTCCAGCAAGG
 GAGACAATATTACTAAGGAAGGTTTAGAAATTGGGAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCC
 AGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAA
 ATGAAGCCATGTACGGGTCTTTTGAAGACCTACGAATATAAAGGGGGCCAAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAG
 ATGTTGAGACTGAGGAGCGAGATATTGATATGGGATTTTGTGATGAGAGAAGCTGCTGACGCTTTTTCGCTGAAATTTGCTTTA
 CACGGAATGTCAATTTAGAACTCCAGACAGCCATGTAGTATACAACCCAGGAACCTTCTGATGGTAACTCTCATGCAATTTGGGT

CAACAACCCATGCCTAACAGACCCCAATTACATTGGCTTCAGGGATAACTTTGTAGGCCTAATGTACTACAACAGTACTGGAAATATGG
GAGTTTTGGCCGGCAAGCATCAAACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGGTTGA
TTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTCTGATTATTGAAATCAT
GGCGTCGAGGATGAACCTAATTACTGTTTTCTCTGGATGGCATAGGACCAGGGAACAAATATCAAGGCATTAAACCTAGAGACA
CTGCATGGGAAAAGATACTAAAGTTTCTACAGCTAATGAAATAGCCATAGGCAACAACTCTGGCTATGGAAATTAATATCCAAGCTAA
TCTTTGGAGAACCTTTCTGACTCCAATGTGGCTTTGACTCTCCAGATGTTTACAAGTACACGCCAACTAACATTACTTGGCCCGCT
AACACCAACAGTTTATGAGTACATGAACGGGCGAGTGGTATCCCCATCTCTGGTCGATTATACATCAACATTGGCGCCAGGCTGGTCTC
TTGACCCAAATGGACAATGTGAATCCATTCAACCACCACCGCAATGCTGGTCTGCGCTACCGGTCCATGCTTCTGGGCAATGGTCTGTA
TGTGCTTTCCACATACAGTGCCTCAAAAATTTCTTGTCTGCAAGAACCTACTTCTTCTACCTGGCTCCTACACCTATGAGTGGAAAC
TTCAGAAAGGATGTGAACATGGTCTCTGCAAGTTCCCTTGGAAATGACCTCAGAACAGATGGTGCTACCATAAGTTTACCAGCATCA
ATCTGTATGCGAATCTTCTCCCATGGCTCACAACACAGCTTCCACCCTTGAAGCCATGCTGCGCAACAGGATCAACCTGATCTAGT
CAACGACTACCTCTCTGCAGCTAACATGCTTTACCCCATCCCTGCCAATGCAACCAACATTCCAATTTCCATCCCATCTCGCAACTGG
GCAGCCTTCAGGGGCTGGTCTTCACTAGGCTCAAAACCAAGGAGACTCCATCTCTTGGATCAGGGTTCGATCCCTACTTCTGATATT
CTGGATCTATTCCCTACCTGGATGGCACCTTTTACCTTAACCAACTTTCAAGAAGGTCTCCATCATGTTTGAATCTCCTCAGTCACTG
GCCTGGCAATGACAGGCTGTTGAGCCCAAATGAGTTGAAATCAAGCGCACTGTGGACGGGGAAGGTTACAATGTGGCCCAATGTAAAC
ATGACCAAGAGACTGGTTCCTGGTTGAGATGCTTGCCAACTACAACATTGGCTACCAGGGCTTTTACATCCCTGAGGGATACAAGGATC
GCATGTACTCTTTTTCAGAACTTCCAGCCTATGAGCAGGAGGTTGGTGTAGGTTAATTAACCTGACTACAAAGCCGTCACTT
ACCATATCAACACAACCTCTGGCTTTGTAGGATACCTTGGCCTACTATGAGACAAGGGGAACCTTACCAGCCAAATATCCATAC
CCGCTCATCGGAATCTTCTGCGCTTAAAGTTTACCCAAAAGTTCTGTGCGACAGGACCTATGGCGCATACCTGTTCTCAGCA
ACTTCTATGTCATGGAGCCCTTACGGACCTGGGACAGAACCTGCTCTATGCCAACTCGGCCATGCGCTGGACATGACTTTTGAAGT
GGATCCCATGGATGAGCCACCCTGCTTTATCTTCTTTCGAAGTCTTCGACAAGGTGAGTGCACACAGCCACACCGCGCGTAACT
GAGGCCGTCTACCTGCGACA

1998

>BA3-99

TGCACATCGCCGGACAGGATGCTTCGGAGTACCTGAGTCCGGGTCTGGTGCAGTTTCGCCCCTGCAACAGACACCTACTTCAGTATGGG
GAACAAGTTTAGAAACCCACAGTGGCGCCACCCACGATGTGACACCGACCGTAGCCAGCGACTGATGCTGCGCTTCGTGGCCGTT
GACCGGGAGGACAATACATACTCTTACAAAGTGGGTACACCTCGCCGTGGGCGACAAACAGAGTGGCTTGACATGCGCCAGCATTCT
TTGACATTAGGGGGCTCTTGATAGAGTCTTCAAGCTTATTCGCGCACAGCTTACAATTCACTGGCTCCTAGGCTCCTAGGCGCGCC
TAACACATCTCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTGGCATTGCTTCCACGAAGGGAGACAATATT
ACTAAGGAAGTTTAGAAATGGGAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAG
TTGGAGAAGAATCATGGAAGTATGATGGAACAAATGAAATTTGGAGGTAGAGCTCTTAAACAGCTACTAAATGAAGCCATG
CTACGGGTCTTTTGCAGAGCTTCAAAACATAAAAGGGGCCAAGCTTAAACACAGAAAGTAACACCAACCGGAGAGATGTTGAAGCT
GAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGTCTGCTGACGCTTTTTCGCTGAAATTTGTGCTTTACACGGAAATG
TCAATTTGGAACCTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACCTCATGCAAAATTTGGGTCAACAAGCCAT
GCCTAACAGACCCAAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTAATAACAGTACTGGAATATGGGAGTTTGGCC
GGCCAAAGCATCAAACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATTCTCTGGGAG
ACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTCTGATTATTGAAATCATGGCGTCGAGGA
TGAAGTGCCTAATTACTGTTTTCTCTGGATGGCATAGGACCAGGGAACAAATATCAAGGCATTAAACCTAGAGACCTGCTATGGGAC
AAAGTGTATAATGTTTCTAAATCTAATGAAATAGCCATAGGCAACACAGGCTATGGAATTAATATCAAGCTAACCTTTGGAGAA
GTTTTCTGATACTGCTCAAGTGGTTTGTACCTTCAGGATGTTTACAGTATACACGCAACTAACATTACTCTGCCCGCCAAACAAACAC
CTATGAGTACATGAACGGGCGAGTGGTATCCCATCTCTGGTTGATTATACATCAACATTGGCGCCAGGTGGTCTCTTGACCCAAATG
GACAATGTGAATCCATTCAACCACCACCGCAATGCTGGTCTGCGCTACCGGTCCATGCTTCTGGGCAATGGTGGTATGTGCTTTCC
ACATACAAGTGCCTCAAAAATTTCTTGTCTGCAAGAACCTACTTCTTCTACCTGGCTCCTACACCTATGAGTGGAACTTCAGAAAGGA
TGTGAACATGGCTGCAAGTTCCTTGGAAATGACCTCAGAACAGGTTGGTCTACCATAAGTTTACCAGCATCAATCTGTATGTC
ACCTTCTTCCCATGGCTCACAACACAGCTTCCACCCTTGAAGCCATGCTGCGCAACGATACCAATGATCAGTCAATTCAACGACTACC
TCTCTGACGTAACATGCTTTACCCCATCCCTGCCAATGCAACCAACATTCCAATTTCCATCCCATCTTGCAGCTGGCAAGCTTCAG
GGGTGGTCTTTACACAGACTCAAAACCAAGAACTCCATCTCTTGGATCAGGGTTCGATCCCTACTTCTGTTATTCTGGATCTATT
CCCTACCTGGATGGCACTTTTACCTTAACCACTTTCAAGAAGGTCTCCATCATGTTTGAATCCTCAGTCACTGGCTGGCAATG
ACAGGCTGTTTGTCTCAAAATGAGTTTGAATCAAGCCATCTGGATGGGGAAGGATACAATGTGGGCCAATGCAACATGACCAAGA
CTGGTTCTGTTGATGATGCTTGCCTTCAACATTGGCTACCAGGGCTTTTACATCCCTGAGGGATACAAGATCGCATGTACTCC
TTTTTCAGAACTTCCAGCCTATGAGCAGGCAAGTGGTGTAGAGTTAATTAACCTGACTACAAAGCCGTCACTTACCATATCAAC
ACAACAACTCTGGCTTTGTAGGATACCTTGCCTTACTATAAGACAAGGGGAACCTTACCAGCCAAATTCATCCATCCGCTCATCGG
AATCACTGCGCTTAAAGATTACCCAAAAGTTCTGTGCGACAGGACCATGTGGCGCATACCTTCTCCAGCACTTCTATGTCC
ATGGGAGCCCTTACGGACCTGGGACAGAACCTGCTCTATGCCAACTCGGCCATGCGCTGGACATGACTTTTGAAGTGGATCCCATGG
ATGAGCCACCCCTGCTTTATCTTCTTTCGAAGTCTTCGACGTGGTTCAGAGTGCACACAGCCACACCGCGCGTCATCGAGGCCGTCTA
C

1999

>BA1-99

TCCGGGTCTGGTCCAGTTCGCCCCTGCAACAGACACCTACTTCAGTATGGGGAACCAATTTAGAAACCCACAGTGGCGCCACCCAC
GATGTGACCACCGACCGTAGCCAGCAGTATGCTGCGCTTCGTGCCCGTTGACCGGGAGACCAATACATACTCTTACAAAGTGGGT
ACACCTTCGCGGTGGGCGACACAGAGTGTGACATGGCCAGCACTTTTGACATTAGGGGGGTGCTTGGCAGAGGTCTTCAAGCTT
CAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCTAACACATCTCAGTGGATAGTTACAACGGGAGAAGAC
AATGCCACACATACATATTGGAATTGCTTCCACGAAGGGAGACAAATATTACTAAGGAAGTTTAGAAATGGGAAAGACATTACTG
CAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGATCATGGAATGATATGATGGAACAA
TGAAATTTGAGAGTTAGAGCTTTAAACACAGCTACTAAATGGAAGCCATGCTACGGGTCTTATGCAAGACCTACAACAACTAAAGGG
GGCCAAAGCTAAACACAGAAAGTAAACCAACCGAAGGAGATGTTGAGACTGAGGAGCCAGATATTGATATGGAATTTTGTATGGTA
GAGAAGTCTGCTGACGCTTTTTCGCTGAAATTTGTGCTTTACACGGAAATGTCAATTTAGAACTCCAGACAGCCATGTAGTTTACAA
GCCAGGAACCTCTGATGATACTCTCATGCAATTTGGGTCAACAAGCCATGCTTAAACAGACCCAAATTCATTGGCTTCAGGATAAC
TTGTAGGCTTATGATGATACTCAACAGTACTGGAATATGGGAGTTTGGCTGGCCAAAGCATCAACACTGTCAGGGGTCTTGTG
AGGCCAGAAACCTTGGACTGTCAAACAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATCTTCAGCATGTGGAATCAGGCTGT
GGACAGCTATGATCCCGATGTTCTGATTATTGAAATCATGGCGTCGAGGATGAACTGCCTAATTACTGTTTTCTCTGGATGGCATA
GGACAGGGAACAAATCAAGGCATTAAACCTAGAGACATGCTGGAAGAAAGATACTAAAGTTTCTACAGCTAATGAAATAGCCA
TAGGCAACAATCTGGCTATGGGAAATTAATATCAAGCTAATCTTTGGAGAAGTTTCTGTAATGCTTCAATGCTTCTGTAATCTTCCAGA
TGTTTACAAGTACACGCCAACTAACATTACTCTGCCGCTAACACCAACACCTATGAGTACATGAACGGGCGAGTGGTTTCCCATCT
CTGGTTCGATTATACATCAACATTGGCGCCAGGTGGTCTCTTGACCAATGGACAATGTGAATCCATTTAACACCACCGCAATGCTG
GCCTACGCTACCGGTCCATGCTTCTGGCAATGGCCGTTATGTGCTTTCCACATACAAGTGCCTCAAAAATTTCTTGTCTGCAAGAA
CCTACTTCTTACCTGGCTCCTACACCTATGAGTGGAACTTCAGAAAGGATGTGAACATGGTCTGCAAGATTCCTTGGAAATGAC

CTCAGAACAGATGGTGCTACCATAAGTTTACCAGGATCAACTCCTATGCCACCTTCTTCCCCATGGGTGACAAACACCGGTTCAAATC
TTGAAGCCATGCTGCGCAACGATACCAATGATCAGTCATTCAACGACTACCTCTCTGCAGGTAACATGCTTTACCCCATCCCTGCCAA
TGCAACCAACATTCCAATTTCCATCCCATCTCGCACTGGGAGCCTTACGGGGCTGGTCTTCACTAGGCTCAAAACCAAGGAGACT
CCATCTCTTGGATCAGGGTTTCATCCCTACTTCGTATATTCTGGATCTATTCCCTACCTGGATGGCACCTTTTACCTTAACCACTT
TCAAGAAGGTCTCAATCATGTTTGAATCCTCAGTCAGCTGGCCTGGCAATGACAGGCTGTGTCTCAAATGAGTTTGAATCAAGCG
CACTGTGGATGGGGAAGGATACAATGTGGCCCAATGCAACATGACAAAGACTGGTTCCTGGTTAGATGCTTCCCACTACAACATT
GGCTACCAGGGCTTTTACATCCCTTGGGAATACAAGGATCGCATGTACTCTTTTTCAGAACTTCCAGCCTATGACCAGCCAGTTGT
TTGATGAGGTTAATTACTGACTACAAAGCCGTACCTTACCATATCAACACAACAACTCTGGCTTTGTAGAATACCTTGCAGCTAC
TATGAGACAAGGGGAACCTTACCCAGCCAATTATCCATACCCGCTCATCGGAACTACTGCGCTTAAAGTGTACCCAAAAAAGTTTC
CTGTGCGACAGGACCATGTGGCGCATACCGTTCTCCAGCAACTTCAATGTCCATGGGAGCCCTTACGGACCTGGGACAGAACCTGCTCT
ATGCCAACTCGGCCATGCGCTGGACATGACTTTGAGGTGGATCCCATGGATGAGCCACCCCTGCTTTATCTTCTTTTCGAAGTCTT
CGACGTGGTCAGAGTGACCCAGCCACCGCGCGCTCATCGAGGCGCTCTACCTGCGCACACCGTTCTCGGCCGGCAACGCCACCACA
T

>BA4-99

GCTTGATAGAGTCTAGCTTCAAGCCATATTCCGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGG
ATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAG
AAATTGGGAAGACATTACTGCAGACAACAAGCCATTATGCGGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATG
GACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACAGCTACTAAAAATGAAGCCATGCTACGGGTCTTTTGCA
AGACCTACAACATAAAAGGGGGCCAAAGCTAAAAACAGAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTG
ATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCTGAAATTTGTGCTTTACACGGAAAAATGTCAATTTGGAACCTCC
AGACAGCCATGGGTATACAAGCCAGGAACCTTCTGATGGTAACCTCTCATGCAAAATTTGGGTCAACAAGCCATGCTTAAACAGACCAAT
TACATTGGCTTCAGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAATATGGGAGTTTGGCCGGCCAAAGCATACAAC
TGAATGTGCTGTTGACTTCTGAGGACAGAACTGAACTGAACTTCTGCTTGAATCTCTGGGAGACAGAACTGAACTGAACTTCT
CAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTTCGTATTATTGAAATCATGGCGTCGAGGATGAACTGCCTAATTAC
TGTTTTCTCTGGATGGCATAGGACAGGGAAACAAATATCAAGGCATTAAACCTAGAGACACTGCATGGGAAAAAGATACTAAAGTTT
CTACAGCTAATGAAATAGCCATAGGCAACAATCTGGCTATGGAATTAATATCAAGCTAATCTTTGAGAAGTTTCTGTACTCCAA
TGTGGCTTTTGTGACTTCTGAGATGTTTCAAGTACACGCCAACTTACTCTGCCCCGTAAACACCAACACTGAACTGAACTGAACT
GGGCGAGTGGTTTCCCATCTCTGGTTCGATTACATCAACATTGGCGCCAGGTGGTCTCTTGACCAATGGACAATGTGAATCCAT
TTAACCAACACCGCAATGCTGGGCTACGCTACCGGTCCATGCTTCTGGGCAATGGCCGTTATGTGCTTTCCACATACAAGTGCCTCA
AAAATCTTTGCTGTCAGAACCTACTTCTTCTACCTGGCTCTACCTATGAGTGGAACTTCAAGAAAGATGTGAACATGGTCTCTG
GAAAGTTTGTGCTGACTTCAAGTCTGAGAACAGATGGTGGTACCTAAGTTTCAACAGGATCAATCTCTATGCCACTTCTCCCATGG
GTCAACACACAGGTTCAACTCTTGAAGCCATGGTGGCGCAACGATACCAATGATCACGGATTCAACGAATACCTCTCTGCAGCTAACAT
GCTTTACCCCATCCCTGCCAATGCAACCAACATTTCCATTTTCCATCCCATCTCGCAACTGGGCAAGCTTCAAGGGCTGGTCTTCCAC
AGACTCAAAACCAAGAGACTCCATCTCTGGATCAGGGTTCGATCCCTACTCTGTTTATTCTGGATCTATTCCCTACCTGGATGGCA
CTTTTTCTTCAACCACTTCAAGAAAGTCTCCATCATGTTTACTCTCAGTCAGCTGGCTGGCAATGACAGGCTGTTGTCTCC
AAATGAGTTTGAATCAAGCGCACTGTGGATGGGGAAGGATGTTGGGCCAATGCAACATGACCAAGACTTGGTCTCTGGTTTCAG
ATGCTTGCCAACTACAACATTGGCTACAGGGCTTTTACATCCCTGAGGGATACAAGGATCGCATGTACTCTTTTTCAGAACTTCC
AGCCTATGAGCAGGCGAGGTGGTTGATGAGGTTAATTACACTGACTACAAAGCCGTCACCTTACCATATCAACACAACAACTCTGGCTT
TGATGAGTACCTTGGCCTACTATGAGACAAGGGGAACCTTACCAGCCAATTATCCATACCCGCTCATCGGAATCTGCGCTTAAAG
AGTGTATACCAAAAAAAGTTCTGTGCGACAGAACCATGTGGCGCATACCGTTCTCCAGCAACTTCAATGTCTGAGGAGCCCTTACGG
ACCTGGGACAGAACCTGCTCTATGCCAACTCGGCCATGCGCTGGACATGACTTTTGAAGGTGGATCCCATGGATGAGCCACCCCTGCT
TTATCTTCTTTTCGAAGTCTTCGACGTGGTCAGAGTGCA

>BA496-99

GATACCTGAGTCCGGTCTGGTGCAGTTCGCCCGTGCAACAGACACCTACTTCAGTATGGGGAACAAATTTAGAAACCCACAGTGGC
GCCCACCCAGATGTGACCACCGATCGTAGCCAGCGCTGATGCTGCGCTTCTGTCGCCGTTGACCGGGAAGACAATACCTACTCTTAC
AAAGTTTCGTACACGCTGGCTGTAGGCGACAACAGAGTGCTTGACATGGCCAGCACATTTCTTGACATTCGGGGGGTGCTTGATAGAG
GTCCTAGCTTCAAGCCATATTCCGGCACAGCTTACAATTCAGTCTCTAAGGGCGCGCCCAATACATCTCAGTGGATAGTTACAAC
GAATCGAGCAATGAGTAACTACCAACCAACACATTTGCGCATTTGCTTCCATGAAGGGAGACAATATTCTAAGAAAGGTTTGCAA
ATTGGGAAGACATTACCGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGA
CTGATCTGATGTAACAAATGAAAAGTTTGGTGAAGAGCCCTTAAACAGCTACCAACATGAAGCCATGCTACGGGTCTTTTGCAAG
AACTACAACATAAAAGGGGGCCAAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGGGGTTGAACTGAGGAACAGATATTGAT
ATGGAATTTTTCAGTGGTAGAGAAGTCTTGCAAGTTTACGCTGAAATTTGTCTTTATACGGAAAAATGTCATTTAGAACTCCAG
ACAGCCATGTAGTTTACAAGCCAGGAACCTTCTGATGATACTCTCATGCAAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAATTA
CATTGGCTTCAGGGATACTTTGTAGGCTAATGTACTACAACAGTACTGGAATATGGGAGTTTGGCTGGCCAAAGCATCACAACTG
AATGCACTGGTTGACTTGCAGGACAGAACTGAACTGTCATACAGCTTTTGGCTGATTCTCTGGGAGACAGAACAGATACTTCA
GTATGTGGAACCAAGCTTGTGAGCAGTTATGATCCCGATTTGCGCATTTGTCATTTGAAATTCAGGCATCGAGGATGACTGCTAATTA
TTTTCTCTGGATGGCATTTATCAAGCATTAAAGTATCAAGGCATTAAATGTAACTCAATGCATGGGCAAAAGATGCTAATGTTGCT
ACAGCTAATGAAATAGCCATAGGAAACAACCTAGGCTATGGAATTAATATCCAAGCTAACTTGGAGAAGTTTCTGTACTCCAATG
TGGCTTTGTACCTTCCAGATGTTTAGAACTAAACATTCACTAACATTACTCTGCCCGCAACACCAACCTATGAGTACATGAACGG
GCGAGTGGTATCCCATCTCTGGTTGATTATACATCAACATTTGGCGCCAGGTGGTCTCTTGACCCAAATGGACAATGTGAATCCATT
AACCACCACCGCAATGCTGGTCTGCGCTACCGGTCCATGCTTCTGGGCAATGGTCTGTTATGTGCTTTCCACATACAAGTGCCTCAA
AATTCTTTGCTGTCAAGAACCTACTTCTTCTACCTGGCTCTACACCTATGAGTGGAACTTCAAGAAAGATGTGAACATGGTCTGCA
AAGTTCCCTTGGAAATGACCTCAGAACAGATGGTGTACCATAAAGTTTACCAGCATCAATCTGTATGCCACCTTCTTCCCATGGCT
CACAAACCTGTTCCACCTTGAAGCCATGCTGCGCAACGATACCAATGATCAGTCATTCAACGACTACCTCTCTGAGCTAACTATG
TTTACCCCATCCCTGCCAATGCAACCAACATTTCCAATTTCCATCCCATCTCGCAACTGGGCAACCTTCAAGGGTGGTCTTCCACG
GCTAAACCCCAAGAACTCCTATTTTGGATCAGGGTTTGTACCTATTTGGAATATCTGGATCTATTCCCTACCTGGATGGCCCC
TTTTACCTTAAACCACTTTCAAGAAGGTCTCCATCATGTTTGAATCTCTCAGTCAGCTGGCTGCAATGACAGGCTGTTGAGCCCCAA
ATGAGTTTGAATCAAGCGCACTGTGGACGGGAAGGATACAACGTCAGCAACATGACCAAGAGTGGTCTCTAGTTGATGAT
GCTTGCCAACTACAACATTGGCTACAGGGCTTTTACATCCCTGAGGGATACAAGGATCGCATGTACTCTTTTTCAGAACTTCCAG
CCTATGAGCAGGCGAGGTGGTTGATGAGGTTAATTACACTGACTACAAAGCCGTCACCTTACCATACCAACACAACAACTCTGGCTTTG
TAGGGTACCTTGCACTTACTATGAGACAAGGGGAACCTTACCAGCCAATTATCCATACCCGCTCATCGGAATCTGCGCTTAAAG
TGTTACCAGAAAAAGTTCTGTGCGACAGGACCATGTGGGATCTCCCTTCTCCAGCAACTCATGTCTGAGGAGGGGCTTACCGAC
CTGGGACAGAACATGCTCTATGCCAACTCAGCCCATGCGCTGGACATGACTTTTGAAGGTGGATCCCATGGATGAGCCACCCCTGCTT
ATCTTCTTTTCGAAGTCTTCGACGTGGTCAGAGTGCA

>BA501-99

CCTGAGTCCGGGTCTGGTGCAGTTCGCCCGCGCCACAGACACCTACTTCAATCTGGGAAATAAGTTTAGAAATCCACCGTAGCGCGG
ACCCAGATGTGACCACCGATCGTAGCCAGCGCTCATGTTGCGCTTCTGTCGCCGTTGACCGGAGGACAATACATACTCTTACAAAG

TGCGGTACACACTGGCCGTGGGCGACAACAGAGTGTGGATATGGCCAGCACGTTCTTTGACATTAGGGGCGTGTGGACAGAGGTCC
CAGTTTCAAACCTATTCTGGTACGGTTTACAACCTCTCTGGCTCTTAAAGGCGCTCCAAATACATCTCAGTGGAT

>BA991-99

CTAGCTTCAAGCCATATTCCGGCACAGCTTACAATTCAGTGGCTCCTAAGGGCGCGCCCTACACATCTCAGTGGATAGTTACACCGGG
AGAAGACAATGCCACCACATACACATTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAGAAATTGGGAAAGAC
ATTACTGCAGACTACAAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATCTGATG
GAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACAGCTACTTAAATGAAGCCATGTCTACGGGTCTTTTGCAAGACCTACAAACAT
AAAAGGGGGCCAGCTAAAACAGTAAAGTAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTC
GATGGTAGAGAAGCTGTGACGCTTTTTCGCCTGAAATTTGTGCTTTACACGGAAAAATGTCAATTTGGAAGCTCCAGACAGCCATGTGG
TATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAAATTACATTGGCTTCAG
GGATAACTTTGTAGTCTTATGTACTACACAGTACTGGAATATGGAGTTTGGCCGGCCAAAGCATCACTACTGAATGCAGTGGTT
GACTTGACAGACAGAACTGAAGTGTATATCAGCTTCTGCTTGATTCTCTGGGAGACAGAAGCAGATACCTCAGCATGTGGAATC
CGACTGAGGACAGTTATGATCCGATGCTCGTAATATTGAAATCATAGCATCG

>BA1325-99

GCTTGATAGAGGTCTAGCTTCAAGCCATATTCCGGCACAGCTTACAATTCAGTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGG
ATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAG
AAATTGGGAAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATG
GACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACAGCTACTTAAATGAAGCCATGTCTACGGGTCTTTTGCA
AGACTACAACTAAAAGTGGGCGCAAGCTAAAACAGGAAATATACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTG
ATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTGCTTTACACGGAAAAATGTCAATTTGGAAGCTCC
AGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAAAT
TACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAATATGGGAGTTTGGCCGGCCAAAGCATCACAAC
TGAATGCAGTGGTTGACTTGCAGGACAAAACACTGAAGTGGCATATCAGCTTTTGTGATTCTCTGGGAGACAGAAACAGATACTT
CAGCATGTGGAATCAGGCTGGGACAGCTATTATCCCGATGTTGCGATTATTGAAAAATCAGGCATCGAGGATGAATCGCTTAATTAC
TGTTTTCTCTGGATGGCATAGGACCAGGGACCAATATCAAGGCATTAAACCTAGAGACACTGCATGGGAAAAAGATACTAAAGTTT
CTACAGCTAATGAAATAGCCATAGGCAACAATCTGGCTATGGAATTAATATCCAAGCTAATCCTTGGAAGAAGTTTCTGTACTCCAA
TGTGGCTTTGTACCTTCCAGATGTTTACAAGTACAGGCCAACTAACTACTCTGCCGCTAACACCAACCTATGAGTACATGAAC
GGCGAGTGGTTTCCCATCTCTGGTGCATTATACATCAACATTTGGCGCCAGGTGGTCTCTTGACCAATGGACAATGTGAATGCAG
TTAACCAACCCGCAATGCTGGGCTACGCTACCGGTCCATGCTTCTGGGCAATGGGCGTTATGTGCTTTCCACATACAAGTGCCTCA
AAAATTTCTTGTGTCAAGAACCTACTTCTTCAACTGGGTCTACACCTATGAGTGAAGCTTCAAGAAAGATGTGAACATGGTCTCTG
CAAAGTTCCCTTGGAATGACCTCAGAACAGATGGTGTACCATAGTTTACCAGGATCAAACTCTATGGCACCTTCTTCCCATGG
GTGCAAAAACGGGTTCAAATCTTGAAGCCATGCTGCGCAACGATACCAATGATCAGTCAATCAACGACTACCTCTCTGAAGCTAACAT
GCTTTACCCCATCCCTGCCAATGCAACCAACATTCCAATTTCCATCCATTCTCGCAATTTGGGCAGCCTTCAAGGGCTGGTCTTACC
AGACTCAAAACCAAGGAGACTCCATTTCTGGATCAGGGTCCGATCCCTACTTCGTTTATTCTGGATCTATTCCCTACCTGGATGGCA
CTTTTACCTTAACCACTTTCAAGAAGGTCTCCATCATGTTTGAATCCTCAGTCAGCTGGCCTGGCAATGACAGGCTGTTGTCTCC
AAATGAGTTTGAATCAAGCGCACTGTGGATGGGGAAGGATACAATGTGGCCCAATGCAACATGACCAAGACTGGTTCTGTGGTTTCA
ATGCTTGCCAACTACAACATTGGCTACCAGGGCTTTTACATCCCTGAGGGATACAAGGATCGCATGTACTCTTTTTAGAACTTCC
AGCCTATGAGCAGGCGAGTGGTTGATGAGGTTAATTACACTGACTACAAAGCCGTCACCTTACCATATCAACACAACAACTCTGGCTT
TGATGGATACCTTGCGCTACTATGAGACAAGGGAACTTACCAGCCAATATCCATACCCGCTCATCGGAATCACTGCGGTTAAA
AGTGTATACCAAAAGGTTCTGTGCGACAGGACCATGTGGCGACCATCCGTTCTCCAGCACTTCAATGCTGGGAGCCCTTACCGG
ACCTGGGACAGAACCTGCTCTATGCCAATCGGCCATGCGCTGGACATGACTTTTGGAGTGGATCCCATGGATGAGCCACCCCTGCT
TTATCTTCTTTTGAAGTCTTTCGACGTGGTCAAGTGCACAGCCACACCGCGCGTCTATCGAGGCCGTCTACCTGCGCACACCGGTT
TCGGCCGGCAACGCCACCACATAAG

>BA1390-99

AGCCTCAAGTTGGAGAAGAATCATGGACTGATCTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACAGCTACTAAAAT
GAAGCCATGCTACGGGTCTTTTGAAGACCTACAATATAAAAGGGGGCCAAGCTAAAACAGAAAAGTAACACCAACCGAAGGAGAT
GTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGTGACGCTTTTTCGCCTGAAATTTGTGCTTTACA
CGGAAATGTCAATTTGGAACCTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAATTTGGGTCA
ACAAGCCATGCCTAACAGACCCAAATTACATTGGCTTCAGGGATAACTTTGTAGGTTCTATGTACTACAACAGTATGGAAATATGGGA
GTTTTGGCTGGCCAAAGCATCAAACTGTATGCAGTGGTTGACTCGCAGGACAGAAACACTGAAGTGTATATCAGCGATAGCGTGATT
CATTTGGGAGACAGGAGCAGATACTTACGAGGGGGAATCAGGCTGTAGAGAGCTATTATCGCGATATAGGGATC

>BA1419-99

ACCTGAGTCCGGGTCTGGTGCAGTTTCGCCCCTGCAACAGACACCTACTTCAGTATGGGGAACAAATTTAGAAACCCACAGTGGCGCC
CACCCAGATGTGACCACCGACCGTAGCCAGCGCTGATGCTGCGCTTCGTGCCCCTGACCGGGAAGACAATCCTACTCTTACAAA
GTTTCGCTACACGCTGGCTGTAGGCGACAACAGAGTGTGATGATGGCCAGCACATTCTTTGACATTCCGGGGGTGCTTGATAGAGGT
CTAGCTTCAAGCCATATTCCGGCACAGCTTACAAT

>BA2972-99

GGCCACCCCTCGATGATCCCCAATGGGCAAAAATGCATCGCCGACAGGATGCTTCGGAGTACCTGAGTCCGGGTCTGGTGCAG
TTTCGCCCCTGCAACAGACACCTACTTCAGTATGGGGAACAAGTTTAGAAACCCACAGTGGCGCCACCCACGATGTGACCACGACC
GTAGCCAGCGACTGATGCTGCGCTTCGTGCCCCTGACCGGGAGGACAATACATACTCTTACAAAGTCCGGTACACCTCGCCGTGGG
CGACAACAGAGTGCTTGACATGGCCAGCACATTCTTGACATTAGGGGGGTGCTTGATAGAGGTCTAGCTTCAAGCCATATTCCGGC
ACAGCTTACAATTCAGTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCACATACA
CATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTAGAAATTTGGGAAAGACATTACTGCAGACAACAAGCCCAT
TTATGCCGATTAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGT
AGAGCTCTTAAACAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTACACATA

>BA3106-99

ACCCAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCA
TCACAACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAAGTGTATATCAGCTTTTGTGATTCTCTGGGAGACAGAAGCA
GATACTTCAGCATG

>BA3150-99

GCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGTGACGCTTTTTCGCCTGAAATTTGTGCTTTACACGGAAAAATGTCAAT
TTGGAACCTCCAGACAGCCTGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAATTTGGGTCAACAAGCCATGCCTA
ACAGACCAATTAATTCAGTTGGCTTCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAAATATGGGAAATTTGGCCGGCCA
AGCATCAAACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAAGTGTATATCAGCTTTTGTGATTCTCTGGGAGACAGA
AGCAGATACTTCAGCATGTGGAATCAGCCTGTGGACAGTTATGATCCCGATGTTCTGATTATTGAAATCA

>CONS3187-99

TCGCCGGACAGGATGCTTCGGAGTACCTGAGTCCGGGTCTGGTGCAGTTTCGCCGTGCAACAGACACCTACTTCAGTATGGGGAACAA
 GTTTAGAAACCCACAGTGGCGCCACCACAGATGTGACACCGTAGCCAGCGACTGATGCTGCGCTTCGTGCCCGTTGACCGG
 GAGGACAATACATACTCTTACAAAGTGCAGTACACCTCGCCGTGGGCGACAACAGAGTGTGACATGGCCAGCACATTCTTTGACA
 TTAGGGGGGTGCTTGATAGAGGTCCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACAC
 ATCTCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAG
 GAAGGTTTAGAAATTGGGAAAGACATTACTGCAGACAACAAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAG
 AAGAATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAAATGAAGCCATGCTACGG
 TTCTTTTGCAAGACCTACACATATAAAGGGGGCCAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAGACTGAGGAG
 CCAGATATTGATATGGAATTTTTTGATGGTACAGAAGCTGCTGACGCTTTTCGCCTGAAATTGTGCTTTACACGGAAAATGTCATTT
 TGGAAACTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAA
 CAGACCCAAATTAACCTTCAGGGATAACTTTGTAGGCTTATGTACTACAACAGTACTGGAAATATGGGAGTTTGTGGCGGCCAA
 GCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACCTGAACTGTCTATCATGCTTTTGTCTGATTCTCTGGGAGACAGAA
 GCAGATACCTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTGCGATTATTGAAAATCATGGCACCAGCATGAACT
 GCCTAATTACTGCTTCTCTCTGGATGGCATAGGACCAGGGAACAAATATCAAGGCATTAAACCTAGAGACACTGCATGGGAAAAAGAT
 ACTAAAGTTTCTACAGCTAATGAAATAGCCATAGGCAACAACTCTGGCTATGGAAATTAATATCCAAGCTAATCTTTGGAGAAATTTTC
 TGTACTCCAATGTGGCTTTGTACCTTCCAGATGTTTACAAGTACACGCCAACTAACATTACTCTGCCCGCTAACACCAACCTATGA
 GTACATGAACGGGCGAGTGGTTTCCCATCTCTGGTTCGATTCTATCATCAACATTTGGCGCCAGGTGGTCTCTTGACCCAAATGGACAAT
 GTGAATCCATTTAAACCACCACCGCAATGCTGGCTACGCTACCGGTCATGCTTCTGGGCAATGGCCGTTATGTGCTTTCCACATAC
 AAGTGCCTCAAAAATTTCTTGCTGTCAAGAACCTACTTCTTCTACCTGGCTCTACACCTATGAGTGGAACTTCAGAACGGATGT

>BA3434-99

TTTTTTGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGGAAACTCCAGACAGCC
 ATGTGGTTTACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTGACAGACCCAAATTACATTGG
 CTTACAGGATAAATTTGTAGGCTTATGATGCTTACACAGTACTGGAAATATGGAGTTTGGCGCCGCAAGCATCAACATGAATGCCA
 GTGGTTGACTTGCAGGACAGAAACCTGAACTGTCTATCATGCTTTTGTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGT
 GGAATCAGGCTGTGGACAGCTATGATCCC

2000

>BA3017-00

AGAGGTCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCTAACACATCTCAGTGGATAGTTA
 CAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAGAAATTGG
 GAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGAT
 ATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTCAAGACCTA
 CAAACATAAAGGGGGGCAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCAGATATTGATATGGA
 ATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGGAAACTCCAGACAGC
 CATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAAATTACATTG
 GCTTCAGGGATAAATTTGTAGGCTTATGTAATAACAGTACTGGAATATGGGAGTTTGGCGGCCAAGCATCACAACCTGAATGC
 AGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATCAGCTTTTGTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATG
 TGAATCAGGCTGTGGACAGCTATGATCCCGATGTTGCTATTATTGAAAATCATGGCCCCGAGGATG

2001

>BA6800-01

TTGATAGAGGTCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCTAACACATCTCAGTGGAT
 AGTTACAGCGGGAGAAGACAATGCCACCACATACACATGTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAGAA
 ATTGGGAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGA
 CTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTGAAATGAAGCCATGCTACGGGTCTTTTGAAG
 ACCTAACACATAAAGGGGGCAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCAGATATTGATATGGA
 ATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGGAAACTCCAG
 ACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAAATTA
 CATTGGCTTCAGGGATAAATTTGTAGGCTTAATGTACTACACAGTACTGGAATATGGGAGTTTGGCGGCCAAGCATCACAACCTG
 AATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATCAGCTTTTGTGATTCTCTGGGAGACAGAAGCAGATACTTCA
 GCATGTGAATCAGGCTGTGGACAGCTATGATCCCGATATTGCTATTATTGAAAATCATTTGCATCGA

2002

>BA148-02

CCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGAAGGTTTAGAAATTTGGGAAAGACATTACTGCAGA
 CAACAAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAA
 AAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTACAAACATAAAGGGGGCC
 AAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGA
 AGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGGAAACTCCAGACAGCCATGTGGTATACAAGCCA
 GGAATCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAAATTAATTTGGCTTCAGGATAAATTTG
 TAGGTCTTATGTACTACAACAGTACTGGAATATGGGAGTTTGGCGGCCAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGA
 CAGAAACACTGAACTGTCTATCATGCTTTTGTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGACTGTGGAC
 AGTTATGATCCCGATGTTGCTATTATTGAAAAATCA

>BA1924-02

GGGGGGTGCTTGATAGAGGTTCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACACATC
 TCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAA
 GGTTTAGAAATTTGGGAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAG
 AATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTTAAATGAAGCCATGCTACGGGT
 TTTTGCAAGACCTACAAACATAAAGGGGGCAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCA
 GATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGG
 AAATCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAG
 ACCAAATTACATTGGTTTGGGATAAATTTGTAGGTTCTATGATCTACAACAGTACTGGAATATGGGAGTTTGGCGGCCAAGCATCACA
 TCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATCATGCTTTTGTGATTCTCTGGGAGACAGAAGCA
 GATACTTCAGCATGTGGAATCAGGCTGTGGACAGTTATGATCCCGAT

>BA2026-02

GGGGGGTGCTTGATAGAGGTTCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACACATC
 TCAGTGGATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAA
 GGTTTAGAAATTTGGGAAAGACATTACTGCAGACAACAAGCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAG
 AATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAAATGAAGCCATGCTACGGGT
 TTTTGCAAGACCTACAAACATAAAGGGGGCAAGCTAAAAACAGAAAAGTAAACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCA
 GATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTGTGCTTTACACGGAAAAATGTCAATTTGG
 AAATCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAG
 ACCAAATTACATTGGTTTGGGATAAATTTGTAGGTTCTATGATCTACAACAGTACTGGAATATGGGAGTTTGGCGGCCAAGCATCACA
 TCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATCATGCTTTTGTGATTCTCTGGGAGACAGAAGCA
 GATACTTCAGCATGTGGAATCAGGCTGTGGACAGTTATGATCCCGAT

GATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGG
AAATCCAGACAGCCATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAG
ACCAATATACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCA
TCACAACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCA
GATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTCTGTATGATTGAAATCA

>BA2169-02

GGGCGCGCTAACACATCTCAGTGGATAGTTACAACGGGAGAATACAATGCCACCACATACACATTTGCCATTGCTTCCACGAAGGGA
GACTATATTACTAAGGAAGGTTTGAAGATTTGGGAAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAG
AGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGGAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAAT
GAAGCCATGCTACGGGTCTTTTGAAGACCTACAACATAAAAGGGGGCCAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGAT
GTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACA
CGGAAATGTCAATTTGGAACTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCA
ACAAGCCATGCCTAACAGACCCCAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGA
GTTTTGGCCGGCCAAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATT
CTCTGGGAGACAGAAGCAGATCTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCG

>BA2268-02

GTGGTTGTAGAGGTCCAGCTTCAAGCCATATTCGGGACAGCTTACAATTCAGTGGCTCCTTAGGGCGCGCCTAACACATCTCAGT
GGATAGTTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTT
AGAAATTTGGGAAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCA
TGGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAATGAAGCCATGCTACGGGTCTTTTG
CAAGACCTACAACATAAAAGGGGGCCAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATAT
TGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAAT
CCAGACGCCATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCA
ATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCACA
ACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATAC
TTCAGCATGTGGAATCAGGCTGT

>BA2687-02

GGGTGCTTGTAGAGGTCTAGCTTCAAGCCATATTCGGGACAGCTTACAATTCAGTGGTTCTTAAGGGCGCGCCTACACATCTCA
GTGGATAGGTACAACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGT
TTAGAAATTTGGGAAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAAT
TGCGACTGATATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTTAAATGAAGCCATGCTACGGGTCTTTT
TCAGAGACCTCAAGACCTAAAAAGGGGGCCAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGAT
ATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAA
CTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACC
CAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCA
CACTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATAC
ACTTCAGCATGTGGAATCAGGCTGTGGACAGTTATGATCCCGATGTT

>BA2688-02

ATTTGGCTTTCCCTCCATGAAGGGAGACAATATTACTAAGGAAGGTTTGAAGATTTGGGAAAGACATTACTGCTGACTACAAGCCCAT
TATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAAAAGTTTGGAGGTA
GAGCTCTTAAACCAGCTACTAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTACAACATAAAAGGGGGCCAGCTAAAAACAG
AAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCT
TTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAATCCAGACAGCCATGTGGTATACAAGCCAGGAACCTTCTGATG
GTAATCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTA
CTACAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAA
CTGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGAAGAGTTATGATCCCG
ATGTTCTGATTATTGAAATCATAGCGTCCGTGATGAAGTGC

>BA2966-02

AGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAAAAAT
TGGAGGTAGAGCTCTTAAACCAGCTACTAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTACAACATAAAAGGGGGCCAGCT
AAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTG
CTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAATCCAGACAGCCATGTGGTATACAAGCCAGGAAC
TTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAATTACATTGGCTTCAGGGATAACTTTGTAGGT
CTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAA
ACACTGAATGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTA
TGATCCCGATGTTCTGATTATTGAAATCATAGCGTCCGTGATGAAGTGCCT

>BA3029-02

GAGGTCTAGCTTCAAGCCATATTCGGGACAGCTTACAATTCAGTGGCTCCTAAGGGCGCGCCTACACATCTCAGTGGATAGTTAC
AACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTGAAGATTTGGG
AAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATA
TTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTAC
AAACATAAAAGGGGGCCAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGA
TTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAATCCAGACAGCC
ATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAATTACATTGG
CTTCAGGATAAATTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCACAACCTGAATGCA
GTGGTTGACTTGCAGGACAGAAACACTGAACTGTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATAGTTTCAGCATGT
G

>BA3037-02

GAGGTCTAGCTTCAAGCAATATTCGGGACAGCTTACAATTCAGTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGGATAGTTAC
AACGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTGAAGATTTGGG
AAAGACATTACTGCAGACAACAGCCCATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATA
TTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACCAGCTACTAAATGAAGCCATGCTACGGGTCTTTTGAAGACCTAC
AACATAAAAGGGGGCCAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGA
TTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTCTTTACACGGGAAATGTCAATTTGGAAATCCAGACAGCC
ATGTGGTATACAAGCCAGGAACCTTCTGATGGTAACTCTCATGCAAATTTGGGTCAACAAGCCATGCCTAACAGACCCAATTACATTGG
CTTCAGGATAAATTTGTAGGTCTTATGTAATAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAAGCATCACAACCTGAATGCA

GTGGTTGACTTGCAGGACAGAACTGAACTGTCTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGT
GGAATCAGACTGTGGACAGTTATGATCCCGATG
>BA3270-02
TCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAATATGGGAGTTTGGCCGGCCAAGCATCACAACCTGAATGCAGT
GGTTGACTTGCAGGACAGAACTGAACTGTCTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAA
>BA3282-02
TATAGGTCCTAGCTTCAAGCCATATTCGGGCACAGCTTACAATTCACTGGCTCCTAAGGGCGCGCCTAACACATCTCAGTGGATAGTT
ACACCGGGAGAAGACAATGCCACCACATACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAGAAATTG
GGAAAGACATTACTGCAGACAACAAGCCCATTTATGCCGATTAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGA
TATTGATGGAACAAATGAAAAATTTGGAGGTAGAGCTCTTAAACAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTGCAAGACCT
ACAAACATAAAAGGGGGCCAAGCTAAAAACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGG
AATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTGCTTTACACGGAAAATGTCAATTTGGAACCTCCAGACAG
CCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAATTTGGGTCAACAAGCCATGCTTAACAGACCCAATTACATT
GGCTTCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAAATATGGGAGTTTGGCCGGCCAAGCATCACAACCTGAATG
CAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCAT
GTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTTCG
>BA3387-02
GGCAGAGCTTACAATTCACGGGTTCCTAAGGGCGCGCCCAATACATCTCAGTGGATAGTTACAACGGGAGAAGACAAGCCACCACAA
ACACATTTGGCATTGCTTCCACGAAGGGAGACAATATTACTAAGGAAGGTTTAGAAATTTGGGAAAGACATTACGGCAGACAACAAGCC
AATTTATGCCGATAAAACATATCAGCCAGAGCCTCAAGTTGGAGAAGAATCATGGACTGATATTGATGGAACAAATGAAAAATTTGGA
GGTAGAGCTCTTAAACAGCTACTAAAATGAAGCCATGCTACGGGTCTTTTGCAAGACCTACAAACATAAAAGGGGGCCAAGCTAAAA
ACAGAAAAGTAACACCAACCGAAGGAGATGTTGAAGCTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGA
CGCTTTTTCGCCTGAAATTTGTGCTTTACACGGAAAATGTCAATTTGGAACCTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCT
GATGGTAACCTCTCACTGCAAAATTTGGGTCAACAAGCCATGCTTAAACAGACCCAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTA
TGTAACAACAGTACTGGAATATGGGAGTTTGGCCGGCCAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACAC
TGAACCTGTCTATATCAGCTTTTGTCTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGAT
CCCGATGTTTCGCAATTATTGAAAAATCA
>BA3478-02
AGGAGATGTTGAGACTGAGGAGCCAGATATTGATATGGAATTTTTCGATGGTAGAGAAGCTGCTGACGCTTTTTCGCCTGAAATTTGTG
CTTTACACGGAAAATGTCAATTTGGAACCTCCAGACAGCCATGTGGTATACAAGCCAGGAACCTCTGATGGTAACTCTCATGCAAAAT
TGGGTCAACAAGCCATGCCTAACAGACCCAATTACATTGGCTTCAGGGATAACTTTGTAGGTCTTATGTACTACAACAGTACTGGAAA
TATGGGAGATTTTGGCCGGCCAAGCATCACAACCTGAATGCAGTGGTTGACTTGCAGGACAGAAACACTGAACTGTCTATATCAGCTTTTG
CTTGATTCTCTGGGAGACAGAAGCAGATACTTCAGCATGTGGAATCAGGCTGTGGACAGCTATGATCCCGATGTTTCGTATTATTGA

Primers diseñados

1. ttggctgtggggigaat (av1)
2. ttgagtccagcgagaagagtttt (av2)
3. acagatcgtgcagcgtagg (av3)
4. agctgaaatggacttgcgttgta (av4)
5. agatttattttacttgcgatt (av5r)
6. tgggatggaccttggaccttag (av6)
7. attacttctgggtgggtcttgg (av7r)
8. ggaatccgcaaaactcact (av8)
9. ggatactgagacacccaccgacca (av9r)
10. gggatgtggccaggggttgcg (av10)
11. gggatgtggccaggggttgcg (av11r)
12. ggggaatgtttagcctaccagt (av12)
13. atagggtgtcgcaggggaatgttt (av13r)
14. gccgcgcgcgtactgct (av14)
15. gctactcgtccgccaacaccatc (av15r)
16. ctgtaaagggtgggattgaa (av16)
17. aaatctgtgagctgggacgggtg (av17r)
18. cctgtggcccgaggagcatagttc (av18)
19. tggtagttaggagcgacagc (av19r)
20. aggcgctctccccaagttagatg (av20)
21. gcagacctctcgttccgagg (av21r)
22. ccgcaccgcaggagacgcagac (av22)
23. caaagaggctgtcgtgtcccg (av23r)
24. cgaggaggatttgatattgacagt (av24)
25. tgggtctttcctgtccgcgc (av25r)
26. attctgactcgccatccctattg (av26)
27. tcatctgccattctgaactgcca (av27r)
28. gcttgcctgcacatagta (av28)
29. agcatccctctccactgg (av29r)
30. attgcgcagcgaggtatgagtga (av30)
31. gtgagagcatagtgtcagggcc (av31r)
32. tgaccagcatgaaggggattag (av32)
33. ctgtgcaggatgagagccgat (av33r)
34. tacctcgcgtctactatg (av34)
35. ccagacctcgcgcggcag (av35r)
36. gacgacgcggggtgacatcc (av36)
37. agacgcggtgttagacca (av37)
38. gaccacagccccacgg (av38r)
39. aagttcacgcgaagttg (av39r)
40. gcacttctctgcccgcctg (av40r)

41. ccatccccaagcatcatcctg (av41)
 42. ccgcgcattggcctggacaa (av42)
 43. caagaccccgccagccgact (av43r)
 44. ttttgcgccagatgcatccagt (av44r)
 45. ctaggcggagcgatggtgtat (av45)
 46. caaggcggagggtggtgtc (av46)
 47. tgacctgatgaacaatgcca (av47r)
 48. gtgggacccgaaactaa (av48)
 49. gaagacaaaatcaacacggcctac (av49r)
 50. atcctcatccgcccgcg (av50r)
 51. gttacaaggccctggg (av51r)
 52. caccgcgatgaattggcttag (av52)
 53. gcaatcagcgagtcagtggg (av53)
 54. agcaccatccactactaaccttcc (av54)
 55. caatactaccactccaaaaccgg (av55r)
 56. taataataaaacatggaacaatac (av56)
 57. ttattcgccctagacaaagcaa (av57r)
 58. tgcaacaaaactaaaccaata (av58)
 59. ggttactattatggttatgacaga (av59r)
 60. gcaaggcttctattatgga (av60)
 61. ccacttctgctccccgc (av61r)
 62. attcttattttctattctgccgt (av62r)
 63. ttgcggcacttcttaggcttatt (av63)
 64. acaaccttctactgctcctgccaat (av64r)
 65. cgactcgcccaactacgcatcc (av65)
 66. atttctcccagcagcacctcact (av66)
 67. gaagatgaaagcacctcccaacac (av67r)
 68. taaaactggagccctgctactgc (av68r)
 69. ggaatacaggagttgcccagaag (av69r)
 70. ccgaatggcagggaagt (av70)
 71. tgattgcccgcgaagaagtaccc (av71)
 72. tgggcggtcagaggttt (av72r)
 73. ttacgctcgaaaatctacaag (av73)
 74. caagatcatcgctcagggtcaaga (av74r)
 75. tggaagactgatggcacaacc (av75)
 76. ccggcctcatctggacgaatagc (av76r)
 77. cggaagccggtgtctcagc (av77)
 78. gtccggtcctgattacttcttc (av78)
 79. ggcatgcactccccacaaac (av79r)
 80. ctactgcggaggaggagag (av80)
 81. gatcctctctctccgcagta (av81r)
 82. atgcccgcacctaatgac (av82)
 83. attctatcgccgcttcttttc (av83r)
 84. ccccatgatcccaggctcgtctct (av84r)
 85. tcgtacgaaggaggaggac (av85r)
 86. cgccggcgcccaccatta (av86)
 87. atctcgcccagtaataaac (av87)
 88. cacaaactccgcccagccatcat (av88r)
 89. gccgcctgctcaagaata (av89r)
 90. atggcaagcgtagtagaataac (av90)
 91. tgcggaacctttgacactgc (av91)
 92. gcaaggaggcagcggaacatt (av92r)
 93. aagcaattgcgggtggtttt (av93)
 94. gtgcggagccaaggatgaaaaa (av94r)
 95. accaggctgtgcagaaaaagataa (av95)
 96. ccagtgcgctagccaggta (av96)
 97. cctggctacggcgactggttc (av97r)
 98. gctgctgtgacggcgtgtaatgac (av98r)
 99. taggagccaggagttgtaagc (av99r)
 100. gcggcatcatcgaggcagtcata (av100)
 101. ttccagcgtlaatacagcggtcctt (av101r)
 102. gaccgctgtattacgtggaaaaa (av102)
 103. ttccgaaltcaagtggcaacc (av103r)
 104. gaatgacccgcaggcagtgatacg (av104)
 105. gcacagccggcatcattcacac (av105)
 106. caatgacgcccgtgtgtgt (av106r)
 107. tatccagcctgtacttctccat (av107)
 108. ccgattagcttgccttggtgtct (av108r)
 109. gggggtatcgaggaggaggact (av109)
 110. gaccgcccacaggagcacacc (av110r)
 111. gcccaaaacacagagcagata (av111)
 112. agtcgggtgtctgatccctatttc (av112r)
 113. cgattagcggcgaggagact (av113r)
 114. cgccgcatgccctactcaat (av114)
 115. ggaagcccgcgtgattac (av115)
 116. tggcacaggtagacatagacacaa (av116r)

117. agcagggtggcgggcaggatc (av117r)
 118. ctgccctccgactttgtcctctc (av118)
 119. gaggaggcagaggagggtggaagaa (av119)
 120. tgtttgcggcggttacttctcc (av120r)
 121. cagcaaatccgggcagtcctcg (av121)
 122. gccgcgctgtctttatctgtcg (av122r)
 123. aaaaaggcgggaattacatcat (av123)
 124. ggcggggcggtggtgtaag (av124r)
 125. acatcaatcgtagccgtccac (av125r)
 126. aggacaaaacatggctactggtg (av126)
 127. gtagttatttgcctcccccttcc (av127)
 128. laaaagcgctccagccaaaactca (av128)
 129. ttccggaagggtatattgggtgc (av129r)
 130. aagctcgggtttcatttcctcac (av130)
 131. tgcgccagacatcaccttac (av131r)
 132. tccatcgcatctaactgttctgag (av132)
 133. tcttcgctcttccctctcctct (av133r)
 134. aaaatatctgtcctctgtgtca (av134)
 135. gcgtcccctgctctgttcc (av135r)
 136. acaagccacagggctctcc (av136)
 137. accctgtggctgtctgtgatg (av137r)
 138. acaaagtaaaaggcacaggagaat (av138)
 139. ggaagttacgcctgttgggattg (av139r)
 140. tcgtgtggtgattggctgtg (av140)
 141. atggtgccaatatgaaatgaggt (av141r)
 142. tcaagcgggttaaaaggacaaaa (av88b)
 143. ccttatagatggaatgggtgcc (itr1)
 144. ggcaccattccatctataagg (itr1r)
 145. cggggggtgcttgatagaggtccta (hx4)
 146. ccagctacccaaatgaagcc (hx5r)
 147. gatgttgagactgaggagccag (hx6r)
 148. ggggaaccttaccagccaattatcc (hx7)
 149. cgtggtcagagtcaccagccaca (hx9r)
 150. gttgcttcaagatggccaccc (hxs1)
 151. agggcgcgccaatacat (hxs2r)
 152. gatgttgagactgaggagc (hxs3)
 153. catcgaggatgaactgcct (hxs4)
 154. ggaccagcgaaaacatatc (hxs5r)
 155. gctaccggtccatgcttctg (hxs6r)

Fe de erratas

Página	Línea	Dice	Debe decir
5	8	infecta	infectan
16	9	protein	proteína
17	2	descripto	descripto
	10	VA RNA genes	genes VA RNA
18	14	intemedias	intermedias
	15	to	a
23	1	response	-
	6	estabilizand	estabilizando
25	5	apoptotica	apoptótica
	8	.	-
26	12	reconida	reconocida
31	12	monomeros	monómeros
35	13	(13)	Lí y Waddell, 1999
37	7	interferon	interferón/interferones
	14	require	requiere
39	15	protína	proteína
45	10	ha	han
62	20	gráfico 1	-
73	5	portenciales	potenciales
75		ref Tabla 10	va en página 76
80	8	Figura	-
97	5 y 8	Bg /II	Bgl II
114	2	-	Altschul et al., 1997.Nucleic Ac Res 25:3389-3402
117	4		Smith et al., 1997. J Gen Virol 78:1511-1519

FCE y N BIBLIOTECA

**Secuencia nucleotídica completa del Adenovirus 7h: análisis de los
polimorfismos del gen del hexón de los Adenovirus circulantes en
Buenos Aires entre los años 1990-2002.**

Tesista: Paola Roxana Barrero

Licenciada en Ciencias Biológicas

Director: Dra. Alicia Susana Mistchenko

Laboratorio de Virología, Hospital de Niños Dr. R. Gutiérrez

Cuadernillo complementario

7142
V.2

Contenido

- **Discusión de los objetivos planteados y alcance de resultados obtenidos**
- **Figuras con mayor definición**
- **Leyendas para figuras**

Discusión de los objetivos planteados y alcance de resultados obtenidos

Discusión de los objetivos planteados y alcance de resultados obtenidos

Esta Tesis es el producto de cuatro años de investigación en la diversidad genética de los Adenovirus.

El trabajo fue abordado en dos etapas, en la primera se estudiaron las *variaciones del gen del hexón*. Esto ha permitido localizar marcadores para la diferenciación filogenética de los Adenovirus asociados a evolución fatal. La *aplicación de estos marcadores en muestras de pacientes con infección respiratoria* permitirá evaluar la aplicabilidad clínica de los mismos y ahondar en su importancia en la patogenicidad.

En la segunda etapa el trabajo estuvo dirigido a la secuenciación completa del genoma del Adenovirus 7h, cuyo tamaño y complejidad es notable. A nivel internacional es uno de los primeros genomas de Adenovirus humanos reportados, y el primero en nuestro país.

Teniendo en cuenta el impacto de la infección por Adenovirus, el trabajo tendrá numerosas aplicaciones futuras.

A continuación se discuten los alcances de los objetivos propuestos.

Polimorfismos Hexón AV7h

¿Por qué utilizar herramientas moleculares en el hospital público?

Para enriquecer los métodos diagnósticos con herramientas de última generación, agregando información más detallada al dato obtenido en un diagnóstico rápido por inmunofluorescencia. Como hemos observado los anticuerpos monoclonales que se usan para diagnóstico no detectan en un 100% el amplio rango de genotipos que circulan en una población pediátrica.

¿Por qué realizar estudios de epidemiología molecular en AV?

Porque hay restricción temporal y geográfica en la circulación de otros virus respiratorios y es importante conocer la incidencia de patógenos, más aún cuando no cuentan con tratamiento y pueden generar enfermedades crónicas como enfermedad pulmonar obstructiva crónica (EPOC) en fumadores y asma y enfermedad pulmonar crónica postviral en niños. En nuestro estudio predominó de la subespecie B1 y el AV7 fue el de mayor prevalencia, dato importante en pediatría dada su asociación con enfermedad severa o fatal.

¿Por qué analizar la filogenia de los AV7h?

Si bien el análisis de las regiones hipervariables en el período 1990-2002 posicionó a las muestras analizadas en el *cluster* GTC2 junto al AV7h, el análisis filogenético del hexón completo permitió agrupar las muestras fatales y las separó de las graves y leves. Esto nos permitió identificar probables marcadores moleculares fuera de HVRs que son las regiones usadas para genotipificar los AV. La presencia de estos marcadores moleculares, fáciles de detectar puede sumar un valor predictivo y pronóstico en una patología endémica de alto impacto en pediatría. Estos datos constituyen un primer paso

hacia la posible elucidación de las relaciones entre la evolución clínica y el genotipo.

Genoma AV7h

¿Por qué secuenciar el genoma AV7h por solapado de productos de PCR?

Porque la estrategia de biblioteca genómica y *shotgun* o caminata cromosómica no se podía completar debido a que era necesario terminar de analizar una secuencia y a partir de ella genera *primers* específicos para completar el clon de tamaño molecular grande y por falta de templados que funcionaran como moldes alternativos de *primers* en bases de datos públicas.

La secuencia de AV11p está disponible en bases de datos públicas desde octubre 2002. A partir de allí la estrategia de solapado de amplicones resultó una metodología aplicable sin inconvenientes en el Laboratorio.

¿Por qué secuenciar el genoma AV7h?

Porque

- AV7h es reconocido por su asociación con enfermedad severa y su distribución geográfica se está ampliando.
- En nuestro país se detecta en niños y está asociado a enfermedad de evolución severa y/o fatal.
- La infección persistente por AV se relaciona con enfermedad pulmonar obstructiva crónica (EPOC) en adultos fumadores.
- En niños, la infección persistente por AV se asoció con asma y con enfermedad pulmonar crónica postviral.

- Se conocen sólo 4 genomas de AV humanos, y se puede encarar desde el conocimiento estrategias terapéuticas tanto farmacológicas como preventivas.

¿Por qué buscamos entre los mimotopes péptidos E1a, pentón, fibra y hexón?

Los componentes de la cápside son los más abundantes y se encontró DNA E1A en células epiteliales de pacientes adultos fumadores con enfermedad pulmonar obstructiva crónica. Si bien no encontramos mimotopes en estos péptidos, su búsqueda exhaustiva resultaba necesaria.

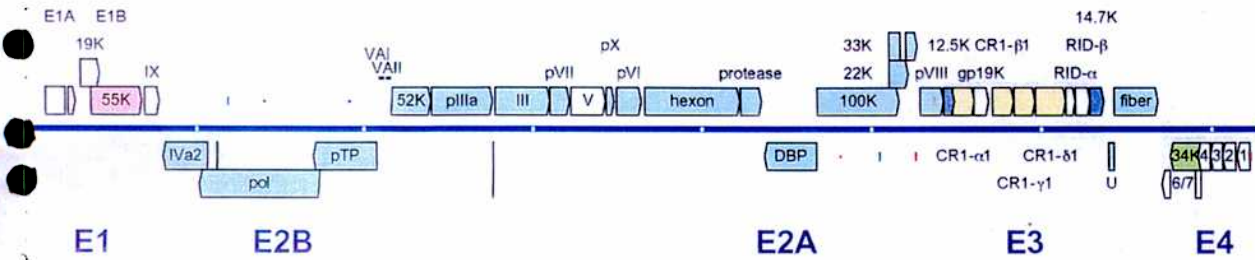
¿Por qué postular como mimotopes ORFs presentes en el genoma viral?

Porque

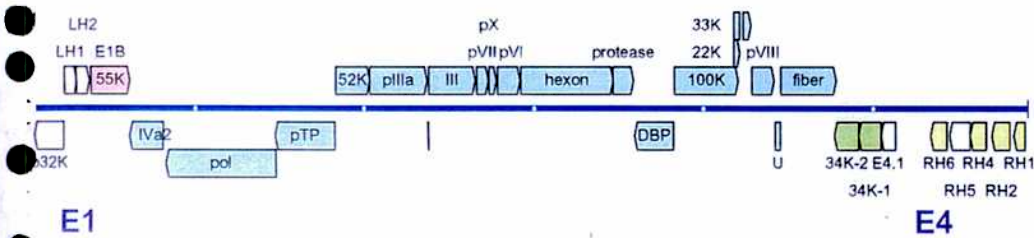
- Existe expresión preferencial de péptidos virales por *shutoff* celular
- Así habría expresión preferencial péptidos miméticos que no se encuentran dentro de ninguna proteína viral ni forman parte del virión
- Habría inducción de autoanticuerpos anti-elastina y miosina, superando la tolerancia al ser expresados como proteínas extrañas
- En la patogénesis de AV la expresión de las proteínas de músculo liso y matriz extracelular están aumentadas en pulmón
- Existe una hiperreactividad generalizada, dado un estado de alerta inmune.
- El daño al pulmón es crónico, disminuyendo su funcionalidad progresivamente de manera independiente de la presencia actual de infección viral.

Figuras con mayor definición

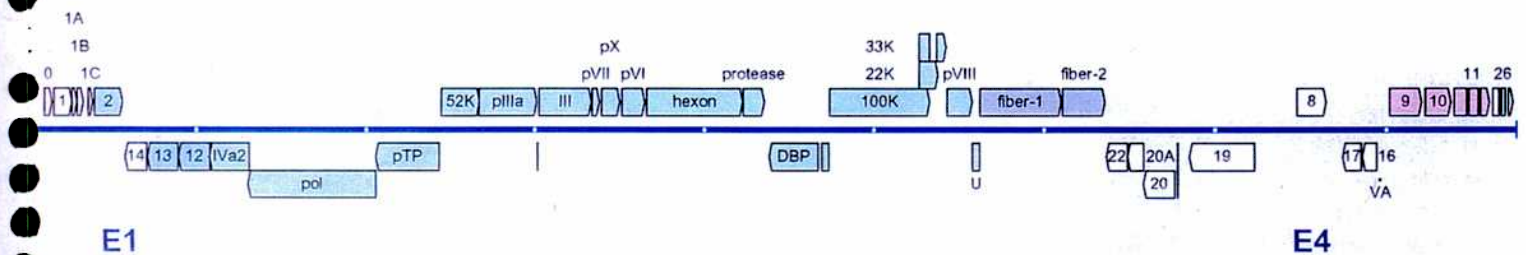
Mastadenovirus: SAdV-25 (HAdV-E)



Atadenovirus: OAdV-7 (OAdV-D)



Aviadenovirus: FAdV-1 (FAdV-A)



Siadenovirus: TAdV-3 (TAdV-A)

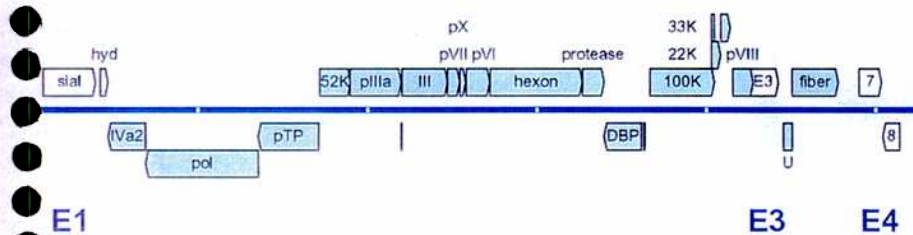


FIGURE 2

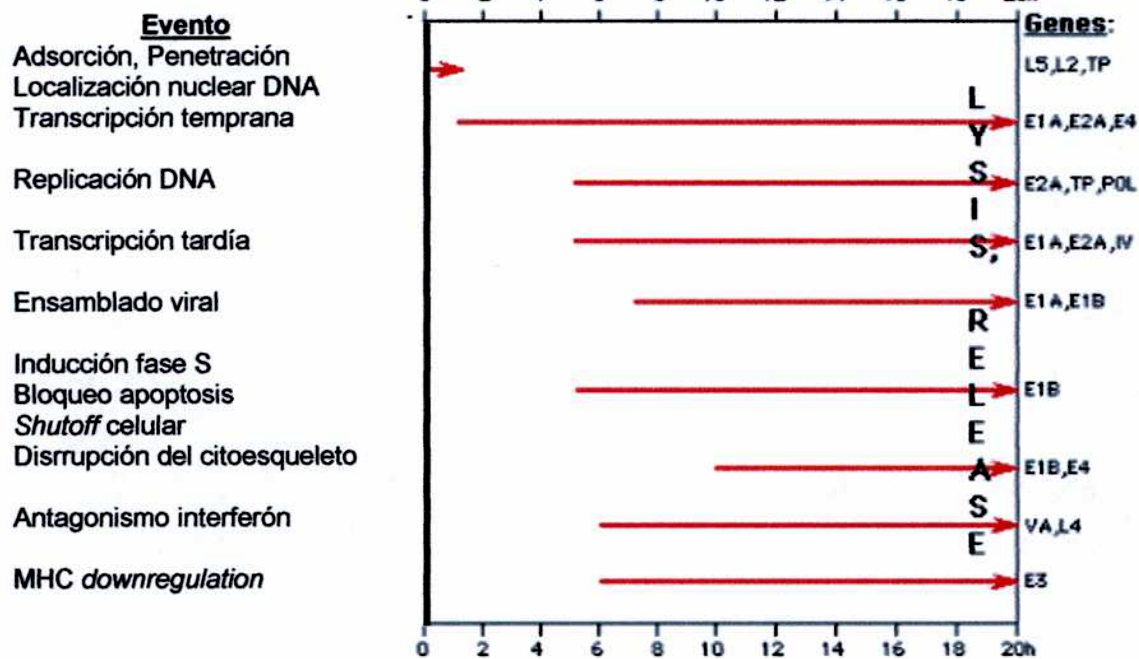


FIGURA 4

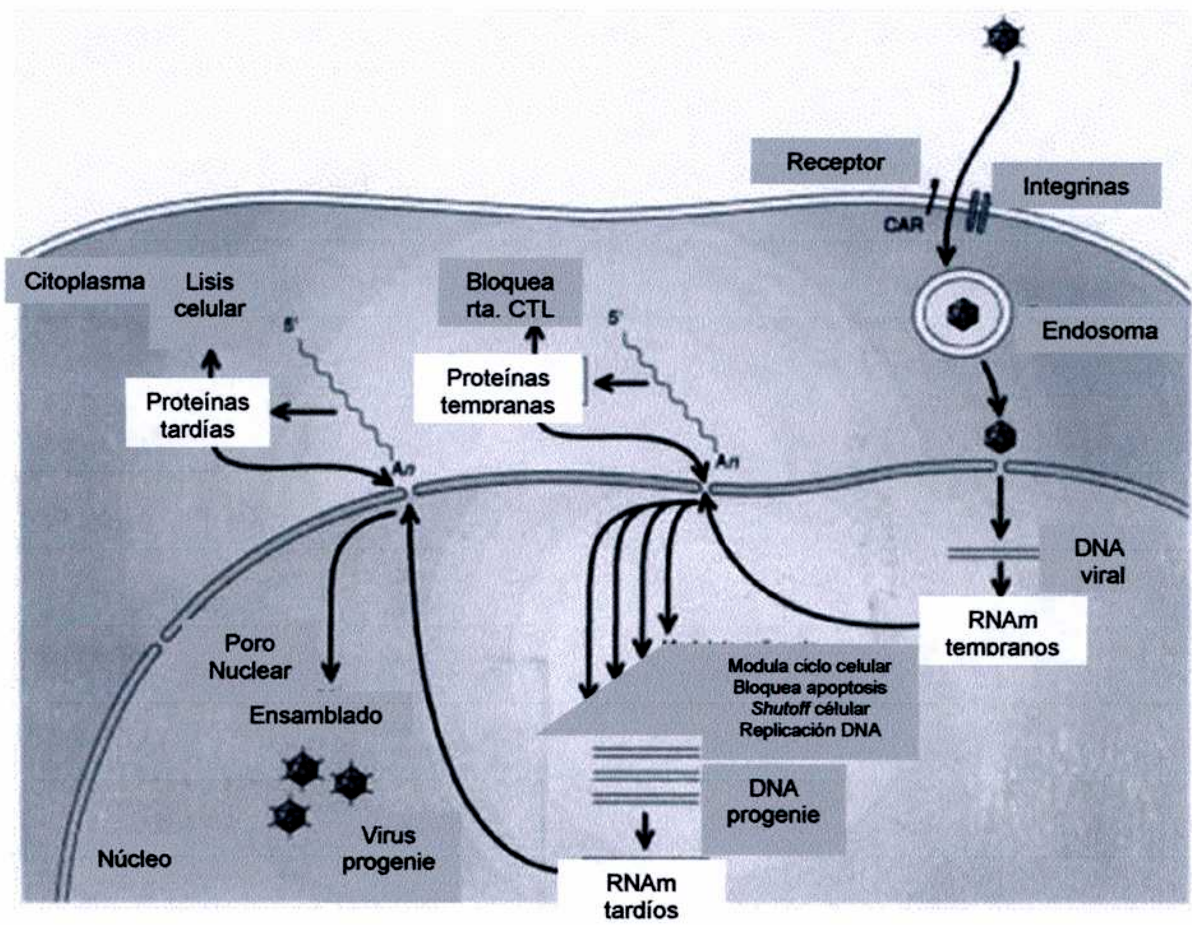


FIGURA 5

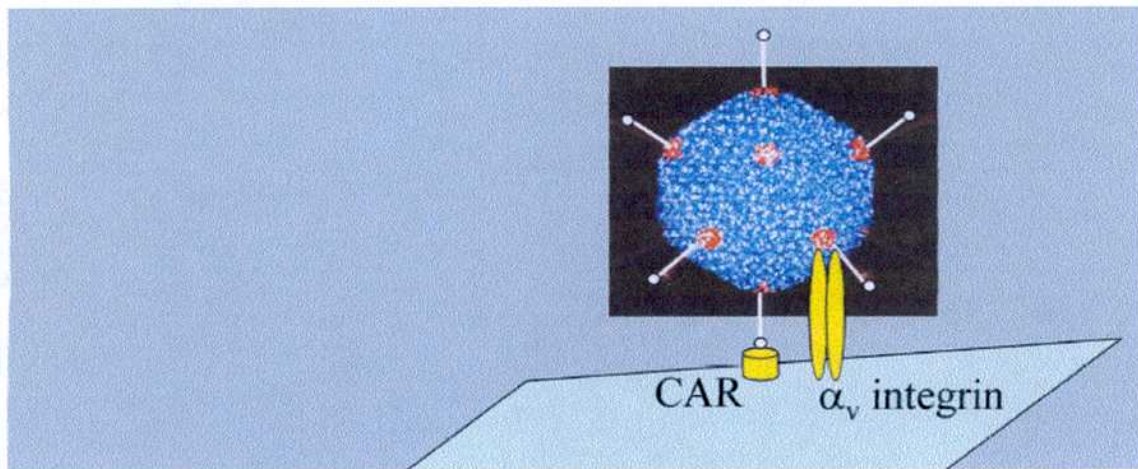


Figura 6

Cara apical

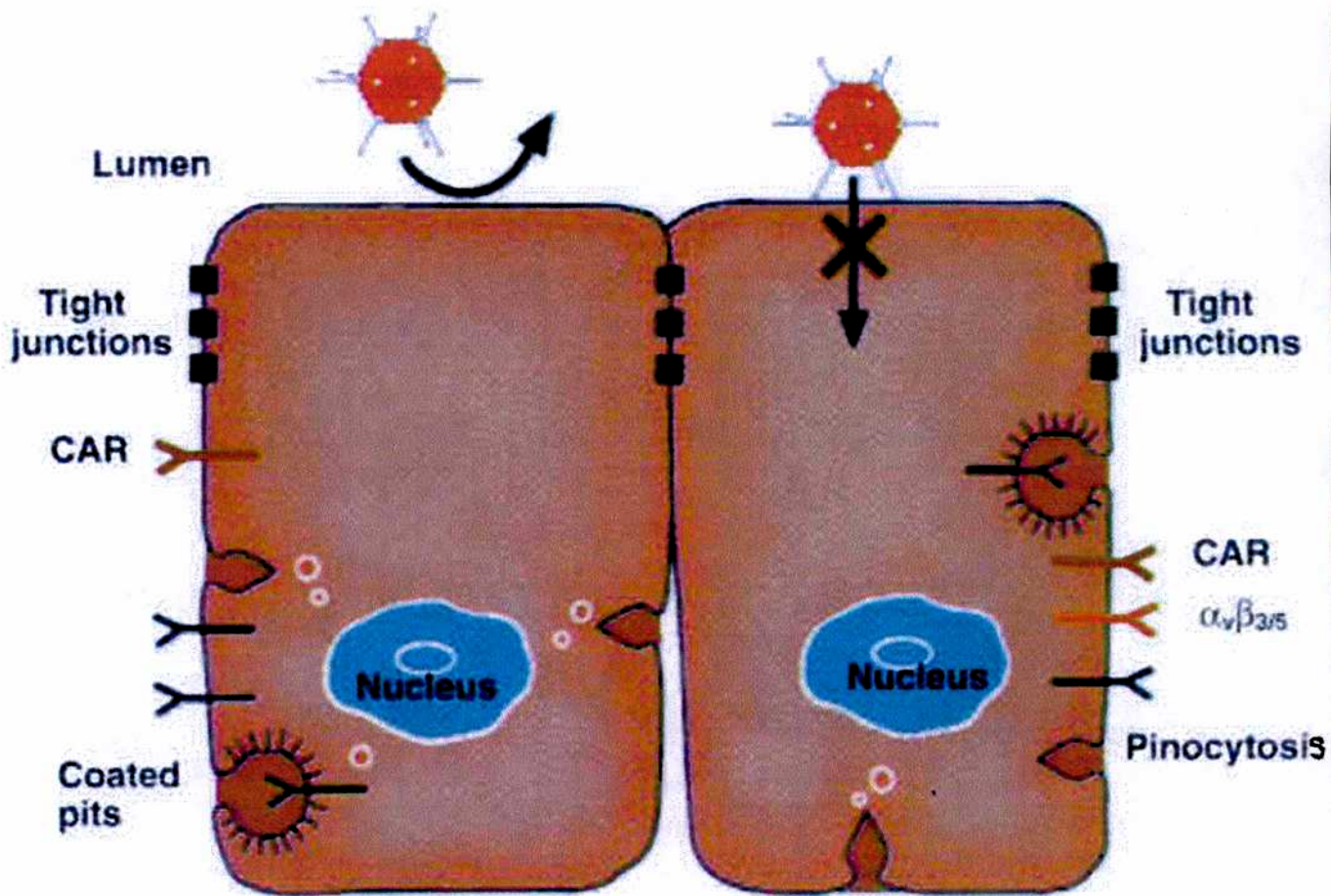


Figura 7

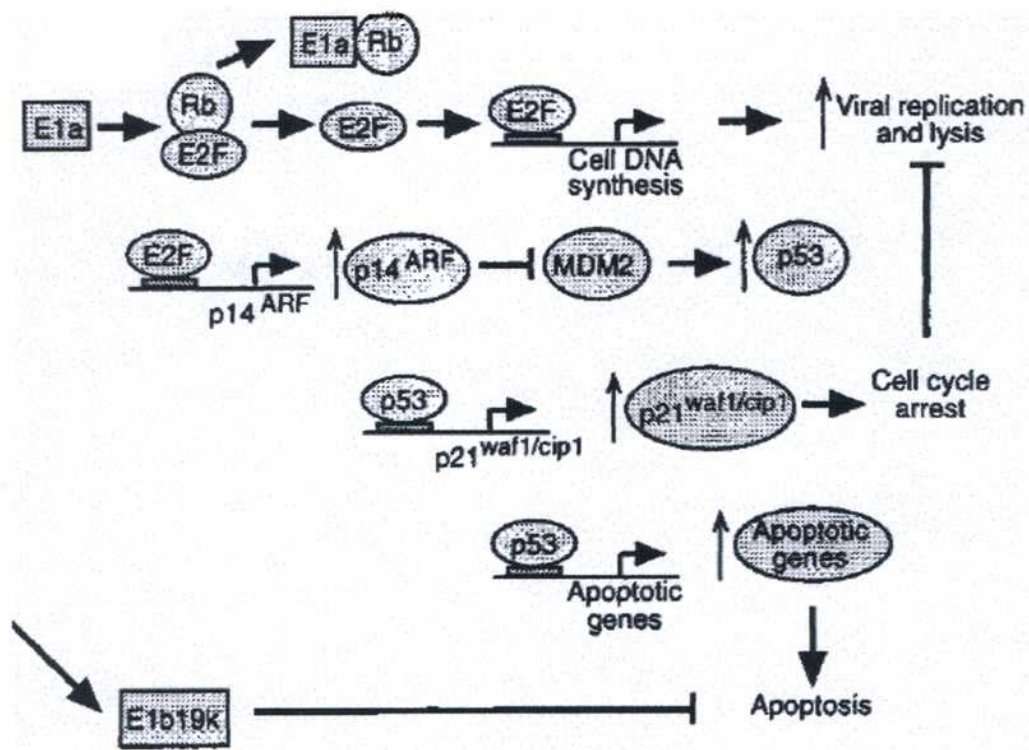


Figura 14

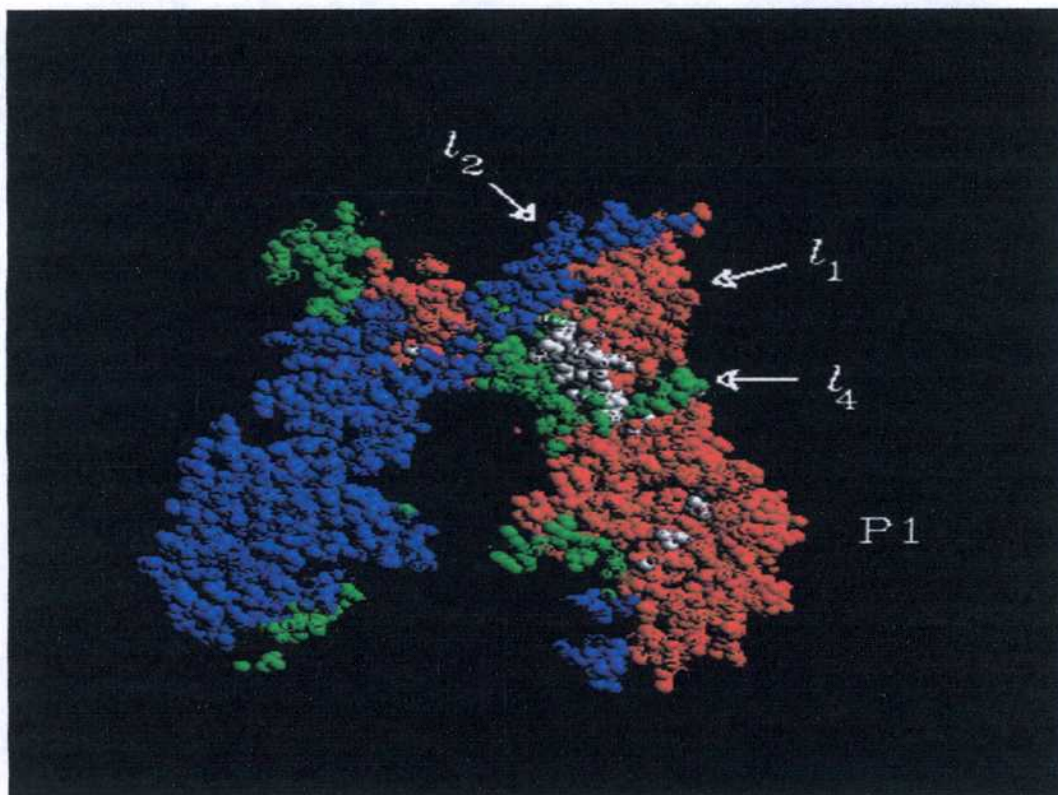


Figura 19

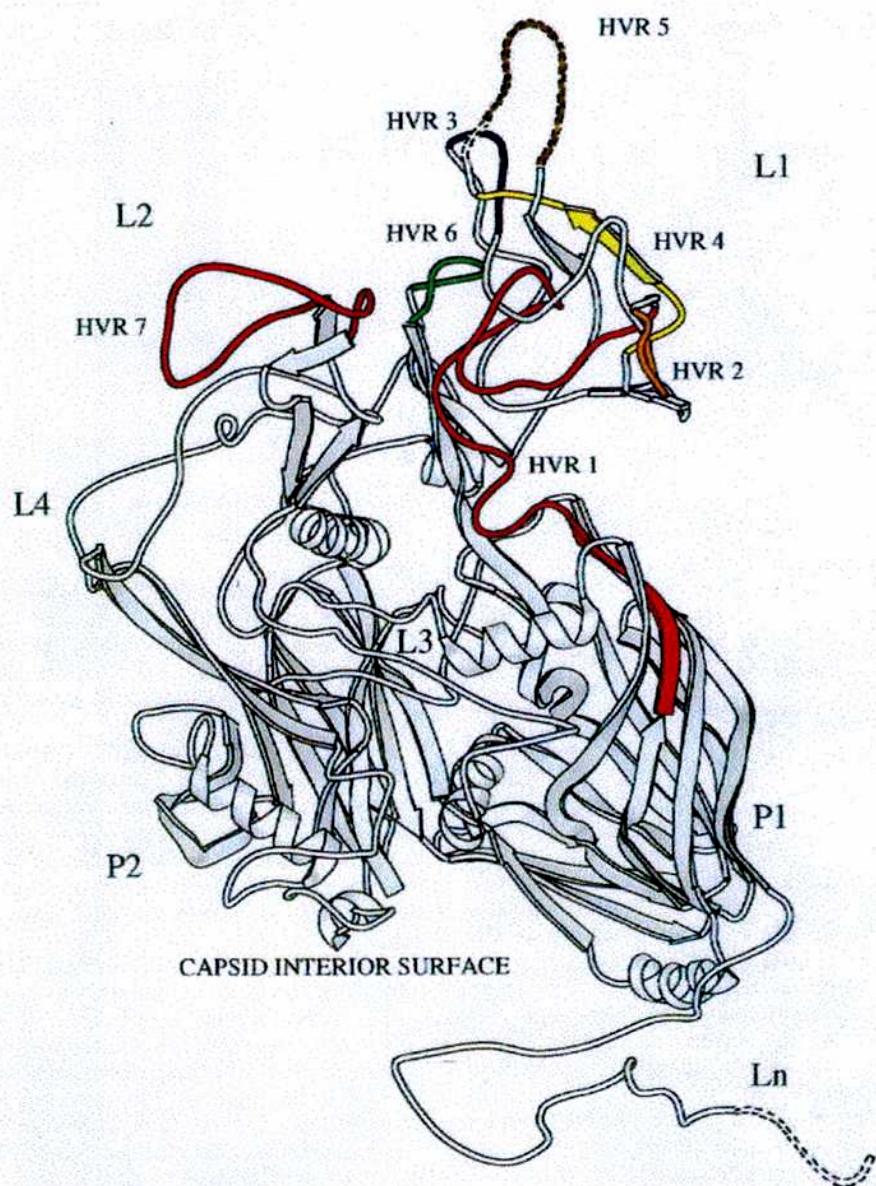


FIGURA 20 b.

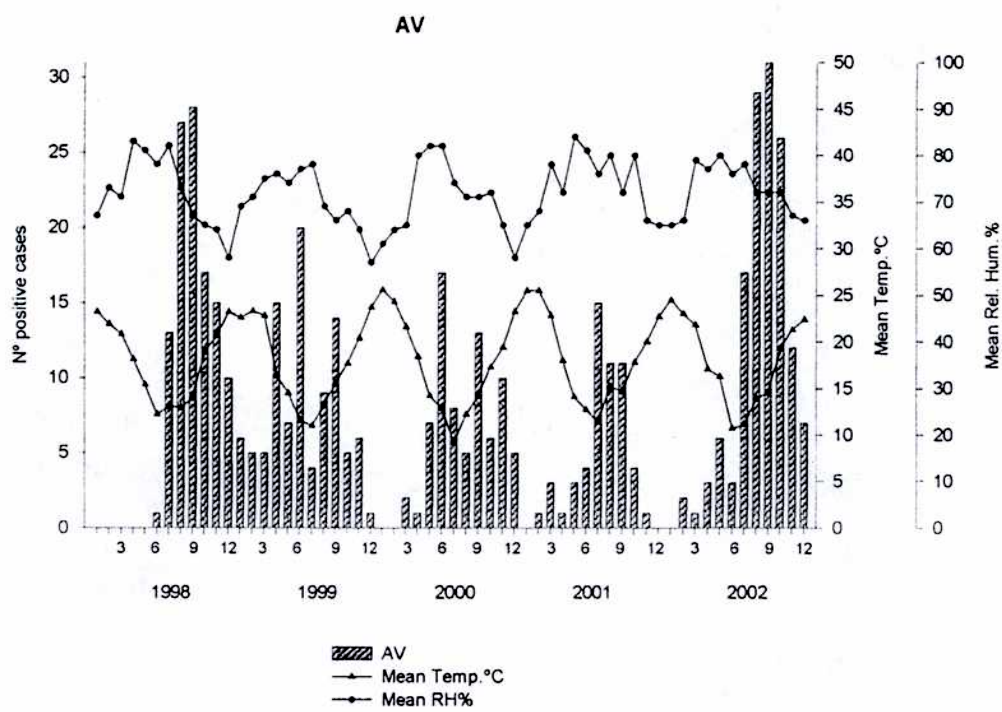


FIGURA 25

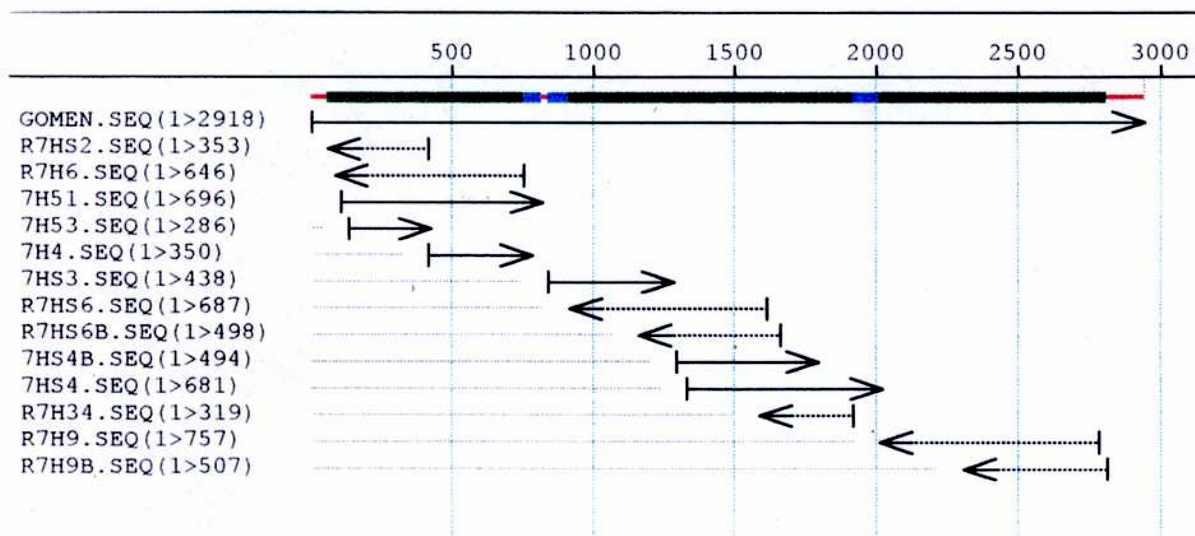


FIGURA 27

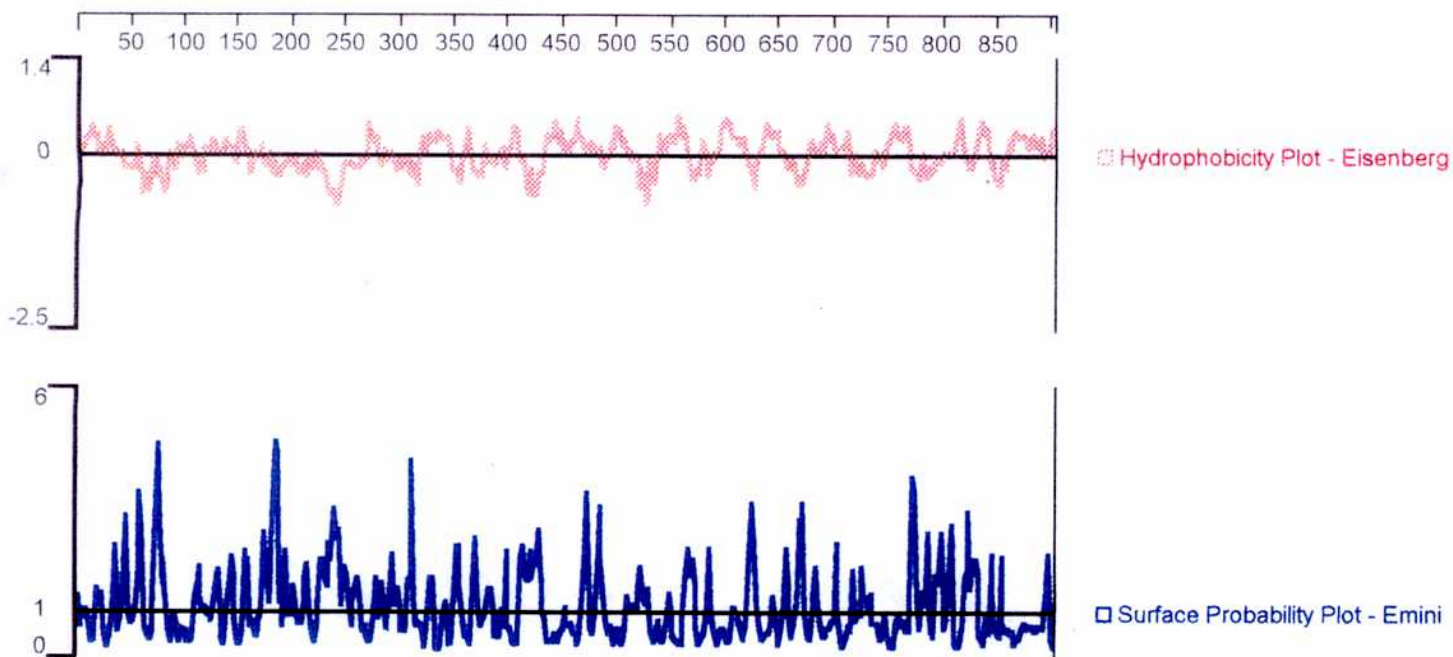


FIGURA 30 a -

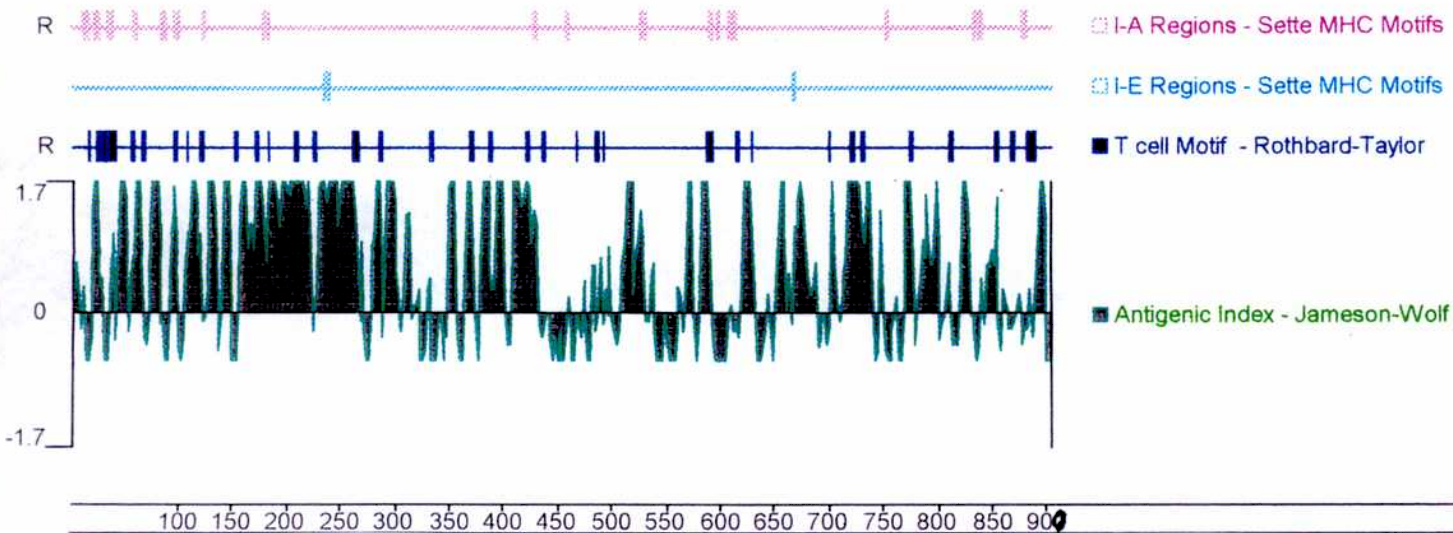


FIGURA 33

Mastadenovirus: SAdV-25 (HAdV-E)

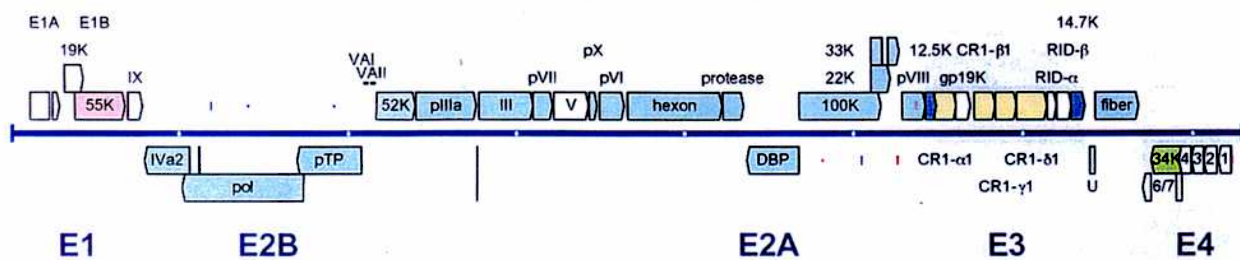


FIGURA 43

Leyendas

Leyendas para las Figuras

Figura 1: Filogenia de la proteína hexón de la familia *Adenoviridae*. Árbol de distancia con raíz en el esturión. Los géneros se marcan en diferentes colores y se destacan en verde los virus que pertenecen a la misma especie. Abreviaturas B: bovino, C: canino, D: pato, E: equino, F: aviar, Fr: rana, H: humano, M: murino, O: ovino, P: porcino, Po: zarigüeyas, Sn: serpiente, T: pavo, TS: musaraña. Tomado de Davison *et al.*, 2003.

Figura 2: Organización genómica de los diferentes géneros de la familia *Adenoviridae*. Los genes comunes se marcaron en celeste, los específicos en otro color. Abreviaturas F: aviar, H: humano, O: ovino, T: pavo. Modificado de Davison *et al.*, 2003.

Figura 3: Estructura del virión de Adenovirus. Se señala la localización de los componentes de la cápside del core y cemento. Modificado de Russell, 2000.

Figura 4: Estructura del virión de Adenovirus. Microscopía electrónica (x 500,000) mostrando la simetría icosaédrica y las fibras que protruyen de la cápside. Tomado de Valentine & Pereira, 2003.

Figura 5: Ciclo replicativo de Adenovirus (ver texto). Modificado de Shenk, 2001.

Figura 6: Entrada a la célula de los Adenovirus. Ilustración esquemática de la interacción de AV2 con los receptores celulares involucrados en la infección. La fibra (blanco) se une con alta afinidad con el receptor conocido como CAR (*coksackie-adenovirus receptor*) y luego se promueve la internalización por la unión de la base del pentón (rojo) con integrinas α_v . Modificado de Nemerow & Stewart, 1999.

Figura 7: Ubicación de los receptores para Adenovirus los receptores para Adenovirus en las caras basolaterales de las células epiteliales. Se muestra que el virus entra a la célula vía endocitosis mediada por receptor. Los complejos AV–receptor CAR (*coksackie-adenovirus receptor*) difunden en vesículas recubiertas (*coated pits*). Se muestra también la localización de las integrinas $\alpha_v\beta_{3/5}$. Modificado de Zhang et al., 2002.

Figura 8: Penetración a la célula y *uncoating* de Adenovirus. Las membranas de las vacuolas se rompen por la actividad tóxica del pentón liberándose así la partícula en el citoplasma. El *uncoating* sigue una secuencia ordenada: los pentones, hexones. El core migra al núcleo. Modificado de Cann, 1997.

Figura 9: Genoma de Mastadenovirus. Los genes comunes a todos los géneros se marcaron en celeste, los específicos en otro color. Las regiones codificantes se muestran como flechas. Modificado de Davison et al., 2003.

Figura 10: Mapa de transcripción de Adenovirus. Se muestran los genes tempranos (E=Early) y las cinco familias de genes tardíos (L=Late). Cada gen transcrito por la RNA polimerasa II da lugar a múltiples RNAs que se diferencian por *splicing* alternativo y en algunos casos por el uso de diferentes sitios poly(A). Modificado de Cann, 1997.

Figura 11: Expresión del genoma en Adenovirus. Los transcritos tempranos se marcan en verde, los tardíos en azul, en marrón la ubicación de los genes VA RNA y en rojo las proteínas que codifica cada uno. Las flechas indican la dirección de la transcripción. MLP: *Major late promoter*, promotor tardío mayor. Modificado de Russel, 2000.

Figura 12: Unión de proteínas celulares a E1A (ver texto). CR1, CR2 y CR3 son regiones presentes en un amplio rango de adenovirus. PLDLS y LXCXE son

motivos de reconocimiento. E2F-1. *homology* se refiere a secuencias que desplazan E2F-1 de Rb y p300, *TATA-box-binding protein* (TBP) y *TBP-associated protein* (TAF). Modificado de Russell, 2000.

Figura 13: Interacción de E1A con NF- κ B. E1A bloquea la vía de NF- κ B, impidiendo su traslado al núcleo. I κ B factor inhibitorio asociado, IKK complejo kinasa. Modificado de Hiscott et al., 2001.

Figura 14: Interacción de E1A y E1B con célula hospedador. E1A desplaza E2F de Rb y se induce p53 aumentando la replicación viral y lisis. p21^{waf1/cip1} lleva al arresto del ciclo celular. E1B inhibe la apoptosis mediada por p53. Modificado de Dix et al., 2001.

Figura 15: Proteínas y funciones E3. Se muestran los distintos productos de E3 (ver texto). Las flechas horizontales indican la dirección de la transcripción. Modificado de Russell, 2000.

Figura 16: Proteínas y funciones de E4. Se muestran los distintos marcos de lectura (ORFs) de E4 y sus funciones conocidas. Modificado de Russell, 2000.

Figura 17: Mecanismos de interacción del Adenovirus con la célula. Se ilustran algunos sitios de acción de las proteínas virales (rojo) en algunas vías celulares (amarillo). Se muestra una partícula viral (verde) uniéndose al receptor. Modificado de Russell, 2000.

Figura 18: Infección lítica por Adenovirus. Las células infectadas con Adenovirus tienen afinidad por la Hematoxilina (colorante azul-purpúreo). Las células infectadas se redondean y la monocapa se desintegra. Las inclusiones basofílicas oscuras en el núcleo representan el cúmulo de proteínas virales en el sitio de ensamblado del virus.

Figura 19: Estructura del Hexón de Adenovirus. Modelo del monómero indicando los *loops* L1, L2 y L4 hacia la superficie y los pedestales P1 y P2 internos. Tomado de San Martin & Burnett (2003).

Figura 20: Hexón de Adenovirus. a- Esquema del perfil hidropático del Hexón de Adenovirus. Se muestra el número de residuos en aminoácidos y la localización de los *loops* y los pedestales en la secuencia primaria. b- Esquema de cintas del monómero del hexón de Adenovirus. Se muestran las regiones hipervariables (HVR). Abreviaturas: L (*loop*), P (pedestal), Pc (conector del pedestal), Ln (*loop* amino terminal), Δ = deleción. Tomado de Crawford-Mikszá et al., 1996.

Figura 21: Esquema de la estrategia de secuenciación del gen hexón. Se delinearon las regiones que rodean al hexón y se muestran las siete regiones hipervariables, los *primers* y las regiones secuenciadas. Tomada de Takeuchi et al., 1999.

Figura 22: Distribución por mes AV positivos 1999-2002. Se grafican los adenovirus positivos detectados por inmunofluorescencia por mes.

Figura 23: Distribución por edad AV positivos 1999-2002. Se grafican los adenovirus positivos detectados por inmunofluorescencia por grupo etario divididos en 1-3, 4-6, 7-9, 10-12, 13-15, 16-18, 19-21, 22-24 y >30 meses.

Figura 24: Causa de Internación por AV muestras 1999-2002. Se graficaron las causas de internación en 321 pacientes.

Figura 25: Relación entre datos meteorológicos y frecuencias de AV 1998-2002. Se graficaron los adenovirus positivos detectados por inmunofluorescencia por mes junto con los datos de temperatura media (°C) y humedad relativa (%) promedio mensual. Modificado de Viegas et al., 2003.

Figura 26: *Primers* usados para amplificación y secuenciación del hexón. Los *primers* señalados en negro fueron tomados de Takeuchi et al., 1999; en azul se señalan los *primers* diseñados de novo para Av7h. Se muestran las regiones que rodean al hexón y las siete regiones hipervariables.

Figura 27: Esquema del armado de la secuencia del gen hexón Av7h. Las flechas indican la dirección de los primers. Se marca en la parte superior en verde las regiones leídas 4 veces, en azul las leídas 2 veces y en rojo las leídas 1 vez. La escala indica el número de pares de bases.

Figura 28: Composición nucleotídica hexón AV7h. Se grafica el número de bases obtenido en la secuencia consenso.

Figura 29: Composición aminoacídica hexón AV7h. Se discriminaron en básicos (Lis, Arg), ácidos (Asp, Glu), hidrofóbicos (Ala, Ile, Leu, Phe, Trp, Val) y polares (Asn, Cys, Gln, Ser, Thr, Tyr).

Figura 30: Esquema del perfil hidropático del Hexón de AV7h. Se muestra el número de residuos en aminoácidos. a- Ploteo de hidrofobicidad en rojo y de superficie (azul). Se tomó como valor de corte 1. b- Esquema del ploteo, se muestra la localización de las regiones hipervariables.

Figura 31: Comparación de las secuencias inferidas de aminoácidos. Cepas Gomen (Ad7p X76551) Ad7h HVR (X76551) y 87-922 (cepa aportada por este trabajo). Los puntos (.) significan igual residuo, los guiones (-) no hay secuencias disponible y los cambios de aminoácidos están indicados. Se indican también las 7 HVRs y los *loops*.

Figura 32: Diversidad hexón AV7h vs. Gomen. Se grafica el número promedio de diferencias por sitio (Pi) con respecto a la cepa prototipo del serotipo 7

(Gomen) avanzando de a 10 nucleótidos en ventanas de 50 bases de longitud. Se tomaron como valores importantes los >1.

Figura 33: Localización gráfica de las regiones antigénicas Hexón AV7h. Se muestra el número de residuos en aminoácidos. El índice antigénico integra hidrofiliidad, probabilidad de exposición en superficie y estructura secundaria. Se dedujeron los epitopes B y T.

Figura 34: PCR especie específica. Los productos de PCR se corrieron en gel de agarosa al 2% en buffer TAE teñido con Bromuro de Etidio. Calle 1: marcador de 100 pb., calles 2 a 9, 11, 13 a 18 producto a cortar con *Taq I*, calles 10 y 12 especie B:2 (290 pb.), calle 19 control negativo y calle 20 control positivo.

Figura 35: Restricción con *Taq I* de los productos de 420 a 510 pb.. Los productos de restricción se corrieron en geles de poliacrilamida nativa al 10% en buffer TBE. Calle 1 marcador de 100 pb, calle 9 marcador de 25 pb, calle 15 marcador de 10 pb; calle 2 AV7h calles 3 a 7 y 12 patrón similar a AV7h, calles 8 y 11 patrón símil AV7g y p, calle10, control negativo, calle 13 patrón especie C, calle 14 AV2.

Figura 36: Alcance de las secuencias AV7 1990-2002 medido en aminoácidos. Se graficó el comienzo y el alcance de cada secuencia obtenida.

Figura 37: Inferencia filogenética hexones completos 1990-1999: La escala indica en el árbol los % de distancia y los números las sustituciones. Se utilizó la tabla de residuos PAM250. El código de las muestras es el usado en la Tabla 13.

Figura 38: Inferencia filogenética de las regiones hipervariables de los hexones 1990-2002. El árbol de distancia se construyó con 953 nucleótidos y se

tomaron como referencias secuencias disponibles de las distintas especies de Adenovirus humanos (A-F) y como grupo externo un adenovirus bovino. Se separan claramente las especies de Adenovirus humanos A (AV12 y 18), C (AV1, 2 y 6), D (AV8, 9, 10, 15, 19, 22, 24, 26, 37, 46 y 48), E (AV4) y F (40-41). Dentro de la especie B se separa la subespecie B2 (AV11 y 14) de la B1 (AV3 y 7). Dentro de la subespecie B1, podemos separar los clusters genómicos GTC1 (AV7p1 y Gomen) del GTC2 (AV7h y todas las muestras analizadas 1990-2002).

Figura 39: Corte de restricción para DNA genómico AV7h. Los fragmentos de restricción obtenidos con enzimas de restricción *Bam*H I, *Bgl* II, *Hind* III y *Sma* I se visualizaron en gel de agarosa al 1% en buffer TAE y se tiñeron con Bromuro de etidio.

Figura 40: Cobertura de genoma AV7h.. Se tomó para cada *primer* (eje y) su punto de origen putativo (referido a la cepa AV11p) y la longitud leída (eje x). Se graficaron las 155 secuencias consenso derivadas de los 155 *primers* diseñados *ad hoc*. Se adicionaron 5 secuencias derivadas de la biblioteca genómica y 4 de *primers* de diagnóstico. Se puede apreciar que la redundancia es alta y la cobertura es total

Figura 41: Esquema de los contigs del armado del genoma AV7h. Se señalan en celeste los *contigs* cuyos *primers* fueron diseñados en base a secuencias de AV7 Gomen y en naranja los que fueron diseñados en base a AV11p. Se muestra el alcance de cada uno de ellos (en pb.), cómo se solapan y se resalta el número de *primers* utilizado para armar cada *contig*.

Figura 42: Marcos de lectura abiertos (ORFs) del genoma de AV7h. Se utilizaron dos programas para predecir los marcos de lectura probables (*ORF*

Finder (NCBI) y *GeneMark* v2.4. Se analizaron los seis marcos de lectura probables (+1, +2, +3; -1, -2, -3) tomando como mínima proteína posible un péptido de 43 aa.. Se marcan en verde los marcos de lectura confirmados por homología con proteínas de AV.

Figura 43: Esquema del genoma de Mastadenovirus. Los genes comunes a todos los géneros se marcaron en celeste, los específicos en otro color. Las regiones codificantes se muestran como flechas. Modificado de Davison et al., 2003.

Figura 44: Mapa de transcripción de AV7h. Se señalan los genes en gris y las proteínas en rosa.

Figura 45: Relaciones filogenéticas entre AV humanos. Se usó el modelo de evolución F81+G con tasas de sustituciones equiprobables para construir un árbol por el método de *Maximum Likelihood* que genera la filogenia más probable. Se tomó como referencia el prototipo de cada especie reconocido por ICTV. HAVA-D: AV especie A-B-C-D. Se tomó como grupo externo un AV bovino (BAVB).