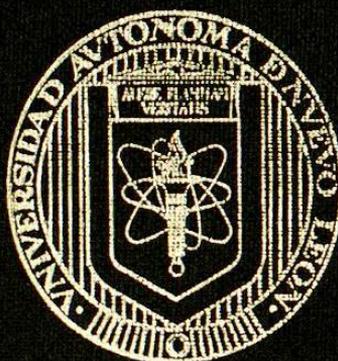


UNIVERSIDAD AUTONOMA DE NUEVO LEON

FACULTAD DE MEDICINA



**EVOLUCION MOLECULAR DE LA FAMILIA
GH/CSH EN PRIMATES**

Por

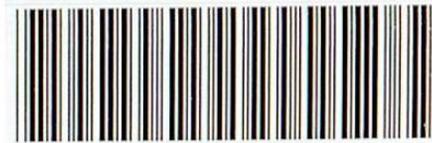
M.C. RAFAEL GONZALEZ ALVAREZ

**Como requisito parcial para obtener el Grado de
DOCTOR EN CIENCIAS con Especialidad en
Biología Molecular e Ingeniería Genética**

Febrero, 2006

TD
QH390
.G6
2006
c.1

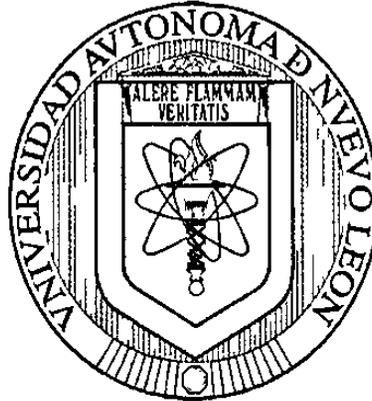
M.C. RAFAEL GONZALEZ ALVAREZ



1080128443

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE MEDICINA



EVOLUCIÓN MOLECULAR DE LA FAMILIA GH/CSH EN
PRIMATES

Por

M.C. RAFAEL GONZÁLEZ ALVAREZ

Como requisito parcial para obtener el grado de
DOCTOR EN CIENCIAS con Especialidad en
Biología Molecular e Ingeniería Genética

Febrero del 2006

①H390

•G6

2006

c.1



EVOLUCIÓN MOLECULAR DE LA FAMILIA GH/CSH EN PRIMATES

Aprobación de la Tesis:



DR. HUGO ALBERTO BARRERA SALDAÑA
Director de Tesis



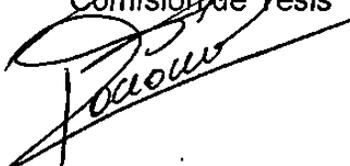
DRA. AGNES REVOL DE MENDOZA
Co-Director de Tesis



DRA. HERMINIA G. MARTÍNEZ RODRIGUEZ
Comisión de Tesis



DRA. DIANA RESENDEZ PÉREZ
Comisión de Tesis



DRA. ROCIO ORTÍZ LÓPEZ
Comisión de Tesis



DR. DIONICIO A. GALARZA DELGADO
Subdirector
de Estudios de Posgrado

El presente trabajo titulado “Evolución Molecular de la Familia GH/CSH en primates” se llevó a cabo por el M.C. Rafael González Álvarez en tres diferentes centros de investigación. El Laboratorio de Biología Molecular de la Unidad de Laboratorios de Ingeniería y Expresión Genéticas del Departamento de Bioquímica de la Facultad de Medicina de la Universidad Autónoma de Nuevo León. El Laboratorio de Evolución Molecular del Centro de Ciencias Genómicas de la Universidad Nacional Autónoma de México, ubicado en la ciudad de Cuernavaca, Morelos. El Human Genetics Center de la School of Public Health del Centro Médico de la Universidad de Texas en Houston. Bajo la dirección del Dr. Hugo Alberto Barrera Saldaña y la co-dirección de la Dra. Agnès Revol de Mendoza.

DEDICATORIA

A mi esposa Ana Isabel Rangel Pérez, a mis hijos Karen Arcelia y Eric Rafael; ya que sin el apoyo de ellos no me hubiera sido posible realizar este trabajo. A los tres gracias por esperar mis largas y constantes ausencias.

Muy especialmente a mis padres Rafael y Arcelia por su apoyo incondicional que recibo desde 1999, año en que me aventure a venir a Monterrey a realizar los estudios de Maestría. De igual manera gracias a ellos por esperar ansiosamente cada verano mi regreso a casa.

A mi abuela materna Irene Gómez de Alvarez, quien siempre espero la llegada de este día, sin embargo se me adelanto en el camino de la vida. Desde donde quiera que te encuentres, te digo que ya termine y este trabajo te lo dedico a ti abuelita.

A ti, mi hermana por estar siempre conmigo a pesar de la distancia que siempre nos mantuvo separados.

No podían faltar mis hermanos, Aarón y Fidel de quienes siempre he recibido su cariño, amistad y apoyo.

AGRADECIMIENTOS

A Dios por la vida y por haberme permitido alcanzar una nueva meta en la vida.

Al CONACYT y a la UANL por el apoyo económico que me permitió realizar esta aventura de mi vida.

Al Dr. Hugo A. Barrera Saldaña por haberme dado la oportunidad de pertenecer a su selecto grupo de investigación, motivarme a ver siempre hacia adelante. Por su apoyo para asistir a congresos y cursos, claro esta por haberme obsequiado la magnifica oportunidad de realizar una aventura científica en Houston. Mis mas sincero agradecimiento.

A la Dra. Agnès Revol de Mendoza, quien además de brindarme su amistad incondicional siempre estuvo dispuesta a contestar todas mis dudas y brindarme su valiosa asesoría.

A los Doctores Herminía G. Rodríguez Martínez, Rocío Ortiz López y Diana Résendez Pérez por la revisión de la presente tesis.

Al Zoológico Guadalajara y al Departamento de Farmacología de ésta Facultas quien muy amablemente nos proporcionaron las muestras de sangre de mono araña y mono rhesus, respectivamente. Muchas gracias, de no ser por su valiosa ayuda no hubiera sido posible la realización de este trabajo.

Al rector de nuestra Universidad, Ing. José Antonio González Treviño y al Dean de la Graduate School of Biomedical Sciences, Dr. George Stancel por el apoyo económico que recibí para realizar mi aventura científica en Houston, TX.

A mis amigos y compañeros del Human Genetics Center de la UT, el Dr Yun-Xin Fu, Sara A. Barton y Xioaming Liu. Gracias por su paciencia, enseñanzas y su amistad.

Al grupo de Centro de Ciencias Genómicas de Cuernavaca, Mor. Los Doctores Guillermo Dávila y Víctor González. Mis amigos Luís Fernández e Ismael Hernández, gracias por darme el tour turístico en cada una de mis visitas esa hermosa ciudad.

A todo el personal de la ULIEG y el Departamento de Bioquímica.

Mis compañeros de generación, Iván, Marcos, Diana, Gloria, Carmén y Enrique, con quienes pase gratos momentos durante mi estancia en Monterrey.

A todos mis amigos y compañeros de las generaciones posteriores a la mía. Como son tantos, por no decir todos preferí omitir nombres como medida de precaución y evitar que se me pase mencionar alguno de ellos. A todos ustedes, muchas gracias por compartir todo este tiempo en la ULIEG.

ÍNDICE GENERAL

LISTA DE FIGURAS	i
LISTA DE CUADROS	ii
NOMENCLATURA	iii
RESUMEN	v
CAPITULO I INTRODUCCIÓN	1
1.1 Las familias multigénicas	1
1.2 Bases moleculares de la evolución	3
1.3 La familia GH/PRL/CSH	4
1.4 Organización del locus GH en el humano	7
1.5 Función de las hormonas GH y CSH en primates	8
1.6 Regulación de la expresión de los genes GH/CSH en el humano	9
1.7 Hipótesis actual sobre la formación del locus hGH/hCSH	11
1.8 Taxonomía y evolución paleontológica de los primates	12
1.9 Situación de la genómica comparativa del locus GH al iniciar el presente trabajo	13
1.10 Justificación del trabajo	14
CAPITULO II OBJETIVOS	15
2.1 Objetivo general	15
2.2 Objetivos específicos	15
CAPITULO III ESTRATEGIA GENERAL	16
CAPITULO IV MATERIAL Y MÉTODOS	17
4.1 Material	17
4.1.1 Origen de los reactivos	17

4.1.2 Material Biológico	17
4.2 Equipo	18
4.3 Métodos	20
4.3.1 Extracción de DNA genómico	20
4.3.2 Amplificación de los genes y regiones intergénicas	20
4.3.2.1 Condiciones de amplificación de los genes del mono rhesus	21
4.3.2.2 Condiciones de amplificación de los genes del mono araña	22
4.3.2.3 Condiciones de amplificación de las regiones intergénicas del mono rhesus	23
4.3.3 Clonación de los genes y regiones intergénicas	24
4.3.3.1 Extracción del DNA plasmídico	24
4.3.3.2 Selección de las clonas recombinantes	25
4.3.4 Secuenciación de los genes y regiones intergénicas	25
4.3.4.1 Secuenciación de los genes de los monos rhesus y araña	26
4.3.4.2 Secuenciación de las regiones intergénicas del mono araña	26
4.3.4.3 Secuenciación de las regiones intergénicas del mono rhesus	27
4.3.5 Análisis de las secuencias	30
4.3.5.1 Ensamble de los <i>loci</i> GH/CSH	30
4.3.5.2 Análisis de las proteínas	30
4.3.5.3 Análisis del promotor proximal	31
4.3.5.4 Análisis del promotor distal	31
4.3.5.5 Identificación de los elementos repetitivos	32
4.3.5.6 Estimación de los índices evolutivos (K) en las regiones codificantes	32
4.3.5.7 Prueba de hipótesis al reloj molecular	32

4.3.5.8 Pruebas de selección	32
4.3.5.9 Identificación de los eventos de conversión génica	33
4.3.5.10 Estimación de los tiempos de duplicación	33
CAPITULO V RESULTADOS	34
5.1 Extracción del DNAg de los monos rhesus y araña	34
5.2 Clonación y secuenciación de los genes de los loci GH de los monos rhesus y araña	34
5.2.1 Clonación y secuenciación de los genes del locus GH del mono rhesus	34
5.2.2 Clonación y secuenciación de los genes del locus GH del mono araña	36
5.3 Amplificación, clonación y secuenciación de las regiones integénicas de los loci GH de los monos rhesus y araña	38
5.3.1 Amplificación, clonación y secuenciación de las regiones intergénicas del locus GH del mono rhesus	38
5.3.2 Subclonación y secuenciación de las regiones intergénicas del mono araña	39
5.4 Análisis de las secuencias	40
5.4.1 Ensamble de los loci GH/CSH	41
5.4.2 Análisis de las proteínas	41
5.4.3 Análisis del promotor proximal	44
5.4.4 Análisis de otros elementos regulatorios	45
5.4.5 Elementos de DNA repetitivo	46
5.4.5 Estimación de los índices evolutivos de las proteínas	48

5.4.6 Pruebas de selección	48
5.4.7 Pruebas de hipótesis al reloj molecular	49
5.4.8 Identificación de los eventos de conversión génica	50
5.4.9 Detalles de la comparación de los loci GH/CSH	51
5.4.10 Una nueva hipótesis de la evolución del locus GH/CSH	53
5.4.11 Estimación de los tiempos de duplicación	53
CAPITULO VI DISCUSIÓN	55
6.1 Obtención de los genes GH/CSH y sus regiones intergénicas	56
6.2 Análisis de las proteínas	57
6.3 Surgimiento de la expresión diferencial de los genes relacionados a GH	60
6.3.1 Análisis del promotor proximal	60
6.3.2 El elemento P	61
6.3.2 El potenciador	61
6.4 Expansión de las regiones intergénicas entre los MNM y MVM	62
6.5 Detalles de la comparación de los loci	63
6.6 Una nueva hipótesis sobre el origen del locus GH/CSH en el humano	64
6.6.1 Evidencias clínicas que soportan la nueva hipótesis	66
CAPITULO VII CONCLUSIONES	67
CAPITULO VIII PERSPECTIVAS	68
CAPITULO IX BIBLIOGRAFÍA	69

LISTA DE FIGURAS

1. Modos de evolución de las familias multigéncias	2
2. Organización del locus hGH/hCSH	7
3. Hipótesis actual sobre la formación del locus hGH/hCSH	12
4. Evolución de los primates	13
5. Estrategia general	16
6. Iniciadores usados en los ensayos de PCR	21
7. Estrategia de amplificación de los genes del mono araña	22
8. DNAg de los monos rhesus y araña	32
9. Amplificación de los genes del mono rhesus a partir de DNAg	35
10. Amplificación del sexto gen del mono rhesus a partir de un BAC	36
11. Amplificación de los genes del mono araña	36
12. Caracterización de las clonas de los genes del mono araña	37
13. Resultados de la secuenciación de los genes del mono araña	37
14. Amplificación de las regiones intergénicas del mono rhesus	38
15. Caracterización de las intergénicas del mono araña	37
16. Banco de 2 – 4 kpb de una intergénica del mono rhesus	39
17. Organización de los loci GH/CSH	41
18. Alineamiento de las proteínas	42
19. Árbol filogenético de las proteínas GH/CSH de primates	43
20. Alineamiento del promotor proximal	44
21. Alineamiento del elemento P	45
22. Alineamiento del potenciador	46
23. Anatomía de las regiones intergénicas	47
24. Organización de los loci GH/CSH analizados	52
25. Nueva hipótesis sobre el origen del locus GH/CSH en el humano	53

LISTA DE CUADROS

1. Oligonucleotidos utilizados en este trabajo	21
2. Condiciones de amplificación de los genes GH/CSH del mono rhesus	22
3. Composición de las mezclas de reacción del PCR largo	22
4. Programa de amplificación del PCR largo	24
5. Secuencias utilizadas en este estudio	31
6. Secuencias generadas en este trabajo	40
7. Cambios aminoacidicos entre las proteínas	41
8. Índices evolutivos de las proteínas relacionadas a GH del mono rhesus	48
9. Pruebas de selección	49
10. Resultados de las pruebas de hipótesis del reloj molecular	50
11. Identificación de los eventos de conversión génica	50
12. Segmentos utilizados para estimar los índices de evolución	54
13. Número de diferencias entre los segmentos utilizados para estimar los índices, así como de los bloques de duplicación	54
14. Tiempos de duplicación, en millones de años (MYA)	54

NOMENCLATURA

$^{\circ}\text{C}$	Grados centígrados
A_{260}	Absorbancia a 260 nano metros
AMPc	Adenosina monofosfato cíclico
BAC	Cromosoma artificial bacteriano
CSH	Chorionic Somatommamotropin Hormona
dATP	Desoxiadenosina trifosfato
d_N	Número de sustituciones no sinónimas por sitio no sinónimo
DNA	Ácido desoxirribonucleico
DNAc	DNA complementario al RNA
DNAg	DNA genómico
dNTPs	Desoxirribonucleosidos trifosfato
d_S	Número de sustituciones sinónimas por sitio sinónimo
EDTA	Ácido etilen-diamino-tetra-acético
GH	Hormona del crecimiento
HPLC	Cromatografía de líquidos de alta resolución
IGF-1	Factor de crecimiento semejante a la insulina tipo I
K	Índice de sustitución sitio/año
K_A	Índice de sustitución sitio/año por sitio no sinónimo
kDa	Kilo Dalton
kpb	Kilo par de base
K_S	Índice de sustitución sitio/año por sitio sinónimo
kV	Kilo volts
LCR	Región controladora del locus
M	Concentración molar
mg	Miligramos
min	Minutos
ml	Mililitros
MNM	Monos del nuevo mundo
MVM	Monos del viejo mundo
MYA	Millones de años atrás

ng	Nanogramos
pb	Par de base
PCR	Reacción en cadena de la polimerasa
PL	Lactógeno placentario
PRL	Prolactina
PRP	Proteínas relacionadas a prolactina
psi	Unidades de presión
r	Índice de sustituciones nucleotídicas
RNA	Ácido ribonucleico
RNA _m	RNA mensajero
RNA _r	RNA ribosomal
rpm	Revoluciones por minuto
SDS	Duodecil sulfato de sodio
SL	Somatolactina
SNP	Polimorfismo de un solo nucleotido
TAE	Solución amortiguadora de Tris, ácido acético y EDTA
TE	Solución amortiguadora de Tris y EDTA
X	Número de veces la concentración
μF	Micro Faradios
μl	Micro litros
μM	Concentración micro molar
Ω	Ohms

RESUMEN

Facultad de Medicina de la UANL

M.C. Rafael González Álvarez

Titulo del Estudio: Evolución Molecular de la Familia GH/CSH en Primates

Numero de páginas: 74

Candidato al grado de Doctor en Ciencias
con especialidad en Biología Molecular e
Ingeniería Genética

Mientras que la mayoría de los mamíferos incluyendo los prosimios, tienen una sola copia del gen de la hormona del crecimiento (GH), los monos antropoides un locus de genes relacionados a GH. A lo largo de la evolución de los antropoides [monos del nuevo mundo (MNM), monos del viejo mundo (MVM) y grandes monos], dos características se observan en sus loci GH: primera, la aparición de los genes chorionic somatomammotropin hormona (CSH) dentro del linaje de los MVM, segundo, la extensión de las regiones intergénicas de los loci GH en dichos monos. En relación con esta extensión de los loci GH, las regiones intergénicas de los MNM poseen longitudes homogéneas (3.5 kpb), en contraste con las longitudes heterogéneas (6 y 13 kpb) que se observan en los MVM y grandes monos. A la fecha, no se ha descrito la organización de ningún locus GH en antropoides no humanos. Aquí, describimos la organización genómica del locus GH del mono del mono rhesus. Dicho locus, consiste de seis genes relacionados a GH, separados por cinco regiones intergénicas. El gen del extremo 5' (GH-1) codifica para la GH que se expresa en la pituitaria. Es seguido por los genes CSH-1, GH-2, CSH-2, CSH-3 y CSH-4. Las cinco regiones intergénicas tienen longitudes heterogéneas y también presentan más o menos la misma distribución de Alu observada para el locus GH del humano. Para analizar los acontecimientos que contribuyeron a la extensión de las regiones intergénicas de los loci GH y de la aparición de los elementos reguladores, cinco regiones intergénicas del mono de la araña fueron secuenciadas. Nuestros resultados sugieren que las regiones intergénicas largas (13 kpb) del locus GH del mono rhesus tienen un antepasado común con las regiones intergénicas de 3.5 kpb del mono araña. Este incremento en la longitud de las regiones intergénicas del mono rhesus, es debido a una inserción (~ 8,7 kpb) en su extremo 3'. Interesantemente, en este inserto, descubrimos un elemento de DNA que se asemeja al potenciador de los genes CSH del humano. Por otra parte, observamos que las regiones intergénicas cortas (6 kpb) aumentaron su tamaño mediante un evento independiente de recombinación.



Dr. Hugo A. Barrera Saldaña
Director



Dra. Agnès Revol de Mendoza
Co-Director

CAPITULO I

INTRODUCCION

1.1 Las familias multigénicas

Se denomina familia multigénica a un conjunto de genes que provienen de la duplicación y variación de un gen ancestral y que aún comparten una similitud en sus secuencias nucleotídicas de al menos el 50%. Sus miembros pueden estar juntos integrando un locus multigénico o dispersos en cromosomas diferentes, o una combinación de ambos. Los miembros de una familia de genes suelen tener funciones relacionadas, traslapadas e incluso idénticas, aunque puedan expresarse en momentos o en tipos celulares diferentes (Singer, 1982).

La evolución de las familias multigénicas ha sido tema de gran interés por muchos años. Su estudio antes de 1970 abrumadoramente se centraba en las cadenas α , β , γ y δ de la hemoglobina y en la mioglobina (Ingram, 1961). Los genes que codifican estos polipéptidos o proteínas están filogenéticamente relacionados, han divergido gradualmente y las copias que han emergido a lo largo de la evolución han adquirido nuevas funciones, tal como es el caso de la hemoglobina y mioglobina. A este modo de evolución se le conoce como “evolución divergente” (Figura 1A).

Fue alrededor de 1970 cuando varios investigadores observaron que los RNAs ribosomales (rRNAs) en *Xenopus* eran codificados por una gran cantidad de genes en *tandem* y que las secuencias de las regiones intergénicas eran más similares dentro de la

misma especie que entre dos especies (Brown et al., 1972). Estas observaciones eran difíciles de explicar por el modelo de la evolución divergente. Por ello, un nuevo modelo llamado “evolución concertada” fue propuesto (Figura 1B). Este asume que todos los miembros de una familia multigénica evolucionan de una manera paralela, más que de manera independiente. Es decir, si ocurre una mutación, ésta se transmite y mantiene en todos los miembros de la familia.

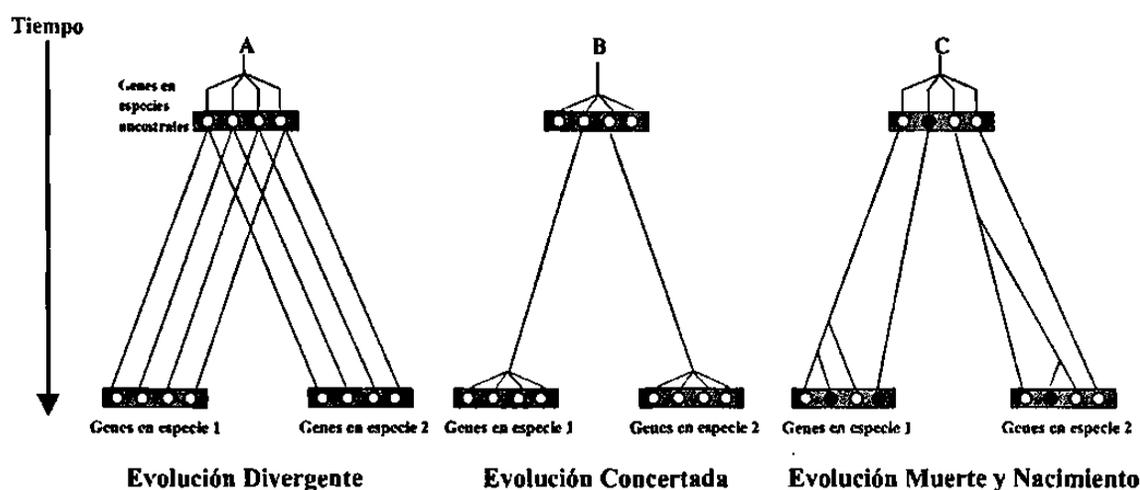


Figura 1.- Modos de evolución de las familias multigénicas. Los círculos en blanco representan genes funcionales, mientras que los negros corresponden a pseudogenes. Tomado de Nei et al., 2005.

Más adelante, conforme las secuencias de DNA y de aminoácidos llegaron a estar disponibles, la aplicabilidad de la evolución concertada a algunas familias multigénicas fue cuestionada y otro modelo llamado la “evolución por nacimiento y muerte” fue propuesto (Nei et al., 1997). En este último, algunos de los genes nuevos son resultado de las duplicaciones y se mantienen en el genoma durante mucho tiempo, mientras que otros se suprimen o acumulan mutaciones deletéreas que los convierten en pseudogenes, es decir, genes cuyas mutaciones les impiden sintetizar una proteína activa (Figura 1C). El modelo de “evolución por nacimiento y muerte” se aplica a la mayoría de las familias multigénicas involucradas en el sistema inmune, tales como las inmunoglobulinas y los

antígenos del complejo mayor de histocompatibilidad (MHC) (Hughes et al., 1990; Ota et al., 1994), así como a los genes relacionados con la resistencia a enfermedades (Zhang et al., 2000). En resumen, todas las familias multigénicas pueden seguir cualesquiera de las rutas descritas en la figura 1.

La familia multigénica sobre cuya evolución versa este trabajo es la de la hormona del crecimiento, comúnmente referida como GH por sus siglas del inglés (growth hormone), misma que se describe a continuación.

1.2 Bases moleculares de la evolución

La principal causa de evolución es la mutación. Una mutación en un gen o cualquier secuencia de DNA es causada por una sustitución, inserción o deleción de uno o varios nucleótidos, así como por eventos de recombinación y conversión génica. Las mutaciones se fijan en la población mediante la selección natural. El nuevo gen mutante produce un nuevo carácter morfológico o fisiológico, el nuevo carácter es heredado a la descendencia. Si se es capaz de construir un árbol filogenético para un grupo de especies, se puede definir cuando apareció la mutación que origino dicho carácter. Esta información es útil para el entendimiento de los mecanismos de evolución de cualquier carácter. Mediante la comparación de las condiciones ambientales de donde vive la especie con el carácter de interés con aquellas especies que carecen del carácter, se logra el entendimiento de los mecanismos de evolución que actúan sobre dicho carácter. Dichos mecanismos son la selección positiva, purificación de la selección o evolución neutral. Mediante la identificación de los genes involucrados en la adquisición de nuevos caracteres, los mecanismos de evolución que actúan sobre ellos es la manera de inferir el

tipo de mutación que origino el nuevo cambio morfológico o fisiológico (Nei et al., 1987; Clark et al., 1999).

Sin embargo, el estudio de los cambios en el DNA es más complicado de lo que aparenta. Debido a que existen varios “tipos” de DNA, como: regiones codificantes, regiones no codificantes, regiones intergénicas, secuencias repetitivas y secuencias reguladoras. De ahí la importancia de identificar previamente el tipo de DNA que se desea analizar. Si se desea analizar el DNA codificante, se debe considerar los patrones de sustitución nucleotídica en el primero, segundo y tercer codón. Además de precisar si dichos cambios repercuten en el cambio de aminoácido (sustituciones no sinónimas) o se se trata simplemente de un cambio sin sentido (sustitución sinónima). Para el caso de las regiones no codificantes, generalmente se aplica la “hipótesis del reloj molecular”. Está hipótesis establece que el índice de sustituciones es aproximadamente constante a lo largo del tiempo. Aunque estrictamente hablando ninguna secuencia de DNA evoluciona a un índice constante, debido a que se ganan nuevas funciones, particularmente cuando el número de genes por genoma aumenta. También cabe destacar que el mecanismo de reparación del DNA repercute considerablemente en la aplicación estricta de esta hipótesis (Nei, 2000).

1.3 La familia GH/PRL/CSH

La familia multigénica de la hormona del crecimiento (GH), se incluye a su vez en una superfamilia que también abarca a la prolactina (PRL). Dependiendo del orden de vertebrados que se trate, también incluye en los primates a los genes de la somatomotropina coriónica hormona (CSH), anteriormente referida como Lactógeno Placentario (PL, como sigue llamándose en mamíferos no primates), las proteínas

relacionadas a prolactina (PRP) y la somatolactina (SL). Esta super familia representa uno de los grupos proteicos fisiológicamente más diversos que han evolucionado por duplicación génica (Rubin et al., 1994). Una breve descripción de estas hormonas sigue a continuación:

1. GH. Esta hormona se produce en la hipófisis de todos los mamíferos. En el caso de los primates, existen dos genes, el de la GH-1 con expresión en la hipófisis y el gen GH-2 cuya expresión se restringe a la placenta. En general, las GH presentan gran similitud en sus secuencias, son de tipo globular, constan de 191 aminoácidos en su forma madura, con un peso molecular de 22 kDa, poseen cuatro hélices alfa y dos puentes disulfuro y por lo general carecen de residuos de azúcares. Recién sintetizadas presentan un péptido señal de 26 aminoácidos para dirigir el transporte de la pre-hormona a través de la membrana del retículo endoplásmico rugos hacia el aparato de Golgi, en donde son secretadas al torrente sanguíneo en forma madura (Handwerger et al., 2000).

2. PRL. Es la hormona hipofisiaria que más funciones tiene por sí sola. Originalmente, fue caracterizada por su habilidad de estimular el crecimiento y desarrollo de la glándula mamaria, así como la producción de leche (Goffin et al., 1999). Estudios subsecuentes le atribuyeron a lo largo de los vertebrados aproximadamente 85 funciones diferentes agrupadas en seis categorías: i) equilibrio hídrico y electrolítico, ii) crecimiento e desarrollo, iii) endocrinología y metabolismo, iv) cerebro y conducta, v) reproducción e vi) inmunoregulación y protección.

En roedores, el gen de la PRL además de expresarse en hipófisis, se expresa en varios sitios como: cerebro, decidua, miometrium, glándula lacrimal, timo, baso, linfocitos, células linfoides de medula ósea, células epiteliales de glándula mamaria, tumores de

glándula mamaria, fibroblastos de la piel y glándulas sudoríparas. Adicionalmente, la PRL puede encontrarse en diferentes fluidos corporales como: líquido amniótico, lágrimas, leche y sudor (Ben-Jonathan et al., 1996).

3. CSH. En varios monos del Viejo Mundo (MVM), el chimpancé y el humano además de genes relacionados a GH en sus loci aparecen genes del tipo CSH (Chen et al., 1989; Golos et al., 1993; Revol et al., 2004, Ye et al., 2005). Aunque varios genes relacionados a GH se han descrito y analizado en primates (Revol et al., 2004; Ye et al., 2005; Li et al., 2005) la familia GH sólo ha sido estudiada en detalle en el humano (ver el cuadro 5 en material y métodos).

4. PLP. En roedores se han identificado 15 genes, los cuales se expresan en pituitaria, útero y placenta. La mayoría de ellas aumenta su producción en forma gradual desde el implante del embrión hasta el parto (Swenne et al., 1987). Al igual que en roedores, existen reportes de PLP en rumiantes. Estas proteínas muestran una similitud aminoacídica entre 43 a 51% con respecto a la propia PRL. Las células responsables de la expresión de estos genes son las células coriónicas binucleadas, pero su función continúa siendo desconocida (Anthony et al., 1995).

5. SL. Ha sido reportada en diferentes grupos de peces y en un anfibio, pero no en otra clase de vertebrados (Ameniya et al., 1998), aunque no se descarta la posibilidad de que esté presente en otras especies. El tamaño de esta hormona puede variar de 24.5 a 28 kDa. A diferencia de la GH que se encuentra en forma no glicosilada, la SL, al igual que la PRL, se puede encontrar en las formas glicosilada y no glicosilada, dependiendo de la especie. Las características moleculares y bioquímicas de esta hormona se han estado esclareciendo, pero su función fisiológica aún no está bien elucidada. Se han sugerido

diversas funciones para la SL, incluyendo su participación en la maduración sexual, regulación de calcio, respuesta al estrés, regulación ácido-base, metabolismo de las grasas y adaptación a los cambios de salinidad (Kaneko et al., 1996).

1.4 Organización del locus GH en el humano

En el humano, el locus GH está constituido por cinco genes y localizado en un área de ~50,000 pb en la banda q22.4 del cromosoma 17 (Figura 2), (Harper et al., 1982).

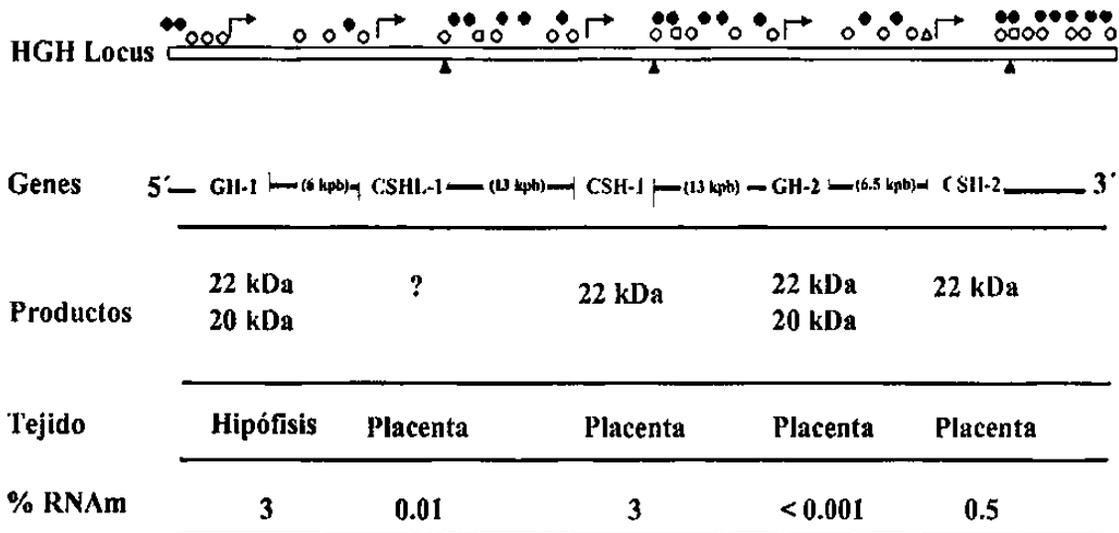


Figura 2. Organización del locus hGH/hCSH. Se muestran los productos de los genes, el tejido donde se expresan y el porcentaje del RNAm en dicho tejido. Las secuencias Alu están distribuidas en las regiones que flanquean a los genes, representadas aquí por círculos rellenos para el caso de los miembros con la misma orientación que la de los genes, mientras que con los círculos abiertos señalan secuencias Alu en la dirección opuesta y los semicírculos identifican a los miembros Alu truncados. Se encuentra una sola copia de las secuencias repetitivas de la familia Kpn en la región promotora distal del hCSH-2, representada por un triángulo en blanco, mientras que el potenciador o enhancer se indica debajo de los genes con un triángulo oscuro. Los genes están separados por regiones intergénicas (de 6 a 13 kpb).

Los genes del locus hGH/hCSH comparten una similitud nucleotídica de entre 91-95% (Chen et al., 1989). Dos de ellos codifican para GHs: el gen hGH-1 (normal) con expresión específica en los somatotrofos de la glándula pituitaria y el gen hGH-2, el cual codifica para una variante placentaria (Frankenne et al., 1987). De los tres genes restantes, también de expresión placentaria, hCSH-1 y hCSH-2 codifican para una misma hormona madura (Barrera-Saldaña et al., 1983). Por último, el gen hCSHL-1 se

considera un pseudogen, pues contiene una mutación en el segundo intrón que evita el procesamiento correcto de sus transcritos (Reséndez-Pérez et al., 1990).

Los cinco genes están separados por cuatro regiones intergénicas: dos de 13 kpb, una de 6.5 kpb y una de 6 kpb (Chen et al., 1989). Se conoce la secuencia nucleotídica de cada una de ellas y se han identificado tanto secuencias repetitivas, como elementos que participan en la regulación de la expresión de los genes. Dentro del locus se encuentran 48 copias de secuencias repetitivas de la familia Alu y una de la familia Kpn (Figura 2). Las inserciones Alu ocurren en ambas orientaciones y tres de ellas están truncadas. El único miembro de la familia Kpn se encuentra frente al último gen del locus (hCSH-2). Cabe señalar, que más del 20% de las secuencias nucleotídica del locus hGH corresponden a elementos Alu, cuyo patrón de distribución semejante para los cinco genes, sugiere su posible implicación en la ruta evolutiva que dio forma al locus (Chen et al., 1989).

1.5 Función de las hormonas GH y CSH en primates

Las GHs de los mamíferos son en su forma madura polipéptidos típicamente de 191 aminoácidos y 22 kDa de peso molecular. Son producidas y secretadas por los somatotrofos de la hipófisis anterior y desempeñan un papel esencial en el crecimiento lineal postnatal y en numerosos sistemas y órganos, participando en procesos tan diversos como la cicatrización de heridas, la homeostasis de la glucosa, el desarrollo del cerebro, la neurogénesis, la neuroprotección, el aprendizaje y la memoria (Sonntag et al., 2005). Por su parte, los niveles de GH-2 en la placenta se asocian con los índices de crecimiento intrauterino, peso al nacer y con los niveles fetales del factor de crecimiento parecido a la insulina I (IGF-I) (Verhaeghe et al., 2002; Chellakooty et al., 2004). Un

papel similar se le adjudica a los hCSH, cuyos niveles se le han asociado con el buen funcionamiento placentario y a los niveles fetales de IGF-1 (Handwerger et al., 2000; Karabulut et al., 2001). Niveles bajos de hCSH se asocian al estado hipoglicémico, crecimiento intrauterino retardado (Bagga et al., 1990) y en algunos casos con deficiencias neurológicas (Paci et al., 1989). Es de interés señalar que diversos genotipos del polimorfismo SNP D175254, localizado a ~ 1.4 kb del promotor de hCSH-1, se han asociado a los pesos corporales bajos en el nacimiento e infancia, así como a diversos síndromes metabólicos como la diabetes, la hipertensión y las enfermedades cardíacas del feto (Day et al., 2004).

1.6 Regulación de la expresión de los genes GH/CSH en el humano

Las 500 pb de la región promotora de todos los miembros de esta familia mutigénica tienen una similitud nucleotídica de más del 90% y son suficientes para dirigir la expresión del gen hGH-1 en los somatotrofos de la pituitaria anterior. Esta es controlada por el factor transcripcional Pit-1/GHF1 (Karin et al., 1990), el cual presenta dos dominios conservados entre los factores de transcripción específicos de tejido: un dominio homeobox (HB) y un dominio POU, ambos ubicados en el extremo carboxilo terminal de la proteína y que intervienen en la unión de Pit-1 al DNA. Este factor se une a los promotores de los genes GH/CSH en dos sitios (a -92 y -130 nucleótidos del sitio de inicio de la transcripción) que se encuentran altamente conservados entre los promotores de todos los genes de esta familia, a excepción del hGH-2. Existe una región (entre 1,500 y 2,000 pb río arriba del sitio de inicio de la transcripción), a la cual se unen proteínas que interfieren con la unión de Pit-1 a los genes placentarios, impidiendo así su expresión en la hipófisis (Nachtigal et al., 1993). Por ser Pit-1 un factor específico de

tejido hipofisiario (Ingraham et al., 1988), aparentemente no está involucrado en la inducción de la expresión de los genes placentarios.

Otros factores han sido involucrados en la regulación de la expresión del gen hGH-1 en la hipófisis, potenciando la acción de Pit-1. Algunos de ellos son: AP-2 (Lefevre et al., 1987), NF1 (Stephanou et al., 1995) y USF (Lemaigre et al., 1990; Courtois et al., 1990). Estos factores se unen al promotor en una región de 290 a 255 pb río arriba del sitio de inicio de la transcripción. Otro factor relacionado con Pit-1 es Zn-15 (descrito en rata), el cual funciona como potenciador de Pit-1 (Lipkin et al., 1993).

La expresión específica de los genes hCSHs en la placenta se atribuye a un elemento potenciador localizado río abajo de los genes hCSHs (Jacquemin et al., 1994), así como a la interacción de un elemento iniciador específico de tejido (InrE), que se localiza entre la caja TATA y el sitio de inicio de la transcripción de dichos genes (Jiang et al., 1995). El factor Sp1 juega un papel muy importante en la regulación de los genes hCSHs y hGH-2, ya que la delección de su sitio de unión en los promotores de estos genes, disminuye de manera significativa su expresión en células BeWo y hasta en un 50% en células JEG-3 (Karin et al., 1990; Fitzpatrick et al., 1990; Barboza-Cerda., 1998).

Los genes GHs y CSHs están también bajo el control de reguladores metabólicos como hormonas (T₃, glucocorticoides), ácido retinoico y AMPc, entre otros, ubicándose las secuencias de respuesta a estos mediadores en la parte distal del promotor (Stephanou et al., 1994; Stephanou et al., 1995).

Todos los elementos hasta aquí mencionados, corresponden a factores transcripcionales que reconocen a los promotores distal y proximal; sin embargo estos resultados son insuficientes para explicar la regulación fina que sucede *in vivo*. Esta observación hizo

suponer la existencia de un elemento regulador localizado fuera del locus. Mediante ensayos de hipersensibilidad a la DNasaI fue posible identificar a 14.5 kpb río arriba del promotor de hGH-1 una Región Controladora del Locus (LCR “Locus Control Region” por sus siglas en inglés), como la responsable de afinar el patrón de expresión de hGH-1 en la hipófisis y del resto de los genes en la placenta (Su et al., 2000).

1.7 Hipótesis sobre el origen del locus hGH/hCSH

El análisis de las secuencias del locus hGH/hCSH evidenció que la expansión del mismo se generó por una serie de duplicaciones mediadas por las secuencias repetitivas del tipo Alu. Los genes y las secuencias flanqueantes se produjeron mediante tres eventos independientes de duplicación (Figura 3). A partir de un gen ancestral (similar al actual gen hGH-1). Una primera duplicación y el acumulo de variaciones resulto en el surgimiento de los precursores de los genes hGHs y hCSHs (primera duplicación). Una segunda duplicación dio origen a los precursores de los genes hGH-2, hCSH-1 y hCSH-2. El tercer y último evento se produjo mediante una duplicación del fragmento que contenía al gen hCSH-1 para generar a los actuales genes hCSHL-1 y hCSH-1 (Hirt et al., 1987; Chen et al., 1989).

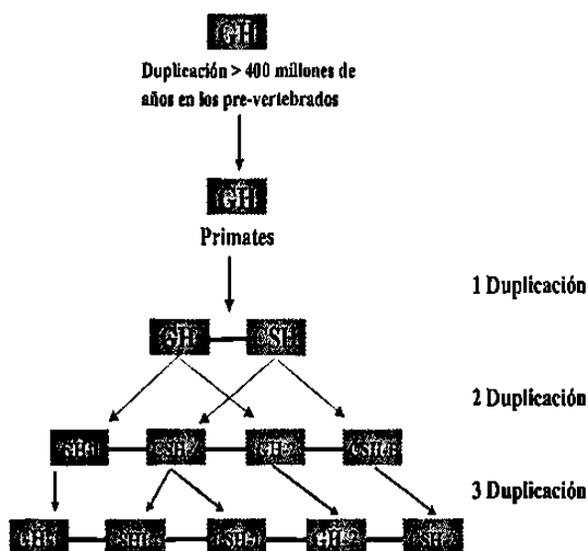


Figura 3. Hipótesis sobre la formación del locus hGH/hCSH. Se esquematizan eventos de duplicación que dieron origen a la familia multigénica en primates. Los CSH se generaron a partir del gen de la GH.

1.8 Taxonomía y evolución paleontológica de los primates

La línea de los primates parece haber empezado en la musaraña arborícola, animal semejante a la ardilla, de caracteres intermedios entre insectívoros y primates (Figura 4).

La evolución de los primates se caracteriza en general por su adaptación a la vida arbórea y solo al llegar a los grandes monos y al hombre, la misma ha sido invertida. Los primates son mamíferos no especializados. Entre sus rasgos de adaptación a la vida arbórea cabe citar las manos y pies prensiles con pulgares y dedos gordos oponibles, dedos terminados en uñas planas, extremidades superiores e inferiores móviles, flexibles y largas; cerebro bien desarrollado y visión binocular. Los dos subórdenes de primates son: los prosimios y los antropoides. Los lémures y los tarsios, que viven actualmente en África y Asia sudoriental son prosimios que han experimentado cambios muy discretos en relación a los de las formas originales. A partir de los prosimios, evolucionaron en el oligoceno tres grupos principales, *Ceboidea*, o monos del nuevo mundo (MNM), *Cercopithicoidea*, o monos del viejo mundo (MVM) y *Hominoidea*, que incluye al

hombre y a los grandes monos. Los cébidos representan un grupo de primates que quedaron aislados en América del Sur durante el terciario y desde entonces siguieron una evolución independiente a la de los otros primates (Ville, 1996).

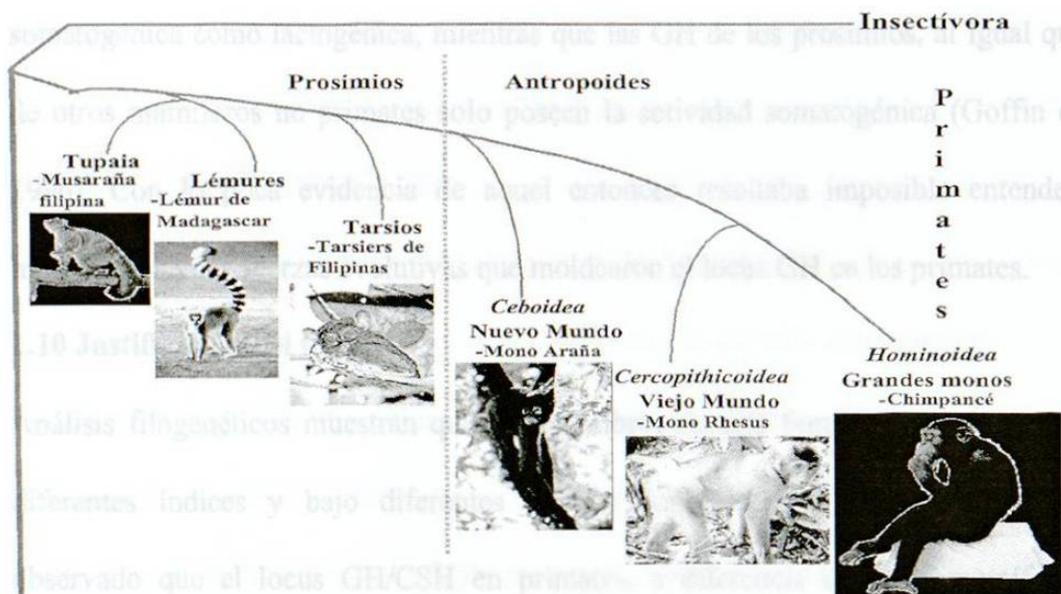


Figura 4. Evolución de los primates. Se muestra en forma esquematizada la evolución de los primates (prosimios y antropoides), según datos paleontológicos. Divergieron los monos del nuevo mundo, los monos del viejo mundo y finalmente los grandes monos y el humano.

1.9 Situación de la genómica comparativa del locus GH al iniciar el presente trabajo

Antes de comenzar el presente trabajo, prácticamente sólo se conocía la organización del locus GH/CSH en el humano, mientras que para el resto de los primates únicamente se contaba con datos aislados. Para los prosimios ya era conocido que solo poseen un gen de GH en sus genomas (Adkins et al., 2001), en los MNV se tenía la secuencia completa de ocho genes (Wallis et al., 2002), del mono araña (otra especie de los MNM) se contaba con la secuencia de cuatro genes y clonas de regiones intergénicas (Esquivel-Escobedo, 1999). Para el mono rhesus, especie representativa de los MVM se contaba

con la secuencia de los cDNAs de la GH-1, tres CSH y la secuencia completa de la GH-2 (Golos et al, 1993). Finalmente, del chimpancé, especie representativa de los grandes monos, se disponía de la secuencia completa de dos GHs y cuatro CSHs (Revol et al., 2004). Adicionalmente ya se sabía que las GH de primates poseen actividad tanto somatogénica como lactogénica, mientras que las GH de los prosimios, al igual que las de otros mamíferos no primates solo poseen la actividad somatogénica (Goffin et al., 1996). Con lo poca evidencia de aquel entonces resultaba imposible entender los mecanismos y las fuerzas evolutivas que moldearon el locus GH en los primates.

1.10 Justificación del trabajo

Análisis filogenéticos muestran que los miembros de esta familia evolucionaron bajo diferentes índices y bajo diferentes fuerzas de selección. Adicionalmente se ha observado que el locus GH/CSH en primates, a diferencia de otros mamíferos, se caracteriza por i) la actividad lactogénica de la GH expresada en hipófisis, ii) el aumento en el número de genes, de 1 a 8; iii) la especificidad tisular de la expresión de dichos genes y iv) la aparición de los CSHs. Sin la disponibilidad de la secuencia de otros loci y el análisis de los elementos reguladores de la transcripción, es difícil determinar el significado biológico de esta inusual variación en el contenido de genes, índices de evolución, fuerzas de selección y expresión diferencial en espacio y tiempo de los genes. Por tal razón, se propuso caracterizar los loci GH de especies representativas de diferentes subórdenes de primates, para tratar de reconstruir la historia evolutiva de estos genes. La especie seleccionada dentro de los monos del Nuevo Mundo es el mono araña (*Ateles geoffroyi*) y de entre los monos del Viejo Mundo es el mono rhesus (*Macaca mulatta*).

CAPITULO II

OBJETIVOS

2.1 OBJETIVO GENERAL

Inferir la historia evolutiva del locus GH en primates.

2.2 OBJETIVOS ESPECÍFICOS

1.- Establecer la composición y organización de los loci GH.

a) Aislar las regiones intergénicas y genes del mono rhesus (*Macaca mulatta*).

b) Aislar los genes faltantes del mono araña (*Ateles geoffroyi*).

c) Clonar y secuenciar los genes y regiones intergénicas aisladas.

2.- Realizar el análisis detallado de las secuencias codificantes y no codificantes.

a) Para las codificantes, determinar el número de sustituciones/sitio/año (K), el número de sustituciones sinónimas (K_S), no sinónimas (K_A) y la relación K_A/K_S .

b) Para las no codificantes, estimar el índice de sustituciones nucleotídicas (r).

c) Construir los árboles filogenéticos correspondientes.

3.- Realizar el análisis comparativo de las regiones intergénicas para:

a) Analizar las regiones intergénicas de los loci GH, en cuanto a su patrón de secuencias Alu y los elementos de regulación de la transcripción.

4.- Con todo lo anterior, inferir la historia evolutiva de la familia GH en ambos primates y relacionarla ésta con la del humano y chimpancé.

CAPITULO III

ESTRATEGIA GENERAL

La estrategia general (Figura 5) comprendió varias etapas. Primero se extrajo el DNA genómico de los mono rhesus y araña a partir de sus sangres periféricas. Adicionalmente se obtuvo una librería genómica en BACs de mono rhesus. Mediante iniciadores consenso se amplificaron los genes y las regiones intergénicas. Luego estos productos amplificados se clonaron y secuenciaron. Finalmente, Con las secuencias de los genes se determinaron varios parámetros evolutivos, mientras que el análisis de las regiones intergénicas consistió en determinar el patrón de las secuencias Alu e identificar los elementos involucrados en la regulación transcripcional.

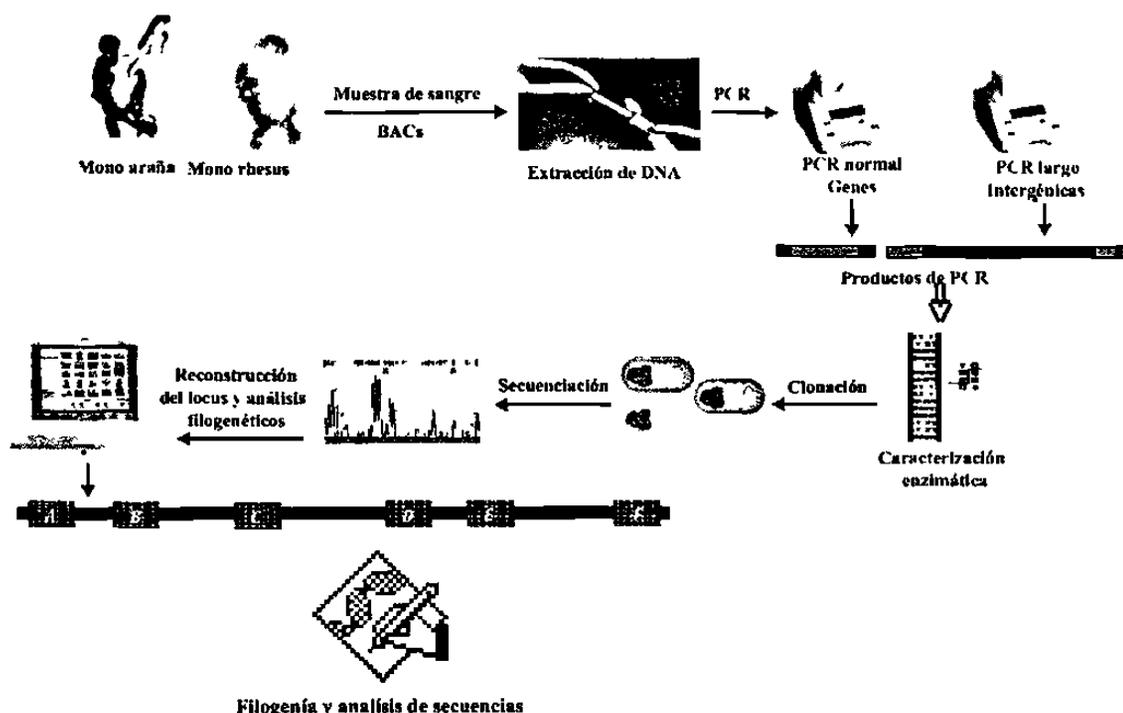


Figura 5.- Estrategia general. Comprendió las etapas de extracción del DNA genómico, purificación de los BACs, amplificación de los genes y regiones intergénicas, clonación, secuenciación y análisis bioinformático de las secuencias.

CAPITULO IV

MATERIAL Y MÉTODOS

4.1 Material

4.1.1 Origen de los reactivos

Los reactivos utilizados para la Reacción en Cadena de la Polimerasa (PCR) fueron obtenidos de las siguientes casas comerciales: la enzima DNA polimerasa *Taq* y el $MgCl_2$ de Promega Corporation (Madison, WI, EUA), la enzima DNA polimerasa *rTth* y el $Mg(OAc)_2$ de Applied Biosystems (Foster City, CA, EUA) y los dNTPs de GIBCO-BRL (Gaithersburg, MD, EUA). Las enzimas de restricción y los diversos iniciadores que se utilizaron fueron de GIBCO-BRL y New England Biolabs (Beverly, MA, EUA). La enzima RNasa y la kanamicina provinieron de SIGMA (St. Louis, MO, EUA). El vector de clonación pTOPO-XL-3.5 fue de Invitrogen (La Jolla CA, EUA). La secuenciación se llevó a cabo con los estuches comerciales de LiCor (LiCor, Lincoln NE) y ABI Prisma (Perkin Elmer, Cetus, Norwalk, CT). Los reactivos para elaborar las soluciones amortiguadoras, medios de cultivo, geles, etc., se obtuvieron de las compañías Sigma Chemical (St. Louis, Missouri, EUA), Merck (Monterrey, México), Aldrich Chemical Company (Milwaukee, WI, EUA), DIFCO Laboratories (Detroit, MI, EUA).

4.1.2 Material Biológico

La muestra de sangre de mono rhesus fue obtenida del Zoológico Guadalajara y del Departamento de Farmacología y Toxicología de la Facultad de Medicina de la UANL. Los cinco plásmidos cuyos insertos contienen las regiones intergénicas del mono araña

habían sido previamente obtenidos en este laboratorio (Esquivel Escobedo, 1999). Nueve Cromosomas Artificiales Bacterianos (BACs) cuyos insertos incluyen al locus GH/CSH completo o parte de este, del mono rhesus fueron obtenidos por Irám Rodríguez con el apoyo del Dr. Pieter de Jong del Children Oakland Hospital Research Institute, en Oakland California, EUA.

4.2 Equipo

Se utilizaron los siguientes equipos: un horno de microondas LG modelo MS-70IC, microcentrífugas Eppendorf modelos 5402 y 5415 (Brinkmann Instruments, Inc. NY EUA), fuentes de poder ISCO 494 (ISCO, Nebraska, EUA), EC 600 (E-C Aparatus Corporation, St Petersburg, Florida) y BRL 250 (GIBCO-BRL), centrífuga Beckman allegra TM 21R modelo 367570, centrífuga Beckman modelo J2-M1, un agitador de tubos Barnstead Termolyne modelo Maxi Mix (Bubuque, IOWA, EUA), una placa agitadora con control de temperatura marca Corning modelo PC320 (N.Y. EUA), un secuenciador automático dNA sequencer Long Reader 4200 (LI-COR, CA, EUA), un secuenciador automático ABI Prisma (Perkin Elmer, Cetus, Norwalk, CT), un congelador So Low Enviromental (Cincinnati, Ohio, EUA), un termomixer Eppendorf (Westbury, NY, EUA), un baño de agua 188 (Precision, Chicago, EUA), una placa de agitación modelo 3520, una incubadora Shell Lab modelo 1525 y un baño de agua giratorio (New Brunswick Scientific Co. inc. Edison, N.J. EUA). El seguimiento del crecimiento bacteriano y la cuantificación del DNA se realizó con un espectrofotómetro de la marca Eppendorf modelo BioPhotometer (Westbury, NY, EUA). El pulso eléctrico para la transformación de las bacterias se hizo en un electroporador y celdas de 0.2 mm de la marca BIO-RAD (Hercules, CA, EUA) modelo Gene Pulser II. Los reactivos se almacenaron en un congelador marca TorRey (Mty, Mex). Los termocicladores utilizados fueron Perkin

Elmer 2400 (Foster City, CA, EUA) y RoboCycler 96 (Stratagene, La Jolla, CA, EUA).

La balanza analítica fue de la marca OHAUS modelo AP110S y granataria de la marca Sartorius modelo 1206MP (Göttingen, Alemania).

El procesamiento de los datos se realizó en computadoras IBM modelo NetVista, Macintosh modelo iBook G3 y SUN modelo Ultra 30. Los programas computacionales utilizados fueron el procesador de textos de Microsoft Word para Mac versión 2004 (©1983-2004 Microsoft Corporation) y el procesador de gráficos Microsoft Power Point para Mac versión 2004 (©1983-2004 Microsoft Corporation).

Los programas de análisis de Biología Molecular utilizados fueron: Amplify versión 1.2b (Bill Engels 1992 University of Wisconsin Genetics M. Madison, WI, EUA), DNA Strider TM 1.3 (Ch Merk and C.E.A. 1999, Service de Biochimie-Département de Biologie, Institut de Recherche Fondamentale, CEA, France), OLIGO versión 4.0 (© 1992 Wojciech Rychlik, National Bioscience Inc., Plymouth, MN, EUA), CLUSTAL X versión 1.83 (Higgins et al., 1988), CONSED version 14.0 (University of Washington, Seattle, EUA) (Gordon et al., 1998), DNAMAN versión 5.2.2 Demo (©1994-2001 Lynnon Corporation) y MEGA 3.0 (Kumar et al., 2004). Para la cuantificación del DNA y la determinación de los tamaños se utilizó el software LabWorks versión 4.5.00.0 (©1996-2003, Media Cybernetics, Inc.) y el equipo UVP Mini Darkroom de Bioluminescence Systems (Upland, CA, EUA). Los programas utilizados vía Internet fueron VISTA (<http://gsd.lbl.gov/vista/index.shtml>), RepeatMasker (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>), MEME (<http://meme.sdsc.edu/meme/website/meme.html>) y BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>).

4.3 Métodos

4.3.1 Extracción de DNA genómico

La extracción del DNA genómico se realizó a partir de sangre periférica anticoagulada con EDTA por la técnica del fenol-cloroformo (Sambrook et al., 2001).

A 1 ml de sangre se le agregaron 400 µl de buffer de lisis TSNT (2% tritón 100X, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl pH 8 y 1 mM EDTA disódico) directamente mezclando ligeramente hasta lograr la lisis total. Los ácidos nucleicos se extrajeron con 3 ml de fenol y 1 ml de SEVAG (cloroformo:alcohol isoamílico 29:1) y 2 ml de TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8), mezclando por inversión después de agregar cada uno de los reactivos y finalmente 5 min en el vortex. Posteriormente, el tubo se centrifugó por 15 min a 5,000 rpm. La fase acuosa se transfirió a un tubo cónico nuevo de 15 ml y el DNA se precipitó con 0.1 volúmen de acetato de sodio 3 M pH 5.3 y 2 volúmenes de etanol al 100%, mezclando por inversión e incubando durante 1 hora a -70°C. Después de centrifugar por 20 min a 5,000 rpm, la pastilla obtenida se lavó dos veces con un volumen de etanol al 70% y el tubo se centrifugó nuevamente. El sobrenadante se eliminó con una micropipeta para en seguida dejar secar al aire la pastilla, la cual se resuspendió finalmente en 200 µl de TE 1 X y se cuantificó en el espectrofotómetro y conservó a -20°C hasta su uso.

4.3.2 Amplificación de los genes y regiones intergénicas

Para la amplificación de los miembros de los loci GH/CSH de los monos rhesus y araña se utilizó una serie de iniciadores consenso (cuadro 1), ubicados en regiones sobrepuestas entre los genes y regiones intergénicas (Figura 6) disponibles en nuestro laboratorio, a excepción de los iniciadores 5GH-2 y T-cam-1 que se diseñaron específicamente para este trabajo.

Cuadro 1. Oligonucleótidos utilizados en este trabajo

Nombre	Secuencia de 5' a 3'
D11	ggg tgc tcg agg act ggc tat c
5IP2	ctt ctc tgc gtt cag gtt gac ca
Pit OZ	gag gag ggg cta aat tat cca tta gca ca
MV5	cac ctg ggc agt tag aag cca c
139	ggc act aga gtg gca ctt tcc agg
403	ccg aat tcc agg aga ggc act ggg ga
AGN	tct gcc tct gca ctt aat agc c
D10	gaa gca gga gag cag ccc gta gtt
IGP IV	aac gta gac ccg cac cgc atc tcc ctg c
5 GH-2	ggc ttc gag ggg tga ggg aga gaa act g
5Ed1	aac aga agt gga gag gtg aga gga gac act
Prim	gac tag ttc ggg gag ttg ggc ctt ggg at
Tcam-1	gca aac aga gct ggc tgg aag gga tg

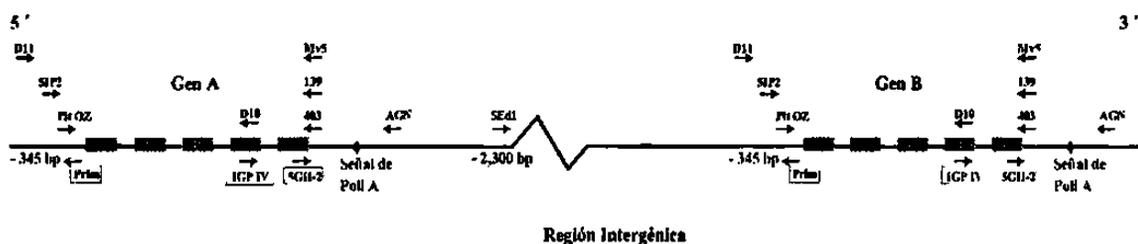


Figura 6. Iniciadores usados en los ensayos de PCR. Las cajas indican la estructura exón-intrón típica de los genes relacionados a GH. Con los iniciadores Prim, IGP IV y 5GH-2 se amplificaron las regiones intergénicas del mono rhesus, mientras que los genes con las diferentes combinaciones del resto de los oligos. La posición del oligo T-cam-1 se muestra en la figura 7 C.

4.3.2.1 Condiciones de amplificación de los genes del mono rhesus

Se realizaron múltiples ensayos a fin de optimizar las condiciones de PCR a las cuales amplificaron los genes GH/CSH de esta especie, quedando como se indican en el cuadro 2. Después de un paso inicial de desnaturalización a 94 °C por 5 minutos, las amplificaciones se llevaron a cabo por 30 ciclos, los cuales incluyeron un paso inicial de desnaturalización a 94 °C por 1 minuto, un segundo paso de apareamiento de 55 °C por 1 minuto y un tercer paso de extensión a 72 °C por un minuto. Terminados los 30 ciclos, se dio un paso de extensión a 72 °C durante 5 minutos. El producto amplificado en

cuestión se visualizó en un gel de agarosa al 1% y fue cortado con diferentes enzimas de restricción para corroborar su identidad.

Cuadro 2. Condiciones de amplificación de los genes GH/CSH del mono rhesus

Componente de la PCR	Concentración final
Buffer 10 X	1 X
MgCl ₂ 25 mM	1.5 mM
dNTPs 10 mM	200 μM
Oligo sentido 5 μM	0.3 μM
Oligo anti-sentido 5 μM	0.3 μM
Taq polimerasa 5 U/μl	1 U/50 μl
DNAg ó BAC	300 ng ó 10 ng
Agua Mili Q	cbp 50 μl

4.3.2.2 Condiciones de amplificación de los genes del mono araña

Al inicio de este trabajo se conocía la secuencia completa de cuatro genes GH y la parcial (los extremos) de cinco regiones intergénicas, material suficiente para proponer una organización hipotética, aunque incompleta del locus GH de esta especie de MNM (Figura 7 A).

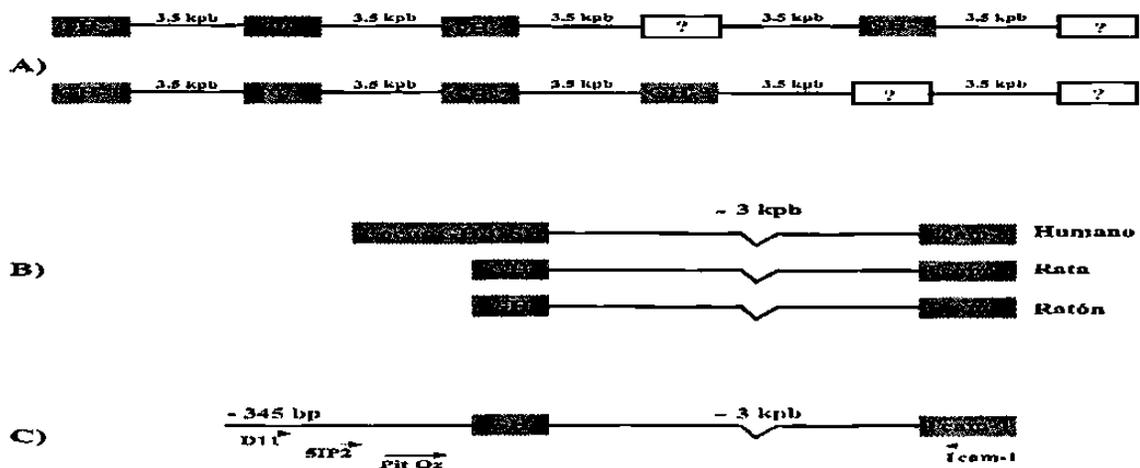


Figura 7. Estrategia de amplificación de los genes del mono araña. A) Posibles organizaciones del locus GH del mono araña. B) Representación del gen TCAM-1 en las diferentes especies. C) Posición de los oligos utilizados en la PCR.

Además se encontró que en el humano (AB026156), rata (NM_021673) y ratón (NM_029467), al extremo 3' del los genes GH se encontraba el gen de la Molécula de Adhesión Testicular 1 (TCAM-1), (ver figura 7 B), (Sakatani et al., 2000). Debido a que con los iniciadores previamente descritos no era posible amplificar el o los genes faltantes, se procedió a identificar las secuencias conservadas en este gen para las tres especies y diseñar un oligo anti sentido TCAM-1 (Cuadro 1). Los ensayos de PCR se practicaron de manera similar como se describió en la sección 4.3.2.3, usando las combinaciones con los oligos sentido: D11, 5IP2 y Pit OZ (Figura 7 C).

4.3.2.3 Condiciones de amplificación de las regiones intergénicas del mono rhesus

La estandarización del PCR largo se realizó variando las concentraciones de DNAG o BAC, Mg⁺⁺, temperatura de alineamiento de los iniciadores y los tiempos de extensión, logrando así establecer las condiciones de reacción aquí descritas. El cuadro 3 se describe la composición de las dos mezclas de reacción. Arriba de la mezcla inferior se depositó una perla de cera a cada tubo, para dar el “Hot Start” (5 minutos a 94 °C) y se agregó la mezcla superior. El programa de amplificación se detalla en el cuadro 5. Los productos amplificados junto con un marcador de peso molecular se corrieron en un gel de agarosa al 0.8% para estimar su tamaño, posteriormente se caracterizaron con diferentes enzimas de restricción para identificar los productos coamplificados.

Cuadro 3. Composición de las mezclas de reacción del PCR largo

Componente de la PCR	Concentración final	Componente de la PCR	Concentración final
Buffer II XL 3.3 X	1 X	Buffer II XL 3.3	1 X
dNTPs 10 mM	200 µM	rTth polimerasa 2 U/ µl	1 U
Oligo sentido 5 µM	0.4 µM	DNAG ó BAC	100 ng ó 10 ng
Oligo anti-sentido 5 µM	0.4 µM	Agua Mili Q	cpb 30 µl
Mg(OAc) ₂ 25 mM	1.5 mM		
Agua Mili Q	cbp 20 µl		
Mezcla inferior		Mezcla superior	

Cuadro 4. Programa de amplificación del PCR largo

1. Desnaturalización previa	80 ⁰ C por 5 minutos "Hot Start"
2. Desnaturalización inicial	94 ⁰ C por 5 minutos
3. 12 ciclos	94 ⁰ C por 30 segundos 68 ⁰ C por 10 minutos
4. 16 ciclos	94 ⁰ C por 30 segundos 68 ⁰ C por 10 minutos con incrementos de 15 segundos cada ciclo
5. Extensión prolongada	72 ⁰ C por 15 minutos

4.3.3 Clonación de los genes y regiones intergénicas

Los productos amplificados de los dos monos se clonaron el vector comercial pTOPO-XL-3.5., de acuerdo a las condiciones del fabricante. Previo a la ligación, a los productos amplificados con la enzima DNA polimerasa *rTth* se les adicionó un residuo de dATP a sus extremos. La purificación y ligación de los amplicones se realizó siguiendo las instrucciones del kit. Las condiciones de electroporación fueron: 200 Ω , 25 μ F y 2.5 kV. Después del pulso, las bacterias se pasaron a medio SOC (incluido en el kit), se llevaron al termomixer a 37⁰C con agitación (800 rpm) durante una hora. 50 μ l de la suspensión bacteriana se sembraron en placa en LB suplementado con kanamicina (50 mg/ml), la cual se incubó a 37⁰C durante 16 horas. Pasado este tiempo, se pasaron las colonias a tubos con LB igualmente suplementado con kanamicina y se pusieron a crecer a 37⁰C en agitación constante durante 18 horas. Finalmente, se procedió a la extracción del DNA plasmídico.

4.3.3.1 Extracción del DNA plasmídico

Se centrifugaron 3 ml del medio durante 5 min a 5,000 rpm y se removió el medio por aspiración, manteniendo el tubo en hielo. El DNA plasmídico se extrajo por el método de la lisis alcalina (Sambrook et al., 2001). Las células se resuspendieron en 200 μ l de solución I (glucosa a 50 mM, Tris-HCl a 25 mM pH 8, EDTA a 10 mM pH 8 y RNAasa

a1 mg/ml) con el vortex. Se agregaron 200 µl de solución II (0.2 ml de NaOH a10 N, 1.0 ml de SDS a 10% y 8.8 ml de agua) recién preparada. La solución se mezcló por inversión y se incubó en hielo por 5 min. Pasado este tiempo se agregaron 200 µl de solución III (acetato de amonio a 7.5 M), se mezcló por inversión y se incubó por 5 min en hielo. El tubo se centrifugó a 14,000 rpm por 10 minutos, se recuperó el sobrenadante en un tubo nuevo donde se le hizo una extracción fenol/Sevag (½ volumen de fenol y ½ volumen de Sevag). Se centrifugó a 14,000 rpm por 2 min, el sobrenadante se transfirió un tubo donde el DNA plasmídico se precipitó con 2 volúmenes de etanol al 100% frío. Después de centrifugar a 14,000 rpm por 10 min, la pastilla se lavó dos veces con 1 ml de etanol al 70% frío y se dejó secar a temperatura ambiente durante 30 minutos. Finalmente, se resuspendió en 50 µl de TE.

4.3.3.2 Selección de las clonas recombinantes

Los plásmidos se corrieron en geles de agarosa al 0.8%. En el caso de los genes, las clonas se digirieron con *Eco* RI que libera el inserto. La digestión se corrió junto con el amplicón que había originado la clona y los insertos que migraron junto con el producto amplificado se secuenciaron. Las regiones intergénicas se caracterizaron con las enzimas *Eco* RI + *Pst* I y aquellas que mostraron patrón diferente de corte se secuenciaron.

4.3.4 Secuenciación de los genes y regiones intergénicas

La secuenciación de los genes se realizo mediante la secuenciación clásica en geles de poliacrilamida, mientras que las regiones intergénicas se secuenciaron en un secuenciador automatizado de capilares, como a continuación se describe.

4.3.4.1 Secuenciación de los genes de los monos rhesus y araña

La secuencia se determinó por el método de la secuenciación cíclica adaptado a partir del enzimático de Sanger (Sanger et al., 1977). Los iniciadores que se utilizaron fueron el M13 reverso y forward que se aparean en el plásmido pTOPO XL 3.5 y los iniciadores internos Pit Oz y D10 (Cuadro 1). La secuenciación enzimática se realizó con el Kit de LiCor. Los fragmentos sintetizados se marcaron al incorporar el análogo de nucleótido trifosfatado portador de un cromógeno, y se resolvieron en un gel de poliacrilamida al 3.7 % con urea al 50 %, utilizando el secuenciador LiCor DNA 2000.

4.3.4.2 Secuenciación de las regiones intergénicas del mono araña

El tamaño de los insertos se encontró en ~ 3.5 kpb, resultando demasiado grande para ser secuenciados de manera clásica, pero muy pequeños para secuenciarse por el método automatizado de escopetazo (shot-gun), por lo que se decidió hacer deleciones a los insertos en los plásmidos. Se digirieron 100 ng de cada plásmidos con las diferentes enzimas del sitio de policlonación del vector pBS-SK+ en un volumen de 20 μ l de acuerdo a las instrucciones del proveedor. 17 μ l de las digestiones se corrieron en un gel de agarosa al 1%; el resto de la digestión se ligó a 16 $^{\circ}$ C durante 12 horas de acuerdo a las condiciones del fabricante. Los productos de las ligaciones se introdujeron en bacterias *E. coli* como se describió en la sección 4.3.3, solo que el marcador de selección fue ampicilina a 40 mg/ml. La secuencia se determinó con los iniciadores M13 reverso y forward. La secuenciación enzimática se realizó con el kit Perkin Elmer (Cetus, Norwalk, CT, EUA) utilizando el secuenciador ABI Prisma Automatic Sequencer, siguiendo el mismo protocolo de secuenciación que se utilizó para los bancos de las regiones intergénicas del mono rhesus, que a continuación se describe.

4.3.4.3 Secuenciación de las regiones intergénicas del mono rhesus

Para este tipo de insertos se siguió la técnica de secuenciación directa a gran escala, mejor conocida por su nombre en inglés como “shot gun”. 5 mg de cada plásmido contenido en 1 ml de solución (25% glicerol en TE) se fragmentaron por nebulización a fin de tener fragmentos al azar de todos los tamaños posibles del plasmido en cuestión. Para esto se aplicó un flujo de N₂ a una presión de 14 psi durante 30 segundos para el plásmido cuyo inserto era de 6 kpb y de 45 segundos para los dos plásmidos con inserto de 13 kpb. El DNA fragmentado se precipitó agregando 1.5 volúmenes de etanol al 100% y 0.3 volúmenes de acetato de sodio al 2.5 M. La precipitación se favoreció incubando a -70 °C durante 1 hora y centrifugando a 14,000 rpm durante 30 minutos. La pastilla se lavó dos veces con 500 µl de etanol al 70%, se secó y resuspendió en 25 µl de agua grado HPLC. Los extremos asumidos heterogéneos, se repararon con una mezcla del fragmento Klenow de la DNA polimerasa I y la DNA polimerasa del bacteriofago T4, a fin de asumir que fueran romos. El DNA reparado se corrió en un gel de agarosa de bajo punto de fusión al 1% utilizando buffer TAE 1 X.

Para cada plásmido se purificaron dos regiones del gel correspondiente a los tamaños de 1 - 2 kpb y la otra de 2 - 4 kpb. Esto a fin de asegurarse de tener una mayor cobertura de los insertos al momento de hacer los ensamblés. De esta manera se generaron seis bancos, tres de cada tamaño. La purificación del DNA a partir de las fragmentaciones se realizó mediante columnas QIAGEN (Valencia, CA, EUA), siguiendo las instrucciones del kit. El DNA purificado junto con el vector pZero de Invitrogen, previamente linearizado con la enzima de restricción *Eco* RV, se cuantificaron al espectrofotómetro (A₂₆₀), para calcular una relación vector:inserto de 10:120 ng y someter éstos a ligación a 16 °C por 16 horas de acuerdo a las instrucciones del proveedor de la enzima. La

ligación se precipitó con 400 µl de butanol, lavando la pastilla dos veces con etanol al 100% y resuspendiéndola en 5 µl de agua grado HPLC.

La electroporación de las bacterias *E. coli* TOP 10 “One Shot” de Invitrogen se hizo con 1 µl de la ligación. Después del pulso se transfirieron las bacterias a 950 µl de medio SOC y se incubaron a 37 ° C por 5 minutos. La siembra se hizo en cajas de Petri con medio LB + kanamicina con 100 µl de bacterias, mismas que se incubaron a 37 ° C por 12 horas. Se recuperaron 96 colonias para cada uno de los dos tipos de bancos (1 - 2 kpb y 2 - 4 kpb) y se pusieron a crecer en 1 ml de medio SOB + kanamicina 40 µg/ml a 37 ° C por 22 horas, en agitación constante (300 rpm). La extracción del DNA plasmídico se practicó mediante la técnica de lisis alcalina (Sambrook, 2001), seguido de una purificación mediante filtración en columna de Millipore (Billerica, MA, EUA). La precipitación de los plásmidos se hizo con 150 µl de isopropanol, lavando la pastilla dos veces con 300 µl de etanol al 70% y resuspendiéndola en 38 µl de agua grado HPLC.

Se corrieron 3 µl de cada uno de los 576 plásmidos en geles de agarosa al 1% para evaluar la cantidad y calidad del DNA plasmidico. Una vez que se terminó la construcción de los seis bancos (dos por cada plásmido,), se procedió a la secuenciación de acuerdo al siguiente protocolo (mismo que se usó para las intergénicas del mono araña):

- a) Se transfirió 1 µl de DNA a cada uno de los 96 pozos de la placa de microtitulación
- b) Se adicionaron 5 µl a cada pozo de agua grado HPLC.
- c) Se preparó la mezcla de reacción para 1 placa de 96 pozos, que a continuación se describe:

Kit ET Terminador	384 μ l
Oligo T7 10 pmol	24 μ l
Agua miliQ	76 μ l

Se depositaron 4.25 μ l de mezcla de reacción en cada pozo.

- d) Enseguida se procedió con el protocolo de PCR cuyos 30 ciclos consistieron cada uno de 95 $^{\circ}$ C por 20 segundos, 50 $^{\circ}$ C por 15 segundos y 60 $^{\circ}$ C por 1 minuto.
- e) Terminada la PCR se procedió con la precipitación, para la cual se preparó la siguiente mezcla por placa de 96 pozos: 2.86 ml de etanol al 100% y 220 μ l de acetato de amonio al 7.5 M.

De la mezcla anterior se depositaron 27 μ l en cada pozo, para proceder como sigue:

- f) Se incubó por 45 minutos a temperatura ambiente, cubriendo de la luz.
- g) Se centrifugó por 30 minutos a 5,000 rpm.
- h) Se desechó el sobrenadante.
- i) Se lavó con 100 μ l de etanol al 70%
- j) Se volvió a centrifugar durante 15 minutos a 5,000 rpm
- k) Se desechó el sobrenadante y se puso a secar durante 3 minutos en el Savant
- l) Finalmente se resuspendió en 5 μ l de "loading solution" y se procedió a secuenciar.

Una vez que se obtuvieron las secuencias de los 6 bancos (3 de 1 - 2 kpb y 3 de 2 - 4 kpb), se procedió a ensamblar las secuencias individuales para reconstruir la secuencia de los insertos, con el programa CONSED diseñado para este fin (Gordon et al., 1998).

4.3.5 Análisis de las secuencias

Conforme se iban obteniendo las secuencias, éstas se mandaban comparar en el BLAST con las reportadas en el GenBank, para verificar la identidad de las mismas, así como para identificar las clonas nuevas de las repetidas.

4.3.5.1 Ensamble de los loci GH/CSH

Las secuencias de los genes y las regiones intergénicas fueron comparadas por ambos extremos en el programa CLUSTAL X versión 1.83 (Higgins et al., 1988), para determinar la organización de los mismos.

4.3.5.2 Análisis de las proteínas

Las secuencias aminoacídicas se dedujeron de las nucleotídicas y con las primeras se hizo el alineamiento en el programa DNAMAN para identificar los aminoácidos clave en la unión al receptor (Liu et al., 2001). Finalmente se construyó el árbol filogenético de las proteínas reportadas en este trabajo junto con las previamente descritas (Cuadro 5), por el método de Neighbor-Joining (Saitou et al., 1987) con un valor de Bootstrap de 1000 réplicas, en la selección correspondiente del DNAMAN.

Cuadro 5. Secuencias utilizadas en este estudio

Especie	No. de Acceso	Especie	No. de Acceso
Grandes monos:		Monos del Nuevo Mundo (MNM)	
Humano (<i>Homo sapiens</i>)		Mono araña (<i>Ateles geoffroyi</i>)	
GH/CSH locus	J03071	GH-1	AAL72286
Chimpancé (<i>Pan troglodytes</i>)		GH-2	AAL72287
Proyecto del genoma	NW_118509	GH-3	AY435434
GH-1	AAL72285	Intergénica 1	DQ002808
GH-2	AF374233	Intergénica 2	DQ002809
CSH-A	AAN84505	Intergénica 3	DQ002810
CSH-B	AAN84506	Intergénica 4	DQ002811
CSH-C	AAN84507	Intergénica 5	DQ002812
CSH-D	AAN84508	Marmoset (<i>Callithrix jachus</i>)	
Monos del Viejo Mundo (MVM)		GH-1	AJ297563
Mono rhesus (<i>Macaca mulatta</i>)		GH-5	AJ489810
GH-1	DQ002799	GH-6	AJ489811
GH-2	DQ002800	Mono ardilla (<i>Saimiri boliviensis</i>)	
CSH-1	DQ002801	GH	AF339060
CSH-2	DQ002802	Prosimios	
CSH-3	DQ002803	Gálago de Senegal (<i>Galago senegalensis</i>)	
CSH-4	DQ002804	GH	AF292938
Intergénica corta 1	DQ002805	Loris pigmeo (<i>Nycticebus pygmaeus</i>)	
Intergénica larga 1	DQ002806	GH	AJ297562
Intergénica larga 3	DQ002807	Grupo fuera	
GH-1*	AAA18842	Cerdo (<i>Sus scrofa</i>)	
GH-2*	AAA03391	GH	NP_999034
CSH-1*	AAA18839		
CSH-2*	AAA18840		
CSH-3*	AAA18841		

* Aquellos reportados previamente (Golos et al., 1993). En negritas se muestran las secuencias generadas en este trabajo.

4.3.5.3 Análisis del promotor proximal

El análisis se centró en identificar los sitios de unión a factores transcripcionales descritos como responsables de la transcripción basal del gen hGH-1 y se enfocó en la región que va del nucleótido -147 al exón I. Dentro de estos destacan Pit-1 proximal y distal, la caja TATA y los elementos SP1 e InrE. Las secuencias de cada uno de los promotores proximales se alineó en el programa DNAMAN.

4.3.5.4 Análisis del promotor distal

El análisis consistió en identificar a los dos elementos claves de la especificidad tisular: el inhibidor o elemento P que reprime la expresión de los genes placentarios en la hipófisis (ubicado 2.5 kpb río arriba de dichos genes) y el potenciador o “enhancer” que incrementa la expresión de los genes CSHs en la placenta (también ubicado a 2.5 kpb

pero río debajo de los genes CSHs). Las secuencias de cada uno de los promotores proximales se alineó en el programa DNAMAN.

4.3.5.5 Identificación de los elementos repetitivos

Esta búsqueda se hizo cargando las secuencias de las regiones intergénicas de los monos rhesus y araña en el servidor Web RepeatMasker.

4.3.5.6 Estimación de los índices evolutivos (K) en las regiones codificantes

Los índices evolutivos (aminoácido/sitio/año) de las proteínas relacionadas a GH del mono rhesus fueron determinados por el método Z de una cola implementado el software MEGA 3.0 (Kumar et al., 2004), usando 23 millones de años como el tiempo de divergencia (MYA) entre las familias *Cercopithecoidea* y *Hominoidea* así como 6 MYA entre el humano y el chimpancé (Stauffer et al., 2001).

4.3.5.7 Prueba de hipótesis al reloj molecular

Los pruebas de hipótesis al reloj molecular de los intrones y regiones intergénicas fueron determinados por los métodos de Tajima implementado el software MEGA 3.0 (Kumar et al., 2004) y el “Two Cluster Test” bajo los modelos de Kimura dos parámetros y Jukes Cantor implementados en el software LINTREE (Takezaki et al., 1995).

4.3.5.8 Pruebas de selección

Las substituciones sinónimas (d_N) y no sinónimas (d_S) fueron calculadas por el método N-G (Zhang et al., 1998). Se probaron las hipótesis de evolución neutral ($d_N = d_S$), evolución positiva ($d_N > d_S$) y purificación de la selección ($d_N < d_S$) por los métodos Z de una cola y “Fisher’s Exact Test”, implementados en el software MEGA 3.0 (Kumar et al., 2004).

4.3.5.9 Identificación de los eventos de conversión génica

El análisis estadístico para detectar los eventos de conversión de genes fue realizado en el software GENECONV 1.81 (Sawyer, 1989).

4.3.5.10 Estimación de los tiempos de duplicación

Los loci GH/CSH disponibles (tabla 5) y los deducidos en este trabajo, se alinearon en el programa CLUSTAL X versión 1.83 (Higgins et al., 1988). Los posibles bloques de duplicación se dedujeron por similitud de secuencia. De manera visual se contabilizó únicamente el número de diferencias (no inserciones y/o deleciones) entre el par de secuencias a comparar. Con el número de diferencias, se estimó el índice de evolución (fórmula 1) y con éste se estimó el tiempo de duplicación (fórmula 2).

$$1) \text{Índice}_{\text{sec 1/sec 2}} = \frac{\# \text{ de diferencias}}{2 \text{ (bp) tiempo}}$$

$$2) \text{Años} = \frac{\# \text{ de diferencias}}{2 \text{ (índice) (bp)}}$$

CAPITULO V

RESULTADOS

5.1 Extracción del DNAg de los monos rhesus y araña

El DNAg extraído de sangre periférica fresca de los monos resultó de buena calidad, como se deduce del perfil electroforético en la figura 8. Se obtuvieron aproximadamente 10 mg de DNA a partir de 1 ml de sangre de cada mono.

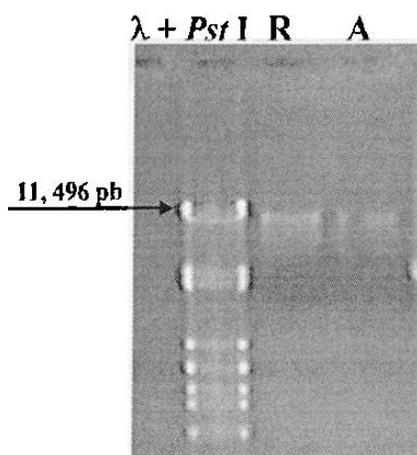


Figura 8. DNAg de los monos rhesus y araña. Se depositaron 100 ng de DNAg extraído de sangre periférica de cada mono, en los carriles R (rhesus) y A (araña), en un gel de agarosa al 1%.

5.2 Clonación y secuenciación de los genes de los loci GH de los monos rhesus y araña

A medida que se iban amplificando los miembros de los loci, estos se clonaban y secuenciaban como parte de su identificación.

5.2.1 Clonación y secuenciación de los genes GH del mono rhesus

Siguiendo las instrucciones descritas en la sección 4.3.2.1; se obtuvieron diferentes productos de PCR con los iniciadores para genes (Figura 6). Su caracterización

enzimática y secuenciación reveló que consistía en realidad de cinco genes. Las diferentes combinaciones de oligos no amplificaban preferentemente uno u otro gen. Por comodidad, se decidió mantener la identidad de los genes, como sigue: D11 – AGN fue el *CSH-1*, D11 – MV5 la *GH-2*, D11 – 403 el *CSH-3*, 5IP2 – D10 la *GH-1*, Pit OZ – 139 el *CSH-3* y 5Ed1 – MV5 el *CSH-4*, como se muestra en la figura 9.

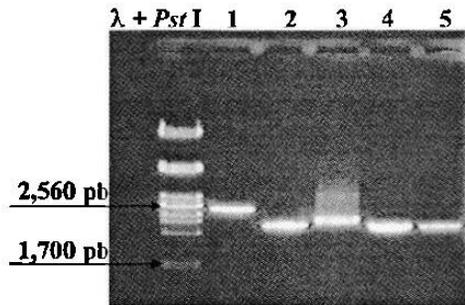


Figura 9. Amplificación de los genes del mono rhesus a partir de DNAg. Se muestra un gel típico de las PCRs exitosas con los diferentes juegos de iniciadores consenso. En el carril 1 se observa el producto amplificado con los oligos D11-AGN, en el 2 D11-MV5, en el 3 D11-403, en el 4 5IP2-D10 y en el 5 Pit Oz – 139.

Dado que al inicio del presente trabajo se tenía la sospecha de que seis genes y no cinco como encontramos empleando el DNAg podrían integrar el locus GH/CSH del mono rhesus (Golos et al., 1993), se intentó amplificar el sexto gen a partir de un BAC, utilizando los iniciadores 5Ed1-MV5. Se estandarizaron las condiciones de PCR y un nuevo producto fue amplificado, como se muestra en la figura 10. El producto amplificado se clonó y su identidad se comprobó por secuenciación, demostrando que no correspondía a ninguna de las cinco secuencias obtenidas anteriormente y por tanto seis genes fueron clonados y secuenciados.

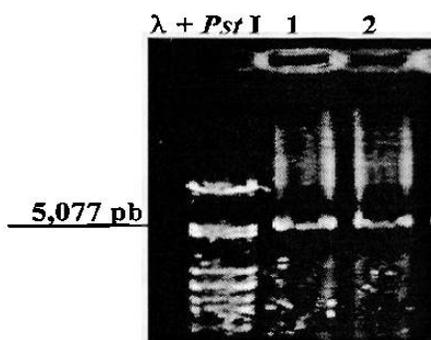


Figura 10. Amplificación del sexto gen del mono rhesus a partir de un BAC. El producto amplificado se resolvió en un gel de agarosa al 1%. En ambos carriles (1 y 2) se observa el producto amplificado con los oligos 5Ed1-MV5.

5.2.2 Amplificación de los genes GH del mono araña

A partir de DNAg y utilizando los diferentes juegos de iniciadores descritos en la figura 7C, se estandarizaron las condiciones de PCR para amplificar los dos genes faltantes, descritos en la figura 7A. Las condiciones descritas en la sección de material y métodos permitieron obtener diferentes amplicones. Idealmente se esperaba una sola banda de ~ 3 kpb, pero ésta no se logró amplificar, aún y cuando se ensayaron diferentes condiciones de la PCR, como se muestra en la figura 11.

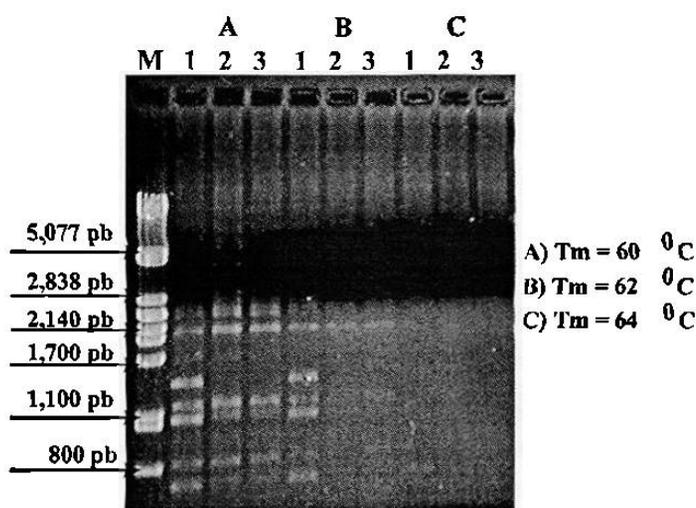


Figura 11. Amplificación de los genes del mono araña. A partir de DNAg y con los iniciadores descritos en la figura 7 C se realizaron los ensayos de amplificación. En los carriles 1 se muestran los productos amplificados con los oligos D11 – T-cam-1, en los 2 5IP2 – T-cam-1, en los 3 Pit Oz – T-cam-1. Los incisos A, B y C indican las Tm a las cuales las PCRs fueron hechas, respectivamente. El marcador (M) fue $\lambda + Pst I$.

De los múltiples amplicones obtenidos con los diferentes iniciadores sentido, más el TCAM-1 antisentido, se clonaron aquellos obtenidos con el juego 5IP2 – TCAM-1 a 60 °C. Las clonas se digirieron con *Eco* RI (Figura 12) y las que mostraron patrones diferentes de corte se secuenciaron (Figura 13).

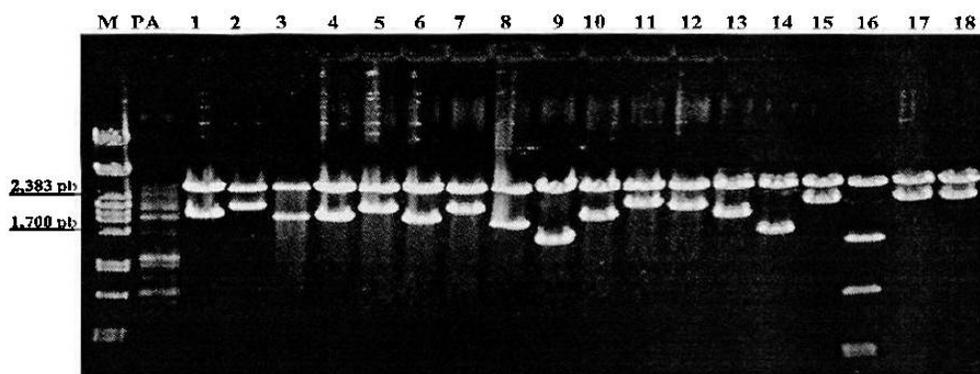


Figura 12. Caracterización de las clonas de los genes del mono araña. Se muestran 18 clonas al azar digeridas con *Eco* RI. La banda común a la altura de ~ 3.5 kpb corresponde al vector, el resto de las bandas a los insertos cuyos tamaños corresponden a alguna de las bandas del PCR que originó las clonas. Sólo la clona 16 tiene sitios *Eco* RI internos a juzgar por sus bandas. Una clona de cada patrón de corte se secuenció para verificar su identidad. El marcador de peso molecular (M) fue $\lambda + Pst$ I.

Sequences producing significant alignments:	Score	E
	(bits)	Value
gi 11415202 gb AC023105.7 Homo sapiens BAC clone GS1-13M19...	209	6e-51
gi 46488806 gb AY429544.1 Homo sapiens BMGCRv_4 mRNA seque...	201	2e-48
gi 40556308 gb AC009127.10 Homo sapiens chromosome 16 clon...	201	2e-48
gi 22655813 gb AC026634.11 Homo sapiens chromosome , clone...	197	3e-47
gi 21747438 gb AC011461.5 Homo sapiens chromosome 19 clone...	197	3e-47
gi 21263349 gb AC105383.3 Homo sapiens BAC clone RP11-9G1 ...	197	3e-47
gi 21211761 emb AL590080.25 Human DNA sequence from clone ...	197	3e-47
gi 19225048 gb AC010853.11 Homo sapiens chromosome 18, clo...	197	3e-47
gi 7768676 dbj AP001711.1 Homo sapiens genomic DNA, chromo...	197	3e-47
gi 37537368 dbj BS000101.1 Pan troglodytes chromosome 22 c...	197	3e-47
gi 4760435 gb AC006441.13 AC006441 Homo sapiens chromosome ...	197	3e-47

Figura 13. Resultados de la secuenciación de los genes del mono araña. Cada una de las secuencias obtenidas por ambos extremos de las clonas seleccionadas para su secuenciación se enviaron a comparar al GenBank mediante el programa BLAST. Se muestra el resultado de dicha comparación de una de las secuencias. El resultado fue similar para el resto. De todas las clonas secuenciadas ninguna resultó tener inserto relacionado con algún miembro de la familia GH. A la fecha se desconoce la causa del porque no fue posible amplificar algún miembro relacionado a la familia GH.

5.3 Amplificación, clonación y secuenciación de las regiones intergénicas de los loci GH de los monos rhesus y araña

A medida que se iban amplificando los miembros de los loci, estos se clonaban y secuenciaban como parte de su identificación.

5.3.1 Amplificación, clonación y secuenciación de las regiones intergénicas del locus GH del mono rhesus

Dado que se clonaron seis genes, forzosamente cinco regiones intergénicas debían ser amplificadas. Por medio del PCR largo, con los oligos Prim – IGP IV y utilizando DNAg como templado, se obtuvieron dos largas y una corta (Figura 14 A), mientras que a partir de los BAC, se amplificó la segunda intergénica corta. (Figura 14 B). Adicionalmente, con los BAC, pero con el juego de oligos 5GH-2 – Prim se amplifico la tercera intergénica larga (Figura 14 C).

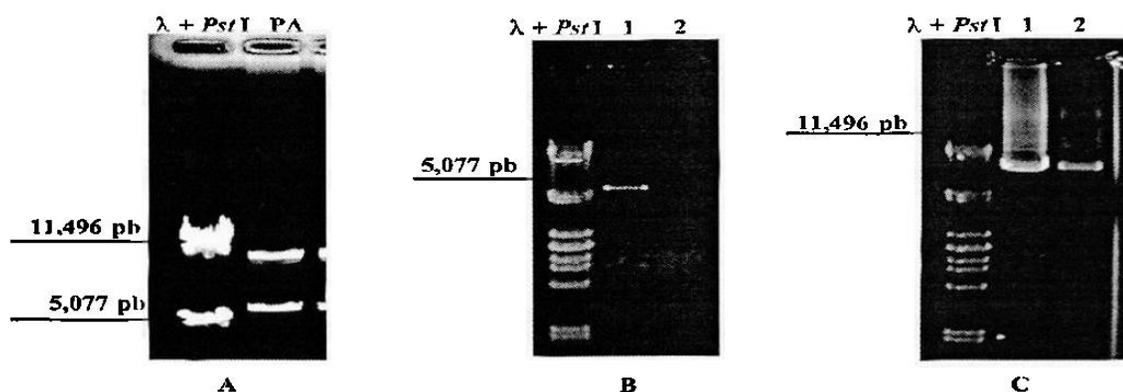


Figura 14. Amplificación de las regiones intergénicas del mono rhesus. Los productos de amplificación se resolvieron en geles de agarosa al 0.8%. A) Se observan los productos amplificados (PA) a partir de DNAg y con los oligos Prim – IGP IV. B) Se muestra el producto de amplificación de la segunda intergénica corta, con los oligos Prim – IGP IV. En el carril dos, se muestra el testigo negativo de PCR. C) Producto amplificado correspondiente a la tercera intergénica larga con los oligos 5GH-2 – Prim, en los dos carriles.

Debido al tamaño de los insertos de este tipo de clonas, se recurrió a la técnica de secuenciación a gran escala, como se describió en la subsección 4.3.4.3 de la sección

Material y Métodos. De las cinco regiones intergénicas de este mono, sólo se secuenciaron en su totalidad tres, dos largas (~ 14 kpb) y una corta (~ 6 kpb). Solo se secuenciaron los extremos de las otras dos regiones intergénicas. En la figura 15 se ejemplifica la construcción de uno de los seis bancos que se hicieron para secuenciar estos tres plásmidos.

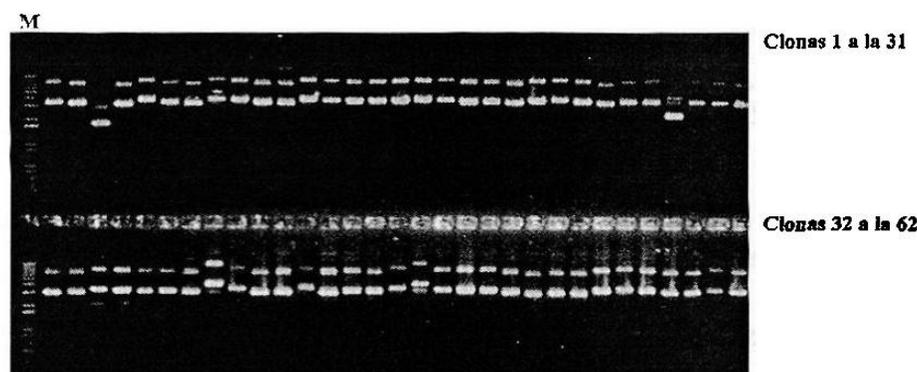


Figura 15. Banco de 2 – 4 kpb de una intergénica del mono rhesus. Gel de agarosa al 1% de 62 subclonas utilizadas para la secuenciación a gran escala. El marcador de peso molecular (M) fue una escalera de 1 kpb. Se corrió el gel para estimar la calidad y cantidad de DNA.

Se levantaron tantas clonas como cuantas fueron necesarias para reconstruir la secuencia original del plásmido que las generó. Las secuencias individuales (al menos tres de cada una) se alimentaron al programa CONSED y éste dio la secuencia completa de cada plasmido.

5.3.2 Subclonación y secuenciación de las regiones intergénicas del mono araña

En este caso ya se contaba en el laboratorio con cinco plásmidos cuyos insertos correspondían a las intergénicas de este mono (Esquivel-Escobedo, 1999) llamados arbitrariamente como E1, E5, E6, E10 y E11. Por lo que se procedió a hacerles hicieron deleciones con las enzimas del sitio de policlonación del vector pBS-SK+, para dividir en fragmentos que facilitarían su secuenciación, como se muestra en la figura 16.

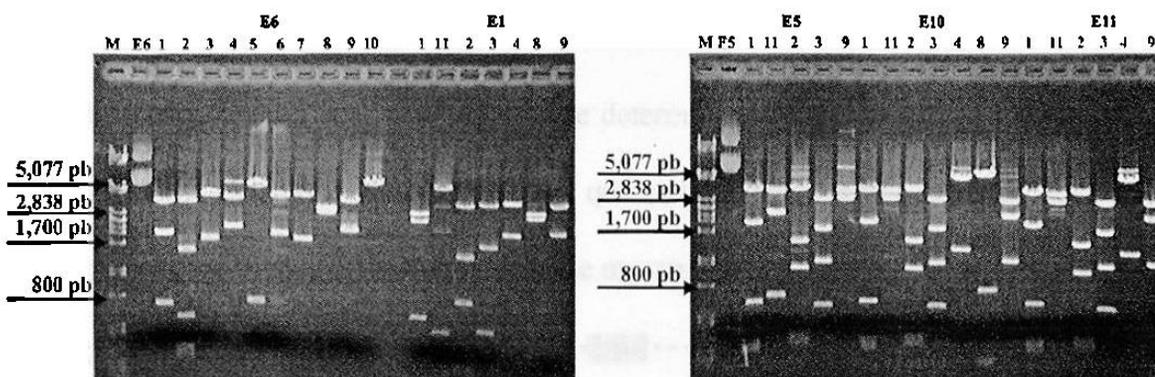


Figura 16. Caracterización de las intergencias del mono araña. El número en los carriles indica la enzima con la cual fueron cortados los diferentes plásmidos. 1 = *Pst* I, 2 = *Sac* I, 3 = *Bam* HI, 4 = *Xba* I, 5 = *Sma* I, 6 = *Acc* I, 7 = *Kpn* I, 8 = *Hind* III, 9 = *Hinc* II, 10 = *Xho* I y 11 = *Eco* RI. El marcador de peso molecular (M) utilizado fue $\lambda + Pst$ I.

Previo a correr el gel se guardó una alícuota de cada digestión. Ésta se ligó nuevamente y se introdujo en bacterias por electroporación. Se hizo la selección de las clonas deletadas y éstas se secuenciaron. La reconstrucción de la secuencia original de los insertos se hizo de manera manual, utilizando los sitios de restricción como marcadores. Se obtuvo la secuencia íntegra de los cinco plásmidos, mismas que se ingresaron al GenBank, siendo los números de acceso los que se indican en el cuadro 5 de Material y Métodos.

5.4 Análisis de las secuencias

Al final de la etapa de clonación y secuenciación de los miembros de los loci GH/CSH de ambos monos se obtuvieron las secuencias que se describen en el cuadro 6.

Cuadro 6. Secuencias generadas en este trabajo.

Inserto	Rhesus	Araña
Genes	Seis	Cuatro
Intergénicas	Cinco	Cinco