

UNIVERSIDAD AUTONOMA DE BAJA CALIFORNIA

FACULTAD DE CIENCIAS

**PATRONES DE EVOLUCION MOLECULAR EN SECUENCIAS DE EXONES E
INTRONES EN GENES DE PROTEINAS ACROSOMALES EN EL
ESPERMATOZOIDE DE ABULON (*Haliotis*)**

TESIS PROFESIONAL

QUE PRESENTA

REFUGIO ROBLES SIKISAKA

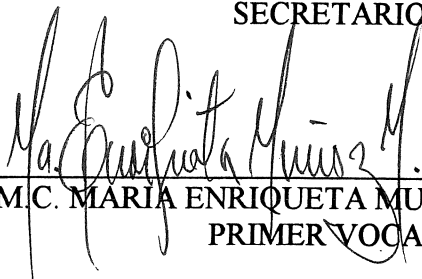
APROBADO POR:



DR. CARLOS MÁRQUEZ BECERRA
PRESIDENTE DEL JURADO



DRA. MEREDITH GOULD CHAMBERS
SECRETARIO



M.C. MARIA ENRIQUETA MUÑOZ MARQUEZ
PRIMER VOCAL

DEDICATORIA

A Dios,

por darme la oportunidad de vivir
y permanecer siempre a mi lado.

A mis Padres,

Rosalinda y Guillermo,
Por apoyarme incondicionalmente en todo momento
Gracias por su enorme esfuerzo y dedicación con
los que me han ayudado a alcanzar este escalón en mi vida.

A mis hermanos,

Marcos, Noemí, Abraham y Midori
A Lorena y mis sobrinas Lorie y Jubilee
Por estar siempre presentes, por compartir los momentos mas
importantes para mi.

A mis amigos y compañeros de clases,

Araceli, Paco, Gerardo y Annouck,
Que con su alegría y amistad a lo largo de toda la carrera,
hicieron que el aprender a veces no fuera tan difícil.

A mis amigos,

que aunque a veces lejos, siempre me acompañaron
Alejandra, Norma, Berenice, Jesús S. , Isela,
Jenny, Liza, Gabriel y Francisca, Silvia y Eddie M.
mil gracias por enseñarme
el valor de la amistad.

AGRADECIMIENTOS

A los Profesores Meredith Gould y José Luis Stephano,
Por brindarme su ayuda y conocimientos
por inculcarme la importancia del trabajo para lograr mis metas.

Al Dr. Edward C. Metz,
por ser más que un profesor un amigo,
Que tuvo la paciencia e interés para enseñarme
cómo desempeñar mi trabajo en laboratorio.

Al Dr. Victor D. Vacquier,
por aceptarme y permitir ser parte de su
laboratorio a lo largo de la realización de este trabajo.

Al Profesor Carlos Márquez,
Por su accesibilidad, interés y colaboración
Para que este trabajo se realizará. Mil gracias.

A mis amigos y compañeros de laboratorio,
Por compartieron su tiempo, conocimientos y amistad
durante la realización de este trabajo, Willie, Kathryn, Joseph,
Johnny, Jun, Priya, Dirk, Elena, Mariah, y Anna.

A la Facultad de Ciencias de la UABC,
Todos sus directivos, docentes y empleados
por su apoyo durante mis estudios.

INDICE DE CONTENIDO

	Página
INTRODUCCION	1
OBJETIVOS	11
ANTECEDENTES	
a) Ensayos de Especificidad de Especie en Lisina y 18 kDa	12
b) Estudios de ADN codificante de lisina y 18 kDa	12
c) Análisis de Secuencias	14
d) Otras Comparaciones de Secuencias de Exones e Intrones	17
METODOS	
I) Extracción de ADN genómico y Preparación del Template para PCR	19
II) Reacción en Cadena de la Polimerasa (PCR)	20
IIa) Preparación de la Mezcla para PCR	21
IIb) Condiciones de PCR	22
III) Detección y Aislamiento de los Productos de PCR	24
IIIa) Electroforesis en Gel de Agarosa al 1%	24
IIIb) Detección de Bandas y Purificación de ADN	24
IIIc) Cuantificación de ADN	25
IV) Reamplificación de ADN	25
V) Diseño de "Primers" Específicos para Intrones	26
VI) Secuenciación Automática de ADN	26
VII) Análisis de Secuencias	27
RESULTADOS	
I) Productos de PCR	29
II) Posiciones de Intrones	29
III) Secuencia de Intrones	33
IIIa) Alineación de Secuencias de Intrones	33
IV) Análisis de las Secuencias de Intrones	40

IVa) Distancias Genéticas	40
IVb) Ventanas Deslizantes	41
IVc) Marcos de Lectura Abiertos y Secuencias de Repeticiones Invertidas	47
IVd) Inserciones y Deleciones	47
DISCUSION	
I) Posiciones de Intrones	52
II) Evolución Molecular en Exones e Intrones	53
CONCLUSIONES	58
COMENTARIOS FINALES	59
BIBLIOGRAFIA	60

INDICE DE FIGURAS

Página

FIGURAS

Figura 1: Localización de las proteínas lisina y 18 kDa en el espermatozoide de abulón.....	3
Figura 2: Alineación de la secuencia de aminoácidos de lisina y 18 kDa en abulón rojo.....	6
Figura 3: Ensayo de disolución de capas vitelinas específica de especie.....	13
Figura 4: Esquema del ADNc de lisina mostrando regiones conservadas.....	15
Figura 5: Esquema de la región codificante de lisina con posición de oligonucleótidos empleados para multiplicar secuencias de intrones.....	23
Figura 6: Gel de agarosa que muestra productos de PCR.....	30
Figura 7: Alineación de la secuencia de aminoácidos de lisina y 18 kDa incluyendo posiciones de intrones.....	32
Figura 8: Posición y nombre de oligonucleótidos intron-específicos para especies de abulón rojo, rosa y azul.....	34
Figura 9: Alineación de la secuencia parcial del intron 1 utilizada para el cálculo de distancias genéticas.....	36
Figura 10: Alineación de las secuencias del exon 4 e intron 4 utilizada para el cálculo de distancias genéticas.....	37
Figura 11a: Ventana deslizante para el análisis de variación de nucleótidos en el intron 1 en las tres especies.....	43
Figura 11b: Ventana deslizante para el análisis de variación de nucleótidos en el intron 1 en las tres especies.....	44
Figura 12a: Ventana deslizante para el análisis de variación de nucleótidos en el exon 4 e intron 4 en las tres especies.....	45

Figura 12b: Ventana deslizable para el análisis de variación de nucleótidos en el exon 4 e intron 4 en las tres especies	46
Figura 13a: Resultados de la búsqueda de ORFs en la secuencia del intrón 4 en abulón rojo.....	48
Figura 13b: Resultados de la búsqueda de ORFs en la secuencia del intrón 4 en abulón rosa.....	49
Figura 13c: Resultados de la búsqueda de ORFs en la secuencia del intron 4 en abulón azul.....	50

INDICE DE TABLAS

Tabla I: Comparación entre especies del porcentaje de identidad y proporción dn/ds de las proteínas de lisina y 18 kDa.....	16
Tabla II: Comparación de la variación de dn , ds , y dn/ds en exones y d en intrones en dos especies de roedores.....	18
Tabla III: Nombre y secuencia de bases de los oligonucleótidos intron-específicos.....	35
Tabla IV: Comparación entre especies de dn y ds en exones y d en los intrones 1 y 4 de los genes de lisina y 18 kDa...	54

INTRODUCCIÓN

Las proteínas de reconocimiento de gametos son moléculas ubicadas en las superficies del espermatozoide y del ovocito que median la unión específica de especie de ambos gametos durante el proceso de fertilización. Estas proteínas son un importante objeto de estudio debido a que pueden conducir al establecimiento del aislamiento reproductivo entre especies, ya que un cambio en dichas proteínas podría restringir la libre interacción de los gametos, generando así barreras para la reproducción. Como consecuencia de tal fenómeno se puede promover la formación de especies. (Palumbi, 1994; Metz, 1994; Lee *et al.* 1995).

Las evidencias de que las proteínas de gametos pueden estar involucradas en el mecanismo de especiación en el mar provienen de estudios realizados en abulón y erizo de mar, que son los únicos dos grupos en los cuales las secuencias del ácido desoxirribonucleico codificante (ADNc) son conocidas para la misma proteína de reconocimiento en varias especies congénicas (Lee y Vacquier, 1991; Metz y Palumbi, 1996). Las proteínas de reconocimiento de gametos que han sido estudiadas a nivel de secuencias de nucleótidos en invertebrados marinos son: la proteína bindin presente en el espermatozoide del erizo de mar (Vacquier *et al.* 1995) y la proteína lisina, presente en el espermatozoide de abulón (Lee y Vacquier, 1992; Lee *et al.* 1995).

El proceso de fertilización en abulón incluye varias etapas:

- 1) La reacción acrosomal del espermatozoide libera las proteínas ubicadas en una vesícula localizada en la parte anterior del espermatozoide llamada acrosoma; 2) las proteínas del acrosoma interaccionan con la capa externa del ovocito, también llamada

capa vitelina; 3) el espermatozoide penetra la capa vitelina; 4) las membranas plasmáticas de ambos gametos se fusionan. Las proteínas que median las etapas de este proceso son llamadas proteínas de fertilización. Estas proteínas son importantes ya que son responsables en gran medida del éxito reproductivo del organismo.

En el acrosoma del espermatozoide de abulón se encuentran dos proteínas involucradas en el proceso de fertilización, una llamada lisina que es una proteína de 16 kilodaltones (kDa) y la otra de 18 kDa (Fig.1). La proteína lisina, está localizada en la parte posterior del acrosoma y crea un orificio en la capa vitelina mediante un mecanismo no enzimático (Lewis *et al.* 1982). Dicho mecanismo rompe los enlaces de hidrógeno que mantienen unidas las moléculas de glucoproteínas filamentosas de la capa vitelina, creando así un orificio a través del cual el espermatozoide pasa hasta llegar a la membrana plasmática del ovocito. La interacción de lisina con la capa vitelina es un proceso específico de especie, es decir, el espermatozoide fertiliza a los ovocitos de la misma especie más efectivamente que a los ovocitos de diferente especie (Vacquier *et al.* 1990; Vacquier y Lee, 1993). Es este proceso en particular el que hace a esta proteína desempeñar un papel importante como proteína de reconocimiento de gametos en abulón. La proteína de 18 kDa, está localizada en la parte anterior de la vesícula acrosomal, ésta cubre la superficie del acrosoma y su función parece estar relacionada con la fusión de los gametos (Swanson y Vacquier, 1995b), aunque esta actividad no se ha probado que sea específica de especie (Vacquier *et al.* 1997).

La comparación de las tasas de sustitución de nucleótidos en diferentes regiones del ADN puede ayudar a entender los mecanismos de evolución molecular. Varios estudios han comparado la divergencia de la secuencia de los nucleótidos en genes

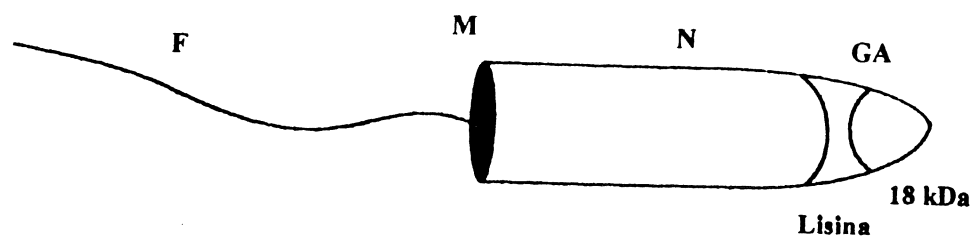


Fig. 1: Localización de las proteínas lisina y 18 kDa en el espermatozoide de abulón. Gránulo acrosomal (GA), núcleo (N), mitocondria (M) y flagelo (F). Adaptado de Vacquier y Lee (1993).

codificantes de proteínas acrosomales en espermatozoides de invertebrados marinos que desovan directamente en el agua de mar. Uno de estos estudios se ha realizado en la proteína bindin la cual es una proteína que media la unión específica de especie del espermatozoide al ovocito en erizo de mar (Moy y Vacquier, 1979). Los análisis de las secuencias del ADNc de bindin en diferentes especies han mostrado que ésta se ha diferenciado extensamente entre especies, aunque comparten un dominio idéntico (Vacquier *et al.* 1995). Las regiones que flanquean este dominio contienen elementos repetitivos y son extremadamente divergentes entre especies (Vacquier *et al.* 1995). Estos elementos repetitivos múltiples revelan numerosas inserciones y deleciones (indels) y rearrreglos cuando se comparan diferentes especies o géneros (Metz *et al.*; 1998). Metz y Palumbi (1996), encontraron en secuencias codificantes de bindin en tres especies de erizo del género *Echinometra* una región de 40 codones que mostraba un patrón inusual de variación de nucleótidos, en el cual la proporción de sustituciones de nucleótidos que alteran aminoácidos es mayor que la proporción de sustituciones que no alteran los aminoácidos de la proteína. Este patrón de variación sugiere la acción de selección positiva favoreciendo la diversificación de bindin. Sin embargo, Metz *et al.* (1998), no encontraron este patrón en las secuencias de bindin del género *Arbacia*.

En el análisis de las secuencias codificantes de lisina y 18 kDa de diferentes especies de abulón éstas han mostrado ser altamente divergentes cuando se comparan entre especies. Las proteínas 18 kDa son aproximadamente 2.5 veces más divergentes que las proteínas de lisina. La proporción de cambios de aminoácidos en 18 kDa puede ser la más alta descubierta en proteínas con secuencias completas de aminoácidos (Vacquier *et al.* 1997). Sin embargo, a pesar del alto grado de variación existe

conservación del tamaño y polaridad de los aminoácidos en ambas proteínas. Más aún, los análisis de la estructura secundaria de ambas proteínas deducida de su ADN codificante muestran que la posición y longitud de las hélices alfa de lisina y 18 kDa coinciden excelentemente cuando son comparados, demostrando una estructura secundaria similar (Fig. 2). Estos resultados apoyan la hipótesis de que estas proteínas acrosomales son homólogas y surgieron por duplicación de un gen ancestral (Vacquier *et al.* 1996).

La divergencia evolutiva del ADN codificante puede ser evaluada por las tasas de dos tipos de sustitución de nucleótidos o cambios de nucleótidos: Sustitución no sinónima por sitio no sinónimo (*dn*) y sustitución sinónima por sitio sinónimo (*ds*). En *dn* un nucleótido es sustituido por otro el cual provoca que el codón codifique un aminoácido diferente al original, es decir, es una sustitución que altera aminoácidos. En *ds* un nucleótido es sustituido por otro pero el codón continúa codificando el aminoácido original, este tipo de sustitución es también llamada sustitución silenciosa. Los valores obtenidos de dichos análisis pueden dar una indicación de la manera en la que la selección está actuando debido a que las sustituciones sinónimas tienen poco o ningún efecto en la proteína codificada, es por ello que éstas proveen una medida de evolución neutra contra la cual las sustituciones no sinónimas pueden ser comparadas (Hughes y Nei, 1988).

Cuando *ds* es mayor que *dn* la selección está actuando en contra de las mutaciones no sinónimas, en este caso, un cambio en la secuencia de nucleótidos es desventajoso para la función de la proteína por lo que hay una tendencia de conservar la secuencia de nucleótidos. Cuando la secuencia de una proteína presenta este tipo de

patrón se dice que se presenta una "limitación funcional", debido a que la función apropiada de la proteína depende de la conservación de la secuencia de nucleótidos original (Li y Graur, 1991).

Cuando dn es igual a ds , no hay selección en contra o en favor del cambio de aminoácidos de la secuencia de la proteína, en otras palabras, se presenta una evolución neutra. Ejemplos de secuencias que presentan evolución neutra son las secuencias no codificantes tales como pseudogenes, regiones no traducidas y la mayoría de intrones. Debido a que estas secuencias no codifican ningún producto proteico estas regiones no están sujetas a una limitación funcional y son libres de acumular mutaciones al azar. En estas secuencias no codificantes, todas las sustituciones de nucleótidos son silenciosas, esto es, cualquier nucleótido puede ser sustituido debido a que carece de función, por lo tanto, solo es posible calcular sustituciones sinónimas por sitios sinónimos (Li y Graur, 1991). La divergencia o diferenciación en secuencias no codificantes puede ser comparada con las sustituciones sinónimas de secuencias codificantes.

La selección positiva es otro tipo que favorece a los nuevos alelos mutantes que son ventajosos aumentando así la diversidad a nivel de aminoácidos. Este tipo de selección se ha encontrado en proteínas involucradas en procesos de reconocimiento extracelular, como las proteínas del sistema inmune (Hughes y Nei, 1988). La actividad de la selección positiva puede ser entendida en términos de interacciones "huésped-patógeno" en proteínas. Estas interacciones pueden ser consideradas como una "carrera de armas", esto es, un cambio en el ligando de una proteína confiere una ventaja selectiva a su poseedor (huésped), la que a su vez causa en los patógenos una presión para adaptarse a dicho cambio, de tal modo que sean capaces de interactuar con su

huésped. De esta manera se conduce a un proceso evolutivo rápido. En este caso las sustituciones que alteran aminoácidos son ventajosas para la función de la proteína, por lo que éstas exceden la tasa de sustituciones sinónimas o silenciosas, donde dn es mayor que ds (Hughes y Nei, 1988).

En el abulón, las secuencias codificantes de lisina y 18 kDa han sido extensamente estudiadas. La selección positiva promueve los cambios de aminoácidos en ambas proteínas alcanzando uno de los niveles más altos que se conocen (Vacquier *et al.* 1993; Swanson y Vacquier, 1995). Aún se desconocen las fuerzas selectivas causantes de este tipo de evolución, sin embargo, existen varias hipótesis para explicar este fenómeno: 1) Reforzamiento: la selección favorece la diferenciación de los sistemas de reconocimiento en los gametos para hacerlos más específicos y prevenir la hibridización. 2) Diversificación funcional de las dos proteínas acrosomales: después de su origen por duplicación del gen ancestral, la selección ejerce una presión para generar en las dos proteínas recién surgidas funciones diferentes en cada una de ellas. 3) Antagonismo sexual: los ligandos en la superficie de los ovocitos son alterados para reducir la polispermia, a su vez la selección ejerce una presión en el espermatozoide para alterar sus proteínas de reconocimiento de manera que éstas puedan ser afines a los ligandos cambiantes de los ovocitos. 4) Selección sexual: en este caso la selección actúa en las variantes de una proteína en el espermatozoide las cuales tienen diferentes afinidades a los ligandos de la superficie del ovocito. 5) Autorreconocimiento-No Autorreconocimiento y Defensa Inmunológica: la selección actúa en las proteínas de la superficie del ovocito de manera que la colonización por microbios pueda ser evitada y eventualmente ésta causa una presión selectiva en las proteínas del espermatozoide para

que se logren adaptar a los ligandos modificados de la superficie de los ovocitos, (Vacquier *et al.* 1997).

A pesar de que los ADN codificantes de lisina y 18 kDa han sido ampliamente estudiados se conoce poco acerca de la estructura genética de estas proteínas incluyendo sus intrones. Un intrón es una secuencia transcrita que es eliminada durante el proceso de maduración del ARNhn (heterogéneo nuclear) para dar lugar al ARNm (mensajero), por lo que ésta no posee ninguna información codificante de polipéptidos. Un patrón común encontrado en la mayoría de los genes es una alta variación de las secuencias no codificantes. La comparación de las secuencias de ADN de diversos organismos ha mostrado que la tasa de sustitución de nucleótidos varía considerablemente dependiendo de la región de ADN examinada y que entre mayor es la importancia de la función de la región menor es la tasa de sustitución, esto es, si la función es central para la célula, tiene menos cambios permitidos (Kimura y Ohta, 1974). La relación inversa de la limitación funcional y la tasa de sustitución de nucleótidos se basa en la observación de secuencias de intrones, pseudogenes y regiones flanqueantes. Un ejemplo de la variación en secuencias de intrones se encuentra en las comparaciones hechas en el gen humano de la enzima dihidrofolato reductasa (DHFR), y en genes homólogos de mamíferos (Yang *et al.* 1984). Se han encontrado intrones en los genes de la proteína bindin en erizo de mar en los cuales se ha detectado una divergencia elevada también en diferentes especies (Metz y Palumbi, 1996; Metz *et al.* 1998). Se ha sugerido que la rápida acumulación de cambios de nucleótidos es un patrón general en la evolución de intrones. Debido a la información que nos puede brindar el estudio de intrones es interesante el planteamiento de numerosas preguntas, por ejemplo: ¿Se presenta algún patrón de

evolución molecular en regiones no codificantes de proteínas bajo selección positiva? El estudio de este proceso evolutivo en proteínas de reconocimiento de gametos tales como lisina y 18 kDa en invertebrados marinos será útil para el entendimiento de cómo este tipo de evolución está involucrada en el establecimiento de barreras de aislamiento reproductivo y la especiación en organismos invertebrados marinos.

Los patrones de evolución molecular en intrones de lisina y 18 kDa no han sido estudiados previamente. Por lo tanto, con el fin de incrementar el entendimiento sobre la evolución de las proteínas de reconocimiento de gametos, en este estudio se secuenciaron y analizaron varias regiones de intrones de abulón rojo (*Haliotis rufescens*), abulón rosa (*H. corrugata*) y abulón azul (*H. fulgens*).

OBJETIVOS

- a) Obtención y secuenciación de regiones no codificantes de lisina mediante las técnicas de PCR y secuenciación automatizada.
- b) Identificar las posiciones de intrones en las secuencias codificantes de las proteínas de lisina y 18 kDa de las especies de abulón, *H. rufescens*, *H. corrugata* y *H. fulgens*.
- c) Comparar la evolución molecular en las secuencias codificantes versus secuencias no codificantes de los genes de lisina y 18 kDa.
- d) Medir patrones de evolución molecular en intrones como mutaciones puntuales, inserciones, deleciones y secuencias repetitivas.

ANTECEDENTES

Ensayos de especificidad de especie en lisina y 18 kDa:

Vacquier *et al.* (1990) desarrollaron ensayos para determinar la especificidad de especie de lisinas de abulón rosa y rojo para disolver capas vitelinas aisladas y marcadas con I¹²⁵. Los resultados mostraron que la lisina de abulón rosa disuelve capas vitelinas de abulón rosa y rojo, pero es necesario una cantidad mayor de lisina de abulón rosa para disolver 50% de capas vitelinas de abulón rojo que la utilizada para disolver 50% de capas vitelinas de abulón rosa. Otro tipo de ensayo es el espectrofotométrico en el cual se abordó la disolución específica de especie de manera cuantitativa. Se purificaron lisinas de abulón negro (*H. cracherodii*), rosa (*H. corrugata*), rojo (*H. rufescens*) y midieron su capacidad de disolución de las capas vitelinas de abulón negro, rosa, rojo y pinto (*H. kamtschatkana*) de manera cruzada, esto es, una especie contra sí misma y contra las otras. Los resultados mostraron que aunque la mayoría de los ensayos de disolución entre diferentes especies no mostraron una especificidad de especie absoluta estos claramente demostraron una selectividad de especie cuantitativa en la acción de las lisinas, como se muestra en la fig. 3, Vacquier y Lee (1993).

Estudios de ADN codificante de lisina y 18 kDa:

Vacquier *et al.* (1990) obtuvieron por primera vez el ADN codificante a partir del ARN mensajero maduro del abulón rosa y rojo mediante clonación. Posteriormente establecieron su secuencia. Los ADN codificantes de otras especies de abulón fueron amplificados por la técnica de la Reacción en Cadena de la Polimerasa (PCR) usando

Fig. 3a
C.V. de Abulón Negro

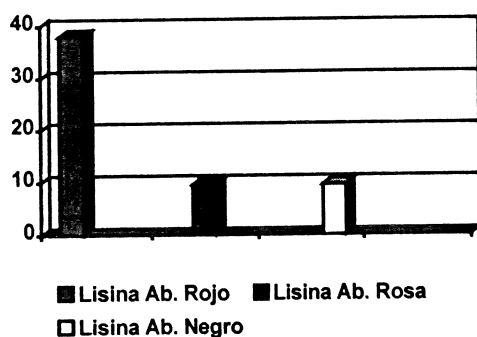


Fig. 3b
C.V. de Abulón Rojo

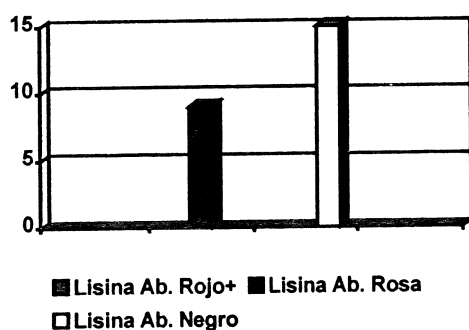
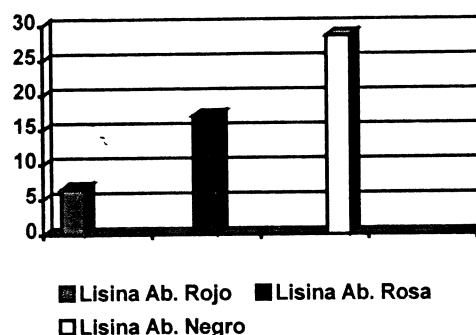


Fig. 3c
C.V. de Abulón Rosa

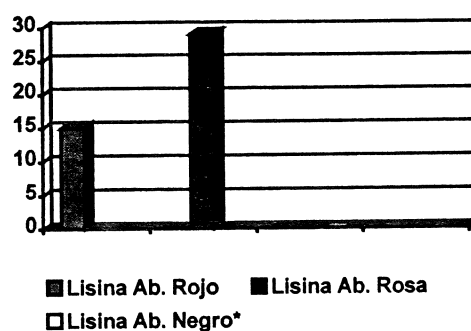


Fig. 3d
C.V. de Abulón Pinto

Fig. 3: Resultados de los ensayos de disolución de capas vitelinas específica de especie que muestran la cantidad de microgramos necesarios para disolver: **Fig. 3a:** 50% de capas vitelinas (C.V.) de abulón negro; **Fig. 3b:** 50% de C.V. de abulón rojo; **Fig. 3c:** 50% de C.V. de abulón rosa, (+) la más alta cantidad de lisina ($45\mu\text{g}$) disolvió <20% de las C.V.; **Fig. 3d:** 50% de C.V. de abulón pinto, (*) la lisina de abulón negro es inefectiva para disolver C.V. de abulón pinto. (Datos adaptados de Vacquier y Lee, 1993).

"primers" (oligonucleótidos de ADN) universales diseñados a partir de regiones conservadas del ADNc de abulones rojo y rosa (Fig.4). Lee *et al.* (1995) y Lee y Vacquier (1995) secuenciaron y analizaron el ADNc de 27 especies diferentes de diversas partes del mundo. Las secuencias de aminoácidos deducidas de las 27 lisinas varían en longitud desde 126 a 138 residuos.

Swanson y Vacquier (1995), aislaron el ADNc de las proteínas 18 kDa y obtuvieron las secuencias de aminoácidos de cinco especies de abulón de las siete reportadas hasta el presente en California. A partir de la secuencia del producto de PCR del abulón rojo se obtuvo la secuencia de aminoácidos de la cual se diseñaron los "primers" para amplificar el ADNc de las otras cuatro especies encontrando que las proteínas maduras de 18 kDa varían en longitud desde 132 a 146 residuos.

Análisis de Secuencias:

A partir de la alineación de las secuencias de aminoácidos deducidos de las proteínas y 18 kDa se han determinado mediante el método de Nei y Gojobori el número de sustituciones no sinónimas por 100 sitios no sinónimos (*dn*), y el número de sustituciones sinónimas por 100 sitios sinónimos (*ds*) (Lee *et al.* 1995 y Swanson y Vacquier, 1995). La tabla I muestra los valores de *dn* y *ds* obtenidos en lisina y 18 kDa en comparaciones de las siete especies de abulón reportadas para California. Cabe destacar que la proporción *dn/ds* varía desde 1.08 a 3.82 en las comparaciones de lisina en tanto que en las comparaciones de 18 kDa la proporción tiene un rango más

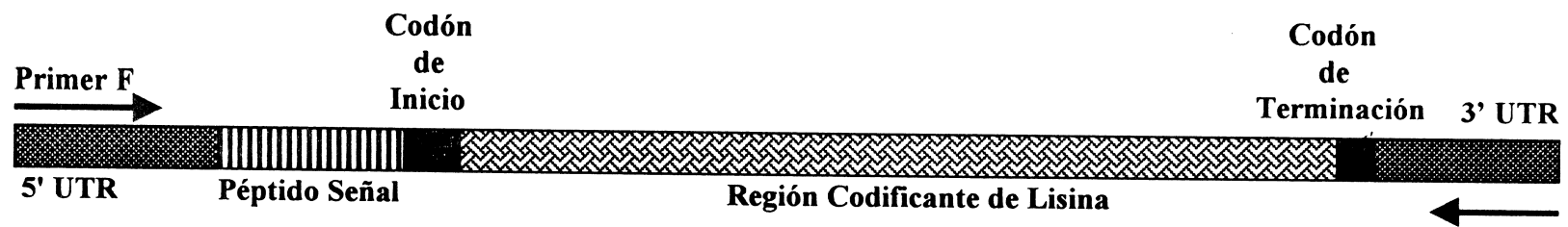


Fig. 4: Esquema del ADNc de lisina, donde se muestran las regiones conservadas: regiones no traducidas 5' y 3' (UTRs), péptido señal, codón de inicio y de terminación. Los oligonucleótidos ("primers") F y R fueron diseñados a partir de los UTRs con el propósito de obtener la región codificante de lisina. Tomado de Lee y Vacquier (1992).

Tabla I: Porcentaje de identidad y proporción **dn/ds** de las proteínas de lisina y 18 kDa comparadas en diferentes especies Hr: *H. rufescens*; Hs: *H. sorenseni* ; Hk: *H. kamtschatkana*; Hc: *H. corrugata*; Hf: *H. fulgens*. Datos tomados de Vacquier *et al.* 1997.

Comparación de Especies	Porcentaje de Identidad	Proporción dn/ds
Lisina		
Hr-Hs	92	3.62
Hr-Hk	83	3.82
Hr-Hc	80	1.38
Hr-Hf	67	1.17
Hs-Hk	86	2.55
Hs-Hf	67	1.20
Hs-Hc	84	1.38
Hk-Hc	74	1.85
Hk-Hf	65	1.08
Hc-Hf	60	1.50
18 kDa		
Hr-Hs	87	4.50
Hr-Ha	75	3.78
Hr-Hc	31	0.95
Hr-Hf	34	1.04
Hs-Ha	77	4.67
Hs-Hf	35	0.92
Hs-Hc	33	0.97
Ha-Hc	31	0.89
Ha-Hf	35	1.02
Hc-Hf	27	0.76

amplio de variación que va desde 0.76 a 4.67, esto significa que la selección positiva actúa más significativamente en las proteínas de 18 kDa.

Otras Comparaciones de Secuencias de Exones e Intrones:

Para tener una referencia de la variación de nucleótidos en las secuencias de exones e intrones en diferentes proteínas, la tabla II presenta varias comparaciones de genes de ratón (*Mus musculus*) y rata (*Rattus norvegicus*) donde se calculan valores de dn y ds en exones y la tasa de sustitución de nucleótidos total, d , en secuencias de intrones de los mismos genes.

En la gran mayoría de los genes la tasa de sustituciones sinónimas excede la tasa de sustituciones no sinónimas. La tasa de sustitución de nucleótidos en intrones es también variable entre genes pero ha mostrado ser muy similar a la tasa de sustituciones sinónimas en exones (Li y Graur, 1991; Hughes y Yeager, 1997).

Tabla II: Variación de nucleótidos en exones (sustituciones no sinónimas por 100 sitios sinónimos, **dn**; sustituciones sinónimos, **ds**; y proporción **dn/ds**) comparados con sustitución de nucleótidos por 100 sitios en intrones (**d**) en genes de ratón (*Mus musculus*) y rata (*Rattus norvegicus*). Datos tomados de Hughes y Yeager, 1997.

Gen	dn	Exones	dn/ds	Intrones
		ds		d
Citocromo P-1-450	2.6	13.8	0.18	16.3
Hormona Liberadora de Gonadotropina	5.1	16.4	0.31	16.7
Cadena Ligera C-K de Inmunoglobulina	10.9	19.1	0.57	17.2
Proteasa de Célula Cebada	14.7	26.5	0.55	25.4
Periferina	2.3	20.8	0.11	15.4
Preprosomatostatina	6.7	15.7	0.42	15.4
Receptor V-β de Célula T	9.9	24.4	0.40	11.6
Péptido Natriurético	12.2	21.4	0.57	20.4
Renina	10.5	25.8	0.40	23.0

MÉTODOS

Se colectaron tres especies de abulón en aguas costeras de California, abulón rojo (*H. rufescens*), abulón rosa (*H. corrugata*) y abulón azul (*H. fulgens*) los cuales fueron proporcionados por el Dr. Victor D. Vacquier del Instituto de Oceanografía Scripps (SIO).

I.- Extracción de ADN genómico y Preparación de Template para PCR:

El ADN genómico se extrajo a partir de uno o dos tentáculos del epipodio del abulón los cuales fueron desprendidos con un hemostato previamente esterilizado con etanol al 95% y secado a la flama. Este procedimiento no causó daños fatales a los organismos.

Los tentáculos fueron colocados en un tubo de centrífuga de 1.5 ml manteniendolos en hielo, se removió el exceso del agua de mar, luego los tentáculos se lavaron con 150 µl de solución de Tris 10 mM- EDTA (TE) 10 mM pH 7.6, se agitaron en "vortex" por 2 o 3 seg y se centrifugó por 5 seg a 14 000 rpm y se le agregaron 150 µl de TE. El extremo de una punta de pipeta de plástico fué derretida y usada como instrumento para macerar los tentáculos dentro de los tubos. Después de macerados, los tentáculos se agitaron en vortex brevemente, se les agregó 100 µl de una suspensión al 10% de Chelex 100 (Sigma) en agua destilada y se hirvieron por 5 minutos. Chelex es una resina quelante que remueve cationes polivalentes tales como Mg^{++} y Ca^{++} , que funcionan como cofactores para las enzimas degradantes de ADN ayudando a mantener la integridad del ADN. Inmediatamente después de que los tentáculos hirvieron se

enfriaron en cama de hielo durante 2 minutos, después fueron agitados con vortex y centrifugados por 30 segundos a 14 000 rpm. El sobrenadante de cada muestra de tentáculos se almacenó a -20°C hasta su uso como template en reacciones de PCR subsecuentes.

Las muestras de ADN genómico altamente purificado se obtuvieron a partir de espermatozoides de abulón proporcionados gentilmente por el laboratorio del Dr. Victor D. Vacquier.

II.- Reacción en Cadena de la Polimerasa (PCR):

La reacción en cadena de la polimerasa es una técnica molecular que permite la síntesis de una determinada región de ADN, usando moléculas de ADN desnaturalizadas como templates y dos "primers". Cada "primer" es complementario en secuencia y flanquea la región a amplificarse de cada hebra. La síntesis de ADN esta dirigida por una polimerasa termoestable. La reacción generalmente es repetida durante 20 a 35 ciclos, lo que es determinado por el usuario y en cada ciclo el ADN se amplifica en forma exponencial.

Cada ciclo de PCR está compuesto por tres etapas principales:

1) Desnaturalización: Debido a que la reacción requiere como template o molde ADN de hebra sencilla, es necesario separar las dos hebras de la doble hélice de ADN mediante la aplicación de altas temperaturas. El tiempo dependerá del tipo de ADN usado como template de manera que se obtenga una desnaturalización eficiente.

2) Alineamiento: Una vez que las hebras de ADN están separadas, la temperatura debe disminuirse a un nivel que permita el apareamiento de los "primers" a la secuencia específica complementaria en el template de ADN.

3) Extensión: Esta etapa involucra la elevación de la temperatura desde aquella relativamente baja de apareamiento hasta una temperatura donde la ADN polimerasa trabaja más efectivamente. La ADN polimerasa comienza a sintetizar la nueva hebra de ADN en la terminación 3' del "primer".

IIa.- Preparación de la mezcla de PCR:

La mezcla para cada reacción de PCR efectuada fué preparada de acuerdo a lo sugerido por Kessing *et al.* (1989), la cual consta de:

50 µl de 10x Taq Extender Buffer (Stratagene)

50 µL 8 mM dNTPs (2.0mM de dATP, dGTP, dCTP, dTTP)

20 µL 10mM del "primer" 1 (Tm: 50-70°C) (Operon)

20 µL 10mM del "primer" 2 (Tm:50-70°C) (Operon)

360 µL de Agua destilada, filtrada a través de membranas milipore y esterilizada con luz UV.

1.8 µL (9 unidades) de Taq Extender (Stratagene)

1.8 µL (9 unidades) de Taq ADN polimerasa (Sigma)

Taq Extender es un aditivo de PCR el cual se agregó a la mezcla de PCR con el fin de incrementar la eficiencia con la que la Taq ADN polimerasa desarrolla las reacciones de extensión en segmentos específicos de ADN en cada ciclo de PCR.

Esta mezcla de reacción hace 500 μL , la cual fué dividida en 10 tubos de PCR conteniendo cada uno 50 μL , a estos se les agregó de .026 a .038 ng de template de ADN genómico y 1 o 2 gotas de aceite mineral para prevenir la evaporación de la muestra.

IIb.- Condiciones de PCR:

- 1) Desnaturalización : 94°C por 30 segundos.
- 2) Alineación: 55°C por 30 segundos.
- 3) Extensión: 72°C por 3 minutos.

Número de Ciclos: 35

La temperatura de alineación utilizada fué menor que la T_m requerida para algunos "primers" diseñados, sin embargo, es una temperatura promedio que demostró ser efectiva en las reacciones de amplificación y secuenciación.

En los casos en los que no se disponía de suficiente template de ADN el número de ciclos era incrementado de manera que se obtuviera una mayor amplificación.

Con el fin de saber si existía alguna secuencia de intrones en el gen de lisina se corrieron PCRs usando los dos pares de "primers", rLA-rLC y r4-r12, diseñados a partir del ADNc de esta proteína. Estos "primers" fueron proporcionados por el laboratorio del Dr. Victor D. Vacquier, en donde se utilizaron en estudios previos del gen de lisina de abulón. La figura 5 muestra esquemáticamente la ubicación de los pares de "primers" en la secuencia codificante de lisina.

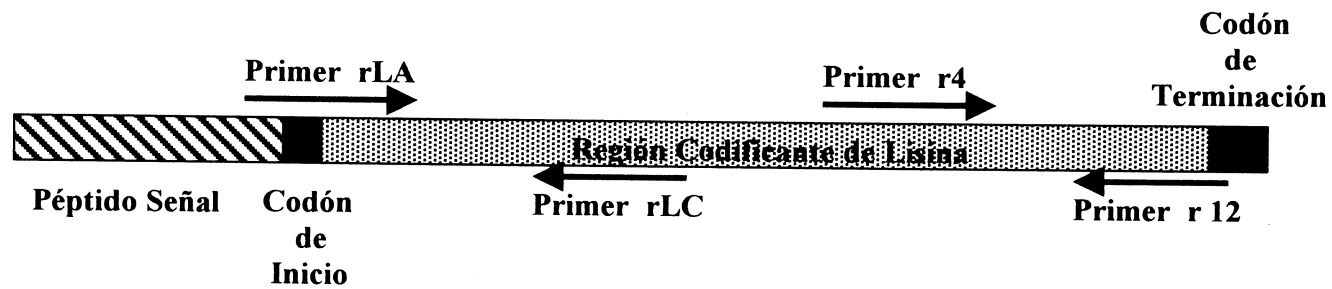


Fig. 5: Esquema de la región codificante de lisina que muestra la posición de los dos pares de oligonucleótidos empleados para multiplicar las secuencias de los intrones. Las secuencias de los "Primers" se muestran en la Tabla III.

III.- Detección y Aislamiento de los Productos de PCR:

IIIa. Electroforesis en Gel de Agarosa al 1%:

Los productos de PCR fueron detectados mediante electroforesis en gel de agarosa al 1% (1gr de agarosa Ultra Pura en 100 ml de Tris Borato-EDTA (TBE) al 0.5% con 3 μ L de bromuro de etidio 10 ng/ml), (Sambrook *et al.* 1989)

El gel se colocó en una cámara de electroforesis que contenía solución de TBE al 0.5% con 6 μ L de bromuro de etidio.

Se tomaron 20 μ L de las muestras de los productos de PCR los cuales se tiñeron con 2-3 μ L de azul de bromofenol 6x (Sambrook *et al.* 1989), y posteriormente fueron cargados en los pozos del gel de agarosa evitando contaminación de una muestra con otra. Para tener una referencia del peso molecular de los productos se corrió en el mismo gel un marcador de peso molecular estándar de 500 pb (Mc Frugals Inc.). En el corrimiento de los geles se aplicaron 110 voltios durante 45 minutos.

IIIb.- Detección de Bandas y Purificación de DNA:

Después de realizada la electroforesis, se detectaron las bandas en el gel utilizando un transiluminador de luz UV (260 nm). El gel fué expuesto al menor tiempo posible para la observación y corte de las bandas, con ello se redujo el riesgo de degradación del ADN.

El ADN de las bandas se extrajo y se purificó con el Kit de Extracción de Gel (Quiagen Inc. 1997), que es un sistema que emplea la propiedad del ADN de adsorberse a membranas de sílice en presencia de altas concentraciones de sales caotrópicas y pH

menores de 7.5 y eluirse en condiciones básicas y a concentraciones bajas de sales. Las muestras de ADN obtenidas fueron almacenadas a -20°C, para usos posteriores.

IIIc.- Cuantificación de ADN:

Una de las etapas previas a la secuenciación de ADN es su cuantificación la cual se determinó mediante espectrofotometría con luz UV, utilizando procedimientos estándares (Sambrook *et al.* 1989).

$$(A_{260})(F.D)(K) = X \mu\text{g/ml DNA}$$

A_{260} = Lectura a 260 nm

F.D. = Factor de Dilución

K = 50 $\mu\text{g/ml}$

IV.- Reamplificación de ADN:

En los casos en que las bandas de ADN resultaban ser demasiado tenues y/o las concentraciones de ADN eran demasiado bajas para permitir la secuenciación del ADN, entonces se efectuaban reamplificaciones del producto deseado de manera que se incrementara la concentración de la muestra. Debido a que la técnica de PCR es muy sensible, en las reamplificaciones se utilizaban concentraciones de .0026 a .0038 ng del template de ADN. La preparación de la mezcla y las condiciones de PCR, extracción y purificación de ADN fueron las mismas que las previamente descritas.

V.- Diseño de "Primers" Específicos para Intrones:

Una vez que se detectaron intrones en el gen de lisina, se diseñaron "primers" específicos para los intrones de manera que se permitiera el avance en la secuenciación. En el diseño se consideraron los siguientes aspectos, los cuales son recomendados por Kessing *et al.* (1989):

- 1) Longitud: Los "primers" se diseñaron con una longitud aproximada de 20 a 24 pb, los cuales mostraron una especificidad adecuada en las reacciones de secuenciación.
- 2) Composición de nucleótidos: Se diseñaron de manera que poseyeran una cantidad similar de bases de Guanina+Citosina y de Adenina+Timina, con una T_m que oscilaba entre 50°C - 70°C. Para disminuir problemas de mal apareamiento o dimerización se evitaron repeticiones internas de bases y la presencia de bases apareables en los extremos.

El programa de computadora Amplify (Engels, 1993) fué usado para confirmar y optimizar la efectividad de los "primers". Este programa permite: detectar la posibilidad de aparición de dímeros, apareamientos múltiples de "primer"-templete, determinar la dirección y la posición de los "primers" en la secuencia y los posibles apareamientos entre los pares de "primers" (autocomplementaridad).

VI.- Secuenciación Automática de ADN:

El proceso de secuenciación y el registro de lecturas se llevó a cabo en un secuenciador automático del Instituto de Oceanografía Scripps, de la Universidad de California en San Diego.

La secuenciación automatizada es una técnica que permite la determinación de nucleótidos de un fragmento de ADN. Este fragmento es amplificado por una ADN polimerasa que coloca cuatro terminadores de color a la hebra que está siendo sintetizada. Cada terminador es específico para cada una de las bases de ADN. Una vez que el terminador está incorporado a la nueva hebra de ADN la síntesis es suspendida. La detención ocurre al azar, de manera que existen hebras que continúan polimerizándose; al final de los ciclos de secuenciación es posible encontrar hebras de ADN de todos los tamaños. Los fragmentos de ADN son corridas en un gel y los terminadores de color son detectados por un rayo láser el cual excita la fluorescencia particular de cada terminador identificando así a cada una de las bases. Finalmente, la secuencia completa de nucleótidos de la hebra de ADN es grabada en un archivo de computadora, llamado cromatograma, el cual registra la intensidad de la fluorescencia de cada base.

Para ejecutar la secuenciación se prepararon de 50 a 100 ng de muestra de ADN de las diferentes especies de abulón, usándose el kit ABI PRISM FS (Perkin Elmer Co., 1995). Las condiciones de secuenciación y precipitación de ADN fueron seguidas de acuerdo al protocolo del fabricante.

VII.- Análisis de Secuencias:

Para el análisis de las secuencias se utilizaron varios programas de computadora, éstas fueron alineadas usando el programa "Seq Ed" (Applied Biosystems Software). Las sustituciones silenciosas de nucleótidos de secuencias no codificantes y las distancias genéticas de las secuencias de diferentes especies fueron calculadas usando el programa "Molecular Evolutionary Genetic Analysis" (MEGA) (Sudhir *et al.* 1993) . Las

comparaciones de la variación de secuencias en regiones codificantes y no codificantes fueron representadas gráficamente usando el programa "DNA Sequence Polymorphism" (DNAsp) (Rozas y Rozas,1997) y los análisis de la estructura secundaria en secuencias no codificantes fueron elaborados usando el programa "Genetics Computer Group" (GCG, 1991).

RESULTADOS

I) Productos de PCR:

Lisina:

En abulón rojo el producto de PCR obtenido al utilizar el par de "primers" rLA-rLC (Fig. 5) fué de aproximadamente 2800 pb (Fig. 6), en abulón rosa el tamaño del fragmento fué aproximadamente de 2500 pb. En abulón azul éstos oligonucleótidos sintetizaron una banda muy pequeña y difusa de aproximadamente 4500 pb que no se amplificaba consistentemente en reacciones subsecuentes, por lo que se optó por utilizar un nuevo "primer" que tuvo una mayor amplificación y se presentaba consistentemente en cada una de las reacciones efectuadas. El par de "primers" que funcionó fué 16.Aintbf - rLC, el cual produjo una banda de aproximadamente 3000 pb.

El par r4 - r12 (Fig.5) sintetizó un producto de PCR de aproximadamente 1500 pb en el abulón rojo (Fig. 6) y un fragmento de aproximadamente 750 pb en el abulón rosa. En el abulón azul este par de "primers" no amplificó ningún fragmento, por lo que se diseñó un nuevo oligonucleótido equivalente en posición al r12 a partir de la secuencia de ADNc del abulón azul, el par de "primers" utilizado fué r4 - Gr12, el cual produjo un fragmento de aproximadamente 1600 pb.

II.- Posiciones de Intrones:

Los productos de PCR fueron secuenciados con los mismos "primers" utilizados en la amplificación de fragmentos de ADN genómico. La secuencia obtenida fué comparada con la de ADNc de manera que se pudieran identificar las regiones donde

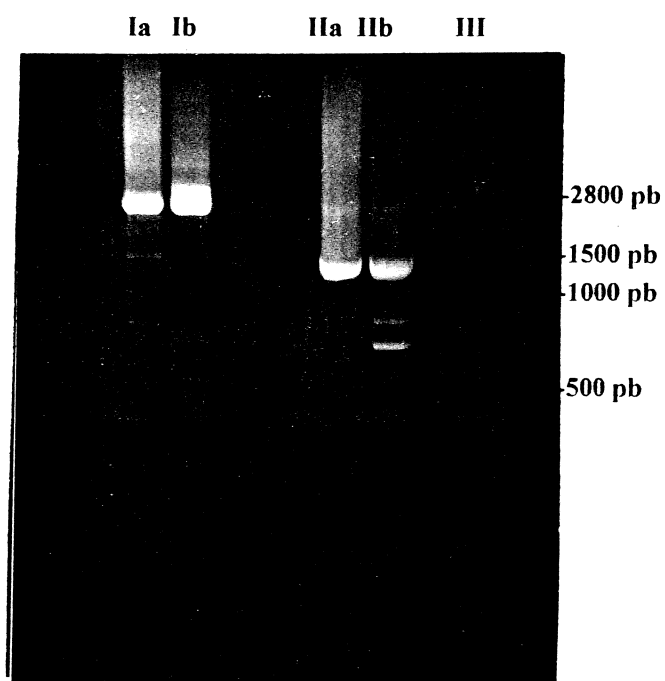


Fig. 6: Gel de agarosa que muestra los productos de PCR obtenidos a partir del ADN genómico de abulón rojo, utilizando los pares de "primers", rLA-rLC y r4-r12. Par rLA-rLC: **Ia)** Reamplificación del producto. **Ib)** Amplificación original. Tamaño aproximado del producto: 2800 pb. Par r4-r12: **IIa)** Reamplificación del producto. **IIb)** Amplificación original. Tamaño aproximado del producto: 1500 pb. **III)** Estándar de peso molecular: 0.5 kb.

terminaban los exones y empezaban los intrones, de esta manera se determinaron las posiciones de los intrones. El comienzo de cada intrón fué identificado por la presencia del dinucleótido GT y el término por la presencia del dinucleótido AG (Li y Graur, 1991). Se encontraron dos intrones en la región comprendida por el par de oligonucleótidos rLA - rLC, y un intrón en la región comprendida por el par r4 - r12. Se diseñaron ""primers" de las regiones adyacentes a las secuencias de los "primers" rLC y r4 con el fin de determinar la presencia de intrones en la región media del ADN codificante de lisina. La mayoría de ellos no amplificaron ningún fragmento de ADN. Aquellos que funcionaron produjeron bandas muy tenues de alto peso molecular, que cuando se secuenciaron mostraron cromatogramas con picos altamente traslapados, haciendo las secuencias ilegibles. Los PCRs de larga extensión realizados para amplificar esta región producen una banda de alto peso molecular, lo cual sugiere la presencia de un intrón de gran tamaño en esta región media. Suponiendo la existencia de este intrón en la región media del ADNc, se convino en denominársele el número 3, la figura 8 muestra la ubicación de este intrón aun no determinado.

En la proteína 18 kDa se encontraron dos intrones en la región rLA-rLC, uno en la región comprendida por el par de "primers" r4 -r12, y un intrón más en la región comprendida por el par rLC y r4 (región media). La figura 7 muestra las posiciones de los intrones en las regiones codificantes de lisina y 18 kDa los cuales fueron detectados a partir de las reacciones de secuenciación realizadas. Cabe señalar que los intrones comparten posiciones idénticas cuando sus secuencias de aminoácidos son alineadas de acuerdo a su estructura secundaria.

Fig. 7: Alineación de la secuencia de aminoácidos de lisina y 18 kDa con las posiciones de los intrones indicadas por flechas. El signo de interrogación sugiere la presencia de un intrón más en el gen de lisina.

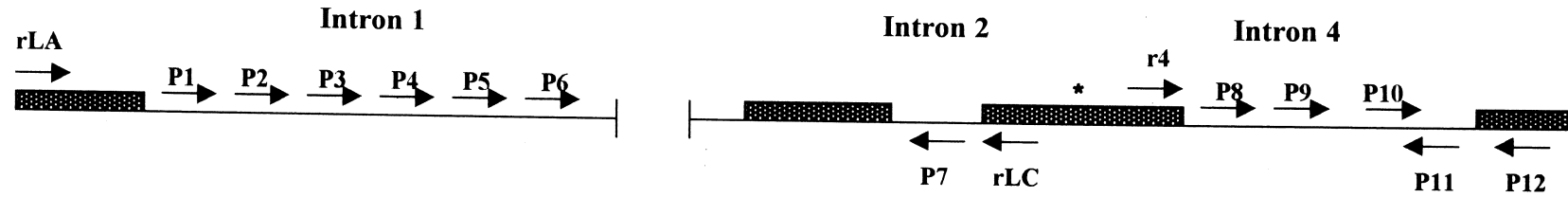
III.- Secuencia de Intrones:

El diseño de "primers" específicos para intrones permitió obtener las secuencias de éstos en las tres especies de abulón. La figura 8 muestra el nombre y la localización de los oligonucleótidos usados en cada especie y la tabla III señala la secuencia de los "primers" específicos de intrones.

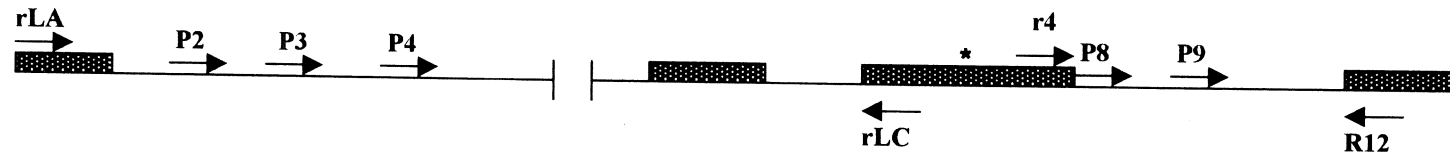
IIIa) Alineación de Secuencias de Intrones:

Las secuencias de intrones obtenidas fueron alineadas usando el programa de computadora "Seq Ed" (Applied Biosystems, Inc., 1995) mediante la alineación de las secuencias que primero se obtuvieron y luego se fueron añadiendo las nuevas secuencias obtenidas de manera progresiva. La figura 9 muestra la alineación de las secuencias parciales obtenidas del intrón 1 y la figura 10 muestra las secuencias totales del exón 4 e intrón 4 de las tres diferentes especies. En el intrón 1 del abulón rojo 2300 pb del total de 2800 pb que mostró tener la banda amplificada, fueron secuenciadas, y 1200 pb del total de 2500 pb indicadas por el tamaño de la banda en el mismo intrón en abulón rosa. En el abulón azul no fué posible determinar la unión de exón-intrón debido a que el "primer" rLA no funcionó en esta especie y un "primer" específico de intrón, 16.Ainbf, fué utilizado en su lugar. En el abulón azul 920 pb fueron secuenciadas del total de las 3000 pb indicadas por el tamaño de la banda amplificada. De la alineación de secuencias de los intrones 1 y 4 es evidente que estos son inusualmente similares en las tres especies de abulón.

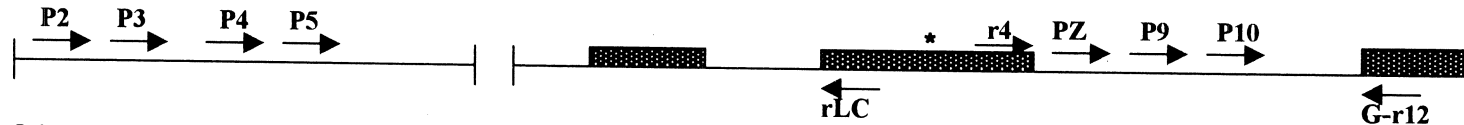
Abulón Rojo



Abulón Rosa



Abulón Azul



Oligonucleótidos:

P1) 16.Aint1f P3) 16.Aintcf P5) 16.Aintef P7) 16.Cint2r P9) 16.4intbf P11) 16.12int4r
P2) 16.Aintbf P4) 16.Aintdf P6) 16.Aintff P8) 16.4int3f P10) 16.4intcf PZ) Gr-4intzf

Fig. 8: Posición y nombre de los oligonucleótidos ("primers") usados para amplificar las secuencias de los intrones del gen de lisina de las especies de abulón rojo, rosa y azul. El asterisco (*) indica la región media de los exones que puede contener un tercer intron.

Tabla III: Nombre y secuencia de bases de los primers utilizados en la amplificación de intrones del gen de lisina.

PRIMER	SECUENCIA
rLA	5'-AAG ATG AAG CTG TTG GTG CT-3'
16.Aint1f	5'-GTT TGA ACG CTT TCC ATA ACC-3'
16.Aintbf	5'-TCA GGA TTT ATC ATC ATA GGA CA-3'
16.Aintcf	5'-GAT TGG TTC GAG ATT GGA GAC-3'
16.Aintdf	5'-GTT TTC TGT CTC GAA GTC GTC G-3'
16.Aintef	5'-CGA TCA AAT ATC GGA CGA GGG-3'
16.Aintff	5'-ATT TAT GTA GGT GTT ATC GAT-3'
rLC	5'-ATA GAC GTT ACA TGC AGA CCC AC-3'
16.Cint2r	5'-AAA AAT CAC AAC TCT TAC TCG-3'
r4	5'-GTT AGG TCG CGA AAT TGG A-3'
16.4int3f	5'-CAC CAC AATAAT GAA GCA CTC-3'
16.4intbf	5'-CTG GAC TAT TGC TGA ATG CG-3'
16.4intcf	5'-AAG TGG GAT CTG GTC GTC AT-3'
16.4intzf	5'-ATC ATAT CAG CTA GGA CTG-3'
r12	5'-ATG CGC CCG GCA GAT GTC CCA GTT-3'
16.12int4r	5'-CAA ACA AGT ATG GAC TGG GCT-3'
G-r12	5'-AAC TGG GAC ATC TGCCGGGCG CAA-3'
G-4int2f	5'-ATC ACT TGA CTC TGC CGA CA-3'

1

Azul	GGGGATAC?C	CTATAAATAT	CGCCAATTGT	TTTCTGTCTC	GAAGTCGTCG	GGAAATAATT
Rosa	G..A.	A....T....	C..			
Rojo		A.....	?			
Azul	TTCAGTAGAT	CTCCGAGACA	CTTTTAGATT	TGCCGATAAT	GATTTCCGCC	AGCGTTGACA
Rosa	A.....	T.....	...??...	G..		
Rojo	A.....	T.....				
Azul	CCCGTCGTGA	ATTTTGAACG	TACCGGTTGC	ATTCCATGCT	ATGCGTCGTG	TTCAAAAGAG
Rosa			...T....	C...		
Rojo				C...		
Azul	AAAACAAACA	CGGTCATGTA	AAGTCGATGT	TTAATCAAGA	TGTGGACCAA	CTTGTACAAA
Rosa			A....	TT...	A.....	
Rojo				CTT...		
Azul	GCTATTGTAG	CGCAATAAAT	ATTGTAAATA	GGAAGTTATT	CGGTAGTTCT	ATATTATGCA
Rosa				A....		
Rojo				A....		
Azul	TTGACTAACA	CTTACACATA	GTCGTAGCGC	TACAATCGCA	CTGAGTAAAT	GGTCACTGGT
Rosa	..T.A.....	A...			..A.....	
Rojo	..T.A.....	A...		G.....	..A.....	
Azul	GCATTTACCG	GATACAT?CC	GCCATGTCTG	AAAGAAACAT	ACGTAAGGTG	ATAGAACCAG
Rosa		T..		GT.C	G?.....	?....
Rojo		T..		G???	??.....A	?....
Azul	GGTTTTCTTT	TC?AAACTTT	GC?????ATG	TTAATTTGAC	GTTTTAAATT	GTGTTTTGCT
Rosa	?....	..G..?..?			A...A....?	?.....A..
Rojo		..G..?..?			A.A.A.....	
Azul	CGAGTAAGAG	TTGTGA?TTT	TTTTTCTAAA	CGCGTATATT	ATATATGCTA	TCATCGAGTT
Rosa		A..C	T.	??????????	??????????	????.....
Rojo		?..	T.	..AA.....	...C.....	..C.....
Azul	CCTGTGAGAG	ATATTAAG?G	TGTAAGGCTG	AATGAAATGA	AATGTACTT?	ACCCTGTGTA
Rosa	A	C.	T....		T..GC	G.
Rojo		C.			T.AGC	G.
Azul	TGACGTGTAC	TGCAGTAGAG	ATTTCTTCTC	ATATTATTAT	ATGGCAATAA	GTACTAAATG
Rosa	...G.....		C...A...T			
Rojo	...G.....		??????????	??????????	??????????	???.....
Azul	TGCGTCAATT	ACGAGATAAT	TCATTTTGTA	ATGTAGTTTT	CAATTTAT??	?????ATCTG
Rosa	G...	C			..T...T.CA	TGTAA.....
Rojo	G...	C			T.CA	TGTAA.....

728

Azul TTGTTTAG

Rosa

Rojo

Fig. 9: Alineación de la secuencia parcial del intrón 1 de las especies de abulón rojo, rosa y azul usadas para calcular distancias genéticas. Los puntos indican identidad, y los signos de interrogación inserciones/deleciones o información faltante.

1

Roj	GATATGGCTT	ACTTTTACGA	CTTCCTGAAA	GATAAGAACA	TGATCCCAAA	ATATCTGCCA
Azul	TA.	TT.....A.	TCC	.GG.GA..A.	CC	...CAG....
Rosa	A.	.T.....A.	C	.GG.G.....		

▼

Roj	TACATGGAAG	AAATTAATAG	AAGTAAGTAT	TGGATATTAT	TAAGAATTGA	TCTCAGTATG
Azul	..T....C.A	..C.....GC	CT.....			
Rosa	G.					

Roj	TTAATGGAGA	TGTAAACGTA	AATCCCGATG	TGCAGACAGC	GAAAATCTGC	TACATGCGGG
Azul		..C.....	...A.....			..?????????
Rosa	...?.....	..C.....				

Roj	CTTACACAGT	GATCTCAGTA	TGTTAATGGA	GATGTAAACG	TAAATCCCGA	TGTGCAGACA
Azul	..?????????	???????????	???????????	???????????	???????????	???????
Rosa		..?????????	???????????	???????????	???????????	???????????

Roj	CGAGAAATCT	GCTACATGCG	GGCTTACACA	GTGAAAAAAA	GTAACTGTA	CTGCACAGAT
Azul	???????????	???????????		??	?????.....	
Rosa	???????????	???????????	???????????	?????.....		??...

Roj	AGCTTATTTT	TTCAGTAACA	TAAAAACAGT	AAACATTGTT	CACCACAATA	ATGAAGCACT
Azul		A.....	AC	.C.....	G???	..?????????
Rosa		A.....	G....	A...	G.....	

Roj	CTATAGTCAG	CCAAAAGTTA	ACAATAGTTC	GCAAGTTGGC	GTGGCCTAAG	ACGGTGACCA
Azul	.G..???????	???????????	?????????		G..	
Rosa	T.....	A.		A.....		.A.....

Roj	TGGAAGGAAA	TCAGCGCTTG	CCCGAACCAG	TTTTGTT?AT	CCCGAATGCT	AGTAGAAGCG
Azul	G.....			T..T..		A.....
Rosa			T.....		T..	G...

Roj	TTTTGGGGCT	TTTTGCTAAT	TTTGCTATAT	GTTATCACTT	GACTATTCGG	ACGATCCAGG
Azul	?	???????????	???????????	???.....	C.G.C.	..A....G..
Rosa		.A.....	T.....			

Roj	CAGAATAGGT	CTTCAGCAAC	CCATGATTGC	CGTAAAAGGC	GACTTTACTT	GTCGTAAGAG
Azul	A.....		...C.C....	.C.....	..?????????	???????????
Rosa			C....	...?????	???????????	???????????

Fig. 10: Alineación de las secuencias del exón 4 e intrón 4 de las especies de abulón rojo, rosa y azul usadas en el cálculo de distancias genéticas. Los puntos indican identidad, los signos de interrogación indican inserciones/deleciones o información faltante. Las bases subrayadas indican repeticiones directas. La flecha indica término del exón 4 e inicio del intrón 4.

Rojo GCGACTAACG GGATCGGGTG GTCAGGCTCG CTGACTTGGT TGATGCATG? ???TCATCG
 Azul ???.....C AGTG.....
 Rosa ?????????? ?????????? ?????????? ?????????? ???????????????

Rojo ATCCCAATTG CGTGGATCGA TGTGTATGAT ATCAGTCACT GGATTGCCTG GTCAAGACTC
 AzulATC.....A.....
 Rosa ?????????? ?????????? ?????????? ?????????? ?????????? ??????????

Rojo GATTATTTAC AGACCAACGT CATATAGCTG GACTATTGCT GAGTGCGGCG T?AAAAAACA
 AzulGC...GC...???? ?..?????????
 Rosa ?????????? ?????????? ?????????? ?????????? ?????????? .C.....

Rojo ACAA?CAAAC AAACGGACGA GTAGACCGTC TTGGTTAATT TTGTCCCTGC TGAATCCGAT
 Azul ????.?.... ..?.....TA...
 RosaA..... ????......C.....

Rojo ATCGTGAAAT CACTGTCATT TTATTTATCT CTATATGTTG AGATAGTTAT ATTTCTTTAA
 Azul G.T.....C.. A..... ??????????
 Rosa G..... A.....

Rojo TATAGATCTG TGCCGATTAC GGGTCAGAAT TGATCGCCAG TACCTCATGC TTTTCTAAG
 Azul ?.....G.
 Rosa

Rojo AGCTGAGAAA GTGGGATCTG GTCGTCATGC TCGGTGACTT ATCTGA?TTG CATAACATGG
 AzulG.G.. ..TT..A..C...
 RosaG.... ..T.....T.

Rojo ATCGTTAGTA AA?TTATAAG CCCAGTCCAT ACTTGTTTGC CGATACTTGC AGTAGAAGGC
 AzulA..... ..C.... ?????? ??????????A..
 RosaA.....A..

Rojo GACTATGCTT GTCGTAAGAG GCGACTAACG GGATCGGGTG GTCAGGCTCG TTGACTTGGT
 AzulC
 Rosa

Rojo TACTGCATGT CTCCGACCTA TCCAAATTGC GTGGATCGAT GCTGATGTCA ATCATTGAAT
 Azul .GA..... ..G. ...C.....A.....
 Rosa ..A..... ..G.G. ...C.....

Rojo TGTCTGGTCC AGACTCGTTT ATTTACAAAC CGCCGTCATA TAGCTGGAAT ATTGCTGAGT
 Azul??..???? ??????.....? ??????????
 Rosa AA..... ..A.....

Fig. 10: Continuación.

Rojo GCCGCG??TT TAAACAACAA ACAAAT???? ?????????? ?????????? ??????????
 Azul ?....?AA.. ...?..???? ??????TGAA TACCCCACTA AGACACAGTA TACTGACTCC
 Rosa

Rojo ?????????? ?????????? ?????????? ?????????? ?????????? ??CATACAGC
 Azul GAGCCGACCA GTCCTTGTTT TAGCCCCTTA ATGACGAACG CCAGACAAGG TA.....
 Rosa

Rojo TAGGACTGAA TAGGTGACTA CCGCATAATA CAACAGAAAG TATTCGTTGT TTGAGAATAA
 Azul??????.A.
 Rosa C.....TCG... ..? ??????????

Rojo AGTTCATCCC TTCGTTTAA GTGTATTGTT GTGTACTCAT AATGTACTTT TTGTT?TTAT
 Azul ..G.....A....TC. ...A..... ..G...?
 Rosa ?????????? ??????.... ...?..... ..TC.?.....

Rojo TATTACAGTG CGCCCGGCAG ATGTCCCAGT TAAATACATG GGCAAGTGAA ACCCATGAAG
 Azul ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
 Rosa???? ?????????? ?????????? ?????????? ?????????? ??????????

Rojo AAGACAGCTT CATGCACCGT GCATGTCAAC AGATACAGAA AACATCGTCA AACACGGGGA
 Azul ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
 Rosa ?????????? ?????????? ?????????? ?????????? ?????????? ??????????

1643

Rojo CCTATGACAC CATCCTAACA TGG
 Azul ?????????? ?????????? ???
 Rosa ?????????? ?????????? ???

Fig. 10: Continuación.

IV.- Análisis de las Secuencias de Intrones:

IVa) Distancias Genéticas:

La diversidad de nucleótidos de las secuencias no codificantes de lisina y 18 kDa en las tres especies fueron calculadas usando el parámetro de distancia- p provisto en el programa de análisis de secuencias MEGA. La distancia genética se define como el número acumulado de diferencias de nucleótidos por sitio. El parámetro de distancia- p calcula el promedio del número de diferencias de nucleótidos por cada 100 sitios en dos secuencias. Este es un parámetro "no corregido" y fué utilizado para calcular la diversidad de nucleótidos debido a la similitud de las secuencias de intrones y el tiempo relativamente corto de divergencia evolutiva de las especies de abulón, el cual sugiere que es de 5 a 25 millones de años (Vacquier y Lee, 1993). Por lo tanto, es poco probable que haya ocurrido más de una sustitución en cualquier sitio de la secuencia de nucleótidos. El número de diferencias observado al comparar las secuencias debe ser cercano al número real de sustituciones. De esta manera, es innecesario hacer correcciones para sustituciones múltiples de nucleótidos como se recomienda en secuencias altamente divergentes.

Para calcular las distancias genéticas entre especies en el intrón 1 se utilizó una alineación de 738 nucleótidos de las secuencias totales obtenidas (Fig. 9), se empleó este número de nucleótidos debido a que este es el tamaño de la secuencia que es bastante confiable en las tres especies de manera que se calculara una distancia genética lo más acertada posible. El intrón 4 fué secuenciado completamente en las tres especies. Para la comparación de las secuencias codificantes y no codificantes en las tres especies se utilizaron las secuencias completas del exón 4 e intrón 4. La longitud total de la

secuencia fué de 1643 pb, los primeros 90 nucleótidos es secuencia del exón y los 1533 nucleótidos restantes corresponden a la secuencia del intrón, (Fig. 10). Para obtener las distancias genéticas en los abulones rojo y azul de la proteína 18 kDa se utilizó una porción del intrón 4 previamente secuenciada en el laboratorio del Dr. Vacquier. La tabla IV muestra comparaciones de las distancias en diferentes especies de abulón en las proteínas de lisina y 18 kDa. En las comparaciones hechas en lisina se observa una mayor variación en sustituciones no sinónimas que aquella en sustituciones sinónimas para exones, esta característica es mas acentuada en la proteína 18 kDa, al comparar sustituciones sinónimas de exones contra sustituciones en intrones ambas proteínas coinciden en presentar una mayor variación en exones que en intrones.

IVb) Ventanas Deslizantes:

La comparación de la diversidad de nucleótidos en exones e intrones en las diferentes especies fué representada gráficamente utilizando el programa de computadora denominado DNA sp que utiliza el método de la ventana deslizante (Rozas y Rozas, 1997). En este método se selecciona un segmento de ADN (ventana) de un determinado número de nucleótidos que es deslizado a lo largo de la secuencia dependiendo del tamaño asignado para el espacio entre ventanas, que también es de un número de nucleótidos definido. La diversidad de nucleótidos es calculada en cada ventana y este valor de diversidad es asignado al nucleótido localizado en el punto medio de la ventana. La variación de nucleótidos en intrones en las tres especies fué representada utilizando dos diferentes valores para el tamaño y espacio entre ventanas para ambos intrones, el primer par de valores fué: un tamaño de ventana de 100

nucleótidos y un espacio entre ventanas de 25 nucleótidos, el segundo par de valores fueron 50 nucleótidos para el tamaño de ventana y 10 nucleótidos para el espacio entre ventanas. Cuando el tamaño y el espacio de las ventanas son reducidos la variación de nucleótidos es calculada con mayor frecuencia y más específicamente, por lo que los picos de variación pueden ser detectados con mayor detalle. Las figuras 11a y 11b muestran las dos diferentes gráficas obtenidas para el intrón 1. Las figuras 12a y 12b muestran las gráficas obtenidas de la comparación de la variación de nucleótidos en el exón 4 e intrón 4. La ventana deslizante 100/25 en el intrón 1 muestra picos donde la diversidad de nucleótidos es tan baja como 1.4% y tan alta como 6.5%, mientras que la ventana deslizante 50/10 muestra regiones con 0% de variación y regiones con 8.6% de variación. Independientemente del tamaño de las ventanas se puede observar que la variación de nucleótidos se mantiene baja a lo largo del intrón 1. En la comparación de la variación de nucleótidos en el exón 4 e intrón 4 se encontró un patrón inusual de variación de nucleótidos: altos niveles de variación de nucleótidos dentro de la secuencia del exón, 18% en la ventana deslizante 100/25 y 22% en la ventana deslizante 50/10, la cual disminuye abruptamente en la secuencia del intrón 4. Este mismo patrón de variación se mantiene en las dos diferentes ventanas deslizantes. La variación de nucleótidos del intrón 4 es similar al encontrado en el intrón 1. Los picos de mayor variación encontrados en la secuencia del intrón 4 (13%, 11% y 10%) son considerablemente menores que los encontrados en la secuencia del exón. En el intrón 4 se encuentran también regiones sin variación de nucleótidos las cuales no se encuentran en la región del exón.

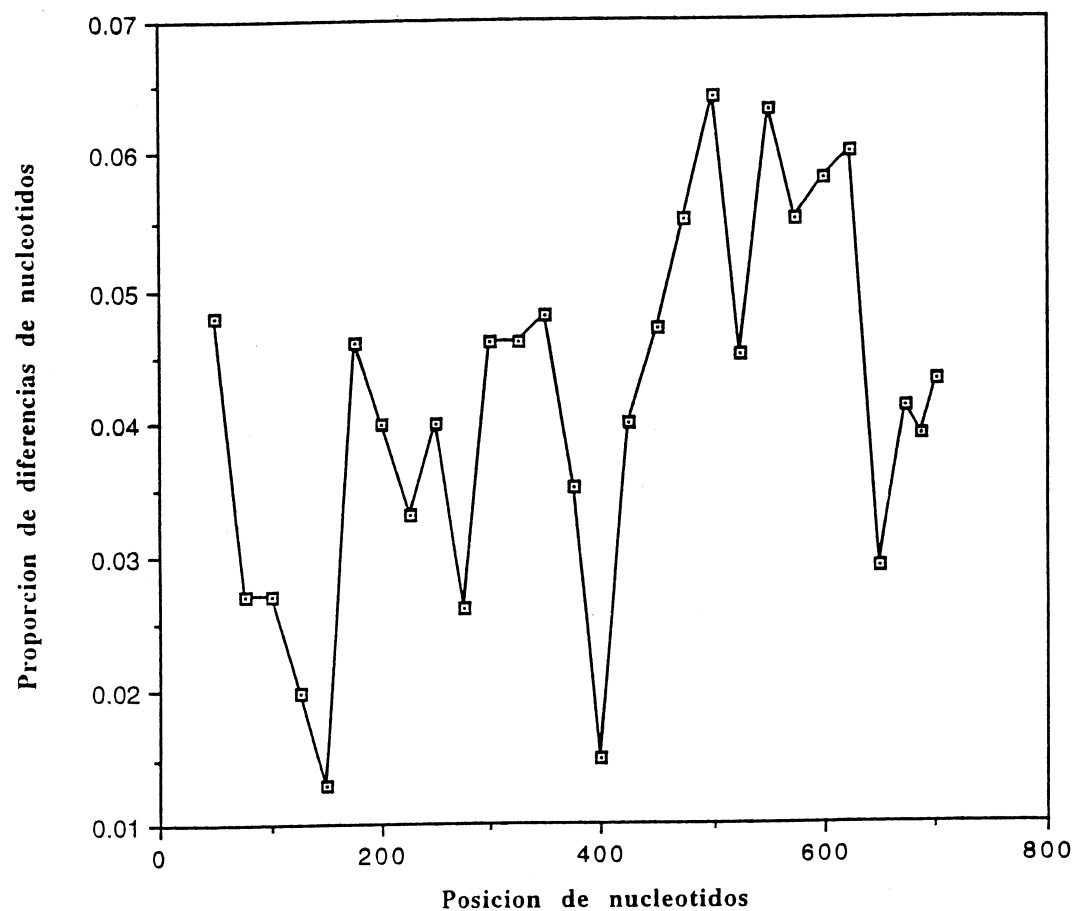


Fig. 11a: Ventana deslizante que muestra proporción de diferencias de nucleótidos en el intron 1 del gen de lisina en especies de abulón rojo, rosa y azul. Tamaño de la ventana: 100 nt; Espacio entre ventanas: 25 nt.

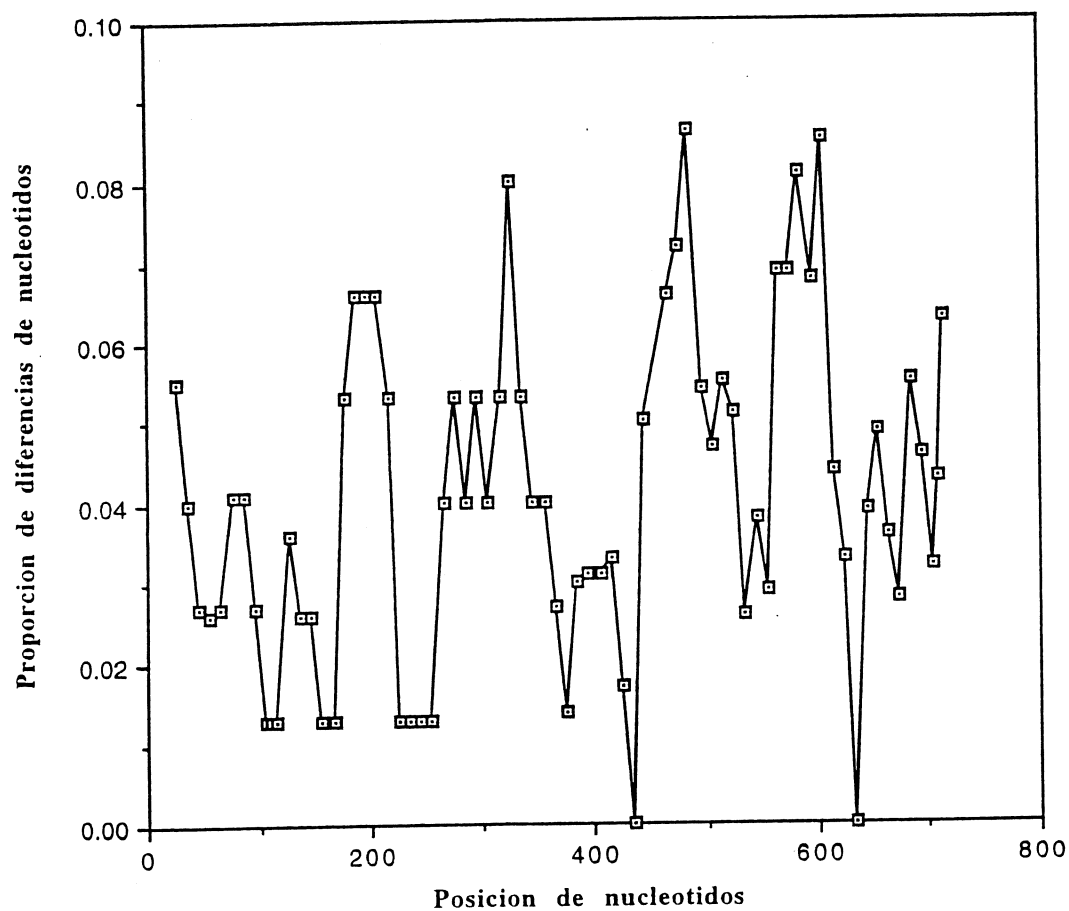


Fig. 11b: Ventana deslizante que muestra proporción de diferencias de nucleótidos en el intron 1 del gen de lisina en especies de abulón rojo, rosa y azul. Tamaño de la ventana: 50 nt; Espacio entre ventanas: 10 nt.

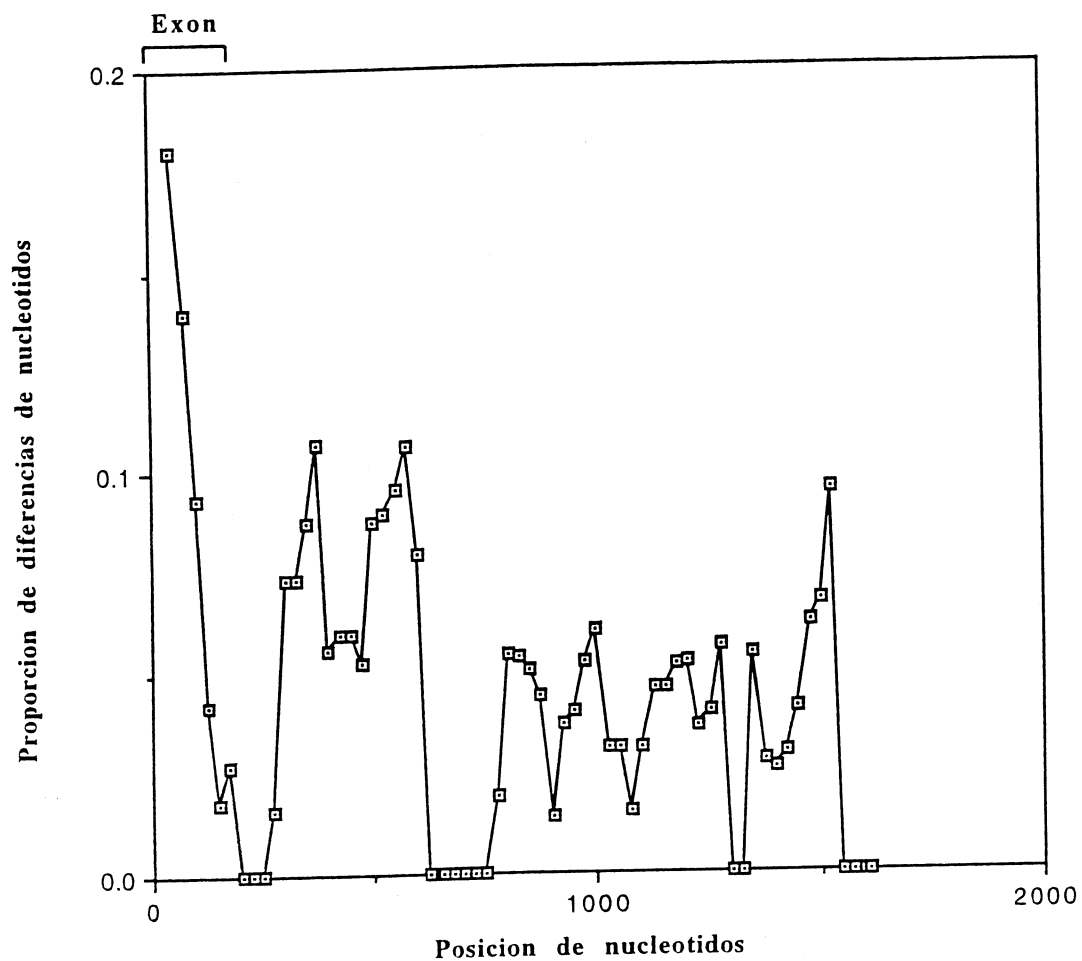


Fig. 12a: Ventana deslizante que muestra proporción de diferencias de nucleótidos en el exon 4 e intron 4 del gen de lisina en especies de abulón rojo, rosa y azul. Tamaño de la ventana: 100 nt; Espacio entre ventanas: 25 nt.

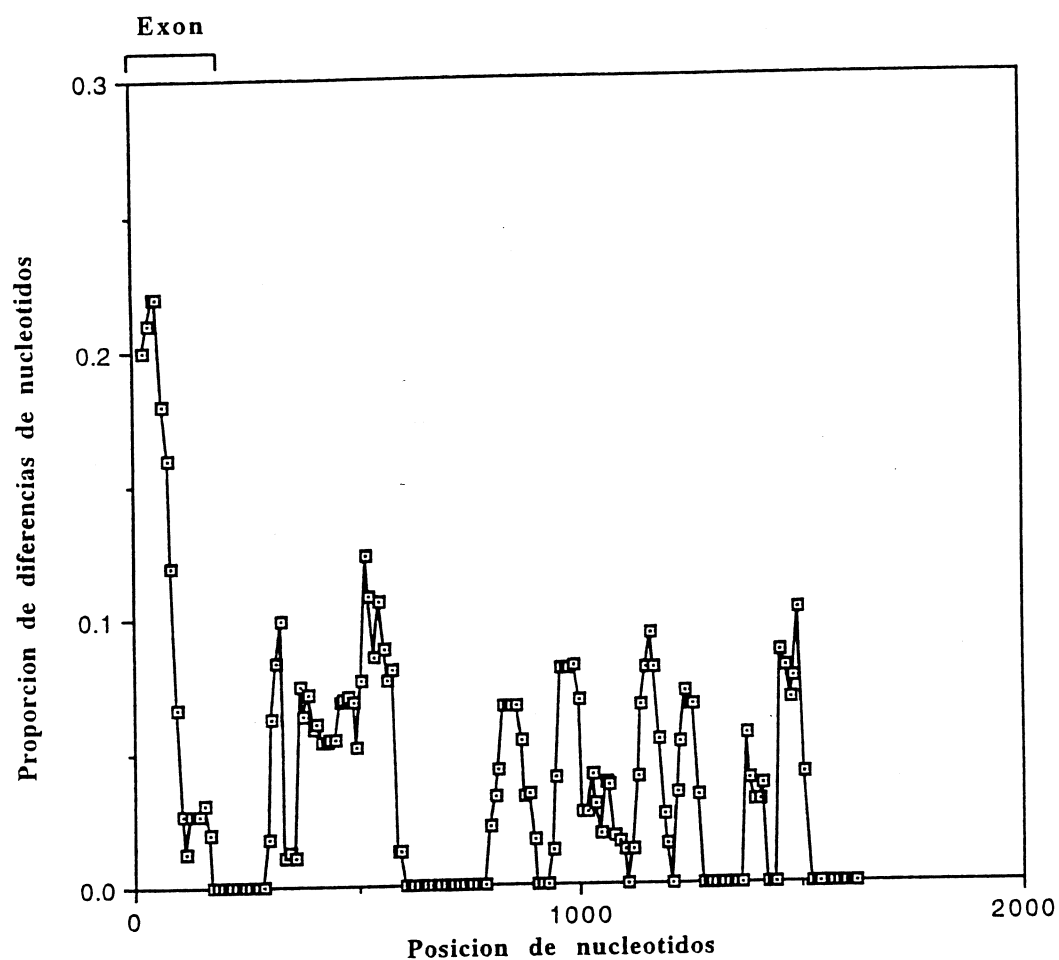


Fig. 12b: Ventana deslizante que muestra proporci3n de diferencias de nucle3tidos en el exon 4 e intron 4 del gen de lisina en especies de abul3n rojo, rosa y azul. Tama1o de la ventana: 50 nt; Espacio entre ventanas: 10 nt.

IVc) Marcos de Lectura Abiertos y Secuencias de Repeticiones Invertidas:

Los marcos de lectura abiertos (ORFs, por sus siglas en inglés) son secuencias potencialmente traducibles a proteínas y las secuencias de repeticiones invertidas son características de las secuencias de ADN que pueden definir estructuras secundarias; estas características están comúnmente relacionadas con la funcionalidad de las secuencias de ADN.

La búsqueda de ORFs en las secuencias del intrón 4 de lisina en las tres especies de abulón se realizó utilizando el programa de computadora GCG. Este programa identifica los codones de iniciación y de terminación en los tres marcos de lectura de cada una de las hebras de la doble hélice de ADN. Se encontraron pequeños marcos de lectura abiertos con poca posibilidad de codificar un producto proteico, para el abulón rojo el marco de lectura abierto de mayor tamaño fué de 98 aminoácidos (aa), para el abulón rosa de 66 aa y para el abulón azul de 95 aa, (ver figuras 13a, 13b y 13c). En el intrón 4 de las proteínas de lisina y 18 kDa de las diferentes especies se encontraron pocas secuencias de repeticiones invertidas de tamaño reducido lo cual indica que existe poca probabilidad de que los intrones desempeñen un papel estructural en dichas proteínas.

IVd) Inserciones y Deleciones:

Las inserciones y deleciones (Indels), o repeticiones directas, en la alineación de los intrones 1 y 4 de las tres especies de abulón fueron pocas y de tamaño reducido. El intrón 4 presenta indels más numerosos y más largos que el intrón 1 cuando las tres

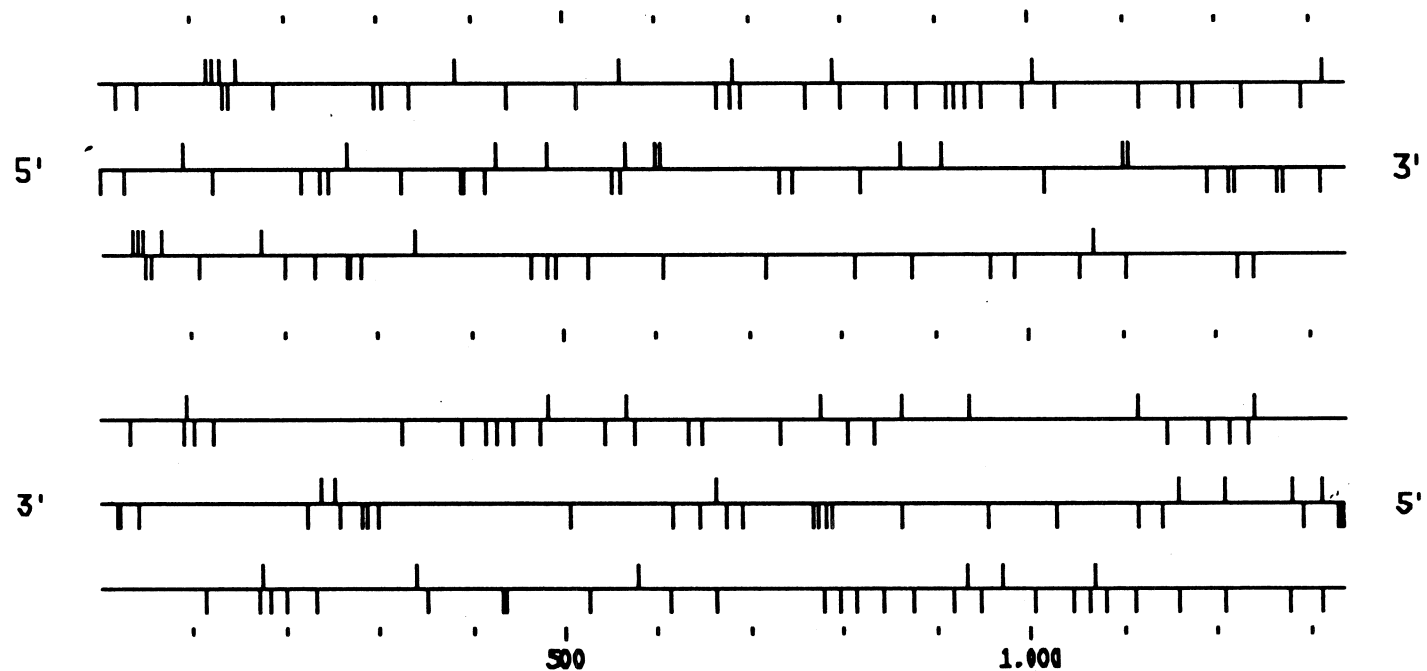


Fig. 13a: Gráfica que muestra los resultados de la búsqueda de marcos de lectura abiertos (ORFs) en la secuencia del intron 4 del abulón rojo, en los tres posibles marcos de lectura para ambas hebras de ADN. Las barras sobre la línea horizontal muestran codones de iniciación y las barras bajo la línea horizontal muestran codones de terminación. Escala: 12 mm: 100 pb.

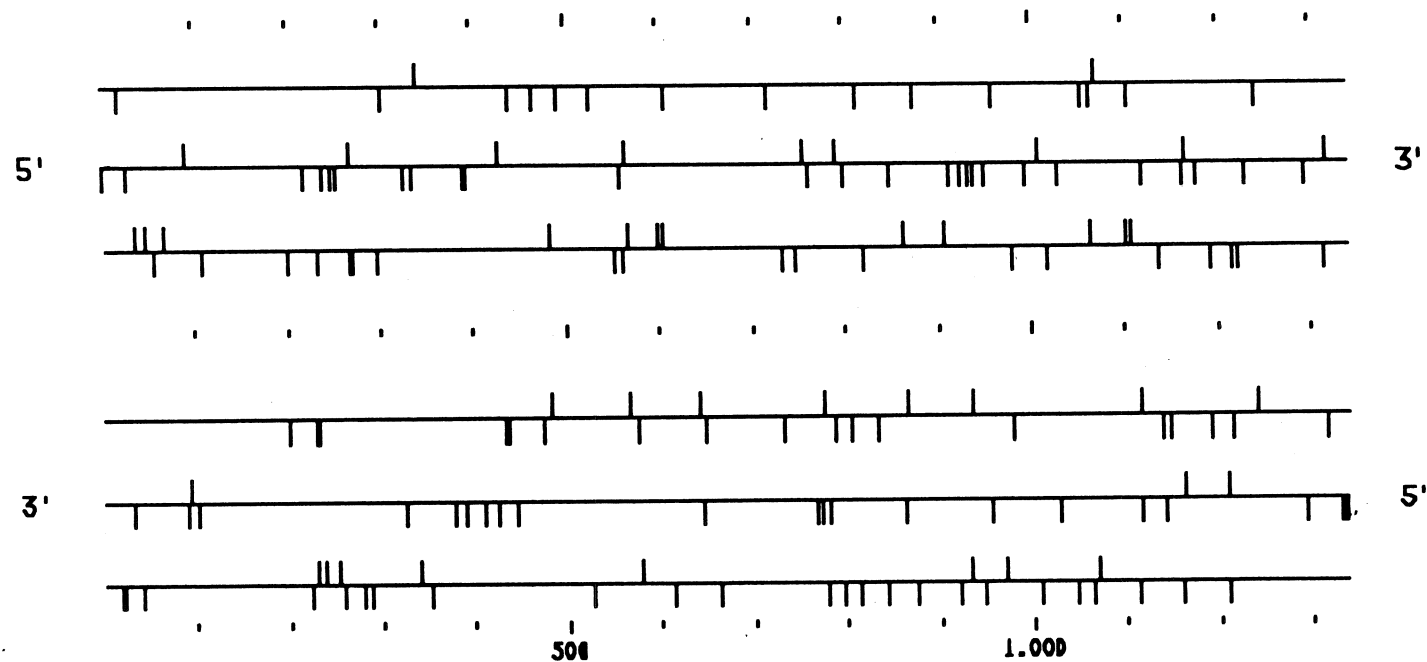


Fig. 13b: Gráfica que muestra resultados de la búsqueda de marcos de lectura abiertos (ORFs) en la secuencia del intron 4 del abulón rosa, en los tres posibles marcos de lectura para ambas hebras de ADN. Las barras sobre la línea horizontal muestran codones de iniciación y las barras bajo la línea horizontal muestran codones de terminación. Escala: 12 mm: 100 pb.

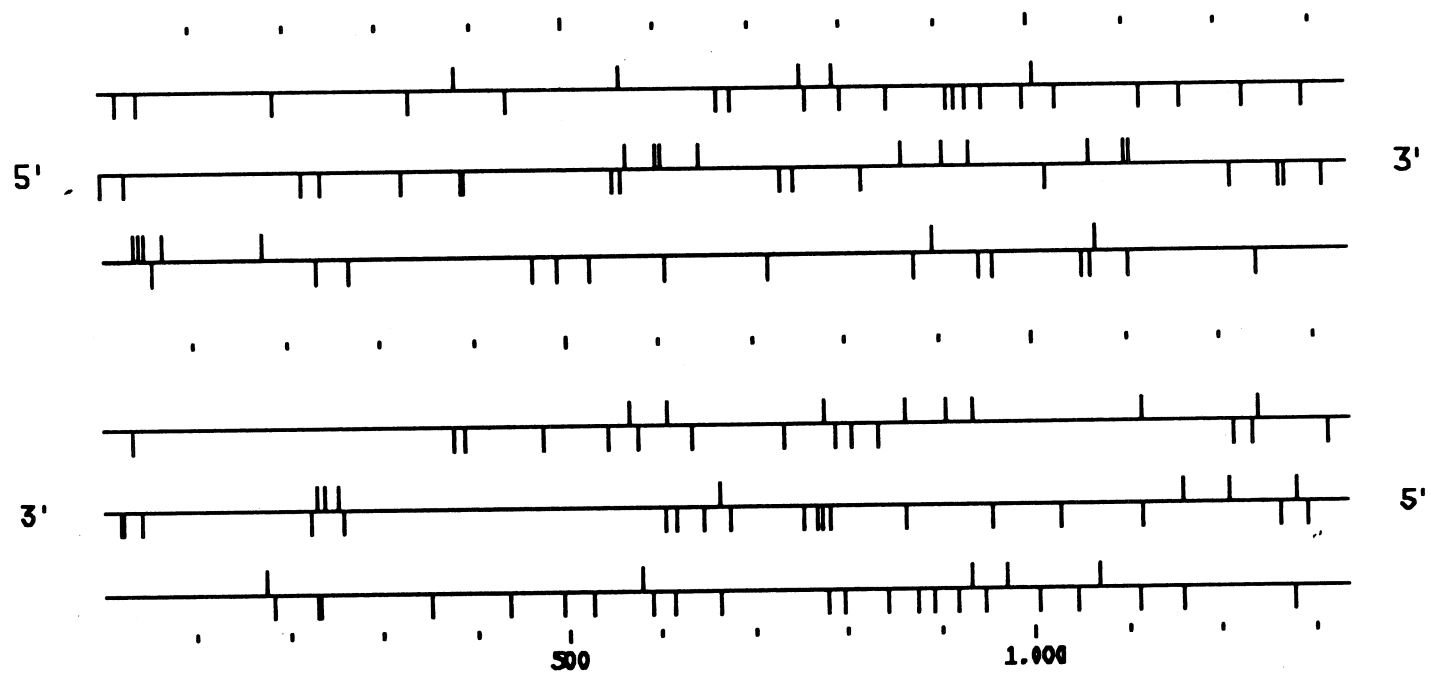


Fig. 13c: Gráfica que muestra resultados de la búsqueda de marcos de lectura abiertos (ORFs) en la secuencia del intron 4 en abulón azul, en los tres posibles marcos de lectura para ambas hebras de ADN. Las barras sobre la línea horizontal muestran codones de iniciación y las barras bajo la línea horizontal muestran codones de terminación. Escala: 12 mm: 100 pb.

especies son comparadas, pero esto puede ser resultado de la diferencia en el tamaño de las secuencias de los ambos intrones (Fig.10).

DISCUSIÓN

I.- Posición de Intrones:

Los intrones encontrados en los genes de lisina y 18 kDa son del tipo clásico, esto es, presentan los sitios conservados de unión exón-intrón (Fig. 10) el dinucleótido GT se presenta al comienzo del intrón y el dinucleótido AG se presenta al final del intrón (Rogers, 1990).

Swanson y Vacquier (1995), sugirieron que las proteínas lisina y 18 kDa provienen de la duplicación de un gen ancestral fundamentando la hipótesis en la semejanza que muestran las estructuras secundarias de ambas proteínas cuando son alineadas, lo cual indica una posible estructura terciaria similar en ambas proteínas. Estos resultados sugieren una homología evolutiva de las dos proteínas, sin embargo, esto no ha sido completamente comprobado. Los resultados de este estudio mostraron que el número y posición de los intrones en los genes de lisina y 18 kDa están conservados lo cual representa un punto de apoyo adicional a la hipótesis antes citada, ya que es poco probable que dos proteínas compartan una estructura secundaria similar, número y posición de intrones solo por azar, sin estar evolutiva y funcionalmente relacionadas. A partir de estos resultados surgió la hipótesis de que el gen ancestral de lisina tenía dos funciones: 1) la de disolver un orificio en la capa vitelina y 2) mediar la fusión de membranas. Después del evento de duplicación del gen, por efectos de la mutación y la selección, los genes originados se diferenciaron hasta especializar las funciones de las proteínas en el proceso de la fertilización de los gametos de abulón.

en moluscos primitivos, (2) el evento genético que condujo a la duplicación de este gen y (3) la localización de ambas proteínas en el genoma de abulón. Estos temas de investigación se espera, ayudarán a entender el proceso y la importancia de los eventos de duplicación de los genes en el surgimiento de nuevas funciones en las proteínas en general.

II.- Evolución Molecular en Exones e Intrones:

Los resultados obtenidos permiten concluir respecto al ADNc de lisina que las distancias genéticas, d , calculadas en las secuencias de los intrones 1 y 4 en las diferentes especies, son más pequeñas que la dn calculada en el ADNc de lisina, como se muestra en la tabla IV.

Las comparaciones efectuadas muestran que hay una menor divergencia en las secuencias de los intrones que en las secuencias de los exones. La diferencia es mucho mayor que la variación de nucleótidos en intrones. Esto también es evidente en el método de ventanas deslizantes (Fig. 12a y 12b), donde el intrón 4 presenta picos más bajos de variación que los que se presentan en el exón 4 en la comparación de las tres especies.

Dicho patrón de variación en secuencias, es inusual, ya que lo comúnmente observado en comparaciones similares, es el patrón opuesto. Esto es, alta variación en secuencias de intrones y baja variación en secuencias de exones debido a que representan las regiones codificantes de los polipéptidos. Generalmente, se supone que los intrones son secuencias no codificantes y la tasa de mutación esperada es alta, mientras que los exones se mantienen bajo una limitante que les restringe la variación

Tabla IV: Comparación entre especies del número de sustituciones no sinónimas (**dn**) y sinónimas (**ds**) por 100 sitios en la secuencia codificante y sustituciones de nucleótidos por 100 sitios (**d**) en los intrones 1 y 4 de los genes de lisina y 18 kDa en abulón rojo (Hr), rosa (Hc) y azul (Hf).

Comparación de Especies	dn en Secuencias Codificantes	ds en secuencias codificantes	d en intrón 1	d en intrón 4
Lisina				
Hr-Hc	14.6±2.3	10.6±3.5	2.98±0.6	3.9±0.6
Hr-HF	24.9±3.2	21.3±5.3	4.2±0.8	7.7±0.9
Hc-Hf	24.0±3.2	16.0±4.4	5.3±0.9	7.4±0.9
18 kDa				
Hr-Hf	81.1±8.5	78.2±15.4	3.4±0.9	6.8±1.5

(Li y Graur, 1991). Debido a que los intrones mutan rápidamente en la mayoría de los genes, dn en exones es menor que d en intrones. En 18 kDa dn es bastante mayor que d (Tabla IV). Este patrón poco común de variación puede ser explicado en términos de una selección positiva que actúa en los sitios no sinónimos de los exones para alterar la secuencia de aminoácidos y de esta manera conducir a un rápido proceso de evolución de estas proteínas.

Existen solo dos casos publicados con un patrón de variación de nucleótidos similar, que se han encontrado en genes que codifican para la fosfolipasa A₂ (PLA₂) presentes en dos diferentes especies de víboras (Nakashima *et al.* 1993 y John *et al.* 1994). En las regiones codificantes de estos genes los niveles de dn son mayores que los niveles de ds , esto es, tal como se observa en la tabla IV para lisina y 18 kDa, estos están sujetos a selección positiva. En este caso, el principio denominado "carrera de armas" puede explicar la acción de la selección positiva en los genes de PLA₂ del veneno de víbora: existe una presión selectiva en los genes que codifican para el veneno de manera que se mantenga su efectividad en contra de la presa, sin embargo, la presa puede desarrollar resistencia al veneno, la cual a su vez, provoca que los genes del veneno se mantengan cambiando. Esto conduce a un rápido proceso evolutivo tanto en los genes del veneno como en los receptores del veneno de la presa. En los genes de PLA₂ el alterar la secuencia de aminoácidos tiene un valor adaptativo, sin embargo, en los genes de lisina y 18 kDa no existe una explicación para la alta variación de nucleótidos presente. Se han propuesto varias hipótesis (Vacquier *et al.* 1997) pero ninguna ha sido probada aún. La pregunta sigue sin resolver: ¿Que ventaja selectiva confiere la rápida evolución de las proteínas de fertilización a los gametos de abulón?

Este tópico es un amplio campo abierto para futuras investigaciones. Su estudio será de ayuda en la determinación de los mecanismos que intervienen en la creación de barreras para la fertilización heteroespecífica, aislamiento reproductivo y especiación de invertebrados marinos en un ambiente que carece de barreras geográficas evidentes.

La selección positiva puede explicar la alta variación de nucleótidos en los exones, pero ¿cómo puede ser explicada la baja variación de nucleótidos en los intrones de estas proteínas? John *et al.* (1994) y Clark *et al.* (1996) han sugerido que la conservación de los nucleótidos en intrones se debe a un importante papel desarrollado por los intrones en la estabilidad del ARNhn y/o en el procesamiento de éste. Si los intrones desempeñan un papel crítico en los procesos de maduración del ARN se esperaría encontrar en estas secuencias una estructura secundaria definida. Con el fin de conocer si este era el caso de las secuencias de intrones de lisina y 18 kDa se realizaron análisis para determinar estructuras secundarias y se encontraron pocas secuencias de repeticiones invertidas, esto significa que existe poca probabilidad de que estos intrones intervengan en la formación de alguna estructura secundaria involucrada en dichos procesos. Hay una posibilidad de que la conservación de nucleótidos en la secuencia de intrones se deba a un papel codificante de algún producto, sin embargo, la búsqueda de ORFs no mostraron evidencia para apoyar esta idea. Otra explicación probable para esta baja variación en intrones es que las tasas de mutación en tales genes son bajas. En los intrones las mutaciones ocurren al azar y algunas se pierden y otras se fijan como en cualquier proceso neutral, sin embargo, en los exones se fijan más mutaciones de las que se pierden debido a que éstas de alguna manera confieren una ventaja selectiva a los gametos en el proceso de fertilización.

En la mayoría de los genes, ds es aproximadamente igual a d (Li y Graur, 1991) debido a que se supone que los cambios en sitios silenciosos y los sitios de los intrones reflejan la tasa de mutación del gen, sin embargo, en las proteínas de lisina y 18 kDa las sustituciones sinónimas (ds) son mayores que d en los intrones. En los genes de PLA2 se puede observar un patrón similar, la variación de nucleótidos en intrones es aproximadamente un cuarto de las sustituciones sinónimas (Nakashima *et al.* 1993). Aún no existe una explicación para este patrón, pero puede estar relacionada con los altos niveles de dn presentes en los exones. Estos resultados indican que esta variación pudiera ser un patrón general de las proteínas sujetas a altos niveles de selección positiva.

El presente estudio muestra que el análisis de la evolución molecular de proteínas que se encuentran bajo la acción de selección positiva es un campo novedoso con características únicas que puede ayudar a entender los procesos evolutivos de proteínas que intervienen en el desarrollo de funciones vitales para el organismo. En las proteínas de fertilización tales como lisina y 18 kDa dicho estudio puede arrojar información valiosa para la comprensión de mecanismos biológicos relevantes como lo son: la especificidad de especie en los procesos de fertilización y la especiación en los invertebrados marinos.

CONCLUSIONES

- 1) Se determinó la presencia de intrones en las proteínas de lisina, se estableció su secuencia y se les analizó con programas computacionales. Estos fueron comparados con los intrones de la proteína 18 kDa. En el gen de la lisina se ubicaron tres intrones. Uno más permanece aún sin determinarse.
- 2) La alineación de las secuencias de aminoácidos de las proteínas lisina y 18 kDa indica que las posiciones de los intrones determinados en lisina coinciden notablemente con tres de los intrones de 18 kDa.
- 3) A partir del análisis citado en el punto anterior, surge la hipótesis de que las proteínas lisina y 18 kDa provienen de un evento mutacional de duplicación de genes.
- 4) La tasa de sustitución de nucleótidos no sinónimos (dn) en exones de lisina y 18 kDa es mayor que la tasa de sustitución de nucleótidos (d) en intrones, lo cual indica un patrón de variación de nucleótidos inusual: mayor variación de nucleótidos en exones que en intrones.
- 5) Los resultados del análisis de la estructura secundaria y la búsqueda de ORFs en el intrón 4 de lisina muestran poca evidencia del desempeño de algún papel funcional de estas secuencias.

COMENTARIOS FINALES

i) Una probable explicación para esta baja variación en intrones es una baja tasa de mutación en el gen de lisina, sin embargo, un mayor número de mutaciones quedan fijas en los exones que en los intrones debido a la acción de la selección positiva, originando este inusual patrón de variación.

ii) Se han encontrado ejemplos de este patrón de variación en los genes de PLA₂ que codifican para el veneno en dos especies de víbora, por lo que se puede proponer que este es un patrón general que se presenta en proteínas con altos niveles de selección positiva.

iii) Las fuerzas selectivas que promueven este patrón en las proteínas de lisina y 18 kDa aún son desconocidas y se requiere el desarrollo de estudios futuros en estas importantes proteínas de fertilización.

BIBLIOGRAFIA

ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit. 1995. Applied Biosystems Inc. Perkin Elmer Corporation. U.S.A.

Clark, A.G., B.G. Leicht y S.V. Muse. 1996. Length variation and secondary structure of introns in the *Mlc1* gene in six species of *Drosophila*. *Mol. Biol. Evol.* 13(3):471-482.

Engels, B. 1993. Amplify Computer Program. University of Wisconsin. Madison, WI. U.S.A.

Genetics-Computer-Group (GCG Computer Program). 1991. Genetics Computer Group Package. Genetics Computer Group. Madison, WI.

Hughes A.L. y M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*. 335:167-170.

Hughes A. y M. Yeager. 1997. Comparative evolutionary rates of introns and exons in murine rodents. *J. Mol. Evol.* 45:125-130.

John T.R., L. Smith y I. Kaiser. 1994. Genomic sequences encoding the acidic and basic subunits of Mojave toxin: unusually high sequence identity of non-coding regions. *Gene* 139:229-234.

Kessing B.C., Mac Intosh C., Martin M.A., Mac Millan W., and Palumbi S. 1989. The simple fool's guide to PCR. Department of Zoology. University of Hawaii. Honolulu. HI.

- Kimura, M. y T. Ohta. 1974. On some principles governing molecular evolution. *Proc. Nat. Acad. Sci. U.S.A.* 71(7):2848-2852.
- Lee, Y.H. y V.D. Vacquier. 1992. The divergence of species-specific abalone sperm lysins is promoted by positive Darwinian selection. *Biol. Bull.* 182:97-104.
- Lee, Y.H., T. Ota y V.D. Vacquier. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* 12(2):231-238.
- Lewis, C.A., D.L. Leighton y V.D. Vacquier. 1982. A protein from abalone sperm dissolves the egg vitelline layer by a non-enzymatic mechanism. *Develop. Biol.* 92:227-239.
- Li, W.-H. y D. Graur. 1991. Fundamentals of molecular evolution. Sinauer Associates. Sunderland Mass. U.S.A. 284 p.
- Metz, E.C., R.E. Kane., H. Yanagimachi y S.R. Palumbi. 1994. Fertilization between closely related sea urchins is blocked by incompatibilities during sperm-egg attachment and early stages of fusion. *Biol. Bull.* 187:23-34.
- Metz, E.C. y S.R. Palumbi. 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Mol. Biol. Evol.* 13(2):397-406.
- Metz, E.C., G. Gómez-Gutiérrez y V.D. Vacquier. 1998. Mitochondrial DNA and bindin gene sequence evolution among allopatric species of the sea urchin genus *Arbacia*. *Mol. Biol. Evol.* 15(2) 185-195.
- Moy, G.W. y V.D. Vacquier. 1979. Immunoperoxidase localization of bindin during the adhesion of sperm to sea urchin eggs. *Curr. Top. Dev. Bio.* 13, 31-34.

Nakashima K., T. Ogawa., N. Oda, M. Hattori, Y. Sakaki., H. Kihara and M. Ohno. 1993. Accelerated evolution of *Trimeresurus flavoviridis* venom gland phospholipase A₂ isozymes. *Proc. Natl. Acad. Sci. U.S.A.* Vol. 90. pp. 5964-5968.

Palumbi, S.R. 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annu. Rev. Ecol. Syst.* 25:547-572.

Quiagen, Inc. 1997. Santa Clarita, CA. U.S.A.

Rogers H. John. 1990. The role of introns in evolution. Federation of European Biochemical Societies. Vol. 268, no.2, 339-343.

Rozas J. y R. Rozas. 1997. DNA Sequence Polymorphism. Version 2.2. Universidad de Barcelona.

Sambrook J., E. F. Fritsch y T. Maniatis. 1989. Molecular Cloning. 2nd Edition. Cold Spring Harbor Laboratory Press. U.S.A.

Sequence Editor (Seq Ed) version 1.0.3. 1995. Applied Biosystems Inc. Perkin Elmer Corporation. U.S.A.

Sudhir K., K. Tamura y M. Nei. 1993. MEGA: Molecular Evolutionary Genetics Analysis, Version 1.0. The Pennsylvania State University, University Park, PA 16802.

Swanson, W.J. y V.D. Vacquier. 1995. Extraordinary divergence and positive Darwinian selection in a fusogenic protein coating the acrosomal process of abalone spermatozoa. *Proc. Natl. Acad. Sci. U.S.A.* 92:4957-4961.

Swanson, W.J. y V.D. Vacquier. 1995b. Liposome fusion induced by a Mr 18 000 protein localized to the acrosomal region of acrosome-reacted abalone spermatozoa. *Biochemistry*, Vol. 34, No. 43: 1402-1408.

Vacquier, V.D., K.R. Carner, y C.D. Stout. 1990. Species-specific sequences of abalone lysin, the sperm protein that creates a hole in the egg envelope. *Proc. Natl. Acad. Sci. U.S.A.* 87:5792-5796.

Vacquier, V.D. and Y-H. Lee. 1993. Abalone sperm lysin: unusual mode of evolution of a gamete recognition protein. *Zygote*, 1:181-196.

Vacquier, V.D., W.J. Swanson y M.E. Hellberg. 1995. What have we learned about sea urchin sperm bindin?. *Develop. Growth Differ.* 37:1-10.

Vacquier, V.D., W.J. Swanson y Y.H. Lee. 1997. Positive Darwinian selection on two homologous fertilization proteins: What is the selective pressure driving their divergence?. *J. Mol. Evol.* 44(Suppl 1): S15-S22.

Yang, J.K., J. Masters y G. Attardi. 1984. Human Dihydrofolate Reductase Gene Organization. *J. Mol. Biol.* 176: 169-187.