CENTRO DE INVESTIGACIONES BIOLOGICAS
DEL NOROESTE. S. C.

Programa de Estudios de Posgrado

"Obtención y Análisis de Secuencias de Genes Expresados en Hepatopáncreas de *Penaeus vannamei* a partir de Oligonucleótidos Diseñados para Identificar Inhibidores de Proteinasas"

T E S I S

Que para obtener el grado de

Maestro en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales
(Orientación en Biotecnología)

Presenta

Juana Meliza Gutiérrez Ayala

La Paz, Baja California Sur, Agosto del 2003.

RESUMEN

En la actualidad, a pesar de la importancia económica que ha representado el cultivo del camarón en las últimas décadas, no se conoce mucho acerca de su organización genómica básica. Entre las tecnologías de la bioquímica aplicables a la acuacultura está la genómica, la cual permite encontrar genes útiles para el mantenimiento y aprovechamiento de los organismos. Las secuencias expresadas, son una herramienta disponible en la investigación para estudios de fisiología, genética, evolución y mapeo genético de una especie, lo cual con respecto al empleo de prácticas tradicionales en acuacultura proporcionan nuevas perspectivas.

Para el presente trabajo se diseñaron seis iniciadores degenerados específicos para detectar transcritos de genes que codifican para inhibidores de proteinasas en el hepatopancreas del camarón blanco *Penaeus vannamei*. Se utilizaron dos metodologías para cumplir nuestros objetivos.

Primeramente, para obtener e identificar los transcritos se sacrificaron camarones obtenidos de un estanque en condiciones optimas de cultivo. Se extrajo ARN de hepatopáncreas y se utilizó como templado en un RT-PCR con todos los iniciadores. Se clonaron los productos amplificados y se secuenciaron clonas positivas. Se obtuvieron 50 transcritos diferentes que codifican al menos para 35 genes distintos. Las proteínas deducidas tuvieron similitudes entre 25% y 50% contra las reportadas en el Banco Genómico de datos del NCBI, sugiriendo la expresión de proteínas de membrana y/o estructurales, además de otras proteínas como la hemicentina, la mucina intestinal de mamíferos, un precursor del factor de coagulación en *Takifugu rubipes*, algunas proteínas de choque térmico no reportadas para la especie en estudio y proteínas con función aún no descrita en *Drosophila melanogaster* y en otros artrópodos. Solo dos secuencias, mostraron identidad con secuencias ya reportadas para *P. vannamei*, una que sugiere una proteína relacionada con el control de la presión de la hemolinfa y la osmorregulación y otra que codifica para la subunidad 18s de ARN ribosomal.

Por otro lado, para determinar la expresión de estos genes bajo condiciones de estrés fisiológico inducido, se sacrificaron camarones sometidos a diferentes períodos de ayuno. Se extrajo ARN de hepatopáncreas y se utilizó como templado en un RT-PCR con solo uno de los iniciadores. La electroforesis mostró variación en la expresión de transcritos. Se purificaron 4 amplificaciones de diferente tamaño en gel de agarosa y se secuenciaron con el iniciador especifico. Se obtuvo una secuencia de 154pb no correspondiente a ninguna reportada para inhibidores de proteinasas. Los tres fragmentos restantes mostraron inespecificidad dentro de las secuencias, la cual puede ser justicada por la elevada variación de transcritos encontrada en los productos clonados a partir de las reacciones de RT-PCR realizadas con los 6 diferentes iniciadores.

Indudablemente la presente investigación, arroja grandes aportes y perspectivas al conocimiento de la composición genética de *P. vannamei* y nos da la pauta para iniciar la caracterización de diferentes genes expresados en el hepatopáncreas de camarón, para elucidar ampliamente en las funciones que este órgano realiza.

Palabras clave: *P. vannamei,* secuencias expresadas, inhibidores de proteinazas

Atentamente: Dr. C. Humberto Mejía Ruiz

ABSTRACT

At the present, although the economic importance that the shrimp culture have represented in the last decades, the knowledge about its basic genomic organization is very poor. The genomic, is one of the biochemical technologies that could be applied in aquaculture. The genomic permits us to find useful genes for the maintenance and exploitation of the organisms. The expressed sequences are an available tool for research in physiology, genetic, evolution and genetic mapping of the species, allowing the improvement of the traditional production practices.

For this work, six degenerated primers were designed to detect gene transcripts that codify for proteinase inhibitors in the hepatopancreas of white shrimp *Pennaus vannamei.* Two different methodologies were followed to achieve our aims. To obtain and identify such transcripts, shrimps grown under optimal conditions were sacrificed for hepatopancreas RNA extraction. The isolated RNA was used as template for cDNA amplification by RT-PCR technique using all the designed primers. The amplified products were cloned and the positive clones were sequenced. We obtained 50 different transcripts that codify at least for 35 distinct genes. The deduced proteins had homologies between 25 % and 35 % with reported sequences from the NCBI Gene Bank. These homologies suggested us the expression of membrane and/or structural proteins, as well as the expression of other proteins like hemicentin, mammal intestinal muccine, a coagulation factor precursor in *Takifugu rubipes*, some no reported heat shock proteins for this specie, and other proteins of still unknown function reported in *Drosophila melanogaster* and other arthropods. Only two of the all obtained sequences showed identity with sequences reported previously in *P. vannamei*, from which, the first seems to be a protein related with control of the blood pressure and osmoregulation, the second one is a protein that codifies for the 18s subunit of the ribosomal RNA.

On the other hand, to determine the expression of these genes under physiological induced stress, a group of organisms were selected from different starvation periods. From the hepatopancreas of these organisms RNA was isolated and used as template for RT-PCR amplification. For this assay only one of the six degenerated primers was used. The agarose electrophoresis, of the amplified products, showed variations among the transcripts expression. From the amplified fragments, 4 fragments of different length were selected and purified. These fragments were sequenced using a specific primer. We obtained a sequence of 154 bp that do not present homology with other proteinase inhibitors. The resting 3 fragments had a high inespecifity degree, which could be justified, by the great transcript variation found in the cloned products from the RT-PCR analysis realized with the 6 different primers.

This research certainly is a great contribution to the knowledge of the shrimp genetics. The genes characterization of the expressed sequences obtained from the hepatopancreas shrimp will be our following objective in order to elucidate the unknown multifunctional role of this organ.

Key words: *P. vannamei,* expressed sequences, proteinase inhibitors

Dr. C. Humberto Mejía Ruiz

DEDICATORIA

A mi esposo, por el apoyo y el sacrificio que significó haber estado separados

AGRADECIMIENTOS

A mi director de tesis y comité tutorial.

A la Doctora Alejandra, por la motivación, disponibilidad y el gran apoyo técnico brindado.

Al CIBNOR y a todas las personas involucradas en el desarrollo de este trabajo en las diferentes áreas de esta institución por la accesibilidad y la disponibilidad de apoyar: Postgrado, Centro de cómputo, Biblioteca, Estanquería y litorales, Biología Molecular y a todos los investigadores, técnicos y estudiantes que formaron y forman parte del edificio de Patología Marina.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado

A mis compañeros de maestría e investigadores involucrados en nuestra preparación académica.

A los amigos que se convirtieron en mi familia: Sandra, Tania, Licha, Christy, Hever, Pedro, José Luis, Jesús, Alberto y por su supuesto.

A mi súper amigo "El Ponchito" Anguiano.

A Susana Ávila y Patty Hinojosa.

A mis hermanitos: Abel, Norma y Sarahí

CONTENIDO

Resumen ii
Abstract iii
Dedicatoria iv
Agradecimientos v
Contenido vi
Índice de Figuras viii
Índice de Tablas viii
Glosario ix
I. INTRODUCCIÓN
1. Generalidades 1
1.1 *Penaeus vannamei*. 1
1.2 Camaronicultura
1.2.1 Historia 5
1.2.2 Producción 6
1.2.3 Investigación y Desarrollo Tecnológico 7
1.3 Enfoque de la Investigación camaronícola actual 9
2. Antecedentes 12
2.1 Alimentación en Peneidos 12
2.2 Hepatopáncreas. Morfología y Función 13
2.3 Inhibidores de proteinasas 16
2.4 Expresión Tejido-Específica 18
2.4.1 Ácido Ribonucleico (ARN) 21
2.4.2 Diseño de oligonucleótidos sintéticos para Transcripción 22
Reversa (RT) y Reacción en Cadena de la Polimerasa (PCR)
DIAGRAMA DE FLUJO (Objetivos, Metodología y Resultados) 25
3. Objetivo 26
II. MATERIALES Y MÉTODOS
Diseño de Oligonucléotidos o Iniciadores 27
Obtención de organismos 27
Disección de organismos, extracción de RNA total y purificación de mensajero 28
RT-PCR 29
Purificación de productos 31
Producción de células competentes 32
Ligación 32
Transformación de células 33
Selección de clonas positivas 33
Extracción y purificación de ADN plasmídico. Método de Lisis Alcalina 34
Preparación de las muestras para secuenciación 35

III. RESULTADOS
Aislamiento de ARN y RT-PCR 36
Amplificación y secuenciación de transcritos obtenidos de Hepatopáncreas de camarones en diferentes períodos de ayuno con iniciadores diseñados para obtener secuencias nucleotídicas codificantes para proteínas inhibidoras de proteinasas tipo Kazal. 37
Secuenciación de transcritos clonados obtenidos a partir de 6 iniciadores diseñados para obtener secuencias nucleotídicas codificantes para proteínas inhibidoras de proteinasas. 41
Análisis de secuencias 42
IV. DISCUSIÓN 52
V. CONCLUSIÓN 62
VI. PERSPECTIVAS 63
VII. BIBLIOGRAFÍA 64
VIII. APÉNDICE
1. Ejemplo de un alineamiento de secuencias de aminoácidos que codifican para inhibidores de proteinasas tipo Kazal en vertebrados para el diseño de iniciadores a partir de una secuencia de una región conservada. 77
2. Secuencias nucleotídicas obtenidas a partir de iniciadores diseñados para inhibidores de proteinasas. 79
3. Iniciadores degenerados. Una guía corta. 86

INDICE DE FIGURAS

Figura 1. Características morfológicas del cefalotorax de *Penaeus vannamei.* 2
Figura 2. Electroforesis en gel de agarosa-TAE 2.5% de un RT-PCR realizado a partir de RNA total extraído de hepatopáncreas de camarón amplificado con diferentes primers diseñados de regiones conservadas de inhibidores de proteinasas. 37
Figura 3. ARN total de hepatopáncreas de *Penaeus vannamei*, en diferentes periodos de ayuno (24, 48 y 72 horas). 38
Figura 4. Electroforesis en gel de agarosa-TAE 2% mostrando la expresión bajo diferentes periodos de ayuno. 38
Figura 5. Alineamiento de secuencias homólogas al fragmento de 154 pb 40
Figura 6. Alineamiento de una secuencia de aminoácidos deducida de un producto de RT-PCR secuenciado a partir de un iniciador degenerado diseñado para un inhibidor de proteinasas tipo kazal. 41
Figura 7. ADN plasmidico recombinante obtenido por el método de lisis alcalina. 42
Figura 8. PCR del plásmido contenido en las clonas positivas. 43

INDICE DE TABLAS

Tabla I. Iniciadores diseñados a partir de un fragmento conservado de la secuencia de aminoácidos reportados para inhibidores de proteinasas. 36
Tabla II. Relación de clonas obtenidas de los productos de RT-PCR con cada iniciador. 43
Tabla III. Relación de clonas seleccionadas para secuenciar. 44
Tabla IV. Secuencia de los iniciadores localizados a los extremos de los amplicones obtenidos. 45
Tabla V. Análisis de secuencias. 46
Tabla VI. Datos bioquímicos acerca del péptido deducido a partir de la secuencia nucleotídica utilizada como iniciador. 59

GLOSARIO

ADN: Ácido desoxirribonucleico. La cadena más grande de moléculas en la mayoría de las células, la cual lleva el mensaje genético y controla las funciones celulares en la mayoría de las formas de vida.
ARNm: Mensajero de ácido desoxirribonucleico. Se transcribe de un gen que codifica para una proteína. La información decodificada en la molécula de ARNm es traducida a un polipéptido de una secuencia de aminoácidos específica por el ribosoma. En eucariotas los ARNm transfieren información genética del ADN a los ribosomas, donde es traducida a proteína.
Clonación: aislamiento y multiplicación de un gen determinado.
Competente: Célula capaz de tomar ADN exógeno y, por tanto susceptible de ser transformada.
Electroforesis: Técnica que separa a las moléculas cargadas y de peso molecular distinto como ADN y ARN o proteínas; con base en la migración relativa en una matriz apropiada (gel de agarosa o poliacrilamida), sometida a un campo eléctrico.
Kilobase (Kb): Unidad de longitud de ácidos nucleicos correspondiente a 1000 nucleotidos. Se abrevia como Kb para ácidos nucleicos de hebra simple y Kbp (kilo pares de bases) para los de doble hebra.
Ligasa: Enzima que puede unir de nuevo un enlace fosfodiester roto de un ácido nucleico.
Oligonucleótido o cebador: Fragmento corto de ADN o ARN alineado a un templado de una sola hebra de ADN proveyendo en su extremo 3' un terminador hidroxilo del cual la ADN polimerasa extiende una nueva hebra de ADN para producir una hebra doble.
PCR (siglas en ingles-Polimerase Reaction Chain): Reacción en Cadena de la Polimerasa
Plásmido: Molécula de ADN extracromosómico que se replica autónomamente.
Promotor: Región reguladora situada a corta distancia del extremo 5' de un gen y que actúa como sitio de unión de la polimerasa de ARN.
Secuenciación: Determinación del orden de los nucleotidos en una molécula de ADN o ARN, o los aminoácidos en una cadena polipeptídica.
Secuencia Río Abajo: secuencias nucleotídicas de un gen localizadas en la región que se transcribe. Se numeran a partir del nucleótido de inicio de la trascripción como +1, +2, y así sucesivamente.
Secuencia Río Arriba: Secuencias nucleotídicas de un gen localizadas anteriormente a la región que se transcribe. Se enumeran desde el nucleótido de inicio de la trascripción como –1, -2, y así sucesivamente, ejemplo, nucleótidos del promotor.
Tm: temperatura de alineamiento, o temperatura a la cual una doble hebra de ADN o una molécula de ARN se desnaturaliza en hebras sencilla. Es característica de cada especie de ADN y da un indicio de su composición de bases. ADNs ricos en pares de bases con G y C con tres puentes de hidrógeno son más resistentes a desnaturalizaciones termales que ADNs ricos en A y T con dos puentes de hidrógeno.

I. INTRODUCCION

1. Generalidades

El camarón blanco del pacífico, es una de las especies acuacultivables más importantes en la región del Mar de Cortés, debido al gran aporte económico que se ha obtenido mediante su cultivo.

1.1 *Penaeus (Litopenaeus) vannamei.*

TAXONOMIA

Pertenece a la clase de los crustáceos que son organismos artrópodos mandibulados con apéndices birrámeos articulados, con dos pares de antenas, y caparazón. Presenta el rostro provisto de 8 a 9 dientes dorsales. En el cefalotórax posee una ranura y carina adrostrales cortas terminándose a nivel del diente epigástrico. El petasma tiene el borde libre del lóbulo lateral elipsoidal, rebasando largamente el lóbulo medial en vista lateral, télico presentado en la parte anterior del esternito XIV con dos prominencias oblicuas proyectadas hacia el interior y en forma de pequeños lóbulos (Hendrickx, 1996).

Los miembros del género *Penaeus* han sido divididos por Pérez Farfante desde 1969 en cuatro: *Penaeus*, *Litopenaeus*, *Fanneropenaeus* y *Melicertus.* Por tanto, el camarón blanco del pacifico, queda clasificado como *Litopenaeus vannamei* (Pérez-Farfante y Kensley, 1997). En el presente documento, salvo en las citas que así lo hayan considerado, se le seguirá nombrando de la manera clásica, ya que a consideración propia, consulta de textos y en apoyo a investigaciones de secuenciación de ADN mitocondrial de varias especies de peneidos publicadas (Chan y Chu, 1996) y no publicadas (Lavery y Chu, s/fecha), que

refutan la propuesta actual, el cambio de género basado en simple experiencia y datos morfológicos no arroja fundamentos claros.

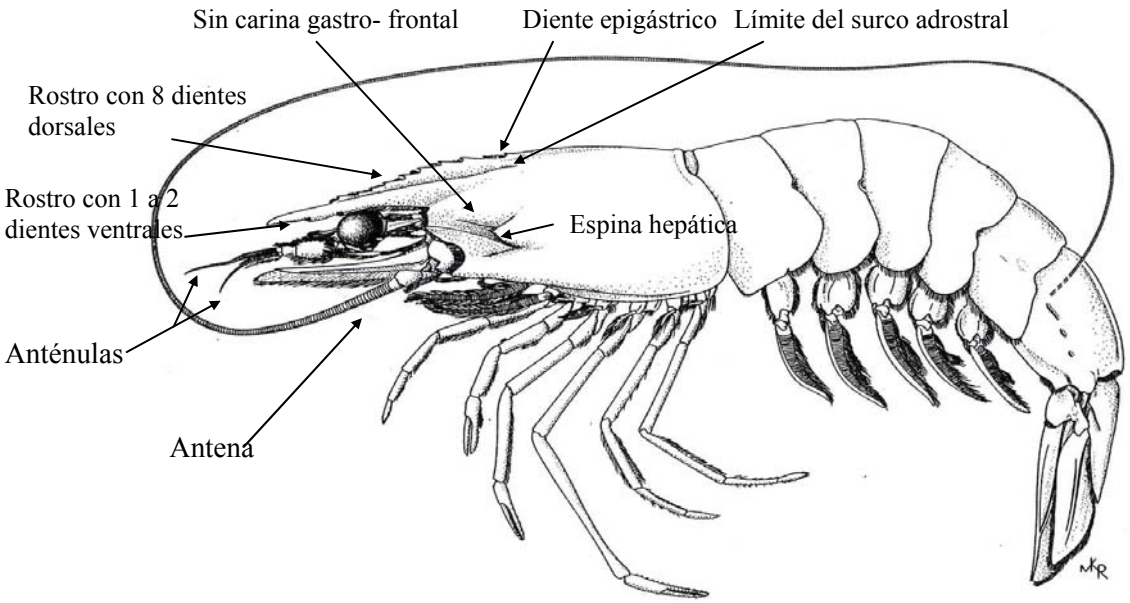


Figura 1. Características morfológicas del cefalotorax de *Penaeus vannamei*

FISIOLOGIA:

El camarón capta el oxígeno contenido en el agua empleando branquias expuestas llenas de sangre (hemocianina), la misma que se encarga de capturar el oxígeno y transportarlo a otras partes del cuerpo, de acuerdo a las necesidades del mismo. A través de las branquias también se atrapan otras sustancias (gases, elementos químicos, etc.) indispensables para la vida del camarón.

La circulación es abierta. La sangre sale del corazón (dorsal), distribuyéndose por el cuerpo donde se extravasa e inunda los tejidos, luego vuelve al corazón pasando por las

branquias, donde se oxigena. El pigmento respiratorio (hemocianina, que contiene cobre) se encuentra disuelto en el plasma de los camarones.

Johansson *et al.*, en 1994 menciona en su artículo que los invertebrados no poseen anticuerpos u otros caracteres del sistema inmune adaptativo de vertebrados, pero tienen sistemas de defensa innatas como los inhibidores de proteinasas que participan en contra de las proteasas de microorganismos o parásitos o regulan reacciones de defensa en el organismo. La idea de la carencia de un sistema inmune en crustáceos no fue apoyada por muchos investigadores y los estudios realizados comenzaron a arrojar avances significativos en este contexto, con enfoque en péptidos antimicrobiales que actúan como antibióticos endógenos, por lo que se consideran elementos clave de inmunidad innata. A la fecha ya se han descrito una amplia variedad de estos péptidos que se han clasificado en tres distintos grupos basados en secuencia de aminoácidos, estructura secundaria y similaridades funcionales (Destoumieux *et al.*, 2000). A este respecto, Gross *et al.*, en el 2001 reporta la presencia de células inmunes (hemocitos) en hemolinfa y en otros tejidos como las branquias. Menciona mecanismos de defensa como fagocitosis, encapsulación, formación de nódulos, coagulación, aglutinación y actividad antimicrobial. Uno de los mecanismos de respuesta inmune en crustáceos decápodos es la cascada profenoloxidasa (Söderhäll *et al.*, 1996).

En las altas salinidades marinas los decápodos son osmoconformes: Adaptan su concentración sanguínea a la del mar. Sin embargo, los que viven en menores salinidades deben mantener una mayor concentración de la sangre que la del medio. Los órganos excretores son las glándulas de las antenas. Se encuentran en la cabeza y desembocan en la

base de las antenas. Sin embargo la mayor parte del nitrógeno se excreta en forma de amoníaco (NH_{3+}) principalmente, óxido de trimetilamina y úrea, en menor proporción los aminoácidos pueden llegar a ser hasta un 10% del nitrógeno excretado (Bucheli, s/fecha).

Los crustáceos decápodos son carnívoros-omnívoros y se alimentan predominantemente de noche. La boca está en posición ventral acompañada de mandíbulas (apéndices masticadores). Posee un esófago corto y un hepatopáncreas que llevan a cabo el proceso de digestión completo.

El sistema nervioso de los decápodos es ganglionar, incluye el encéfalo y los ganglios que se suceden a lo largo del cuerpo hasta el final del abdomen. El sentido de la vista reside en ojos compuestos, las antenas tienen función olfatoria. En la base de las antenas están los órganos del equilibrio (Martínez, 1999).

Los camarones tienen sexos separados y en muchas especies con diferenciación externa. La fecundación es externa. El control de la gametogénesis se efectiva principalmente, a través de influencia de la temperatura y, en algunas especies del fotoperiodo.

Sobre el cuerpo se extiende una capa celular que segrega el caparazón o cutícula (exoesqueleto). Las capas interior y exterior están formadas de quitina (polisacárido aminado) y proteínas ligadas formando un complejo glicoprotéico. En la muda, la capa celular segrega quitinasa y proteasas (fluídos de muda y enzimas), estas secreciones digieren parcialmente la capa interior. La vieja cutícula se rompe, y el animal sale dejando un exoesqueleto completo (Martínez, 1993).

Los cromatóforos son células especializadas en la producción de sustancias químicas causantes del color.

El ciclo biológico de los camarones peneidos puede ser dividido en dos fases: la marina y la estuarina.

La cópula (apareamiento) y el desove ocurren en mar abierto, en aguas de mayor profundidad. Luego pasan a través de un número de estadíos larvales (nauplio, zoea, mysis), antes de alcanzar el estadío de postlarvas. Estas postlarvas se mueven en dirección a la costa, hacia los estuarios de los ríos, donde se desarrollan rápidamente. Las hembras son sexualmente inmaduras cuando salen de los estuarios. El número de huevos por desove varía entre 200,000 y 350,000. Existe evidencia de que las hembras desovan más de una vez. La vida normal de camarón es de 12 meses aproximadamente, pero algunos llegan a los dos años (Martínez, 1999; Bucheli, s/fecha).

1.2 Camaronicultura

1.2.1 Historia

La acuacultura es tan antigua como la propia humanidad. Remontándonos a lo que constituye la primera puesta en marcha de la acuacultura a escala comercial, esta el caso de Java en el año 1200 a.C.-1400 d.C., tiempo en el cual la acuacultura fue el centro económico de la población; esta iniciativa no fue la primera manifestación pero si su primera representación del gran potencial de lo que podría representar en el futuro. (Alatorre, 1998). Las raíces de la camaronicultura a nivel global, se enclavan en las civilizaciones orientales; en el sureste asiático se ha cultivado el camarón por más de cinco

siglos (Shigueno, 1992). Las técnicas modernas de cultivo se iniciaron hace casi 35 años y han tenido rápida aceptación. En Japón en 1934, Motosaku Fujinaga, inició los experimentos para el cultivo controlado de camarón y fue hasta 1959, cuando bajo su dirección comenzaron a operar las primeras granjas de ciclo completo de este crustáceo (CICTUS, 1988).

En nuestro País los orígenes de la acuacultura se remontan a la época prehispánica; sin embargo, prácticamente es hasta finales del siglo pasado, en 1884 cuando Don Esteban Cházari publica el primer tratado de piscicultura, que apunta las bases para iniciar su desarrollo (Martínez, 1999).

1.2.2 Producción

La producción mundial de camarones esta dominada por tres especies principales, todas ellas del Océano Pacífico, *Penaeus monodon* (58.2%) y *Fenneropenaeus chinensis* (9.8%) del Pacífico Occidental y *Litopenaeus vannamei* (13.1%) del Pacífico Oriental, que en conjunto representan mas del 80% de la producción mundial y se consideran especies líderes en el mercado de camarones cultivados. Si bien, *L. vannamei* representa el 13.1% de la producción mundial, en Latinoamérica representa el 91.7% de la producción (Magallón, 1999).

En México el desarrollo camaronícola se ha dado principalmente en la vertiente del Pacifico, sobre todo en el estado de Sinaloa, donde funcionan en la actualidad mas de 100 granjas comerciales principalmente semiintensivas. En la región del estado de Sonora, la camaronicultura también empieza a surgir muy fuerte y ya están funcionando más de 20 granjas comerciales. Otros estados como Oaxaca y Nayarit recién empiezan a implementar

proyectos comerciales. En Baja California Sur, existen también, instituciones que han trabajado exitosamente con el cultivo de camarón blanco (Cortes, 1993), como el laboratorio de producción de poslarvas y la granja de engorda de la sociedad cooperativa de acuicultores de la península de Baja California, localizada en la ciudad de La Paz, en donde en el año 2000 se alcanzaron producciones de camarón a tallas comerciales en cultivos intensivos de hasta 50 ton/ha/año a partir de un laboratorio que produjo 175 millones de postlarvas (Dubost, citado en Latisnere, 2001).

1.2.3 Investigación y Desarrollo Tecnológico

Los primeros estudios enfocados hacia la camaronicultura fueron para investigar el potencial de cultivo de las diferentes especies en diferentes lugares del mundo adoptando mucha de la tecnología japonesa que había implementado con éxito Motosaku Fujinaga., para 1970 investigadores y acuacultores ya tenían probadas docenas de especies de peneidos por su potencial de cultivo. En el proceso, se trabajaron técnicas de crianza y desove de la mayoría de las especies, otros se concentraron en el desarrollo de tecnologías de crecimiento, nutrición y enfermedades. Estos primeros esfuerzos ponen en manifiesto los fundamentos de una industria en expansión. No es hasta la ultima década cuando las investigaciones se empiezan a puntualizar en la detección de nuevas enfermedades virales como la “mancha blanca”, ya que, estas disminuyeron en la producción o provocaron perdidas totales por mortandad masiva de organismos; y en estudios de aspectos fisiológicos a nivel genético y molecular que permitan la estabilidad productiva de esta gran industria a partir de un conocimiento integral de la especie manipulada (Rossemberry, 2003).

En América los estudios de camarón en estanques se iniciaron en los Estados Unidos de América a principios de los años cincuenta (Alfonso *et al.*, 1993). Para 1970 en nuestro país se empieza a generalizar el consumo ocasional de productos generados a través de prácticas de cultivos extensivos y sus beneficios se traducen en mayor consumo del producto y comercialización del mismo. En los años posteriores se empiezan a considerar técnicas innovadoras y se observa la transición hacia la utilización de sistemas semintensivos e intensivos (Martínez, 1999). A la fecha, gracias al trabajo de muchos investigadores y cultivadores de la industria privada alrededor del mundo, la biotecnología y el cultivo de camarón se encuentra totalmente desarrollada, hasta el punto en que se ha completado el control sobre todo el ciclo de vida de algunas especies para alcanzar producciones rentables y predecibles, pero no tan optimas como se esperaría, de acuerdo a la cantidad de organismos con los que se inician los cultivos y los que se obtienen en la cosecha, de manera que estos éxitos parciales han ampliado el campo de investigación dedicada a esta especie.

Desde 1984, el Centro de Investigaciones del Noroeste (CIBNOR) ubicado en La Paz Baja California Sur, ha enfocado, un número sustantivo de sus programas a la evaluación del cultivo de camarón en la zona semiárida del Noroeste de México. Se sabe que especies como *Penaeus stilirostris* y *Penaeus californiensis*, se han cultivado en esta zona, aunque sigue siendo *Penaeus vannamei* la especie predominante en las granjas de cultivo comercial de las costas del Pacifico (Villarreal y Castro, 1993).

1.3 Enfoque de la Investigación científica

Actualmente los estudios se dirigen a la disciplina biotecnológica, la cual hasta hace casi 30 años su concepto estaba basado en la búsqueda para, y descubrimiento de, "biología explotable" que inicia con la detección de materiales biológicos apropiados, seleccionados por algún atributo deseado y culmina con el desarrollo de un producto comercial o proceso; sin embargo, los avances científicos y tecnológicos de la última década están revolucionando las perspectivas para la biología explotable de tal manera que los procesos están teniendo una mayor reevaluación y en muchos casos están siendo suplantadas por nuevas estrategias, de tal manera que la búsqueda y el descubrimiento se están convirtiendo en consecuencia de la revolución bioinformática, cuya estrategia esta basada en la colección de datos y almacenaje, y el significado (recuperación e integración) de las bases de datos para generar conocimiento. Ejemplos de la trascendencia que esta generando la bioinformática en áreas biológicas son, la sistemática, la genómica, la proteómica y la ecología. La Genómica es una disciplina que se fundamenta en la actividad de secuenciar genomas y obtener información de ellas con herramientas computacionales. Por su parte la genómica funcional define el estado del transcriptoma y proteoma de una célula, tejido u organismo bajo ciertas condiciones. El concepto transcriptoma describe los perfiles de transcripción (RNA mensajero), mientras el proteoma describe la traducción (proteína) complemento derivado de un genoma incluyendo modificaciones postraduccionales de proteínas, y proporciona información de la distribución de proteínas dentro de una célula u organismo en tiempo, espacio y respuesta al ambiente (Walter y Gingold, 1997).

El desarrollo de estudios a nivel molecular, ha sido causa de la aparición de una serie de técnicas indistintamente referidas como ingeniería genética, clonación y tecnología del ADN recombinante o manipulación genética in vitro (Walter y Gingold, 1997).

Con el desarrollo de la Ingeniería Genética, se ha abierto un campo nuevo en el estudio de los ácidos nucleicos de los seres vivos. Esto ha sido posible gracias a la clonación molecular, la cual ha permitido aislar en forma pura, reproducible e ilimitada fragmentos específicos de ADN de un organismo dado, al introducir estos en vectores o vehículos moleculares y propagarlos en un organismo alterno, por ejemplo en *Escherichia coli*. Cualquier procedimiento de clonación de ADN requiere de cuatro partes esenciales: un método para generar los fragmentos de secuencias nucleotídicas de interés, las reacciones que van a ensamblar el ADN ajenos al vector, una manera de introducir el recombinante artificial en la célula huésped donde este se pueda replicar, y un método de selección de las clonas de las células receptoras que adquirieron el recombinante (Sambrook *et al.*, 1989). La combinación de conocimientos pasados de bioquímica y fisiología de muchos organismos y nuevos revelaciones dentro de la función biológica derivados de los estudios de genómica y genómica funcional pueden guiar estrategias de búsqueda mas especificas y/o confirmativas (Bull *et al.*, 2000).

El estudio de la genómica del camarón esta incluido en un campo relativamente nuevo conocido como “ingeniería acuacultural” que está creciendo, gracias al desarrollo de técnicas intensivas de acuacultura, las cuales pueden optimizar los resultados de producción, en combinación con la reducción de costos de mano de obra y otros. De todas estas actividades de manipulación, la más importante enfoca la atención a la fisiología de

los animales cultivados, con el objeto de eliminar propagaciones de materia no apropiada y métodos de cultivo erróneos (Van Herp, 1992), por lo que es claro el hecho de que los problemas que ahora confronta la industria de la acuacultura del camarón deba requerir una combinada aplicación de biotecnologías con la finalidad de desarrollar un sistema de cultivo sustentable (Benzie, 1998).

En cultivo de camarones peneidos, las principales aplicaciones de la biotecnología se refieren a: la provisión de semilla con mejores características para el cultivo, la nutrición en las diferentes fases de la producción, la solución de aspectos críticos de la alimentación de grupos sensibles (primeros estadios, reproductores, cultivos intensivos) (Kureshy y Davis, 2000; Molina *et al.*, 2002; Molina *et al.*, 2000; Vega-Villasante *et al.*, 2000; Guarna y Borowsky, 1993; Fernández-Luna *et al.*, 1997), en la prevención y el enfrentamiento de las enfermedades (Arce *et al.*, 1999), en el mantenimiento de buenas condiciones del ambiente de cultivo, particularmente en la intensificación y en nuevas estrategias de producción (Berger, 2000).

Las técnicas de Biología Molecular en años recientes han permitido entender muchos de los procesos y estructuras subcelulares de una gran variedad de organismos, sin embargo en camarón la práctica de estas técnicas es aun reciente. Muy poco se ha hecho por elucidar el comportamiento de los genes que participan en el proceso de digestión y su control (Martínez, 1999).

La expresión de genes nos permite detallar conocimientos de procesos biológicos que la bioquímica, Biología celular, fisiología y farmacología tradicionalmente nos proporcionan, pero de una manera alterna nos da también la oportunidad de obtener un conocimiento

comprensivo y holístico de sistemas complejos y del estrecho espacio que existe entre secuencia y función.

2. Antecedentes

El éxito en el cultivo de las diferentes especies de camarón, depende en gran parte de una adecuada nutrición y un buen manejo del alimento. La nutrición de camarones implica procesos químicos y fisiológicos que proveen nutrientes al animal para sus funciones normales, de mantenimiento y crecimiento (Ceccaldi, 1997).

2.1 Alimentación en Peneidos

Los crustáceos decápodos son carnívoros-omnívoros y se alimentan predominantemente por la noche. Poseen un esófago corto dividido en dos cámaras. La primera cámara es el estómago gástrico donde el alimento es triturado gracias a un sistema llamado molino gástrico. En la segunda cámara (el estómago pilórico), ocurre la digestión del alimento, ya que en él desembocan las secreciones del hepatopáncreas. El esófago, el estómago gástrico, y la mitad del estómago pilórico están formados en gran parte, de quitina. El alimento troceado en la boca pasa al estómago gástrico, este es de paredes musculosas y formadas interiormente de quitina. Aquí el alimento es triturado y mezclado a 7-8 pH con las enzimas digestivas procedentes de las secreciones de las paredes del estómago dando una consistencia fina. De allí el alimento pasa al estómago pilórico, formado de paredes plegadas que constituyen un verdadero filtro que conduce al alimento a las secreciones del hepatopáncreas. Este segrega enzimas que digieren proteínas

(proteasas), carbohidratos (amilasas, células, quitinasas) y lípidos (lipasa). El alimento pasa al intestino en cuya primera parte es probablemente absorbido, el alimento ya digerido pasa a formar las heces (Dall *et al.*, 1990).

2.2 Hepatopáncreas. Morfología y función

Las investigaciones acerca de las funciones de la glándula digestiva en decápodos es difícil. El órgano es extremadamente delicado; cualquier trauma, incluyendo la muerte es rápidamente seguido por autolisis por las enzimas digestivas en las células.

La glándula digestiva (hepatopáncreas), posee de 2-3 lóbulos bien diferenciados. Se comunica con el estómago a través de dos ductos principales y estos se dividen en conductos secundarios y terciarios que terminan en túbulos pequeños o divertículos. Estos túbulos constituyen la masa glandular. Los túbulos están conformados por diferentes tipos de células que llevan la nomenclatura siguiente: E (embriónicas), R (de Restzellen o de almacenaje), F (fibrilares), B (de Blazenzellen o vacuolares) y M (de Midget); y cumplen funciones variadas. Así las células F sintetizan gránulos zimógenos, los cuales son exocitosados hacia el lumen del túbulo y transportada hacia el estómago. El material soluble es absorbido por las células R, pero las principales funciones digestivas y absortibas de la glándulas son llevadas a cabo por las células B, las cuales ingieren los nutrientes por endocitosis y las procesan en una gran vacuola digestiva. Las funciones de almacenamiento del hepatopáncreas son llevadas a cabo por las células M y R. Las células R almacenan lípidos, glucógeno y elementos traza; mientras que las células M almacenan materiales glico-proteícos y en menor cuantía, glucógeno. Casi al final del ciclo digestivo,

las células E se dividen por mitosis para reemplazar la perdida degenerativa de las células R y B. La perdida de células B involucra la eliminación de productos de desecho mediante secreción halocrina siguiendo la fase terminal de digestión intra-celular y asimilación. El hepatopáncreas parece ser que detoxifica los metales pesados almacenándolos en una forma insoluble en las vacuolas suprarrenales de las células R (Dall, 1992; Dall *et al.*, 1990; Talavera *et al.*, 1998).

Es indudable el avance en los conocimientos sobre el proceso de la digestión del camarón y la participación del hepatopáncreas con funciones como la absorción de nutrientes del alimento digerido, absorción de iones y control del flujo neto de agua entre el lumen del intestino medio y la hemolinfa, excreción de iones y secreción de membrana peritrófica (Lovett y Felder, 1990).

Para que la digestión química se lleve a cabo se requiere de la acción de una serie de enzimas digestivas. Desde el punto de vista fisiológico el estudio de las enzimas digestivas del camarón ha representado una importante herramienta, ya que se pueden describir patrones de conducta alimentaria de los animales, como la habilidad para hidrolizar materiales específicos de la dieta, la respuesta a diferentes fuentes de nutrientes, así como el nivel de inclusión de estos, la contribución de las enzimas bacterianas, los ciclos de secreción y los cambios que los animales sufren durante su crecimiento y maduración.

Los mecanismos de control de las funciones de las enzimas digestivas son diversos. Uno de estos puede darse a nivel del núcleo de las células que sintetizan las enzimas. Sin embargo, la información que se tiene al respecto es muy pobre. Se piensa que la

composición de la dieta tiene influencia sobre la expresión de los ácidos nucleicos, particularmente a nivel de ARNm que codifica las enzimas digestivas (Martinez, 1999).

La producción de enzimas digestivas en el hepatopáncreas de los crustáceos se encuentra controlada, en parte, por hormonas del pedúnculo ocular (Ceccaldi, 1997).

En *Litopenaeus vannamei* al igual que en otras especies de peneidos, se ha determinado la presencia de enzimas endógenas tales como: tripsina, quimotripsina, aminopeptidasa, lipasas, carbohidratasas, carboxipeptidasa A y B. La colagenasa quitinasa y celulasa, son enzimas también identificadas pero se cree que son de fuente exógena (Molina *et al.*, 2002).

La expresión de las enzimas es afectada por una serie de factores limitantes como son: parámetros físico-químicos del agua (pH, oxigeno, salinidad y temperatura), edad y tamaño del camarón, cambios ontogénicos, ayuno, ingredientes de la dieta, nivel y fuente proteica, nivel y tipo de aglutinantes, aditivos promotores del crecimiento, cantidad y frecuencia de alimentación, ritmo circadiano, ciclo de muda, e incluso ha sido reportado que el agua ejerce un efecto estimulante sobre la actividad enzimática digestiva. A partir de esa diversidad de variables es importante optimizar la digestión y absorción de nutrientes en la dieta tomando en cuenta todos los factores antes mencionados para aprovechar al máximo la capacidad de las enzimas digestivas (Molina *et al.*, 2002)

2.3 Inhibidores de proteinasas

FUNCION, ESTRUCTURA Y CLASIFICACION

Las proteínas inhibidoras de proteinasas son ubicuas, están presentes en formas múltiples tanto en tejidos de animales, de plantas como en microorganismos. Su función fisiológica en el sentido amplio es *la eliminación de proteólisis* no deseada (Laskowski y Kato, 1980). Sin embargo, la función detallada no ha sido aún elucidada con la excepción de las macroglobulinas, las cuales inhiben proteinasas de todas clases.

Los inhibidores de proteínas individuales inhiben únicamente a través de una clase de mecanismo simple. De estos los mas estudiados son los inhibidores tipo serin proteinasas.

Hans Neurath (Laskowski Jr. y Quasim, 2000), fué de los primeros científicos que reconoció que las proteinasas no únicamente actúan como enzimas digestivas, si no que también cumplen con otras numerosas funciones en el organismo. La versatilidad de las proteinasas ha sido un punto de su investigación, sus revisiones y lecturas. Hans también reconoció que mientras las proteinasas son altamente benéficas, también son peligrosas. Para limitar este peligro, deben ser estrictamente controladas en tiempo y lugar. Mientras muchos mecanismos han sido apoyados para el control de las proteinasas, son dos los que predominan. El primero es que casi todas las proteinasas son sintetizadas como precursores inactivos llamados zimógenos o proproteínas que son almacenadas y activadas en demanda. Otro mecanismo de control es necesario para inactivar las enzimas una vez activadas. Este procedimiento es consumado por la ubicua presencia de proteínas

inhibidoras de proteinasas. Estos inhibidores forman complejos inactivos o menos activos con su complemento enzimático.

No todos los miembros de las familias de inhibidores de proteinasas, son inhibidores eficientes. Muchos han sido reconocidos para realizar otras funciones fisiológicas (Laskowski Jr. y Quasim, 2000).

Los inhibidores de proteinasas a menudo se encuentran como mayores componentes de citoplasma, fluidos intercelulares y secreciones. Aparte de estar involucrados en los procesos fisiológicos de digestión, actúan como reguladores; son importantes factores en el control de una variedad de eventos críticos asociados con coagulación, fibrinólisis, reacciones inflamatorias y fagocitosis; se ha encontrado que componentes de la pared celular de microbios patógenos activan el sistema inmune en invertebrados (Liang y Soderhall, 1995). Sin embargo, el significado fisiológico de los inhibidores de proteinasas es muy amplio y no está limitado a las funciones citadas. Los inhibidores pueden ser usados en el control de plagas (Lawrence *et al.*, 2001), prevención y tratamiento de enfermedades como el cáncer y SIDA, hipertensión, parásitos invasivos y eliminación de actividad proteasa no deseada en la industrialización de alimentos (García-Carreño, 1995).

La importancia del control de la actividad proteolítica por inhibidores en procesos fisiológicos está demostrada por el hecho de que las moléculas inhibidoras exceden el 10% del total de la proteína en el plasma humano (García-Carreño, 1996).

Los inhibidores conocidos no son homólogos pero consisten de varias familias, las cuales surgieron por evolución convergente más que por divergencia y pueden ser

identificadas por la relación topológica que presentan entre sus puentes disulfuro y la localización de su sitio activo.

Una de las características de los inhibidores es la presencia de varios sitios activos homólogos en la misma cadena polipeptídica, la cual surge por elongación de genética por duplicación repetida. En algunos casos, existe una interhomología, donde dos regiones homólogas, están interconectadas por la región de los puentes disulfuro. En muchos otros casos, la duplicación de genes da origen a diversos dominios en bloques con todos los puentes disulfuro intradominio y con los dominios conectados únicamente por péptidos cortos. Existen otros casos en los cuales se habla de dos dominios en bloque en la misma cadena polipeptídica, que pertenecen a familias de inhibidores diferentes (Laskowski y Kato, 1980).

Hasta el año 2000 se habían reconocido 18 familias de inhibidores de proteinasas, es probable que más familias sean descritas en el futuro por la gran importancia que ha manifestado los estudios de secuenciación genómica y seguramente establecerá nuevos parámetros para determinar similitudes y reconformar o confirmar clasificaciones, ya que, a la fecha se ha notado que en algunos casos la familia asignada es arbitraria, particularmente real para inhibidores de plantas (Laskowski Jr. y Quasim, 2000)

2.4 Expresión Tejido-Especifica

Las células de dos tejidos diferentes son especializadas para realizar funciones diferentes en un organismo. Aunque podemos reconocer células de diferentes tejidos por su fenotipo no se conoce con exactitud como hacen las células para funcionar como músculo

liso, como neurona y otras como gónadas. Lo que sabemos es que el papel de una célula está determinado por las proteínas que ésta produce, la cual en turno depende de su expresión de sus genes. Experimentos de hibridación comparativa revelan genes que son preferencialmente expresados en tejidos específicos. Algunos de estos genes implementan comportamientos que distinguen el tipo de célula en un tejido, mientras otros actúan controlando genes asegurando que la célula únicamente realice funciones de su tipo (Buhler, 2002).

El descubrimiento de nuevos genes tradicionalmente había sido una tarea laboriosa, requiriendo de meses o algunos años de trabajo. Ahora, en la era de la secuenciación a gran escala del genoma, identificar nuevos genes requiere muy poco tiempo, unos pocos minutos para realizar la búsqueda por computadora en una base de datos para el análisis de secuencias. La base de datos con la más alta tasa de crecimiento, tiene una división en el GeneBank, llamada dbEST, o base de datos de secuencias blanco expresadas (http://www.ncbi.nlm.nih.gob/dbEST). Estas secuencias son cortas, unos pocos cientos de pares de bases de longitud, las cuales son parciales y derivadas de secuenciación de paso simple del inserto de una clona seleccionada al azar (Wolfsberg, 1997).

Basado en la información recopilada a la fecha, se han desarrollado estudios para evaluar el efecto de diferentes factores que afectan la actividad enzimática del hepatopáncreas en diversas especies de peneidos (Molina *et al.*, 2002).

Las investigaciones en genómica funcional, la cual utiliza la combinación de análisis computacionales y análisis basados en expresión de información de gran cantidad

de secuencias, está surgiendo como una poderosa herramienta para acelerar, apoyar y reforzar el conocimiento bioquímico recopilado a la fecha acerca del metabolismo celular (Lange *et al.*, 2000) en organismos completos y tejidos especializados como el hepatopáncreas del camarón.

El análisis de secuencias expresadas en tejidos específicos ha permitido en los últimos años incrementar los conocimientos que se tenían acerca de moléculas de interés, de esta manera, actualmente se conoce, por medio de estudios en Biología Molecular, que la maquinaria nuclear de las células del hepatopáncreas de camarón, sintetiza también proteínas ajenas a la función digestiva.

Gross *et al.*, en el 2001, reportó en un estudio realizado con la finalidad de encontrar genes con función inmune por ESTs en hepatopáncreas de los camarones *Litopenaeus vannamei y L. satiferus*, que del 4 al 5% de las secuencias expresadas analizadas por PCR y secuenciación sencilla estuvieron relacionadas con funciones inmunes (proteínas unidas a β-1,3 glucanos y lectinas). Entre otras, encontraron proteínas de choque térmico (HSPs: HSP 10, HSF 3 y HSF 5), enzimas oxidativas (peroxinectina y superoxido dismutasa), otros efectores inmunes (ferritina).

Para la obtención de secuencias expresadas es necesario tener conocimiento de terminología y diferentes técnicas en Biología Molecular para el análisis.

2.4.1 Ácido Ribonucleico (ARN)

Una célula típica de mamífero contiene aproximadamente 10^{-5} μg de ARN, 80-85% del cual es ARN ribosomal (especies 28s, 18s, 5.8s y 5s). Mucho del resto consiste de una variedad de especies de bajo peso molecular (ARN de transferencia y ARN pequeño nuclear). Estos ARNs abundantes son de tallas y secuencias definidas y pueden ser aislados en forma virtualmente puras, por gel de electroforesis, centrifugación por gradiente de densidad, cromatografía de intercambio aniónico o HPLC. En contraste, el ARN mensajero (ARNm) sintetizado se encuentra en proporciones del 1-5% del ARN total de la célula, es heterogéneo en talla- de unos pocos cientos de bases a varias kilobases de longitud- y secuencia. Sin embargo, muchos ARNsm poseen en sus terminales 3' un tracto de residuos de ácido poliadenílico que es generalmente largo y permite su purificación por cromatografía de afinidad en oligo (dT) celulosa. El resultado heterogéneo de la población de moléculas colectivamente codifican todos los polipéptidos sintetizados por la célula, en ese tiempo y espacio específico.

Los residuos de ribosa poseen grupos hidroxilo en las posiciones 2' y 3', por lo que, el ARN es mucho más reactivo que el ADN y es rápidamente atrapado y procesado por las ribonucleasas. Las ribonucleasas son liberadas de la célula durante la lisis y están presentes en la piel, se requiere vigilancia constante para prevenir contaminación de cualquier fuente o material requerido. El problema consiste en que no existen métodos simples para inactivar ribonucleasas. Son resistentes a la ebullición prolongada, a desnaturalizantes leves y tienen la capacidad de replegarse rápidamente cuando se desnaturalizan ya que poseen

puentes disulfuro entre sus cadenas. Como muchas ADNasas, las ARNasas no requieren cationes divalentes para su actividad y no pueden ser inactivadas por la inclusión de ácido etilendiaminotetracetico (EDTA) u otro quelante de iones metal en soluciones buffer (Sambrook, 2001).

La conversión enzimática de poli(A)+ ARNm a ADN complementario (ADNc) de doble cadena y la inserción de éste, a vectores procariotas y eucariotas son herramientas fundamentales de Biología Molecular.

2.4.2 Diseño de oligonucleótidos sintéticos para Transcripción Reversa (RT) y Reacción en Cadena de la Polimerasa (PCR)

El análisis de ARN con iniciadores oligonucleótidicos (primers), se realiza mediante una preparación de poli(A)+ ARN, el cual es hibridado con un iniciador oligodesoxiribonucleótido de cadena sencilla complementario al ARN blanco. La enzima transcriptasa reversa, se usa para extender el iniciador. El ADNc resultante es complementario al ARN templado y es igual en longitud a la distancia entre el extremo 5' del primer y el terminal 5' del ARN. Casi todos los ensayos de extensión, utilizan iniciadores oligonucleótidos sintéticos, de 20 a 30 nucleotidos de longitud, idealmente se utilizan dos primers que hibriden regiones de ARNm separadas por una distancia conocida (Sambrook y Russell, 2001).

El aislamiento de secuencias desconocidas relacionadas a secuencias conocidas es un método poderoso para investigación de la función biológica. La secuencia de una proteína desconocida en un organismo puede ser homóloga a otras proteínas de organismos

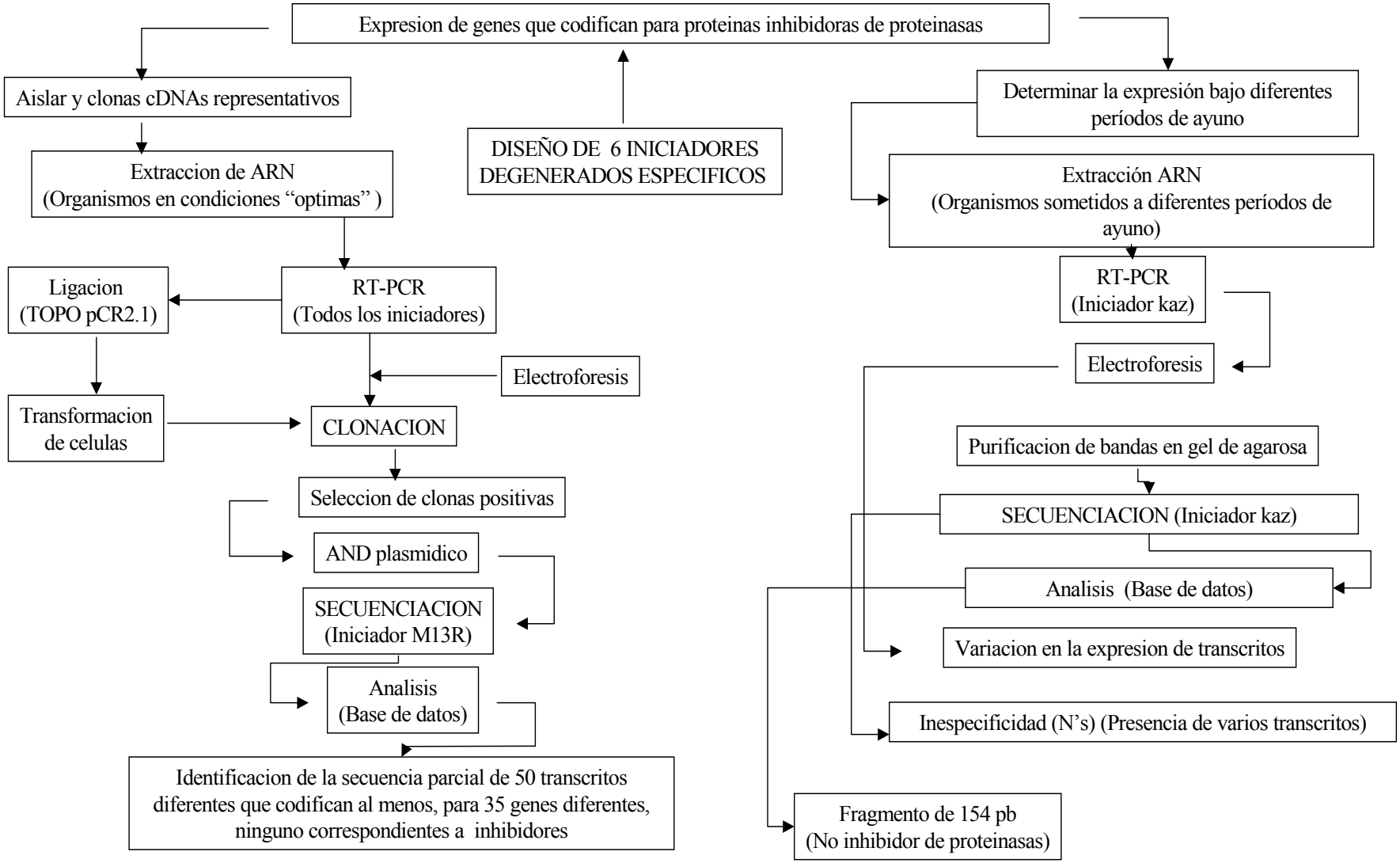
diferentes, o muy relacionadas a proteínas de secuencias conocidas que pertenezca a una familia multigénica en el mismo organismo. En muchos casos, la hibridación de baja astringencia o métodos de Reacción en Cadena de la Polimerasa (PCR) tienen éxito en la obtención de tales genes deseados. Sin embargo, conforme el grado de similitud de la proteína disminuye, el aislamiento del gen resulta más difícil.

Muchas aplicaciones de la PCR están basadas en el diseño de oligonucleótidos o iniciadores que se aparean precisamente en una secuencia blanco conocida. Sin embargo, en algunas situaciones, los iniciadores son diseñados para secuencias desconocidas, como cuando se prueban genes aislados codificantes de una proteína que pertenece a una familia conocida. En tales casos, los iniciadores diseñados están usualmente basados en la trascripción reversa de secuencias múltiplemente alineadas a través de regiones conservadas de proteínas. Se aplican varias reglas para solucionar estos problemas, pero frecuentemente las fallas para amplificar una secuencia blanco deseada, son atribuidas al inadecuado diseño de primers. El diseño de primer puede ser difícil, por la degeneración de codones y la degeneración adicional necesaria para representar codones múltiples en una posición en el alineamiento. Tales degeneraciones complican las pruebas para encontrar temperaturas de alineamiento compatibles. La necesidad de localizar regiones de secuencias de alta conservación conteniendo codones de baja degeneración limita la detección por PCR de secuencias desconocidas bastante parecidas, por tal razón el mejoramiento en el diseño de iniciadores es extensamente aplicable (Timothy *et al.*, 1998).

Para aislar secuencias distantemente relacionadas, dos estrategias han sido previamente empleadas. Una, es la síntesis de un juego de oligonucleótidos degenerados conteniendo muchos o todas las secuencias de nucleótidos posibles implícitas en un alineamiento múltiple. Un problema con este método es que a medida que la degeneración incrementa para acomodar más genes divergentes, la concentración de cualquier iniciador simple disminuye. La segunda estrategia es diseñar un iniciador consenso simple a partir de regiones altamente conservadas. El iniciador consenso se diseña escogiendo los nucleótidos más comunes en cualquier posición de una secuencia nucleotídica múltiplemente alineada.

La técnica de análisis serial de expresión de genes (SAGE), permite la construcción de un perfil de expresión comprensible en el cual cada ARNm es definido por un iniciador específico y es sometido al método basado en la técnica de RT-PCR, análisis rápidos de secuencias blanco SAGE desconocidas, para analizar la expresión del gen correspondiente. Estas investigaciones pueden ser usadas como un método de detección para investigar si el gen está siendo diferencialmente expresado o no en diferentes tipos de células (Van Der Berg *et al.*, 1999).

DIAGRAMA DE FLUJO
(Objetivos, Metodología y Resultados)



3. Objetivos

GENERAL

Detectar transcritos de genes que codifican para inhibidores de proteinasas en el hepatopáncreas del camarón blanco *Penaeus vannamei*, con la finalidad de ampliar el conocimiento que se tiene acerca de la estructura y función de genes que intervienen en la regulación de diferentes procesos fisiológicos.

ESPECIFICOS

Obtener secuencias expresadas en hepatopáncreas de camarón blanco *Penaeus vannamei*, a partir de oligonucleótidos diseñados para detectar inhibidores de proteinasa.

Determinar la expresión de transcritos en hepatopáncreas de camarones sometidos a diferentes periodos de ayuno.

Identificar secuencias expresadas por análisis de homología utilizando la base de datos del GenBank.

Determinar la posible función de las secuencias nucleotídicas obtenidas, a partir de la homología con proteínas reportadas para otras especies de organismos obtenida a partir del análisis en la base de datos.

II. MATERIALES Y METODOS

Diseño de Oligonucleótidos o Iniciadores

En la actualidad en el GenBank existen secuencias nucleotídicas y de aminoácidos de muchos inhibidores de proteinasas de diversos organismos, lo cual nos sirvió como antecedente y herramienta para detectar regiones conservadas de aminoácidos dentro de las secuencias de proteínas reportadas en el programa Swissprot para el diseño de cebadores. Para la realización de presente trabajo se diseñaron iniciadores (Tabla I), y el procedimiento para ello, consistió, en alinear secuencias de inhibidores de proteinasas en grupos, de acuerdo a diferentes familias (Apéndice 1), posteriormente se detectaron las regiones mas conservadas y se seleccionaron secuencias entre siete y diez aminoácidos para el diseño de los oligonucleótidos, para lo cual se consideraron las características usuales para su elaboración (Ver apéndice 3), las sustituciones de aminoácidos en las posiciones mas variables de la secuencia seleccionada, y la tabla de codones mas usados por *Penaeus vannamei* (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Litopenaeus+vannamei+[gbinv]).

Los iniciadores diseñados se sintetizaron en los laboratorios de Operon Technologies.

Obtención de organismos

De acuerdo a lo planteado en antecedentes, con el conocimiento de que las proteínas inhibidoras de proteinasas, están relacionadas con el control de la digestión, actuando directamente sobre las proteínas que degradan el alimento inactivando su acción, se planteó el desarrollo de un bioensayo, sometiendo camarones de la especie de interés a diferentes

periodos de ayuno, con la intención de caracterizar y detectar diferencias en la expresión de estos genes.

Veinte organismos vivos de aproximadamente 12 gramos de peso, de la especie *Penaeus vannamei,* sin tomar en cuenta sexos, se obtuvieron de los estanques ubicados en las instalaciones del CIBNOR, La Paz B.C.S., para realizar un bioensayo; se aclimataron a las nuevas condiciones y se colocaron cinco en cada uno de cuatro tanques de plástico de 300 litros llenados a ¾ partes de su capacidad con agua de mar oxigenada con bomba de aire; se les hizo recambio matutino diario de agua y se alimentaron con una formula peletizada (Purina, con 35% de proteína), por la mañana y por la tarde, suministrando aproximadamente 6 gramos de alimento en cada ocasión (correspondiente al 10% de su biomasa). Los camarones, se alimentaron todos el primer día y se sacrificaron los de tres tanques en diferentes periodos 24, 48 y 72 horas respectivamente después de haberle agregado por última vez alimento. Los organismos control se sacrificaron al cuarto día, dos horas después de su última alimentación.

Disección de organismos, extracción de RNA total y purificación de mensajero

Se sacrificaron los organismos después de cuatro días de su mantenimiento en condiciones controladas y se les extrajo el hepatopáncreas del cual inmediatamente se aisló ARN total basándonos en la técnica del *Kit Stratagene RNA isolation* (Protocolo en línea: http://www.stratagene.com/manuals/200345.pdf) modificando proporcionalmente las cantidades de los reactivos utilizados, partiendo de fracciones de 100mg de tejido. Se cuantificó concentración espectrofotométricamente y se visualizó la calidad mezclando 5µl de la muestra con 1µl de buffer de carga (LB[50% glicerol, 0.1 M EDTA pH 8.0, 0.25% azul de

bromofenol, 0.25% xilencianol) de cada una de la muestras e inyectándolas en los pozos de un gel de agarosa al 1% con TAE (Tris, Ácido acético, EDTA [ethylenediaminetetraacetic acid]) desnaturalizante con tiocianato de guanidina (0.06 gramos/100mL), preteñido con 1µl de Bromuro de Etidio (10mg/mL). La electroforesis se llevo a cabo a 70V/ 1 hora. Los resultados fueron documentados, mediante un sistema de fotodocumentación UVITEC (UVP Inc.) Se purificó el ARN mensajero utilizando el Kit Oligotex de Quiagen (Protocolo en-línea: http://www.qiagen.com/literature/Handbooks/PDF/RNA/INT/RNY_OTX/1019295_HB_OTX_prot01.pdf), partiendo de 250 µg de ARN total. La técnica Oligotex esta basada en el hecho de que el ARN mensajero puede ser purificado hibridando la cola poli-A a un oligómero dT unido a una matriz de fase sólida. Las especies de ARN ribosomal y de transferencia sin una cola poli-A no se unen al oligo dT y son arrastradas fácilmente. Ya que la hibridación requiere altas concentraciones de sal, el ARN mensajero puede ser fácilmente separado disminuyendo la fuerza iónica y desestabilizando los híbridos dT:A.

RT-PCR

El ARN mensajero obtenido se sometió a la técnica de RT-PCR basada en la amplificación de una cantidad pequeña de moléculas blanco de RNA (tanto ARNm como RNA total) con gran especificidad, mediante la transcripción reversa del ARN a ANDc (complementario), que es posteriormente amplificado. De esta forma, se puede medir la expresión génica mediante una técnica alternativa al Northern blot o Dot Blot, ensayos de protección de ARNasa, hibridación in situ, y ensayos de nucleasa S1. La RT-PCR es más sensible que estas técnicas, y requiere menos RNA. Se utilizó el Kit Promega Access RT-

PCR System and Access RT-PCR Inductory System (Protocolo en línea: http://www.promega.com/tbs/tb220/tb220.pdf), siguiendo las instrucciones del proveedor para amplificar el ARN control, utilizando los cebadores ya descritos como iniciadores rio arriba y un cebador poliT como iniciador rio abajo y las siguientes condiciones de temperatura para obtener los fragmentos de interés:

Secuencia del Primer Reverso:

PoliT (16mero) 5' TTT-TTT-TTT-TTT-TTT-T

Condiciones de Tiempo y Temperatura:

SÍNTESIS DE LA PRIMERA HEBRA:

Retrotranscripción	1 ciclo/42ºC/45min
Inactivación de la AMV Retrotranscriptasa y desnaturalización RNA/ADNc/primers	1 ciclo/94ºC/2min

SÍNTESIS DE LA SEGUNDA HEBRA DEL CDNA Y AMPLIFICACIÓN POR PCR:

Desnaturalización	94ºC/30seg	40 ciclos
Alineamiento	55ºC/1min	
Extensión	68ºC/2min	
Extensión final	1 ciclo/68ºC/7min	

Preservación 1 ciclo/4ºC/α

Los ensayos de RT-PCR se realizaron en un termociclador *GenAmp PCR System 9700, Perkin Elmer*.

Purificación de productos

Los fragmentos de cDNA se purificaron con el *Kit Gene Clean* (Protocolo en línea: http://www.qbiogene.com/literature/protocols/dna-kits/pdf/geneclean.pdf , pag 15-21). Se cortaron las bandas obtenidas en un gel de agarosa-TAE (Tris, Acido acético, EDTA) al 2.5% teñido con Bromuro de Etidio con un bisturí usando luz ultravioleta de onda larga. Se determinó el volumen del trozo de gel y se transfirió a un tubo eppendorf y se siguió el protocolo del proveedor; el cual, esta basado en el principio general de que el ADN se une a la sílica en concentraciones altas de sales caotrópicas y eluye en bajas. El mecanismo de unión DNA-Sílica aun no ha sido descrito, pero pudiera involucrar la disrupción de las sales caotrópicas de la estructura del agua alrededor de la silica cargada negativamente permitiendo la formación de un puente entre un cation y el esqueleto de fosfatos del DNA cargado negativamente. Cuando las sales son removidas, la matriz se sílica se rehidrata y la atracción entre la matriz y el DNA se rompe. Alternativamente se purificaron los productos de RT-PCR en columnas de purificación del Kit Quiagen (Protocolo en línea: http://www.qiagen.com/literature/Handbooks/PDF/DNA_cleanup/INT/QQSpin/1021422_HBQQSpin_prot01.pdf).

Producción de células competentes.

Método de Cloruro de Calcio (Sambrook, et.al. 1989) para bacterias.

Se estrió una cepa de *Escherichia coli* INVαF' de un stock congelado a –80ºC sobre la superficie de una placa de agar LB (Luria-Bertani) se incubó la placa durante la noche a 37ºC. Se transfirió una de las colonias a 5 mL de medio LB, se dispersaron bien las células con ayuda de un vortex y se diluyó el cultivo en 45 mL de medio LB en un matraz de 200mL. Se incubaron a 37ºC con agitación orbital constante (Incubadora New Brunswick Scientific Co. Inc. Edison NJ, USA) de 225 rpm durante 3 horas o hasta alcanzar una OD_{600} = 0.5. Las células se transfirieron a tubos Falcon 2070 estériles, se enfriaron en hielo por 10 minutos y se recuperaron por centrifugación a 4000 rpm a 4ºC por 10 minutos (Centrífuga BECKMAN J2-HS). Se descartó el sobrenadante, el botón celular se resuspendió en 10 mL de $CaCl_2$ 0.1 M frío. Las células se recuperaron por centrifugación a 4000rpm/4ºC/10 minutos. El sobrenadante se descartó y el botón celular se resuspendió en 2 mL de CaCl2 0.1 M frío. Las células se distribuyeron en fracciones de 50µl en tubos eppendorf de 1.5 mL estériles manteniéndose a 4ºC durante un periodo no mayor de una hora. Las células no utilizadas en el momento se les agregó DMSO (Dimethyl Sulfoxide) a una concentración final del 7%, se almacenaron a –80ºC, para su posterior uso.

Ligación

Se ligaron los fragmentos de cDNA obtenidos por la técnica de RT-PCR en el vector pGEM T easy de Promega (Protocolo en linea: http://www.promega.com/tbs/tm042/tm042.pdf) y

TOPO2.1pCR Invitrogen (Protocolo en línea: http://www.invitrogen.com/content/sfs/manuals/topota_man.pdf) de acuerdo al protocolo proporcionado por el proveedor.

Transformación de células. Método de choque térmico (Sambrook *et al.*, 1989).

A cada 50µl de células competentes, se agregó 2µl del cDNA ligado al vector. Se incubó en hielo durante 30 minutos y posteriormente se les dió un choque térmico a 37ºC durante 30 segundos en baño Maria en un termoblock (Sybron Thermolyne), y se colocaron inmediatamente en hielo y se les agregó 250µl de medio SOC. Se incubaron a 37ºC durante una hora con agitación orbital constante a 225 rpm. Se centrifugaron e inocularon 10 y 50µl respectivamente de células transformadas sobre placas LB con ampicilina (100mg/mL), IPTG (isopropyl-beta-D-thiogalactopyranoside) y Xgal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) distribuyéndose gentilmente sobre la superficie de agar. Se incubaron las placas a 37ºC durante la noche en una estufa (VWR Scientific).

Selección de clonas positivas

Para verificar la presencia del inserto de cDNA, incluso en colonias azules, se picaron cada una de las clonas obtenidas con un palillo de madera y se resuspendió su contenido en 10.5µl de agua destilada estéril, se incubaron las células a 94ºC/10 minutos para permitir que el AND plasmídico quede en resuspensión y la maquinaria enzimática de la bacteria inactiva, después de la lisis, se agregó la mezcla maestra para PCR la cual contenía para cada reacción: 1.25µl de Buffer 10x, 0.66µl de dNTPs (2.5mM), 0.13µl de cada uno de los cebadores (sentido y antisentido [M13* específicos para el vector TOPO2.1pCR y, T7 y

SP6** para pGEM-T easy])a una concentración de 250µg/mL y 0.05µl de Taq DNA polimerasa (Roche); y se llevó a cabo la reacción de PCR con las siguientes condiciones de temperatura: Un ciclo a 94ºC/2 minutos(Desnaturalización inicial); 30 ciclos de 3 temperaturas a 94ºC/1 minuto (desnaturalización), 55ºC/1 minuto (alineamiento), 72ºC/1 minuto (extensión); un ciclo de 72ºC/7 minutos (extensión final); 4ºC por tiempo indefinido para almacenar la muestra para procesos posteriores. Los resultados se revelaron a partir de una electroforesis en un gel de agarosa al 1% con Buffer TAE, teñido con Bromuro de Etidio y expuesto a luz ultravioleta para su visualización. *proporcionados por el proveedor del kit de ligación TOPO2.1pCR. **Elaborados en el laboratorio de síntesis del Instituto de Biotecnología de Cuernavaca, Mor.

Extracción y purificación de ADN plasmídico. Método de Lisis Alcalina

(Modificación Sambrook *et al.*, 1989 para la obtención de una muestra pura, lista para secuenciación)

Se picaron colonias transformadas con un palillo de madera estéril y se inocularon cada una de ellas en 3mL de medio LB c/ ampicilina (antibiótico o componente de selección). Se incubaron aprox. 16hr/37ºC/225rpm. Se recuperó aproximadamente 1.6 mL en un tubo eppendorf y se centrifugó a 14 000rpm/1min. Se decantó el sobrenadante y se resuspendió el botón celular por pipeteó en 200µl de solución I (50mM TRIS-HCl pH 8 y 10mM EDTA). Se agregaron 200µl de solución II (200mM NaOH y 1% SDS mezclado a temperatura ambiente) recién preparada. Se mezcló por inversión de 6-8 veces. No dejando más de 5 min. en esta solución, para evitar que el plásmido se contaminara con ADN genómico. Se agregaron 300µl de solución III (Acetato de Potasio 3M pH 5.5), se mezcló

por inversión 5 veces. Se centrifugó 14000rpm/10min. Se recuperó el sobrenadante y se agregaron 300µl de isopropanol, se mezcló por inversión 1 minuto a Temperatura ambiente y se centrífugó de nuevo por un minuto. Se eliminó el sobrenadante. Se lavó el botón de ADN plasmídico con 500µl de etanol al 70% (sin resuspender). Se secó a temperatura ambiente y se resuspendió en 50µl de agua destilada estéril. Se agregó 1µl de RNAsa A a 2 µg/ µl. Se incubó 30minutos a 37ºC. Previniendo un problema en la secuenciación, alternativamente se utilizó el Kit de sigma GenElute Plasmid Extraction (Protocolo en línea: http://www.sigmaaldrich.com/sigma/bulletin/pln350bul.pdf). Se midió concentración en espectrofotómetro Beckman DU-600 a 260 y 280 nm de absorbancia y se corrió una electroforesis en gel de agarosa al 1% con TAE.

Preparación de las muestras para secuenciación

Los plásmidos obtenidos a partir de clonas positivas, es decir con los fragmentos de interés, se secuenciaron, en el Laboratorio de Biología Molecular del CIBNOR, mediante secuenciación automática en un equipo ABI PRISM 310 (Perkin Elmer) basado en método de Sanger & Co. desarrollado en 1977 (Ausubel *et al.*, 1997), usando Bigdye Terminator cycle sequencing Reaction Kit, y particularmente para el presente trabajo, los cebadores antisentido M13 y SP6 respectivamente dependiendo del vector; y en el Departamento de Secuenciacion de ADN de MacroGen (Corea, Japón) con un equipo ABI PRISM® 3700 DNA Analyzer.

El análisis de secuencias se realizó mediante el uso de diferentes programas (Edit Seq, Meg Align, Gen Quest) del software DNAStar Inc., y se confrontó por análisis BLAST disponible en la base de datos mundial del NCBI.

III. RESULTADOS

Aislamiento de ARN y RT-PCR

La amplificación de secuencias expresadas se realizó con el uso de iniciadores degenerados (Tabla I) y de ARN mensajero obtenido de hepatopáncreas de camarón.

Los productos de amplificación (Figura 1), nos proporcionaron evidencia de la expresión de transcritos de diferente tamaño en el órgano de interés en cada una de las reacciones que contenían los distintos iniciadores.

Tabla I. Oligonucleótidos diseñados a partir de un fragmento conservado de la secuencia de aminoácidos reportadas para inhibidores de proteinasas

DESCRIPCIÓN DE LOS INICIADORES DISEÑADOS	SECUENCIA DE AMINOÁCIDOS	SECUENCIA SUGERIDA DE NUCLEÓTIDOS DE ACUERDO A LA FRECUENCIA DE CODONES QUE CONFORMAN LOS TRANSCRITOS EN *Penaeus vannamei*
Región aminoconservada de IPK* de anélidos (ANEK5)	WTKCGGC	5'TGG ACW AAG TGC GGS GGS3'
Región amino conservada de IPK en insectos (INSK5)	YNADTKKC	5'TAC AAC GCY GAC ACK AAG AAG TGC3'
Región amino conservada de IT** en páncreas de mamíferos (PANK5)	FYNACTKFC	5'CTA CAA CCG TAA GAC TAA GTC TGC3'
Región amino conservada de IPK en diferentes invertebrados (INVK5)	YDSESGKC	5'TAC GAT WSC GAG WSC GGC AAG TGC3'
Región amino conservada de IPK de animales (INHT5)	CLEPKVVGPC	5'TGC CTT GAA CCK AAG GTS GTS GGS CCY TGC3'
Región amino conservada de IPKz*** de vertebrados (KAZ5)	YDPVCGSD	5'TAC GAY CCK GTS TGC GGS W3'

*Inhibidor de proteinasas tipo Kunitz, **Inhibidor de tripsina, ***Inhibidor de proteinasa tipo Kazal.

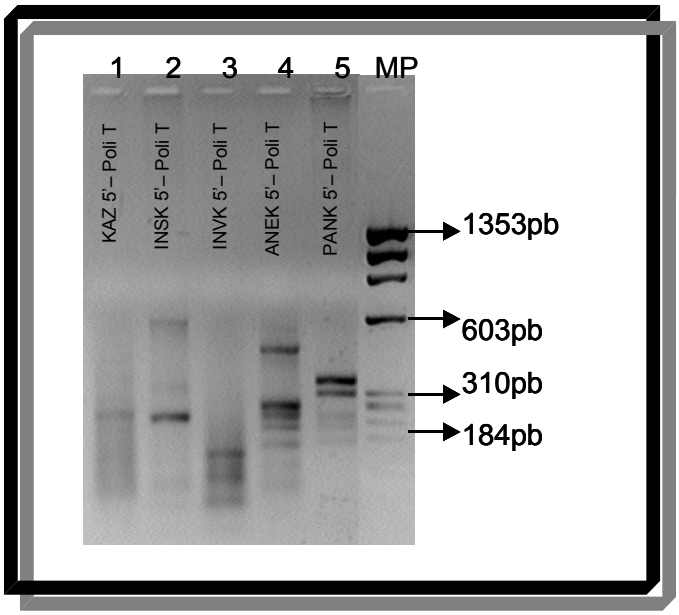


Figura 2. Electroforesis en gel de agarosa-TAE 2.5% de un RT-PCR realizado a partir de RNA total extraído de hepatopáncreas de camarón amplificado con diferentes iniciadores diseñados de regiones conservadas de inhibidores de proteinasas: (1) tipo Kazal de vertebrados, (2)Tipo Kunitz de insectos, (3)tipo Kunitz de invertebrados-cnidarios y moluscos-, (4) tipo Kunitz de Anélidos, (5) inhibidores de tripsina en páncreas de mamíferos, (MP) Marcador de Peso Molecular

Amplificación y secuenciación de transcritos obtenidos de Hepatopáncreas de camarones en diferentes períodos de ayuno con iniciadores diseñados para obtener secuencias nucleotídicas codificantes para proteínas inhibidoras de proteinasas tipo Kazal.

Con la finalidad de observar alguna diferencia en la expresión de tales transcritos bajo estrés fisiológico inducido como el ayuno, se realizó un bioensayo (descrito en metodología) del cual se obtuvieron ARNs mensajeros de organismos tratados bajo diferentes condiciones (Figura 3) y se sometieron a RT-PCR con el iniciador KAZ 5´ y poliT

Se obtuvieron diferencias de expresión, de acuerdo a lo observado en la electroforesis de los productos de amplificación en un gel de agarosa. (Figura 4).

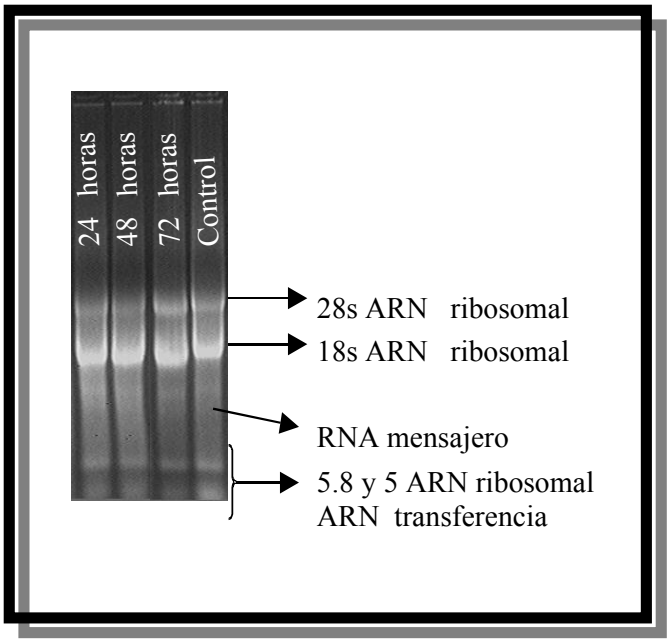


Figura 3. ARN total de hepatopáncreas de *Penaeus vannamei*, en diferentes períodos de ayuno (24, 48 y 72 horas). El control no se sometió a ayuno.

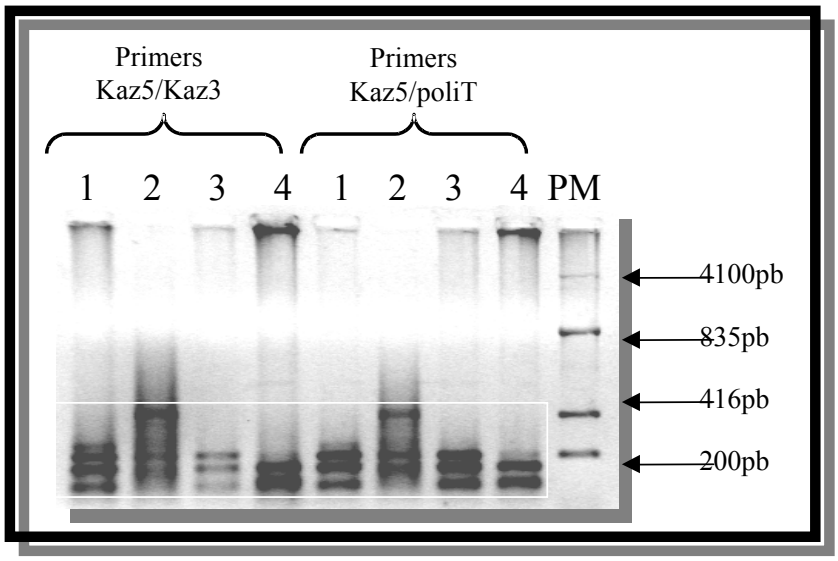


Figura 4. Electroforesis en gel de agarosa-TAE 2% mostrando la expresión bajo diferentes períodos de ayuno, (1) organismos sin ayuno-sacrificados dos horas después de ser alimentados-, (2) 24 horas, (3) 48 horas, (4) 72 horas. Las amplificaciones son del Peso Molecular esperado ya que los genes que codifican para inhibidores de proteinasas están entre 120 y 300 nucleótidos.

La clonación en un vector plasmídico, no resultó eficiente al realizarse en repetidas ocasiones con cada una de las cuatro bandas de diferente tamaño amplificadas por RT-PCR y puificadas con Gene Clean, por lo que, se secuenciaron directamente con el iniciador

específico KAZ 5´ y se detectaron datos analizables solo de la banda de 154pb, de la cual se presenta la secuencia (Región del iniciador subrayada):

GTTTCGGGANCCGTNCNTGCATGGTGANGTNCGCNGNGCAANGTCGANATACA
NTCGTAGNAACATGTACACGTTGCACATGCACCCGCACACCGNGATCGTATAG
GCNTTGGANCNGCACACCGGATCGTATCGCCGCACACCGGATCGTA

Los resultados obtenidos, se tradujeron a su respectiva secuencia de aminoácidos y se obtuvo lo siguiente (Región del iniciador subrayada):

LRSGVRRYDPVCRSNAYTIPVCGCMCNVYMFLRLYPRPCPRAPHHARTAPET

La información de la secuencia, nos proporcionó homología con diversos fragmentos de diferentes proteínas, entre ellos un inhibidor de proteinasas de rata (Figura 5).

Con la finalidad de corroborar que la secuencia obtenida pertenecía o no a un fragmento de una proteína inhibidora de proteinasas tipo kazal (de acuerdo a las caracteristicas de diseño del iniciador). La secuencia, fue analizada entonces en el programa MegAlign (del Software DNAStar Inc), contra las secuencias reportadas para inhibidores de proteinasas tipo kazal en *Pascifastacus leniusculus* y *Rattus norvegicus* (Figura 6). El resultado del análisis muestra homología entre las secuencias comparadas, pero nos indica claramente que nuestro fragmento no representa a un inhibidor ya que, el patrón de secuencia no presentó las relaciones topológicas de cisteínas, característicos de los inhibidores de proteinasas tipo kazal [CX (7)-CX (6) YX (3)-CX (2,3) C].

FIGURA 5: Porcentajes y nombre de la proteína con las que presento identidad nuestro fragmento de 154pb al ser alineada en la base de datos. No se detecta la identidad u homología con proteínas inhibidoras de proteinasas. Las letras coloreadas representan los aminoacidos idénticos a los de nuestra secuencia.

```
                                                                  1 [        .         .         .         .         : ] 52
                                                        100.0%    LRSGVRRYDPVCRSNAYTIPVCGCMCNVYMFLRLYPRPCPRAPHHARTAPET
 1 DTDP-glucose-4,6-dehydratase                          37.5%    VRAGVERFVHVSTDEVY-LPVCVTRCS----LKVQPVPAPAADGHALT----
 2 Hypothetical protein (Fragment).                      32.4%    ------------REPRYTLLVAACICHAVILLQLwpQPFPKSTQHSKKA---
 3 Dolichyl-P-Man Man(5)GlcNAc(2)-PP:                    32.4%   ------------ REPRYTLLVAACICHAVILLQLwpQPFPKSTQHSKKA---
 4 Submaxillary gland androgen regul                     27.3%    ---GPRRHDP-------------------------PPPCPPVPPHPR-----
 5 Submaxillary gland androgen regul                     27.3%    ---GPRRHDP-------------------------PPPCPPVPPHPR-----
 6 P0007F06.1 protein (OJ1125_C04.11 pro                 28.6%    ---GIARSSRQCRENSST-----------------PQPCPQEGYHTRKRPES
 7 Hypothetical protein (Fragment).                      30.2%    -RAGSARQPPPCRSS----------------LALRPGPCWKRLH--------
 8 Androgenic gland hormone precursor.                   31.1%    ------RSDVLCGDIRFTVQ---CICNeyFPTERLDKPCPWPNREKRSAPE-
 9 50S ribosomal protein L44E.                           30.6%    ----VRRYCPYCKKHTIHI---------------FPRPLP------------
10 Hypothetical protein FLJ34218.                        33.3%    ----------LCMTVAYtlGACVCVC-VCLYVCVHVCECARAPRRVRPA---
11 Serrano protein.Drosophila melanogaster.              32.7%    LSKGLARFEPICES-------CGSILNVHGYLSS-PHPLPLPLPHSRSS---
12 CG12758-PA. Drosophila melanogaster (Fruit fly).      32.7%    LSKGLARFEPICES-------CGSILNVHGYLSS-PHPLPLPLPHSRSS---
13 Endosperm tissue precursor. Sorghum vulgare           40.6%    -------------SAASTLTTGGCGCHPHPTLPPHPHPCPTYPPH-------
14 GATA-4 zinc-finger transcription factor. Rabbit       50.0%    ------------RRNAEGKPVCNA-CGLYMKLHGVPRP--------------
15 Hypothetical protein. Bacillus anthracis.             50.0%    ---------------AYQVPKCDCGNNlyMFDKLY-----------------
16 Hypothetical protein. Streptomyces avermitilis.       44.4%    -----------CPTDGYHVPECTCACCVY-----------------------
17 Hypothetical protein FLJ38312. Human                  41.9%    ----------VCIYMCVCVPVCMCVY-VCVYVRI-PRPCQR-----------
18 1500010O20Rik protein. Mus musculus                   32.7%    LRKGKRENVNSWRTTGPTHPVPGCAvsPHRHPHPLPRVLPSFPQHNRANPRS
19 AOH1 homolog (Fragment). Mus musculus                 38.2%    ----ISRYDPISkaTACLVPICGMVMSIYTLLRNHPEP--------------
20 Conserved hypothetical protein. Neurospora crassa.    46.5%    LRSGTRRYDEWC----YTSPIW-----TYL-LALLPRPTPDFP---------
21 Predicted protein. Neurospora crassa                  46.5%    LRSGTRRYDEWC----YTSPIW-----TYL-LALLPRPTPDFP---------
22 HypD protein. Rhodobacter sphaeroides                 36.8%    ---GIHAYDFVARDYSKPLVVAGCACGAVMTGRIKPPACPQ-----------
23 T-cell immune regulator 1 transcript variant3. Human. 30.3%    ----------------HTIEFcgCVSNTASYLRLWALSLAHARWNSRTS---
24 Hypothetical protein. Pirellula sp                    28.9%    -----RKHDPILANGRIQFKTCGLMLAVGQLISLRSRSLGRCP---------
25 Hypothetical protein BH1870. Bacillus halodurans      36.4%    ----IQWYEPTANSSAYnaPLC---CGRPMVLSLSPGPApkYAHHFKT----
26 Toll interacting protein. Mus musculus                35.3%    ---GMTRMDPYCR-----VPIADLCCAIYGNTTLHPR---------------
27 A disintegrin and metalloprotease domain 22. Mouse.   35.9%    ----------VCSGNGVCSNELKCVCN-----RKYPYPMPPLPDEGKTA---
28 Hypothetical protein. Oriza sativa                    30.8%    ---------PVCWSN--------------------PNLCHNLPHHAST----
29 Pancreatic secretory trypsin inhibitor I precursor    47.6%    -----RIYDPVCGTNGITYPSECSLC--------------------------
 (PSTI-I) (Cholecystokinin-releasing peptide) (Monitor
 peptide). Rattus norvegicus
30 CG31758-PA. Drosophila melanogaster                   60.0%    -----RNYEPVCGSNLVTYP--------------------------------
   consensus/100%            ..................................................
   consensus/90%             ..................................................
   consensus/80%             ...................................s.............
   consensus/70%                ..........t...................t..P.sh..........
```

```
   XXXXX-------------------------XCPXX---CPRXYDPV Majority
            10        20        30        40
1  ------------------------------RCPSI---CPLNYKPV Plkazinh1.PRO
1  LRSGV-----------------------------------RRYDPV KAZvannamei.PRO
1  MKVAIIFLLSALALLSLAGNPPAEVNGKTPNCPKQIMGCPRIYDPV inhratPS1.PRO

   CGSNXXTYP-XC---XXXCXXXXXXRLYXXPCXXXPXHXXRXXXP  Majority
    50        60        70        80        90
14 CGSDLKTYGNSCQLNAAICRNPSLKKLYDGPCIDKPQPQCPSICP  Plkazinh1.PRO
12 CRSNAYTIP-VC---GCMCNVYMFLRLYPRPCPRAPHHAARLPKP  KAZvannamei.PRO
47 CGTNGITYP-------SECSLCFENRKFG-----TSIHIQRRGTC  inhratPS1.PRO
```

Figura 6 Alineación de la secuencia de aminoácidos deducida del fragmento obtenido de 154pb contra las secuencias reportadas del langostino europeo *Pascifastacus leniusculus* (arriba) y PS1 de rata (abajo) como inhibidores de proteinasas tipo Kazal. Las posiciones de las cisteínas 1, 2, 3 y 4 (mostradas con flechas) están conservadas en todos los inhibidores de proteinasas tipo kazal de acuerdo a la fórmula: [CX (7)-CX (6) YX (3)-CX (2,3) C], nuestra secuencia presenta esta homología solo parcialmente, lo que no proporciona evidencia de que la secuencia objeto de análisis sea un inhibidor.

Secuenciación de transcritos clonados obtenidos a partir de 6 iniciadores diseñados para obtener secuencias nucleotídicas codificantes para proteínas inhibidoras de proteinasas.

Dada la diversidad de transcritos que se dedujo se habían obtenido por la mala calidad en las secuencias, la gran cantidad de inespecificidades detectadas (N´s) y dada las complicaciones para clonar los fragmentos purificados de los productos del bioensayo, se suspendieron los experimentos con ARNm obtenidos de camarones en diferente período de ayuno.

Con el objeto de detectar por secuenciación un transcrito codificante para un inhibidor de proteinasas, se clonaron, los productos de RT-PCR obtenidos y representados en la Figura 2 para su posterior manipulación en forma de plásmido (Figura 7).

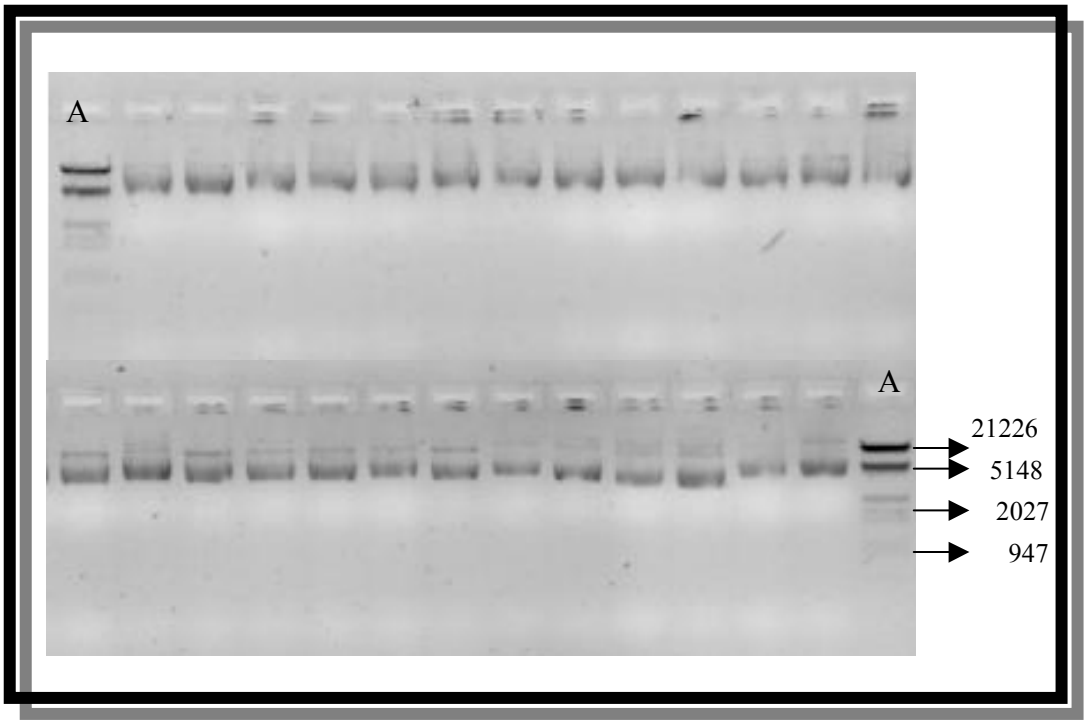


Figura 7. ADN plasmídico extraído de clonas positivas. Material para secuenciación. A: Marcador de Peso Molecular, Lambda DNA/EcoRI+Hind III

Una vez obtenidas las clonas positivas de la ligación de nuestro fragmento en el plásmido, se sometieron a PCR para comprobar la presencia del inserto deseado. (Figura 8). Se obtuvo gran diversidad de clonas en cuanto a tamaño del inserto, lo cual se esperaba dada la inespecificidad y variedad de productos clonados y amplificados por RT-PCR.

Se seleccionaron para secuenciación 65 clonas de las totales obtenidas (Tabla II), de acuerdo al tamaño del inserto obtenido, tratando de evitar en gran medida los insertos de pesos moleculares visiblemente homólogos. (Ver flechas en la figura 8).

Análisis de secuencias

Una vez obtenidas las secuencias se hizo un análisis de comparación entre el tamaño del fragmento esperado de acuerdo al peso molecular estimado en un gel de agarosa que se muestra en la figura 8, y el tamaño del fragmento obtenido de la secuenciación de acuerdo al número de bases (Tabla III).

TABLA II. Relación de clonas positivas obtenidas y clonas seleccionadas para secuenciación de cada iniciador. Se obtuvo mayor diversidad de tamaños en los fragmentos clonados a partir de los productos de RT-PCR realizados con el iniciador KAZ, por lo que el número de clonas seleccionadas para secuenciación también fué más alto

NOMBRE DEL OLIGONUCLEÓTIDO USADO COMO INICIADOR	Número de Clonas obtenidas	Número de Clonas secuenciadas
ANEK	19	5
INHT	45	5
INSK	30	7
INVK	15	3
PANK	43	7
KAZ	78	38
TOTAL	230	65

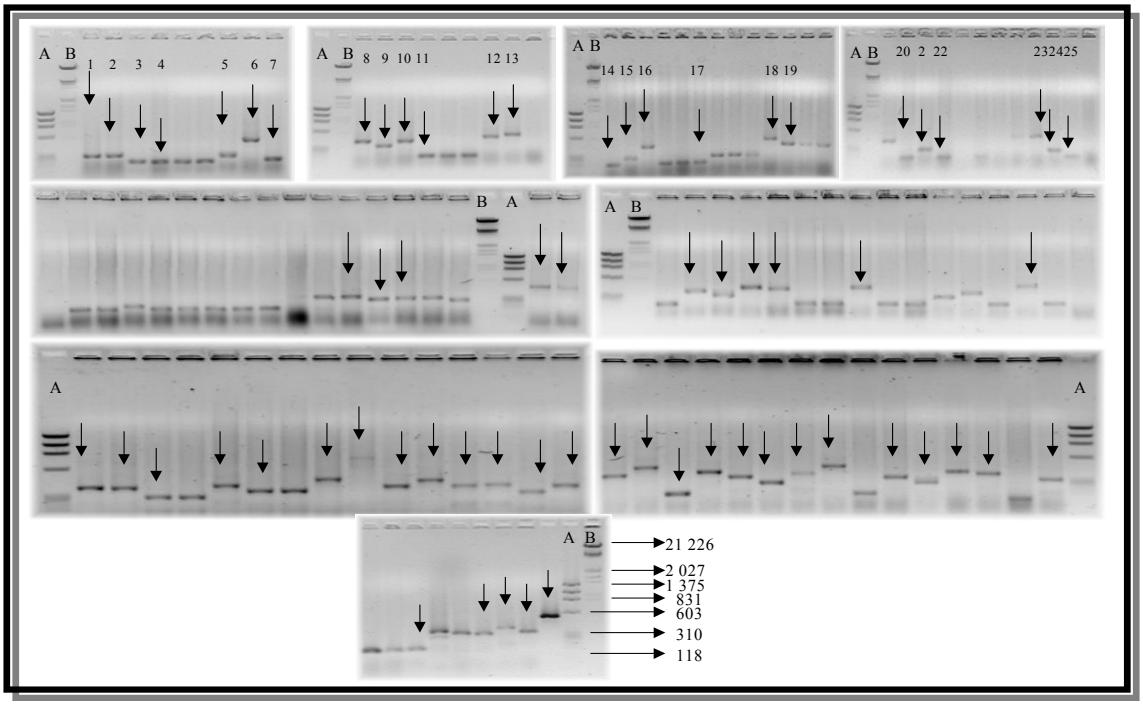


Figura 8. Productos de PCR con iniciadores específicos para el vector TOPO2.1 pCR, M13, a partir de clonas positivas. Las flechas indican las muestras seleccionadas para secuenciación. A y B, marcadores de Peso Molecular ØX174 DNA/Hae III y Lambda DNA/EcoRI+ Hind III repectivamente. Relación de números, ver Tabla III.

Tabla III. Relación de nombres de las clonas seleccionadas para secuenciar. La primer columna contiene el número de referencia del gel (Figura 6) y la segunda el nombre asignado de acuerdo al iniciador utilizado y el

número de la clona. El tamaño esperado del inserto se describe en kilobases, porque fué como lo requirió el laboratorio responsable de la secuenciación, el tamaño obtenido, se reporta de la misma manera para homogeneizar datos.

No	Nombre de la muestra	Tamaño esperado del inserto (Kb)	Tamaño obtenido del inserto (Kb)
1	KAZ 1	0.2	0.090
2	KAZ 2	0.2	0.080
3	KAZ 3	0.15	s/inserto
4	ANEK 3	0.15	s/inserto
5	ANEK 4	0.1	s/inserto
6	ANEK 5	0.15	0.055
7	ANEK 6	0.3	0.200
8	INH 1	0.2	0.055
9	INH 2	0.4	0.095
10	INH 3	0.15	s/inserto
11	INH 4	0.1	s/inserto
12	INH 8	0.4	0.300
13	INSK 1	0.25	0.220
14	INSK 2	0.2	0.140
15	INSK 3	0.25	0.190
16	INSK 4	0.3	0.220
17	INSK 8	0.1	s/inserto
18	INVK 3	0.1	s/inserto
19	INVK 4	0.2	0.090
20	INVK 5*	0.1	---------
21	PANK 1	0.3	0.280
22	PANK 2	0.4	0.320
23	PANK 5	0.3	0.280
24	PANK 6	0.2	0.068
25	PANK 7	0.1	s/inserto
26	K32	0.15	0.130
27	K33	0.15	0.120
28	K41	0.15	0.100
29	K52	0.2	0.145
30	A3	0.25	0.210
31	IS1	0.15	0.125
32	IS3	0.2	s/inserto
33	P2	0.25	0.220
34	P3	0.25	0.280
35	KA1	0.15	0.090
36	KA2	0.15	0.125
37	KA3	0.1	0.090
38	KA5	0.15	0.115
39	KA6	0.15	0.110
40	KA8	0.15	0.090
41	KA9	0.4	0.200
42	KA10	0.15	0.090
43	KA11	0.15	0.180
44	KA12	0.15	0.140
45	KA13	0.2	0.265
46	KA14	0.15	0.560
47	KA15	0.15	0.180
48	KA16	0.15	0.245
49	KA17**	0.3	----------
50	KA18	0.1	0.570
51	KA19	0.25	0.120
52	KA20	0.2	0.180
53	KA21	0.15	0.180
54	KA22	0.2	0.270
55	KA23	0.25	s/inserto
56	KA25	0.15	0.220
57	KA26	0.15	0.200
58	KA27	0.15	0.145
59	KA28	0.2	0.085
60	KA30	0.15	s/inserto
61	KA33	0.2	s/inserto
62	KA36	0.2	0.145
63	KA37	0.2	0.085
64	KA38	0.2	0.200
65	KA39	0.2	s/inserto

*Resultado no enviado. **Resultado no analizable, mucha inespecificidades(N'

Se obtuvieron 50 secuencias analizables a partir de las clonas seleccionadas (Apéndice 2), en las cuales se detectó, la región del iniciador 5´en ambos extremos de cada uno de los fragmentos (Tabla IV). Se hizo una comparación entre secuencias con el programa GenQuest, para detectar repetición de secuencias, e individualmente por Blast con las secuencias del GenBank (nr y dbEST) del NCBI, obteniéndose los resultados observados en la Tabla V.

TABLA IV. Secuencia del iniciador KAZ 5´ localizado a ambos extremos de los amplicones obtenidos. El oligonucleótido específico actuó como iniciador en ambos sentidos, lo cual se puede comprobar al deducir la secuencia Río abajo complementaria, que es la misma que que el iniciador Río arriba con pequeñas diferencias en los lugares de degeneración. *K= G o T, S= C o G, W= A o T

SECUENCIA (NOMBRE)	INICIADOR RIO ARRIBA	INICIADOR RIO ABAJO	INICIADOR RIO ABAJO COMPLEMENTARIO
KA18(570)	--CGATCCGGTGTGCGGCA	ACCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGT
KA14(560)	--------GGTGTGCGGCA	ACCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGT
K52(145)	TACGATCCTGTGTGCGGGT	AGCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGCT
KA36(145)	TACGATCCTGTGTGCGGGT	AGCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGCT
KA8(90)	TACGATCCGGTCTGCGGGT	ACCCGCACACAGGATCGTA	TACGATCCTGTGTGCGGGT
KA10 (90)	TACGATCCGGTCTGCGGGT	ACCCGCACACAGGATCGTA	TACGATCCTGTGTGCGGGT
K33(120)	TACGATCCGGTGTGCGGCT	ACCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGT
KA1(90)	TACGATCCGGTCTGCGGGT	AGCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGCT
KA12(140)	TACGATCCGGTCTGCGGGT	TCCCGCACACTGGATTGTA	TACAATCCAGTGTGCGGGA
KA19(120)	TACGATCCGGTGTGCGGGT	ACCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGT
KA6(110)	TACGATCCAGTGTGCGGCA	ACCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGT
KAZ2(80)	T-CGATCCGGTCTGCGGGT	AGCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGCT
KAZ1(90)	TTCGATCCGGTCTGCGGGT	AGCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGCT
KA32(130)	TACGATCCGGTGTGCGGGT	ACCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGT
KA13(265)	TACGATCCGGTGTGCGGCA	TCCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGA
KA2(125)	TACGATCCGGTGTGCGGGA	TGCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGCA
KA3(90)	TACGATCCGGTCTGCGGCT	ACCCGCACACAGGATCGTA	TACGATCCTGTGTGCGGGT
KA9(200)	TACGATCCTGTGTGCGGCA	ACCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGT
KA27(145)	TACGATCCGGTGTGCGGGA	TGCCGCACACAGGATCGTA	TACGATCCTGTGTGCGGCA
KA26(200)	TACGATCCTGTGTGCGGCA	TCCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGA
K41(100)	TACGATCCGGTCTGCGGGA	AGCCGCAGACAGGATCGTA	TACGATCCTGTCTGCGGCT
KA11(180)	TACGATCCAGTGTGCGGGA	TCCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGA
KA20(180)	TACGATCCTGTGTGCGGGA	TGCCGCAGACTGGATCGTA	TACGATCCAGTCTGCGGCA
KA38(200)	TACGATCCGGTCTGCGGCT	TCCCGCACACAGGATCGTA	TACGATCCTGTGTGCGGGA
KA25(220)	TACGATCCTGTGTGCGGCA	TGCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGCA
KA22(270)	TACGATCCTGTGTGCGGGA	TCCCGCAGACAGGATCGTA	TACGATCCTGTCTGCGGGA
KA16(245)	TACGATCCGGTCTGCGGCT	TCCCGCAGACCGGATCGTA	TACGATCCGGTCTGCGGGA
KA37(85)	TACGATCCTGTGTGCGGCT	ACCCGCAACACCGGATCGA	T-CGATCCGGTGTTGCGGG
KA28(85)	TACGATCCTGTCTGCGGGT	TGCCGCACACCGGATCA-A	T-TGATCCGGTGTGCGGCA
KA5(115)	TACGATCCTGTCTGCGGCA	ACCCGCACACAGGATCG-A	T-CGATCCTGTGTGCGGGT
KAZ1(90)	TTCGATCCGGTCTGCGGGT	AGCCGCACACCGGATCGTA	TACGATCCGGTGTGCGGGT
KA15(180)	TACGATCCTGTCTGCGGGA	AGCCGCACACTGGATCGTA	TACGATCCAGTGTGCGGGT
SECUENCIA CONSENSO	TACGATCCKGTSTGCGGSW*		TACGATCCKGTSTGCGGSW*

Tabla V. Análisis de secuencias. Deducción de la relación secuencia-función de los amplicones obtenidos a partir de iniciadores diseñados a partir de secuencias peptídicas conservadas de proteínas inhibidoras de proteinasas reportadas en las bases de datos de la WWW, de acuerdo a la homología que estas presentaron mediante el uso del BLASTX. BLASTN y dbEST (solo en las secuencias donde se especifica).

No se tomó ningún criterio en el ordenamiento de las secuencias dentro de la tabla. Se desglosaron primero todas las secuencias KAZ, por ser de las que teníamos mas representantes; posteriormente las obtenidas con los otros iniciadores en orden alfabético: ANEK, INHT, INSK, INVK y PANK.

INSERTO NOMBRE Y TAMAÑO	CLONAS HERMANAS	BLASTX (FRAME, NOMBRE Y TAMAÑO DE LA PROTEÍNA	ESPECIE Y POSICIÓN DEL ALINEAMIENTO QUERY/SBJCT	ID %	(+) %	FUNCIÓN DE LA PROTEÍNA HOMÓLOGA	OBSERVACIONES EN BASE A LA PRESENCIA DEL INICIADOR KAZ 5
KA18 (570)	KA14 (560)	FRAME +2 producto de proteína no nombrada , 350	*Homo sapiens* 173-472 /166-266	30	46		No iniciador
		Precursor del adaptador transmembranal tipo I de Retículo Endoplásmico, 350		30	45		
		Receptor de Renina, 350	173-472 /166-266	30	46	Control de la presión sanguínea y osmorregulación (Nguyen et al, 2002)	
K52 (145)	KA36 (145)	EST de branquia	*Penaeus vannamei*				blastx (dbEST) 7 bases del iniciador
		FRAME +3 producto de proteína no nombrada, 201	*Oriza sativa* (*japónica* cultivar-group) 15-140/24-65	45	47		
		FRAME -1 proteina como-extensina, 1016	*Zea mays* 17-124/948-913	44	66	Proteínas estructurales, relacionadas con pared celular en plantas (Bucher et al, 2002; Destoumieux et al, 2000)	5 bases del iniciador
KA8 (90)	KA10 (90)	----------					No alineamiento
K33 (120)	-------	FRAME +1 Posible proteína membrana, 262	*Mycobacterium bovis* subsp. Bovis 13-84/230-253	50	66	Transporte	10 bases del iniciador
KA1 (90)	KA12 (140)	FRAME -2 similar a mucina, 236	*Rattus norvegicus* 17-142/210-175	42	50	Protección contra abrasión por alimento	6 bases del iniciador

			44-142/149-119 44-142/151-181 17-142/138-103 44-142/77-47 44-142/173-143	45 45 42 45 42	51 51 50 48 48	y parásitos invasivos (Wang y Granados, 1997; Shen et al, 1999; Carraway et al, 1999; Komatsu et al, 1999)	
KA19 (120)	KA6 (110)	FRAME+1 proteína hipotética, 229	*Rhodobacter sphaeroides* 28-96/97-119	52	63	Función no descrita	KA19 no blastx, pero homologa con KA6 No primer
KAZ2 (80)	KAZ1 (90)	----------					No alineamiento
KA32 (130)	-------	FRAME +2 proteína hipotética, 254	*Rhodopseudomonas palustris* 1-123/6-45	52	54	Función no definida	Incluye iniciador completo
KA13 (265)	-------	FRAME –3 precursor del factor IX de coagulación , 537	*Takifugu rubripes* 5-130/339-296	47	68	Proteína del plasma Vitamina K dependiente que participe en la ruta intrínseca de la coagulación sanguínea, convirtiendo al factor X a su forma activa en presencia de iones Ca++, fosfolípidos y factor VIIIa. (EBI SRS server., 2003)	Incluye iniciador completo
KA2 (125)	-------	FRAME+1 CG4090-PA, 2112 ENSANGP00000022434, 260	*Drosophila melanogaster* 22-75/1768-1785 16-123/141-181	61 39	72 53	Función no conocida	No iniciador
KA3 (90)	-------	----------					No alineamiento
KA9 (200)		FRAME +1 similar to Galbeta1-3GalNAc 3-sulfotransferase, 47 3'-	*Homo sapiens* *Rattus norvegicus* 110-156/269-317 1-180/230-289	63 50	71 61	Sulfatación de sustratos aceptores (Fukuda, 2002; Honke, 2003)	No iniciador Incluye región del iniciador

		phosphoadenylylsulfate: galactosylceramide 3', 423					
KA26 (200)		FRAME-1 AT01956p, 651	*Drosophila melanogaster* 22-165/204-157	12	25		No iniciador
K41 (100)	----------						No alineamiento
KA11 (180)	KA20 (180)	Secuencia parcial de la subunidad 18s del ARN ribosomal, 8119	*Litopenaeus vannamei* 22-162/7312-7452	99		estructura de los ribosomas y en la formación del enlace peptídico que ensambla a los aminoácidos durante la síntesis de proteínas	No iniciador alineamiento en blastn
KA38 (200)		FRAME +2 Proteína Fot-G, 398	*Escherichia coli* 14-181/174-231	37	53	Biogénesis fimbrial (Honarvar, 2003)	10 bases del iniciador
KA25 (220)		FRAME = ENSANGP00000009992, 572	*Anopheles gambiae* 1-207/128-197	67	80	Función no conocida	Incluye región del iniciador
		Aldehído dehidrogenasa type III CG11140-PA, 561	*Drosophila melanogaster* 25-207/96-156	63	75	Metabolizacion de aldehídos	
KA22 (270)		FRAME –1 Precursor de Fasciclina II, 898	*Schistocerca americana* *Drosophila melanogaster* 55-261/171-103	40	60	Adhesión celular y señalización. (Hu et al, 1997)	No iniciador
		similar a hemicentina, 4786	*Homo sapiens, Rattus norvegicus* 252-55/169-241			Diferenciación de células epiteliales especializadas (Vogel, 2001)	
			252-34/1016-1089	40	51		
			258-55/641-709	29	44		
			258-10/266-340	28	44		
			252-46/1944-2015	36	43		
			249-46/1110-1176	29	48		
			249-55/2927-2990	31	52		
			255-58/829-895	31	45		
			249-34/1842-1924	29	44		
			252-46/549-618	28	42		

			249-46/737-806	27	44		
			249-46/2227-2294	34	47		
			252-46/2602-2671	30	49		
			249-55/2132-2199	30	45		
			246-46/1670-1740	29	42		
			249-46/924-991	29	42		
			129-46/2458-2485	28	47		
			249-46/1202-1268	39	57		
			249-37/360-434	27	45		
			249-46/2040-2107	29	41		
			258-55/1573-1641	28	43		
				26	42		
KA16 (245)		FRAME –3 ascites sialoglicoproteina 1 – de rata	*Rattus norvegicus* 233-12/779-851	28	41	Protección (Komatsu, 1999)	Codón stop intermedio Incluye iniciador
KA37 (85)	----------						No alineamiento
KA28 (85)		FRAME –2 A Chain, Formate DehydrogenaseN, 1015	*Escherichia coli* 7-78/402-379	83	87	Respiración anaeróbica, proteína de membrana (Hafid y Mandrand, 1998)	Incluye región del iniciador
KA5 (115)		FRAME +3 Putative hidroxiprolina rich glycoprotein, 464	*Oryza sativa* (japonica cultivar-group) 21-86/360-381	54	59	Extensina. Encontrada en la pared celular producen respuesta a la agresión (Ahn et al, 1998)	No iniciador
		possible G-protein receptor, 581	*Homo sapiens* 12-86/470-496	51	59	Transportadora de membrana (Joost y Methner, 2002)	
KA27 (145)		Precursor de peritrofina ovárica 1, 277	*Penaeus monodon* 117-28/117-146	36	60	Componente precursor de bastones corticales de la capa gelatinosa de los huevos del camarón, se glicosila y se une a quitina, justo antes del desove. (Khayat et al, 2001).	No iniciador
KA21 (90)	KA15 (180)	FRAME -2	*Mus musculus*			Inhibición del proceso	No iniciador

		esfingomielina fosfodiesterasa 1, ácido lisosomal, 627	165-25/155-203	34	42	de apoptosis celular (Sillence, 2001)	
A3 (210)	ANEK6 (200)	FRAME +2 Proteína hipotética, 402	*Plasmodium yoelii yoelii* 110-193/257-284	64	82		No iniciador
		similar a proteína CDC27Hs (Cell division cycle protein 27 homolog), 450	*Homo sapiens* 33-134/105-138	64	73	División celular ciclo 27 (Locus link, 2003)	
ANEK5(55) INH2 (95)	---------- INH8 (300)	FRAME –1 KIAA0995 proteína , 1014	*Homo sapiens* 291-22/440-530	26	43	Similar a Pecanex de *Drosophila melanogaster* involucrada en la neurogénesis (EMBL Bioinformatic Harvester, 2003)	No alineamiento No iniciador
		RPGR proteína (Regulador de Retinitis pigmentosa GTPasa), 1001	*Mus musculus* 258-25/704-775	25	43	Distrofias retinales. (Roepman et al, 2000; Hong et al, 2000)	
INH1 (55) INSK1 (220)	---------- INSK4 (220) INSK3 (190) P2 (220)	FRAME –2 Proteína de choque térmico 83, 370	*Drosophila melanogaster* 16-141/315-274	92	97	Respuesta al calor o a agentes tóxicos.Protección celular, interactúa con receptores de hormonas esteroideas y regula su actividad. Homologa a HSP90 de mamíferos. (Ding,	No alineamiento 12 bases del iniciador

						1993; Gross, 2001)	
IS1 (125)		FRAME +1 Proteína hipotética, 138	*Burkholderia fugorum* 43-108/2-23	50	68	Función no conocida	No iniciador
INSK2 (140)		FRAME -3 DELTEX 3, 347	*Mus musculus* 1-138/169-123	40	51	Regulador de la Ruta Notch (Diederich et al., 1994)	Codón stop Iniciador completo
INVK4 (90)	----------						No alineamiento
P3 (280)	PANK1 (280) PANK5 (280)	FRAME –2 receptor de tirosina kinasa, 1348	*Hydra vulgaris* 17-211/198-134 17-211/335-272 17-233/464-401	26 26 23	55 46 44	Morfogénesis neural y epitelial. Son proteínas transmembranales reguladores clave en procesos celulares como el ciclo celular, migración, metabolismo, sobrevivencia, proliferación y diferenciación celular. (Bell et al., 2000; Grassot et al., 2002)	12 bases del iniciador
		Receptor de manosa, C tipo 2	*Homo sapiens* 229-557/485-557 211-17/346-418	29 28	39 37		
PANK2 (320)		FRAME +1 Proteína de recombinación meiótica, mei-218 - fruit fly, 1186	*Drosophila melanogaster* 106-279/657-715	34	54	Participa en la oogénesis, desarrollo de oocitos y entrecruzamiento de gametos. (Manheim, 2002; McKim, s/fecha)	No iniciador
		CG33067-PA,1186	106-279/657-715	34	54		
PANK6 (68)	----------						No alineamiento

IV. DISCUSIÓN

El objetivo central del estudio del genoma es identificar y clasificar todos los genes de una especie en particular. La genómica funcional da a conocer los roles precisos de estos genes, incluyendo funciones únicas y redundantes. (Fernández *et al.*, 2002).

La industria camaronera en los últimos años ha enfrentado graves problemas que han amenazado la producción, lo que ha llevado a desarrollar estrategias para contrarrestar estas amenazas, como la generación de cultivos en ciclo cerrado y programas de selección y mejoramiento genético. Los programas de mejoramiento basados en la cría de individuos con características fenotípicas (características externas observables) de interés comercial, para incrementar peso o resistencia a enfermedades, requieren mucho tiempo para la obtención de resultados, pero pueden ser acelerados al incluir técnicas moleculares para su desarrollo. Sin embargo, la información disponible acerca de la composición genética en camarones es fragmentada y las tecnologías dirigidas a solucionar problemas básicos como el mapeo genético aún no están disponibles o tienen un desarrollo muy limitado.

Se estima que miles de genes son los que conforman el genoma de los organismos, el cual es como una huella digital para cada uno de ellos. Estos genes, pueden ser expresados en contextos temporales y espaciales específicos. En estudios pasados, las respuestas de las células y de los organismos eran estudiadas en un contexto limitado, por ejemplo, un gen o ruta metabólica, en un tiempo. El objetivo de la nueva Genómica funcional, es brindar nuevas tecnologías para ampliar el estudio de la expresión de genes a gran escala. El alcance de esto no rendirá conocimientos detallados de los procesos biológicos como los que la bioquímica, la biología celular, la fisiología y la farmacología,

tradicionalmente han aportado, pero dará a los biólogos un nuevo, comprensivo y holístico conocimiento de sistemas complejos y describirá la relación entre secuencia y función.

Es posible pensar esto, por los estudios de expresión que se han estado generado en los últimos años con el análisis de datos de secuencias obtenidas de ADNc de muchos organismos. De esta manera muchos esfuerzos están dirigidos a desarrollar tecnologías que permitan estudiar perfiles de ARNm a gran escala. (Brownstein, 1998), iniciando por la estandarización de sus métodos de extracción y el cuidado hacia cualquier contaminación, para conservar la integridad de la molécula, que desde siempre han sido los puntos de partida metodológicos clave para el desarrollo de trabajos de expresión génica en cualquier organismo.

Otro aspecto metodológico clave para alcanzar los objetivos de obtener secuencias expresadas, es el planteamiento de cómo y a partir de qué, obtenerlas. En la actualidad, existe un gran interés por la generación de información genética acerca de la expresión de genes tejido específica, por lo que el estudio del transcriptoma ha tomado gran importancia y las metodologías iniciales mas comunes es la práctica de obtener librerías de ADNc, para la obtención de ESTs, que son fragmentos cortos de ADNc, secuenciadas por paso simple generadas a partir de clonas seleccionadas al azar en una librería. Algunos ejemplos de trabajos que se han hecho acerca de la identificación de ESTs y su análisis y manejo en base de datos son los de Wolfsberg y Lansman, 1997; Verdum *et al.*, 1998; Scharfe *et al.*, 1999; White *et al.*, 2000; Reymond *et al.*, 2000, Liang *et al.*, 2000; Lange *et al.*, 2000; Kawashima, 2000; Skarabanek y Campagne, 2001 y McCarter *et al.*, 2003.

En el presente trabajo, la metodología para la obtención de transcritos en hepatopáncreas de camarón fue con el diseño y uso de iniciadores degenerados (Apéndice 3), con la finalidad de detectar inhibidores de proteinasas. El resultado arrojó evidencia de la presencia de diferentes transcritos que se supuso podían corresponder a la expresión de diferentes fragmentos correspondientes a distintos dominios de proteínas inhibidoras de proteinasas basándonos en el conocimiento teórico de que los inhibidores poseen varios dominios que pudieron haberse originado por duplicación repetida de la secuencia (Laskowski y Kato, 1980), lo que sugiere que los iniciadores pudieron haberse alineado en cualquiera de sus dominios dada la alta homología en su secuencia que presentan entre ellos, resultando la amplificación de transcritos de diferentes tamaños.

Con la idea de observar si la amplificación de las bandas observadas (Figura 2), se mantenía constante al someter al camarón a un estrés fisiológico inducido como el ayuno, se programó un bioensayo sometiendo a organismos a diferentes tratamientos para su posterior aislamiento y análisis de ARN de hepatopáncreas por RT-PCR con solo uno de los iniciadores (KAZ). Los experimentos derivados del bioensayo, se suspendieron por la gran cantidad de inespecificidades (N´s) obtenidas en las secuencias de los productos de la figura 4 y la baja eficiencia de la clonación de tales productos, como alternativa para secuenciar fragmentos únicos. La eficiencia en la clonación disminuye fuertemente en productos expuestos a luz ultravioleta, el alto grado de exposición de los ácidos nucleicos a tal efecto físico, provoca la formación de dímeros pirimídicos que impiden la ligación del fragmento al vector.

El análisis de la única secuencia analizable (banda de 154pb), derivada del bioensayo muestra homología con fragmentos de proteínas con alto contenido de cisteínas, un aminoácido de gran importancia para la conformación de estructura terciaria de la proteína, lo cual pudiera sugerir la presencia de pequeñas secuencias altamente conservadas ricas en cisteínas en muchas proteínas, que no están relacionadas en cuanto a la función que desempeñan en los organismos.

Por la gran inespecificidad en la secuencia, se sospechó de la presencia de fragmentos diferentes correspondientes a distintas proteínas en la misma banda, por lo que se requería de la clonación, para asegurar la secuenciación de solo un transcrito a la vez, predecir parte de su estructura molecular y función y detectar la presencia de algún inhibidor de proteinasas.

Aunque evidentemente los diferentes períodos de ayuno manifestaron la expresión de diferentes proteínas, lo cual implica una posible adecuación en la maquinaria fisiológica del camarón manifestada mediante la expresión por requerimiento de diferentes proteínas para mantenerla estable. La estructura y función de tales proteínas a la fecha se desconoce, por lo que esta tesis pretende ser un estudio preliminar aportando fragmentos clonados y secuenciados derivados de transcritos obtenidos de hepatopáncreas de camarones mantenidos en condiciones óptimas para su desarrollo que permitan caracterizar genes completos para su posterior análisis de expresión bajo diferentes condiciones que permitan elucidar un poco acerca de la fisiología de éste órgano.

La razón por la que se pensó en la presencia de diferentes transcritos codificantes para diferentes proteínas, es por el hecho de haber utilizado iniciadores degenerados, ya

que su especificidad va a depender del grado de degeneración del iniciador. Aunque el uso de iniciadores degenerados es una importante herramienta en prácticas de Biología Molecular para la detección de genes (altamente conservados en muchos organismos) en especies recientemente estudiadas (Yuen *et al.*, 2001).

A la fecha, la mejor ruta para clonar un gen y para análisis genómico de librerías de ADNc por hibridación (Lathe, 1985), cuando únicamente una secuencia fragmentaria de la proteína se conoce, es utilizar secuencias de aminoácidos para diseñar oligonucleótidos que puedan ser usados para hacer pruebas para analizar librerías o para amplificar una secuencia blanco por PCR. Un problema obvio con ambos métodos es la degeneración del código genético. Éste tiene 64 posibles codones y únicamente 20 aminoácidos; muchos aminoácidos son representados por más de un codón. Por lo que no hay reglas para garantizar la selección de un codón correcto; pues cierta secuencia de aminoácidos puede ser codificada por un gran número de diferentes oligonucleótidos. Al menos que el diseñador corra con suerte y escoja el codón correcto en todas las posiciones, el oligonucleótido no formará un híbrido perfecto con el gen blanco o ADNc (Sambrook y Russell, 2001)

De acuerdo a lo anterior, y centrando la discusión en los resultados de nuestro trabajo, tenemos entonces que el uso de iniciadores degenerados en métodos de RT-PCR en un solo paso, nos genero amplicones que representan diferentes genes, que poseían dentro de su secuencia regiones muy similares a la de nuestro iniciador 5', el cual, por el requerimiento de una alta temperatura de alineamiento (no compatible con nuestro iniciador poliT utilizado como iniciador en sentido reverso) fungió como iniciador en

ambos extremos de la secuencia blanco (Tabla V), restringiendo el tamaño de los amplicones a la posición de dichas secuencias oligonucleotídicas en los diferentes blancos, presumiblemente correspondientes a diferentes genes.

La enzima transcriptasa reversa es sensible a altas temperaturas, por lo que el uso de iniciadores específicos con altas temperaturas de alineamiento (Tm) para llevar a cabo la Trascripción Reversa es complicado cuando se realiza en un solo paso, junto con la PCR, ya que se requieren solo de un par de iniciadores específicos o una combinación específico-poliT (por las características de que todos los ARNsm eucariotas poseen un tracto de adeninas). La realización separada de ambos métodos; RT y PCR, el primero con el uso de poliT como iniciador, y el segundo con ambos iniciadores específicos, es al igual una práctica azarosa, ya que se tiene evidencia de que generan una alta frecuencia de ADNc truncados a partir de la iniciación de la transcripción reversa en regiones internas de la cadena nucleotídica (Nam *et al.*, 2002), lo cual no garantiza la obtención integra del transcrito que se desea obtener.

Las secuencias expresadas encontradas en el presente trabajo son claramente diferentes, lo cual nos refleja un rol multifuncional del hepatopáncreas del camarón y no solo digestivo. Si bien es cierto las secuencias aquí reportadas no representan genes completos, por las características de la metodología utilizada, tenemos indicios de una buena parte de las secuencias expresadas en el hepatopáncreas del camarón, aún no reportadas en las bases de datos, a excepción de los fragmentos KA14 y KA18 que se reportó como un EST de branquia de *P. vannamei* , KA11 y KA20 que resultaron ser

fragmentos de ADNc correspondientes al gen parcial que codifica para la subunidad 18s de ARN ribosomal en esta misma especie.

La mayoría de los fragmentos de ADNc obtenidos mostraron, por lo menos en parte, homología con secuencias reportadas en las bases de datos del GenBanK (nr y EST), para otras especies ajenas, por lo menos taxonómica y morfológicamente, por lo cual la función propuesta para las secuencias obtenidas es una suposición subjetiva acerca de lo que realmente podría representar cada uno de los fragmentos de genes reportados en este trabajo, ya que en su mayoría se alinearon solo secuencias cortas de estos fragmentos y manifestaron porcentajes de identidad menores al 50%, o si mayores, representando menos del 50% de la longitud total del inserto.

Indudablemente, nuestro trabajo representa una importante contribución al conocimiento de la composición genética de *P. vannamei* y nos da la pauta para iniciar la caracterización de diferentes genes expresados en el hepatopáncreas de camarón, para elucidar ampliamente las funciones reales que este órgano realiza.

Las clonas Kaz y Pank presentan una secuencia de aminoácidos relacionada con varias proteínas de membrana de diferentes especies de vertebrados e invertebrados, lo cual sugiere que hay aminoácidos hidrofóbicos en su composición. La secuencia consenso (**YDPVCGSD, en el caso de clonas Kaz**), en la que está basado el diseño del primer KAZ5, está compuesta de tres aminoácidos polares, 2 acídicos y uno hidrofóbico lo que le dan una carga neta de –2.2 y un punto isoleléctrico de 3.232 (Tabla V); tales características no parecen estar relacionadas con las proteínas que alinean en el análisis BLASTX.

Tabla VI. Datos bioquímicos acerca del péptido deducido a partir de la secuencia nucleotídica utilizada como iniciador

Información del Péptido deducido	**YDPVCGSD**
Número de aminoácidos	8
Peso Molecular	853.88 Daltons
Aminoácidos fuertemente Básicos (+) (K,R)	0
Aminoácidos fuertemente Acídicos (-) (D,E)	2
Aminoácidos Hidrofóbicos (A,I,L,F,W,V)	1
Aminoácidos Polares (N,C,Q,S,T,Y)	3
Punto Isoeléctrico	3.232
Carga a PH 7.0	-2.121

Las proteínas homólogas reportadas por el BLASTX respecto a cada clona, aunque deben ser consideradas con reserva, es posible advertir que el alto porcentaje de identidad entre sus secuencias establece un primer acercamiento en la identificación del producto codificado. Por ejemplo, KA18 y KA14 podrían estar involucrados en la regulación de la presión de la hemolinfa del camarón, ya que la proteína homóloga tiene esta función en la sangre humana. A su vez KA52 es KA36, su función podría estar relacionada con proteínas de membrana como las extensinas con función de reconocimiento. En el caso de K1 y K12 representan un gen que esta relacionado con las mucinas de rata, éstas se encuentran en el sistema digestivo, lo cual puede ser altamente significativo para su identificación definitiva, ya que en éste caso recordemos que se aisló de RNAm de hepatopáncreas. Otro caso interesante del análisis son las clonas KA2, KA26, PANK2 y las obtenidas con los iniciadores ANEK (A3 y ANEK6), pues la contraparte homóloga se encuentra en insectos como *Drosophila melanogaster* y *Plasmodium yoelii,* parientes cercanos del camarón, por el momento no hay descripción en éstas especie de dichas proteínas, pero el hecho de que la homología tenga una identidad del 61% para KA2 en el nivel de secuencia de ADN atrae la atención. El precursor del factor coagulation IX del pez *Takifugu rubripes*, es la

proteína relacionada en un 52% de identidad con el producto de la clona KA13, aunque su expresión podría esperarse mas en la hemolinfa, los 265 nucleótidos que comprende este inserto sugieren que en el hepatopáncreas podría tener un papel funcional importante. De igual forma la clona KA22 presenta una alta similitud en varias regiones de la secuencia proteica de la hemicentina, la cual tiene la función de organizar y orientar las células del tejido epitelial de ratas y el hombre (Vogel y Hedgecock, 2001). Las identidades van desde 27% hasta 3l y 40% en casi 70 aminoácidos alineados deducidos de la secuencia del inserto de ADN complementario.

Las secuencias obtenidas con el iniciador INHT mostraron una ligera homología con la proteína RPGR. La retinitis pigmentosa es un grupo clínica y genéticamente heterogéneo de distrofias retinales que se caracteriza por producir oscurecimiento y constricción progresiva del campo visual. Estudios de la expresión de sus genes han detectado e identificado splicing alternativo, lo cual da origen a varias isoformas que al ser comparadas en las bases de datos con otras proteínas se detectó significante homología (33-40%) con dominios de proteínas involucradas en el tráfico vesicular a través de membrana en conocidos procesos celulares; de estas isoformas, algunas fueron mayormente expresadas en hígado y riñón (Roepman *et al.*, 2000; Hong *et al.*, 2000).

Se encontró un importante porcentaje de identidad de los insertos obtenidos con INSK con la proteína 83 de Choque térmico de *Drosophila melanogaster*, no reportada a la fecha para tener homología con secuencias de *Penaeus vannamei,* como es el caso de la HSP A Y B identificadas en hepatopáncreas y HSP70 que se encontró en hemocitos. Ésta última que funciona como “chaperona molecular” para facilitar la conformación estructural

madura de las proteínas, previene la desnaturalización de las proteínas bajo condiciones estrés y renaturaliza proteínas desnaturalizadas. Sus niveles incrementan de 3 a 15 veces en respuesta al estrés celular. Entre otras funciones, actúa como ligando para el receptor 4 Toll-like en macrófagos de ratón (Gross, 2001). Tanto la HSP83, como mei-218 (homología con PANK2), son proteínas que intervienen en la oogénesis en *D. melanogaster* (Manheim *et al.*, 2002).

V. CONCLUSIÓN

En el presente trabajo se analizaron 50 secuencias nucleotídicas a partir de clonas con amplicones o fragmentos ligados obtenidos con iniciadores diseñados para reconocer genes que codifican para proteínas inhibidoras de proteinasas de diferentes organismos ya caracterizadas y reportadas en la bases de datos. De estas clonas, se detectaron 35 fragmentos con secuencia nucleotídica diferente, lo que nos proporciona bases para determinar que contamos con una variedad de transcritos o parte de genes que codifican para diferentes proteínas expresadas en el hepatopáncreas del camarón blanco *Penaeus vannamei* .

Los transcritos obtenidos, sin duda alguna, nos da la pauta para continuar el estudio de los genes que se expresan y participan en los diferentes procesos fisiológicos, además de la digestión. El uso de técnicas en Biología Molecular nos permitió, hacer un bosquejo en la versatilidad funcional del hepatopáncreas y tener un acercamiento al conocimiento de la amplia gama de proteínas que en él se sintetizan, para llevar a cabo diferentes procesos biológicos importantes para el desarrollo del organismo.

Dentro de los análisis realizados, en base a las secuencias obtenidas, no se detectó la presencia de fragmentos que determinaran la homología con proteínas inhibidoras de proteinasas. Si bien es cierto que no todas las clonas obtenidas fueron analizadas, la probabilidad de haber amplificado una proteína de este tipo se ve muy disminuida, debido a que después del análisis de los amplicones secuenciados detectamos que las condiciones

que predominaron durante la reacción de RT-PCR, no fueron las mas propicias para haber pretendido la obtención de dichas proteínas, por lo que si se desea hacerlo, se debera replantear el diseño experimental, con nuevas estrategias.

VI. PERSPECTIVAS

La obtención de una amplia variedad de secuencias nucleotídicas diferentes, presumiblemente codificantes para diferentes proteínas expresadas en el hepatopáncreas del camarón blanco del Pacífico, nos proporciona herramientas para la caracterización de genes aún no descritos en este organismo, lo cuales pueden servir como modelo para ampliar el conocimiento que se tiene acerca de la expresión de genes, en diferentes estados fisiológicos del camarón.

Actualmente contamos con ARN total extraído a partir de camarones sometidos a diferentes períodos de ayuno, el cual puede ser, a corto plazo, una herramienta importante para estudios de caracterización de los genes completos a los que pertenecen las secuencias obtenidas en el presente trabajo, y determinar si tenemos genes constitutivos o inducibles e investigar aspectos cualitativos o cuantitativos de expresión.

Contamos también con mas de 150 clonas, obtenidas bajo las mismas condiciones ya descritas en ésta tesis, que aún no han sido secuenciadas para su posterior análisis, lo cual podría generarnos nueva información que nos permita incrementar el conocimiento que se está generando, mediante estudios basados en biología molecular, en apoyo a la información que se está aportando con otras disciplinas en la investigación.

BIBLIOGRAFÍA

Ahn J. H., Y. Choi, S. Kim, Y. M. Kwon, Y. D. Choi, y J. S. Lee. 1998. Expression of a Soybean Hydroxyproline-Rich Glycoprotein Gene Is Correlated with Maturation of Roots. *Plant Physiol*. February; 116 (2): 671–679

Alatorre M. R. 1998. La acuicultura en México: un sector de futuro. Edit. Panorama Acuicola, México, D.F.

Alfonso E., L. Ramos, E. Díaz-Iglesia, T. García y C. Rosas. 1993. Manual del II Curso Internacional de Producción de Postlarvas de Camarones Peneidos del Atlántico de América. (Ciudad de Campeche, México). Editorial Facultad de Ciencias UNAM: 1-133p.

Arce S., M. Moss y B. Argue. 1999. Artificial insemination and spawning of pacific white shrimp Litopenaeus vannamei: implications for a selective breeding program. In Spawning and maturation of aquaculture species. UJNR Aquaculture Techcnical report N.28: 5-8.

Ausubel F.M, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl. 1997. Current Protocols in Molecular Biology, John Wiley & Sons, NY. Capítulo 7. 1-62p.

Bateman A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall, E.L. Sonnhammer. 2002. The Pfam Protein Families Database. Nucleic Acids Research 30(1):276-280. Secuencias en línea: Protein families database of alignments and HMMs. Instituto Sanger. Disponible en: http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00050

Bell C. A., A. T. John, C. H. Kristen, A. N. Meyer, S. C. Robertson, y D. J. Donoghue. 2000. Rotational Coupling of the Transmembrane and Kinase Domains of the Neu Receptor Tyrosine Kinase. *Mol Biol Cell*; 11 (10): 3589–3599.

Benzie J.A.H. 1998. Peneid Genetics and Biotechnology. *Aquaculture*. 164: 23-47

Benoit S., H. Abaibou, y M. A. Mandrand-Berthelot. 1998. Topological Analysis of the Aerobic Membrane-Bound Formate Dehydrogenase of ***Escherichia coli***. *J Bacteriol*. December; 180 (24): 6625–6634

Berger C. 2000. Aportes de la Bio-Tecnología a la Alimentación y a la Inmuno Estimulación de Camarones peneidos. En: Cruz -Suárez, L.E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo, R. , (Eds.). Avances en Nutrición Acuícola. Memorias del V Simposium Internacional de Nutrición Acuícola. 19-22 Noviembre, 2000. Mérida, Yucatán.

Bones-Atle M. 2003. Degenerate PCR, a short guide. The Plant Group Home Page, Norwegian University of Science and Technology. En linea, disponible en: http://boneslab.chembio.ntnu.no/DegPCRshortguide.html

Brownstein M. J., J. M.Trent, M. S. Boguski. 1998. Functional Genomics. Trends Guide to bioinformatics. Elsevier Science. 27-29p. En línea, disponible en: http://www.cifn.unam.mx/moreno/Vallarta2002/TrendsGuide2BioInf1998.pdf

Bucheli P. (s/fecha). Diplomado: Consideraciones gerenciales para proyectos de acuacultura. Módulo: Consideraciones técnicas. Universidad Jefferson. Acuática SA. Disponible en: http://www.sica.gov.ec/agronegocios/Biblioteca/Ing%20Rizzo/perfiles_productos/camaron.pdf.

Bucher M., S. Brunner, P. Zimmermann, G. I. Zardi, N. Amrhein, L. Willmitzer, and J. W. Riesmeier. 2002. The Expression of an Extensin-Like Protein Correlates with Cellular Tip Growth in Tomato. *Plant Physiol*. 128: 911-923.

Buhler J. 2002. Anatomy of a Comparative Gene Expression Study. Información en línea, disponible en: http://www.cs.wustl.edu/~jbuhler/research/array/

Bull A., A.C. Ward y M. Goodfellow. 2000. Search and Discovery Strategies for Biotechnology: The Paradigm Shift, *Microbiology & Molecular Biology Reviews*, Vol 64, 573-606, ISSN: 1092-2172.

Carraway K. L., S. A. Price-Schiavi, X. Zhu, y M. Komatsu. 1999. Membrane Mucins and Breast Cancer. Conference 3th, Preview: Joint Cancer Conference 2000. Translational Research. Cancer control. Journal of the moffitt cancer control from the *Department of Cell Biology and Anatomy, University of Miami School of Medicine, Miami, Fla*.

Chan T.Y. y Chu, K.H. 1996. On the different forms of *Panulirus longipes femoristriga* (von Martens, 1872) (Crustacea, Decapoda, Palinuridae), with description of one new species. *Journal of Natural History* 30:367-387.

Ceccaldi H.J. 1997. Anatomy and physiology of the digestive system. In: D'Abramo, L.R., Conklin, D.E. y Akiyama, D.M. (eds.) *Crustacean Nutrition* Vol. VI: 261-291.

CICTUS. 1988. El cultivo del camarón azul (*Penaeus stylirostris* Stimpson). *Centro de Investigaciones Científicas y Tecnológicas de la Universidad de Sonora. México.* 45 p.

Cortés J.E. 1993. Evaluación del efecto de la calidad y cantidad de proteínas y raciones comerciales peletizadas en el requerimiento y la supervivencia del camarón blanco *Penaeus vannamei*. Tesis de Licenciatura. UABCS. 86p.

Dall W., J. Hill, P.C. Rothlisberg, y D.J. Staples. 1990. The Biology of Penaeidae. Advances in Marine Biology, Volume 27. Blaxter J.H.S. and Southward A.J. (Eds.). Academic Press, New York, USA. 489 p.

Dall W. 1992. Feeding, digestion and assimilation in Penaeidae. En: Allan, G. L. y Dall, W. (Eds.). Proceedings of the Aquaculture Nutrition Workshop (Salamander Bay, NSW: April 15-17, 1991), p.57-63.

Destoumieux D., M. Muñoz, P. Bulet, y E. Bachère. 2000. Penaeidins, a family of antimicrobial peptides from penaeid shrimp (Crustacea, Decapoda) .*Cellular and Molecular Life Sciences (CMLS)*. Issue: Volume 57, Numbers 8-9. August. Pages: 1260 – 1271.

Diederich R.J., K. Matsuno, H. Hing y S. Artavanis-Tsakonas. 1994. Cytosolic interaction between deltex and Notch ankyrin repeats implicates deltex in the Notch signaling pathway. *Development*, Vol 120, Issue 3 473-481.

Ding D., S.M. Parkhurst, S.R. Halsell y H.D. Lipshitz. 1993. Dynamic Hsp83 RNA localization during Drosophila oogenesis and embryogenesis. *Mol. Cell. Biol*., 3773-3781, Vol 13, No. 6

EBI SRS server. 2003. protein sequence database. Entry name : FA9_MOUSE. En línea: http://srs6.ebi.ac.uk/srs6bin/cgi-bin/wgetz?-newId+-e+[SWall-ACC:P16294]

EMBL Bioinformatic Harvester. 2003. Human protein: AL136581 - Hypothetical protein FLJ12799. -. Similar to PECANEX. En línea: http://www.embl-heidelberg.de/gfp-cdna/bioinformatics/AL136581.htm

Fernándes J., V. Brendel, X. Gai, S. Lal, V. L. Chandler, R. P. Elumalai, D. W. Galbraith, E. A. Pierson, y V. Walbot. 2002. Comparison of RNA Expression Profiles Based on Maize Expressed Sequence Tag Frequency Analysis and Micro-Array Hybridization. *Plant Physiol* 128, 896-910.

Fernández-Luna I., Oliva M., Carrillo O. y A. Van-Wormhoudt. 1997. Digestive enzyme activities of ***Penaeus notialis*** during reproduction and moulting cycle. Comp. Biochem. Physiol. 118A : 1267-1271

Fukuda M. 2002. Roles of mucin-type O-glycans in cell adhesion. Biochim Biophys Acta; 1573(3):394-405. Review.

Garcia-Carreño F.L., Navarrete del Toro M.A., Díaz-López M., Hernández-Cortés M.P. y Ezquerra J.M. 1995. Proteinase Inhibition of Fish Muscle Enzimes Using Legume Seed Extracts. *Journal of Food Protection*, Vol 59, No. 3:312-318.

Garcia-Carreño F.L. 1996. Proteinase inhibitors. ***Reprints from Elsevier Trends Journals*. 7:197-204**.

Grassot J., G. Perrière y G. Mouchiroud. 2002. RTKdb: database of receptor tyrosine kinase. En linea: http://www.irisa.fr/manifestations/2002/jobim/papiers/O-p199_012.pdf.

Gross P.S., T.C. Bartlett C.L. Browdy R.W. Chapman y G.W. Warr. 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White Shrimp, *Litopenaeus vannamei*, and the Atlantic White Shrimp, *L. setiferus*. *Dev Comp Immunol*. 25(7):565-77.

Guarna M.M. y Borowsky R.L. 1993. Genetically controlled food preference : Biochemical mechanisms. *Proceedings of the National Academy of Science, U.S.A.*,90:5257-5261

Hendrickx M. 1996. Los camarones peneoidea bentónicos del Pacífico Mexicano. Comisión Nacional para la Biodiversidad Universidad Nacional Autónoma de México, México. 35-38 p.

Honarvar S., B.K. Choi y D.M. Schifferli. 2003. Phase variation of the 987P-like CS18 fimbriae of human enterotoxigenic *Escherichia coli* is regulated by site-specific recombinases. *Mol Microbiol*; 48(1):157-71.

Hong D. H., B. S. Pawlyk, J. Shang, M. A. Sandberg, E. L. Berson, y T. Li. 2000. A retinitis pigmentosa GTPase regulator (RPGR)- deficient mouse model for X-linked retinitis pigmentosa (RP3) *PNAS*, March 28, vol. 97, no. 7, 3649-3654.

Honke K. 2003. Biological Function of Sulfoglycolipidse. En: Beyond Glycogenes, por Naoyuki Taniguchi. Disponible en:
http://www.glycoforum.gr.jp/science/glycogenes/glycogenesE.html

Hu S., M. Sonnenfeld, S. Stahl y S. T. Crews. 1998. Midline Fasciclin: A *Drosophila* Fasciclin-I-Related Membrane Protein Localized to the CNS Midline Cells and Trachea. John Wiley & Sons, Inc. *J Neurobiol* 35: 77–93.

Johansson, M.W., Keyser, P., y Söderhäll, K. 1994. Purification and cDNA cloning of a four-domain Kazal proteinase inhibitor from crayfish blood cells. Eur. J. Biochemistry. 223:389-394.

Joost P. y A. Methner. 2002. Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands. *Genome Biol*. 3 (11).

Kawashima T., S. Kawashima, Y. Kohara, M. Kanehisa y K. W. Makabe. 2002. MAGEST: Maboya Gene Expression patterns and Sequence Tags**.** *Nucleic Acids Research*, Vol. 30, No. 1 119-120.

Khayat M., **P. J. Babin**, **B. Funkenstein**, **M. Sammar**, **H. Nagasawa**, **A. Tietz** y **E. Lubzens. 2001.** Molecular Characterization and High Expression During Oocyte

Development of a Shrimp Ovarian Cortical Rod Protein Homologous to Insect Intestinal Peritrophins. *Biology of Reproduction* 64, 1090-1099.

Komatsu M., L. Yee, y K. L. Carraway. 1999. Overexpression of Sialomucin Complex, a Rat Homologue of MUC4, Inhibits Tumor Killing by Lymphokine-activated Killer Cells. *Cancer Res* 59: 2229-2236.

Kureshy N. y D. A. Davis. 2000. Metabolic requirement for protein by Pacific white shrimp *Litopenaeus vannamei*. Cruz -Suárez, L.E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo. R, (Eds.). Avances en Nutrición Acuícola V. Memorías del Quinto Simposium Internacional de Nutrición Acuícola. 19-22 de Noviembre de 2000. Mérida, Yucatán, México.

Lange B., M., M.R. Wildung, E. J. Stauber, C. Sanchez, D. Pouchnik, y R. Croteau. 2000. Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Plant Biology PNAS* 97: 2934-2939.

Laskowski Jr M. y M.A. Qasim. 2000. Review: What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? *Biochimica et Biophysica Acta* 1477, 324-337

Laskowski,M.J. y Kato,I.. 1980. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* 49, 593-626.

Lathe R. 1985. Synthetic oligonucleotide probes deduced from amino acid data: theoretical and practical considerations. *J. Mol. Biol.* 183:1-12.

Lavery S. and Chu, K.H. (no publicado) Phylogenetic relationships within the genus Penaeus. Harvard University's Biological Laboratories. Artículo en preparación. Publicación referida. Disponible en: http://www.oeb.harvard.edu/palumbi/people%20pages/Shane.html. Información a cerca del

artículo, en: Rosenberry B. 2003. Production by species. Free Reports on Shrimp Farming. Disponible en: http://members.aol.com/brosenberr/Species.html

Lawrence P. K., J. Nirmala y K. R. Koundal. 2001. Nucleotide sequence of a genomic clone encoding a cowpea (*Vigna unguiculata L.*) trypsin inhibitor. *Electronic Journal of Biotechnology* 4(1).

Lehnert S.A., K.J. Wilson, K. Byrne y S.S. Moore. 1999. Tissue-specific Expressed Sequence Tags from the Black Tiger Shrimp *Penaeus monodon*". *Marine Biotechnology 1*, 465-476.

Liang F., I. Holt, G. Pertea, S. Karamycheva, S. L. Salzberg, y J. Quackenbush. 2000. An optimized protocol for analysis of EST sequences. *Nucleic Acids Res*. 28 (18): 3657–3665.

Liang Z, y Söderhäll K. 1995. Isolation of cDNA encoding a novel serpin of crayfish hemocytes. *Comp. Biochem. Physiol*. Vol. 112B, No. 2:385-391.

Locus link. 2003. NCBI. CDC27: cell division cycle 27. En linea: http://www.ncbi.nlm.nih.gov/LocusLink/LocRpt.cgi?l=996

Lovett D. L. y D. L. Felder. 1990. Ontogenetic Changes in Enzyme Distribution and Midgut Function in Developmental Stages of *Penaeus setiferus* (Crustacea, Decapoda, Penaeidae). *The Biological Bulletin*, Vol 178, Issue 2 160-174.

Magallón F. 1999. Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo. Subprograma IIA; Proyecto II3. 52-62p.

Manheim E.A., K.J. Janet, D. Danielle, y K.S. McKim. 2002. Cytoplasmic Localization and Evolutionary Conservation of MEI-218, a Protein Required for Meiotic Crossing-over in *Drosophila*. *Mol Biol Cell*. 13 (1): 84–95

Martínez L. R. 1993. Camaronicultura: bases técnicas y científicas para el cultivo de camarones peneidos. Centro de Investigaciones Científicas y Tecnológicas de Sonora. AGT editor, SA, México

Martínez L.R. 1999. "CULTIVO DE CAMARONES PENEIDOS: PRINCIPIOS Y PRACTICAS". AGT Editor SA. 1ª Ed. Mexico

McCarter J. P., M. D. Mitreva, J. Martin, M. Dante, T. Wylie, U. Rao, D. Pape, Y. Bowers, B. Theising, C. V. Murphy, A. P. Kloek, B. J. Chiapelli, S. W. Clifton, D. Mck Bird y R. H. Waterston. 2003. Analysis and functional classification of transcripts from the nematode *Meloidogyne incognita. Genome Biology* 4(4):R26

McKim K.S. (s/fecha). Molecular Genetics of Meiotic Recombination. Institute of Microbiology-Waksman. Departament of Genetics. Disponible en: http://waksman.rutgers.edu/Waks/AnnualReport01/McKim.htm

Molina C., Cadena, E., y F. Orellana. 2000. Alimentación de camarones en relación a la actividad enzimática Como una respuesta natural al ritmo circadiano y ciclo de muda. En: Cruz-Suárez, L.E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo. R, (Eds.). Avances en Nutrición Acuícola V. Memorias del V Simposium Internacional de Nutrición Acuícola. 19-22 Noviembre, 2000. Mérida, Yucatán, Mexico.

Molina-Poveda C., V. Escobar, J. Gamboa-Delgado, E. Cadena, F. Orellana, y R. Piña. 2002. Estrategia de alimentación de acuerdo a la demanda fisiológica del juvenil *Litopenaeus vannamei* (Boone). In: Cruz-Suárez, L. E., Ricque-Marie, D., Tapia-Salazar, M., Gaxiola-Cortés, M. G., Simoes, N. (Eds.). Avances en Nutrición Acuícola VI. Memorias del VI Simposium Internacional de Nutrición Acuícola. 3 al 6 de Septiembre del 2002. Cancún, Quintana Roo, México

Nam D. K., S. Lee, G. Zhou, X. Cao, C. Wang, T. Clark, J. Chen, J. D. Rowley, y S. M. Wang. 2002. Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription**.** *PNAS* 2002 99: 6152-6156.

Nguyen G., F. Delarue, C. Burcklé, L. Bouzhir, T. Giller y J. D. Sraer. 2002. Pivotal role of the renin/prorenin receptor in angiotensin II production and cellular responses to rennin. *J Clin Invest*, June, Volume 109, Number 11, 1417-1427.

Pérez-Farfante I. y B. Kensley. 1997. Penaeoid and sergestoid shrimps and prawns of the world: keys and diagnoses for the families and genera. *Mémoires du Muséum national d'Histoire Naturelle,* 175: 1-233.

Reymond P., H. Weber, M. Damond y E.E. Farmer. 2000. Differential gene expression in response to mechanical wounding and insect feeding in Arabidopsis. *Plant Cell* 12: 707-719.

Roepman R., N. Bernoud-Hubac, D. E. Schick, A. Maugeri, W. Berger, H.H. Ropers, F. P.M. Cremers y P. A. Ferreira. 2000. The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Human Molecular Genetics*, Vol. 9, No. 14 2095-2105.

Rose T.M., E.R. Schultz, J.G. Henikoff, S. Pietrokovski, C.M. McCallum y S. Henikoff. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res*. 26, 1628-1635.

Rosenberry B. 2003. A Brief History of Shrimp Farming. (Updated November 18, 2002). Free Reports on Shrimp Farming. Disponible en: http://www.shrimpnews.com/History.html

Rosenberry B. **2003**. Production by species. Free Reports on Shrimp Farming. Disponible en: http://members.aol.com/brosenberr/Species.html

Sambrook J., E.F. Fritsch, y T. Maniatis. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sambrook J. y D.W. Russell. 2001. Molecular Cloning: *A Laboratory Manual,* 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Scharfe C., P. Zaccaria, K. Hoertnagel, M. Jaksch, T. Klopstock, R. Lill, H. Prokisch, K. D. Gerbitz, H. W. Mewes, y T. Meitinger. 1999. MITOP: database for mitochondria-related proteins, genes and diseases. *Nucleic Acids Res*. 27 (1): 153–155.

Shen Z., G. Dimopoulos, F. C. Kafatos, y M. Jacobs-Lorena. 1999. A cell surface mucin specifically expressed in the midgut of the malaria mosquito *Anopheles gambiae* Proc. Natl. Acad. Sci. USA. May 11; 96 (10): 5610–5615.

Shigueno K. 1992. Shrimp culture industry in Japan. 641-652p. In: A.W. Fast and L.J. Lester (Editors), Marine Shrimp Culture: Principles and Practices. *Elsevier, Amsterdam*.

Sillence D. J. 2001. Apoptosis and signalling in acid sphingomyelinase deficient cells. *BMC Cell Biol*. ; 2 (1): 24.

Skrabanek L. y F. Campagne. 2001. TissueInfo: high-throughput identification of tissue expression profiles and specificity. *Nucleic Acids Res*. 29: e102.

Söderhäll K., L. Cerenius, y M.W. Johansson. 1996. The prophenoloxidase activating system in invertebrates. In: *New Directions in Invertebrate Immunology.*(Söderhäll, K. Iwanaga, S. & Vasta, G.R., eds), 229–253 p. SOS Publications. Fair Haven.

Talavera V., D. Sánchez y L. M. Zapata. 1998. Descripción y funciones de la glándula media o hepatopáncreas. BOLETÍN NICOVITA. Camarón de Mar. Edición Tumpis. Vol 3, Ejemplar 10, Lima, Perú. En-línea, disponible en: http://www.nicovita.com.pe/pdf/esp/boletines/bole_9810_02.pdf

Timothy M.R., E.R. Schultz, J.G. Henikoff, S. Pietrokovski, C.M. McCallum y S. Henikoff. 1998. Consensus degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Reseca, Vol. 26, No. 7. 1628-1635.

Van den Berg A., J. Van der Leij y S. Poppema. 1999. Serial analysis of gene expression: rapid RT-PCR analysis of unknown SAGE tags. *Nucleic Acids Research*. 27(17):e17.

Van Herp F. 1992. Inhibition and stimulating neuropeptides controlling reproduction in Crustacea. *Invertebrate Reproduction and Development* 22: 21-30.

Vega-Villasante F., H. Nolasco-Soria, R. Civera-Cerecedo, R. González-Valdés, y M. Oliva-Suárez. 2000. Alternativa para la alimentación del Camarón en Cultivo: El Manejo de la Muda. In: Cruz -Suárez, L.E., Ricque-Marie, D., Tapia-Salazar, M., Olvera-Novoa, M.A. y Civera-Cerecedo, R. (Eds.). Avances en Nutrición Acuícola V. Memorias del V Simposium Internacional de Nutrición Acuícola. 19-22 Noviembre, 2000. Mérida, Yucatán.

Verdun R. E., N. Di Paolo, T. P. Urmenyi, E. Rondinelli, A. C. C. Frasch, y D. O. Sanchez. 1998. Gene Discovery through Expressed Sequence Tag Sequencing in ***Trypanosoma cruzi.*** *Infect. Immun*. 66:5393-5398.

Villarreal H. y P. M. Castro. 1993. Preliminary studies on the effect of protein content on the growth of *Penaeus vannamei* at marine salinities. *Revista de Publicaciones Biológicas de la Univesrsidad Autónoma de Nuevo León.*

Vogel B. E. y E. M. Hedgecock. 2001. Hemicentin, a conserved extracellular member of the immunoglobulin superfamily, organizes epithelial and other cell attachments into oriented lineshaped junctions. *Development* 128, 883-894.

Walker J.M. y E. B. Gingold. 1997. Biología molecular y biotecnología. 2ª ed. Zaragoza: Acribia, DL 475 p.

Wang P. y R. R. Granados. 1997. An intestinal mucin is the target substrate for a baculovirus enhancin. Proceedings of the National Academy of Sciences of USA 94: 6977-6982.

White J. A., J. Todd, T. Newman, N. Focks, T. Girke, O. Martínez de Ilárduya, J. G. Jaworski, J. B. Ohlrogge, y C. Benning. 2000. A New Set of Arabidopsis Expressed Sequence Tags from Developing Seeds. The Metabolic Pathway from Carbohydrates to Seed Oil. *Plant Physiol*.; 124: 1582-1594 .

Wolfsberg T.G. 1997. Landsman D. A comparison of expressed sequence tags (ESTs) to human genomic sequences. *Nucleic Acids Research*, 25:1626-1632.

Yuen Peter S. T., K. M. Brooks y Y. Li. **2001**. RNA: a method to specifically inhibit PCR amplification of known members of a multigene family by degenerate primers. *Nucleic Acids Research*, Vol. 29, No. 6 e31.

Apéndice 1: Ejemplo de un alineamiento de secuencias de aminoácidos que codifican para inhibidores de proteinasas tipo Kazal en vertebrados. El oligonucleótido consenso se diseñó a partir de la secuencia de una región conservada (señalada en azul).

```
>SPRC_BOVIN/95-149 CQDPTS...........CPAPIGEFEKVCSNDNKTFDSSCHFFATKCTLEGTKKGHKLHLDYIGPC
>SPL1_RAT/426-479  CQDPET...........CPP.AKILDQACGTDNQTYASSCHLFATKCMLEGTKKGHQLQLDYFGAC
>SPL1_COTJA/468-521CQDPAA...........CPS.TKDYKRVCGTDNKTYDGTCQLFGTKCQLEGTKMGRQLHLDYMGAC
>ISK4_PIG/37-86    CEHMTESPD........CS...RIYDPVCGTDGVTYESECKL....CLARI.ENKQDIQIVKDGEC
>AGRI_RAT/385-429  CQRV.............CS...GIYDPVCGSDGVTYGSVCELESMAC.....TLGREIQVARRGPC
>AGRI_CHICK/374-418CQQV.............CQ...GRYDPVCGSDNRTYGNPCELNAMAC..... LKREIRVKHKGPC
>IPK1_CANFA/10-57  CNLKVNG..........CN...KIYNPICGSDGITYANECLL....CLENK.KRQTSILVEKSGPC
>IPK1_BOVIN/9-56   CTNEVNG..........CP...RIYNPVCGTDGVTYSNECLL....CMENK.ERQTPVLIQKSGPC
>IPK2_RAT/32-79    CPNTLVG..........CP...RDYDPVCGTDGKTYANECIL....CFENR.KFGTSIRIQRRGLC
>IPK1_RAT/32-79    CPKQIMG..........CP...RIYDPVCGTNGITYPSECSL....CFENR.KFGTSIHIQRRGTC
>AGRI_RAT/515-559  CPR..............CE..HPPPGPVCGSDGVTYLSACELREAAC.....QQQVQIEEAHAGPC
>AGRI_CHICK/504-548CPR..............CEQ..QPLAQVCGTDGLTYDNRCELRAASC.....QQQKSIEVAKMGPC
>AGRI_RAT/450-494  CPSE.............CV...ESAQPVCGSDGHTYASECELHVHAC.....THQISLYVASAGHC
>AGRI_CHICK/439-483CPTE.............CV...PSSQPVCGTDGNTYGSECELHVRAC.....TQQKNILVAAQGDC
>AGRI_RAT/818-864  CPTLT............CP..EANSTKVCGSDGVTYGNECQLKAIAC.....RQRLDISTQSLGPC
>AGRI_CHICK/805-851CPSPL............CS..EANMTKVCGSDGVTYGDQCQLKTIAC.....RQGQLITVKHVGQC
>IAC2_HUMAN/36-84  CSQYRLPG.........CP...RHFNPVCGSDMSTYANECTL....CMKIR.EGGHNIKIIRNGPC
>AGRI_CHICK/81-126 CKKTA............CP...VVVAPVCGSDYSTYSNECELEKAQC.....NQQRRIKVISKGPC
>FSA_HUMAN/118-164 CAPD.............CS.NITWKGPVCGLDGKTYRNECALLKARC.....KEQPELEVQYQGRC
>AGRI_RAT/600-645  CDFS.............CQ..SVPRSPVCGSDGVTYGTECDLKKARC.....ESQQELYVAAQGAC
>AGRI_CHICK/588-633CDFT.............CL..AVPRSPVCGSDDVTYANECELKKTRC.....EKRQNLYVTSQGAC
>IOVO_COTJA/70-127 CSRYPNTTSEDGKVTILCT...KDFSFVCGTDGVTYDNECML....CAHNV.VQGTSVGKKHDGEC
>IOVO_CHICK/94-151 CSSYANTTSEDGKVMVLCN...RAFNPVCGTDGVTYDNECLL....CAHKV.EQGASVDKRHDGGC
>IOVO_COTJA/5-62   CSRFPNTTNEEGKDEVVCP...DELRLICGTDGVTYNHECML....CFYNK.EYGTNISKEQDGEC
>IOVO_CHICK/29-86  CSRFPNATDKEGKDVLVCN...KDLRPICGTDGVTYTNDCLL....CAYSI.EFGTNISKEHDGEC
```

```
>IOV7_CHICK/225-282CDQYPTRKTTGGKLLVRCP...RILLPVCGTDGFTYDNECGI....CAHNA.QHGTEVKKSHDGRC
>IOV7_CHICK/159-216CSKYPSTVSKDGRTLVACP...RILSPVCGTDGFTYDNECGI....CAHNA.EQRTHVSKKHDGKC
>IOV7_CHICK/94-150 CSPYLQV.VRDGNTMVACP...RILKPVCGSDSFTYDNECGI....CAYNA.EHHTNISKLHDGEC
>IOV7_CHICK/28-85  CSLYASGIGKDGTSWVACP...RNLKPVCGTDGSTYSNECGI....CLYNR.EHGANVEKEYDGEC
>IOV7_CHICK/356-413CSKYKTSTLKDGRQVVACT...MIYDPVCATNGVTYASECTL....CAHNL.EQRTNLGKRKNGRC
>AGRI_RAT/167-212  CPTT.............CF..GAPDGTVCGSDGVDYPSECQLLSHAC.....ASQEHIFKKFNGPC
>AGRI_CHICK/156-201CPAS.............CS..GVAESIVCGSDGKDYRSECDLNKHAC.....DKQENVFKKFDGAC
>IPK1_ANGAN/12-61  CGEMSAMHA........CP...MNFAPVCGTDGNTYPNECSL....CFQRQ.NTKTDILITKDDRC
>FSA_HUMAN/192-239 CNRI.............CPEPASSEQYLCGNDGVTYSSACHLRKATC.....LLGRSIGLAYEGKC
>FSA_HUMAN/270-316 CDEL.............CPD.SKSDEPVCASDNATYASECAMKEAAC.....SSGVLLEVKHSGSC
>AGRI_RAT/311-356  CDRVT............CD...GSYRPVCAQDGHTYNNDCWRQQAEC.....RQQRAIPPKHQGPC
>AGRI_CHICK/299-344CDRIT............CD...GTYRPVCARDSRTYSNDCERQKAEC.....HQKAAIPVKHSGPC
>IPSG_MELME/1668   CSKYNAKGS.....QFACS...RHLDPVCGTDHRTYSNECMF....CMLTQ.NKRFSVRILQDNNC
>IPSG_CANFA/12-64  CSNYKGKGS.....QIACP...RLHQPICGTDHKTYSNECMF....CALTL.NKKFEVRKLQDTAC
>IPSG_FELCA/8-60   CSQYNRKGS.....GITCS...KEWKPICGIDHKTYSNECMF....CQLNQ.NKRFQLRKLHDNKC
>IPSG_FELCA/64-111 CTNYSA..........ICT...MEYFPLCGSDGQEYSNKCLF....CNEVV.KRRGTLFLAKYGQC
>IPSG_CANFA/68-115 CTEYSDM..........CT...MDYRPLCGSDGKNYSNKCSF....CNAVK.KSRGTIFLAKHGEC
>IOVO_FRACO/5-53   CSEYPKPG.........CT...MEYRPVCGSDNITYGNKCNF....CNAVV.KSNGTLTLSHFGKC
>IOVO_ABUPI/8-56   CSDHPKP.........ACL...QEQKPLCGSDNKTYDNKCSF....CNAVV.DSNGTLTLSHFGKC
>IOVO_CASCA/6-54   CSEYPKP.........VCS...PEYMPLCGSDSKTYNNKCDF....CSAVV.ESNGTLTLGHFGKC
>IOVO_EUDEL/5-51   CSGYPKP.........ACT...LEFFPLCGSDNQTYSNKCAF....C.NA. AVEKNVTLNHIGEC
>IOV7_CHICK/422-472CREFQKV.......SPICT...MEYVPHCGSDGVTYSNRCFF....CNAYV.QSNRTLNLVSMAAC
>IAC2_BOVIN/7-57   CAEFKDP.......KVYCT...RESNPHCGSNGETYGNKCAF....CKAVM.KSGGKINLKHRGKC
>IACA_PIG/7-57     CNVYRSH.......LFFCT...RQMDPICGTNGKSYANPCIF....CSEKG.LRNQKFDFGHWGHC
>IAC1_BOVIN/14-61  CKVYTEA..........CT...REYNPICDSAAKTYSNECTF....CNEKM.NNDADIHFNHFGEC
>AGRI_RAT/237-284  MLLRPEN..........CP...AQHTPICGDDGVTYENDCVMSRIGA.....TRGLLLQKVRSGQC
>AGRI_CHICK/227-273 LSRPEN.........CP...SKREPVCGDDGVTYASECVMGRTGA.....IRGLEIQKVRSGQC
OLIGONUCLEÓTIDO CONSENSO                   YDPVCGSD
```

Bateman, 2002.

Apéndice 2: Secuencias nucleotídicas obtenidas a partir de cebadores diseñados para inhibidores de proteinasas.

KA18(570)

CCTTCGATCCGGTGTGCGGCAGGAGGGATCAACCTACAGCCTGAAGGAAGACACAACTCTCGATGATGTTTTCCAGCGCATGAAGGC
TGTCATAGGATACAGAGCTCAGAGGGAAACACTCTTCAAACGCATCACTGTGGATGACAATTTGGATGAAGACAAATACCATGGGTC
CACCCTTTCATCCGAGGTACTTAGTCCCAATGAGGAGCCTGATTCAACCTTCTTGCGAGAGATGGATGCTCTTGTGGATGCTGCAGA
GAAGGTTGGGGATATAACAACCACAGCCCATGGTGGCCAGGACATCATCTTCCTGCAGGTCAATTCTCTGGCACCACTAGTCAAAAC
ATATGGGCCAGAATCAGCAAAGGTTAAGGAAGCAGAAAAGATCCTGAGAAAACAGCTTGTCCAAGTCACAGACATGATGCGCCACAT
TTATAATGACAGAGTCTTGGTTGCTACGGCTACAGTGGAGCAGCTAGAGGAACTCTCACGATCTTCACGTTCCATTCTCCAGAGTGA
GGATGTGAACTTGGAAGACTACAACCTTGCCACCCGCACACCGGATCGTAAAGG

KA14(560)

GGTGTGCGGCAGGAGGGATCAACCTACAGCCTGAAGGAAGACACAACTCTCGATGATGTTTTCCAGCGCATGAAGGCTGTCATAGGA
TACAGAGCTCAGAGGGAAACACTCTTCAAACGCATCACTGTGGATGACAATTTGGATGAAGACAAATACCATGGGTCCACCCTTTCA
TCCGAGGTACTTAGTCCCAATGAGGAGCCTGATTCAACCTTCTTGCGAGAGATGGATGCTCTTGTGGATGCTGCAGAGAAGGTTGGG
GATATAACAACCACAGCCCATGGTGGCCAGGACATCATCTTCCTGCAGGTCAATTCTCTGGCACCACTAGTCAAAACATATGGGCCA
GAATCAGCAAAGGTTAAGGAAGCAGAAAAGATCCTGAGAAAACAGCTTGTCCAAGTCACAGACATGATGCGCCACATTTATAATGAC
AGAGTCTTGGTTGCTACGGCTACAGTGGAGCAGCTAGAGGAACTCTCACGATCTTCACGTTCCATTCTCCAGAGTGAGGATGTGAAC
TTGGAAGACTACAACCTTGCCACCCGCACACCGGATCGTAAAGG

K52(145)

CCTTTACGATCCTGTGTGCGGGTGGCGGGGTGTTGAGGCAAGAACCCGAAGGGCGCTACATCAGGGGTGACGGCGGGTCAGACAGCG
AGGCGGGGCGAGGACGGGACTGCAGCGGAGTGGGCAGCCGCAGACCGGATCGTAAAGG

KA36(145)

CCTTTACGATCCTGTGTGCGGGTGGCGGGGTGTTGAGGCAAGAACCCGAAGGGCGCTACATCAGGGGTGACGGCGGGTCAGACAGCG
AGGCGGGGCGAGGACGGGACTGCAGCGGAGTGGGCAGCCGCAGACCGGATCGTAAAGG

KA8(90)

CCTTTACGATCCGGTCTGCGGGTGGACACTTTTGTCGAAAGTTACCGGCCACGAATTAAGTATCAGCTTGTACCCGCACACAGGATC
GTAAAGG

KA10(90)
CCTTTACGATCCGGTCTGCGGGTGGACACTTTTGTCGAAAGTTACCGGCCACGAATTAAGTATCAGCTTGTACCCGCACACAGGATC
GTAAAGG
K33(120)
CCTTTACGATCCGGTGTGCGGCTGGTGACGTCACACACGGCACGGCTGGGCACTCCCCCGGCACGATGGCCGCTGCATCGCCACACC
GCCGCCATCCACCCGCAGACCGGATCGTAAAGGG
KA1(90)
CCTTTACGATCCGGTCTGCGGGTGGAGGGGTATCGGTATTAATTTGTGTTGTGCTTTTTGGGGTGTACCAGCCGCACACCGGATCGT
AAAGG
KA12(140)
CCTTTACGATCCGGTCTGCGGGTGGAGGGGTATCGGTATTAATTTGTGTTGTGCTTTTTGGGGTGTACCAGCCGCACACCGGATCGT
AAGCCGCAGACCTCCCGCACACCGGATCGTACCCCCTCCCGCACACTGGATTGTAAAGG
KA19(120)
CCTTTACGATCCGGTGTGCGGGTAGCGTACCGGGGGAGTGCCCAGCCGTGCCGTGTGTGACGTCACCAGCCGCAGACCGGATCGTAA
GCCGCAGACCACCCGCAGACCGGATCGTAAAGG
KA6(110)
CCTTTACGATCCAGTGTGCGGCACGATGGGCGCTGCATTACGATCCGGTCTGCGGCTGGCAAGTGGCAGGTGAGGCTCGCGGTGCGC
TACCCGCACACCGGATCGTAAAGG
KAZ2(80)
CCTTCGATCCGGTCTGCGGGTGGAGGGAGAGGGAGAGCTGTTATATTGTCTCTTTGCCAGCCGCACACCGGATCGTAAAGG
KAZ1(90)
GCAGAATTCGCCCTTCGATCCGGTCTGCGGGTGGAGGGAGAGGGAGAGCTGTTATATTGTCTCTTTGCCAGCCGCACACCGGATCGT
AAAGG
KA32(130)
CCTTTACGATCCGGTGTGCGGGTCGACCACGGACACCTTCGAGTCGAGCGTGGCGGCGTTGAACCAGATCGTGTTCTTCTCCACGCC
GTATTTCATGCAGACCGAGCACCCGCAGACCGGATCGTAAAGGG
KA13(265)
CCTTTACGATCCGGTGTGCGGCAGTCAGGACCCATTCCGTGCTGACGAGGGTGCCTCCGCATAAGGGCCAACTCTCCTGGAAGTTGG
TGGCTTCGATGCTCACCAACCACGGGATTTCTCCTGGAAATGCAACGTGACCTCCGGCAATCTTGGTGTCTTGTCTTGCACGCTCAG
CAAGAACGCAGCCGAAGAGACACAGGAACGCGATTGTGAACTTCATATTACTTTCCTTACTGAATGTGTGTCCCGCAGACCGGATCG
TAAAGG

KA2(125)
CCTTTACGATCCGGTGTGCGGGACCCCACCGACTGCACCAGGTACTACGTCTGCGTCGACACCTCCGGCACTGGCGTGTTATTGCCA
TCGATCGACCCCGTGGAGTGCCGCAGACCGGATCGTAAAGG
KA3(90)
CCTTTACGATCCGGTCTGCGGCTGGGGGATTTAACATGCGACTCGTCACGGAGTTTCATGCGGCGCCTGTACCCGCACACAGGATCG
TAAAGG
KA9(200)
CCTTTACGATCCTGTGTGCGGCAGGCGATCGAGGAGCTGGACAGCGTGTTCGACCTGGTGATGGTGGCGGAGCGGATGGACGAGTCG
CTGGTCCTGCTGCGCCACCTGCTGTGCTGGAGCCTCCACGACGTGGTGGTGTTCACCAAGAACGCGCGCCGCCAGGAGGTCAAGCCC
ACGCTCGACCCGCACACCGGATCGTAAAGG
KA27(145)
CCTTTACGATCCGGTGTGCGGGAGGAACGGGCAGTTGTTGGGGGTGACGCAGCGCATGGTCTCCGGGTTGAGCACCATGCCCGAGGG
GCACCGGAAAGGGTCGACGGTGTGGCCGATATCGCTGCTGCCGCACACAGGATCGTAAAGG
KA26(200)
CCTTTACGATCCTGTGTGCGGCAGGTGTTCAGTGGAGAGGCAGTTGTAGAGAGTCTTAAAATAAGCCTTCAGTTTGCTGCCGAGCCT
CTCCTCCTGCCAGTCCCTGAAGTTGTGCTCACACATCCAGTACAAGATGAACCGGAGATGCTGCGTCTTGAGGCCGGGGATCTCCTG
CAGGTGGTCCCGCACACCGGATCGTAAAGG
K41(100)
CCTTTACGATCCGGTCTGCGGGAGGGGGAGGTACGATCCTGTGTGCGGGTAGCGCACCGCGAGCCTCACCTGCCACTTGCCAGCCGC
AGACAGGATCGTAAAGG
KA11(180)
CCTTTACGATCCAGTGTGCGGGACGCGATCACGTCGCCTAACGCTCAGCGCGGTTTCCCCCTTGAGAGGCTTTCAGGGGTAAAACCT
GCCGGACGTGGCTGCCCTCCCCTGGTGGCCTGTCGGGCAGGAGGTGCACCATTGGTCCGGACCCACCGTTCCTCCCGCAGACCGGAT
CGTAAAGG
KA20(180)
CCTTTACGATCCTGTGTGCGGGAGGAACGGTGGGTCCGGACCAATGGTGCACCTCCTGCCCGACAGGCCACCGGGGGAGGGCAGCCA
CGTCCGGCAGGTTTTACCCCTGAAAGCCTCTCAAGGGGGAAACCGCGCTGAGCGTTAAGCGACGTGATCGCGTGCCGCAGACTGGAT
CGTAAAGG
KA38(200)

CCTTTACGATCCGGTCTGCGGCTCGATGACCTCCTGCAGCCGTGTCACCACTTTGAGGTCAACCTCAGGTGCCGAACCGTTGGAGTG
GATGACTGTTCCCTCGGACTTGGCGAGCGACGTGGAGAGTTTCCTGTTCTCGTCCTGGAGGCGGGACACCATCGCCACCAGCTCATG
GCTCTCCTCCCGCACACAGGATCGTAAAGG

KA25(220)

CCTTTACGATCCTGTGTGCGGCAGGACAAGTGCTTCATCTACAAGGATCCTTATGGCGTCGTCCTCGTCATGGGGGCGTGGAACTAC
CCGCCGCAGCTCAGTCTCATTCCTGCGGCAGGTGCCATCGCTGCAGGTAACACAGTGGTAATTAAGCCTTCCGAGGTCGCCCCCGCC
ACCGCCGCCGCCATCGCCAAGCTCGTGCCGCAGACCGGATCGTAAAGG

KA22(270)

CCTTTACGATCCTGTGTGCGGGAGGGGGAGGTACGATCCTGTGTGCGGCAGCAGGTCCTTCAGCCAGTCCACCTGCGGCGGGGGCTT
GGACTTGGGGTTGCAGCGGATCTTGGACTCCTCGTTGATGATGGGCGTTTGGACGACGGGCGTGTCGCCGAAGTCGATGTCCACGAA
GAAGTCCACCATGAGCTTCACCGACAGCTTCTGGTTGCCGGCGTAGACAGCGGTGCACTCGTACTCTCCCGAGTCCTCCTCCCGCAG
ACAGGATCGTAAAGG

KA16(245)

CCTTTACGATCCGGTCTGCGGCTGCGTGGGCCGGAAGCCCGGGTTGAGGTACGAGAAGGGGCGAAGCTCCGTCGGCACCTTCGGCAG
CTCTCTCGGCGCGGGCGCAGGCTGGTTGGGATACACCGGCGGATGGTCAGGGTCAAAGGCCGGGTTGACATGGCCATCGGCGACGTA
TGGCGGCGCATTTCCATGGTACTGGGTCACTCCTTGTTTAAGGACGTTGTTCCCGCAGACCGGATCGTAAAGG

KA37(85)

CCTTTACGATCCTGTGTGCGGCTGGCGNTCAACGACAGACACAGTCGCGNCCACTACCCCCCCCCACCCGCAACACCGGATCGAAGG

KA28(85)

CCCTTTACGATCCTGTCTGCGGGTGGAGGCCAGCACTTCACACACTTTCAGGAAGTCGGCTTTTGGCGTGCCGCACACCGGATCAAG
G

KA5(115)

CCTTTACGATCCTGTCTGCGGCAGCACCTACCCATCCTCACCACTCTCCCTCTCTACCGCAACCCTGCCCAGGTTTTCTCCCAGCGG
CCTGATGCACCCGCACACAGGATCGAAGG

KA21(90)

GCAGAATTCGCCCTTCGATCCGGTCTGCGGGTGGAGGGAGAGGGAGAGCTGTTATATTGTCTCTTTGCCAGCCGCACACCGGATCGT
AAAGG

KA15(180)

CCTTTACGATCCTGTCTGCGGGACGGACAAATGTAGGATCCTCACCACGCTATCTGTTGGCCCAGTCGGCTCCGGGTGCGAGGGCGC
GGGTTTGTCTCCGGGGATGGACACCGTCCAGGGGCTCAGCTGGTCGTGCTCGCAGTGGCCGCCCAGTAACCAGCCGCACACTGGATC
GTAAAGG

A3(210)
CCTTTGGACTAAGTGCGGCGGGTCGCTAGTCAGCATCGTTTAAGGTTAGAACTAGGGCGGTATCTGATCGCCTTCGAACCTCTAACT
TTCGTTCTTGATCAATGAAAACATCCTTGGCAGATGCTTTCGCAGTAGGTCGTCATGCGACGGTCTAAGAATTTCACCTCTAGCGTC
GCAGTACGAATGCCCCCGCACTTAGTCCAAAGG
ANEK6(200)
CCTTTGGACTAAGTGCGGCGGGTCGCTAGTCAGCATCGTTTAAGGTTAGAACTAGGGCGGTATCTGATCGCCTTCGAACCTCTAACT
TTCGTTCTTGATCAATGAAAACATCCTTGGCAGATGCTTTCGCAGTAGGTCGTCATGCGACGGTCTAAGAATTTCACCTCTAGCGTC
GCAGTACGAATGCCCCCGCACTTAGTCCAAAGG
ANEK5(55)
CCTTTGGACTAAGTGCGGCGGCAGCAGGAACCCCCCCCCCCGCACTTTGTCCAAAGGG

INH2(95)
CCTTTGCCTTGAACCGAAGGTGGTCGGGCCTTGCTGCAATCGTTTAAGACTGCCTTGCTTTGATCCTCCTAGCCTAAATGTAATTGG
TATAGAAC
INH8(300)
CCTTTGCCTTGAACCGAAGGTGGTCGGGCCTTGCTGCAATCGTTTAAGACTGCCTTGCTTTGATCCTCCTAGCCTAAATGTAATTGG
TATAGAACGAGAATTCTTACTCATATCTTTACTATTATCATCTTCCAATGCCTCTAGCTCACAATTCTTTTCCACCCACTGTGTTGA
AACCAATCCAGTGGTGTCATCAGATTCATTCTTAGAACTTTTTAATGTACTGGCTTCCTCAATTGCCTTTTTCTGACTATATTCATT
ATCTAGGTTTAATGCGGAGCTTGGCAAGGCCCCACGACCTTAGGTTCAAGGCAAAGG
INH1(55)
CCCTTTGCCTTGAACCGAAGGTCGTNGGGCCCTGCCACCTTAGGTTCAAGGCAAAGG

INSK1(220)
CCTTTACAACGCTGACACGAAGAAGTGCTTCACGGCCAGGTGGTCCTCCCAGTCGTTGGTCAACGACTTGTAGAACTCGCCGTACTC
CTCCTTCGAGATGTCGTCGGGGTTGCGCGTCCACAGGGGCTTCGTCTTGTTCAGCTCCTCGTCCTCCGTGTACTTCTCCTTCACCGT
CTTCTTTTTCTTCTTGTCATCGCACTTCTTCGTGTCGGCGTTGTAAAGGG
INSK4(220)

CCTTTACAACGCTGACACGAAGAAGTGCTTCACGGCCAGGTGGTCCTCCCAGTCGTTGGTCAACGACTTGTAGAACTCGCCGTACTC
CTCCTTCGAGATGTCGTCGGGGTTGCGCGTCCACAGGGGCTTCGTCTTGTTCAGCTCCTCGTCCTCCGTGTACTTCTCCTTCACCGT
CTTCTTTTTCTTCTTGTCATCGCACTTCTTCGTGTCGGCGTTGTAAAGG

INSK3(190)

CCTTTACAACGCTGACACGAAGAAGTGCTTCACGGCCAGGTGGTCCTCCCAGTCGTTGGTCAACGACTTGTAGAACTCGCCGTACTC
CTCCTTCGAGATGTCGTCGGGGTTGCGCGTCCACAGGGGCTTCGTCTTGTTCAGCTCCTCGTCCTCCGTGTACTTCTCCTTCACCGT
CTTCTTTTTCTTCTTGTCATCGCA

IS1(125)

CCTTTACAACGCCGACACGAAGAAGTGCTCCAGAAGGGGTCCCGCCCCCGAGCATGGTCATGCAGGTGGCCATATCCAGCACGGTCC
TCACAGGCTTGCACTTCTTCGTGTCAGCGTTGTAAAGG

INSK2(140)

CCCTTTACAACGCTGACACGAAGAAGTGCAAGTCTCGCAGTCGTTCAAAATCTGCTGAGGCAGAATCAGAGTCAGAGGATGAGAAGT
AGGAGCCAGAGTAGAGGCCCCTGCTGGTCATGGAGCACTTCTTAGTGTCGGCGTTGTAAAGG

INVK4(90)

CCTTCGATTGCGAGAGCGGCAAGTGCTGTCTTAGATCCCGGTCACTTGGGCGCTCTTGTCTCACGCACTTGCCGGACTCGCTATCGT
AAGG

P3(280)

CCTTCTACAACCGTAAGACTAAGTCTGCCCCTTTACTCTTGCAGTCGTCTCTCGCTTCGCTGAACGTGAGAGTGTAGTACGAATAGT
CAGTGTTGGAATACCAGTAGCAATAGTCGCCGAAGGTCTGCCAGCCTTCTTCACAACCGACCGTGTGCGTAGGGAAGAACTCGCAGC
CAAAGGCCCTGCTGGAGGAACACGGCATATTGTTCCACTTGGCGTGCGTGTCCCAGCCCACCGCGTCGATGATCTCCGTGCAGACTT
AGTCTTACGGTTGTAGAAGG

PANK1(280)

CCCTTCTACAACCGTAAGACTAAGTCTGCACGGAGATCATCGACGCGGTGGGCTGGGACACGCACGCCAAGTGGAACAATATGCCGT
GTTCCTCCAGCAGGGCCTTTGGCTGCGAGCTCTTCCCTACGCACACGGTCGGTTGTGAAGAAGGCTGGCAGACCTTCGGCGACTATT

GCTACTGGTATTCCAACACCGACTATTCGTACTACACTCTCACGTTCAGCGAAGCGAGAGACGACTGCAAGAGTAAAGGGGCAGACT
TAGTCTTACGGTTGTAGAAGG

PANK5(280)

CCTTCTACAACCGTAAGACTAAGTCTGCACGGAGATCATCGACGCGGTGGGCTGGGACACGCATGCCAAGTGGAACAATATGCCGTG
TTCCTCCAGCAGGGCCTTTGGCTGCGAGTTCTTCCCTACGCACACGGTCGGTTGTGAAGAAGGCTGGCAGACCTTCGGCGACTATTG
CTACTGGTATTCCAACACTGACTATTCGTACTACACTCTAACGTTCGGCGAAGCGAGAGACGACTGCAAGAGTAAAGGGGCAGACTT
AGTCTTACGGTTGTAGAAGG

PANK2(320)

CCTTCTACAACCGTAAGACTAAGTCTGCTTCGTCCTAAGTTATCCCTGACAAGCCAGAAGTCAGGCAAACTAAAGGACTCCAAAGGC
GTGCAACACAAAGGCTACTCAAAAGACCTGATACAGAACATTGGTTTACTAGCAAGAGAAGGTTCTGCACAGGTGAGTACTACGAAC
TTGATTCATGACCTTGTGACATCAGCTCCACCCCCAATACCTCAATGTTTGTCAACAAAAGTGATAGAAGGTAGCTGCCAGATTGCA
TCATCCTTCCTGAGTCCCCAGGCACTTGTTTTGGCAGACTTAGTCTTACGGTTGTAGAAGG

PANK6(68)

CACCATTATGAGCGCACTACAGGTTCAGCAGATGGCCCCAGCAGACTTAGTCTTACGGTTGTAGAAGG

P2(220)

CCTTTACAACGCTGACACGAAGAAGTGCTTCACGGCCAGGTGGTCCTCCCAGTCGTTGGTCAACGACTTGTAGAACTCGCCGTACTC
CTCCTTCGAGATGTCGTCGGGGTTGCGCGTCCACAGGGGCTTCGTCTTGTTCAGCTCCTCGTCCTCCGTGTACTTCTCCTTCACCGT
CTTCTTTTTCTTCTTGTCATCGCACTTCTTCGTGTCGGCGTTGTAAAGG

Apéndice 3: INICIADORES DEGENERADOS. Una Guía corta. (Modificado y traducido; de Bones, 2003)

Iniciadores degenerados en la PCR

La PCR utilizando iniciadores degenerados, es idéntica a la PCR ordinaria, con la diferencia de que en lugar de utilizar iniciadores específicos para una secuencia nucleotídica, se utilizan iniciadores mezclados. Es decir, si no se conoce exactamente la secuencia del gen que se desea amplificar, el inserto varía ("wobbles") en la región donde se alinea el iniciador donde este tiene más de una posibilidad. Por lo tanto, si se tiene la secuencia primaria de una proteína, se puede traducir a su correspondiente secuencia nucleotídica.

Ejemplo del diseño de un iniciador degenerado para PCR a partir de una secuencia de aminoácidos:

Trp Asp Thr Ala Gly Gln Glu

5' TGG GAY ACN GCN GGN CAR GA 3' Este da una mezcla de 256 iniciadores diferentes.

Donde la Y=C o T, R=G o A, N=G, A, T o C.

Mientras más variaciones ("wobbles") introduzcas en el iniciador, más degenerado será éste. La degeneración del iniciador es producida durante la síntesis del ADN, de tal manera, que no se necesitan 256 iniciadores diferentes para obtener una secuencia con tal variabilidad, que sería una pérdida de tiempo de trabajo y además, muy caro.

¿Porqué usar iniciadores degenerados para la PCR?

Los iniciadores degenerados, son una poderosa herramienta para detectar "nuevos" genes o familias de genes. Muchos genes forman familias con similaridades estructurales. Por alineamiento, las secuencias de proteínas de un número de proteínas relacionadas se pueden encontrar regiones en las cuales la secuencia es conservada o variable. Basados en esta información se pueden encontrar partes conservadas que pueden servir como punto de partida para el diseño de iniciadores degenerados.

Se pueden "resolver" problemas con el uso de iniciadores degenerados.

1. Si se tiene aislada una proteína y se tiene la secuencia de algunos aminoácidos de ésta. Se puede encontrar el gen correspondiente!! (Porqué no probar con iniciadores degenerados!).

2. Si se encontró un gen en algún organismo y se busca clonar el gen homólogo de otra especie. (Se puede probar con hibridación de baja astringencia, pero, ¿cuantos falsos positivos tendrían que secuenciarse antes de encontrar la correcta?).

3. Si se encontró un gen de interés en algún organismo y se busca encontrar si existe el homologo en otra especie.

4. Estudios filogenéticos y evolucionarios de genes. Se pueden encontrar genes ortologos específicos de un número de especies relacionadas y compararlos. (Este tipo de información puede revelar sitios activos potenciales, regiones regulatorias y mucho mas).

5. En estudios de familias de genes. (Cantidad de miembros e una familia y detección de diferentes en comparaciones con las obtenidas en especies cercanas o alejadas filogenéticamente).

COMENTARIOS TECNICOS

Requerimientos. (Qué tipo de información acerca de la secuencia necesito para empezar).

- Dos bloques de aminoacidos/nucleotides conservados. La longitud de los iniciadores debe ser como mínimo de 20 pb.

La secuencia no tiene que estar 100% conservada. Una secuencia parcialmente conservada puede ser suficiente. Ejemplo de sustituciones comunes encontradas son Glu <--> Asp and Arg <-->Lys. Si se usa el "codón" degenerado GAN, este se convierte en ambos Glu y Asp. Es similar si se utiliza el codón MGN, (M=C o A), donde este será un aminoácido básico (Arg or Lys), el codón MGN se convierte parcialmente en el codón Lys (AAR). Si este es un residuo de Lys, tendremos un error en G/T en la segunda base. Esto normalmete no es un problema si el error ocurre al centro del iniciador. (Recuerda tu bioquímica, la forma enol de la timidina puede aparearse con la guanine).

- La parte N-terminal de una proteína (obtenida por secuenciación) a menudo da, suficiente información de la secuencia para el diseño de iniciadores degenerados para PCR.

Si la secuencia N-terminal es de 20 a 30 Aminoácidos, a menudo es posible hacer dos iniciadores degenerados, hacer una PCR y obtener un fragmento de cADN de 50-90 pb el cual puede ser usado como sonda para obtener por hibridación en una librería de cADN. Alternativamente se puede realizar un RACE 3', para amplificar el resto de el cADN.

ADNc o ADN genómico?

- En general es mejor trabajar con ADNc que con AND como templado. El AND es de menor complejidad (En eucariotes solo un pequeño porcentaje del ADN genómico codifica proteínas).

- En ADNc la talla de los fragmentos de PCR es “predecible”, porque este no posee introns.

- Cuando no se está seguro si un gen no se expresa en un tejido o estado de desarrollo determinado se utilice ADNc. El ADN genómico se puede usar solo como material de inicio. La desventaja es que para examinar lote de amplificaciones de AND inespecíficas antes de encontrar el gen deseado.

¿Cómo un iniciador degenerado puede ser y hacer su función?

Degeneraciones de 1000 - 10.000 veces no son tan comunes.
La degeneración de los iniciadores puede "bajar" la variación sustituyendo cuatro bases con inosinas. (Ejemplo: GGI en lugar de GGN).
Ej: Secuencia: CVGG(M/L)NRRP (encontrada en proteínas p53).
Sin Inopinas obtendriamos 131072 secuencias iniciadoras posibles.
5' TGY GTN GGN GGN MTN AAY MGN MGN CC 3'
Con inosinas, 512.
5' TGY GTI GGI GGI MTI AAY MGN MGN CC 3'

Cómo escoger las condiciones de PCR.

- Probar "condiciones estandar" con bajas temperaturas de alineamiento, 35-50 ciclos.
- Si las condiciones estandar arroja resultados negativos, correr los primeros 4 ciclos de 5-10ºC mas abajo de lo recomendado. (Se extenderán muchas secuencias, pero se puede tener la esperanza de que ahi esta el gen de interés).
- Si los iniciadores son muy degenerados (512 posibilidades o mas), una inhibición competitiva puede dar problemas (Los iniciadores se unen al templado correcto, pero no son extendidos por la polimerasa, por la inestabilidad de los extremos 3'). Esto significa que los primeros ciclos de la PCR son muy ineficientes y algunas veces se requiere del corrimiento de 50 ciclos, para percibir apenas la banda correspondiente al gen de interés.
- Recuerda no utilizar una ADN polimerasa con actividad exonucleasa 3-->5', (Los iniciadores podrían degradarse).

Qué tipo de genes son "fácil" de encontrar con iniciadores degenerados?

Muchas proteínas tienen similaridades estructurales con otras proteínas, y a menudo tienen un origen evolucionario común.

Las proteínas con antiguas secuencias conservadas (Ancient Conserved Motifs_ACM's-), son en general "fácil" de encontrar.
Mas de 500 familias de proteínas de ACM's son conocidas (algunas de estas familias son enormes: Ser- Thr- Tyr- kinasas en humanos, alrededor de 1000 genes). En el año 2000 la secuencia genomita de 8 especies eucariotas son conocidas (Humano, *Drosophila melanogaster* (mosca de la fruta), *Anopheles gambiae* (el mosquito), *C. elegans* (nematodo), *S. cerevisiae* (levadura), *Schizosaccharomyces pombe* (levadura), *Arabidopsis thaliana* (planta), *Plasmodium falciparum* (protista) y *Giardia lamblia*). En adición, 10 de los genomas bacteriales mas estudiados están completes. Estos genomas brindan información de la evolución de varias familias de genes y pueden ser utilizados como punto de partida para encontrar genes, por su falta de estudio, hasta en los organismos más oscuros. Iniciando con el alineamiento de una proteína de interés. Incluyendo tantas proteínas como se puedan encontrar. Si la proteína no es conservada, se puede probar buscando regiones conservadas de aminoácidos y si se conoce la secuencia de la proteína de una especie cercanamente emparentada, se puede usar esta como guía. Algunas veces se puede jugar con la secuencia y correr con gran suerte.

Implicaciones:

Por el uso de iniciadores degenerados, a partir de secuencias de organismos bien conocidos, se pueden encontrar muchos genes de especies apenas estudiadas. Los problemas pueden surgir si se quiere probar la captura de genes de la rápida evolución. Si no es el caso, se puede tener la seguridad de encontrar lo que se busca con el uso de iniciadores degenerados. El caso puede estar dado por la presencia de una fuerte región, como lo que se observa para los genes en protistas como las criptomonas, donde muchos genes tienen una deriva genética masiva, y tienen cambiada una región comparada a otros eucariotas. Aparte de esto, las limitaciones son, en general relativamente pocas.

Otras limitaciones:

1. Los aminoácidos conservados que se prueben para diseñar después los iniciadores están compuestos principalmente de Ser, Arg y Leu. (Estos son los aminoácidos que pueden dar mas variaciones o "wobbles" en una secuencia nucleotídica). Estos pueden, algunas veces, dar iniciadores, los cuales amplificarán cualquiera de las degeneraciones encontradas en un blanco, lo cual se traduce en un alto contenido de "basura" o productos inespecíficos.

2. La región a amplificar es grande. Como una regla, evitar productos de PCR mayores de 1000 pb.

3. El organismo del cual se pretende amplificar un fragmento, tiene alto contenido de GC. Esto a menudo acarrea errores, de manera que al final de la amplificación, se podría obtener un lote de fragmentos incorrectos.

4. El organismo del cual se pretende amplificar un fragmento, tiene bajo contenido de GC y además se diseña un iniciador corto. (Alto contenido de TA, da una temperatura de alineamiento baja).

5. El gen que tu buscas no existe en el organismo que tu seleccionaste. Hay pocos ejemplos de "genes dinosaurios", los cuales tienden a desaparecer en ciertos linajes durante la evolución. (por ejemplo, los genes *Rac* en *S. cerevisiae*).