

Maestría en Bioinformática y Biología de Sistemas

Título del Trabajo

Análisis del transcriptoma de *Delphacodes kuscheli*, vector del virus del Mal de Río Cuarto

Trabajo Final de la Alumna

Magali Piccione

El trabajo ha sido presentado como requisito
para la obtención del título de

Magíster en Bioinformática y Biología de Sistemas

María Inés Catalano
Directora

Agustina Pascual
Co-directora

Pergamino, 21 de julio de 2026

Análisis del transcriptoma de *Delphacodes kuscheli*, vector del virus de Mal de Río Cuarto

Trabajo final de la alumna:

Magali Piccione

Aprobada por el Tribunal Evaluador:

Dra. María E. Brentassi
Evaluador/a

Dra. Gabriela P. Fernández
Evaluador/a

Dra. Lucila M. Traverso
Evaluador/a

María Inés Catalano
Directora

Agustina Pascual
Co-directora

**Escuela de Ciencias Agrarias, Naturales y Ambientales, Universidad Nacional del
Noroeste de la Provincia de Buenos Aires**

Pergamino, 21 de julio de 2026

Resumen

La creciente disponibilidad de métodos asequibles para la secuenciación de genomas completos, junto con el desarrollo de grandes proyectos de anotación genómica, ha generado una gran cantidad de datos que deben convertirse en información funcional. En este contexto, el análisis comparativo resulta fundamental tanto para comprender procesos evolutivos como para identificar genes en organismos previamente desconocidos.

En este trabajo, se propuso caracterizar el transcriptoma del vector del patógeno causal de la enfermedad Mal de Río Cuarto (MRC), *Delphacodes kuscheli* Fennah (Hemiptera: Delphacidae). Para ello, se recolectaron ejemplares de todos los estadios del ciclo de vida del insecto, a partir de los cuales se realizó la extracción de ARN total y, posteriormente, se obtuvo el perfil de expresión génica mediante secuenciación de ARN (ARN-Seq) utilizando la tecnología Illumina.

Las lecturas crudas obtenidas fueron pre-procesadas para evaluar su calidad y posteriormente ensambladas *de novo* en transcritos, sin la necesidad de contar con un genoma de referencia. Las métricas de cobertura y completitud alcanzadas permitieron proseguir con la anotación génica, la cual se contrastó con diversas bases de datos asociadas a procesos como la ovogénesis, ecdisis y olfacción. De este modo, se logró inferir la identidad de un amplio repertorio de genes conservados en el insecto plaga.

Este trabajo sienta las bases para futuros estudios funcionales y para el desarrollo de estrategias de manejo de plagas y control de vectores de enfermedades.

Palabras claves: Insecto plaga, Auchenorrhyncha, transcriptoma, ensamble *de novo*.

Agradecimientos

A Gabriela Fernández, por su ayuda durante la carrera, y hacer posible su realización.

A Agustín Baricalla, mi consultor, por su paciencia, y estar siempre dispuesto a explicarme, enseñarme y ayudarme, así como también por compartir su sabiduría conmigo.

A mis directoras, María Inés Catalano y Agustina Pascual, por la oportunidad de participar en este proyecto.

A mis compañeros por los buenos momentos vividos durante la carrera y especialmente por compartir esos momentos difíciles de cursada, con la misma preocupación y motivación. Especialmente a Yael Nazar, gran compañero de cursada y consultor.

Gracias a todos aquellos que de una u otra forma me acompañaron en este camino. A mis compañeros de trabajo que siempre preguntan, escuchan, levantan el ánimo y me apoyan. A todos los que me han dado fuerza a lo largo del camino, como Isa, Guillermina, Francesco, Elena y Abel por su empuje constante y energía; y especialmente a Carlos y Miguel, mi papá adoptivo, por estar siempre pendientes y predispuestos a ayudarme en todo. Gracias por estar.

A mi gran amigo Matías Porta, por su empuje para seguir adelante en todo, por ayudarme a no rendirme nunca y pelear para salir adelante, por su apoyo incondicional.

Aunque no es tan encantador como cree, es la voz de mi conciencia. A mi némesis, Samuel Jurado, que me ha apoyado mucho en todo momento, y con mucha paciencia...gracias por sus aportes, impulso e insistencia para seguir adelante. ¡¡Infinitas gracias!!

A mi hermana y a mi mamá, las dos grandes mujeres de mi vida; que siempre me dan fuerzas para seguir y me enseñan que no importa lo que pase, siempre hay que tirar para adelante. ¡¡Gracias por todo!!

A Leo, mi pilar y sostén. Por su paciencia y aguante. Por su ayuda durante todo el proceso. Por su fuerza y empuje para que siempre me supere y siga avanzando. ¡¡Gracias por estar siempre!!

Simplemente....

¡¡¡Gracias Totales!!!

Índice

RESUMEN	3
AGRADECIMIENTOS	5
ÍNDICE.....	7
ABREVIATURAS	10
INTRODUCCIÓN	12
1. TECNOLOGÍA DE SECUENCIACIÓN MASIVA DEL ARN (ARN-SEQ)	13
2. ANÁLISIS DE DATOS DE ARN-SEQ	15
3. ANOTACIÓN GÉNICA MEDIANTE SIMILITUD DE SECUENCIA	16
4. <i>DELPHACODES KUSCHELI</i>	17
4.1 <i>Procesos funcionales en insectos</i>	18
4.1.1 Ecdisis	19
4.1.2 Ovogénesis	20
4.1.3 Olfacción	20
HIPÓTESIS Y OBJETIVOS	22
HIPÓTESIS.....	23
OBJETIVO GENERAL	23
<i>Objetivos Específicos</i>	23
MATERIALES Y MÉTODOS	24
1. MANTENIMIENTO DE LA COLONIA DE <i>DELPHACODES KUSCHELI</i>	25
2. EXTRACCIÓN DE ARN	25
3. SECUENCIACIÓN DE ARN	25
4. ANÁLISIS BIOINFORMÁTICO	26
4.1 <i>Análisis de calidad y limpieza de las lecturas</i>	26
4.2 <i>Ensamble de novo</i>	27
4.3 <i>Análisis de la calidad y completitud del ensamble</i>	27
5. CARACTERIZACIÓN DEL ENSAMBLE DE <i>NOVO</i>	28
6. ANOTACIÓN GÉNICA	28
RESULTADOS.....	30
1. ANÁLISIS BIOINFORMÁTICO	31
1.1 <i>Análisis de calidad y limpieza de las lecturas</i>	31
1.2 <i>Ensamble de novo y análisis de calidad</i>	34
1.3 <i>Anotación Génica</i>	44
DISCUSIÓN	52
1. SECUENCIACIÓN Y ENSAMBLE	53
2. ANOTACIÓN GÉNICA	57
2.1 <i>Ecdisis</i>	57
2.2 <i>Ovogénesis</i>	58
2.3 <i>Olfacción</i>	59
CONCLUSIONES FINALES	60
BIBLIOGRAFÍA	62
ANEXO I: COMANDOS	83
ANEXO II: FIGURAS	91

ANEXO III: BASES DE DATOS GENERADAS	93
ECDISIS	94
OVOGÉNESIS	102
OLFACCIÓN	102
ANEXO IV: SECUENCIAS ANALIZADAS	141
ECDISIS	142
OVOGÉNESIS	143
OLFACCIÓN	159

Abreviaturas

.gz	Formato de compresión GZIP
µg	Microgramo
aa	Aminoácidos
ADN	Ácido Desoxirribonucleico
ADNc	ADN copia
A-P	Anteroposterior
ARN	Ácido Ribonucleico
ARNm	ARN Mensajero
ARNr	ARN ribosomal
ARN-Seq	Secuenciación Masiva del ARN
ATP	Adenosín trifosfato
cola poly-A	ARN poliadenilado
D-V	Dorsoventral
EDTA	Ácido Etilendiaminotetraacético
GB	Gigabyte
kb	Kilobyte
MRC	Mal de Río Cuarto
MRCV	Virus del Mal de Río Cuarto
NGS	Secuenciación de Próxima Generación
nt	Nucleótidos
pb	Pares de Bases
PCR	Reacción en Cadena de la Polimerasa

Introducción

1. Tecnología de secuenciación masiva del ARN (ARN-Seq)

La tecnología de secuenciación de próxima generación (NGS, del inglés *Next Generation Sequencing*) se desarrolló originalmente para permitir la secuenciación masiva y simultánea de genomas completos, proporcionando información específica sobre la estructura, función y dinámica del ADN (Satam *et al.*, 2023). La gran cantidad de datos generados por NGS impulsó el desarrollo de herramientas bioinformáticas, esenciales para su procesamiento y análisis eficiente. Esta combinación tecnológica y computacional permitió la extensión de NGS a diversos enfoques ómicos, tales como la transcriptómica, la epigenómica y la metagenómica, contribuyendo de manera sustancial al avance en la comprensión de la genética y la diversidad biológica (Wesolowski *et al.*, 2013; Levy & Myers, 2016).

La transcriptómica, definida como la secuenciación del conjunto completo de moléculas de ARN mensajeros (ARNm) de una muestra biológica, ya sea a nivel celular, tisular u orgánico, ha surgido como una herramienta fundamental en la investigación biológica (Morozova *et al.*, 2009; Wang *et al.*, 2009; Bramlett *et al.*, 2011). A diferencia del genoma, que constituye el complemento genético total de un organismo, el transcriptoma refleja la expresión génica en un momento y contexto determinado, modulada tanto por factores internos como externos (Velculescu *et al.*, 1997). La secuenciación de ARN transformó el análisis de transcriptomas, al permitir obtener una cobertura completa de los transcriptos, incluso en organismos sin genoma previamente caracterizado. Además, permite detectar un amplio rango de expresión con alta reproducibilidad y eficiencia (Nagalakshmi *et al.*, 2008; Vera *et al.*, 2008; Malone & Oliver, 2011; McGettigan, 2013; Kukurba & Montgomery, 2015; Ambardar *et al.*, 2016). El análisis de transcriptomas resulta esencial para interpretar los elementos funcionales de un genoma, revelar los constituyentes moleculares de células y tejidos (Wang *et al.*, 2009), y abordar diversas problemáticas biológicas de interés (Denoeud *et al.*, 2008; Wang *et al.*, 2008; Maher *et al.*, 2009; Yassour *et al.*, 2009; Berger *et al.*, 2010; Blekhman *et al.*, 2010; Guttman *et al.*, 2010; Mortazavi *et al.*, 2010; Trapnell *et al.*, 2010; Wilhelm *et al.*, 2011).

Las tecnologías de secuenciación han evolucionado de manera significativa a lo largo de las últimas décadas, y se suelen clasificar en tres generaciones principales: primera, segunda y tercera (López de Heredia, 2016; Rubio *et al.*, 2020). La primera generación, representada por el método de Sanger, se basa en la terminación de cadenas, y constituyó el primer enfoque exitoso para la secuenciación (Sanger *et al.*, 1977). Posteriormente, la llegada de la tecnología de segunda generación o NGS, transformó radicalmente el proceso, permitiendo obtener un gran volumen de secuencias en menos tiempo y a un costo relativamente bajo (Metzker, 2005; 2010).

Entre estas se destacan la pirosecuenciación (454 *Sequencing* de Roche) y el PCR en puente (*Bridge* PCR de Illumina) (Nagalakshmi *et al.*, 2008), siendo este último el más utilizado en la actualidad a nivel global (Goodwin *et al.*, 2016; Valderrama Martín *et al.*, 2020). A pesar de los avances, estas tecnologías presentan la limitación de generar secuencias (lecturas) relativamente cortas, lo que requiere posteriores procesos de ensamblajes complejos (Blencowe *et al.*, 2009; Wolf, 2013). En contraste, la tecnología de tercera generación constituye un avance adicional al posibilitar la secuenciación en tiempo real de moléculas únicas de ADN (*single molecule real time sequencing*), permitiendo la obtención de secuencias largas (Bautista Moreno, 2010; Rodríguez-Alonso & Shishkova, 2018). Entre estas, la plataforma de *nanopore* (Oxford Nanopore Technologies, ONT) se destaca por alcanzar lecturas de hasta 23 kilobytes (kb) (Wang *et al.*, 2023) y, aunque aún se encuentra en desarrollo, ha demostrado ser particularmente útil en la secuenciación de genomas de pequeño tamaño (López de Heredia, 2016).

Las tecnologías de secuenciación de segunda generación, ya sea para la obtención de ADN o ARN, e independientemente de la plataforma utilizada, comparten una serie de etapas fundamentales que las caracterizan. En cuanto a la secuenciación de ARN (ARN-Seq), inicialmente, se extrae el ARN total de una muestra biológica de interés y se procede al enriquecimiento de ARNm. Este paso puede realizarse mediante la selección positiva de los transcritos poliadenilados (poly(A)) o a través de la depleción de ARN ribosomal (ARNr). Posteriormente, las moléculas de ARNm se fragmentan en unidades más pequeñas y se utilizan para sintetizar moléculas de ADN complementario (ADNc). A estos fragmentos de ADNc se les añaden adaptadores de secuencia conocida en sus extremos, los cuales actúan como iniciadores para la amplificación mediante PCR (Fedurco *et al.*, 2006; Rodríguez-Alonso & Shishkova, 2018), generando así la biblioteca de secuenciación, que constituye el material de entrada para el secuenciador. Las lecturas resultantes pueden ser *single-end*, secuenciadas desde un único extremo o *paired-end*, secuenciadas desde ambos extremos, y con longitudes variables entre 25 y 300 pares de bases (pb). Dado que estas lecturas son más cortas que los transcritos originales, se requiere un paso posterior de ensamblado para reconstruir los transcritos completos (Fedurco *et al.*, 2006; Wang *et al.*, 2009; Wesolowski *et al.*, 2013; Rodríguez-Alonso & Shishkova, 2018).

La reconstrucción del transcripto presenta un desafío operacional debido a múltiples fuentes de variabilidad, entre ellas, la introducida durante la preparación de la biblioteca, la tasa de error inherente al secuenciador y la variabilidad biológica propia de cada muestra (Metzker, 2010; Wesolowski *et al.*, 2013). Ante esto, se requieren métodos eficientes, capaces de reconstruir transcritos de diferentes tamaños, con niveles de expresión e isoformas variables, y

garantizar resultados de calidad. En este contexto, se han desarrollado diversas herramientas bioinformáticas que permiten superar dichas dificultades y enfrentar de manera eficiente los distintos desafíos presentes durante el análisis transcriptómico (Haas & Zody, 2010).

2. Análisis de datos de ARN-Seq

El análisis de las lecturas generadas por ARN-Seq requiere estrategias computacionales específicas para procesar el gran volumen de lecturas y enfrentar los desafíos asociados al ensamblado y la interpretación biológica (Garber *et al.*, 2011). El primer paso consiste en el control de calidad de las lecturas, que permite identificar y eliminar secuencias no informativas, como adaptadores, bases ambiguas o de baja calidad (Kukurba & Montgomery, 2015). Para esto, existen diversas herramientas bioinformáticas de acceso abierto como FastQC (Wingett & Andrews, 2018), Trimmomatic (Bolger *et al.*, 2014), Blast (Altschul *et al.*, 1990) y Fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), las cuales son ampliamente utilizadas para esta etapa de pre-procesamiento, asegurando que las lecturas sean confiables para los pasos posteriores.

En ausencia de un genoma y/o transcriptoma de referencia, la reconstrucción *de novo* constituye la estrategia más adecuada (Conesa *et al.*, 2016). Este enfoque, ampliamente utilizado en organismos no modelo o en especies con información genómica incompleta o fragmentada (Conesa *et al.*, 2016; Rodríguez-Alonso & Shishkova, 2018), se basa en algoritmos como los grafos de *De Bruijn*, que dividen las lecturas en fragmentos solapantes (*k-meros*) para generar transcritos completos (Pevsner, 2009).

Brevemente, los *k-meros*, de una longitud fija de *k* nucleótidos, se conectan a través de un solapamiento exacto de un largo de *k-1* nucleótidos desde sus extremos (Figura 1), lo que permite reconstruir secuencias más largas a partir de un *k-mero* inicial, elegido aleatoriamente (Pevsner, 2009). Cada camino en el grafo representa un posible transcripto, y cuando un extremo admite más de un *k-mero* que puede ser alineado, se genera una burbuja, y se continúa la extensión de la secuencia en las dos posibles variantes, representando isoformas, variantes alélicas o errores de secuenciación (De Bruijn, 1946; Good, 1946). Por ello, resulta esencial disponer de lecturas limpias para evitar solapamientos falsos que desvíen la reconstrucción por caminos incorrectos (Grabherr *et al.*, 2011). Entre los programas disponibles para la reconstrucción de transcritos, Trinity se destaca por su gran flexibilidad y versatilidad, aunque también existen alternativas como Oases o Trans-AbySS (Grabherr *et al.*, 2011; Palacio, 2018).

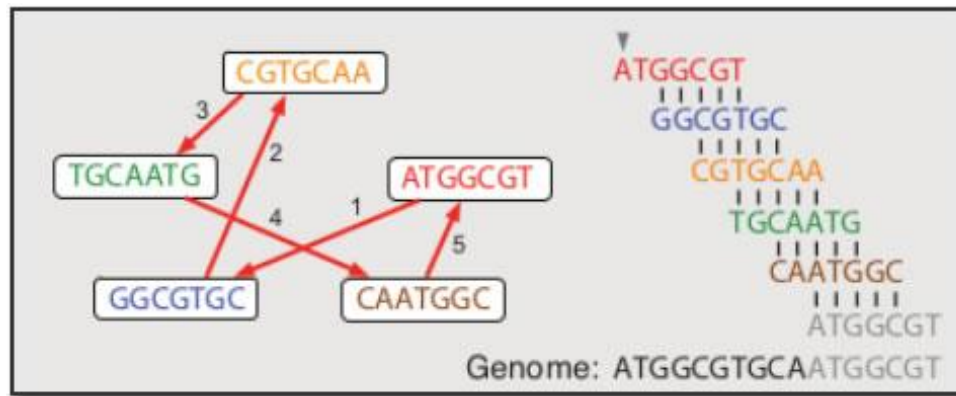


Figura 1. Representación gráfica del proceso de ensamblado. Cinco lecturas de longitud $k\text{-mer}$ 7 muestran los posibles solapamientos como consecuencia de su coincidencia de base. Los números especificados en las flechas rojas determinan el camino de reconstrucción del transcripto. En la derecha de la figura se observan las lecturas alineadas que constituyen la secuencia génica (extraído de Compeau *et al.*, 2011).

Una vez finalizado el ensamblado, se procede a evaluar la calidad e integridad del conjunto de datos mediante diversas métricas. Entre ellas se incluyen el porcentaje de genes codificantes de proteínas, que refleja la completitud del ensamblado, así como las métricas de cobertura y estadísticas descriptivas relacionadas con la longitud y la cantidad de transcriptos generados, las cuales definen al transcriptoma ensamblado (Seemann & Gladman, 2012; Cock *et al.*, 2015; Simão *et al.*, 2015). La obtención de métricas robustas establece una base sólida para avanzar en la identificación de los genes que caracterizan el ensamblado (Anders & Huber, 2010; McIntyre *et al.*, 2011; Wang *et al.*, 2011).

3. Anotación génica mediante similitud de secuencia

La genómica comparada es una disciplina que permite comparar las secuencias génicas entre diferentes especies para inferir relaciones evolutivas y funcionales (Haubold & Wiehe, 2004). Esta disciplina parte de la premisa de que las características compartidas por diferentes organismos se encuentran codificadas en regiones genómicas conservadas a lo largo de la evolución (Hardison, 2003; Jiang *et al.*, 2013). Si bien la función de un gen se determina experimentalmente, la identificación de secuencias ortólogas constituye una aproximación, dado que genes con similitud de secuencia suelen desempeñar funciones biológicas afines (Aguilar-Bultetl & Falquet, 2015).

La comparación de genomas ha sido clave para comprender procesos biológicos

fundamentales (Tatusov *et al.*, 1997; Oshlack *et al.*, 2010), y ha permitido revelar la genética que caracteriza diversos organismos modelo, evidenciando la conservación de componentes genéticos incluso entre especies filogenéticamente distantes (Cloonan *et al.*, 2008; Marioni *et al.*, 2008; Morin *et al.*, 2008; Oshlack *et al.*, 2010; Ewen-Campen *et al.*, 2011; Jiang *et al.*, 2013). Para este tipo de análisis, una de las herramientas más utilizadas es Blast (*Basis Local Alignment Search Tool*), que mediante un enfoque heurístico permite identificar secuencias similares y, con ello, detectar si los genes de interés se encuentran representados en ambas especies (Altschul *et al.*, 1990). La elección de la base de datos de referencia para estas búsquedas depende del objetivo del estudio, pudiendo abarcar desde genomas completos hasta bases de datos específicas, disponibles en repositorios públicos o generadas manualmente (McCartney *et al.*, 2024).

El uso de la similitud de secuencia ha sido particularmente relevante en el estudio de transcriptomas de insectos, donde ha permitido generar información clave tanto para la comprensión de su biología y evolución, así como para aplicaciones en salud y agricultura (Ellegren *et al.*, 2012; Lamichhaney *et al.*, 2012; Jiang *et al.*, 2013; Ekblom & Wolf, 2014; Wanninger, 2015). En este sentido, se destacan numerosos estudios en distintos órdenes tales como Diptera (Etebari *et al.*, 2017; Santana *et al.*, 2019), Hemiptera (Ewen-Campen *et al.*, 2011; Ons *et al.*, 2016; de Carvalho *et al.*, 2017; Lavore *et al.*, 2018; Palacio, 2018; Singh *et al.*, 2019; Cagliari *et al.*, 2020; Pascual & Rivera-Pomar, 2022) y Lepidoptera (Perera *et al.*, 2015).

En este contexto, la anotación por similitud de secuencia constituye una herramienta valiosa para identificar regiones funcionales y sentar las bases para la predicción de funciones biológicas (Hardison, 2003). No obstante, este enfoque presenta limitaciones, especialmente en organismos no modelo, debido a la escasa disponibilidad de anotaciones específicas, lo que restringe el alcance de las inferencias. En estos casos, las predicciones dependen de la información proveniente de especies previamente caracterizadas y de la calidad de las bases de datos utilizadas (Conesa *et al.*, 2016).

4. *Delphacodes kuscheli*

Los insectos son el grupo de animales con mayor éxito evolutivo, principalmente por su abundancia, diversidad y amplio espectro de hábitats que ocupan (Purvis & Hector, 2000; Stork, 2018). Entre ellos, el orden Hemiptera destaca por su riqueza taxonómica, con más de 110.000 especies descritas (Berenbaum, 1996; Bartlett *et al.*, 2018). Un suborden particularmente relevante es Auchenorrhyncha, el cual está integrado exclusivamente por especies fitófagas,

muchas de las cuales representan plagas agrícolas responsables de importantes pérdidas económicas (Maramorosch & Harris, 1979; Nault & Ammar, 1989; Virla *et al.*, 2004; Bartlett *et al.*, 2018).

La familia Delphacidae, dentro de Auchenorrhyncha, reúne numerosas especies plaga que ocasionan daños a los cultivos a través de la oviposición, la alimentación y la transmisión de fitopatógenos (Wilson, 2005).

Dentro de ellos, se destaca *Delphacodes kuscheli* Fennah como vector del virus del Mal de Río Cuarto (MRCV), causando la enfermedad Mal de Río Cuarto (MRC), que afecta el cultivo de maíz en Argentina (Remes Lenicov *et al.*, 1985). Esta enfermedad es ocasionada por un virus de la familia de *Fijivirus* (Conci & Marzachí, 1993), cuya transmisión es persistente y propagativa, tanto a través de ninfas como de adultos (Truol *et al.*, 2001).

El ciclo de vida de *D. kuscheli* tiene una duración de varias semanas, y comprende la etapa de huevo (7,5 a 10 días), 5 estadios ninfales (I a V, con una duración aproximada de 3 a 5 días cada uno) y el estado adulto (1 a 3 semanas). La duración de cada estado y estadios ninfales depende de la planta huésped en la cual se encuentre, variando también según el sexo (Virla & Remes Lenicov, 1991).

Los principales hospedantes de *D. kuscheli* son gramíneas y cereales de invierno, como avena y trigo, en los cuales se reproduce y desarrolla (Remes Lenicov *et al.*, 1991; Ornaghi *et al.*, 1993). Sin embargo, en ausencia de estos hospedantes, tanto ninfas como adultos se alimentan sobre plantas jóvenes de maíz, lo que favorece la transmisión del patógeno a este cultivo (Remes Lenicov *et al.*, 2008).

A pesar de su importancia biológica y económica, los elementos genéticos que caracterizan a la especie son aún limitados en comparación con otros hemípteros. Estudios recientes han aportado información relevante mediante el análisis transcriptómico del cuerpo graso, lo que permitió identificar genes involucrados en diversas vías metabólicas (Pascual *et al.*, 2025). En paralelo, la caracterización de su microbioma reveló la presencia de una comunidad de simbiosis tipo levadura (YLS) que contribuyen a funciones metabólicas esenciales, evidenciando una estrecha relación funcional entre el hospedador y su simbionte (Brentassi, *et al.*, 2020; Pascual *et al.*, 2025).

4.1 Procesos funcionales en insectos

El gran éxito evolutivo de los insectos se debe, en gran medida, al desarrollo y

establecimiento de procesos claves para su supervivencia, como la muda o ecdisis, la detección de compuestos volátiles mediante la olfacción, y su elevada capacidad reproductiva (Andlla *et al.*, 2018).

4.1.1 Ecdisis

La ecdisis o muda es el proceso mediante el cual los insectos reemplazan su antigua cutícula por una nueva, lo que posibilita el incremento en tamaño corporal (Wigglesworth, 1934; Reynolds, 1980). La cutícula, que constituye la capa más externa del insecto, es una estructura multifuncional formada por proteínas fuertemente entrelazadas, y se encuentra organizada en dos capas: la epicutícula, la más externa y rica en lípidos, y la procutícula. Esta, a su vez, se divide en la exocutícula, esclerotizada y rígida, y la endocutícula, capa interna y más flexible. La cutícula actúa como barrera frente a factores ambientales, la pérdida de agua y como defensa ante sus depredadores, además de permitir la inserción de la musculatura responsable de la locomoción (Gangishetti *et al.*, 2012; Moussian, 2013).

El proceso de muda puede dividirse en cuatro etapas principales: la apósis o premuda, durante la cual el insecto separa la cutícula vieja de su epidermis y comienza a formar una nueva; la muda o ecdisis propiamente dicha, donde se produce el desprendimiento de la cutícula vieja; la postmuda o esclerotización, en la que la nueva cutícula se endurece progresivamente; y la intermuda, etapa en la que el insecto se alimenta y crece (Schwartz, 1977; Truman & Riddiford, 1999).

La formación de la nueva cutícula y eliminación de la vieja está regulada hormonalmente. Dos familias de hormonas regulan de manera central la ecdisis, las ecdisteroides y la hormona juvenil (JH). Las primeras inducen la producción de una nueva cutícula, mientras que la segunda modula el proceso; en particular, la disminución en sus niveles desencadena la transición de larva a adulto (Truman & Riddiford, 1999). Además, múltiples genes participan en la coordinación del proceso, actuando como reguladores, a través de la represión o activación de la expresión génica, desencadenando respuestas celulares específicas (Doctor *et al.*, 1985; Dubrovsky, 2005; Soares *et al.*, 2007; Charles, 2010; Gangishetti *et al.*, 2012; Planelló *et al.*, 2015). Entre ellos se destacan, por un lado, los relacionados con la regulación de la ecdisis, como corazonina (Crz), la hormona de eclosión (*eclosion hormone*, EH), la hormona disparadora de la ecdisis (*ecdysis-triggering hormone*, ETH) y el péptido cardioactivo de crustáceos (*crustacean cardioactive peptide*, CCAP) (Lorenz *et al.*, 1989; Zitnan *et al.*, 1996; Ewer *et al.*, 1997; Kingan *et al.*, 1997; Helfrich-Förster, 2004; Truman, 2005; Planelló *et al.*, 2015; Daubnerová & Žitňan, 2021); y por otro lado, los que

actúan posteriormente al proceso de ecdisis como *bursicon* (*Burs*) y *partner of bursicon* (*Pburs*) (Luo *et al.*, 2005; Yu *et al.*, 2021).

4.1.2 Ovogénesis

La ovogénesis es un proceso esencial que implica la formación y desarrollo del ovocito, el cual debe ocurrir en el momento adecuado y siguiendo una secuencia precisa para asegurar el inicio exitoso del desarrollo embrionario. Durante este proceso, no solo se produce el establecimiento de los ejes corporales del futuro embrión (Nüsslein-Volhard & Roth, 1989; Schüpbach & Roth, 1994; Speder *et al.*, 2006; Coutelis *et al.*, 2008), sino que también se lleva a cabo la vitelogénesis (Atella *et al.*, 2005) y se desarrolla la capa protectora del mismo, denominada corion (Papantonis *et al.*, 2015).

El conocimiento sobre la red que regula el proceso de ovogénesis se encuentra caracterizado en el díptero *Drosophila melanogaster* (Brody, 1999), y en los últimos años se han generado avances significativos en otras especies, como en el coleóptero *Tribolium castaneum* (Bäumer *et al.*, 2012; Schmitt-Engel *et al.*, 2015), el hemíptero *R. prolixus* (Pascual *et al.*, 2021) y el delfácido *Nilaparvata lugens* (Zhang *et al.*, 2015), aunque este sigue siendo muy limitado.

La sucesión de eventos que conforman la ovogénesis refleja la compleja interacción entre el ovocito y las demás células que lo acompañan (Lynch, 2019). Se establecen diferencias entre los polos opuestos del ovocito, lo que permite establecer la polaridad y definir los ejes corporales que definen la anatomía del futuro embrión (Roth, 2003). En este contexto, los genes involucrados en dicho proceso pueden clasificarse según su función en los relacionados a la progresión de la ovogénesis y desarrollo del ovocito, y los genes responsables del establecimiento del eje anteroposterior (A-P) y del eje dorsoventral (D-V) del futuro embrión (Manseau & Schüpbach, 1989; Nüsslein-Volhard & Roth, 1989; Hay *et al.*, 1990; Kelley, 1993; Goode *et al.*, 1996; Cox *et al.*, 2000).

4.1.3 Olfacción

La detección de señales químicas del medio ambiente, se da por medio de la olfacción (Chapman, 1998; Guidobaldi & Guerenstein, 2012; Liu *et al.*, 2022), y el circuito neural específico activado es el determinante de la respuesta conductual (Stocker *et al.*, 1983; Carlson, 1996). Los insectos se basan en gran medida en sus estímulos olfatorios tanto para la búsqueda de alimento como de pareja (Guidobaldi & Guerenstein, 2012). El reconocimiento de los olores comienza en las sensilas olfativas, las cuales poseen células sensoriales especializadas que transmiten

señales químicas hacia el sistema nervioso central (Levy *et al.*, 1991; Parmentier *et al.*, 1992; Ben-Arie *et al.*, 1994; Liu *et al.*, 2022).

Los odorantes externos son reconocidos por proteínas de unión a olores (OBP) y proteínas quimiosensoriales (CSP), los cuales son transferidos a través de la linfa de cada sensila a los receptores quimiosensoriales, receptores del olor (OR) (Smart *et al.*, 2008; Liu *et al.*, 2022) y receptores ionotrópicos (IR), los cuales son los encargados de otorgar especificidad a la respuesta (Vosshall *et al.*, 1999; Galizia & Sachse, 2010). Junto a los OR se encuentra uno de los co-receptores más conservados, el co-receptor del receptor olfativo (Orco), esencial para desencadenar la respuesta (Harini & Sowdhamini, 2012; Liu *et al.*, 2022).

Aunque el circuito de olfacción básico se conserva entre especies, la diversidad en los repertorios génicos presentes provoca que las respuestas olfatorias varíen, adaptándose a sus diferentes nichos ecológicos y comportamientos específicos (Hallem & Carlson, 2006; Bohbot *et al.*, 2007; Bohbot *et al.*, 2011). En este sentido, estudios en delfácidos como *Sogatella furcifera* y *Nilaparvata lugens* evidencian que esta diversificación génica se asocia a sistemas olfatorios especializados, orientados principalmente a la detección de volátiles de plantas hospedadoras, lo que refleja una adaptación funcional a su ecología específica (Hu *et al.*, 2019; Liu *et al.*, 2022).

Zhou *et al.*, (2014) analizaron el transcriptoma *de novo* de las antenas de *N. lugens* pudiendo identificar varios genes relacionados con la olfacción, entre ellos OBPs, IRs, ORs y CSPs, los cuales permiten al insecto detectar compuestos volátiles emitidos por plantas hospedadoras y feromonas. Esto sugiere que el sistema olfativo de este delfácido está altamente especializado para identificar su planta hospedadora de preferencia, detectar señales químicas del entorno y facilitar comportamientos claves como la alimentación, dispersión y reproducción.

En base a lo explicado hasta el momento, en este trabajo nos propusimos generar información transcriptómica de *D. kuscheli*, e inferir la identidad de los genes involucrados en algunos de los procesos más importantes para la supervivencia del insecto: ecdisis, ovogénesis y olfacción. El conocimiento de la genética de *D. kuscheli* no sólo permitirá esclarecer aspectos biológicos fundamentales de la especie, sino que también podría ser crucial para diseñar nuevas estrategias destinadas a reducir la transmisión del virus del Mal de Río Cuarto.

Hipótesis Y Objetivos

Hipótesis

Delphacodes kuscheli presenta un repertorio genético característico, cuyo correcto funcionamiento es esencial para su desarrollo y supervivencia, y se encuentra conservado evolutivamente.

Objetivo General

- Generar información genética relevante sobre *Delphacodes kuscheli* mediante la caracterización de su transcriptoma.

Objetivos Específicos

1. Generar un transcriptoma representativo de *D. kuscheli* considerando los diferentes estados de desarrollo que conforman su ciclo de vida.
2. Evaluar la calidad y completitud del transcriptoma generado.
3. Realizar la anotación génica, en base a similitudes con otras especies de insectos, de las familias de genes involucradas en los procesos vitales para el desarrollo del insecto:
 - ✓ Ecdisis
 - ✓ Ovogénesis
 - ✓ Olfacción

Materiales

Y

Métodos

1. Mantenimiento de la colonia de *Delphacodes kuscheli*

Los ejemplares de *D. kuscheli* se obtuvieron a partir de una colonia establecida en el Centro de Bioinvestigaciones (UNNOBA-CICBA), la cual fue mantenida en condiciones controladas con un fotoperíodo de 16:8 hs luz: oscuridad a una temperatura de 24°C y una humedad relativa de 70%. Los insectos se colocaron junto a su planta hospedera preferencial (avena) en jaulas de varillas de aluminio cubiertas con *voile*. En estas condiciones, su ciclo de vida pasa por una etapa de huevo, con una duración promedio de 7 a 10 días; cinco estadios ninfales: I, II, III, IV y V, los cuales poseen una duración promedio de 4; 2,5; 3; 4 y 5,5 días respectivamente. Luego emergen los adultos, que presentan un período de pre-oviposición de unos 4 días aproximadamente, y una longevidad de 1 a 3 semanas (Virla & Remes Lenicov, 1991).

2. Extracción de ARN

La extracción de ARN a partir de los ejemplares de *D. kuscheli*, se llevó a cabo seleccionando muestras independientes para cada etapa del ciclo de vida: huevo, ninfa I, ninfa II, ninfa III, ninfa IV, ninfa V y ejemplares adultos (tanto machos como hembras), con el objetivo de garantizar una representación completa de todo el ciclo de vida. La identificación de los estadios ninfales se basó en Remes Lenicov *et al.*, (2008). Se colectó una cantidad aproximada de 10 individuos adultos y un volumen similar de los distintos estadios ninfales (el número de ejemplares varía de acuerdo al tamaño del estadio). Esto se realizó en base a que sabíamos que por extracciones previas esto representaba un volumen suficiente para la obtención de la cantidad requerida de ARN.

Para la extracción de ARN total, se utilizó el reactivo comercial TRIzol (*Thermo Fisher Scientific*), siguiendo las especificaciones recomendadas por el fabricante. El ARN total se cuantificó por fluorescencia utilizando QUBIT 2.0 (Invitrogen), su integridad se evaluó mediante un gel de agarosa 1 % en *Buffer* TBE (1x). La extracción de ARN se realizó de manera independiente para cada estadio, y luego se tomó 1 µg total de cada uno de ellos para generar un único molde de secuenciación.

3. Secuenciación de ARN

La biblioteca fue preparada a través del servicio de secuenciación de MacroGen (Corea) con el *kit* de preparación de muestras *TruSeq Stranded mRNA* LT de Illumina. La misma fue

secuenciada con la plataforma NovaSeq 6.000 de Illumina. Los parámetros establecidos para la secuenciación fueron lecturas del tipo *paired-end* (R1: *forward* y R2: *reverse*) de 100 pb de longitud.

4. Análisis Bioinformático

4.1 Análisis de calidad y limpieza de las lecturas

Las lecturas crudas obtenidas en formato fastq fueron analizadas con el programa FastQC v.0.11.9 (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Se observó el número total de secuencias obtenidas y su longitud final.

La búsqueda de secuencias adaptadoras y contaminantes se realizó utilizando la herramienta Blastn v.2.9.0+ (Altschul *et al.*, 1990) ejecutada desde un entorno de sistema operativo Ubuntu 20.04 LTS, (Anexo I: Comando 1). Esto consistió en una búsqueda por similitud entre las lecturas crudas y la base de datos de nucleótidos UniVec, descargada de NCBI (<ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec>). Como resultado, se obtuvo una lista de secuencias que dieron *hit*, indicando con qué longitud y en qué posición. El siguiente paso consistió en ejecutar un *script* escrito en lenguaje *Python* (Anexo I: Comando 2), para eliminar de nuestras lecturas las secuencias contaminantes encontradas. Se estableció como parámetros de remoción la longitud de la similitud y posición a lo largo de la lectura; si el *hit* se encontraba dentro de las primeras 40 pb, estas eran eliminadas conservando el resto de la secuencia, en cambio, si el *hit* se encontraba dentro de las últimas 30 pb, se eliminaban conservando las primeras 70 pb, y, por último, si el *hit* se encontraba entre las bases 40 y 70 pb, se eliminaba la secuencia completa. Al eliminar secuencias en uno de los dos archivos procesados, se generaba una ruptura en el emparejamiento entre estos (aparición de lecturas *singles*), ya que quedaba la secuencia recíproca en el otro archivo sin su pareja. Para sincronizarlos, se utilizó la herramienta SeqKit v2.5.1 (Shen *et al.*, 2024) (Anexo I: Comando 3). Luego, mediante el *software* Trimmomatic v.0.39 (Bolger *et al.*, 2014) ejecutado desde el mismo entorno (Anexo I: Comando 4), se realizó un último paso de filtrado y limpieza para asegurar que los archivos R1 y R2 quedaran sin ningún tipo de contaminantes. Los parámetros de *trailing*, *leading*, *minlen* e *Illuminaclip* utilizados para su ejecución fueron los recomendados por Bolger *et al.*, (2014), para descartar lecturas de muy baja calidad, de corta longitud o adaptadores remanentes. Además, se incluyó el parámetro *keepBothReads* para mantener la condición *paired-end*. Mediante FastQC se corroboró nuevamente la calidad de las secuencias obtenidas (Anexo II: Figura 1), considerándose como

métrica de buena calidad valores iguales o mayores a un puntaje de calidad Q30, el cual se determina a partir de la probabilidad predicha de incorporar una base incorrecta.

4.2 Ensamble de *nov*

Las lecturas limpias fueron utilizadas para la reconstrucción de los transcriptos mediante un ensamble de *nov*, utilizando el programa Trinity v.2.15.1, estableciendo datos *paired-end*, el parámetro *Strand specific data* activado en el modo *Reverse-Forward* y el parámetro mínimo de cobertura (*min_kmer_cov*) de 2 (Grabherr *et al.*, 2011). Quedando el resto de los parámetros por *default*.

4.3 Análisis de la calidad y completitud del ensamble

Para remover la redundancia en los transcriptos ensamblados, se utilizó un *script* en lenguaje *perl* del paquete Trinity (https://github.com/trinityrnaseq/trinityrnaseq/blob/master/util/misc/get_longest_isoform_seq_per_trinity_gene.pl), el cual selecciona y conserva la isoforma más larga por cada gen/transcripto definido por Trinity como secuencia representativa. Este procedimiento permitió generar un ensamble no redundante, compuesto por transcriptos únicos, reduciendo y/o eliminando la presencia de secuencias asociadas a la fragmentación del ensamble y de posibles artefactos de secuenciación, así como también eliminando isoformas alternativas al conservar únicamente la secuencia más larga.

Se utilizó el programa Fasta Statistics v.2.0. (Seemann & Gladman, 2012), para reportar y comparar la estadística característica de cada ensamble. Por otro lado, se evaluó la calidad y el grado de completitud de ambos transcriptomas generados, utilizando el programa BUSCO v.5.4.6 (Simão *et al.*, 2015), analizando la ortología a nivel de Eucariota, Arthropoda, Insecta y Hemiptera en cada caso. A partir de este paso, sólo se siguió trabajando con el ensamble formado por transcriptos únicos.

Para predecir las regiones codificantes de los transcriptos reconstruidos se utilizó TransDecoder v.5.5.0 (Haas *et al.*, 2013), el cual se ejecutó con todos los parámetros por *default*, excepto *strand-specificity*, el cual fue activado.

Se realizó una búsqueda de similitud de secuencia utilizando la herramienta heurística Blastx v.2.10.1 y Blastp v.2.10.1 (Altschul *et al.*, 1997). Para ello se emplearon diferentes bases de datos obtenidas desde los repositorios UniProt (2023, <https://www.uniprot.org/>) e InsectBase

2.0 (2023, <https://doi.org/10.1093/nar/gkab1090>). En particular, de Uniprot se descargaron las bases UniProt total (569.793 secuencias), Uniprot Arthropoda (15.579 secuencias) y Uniprot Insecta (11.019 secuencias), con el fin de abarcar un amplio espectro de ortologías. Complementariamente, se utilizaron las proteínas predichas de especies filogenéticamente cercanas, todas ellas insectos plagas y con su genoma secuenciado: *Laodelphax striatellus* (46.351 secuencias), *Nilaparvata lugens* (24.901 secuencias) y *Sogatella furcifera* (20.978 secuencias), obtenidas de InsectBase 2.0, con el fin de explorar la similitud de secuencia entre especies de la familia Delphacidae. Cada búsqueda se realizó estableciendo que el valor mínimo de cobertura de los *hits* fuera del 30% y el punto de corte del *e-value* $\leq 1 \times 10^{-5}$.

5. Caracterización del ensamble de novo

Las predicciones proteicas se caracterizaron mediante el programa Hmmscan v.3.3.2 del paquete HMMER (Eddy, 1995; Finn *et al.*, 2014), utilizando la base de datos *Pfam* (2015, <https://pfam.xfam.org/>). Además, se analizaron para predecir la secuencia de péptido señal y regiones transmembrana, utilizando SignalP v.5.0 (Almagro Armenteros *et al.*, 2019) desde Ubuntu 20.04 LTS (Anexo I: Comando 5) y TMHMM 2.0 v.0.0.17 (Cock *et al.*, 2013).

Para realizar la anotación funcional del transcriptoma generado se utilizó el *software* EggNOG en su versión *on line* (<http://eggno5.embl.de/#/app/home>), que permite determinar las diferentes funciones y términos GO correspondiente a cada transcripto.

En base a los reportes obtenidos se utilizó el paquete bioinformático de anotación Trinotate v.3.2.2 (Grabherr *et al.*, 2011). El programa se ejecutó de manera tal que se incluyeron en la anotación los resultados obtenidos tras la contrastación con las diferentes bases de datos previamente mencionadas de Uniprot (Anexo I: Comando 6), el cual permitió generar un reporte integral del análisis realizado sobre el transcriptoma reconstruido.

Para visualizar los resultados en una interfaz gráfica se implementaron *scripts* en lenguaje R (Anexo I: Comando 7), que permitieron observar la distribución de los géneros más representados en cada búsqueda realizada.

6. Anotación Génica

Para realizar la anotación génica dirigida a procesos específicos, se utilizaron bases de datos proteicas, previamente utilizadas y desarrolladas en el laboratorio (Andrada, 2021; Crespo, 2021; Pascual & Rivera-Pomar, 2022). Estas bases de datos no representan la totalidad de los

genes involucrados en cada proceso, sino un conjunto de genes clave de referencia anotados. Todas las secuencias utilizadas se encuentran identificadas mediante sus respectivos números de acceso en las bases de datos de origen (NCBI, UniProt, Flybase o VectorBase), permitiendo su trazabilidad (Anexo III- Ecdisis, Ovogénesis y Olfacción)

La base de datos referente a la ovogénesis se compone de 2.583 secuencias correspondientes a la especie *Drosophila melanogaster* (Pascual & Rivera-Pomar, 2022).

En cuanto a la ecdisis, se amplió la base de datos generada previamente (Andrada, 2021) con secuencias proteicas obtenidas de NCBI (<https://www.ncbi.nlm.nih.gov/protein>) correspondientes a múltiples órdenes de insectos. Debido a la limitada disponibilidad de anotaciones en Hemiptera, la inclusión de especies filogenéticamente diversas permitió aumentar la cobertura y sensibilidad en la detección de genes conservados. Se incluyeron insectos pertenecientes a los órdenes Blattodea, Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera, Orthoptera y Thysanoptera. Esta base de datos quedó conformada por 114 secuencias correspondientes a proteínas clave del proceso: *Burs*, *Crz*, *EH*, *ETH*, *Pburs* y *CCAP*.

Para la base de datos de olfacción, se incorporaron nuevas anotaciones de especies hemípteras obtenidas de NCBI y VectorBase (<https://vectorbase.org/>), con el objetivo de mejorar la representación de este orden. Dichas anotaciones correspondieron a las especies: *Acyrtosiphon pisum*, *Adelphocoris lineolatus*, *Aphis gossypii*, *Apolygus lucorum*, *Cyrtorhinus lividipennis*, *Diaphorina citri*, *Halyomorpha halys*, *Laodelphax striatellus*, *Lygus lineolaris*, *Nasonia vitripennis*, *Nezara viridula*, *Nilaparvata lugens*, *Rhodnius prolixus* y *Sogatella furcifera*. La base de datos final incluyó 342 secuencias correspondientes a las familias IR, OR, Orco y OBP.

Utilizando las diferentes bases de datos, se realizó una búsqueda de similitud mediante Blastx, estableciendo un umbral de $e\text{-value} \leq 1 \times 10^{-5}$, y dejando el resto de los parámetros por *default*.

Los programas Trinity, Filter Fasta, BUSCO, Fasta *Statistics*, TransDecoder, Blast, Hmmscan, TMHMM y Trinotate fueron utilizados desde la plataforma web Galaxy (<https://usegalaxy.org/> y <https://usegalaxy.eu/>).

Resultados

1. Análisis Bioinformático

1.1 Análisis de calidad y limpieza de las lecturas

Como resultado de la secuenciación de ARN (ARN-Seq), se obtuvieron dos archivos, R1: *forward* y R2: *reverse*, dado por la naturaleza de la secuenciación *paired-end*. Ambos comprimidos en formato .gz, de un peso de 5.1 GB.

El análisis realizado a través de FastQC permitió identificar la presencia de 105.896.152 lecturas totales en cada uno de los archivos generados, todas con una longitud de 101 pb. Además, se observó un porcentaje de GC (Guanina-Citocina) del 42% para R1 y 43% para R2. Pese a la distribución normal esperada se detectaron picos adicionales, los cuales evidenciarían la presencia de secuencias contaminantes (Figura 2 A-B).

Por su parte, los reportes de calidad de secuencia por base indicaron valores superiores a Q30 a lo largo de toda la longitud de secuencia, lo que predice que la probabilidad de que se haya añadido una base errónea en las lecturas fue menor a 1 en 1.000 pb (Figura 2 C-D). Esto se complementa con la obtención de lecturas sin presencia de N en sus secuencias (Figura 3 A-B). Por último, se observa la ausencia de secuencias adaptadoras reconocidas por FastQC en las lecturas generadas (Fig. 3 C-D).

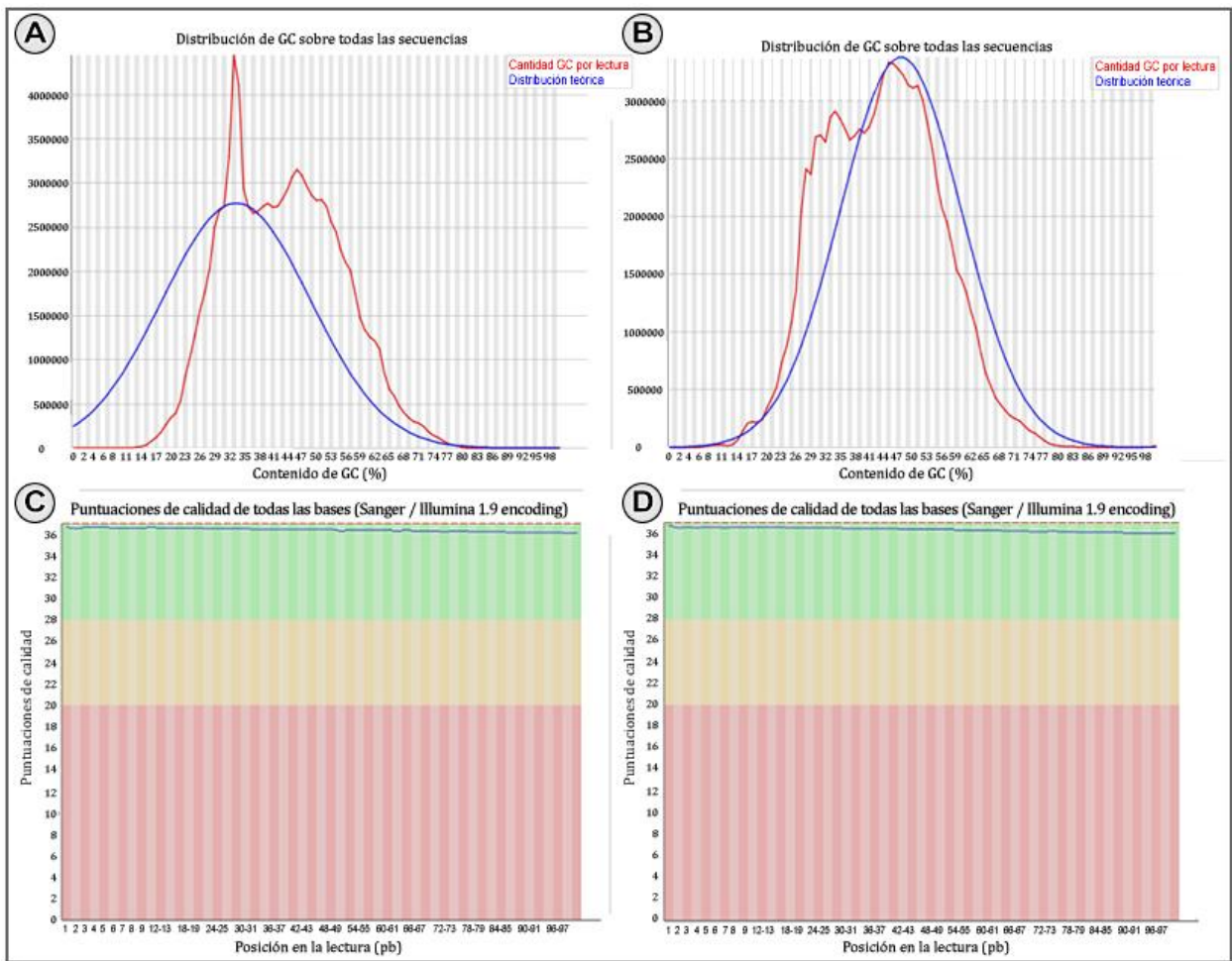


Figura 2. Análisis de calidad de las lecturas obtenidas mediante FastQC. Contenido de GC por lectura (rojo) vs. la distribución normal modelada (azul), para *forward* (A) y *reverse* (B). En el eje X se representa el porcentaje de contenido GC. En el eje Y se representa la distribución de GC sobre las secuencias. Distribución de *Phred Score* en función de la posición a lo largo de cada lectura, para *forward* (C) y *reverse* (D). En el eje X se representan las distintas posiciones de la lectura en pares de bases. En el eje Y se representan los valores de calidad posibles, abarcando valores de 0 a 36 (Valor Q).

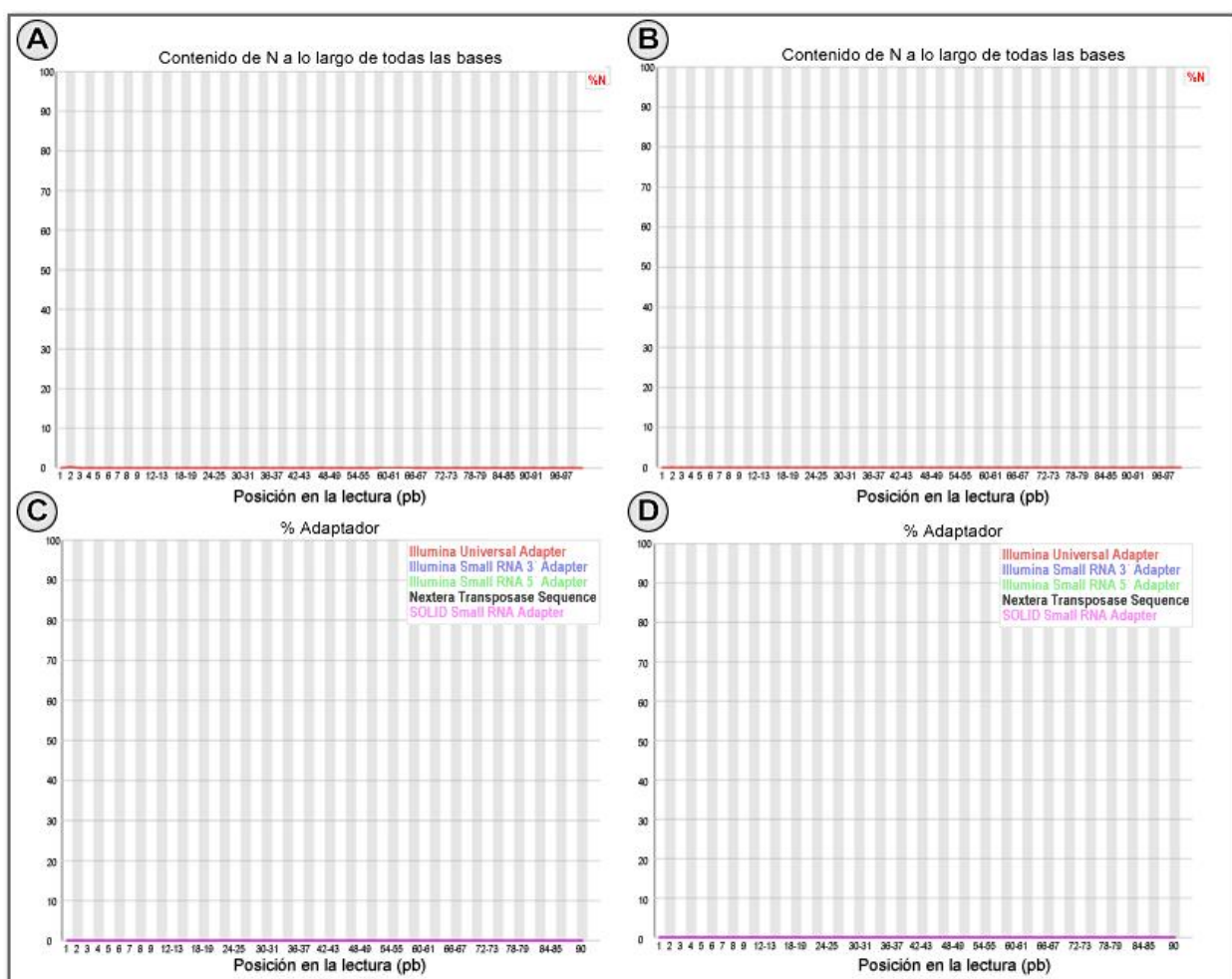


Figura 3. Análisis de calidad de las lecturas obtenidas mediante FastQC. Porcentaje de N presente a lo largo de la secuencia, para *forward* (A) y *reverse* (B). En el eje Y se muestra el % de N, y en el eje X se representan las distintas posiciones de las lecturas en pares de bases. Distribución porcentual de la presencia de adaptadores en las lecturas, para *forward* (C) y *reverse* (D). En el eje Y se muestra su porcentaje, y en el eje X se representan las distintas posiciones de las lecturas en pares de bases.

Se complementó la búsqueda de secuencias adaptadoras utilizando la base de datos de UniVec, lo que permitió identificar 15.244 y 15.628 lecturas contaminantes en R1 y en R2, respectivamente, las cuales no habían sido reportadas por FastQC. Estas lecturas fueron eliminadas de los sets de datos, generando lecturas no apareadas ("*singles*"), las cuales también fueron descartadas mediante SeqKit. Como resultado, se obtuvieron 105.875.767 lecturas finales para cada uno de los archivos (Tabla 1), que se utilizaron como material de partida para proseguir con el correspondiente ensamble del transcriptoma.

Tabla 1. Datos de secuenciación. Lecturas crudas y lecturas finales obtenidas tras la eliminación de secuencias inespecíficas. Entre paréntesis, se indica el porcentaje de retención tras el procesamiento realizado.

Lecturas Crudas	Lecturas Finales
R1 105.896.152 (100%)	R1 105.875.767 (99,98%)
R2 105.896.152 (100%)	R2 105.875.767 (99,98%)

1.2 Ensamble de *novο* y análisis de calidad

La reconstrucción del transcriptoma generó un total de 140.247 transcriptos totales, de los cuales 69.314 fueron transcriptos únicos, con una longitud mayor a 200 pb. Ambos conjuntos fueron analizados con el fin de reportar y comparar las métricas de calidad del ensamble.

En el conjunto total de transcriptos generados, el contenido de GC fue del 37,20%, mientras que en los transcriptos únicos alcanzó el 37,72%. La longitud media obtenida fue de 808 pb en el conjunto total y 669 pb en los únicos. Del mismo modo, el valor de N50, como estadístico de calidad del ensamble, fue de 1.295 nt y 1.121 nt, respectivamente (Tabla 2).

El grado de completitud del transcriptoma mediante el análisis del porcentaje de representatividad de ortólogos reportó valores superiores al 80% para cada una de las cuatro bases de datos consultadas (Tabla 2, Figura 4). En particular, se observó un mayor porcentaje de ortólogos de copia única al realizar la búsqueda en el conjunto de transcriptos únicos. En contraste, el análisis realizado con el *set* completo de transcriptos evidenció un elevado número de ortólogos duplicados con porcentajes cercanos al 50% en cada base consultada. Además, se observó que el número de ortólogos identificados aumentó con la especificidad del nivel evolutivo considerado (Figura 4).

Cabe aclarar que, en base a las métricas obtenidas, se prosiguió a realizar los análisis subsiguientes considerando el número de transcriptos únicos reconstruidos.

Tabla 2. Métricas de calidad del ensamble de *novo*. Se muestran las estadísticas básicas (% GC, longitud media y N50) correspondiente a los conjuntos de transcritos totales y únicos. Además, se presentan los porcentajes de ortólogos identificados mediante BUSCO en cuatro niveles filogenéticos: Eucariota, Arthropoda, Insecta y Hemiptera. (nt = nucleótidos).

	Fasta Statistics			BUSCO (%)			
Ensamble	%GC	Longitud media (nt)	N50 (nt)	Eucariota	Arthropoda	Insecta	Hemiptera
Transcriptos totales	37,20	808	1.295	91,00	87,30	83,90	80,10
Transcriptos únicos	37,72	669	1.121	88,70	85,90	82,40	78,50

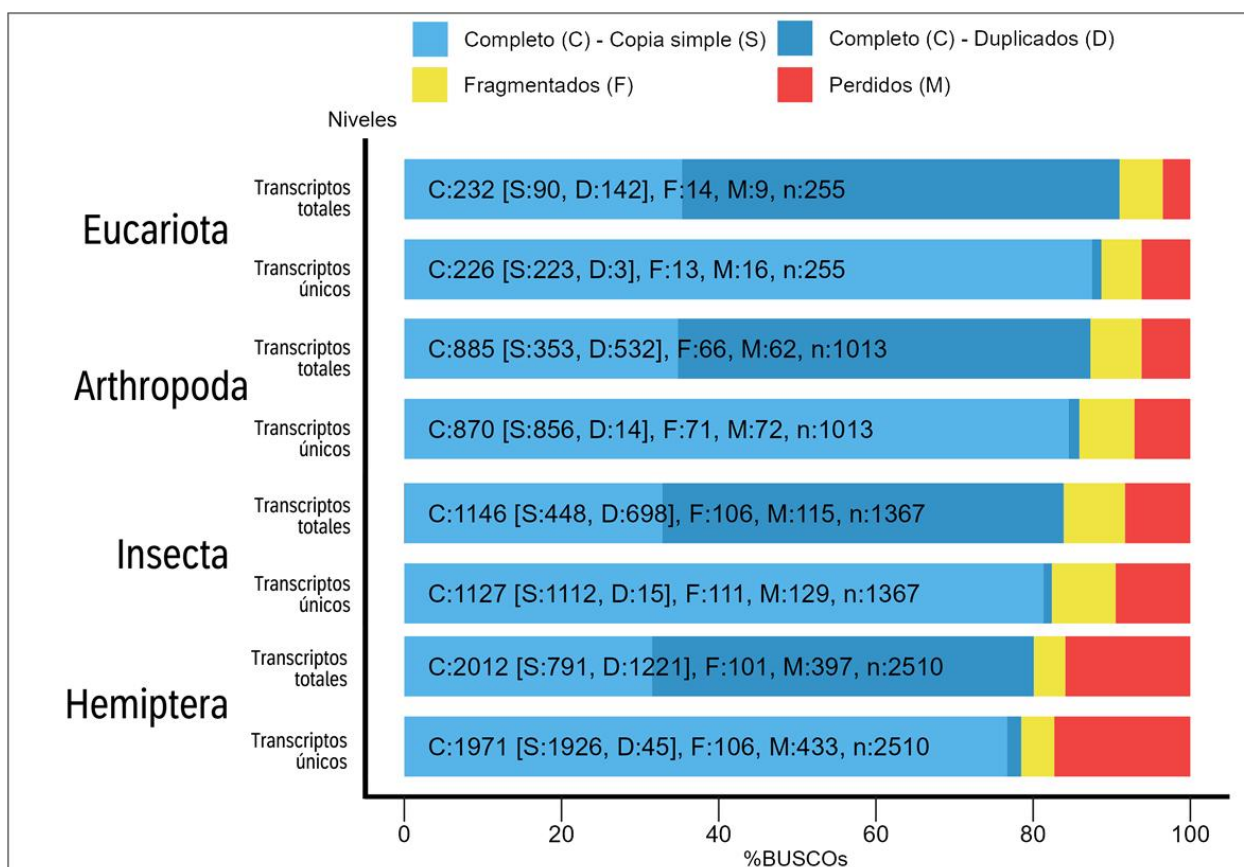


Figura 4. Análisis de completitud del ensamble transcriptómico mediante BUSCO. Se muestra el porcentaje de ortólogos BUSCO identificados como completos (de copia simple y duplicados), fragmentados o perdidos, a partir de los conjuntos de transcritos totales y únicos, en distintos niveles filogenéticos: Eucariota, Arthropoda, Insecta y Hemiptera.

Por otro lado, se realizó un análisis de ortología, tanto a nivel de transcriptos como de predicciones proteicas (18.607 secuencias totales predichas), utilizando las bases de datos Uniprot, Uniprot Arthropoda y Uniprot Insecta. En la Tabla 3, se muestran las métricas obtenidas para cada una. En cuanto a la búsqueda de predicciones proteicas, se identificaron 7.302 hits contra la base de datos Uniprot, 6.860 contra Uniprot Arthropoda y 6.820 contra Uniprot Insecta. En cambio, cuando la búsqueda de predicciones proteicas se realizó a nivel nucleótido, se reportó un mayor número de hits, con 8.135, 7.865 y 7.848 transcriptos en cada base de datos, respectivamente. Respecto a la búsqueda específica entre los delfácidos, se identificaron a nivel proteico 15.027 transcriptos por ortología contra *L. striatellus*, 14.060 transcriptos contra *N. lugens* y 13.810 contra *S. furcifera*. De igual manera, se observó un mayor número de transcriptos identificados cuando la búsqueda se realizó a nivel nucleótido, alcanzando 21.147, 19.438 y 20.376 transcriptos respectivamente (Tabla 3).

Los porcentajes de similitud de secuencia en base a las predicciones proteicas (18.607 secuencias totales predichas – 100%) que reportó *D. kuscheli* frente a las bases de datos de Uniprot (39,24 %), Arthropoda (36,87 %) e Insecta (36,65 %) resultaron similares entre sí. Asimismo, se observaron valores próximos al compararlo con otros delfácidos, alcanzando un 80,76 % con *L. striatellus*, 75,56 % con *N. lugens* y 74,22 % con *S. furcifera*, evidenciando un porcentaje de similitud notablemente más alto al comparar directamente con especies del mismo grupo (Tabla 3).

Tabla 3. Análisis de ortología. Número de Blast *Hits* únicos obtenidos al realizar la búsqueda de ortólogos a nivel de transcriptos (Blastx) y predicciones proteicas (Blastp) contra las bases de datos Uniprot, Arthropoda e Insecta. A nivel de delfácidos, se utilizaron como bases de datos las especies *L. striatellus*, *N. lugens* y *S. furcifera*.

		<i>Delphacodes kuscheli</i> vs.					
		Uniprot	Uniprot Arthropoda	Uniprot Insecta	<i>L. striatellus</i>	<i>N. lugens</i>	<i>S. furcifera</i>
<i>Hits</i>	<i>Blastp</i>	7.302	6.860	6.820	15.027	14.060	13.810
	<i>Blastx</i>	8.135	7.865	7.848	21.147	19.438	20.376

En cuanto a los géneros más representados en la búsqueda en la base de datos Uniprot, se identificaron Homo, Sophophora y Mus (pertenecientes al orden Primates, Diptera y Rodentia,

respectivamente). En las búsquedas en Uniprot Arthropoda y Uniprot Insecta, el género más representado fue *Drosophila*, seguido por *Anopheles* y *Aedes* (pertenecientes al orden Diptera) (Figura 5 A, B, C).

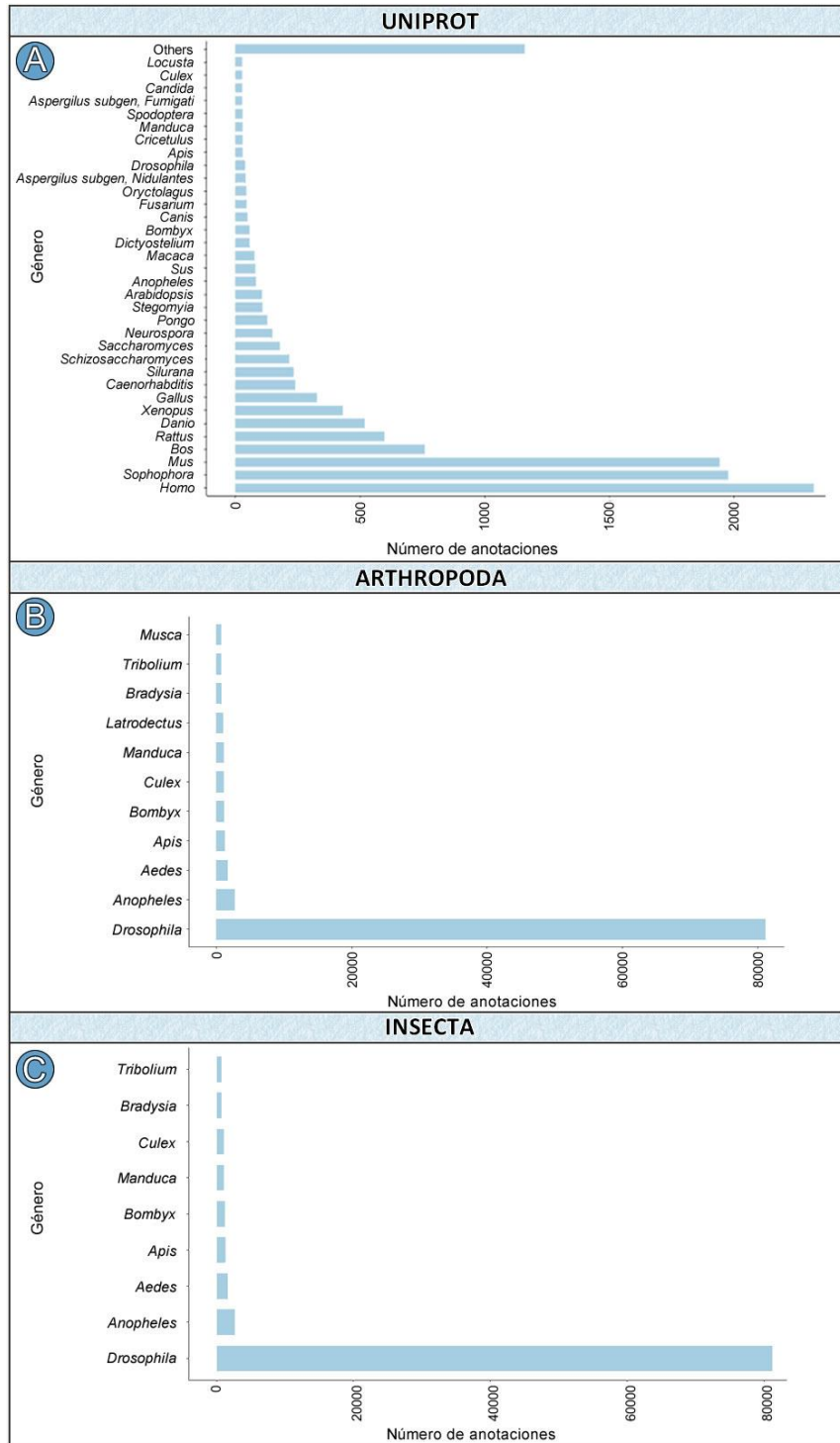


Figura 5. Correspondencia de géneros. Transcripts más frecuentes anotados a nivel género mediante Blastp y Blastx, contra Uniprot (A), Arthropoda (B) e Insecta (C). Se observa en el eje X el número de transcripts correspondiente a de cada uno de los géneros representados en el eje Y.

Además, se realizó una anotación funcional utilizando la base de datos de ontología génica (GO), estructurando las identidades en tres categorías principales. Dentro de procesos biológicos (Figura 6), se destacaron la regulación de la transcripción por la ARN polimerasa II (GO: 0006357), regulación positiva de la transcripción por la ARN polimerasa II (GO: 0045944), desarrollo de organismos multicelulares (GO: 0007275), la división celular (GO: 0051301), transporte de proteínas (GO: 0044861, GO: 0032596) y la regulación negativa de la transcripción de ADN por ARN polimerasa II (GO: 000122). Respecto a los componentes celulares (Figura 7), los términos más representados fueron los componentes citoplasmáticos (GO: 0005737), nucleares (GO: 0005634) y del citosol (GO: 0005829), así como los componentes integrales de la membrana (GO: 0016021), membrana plasmática (GO: 0005886), nucleoplasma (GO: 0005654) y mitocondrias (GO:0005739), como los más frecuentes. Por último, en la categoría de funciones moleculares (Figura 8), predominaron aquellas relacionadas con la unión de iones metálicos (GO: 0046872), unión de ATP (GO: 0005524), de ARN (GO: 0003723) y de ADN (GO: 0003677), de proteínas idénticas (GO: 0042802) y de iones de zinc (GO: 0008270). En suma, se realizó una caracterización del transcriptoma en base a la categorización obtenida con EggNOG, donde los transcritos se agruparon en un total de 24 categorías. La correspondiente a función desconocida (S) fue la más asignada, seguida por las categorías de mecanismo de transducción de señales (T); modificación postraduccional, recambio proteico, chaperonas (O); transcripción (K); traducción, estructura ribosómica y biogénesis (J); tráfico intracelular, secreción y transporte vesicular (U) (Figura 9).

Además, en las predicciones proteicas se identificaron 10.017 dominios proteicos, 4.370 proteínas transmembranas y 7.843 péptidos señal.

En la tabla 4 se muestran las métricas y características biológicas del transcriptoma generado. En conjunto, estos proporcionan una base sólida para la inferencia y anotación génica del transcriptoma reconstruido.

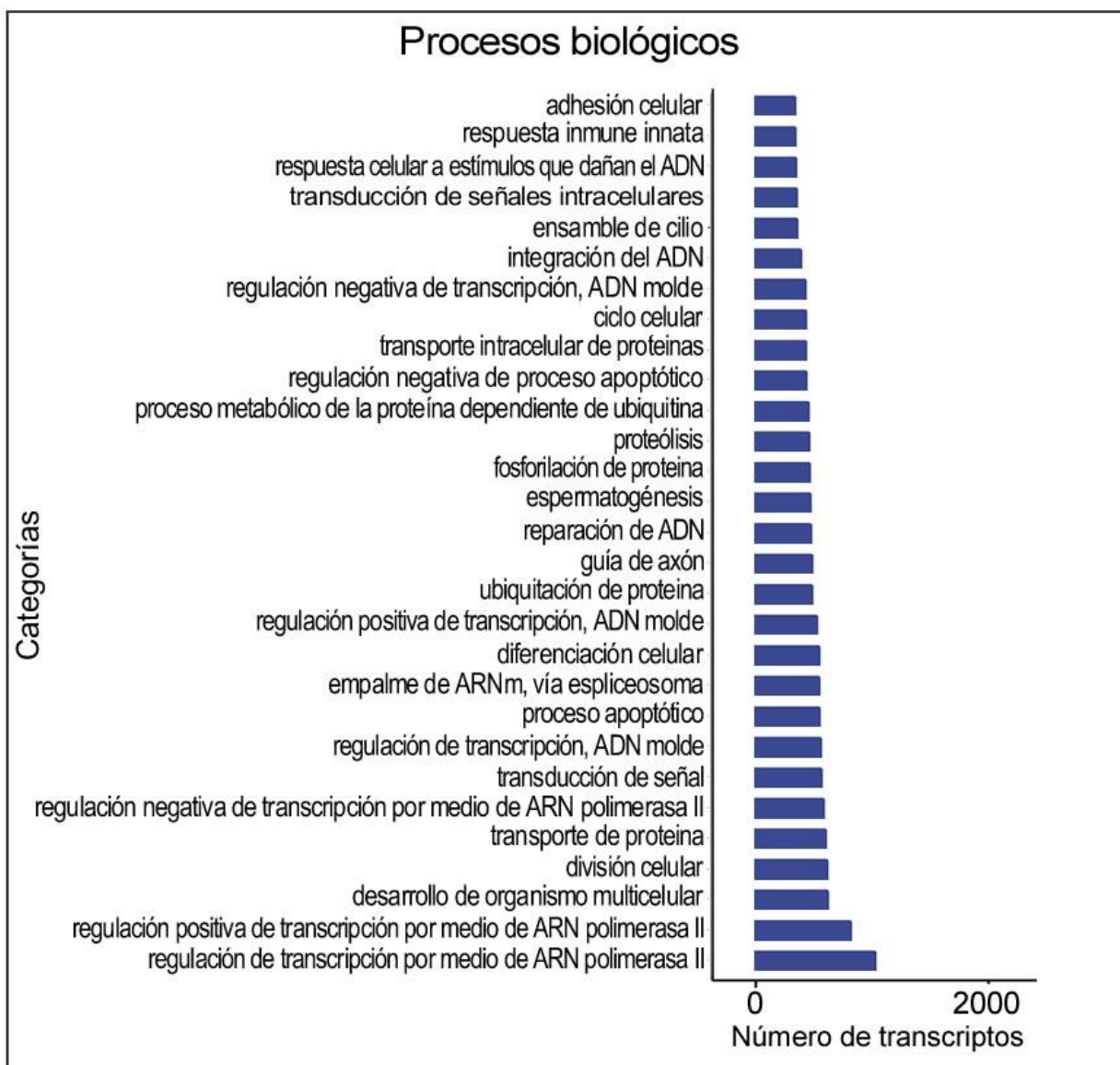


Figura 6. Anotación funcional. Anotación ontológica mediante la base de datos *Gene Ontology* (GO), en función de los procesos biológicos al que pertenecen los transcritos. En el eje X, se representa la cantidad de transcritos que pertenecen a cada uno de los términos GO asociados a los procesos biológicos representados en el eje Y.

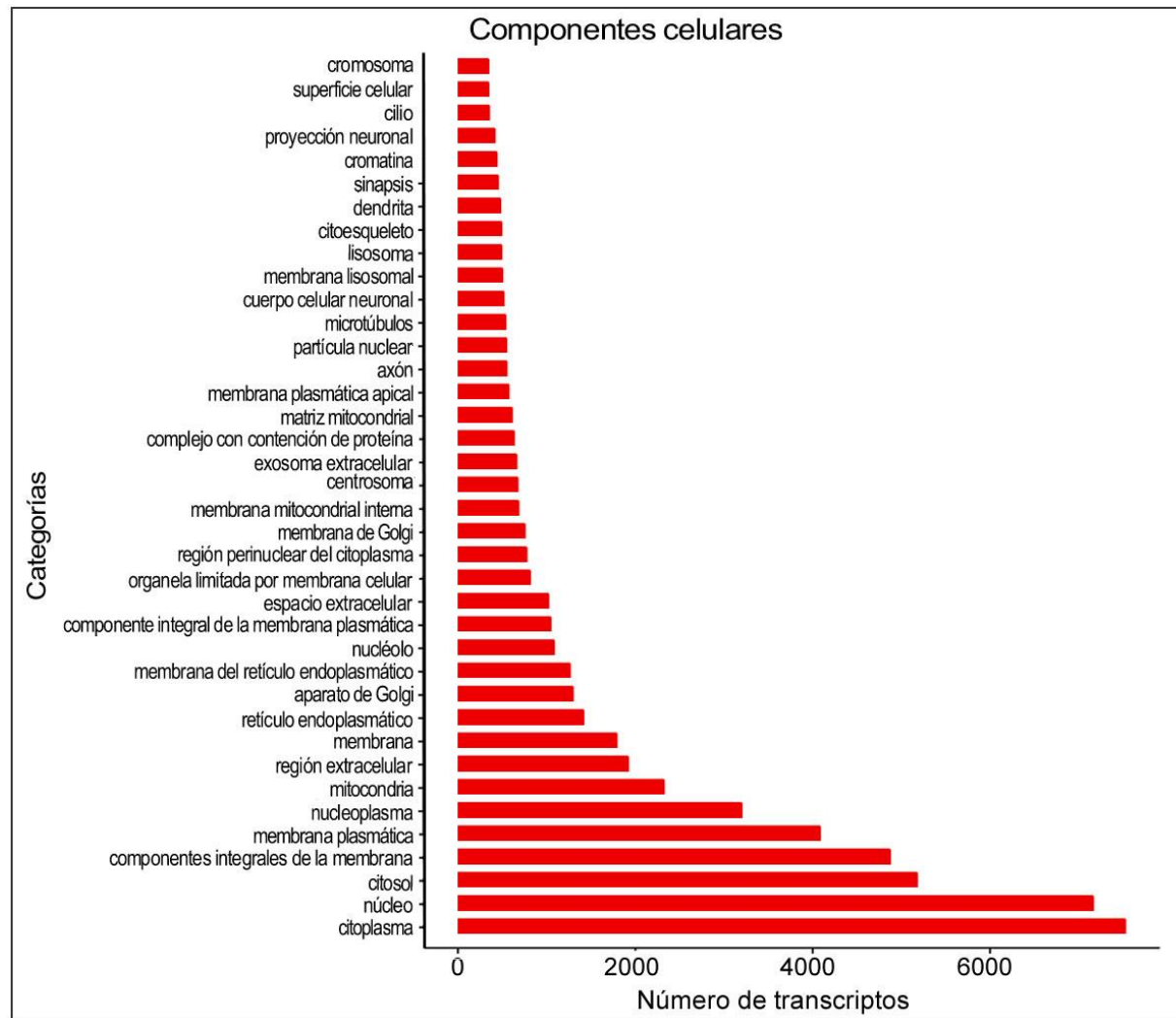


Figura 7. Anotación funcional. Anotación ontológica mediante la base de datos *Gene Ontology* (GO), en función de los componentes celulares al que pertenecen los transcritos. En el eje X se representa la cantidad de transcritos que pertenecen a cada uno de los términos GO asociados a los componentes moleculares representados en el eje Y.

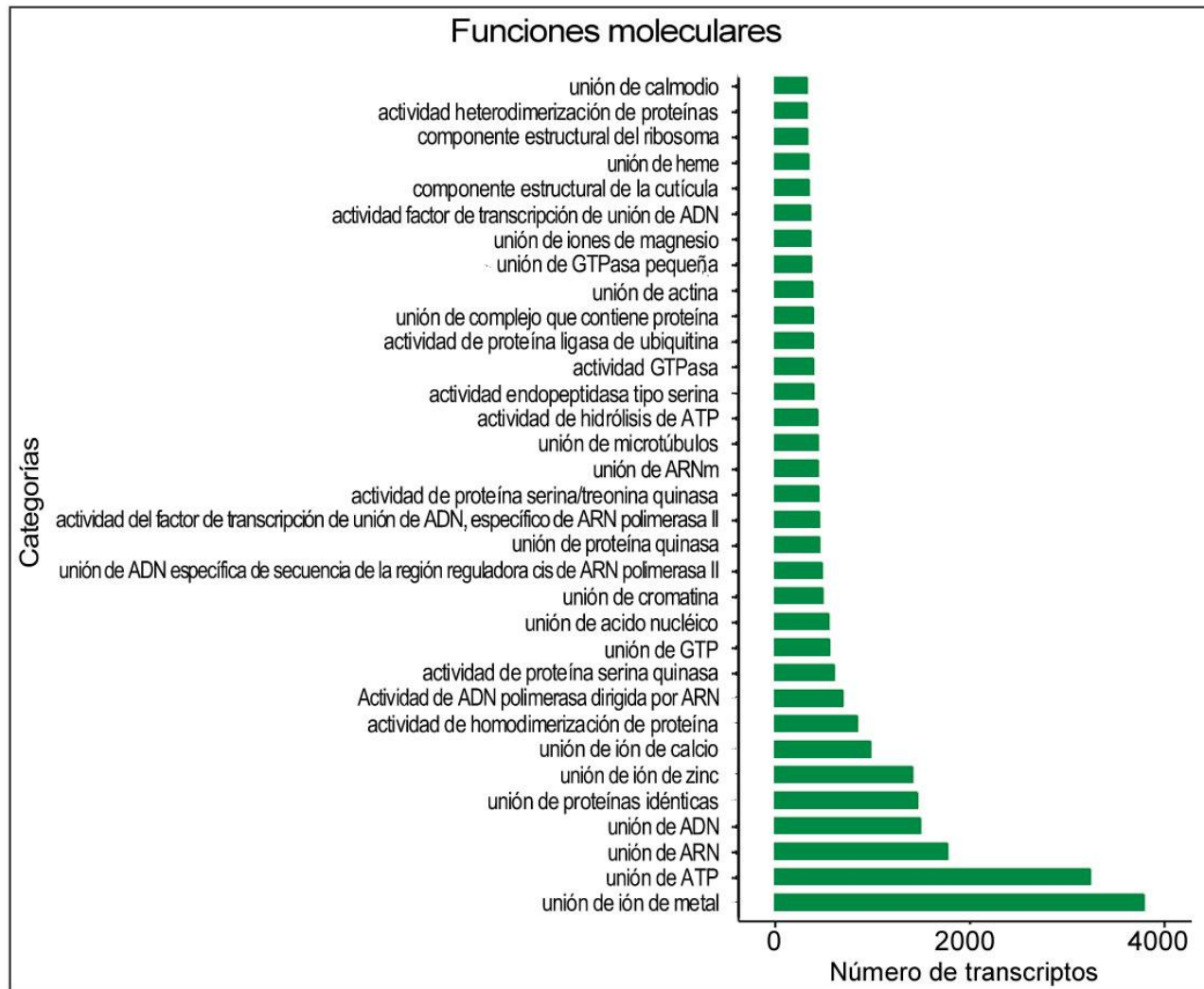


Figura 8. Anotación funcional. Anotación ontológica mediante la base de datos *Gene Ontology* (GO), en función de las funciones moleculares a la que pertenecen los transcritos. En el eje X se representa la cantidad de transcritos que pertenecen a cada uno de los términos GO asociados a las funciones moleculares representadas en el eje Y.

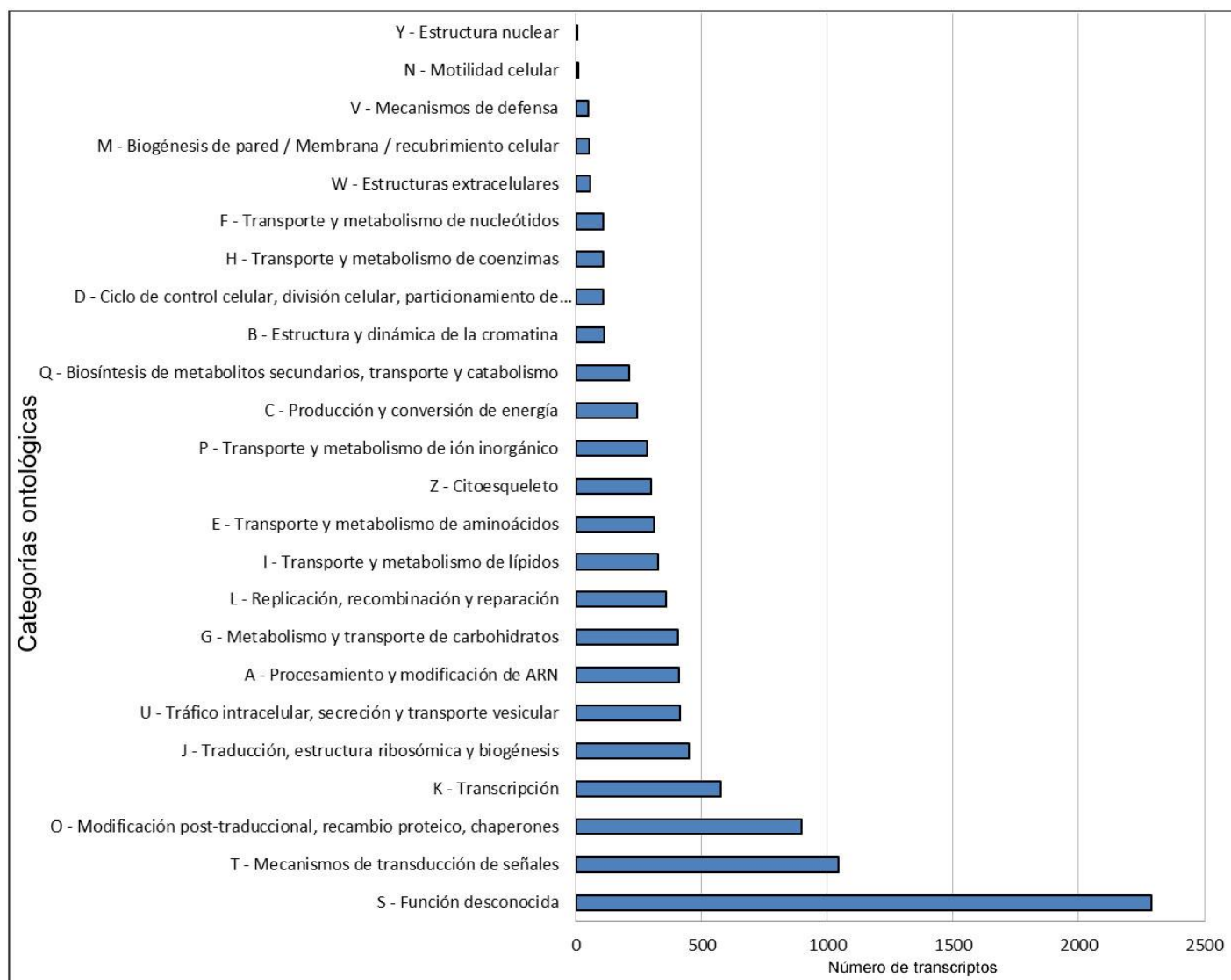


Figura 9. Anotación de la categoría funcional. Caracterización de los transcritos obtenida mediante EggNOG. En el eje X se representa la cantidad de transcritos que pertenecen a cada una de las distintas categorías funcionales representadas en el eje Y.

Tabla 4. Resumen de métricas y resultados del transcriptoma. Se listan las principales características biológicas obtenidas mediante Trinotate, incluyendo predicciones proteicas, búsqueda por similitud (*Blast*), identificación de proteínas transmembrana, péptidos señal, dominios proteicos y anotaciones funcionales a partir bases de datos especializadas.

	Únicos
Transcriptos totales	140.247
Transcriptos únicos	69.314
Predicciones proteicas	18.607
Blastx vs. Uniprot	8.135
Blastp vs. Uniprot	7.302
Blastx vs. Arthropoda	7.865
Blastp vs. Arthropoda	6.860
Blastx vs. Insecta	7.848
Blastp vs. Insecta	6.820
Blastx vs. <i>L. striatellus</i>	21.147
Blastp vs. <i>L. striatellus</i>	15.027
Blastx vs. <i>S. furcifera</i>	20.376
Blastp vs. <i>S. furcifera</i>	13.810
Blastx vs. <i>N. lugens</i>	19.438
Blastp vs. <i>N. lugens</i>	14.060
Proteínas transmembrana	4.370
Dominios proteicos	10.017
Ontología genética Pfam	6.100
Péptido señal	7.843

1.3 Anotación Génica

Los 69.314 transcriptos reconstruidos se compararon con bases de datos específicas para caracterizar genéticamente los procesos vitales para la supervivencia del insecto. Se muestran los mejores *hits* para cada uno de los procesos analizados, según posean mejor cobertura y mejor *e-value*. En relación con el proceso de ecdisis, se pudo inferir la identidad de seis genes, tales como el gen de la ETH (TRINITY DN1204 c0 g1 i1), de la Crz (TRINITY DN69828 c0 g1 i4), de

la EH (TRINITY DN10952 c0 g1 i1), del CCAP (TRINITY DN7706 c0 g1 14), el gen de *Burs* (TRINITY DN5062 c0 g1 i2) y de *Pburs* (TRINITY DN9187 c0 g1 i1) involucrados en la regulación de la ecdisis (Tabla 5, Anexo IV-Ecdisis). Todos presentaron altos porcentajes de identidad, entre el 73% y el 100%, excepto la inferencia del gen EH que reportó un 44% de similitud de secuencia con la del díptero *A. aegypti*.

En cuanto al proceso de ovogénesis, se identificaron un total de 32 genes involucrados en las diferentes etapas que caracterizan dicho proceso, como los que están activos durante la diferenciación celular (entre ellos *armadillo* (TRINITY DN6159 c0 g1 i19), *piwi* (TRINITY DN185 c0 g1 i9) y *vasa* (TRINITY DN3339 c0 g1 i1)), los que representan un rol en la determinación de la polaridad anteroposterior (A-P) (algunos de los cuales son *bicaudal* (TRINITY DN4949 c0 g1 i1) y *egghead* (TRINITY DN19137 c0 g1 i1)) y en la polaridad dorsoventral (D-V) (como *cappuccino* (TRINITY DN5275 c0 g3 i3) y *squid* (TRINITY DN4198 c0 g1 i7)). Respecto a los genes con funciones más específicas en el proceso, como *maelstrom* (TRINITY DN1714 c0 g1 i4), *Rac* (TRINITY DN786 c0 g1 i4) y *quail* (TRINITY DN2076 c0 g1 i19), se identificaron con un porcentaje de identidad variado, entre el 23% y 76%, aproximadamente. Por otro lado, no se pudieron identificar *pumilio*, *gurken*, *bicoid* y *oskar* mediante esta búsqueda. En este proceso, en algunos casos se identificaron varios transcritos por genes, con porcentajes de identidad que oscilaron entre el 21 % y el 85% (Tabla 6, Anexo IV-Ovogénesis).

Finalmente, en cuanto al proceso de olfacción, se han encontrado varios transcritos correspondientes a los diferentes receptores, entre ellos 25 receptores ionotrópicos (IR), 30 receptores olfativos (OR) y un correceptor de receptor olfativo (Orco) (TRINITY DN24825 c0 g2 i1). Respecto a las proteínas solubles, se identificaron 12 proteínas de unión a odorantes (OBP) (Tabla 7, Anexo IV-Olfacción). Los porcentajes de identidad fueron muy variados, abarcando desde el 20% hasta el 98%, siendo, en líneas generales, los OBP los que presentaron los mayores porcentajes de identidad.

Tabla 5: Anotación Génica. Identidad de los transcritos obtenidos mediante similitud de secuencia contra la base de datos generada para el proceso de ecdisis. Visualización resumida de los transcritos y sus métricas que hacen mejor *hit* con los distintos genes estudiados en el proceso de ecdisis.

El * a la izquierda significa que se trata de una secuencia parcial 5'. La ausencia de * significa que es completa.

				BLAST			
Genes	ID	Long. Trans cripto (pb)	Long. Prot. (aa)	Especie - Gen ID (Mejor hit)	% ID	Long. <i>Hit</i> (aa)	e-value
Regulación de la Ecdisis							
BURSICON- ALPHA (<i>Burs</i>)	TRINITY_DN5062_c0_g1_i2	1.193	160	<i>N. lugens</i> (BAO00937.1)	95,60	159	4.23e-108
CORAZONINA (<i>Crz</i>)	TRINITY_DN69828_c0_g1_i4	1.026	*132	<i>N. lugens</i> (BAO00944.1)	73,75	80	1.58e-23
HORMONA DE ECLOSIÓN (<i>EH</i>)	TRINITY_DN10952_c0_g1_i1	476	*117	<i>A. aegypti</i> (XP_001647653.1)	44,23	52	1.13e-12
HORMONA DISPARADORA DE LA ECDISIS (<i>ETH</i>)	TRINITY_DN1204_c0_g1_i1	907	215	<i>N. lugens</i> (BAO00949.1)	77,36	212	3.89e-121
PARTNER OF BURSICON- BETA (<i>Pburs</i>)	TRINITY_DN9187_c0_g1_i1	773	139	<i>A. pisum</i> (XP_008179450.1)	77,48	111	1.17e-62
PÉPTIDO CARDIOACTIVO DE CRUSTÁCEOS (<i>CCAP</i>)	TRINITY_DN7706_c0_g1_i14	1.639	*159	<i>N. lugens</i> (BAO00946.1)	92,59	135	1.77e-73

Tabla 6: Anotación Génica. Identidad de los transcritos obtenidos mediante similitud de secuencia contra la base de datos generada para el proceso de ovogénesis. Visualización resumida de los transcritos y sus métricas que hacen mejor *hit* con los distintos genes estudiados en el proceso de ovogénesis.

El * a la izquierda significa que se trata de una secuencia parcial 5'. El * a la derecha significa que es parcial 3'. La ausencia de * significa que es completa.

				BLAST			
Genes	ID	Long. Transcrito (pb)	Long. Prot. (aa)	Especie - Gen ID (Mejor hit)	%ID	Long. Hit (aa)	e-value
Diferenciación de las células madre y germinales							
<i>ARMADILLO (arm)</i>	TRINITY_DN6159_c0_g1_i19	4.361	815	<i>D. melanogaster</i> (FBpp0089033)	85,19	729	0.0
<i>AUBERGINE (aub)</i>	TRINITY_DN185_c0_g1_i9	4.065	830	<i>D. melanogaster</i> (FBpp0079754)	25,82	457	1.78e-29
<i>P-ELEMENT INDUCED WIMPY TESTIS (piwi)</i>	TRINITY_DN185_c0_g1_i9	4.065	830	<i>D. melanogaster</i> (FBpp0079755)	21,69	807	1.62e-30
<i>PUMILIO (pum)</i>	-	-	-	-	-	-	-
<i>TWIN</i>	TRINITY_DN11696_c0_g1_i1	1.987	601	<i>D. melanogaster</i> (FBpp0083952)	26,51	347	2.31e-25
<i>VAS (vasa)</i>	TRINITY_DN3339_c0_g1_i1	3.305	1.013	<i>D. melanogaster</i> (FBpp0401445)	39,75	395	9.25e-89
<i>β SPECTRIN (β Spec)</i>	TRINITY_DN572_c0_g1_i18	5.108	*1.048	<i>D. melanogaster</i> (FBpp0306508)	76,70	837	0.0
Determinación y formación del eje A-P							
<i>BICAUDAL (Bic)</i>	TRINITY_DN4949_c0_g1_i1	1.146	184	<i>D. melanogaster</i> (FBpp0086896)	71,35	171	6.58e-61
<i>BICAUDAL C (BicC)</i>	TRINITY_DN11822_c0_g1_i1	2.525	692	<i>D. melanogaster</i> (FBpp0080363)	45,79	356	3.85e-97
<i>BICAUDAL D (BicD)</i>	TRINITY_DN45988_c0_g1_i1	967	286	<i>D. melanogaster</i> (FBpp0296968)	64,49	276	1.20e-100
<i>BICOID (bcd)</i>	-	-	-	-	-	-	-
<i>DYNEIN HEAVY CHAIN 64C (Dhc64C)</i>	TRINITY_DN50570_c0_g1_i1	4.696	*1.440	<i>D. melanogaster</i> (FBpp0304986)	78,09	1.442	0.0
<i>EGGHEAD (egh)</i>	TRINITY_DN19137_c0_g1_i1	726	*242*	<i>D. melanogaster</i> (FBpp0309090)	71,97	239	4.92e-116
<i>EXUPERANTIA (exu)</i>	TRINITY_DN8816_c0_g1_i1	2.249	404	<i>D. melanogaster</i> (FBpp0085555)	41,30	184	5.77e-41
<i>GURKEN (grk)</i>	-	-	-	-	-	-	-
<i>INTEGRIN ALPHAPS5 SUBUNIT (ItgaPS)</i>	TRINITY_DN39338_c0_g1_i2	2.386	*732	<i>D. melanogaster</i> (FBpp0423081)	25,09	562	1.23e-27
<i>OSKAR (osk)</i>	-	-	-	-	-	-	-

Establecimiento de la polaridad D-V							
18 WHEELER (18w)	TRINITY_DN7562_c0_g4_i1	3.530	*1.003	<i>D. melanogaster</i> (FBpp0085620)	54,08	980	0.0
BRAINIA (brn)	TRINITY_DN915_c0_g2_i2	2.518	568	<i>D. melanogaster</i> (FBpp0070606)	21,19	302	1.15e-09
CAPPUCCINO (capu)	TRINITY_DN5275_c0_g3_i3	3.688	1.062	<i>D. melanogaster</i> (FBpp0288516)	25,12	414	4.81e-23
CORNICHON (cni)	TRINITY_DN23649_c0_g1_i1	1.029	145	<i>D. melanogaster</i> (FBpp0080406)	76,39	144	8.83e-72
DYNACTIN 1 (DCTN1)	TRINITY_DN22476_c0_g1_i1	1.749	*584*	<i>D. melanogaster</i> (FBpp0075498)	49,72	529	1.08e-136
oo18 RNA-BINDING PROTEIN (orb)	TRINITY_DN2621_c0_g1_i1	2.796	*332	<i>D. melanogaster</i> (FBpp0290706)	52,08	240	2.32e-72
SQUID (sqd)	TRINITY_DN4198_c0_g1_i7	2.715	856	<i>D. melanogaster</i> (FBpp0082320)	25,32	158	4.08e-07
Otros genes involucrados en la ovogénesis							
DARKENER OF APRICOT (Doa)	TRINITY_DN207_c0_g1_i2	6.303	1.955	<i>D. melanogaster</i> (FBpp0289031)	24,82	270	1.71e-13
DYNEIN LIGHT INTERMEDIATE CHAIN (Dlic)	TRINITY_DN1812_c0_g1_i8	3.289	525	<i>D. melanogaster</i> (FBpp0073299)	57,23	491	1.23e-179
EGG-DERIVED TYROSINE PHOSPHATASE (EDTP)	TRINITY_DN2829_c0_g1_i1	2.297	653	<i>D. melanogaster</i> (FBpp0086114)	49,69	322	3.70e-103
ENDOPHILIN B (EndoB)	TRINITY_DN37749_c0_g1_i2	921	285*	<i>D. melanogaster</i> (FBpp0085743)	63,18	277	2.46e-124
G PROTEIN-COUPLED RECEPTOR KINASE 2 (Gprk2)	TRINITY_DN207_c0_g1_i2	6.303	1.955	<i>D. melanogaster</i> (FBpp0306622)	23,33	270	2.61e-15
HEAT SHOCK PROTEIN 60C (Hsp60C)	TRINITY_DN10510_c0_g1_i1	2.621	575	<i>D. melanogaster</i> (FBpp0078744)	75,45	554	0.0
MAELSTROM (mael)	TRINITY_DN1714_c0_g1_i4	2.333	507	<i>D. melanogaster</i> (FBpp0088574)	27,86	341	1.01e-25
PEROXINECTIN-LIKE (Pxt)	TRINITY_DN1715_c0_g1_i5	4.058	784	<i>D. melanogaster</i> (FBpp0082932)	38,81	603	1.45e-124
QUAIL (qua)	TRINITY_DN2076_c0_g1_i19	2.911	850	<i>D. melanogaster</i> (FBpp0080550)	30,64	904	1.58e-124
RAC	TRINITY_DN786_c0_g1_i4	3.158	524	<i>D. melanogaster</i> (FBpp0305023)	27,87	122	9.09e-06
SER/THR-RICH CASPASE (Strica)	TRINITY_DN3046_c0_g1_i2	2.486	619	<i>D. melanogaster</i> (FBpp0307275)	30,20	202	2.43e-12

Tabla 7: Anotación Génica. Identidad de los transcritos obtenidos mediante similitud de secuencia contra la base de datos generada para el proceso de olfacción. Visualización resumida de los transcritos y sus métricas que hacen mejor *hit* con los distintos genes estudiados en el proceso de olfacción.

El * a la izquierda significa que se trata de una secuencia parcial 5'. El * a la derecha significa que es parcial 3'. La ausencia de * significa que es completa. Un * a ambos lados significa que es interna.

				BLAST			
Genes	ID	Long. Transcritos (pb)	Long. Prot. (aa)	Especie - Gen ID (Mejor hit)	% ID	Long. Hit (aa)	e-value
CORRECEPTOR DE RECEPTORES DE ODORANTES (ORco)	TRINITY_DN24825_c0_g2_i1	1.091	279	<i>R. prolixus</i> (RPRC000476)	26,57	143	3.57e-07
PROTEÍNA DE UNIÓN A ODORANTES (OBP)	TRINITY_DN11975_c0_g1_i5	812	143	<i>L. striatellus</i> (AGZ04925.1)	92,25	142	1.81e-93
	TRINITY_DN14064_c0_g1_i1	614	138	<i>L. striatellus</i> (AGZ04921.1)	91,30	138	1.69e-90
	TRINITY_DN1864_c0_g1_i1	1.422	148	<i>S. furcifera</i> (AHB59656.1)	96,60	147	3.06e-77
	TRINITY_DN206_c0_g1_i10	966	*268	<i>S. furcifera</i> (AHB59663.1)	45,95	111	7.11e-27
	TRINITY_DN3903_c0_g1_i1	658	145	<i>L. striatellus</i> (AGZ04924.1)	91,20	125	5.84e-71
	TRINITY_DN43095_c0_g1_i1	897	*193	<i>N. lugens</i> (ACI30679.1)	83,67	147	2.58e-75
	TRINITY_DN5192_c0_g1_i6	968	179	<i>N. viridula</i> (QCZ25113.1)	31,78	129	2.06e-06
	TRINITY_DN5517_c0_g2_i1	1.311	197	<i>L. lineolaris</i> (AHF71059.1)	61,29	155	3.46e-67
	TRINITY_DN5568_c0_g1_i1	750	*201	<i>N. lugens</i> (AGZ04895.1)	84,24	165	1.28e-101
	TRINITY_DN5926_c0_g1_i1	883	259	<i>N. lugens</i> (AGZ04896.1)	83,33	246	1.17e-120
	TRINITY_DN6387_c0_g1_i4	1.653	*245	<i>S. furcifera</i> (AGZ04907.1)	98,37	184	4.33e-118
	TRINITY_DN6777_c0_g1_i1	681	146	<i>S. furcifera</i> (AGZ04902.1)	96,21	132	9.70e-93
RECEPTORES IONOTRÓPICOS (IR)	TRINITY_DN11355_c0_g1_i14	1.427	*459	<i>N. lugens</i> (XP_022187223_1)	27,47	91	4.40e-07
	TRINITY_DN17880_c0_g1_i5	2.017	*423	<i>D. citri</i> (XP_026679576_11)	33,33	132	1.16e-19
	TRINITY_DN19602_c0_g1_i2	1.376	414	<i>A. pisum</i> (XP_029345828_1)	20,82	365	5.60e-16
	TRINITY_DN2107_c0_g1_i1	3.099	926	<i>N. lugens</i> (XP_022200713_1)	58,74	744	0.0
	TRINITY_DN40231_c0_g1_i4	1.022	*341*	<i>H. halys</i> (XP_014272502_1)	21,93	301	4.70e-07
	TRINITY_DN45996_c0_g1_i6	2.571	761	<i>A. pisum</i> (XP_029346343_1)	61,37	717	0.0
	TRINITY_DN7133_c0_g1_i5	3.237	*727	<i>N. lugens</i> (XP_022200714_1)	55,09	786	0.0
	TRINITY_DN95621_c0_g1_i1	741	*248*	<i>N. lugens</i> (XP_022200717_1)	86,49	37	1.86e-14
	TRINITY_DN11316_c0_g1_i8	1.590	199	<i>R. prolixus</i> (RPRC017345)	32,32	198	2.83e-15
	TRINITY_DN11355_c0_g1_i14	1.427	*459	<i>R. prolixus</i> (RPRC000379)	21,37	454	2.43e-06

	TRINITY_DN13434_c0_g1_i2	1.880	*470	<i>N. lugens</i> (XP_022203339_1)	91,58	463	0.0
	TRINITY_DN14179_c0_g1_i1	1.360	*415	<i>R. prolixus</i> (RPRC000379)	24,39	246	1.56e-06
	TRINITY_DN14755_c0_g1_i1	435	110*	<i>R. prolixus</i> (RPRC002763)	61,54	104	1.06e-41
	TRINITY_DN17880_c0_g1_i5	2.017	*423	<i>A. pisum</i> (XP_008183092_2)	84,69	405	0.0
	TRINITY_DN19602_c0_g1_i2	1.376	414	<i>A. lineolatus</i> (APZ81410_1)	24,42	344	2.37e-23
	TRINITY_DN20443_c0_g1_i1	507	*169*	<i>C. lividipennis</i> (AXU25111_1)	32,48	117	1.45e-12
	TRINITY_DN2107_c0_g1_i1	3.099	926	<i>A. lineolatus</i> (APZ81411_1)	30,30	439	6.03e-48
	TRINITY_DN25091_c0_g2_i1	810	*240	<i>A. lineolatus</i> (APZ81415_1)	26,98	215	3.35e-16
	TRINITY_DN3618_c0_g1_i6	3.852	*1.016	<i>R. prolixus</i> (RPRC002763)	58,75	160	1.11e-61
	TRINITY_DN40231_c0_g1_i4	1.022	*341*	<i>H. halys</i> (XP_024214423_1)	20,79	303	7.05e-07
	TRINITY_DN43282_c0_g1_i1	600	200*	<i>A. lineolatus</i> (APZ81410_1)	38,61	202	1.24e-37
	TRINITY_DN45996_c0_g1_i6	2.571	761	<i>A. pisum</i> (XP_008183092_2)	35,16	512	1.81e-73
	TRINITY_DN7133_c0_g1_i5	3.237	*727	<i>N. lugens</i> (XP_022187829_1)	28,32	452	1.61e-44
	TRINITY_DN7270_c0_g1_i1	3.219	990	<i>R. prolixus</i> (RPRC002763)	28,11	153	9.83e-08
	TRINITY_DN85391_c0_g1_i1	647	*180	<i>C. lividipennis</i> (AXU25112_1)	52,26	155	1.51e-43
RECEPTORES OLFATIVOS (OR)	TRINITY_DN27871_c0_g2_i2	626	*167	<i>N. lugens</i> (XP_022200348.1)	29,81	161	1.36e-16
	TRINITY_DN77195_c0_g1_i1	501	*139	<i>N. lugens</i> (XP_022200348.1)	30,44	115	2.42e-17
	TRINITY_DN10113_c0_g1_i1	563	*142	<i>N. lugens</i> (XP_022204744.1)	42,75	138	3.35e-41
	TRINITY_DN4341_c0_g3_i3	969	109	<i>N. lugens</i> (XP_022204744.1)	41,30	138	5.27e-35
	TRINITY_DN11791_c0_g1_i1	1.217	*286	<i>N. lugens</i> (XP_022200228.1)	45,66	265	9.72e-78
	TRINITY_DN12315_c0_g2_i2	992	*273	<i>A. pisum</i> (XP_016660447.1)	28,74	167	4.65e-07
	TRINITY_DN12588_c0_g1_i3	817	255	<i>H. halys</i> (XP_024214259.1)	25,00	152	1.64e-09
	TRINITY_DN38863_c0_g1_i1	734	*200	<i>H. halys</i> (XP_024214259.1)	25,75	167	2.04e-06
	TRINITY_DN12954_c0_g1_i2	1.356	*414	<i>A. lineolatus</i> (APZ81474.1)	24,07	162	2.90e-06
	TRINITY_DN13549_c0_g1_i2	1.191	*374	<i>A. lineolatus</i> (APZ81500.1)	28,31	166	5.13e-07
	TRINITY_DN58542_c0_g1_i1	919	279	<i>N. lugens</i> (XP_022192649.1)	26,47	170	1.83e-06
	TRINITY_DN14580_c0_g2_i1	629	*202	<i>A. gossypii</i> (XP_027836930.1)	25,49	153	2.57e-06

TRINITY_DN64921_c0_g1_i1	1.142	*349	<i>H. halys</i> (XP_024220181.1)	25,77	97	1.58e-06
TRINITY_DN15317_c0_g1_i2	617	*136	<i>A. gossypii</i> (XP_027842896.1)	23,53	102	9.08e-06
TRINITY_DN18132_c0_g1_i3	990	295	<i>A. lineolatus</i> (APZ81502.1)	28,86	149	7.75e-07
TRINITY_DN20422_c0_g1_i1	627	*199	<i>N. lugens</i> (XP_022207573.1)	64,23	137	9.06e-55
TRINITY_DN20522_c0_g1_i1	615	*178	<i>R. prolixus</i> (RPRC000235)	25,29	170	1.06e-07
TRINITY_DN20817_c0_g1_i1	457	*153*	<i>C. lividipennis</i> (AXU25094.1)	20,14	144	9.48e-07
TRINITY_DN24825_c0_g2_i1	1.091	279	<i>C. lividipennis</i> (AXU25099.1)	21,66	277	1.04e-07
TRINITY_DN27120_c0_g1_i1	562	*134	<i>H. halys</i> (XP_024217960.1)	27,85	79	1.85e-06
TRINITY_DN53574_c0_g1_i1	551	*144	<i>H. halys</i> (XP_024217960.1)	25,93	81	1.36e-07
TRINITY_DN74751_c0_g1_i1	361	*120*	<i>A. lucorum</i> (AQM56101.1)	25,58	86	2.18e-07
TRINITY_DN34345_c0_g1_i1	731	*142	<i>N. lugens</i> (XP_022199828.1)	25,66	113	1.54e-07
TRINITY_DN43788_c0_g1_i1	455	*111	<i>N. lugens</i> (XP_022193747.1)	64,36	101	1.50e-46
TRINITY_DN3941_c4_g1_i2	961	*286	<i>H. halys</i> (XP_024214894.1)	28,75	80	5.05e-06
TRINITY_DN53423_c0_g2_i1	720	*217	<i>D. citri</i> (XP_026689162.1)	28,57	126	2.18e-06
TRINITY_DN5629_c2_g1_i1	960	*293	<i>C. lividipennis</i> (AXU25102.1)	24,84	157	5.79e-08
TRINITY_DN7031_c0_g1_i1	1.008	*319	<i>A. lineolatus</i> (APZ81479.1)	24,84	153	3.55e-09
TRINITY_DN76106_c0_g1_i1	493	*148	<i>A. pisum</i> (AQS60742.1)	29,46	129	8.55e-08
TRINITY_DN8662_c0_g1_i1	1.439	*400	<i>R. prolixus</i> (RPRC000616)	22,68	313	6.20e-06

Discusión

1. Secuenciación y ensamble

La técnica de ARN-Seq permite generar grandes volúmenes de datos a partir de organismos o tejidos específicos (Trapnell *et al.*, 2010; Satam *et al.*, 2023). La cantidad de lecturas necesarias para un experimento de secuenciación depende del objetivo experimental y de la naturaleza de la muestra. Según el consorcio ENCODE (Instituto Nacional de Investigación del Genoma Humano, NHGRI), para el análisis de genes altamente expresados se recomiendan entre 5 y 25 millones de lecturas por muestra; para evaluar la expresión génica global o analizar *splicing* alternativo, entre 30 y 60 millones de lecturas; y para la identificación y cuantificación precisa de nuevos transcritos, un mínimo de 100-200 millones de lecturas de más de 70 pares de bases. En nuestro estudio, se obtuvieron 106 millones de lecturas de 101 pb con altos valores de cobertura, lo que asegura una profundidad adecuada para la identificación de nuevos transcritos. Esto es similar a los resultados reportados en estudios previos en insectos plaga como en *Dalbulus maidis* con lecturas con una longitud de 93 pb (Palacio, 2018) o en *Nezara viridula* con lecturas de 100 pb de longitud (Lavore *et al.*, 2018).

Si bien la generación masiva de transcritos mediante ARN-Seq puede producir errores de secuenciación (Fang *et al.*, 2012; Goodwin *et al.*, 2016), los análisis de calidad indicaron que en nuestro estudio no fueron significativos; la elección de la plataforma Illumina, reconocida por su alta precisión y confiabilidad (Metzker, 2010), contribuyó a minimizar el posterior sesgo en la secuenciación y obtención de lecturas.

Aun teniendo lecturas de buena calidad según FastQC, se realizó una búsqueda de similitud contra la base de datos UniVec, ya que la presencia de secuencias contaminantes es un problema relevante en los análisis transcriptómicos. En especies sin un genoma o transcriptoma de referencia disponible, la detección y eliminación de dichas secuencias resulta particularmente desafiante, ya que no es posible discriminar mediante alineamiento directo cuáles secuencias pertenecen al organismo estudiado y cuáles provienen de fuentes externas. Su omisión puede sesgar el número total de transcritos en los ensamblados *de novo* y aumentar la frecuencia de errores, lo que reduce tanto la calidad global del ensamblado como el poder de los análisis posteriores (Davidson & Oshlack, 2014). En esta línea, Alvarez & Landsman (2024), comprobaron que al realizar un ensamble *de novo* de lecturas de *Solanum lycopersicum* (orden Solanales), la presencia de contaminantes afectó significativamente el resultado, generando transcritos erróneos que comprometieron la calidad y completitud del transcriptoma obtenido. En este sentido, la búsqueda de similitud contra bases de datos especializadas como UniVec representa una estrategia clave, ya que permite identificar segmentos de origen vectorial introducidos

durante la construcción de la biblioteca.

La reconstrucción de transcriptos constituye un paso fundamental para obtener información biológica significativa (Istrail *et al.*, 2004; Grabherr *et al.*, 2011; Ono *et al.*, 2015). Los parámetros utilizados para generar el ensamble de *novo* fueron escogidos en base a estudios previos en insectos plaga como *D. maidis* (Palacio, 2018) y *N. viridula* (Perez-Gianmarco, 2018). Sin embargo, es importante destacar que el rendimiento de cualquier ensamblador puede variar según las características específicas de la muestra, como la complejidad del transcriptoma o la calidad del ARN, lo que refuerza la necesidad de interpretar las métricas dentro del contexto biológico de cada organismo.

En el caso de *Delphacodes kuscheli*, las métricas obtenidas se encuentran dentro de los rangos reportados para otras especies de insectos plaga (Xue *et al.*, 2010; Palacio, 2018; Perez-Gianmarco, 2018; Jin *et al.*, 2021), lo que indica que el transcriptoma reconstruido es sólido y confiable. La utilización de secuenciación *paired-end* contribuyó a este resultado, al proporcionar información bidireccional de los transcriptos y mejorar la capacidad de ensamble y la resolución de regiones complejas. Esta estrategia resulta particularmente valiosa en transcriptomas de organismos no modelo, donde la ausencia de un genoma de referencia disponible exige maximizar la precisión de las lecturas. Estudios previos han demostrado que las lecturas *paired-end* facilitan la reconstrucción de transcritos completos y la resolución de regiones repetitivas, en comparación con los ensambles basados únicamente en lecturas *single-end* (Zhao *et al.*, 2011). En estudios comparativos, los ensambles derivados de datos *paired-end* tienden a presentar un mayor N50, mayor longitud media y una mejor representación de isoformas que aquellos contruidos únicamente con lecturas *single-end* (Zhao *et al.*, 2011; Yang & Smith, 2013). En nuestro trabajo, esta ventaja se refleja en métricas robustas, incluyendo un alto valor de N50, que confirma la calidad del ensamble, de manera comparable a lo reportado previamente en *N. viridula* (Perez-Gianmarco, 2018).

Respecto a las métricas de completitud, el análisis de búsqueda de ortólogos evidenció una alta conservación génica, con valores de porcentaje de cobertura superiores al 80%. No obstante, se observó una proporción considerable de ortólogos duplicados, que podrían reflejar limitaciones algorítmicas del ensamblador o representar isoformas genuinas. La calidad de un transcriptoma de *novo* puede verse afectada por diversos factores, incluidos artefactos del ensamble como transcriptos redundantes no correspondientes a isoformas reales, los cuales aumentan el número de transcriptos y reducen las métricas de calidad. Estudios previos demostraron que los ensambles de *novo* generados con Trinity producen una gran cantidad de

isoformas para cada transcripto; por ello, cuando el objetivo es evaluar la presencia o ausencia de genes, se recomienda conservar la isoforma de mayor cobertura para minimizar errores y generar un ensamble con genes comparables entre especies (Yang & Smith, 2013). En nuestro estudio se aplicó esta estrategia, dado que la selección de los transcriptos más largos como representantes por gen (locus) permite recuperar secuencias más completas y con mayor contenido informativo. No obstante, es importante considerar que la isoforma más larga no necesariamente representa la variante biológicamente funcional o predominante, y que este criterio implica la pérdida de isoformas alternativas (Yang & Smith, 2013; Yan *et al.*, 2026).

Independientemente del nivel de ortología analizada, las métricas de cobertura obtenidas en nuestro estudio resultan significativas. De acuerdo con Simão *et al.*, (2015), un ensamble puede considerarse íntegro cuando presenta un porcentaje de representación superior al 60%. En nuestro caso, los valores oscilaron entre un 80% y un 90%, lo que evidencia una alta completitud. Estos niveles se encuentran dentro del rango reportado en otros transcriptomas de insectos; por ejemplo, en el ortóptero *Shirakiacris shirakii* se reportó un 78% de genes conservados (Qiu *et al.*, 2016), mientras que en el hemíptero plaga *Phenacoccus solenopsis* se alcanzó un 82% (Singh *et al.*, 2019). Valores similares se han observado en especies hemípteras de interés agronómico, como *D. maidis* y *N. viridula*, con aproximadamente un 85% de cobertura de genes conservados (Palacio, 2018; Perez-Gianmarco, 2018). En hemípteros de relevancia médica, como *Rhodnius montenegrensis* y *Rhodnius robustus*, de Carvalho *et al.*, (2017) también reportaron más del 80%, mientras que en el hemíptero *Oncopeltus fasciatus* se alcanzaron niveles superiores al 90% (Panfilio *et al.*, 2019; Baricalla, 2020). En conjunto, estas comparaciones sugieren que nuestro transcriptoma es representativo, respaldando su utilidad para la inferencia de genes implicados en procesos biológicos de interés.

Si bien el transcriptoma generado se obtuvo a partir de un *pool* de todos los estados de desarrollo de *D. kuscheli*, y presentó altos porcentajes de cobertura, el análisis con BUSCO reportó la presencia de secuencias faltantes. Esto no necesariamente refleja su ausencia en el organismo, sino que puede atribuirse a diversos factores, como una baja o nula expresión de estos genes en las condiciones de muestreo, una expresión tejido específica, o limitaciones inherentes al ensamblado y al filtrado inicial de lecturas (Waterhouse *et al.*, 2017).

Al comparar el transcriptoma de *D. kuscheli* mediante alineamientos locales frente a diferentes bases de datos, se observó que el número de transcriptos detectados fue consistentemente mayor cuando el alineamiento se realizó a nivel nucleotídico en comparación con las predicciones proteicas. Esta diferencia probablemente se deba a que Blastx traduce la

secuencia en seis marcos de lectura para generar las secuencias de aminoácidos necesarios para la búsqueda, generando potenciales redundancias, frente al único considerado en las predicciones proteicas, tal como se ha reportado por Cock *et al.*, (2015). En cuanto a los géneros encontrados, la predominancia de los bien estudiados, como *Homo*, *Sophophora*, *Mus*, *Drosophila*, *Anopheles* y *Aedes*, evidencia el sesgo inherente de las bases de datos, donde la riqueza de anotaciones favorece la identificación de estas especies (Uniprot 2023, <https://www.uniprot.org/>). Sin embargo, al comparar con los delfácidos se observó un elevado número de transcritos conservados, consistente con la estrecha relación filogenética dentro de este grupo (Glover *et al.*, 2019). La mayor coincidencia observada fue con *Laodelphax striatellus*, un *planthopper* polífago vector de *Rice black-streaked dwarf virus*, que afecta al cultivo del arroz (Zhu *et al.*, 2017). Esta conservación entre delfácidos, también fue evidenciada recientemente en un análisis transcriptómico del cuerpo graso de *D. kuscheli* (Pascual *et al.*, 2025), reforzando la hipótesis de que la cercanía filogenética y la superposición en la selección de plantas hospedantes contribuyen a patrones de expresión génica similares entre las especies (Huang *et al.*, 2020). No obstante, no puede descartarse que la mayor similitud observada se deba, al menos en parte, a la mayor cobertura y anotación presente en el genoma de *L. striatellus* en comparación con las demás especies analizadas (Zhu *et al.*, 2017).

Al realizar la anotación funcional del transcriptoma mediante GO, se observó que las categorías más representadas corresponden a procesos relacionados con la regulación transcripcional y a funciones de unión como las de iones metálicos, de ATP, de ARN y ADN, de proteínas idénticas y de iones de zinc, coincidiendo con lo reportado en otras especies de insectos como *O. fasciatus* (Baricalla, 2020). Estas dos funciones están altamente relacionadas entre sí, ya que la regulación transcripcional controla la transcripción de genes a través de la unión de proteínas mediadas por señales intra y extracelulares (Casamassimi & Ciccodicola, 2019). Por otro lado, el análisis por EggNOG identificó 24 categorías funcionales, destacándose nuevamente aquellas relacionadas con procesos transcripcionales. A diferencia de otras especies, como *Sogatella furcifera* (Jin *et al.*, 2021), la categoría referente a solo predicción de función general (R) no fue encontrada, aunque sí se observaron proteínas con función desconocida (S). Entre las características más representadas se encuentran transcripción (K), tráfico intracelular, secreción y transporte vesicular (U), modificación postraducciona, recambio proteico, chaperonas (O) y replicación, recombinación y reparación (L), coincidentes con las más representadas encontradas en *O. fasciatus* (Baricalla, 2020). Dado que en este estudio se trabajó con material genético proveniente de un *pool* de insectos en todos sus estados de desarrollo, no es raro que los

procesos y funciones más representados correspondan a funciones transcripcionales. Durante todo el ciclo de vida se activan diversos factores de transcripción, lo que aumenta la representación global de genes con funciones reguladoras de transcripción, como se ha observado en *Bactrocera dorsalis* (Geib *et al.*, 2014).

En conjunto, nuestros resultados permiten afirmar que el transcriptoma de *D. kuscheli* generado en este trabajo es de buena calidad y cobertura, lo que establece una base sólida para proseguir con la inferencia y anotación génica.

2. Anotación Génica

En este trabajo, se decidió hacer foco en la caracterización génica de los procesos de ecdisis, ovogénesis y olfacción. Estos son algunos de los procesos vitales para el desarrollo y supervivencia de *D. kuscheli*.

2.1 Ecdisis

El proceso de ecdisis, que implica la muda periódica del exoesqueleto, está regulado por una compleja cascada hormonal (Malhotra & Basu, 2023). En este estudio se identificaron seis genes asociados a dicha regulación, cuatro pertenecientes a su fase inicial y dos a los eventos posteriores a la muda. Dentro del grupo de genes involucrados en la activación de la ecdisis, se detectaron transcritos con alta similitud a los genes que codifican la hormona disparadora de la ecdisis (ETH), la corazonina (Crz), la hormona de la eclosión (EH), y el péptido cardioactivo de crustáceos (CCAP). Los niveles de identidad obtenidos fueron elevados en la mayoría de los casos, lo que respalda la confiabilidad de las anotaciones y sugiere una conservación evolutiva de estos genes. Esto concuerda con estudios previos que muestran conservación de secuencias y funciones de los genes ETH/ETHR, EH, CCAP, en holometábolos y hemimetábolos (Malhotra & Basu, 2023). Se destaca el caso de la hormona de eclosión (EH), donde sólo se obtuvo ~44 % de similitud de secuencia; esta divergencia podría deberse a dos factores independientes: por un lado, a que la secuencia reconstruida corresponde a un transcripto parcial del ensamble, y por otro, a una variación natural entre especies, como ya se ha reportado en el lepidóptero *Heortia vitessoides*, en la cual pese a esa variación el silenciamiento de EH lleva a defectos en la pupación y eclosión, indicando que su función reguladora en la ecdisis se mantiene conservada (Li *et al.*, 2020).

Además, se identificaron los genes *bursicon* (*Burs*) y *partner of bursicon* (*Pburs*), los cuales podrían tener un rol en la expansión alar y el esclerosamiento de la nueva cutícula; como

ha sido descrito en *Drosophila melanogaster* (Luo *et al.*, 2005).

2.2 Ovogénesis

Muchos de los genes que se expresan en ovarios pueden presentar diferentes mecanismos de regulación como consecuencia de la diversidad inherente al tipo de ovario que presente el insecto (Brody, 1999; Atella *et al.*, 2005; Lynch & Roth, 2011; Leyria *et al.*, 2020).

En este trabajo fue posible inferir la identidad de varios genes pertenecientes al proceso de ovogénesis, abarcando los estadios de diferenciación de las células madres y germinales tales como *armadillo* (*arm*) (Nüsslein-Volhard *et al.*, 1987) y *vasa* (*vas*) (Hay *et al.*, 1990), determinación y formación del eje anteroposterior (A-P) como los genes *Bicaudal* (*Bic*, *BicC* y *BicD*) (Mohler & Wieschaus, 1986) y *egghead* (*egh*) (Goode *et al.*, 1996), establecimiento de la polaridad dorsoventral (D-V) entre ellos *cappuccino* (*capu*) y *squid* (*sqd*) (Manseau & Schüpbach, 1989). Como era de esperar, *bicoid* (*bcd*), *oskar* (*osk*) y *gurken* (*grk*) no se identificaron en el transcriptoma de *D. kuscheli*. Esto es coincidente con lo reportado en otros insectos como *Apis mellifera* (Hymenoptera), *Tribolium castaneum* (Coleoptera), *Bombyx mori* (Lepidoptera), *Oncopeltus fasciatus*, *Acyrtosiphon pisum*, *Rhodnius prolixus* y *Dalbulus maidis* (Hemiptera), evidenciando la presencia del gen *bicoid* únicamente en la mosca de la fruta *D. melanogaster* (Bonneton, 2003; Dearden *et al.*, 2006; Shigenobu *et al.*, 2010; Palacio, 2018; Pascual & Rivera-Pomar, 2022). Una situación similar se reportó para los genes *oskar* y *gurken* (Dearden *et al.*, 2006; Shigenobu *et al.*, 2010; Ewen-Campen *et al.*, 2011).

Dentro de los hemípteros se realizó una caracterización *in silico* de un gran repertorio de genes vinculados a este proceso, encontrándose en este trabajo 13 genes en común con *R. prolixus*, los cuales fueron: el gen *armadillo* dentro de los genes de segmentación, los genes *cappuccino*, *cornichon* y *squid* dentro de los genes involucrados en el establecimiento de la polaridad dorsoventral y los genes *Bicaudal C*, *Bicaudal D*, *egghead*, *piwi*, *aubergine*, *exuperantia*, *maelstrom* y *vasa*, pertenecientes a la ovogénesis (Pascual & Rivera-Pomar, 2022). Algunos de los mismos fueron validados experimentalmente, como por ejemplo *BicC*. No sólo se evidenció conservación a nivel de secuencia, sino que también se confirmó el rol durante la ovogénesis tanto en los hemípteros plagas *D. maidis* y *Nilaparvata lugens* como en la chinche de relevancia sanitaria *R. prolixus* (Zhang *et al.*, 2015; Pascual *et al.*, 2021; Dalaisón-Fuentes *et al.*, 2022). Sin embargo, en este estudio, no se encontró el gen *pumilio*.

2.3 Olfacción

El proceso de olfacción es de gran importancia en insectos, ya que lo utilizan para localizar plantas hospedadoras, encontrar parejas y detectar feromonas o señales químicas (Zhou *et al.*, 2014). Por medio de búsquedas por similitud de secuencia, en este trabajo hemos logrado identificar 25 receptores ionotrópicos (IR), 30 receptores olfativos (OR), 12 proteínas de unión a odorantes (OBP) y un correceptor de receptor olfativo (Orco) como se ha hallado en estudios previos en el género *Drosophila*, y las especies *Aedes aegypti*, *R. prolixus* y *D. maidis* (Etebari *et al.*, 2017; Latorre-Estivalis *et al.*, 2017; Crespo, 2021; Rondón *et al.*, 2022). Es de destacar que, si bien estas proteínas se encuentran altamente conservadas entre especies y órdenes, pueden ser muy divergentes llegando a conservarse un 10-15% únicamente entre especies (Pelosi *et al.*, 2017), sin embargo, en este estudio los porcentajes de identidad encontrados para estas proteínas son elevados, incluso cuando se compara con estudios previos similares, por ejemplo, en *D. maidis* (Crespo, 2021).

En este trabajo se evidencia una alta similitud entre *D. kuscheli* y los distintos hemípteros contrastados, destacándose los delfácidos *L. striatellus* y *S. furcifera* (mayor al 90%) como las especies más cercanas, lo que sugiere una conservación evolutiva de estos genes. Esto concuerda con lo observado en *D. maidis* (Crespo, 2021).

Es importante destacar que la imposibilidad de identificar determinados genes, como por ejemplo el gen *pumilio*, no implica necesariamente su ausencia en la especie. Esta limitación puede deberse a una divergencia evolutiva, así como a la presencia de alguna mutación, por ejemplo, presente en las secuencias a comparar, lo que impide inferir su identidad, o bien a que no haya sido posible reconstruir el transcripto durante el proceso de ensamblado. Otra posibilidad es que algunos genes se expresen en niveles muy bajos siendo necesario aplicar un análisis de enriquecimiento para poder identificarlos, como por ejemplo realizar un enfoque cuantitativo (Reimand *et al.*, 2019).

Cabe destacar que este trabajo aporta información nueva y relevante sobre el repertorio génico de *D. kuscheli*; dada la importancia de los procesos funcionales elegidos, esenciales para su supervivencia, podría ser útil para el desarrollo de estrategias de manejo y control de plagas.

Conclusiones

Finales

- La reconstrucción del transcriptoma de *Delphacodes kuscheli* a través de un ensamble de *novo* presentó buenos niveles de cobertura, completitud y calidad. Esto permitió generar un transcriptoma confiable, para continuar con los demás objetivos.
- La similitud de las secuencias transcriptómicas de *D. kuscheli* y las de otras especies de delfácidos respalda la existencia de una relación filogenética cercana entre ellas.
- Se identificaron en *D. kuscheli* genes asociados a la ecdisis, ovogénesis y olfacción a partir de su homología con secuencias previamente descritas en otras especies de insectos.

Bibliografía

- Aguilar-Bultet L., Falquet L., 2015. *Secuenciación y ensamblaje de novo de genomas bacterianos: una alternativa para el estudio de nuevos patógenos*. Rev. Salud Anim., Vol. 37, N.º 2, pp. 125-132.
- Almagro Armenteros J.J., Tsirigos K.D., Sønderby C.K., Petersen T.N., Winther O., Brunak S., von Heijne G., Nielsen H., 2019. *SignalP 5.0 improves signal peptide predictions using deep neural networks*. Nature Biotechnology, Vol. 37, N° 4, pp. 420-423. <https://doi.org/10.1038/s41587-019-0036-z>
- Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J., 1990. *Basic local alignment search tool*. Journal Molecular Biology, Vol. 215, pp. 403-410. doi: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., 1997. *Gapped BLAST y PSI-BLAST – a new generation of protein database search programs*. Nucleic Acids Res., Vol. 25, pp. 3389-3402.
- Alvarez R.V., Landsman D., 2024. *GTax: improving de novo transcriptome assembly by removing foreign RNA contamination*. Genome Biology, Vol. 25, N° 12. <https://doi.org/10.1186/s13059-023-03141-2>
- Ambardar S., Gupta R., Trakroo D., Lal R., Vakhlu J., 2016. *High throughput Sequencing: An overview of sequencing chemistry*. Indian Journal of Microbiology, Vol. 56, N° 4, pp. 394-404.
- Anders S., Huber W., 2010. *Differential expression analysis for sequence count data*. Genome Biol., Vol. 11, R106. <https://doi.org/10.1186/gb-2010-11-10-r106>
- Andrada N.L., 2021. *Análisis transcriptómico de genes implicados en la muda de Dalbulus maidis*. Tesis Licenciatura en Genética, Escuela de Ciencias Agrarias, Naturales y Ambientales. Universidad del Noroeste de la Provincia de Buenos Aires.
- Atella G.C., Gondim K.C., Machado E.A., Medeiros M.N., Silva-Neto M.A., Masuda H., 2005. *Oogenesis y egg development in triatomines: a biochemical approach*. An Acad Bras Cienc., Vol. 77, pp. 405-30.
- Baricalla A.A., 2020. *Genómica, transcriptómica y proteómica de neuropéptidos de Oncopeltus fasciatus (Hemiptera: Lygaeidae)*. Tesis Magister en Bioinformática y Biología de Sistemas, Universidad del Noroeste de la Provincia de Buenos Aires.
- Bartlett C.R., Dietz L.L., Dmitriev D.A., Sanborn A.F., Soulier-Perkins A., Wallace M.S., 2018. *The Diversity of the True Hoppers (Hemiptera: Auchenorrhyncha)*. In: Footitt

- RG, Adler PH (Eds) *Insect Biodiversity: Science y Society*. Vol. II. Wiley Blackwell, Chichester, UK, pp. 501–590. <https://doi.org/10.1002/9781118945582.ch19>
- Bäumer D., Ströhlein N.M., Schoppmeier M., 2012. *Opposing effects of Notch-signaling in maintaining the proliferative state of follicle cells in the telotrophic ovary of the beetle Tribolium*. *Front Zool.*, Vol. 9, 15. <https://doi.org/10.1186/1742-9994-9-15>
- Bautista Moreno R., 2010. *Las tres generaciones de la secuenciación*. ResearchGate, Vol. 3, Nº 128. <https://www.researchgate.net/publication/44708460>
- Ben-Arie N., Lancet D., Taylor C., Khen M., Walker N., Ledbetter D. H., Carrozzo R., Patel K., Sheer D., Lehrach H., *et al.*, 1994. *Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire*. *Hum. Mol. Genet.*, 3, pp. 229-235.
- Berenbaum M.R., 1996. *Bugs in the System*, in: Wesley., N.Y.A. (Ed.).
- Berger M.F., Levin J.Z., Vijayendran K., Sivachenko A., Adiconis X., Maguire J., Johnson L.A., Robinson J., Verhaak R.G., Sougnez C., Onofrio R.C., Ziaugra L., Cibulskis K., Laine E., Barretina J., Winckler W., Fisher D.E., Getz G., Meyerson M., Jaffe D.B., Gabriel S.B., Lyer E.S., Dummer R., Gnirke A., Nusbaum C., Garraway L.A., 2010. *Integrative analysis of the melanoma transcriptome*. *Genome Research*, Vol. 20, pp. 413-27. doi: [10.1101/gr.103697.109](https://doi.org/10.1101/gr.103697.109)
- Blekhman R., Marioni J.C., Zumbo P., Stephens M., Gilad Y., 2010. *Sex-specific y lineage-specific alternative splicing in primates*. *Genome Research*, Vol. 20, Nº 2, pp. 180-9. doi: [10.1101/gr.099226.109](https://doi.org/10.1101/gr.099226.109)
- Blencowe B.J., Ahmad S., Lee L.J., 2009. *Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes*. *Genes Dev.*, Vol. 23, pp. 1379–86. <https://doi.org/10.1101/gad.1788009>
- Bohbot J., Pitts R.J., Kwon H.-W., Rützler M., Robertson H.M., Zwiebel L.J., 2007. *Molecular characterization of the Aedes aegypti odorant receptor gene family*. *Insect Molecular Biology*, Vol. 16, Nº 5, pp. 525-537. <https://doi.org/10.1111/j.1365-2583.2007.00748.x>
- Bohbot J., Fu L., Le C., Chauhan C.L., Dickes J.C., 2011. *Multiple activities of insect repellents on odorant receptors in mosquitoes*. *Medical and Veterinary Entomology*, Vol. 25, Nº 4, pp 436-444. <https://doi.org/10.1111/j.1365-2915.2011.00949.x>

- Bolger A.M., Lohse M., Usadel B., 2014. *Trimmomatic: A flexible trimmer for Illumina Sequence Data*. Bioinformatics, Vol. 30, N° 15, pp. 2114-2120.
<https://doi.org/10.1093/bioinformatics/btu170>
- Bonneton F., 2003. *Divergence évolutive extrême d'un gène homéotique: le cas bicoid* [Extreme divergence of a homeotic gene: the bicoid case]. Med Sci (Paris) Vol. 12, pp. 1265-70. <https://doi.org/10.1051/medsci/200319121265>
- Bramlett K., Lea K., Qu L., Whitley P., Schageman J., Gu J., 2011. *Transcriptome Analysis Using Next-Generation Sequencing Technology*. Journal of Biomolecular Techniques, Vol. 22, pp. 44-45.
- Brentassi M.E., Medina R., de la Fuente D., Franco M.E.E., Toledo A.V., Saparrat M.C.N., Balatti P.A., 2020. *Endomycobiome associated with females of the planthopper Delphacodes kuscheli (Hemiptera: Delphacidae): A metabarcoding approach*. Heliyon, Vol. 6, N° 8. <https://doi.org/10.1016/j.heliyon.2020.e04634>
- Brody T., 1999. The Interactive Fly: gene networks, development y the Internet. Trends Genet., Vol. 15, pp. 333-4.
- Cagliari D., Pinto Dias N., dos Santos E.A., Neutzling Rickes L., Schmitt Kremer F., Farias J.R., Lenz G., Manzano Galdeano D., Mello Garcia F.R., Smagghe G., João Zotti M., 2020. *First transcriptome of the Neotropical pest Euschistus heros (Hemiptera: Pentatomidae) with dissection of its siRNA machinery*. Scientific Report, Vol.10, N° 4856. <https://doi.org/10.1038/s41598-020-60078-3>
- Carlson J.R., 1996. *Olfaction in Drosophila: from odor to behavior*. Trends Genet., Vol. 12, pp. 175-160. [https://doi.org/10.1016/0168-9525\(96\)10015-9](https://doi.org/10.1016/0168-9525(96)10015-9)
- Casamassimi A., Ciccodicola A., 2019. *Transcriptional Regulation: Molecules, Involved Mechanisms, y Misregulation*. Int J Mol Sci., Vol. 20, N° 6, pp. 1281.
<https://doi.org/10.3390/ijms20061281> PMID: 30875728; PMCID: PMC6471904.
- Chapman R.F., 1998. *The Insects. Structure y Function*, 4th edn. Cambridge University Press, Cambridge.
- Charles J.P., 2010. *The regulation of expression of insect cuticle protein genes*. Insect Biochem Mol Biol., Vol. 40, N° 3, pp. 205-13. doi: [10.1016/j.ibmb.2009.12.005](https://doi.org/10.1016/j.ibmb.2009.12.005)
- Cloonan N., Forrest A.R., Kolle G., Gardiner B.B., Faulkner G.J., Brown M.K., Taylor D.F., Steptoe A.L., Wani S., Bethel G., Robertson A.J., Perkins A.C., Bruce S.J., Lee C.C., Ranade S.S., Peckham H.E., Manning J.M., McKernan K.J., Grimmond S.M., 2008. *Stem cell transcriptome profiling via massive-scale mRNA sequencing*. Nature

- Methods, Vol. 5, N° 7, pp. 613-9. doi:[10.1038/nmeth.1223](https://doi.org/10.1038/nmeth.1223)
- Cock P.J.A., Grünig B.A., Paszkiewicz K., Pritchard L., 2013. *Galaxy tools y workflows for sequence analysis with applications in molecular plant pathology*. PeerJ 1: e167. <https://doi.org/10.7717/peerj.167>
- Cock P.J.A., Chilton J.M., Gruening B., Johnson J.E., Soranzo N., 2015. *NCBI BLAST+ integrated into Galaxy*. GigaScience, Vol. 4, N° 39. <http://dx.doi.org/10.1186/s13742-015-0080-7>
- Compeau P.E., Pevzner P.A., Tesler G., 2011. *How to apply de Bruijn graphs to genome assembly*. Nat Biotechnol., Vol. 29, N° 11, pp. 987-91. <https://doi.org/10.1038/nbt.2023>
- Conci L.R., Marzachi C., 1993. *Avances en la caracterización molecular del genoma del virus causal del “Mal de Río Cuarto”*. Pp. 27-28 in *Actas del workshop “Mal de Río Cuarto del Maíz”*. INTA, Córdoba, Argentina.
- Conesa A., Madrigal P., Tarazona S., Gomez-Cabrero D., Cervera A., McPherson A., Szczesniak M.W., Gaffney D.J., Elo L.L., Zhang X., Mortazavi A., 2016. *A survey of best practices for RNA-seq data analysis*. Genome Biology, Vol. 17, N° 13.
- Coutelis J.B., Petzoldt A.G., Speder P., Suzanne M., Noselli S., 2008. *Left-right asymmetry in Drosophila*. Semin. Cell Dev. Biol., Vol. 19, pp. 252-62.
- Cox D.N., Chao A., Lin H., 2000. *piwi encodes a nucleoplasmic factor whose activity modulates the number y division rate of germline stem cells*. Development, Vol. 127, N° 3, pp. 503-14. <https://doi.org/10.1242/dev.127.3.503>
- Crespo M., 2021. *Identificación y caracterización funcional del comportamiento olfatorio de Dalbulus maidis*. Tesis Licenciatura en Genética, Escuela de Ciencias Agrarias, Naturales y Ambientales. Universidad del Noroeste de la Provincia de Buenos Aires.
- Dalaisón-Fuentes L.I., Pascual A., Gazza E., Welchen E., Rivera-Pomar R., Catalano M.I., 2022. *Development of efficient RNAi methods in the corn leafhopper Dalbulus maidis, a promising application for pest control*. Pest Manag Sc, Vol. 78, N° 7, pp. 3108-3116. <https://doi.org/10.1002/ps.6937>
- Daubnerová I., Žitňan D., 2021. *Handbook of Hormones* (Second Edition).
- Davidson N.M., Oshlack A., 2014. *Corset: enabling differential gene expression analysis for de novo assembled transcriptomes*. Genome Biol., Vol. 15, N° 410. <https://doi.org/10.1186/s13059-014-0410-6>
- de Bruijn N., 1946. *A Combinatorial Problem*. Koninklijke Nederlyse Akademie van

- Wetenschappen, Vol. 49, N° 7, pp. 758–764.
- de Carvalho D.B., Congrains C., Chahad-Ehlers S., Pinotti H., de Brito R.A., da Rosa J.A., 2017. *Differential transcriptome analysis supports *Rhodnius montenegrensis* y *Rhodnius robustus* (Hemiptera, Reduviidae, Triatominae) as distinct species*. PLOS ONE, Vol. 12, N° 4, e0174997. <https://doi.org/10.1371/journal.pone.0174997>
- de Loof A., Briers T., Huybrechts R., 1984. *Presence y function of ecdysteroids in adult insects*. Comp. Bioch. y Phys. Part B: Comp Bioch., Vol. 79, N° 4, pp. 505-509.
- Dearden P.K., Wilson M.J., Sablan L., Osborne P.W., Havler M., McNaughton E., Kimura K., Milshina N.V., Hasselmann M., Gempe T., Schioett M., Brown S., Elsik C., Holly P., Kadowaki T., Beye M., 2006. *Patterns of conservation y change in honey bee developmental genes*. Genome Res., Vol. 16, pp. 1376–1384.
- Denoeud F., Aury J.M., Da Silva C., Noel B., Rogier O., Delledonne M., Morgante M., Valle G., Wincker P., Scarpelli C., Jaillon O., Artiguenave F., 2008. *Annotating genomes with massive-scale RNA sequencing*. Genome Biology, Vol. 9, R175. <https://doi.org/10.1186/gb-2008-9-12-r175>
- Dewey E.M., McNabb S.L., Ewer J., Kuo G.R., Takanishi C.L., Truman J.W., Honegger H.W., 2004. *Identification of the gene encoding bursicon, an insect neuropeptide responsible for cuticle sclerotization y wing spreading*. Curr. Biol., Vol. 14, N° 13, pp. 1208-13. <https://doi.org/10.1016/j.cub.2004.06.051> PMID: 15242619.
- Doctor J., Fristrom D., Fristrom J.W., 1985. *The pupal cuticle of *Drosophila*: biphasic synthesis of pupal cuticle proteins in vivo y in vitro in response to 20-hydroxyecdysone*. J. Cell Biol., Vol. 101, N° 1, pp. 189-200. doi:[10.1083/jcb.101.1.189](https://doi.org/10.1083/jcb.101.1.189)
- Dubrovsky E.B., 2005. *Hormonal cross talk in insect development*. Trends in Endocrinology & Metabolism, Vol. 16, N° 1, pp. 6 – 11. doi: <https://doi.org/10.1016/j.tem.2004.11.003>
- Eddy S.R., 1995. *Multiple alignment using hidden Markov models*.
- Ekblom R., Wolf J.B.W., 2014. *A field guide to whole-genome sequencing, assembly y annotation*. Evol. Appl., Vol. 7, pp. 1026–42. <https://doi.org/10.1111/eva.12178>
- Ellegren H., Smeds L., Burri R., Olason P.I., Backström N., Kawakami T., Künstner A., Mäkinen H., Nadachowska-Brzyska K., Qvarnström A., Uebbing S., Wolf J.B., 2012. *The genomic landscape of species divergence in *Ficedula flycatchers**. Nature, Vol. 491, pp. 756-760.

- Etebari K., Hegde S., Saldaña M.A., Widen S.G., Wood T.G., Asgari S., Hughes G.L., 2017. *Global Transcriptome Analysis of Aedes aegypti Mosquitoes in Response to Zika Virus Infection*. mSphere, Vol. 2, N° 6. doi: [10.1128/mSphere.00456-17](https://doi.org/10.1128/mSphere.00456-17)
- Ewen-Campen B., Shaner N., Panfilio K.A., Suzuki Y., Roth S., Extavour C.G., 2011. *The maternal y early embryonic transcriptome of the milkweed bug Oncopeltus fasciatus*. BMC Genomics, Vol. 12, N° 61. <https://doi.org/10.1186/1471-2164-12-61>
- Ewer J., Gammie S.C., Truman J.W., 1997. *Control of insect ecdysis by a positive-feedback endocrine system: roles of eclosion hormone y ecdysis triggering hormone*. J. Exp. Biol., Vol. 200, pp. 869-81. <https://doi.org/10.1242/jeb.200.5.869>
- Fang Z., Martin J.A., Wang Z., 2012. *Statistical methods for identifying differentially expressed genes in RNA-Seq experiments*. Cell, Bioscience, Vol. 2, N° 26, pp. 1-8. <https://doi.org/10.1186/2045-3701-2-26>
- Fedurco M., Romieu A., Williams S., Lawrence I., Turcatti G., 2006. *BTA, a novel reagent for DNA attachment on glass y efficient generation of solid-phase amplified DNA colonies*. Nucleic Acids Res., Vol. 34, e22.
- Finn R.D., Bateman A., Clements J., Coghill P., Eberhardt R.Y., Eddy S.R., Heger A., Hetherington K., Holm L., Mistry J., Sonnhammer E.L., Tate J., Punta M., 2014. *Pfam: the protein families database*. Nucleic acids research, Vol. 42, pp. 222-230. <https://doi.org/10.1093/nar/gkt1223>
- Galizia C.G., Sachse S., 2010. *Odor coding in insects*. The Neurobiology of Olfaction, pp. 35-70.
- Gangishetti U., Veerkamp J., Bezdan D., Schwarz H., Lohmann I., Moussian B., 2012. *The transcription factor Grainy head y the steroid hormone ecdysone cooperate during differentiation of the skin of Drosophila melanogaster*. Insect Mol. Biol., Vol. 21, N° 3, pp. 283-95. doi: [10.1111/j.1365-2583.2012.01134.x](https://doi.org/10.1111/j.1365-2583.2012.01134.x)
- Garber M., Grabherr M.G., Guttman M., Trapnell C., 2011. *Computational methods for transcriptome annotation y quantification using RNA-seq*. Nat. Methods., Vol. 8, pp. 469-77.
- Geib S.M., Calla B., Hall B., et al., 2014. *Characterizing the developmental transcriptome of the oriental fruit fly, Bactrocera dorsalis (Diptera: Tephritidae) through comparative genomic analysis with Drosophila melanogaster utilizing modENCODE datasets*. BMC Genomics, Vol. 15, N° 942. <https://doi.org/10.1186/1471-2164-15-942>

- Glover N., Dessimoz C., Ebersberger I., Forslund S.K., Gabaldón T., Huerta-Cepas J., Martin M.J., Muffato M., Patricio M., Pereira C., da Silva A.S., Wang Y., Sonnhammer E., Thomas P.D., 2019. *Advances y applications in the quest for orthologs*. Mol. Biol. Evol., Vol. 36, N° 10, pp. 2157-64.
<http://dx.doi.org/10.1093/molbev/msz150>
- Good I.J., 1946. *Normal Recurring Decimals*. J. London Math. Soc., Vol. s1-21, N° 3, pp. 167–169. <https://doi.org/10.1112/jlms/s1-21.3.167>
- Goode S., Melnick M., Chou T.B., Perrimon N., 1996. *The neurogenic genes egghead y brainiac define a novel signaling pathway essential for epithelial morphogenesis during Drosophila oogenesis*. Development, Vol. 122, N° 12, pp. 3863-79.
<https://doi.org/10.1242/dev.122.12.3863>
- Goodwin S., McPherson J.D., McCombie W.R., 2016. *Coming of age: ten years of next generation sequencing technologies*. Nat. Rev. Genet., Vol. 17, pp. 333–351.
<https://doi.org/10.1038/nrg.2016.49>
- Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., Di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A., 2011. *Full-length transcriptome assembly from RNA-Seq data without a reference genome*. Nature Biotechnology, Vol. 29, pp. 644–52.
<https://doi.org/10.1038/nbt.1883>
- Guidobaldi F., Guerenstein P., 2012. *El sistema olfativo de los insectos*. Temas Selectos en Ecología Química de Insectos, pp. 46-71.
- Guttman M., Garber M., Levin J.Z., Donaghey J., Robinson J., Adiconis X., Fan L., Koziol M.J., Gnirke A., Nusbaum C., Rinn J.L., Lyer E.S., Regev A., 2010. *Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs*. Nature Biotechnology, Vol. 28, N° 5, pp. 503-10.
<https://doi.org/10.1038/nbt.1633>
- Haas B.J.; Zody M.C., 2010. *Advancing RNA-Seq analysis*. Nature Biotechnology, Vol. 28, N° 5, pp. 421– 423. <https://doi.org/10.1038/nbt0510-421>
- Haas B., Papanicolaou A., Yassour M., et al., 2013. *De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation y*

- analysis*. Nature Protocols, Vol. 8, pp. 1494–1512.
<https://doi.org/10.1038/nprot.2013.084>
- Hallem E.A., Carlson J.R., 2006. *Coding of odors by a receptor repertoire*. Cell, Vol. 125, N° 1, pp. 143-60. <https://doi.org/10.1016/j.cell.2006.01.050>
- Hardison R.C., 2003. *Comparative genomics*. PLoS Biol., Vol. 1, E58.
- Harini K., Sowdhamini R., 2012. *Molecular Modelling of Oligomeric States of DmOR83b, an Olfactory Receptor in D. Melanogaster*. Bioinform. Biol. Insights., Vol. 6, pp. 33-47. <https://doi.org/10.4137/bbi.s8990>
- Haubold B., Wiehe T., 2024. *Comparative genomics: methods y applications*. Naturwissenschaften, Vol. 91, N° 9, pp. 405-21. <https://doi.org/10.1007/s00114-004-0542-8> PMID: 15278216.
- Hay B., Jan L.Y., Jan Y.N., 1990. *Localization of vasa, a component of Drosophila polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity*. Development, Vol. 109, N° 2, pp. 425-33. <https://doi.org/10.1242/dev.109.2.425>
- Helfrich-Förster C., 2004. *Neurobiology of the fruit fly's circadian clock*. Genes Brain Behav., Vol. 4, N° 2, pp. 65-76. doi: <https://doi.org/10.1111/j.1601-183x.2004.00092.x>
- Hu K., Liu S., Qiu L., Li Y., 2019. *Three odorant-binding proteins are involved in the behavioral response of Sogatella furcifera to rice plant volatiles*. PeerJ. 7:e6576. <https://doi.org/10.7717/peerj.6576>
- Huang W., Reyes-Caldas P., Mann M., Seifbarghi S., Kahn A., Almeida R.P.P., Béven L., Heck M., Hogenhout S.A., Coaker G., 2020. *Bacterial Vector-Borne Plant Diseases: Unanswered Questions y Future Directions*. Molecular Plant, Vol. 13, pp. 1379-1393. <https://doi.org/10.1016/j.molp.2020.08.010>
- Istrail S., Sutton G.G., Florea L., Halpern A.L., Mobarry C.M., Lippert R., Walenz B., Shatkay H., Dew I., Miller J.R., Flanigan M.J., Edwards N.J., Bolanos R., Fasulo D., Halldorsson B.V., Hannenhalli S., Turner R., Yooseph S., Lu F., Nusskern D.R., Shue B.C., Zheng X.H., Zhong F., Delcher A.L., Huson D.H., Kravitz S.A., Mouchard L., Reinert K., Remington K.A., Clark A.G., Waterman M.S., Eichler E.E., Adams M.D., Hunkapiller M.W., Myers E.W., Venter J.C., 2004. *Whole-genome shotgun assembly y comparison of human genome assemblies*. Proc. Natl. Acad. Sci U S A, Vol. 101, N° 7, pp. 1916-21. <https://doi.org/10.1073/pnas.0307971100> PMID: 14769938; PMCID: PMC357027.

- Jiang Y., Gao X., Liu S., Zhang Y., Liu H., Sun F., Bao L., Waldbieser G., Liu Z., 2013. *Whole genome comparative analysis of channel catfish (*Ictalurus punctatus*) with four model fish species*. BMC Genomics, Vol. 14, pp. 780.
- Jin J.X., Ye Z.C., Jin D.C., Li F.L., Li W.H., Cheng Y., Zhou Y.H., 2021. *Changes in Transcriptome y Gene Expression in Sogatella furcifera (Hemiptera: Delphacidae) in Response to Cycloxaprid*. Journal of Economic Entomology, Vol. 114, N° 1, pp. 284–297 <https://doi.org/10.1093/jee/toaa238>
- Kelley R.L., 1993. *Initial organization of the Drosophila dorsoventral axis depends on an RNA-binding protein encoded by the squid gene*. Genes Dev., Vol. 7, N° 6, pp. 948-60. <https://doi.org/10.1101/gad.7.6.948>
- Kingan T.G., Gray W., Zitnan D., Adams M.E., 1997. *Regulation of ecdysis-triggering hormone release by eclosion hormone*. J. Exp. Biol., Vol. 200, pp. 3.245-56. doi: [10.1242/jeb.200.24.3245](https://doi.org/10.1242/jeb.200.24.3245)
- Kukurba K.R., Montgomery S.B., 2015. *RNA Sequencing y Analysis*. Cold Spring Harb Protocols, Vol. 11, pp. 951-69. doi:[10.1101/pdb.top084970](https://doi.org/10.1101/pdb.top084970). PMID: 25870306; PMCID: PMC4863231.
- Lamichhaney S, Barrio AM, Rafati N, Sundström G, Rubin CJ, Gilbert ER, Berglund J, Wetterbom A, Laikre L, Webster MT, Grabherr M, Ryman N, Andersson L, 2012. *Population-scale sequencing reveals genetic differentiation due to local adaptation in Atlantic herring*. Proceedings of the National Academy of Sciences, Vol. 109, pp. 19345-19350.
- Latorre-Estivalis J.M., Robertson H.M., Walden K.K., Ruiz J., Gonçalves L.O., Guarneri A.A., Lorenzo M.G., 2017. *The molecular sensory machinery of a Chagas disease vector: expression changes through imaginal moult y sexually dimorphic features*. Sci Rep., Vol. 7, N° 40049. <https://doi.org/10.1038/srep40049>.
- Lavore A., Perez-Gianmarco L., Esponda-Behrens N., Palacio V., Catalano M.I., Rivera-Pomar R., Ons S., 2018. *Nezara viridula (Hemiptera: Pentatomidae) transcriptomic analysis y neuropeptidomics*. Scientific Reports, Vol. 8, N° 17244. <https://doi.org/10.1038/s41598-018-35386-4>
- Levy N.S., Bakalyar H.A., Reed R.R., 1991. *Signal transduction in olfactory neurons*. J. Steroid Biochem. Mol. Biol., Vol. 39, pp. 633-637. doi: [10.1016/0960-0760\(91\)90262-4](https://doi.org/10.1016/0960-0760(91)90262-4)

- Levy S.E., Myers R.M., 2016. *Advancements in Next-Generation Sequencing*. Annu. Rev. Genomics Hum. Genet., Vol. 17, pp. 95-115. <https://doi.org/10.1146/annurev-genom-083115-022413>. PMID: 27362342.
- Leyria J., Orchard I., Lange A.B., 2020. *What happens after a blood meal? A transcriptome analysis of the main tissues involved in egg production in Rhodnius prolixus, an insect vector of Chagas disease*. PLoS Negl. Trop. Dis., Vol. 14, N° 10. <https://doi.org/10.1371/journal.pntd.0008516> PMID: 33057354; PMCID: PMC7591069.
- Li Z., Ye Q., Lyu Z., Lin T., 2020. *RNA interference of eclosion hormone gene reveals its roles in the control of ecdysis behavior in Heortia vitessoides Moore*. Arch Insect Biochem Physiol, Vol. 105, N° 1. <https://doi.org/10.1002/arch.21726>
- Liu X.Y., Yan R., Chen S.J., Zhang J.L., Xu H.J., 2022. *Orco mutagenesis causes deficiencies in olfactory sensitivity and fertility in the migratory brown planthopper, Nilaparvata lugens*. Pest Manag Sci. Vol. 79, N° 3, pp. 1030-1039. <https://doi.org/10.1002/ps.7286>
- López de Heredia U., 2016. *Las técnicas de secuenciación masiva en el estudio de la diversidad biológica*. ResearchGate, Munibe, Ciencias Naturales, N° 64, pp. 7-31. <https://doi.org/10.21630/mcn.2016.64.07>
- Lorenz L.J., Hall J.C., Rosbash M., 1989. *Expression of a Drosophila mRNA is under circadian clock control during pupation*. Development., Vol. 107, N° 4, pp.869-80. doi: <https://doi.org/10.1242/dev.107.4.869>
- Luo C., Dewey E. M., Sudo S., Ewer J., Yu Hsu S., Honegger H., Hsueh A.J.W., 2005. *Bursicon, the insect cuticle-hardening hormone, is a heterodimeric cystine knot protein that activates G protein-coupled receptor LGR2*. Biological Science, Vol. 102, N° 8, pp. 2820-2825. <https://doi.org/10.1073/pnas.0409916102>
- Lynch J.A., Roth S., 2011. *The evolution of dorsal-ventral patterning mechanisms in insects*. Genes Dev., Vol. 25, pp. 107-18.
- Lynch J.A., 2019. *Evolution of maternal control of axial patterning in insects*. Current Opinion in Insect Science, Vol. 31, pp. 37-42.
- Maher C.A., Kumar-Sinha C., Cao X., Kalyana-Sundaram S., Han B., Jing X., Sam L., Barrette T., Palanisamy N., Chinnaiyan A.M., 2009. *Transcriptome sequencing to detect gene fusions in cancer*. Nature, Vol. 458, pp. 97-101. doi: [10.1038/nature07638](https://doi.org/10.1038/nature07638)

- Malhotra P., Basu S., 2023. *The Intricate Role of Ecdysis Triggering Hormone Signaling in Insect Development y Reproductive Regulation*. Insects, Vol. 14, N° 8, pp. 711. <https://doi.org/10.3390/insects14080711>
- Malone J.H., Oliver B., 2011. *Microarrays, deep sequencing y the true measure of the transcriptome*. BMC Biol., Vol. 9, N° 34. <https://doi.org/10.1186/1741-7007-9-34>
- Manseau L.J., Schüpbach T., 1989. *cappuccino y spire: two unique maternal-effect loci required for both the anteroposterior y dorsoventral patterns of the Drosophila embryo*. Genes Dev., Vol. 3, N° 9, pp. 1437-52. <https://doi.org/10.1101/gad.3.9.1437>
- Maramorosch K, Harris K., 1979. *Leafhopper Vectors y Plant Disease Agents*. London: Academic.
- Marioni J.C., Mason C.E., Mane S.M., Stephens M., Gilad Y., 2008. *RNA-seq: an assessment of technical reproducibility y comparison with gene expression arrays*. Genome Research, Vol. 18, N° 9, pp. 1509-17. doi:[10.1101/gr.079558.108](https://doi.org/10.1101/gr.079558.108)
- McCartney N., Kondakath G., Tai A., Trimmer B.A., 2024. *Functional annotation of insecta transcriptomes: A cautionary tale from Lepidoptera*. Insect Biochem Mol Biol., Vol. 165:104038. <https://doi.org/10.1016/j.ibmb.2023.104038>
- McGettigan P.A., 2013. *Transcriptomics in the RNA-seq era*. Current Opinion in Chemical Biology, Vol. 17, N° 1, pp. 4-11. <https://doi.org/10.1016/j.cbpa.2012.12.008>.
- McIntyre L.M., Lopiano K.K., Morse A.M., Amin V., Oberg A.L., Young L.J., Nuzhdin S.V., 2011. *RNA-seq: technical variability y sampling*. BMC Genomics, Vol. 12, N° 293. <https://doi.org/10.1186/1471-2164-12-293>
- Metzker M.L., 2005. *Emerging technologies in DNA sequencing*. Genome Res., Vol. 15, pp. 1767–1776. <https://doi.org/10.1101/gr.3770505>
- Metzker M.L., 2010. *Sequencing technologies - the next generation*. Nat. Rev. Genet., Vol. 11, N° 1, pp.31-46. <https://doi.org/10.1038/nrg2626>
- Mohler J., Wieschaus E.F., 1986. *Dominant maternal-effect mutations of Drosophila melanogaster causing the production of double-abdomen embryos*. Genetics, Vol. 112, N° 4, pp. 803-22 <https://doi.org/10.1093/genetics/112.4.803>
- Morin R., Bainbridge M., Fejes A., Hirst M., Krzywinski M., Pugh T., McDonald H., Varhol R., Jones S., Marra M., 2008. *Profiling the HeLa S3 transcriptome using randomly primed cDNA y massively parallel short-read sequencing*. Biotechniques, Vol. 45, N° 1, pp. 81-94. doi:[10.2144/000112900](https://doi.org/10.2144/000112900)
- Morozova O., Hirst M., Marra M., 2009. *Applications of New Sequencing Technologies for*

- Transcriptome Analysis*. Annual Review of Genomics y Human Genetics, Vol. 10, pp. 135-151.
- Mortazavi A., Schwarz E.M., Williams B., Schaeffer L., Antoshechkin I., Wold B.J., Sternberg P.W., 2010. *Scaffolding a Caenorhabditis nematode genome with RNA-seq*. Genome Research, Vol. 20, pp. 1740-7.
- Moussian B., 2013. *The arthropod cuticle*. *Arthropod Biology y Evolution: Molecules, Development y Morphology*. Springer, pp: 171-196. doi:[10.1007/978-3-642-36160-9_8](https://doi.org/10.1007/978-3-642-36160-9_8)
- Nagalakshmi U., Wang Z., Waern K., Shou C., Raha D., Gerstein M., Snyder M., 2008. *The transcriptional landscape of the yeast genome defined by RNA sequencing*. Science, Vol. 320, pp. 1344-1349. doi:[10.1126/science.1158441](https://doi.org/10.1126/science.1158441).
- Nault L.R., Ammar E.D., 1989. *Leafhopper y Planthopper Transmission of Plant Viruses*. Annu. Rev. Entomol., Vol. 34, pp. 503–529.
<https://doi.org/10.1146/annurev.en.34.010189.002443>
- Nüsslein-Volhard C., Frohnhofer H.G., Lehmann R., 1987. *Determination of anteroposterior polarity in Drosophila*. Science, Vol. 238, Nº 4834, pp. 1675-81.
<https://doi.org/10.1126/science.3686007>
- Nüsslein -Volhard C., Roth S., 1989. *Axis determination in insect embryos*. Ciba Found Symp., Vol. 144, pp. 37-55; discussion pp. 55-64, pp. 92-8.
- Ono H., Ishii K., Kozaki T., Ogiwara I., Kanekatsu M., Yamada T., 2015. *Removal of redundant contigs from de novo RNA-Seq assemblies via homology search improves accurate detection of differentially expressed genes*. BMC Genomics, Vol.16, Nº 1031. <https://doi.org/10.1186/s12864-015-2247-0>
- Ons S., Lavore A., Sterkel M., Wulff J. P., Sierra I., Martínez-Barnetche J., Rodríguez M. H., Rivera-Pomar R., 2016. *Identification of G protein coupled receptors for opsines y neurohormones in Rhodnius prolixus*. Genomic y transcriptomic analysis. Insect Biochemistry y Molecular Biology, Vol. 69, pp. 34–50.
<https://doi.org/10.1016/j.ibmb.2015.05.003>
- Ornaghi J., Boito G., Sanchez G., March G., Beviacqua J., 1993. *Studies on the population of Delphacodes kuscheli Fennah in different years y agricultural areas*. J. Gen. Breed. Vol. 47, pp. 227–282.
- Oshlack A., Robinson M.D., Young M.D., 2010. *From RNA-seq reads to differential expression results*. Genome Biol., Vol. 11, pp. 220.

- Palacio V.G., 2018. *Secuenciación y análisis del transcriptoma de Dalbulus maidis*. Tesis Magister en Bioinformática y Biología de Sistemas, Universidad del Noroeste de la Provincia de Buenos Aires.
- Panfilio K.A., Vargas Jentzsch I.M., Benoit J.B. *et al.*, 2019. *Molecular evolutionary trends y feeding ecology diversification in the Hemiptera, anchored by the milkweed bug genome*. Genome Biol. Vol. 20, N° 64. <https://doi.org/10.1186/s13059-019-1660-0>
- Papantonis A., Swevers L., Iatrou K., 2015. *Chorion genes: a landscape of their evolution, structure, y regulation*. Annu. Rev. Entomol., Vol. 60, pp. 177-94.
- Parmentier M., Libert F., Schurmans S., Schiffmann S., Lefort A., Eggericks D., Ledent C., Molleareau C., Gerard D., *et al.*, 1992. *Expression of members of the putative olfactory receptor gene family in mammalian germ cells*. Nature, Vol. 355, pp. 453-455. doi: [10.1038/355453a0](https://doi.org/10.1038/355453a0)
- Pascual A., Vilardo E.S., Taibo C., Sabio Y García J., Pomar R.R., 2021. *Bicaudal C is required for the function of the follicular epithelium during oogenesis in Rhodnius prolixus*. Dev Genes Evol., Vol. 231, N° 1-2. pp. 33-45. <https://doi.org/10.1007/s00427-021-00673-0>
- Pascual A., Rivera-Pomar R., 2022. *Dynamics of maternal gene expression in Rhodnius prolixus*. Scientific Reports, Vol. 12, N° 6538. <https://doi.org/10.1038/s41598-022-09874-7>
- Pascual A., Calabresi F., de la Fuente D., Catalano M.I., Brentassi M.E., 2025. *Transcriptome Analysis of the Fat Body of the Maize Pest Delphacodes kuscheli (Hemiptera: Delphacidae) Reveals Essential Roles of Fungal Endosymbionts*. Microb. Ecol., Vol. 88 (1), N° 74. <https://doi.org/10.1007/s00248-025-02572-7>
- Pelosi P., Iovinella I., Zhu J., Wang G., Dani F., 2017. *Beyond chemoreception: diverse task of soluble olfactory proteins in insects*. Biological Reviews. <https://doi.org/10.1111/brv.12339>
- Perera O.P., Shelby K.S., Popham H.J.R., Gould F., Adang M.J., Jurat-Fuentes J.L., 2015. *Generation of a Transcriptome in a Model Lepidopteran Pest, Heliothis virescens, Using Multiple Sequencing Strategies for Profiling Midgut Gene Expression*. PLOS ONE, Vol. 10, N° 6, e0128563. <https://doi.org/10.1371/journal.pone.0128563>
- Perez-Gianmarco L., 2018. *Análisis Transcriptómico de los Genes Implicados en la Supervivencia de Nezara Viridula, Plaga de Cultivos Agrícolas en Argentina*. Tesis

Magister en Bioinformática y Biología de Sistemas, Universidad del Noreste de la Provincia de Buenos Aires.

Pevsner J., 2009. *Bioinformatics y Functional Genomics*. New Jersey, USA: Wiley Blackwell.

Planelló R., Herrero Ó., Gómez-Sande P., Ozáez I., Cobo F., Servia M.J., 2015. *Ecdysone Related Biomarkers of Toxicity in the Model Organism Chironomus riparius: Stage y SexDependent Variations in Gene Expression Profiles*. PLoS ONE 10(10): e0140239. <https://doi.org/10.1371/journal.pone.0140239>

Purvis A., Hector A., 2000. *Getting the measure of biodiversity*. Nature, N° 405, pp. 212-9. <https://doi.org/10.1038/35012221>. PMID: 10821281.

Qiu Z., Liu F., Lu H., Yuan H., Zhang Q., Huang Y., 2016. *De Novo Assembly y Characterization of the Transcriptome of Grasshopper Shirakiacris shirakii*. Int J Mol Sci., Vol. 17, N° 7, pp. 1110. <https://doi.org/10.3390/ijms17071110>

Reimand J., Isserlin R., Voisin V., Kucera M., Tannus-Lopes C., Rostamianfar A., Wadi L., Meyer M., Wong J., Xu C., Merico D., Bader G.D., 2019. *Pathway enrichment analysis y visualization of omics data using g: Profiler, GSEA, Cytoscape y EnrichmentMap*. Nat Protoc., Vol. 14, N° 2, pp. 482-517 <https://doi.org/10.1038/s41596-018-0103-9> PMID: 30664679; PMCID: PMC6607905.

Remes Lenicov A.M., Teson A., Dagoberto E., Huguet N., 1985. *Hallazgo de uno de los vectores del Mal de Río Cuarto en maíz*. Gaceta Agropecuaria, Vol. 5, pp. 251-258.

Remes Lenicov A.M., Virla E., Dagoberto E., 1991. *Cambios estacionales en la población del vector del “Mal de Río Cuarto” del maíz (Delphacodes kuscheli Fennah, 1955) en cultivos de avena y sus malezas circundantes en Sampacho, Córdoba (Insecta-Homoptera Fulgoroidea)*. Actas del Taller de actualización sobre Mal de Río Cuarto. INTA-CIMMYT, pp. 116- 129.

Remes Lenicov A.M., Brentassi M.E., Toledo A. V., 2008. *Description of the immature stages of Delphacodes kuscheli Fennah (Hemiptera: Delphacidae), vector of “Mal de Río Cuarto virus” on maize in Argentina*. Studies on Neotropical Fauna y Environment, Vol. 43, N° 1, pp. 25–33.

Reynolds S.E., 1980. *Integration of behavior y physiology in ecdysis*. Adv. Insect Physiol., Vol. 15, pp.:475–595. [https://doi.org/10.1016/S0065-2806\(08\)60144-7](https://doi.org/10.1016/S0065-2806(08)60144-7)

Rodríguez-Alonso G., Shishkova S., 2018. *Estudio del transcriptoma mediante RNA-SEQ con énfasis en las especies vegetales no modelo*. Revista de Educación Bioquímica

- (REB), Vol. 37, N.º 3, pp. 75-88.
- Rondón J.J., Moreyra N.N., Pisarenco V.A., Rozas J., Hurtado J., Hasson E., 2022. *Evolution of the odorant-binding protein gene family in Drosophila*. Front. Ecol. Evol., Sec. Chemical Ecology, Vol. 10. <https://doi.org/10.3389/fevo.2022.957247>
- Roth S., 2003. *The origin of dorsoventral polarity in Drosophila*. Philos Trans R Soc Lond B Biol Sci. 358, pp. 1317-29; discussion 1329.
- Rubio S., Pacheco-Orozco R.A., Gómez A.M., Perdomo S., García-Robles R., 2020. *Secuenciación de nueva generación (NGS) de ADN: presente y futuro en la práctica clínica*. Universitas Medica, Vol. 61, N.º 2. <https://doi.org/10.11144/Javeriana.umed61-2.sngs>
- Sanger F., Nicklen S., Coulson A.R., 1977. *DNA sequencing with chain-terminating inhibitors*. Proc. Natl. Acad. Sci. U.S.A., Vol. 74, pp. 5463-5.
- Santana R.A.G., Oliveira M.C., Cabral I., Silva Junior R.C.A., Teixeira de Sousa D.R., Ferreira L., Lacerda M.V.G., Monteiro W.M., Abrantes P., das Graças Vale Barbosa Guerra M., Silveira H., 2019. *Anopheles aquasalis transcriptome reveals autophagic responses to Plasmodium vivax midgut invasion*. Parasites Vectors, Vol.12, pp. 26. <https://doi.org/10.1186/s13071-019-3506-8>
- Satam H., Joshi K., Mangrolia U., Waghoo S., Zaidi G., Rawool S., Thakare R.P., Banday S., Mishra A.K., Das G., Malonia S.K., 2023. *Next-Generation Sequencing Technology: Current Trends y Advancements*. Biology (Basel), Vol. 12, N° 7, pp. 997. <https://doi.org/10.3390/biology12070997>. PMID: 37508427; PMCID: PMC10376292.
- Schmitt-Engel C., Schultheis D., Schwirz J., et al., 2015. *The iBeetle large-scale RNAi screen reveals gene functions for insect development y physiology*. Nat. Commun, Vol. 6, pp. 7822. <https://doi.org/10.1038/ncomms8822>
- Schüpbach T., Roth S., 1994. *Dorsoventral patterning in Drosophila oogenesis*. Curr. Opin. Genet. Dev., Vol. 4, pp. 502-7.
- Schwartz V., 1977. *Embriología animal comparada*; Ediciones Omega S.A.; pp. 402.
- Seemann T., Gladman S., 2012. *Fasta Statistics: Display summary statistics for a fasta file*. <https://github.com/galaxyproject/tools-iuc>
- Shen W., Sipos B., Zhao L., 2024. *SeqKit2: A Swiss Army Knife for Sequence and Alignment Processing*. iMeta, Vol. 3, N° 3. <https://doi.org/10.1002/imt2.191>

- Shigenobu S., Bickel R.D., Brisson J.A., Butts T., Chang C.-c., Christiaens O., Davis G.K., Duncan E.J., Ferrier D.E.K., Iga M., Janssen R., Lin G.-W., Lu H.-L., McGregor A. P., Miura T., Smagghe G., Smith J.M., van der Zee M., Velarde R.A., Wilson M.J., Dearden P.K., Stern D.L., 2010. *Comprehensive survey of developmental genes in the pea aphid, Acyrthosiphon pisum: frequent lineage-specific duplications y losses of developmental genes*. Insect Molecular Biology, Vol. 19, pp. 47–62.
- Simão F.A., Waterhouse R.M., Ioannidis P., Kriventseva E. V, Zdobnov E.M., 2015. BUSCO: assessing genome assembly y annotation *completeness with single-copy orthologs*. Bioinformatics, Vol. 31, pp. 3210-2.
<https://doi.org/10.1093/bioinformatics/btv351>
- Singh S., Gupta M., Pandher S., Kaur G., Goel N., Rathore P, 2019. *Using de novo transcriptome assembly y analysis to study RNAi in Phenacoccus solenopsis Tinsley (Hemiptera: Pseudococcidae)*. Scientific Reports, Vol. 9, N° 13710.
<https://doi.org/10.1038/s41598-019-49997-y>
- Smart R., Kiely A., Beale M., Vargas E., Carraher C., Kralicek A.V., Christie D.L., Chen C., Newcomb R.D., Warr C.G., 2008. *Drosophila odorant receptors are novel seven transmembrane domain proteins that can signal independently of heterotrimeric G proteins*. Insect Biochem. Mol Biol., Vol. 38, N° 8, pp. 770-80.
<https://doi.org/10.1016/j.ibmb.2008.05.002>
- Soares M.P.M., Elias-Neto M., Simões Z.L.P., Bitondi M.M.G., 2007. *A cuticle protein gene in the honeybee: Expression during development y in relation to the ecdysteroid titer*. Insect Biochemistry y Molecular Biology, Vol. 37, N° 12, pp. 1272-1282.
<https://doi.org/10.1016/j.ibmb.2007.07.014>
- Speder P., Adam G., Noselli S., 2006. *Type ID unconventional myosin controls left-right asymmetry in Drosophila*. Nature, Vol. 440, pp. 803-7.
- Stocker R.F., Singh R.N., Schorderet M., *et al.*, 1983. *Projection patterns of different types of antennal sensilla in the antennal glomeruli of Drosophila melanogaster*. Cell, Vol. 232, pp. 237–248. <https://doi.org/10.1007/BF00213783>
- Stork N.E., 2018 *How many species of insects y other terrestrial hemipodops are there on earth?* Annu. Rev. Entomol., Vol. 63, pp. 31–45.
- Tatusov R.L., Koonin E.V., Lipman D.J., 1997. *A genomic perspective on protein families*. Science, Vol. 278, pp. 631-7.
- Trapnell C., Williams B.A., Pertea G., Mortazavi A., Kwan G., van Baren M.J., Salzberg

- S.L., Wold B.J., Pachter L., 2010. *Transcript assembly y quantification by RNA-Seq reveals unannotated transcripts y isoform switching during cell differentiation*. Nature Biotechnology, Vol. 28, Nº 5, pp. 511-5.
- Truman J.W., Riddiford L.M., 1999. *The origins of insect metamorphosis*. Nature, Vol. 30; Nº 401(6752), pp. 447-52. doi: <https://doi.org/10.1038/46737>.
- Truman J.W., 2005. *Hormonal control of insect ecdysis: endocrine cascades for coordinating behavior with physiology*. Vitam. Horm., Vol. 73, pp.1-30. [https://doi.org/10.1016/s0083-6729\(05\)73001-6](https://doi.org/10.1016/s0083-6729(05)73001-6) PMID: 16399406.
- Truol G.A., Usugi T., Hirao J., Arneodo J.D., Giménez Pecci M.P., Laguna I.G., 2001. *Transmisión experimental del virus del Mal de Río Cuarto por Delphacodes kuscheli*. Fitopatología Brasileira, Vol. 26, pp.39-44.
- Valderrama Martín J.M., Ortigosa F., Cañas R.A., 2020. *Métodos de Secuenciación: Tercera Generación*. Encuentros en la Biología, Vol. XIII, Nº 175, pp. 15-21.
- Velculescu V.E., Zhang L., Zhou W., Vogelstein J., Basrai M.A., Bassett D.E., Jr. Hieter P., Vogelstein B., Kinzler K.W., 1997. *Characterization of the yeast transcriptome*. Cell, Vol. 88, pp. 243-51. [https://doi.org/10.1016/s0092-8674\(00\)81845-0](https://doi.org/10.1016/s0092-8674(00)81845-0)
- Vera J.C., Wheat C.W., Fescemyer H.W., Frilander M.J., Crawford D.L., Hanski I., Marden J.H., 2008. *Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing*. Molecular Ecology, Vol. 17, Nº 7, pp. 1636-47. doi:[10.1111/j.1365-294X.2008.03666.x](https://doi.org/10.1111/j.1365-294X.2008.03666.x)
- Virla E., de Remes Lenicov A., 1991. *Ciclo de vida de Delphacodes kuscheli criado sobre distintos hospedantes en condiciones de laboratorio (Insecta-Homoptera-Delphacidae)*. Taller de actualización sobre “Mal de Río Cuarto”. INTA. Pergamino. Bs. As. INTA-CIMMYT, pp. 104-115.
- Virla E., Díaz C., Carpane P., Laguna I., Ramallo J., Gerónimo Gómez L.G.P., 2004. *Evaluación preliminar de la disminución en la producción de maíz causada por el “Corn Stunt Spiroplasma” (CSS) en Tucumán, Argentina*. Bol. San. Veg. Plagas, Vol. 30, pp. 257-267.
- Vosshall L.B., Amrein H., Morozov P., Rzhetsky A., Axel R., 1999. *A spatial map of olfactory receptor expression in the Drosophila antenna*. Cell, Vol. 96, pp. 725-736.
- Wang E.T., Syberg R., Luo S., Khrebukova I., Zhang L., Mayr C., Kingsmore S.F., Schroth G.P., Burge C.B., 2008. *Alternative isoform regulation in human tissue transcriptomes*. Nature, Vol. 456, pp. 470-6.

- Wang F., Xu Y., Wang R., Zhang B., Smith N., Notaro A., Gaerlan S., Kutschera S., Kadash-Edmondson K. E., Xing Y., Lin L., 2023. *TEQUILA-seq: a versatile y low-cost method for targeted long-read RNA sequencing*. Nat. Commun, Vol. 14, N° 4760. <https://doi.org/10.1038/s41467-023-40083-6>
- Wang Y., Ghaffari N., Johnson C.D., Braga-Neto U.M., Wang H., Chen R., Zhou H., 2011. *Evaluation of the coverage y depth of transcriptome by RNA-Seq in chickens*. BMC Bioinformatics, Vol. 12 Suppl. 10, N° 5. <https://doi.org/10.1186/1471-2105-12-S10-S5>
- Wang Z., Gerstein M., Snyder M., 2009. *RNA-Seq: a revolutionary tool for transcriptomics*. Nature Reviews Genetics, Vol. 10, pp. 57-63.
- Wanninger A., 2015. *Evolutionary Developmental Biology of Invertebrates 5*, in: Wanninger, A. (Ed.), Ecdysozoa III: Hexapoda. Springer-Verlag Wien, pp. VII, 219.
- Waterhouse R.M., Seppey M., Simão F.A., Manni M., Ioannidis P., Klioutchnikov G., Kriventseva E.V., Zdobnov E.M., 2017. *BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics*. Mol. Biol. Evol. Vol. 35, N° 3, pp. 543–548 <https://doi.org/10.1093/molbev/msx319>
- Wesolowski S., Birtwistle M.R., Rempala G.A., 2013. *A Comparison of Methods for RNA-Seq Differential Expression Analysis y a New Empirical Bayes Approach*. Biosensors (Basel), Vol. 3, N° 3, pp. 238-58. <https://doi.org/10.3390/bios3030238>
- Wigglesworth V.B., 1934. *The physiology of ecdysis n Rhodnius prolixus (Hemiptera). II. Factors controlling moulting y metamorphosis*. Quarterly Journal of Microscopic Science, Vol. 77, pp. 191-222.
- Wilhelm B.T., Briau M., Austin P., Faubert A., Boucher G., Chagnon P., Hope K., Girard S., Mayotte N., Landry J.R., Hébert J., Sauvageau G., 2011. *RNA-seq analysis of 2 closely related leukemia clones that differ in their self-renewal capacity*. Blood, Vol. 117, N° 2, pp. 27-38. <https://doi.org/10.1182/blood-2010-07-293332>
- Wilson S.W., 2005. *Keys to the families of Fulgoromorpha with emphasis on planthoppers of potential economic importance in the southeastern United States (Hemiptera: Auchenorrhyncha)*. Florida entomologist, Vol. 88, pp. 464-481.
- Wingett S. W., Andrews S., 2018. *FastQ Screen: A tool for multi-genome mapping y quality control*. F1000Research, 7, 1338. <https://doi.org/10.12688/f1000research.15931.2>

- Wolf J., 2013. *Principles of transcriptome analysis y gene expression quantification: An RNA-seq tutorial*. Molecular Ecology Resources, Vol. 13, Nº 4, pp. 559- 572.
<http://doi.org/10.1111/1755-0998.12109>
- Xue J., Bao Y.Y., Li B.L., Cheng Y.B., Peng Z.Y., Liu H., Xu H.J., Zhu Z.R., Lou Y.G., Cheng J.A., Zhang C.X., 2010. *Transcriptome analysis of the brown planthopper Nilaparvata lugens*. PLoS One, Vol. 5, Nº 12.
<https://doi.org/10.1371/journal.pone.0014233>
- Yan F., Baldoni P.L., Lancaster J., Ritchie M.E., Lewsey M.G., Gouil Q., Davidson N.M., 2026. *A comprehensive evaluation of long-read de novo transcriptome assembly*. Genome Biol., <https://doi.org/10.1186/s13059-026-04001-5>
- Yang Y., Smith S.A., 2013. *Optimizing de novo assembly of short-read RNA-seq data for phylogenomics*. BMC Genomics, Vol. 14, Nº 328. <https://doi.org/10.1186/1471-2164-14-328>
- Yassour M., Kaplan T., Fraser H.B., Levin J.Z., Pfiffner J., Adiconis X., Schroth G., Luo S., Khrebtkova I., Gnirke A., Nusbaum C., Thompson D.A., Friedman N., Regev A., 2009. *Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing*. Proceedings of the National Academy of Sciences of the United States of America, Vol. 106, pp. 3264-9. <https://doi.org/10.1073/pnas.0812841106>
- Ylla G., Piulachs M.D., Belles X., 2018. iScience, Vol. 4, pp. 164-179.
[https://www.cell.com/iscience/fulltext/S2589-0042\(18\)30072-5](https://www.cell.com/iscience/fulltext/S2589-0042(18)30072-5)
- Yu Y., Zhan Q., Yang T., Teng W., 2021. *The Influence of Autoclaving on Diamond Burs Fabricated via Different Methods*. Int. J. Prosthodont., Vol. 34, Nº 4, pp. 451–457.
<https://doi.org/10.11607/ijp.6231>
- Zhang B. X., Huang H. J., Yu B., Lou Y. H., Fan H. W., Zhang C. X., 2015. *Bicaudal-C plays a vital role in oogenesis in Nilaparvata lugens (Hemiptera: Delphacidae)*. Journal of Insect Physiology, Vol. 79, pp. 19–2.
- Zhao Q.Y., Wang Y., Kong Y.M., Luo D., Li X., Hao P., 2011. *Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study*. BMC Bioinformatics, Vol. 12, Suppl. 14 (Suppl 14): S2. <https://doi.org/10.1186/1471-2105-12-s14-s2>
- Zhou S.S., Sun Z., Ma W., Chen W., Wang M.Q., 2014. *De novo analysis of the Nilaparvata lugens (Stål) antenna transcriptome and expression patterns of olfactory*

genes. Comp Biochem Physiol Part D Genomics Proteomics, Vol. 9, pp. 31-39.

<https://doi.org/10.1016/j.cbd.2013.12.002>

Zhu J., Jiang F., Wang X., Yang P., Bao Y., Zhao W., Wang W., Lu H., Wang Q., Cui N., Li J., Chen X., Luo L., Yu J., Kang L., Cui F., 2017. *Genome sequence of the small brown planthopper, Laodelphax striatellus*. GigaScience, Vol. 6, N° 12, pp. 1-12.

<https://doi.org/10.1093/gigascience/gix109>

Zitnan D., Kingan T.G., Hermesman J.L., Adams M.E., 1996. *Identification of ecdysis-triggering hormone from an epitracheal endocrine system*. Science, Vol. 271, N° 5245, pp. 88-91. doi: <https://doi.org/10.1126/science.271.5245.88>

Anexo I:

Comandos

```
blastn -query R2_1.fasta -db UniVec/UniVec -reward 1 -penalty -5 -gapopen 3 -gapextend 3
-dust yes -soft_masking true -evaluate 0.00001 -searchsp 17500000000000 -outfmt 6
-out R2_1_hits_univec.fasta
```

Comando 1. Blastn vs. UniVec. Este comando se ejecuta en busca de secuencias contaminantes para ambos archivos, *forward* y *reverse*.

```
1 #!/usr/bin/python
2
3 # Se importan librerías de la biblioteca BioPython para procesar las secuencias.
4 from Bio import SeqIO
5 from Bio.Seq import Seq
6 from Bio.SeqRecord import SeqRecord
7
8 # Archivo Input de datos crudos.
9 fasta_file = "R2_1.fastq" # También para R2_2.fastq
10
11 # Archivo de Hits contra univec. Secuencias sin repetición.
12 univec_file = "R2_1_hits_univec_unique.txt" # También para R2_2_hits_univec_unique.txt
13
14 # Descarte. Secuencias que por como dieron Hit, se eliminan del archivo a usar para ensamblar.
15 result_file_descarte = "R2_1_Descarte.fastq" # Con R2_2.fastq, resulta R2_2_Descarte.fastq
16
17 # Secuencias que se mantienen para realizar el ensamble posterior (No dieron Hit, o el Hit que
18 # dieron fue eliminado).
19 result_file_limpias = "R2_1_Limpias.fastq" # Con R2_2.fastq, resulta R2_2_Limpias.fastq
20 # ----- Comienzo ----- #
21
22 # Se crea un set de datos y el archivo de hits se vuelca en él.
23 univec = set()
24 univec_aux = set()
25 format = "fastq"
26 mark_seq = "@"
27
28 # El archivo de hits contra UniVec se lleva a memoria para mas fácil acceso.
29 with open(univec_file, "r") as f:
30     univec = [line.strip().split("\t") for line in f]
31
32 f.close()
33
34 # Se abre el archivo a procesar (Datos crudos en formato fastq).
35 fasta_sequences = SeqIO.parse(open(fasta_file), format)
36
37 file_clean = open(result_file_limpias, "w")
38 file_dispose = open(result_file_descarte, "w")
39
40 # Set auxiliar para conservar hits que no se ignoran.
41 hits_parciales = ()
42
43 # Inicialización de contadores para mostrar log.
44 contador1 = 0 # Contador de secuencias limpias.
45 contador2 = 0 # Contador de secuencias cortadas al inicio.
46 contador3 = 0 # Contador de secuencias cortadas al final.
47 contador4 = 0 # Contador de secuencias descartadas.
48 contador5 = 0 # Suma total de secuencias que se usarán para el ensamble.
49 contador6 = 0 # Contador total.
50
51 # Recorrido de secuencias de los datos crudos.
52 for seq in fasta_sequences:
53
54     # Obtiene datos de interés de la lectura.
55     aux_seqid = seq.id
```

```

56     aux_seqid = mark_seq + aux_seqid.strip()
57     aux_seq = str(seq.seq)
58     aux_seq = aux_seq.strip()
59     aux_seqdescription = seq.description
60     aux_seqdescription = mark_seq + aux_seqdescription.strip()
61
62     len_seq = len(aux_seq) # Cadena ATCG
63
64     contador6 = contador6 + 1
65
66     Encontro_en_univec = "N"
67     # Recorrido de hits contra UniVec.
68     for line_univec in univec:
69         # Descompone la lectura.
70         qsecid = line_univec[0] # ID de secuencia
71         length = line_univec[3] # Longitud del hit
72         qstart = line_univec[6] # Inicio del hit
73         qend = line_univec[7] # Fin del hit.
74
75         # Se verifica si la secuencia de los datos crudos
76         # hace match contra UniVec.
77         if seq.id == qsecid:
78
79             #print("secuencia " + seq.id + " match Univec")
80             Encontro_en_univec = "S"
81
82             # Verifica si el fin del hit es menor a 40, en este caso, se
83             # recortan las primeras 39 PB.
84             if int(qend) < 40:
85                 contador2 = contador2 + 1
86                 contador5 = contador5 + 1
87                 sub_seq = seq[int(qend):int(len_seq)]
88
89                 # Se escribe la secuencia recortada en archivo "Clean".
90                 SeqIO.write(sub_seq, file_clean, format)
91
92             # Verifica si el inicio del hit es mayor a 70, en este caso, se
93             # recortan las últimas 30 PB.
94             elif int(qend) >= 40 and int(qstart) > 70:
95                 contador3 = contador3 + 1
96                 contador5 = contador5 + 1
97                 sub_seq = seq[0:(int(qstart)-1)]
98
99                 # Se escribe la secuencia recortada en archivo "Clean".
100                 SeqIO.write(sub_seq, file_clean, format)
101
102             else:
103                 # Si el hit no cumplió las condiciones para ser editado,
104                 # se descarta la lectura.
105                 contador4 = contador4 + 1
106
107                 # Se escribe la secuencia en "Dispose".
108                 SeqIO.write(seq, file_dispose, format)
109
110             if Encontro_en_univec == "N":
111                 contador1 = contador1 + 1
112                 contador5 = contador5 + 1
113
114                 # Si no está en el archivo de hits contra UniVec, se graba como Limpia.
115                 SeqIO.write(seq, file_clean, format)
116
117 file_dispose.close()
118 file_clean.close()
119
120 # Imprime log.
121 print("")
122 print("Proceso finalizado...")
123 print("")
124 print("-----|")
125 print("Total de secuencias : " + str(contador6).rjust(9, " ") + " |")
126 print("-----|")
127 print("Secuencias sin contaminantes : " + str(contador1).rjust(9, " ") + " |")
128 print("Secuencias cortadas al inicio : " + str(contador2).rjust(9, " ") + " |")
129 print("Secuencias cortadas al final : " + str(contador3).rjust(9, " ") + " |")
130 print("-----|")
131 print("Secuencias limpias : " + str(contador5).rjust(9, " ") + " |")
132 print("Secuencias descartadas : " + str(contador4).rjust(9, " ") + " |")
133 print("-----|")

```

Comando 2. Limpieza de contaminantes. Script codificado en lenguaje *Python* utilizado para la remoción de secuencias contaminantes de los archivos *forward* y *reverse*, halladas en la base de datos de UniVec.

```
seqkit pair -1 R2_1_Limpas.fastq -2 R2_2_Limpas.fastq -u
[INFO] 105880302 paired-end reads saved to R2_1_Limpas.fastq and R2_2_Limpas.fastq
[INFO] 606 unpaired reads saved to R2_1_Limpas.unpaired.fastq
[INFO] 222 unpaired reads saved to R2_2_Limpas.unpaired.fastq
```

Comando 3. SeqKit. Este comando se ejecuta para recuperar la condición de *paired-end* entre ambos archivos, *forward* y *reverse*.

```
java -jar /usr/share/java/trimmomatic-0.39.jar
PE R2_1_Limpas.fastq R2_2_Limpas.fastq out_forward_paired.fq.gz out_forward_unpaired.fq.gz out_
reverse_paired.fq.gz out_reverse_unpaired.fq.gz ILLUMINACLIP:TruSeq3-PE.fa:2:30:10:2:keepBothRead
s LEADING:3 TRAILING:3 MINLEN:36
TrimmomaticPE: Started with arguments:
R2_1_Limpas.fastq R2_2_Limpas.fastq out_forward_paired.fq.gz out_forward_unpaired.fq.gz out_re
verse_paired.fq.gz out_reverse_unpaired.fq.gz ILLUMINACLIP:TruSeq3-PE.fa:2:30:10:2:keepBothReads
LEADING:3 TRAILING:3 MINLEN:36
Multiple cores found: Using 3 threads
Using PrefixPair: 'TACACTCTTTCCTACACGACGCTCTTCGATCT' and 'GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT'
ILLUMINACLIP: Using 1 prefix pairs, 0 forward/reverse sequences, 0 forward only sequences, 0 reve
rse only sequences
Quality encoding detected as phred33
Input Read Pairs: 105880302 Both Surviving: 105875767 (100,00%) Forward Only Surviving: 4535 (0,0
0%) Reverse Only Surviving: 0 (0,00%) Dropped: 0 (0,00%)
TrimmomaticPE: Completed successfully
```

Comando 4. Ejecución de Trimmomatic. Este comando se ejecuta para eliminar las secuencias *singles* producidas por la eliminación de contaminantes de los archivos *forward* y *reverse*.

```
signalp-5.0b/bin# ./signalp -format short
-org euk -fasta ../../TransDecoder_Longest_ORFs_PEP.fasta -prefix signalp.out
SignalP-5.0. Starting fasta file reading...
Total proteins read: 18607.
Organism: euk.
Starting protein prediction...
Temporary directory: /tmp/signalp5.792187987
Protein prediction done!
Generating output files...
Completed.
```

Comando 5. Ejecución de SignalP. Este comando se ejecuta en busca de péptido señal.

A	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastp_TransDecoder_Longest_ORFs_vs_Laodelphax_striatellus.tabular --prog blastp --dbtype Lstriatellus</code>
B	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastp_TransDecoder_Longest_ORFs_vs_Nilaparvata_lugens.tabular --prog blastp --dbtype Nlugens</code>
C	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastp_TransDecoder_Longest_ORFs_vs_Sogatella_furcifera.tabular --prog blastp --dbtype Sfurcifera</code>
D	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastp_TransDecoder_Longest_ORFs_vs_Uniprot_Insecta.tabular --prog blastp --dbtype Insecta</code>
E	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastp_TransDecoder_Longest_ORFs_vs_uniprotkb_arthropoda.tabular --prog blastp --dbtype Arthropoda</code>
F	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastx_Trinity_Assembled_Transcripts_sin_isoformas_vs_Laodelphax_striatellus.tabular --prog blastx --dbtype Lstriatellus</code>
G	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastx_Trinity_Assembled_Transcripts_sin_isoformas_vs_Nilaparvata_lugens.tabular --prog blastx --dbtype Nlugens</code>
H	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastx_Trinity_Assembled_Transcripts_sin_isoformas_vs_Sogatella_furcifera.tabular --prog blastx --dbtype Sfurcifera</code>
I	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastx_Trinity_Assembled_Transcripts_sin_isoformas_vs_Uniprot_Insecta.tabular --prog blastx --dbtype Insecta</code>
J	<code>perl Trinotate Trinotate_sqlite_database.sqlite LOAD_custom_blast --outfmt6 blastx_Trinity_Assembled_Transcripts_sin_isoformas_vs_uniprotkb_arthropoda.tabular --prog blastx --dbtype Arthropoda</code>
K	<code>perl Trinotate Trinotate_sqlite_database.sqlite report > trinotate_report_complete.txt</code>

Comando 6. Carga de datos a Trinotate. Este comando se ejecuta para agregar a la base de datos SQLite, resultante de la ejecución de Trinotate, los resultados de *Blastp* y *Blastx* contra las bases de datos Artrópoda e Insecta, y contra las proteínas predichas correspondientes a los genomas de las especies *L. striatellus*, *N. lugens* y *S. furcifera*. Agregado de resultado de *Blastp* vs. *L. striatellus* (A), vs. *N. lugens* (B), vs. *S. furcifera* (C), vs. Insecta (D) y vs. *Arthropoda* (E), y de *Blastx* vs. *L. striatellus* (F), vs. *N. lugens* (G), vs. *S. furcifera* (H), vs. Insecta (I) y vs. *Arthropoda* (J). Con los datos agregados a la base de datos SQLite, se genera el reporte completo (K).

A

```
# ==== Read trinode format
file <- "trinode_report_complete.txt"
x <- read_trinode(file)
summ <- summary_trinode(x)

# La información de Trinode se descarga a csv para exposición.
trinode_tabla <- data.frame(summ)
view(trinode_tabla)
write.csv(trinode_tabla, "trinode_tabla.csv")

write.table(summ, file=paste0(outpath, "/", file, ".resumen.tsv"), sep="\t", row.names = F, col.names = T)
```

B

```
data(cogs)
download.file("https://raw.githubusercontent.com/RJEGR/infovis/master/NOG.annotations.tsv", "NOG.annotations.tsv")
egg <- read.table("NOG.annotations.tsv", sep="\t", stringsAsFactors=FALSE, quote="")
names(egg) <- c("db", "nog", "proteins", "species", "class", "description")

head(egg)

#plot_NOGs(x, by="gene_id")

png(filename = paste0(file, '_plot_NOGs1.png'), height = 30,width = 40,res = 300,units = "cm")
plot_NOGs(x, by="gene_id")
dev.off()
```

C

```
go <- split_GOP(x)
goP <- summary_GO(go)

plotgos <- head(goP[order(-goP$transcripts),], 100)

bar <- ggbarplot(plotgos, "name", "transcripts",
  fill = "ontology",
  color = "ontology",
  xlab = "Nombre",
  ylab = "Número de transcritos",
  palette = "aaas", # jco journal color palett. see ?ggpar
  sort.val = "asc", # sort the value in descending order
  sort.by.groups = TRUE,
  title = "Mayores ontologías (top 100)")

png(filename = paste0(file, '_plot_NOGs.png'), height = 28,width = 46,res = 600,units = "cm")
bar + theme(axis.text.x = element_text(angle = 75, hjust = 1, size = 10))
dev.off()
```



```
# Elección de columnas de datos a graficar.
# Uniprot ==> sprot_Top_BLASTX_hit
# Arthropoda ==> Arthropoda_BLASTX
# Insecta ==> Insecta_BLASTX
# S. furcifera ==> Sfurcifera_BLASTX
# L. striatellus ==> Lstriatellus_BLASTX
# N. lugens ==> Nlugens_BLASTX
blastx <- split_blast(x, "sprot_Top_BLASTX_hit")
sblastx <- summary_blast(blastx)

# Uniprot ==> sprot_Top_BLASTX_hit
# Arthropoda ==> Arthropoda_BLASTX
# Insecta ==> Insecta_BLASTX
# S. furcifera ==> Sfurcifera_BLASTX
# L. striatellus ==> Lstriatellus_BLASTX
# N. lugens ==> Nlugens_BLASTX
blastp <- split_blast(x, "sprot_Top_BLASTP_hit")
sblastp <- summary_blast(blastp)
# Se agrega la lectura de la columna "Genus" separada.
trinotate_report_blastx.genus <- read.delim("trinotate_report_complete.txt.blastx - genus.tsv", quote="", na = c("ERROR"))
trinotate_report_blastx.genus <- na.omit(trinotate_report_blastx.genus)
view(trinotate_report_blastx.genus)

xgenus2 <- data.table(table(trinotate_report_blastx.genus$genus))
view(xgenus2)
names(xgenus2) <- c("Genus", "Number")
xgenus2 <- xgenus2[order(-xgenus2$Number), ]
xgenus2[(Number <= round(mean(xgenus2$Number))), genus := "others"]
length(table(xgenus2$genus))

plot2 <- rbind(genus, xgenus2)
ggbarplot(plot2, x = "Genus", y = "Number",
  fill = "Type", # change fill color by cyl
  color = "Type", # Set bar border colors to white
  palette = "Paired", # jco journal color palett. see ?ggpar
  group = "Type",
  sort.by.groups = TRUE, # Sort inside each group
  x.text.angle = 90, # Rotate vertically x axis texts
  xlab = "Género",
  ylab = "Número de anotaciones", # Number of annotations
  rotate = TRUE,
  ggtheme = theme_classic(),
  #facet.by = "Type"
) + theme(axis.text.y = element_text(hjust = 1, size = 8))
```

E

```
xy <- data.frame(Identity=blastx$identity, Type="Transcript")
y <- data.frame(Identity=blastp$identity, Type="ORF")

data <- rbind(xy,y)

ggplot(data, aes(Identity, fill = Type)) +
  geom_histogram(bins = 100, alpha = 0.7, aes(y = ..density..), position = 'identity') +
  theme_classic() +
  scale_fill_brewer(direction = -1, palette = "Paired") +
  stat_function(fun=dnorm,
               color="red",
               args=list(mean=mean(data$Identity, na.rm = TRUE),
                        sd=sd(data$Identity, na.rm = TRUE))) +
  scale_x_continuous("Identidad del alineamiento") # Identity of the alignment
geom_histogram(binwidth = 2.5) +
  geom_density(aes(y=2.5 * ..count..)) #<--- para emparejar ambos, histograma y densidad
x <- blastx$identity
y <- blastp$identity

h1 <- hist(x, breaks=100, col=rgb(1,0,0,1/4),
          xlab="Identidad del alineamiento (Blastp hits)",
          ylab="Frecuencia",
          main="Distribución de la identidad")
```

Comando 7. Gráficos de Trinotate. Mediante este script en lenguaje R, se realizaron varios gráficos y tablas a partir de la información obtenida de Trinotate. Se observa el script realizado para la base de datos de Uniprot, realizando la misma ejecución para las bases de datos de *Arthropoda* e *Insecta*, y para las proteínas predichas para las especies *L. striatellus*, *N. lugens* y *S. furcifera*, modificando determinados parámetros convenientemente para procesar los datos deseados. (A) Generación de tabla con los resultados obtenidos de Trinotate. (B) Categoría funcional EggNOG. (C) Función ontológica (GO). (D) Transcritos encontrados mediante *Blastx* y *Blastp*. (E) Distribución de identidad.

Anexo II:

Figuras

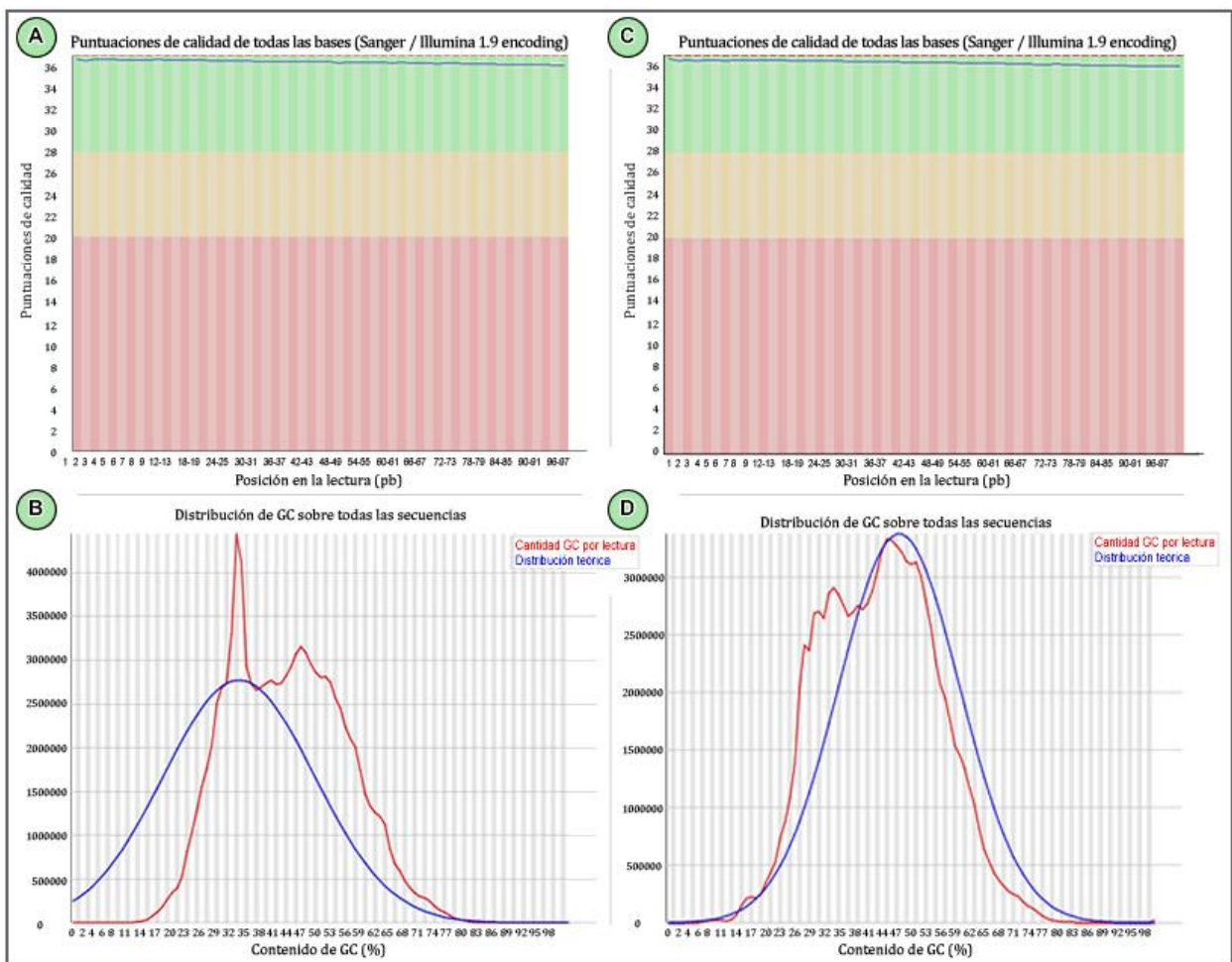


Figura 1. Análisis de calidad de las lecturas luego de la eliminación de secuencias contaminantes, mediante FastQC. Distribución de *Phred Score* en función de la posición a lo largo de cada lectura, para *forward* (A) y *reverse* (C). En el eje X se representan las distintas posiciones de la lectura en pares de bases. En el eje Y se representan los valores de calidad posibles, abarcando valores de 0 a 36 (Valor Q). Contenido de GC por lectura (rojo) vs. la distribución normal modelada (azul), para *forward* (B) y *reverse* (D). En el eje X se representa el porcentaje de contenido GC. En el eje Y se representa la distribución de GC sobre las secuencias.

Anexo III:

Bases de Datos

Generadas

Ecdisis

>sp|P85316|BURS_AEDAE

MKSSVCVLLKVLACILLPGGLNAQKESNDDIQHYTADDCQVTPVIHVLQYPGCVPKPIPSFACIGRCASYIQVSGSKI
WQMERSMCCQESGEREASVSLFCPKAKNGEKKFKVSTKAPLECMCRPCTGIEDANVIPQELAAFADDGSLAGYFQK
SQFKAEE

>sp|Q66Q82|BURS_ANOGA

MKSTFLVVLELAFFLLPGRVLYAQKDESDGGSHYSSDDCQVTPVIHVLQYPGCVPKPIPSFACIGRCASYIQVSGSKI
WQMERSMCCQESGEREASVSLFCPKAKNGEKKFKVSTKAPLECMCRPCTGIEDANVIPQELTSFADEGTLTGYPQK
SHYKSIE

>sp|P85315|BURS_CULPP

MISSPSTPATFAAGSLVLLCLVLGGGHFALAQKEGNDDIQHYTADDCQVTPVIHVLQYPGCVPKPIPSFACVGRASY
IQVSGSKIWQMERSMCCQESGEREASVSLFCPKAKNGEKKFKVSTKAPLECMCRPCTGIEDANVIPQELAAFADDG
TLTSYFQKQQLRNSE

>sp|Q4FCM6|BURS_MANSE

MYALDFLFIAFVYFAACHIQEKPVRAQEVQLPPGQECQMTPIHILKHRGCKPKAIPSFACIGKCTSYVQVSGSKIWQ
MERTCNCCQEAGEREATVVLYCPDAKNEERRFRKVSTKAPLECMCRPCGSIEESAIIPQEVAGYAEEGPLYNHFRKSF

>sp|A0A143WDY5|BURS_BLAGE

MACQKSTSTQQIVTGVLVALVYVVLVGAMDECQVTPVIHVLQYPGCVPKPIPSFACTGRCSSYLQVSGSKIWQMERSM
MCCQESGEREASVSLFCPKAKAGERKFRKVTTKAPLECMCRPCTTVEESAVIPQEIAGYADEGPLSNHFRKSL

>sp|A0A3Q8KJL9|BURS_NEZVI

MRIFAFLAQLLVIAAYCQNDGTRKPADECQVTPVIHVLQYPGCVPKPIPSFACTGRCSSYIQVSGSKIWQMERSMCCQ
ESGEREASVTLFCPKAKPGEKKFRKVMTKAPLECMCRPCTSVESAVIPQEIAGLTDEGPLNQHFVKPQ

>sp|A0A4Y2SU0|BURS_ARAVE

MGYKLGIIIDTCFFICLLACFWCTTFMPAVKARGECHLRPVIHVLQOPGCIPKPIPSFACQGSCSSYVQVSGSKFWQ
VERSCMCCQEMGEREASIGIFCPKASPRFRKIVTRAPVECMCRPCTAVDESAIQPQEMQLIIADPQADFSKGMPLMED
YHLLNELKNGKDIPVV

>sp|UniRef100_UPI0007E5FF76|BURS_DFIC

MLRHLLRHENNKVFLVILLYCVLVSILKLCTAQAEENSLTATDNDITHLGDDCQVTPVIHVLQYPGCVPKPIPSFACVG
RCASYIQVSGSKIWQMERSMCCQESGEREAAVSLFCPKVKPGERKFKKVLTKAPLECMCRPCTSVESAVIPQEIAG
YSDEGPLNNHFRRIALQ

>sp|U3U8R1|BURS_NILLU

MTSTIQILMHGAVVIAMMLAESFQEDAKPAVSSDECQVTPVIHVLQYPGCVPKPIPSFACTGKCSSYLQVSGSKIWQM
ERSCMCCQESGEREASVSLFCPKAKPGEKKFRKVNTKAPLDCMRPCTGVEETSVIPQEIAGYADEGPMNNHFRHLSS
R

>sp|A0A345BED3|BURS_LAOST

MTCTIQIFIHGAVVLAMMLAESFQDEAKPPAGAVSSDECQVTPVIHVLQYPGCVPKPIPSFACTGKCSSYLQVSGSKI
WQMERSMCCQESGEREASVSLFCPKAKPGEKKFRKVNTKAPLDCMRPCTGVEETSVIPQEIAGYADEGPMNNHFRHL
HSSR

>Burs_Dalbulus_Maidis

CELFTYGVIVKVLVALLWACVLGADSPKSSPVSSDECQVTPVIHVLQYPGCVPKPIPSFACTGRCSSYLQVSGSKIWQM
ERSCMCCQESGEREASVSLFCPKAKAGERKFRKVITKAPLECMCRPCTGVEESAVVPQEIAGYTDEGPLNAHFQKSQ

>sp|A0A1B6GB49|BURS_9HEMI

MMGFVHTFVDSVLASILLTYVVTVTENDASKPSDECQVTPVIHVLQYPGCVPKPIPSFACTGRCSSYLQVSGSKIWQM
ERSCMCCQESGEREASVSLFCPKAKAGERKFRKVITKAPLECMCRPCTGVEESAVVPQEIAGYTDEGPLNAHFQKS

>sp|F1AQ58|BURS_PROCL

MGRISWVLVVMGMSMGVVSGDECSLTPVIHILSYQGCTSKPIPSFACQGRCTSYVQVSGSKLWQTERSCMCCQESGE
REASVILNCPNARKGEPKQKKILTRAPIDCMCRPCTDVEEGTVLAQEIANFIHDSMPGNVPFLK

>sp|A0A4D6BNJ4|BURS_NEPNO

MGGMSWVLMVLGVATVVWSDECSLTPVIHILSYPGCVSKPIPSFACQGRCTSYVQVSGSKLWQTERSCMCCQESGERE
ASVVLNCPKVRKGEPTRRKILTRAPIDCMCRPCTDVEEGTVLAQEIANFIHDSMPGPNVFLK

>sp|A0A212ES77|BURS_DANPL

MFVMRICFVSVFFLICYLGTKDVNGQEIQLPPGQECQMTFVIHVLQHPGCVPKPIPSYACIGKCTSYVQVSGSKIWQM
ERSCNCCQESGEREASVVLFCPKAKSEEKFRKRVSTKAPLECMCRPCGSVEESSVIPQEIAGYADEGLLQSHFRNSL

>sp|A0A0H3YID2|BURS_LOCM

RPDASQGGAEQAARGPADECQLTPVIHVLQYPGCVPKPIPSFACTGRCSSYLQVSGSKIWQMERSMCCQESGEREA
SVSLFCPKAKAGERKFRKRVSTKAPLECMCRPCTGVEESAVIPQEIAGYPDNGPLTAHFRKSQ

>sp|Q8WRC7|CCAP_MANSE

MTVSRVCLLLLVALVYLDCCYAASIPRNFDPRLSEEIVMAPKKRPFPCNAFTGCGRKRSQGPPGMPAQDLRTKQYLDEE
ALGSILDSESAIDELSRQILSEAKLWEAIQEASAEIARRKQKEAYIQ

>sp|Q9VCW0|CCAP_DROME

MRTSMRISLRLLALLACAICSQASLERENNEGTMANHKLSGVIQWKYEKRPFCNAFTGCGKRRTYPSYPPFSLFKRN
EVEEKPYNNEYLSEGLSDLIDINAEPAVENVQKQIMSQAKIFEAIKEASKEIFRQKNKQKMLQNEKEMQQLEERESK

>sp|Q75UG5|CCAP_PERAM

MQMYHVVLCGSLAILLVILDIPQASCDDVVIQKRQVDPAMDRLLDPKRKRPFPCNAFTGCGKKRSDESMGTLVEMNSE
PAVDDLRSQILSEVKLWEAIQEARVELLRRLRRQEQQLQQQSNQFGAGMDRPLPLPIAGYRRKRFPADPESQAPAPHSNLPR
ATSQEQEEITKPWSR

>sp|A0A1S3DKB5|CCAP_DIACI

MRRYMLVLVSSILTIFILIFIPTVKSDDVIMQKRDFERGYDMLDPKRKKPFPCNAFTGCGKKRSASESMGTLVEMNSEP
AVEELSRQILSEAKLWEAIQEARLELLRRREKHEQAERQMOMQMESNMLGDNRPMSFTGFRKRRSASSNC

>sp|A0A0S2PZN7|CCAP_PENVA

MSTMNWCGRAGVIAAASLLLLLLAAHSHAGPVAKRDIADLLDGKDKRPFPCNAFTGCGKKRSEPKVEPLSSGAELDALA
KHVVAEAKLWEQLQNKMEVMRSLAARMEEHPIFRKRSPAPAHAPQAAAPSQNQQMDTDKQ

>sp|Q7PM76|CCAP_ANOGA

MTHRTAAKLLAVVSLFCVLQMLECGVVDRQPRAYKQYNTEPQKRPFPCNAFTGCGKKRSSASSPPTAAAAAAAAMLQ
RHLQTMDPGRKDRMGSDFDPNEEISSLLDLNTEPAVEDLLRQIMSEAKLWEAIQEANREIYLQKSGMKDQRNDFPLT
FSTQ

>sp|A0A034VGY1|CCAP_BACDO

MFAWKLWSMIWLLCLHSTLQNMNDNDLNHKLRLGIIQWKYEKRPFCNAFTGCGKKRSSLPYPLWKRADLDDKMYNNED
MSEGLSDLIDINAEPAVENVQKQIMSQAKIFEAIKEASKEIFRQKNMQKQRQRETNPETIERV

>sp|U3U430|CCAP_NILLU

MQPHVAVICAMAALLCLLTQPTNCDDVIIQKRVDSDLAIFDPKRKKPFPCNAFTGCGKRKRSDESMATLVDLNSEPAV
EELSRQILSEAKLWEAIQEARQELLRRRQEGKLPPIPYTLTGMRKKRSPEAPLTN

>sp|A0A345BED7|CCAP_LAOST

MERAQVKSKMQLHVALCCAMAALLRLLTQPTNCDDVIIQKRVDSDLAIFDPKRKKPFPCNAFTGCGKRKRSDESMATL
VDLNSEPAVEELSRQILSEAKLWEAIQEARQELMRRRQEGKLPPIPYTLTGMRKKRSPEKPVTK

>sp|Q1WK20|CCAP_HOMGA

MTNMSWCGRVGILGVTTVLLLVLLAAHAQAGPVAKRDIQDLLEGKDKRPFPCNAFTGCGKKRSDPSMEGLASSSELDAL
AKHVLAEAKLWEQLQSKMEMMRSYASRMENHPVYRRKRSTPHTQPRQHLTSTPQQKVETEKQ

>sp|ABB46293.1|CCAP_FAIMM

MSTISTHGRAGVMVLTALLLVLAHAHAGPLVKRDIQDLLEGKDKRPFPCNAFTGCGKKRSDPGLEGVASSSELDALA
KHVLAEAKLWEQLQNKMEVMRSLAARMENHPLYRRRRSAPQQPRHLLTTTPKQKVESEKQ

>sp|D2J4U0|CCAP_RHOPR

MQLLVPCFLLFTALVSAVLTDDVFLQKRLYFPGEIAEPIDPKMKKPFPCNAFTGCGKKRSDPSMATLVDLNSEPAVEEL

SRQILSEAKLWEAIQEARMELLNRKQQQSDRIPLQPLPLTTIRKRSHYLYT

>sp|D6WA20|CCAP_TRICA

MTTAKLFVICIFAALAIETHSRFLPKSISKNLGATERVLEPKKRPFCNAFTGCGKRKSNLPALPEQSEVV DENLGSLLELSAEPAVEDLSRQIMSEAKLWEAIQEANMELHRRRQESAESSEEDA AVPARSATASCALPPCYI

>sp|A0A0K8V341|CCAP_BACLA

MFAWKLWSMIWLLCLHSTLQIEMDSDDL NHKLRGIIQWKYEKRPFCNAFTGCGKKRSSLPYPLWKRADLDDVKMYNNE DMSEGLSDLIDINAEP AVENVQKQIMSQA KIFEAIKEASKEIFRQKNMQKLRQRET NPEALGV

>sp|A0A164V9U8|CCAP_9CRUS

TRPLFYSLVVLAWLVISLYVSNSLSQPLTANQNDSDSIEDVEQWSYKEKRPFCNAFAGCGKRKSMVSKENKHSAYGKA PSNHPRPLNTNTKLENIVARLQNKRAIAQLDDTEYY

>sp|PSN54967.1|CCAP_BLAG

MERLLDPKRRPFCNAFTGCGKKRSDSMGTLVEMNSEPAVEELSRQILSEAKLWEAIQEAEFFRRRQEQVRILFL KMLVLYKVEQFIAYGLIEEEEEEDDDDDDDDDDDDEDVYLFQQSNAMERPLPLPIAGYRKKRFAAADRTSQVEL NMPMWAKPYRAAVAIG

>sp|E3VW01|CCAP_RHOPR

MQLLVPCFLLFTALVSAVLTDDVFLQKRVYFPGEIAEPIDPKMKKPFPCNAFTGCGKKRSDSMATLVDLNSEPAVEEL SRQILSEAKLWEAIQEARMELLNRKQQQSDRIPLQPLPLTTIRKRSHYLYT

>CCAP_Dalbulus_maidis

MYPPAVVLSCVVAVVLLSHSDSALADDVIIQKRVYDGD LA AIFDPKKKKPFPCNAFTGCGKRKRADESMATLVDLNSEPA VEELSRQILSEAKLWEAIQEARVELLRHQQAERERYAAKPFPLSSLKKRSIDYSQS

>sp|A0A1B6GJ26|CCAP_9HEMI

MYPQVLVATSVVTAII LLSHYTVADDVIIQKRVYDGD LAAMFDPKKKKPFPCNAFTGCGKKRSDSMATLVDLNSEPAV EELSRQILSEAKLWEAIQEARVELLRKHQDQAERERYLAKPFPLSSLKKRSIDYNQS

>sp|A0A1B6I6M5_9HEMI|CCAP_HL

GIMYPQVLVATSVVTAII LLSHYTVADDVIIQKRVYDGD LAAMFDPKKKKPFPCNAFTGCGKKRSDSMATLVDLNSEP AVEELSRQILSEAKLWEAIQEARVELLRHQQDQAERERYLAKPFPLSSLKKRSIDYNQS

>sp|A0A1B6MCB7_9HEMI|CCAP_GRCPAT

QLLVATSVVTAII LLSHQTVADDVIIQKRVYDGD LAAMFDPKKKKPFPCNAFTGCGKKRSDSMATLVDLNSEPAVEEL SRQILSEAKLWEAIQEARVELLRHQQDQAERERYLAKPFPLSSLKKRSIDYNQS

>sp|QBX89036.1|Crz_NENOR

MGMVVVVVMVVVVVFAVTLAAQTFQYSRGWTNGRKRS DPNVGVTELLADPPHLSAHSHPHPTHTLPKNIEERLR ALEAGLNAV LKANSVNFSPGGDEEYYAEN

>sp|KAE8742130.1|Crz_FRAOC

MIVLLAAAMAALLADCAVGQTFQYSRGWKS GKRAVGAAGAALPPSGPPSVSLGGRPADDILAEDLPLYRYFLEGRGE HRVPYDPWRLVEDPVLHLVRRPGLPQSSEQAGEHDER

>sp|Q5W1L4|Crz_DROME

MLRLLLLPLFLFTLSMCMGQTFQYSRGWTNGKRSFN AASPLLANGHLHRASELG L TDLYDLQDWSSDRRLERCLSQLQ RSLIARNCVPGSDFNANRVPDPENSAHPRLSNSNGENVLYSSANIPNRHRQSNELLEELSAAGGASAEPNVFGKH

>sp|AVA26882.1|Crz_CARMA

MVRRVTVLVLVASLALAAQTFQYSRGWTNGRKRSND LGTVVGLGSGRRTGVNLLPAEARRSLQQQTTPHRSLPRDVE ERLRAVEAGVSALLHAAQQNPEAPPATPGEQDYQAQ

>sp|Q5W1L3|Crz_Dvir

MLRLLLLPLFLFTLSMACMGQTFQYSRGWTNGKRAPP AALVTNGHNGLLDIYDIQDRPTDIKLERCLLQLQH FVGNA LLHRSFANGLAYSASRPDPETDVR SINIHSRPGSGNNNIENSLYPNVNHRQSNELFEALNAPGPDAVEPNDYGKH

>sp|U5LKT5|Crz_MEGSC

MIRMFVVPLLFLGICLSMCMGQTFQYSRGWTNGKRAS SDMDVLNPFNVGDSVIENKLERCLMLLQRCSHLSLKGGV NKK

FYKNGSPEMFDELSTPDTK

>sp|BAV78800.1|Crz_PSIV

MWFRLLQSMILLIALLIGSALAQTFQYSRGWTNGKRSFPGSQTACQIQRLRAMLQGKPFPPSYHFFCDLYHLPEEDMKME
QLDKTQLRNTGAEDAIEK

>sp|BAO00944.1|Crz_NILLU

MLLSRLPAALLLVLGCMVCALVAQTFQYSRGWTNGRKRRALLQPPTPLQLQAQAAARARDIDCQLQRLRAMFDGRTDLE
PPCTLSCLLHSCAPQQSNAQQL

>sp|XP_021191518.1|Crz_HELAM

MVTNTTLLLLIFVTIASVTAQTFQYSRGWTNGKRDGHKRTDDIREIAMNIDKFLSPCQINKLKYLLEGKPVTEKLLVPC
DILDDEESPKRYTERTLDALVEAFH

>Crz_Dalbulus_maidis

MVSRHPTSTLLMLCCLTGCVLLAQTFQYSRGWTNGKKRSGLQPSPEHAACQLQKLRIILEGKNTFQLYWPCDWQQPSPD
YEGHHIPPPPLMDVDNN

>sp|XP_069956933.1|Crz_CHEQU

MEKRNSQVVLMMVVLVVALTVSLTAAQTFQYSRGWTNGKRSDDPSVGVRDVTDLLADSTHRLPSHRPLPPTHALPKNL
EERLRALELGLNAALKASATFPPAADDQYYSDN

>Crz_Ixodes_scapularis

MSHYLGLSAVLLVCLAVTAYSQTFQYSRGWTNGKRRVAEMPLVGVPPLRASNDHRALDEVLSKFTPRDRIVLERLGHMV
RVLDHAEEEEQQED

>sp|ALM30312.1|Crz_CHISP

MAPNLSLLLLFVALSSVSAQTFQYSRGWTNGKRDGMKRDAAAAREMAALERIMSPCQMKNLKNLLEGRPLNDRLFGPC
DYFEEEEIQPKRSKTDHSADSLYEAFO

>Crz_Blattella_germanica

MQSHNRCRSHKFFRILLILSCLTGAVLAQTFQYSRGWTNGRKRRAGPLLVPASAASGLLQDADESNPCSQLQRIKFLLLGA
RNPQQIFFPCDTWREVSESASDNTSERFKRRAARDAATLNASDEMNHEN

>Crz_Zootermopsis_nevadensis

MHLNSTASSSRCRSRMTGLLLIFCCLTGSIILAQTFQYSRGWTNGKRSRGPSPQMLIPSSASGERLQNTDESSAISNP
CSQLQRIKFLLLGARNPQQFYFPCETWTDIFETPSEEVSEFRRAKAHQDFAEGNNIEGN

>Crz_Oncopeltus_fasciatus

MIWVRQLLFLLAIAILAGSALSQTFQYSRGWTNGKRRAGQTACQLQRLRAVLEGRAPSADPSYLLCELYRSPQEELRLD
HLGKTQPREPIPVDAIEK

>Crz_Nezara_viridula

MWFRLLQSLLLVALLIGSALAQTFQYSRGWTNGKRSFPGSPTVCQIQRLRAMMQGKPFPPSYHLLCDLYRLPEEDLKMD
QLDKTQLRNPGAEDTIEK

>sp|Q07892|EH_DROME

MNCKPLILCTFVAVAMCLVHFGNALPAISHYTHKRFDMSGGIDFVQVCLNNCVQCKTMLGDYFQGQTCALSCLKFKGK
AIPDCEDIASIPFLNALE

>sp|XP_030031898.2|EH_MANSE

MAGKVTVAFFMFAMIAFLANFGYVECNPAIATGYDPMEICIENCAQCKKMLGAWFEGPLCAESCIKFKGKLIPECEDF
ASIPFLNKL

>sp|NP_001037307.1|EH_BOMMO

MANKLTAVIVVALAVAFMVNLDYANCSPIASSYDAMEICIENCAQCKKMFGPWFEGSLCAESCIKARGKDIPECESF
ASIPFLNKL

>sp|A0NCV3|EH_ANOGA

MVSGKTVFLGVCLLVFVVANNAPVKASPQIELMGGYDIIGICITNCAQCKKMYGAFFEGHLCAEACVQFKGKTIPDCE
DLSSIAPFLNKMN

>sp|XP_016034272.1|EH_DSIM

MNCKPLILCTFVAVAMCLVHFGNALPAISHYTHKRFDMSGIDFIQVCLNNVCQCKTMLGDYFQGQTCALSCLKFKGK
AIPDCEDIASIAFFLNALE

>sp|XP_011192602.1|EH_ZEUCU

MFATTKSTIFVA AVAVVLICALSSTVEGLPSIGHYGKRFDMMSGIDFIQICLNNCAQCKKMFGDYFQGQTCAESCLKF
KGKAIPDCEDIGSIAPFLNAIA

>sp|XP_032295022.1|EH_DROVI

MCNNRTL LLLSIVLTALCLLQCGEALPSIGHYGKRFDPMSGIDFIQICLNNCAQCKKMFGDYFQGQTCAESCLKFKGK
AIPDCEDIGSIAPFLNALE

>EH_Culex quinquefasciatus

MSPMKFAPIAIVLA AFVVILAFNTSEANPQIDLLGGYDLIGVCITNCAQCKKMFGEFFEGHLCAEACIQFKGKMVPDC
EDINSIAFFLNKLN

>sp|XP_002069674.1|EH_DROWI

MYSSKMTTLM TILLSSYCILSLMDAGEALPSIGHYGKRFDPMSGIDFIQICLNNCAQCKKMFGDYFQGQTCAESCLKF
KGKAIPDCEDIGSIAPFLNALE

>sp|XP_026686679.1|EH_DIACI

MLARPSVSNCKKGADRDREELMSRWFTLVNKKNALLRRQMQLNILEKEDDLERKFELNLELRSILSIEEWQKTEDQ
KKRESLLLSELVNIVNKRDELVHHLDSQERAIEDDDEIERDVSRAGMGQRNKNCVIQ

>sp|XP_013114896.1|EH_STOCA

MFSSKSIAATFALVAIGFAFMQFSQAMPSISHYGKRFDSMSGIDYIQICLNNCAQCKKMFGDYFQGQTCAESCLKFKG
KAIPDCEDLPSIAFFLNTLE

>sp|XP_018790275.1|EH_BACLA

MFSINKSTVAFVFI AVVLISVFSSTVEGIPSIHYGKRFDTMSGIDFIQICLNNCAQCKKMFGDYFQGQTCAESCLKF
KGKAIPDCEDIGSIAPFLNALE

>sp|CRL00034.1|EH_CLUMA

MFSWLKKEKESA EVVETALNEMKKIYKTKLLPLEEYYCFHEFHSPKLEDPDFDAKPMILLVGQYSTGKTTFIRYLLER
DFPGIRIGPEPTTDRFIAVMYDDKEGVVPGNALVVDPNKQFRPLSKYGNAFLNRFQCSTVASPVLSAISIVDTPGILS
GEKQVRVDRGYDFTGVLEWFAERVDRIILLFDAHKLDISDEFRRSIEALRGHDDKIRIVLNKADMVDHQQLMRVYGALM
WSLGKVLQTPEVARVYIGSFWDQPLRYDVNRRLFEDEEQDLFRDLQSLPRNAALRKLNLDLIKRRARLAKVHAYII SELR
KDMPSVFGKDSKKKDLIKNIGTIYDRLQREHQISP GDFPDIKKMQEVLQNHDFS KFHSLKPHLIETVDRMLATDIARL
MSMPIQEESEMVSEPMIKGAFEGIEDHISPFGYKRGEIDAGSGEVEWICGRDRARTDPLFESLNPIDGKISGAAAK
SELIKSKLPNSVLSKIWKLSDVDCDGF LDMEEFALAMHLINVKIDGNELPNILPDHLVPPSHKN

>sp|PSN35086.1|EH_BLAG

MVWKIILFKPRTACEVHRRFLSWFSGRKREPETFTNVVDGLKKVYRTKLLPLEETYL FHDHSPKLNDAFDSDKPIIL
FMGQYSTGKTTLIKYLLESDFPGMRIGPEPTTDRFIAVMHDENEGVIPGNALVVDPKKQFRPLSKFGNAFLNRFQCST
VNSPVLKTI SVIDTPGILSGEKQRIDRGYDFTGV LQWFAERVDRIVLLFDAHKLDISDEFRRSIEALHGHDDKIRIVL
NKADMVDHQQLMRVYGALMWSLGKILYTP EVARVYIGSFWDQPLRNDLNRRLFEDEEQDLFKDIQTLP TNAALRKLN
LIKRRARLAKAHAYII STLYSEMP TLF GKDKKME LISNLDSIYQQLERKYLISP GDFPDVRKMQELLVNHDFS KFTSL
NFNLLETVDGMLANDIAQLMSMIPEKASTNSESFVEGGAFDVVG DRES PFYKRGEGVDAGSGDLEWII TKKKA EYD
GIFKTLKPTEMVKSKLPNVVLGKIYKLADVDQDGLLDADEF AVMMHLISVKLEGYDLPTELPVHLIPPAKRHRPEPIS
QLQKHS

>EH_Culicoides sonorensis

MAKSSKLAVIVAVVILSLMLTTTQT EANPGLDLMGGFDFLGVCIRNCAQCKKMFGNYFEGERCAEACVKFKGKIIPDC
EDIASIAFFLNKIEA

>EH_Nilaparvata lugens

MASKLLLAGLLLLVVALAVVEGNNIGVCIRNCAQCKKMFGAYFEGQLCADACVKFKGKLLPDCE DVASIAFFLNKLE

>sp|XP_064212533.1|EH_TRICA

MDSGSRNFLVLLLLFASLLVVDANPIGVCIRNCAQCKKMFGPYFEGQLCADACVKFKGKIIPDCEDITSIAFFLNKF

E

>EH_Aedes_aegypti

MFKVKIPSFISIICIAAIATIAFVGSAAENPQLGGYDMLSVCIINNCAQCKRMFGFEFFEGRLCAEACIQFKGKMVPDCED
INSIAPFLTKLN

>sp|GBP02493.1|EH_EUMVA

MHDDKEGVIPGNALVVDPPKQFRPLSKYGN AFLNR FQCSTVNSPVLQAISIVDTPGILSGEKQRIDRGYDFTGVLEWF
AERVDRIILLFDAHKLDISDEFRRSIEALKAMTTKFV

>EH_Rhodnius_prolixus_Fragment

NIINSKLLCRYTMKLLLVILLTSFLAEISGRQIGVCIRNCAQCKKMFGVYFEGQMCADTCLKYKGKLI PDCEDIAS I
GPFLNKLI

>EH_Apis_mellifera

MTVKENAEVRSGYIDDMNTSIRSILLVFLVYSMV FVLVAGSVPFERAIGVCIRNCAQCKKMFGPYFLGQKCADSCF
KNKGKLI PDCEDEDSIQPFLQALDSEY

>EH_Laodelphax_striatellus

MASNGSVKM LLAGLLLLVVALAVVEGNNIGVCIRNCAQCKKMFGAYFEGQLCADACVKFKGKLL PDCEDVAS IAPFLN
KLE

>EH_Dalbulus_Maidis

MASKSLVCLLLLLPLLAALFLSASASAVDIGTCIRNCAQCKKMFGAYFEGQLCADACVKFKGKLI PDCEDVAS IAPFL
NKFE

>sp|EFX73380.1|ETH_DAPPU

MYRELIMIKGLFLTWLLLASALSDPSPEPFNPNYNRFRQKI PRIGRRGEGIIAEYMNSESFPHEGSLSNFFLKASKAV
PRLGRRKDISTESGRAAMVGEEPFGRISNEIPI MNQKQDLWPNMNINELTGALNKELNYPGPRI PKDLQDNYIQDLIH
SWINQYENLNEN

>ETH_Tribolium_castaneum

MYLAIKNYVLKAAKNVPRIGRSNTNKN TNIDEMGKFFMKASKSVPRIGRRNENFDYGQPIVKRDEVPIWSDIADRFEY
DPEILTSPEILEQLEMGGDDPSVYEW EKIRTKRDSHKPHPKFYVVM

>ETH_Asbolus_verrucosus

MLRYKTLTVVFFVNILL DSTLAEGNYYLQAAKNVPRIGRSSSNKNSNIDQMKGFFMKVSKSVPRIGRRNEDFDYDQA
IDRREQSPTWSDIADRFEYEPEVLTSPEVLEHLEMRDDPSAYDLDKIRLKRDTSHKKHPRFQY

>sp|AD45613.1|ETH_MANDU

MAFRVTKVLTALCLVCLFLQVESSFIKPNNVPRVGRSNEAISPFDQGMGVIKTNKNIPRMGRRNYS ENRFDIPKL
YPWRAENTELYEDDAQPTNGEEINGFYGKQRENMKR

>sp|NP_001165743.1|ETH_BOMMO

MTSKLTMM LFTLSLMFIAGLDGSFIKPNNVPRVGRSNEAFDEDDVMGYVIKSNKNIPRMGRRNYS GNHFDIPKVYSLP
FEFYGDNEKSLNNDDAEEYYAKKM GSMKK

>ETH_Acyrtosiphon_pisum

PLKYRSLNTNAGQDVYSGSCWTATVVESTHHHESTFSRISALNRKSIYRITMSGYLAIIVLVCLQILRVMSINEFPEK
KVQNIWLADLDDKQIASRIERSDQFETASDVLMDASVYPKITRRGFAGEEFFLKASKSVPRIGRRNNDIQETPKRSL
SKDQVNMVEYW PYLQPN DINDLTRKHDFDLPYNCQQLDAKTIFLVDMYNFVCNDQFYCCAPAKRAIANSPNSNSL

>ETH_Sipha_flava

MSGYLTIIIGLVCLHALSAISVNEYTDKKLQNIWLTDMDDKQPSSRFERSDQFETASDIILKDASAYPKITRRGFAGEE
FFLKASKSVPRIGRRNNDIKESPKRSTNKDQMMNVEYWPFYQTNDINDSTKKHDYDVTYGCQQLDAKTIFLVDMYNFV
CNDQLYCCGPAKRTMMNSPSVNYL

>sp|BAO00949.1|ETH_NILLU

MKTMSTVQNRFITEGVFILISATII TASSLHNEESIDVHSKLNRAGAFLPTEPEESSAHTMQRRNDDFFLKATGKSGPY
LPRIGRRNIIYPEKKEDVESPFVVSRRNDDFFLKAHKSI PRIGRRNSYSAPQTPTEDQTPYSQRIKKFDSEVMEKSEAS
AWPWFRAPGLYVPQKRAIYFTENPAYFGPGVILWDKDEHPSLLEDSEANFGTGLD

>ETH_Laodelphax_striatellus

MYTLHNRFTVEGVLLIILSATLLSVFSLHSEESIDMHSKYNRAGGLLATTAEQPSSHTMQRRNDDFFLKVNKGNGNKIPR
MGRRNIVYPEKKEELDSPFVVNRNDDFFLKAHKSIPRIGRRNSNSEPQSTLSDDQATYNQRVSKKFDSEGMEKAEASA
WPWFRAPGLYVPQKRAIYFTENPAYFGPGVILWDAEEHPSLLEDSEANFGSID

>ETH_Oncopeltus_fasciatus

MSTVTGKREAEGFTVLCLTASCQSEPLEDVSEDFVPLLMRRNDDFFLKAAKSIPRIGRRNDDFFTKAVKSVPRIGRRND
FFMKAAKSIPRIGRRDLTTVSILI

>ETH_Nezara_viridula

MSHLIRFILLLVGALLIITESTCGSIMDDSSSEDMIPVMIRRNDDFFLKAAKSVPRIGRRNDDFFFAKNLKSVPRI
FFLKAAKSVPRIGRRDEITSVEKEASVAWPWFRDHDFIPRVMKKDLGDEDVPVDEPSKEKEEPRLQAIV

>ETH_Apis_cerana_cerana

MTSLRFLGFVRCFFVGAIWIAILTLICENVSTADEVPAFFLKIAKNIPRVGRSEGYDDFFKSRKNLPKLGNDGQSE
SWPQYSLDEPFSSRPIKRRVDYPSVDDWSWQHFLAIEGPRELWRTLAGEYKSDTSDEYVCKNY

>sp|NP_001136079.1|ETH_APIME

MTSLRSLGFGVRCFFVGAIWISILALLICENASMADEVPAFFLKIAKNIPRVGRSEGYDDFFKSRKNLPKLGNDGQSE
SWPQYSDVDEPFSSRPIKRRVDYPSVDDWSWQHFLAIEGPRELWRTLAGEYKSDTSDDVDNEIWKRRKRTGNEPMASEEN

>sp|XP_003400029.1|ETH_BOMTE

MTSLRSLGFSRRFIVGALTGVGLALLACENLSKADEVPAFFLKIAKNIPRVGRSEGYNDFLKSRRNIPQVSGYNSRAQ
AESWAPYASDKTFSRPIKRRVDYPSIDDAWAWQHFLAIEGPRELWRTLAGEYKSDTSDDLVLNLFNKKRGQRRMETEEE
NGK

>sp|XP_012248764.1|ETH_BOMIM

MTSLRSLGFSRRFIVGALTGVGLALLACENLSKADEVPAFFLKIAKNIPRVGRSEGYNDFLKSRRNIPQVSGYNSRAQ
AESWAPYASDKTFSRPIKRRVDYPSIDDAWAWQHFLAIEGPRELWRTLAGEYKSDTSDAVDNDVWKRKRTGNEPVIS
EDN

>sp|XP_017788897.1|ETH_9HYME

MTTSRSLGSTRYFFVYVLLIGVLALFVGANASAADEVPAFFLKIAKNIPRVGRSDGYDDFLKSRRNIGKIGGLDADGQ
SESWSYSGNDQPFSSRLIKRRVDYPALDDAWSWQHFLAIEGPRELWRTLAGEYTKDITDDIDNEIWERKRTGNEQALA
VAEEN

>sp|ARK38543.1|ETH_SCHGR

MNLALSAALHNASEGAAGVSPFLPASFFNATSWWSAGAGANGSVASAPDVGDPQVPAYIRTASTVACVIIMSIGVAGN
AMVPAVILRSRDMRNSTNIFLVNLSVADLLVLLVCTPTVLVEVHSRPEVWLLGETMCKAVPFVELTVAHASVLTILAI
SFERYYAICEPLRAGYVCTKARALLICLLAWALAALLTSPVILVAEYGTTEEYFDGSMVPVCFMSMVSFWPAFFFLATI
AVFFLVPLIILVVLYGVIAHLMANPRIMSVHQQQLHYRYRKQVVLMLGTVVAAFFVCLLPFRAFTLWIIAAPPETVVS
IGYENHYGLLCLCRMHYLNSAVNPILYNLMSSKFRDGFSSRLCCRCGCAGGPLGRRGTGLSSSATATTVRSSAGDSLW
RRPGRASSSSSRRAERRLLHHIARECSVSVHNGVCAVKAGRSKDVSYL

>ETH_Blattella_germanica

MGKYCLSLGLCARIYAVLVVVSTLAVGATGEDGAGTNFFLKASKSVPRIGRRSEYDNFFLKASKSVPRIGRRRELSP
LTEGRDWGNVPWFRTSDNIPGPSRRADYIIEGGPPAHPLSWNDVEKTMEESPELWKPDLWRKNNENFPLRDEFDIQH
IVRRSTGPFNKDNTKRETNDQNLTETI

>sp|XP_308702.4|ETH_ANOGA

MKLVVFFALICYAIGSESPGFFIKLSKSVPRIGRRGDLENFFLKQSKSVPRIGRRAGGFHMAAVAPQEAHGSSWFERY
MKTAKRPNGGSLVPDDFGKSYKKIRPLDLNHIIDILSDDLFFGSDLKFISWEVLDEALEEDPELLQKLASLARDKEV
QQLKLLGEQDDELVKYVPIDRNEINGASSRAYMYRLDRTPNKERLEYQQ

>sp|XP_021698718.1|ETH_AEDAE

MMKFILLLVALTVAVADETSGFFIKLSKSVPRIGRRGDFENFFLKQSKSVPRIGRRSGLFTEHSPAQDNSNWFERFM
RPAKRPISTPQTENGKLYSKVRPLDLHHIIDILSDDLFFGSDLKFISWDVLDEALEEDPELMGKLSALARDKEVHQLK
QLLTENKDELQYVPLESGTSDMHSNKAYIYRTDRDEQNAELNKEKIEYQQ

>ETH_Musca_domestica

MNYFFLISYTLIFFLGLPEQSLCKADDS PGFFVKIAKNVPRLGRRSDKPFTRRFFRNYKII PRIGRSVEELPKSDSLN
NWLDTPTFTVHMEKRLHYSPASRFVDQDLSLLQPVNAATLKELIDSHTIARENVKFIHWKDFDRALQMDPELFGKISSI
GRTPDQHLKDNLSKENFTPV LKNMMSDGD TDFVLYTQ GENDQGVYAPEFLRYNIM

>sp|XP_002004620.1|ETH_DROMO

MNRLTILFAVLLLAELWIAGYADESPGFFLKITKNVPRLGKRGEAF LMKNMKTIPRIGRSDPEASITPLLAWLWNLDA
DPQFSKRRLPSSGAAGSVEHEL SVVQPVNSNTLFELLDRNAIPSEHVKFVHWKDFDRALQVDTELYTKI IHLGRDPDQ
RLKQDLSFN SYIPIFGSDDAKSND FMLYNKDDRDLYGGYDRNFMQYNRL

>sp|NP_524699.1|ETH_DROME

MRIITVLSVSLVLGLVAISQADDSSPGFFLKITKNVPRLGKRGENFAIKNLKTIPRIGRSEHSSVTPLLAWLWDLDT
PSKRRLPAGESPAKEQELNVVQPVNSNTLLELLDNNAIPSEQVKFVHWKDFDRALQADADLYSKVIQLGRRPDQHLKQ
TLSFGSFVPIFGDEQNPD FMMYKNNEDQELYGGGNRYDRQFLKYNIL

>ETH_Dalbulus_maidis

MACYTLLLECVVVCSLLTISESSPQSQFDDQFDPFDDSFYEPGPPYLGDIVKNVPRVGRNRNDFFLKAAKSVPRIGRRD
GDDVKPSQINKNFPRVERRNDFFLKAAKSVPRIGRRNDFFLKAAKSVPRIGRRQDRLDKRGALTNAMEGGEMAAPWF
RNPELMPQAFKRSAPFQGP DQEE SDHGLVFYKINLHREVEEPRLQAEV

>Pob_Acyrtosiphon_pisum

MYTTQLFIISIAIIIVHNNVLAEDNPEECETLPSELHIIKEEFDELGRLQRTCNGDIAVNKCEGACNSQVQPSVITPS
GFLKECYCCRETFLRERMIALTHCYDPDGVRLTSDKLATLEVKLKEPADCKCFKCGDFSG

>Pob_Drosophila_melanogaster

MHVQELLFVAAILVPQCLRALRYSQGTGDENCETLKSEIHLIKEEFDELGRMQRTCNA DVI VNKCEGLCNSQVQPSVI
TPTGFLKECYCCRESFLKEKVITLTHCYDPDGT RLTSPEMGSM DIRLREPTECKCFKCGDFTR

>Pob_Apis_mellifera

MKENFSIMFIHSIFLILIIIFIYSNETIAQVTDDENCETLQSEVHITKDEYDEIGRLKRTCSGDISVTKCEGFCNSQVQ
PSVASTTGFSKECYCCRESYLKERHITLHHCYDADGIKLMNEENGVM EIKIREPVECKCIKCGDISQ

>Pob_Anopheles_gambiae

MCNSVRTALAASNCCSIVLCCVLLLTLTVAVTAQHNQADETCETLPSEIHLIKEEYDELGRLYRTCNGDVT VNKCE
GKCNSQVQPSVITATGFLKECYCCRESFLRERQLQLTHCYDPDGV RMTDHE SATMEIRLKEPVDCKCFKCGEMVR

>Pob_Bombyx_mori

MNIMITKIFFLVQLFYIVVSKSSAEENCETVASEVHVTKEEYDEMGRLLRSCSGEVSVNKCEGMCNSQVHPSISSPTG
FQKECFCCREKFLRERLVTLTHCYDPDGIRFEDEENALMEVRLREPDECECYKCGDFSR

>Pob_Atta_colombica

MLVASKSLVLVYVTISDAEITEIDDESCETLQSEIRITKDEYDELARLRRTC SGDVSVTKCEGFCNSQECYCCRES
YLKERFVILNHCYNADGIKLIGTDDEVMEIKIREPAECKCIKCGDFAR

>Pob_Tribolium_castaneum

MFDKII LCLVYATCVYSVSEISEETCETLMSDINLIKEEFDELGRLQRICNGEVA VNKCEGSCKSQVQPSVITPTGFL
KECYCCRESFLRERTITLTHCYDPDGVRLTAETVNSMDVKLREPAECKCYKCGDFSR

>Pob_Bactrocera_dorsalis_Fragment

VPGNSHNAYPPSGRVVRRKCGEVPNMSTRYCSSSIYGLLSWLLVLLFVLSRLERRVAALRYAQGTGDENCETLKSEI
HLIKEEFDELGRMQRTCNA DVI VNKCEGLCNSQVQPSVITPTGFLKECYCCRESFLKEKIITLTHCYDPDGT RLSTPE
MATMDIRLREP TDCKCFKCGDFTR

>Pob_Araneus_ventricosus

MTSNFANCLLITWYTFLLINRVCSREQNEVEVAEVVAAMKDACETLPSTIQVTKEEYDDLGNVVRSC EGTIGVTKCEG
TCTS QIQPSVVTPTGFLKDCHCCRETWMRQREILLTDCYTQDGKRIFGQQGTMVIHLKEPEACACHKCGL

>Pob_Dalbulus_maidis

MKAGVRGVSVLLLGAMLSVTRSEGE EACETLPSEIHLIKEEFDELGRLQRTCNGDVA VNKCEGACNSQVQPSVITPTG
FLKECYCCRESFLRERMITLTHCYDPDGMRLTSQAQASMEIKLREPADCKCYKCGDYSR

>Pob_Locusta_migratoria

RPQPAGGGEEACETLPSEIHIIKEEFDELGRLQRTCTSEVGVNKEGACNSQVQPSVTTPTGFLKECYCCRESFLRER
TVTLSHCYDPDGARLTAEGLATMDIRLREPADCKCFKCGDFSR

>Pob_Lygus_hesperus

MEIVPLAVVLI FVG VVKSHNNDEACETLPSEIHIIKEEFDELGRLSRTCNGDI AVNKEGACTS QVQPSVITPTGFLK
ECYCCRESFLRERIVTLTHCYDPDGVRLEKEGVATMDIKLKEPSDCKCFKCGDYSR

>Pob_Acyrtosiphon_pisum

MYTTQLFI IISIAIIIVHNNVLAEDNPEECETLPSELHIIKEEFDELGRLQRTCNGDI AVNKEGACNSQVQPSVITPS
GFLKECYCCRETFLRERMIALTHCYDPDGVRLTSDKLATLEVKLKEPADCKCFKCGDFSG

Ovogenésis

La base de datos referente a la ovogenésis se compone de 2.583 secuencias correspondientes a la especie *Drosophila melanogaster* (Pascual & Rivera-Pomar, 2022). Todas las secuencias fueron descargadas de Flybase.

Olfacción

>ADK73604.1_odorant_binding_protein_Nasonia_vitripennis

MKIVVLCLVVL SAVACVSAGYREYQNA CLDENGLTKEEFYAMKRNQDPRSGCVTACIMKKNGSMKHGIIDARGIKRRM
RTLLAPFISKDKLYEKIDYCVDEAENHVGVCEKAYVLQKCLRTPRANNVQGERQK MID

>APX56329.1_odorant-binding_protein_30_Halyomorpha_halys

MKPLHLVCLIGAF AFASGV DIPDELKEMA QMVHDQCIGETGASADAIEGTTKGVFPADDQKLKCYLKCIYSNMGAISD
EGELDAEAFGSIMPEELGRVLNPMINKCKDTTGPDGCELA FNFNICLYNADPKNYLVI

>AOV87046.1_odorant-binding_protein_29_Halyomorpha_halys

MVLIMS NYILALV LAVVSVDG SVGGLOGELSEKELLGLADMYHHSKAI SKRSLRNTGDMDDHRYH KHRHMFDPYKHCCG
ETDRNNIRRK EAWTVVKDCVEQVKTSNELEEFVNPI TDEFSCERLKQEKNRHYCI ADCVLKNT EALKDGTVNRD KA
KEFFT KGISAEWLKEIAENAVEKCAEQKVFTSNGEQLDCSPQAVNLKHC FWKEMILTCPK EHFKKSPYCSRVKEAFEE
MEN

>AOV87045.1_odorant-binding_protein_28_Halyomorpha_halys

MHLYLAGICSL LISLEIVSVANGAATPDECNDTG GENTGMCKCHKREPESKEKQE EYENNMKNCF SKYFGDIQDPPPE
DVKISLMECVGECFFNASGLLTPDLKLSQEKLNAKDDIVDPDPKWK AIDEAAVKSCFDKSVAEGNGSAKCKSGSFQFL
RCVLRERFLNCPKDVWTDGDECNNYRKVVEKCPNVI PGVA

>AOV87044.1_odorant-binding_protein_27_Halyomorpha_halys

MNAVLCITTLAAISVSGAMAPEYKKKLLTSLSDCMEKHNV DKEEIMQVIKNEKLPEDKEKQCLFGCLIDNMGYVKDL
TLDWNLIKASNPEKFRDPEDVEKANTVTDTC AKTLSEKQPSLCELGVNAIKCIHQESLKIKLPKPEFSFD

>AOV87043.1_odorant-binding_protein_26_Halyomorpha_halys

MKLFESILLTAIVIAVSVSLIGCHEDIVSIKKMAISRCSEIHSISKEAVKSIMNNEEVPNDYEIKCWLSCVLKTLGML
KDGKIMWETCRNITKYGFSEEDKAKVDKITEICQAEVPQEEKDECQLAYSAAVCKMNNWKKLGLPKGNWEE

>AOV87042.1_odorant-binding_protein_25_Halyomorpha_halys

MKGITTALALT LVLATIEAEENVKAKIAKVFN TCKEKNPVTEDEIGA FRKADIGFSYSHDAKCM LACMLEEGKMLRDG
KYMKENALIMADV FHIDDFDEAAKARKVIENCAIEVPEVGPDPCEFAFKMAVCGANESKKLGMDHDFFE

>AOV87041.1_odorant-binding_protein_24_Halyomorpha_halys

MRLSVAVIVSFAIITSSLVLIQCQEEDIMEKKTEAITQCTAEHNVSK EAVKSI SINDNVPNDPEVKCWLRCIFEKLG V
MKEGQIDWERCKFITKHCLSNEQDKAKVDQITEICAEVPKDEKDECQLAYSAMVCKMKNNWKNFGLQ

>AOV87040.1_odorant-binding_protein_23_Halyomorpha_halys

MNRVVIVVALILACIFNCHGQANPEIKA AVNECRSQHNIEAGQIKDAINNKKI PETEHGQCFM SCVMKKMGVLKNGKI
DLDRVSELINNKFQDQENREKAYEIAKRCANVKS PDGKECSQASEMAKCALKN AIELKMEVPKGSF

>AOV87039.1_odorant-binding_protein_22_Halyomorpha_halys

MFIVLLFLSSAYHSISFINAQPDYSRKASSFKTNFM ESSKYCSNIYNTSVF DVLPFGQSDNGNLEAKCFL LCLLQR
YQ MIDKGGAFHRDKVDRFLDSMPDSKFKSSLKS FKEKCLKEDSQNQ CQKAYKFSACFYNGIDKKHES

>AOV87038.1_odorant-binding_protein_21_Halyomorpha_halys

MKYLVLCCVLSYAYAASVDTSKAILQKQFDVFRKCMVVGNTVESEVEVIFKKGEVPEGRKIKCLLGCFMKGMYLTDD
GKIDWSIMDDINRKEYVNPDEVKRALEISALCSKSVFQNLGNLCDAGYAGTKCFLDEAVKRKLT LHGPKDAE

>AOV87037.1_odorant-binding_protein_20_Halyomorpha_halys
MLSKGTVLVFAAINLMMVTAYREELKRTISGCQNGAEVTDDELEEFKPLIPKNDEEKCLMACVFKSFSGVIVDGTDFP
KLALAVAKDLLKSDPEKVKKITAVVEHCGDDIPKKMDNECELAGEIMKCYVKWEKEVGLA

>AOV87036.1_odorant-binding_protein_19_Halyomorpha_halys
MQPISSTIVFVVCIALAQGAAVQDQVTAIAKKCASEHKMSEEQAKIIFSLVVPKTPPEENCYMECIYIGAGVMKDGKYS
VEGAKKLADQRFTGADEKTA AVKLIETCKSVSAGSEKCGLAKSVRECFVKNGNNISFFPPPK

>AOV87035.1_odorant-binding_protein_18_Halyomorpha_halys
MNRINLLLLLVVLVASAVGWFPQPPPPPPGGEDVPEECRPKPPPNPGKERAVCCDMPHPIAKDSDQFKTLFEECKEQIK
SEHPEAMPHPHHPHPPPPPPPAAGGPPPPPPHGGPHLFMTCMDECVFNKSGLLTDSKLNQEALTKMVDSFAPDGS
WKTIALEATKFCYEKTSNETKTEDKDAKCSVSHDFVKCMIKQMYLKCP SDKWQNNDDCNAEKQLIEKCPNLMFPFPH
PPMKKKE

>AOV87034.1_odorant-binding_protein_17_Halyomorpha_halys
MNAALCLTTLAAVCLSGAATPEYKAKVITSVTACAKEYNVELKDIEIMKQSKLPETKEQKCVVGCFFEKMDYVTDH
KVDWEKVKALNPQKYDTPELVEKINQVTDTC SKVVTEKSTDICELGVP AVKCLKEESEK VQLPKPEVKFDSA

>AOV87033.1_odorant-binding_protein_16_Halyomorpha_halys
MKHILGVGLMTVFLLLAGKGRSSPSLGHTRNAMGGYNMNYDSEEEAASSYLKGD SIFKYSPQEETIAKILKSFSGCQA
KHPLSADELTA LKNQDTEMHFSHDAKCFACMFKAADLLKNGRIIPENLIEYFNEDSIGVSN SREIIEYCNQEVSESE
PDLCEYVYKMTKCTWKILTS

>AOV87032.1_odorant-binding_protein_15_Halyomorpha_halys
MNIYIVNFIIILTSISFTFASTVKPGIISQAFQIVGRCNKEHPLDMKEMEKAVKNFELPSSEEGKCFISCFLEGFGLVT
DGQINLDRSLEFNKMQFHDPDNLEKANAI SAACKDEMSTSTEKGCDFAIAASKCMLEKSKEMGFRFFQFHS

>AOV87031.1_odorant-binding_protein_14_Halyomorpha_halys
MKSFFVVIALSALIALAKGDDVKEKFKKTA EKCKEKHAVTEEEIEKLKAKDTEFQYSHVAKCYMACFLEEGKILQNGK
YNKENALVMTDVIHKENPEQA AKGKEI IETCAKQYPEVGEDQCEFAYQVS VCAAKEAKKVGLENTDFFAK

>AOV87030.1_odorant-binding_protein_13_Halyomorpha_halys
MQNHRVLFLLYALLAILLNESNGVEEQAECKDKSNLQVLGRCCNIPNEELLSEEEKAVGTMCDKTVYGDRLTSP TTEE
EQLELYECMDEC FYNGTHLLTADMKLNETALIAEYDVSTENFSHWKEPITKAVKSCLEKNVVSEVKPNAKCKSGSYQF
SQCFGLKLYLYCPPEDWKVKDEGCNKAREVLMNC

>AOV87029.1_odorant-binding_protein_12_Halyomorpha_halys
MNSLLFLTILSAVCLSWAASDEYKMKVIASVAACAKEHGANIADVIEIVKAEKLPTNKAQKCVAGCFFHKMGYVTDS
KVDWAKVKALNPKKYESKDLIAKADKVVDACSKVVTQKNTDVCELGI PPMKCLLEESKKVQLPKPDIKFKL

>ACI30694.1_odorant-binding_protein_Acyrtosiphon_pisum
MISSTFYLTLLFGIAMLISCGYGRFTTEQIDYYGKACNASEDDL VVVVKS YKVPSSSETGKCLMKCMITKLGLLND DGSY
NKTGMEAGLKKYWSEWSTEKIESINNKCYEEALLVSKEVIATCNYSYTVMACLNKQLDLDKST

>NP_001153532.1_odorant-binding_protein_6_Acyrtosiphon_pisum
MPNILPNLDSTWEKCFETFKQFKDKPETKEYKEMAHGKEPPCLFQCIFMQSGLTTS DGKLNEDAITKKMSEGINNDEK
WKS IWQNSLNKCFDDVKQEDKKQILIMNTPAGRLMKCFLRDMYMSCPKNVWVESSECLNMKDLVQKCPMPPPVFKSP
PKLI

>CAX63068.1_odorant-binding_protein_11_Acyrtosiphon_pisum
MSSSTFYITLLFGIAMLISCGYGIFTEQIDYYGKACNASEDDLIVLKS YKVPSTETGKCLMKCMITKLGLLND DGSY
NKTGMEAGLKKYWSEWATEK IETINEKCYEEGNTATLLYHVAIYFTCVSGDYS DVQLLIHCDGMFEQEVGSRQVNLKL
LIMLKIGLSEPKR

>NP_001153529.1_odorant-binding_protein_3_precursor_Acyrtosiphon_pisum
MISSTFYITLVFGIAMLISCGHGRFTTEQIDYYGKACNASEDDL VVVVKS YKVPPTTETGKCLMKCMITKLGLLND DGSY
NKTGMEAGLKKYWSEWSTEKIESINNKCYEEALLVSKEVVATCNYSYTVMACLNKQLDLDKST

>NP_001153534.1_odorant-binding_protein_8_precursor_Acyrtosiphon_pisum
MFALKVACLCLSVAVVFGENNQQNGPSDRSATIFQSCIAETKLSGDALKGFRSMSIPK TQAEKMMGCLMRKVN VINK
GKFSVEEATKVAQKYGTNEAMMKKAKDLIDVCAKKAQSTTEECALAGIVTTCIVEEAQKAGLSGGPGSR SRRTVSPK
FRDAM

>NP_001153527.1_odorant-binding_protein_1_precursor_Acyrtosiphon_pisum
MLNLKVMFLCLSVIVVYCESDQVPINSSAAVESCLLETNMTRDEFEDMLTSPNARELTILKSHAHKCMFGCVMRK NH
IVNDGVVSKEVLSKYVLNFYGRPDYKRRLI IKDVEHIVDVCAKKVADESETDECELAATLVTCIVLEANKAGLVDDPA

RQI

>NP_001353810.1_odorant-binding_protein_precursor_Acyrtosiphon_pisum

MSSSTFYLTLLFGIAILISCGYGTFTTEQIDYYGRACKASEDDLTVVVKSYKVPSTETGKCLMKCMITKLGLLNDGSGY
NKTGMEAGLKYYSEWSTEKIEAINNKCYEEALLVPKEIVATCDYSYTVMACLNKQMDLDKLT

>NP_001153531.1_odorant-binding_protein_5_precursor_Acyrtosiphon_pisum

MSANSATIKCIAVAAILLQISVIFADAGHHRRGKELLDTEDSDFFRCKQASRKSCCGPENAMKRFGDKDKVAADCECYA
QVAEKFATVTATTPKQDLFSAEAVKITKKKQFCLHECIGKKNLLTEDGSLNKTFIADYAMKSVFKEQWQKQVGVQKAL
DKCLEETYIPWPAEDKENVCNPVYVQFQHCLWLQYESNCPANKIKITKKCEKTRNRYRMQKSTSN

>NP_001153530.1_odorant-binding_protein_4_precursor_Acyrtosiphon_pisum

MRGNYSSMVFLFLFAIGFQDIFCQKQEPGKCRAPDKAPLNLEIIINTCQEEIKSALLQEALDILNDGNVEQNTPNYSS
RSKREAEEDLTNEERRVAGCLLQCVYKKVKAVIDETGFPVVDGLMKLYNEGVDNRNYIATLSAVRHCISIAQQLKQQQ
PSKSFDDGQTCDLAYEMFECVSEKIEENCGVENKSN

>CAX63265.1_odorant-binding_protein_RproOBP6_precursor_Rhodnius_prolixus

MRAVTFCLAAATLLVIAQSESTTTEPETTQDNDKSTTAGASTVSKTPEEIRQKIKEQVEALTEACKTQTKLTPDQTK
IATNMAVPKTEPEKCFLECIYNGIGLTKDGAFNEQGARVLAQQRFTGAPDDLAKANTMIEACTKEVVVKDANEKCGLG
RLVRECFVKNQVKNFFPKP

>CAX63264.1_odorant-binding_protein_RproOBP5_precursor_Rhodnius_prolixus

MTKTKLAELEAKMSECLKNSGVNITTLGSLNEESNNTPKFEKCVFGCFIEMGYGKDTKPLWDIMEEVHKIEYGVK
EDKEKALKIVETCKIIVPEEVEDSCELGFRHSCYTAQARKLGLTLE

>CAX63263.1_odorant-binding_protein_RproOBP4_precursor_Rhodnius_prolixus

MFLYAILLVLTCTVLCTYAALDLKGMTKTKLAKLEAKMSECLKSGVNITTLGSLNEESNNTTKEFRVLCGFTEEM
GYGKDNKPLWDIMEKVHKIEYGAKEYKEKALKIVETCKTIVPKETEDSCELGFMHSCNTAQTKKLGLTLE

>CAX63262.1_odorant-binding_protein_RproOBP2_precursor_partial_Rhodnius_prolixus

FRFRAAARARFWVLDPEKQTNVQQNVTELSKKELSDLEDEVMEKCREKHNVGEEIFSKLEAISQHKELPSSKDFKCM
LGCFAEEMGYAKGAQPQWETMVAHVHKIEYEKEEDANKAIQVVETCKKSVPKEEDI CELGYSMATCYLLESEKLGLDL
V

>AAQ20833.1_odorant-binding_protein_precursor_Rhodnius_prolixus

MKTTLVLAVFLVIAVYVNGAATQSPVKEKVAALAKKCSADHKATPEQTKLVFSQKVPSEVERCLLECXYTGVGVIQG
GEFSAEGSKKLATLRFSDPKERETVNKLVDCTCKEVTKNKDEKCSLGKSVRECFVKHGEKVHFFPSAN

>QCZ25113.1_odorant_binding_protein_57_Nezara_viridula

MKQSIMEMLKLLAFIFVIPACFAEEKTDAKLISKIKNVGECAADSNIIDFLCSEMFEADAKHDDPKYNPCKCMISCV
SKKMNVMSEDGKFKEDESLKIINDLKTPGYKEEALRVYEICKNPEGTNCEAGFNKVKCVMTNSEKARNLTMLQVKTMTQ
GKA

>QCZ25110.1_odorant_binding_protein_53_Nezara_viridula

MQDYRVFVLLYVVFALFLNESYGAEDQVECDNPNSFRVLERCCNNIPDEELLTEEDKAAGLMCDQTVFGNRTAPLTNE
ERLELYECMDECIFYKATQLLTADMKLNETALIAEYDTPLENFPQWKEPMAKAVKTCIEKNVVLVDVKPEAKCKSGAYQF
SNCFGVNLYLNCPPENRKMDDECTKAREILEKC

>QCZ25105.1_odorant_binding_protein_48_Nezara_viridula

MQLILFASALTVAADAQAKQICVAPTTAPHKIEKVLSQCQDEIKYALLQEALSVLGQSVGRQKRETFSGEEKRIA
GCLLQCVYRKMAVDENGFTAPALVQLFTEGVKDRNYLATIQGVQQCLAKEIQQRKYNQTLAQAEGYTCDVAYDMF
MCVSEETLTCGISP

>QCZ25103.1_odorant_binding_protein_46_Nezara_viridula

MAKMFWNIFVAVSLLGYLEARITLPPEIEAFIDKLHAHCIENTSDLTEEDHAAYEVGHKNEALMCYMKCLMLESKWMSS
DGEIQYDSIIQNAHPSYADLLTAAINKCRSIENTGADLCEKSYNFNYCMHQADPENWFLV

>QCZ25102.1_odorant_binding_protein_45_Nezara_viridula

MCGKLVSLLALLAFTVCTFGFELTEEGKKHLQECQGHKHSISPDEVVAAVKDKLPESENGKCFMECVMEKFGVLNDG
KFDPEKAKSIASGKIGNKPELLEKVNTVIDKNAEVSNPCKEYGVGLAKCAMKYAKEMNISEIPIGN

>QCZ25100.1_odorant_binding_protein_43_Nezara_viridula

MQMKMIVLLALFTSASFILAEDIPVENLKDVVMKELHNVKVQVAMECAAVYPVENSTIRQIMTHDGIPDNQEVKWL
CVLRNMGIINNGKVDWKKAIELNKMVLKSDKVIDQIFEICKKEVEQDEDECQLAYNIAICKFKYKLELPLNGTW
EE

>QCZ25099.1_odorant_binding_protein_42_Nezara_viridula

MQCSYFILLAVVCTAHPILSMTLKEINDACMKEHNFDDGLGKMNYNDPAFPKEAKCTLACTFEKEGVLKPDGSIDKQK
DKEALEEIVKDEELKAKLLVAIEECGVTAADKDCETAYEFVKCKKAKTGQH

>QCZ25098.1_odorant_binding_protein_41_Nezara_viridula
MKKISAVIEECAKEHNTQKADIITMLTTQTLPQSQDLKCTVTCYLEKSEYLNKGQLDWMKAKENNKFKYADEESIKKA
DGVADICSKQVNISGLEKCDIGNAVLKCLYEESDKVQLAHPDFLAKSS

>QCZ25096.1_odorant_binding_protein_39_Nezara_viridula
MPTMKKTHLIVLLVVSISKLAKEIDKNCIEITNVIESEVNNYRHFRLPQTKEGKCYGACVMKSVSIINSRGKISSRK
VEQIAKELRNSQEAMRFADIKRCSHSANQLNDECETAYNFMKCLSAFKLEKLAKTGSFISISPPTVSVNLPPVTFVL
FQLSG

>QCZ25095.1_odorant_binding_protein_38_Nezara_viridula
MWKNSAISLLVLLSATPLVLSLSEEMQELANQLHSACVEETGAQESDIPKARDGVFSEDENFKCYIKCLLEQMAVIDD
DGMIDVEAMISVLPEELEEEIAAPVMRKCGTVKGANACENAWLTHKCYKENPEAYFLI

>AMQ76488.1_odorant-binding_protein_35_Apolygus_lucorum
MINVVVFLIGTGIVSGGFMEALIECKQQHHVSKEEAMTGESEEVKCFSECVLKKSGMMSDNNEFDEEKIQAEGARMI
KNDEQKNREFEGAACACIEKVNGENPSEKCAKGHALFKCMKEAMPMSKMRG

>AMQ76479.1_odorant-binding_protein_26_Apolygus_lucorum
MNPTVAIIFTLLVAVVKANTKELSPSEALKQKVQVQCQQEVKATPEQLKIYDNFKDVPKDDVENCLMECMYTKTGIG
ADGKYSVEGFKKLVDMKYKGEENTKARKIAADCEAKAAPKEGKCSMGRAIRECLAAATKENEFFTII

>AEP95759.1_odorant_binding_protein_5_Apolygus_lucorum
MKSFVGLIFAVALVEFASAITKEYHRAVAADACLKHKPSIKESDVQEFLLKHKLPETDDGKCMIACYMEEMNLMD
GKINVEEAKKTNDSKYDGEPDNKLADKLIDHCSSQVSPDGMSKCEYAYQFSKCGLEYGMKNGLTPPKMYEEQRR

>AMQ76491.1_odorant-binding_protein_38_Apolygus_lucorum
MGFKFVKYRSYFFVLVIHIIILCIQIKAKELTDEQKEQIFAEIKNCMESTKLTDEEFESIMAKKELPTSKEGKCFTKCL
MEKMEYLEEGGKINVIQAGLEENMEKESEITKAKEIIQQCADTVPPEDSCEYAYGISQCMYTKMKEAGISGGP

>AMQ76490.1_odorant-binding_protein_37_Apolygus_lucorum
MDTHFGLLIASLAILHTANAVINKDYLEKVVTAKDKCLKEFNVDDSVVEDFIVRYNKPQSESGKCMVACYMEERGMMK
DGKTIITEQVMLDNQEKWIAATHVNMGKEVIDTCDKEVPNEKNDKCDLAVDYMCLVKRGDEAGLPKMDVAQLKH

>AMQ76489.1_odorant-binding_protein_36_Apolygus_lucorum
MKTFFVGLIFAVALVEFASAVSKEYHDKAIAAKNTCAKLHNVDDETIMKFWKAHQLPEKEPETCIIICYMKEMKLVVDG
KVDADAWKASNKEKWDDDEKHVAAADEIVDKCSAEVPPTENECEWGLALTKALKHGKEAGIPPPDMEHPKRR

>AMQ76487.1_odorant-binding_protein_34_Apolygus_lucorum
MEHWKWLALLIFGMVTCVPQLEGAQKSKQPSKAKTKESQVVAARPKDARAAACVTQIGPDEEEESFYRKEIPETDK
GKCLLACYLESKGVLSGGKFSSSGAAKIAARAYPNNAAKTGNVKHILSHCGTIAARETEQCQLAYRLAECTTTLADKF
KL

>AMQ76486.1_odorant-binding_protein_33_Apolygus_lucorum
MHPWKTTCCLIGMTAALMVVTAFAGLPFQNMAMVQCKVKFDVTAEDIQLLKDSKLPSSHSGKCMMACILKKMKVMTKR
GQFDLRNVQKWLNRKYQGQANLAKGNYVAEACANTLPTLGIQDECEMAAEIMTCVVRTKSKLVKKTNLGELPKEVSP

>AMQ76485.1_odorant-binding_protein_32_Apolygus_lucorum
MSGRHSLILVLLAAVTSAEVLTGDGCPKTPKEMKPLYKCCVEMDSNKTISDDQKAAVDSCVNTSKSDSDANKHDCM
IECIFIKLGYMGEDKTINVYVLKEMNSLLPEDFHEQTSKSLATCMGKKFSSTECPSEIDGVMACFSTMVLMNCPAKH
WTDDEECKATRKFFQKCGDSIGYRYD

>AMQ76484.1_odorant-binding_protein_31_Apolygus_lucorum
MFTSATFTVFLFAVTLTRGQIDEDPECRPSGPPGKEPECCTIPMKLFGDEVQEAUVKNCNFDEAGMKRPSGPHGGGSPP
TAEEMAHAHISAHECADECVFKSGNFIKSDGGLDEDAIKAVIAKLFTGDWAPIATAAVNKCLASAKSGVSASAKCKSGA
YQLSKCFQRELFLGCPASLWTESTDCSAIKARITKCPNAKVPIGHHHKKH

>AMQ76483.1_odorant-binding_protein_30_Apolygus_lucorum
MNAHIVLCLVASVFALSQGTPTTPTPATSRRTVAPEDLEQAKSLRKFC TAKTGFTGITTETSKGKDQARTATTTPRP
KTQLEKCYLECLYTGLQLTKDGKFNEPGARALANKRYKNAPEELRKVNSIIDFCITEVVVRDIEEMCALGRLLIKECFS
KYGAKNFPEL

>AMQ76482.1_odorant-binding_protein_29_Apolygus_lucorum
MNRPLLLLTAVALVSGSQQEDCKTAPAGWPRRPPQCCDLFPFLEGMKKEFGSCIRQIGNRQSSAVPTAQAVRDARLCI
EECVYKGLGFMEHNLNKDQILQQLTKGVADKKDWTKPMEDAVKSCHETITKRETPQEGTCKDSAHEFTHCVMRQLFL
SCPASEWNNNDECNLVKSQMCPNIPPPPPPPQGFGRGQPPPPQ

>AMQ76481.1_odorant-binding_protein_28_Apolygus_lucorum
MIIIEICVLTVGISPHFIEGQELPPPGVGNGKTAVFKEFIRTAKYCSSIHETSTVAVLAAILMSEESDDQNGKCFLNC
MLQRYQLMSKQGAYNKDKFKPFLDYIPESRFLQSIKGNLKTICITERDPAPCEKAYKFIKCFYTRARNKDEFGKIQRK

>AMQ76480.1_odorant-binding_protein_27_Apolygus_lucorum
MARKFIKSCYTLVALLVFVGSIHVEAKELTEEQRTQLFEDLKQCKNSTDLSDDEFETIIAKKELPTSEAGKCFTKCLM
EKLDIIEDAEGGKKKISVITMQASLEENMEKEDDIAKGKDIIQKCGDTVEPEDSCAYAYNISKCIYDRMKEAGISQ

>AMQ76478.1_odorant-binding_protein_25_Apolygus_lucorum
MFTSTIFAVFLFSVALTQGGQMDDDPECRPPPPPNKEGSCCTVPRLLDNADKPEVIKKCHDEAGMKRPSGPPGSGTPPT
AEEMAAHKS AHECADECIFKSSNLLKSDGELDQDAIKATTTKMFTGDWSTIASTAVEKCLATAKSEVGASAKCKSGAH
QMVKCFARTMFLNCPASSWTESTECAAAKTRLTKCPNAMPPPPHHSRH

>AMQ76477.1_odorant-binding_protein_24_Apolygus_lucorum
MSTKLRSVGMILAIATHVCAYQEQLKETIKQCQDGREVTDDEVEEFTKPLVPKNQEERCLVACVFKEYKVIIDGHFD
PVNALNVAKMVKDYPEKWKRIKDVIDHCGEDIPTHNDNECDLAGDIMNCEVKYLN SMPKGVSLELLAGSIAATAEP

>AMQ76476.1_odorant-binding_protein_23_Apolygus_lucorum
MYVFTVALSFALLNIVFTHPGHFDEDEPCRPPHSHRHEEKECKTPNLFSKNKDEMHELVHKCFEEAGIKKPHHGHG
PPPPGDEPPPPPPPPHFSKNNTKFECVEQCFLKNLELIDEEGLKIDDFKALVGEKYTGDWASVGSAALEKCLEKTKT
EEKESSKCKAGSKHVLLCIARESFINCPASDWTESEVCSDAKERVVKCPDIPPPMNH

>AMQ76475.1_odorant-binding_protein_22_Apolygus_lucorum
MSLKIHFVFAAIGAACVCAYQDQLKQTIKDCQGGKEVTDEELEEFTKPLIPKNEEERCIMACVMRTYNI INNGHYDP
KIAFGI IKGILKDHPEKLDRIKEVMDHCGEDVPQHMDNECDLAGEIMQCEVKYQKAMGLN

>AMQ76474.1_odorant-binding_protein_21_Apolygus_lucorum
MKFFVVSAAVLVLA AVKAN EKKANEKVTEIFNKCKETWPTVTDDEIEQVKQKQSI PDSKNVKCILACMLKEAKILRD
GEYNKDNAELMADVLYKDEPEHA ESKQI IEMCSSELGKTGDDCEYAYKMSVCASKHAKELGVKTPEF

>AMQ76473.1_odorant-binding_protein_20_Apolygus_lucorum
MYTFKTFVFLTLASYVIAAPPADEPAECKPMKEKEEEEISKCKLEPVTVKEQAAFVDCMKLVKDTDKKGPPKPEGFEC
LDDCILSKTGSLSGDKKIDPAKINAAKTYYTG DWAE PGAKMVEKCLAQVAENKDKTV CSTSGADVYTKCIFRESYIN
CPEKSWTNSDACKANKERVIKCPKTL PYNAEQHKAE TR

>AGZ04910.1_odorant_binding_protein_10_Sogatella_furcifera
MASGLVAAELEKLKNSCLKSGATEDTARRLVGINVV TENHVESCF LTCIYKGLKIVSSDNKFQPD TVKKIAD DHFLG
RNLKVTYQIAD SCKEIKADPADKCSIGASFRNCF SKYGKELGFFPHM

>AGZ04909.1_odorant_binding_protein_9_Sogatella_furcifera
MNTFQKFILSGMVVLGAMLITAEDTTIKIKNQSPHKQQQVY CQAPPTAPERLERIIEQCQDDIKTALLQEALNVLT
DTS PRDLVKKTRSKREVFSGEEKRIAGCLLQCVRKVKAVNDQGMPTV PGLVRLYSEGVD RNYVATVQAVQQCVSA
SQHFRYYNPQVLKEDGYTCDLAYDMFNCVSDKIEAFCGRTP

>AGZ04908.1_odorant_binding_protein_8_Sogatella_furcifera
MERTHVLI IAFAFI PFLSSAMQADFAMMQFPMQGTGT PMIQSIAGELKYCMDVNAEQNSDGL E DYLP LLFNEELPTSL
GQKCF L TCLFNRFGLLKDGFLDAQTAKTLVETFYKDKHDEKTMANIAINVC RVSAVPDILNPCEIGFSLKSCFVDSNK
KGKELRGKN

>AGZ04907.1_odorant_binding_protein_7_Sogatella_furcifera
MEEEMLSIGRVQRDADQTQEVADEYFKCKHRNLKTC CGKINLMKNYGDKGKIY GKQCYEEVVS AFKTNSSSTADDDDS
MMDMFSC EKV KMIKLKHICVHECIGKKT KILKEDGSLNAEEIKQYAREYMFNEEWSKELGERALDKCLTQSYNSVTKM
LDEYEIKCNPTSVQFHHCLWKEIEMTCPE SKVDLKA KCVR LRERLRKQQAAGM

>AGZ04906.1_odorant_binding_protein_6_Sogatella_furcifera
MKCQVFLASFVLVAVFELGYAGLTPEKLKEIKPLIDTCIKESKVEEETLGKLHNGHEIPSSQSGKCFIACMAEHMKLM
KDGKFEPAMTMEFIDKMVQDKVKADEIKKAVDDCFKSVPDGD KCEMAASLATCMKEHHAELAGMN

>AGZ04905.1_odorant_binding_protein_5_Sogatella_furcifera
MLYVS YFVIVTAASSAVITQ IMAADSNNPDMQTVFN NCREEASATEDDIKTFRAQQIPSTTTGKCMLACMFNHSGLMK
EGKYDSEGALKLVGQVFAADPVKLGAKTLINTCSDEVKNENDKCEIASKIADCTVKMTSQVGLS

>AGZ04904.1_odorant_binding_protein_4_Sogatella_furcifera
MCCDLPLVYRGTPELFKACREELGFPDHKPPPPPPSSDGHGPHGHPQRGMCVAECLFNRTG LLENGKINKEALKKALD
EYLKTDGAWKDVATTTLEICYDAQTRGDFKPDNEKFTSGSSEFLKCFTRNLFMDCIPEKWT DSEECKMKMEKIDKCPK
MLPPALFNKRPH

>AGZ04903.1_odorant_binding_protein_3_Sogatella_furcifera
MPPVFR TAHDKFESCLEELSSIFPPPPPPNGHHGPPPPPGGPGHGPPPPPPPGGRRGPPPPGFGGPPGHEPPIFACAHE
CLFNKTG LLENGKLNVEALKKKLEGELGDDEVWKNLLQSI VDKCMESKDPPSNDMCTSGSHELARCVL RDMFMNCPQE
KWKESDDCSNMKMKLEKCPELVPPMAMRLSQPPMP

>AGZ04902.1_odorant_binding_protein_2_Sogatella_furcifera

MSTLLNFVVFVFLVCLCSYSEASPALTEAQIEQVGKAMANMCISSSGVQSRSLITKAMTGEIEDDRKLKCFFGCIMEAVQ
VTNKGKMQPEVLKRRANAMLPKTMREMILPTIDSCSHIENEDKCELAYSIVKCHFSVNGKNPFFNF

>AGZ04901.1_odorant_binding_protein_1_Sogatella_furcifera
MLLEVCRFSVFLLAFSATVYGRFSEEEKQLMNQVHTQCVTETGTSEDLVNKATNGDFAEDENLKCYVKCIWSTLTVMD
DDGNFVDGVLEVMLPADMKDIVMKAMSACIGAGGSPCEKAFAVTKCLYKEAPADFFLP

>AHJ61048.1_odorant-binding_protein_12_Sogatella_furcifera
MRGYIILVLCVLLFMRGMASGLVAAELEKLKNSCLKKS GATEDTARRLVGINVV TENHVESCFLTCTIYKGLKIVSSDNK
FQPD TVKKIAD DHFLGRNLKV TYQIAD SCTKEIKAD PADKCSIGASFRNCF SKYKELGFFPHM

>AHB59663.1_odorant-binding_protein_6_Sogatella_furcifera
MGIYTTNVLITLLGSAVVSGVFVDRENGYRLTRQAPPDDECRPPRPGHNDDGVCCDMPPVFRTAHDKFESCLEELSSI
FPPPPPPNGHHGPPPPPGGPGGHGPPPPPPPGGRRGPPPGFGGPPGHEPPIFACAHECLFNKTGLLENGKLNVEALKKK
LEGLGDDEVWKNLLQSI VDKCMESKDP PSNDMCTSGSHELARC VL RDMFMNCPQEKWKESDDCSNMKMKLEKCP ELV
PPMAMRLSQPPMP

>AHB59662.1_odorant-binding_protein_11_Sogatella_furcifera
MLLEVCRFSVFLLA LSA TVYGRFSEEEKQLMNQVHTQCVTETGTSEDLVNKATNGDFAEDENLKCYVKCIWSTLTVMD
DDGNFVDGVLEVMLPADMKDIVMKAMSACIGAGGSPCEKAFAVTKCLYKEAPADFFLP

>AHB59661.1_odorant-binding_protein_10_Sogatella_furcifera
MSTLLNFVVFVFLVCLCSYSEAI PALTEAQIEQVGKAMANMCISSSGVQSRSLITKAMTGEIEDDRKLKCFFGCIMEAVQ
VTNKGKMQPEVLKRRANAMLPKTMREMILPTIDSCSHIENEDKCELAYSIVKCHFSVNGKNPFFNF

>AHB59659.1_odorant-binding_protein_7_Sogatella_furcifera
MCSLRMKS VSLC ILFVVVAITVSSAQYGP IALSTAQSAATAGLPPSATNKVDTFNRVRREALEANERVVRESSEEEDE
ACKPTKAEHEAKMCCDLPLVYRGTPELFKACREELGFPDHKPPPPPPSSDGHGPHGHPQRGMCAECLFNRTGLLENG
KINKEALKKALDEY LKTDGAWKDVATTTLEICYDAQTRGDFKPDNEKFTSGSSEFLKCFTRNLFMDCIPEKWT DSEEC
KKMKEKIDKCPKMLPPALFNKRPH

>AHB59658.1_odorant-binding_protein_5_Sogatella_furcifera
MVTSGLVQAVIAACLLLV TAYAYDFSDPYFNEHLQSAMEEIMEEEMLSIGRVQRDADQTQEVADEYFKCKHRNLKTCC
GKINLMKNYGDKGKIY GKQCYEEVVSAFKTNSSSTADDDDSMMDMFSC EKVKMIKLKHICVHECIGKKT KILKEDGSL
NAEEIKQYAREYMFNEEWSKELGERALDKCLTQSYNSVT KMLDEYEIKCNPTSVQFHHCLWKEIEMTC PESKVDL KAK
CVRLRERLRKQQAAGM

>AHB59657.1_odorant-binding_protein_4_Sogatella_furcifera
MNTFQKFIFSGMVVL AGAMLITAE DTTIKIKNQOSP HKQQQVYCQAPPTAPERLERIIEQCQDDIKTALLQEALNVLT
DTS PRDLVKKTRSKREVFSGEEKRIAGCLLQCVYRKVKA VDDQGMPTV PGLVRLYSEG VQDRNYYVATVQAVQQCVSA
SQHFRYYNPQVLKEDGYTCDLAYDMFNCVSDKIEAFCGRTP

>AHB59656.1_odorant-binding_protein_3_Sogatella_furcifera
MKASASLT LIFLTLAVFHCSEAKIDKAKKEAAIKKCAETSASDEDVKVRKEHVVPDSEEGKCFIACGFNHYDMLKD
NRINLEGVNAFF EKLYDEQEKRDIAIKAAASCAATESVSGLNDCHVAAKYFACLQRHPDFVKMKDDFDV

>AGZ04899.1_odorant_binding_protein_8_Nilaparvata_lugens
MVTSALMQTATAACLLLV TAYAYDFSDPYFNEHLQSAMEEIMEEEMLSIGRVQRDADQGQEVAD EYFKCKHRNLKTCC
GKINLMKNYGDKGKIY GKQCYEEVVSAFKTNSSSTADDDDSMMDMFSC EKVKMIKLKHICVHECIGKKT KILKEDGTL
NP EEIKQYAREYMFNEEWSKELGEKALDKCLSQT YNSVT KMLDEYEIKCNPSSVQFHHCLWKEIELTC PESKVDL KAK
CVRLRERLRKQQAAGM

>AGZ04898.1_odorant_binding_protein_7_Nilaparvata_lugens
MLLEVCRFSVFLI ALFATVNGRFTEEEKQLMNQVHSCIS ETGTSEDLVTKATTGDFADDDNLKCYVKCIWSTLTVMD
DEGNFVDGVLEVMLPADMKD TVMKAMNACTGVGGATPCEKAFAMTKCLYKEAPSDFFLP

>AGZ04897.1_odorant_binding_protein_6_Nilaparvata_lugens
MSTFHKFVISGMVVL AGALFVTAEDTTIKIKTPSPHKHQVY CQAPPTAPERLERIIEQCQDDIKTALQEALNVLTDT
SPRDLVKKTRSKREVFSGEEKRIAGCLLQCVYRKVKA VDDQGMPTV PGLVRLYSEG VQDRNYYVATVQAVQQCVSASQ
HFRYYNPQVLKEDGYTCDLAYDMFNCVSDKIEAFCGRTP

>AGZ04896.1_odorant_binding_protein_5_Nilaparvata_lugens
MGIYTTNLI FTLLGSAVVSGVFIDGRNEYRLTRQAPPDDECRPPRPGPNEDGVCCDMPPVFRTAHDKFESCLEELSSI
FPPPPPPPHGHGHPPPPPPGARGPPPPPPPGGRRGPPPGFGGPPGHEPPIFACAHECLFNKTGML ENGLNVEALKKK
LEDELGENEVWKNLVQSI VDKCMESKDAPSNE MCTSGSHELARC VL RDMFMNCPQEKWKESDDCSNMKMKLEKCP ELV
PPMAMRLPHPPMP

>AGZ04895.1_odorant_binding_protein_4_Nilaparvata_lugens

MERTSVLIVFTFIPFLSSVLGANFLMMQQSMQGTQMPMIQSIASELKFCMDVNAEQNSDGLNDYLPLLFNEELPSTLG
QKCFLTCLFNRFGLLDGFLDTKTAKNLVETFYADKHDEKTMANIAINVCHVAAVPDALNPCEIGFSLKSCFVDSNKK
GKELRGKN

>ACI30679.1_odorant_binding_protein_1_Nilaparvata_lugens
MKSFIVCIAVSYLLVANIKADEATSSSDAESLITSTTLSPASNESDAARSAIKEQLAKLTESCKTSSQANSDDAKIIG
TESVPKTEGEKCFLQCVYTGFGIVKNDQFSVEGARLLAQKRFGAFPEELEKANQLIETCSKEAVKKDSKDKCPMGFLI
RQCFVKNGQKINFFPKA

>AGZ04928.1_odorant_binding_protein_9_Laodelphax_striatellus
MSRLHKFAISGMAVLGAMLIAAEDTTIKIKNPSPHKQQQVYQCAPPPTAPERLERIEQCQDDIKTALLQEALNVLTD
TSPRDLVKKTRSKREVFSGEEKRIAGCLLQCVYRKVKAVDQGMPTVPGLVRLYSEGVQDRNYVATVQAVQQCVSAS
QHFYYNPQVLKEDGYTCDLAYDMFNCVSDKIEAFCGRTP

>AGZ04927.1_odorant_binding_protein_8_Laodelphax_striatellus
MITSALTPSAVLAACLLLATAYAYDFSDPYFNEHLQSAMEEIMEEEMLSIGRVQRDADQGGQEVADDEYFKCKHRNLKTC
CGKINLMKNYGDGKIYQKQCYEEVVSFAKSNSSSTADDDSMMDMFSCQKVKMIKLKHICVHECIGKTKILKEDGS
LNAAEIKQYAREYMFNEEWSKELGEKALDKCLTQTYNSVTKMLDEYEIKCNPTSVQFHHCLWKEIEMTCPESKVDLKA
KCVRLRERLRKQQAAGM

>AGZ04926.1_odorant_binding_protein_7_Laodelphax_striatellus
MKASAALTFLVFLALAVFHCSEAKIDKAKKEAAIKKCAETSASDEDVKVRKEHVVPESSEEGKCFIACGFNTYDMLKD
NRINLEGVNAFFEKLYDEQEKRDIAIKAAASCAATESISGLNECHVAAKYFACLQRHPDFVKMKEDFDV

>AGZ04925.1_odorant_binding_protein_6_Laodelphax_striatellus
MLYVLYFVIVTSALSAVITQIMAADSSDMLTVFNKCRDKTSATEDDIKTFRAQQIPSTTTGKCMACMFNHSGLMKDG
KYDSEGALKLVGQVFADNPILGKARQLINGCTDEVKKENDKCEIASKIADCTVKMSSQVGLS

>AGZ04924.1_odorant_binding_protein_5_Laodelphax_striatellus
MKCQVLLALFVVAACEVCYAGLTPEKLKELKPLIDTCIKQSKAEETIGKLQNGHEIPSSQTGKCFIACMAEHMKLM
KDGKFEPAMTMEFIDKMOVQDKDAEIKKAVDDCFKSPDGTGDKCEMAAGLATCMKEHHAELAGMN

>AGZ04923.1_odorant_binding_protein_4_Laodelphax_striatellus
MKSLIVCVVVSCLLVANTKADEATSSKPS SSPNAADALIASTTLSPASNETDAARAAIKEQLAKLTESCKTSSQANS
EAKIIGTESVPKTEGEKCFLQCVYTGLGIVKNEQFSVEGAKLLAQKRFGSFPEELEKANQLIETCSKEAVKKDSKDKC
PMGFLIRQCFVKNGQKINFFPKA

>AGZ04922.1_odorant_binding_protein_3_Laodelphax_striatellus
MERTRLFIILAFMPFLPSALGANFAMMQSLQGTNIPMIQSIAGELKFCMDVNSEQNSDGLDDYLPLLFNEELPTTLG
QKCFLTCLFNRFGLLDGFLDTQTAKTLVETFYKDKHDEKTMANIAINVCHVSAVPDVLNPCEIGFSLKSCFVDSNKK
GKELRHKN

>AGZ04921.1_odorant_binding_protein_2_Laodelphax_striatellus
MLLEVCRFVSFLLALSATVNGRFSEEEKQLMNQVHAQCVTETGTAEDIVNKATKGDFTEDDNLKCYVKCIWSTLTVMD
DEGDFDVGVLVLEVMPLPADMKDIVMKAMSACTGAGAGLSPCEKAFAVTKCLYKEAPADFFLP

>AGZ04920.1_odorant_binding_protein_1_Laodelphax_striatellus
MNVKMSALLKVLVFLVYLCYSYGAIPALTEAQIEQVGKAMANMCISSGVQRSKITKAMKEIEDDRKLKCFFGCIM
EAVQVTKNGRMQPEVLKRRANAMLPKTMREMILPTIDSCSHIENEDKCELAYSIVKCHFSVNGKNPFFNF

>ACZ58032.1_odorant-binding_protein_6_Adelphocoris_lineolatus
MGFKFVKYRSYFFVLVIRIILCIQIKAKELTDEQKEQIFAEIKNCMESTKLTDEEFESIMAKKELPTSIEGKCFTKCL
MEKMEYLEEGGKINVIQVAGMEENMEKESEITKAKEVIQQCADSVPPEDSCEYAYGISQCMYNKMKEAGISGS

>ACZ58031.1_odorant-binding_protein_5_Adelphocoris_lineolatus
MVLKMNLLLVVLVMSQVFFSVTEAAMSQAQMKQAMKTVRNMCI PKSGVDKEALAKMVNGEFDESQKLKCYLGCVLGM
MQAVKNNKINLTMVRNQITKMLAPERQORILAAFESCATVTGDDNCGLAFRFAKCIYDTDKEAFIVP

>ACZ58030.1_odorant-binding_protein_4_Adelphocoris_lineolatus
MRIFVIIFTAALTTCVMAGELPEEMKEMAQGLHDSVEETGVDNGLIAPCAKGNFADDAKLRCYFKCVFGNLGVISDEGE
LDAAEAFGSILPDSMQELLPTIKSCGGTTGSDPCDLAMNFNKCLQKADPVNFLVI

>ACZ58029.1_odorant-binding_protein_3_Adelphocoris_lineolatus
MDIRFGFIIACLAILSVANAISKEYSARMIAAKEKQKEFNVTDSVVEDFMKRNIPESKSGKCMVHCIMEEMGMIDD
HKINTEQVKLGKNEKWDPPALVELANQVADTCDQEVFTEGRCKCLVAVEYMMCLATHGDEVGLPHVDFEDSQDS

>ACZ58028.1_odorant-binding_protein_2_Adelphocoris_lineolatus
MSLKIQFFVFAAICAACVCAYQEQLKQTI RDCQDGKEVTDDELEEF TKPLIPRNREEKCIMACVMRTYNIISNGHYDP
KIAFGILKGILKDHPEKLNKIKEVMDHCGEDVPSHMDDECDLAGEIMQCEVKYQKAMGMA

>ACZ58027.1_odorant-binding_protein_1_Adelphocoris_lineolatus
MNSLIPVLLVCAAATRADEQTNAMVAKAFNKCREEFPISDDEIGGVREKTTIPESHNAKCLMACMLREGKMLRDGKY
EKENALIMADVLNKKDDPASADKAKQLVETCAGKVGTDAGGDECEFAVKMAVCAAAEEAKKLGVRRPDF

>AXS78212.1_odorant-binding_protein_27_Adelphocoris_lineolatus
MSSDMANILLVIGAGIAVNEGDAEMEALMECKKDYKVSREQIMSGDSSEEVKCFACELMKKTGGMDEGGNFNTEKI
KEEGRKHAKTDDQKRANDAAMDVKCISETEAANPTGKCEKGFEFFKCVRGEMKSPM

>AXS78225.1_odorant-binding_protein_34_Adelphocoris_lineolatus
MKPQGIDLNCPIYDTLQCFSSLIKHATPSQWSTTSECQGYRAFAAACPICPEDCFAAQVPIGSCNACLTLPRSA

>AXS78224.1_odorant-binding_protein_32_Adelphocoris_lineolatus
MNCPDSSWTKSDLCKANVARLEKCPMSMPFMPVLGLKKN

>AXS78223.1_odorant-binding_protein_29_Adelphocoris_lineolatus
MQLKNTSDSLVGHFLFSIALPNTLLNPSQLPELAITSRQSPDIVSIACFLHAIGSNFLFASKTPSEANRQSWMHSNSL
ESVPLVSFESQVVIVLLLVKGN

>AXS78213.1_odorant-binding_protein_28_Adelphocoris_lineolatus
MYQLTCPVFVIVMYVWNQGSADVGLLDYSAQCVTEAGVTHDEADKIQQGNLPSNQGGKCYVACVLKSLGLVDRRGKIS
AENTNRLIDMYSNEPGDAKDKTKQAVNTCATEANRAWTWSQCEVAYRMMSCLIRTRGSLQQTRSITITLPSITINPP
PLSFTLFSVG

>AXS78210.1_odorant-binding_protein_25_Adelphocoris_lineolatus
MKSLWKCKKTQSSKATHPTSEQMVEIKSCYSNWNHVSVEETNAAPGFDCEVEECVYSKLGFMGTDKTINKEKLLQFQKE
ETHEDFHEAITKSMDMCMGKTFTTKCPSGIDAVIKCEAIQIYLNCPAKHWDNGDDCQETKKLMEKCADVTSMYN

>AXS78209.1_odorant-binding_protein_24_Adelphocoris_lineolatus
MDVMILPVFLIVFIAATASPAVLSAECPQRFKEMRPLSSCKVEMNPNSSYNSTAGEAIIIDKCFGGFNSSATRPTGP
PSGYDCEMECLMIEFGFMGDKDTINKDKIVTSIQDEYSADFQEAANKAIEICMGRKYQTSCTPSGIDGMMECFVQMMML
NCPAKHWSGGEDCKETKQLIEKCGEVLISIFADYDTE

>AXS78208.1_odorant-binding_protein_23_Adelphocoris_lineolatus
MPTFSTLSTFLFALALTYGQIVEDPECRPPHPPGREDSCCILPELFIGGIQDVIRKCHDEAGLKKPSGPHGGGSPPSA
EEMAAHWRAHECADECVLKSQNLTTDGEINLKDIAKAQVVKSFSGDWAKLASDIADKCLASAKGEVTSTATCKSGAGQ
FIYCFRRNLFLQCPSSWTETTDCAAARITKCPNAKIPMGHHRH

>AXS78207.1_odorant-binding_protein_22_Adelphocoris_lineolatus
MNTFAALVLVASVVALSQGNPTTGAPSTTLVSDQPTGSTASGVSKSPPEIKQKIKEQVGALTEACKSQSKITGEQAKI
VATQAIPKTEAEKCFLECIYTGQLTKDGKFNEPAARALAQKRFGNAPDDLQKANSMITACVKEVVVKDTNEKCALGR
LIRECFVKNGAKINFFPKP

>AHF71057.1_odorant-binding_protein_26_Lygus_lineolaris
MNQSSCILTLALTIVAMAVVSGFKELDSVLPQAKQEECRKESNFGELSGDLSQNVQTQELKCFACSLVKLGLMNEKD
GTINTTQLDELIKHTEGKDAADMFKHSVVEPCLKEVNKTADYCEYSFQLVTCGMNKVKPPTTG

>AHF71063.1_odorant-binding_protein_32_Lygus_lineolaris
MEVNAERNVTDEQRXAVRLCSRYTEVESGLAEAGYDCLAECFFIKLGLMGEDKTLNKENILEEVRIQFHEDXVEPARK
ALETCMEKKYNTKCPSGIDGTMQCFTVQLMLNCPXQNWTDGEECKETRTFMEKCGETLNYYD

>AHF71062.1_odorant-binding_protein_31_Lygus_lineolaris
MLTAYMIVATLSVFFFAVALTQGMDEDPDCRPPHPPGKEAQCCPLPDFVGVVDNFHDMHKCSDEAGLRKPSGPPGS
GTPPTAEEMAAHMSAHECADECLFKNTKYLQSNGLDKDAIKASVTKIFTGDWAALASSAADKCLASAKSEVGASAKC
KSGARQMVKCFTRAMFLNCPASSWTESTECAAAKARITKCPNAMVPMIPPKH

>AHF71061.1_odorant-binding_protein_30_Lygus_lineolaris
MNSFTVLCLVASVVALTQGNPTTPNPTSSSHAASVSGGSTVSGVSKSPPEEVKQKIKEQVEALTGACKSQTKITGEQAK
IVATQAIPKTEAEKCFLECIYTGQLTKDGKFNEPAARALAQKRFGNAPEDLTkansmidTCVKEVVVKDLNEKCALG
RLIRECFVKNGAKINFFPKP

>AHF71060.1_odorant-binding_protein_29_Lygus_lineolaris
MNRPLLLLTAVALVGSQQEDCKTAPAGWPKRPPQCCDLFPFLEGMKREFGSCIRQIGNRQSSAVPTAQAVRDARLCI
EECVYKGLGFMEHNLNKDQLLEQLKKGAVAGKKDWEKPMEDAVKSCHEITITKRETPQEGACQDSAHEFTHCVMRQLFL
SCPASEWNNNDECNLVKNRMQACPNIPPPPPPPQGFQGGPPQ

>AHF71059.1_odorant-binding_protein_28_Lygus_lineolaris
MEVAACLVLLAALAALTSVDEKRPCKAPTAPRKLEKVINQCQEEIKYALLQEALSVLGETVSLRTALTRNRSKRE
TFTGEERRIAGCLLQCVYRKMKALDETFPTATGLVKIYSEGVEDRNYLATIQGVQQCLSRELQNRNKNPSIVKAE
YSCDVAIDVFNVCVSEEIEQLCGTSP

MFTTATSTIIIFLFAVALTRGQMDDEPECRPPHPGKAGDCCVQPKLFDEGMPDVIKKCHEEAGVKRPSGPPGSGTPP
TAEEMA AHKSAHECAAECIFKNNNFIKSDGELDXDAIKATVTMKFTGDWATLASTTIDKCLASAKSEVDASPKCKSGA
DOVVRCEGRSLFIGCPASAWTESTECAA EKARLT KCPNAMPPPPHHKH

MADSVGECMKLIKVKPEKGPPVPEGFDCMDTCVFSKLGFIGADNKLDPKLAKKFSELFKGDWSALSSESTLKKCLPMA
DVGKGYCSSGADVFEKFLIRELYMNCPASSWTKSDICKANVERLEKCPNSLPFMNGSGLKNKSSR

MYAFTAALSFFLLNIAFAHPGHGFDEDPECRQPHHHRHEENDCCCKVPSLFSNNKDEMHELHVHKCFEEAGIKKHGPHHEH
HGPPPLEDGPPIPPPPPPFSPKNSDKFDCVEQCFLKNLDLVDDEGLKVDLALVTEKFSGDWASVGSSAIEKCLEK
AKTEENEPSKCKAGSKRVLHCLAREFFMNCPSADWTESEVCLAAKDRVSKCPHSLPPMH

MLTAYMIVATLSVFFFAVALTQGGMEDDPDCRPPHPPGKEAQCCPLPDFVGVVDNFHDMHKCSDEAGLRKPSGPPGS
GTPPTAEEMAHAHSAHECADECLFKNTKYLQSNGELDKDAIKASVTKIIFTGDWAALASSAADKCLASAKSEVGASAKC
KSGARQMVKCFTRAMFLNCPASSWTESTECAAAKARITKCPNAMVPMIPYPQPI SANSTILDSL FACITDGEYMEFQ
SISSGRLG

MKLVKDDKARPPKPEGYECIDDCIMAKNGFLGTDKKIDAAKVNAAAKTSYTGWEAEPGAKMVEKCLAQVSANKEKGEC
 TSGADIEFSICMFRESFINCPKESWTSSETCKANKERLIKCPKSIPIFLNKSAT

MKSFVGLIF¹ALVEFASAVTKE²YH³DR⁴AA⁵DKCAKEHN⁶IK⁷ES⁸EIQ⁹EFVKKHKL¹⁰PETEDGKCM¹¹AC¹²YMEEMKLITD¹³G
KVNVD¹⁴EWKKS¹⁵SNKEKWDEE¹⁶AHVMADEIV¹⁷DKCNEOVSPDGLAKCEYGF¹⁸KLTECGLK¹⁹HRLEKGLPAPNMDDV²⁰KRR

MDGRIQRRGKKNSQLQGCTDYRQFFFCRLMGLDGGGLIERGITFRLVFIVVASVISETFSFLSIFASNCADQKLD CMRV
FIILGLTINVAPFNEYTNRHALDRHLHKFLDRMLLHPSRTGLKEEKIILEDARHKSSKAVNFYIRIFSGYLGTM LFAMPL
AELLMGHSWKKLPIPWISPTDTAKFLLVFGFQSVAVVISNCLGMIMMSFSAITVQMTALFEVLLFSIRHIEDRAAV
NSK CERISYRDAMVACLREDIMFYQDIVRELTSATPHLRNMFMACSVSIPVVMACEAYPLVSGNFVLGALIKSCIFFS
I IILCWAQCKTOMETMDKHKAVARALYETPWYENGLKFRKMYVLTTLVFANQPKYIKARFSIEINANTDTFYSFVS
FNYLNLMRKMN

MVYKSPKFEEREMHEVLRDYYSVLPVLGGFFPFLHTRGWRIVSLIHYTEIMYSVMILAEVNVAYTTYLVDRDPSLLSQS
 LHLFIVGLVAMGCSLFLTLQRRKIKDYTTIFNHQFSMCEYANNFFVVGISKKTSQLNKVRILTWFAMYGSCGVVGILQ
 PVINSFFDIHFEVIQLNGVWMNLPFLMWCPCWDLEKSNTAWFWAFLFQGVYLSYCAVVVTSAIMVQFNVIDRLLDQFKL
 IITYGVKTLDRAKIIFKERYPMLTRKLNKRAYENCIFYDCIKQSVLHHLRVGKEYKSFVELSTVPVAVPFYGGAILLGM
 AMITMTADDDPRIGPKFFFYGMAMVSEFINMYLICQAGNFIDITTEMFTALYQTRWYLWPKRTQKALMIMRFGCIRPL
 KYKCANVLNMLELFSFLVNSAYSIFNLKAASS

MTGGIKQESENGVEVFSTQFKIMEAVGMWSTGPPTRRKEVITKLKVGVSFVLGVSTVALLVKAVVTDNLVDRSEAMD
ICTLTGSACYKLLFFYYYREEFQDFVSWDAALVPHLPPKWTAHMNNFSSILHCATGLTLCISAWSMCPLFKWAYKEIELK
DMTLPMNVYDPDFSQGLLFAFFYLCGVYGLISSTHIYMSCDAFFFFSTAHVANGAYETLKTkFKSMGVGNENSLGDPMH
DELKACVKLHTHIITFIRKTDALFKSLFFVDVLHAIISLSFSLQASESKGIFEKLKMITFVITYCFVHQYLNYSYFGQR
LIDQQESLANTLVAYVPWNEGLTKMKKSFLIVTANAQKTFKLAAWKVYTLQYATFMEFVKMTMIQFYMVLRQVQDESQT
S

MFFKKKKKEVTAYAVKDSKDFSPDERPVFVWLGYVYDQSLLSKVRLVFQVFLLYTAPFHHMMIIVYEDLTVDEIGLALS
TMMYIILGISWPIIMYRSKEILQMGE LLRKGYFEYFDPPTLEEK TILSKANDIIKSRRLSILAYFCAGLGTFIKELS
PDSFRQYHMPYPGWFFEMSSNYRFMAGFLYQLGLCVNTTAAYLSIFLLFVFYSLHFEAQFSLLIQHFRDFTPTAMSP
DFVRTPARRVWTQMRLEKCVQHHLDIIDFHQMLGFFGPTLTFVFRITCTVMLCMHCYMINGSEIEIEELFQLVLGFA
VLYLLFIFSLKGEKVTAMSDLWSQTLYEVEWWNQPVLEVQKTVLLMLVGANRPLLIYGVWTPVLYSHRGVTAIGQETFS
FLNLRMN

MVFGITSDDDVGLSRNTMIEEDITLLQLAALYPANKYYAVVFFCFLLWGTIGVAVSGFTCYSEGDIGFAFECLHQSVVG
 ANMLVQIAAHNYYSATLSKYFRKMDENYYSYNDTMDDDSRKFIADVAHEKKQRKLMFGRVFKVLIVGCVGGVLKSSA
 LHLIRGOGLKEVDGLNWIIEAPLKVYVPYSDOWGPYLIGYACACIVPFEVGFACSGSISCFIRFSEELLHOILVLNV

GLRNTIKRAKHTYKVRYGREYVGTDNKQFEHCLKECMKCSVEHHLVIIQLFADFRKLMHIPLFTVIFDGGALICLAV
VVILTNDGPSVILPIPFVSAELYTYVYCGYGERLTSASFSEVGYQLYLDWYLTEAKVIQPYLEIIRSYSFIPKELS
ALGFATPNHQMFSGILRTAYSFSNVILARD

>AXU25099.1_OR8_Clividipennis

MGFASWIKDARERDTEDIQDVVKYYHVRPVLRIISLSLNTQGYWLHFNLVYGFVYCVLGLLEATLIFSINKFEDN
VDLMIITCHHAVIGGFTLGMNTCFIYKRRNFSKMLEMMKSRLHDYKDDPVLHTTKGRALKAAEFVVEVMSFFPILIV
GFVFGVKRIFTIIFPSDVNVTETGINYLVAHPIYCPTEGFSSVFYVCEIAEVVFIQSASAVASVCSLYVIHCIFLIT
ELDYLRLSVANFESRVKCRKRTDGRKMPFLKQKMQECASDCLVENIKHHHAIKKYFDQLNDMMSTTIFIGYCGTFT
MGLAMMKLFTKKENQKLDVGNTSLIGLLFMFEIVFMLVVSYFGEMITTNHRVFEQLRFMRDLKFTGVKFRKTLMLMT
IGMQKPMQLDAGGFSKCNMETFASVMNVAYSMLNIAAAKNETDK

>AXU25100.1_OR9_Clividipennis

MSGEIDKIYNAPMFTLPSWQALDTADKKIALKKGYSVSGGWIMEFSGLYLQGAARPIIGFMVAITANFMMGSAAYFA
FVLDDHLECASEGLHLLVMTSNMMWITMNLWYRPIVDELYTAIGAGFFSYGDTIDEVTQKEMDNAISEMRKKKKVFM
KVFGGIVMSCGSMFLVKCLVAYFRFGHRVDGEGGTVLRRQVVPVWFFGVVDKWPYGLFACFMAYSMQAWNMLTIVGSVL
PMIAFFEEMIAQLLLVGIGLRRLVRRRAKHVLTNTKYQSNDKILFQKCMEDSLKASIQHHIVVLEMCAKLLQLLMVPLMT
IVFGSALLLCMQGYVIEDTVPLLVKIASVFLMAELIYTFLFISIYGETLTNASEDIGNQLYLANWIQFSNALRPYYA
MIKFRCSRPRISAGGFIVVDLKSFLNVLSTSYTYVNLMLSSRK

>AXU25101.1_OR10_Clividipennis

MSILRQVEDYITRMRELEDEDVRKAIRDNYGPLLRVNNTFSRIDGGWIYVSVLHGIIIFILFVFSYIYLFITCYLVQD
DFVLLGVQFHYLLLAFFGASFVVHLLGCRPDVNQIHKIIGQGFFKYTNDHLWSKEQDEFKKEMRKEQLQLIPFLGLIG
VIGLFIVVLGPLIDNMLGVGYTDDYNGVYMKTPIMYFPFIEIDSMVKHHLATSQOMITVLMALGIIIGVVFIYIVTT
QHMALQKILNTSLGKLNKRIRMILKSRYPKHAIQREDITKDEKYQECVTFLKENIKHHQMIKQFQRR LAPLVGVPL
LAAFLMGTMIMALSLMILVKRGERIGLLMTNTIAMIGE VANLYIACSFGE DILTLSSE LFDKVYFIEWNTKMNVKNRK
IVLSFQGATKDPITLGAALFFSCSMATFSSICNSAYSYSMLTATADQTAEN

>AXU25097.1_OR6_Clividipennis

MKGDGAKVTKPLNDVLKFGGIWFDHDLKYGTLLKWLNFLRHATAFVWVAIETSNLWFA GSSFILTESAIFIPIGFEE
LGFFVVI VSNMSTVN RVVGKYEEKIDRYSDTPWGKKIIDEEVRVYKVFKL GKNLSLVGFFIMYIVVPVSFDAIRAIAG
LPPYLVPLPLNFLLGKQPERNYTYTAIFVSYLFFAIIVPRFVATQSLILFFVLFIVIDVRICIKKIQIMSGNDHKGV
GLQKEWDLKDIIDHHSQILD MVKREYNF IGFM LIAQNGFIALVLC LLSYMMKTS LNAGDLILAVFCGTFEFQVMIISL
LHNGAGYIVEKQSEDLANELYNTAWYKQPPRIRKYVNTMIGQGVKMFNISYHMS PANLEAYLQVLNKSYSFFALINSK
V

>AXU25096.1_OR5_Clividipennis

MEATLKL VVKSILILSGDEKYQLFSLQKALLNLVGLDFFQNGPQIRMIPHVKHKIYPLMFPMWMIISIAILVLKHVSS
NQPLTQFEFIKQINGIIVSMTFFAASF KFFIFFIFGKSFKELFIETEQLGVIHADNVKLQFHIKTCVGYTTLVASNLG
TWVAWFY LKHDDPEFDAAYPWEDKGFGRLLT YFFGFFAATFCAICHTLVDTAFQMMVAGIIIHIDSVAEKL SLVGSGK
IKDRHAMSEIYEKHVKLLRVANNFSRGYSNLFVAQSVYTVGHSCVLLFGAAHVDDRVEIAMSQGTMLMTSYMQLLIYC
YYGELLTSKFSGLLFDSYNNEWYECDMAVKDLGRFALMCHKHISLRGFGNVHPSKSNWLHVS

>AXU25095.1_OR4_Clividipennis

MTKKKVVQPLEHDPSKFSTNRELLENFGLILTDNRSFRPVGSMTLILSGMLHYSLRVILT YHELELLTVMKVLHLTLI
YLLVAVCGTVVL RQSKRFLVLHSIIRNGFYKYPQGLTDEQLMIRSRANEKVLKITKYGLRFFVIATLGSTLKEPLPYD
FQVWTRQYNGWFPFVVD SWSKYVVSQIYHLMAALSAALLGGSVCIVFIAIAEHLLAQLEILSISLKNAIHDVPDPGDS
RANYLAYLKI KSCVKHHHVILRFFIEFQKYYSIPVLSMLAGTTVAMCTIAFVVTDPNSTFGVSA AFLSLMAPEIAFCL
CYCTYGQKITDMGFNVRD TVCNAPWYCHSKKVQLALLMILIKTRTPLTLTAAGLKDCS IKSIGEVRL

>AXU25094.1_OR3_Clividipennis

MSKNKKSVDIAGTYKDVINSAGYKANFGTIMKFGALYNETYLYPVALLIFFIVNTSLLAYTITLATDDVKLMAESAHF
TMLRTTATVVL LNSTYRKPLTNLFLLLGEPNLHDYGGTLTKELKAKIQEKIKICVDRKEFYGKNLARGFFVALVLFA
TRSILEYFGGFLNQPYGEDGINRNLVPPTYMPYDSEGWPYQFALISEAYLVAAACLVAIGNDCSFICFSEEVRLRELQ
IIIIITLGSIKERSEYVRRKS NYKISVKDSVHLCLKHSVKHHQRVIQVFEGFRTYFFFSLFIMLGCGAFLVCLSGLMFI
ADVVDVKTKILFFIFLASELFHIFVFCYYGEHISFTSRKVGESLYDPSTVDISQYVKPFFIMINARCQTPLQLSGGGF
MNADFDTYGNVLRSSSYLNLQAAN

>AXU25093.1_OR2_Clividipennis

MSDRCMSGFPFVSCALCLFMTFYPRVRVSKFRDLLESYEEVRRNAPYSEHAEAKILQQSKIVRSITKVLVMLEFCPMF
IIALILPIVNEMTDFAFGERKLSIPSRFPFDPLASFASFLPVMILHTSGAVFVNFKKIAFEETFFIFLTRHIAMVNNL
KYNLEMILDSL SVDDQGRLTNGEKLNDASWRYENIQKPLKKWAQNHQECLRIFRDLEALYSWPLAIYFALTTLILCTA

AYVTAEINELNHALTGKNTFLLNKTEHGLLKLILTRCRQDSIIKAVSLFPLSVDTFKSVLTTSYSYMMVIKIVNGEED
KS

>AXU25092.1_OR1_Clividipennis

MEDDDPSQALLKAGYHNLM SVACGLRGPRGPLFKRSIQNYLFTFHGYFGAAVSVYLTVEILTYYATATDLEFTKEMFL
LVPLITGTHSNINFFYMEYCRPRFIRMLRMNEEASKDKFYEEGLQEEIKLWAARARKLQPLLYFIVSAPVFPWGMSP
LNPSRRDVAFDLWFPYDINDLKYWVLTLCIQTTAGFYATLQNVMFDAVFFCFTFRILALLVHLKKTFFKHIFHVIHVNA
AGHVFYTDYSGRAAYREEIEDQLTERLTYYWISQHQATMRLLENEALYSIPLLFHFIHAGSVLATGA AVLVKGNLPLG
ELYFVTVHLLGLIATLFVICRIGDLLKTQTSEILDGLAGKNYFMLSQYQRTILKTLLIGSKEPFVVRVVQTFPLNTET
FKSIIATTYSFFTFLQKAKG

>XP_024219388.1_OR4_Hhalys

MDTFRNEKIALRLLGITFPKDLSSPVLKTMVTLYFWLQIAAEGYWAYLYFRGAYLYLHEHFLIRMVGLSAAISVSGTI
AKLIAFKLNSGLLQELFDQLGGFVANTDRGKLVKKTILYESAVAAILVTHNYGYFINSTPTRSTPYYSRYQTGPYEE
VALDYWYTL YCVFVTLTGNVMGDTVFLLIINQICHRLLLLSNTVSVMGKQEKNSTMEKMDMRGGESDQEI IQMCVREH
NTLLKFKNQLEELYNYVIVVQLAYSFSTICLTLFVISQEANVMEAMSELLPQTISLFL EIFLYCWAGQMINNHFEELR
LAVYNSGWYDSCRKEDRSSVLIMQTFTTVPVEVTAVSVFHLNFQLYETICREAFYYTIMSQQLLGQ

>XP_024218985.1_OR82a_Hhalys

MERLYITKFIIILRGHQFATRSR SVMIDEGIEDSDILEGLNVRYLKFIFGLWKVINDYRK TGNRNGIIKFEVFCFILTL
PYVICQYMSYFNIEVDIQKATFLNLYPLPALQMCCRIAVFWFRIDIQCRLYDLIKKNFLNIPEGSRAQIEEIKRISK
VSNMCCMATLIVNATIVSLYVNVNPGISVDYILYHTGSMTEVTTGRKKILGGWYPVPM AESPYEIIIFVYEAICILWGG
LFLAVYFSLYYHVLVCLYAHFTVLGFGQIENLKMKS VKNGGSKSCIKRDQNNNSIVYENLIDIIRDHQKLLRYFEELRT
VYNPLVTLTLGIGVIVLIIGAIQFLLGKSINPGFIFQISEVFALQGLEVCMFCFGSSSIEAAGLTGLAFRQSSGLQFA
IYSSDWYKANVKFRKAAQMLMVRANRGVSLTAIRMPVNVETLMAFEDKLLYAQNA

>XP_024218692.1_OR4-like_Hhalys

MLFCLNFSGMVYVNGGSFLKAYVIYQTCAFIIFGTSLILFAYFEPKPLIAKLQTLHFLIALSQIFCKFLSIALNQRET
VDILDNLEKLSKEAFTDQKNAEILKRANKSSTKLPKIFMYVFCGYPLRTINLLIRYL VFKSPQKFVEASWIPPFLAS
NYLTGLIYQVATCLCTGFTNAAYTGLIIFLTIQLSGQIQILRRAMEDGYPIKECVVLHQKLLRLFFRINRLLNALMLA
EYILSSLYSCISTFLILKSLSSGGKGVANNAYVLVLVLLRVFLICYCGNIIQTQGEKLHSSAYFGEWYRKS LPEKKS
VQLMMSLT MVQPFKYNYRRILTVDMALFLKSIHTTYSYLTMLLSFADHK

>XP_014290613.2_OR45b-like_Hhalys

MVKIIEKELLGDLPRILIKTFLISRDSKKRTGQTFVTISNAIMTLVCISSMYFLGLEKSLEGTASF SAYAFIIATKHVI
YYFRQDEIKRMLNIFARLQKEYKEKWEREIFKR DSEDTWNMVYKFC LIMISYQVFYLMFTISVDYIIGNIFPNFPSVR
VHLPCDGFIEFFEPRTLGRFIIITIPVLMWTAEAITIHVGSETLVFVLMMYTKSELKMIRYRLIIKKQLNRNKSNC SV
NAEMLLWEVIQRHQRALDVLDMKDTLGLPLAIHYTSVSVTL SVIFYCLMTFDERGTLTAKFNGITAIICIGSLLFAL
CYFGESLEEENNEINKRIYDLPWYNESIFVKRAI IIMLRQTQKPFVINYRLTAQLSLQTFMQVINTAYS YLMVLKSTV
R

>XP_024217354.1_OR94a-like_Hhalys

MKPRNCSRNVVTINVIADAQEVIGPAVAAQIALGAILLAAFLFSVMEIKGNISLTVTYIILFNIMASLNFTFCFFGEM
LQTQGNELFECLCLPWNEMTPKVRKDFNVILVLARRPILISYKGLLPVNFESFQFVMNSSYSFLMLLTSTKNANDM

>XP_024217426.1_OR24a-like_Hhalys

MSYQTVIADSDVIHGLNVGYLKFFGIWKTIKEYRACGKRNAI IKIKVFITLLALPYIFCQYLSYFM IKVDIQKATFL
NLHTLPALQMCKILVFWFRLDSQSKLYDLMRKDFISIPKSNTSVTDEIYEKIAKKS NVLCKAAFV VNTSIVVLSIAD
PGISVDYILYHTGNMNTVTTGKKKIMGGWYPLPIDKSPFYEIIFAYEMIVMIWGGFLAVYICMFYQVLMCLY AQFAV
LSHHLTTLKFKLNQKTNRNNSLLYNQLRDIKDHQKLLSYANELRSVYNPLVTMILGMGIFVLVIAVFQFLFGKKGN
PMFIFKFFQFLAYQAIEVSMFCFGSSYIETASSDLQFAIYSSDWYKADVKFRKAAQMMVRSRKGETLTAIRMYVSRL
DKVVAGSGWDAKIAVTSRNVAGSNGNVSSEEGVEGWMP SASIIYALGKTCLDRIDIGVLKANANSKSDLIGITGG

>XP_024217428.1_OR30a-like_Hhalys

MYSGSRWTHQFWFRSQSAMIKEGIKSDSIDGLSLRYLKFFGIWRIVNEYRTTGKKNAIKLHLLGTL LLGIPYVIFQ
FMSYFVIKVDIQKATILNFNAIPALQLCCRMFVFTFCMESQCRLYNVLRKDFLNIPKQNLVKEIFMSISKTSNFCCT
MSLAVNGSIVLFYIIYPGVSVDYILYHTGSM AAVRTGRKKILGGWYPVPIDQSPYIEIVFAYEAI VLLWGGFFLAVYF
VLFYQVLMCLY AQFSALGLQMSSLKIEQYRS DINTSLNHNYISSTVYEELYKLLKDHQKLLRYTEELRN VYNPLVTMT
LGMGILILIIIGAQTGPVYIFRFLQLLAYQGIEVSLFCIGSSSLQTASSDLQFAIYSSDWYKADIQFRKTAQMLMVRA
NKGVTLTAIRMPVNVETLMAILQFTYSVSTLMSRMT E

>XP_024217960.1_OR67c-like_Hhalys

MIRMLSFFVSGKQPVAQEPRTRTLNMLADTISP NFFMQVFATLSIVIFSAFTAVTSNESQHFSA CLSTLATLLTYCWL

GQMTNATEALHFAAYNNKWYNCNRRMKMNLQILSSFTREPFEIRGCSIVRMNLQTFKEVVTDSSYLMILITMDDTS

>XP_024214894.1_OR33a-like_Hhalys
MAGTIVIALSLLSSGSGIELPGTTIASALLCFAEVGFMYLFSVFSQHITDLSLDLRFMTYSTKWYYCNKSTAQSI AIF
QTMTLKPMTLTAGGIVPANMDTFGKVMNSAYSYYNVVNAFNITV

>XP_024220019.1_OR49b-like_Hhalys
MPEHLFARVLQCPARDFQENAVALEKEKPVRCLEMASSLIGKLVWGFVKFAEGDIMQNLF TIFAAGPGMVGLGKFLGI
VFYRKQIKSVWDRLSAMISEADNPKLEHIARNSLKRTWILFIVYLSIFTASVTDWLMRPPAAAILYHKKARIVDTWPV
FLDTWLQWFLSYLLQFPAIVLLGHSNYTYEILYFCTSEVILCHFLLNYKLKHLVLRDNKESTEKLISCVKYHQNLLS
VCNDFKNSTSKVLIWQFLNTVIMLCTGIFILTVLIKQTPYVAVINLFEVCTFEIMSLYLYCWYSNEITFQCSDVSNV
YMAEWINAKPSDKKTM LITMSKAEQPLLFGGILEIRLETFINILKTSFSFYNFLLAQVNSTEK

>XP_024220181.1_OR43a-like_Hhalys
MLS LPCDRGRSWAFNLIKEISSIFIGSAYLTGTIVIALSFISAGSAKELPGTTVGSM L LCAGEISYMF L FSLSGQRII
DLSTELRYKMYNTRWYICSEEIKKSF MIFQEITLKPMTMMVAKVVPANMETFTAVMNASYTYYNLINAVDKK

>XP_024214259.1_OR85b-like_Hhalys
MATKTTEFTVVEVFHWMCDFTIMTSMTM SCLYNPILIRMEARIKAGFFDYGGPLTPEQIKVRATMNRNSQLMSKLY
CSLCYCGLVATYIHVFFAKKEYLLPYPMWFPHEIRNMYQYVLTLIHVFI VAETMTICAFSQLSAFVALSSHLIAQYKI
LIIAIKEIDLIAGTKATPEEKILKMHAKLKTCVKHHVIINKYFDDLHLKLSIPLLT TAIFICLAICTLGFILLSPNVS
VPVIGALIILFLPEMFII VVYCVYGQKVADVCEEFGHTIYYS DWYTKPLSVQRDFLMILIGSRKKRQLTGFG LHEFSM
KGLSEIIKATFSYFNMLRAMN

>ANC27967.1_OR40_Alucorum
MFRGKEKERPYVIKDPKDFDTRDLVFLWLGFVYDGSFFSKVRLTVFVLLLYSAPIHHMLPVILDKSTTTDEVLI ALSI
NMLYVLLCIAWPFMIYRSPEIIRLWSTVRQGFFHYSDPLTSYERTILSKANDLI IKSTRMSLIAYFCAGFGTYLKEMS
PNSMRLYNPPYPGWFPWTINSNFRFAMALLYQLSICLN TTFALEGIFLLFLFHTISFEGQISLLKQHFEDTFPPGLPT
EMTHVPAFKERTLKR LKECVRHHLVIMDFHKRILSYFGICLLVYRAICTIMLCILCYLTTTGIALNKFLQLACLAALI
LYLLFIFCLKGQKVSKMSEIWRETLYEVDWNNHPVEVQKAVLLMLVGAGKTM TVYG VWT PAMYSHEGISAIGQETFSF
FNMLRAMK

>ANE06404.1_OR46_Alucorum
MGYRVYPQQDLSDP SHMFSFQLNALKSTTMWKPDNQKYYIPFMILFAVNVFVLAICTVGLLLKGCSTKDLVDRSEAMD
IFTLTGSALYKMVFFLYHYEQLVDMVTCGLALVRNLPEGWTKNCGLLSRIHYTAGFLVLLI WGLAPILKVMYGETTWA
EMKLPINTYDPFDSTGFLFFLFYITGQYVLVLSAVIYMAADCYLFTSIYVAVGALQYIVDQFENMRDLNNNNKHTVAD
TMHDCLOQECIEIHVHVL DYLRKTDKLFKSMILADVHVHAVISLSFAMLQTSSESKGIFEGVKMVL FVQVC FVHQFLNSHF
QOELIDKQDNLAKQIITDIPWTDASRKFKKSYIIMLTVCVREPFKLSAWN VYFLQYATFLEFSKTM IQYYMVLQEVQDE
AEVS

>AQM56113.1_OR_Alucorum
MEESWLVRYYGGGLGQAEYERVRDFAVSEFTPLVLFLGIFPPTDKMALMSIVVSLSI AVYAFYIVLFTITCSFATDDF
VLWSELIHHTSLMYLGIFIRSVLILEAKEMIKLARDYLDGIYHYEEGYVDPIFQQLQDKSRKLQ RKL FMLPLFIVLVT
GIALGLKPLLDVNEVEPHPKLLENGITYRSLPIFYPFNNTYQVLLMNGALLYFAFLVVVTVIAADLLFIRVSCR
ISLEIAILVESLNLIDKRAKRLYARKYGLNKKNESWPLYQDCIEECIKENVKHHQKIIIFYEQFSAVAAPAIGGGFFT
CTIVLGLGMIVVNMDNVNISDIIAFVGTVFAEMMNAFMISWMSEKIGE QNYELYN AVYNLKWFKWRQSNKKLVITFLD
GTRQPLFLNAFGMATINMEAFGSVVNTAYSFLNLVNASETLEKK

>AQM56101.1_OR_Alucorum
MIAISLLALMSGDGKSSVLLRAALMVVAEVLYLALVCGQSES VQELGESLHEELYNMNWLDLDTPAKKTIMIMIEQSK
RPLVLTAAGLQPLNWEAFTRVMNTAYTYVNLLLAANA

>AQM56102.1_OR_Alucorum
MAKLFGFHIIINADSRKCLSYRMSFPMTTLLMAFNLCMLSGNAIGCLIAALLDSNFDRRLMNVKGMMLLIMIILLAIN E
SFISRNRVNRILDYINRIKSITRYGFKEEEDIMNKAIDDCYKSTKYSTTFFFTNSMLMVTLPPTMAIITGESWKQLPY
PWVITQTNDWLYYSSLVLQVVATALCHGVGAAGFSLLTTMKPLAAAFDKVILGINRIEERAARKMSEEGITYQESMLS
CLKESIAHHQEIVDELLMEKPHLEIMFFAQVTFISIVMACEAYPIIMGIVDVSGLIRGVLF LFIQVMCCGFLNLEFDT
IANKNVEVSEALYGTWPYALGVEYRHVVLSMTFSQNP I WICGMGFFGLRASRATFYSAMVSACNMLNMFRKFA

>AQM56103.1_OR_Alucorum
MVKLFDLAEQEDELMEGVYEKLYGPALQLSLLFPSWKRENLYKTFGII LVYTVTLLVHFYVLSVSVVMLRDDFEAAC
LAFHYWLIFVMVFLSLALINMDRRTFSFAHRCLARDMGNYAAGRIYSESKPLALEARKKKELFRFLVLPGMVVILAAA
LLVVPYLKKINNPPHYNAYGVNMNLPLATHYPFPTDHGILHG VVVLGQLSAAFSLAVIVVSLELLLLFRVSQAIIFEFK
ILQYALETLFERSERLFFQLHPDYYGKLSHMNSNYQKCITRCIQDCVKHHYKIQELLQAYEYVLK WPAALGYGIGTGV

IGLGLVTLLMAKEKGNLENVVIFSLILVAEVLNMYIVSVFGEDITTESAAVRDELYFIEWYKLNIPNRRMMLNFQVGI
TNPVIVKAGGLVALCMDTFSSIMNTSYSFFNLNMNANPLDGNK

>AQM56094.1_OR_Alucorum

MSFLMKFIDSLAENEDDALWDQLHKFYGPILFIAFIFPSWRRSKLPISLAVMSFYAFIFPVHIWLLTVGIIQCVRRDDFN
LASLEFHYWLIIFMFSLVALLMMNGNRNFMISFHRTLTSDVGKYRAGRIYDEKRPLEWEHNKKKQLLKFLSLPTLVVLV
AGFSLLVPLYLQKMDGTVEYNERGANMKLP I PAWYPFPTHEGILSLLAVLGQFMAAGGLATTVATLDIVVFRLTQSLLF
EYQVLRYALETLM PRAKRLYTLKYP AEDMRKLRTNDEAFQRCIGKCLEDCVIHHQDI IKLIKDYKTLVKWPGFMAYGF
GTGVIGLSLVNMLS AKEQGRYEDIVLFFLLALAEVLNMFMLSTFGESITTESKELREQLYFIDWHLLNTSNRKLVLNF
QIGVTHPVIIKVGGLVNVSLDTFSSIMNTSYSFFNLMAQ

>AQM56090.1_OR_Alucorum

MYSPWMANNCIMDTQEYKTARKKDYWYLFEMSGMILDWRPGFYIINLAYVVMVINAVYTCACLIASIFKVDDLIDV
CQYLNLFVGLLFVSVLASLNFQRERTLETCAIGSTEFFDYGVFSRSEEIEQYRKEGRRMKILFMVIPPWLTIIAL
SLMMSGPIDAAFSYPKINATYVNGIYQMAPLKMYLHPIDNEFIRWLTVLSQAACSGNTALVIGCADLIMFNAGQNV
IQLEILNLAVLDTDKRASKLYELKFGREPPSDPVDKTNDRPLMNIYGFCLKQIVDHHKII LRRAEVYHKIVNWP
CGIVV VNGSIVVAMSLLSIMQGGKPSVLVLSFFLI IAEVASIFMVCEIGQSITSQCERLFD SMYAFKWMDCSVEV
NRAINI MKCRFRKPIIMTAGSLTPINRDTFGTMMNTAYSYINLVAASGKEDAD

>AQM56091.1_OR_Alucorum

MSKAKKSAVGTSSYKELEGLERA EALKKGYDENG GAYIRLGGQYVVS RSEIWRPVLFYADTALAMFELVTASYFSVLQ
GDMEAASECFHF IQMIFNMVISANLQYYRKNIDELFTAIGAGFFDYGDTIDPQTKEKMDKHMMDMRANKKFRKVFV
LVVWVLGGCMFIKVIVAHFRFGDTIDGEGGSVSRKHIVAIWLP GIDEWPSYIAMVVAAYLCQVLIMNSIWGFVLPVIC
FAEELNAQLHIVGIGLRHTTARARHILYRKYGEQKSGNLKFKEYEESLREALKSSVQHNNVILEACKCASTLLNPLMT
VMFGTAVLLCMSGFMVEDSVPIIANIISLLFIGGEVIYAYLFCYYGEMITATSLEIGDELYNDDWWEGRDVFRPYMA
MISLRSNRPIKLSAGGF TDVNNAAFSNIISTSYSYFNLMTSKS

>AQM56092.1_OR_Alucorum

MWLREWFKNRSSKTASKEKQVVPVNI DGTPDFASFATCLSYQKFVGLYLDG SVINY LKISIPILFLTTCISFAMADI
LNYKNKNI IWLIENGHWCI VYIAAIFWDTQMGIKSPLLLRMSRSVKSGVYKYAKYDTINREDLEKTNSQVVTTSRFCV
FVYVAVLATLVKPTILEEYDPYRHFFNGWF PFEVNSLWRVSI VRVYELGCAWSAASGVCTFFVTFMAYSYHIEAHLK
LLIQKIEKVDFDPESSEYDI PQLDKKIRECLGHHREILRVFNDFSEFCDPTIGCATLMATFMVCTLLYLMTNPDFDVS
VIVTFSGVVAPEFSLLSIFRLRGQRITDLSNKINEAIYKLKWLDDQDVKVQKNVLMWLRLTSKPLELKSFGYRNVSN
SGIKEVLQTSYTFFNMLKAST

>AQM56093.1_OR_Alucorum

MWKCKWKSSAPPPSLKDEAWKWPHESSMMNWFGWAAEELERPLVVKLLQVMRAILIPSHLIFYGSLLYQTSNEFRQGTI
ISTVKSFAVSGPSTVACFKLYVIVRHRKSLKEITNSMDVMMKGILSRHIPEDLEKEMRSRWTGCRKLYKCCVYFGCSV
TTHASVTPLLQTIAGALLTDDPLPFD SWPYFLMGYYFWALNTFCIGHVLYMF DATWFAMADNLQIHFAVLKNYLENLD
LTKRSDVDLNLCLKNHMEILRLCRIFRRI SRTVIVTTRMCSMLLLCAGTFVLTSAGDEFTSNDRGNLLSTLIYIAAVF
FNYCRCADNIAHQ LDELTTDCYSAQWVNAEKSQKTSILNMMTITRMEPKFCGIASIDLDTFVNVMRGVYSYNNFLTAV
DVGDESETSRTEVNEPL

>XP_022204744.1_OR43a-like_Nlугens

MNIRTCADNSQFLICCMQVTIMLPSCVDAYSILKATELRKAAVSVIMFFFTNLILFTIYNNGQQIFNENDRLRKSLL
EVPWTDKPRWLRQSMHIMLTQANRDVQIKPYGIFVLNYMSFADLMKFTFSVGNVLYTRKQLTSQT

>XP_022200228.1_OR83a-like_Nlугens

MISNQNFQFDTNAYLGLQQSMLRFLSLWPSATEGMLFKLARSFLLIIRVLLMLIASVGQIVQIYYMPDLGKISNIMDIV
TLTTS GSYKAMFVFVRQRDYQKLVDVMQTKFMEKPFFGETKPRKTMRFYLRHTELLSNMYVYSSLIAITLWGLTPLLN
NFIEDEETMNSCVQTSSTLNNTSCNTLYKDLKLPINLWLPHYDISWFPYQILFVVVVFVPLIDSTLFYIMEDAFFFTML
YSVCGQFDLLKESLENMSQHIVKNGRQSDRNSVVNGSHDNYGSNDYTAIDKTQYDEVELHNILKKCIEHHTVLLKFLR
DYEKLSQTMII VDFLHAIISLSFALLEITVSQNIIEYTKMLLFLVICIWHQFLNNHFGELIIQKQMSIADAAYDLPWV
ERSPQFR TSMKIMIMRSQQPVILNGLKMYFLCYKTFMEFTKALISYFMVLR TLQEGK

>XP_022200348.1_OR56a-like_Nlугens

MRRNRLEFENDLNEVEEMSLKGESDLKRS LKYLIMQHQSICNDLKIVDKGLQNFVLACNNSFGIQCISYCTREMDG
FFMKLEYSLLLIWIVW TME MYSALGQQVTDQGENLRSLSSCNWLES PQWCQKSLIMMTKATKTMEIKPYGLSALNL
RSFTHVINAAYS YFNLLQNLK

>XP_022199828.1_OR23a-like_Nlугens

MMDFIYTKLHGDKVDRNQYNSFLKIMKLLLI VSGLRKQSNIKELNLIVMHATMTIASLGKNICMLCALLFHEFVYFFY
GQQLVNKTADIHHDVYFCKWYKKT AQFKSSTSIIQYRTKQPFII MAGKII PVTLASYVACLQA AFSYFNLLKALSHNQ

EG

>XP_022193747.1_OR23a-like_Nlugens

MTIASLGKNICMLCALLFHEFVYFFYGQQLVNKTADIHHDVYFCKWYKKTAAQFKSSTSIIQYRTKQPFIIIMAGKIIIPV
TLASYVACLQAAFSYFNLLKALSHNQEG

>XP_022192649.1_OROr2-like_Nlugens

MQPPHRGTKLALIFTLLAVYFGIEACISLVKNWHQLHSRLLAFEDIVIAIVMTMCCIEFAYIPNHIQCLLNILNDNYF
DGYNEKESEDKRSIITKAVEAELITAKWIEILSYVTAVNVNTALPIFVLLKLLMGESFFENYANLPLPLEYSFTFFDC
TLMTYLVIIYAVEAIWYMSAPAFCLCVYMATVLMRRVRTELDLLVLSLNELDKFVGKGADSIEQDQPRMEQKLKSHL
MSFVRHHYLIQVRNVELNMGLKLTPTMTLQAFCTLISFCLWSIVKGDEIILKIKYFIFLINLLIYFLYADAGQDIIN
GQEQLAITLTFQCEWLDKPKWFKNSIATMMLFVNRPFYVKQFGIYKLDIAYFSRILNAGYSYFNLLNSVH

>XP_022185545.1_OR46a_Nlugens

MLVIAAITIVITKVEMSMETSKFYFMMFIELVNVFIYCLLGLTITEEAKRVGETIYCTPWTASANKNRGIVKFIQLRV
THDVVLKGNGLFSLSLQLYIQIMQTSYSIFNLFNSVDT

>XP_022207573.1_OR7a-like_Nlugens

MIFHRLFKLYQDLTSTIYGVIIGLSMIVLATISLVLTTRTKLFSFETVKFTFLFLELTMVFGYCFMGTSLSSEASSKVS
SALYDCPWFNLPANQKKNLIIFQICTKKKIVLKGNGLFEINLQLFVQIVRTTYSIFNIFSSTER

>APZ81508.1_OR86_Alineolatus

MKTVLVHEDIAKFIHFMAKARTMWYGTTPWDDSLFDRFRKSYTDLKWLFCLLFFFMCGFSLSFADNFGFLDGTFIYWP
ICFMMTLLTSIAKFTARKQDVLMTSLNDHFLRNTEPWMRSIKDDYIKSLWKFIHFFSSYQLLVSILYMVVPFVADLIL
HYGFDYLESPI SMPTPLSPIIKYNNAWNPKHFAVTVVNLWAFVEVVLIVQWFIANFSLTTVFVLTETLIIFKHQVKSIN
FEDEDNWEQQVKNIVDKHNQMIRLCKDLKDYGLGLSSALVCFFTSVLVTFTTFTMYASSDITLRLSYGCGFSLYFSSAL
LNSYLGQKLENESDEVFKALYGLRWYRFKPEARKSLNMMMRQARDPLIIDFHGRYKMNLNSNIS

>APZ81507.1_OR85_Alineolatus

MEEENKARKNKGDNLCVDYRKYTFARLIMIDDGLAERGLTLPLLTIMVINVCMESCSFISIFLSTQASVSLNNVRSF
LLGSSVTLSLFNQYISRKEIARLHDFFEKSFERSFRTDFPEERTILDDARKAATAQLDLVVKIIFACNTVAMIFAQPLGG
LLSGNSWKKLPVDWTFPPSDNEFSFWLIFLFQVTGVCIAHCVGIMIMSFTSITIQTALFDVLVFCIENVEKRATTRS
QQTGRPYHPSLLACLKDDVAHYQQLIRELSSATPYLRNTVLIISAAVPMVMACEAYPLMQGNIILGDLVKSFLFLSIQ
FLCWAQTCRLETMTDRHDAVFNALYDSPWFDTGFEYRKLVFNTMTFTSTHPKYIKARLSNEVTATMATFYSFVMSSFN
LLNMIRNIG

>APZ81506.1_OR84_Alineolatus

MTKKIYEKPERIYWDLSLGFSQLSVFLDLNPPIDSSKFLWHLKRKTAQFFTFIFSPIFISVHIWSILSASPYNLKQISF
DLGFLSHNVQNLKMFFWLTNLTSVRDLCLDFSQFHVNRYPRLSSWVLQKESNVTRKYIERCFLISKGLIIWVFIP
TSIALINIFYLVSGLSAGQETIIPRLSVTQFPFNMSYLGRLFVGTLEYGQMISVFIYYQPIDMFLIAAVNMVTRQYL
ILNFGLEFWMPKDLEAFWGTVEVSVKDTEPPMDMRLFVEDHQRLVRYGQKLREVLNPVLGIVTFDCIILMCNCAIFITKL
LEDQLELLELVLYMTAYITIIILASLVVYFTFSSMSGMLKEAESVFDALYAQDWYRMSTQKKRSMIFIRKQTTITARKI
PMLNLGDMDRATFIEGLRAVYTFYHFAKQFK

>APZ81505.1_OR83_Alineolatus

MSKQSGPLDNIVDYRKHLSYTMFTTESSQGLSKPKVIGTAIMMFIIPLGCFVSVVKSEELKESLENFKGVMAEVMMI
SVLLNETIFGPEREERLFEYLSKCHMTTRYGLAEHTIIRNAGDSARKETFVYALLFIVNCPLMLLTTPVMAAIIAGQS
WKQLPLPWTLPEDDDVVLNAVLLLQIVGLLISHVVAIVMMSVLAITTQLNAQFDVVILGLRRIEDLAAQKATEYGLSH
SDSMLSCIKESVAHHQELIRELLLVKPHLETEFFCQIMTISIIMACEVYPLIKKDLELSDAVRGIWFLIVQVLCTALI
CNKMEVMSNKNVEVCDALYNTCWYDCDIRYRRVVLNAITFSQNPIPIRGKGLGMKACRETFYSAMVSTYNMLNMLRN
AQ

>APZ81504.1_OR82_Alineolatus

MNWFGWAAEEMDRPLVVRLQLIMRAMLIPSHMIFYATLLYQTSHEFREGTIIISTVKSAFVSGPSTVACFKLYVIVRHR
KSLKEITNSMDVMMKGILSRHIPADLEEEMKSRWGSCKRKYKVCVYFGCSVTTHASVTPLLQTIAGALLTDDPLPFDS
WPYFLLGYCWAALNTFCIGHVLYMFDATWFAMADNLQIHVVVLKNYLENLDLTKQGDVDLNLCLKNHIELIRLCRIFR
RISRTVIVTTRMCSMLLLCAGTFVLTSAGSEFTPNDRGNLLSTLIYIAAVFFNYCRCADNIAHQDELTTDCYSAKWV
HADKSQKNSILNMMTITRMEPKFCGIASIDLDTFVNVMRGVYSYNNFLTAVDVGDESSEPNADSNTPL

>APZ81503.1_OR81_Alineolatus

MDRPSDPTSRNGPKFAFDRLEKWQRSYGKLLLYGMSAKRFMKSSSQRAVVYLVVMTLGLYSLHELIAALFFSRSL
EGISHAYIMTYILTFLIQWYMLQHVRSFHENEIYLEKFESTQAHSFADHILDRNMKDFLKYIMFCAGWWATNNFTH
LIGPLIEFSLAFIRTFEFIGISLLPQVFDLPKWGVQVAMYIHNAVMTFSFFMYCGANSLILGTRVLKVKTQCDILNEAL
RNDHGVESNIKAFIKDHIVILRAAKMLNGELADLNLVIFTGSYMEIATQMFTLTLFEPGSGVYFFAVAFDFGSIFLITA

TQCWLSSTLTIALDTVSDGVYDTEWYGRDKNHALNVLIMLQMAQQEKSHRIWFNSFKIDRSAALSLLQSSYAVYTFLM
IVQSN

>APZ81502.1_OR80_Alineolatus

MSVEKNLIKSKAFKDNFGQYMWVWSGIYCGNNVYPGCVLIFYIVHTFFLLYTLILSISKDNEKLMGETAHFTAFRFSAI
LLLLINGLWKKEDEALFLKLCQPQVHDYGNLSDQCRKDIEGARQNCARKDFYGNFMRTVTVALVIFWIRSLMEYF
NGHLDNPKSDDGINSNLPVPTYLPYESHEWPGYHFALVCEVLLVMMSYFLVLGHDCSFICFAEEILRELDILIIITL
TEVERRIDHVRKSSNYKISQEESVRLCLKHSVMHHQKVIHIFEHFQHYCFHSLFFMLSGGAFLICLSSLMFTSDTISLRD
KSVFLMFLGNELFHIFIFCYGGEHIMGKSDDVGNSLYNSSWVGISKHVKPTFMLLNLRCQVPLTLSAGGFMTASFDTY
GNVLRRTAYSYNLLQATN

>APZ81501.1_OR79_Alineolatus

MSGIGRIGDDEIVNGLDIWYLKCSGLWDVFNDYRDSGARNKLFRIWMVITVMFFAPLAFLSTFGPFFVEADLEGLTLI
ILNPMSCSQTVIKFAVLWYGIETQCKLLDLFKNDFLTCTVPPDKKLKASRILTASAKKANILAYLGIFMDAATVAVWNI
LPILRSEFFRVQLGITAFGTPMKHNKILGFWYPVDYDVAPYVQFVYCYEFFTCFWAGFIIALLEGLIVQLILLLTANI
RVLQYLLLEEIKASNSNLNSETLLLYAKEYQKLLVVGDDMRHLYNSLITMQLSTGLIILIIITIFNFFLSSGNGDIVIMF
KFVIYLMYTLVEVALYCYVGSDELTTGDEIGFATYCSSEWYKVGKFRKTLQMMMVRSSRYSMIAIKFGRLYPVNLMALTN
ILQMAYSTSMILLYRITNKEEENRV

>APZ81500.1_OR78_Alineolatus

MSRSGRLKEEDILDGLSLRYLKVSRLWGIILNHHRETGGRSKVLKAWAIGMVI AFTPLLMVTVVGP I FTEKDLEGMTVI
FLNPIATVQMIFKFGIFWFHMEQQTKLLDLMRKDYLSVCPADQKA AVSEIMINAVKDANIY TFFGSRMNIITVSLWNI
LPVLRSEFFRLTLGISIFGTPVDHNKILGFWYPDFDGTGPGNEIVYVYEFILCIWTGFIITLLECLIAQLVVLTLGHL
KAFQYLMENFRSSHDPMDHRKLLSYVKEHQKLSRAGESICELYNFLITVQLATGLVILILGAFNFILGSGKGDDVVM
IKFVVYTYMYTLVEVTVCNAGQNLETTSEDVCF AAYS CDWYEMGVDFRKT LQMMMVRSSQSP IVMKAGKLYPVNLTALN
NILHMAYSTSMILLYRIKVPEEKQVM

>APZ81499.1_OR77_Alineolatus

MDTLIRKNNIDKYL PWHPDLTKVWRWPIARYLNTFGWWS EDYTRPSVKFLFRVIKFTFLAVNVVYLISLTIGVHRQL
QSTTNFMDNMFAFFAATPSYLGAAKIFALIIQRRELRLRIWKNLDDLLKDV LKRGVDDKLDREL RWRMKRCWLMYSIFL
TVGSCITLHWLLRPLVYAMYGERTSIVSTWPSYLENWPLWFATYVFQAMNISSVGHALYIYDNVYFCICENILIHFAI
VKHHLHEMDISTGKPGGITMNF CISHHIRLTSVCMDLRECSKYVIMQQVFWTIFIICPAVFELISGRQTDTTIVVNLL
EITTIMTCILFLYSWYSNEVTLQSSQIYNTCYMSDWVQGTASQRKTLMTMMTRSMKPIIFGGLVNVDLGTFISVLKTT
FSYYQFLDTMDKNKRKLDDTS

>APZ81496.1_OR74_Alineolatus

MGKLEISWPQLNYLA IYGLWPVSVGGIRFRKVQWLWGRLIVTLTVMVIVLEIVNLHELLMANEMDLNDFVITMNDILY
ATQSICILVNILVYKKQFAIMMDELGEMLDHSQLLMKSYAEHVFTKKLRVCRFAFIAFLSGTLLMAHWGISPLLNRIL
YHRNTFVYHVWIPFDVEPWYIYF SVLTAQITIGVSWIMGQPMFASLFISISEHLLGHFDVLRQGLETLDYSAPSASQE
VNRYFEYHQKILKVGYILRSATKLVFLSQFLCVTSIMCLNLYEITYVDVETS RFLNMVEYTLLELLIIALYCSYCNE
LTLRGSEIMSSAYFSGWESASIQDRKSLWIFMTSTKAPLNYGGMVKMDWTT FVNILKTAFSFYNFLGAVRTSKEGQITN
G

>APZ81497.1_OR75_Alineolatus

MGSLLDKFDEWAEVEDEDFWKFIWKSYGPLIQISCMFPSYRRSMLPITLLTFAVHLLLLVPHFLLL VITITYQTIVDWD
LELCSLAIHFSLMIFYAIFTLFYVQYIRYAYSSQAKAMSTDIGNYISGRIYEDKWCVRAKEENWQETLRLLIGPAFLM
LGSGAVLILPYTLKTIKGLDNPYGAGMVNANLPIPAWYFPPTHEGAVHWIAVLYQLVAALSVGVIMMQVLLLFLSNAQ
RLRFELSVVGYSMTSILKRSMKLHLRANPGLNRDEVDMWDPKFQSVVESCLRDSLLHHQTILELLSLFTTQTDTLVLL
GYSVGTASIGMSLFNILRALNTQSYESIMLFTIMIMGETLVMIYILSYIGESITSQTSDLRNQLYYSPWNNFDLRNRKI
FLNFHTAITEPIVITAGGLIPISLDTFSSIMNTAYSCFNLLNTQPEI

>APZ81498.1_OR76_Alineolatus

MVNMQRVREVEDVKTRSGLSKLIGPCGGYRGPLYNEYYSKNIFFRAYVHFTDLAVIINYITMIAAAIHSKSVLEFVMV
GFPISAESLSLFLSYSGYKNKEMTNVLLGFDDCFDDDPYEPHLESEINKAAKYYYHFSRTLLWLQVFTMEIYCFIFP
VTNELMQDYFRPRALPLPSLYPCDWKESRSCFIMIIFIHFLGATYVNWKIIAFGEVFFAMVSRQVALFRHLNHNLNKI
LTAIQVSENGVILYRLNKEADHMHKIKALRKWIKHHQSVMQAQYDRLQALYSWPLFVHFGLVSGALCCSAYATSDETLD
FDANLLCGGFLMGQMLELFYLCRIGDWITTIETNELTLALTGSFTSVLDRIESQMLRIILSRVHRPSVMRAVGLYPLNT
ATLKMLIQSTYSYYTMLKKVNKG

>APZ81495.1_OR73_Alineolatus

MAISDHQIEWLTQEEVLVFDNLTKLLAIVGTTSRSKKWYRMFSLIFWLAAYIPMIVGLTVTAVHYIKDLDFLAYTLHH
LVLITVAVFVTHLVVPLFDTQILYVMNASKNTYNYESDFVENYAKKII RKRIKLSRTITKMIYFGMVWIMLEVQVFFM

IETFFLKSYQTVFPLAIGLDLNRWLVFVPVVLWQELIVYYTTFLPTTLAVLCYTAWSHLDLEMRILTYAIANVQKIVN
EKLSQDGGPGSVSEIYVSYCNHFAXHHADIISEYRLAVSWITPLVFFIGAILFITVGLSLMSDNLGIKLLFWFLVVQ
YIVIIYFWCLISQKIGEKSEEISEVVVSTPWWMMPRSCQSTLLIMTRCKKPLVLSTPLGADVNVESFMEMMKSVMYQAI
SVVYQMKTTSS

>APZ81494.1_OR72_Alineolatus

MKEFKSKTHALLLTMLEVAIIHRNPERSMFSPNGFLIYRVGLCLSYITMITSSLLYLNEEGGNQLMKMVLSSVALQFF
VCSMVLISKKESITKMLLASTEVFSQLTPNDENRIKMDLLDKKSLKFAKFFAFLISMNVISGVMKGLMNSWLSGETHY
PFDLSCLGLSAVPCWFTQLFSTSGTIIYIYCYFAILKLILYQLWGYVDVLSALIRDRPVSASEKDDRDLLKFYCDYGN
FAVSFTSIFGITAFIDMIFTSVRCGLISYYIVMSISENNWSDAAGSMVGLASSFILPYIVCSGEDMDEMNNQIRNGF
LESNWYQCSPQSRKRLLPILTLNNVPIKFQYRQCMHFNMERFMQVLRSSSYSLTTALANFM

>APZ81493.1_OR71_Alineolatus

MEADDGTVNEVLNLSGLINWIGIVGGIRFAQRTFRKKNFQSTIYWFYEIFTDLTVFINILSQIFALMASEKMTERCLI
GFPLVSCSFCFFISNYPRFKRQEFATLVLRYYDDVFPNTQYSDHLEEEIKKAAERVRLTSTILVFLELGPMTFTCLILP
LVNEASGFALGPRKVAIPSLWPWDPLASFKNYIILVFIHAWASIFVNLKKIGFEESFFIFASRQVALLRHLRYNLEKL
FDPLEVNVDGTVNKNEFTSPQRYWMEKLTQWVKDHQHLRLFYELEALYKWPICIFYGATILILCTSTFVTSDNSID
SQTCIICGVFTTGIFCELFFICRMGDHIQIETEELLGLSGKNAFLSNWKEFKYLRMIMTRCQKESVIRAAGGFPLTI
STFKSITTSSSYTLLKEVNGKLTE

>APZ81492.1_OR70_Alineolatus

MKNAEEGQLTINDCILHYINQSVKHHQHTLRMMNAFKKLMFFPLFALVFDGGFILCMSSYMLINNDLSITLRVPMPCV
IAAEATFAFIFCYYGGLLSEANEDVGNGIYDCEKWMYYSKVTSPYAI I IKGVCNIPNELSAGGFTSVNQRAFGSILSA
TYSYIGFMLST

>APZ81491.1_OR69_Alineolatus

MKEEKEDNSSGCTDYRKYIFCRLMCIDDRLVDRGLTWPLLLIMAQTVAMETCSFVSIFKSTEIKNGMDCVRSFLLGNL
VTMVLNFNEFTHRRRLARLHEFLEKSMSTLRGTGLPEEEEEILKNARDQASSNLKMYMMVFFGNIAPMVFAQPLGEWLSGH
SWKRLPIPWSPPSDTDWAFIFIFLQFVGVVMANCLGMVFMSFSSITIQVTALFDVLLLSLRHMEMRAKLKGQIEYL
DYHQSLYSCLRDDVIFYEQLVRELESITPHLRNTFLAFSATVPMIMACEAYPIMLGNFAIADLIKSFVFLAIQFMCWA
QTCMKLETMTDQHVAVFRTLYDTPWYEAGVRFKKLI FMSLTYASPPKYIKARMSNEITATTATFYSFVVSFNLNLI
RKMS

>APZ81490.1_OR68_Alineolatus

MSFVSNILSSWAKDEDELWEVYHKLYGPALELSFTFSPWRRSRLPLTLLIFGYAIIFPFHIALLSVSDICVRDDFN
MASMEFHFWMI FMFSSISILLMNSNRQSMMLHRTLTIGVGKYKAGRIYNEDYPVMLERKKPKQTIKFLFLPGLVMVL
AGLTLPIPVKNMDGTVEYNHRGANMKLP I AACYFPFTHGITYLLAVAGQFMAAGGLALVIATLDLLFRMTQSVIF
EYEVLIYALQTMSSRAKKLYRITYPNSNIKEVRTKDEAFQRCIGECIRDCIVHHDI I KIVKEYMNLVKWPGLLAYGF
GTGVIGLSLVNILSAKEAGNYENIILFMLLMIAEVLNMFMISSFGEAITTESKVLREQLYFIDWPKLDTQNRKMMLNF
QVGINNPVILKVGGIVNVTLDTFSSIMNTSYSFFNLVNAQ

>APZ81489.1_OR67_Alineolatus

MGYREYPKRDLSDPSHLFNFHLSALRVVTMWKLNMDKYHTPFMILFFWNVTVLSICTVGLFIKGCTTDDLVDREAMD
IFTLTGSAMYKMYFIYYHDDLVDMITCGLALTEKLPKGWTKHCTLFKAFHCCGGIFCMTFWGLVPMFKTLLGETTLE
EMKLPINTYYPFEGRFAFSVTYSIAQYALMVSGQIYMAADIYLFSTSIYVAVGALQYISDELEKMNENSNLKIADVGKV
DQTHEHLKECMELHVNVDYIRKTDKLFRRSMILADVHAIISLSFAMLQASEAKGLFEGLKMAVFWVVCFLHQFLNSH
FGQTLIDKQDHLIEKIRIAVPWREASRSFKKSYHIMITSNTNAIKLSAWSAYYLQYATFLEFSKSMISYYMVLRELQD
QEEVP

>APZ81481.1_OR59_Alineolatus

MDDPGGRALEKSGLNNVMSILAGFQGPREVRFKGTKYQHVFTAYSIFGLCVSHYIILICCYLTPLFITDMSFKDMMFFA
VPTITSTFSTLRVYYMMWNRSKIIRLLIKNEEASEDDYYQDKQEEEEIKKWANTAKVLQPILYFAVCAPIAPWGVSPII
NEILGNPWGPRKATIIISWYPYDVQKTHFWVLTVCLQTMAGCYATLSNVMFDAVFICIAARQLALLIHLKNSFSKIFQV
FHVHLNGNVWYSNYNGEVVERKEIEDVVIQRLKYWIKKHQTVLRLLDESQSMYSFPLFLHFCYVGMVLATGSAAVLKG
TLTSMEYYFIGMHLVGLSFTLFIICRIGDYIKIETDEITEGLYGQNYFMLSKEQHMLIKNILTAVKQPFVFTVAQAFP
LSTETFKSVMTTYSFFAMFTQM QHKN

>APZ81480.1_OR58_Alineolatus

MWDWIQLRMFLNWGWPKVIKDPKRVRMMRIYGYCMFGLDSATMIAEFVSLYLAVVNGSFRTGAIVNIVTTTLGATAAM
KIYTLVHHEFISHICNTLEDLNNRAIALMGEESQVTMGTRKRCKLTFVFGSCLFTVCHYNVRPILVYTYGERTI
AMDMWTPWDEQTSLSGWVIVLIYEWIHIAAAMYGMVVFDSFVLCIFEMVLAEFVDVLKIALKKVDFAAEKNAVPIEFICI
QFHQDLLVLIKINDFLIPIQTFQCIMLTLTICFSGFELVSLSDVSMNKAANLLEVLGASTYITFGYCYQCHCITEEC

VEVITTACDNNWFEGSIKDQKSLSILLERAKNPISFGNIIKFDLGCFAIIKTAFSYYQVLEAFDIH

>APZ81478.1_OR56_Alineolatus

MLCCRKKKEKQOPVVQEQNDFVNNYRLFYIGMIQDGSIFSRI RVPIATFLLFYAWFHHLIPLIMNYDEYSFDELMDL
IHLEMVYFLWCIVWPSYIIRSPLLTSLTSKIKSGLYTYS DPLTLEEK TILSTANDAVARITKISVYVYVCGGIGIFLK
GMNKEKMRKLSLPNIGWFFPAINSLSRYAIGCLCQAIMGVNAVSI AIGTFMSFAMFMIHYEAQFKLLRTHLRRSF PKN
IPLRIAQTDKYKEATLRKLKDCYNHHLAILEYHEELLKYYGILLLVFRVAIVLWLCTLAYVTIVDINAHNLLKMMSF
ASTELLFVFLFSFRGQDVTWNYEWREELYSIEWWEQPKVQTNIEVMILGTTEPLLLYGWKIALYSHEGLFAIGNE
SFSFFNMLRAMN

>APZ81479.1_OR57_Alineolatus

MFARP NMEFVQTWHYYGYWNI FLGGKYDSYKKVALVVR LILFTAGAF TALAKISSDGVESVLNGALIYVPLCVYTFVI
PTILFFKRKNLFSLLKVFDECWASLNDEEDRAVFEDYYQKTWK FVRIIHVGLFCQLTYITFPILYSLVFHYIFDVLS
QPYILSYPHMSYLDKNFTWNGEYYAVVFAGVYCIAEIEFALLGWVLMYVVIVGYCYPVLVVTRNAVQAVG SNRSPSSG
DRAMKKAVVAHSLLVKVNQDLKAFLGLPCAFQSIF TSLCLTSLSVFTTIRSSDYFVEGAYGSAIILYSLSCLLYCSLGQ
ILENQSEKLFEDLYNLPWYQFSPLVRKDLNMMLRQAGKPF TIDFHGGYNMNL ENYMEILKSAYSFFTLETLASNT

>APZ81477.1_OR55_Alineolatus

MIRNWQDEEEGQSLLARMGCKFLNGHSIYVGSWVLRFPVR LPLFLYVTCAVGISIKMVLNYDNLVLIIDCAHMMIHMV
VGIQTTLIC TQKGRIMKLKDQLDKFDDVPGEASSTR IKEDYELQVMNLYRTFSRCILVTINIYILFPICKLFT EAG
RAKLSKVLVWQMWLVWPEETWWGFTIIFLFELV TSLFLLLSVMYAVPYLACLGTITVAHCKVLILRLKSLQKRAEERS
GKMSESYEAALNYEIDGCARRLHENLR IANEVADV KYLSSFFYGGMFALCMSGLQAVANENIEESLKFMGVLTGE
LVAIGLATYVSEGIIEGFADVRASIYDLPWYVYPKL SCKRLHLMSTMSSFRGLRTMFGYELTLLHFGDVLNASYKYYN
LLLLSMK

>APZ81474.1_OR52_Alineolatus

MMKNKRESSKPTQKGTMR LGWF SFHFS PQKTSKKKPANTAYHDGVPDFANSVTCFTNVRYSLLYLDGSLVNYLKICI
PIFFLATGTISFAMADIVHNEKNTI WIVENGHWCI IYTAVLFWDAHMGWFSPIILEMNSV KYGVYRYEKEYEANNRM
EFGKNNAWISKMNRVFGVIYIYAVIGTLVKSTILEEQY PFKHLFNGWFPFEINSFMRLNIVRIYELGCAWSAASGAQT
FFMTTMAYTYHVEAHLRLLMQKVAKVLDSQHPEKQIQECLAHHRAILRLFN NLSAFFDPTVG FSTLTATFMVCTLLYL
ITNPDFDMNVI VTF SFLVAPELGLLVSVRIRGQKLT DLSSEVNKTIYDLNWLEQDIK LQKDLLMWLR LRTSKPLELKAF
GYRNVSHAGVKEVLQTSYTF FNMLKAST

>APZ81475.1_OR53_Alineolatus

MAEGEAYFDYMVQILKNYNLWYGYSTASIQGIILTMYSVLQIFLMLGLLSAAISQTYFYGMSYIVHDSAVFLPIGFMG
LVIIIMALNYS AIMKTASKFELFLNSFTVDWEEELIKNHMKDTRTVV TLMVNVISFYVLSTGAMHFNSISLHYFFGI
FSKPVLLPLVLEKFMEGNFQLRPMFFVHVLLSLLYIKIAAS IITVIALNIHFSGSAVGALQVLIKRLENVSKKTADGC
HKLAADSDLRDTIQKHVELLNIVSDMMEWNGFIVSFTLTSCSILFCINAITVKKS IETKEYSSTCVWGSFLLVALAIG
GTMSRGRQEIEKMSEELLRAMYNLPWYRESSKSRNIVLMIAQGNRLISLDYKGLMRVMNMTYSEM VQKAYSFYMILG
SVE

>APZ81476.1_OR54_Alineolatus

MNFSKMMLPLFAKKTRTCGNNRKS VNKKWDAIFKLGLRMLYL GASPDQVWRGERNVLRFGFFPLLILLYSGFGLLIV
VFPFIFESDFMESLENLYLGTYFVAAGFQWAYLQIHL DGIYEHRTKIESFNSTQAI PDVANSIARRNIKNFIKLFWVL
VILWNVGIGLFTIKITMKIITLLTGESDSSKVDSDS FQNSQSWRLTVDFMNAVMTLLSLPISLSLIIISRFLTLEVVT
QCEIIRAAMALDTGKKVMFKGYILDHIRVIKNAKWL IKRLEAVNTANTLLSYLMFALNMFM LTLVEPALYTYGVLALY
FFIDFLQLAAQGWFS TSVKIALESLS DGSYETI WYERDRSNALDVLIVSQMAQQECIQKVLFGSLRIERATAISIVRS
SYSFYTLLMVLQT

>APZ81467.1_OR45_Alineolatus

MICSVYLFLFGVTAIKQIDDFVTATQTVHFALIISGSILAMVTVAIKRKHFIIIVRTMSVKYFDYGD SFIIPEMEEEY
KNTKKQRIIFLTVIPMYFAISTVVLGLGRTIDGYFGSPLNETYVNDVYMLTPEPVWYPMKVDS DILYWFLALS VVSMS
YAACITVAGGDWILFVLYLSITQQFKILIYRMRRINAYAYRLHRKAGGVKL RKSMMSFNS SFLKYFNLCLSKHAEHHS
IIKQFEQLSIIMSWPAGFIFIFG SVIIAMSL LGVSVQGGGKPSILVLATFLTFSEIGEMALFCGLSESIQTLGLSLH
EELYRLNWADADISAKRTIMIMIEQSKRP NVLKAAGLQSLDWM AFSSIINTAYS VNILMAVDA

>APZ81466.1_OR44_Alineolatus

MSSNTVYIIEKSDAAKKATEYVYERLFI FTIFGGFFGFHHTRWWSMFTYTVFTLYYISVISMIVTMSYSSFLNQSNTS
VMSGCLHMTITGLVVCASITLQLARKDLFKLLVNFVLDENLCEYQSNEYFSHLLKKADEKLRSLLILWLLLYGSAAS
IADVFFPVDVYLGYELTLNNVTNVYWRGLPFTLW WPMADANSTFAWMTCFTSQGLYAMFAASLCTGCMVFYAMMCENI
FNHIKLLVFSLEHLDERATLMFKKLYPGKSPKMRDLYDECYACIVQNVKHHHRIVIFKD AVMKAANLP IAMPFFGG
ALLLGLAGINLLSHDDNRIAPKVFFSCLGMTEAGQMFL LCKYGERYQDLSEVLFNASFYTSCFR RSMKCRAMMI FRL

GVSKPMRMTAAKIILLNMETFANLVNSAYSIFNLQSVTS

>APZ81465.1_OR43_Alineolatus

MKSGGRSFNKRRAELLKAYNAIHWTFTKWFYEDNPIEDNVWRDKRLWVHFAFAFICQSNIIIFKVFLVSEENFFAT
LTSLSGLVLLVVRVTVLYQIPTFKQLYFKPEIFNCKHRPTSSLDVLIQTVKHSRKVGIWCMMLFLTDFVSWLVL
PIVPPIIIEIKGTNQTYDELIPQYPSINPVRLTWLTKEAKYAFDLFMGAVNTIPWVGFFVVVYAVVQLFLAQHKIMML
SMTRGPQVPGDAKEPLELRLWIQDHALIRKLVYQLRSTVSPALAGTICANVFTVGLNMLALISSPIGPEAPLYTRYLF
YFSFGTYSAISIFDIFIHCWLSSEISNSGKELNYAIYAGDWNSDLKRPPQDNVIPLMVCNKEIRFTALGLIPVTMTTF
TEVIRISYSYFTILKETGH

>XP_016657134.1_OR43b-like_Apisum

MAHIVDIFQNMGCSDHGYGMVFFNCCELAITLFFTVSTYPTIADPTQNLISIRLYGVLCLLIEAHIFAFIAVRIYHQ
SQHRDMYQHLHGVEIPENYRRKIATVIKHHFIIISNVFVAVSVLYTISLDWVRIGDPFTFPFIDVLPKTTNVTVYVCK
YIVYALPVYFAHLETCFLNVTFMFSVGIVKRHFQILNDQVEEAIVNEDEQKLKIAIKHHQOVLKYFEDMKTVYEKPI
MTIEFCGLYVGLTSCFVIQVIQGFHFIILGLCIVSSIACLMTIIIIYCIYASNMYALHNGILNALFEHRSCYSRNKSF
KRIILIMMTRATIPLEIKAGSVFTINLNLVKILKFAYTVFNVLLSSINRQFKETAI

>XP_029346749.1_OR33b-like_Apisum

MSIFITIFFFLHRKYDEFLSIFRPTMLLQVFVVSIIIFLWFIFLTSTFIEDDFTQYMALTSMEAFGIPFCTFQMYMS
CFVFNTLNKIDSITFALYSSNWTEMDMKFKKLILLTMRMNDAHQQKLQYTKTKIINMEIFYHVCFILPS

>XP_029346718.1_OR2a-like isoform X4_Apisum

MVHISSTVPATPRPTFIPAGLVTSQQEVINRAFKSIGYEENSQSKYYEDFKSILGDQIQNLKIKSYYSIVKPVILA
NVAMSSTSFIIIVTVLIVVSFSKESNQILTIKLGSSAIFICGQFFLYCYLLDSMNLKKESVNFALYSCDWTMMDIKF
KKLLLLTMRMNDANNFMIRASPRKVVNLQMFANVISMSYNIISVMLKSMNSNNQSTE

>XP_029347565.1_OR43b-like_Apisum

MQRIDTINMFLQMTGCTDSKAMLYLTYFEFLITFYLLIAAYASIVHFEQSVTIQLFALLCMLIECVILLNITFRLYHK
NHIREMHQYSRRLGIPDSYRSVINVTIKYHLIASNIFVVPVTYAIFCDSVRVGDPTFPFLDVLPMHTDNLAITYACK
YLVAISVYIAHVELCFINTTFIYVGVKLRLETIVQTIGEAFADNDEQKFYAI IQHQKLLSYFNTMKIVFSKPIL
LSMSFNAIYFGLTTSFVIQAIRGYINQAILSICIASAAAVINITIYTFYGSELMDLHDKILHVLFDNAFFYVSKSFK
SSILIMMTRVTIPLKFTVGYIFTINLNLKILKMSYTVLNVLLSSETIKPHKLS

>XP_029342650.1_OR46a-like_Apisum

MIGVHQLLRPDEYGDNDLYRTAAKVIVGLTLVLQSMQVCRLYLARHDITMFAYIGVMI INGLMCLLKGYMVAADQ
MSATLTAANYAFTKCGGRDPSKRLRLCRARLSAILRTFVGLSFGTLIVWLTMPWFMASDYDDQPFIFWGVVYVIESIILT
VNVFCWTSFDCYLVMTMCFVFEAQFCTMSTGYETLGRRTGAKSSAPQKLGNASTINNVKISDVNYDDLTSHIRDNQNI
IKQYDAFFDVVRPMLVIQIANGMYSIIMLIFLTLVTHLSGYSIFSAPILKFVCGLASLTIELYIICYGFNHIEDGKST
VNFGLYSSNWTEMDLKFKKTLLLAMTLNSAHKRVMKVSPNSIVNLEMFTGVMNMSYSIVSVLLK

>XP_01665750.2_OR2a-like_Apisum

MRSSSATVVDVMLFKAIGLYQLLCPADRGGSVRSRRALMTALGLSFALHSFQVPYLYALNDLQRFAYMAAVIIYGM
MCSFKGYVLVTNADRLWLVLNAADYGYTGCGHRDPSRLRRCRATLSALLRTFVALSYGTLIVWIVLPFFVDEYTGITN
SDGTVTRYRTTIHNMQYPIPLAVYNSRPVWALIYVTELYVCIVNVFIWSLFDYLVMTMCFVLNAQFHTMSAGYGTG
RRTGSSPDDTTFAGVRRIKFDEIESNHYSDLISHIQDNQNLKMFVDFEVVRPVVLVQIANGSYSVISLIFLTALMY
LMGVPVLSAAFLKFICGLISLTIELFIFCYGFNHIE TAKSVLNFGIYSSNWTEMDLTFKKTMLLTMKMNSSHKRAMKV
SPNSAVGLEMFARVMNMSYSTVSVLLNSRS

>XP_016663583.1_OR49a-like_Apisum

MTDIVEIFLQRLGCSDDGDGRGTRRAVFFTYCESAITLFFAVLTYLSLVYSSTGDNLSIRIYGLLCFVIEIYIFAYTA
VRVYHQSEHDMYQRSRQMGIPENYRRKIATVIKYHLVMSNVVLAIPMLCTILLDELRMGDPFTFPFADVMPKNTANV
TVYVCKYIILYALHTYFAHLELCFLNVTFMYSTGVVKRHFQVLDEQVEEAMVTEDEQKLKIAI IHHQOVLKFFKDMKT
YEKPLLLTIEGFGLYIGLCCCAIIQVIQGFVDQIILGLCMASCVAGFMTISLYCICASNMYDLQNGILNSLYEHRACY
SRNKSFKRLNLMIMMTRATIPLEFNVYSLFIVNLNLLVKILKLTYSVLNVLLTTINLKFKETAT

>XP_016660447.1_OR43b_Apisum

MAHIVDIFLQKMVCSDHGRGYGMVFFNYCELAITLFFTVSTYPTIADPTENLSIRLYGVLCLLIEVHIFAFIAVRIYHQ
SQHRDMYQRSQGVIEIPGNRYRRKIATVIKHYFIIISNVFVAVSVLYTILLDWVRIGDPFTFPFIDVLPKTTNVTVYVCK
YIVYALPVYFAHLETCFLNVTFMYSAGVVKRHFQILDEQVEEAIVNEDEQKLKIAIKHHQOVLKYFEDMKTVNEKSIL
VTIEFCGLYVGLTSCFVIQVMQGFHFIILGLCIVSSMACLMTIIIIYCIYASNMYDLHNGILNALFEHRSCYSRNKSF
KRLILIMMTRATIPLEFKAGSVFTINLNLVKILKFAYTVFNVLLSSINRQFKETRM

>ARJ54248.1_OR4_Apisum

MTSIGKKNQRKFYQTLMTLAFFLDTSQYRYISRFVKQFYIFDWMVLVSAAAFTILEGNYRMPFVMELIQYMIVGIFYF

TSIFVVFIIKKEAIMSNYNQCIQTKFIQWSNKRALHSNAAYKRNIKTVKSLSIPLAILSLSIALGPLISTINDIGKLPL
DNRAHFVLFWPTIVDTNKLSMYGIITYTLQVIFTIILYISVLSFNLGYMVFNLNELITQFEMLLNGINDAFKYKMDKQFQ
TLFIDCIRHHQIIKFLDDLKSYFKWMILIEIIVVQVILAILIYNLTKVNASLGKVKIAGSILFNLLPICFHCHVGE
VVLSLHTRLNHNHYNMPWYDMPNKNKQLIVIMLQRTQRDLTLSALFSSERASRSLISKVIKQVYTILNVLLKT

>AQS60754.1_OR42_Apisum

MPNSSEECVMSSMAKCTGLHYIIDPEGPTVGGHNVFHVTVMMVMIGFTVVCLSMCPFGLYYWANDVTQCFILLIYIVN
FSFGCFKAFTLVRHSDDICRCLDVTRFDFSSGAIMSDPDSARFFRKCRDASSTFTGWFAASSHFVLLVWTLPLPFVVVG
KGVEINNRRDGSTSYHFNPNMYFLVSSETYNRLHLVFLVEWAFGLCFVLIMVAFDTFMVTLCAITCQMRGIGNAY
SKLGHDRCATASNVCSDGGIESNKSNEYLRDLKLIKDHQAVLGKMNDFYKIVGPVILPQLIVASFTIIFVSFIITR
NYFNGMLLTSTQSLKMCCFPIFFYQVYYTCHAFGNLSHRKNVMNFALYSSDWTQMEIKFKKLLLLLAMQMDANKLDMK
LTDKLVINLELFTRVINMCYSIFSVLVNSQLKIADKQ

>AQS60753.1_OR39_Apisum

MFSCDFINRTVMNSENLFNGGSAFNLSTYKQLGGYQQLDPKGPPIHYGYHLYRTILKIFLLIVQFITIFGVMGFFIE
MEDTDPGKSNSFELIILTNCSLSSLKIYTLISNSKIWDLFDLTRIDFLRCSRHSKLITKNFVKRCKKSTTITKWIA
RSFLVGLILWLMGPFIANEEHTEPNTVHRHKNIINIKFPVTMKTYNYYFVYFLMEVAVGFCIVYGSVLIDAYLMSFC
WIIISAQYQSVTKAFATFGYNKQGSPPKDIYKDFKSIIDHQNIYLMKMSFYAVVRPITLIHVFAYSCLIMYAYVIVTI
FNSKELFIIAEIMKIVMTVSNVTMEVFIFCYLFELIDNKKEDVNFGLYSCNWTGMDIKFKQLLLMSMKMNANRFLK
ASPDVTINRPFANVIHTCFKIVSVLIQTQSIDLNL

>AQS60750.1_OR31_Apisum

MNPTFKHFFKGDCNTITKPSMETCIDHTCTINLNILKQCGFYQIFDPNSKKIFGWNVYRISFIALTVITQCLIGFGN
CGFLFELEDTTDNIDFLIIFSNSYFCLTEWKVVILIINRKKFLELLDVTDLIFLKSQCRKNIKILCKHRIRALQLT
NLYFKFCIFVIEWIIFFPIMINSFIAHKTENRRLENNVNRYPVDVNTYNKYIILFYVFEIIGVKTVYLVLMVDILL
LSIGWAIVIQYEVLAFAFKNIGYNENLQKDHDHVDVDDYKYFKSILFDQQQLDSKVLYFPVIVKPIVLMHVAINSVLFI
MLSNSFLMVFLSTESFTYKIVNLFKIGTGILYICLQLFLYCHLFDNINLKRKSVNLGIYSCNWTGMDLKFKKLLLLTM
QINDANYITIKASTKTIVNLPFANVLMTSYNIVSVMVKTMSKYRKT

>AQS60751.1_OR37_Apisum

MKWLQDHEVAINALFKRYQFYQIFNPNGSKLLNYDITYKLTNVMFIVAVTTYNIFSAMCFFDTIDTIDSVDLLLMIF
IYSIIISLLKISVLLFNADQIWELFDLTRFDLTSRQCRKNVGILCKYRDRSITITNLYQNYSTMVFIWMITPLVL
NTFVVVGGPNQRYHNIIFNMQYPVSANIYNQYYLFLYLMFIAMGIFVLNYSMIVDNFLISLCWVIAQYEVITTAFEKI
GNDCELTTLQNEKNNSFEAYEDLKSILMDQNKLYIKLKSFYRVVWIIIVIFLIIIDSVLLIILTYSFVMICSSAESFS
IFNILKISTAFFVFVVIQLYLYCYLFDVLDNKKESVNFGLYCCDWTGMDLRFKKLLLLLATKFNNANTLKIKSTPNKIVN
LQLFSSVMTTAFNIVTVMLKTMNGKN

>AQS60748.1_OR23_Apisum

MNLNDEQNIYIVNLKLMKITGFYHLISSRAPKYFGFNVYKVTAIEVMTGIFSIIIMLFLSSYYLNDNTNELMSHFMLVV
AIFFTSLKIFWVSRNSETIWNNDMTNINFLSYTGHHKEILKKARAKSISTTILFVILWSSVTVAWSISPFFVKDVYL
NIKFKDETRFRYNSLNYVPISEEFYNEHFLYFYVVEMLSVVFWGHGTVAYDTFVISICITIAFQLKTIASYSISLN
DKKGDIKNLKNDLEAMFNLKLLIQDQQNMFKKIKEIYKIFEPVTFVQLAAQSMILILQAYMIFINHYNGFSLLSVPI
IKLIVTVAPNIIHLFITCYLYTNINHQQDSMNALYSSDWTAMSINYYKMLLFTMRMNDAEKLLKISLRKIVNLEMF
ASVMHLTYSIISVLAKSYGNTNTK

>AQS60747.1_OR20_Apisum

MRSSSATVVDVMLFKAIGLYQLLCPADRGYSVRSRRALMTALGLSFALHSFQVPYLYALNDLQRFAYMAAVIIYGM
MCSFKGYVLVTNADRLWLVLNAADYGYTGCGHRDPSRLRRCRATLSALLRTFVALSYGTLIVWIVLPFFVDEYTGITN
SDGTVTRYRTTIHNMQYPIPLAVYNSRPVWALIYVTELYVCIVNVFIWSLFDICYLVTCFVLNAQFHTMSAGYGTGIGI
RRTGSSPPDTTFAGVRRIKFDEIESNHYSDLISHIQDNQNLKMFVDFEVRPVVLVQIANGSYSVISLIFLTALMY
LMGVPVLSAAFLKFICGLISLTIELFIFCYGFNHIETAKSVLNFGIYSSNWTGMDLTFKKTMLLTMKMNSSHKRAMKV
SPNSAVGLEMFARVMNMSYSTVSVLLNSRS

>AQS60746.1_OR17_Apisum

MTTTPRVTELTAPASEDLTIVDNRLFKAICLHQILDPTKGGNRYRLAFMVMWVSLSVQIIQLVGLYFAVNDLQRF
FTTTVIFNALLCLSKGYVLVVNADRLRASLEVARYEFTSCGARNQRLVRRSRVLTSTILRTFAVLSWVTCFIWALTPL
FAMDEYLQVTNADGTVSRYRVTIYNVWLPVPATVYNETTVWSLVYAVEVIACFVNVFVSWLLFDSYVVTMCFTFNAQFR
TVSASTTIGHSDSFSPPPHAPGTSDDNNTFNCYDELINRIKDNQSIKIYDDFFEILQPAILFQIIGGSYSVITL
IFLSTLYLMGFSIISIPVLKVFFGFLSVTFELFLYCYVFNHIETKCNMNFGLYSSNWTGMDLKFKKTLLFAMNTNS
SHRRVMKVTPMSIINLEMFANVMNMSYSIVSVLLNSRVQK

>AQS60745.1_OR10_Apisum

MAHIVDIFQNMGCSDHGYGMVFFNCCELAITLFFTVSTYPTIADPTQNLISIRLYGVLCLLIEAHIFAFIAVRIYHQ
SQHRDMYQHLHGVEIPENYRRKIATVIKHHFIIISNVFVAVSVLYTISLDWVRIGDPFTFPFIDVLPKTTNVTYVYCK
YIVYALPVYFAHLETCFLNVTFMFVSVGIVKRHFQILNDQVEEAIVNEDEQKLKIAIKHHQOVLKYFEDMKTVEKPIIL
MTIEFCGLYVGLTSCFVIQVIQGFHQQIILGLCIVSSIACLMTIIIIYCIYASNMALHNGILNALFEHRSCYSRNXSF
KRIILIMMTRATIPLEIKAGSVFTINLNLVLKILKFAYTVFNVLLSSINRQFKETAI

>AQS60742.1_OR2_Apisum

MDVMQKPERFILTPFQKFCIRWSVFFDSSSDRLSRIETVLRITIQFSTIMITSGMTMTSVLIADNKKALESFTYFVICV
FMLAIITFAIRTKRFNRAMLLMVVDEFPGYNRPMPDVLKRKMAAIRTSYGDFTMKVIVSYLTLVLFEIPATAMVPLAA
ASLTDVKLGQSQTQMVVLWFPADTSQVGMVAVSYVIQFLIVVTVKFIITGIMCSFSFFVFSQMISEFQILSAYVEHAVE
IVEYDQSDAKTTEQKLLDHVKNCVMLHDRLIYFKDQNLNESYGYIIILLELMFSTLYFCLSAFNMI FVGNRFVMVKGLLT
LSNYLAELFIFCMYGSVMVEDAHMGLLRASYSVAWYAQPVRFRQSLTMVMSRTQTPLQLTVGKVFIANLPLFLSVLKVS
YSGVNALRAANAK

>AQS60743.1_OR4_Apisum

MTSIGKKNQRKFYQTLMTLAFFLDTSQYRYISRFVKQFYIFDWMVLVSVAFAFTILEGNYRMPFVMEI IQYMIVGFYF
TSIFVVFIIKKEAIMSNYNCIQTKFIQWSNKRALHSNAAYKRNIKTVKSLSIPLAILSLSIALGPLISTINDIGKLPL
DNRAHFVLFWPTIVDTNKLSMYGIITYTLQVIFTIILYISVLSFNLGYMVFLNELITQFEMLLNGINDAFKYKMDKQFQ
TLFIDCIRHHQIIIKFLDDLKSYFKWMILIEIIIVVQVILAILIYNLTKVNASLGKVKIAGSILFNLLPICFHCHVGE
VVLSLHTRLSNHIYNMPWYDMPNKNKQLIVIMLQRTQRDLTLSSALFSSERASRSLISKVIKQVYTILNVLLKT

>XP_027852915.1_OR46a-like_Agossypii

MVLIQIADGSYSIITLIFLTSVLKGYISVSAPILKFVCGLASLTIELYIYCYGFNHIEDGKSTVNFGLYSSNWTEM
DLKFKKTLLLAMTMNSAHKRVMKVSPNSIVNLEMFSRVNMMSYSIVSVLLNY

>XP_027847983.1_OR43b_Agossypii

MVHYMAEMFFKKAICSDDDLHGRDAMRMVFFTYGELAITLFFAVSTYLSIVHSTEDLSVLLYGVLCIIQLLIFAFIS
FRLYHRSHFRDMYQRSRGMEISENSNRKIAAVIKHHLIMPNVFVVISALYKISSDRVHIGDPFTFPFMDVLPPIETTSV
AVYVCKYVVYALPVYLTQIEVCFLYVTYMYSTSIMKSHFQILEKQVEEAMVNKDEHKLMIAIKHHQELLKFFKEMKTV
YEKPIFLIIVSCGLYIGLTSSLIIQVIQGFHQQIILLGICIVSSLECALTIIIYCVYATNLYDLHDIRILNALFQHQLLY
SQNKSFQQLILIMMTRTTIPLEFKAGSIFTVNMMNLLVRILKFAYTVFNVLITSINHQLIKTAV

>XP_027842900.1_OR10a-like_Agossypii

MTVQMFVTTTELITLARTIDIWTFMFFSGLYKWIYMTVFNWEFAHLKFALAQLOTQGSKAYGRSADVFTADYLKQTKKIS
YWYLFSGFVAGFFIIVSPLLTYSKGD RSNFQYYNDPKSYPLSCWMPFTLNENWMFLTIFVCHSIALFLIVLVYLGIDT
YFFGAIYAVGGQIELLNTSLNNNENTLAQFENVSCTNQTSVFTEKQKMKCYSTLRECVKHHILILDYIKNIRKLFSTL
ILVDYLHGITSVTFALFQLTISASVIETISVICFICVSVWHQYLNFFGFEEIQKQLSVCTALYNVPWWRCDKRVQRQL
LTLMLILRSIKPTLITGYMYKLSYESFISFVKALYTYMVLRRVNTKT

>XP_027842896.1_OR67c-like_Agossypii

MVISLMISASVIETISVICFICVSVWHQYLNFFGFEEIQKQLSVCTALYNVPWWRCDKRVQRQLLTLMLILRSIKPTLI
TGYYMYKLSYESFISFVKALYTYMVLRRVNTKT

>XP_027841362.1_OR43b-like_Agossypii

MPRIDAINVFLQMTGCTDSKRMLYLYTYFEFLITLYYFIASYVSIIYYEQSVSIQLFTLLCMLIESYILLNITFRIYHK
NQFREMDQYSKQLGIPDDYQSKINIITMYHLIASNMFVIFPVTYTILYDSVRVGD PFTFPFLDVLPKTRNLLIYACK
YFVYAISVYIAHIELCFINTTFIYYVGVLYKRYLETIVQTIREAVVDNDEQKMKYAVIQHQRLLIYFNTMKIVFTKPIL
QSMSFNAIYFGLTTTLVIQAIRGYINQTIIVSICIASGIAAINITIYTFYGSVLLDLHDEILRVLFDNSFFFYVNKSFK
RSILIMMARATIPLTFTVGYIFTINLNLKIVKMSYTVLNVLLSSEAIKPHKMS

>XP_027841183.1_OR2a-like_Agossypii

MYGSSCTDKIALSSENSATDVVGEICLTPRSDTVLTIVKIGSSVIYLTVRLEFIYCYLFDSINKKRELNVNFSIYSCNWT
KMDLKFKKLLLLTMQMDANQMVMKASPKKIIDLQLFASIMSTS FNMVPLKIKNSEYYKSQ

>XP_027836930.1_OR4-like_Agossypii

MVGHSLFASVDVMQQQPESFVLTPFQKFCIRWSVFFDSTSDRLSRIETVLRVSVQLSTIMITSVLTMTSVLIADNKKAL
ESFTYFVICVFLLAITFAIRTKRFNRAMLLMVVDEFPGYERPMPPDDLKRKISAIRKSYGEFTMKVMVSYLTLVLFEI
PATAMVPLTAARLTDVKLGQSQTQMVVLWFPDGTQIGTYVFSFVIQFLIVMIVKFIITGIMCSFSFFVFSQMISEFQI
LSAYVEHAVEIVEYDLSTGKTTDQKLLDHVKSCVMLHHRLIDFKDQNLNESYGYIIILLELMFSTLYFCLSAFNMI FVGN
RFVIAKGLLTL SNYLAELFIFCMYGSMVVEAHMGLLRASYSAAWYSQPVRFRRLMMVMSRTQTPLQLTIGKVFIANL
PLFLSVLKVSYSGVNALRAANAK

>AQS60756.1_OR5_Agossypii

MPRIDAINVFLQMTGCTDSKRMLYLYTYFEFLITLYYFIASYVSIIYYEQSVSIQLFTLLCMLIESYILLNITFRIYHK

NQFREMDDQYSKQLGIPDDYQSKINIITMYHLIASNMFVIFPVTYTILYDSVRVGDFFTFPFLDVLPIKTRNLVIYACK
YFVYIAISVYIAHIELCFINTTFIYVGVLYRLETIVQTIREAVVDNDEQKMKYAVIQHQRLLIYFNTMKIVFTKPIL
QSMSFNAIYFGLTTTLVIQAIRGYINQTIVSICIASGIAAINITIYTFYGSVLLDLHDEILRVLFDNSFFYVNKSFK
RSILIMMARATIPLTFTVGYIFTINLNLKIVKMSYTVLNVLLSSEAIKPHKMS

>XP_027847498.1_OR33b-like_Agossypii

MDLSNLIVIIQDHQKLTKKIHDIFDEMRFIILFQLLSESILMSLIPLVLFLNSNNGISLTSTESIKLLSAEIVNTGHL
FSACYLFSLLIDIYNDTINFALYNCNWTEMNINFKKLLLFTMQMNNANNFKLNISTNIIIVNLKLFNTLKFVAVASYKTYA
N

>XP_026685329.1_OROr2-like_Dcitri

METSFSTQETEIVEKTYAKAKFIFYVCVTSAGICLSGSILETFFTVSQEELDLLSHIYDRKYPNNRLQTNFWVPDIDD
SEPPYYQIIFVTELYLIYLIILLTILSVSLIPMLVTHIAGQYEILCNRIQSLGHSSIMWNKRPIDDLQDRESLRSIM
QAHKNLISAQSKLQNLYSAGMFIVVFNNIMFALCLYQLTSLNTNTVSKFRVYKLTLEFLTAVVYTYLCHCSEILDD
CNSKLCDAIWNRSRWYHCSRHVQKDLIMFLRRVQRPNHLKCLHGAIILSHVYFLGVVKVYSYFVNCMRFKSRL

>XP_026678730.1_OR2a-like_Dcitri

MISLCLYQIVVSPNKLSPLRYFKFITEFIAITMEYFIICNCSEIMDDCNGLMRSALMNCGWKDCSTSLRRDLCLWRR
VQRSNHLRFYNGAVILSRLFFLQVVKVAYTFTNFMRLKSSHG

>XP_026689162.1_OR22c-like_Dcitri

MEQFKRIVKLHQSLQLQNKVTTCLSSAVWLTVIESNLSFSLCLYQMSVNIDQVSGYRFYKTLALFLGIAFEYYFLCD
SGERLEECSHLLKNSITQCDWHRCSSKTRKGLCILLRRAQRPNHPSFSWGLVVLQREYTLKAARVGYFSFMFMRLKSG

>XP_008482312.1_OR85b-like_Dcitri

MSQPSKVQKNQYMAFNILLKTCFIWPLDCQNRLVLLIYRLMTWSTCFVLAFFVSIGQLLKICVSLDLQELSSTVDVFT
MTASALYKMFVFLMYNMETLLRIVQVVHTQFEEYPIRGHIFTMSSHIGPCSLYSSLYVLAGIISFVWNFAPLFFSPM
CPIIDPTAPLTIKRNKVPINIWLPINIAESPAYEIIFALESYTFMSMTLLYLSIDAFFFYLIHVISGQLKLVDASF
SLFILDRLKHEHATRLRQMSKLPDAGSLTVSKSIRLEDSLQYRHHVLRNRLNSIVKIHATVLELIRTVEILFSSTI
GIDFLHAILSLSFALFQTQMTVAFIDTVKMYLFLVICIVHQFSNNFFGEMLIHFQHIIVNSIYESPWYTGDAAFKRSV
RIITNRSLKPKVLEGFKMFILCLHTFVEVPSSVLWLR

>RPRC000362_OR13_Rprolixus

MFTKFIEYLDNFIEHFDNRGNEICDQTIYKEFKRLLEIMAIYPKLKPRLVIINILNFSFINGLLLFMIFISTVTAVC
AGKEDFPTMAHSIVYVLVTILFLFCYNGMVFNRSSIAIMHTIFGQNVYDYMEQNNDIPAAIKAKYVRMKKVINQIVGV
ASVLLALFIVCVGPLLNTDVTETVQHPENGINYDLPIPLWLPFEASNGFAYFIAVAIITVTFTYVFLINSVVSALF
YNVAMTMDQCLEILIKSINNIEERAINMYKRNNSLKLPSYGLYESLEFQKIIIEICMKQNIQHHQOILRLRKAITKI
SEYCVLVLYVVSIIIALTAYTATTVSNEPGTVLTFGCLTTFLILAIAGDCLMGQS FVNLNEKLRIEISTKWRLYSK
STRQLIGIMQQYTCRDFNIKVAFNVTLDLEFYGSLMNTTYSYLSILIAMNK

>RPRC000616_OR18_Rprolixus

MEKFLQNYFSEDNENLHQFVRKDYGILFQISGFYRRMDTVKECFTTSYVFVMIITMLTYQATLFLITMFHTDNDLKIM
FQSFHFLVLYFLSIFWLFFYLHREKVARIFKAIHHGLTYDEEREIYREKVKNQDYQFYRKLTIKYPPINSIAVALFV
EIGPLIDRFIRNIQYETYSPPGGLSQNIPVHTWIPYGTDTNFKLIFTILVKLVVASWAGLMFAIALTMILESTNMLHLE
MELLIHSLNHLPSRTENLFNKTASKRRKLDITKLREDKEMASCYKRNLRDNAIHYQRIIRHFLDLQAILSVPLLVNF
GSSILVALAGIALVYGRGQVGLIVSVVGIFITESLTLLLACWFAERVQRQSENREALLYATDWIYAPKDAVIYIQIMS
SRALKSLTLTSYGLIPVTFETFANIMSTGYSYFNLIYATSSE

>RPRC000235_OR33_Rprolixus

MDYFTDDKDVEKFLVEEYKYLPLYLALTYASWNRKNVPISLAQYIIYFTIFGFMIFAFTYTAYLCLDDIAVWSQSIHFA
IVGIIICWTICLTGFFSRKEFIISHRNMMSSFCEYQSGDFLKLGLIRRRQKEKKILLIFWPCYYAVCALIATIIAPMIDI
YIGVNSDVTNKNGVFLKLPLALWLPFDVDNSMVLVLVGVFLIGAYCAISALSVSSGVLVATFIKQQLCLQLRLLNYS
KELNRSKDEYKRRSQNPDLYNDEERKYSSCYKECLQKNIKHQNIISAHAQYTYCTVTASFFNVLSVPASGYMCYVF
SNFCYFFCYSKYLTFKIWIFVLIILFRFCTIFCYIKYFLGEEVRVSAYSIIKWYDHENVRVSIMHFGCLIPLEFKA
AHILQLNMQTFATVMNSAYSIFNVMYAFKE

>RPRC000229_OR40_Rprolixus

MVNIEEYFESTEKLYMWSGFQLFQSSGHILRNKIRRQLLIFIISCDLSLLAGAVFSEDLMDMTESLLICIIIFNAGN
KFINILIKELEMVKTFFVQLRKLVEDQLKVNWKKEILVNSKARVDWLQKMFLWNFGICVIAAILNTIVDNIRHPENLRF
PLKMWFLAPNSGLNLTIGTIYQITICSCAGIIIFTFGYMTFTGLIIHISSQIIILSKDMRNISAKANVQKQEMLKNF
RHHQIILRLVKSTNNVLGNVLFLEIILSSMQTCLVAFNCFQRPLNDPKQAFFVPFLLCTTILPASICWCGQLLMTKSD
DLINAIYDEWYAMPKIRRDILMIQIQASNSMTLNARNTFVMAMPFTFLKTLQTAYSYLTMLTATDK

>RPRC000108_OR42_Rprolixus

MAFKEMNMFYRIFHMFVWYAQPTDLYGTLSFIGQVFRHFLFILSWSAEVSVSYLEGFRRSMNGSAFFLIISFVTVFE
IVMIFLKRKEIHCIMNMLEMSANNRSGWEKQYFDQRTVFVKKVVKTYQTMMIVIGLLLYLVFPMVNELLELYGVQPMN
VALPFSSELSIEGPYKKILEKLRTFWGIIWCLHGITFFSCVHSLFFTFTLYTAAEIDLVC EKIRRLSYIDDEVEYYKVDK
ANKNGISLLEIIVHHQFVLKLSKKLINLFLGLIFALENTIAAVSLCLNLITVRFEIMNSKPGVVIISVVSMSFNVIQQT
TVYYFGNYLTERCNVIYERLYELPWYRLPPSVRKTISIMSRQASMP TVINRGLARITLENYAQLMNAAYSIFTVMSS
LLDE

>RPRC001623_OR50_Rprolixus

MEPSTLGSEPLKWLHTTLKIGGFSMNFHDKNKKRLFGHIIQNIVLYFFLSYDLVLIYSDGVINIDIDKFIYLPFFIIN
SILHATLFFKQKSIRHLLGKLKEVMECNKEDRNHFAETYKFVKKLIKMFVTVQITVWSCVYIIPLLQSIASLFYAKNF
KIFLLSPLQRI MGFS PGYFQDWILLICTILWTFCLMCTFFSTFASVDLILIFSYLEIKLFHYEIEITYYTSSLKSDYDS
TYRTIVRRQVQISKLF EALKNAGLYLGLQNTFIAVTICSSSLYAATTGQSLYVILLFCYGTLLFIFNSFLCCYISDRI
ETEGSNICNSIYELPWYEMRPQERKVFVKMFIKASQSMKMNYKGIFLTNLDTF SQIMQASYSYFMLLQQFS

>RPRC002122_OR53_Rprolixus

METEGNSHYSFLHTVSWLLYVGGIKKNEYGDVPITFRSLVLMVWVFALFFAIILKLFCCVASLEL FVAGIMVAAITFH
LVVKQFILI IKRKEIISLIDDL EHC FILARTDPKFIEEKSLSVKKSGNKLIIIMGVSIVIAFGFYINHELTGDTSLP
VEMWTPNGDEGTSFTVLVYQLFMGGGCATELVIIYALFLGLVSTCRMIIDSLKGNFVKGPDEKETLEAFLSRNVLAHA
HLLKVVDINANISPIVFLEILLSSIQCCAGYLANFDSFHNSLYFVYVQATLVMPGMICYLGQEIKNYYQDLFDQLY
DCPWYIATPKQRRYYLMMISMSTSKPLILHFRFLMTFDLERYANVLQAAYSITILQNLN

>RPRC000055_OR54_Rprolixus

MKGQKTYYSFLETVIWLLAITGIRDELGNVPITFRAIAFWFWIFLMIISIIILKMFGGQTLEEF LNVLFVGVIIITVA
IKQANLIIRRKQIFKLLDDIENCLTLARADLQVEKRSRLKKIGKLITIIYILIATGTMF AFHYVHHELTGDKTLPVE
LWTPYSEEGTFTTVIIYQLIIDTGATIGLIVFSALLLTAITFSTMIMDVLKKNLAEGPDDEETLETFLSRVLAHAKL
LNI IKAINVNISPMLFIETVLSSIQCCAGYLVNFRSFHGAMTYIHVSTLTLPGLICYLGQLIKNKNEELFDQIYDS
PWYTAIPNLRRYYLMMISMSSKPLILHFRFLMTFDLERYGNVLQAAYTYVTVLKNLLS

>RPRC002162_OR56_Rprolixus

MIPPSTRLYLLIEMAGFGYRPEHKYPTCMKLYMIFMMITYAINWCIIIVALYYEEDNFNTKAELIHYLLESVIATVKY
LNFLLR FKEVRGLMKFVDEVRESRSDPNNQSFIEESEKSESKALKNCFIVAITVPIIFCTSSSLVYIMISGLKKLPFL
YWVPFDSQRYFWFALAFQLIIYSTAFKVYLGSGFGNILLVMVITSYMKVLQRLLTTKYGETAKRRSAYILHQINRLV
FLTNEFLSGQFLFEISFSAIKLSLRIFAFIKVFSTRKQDVVVSFGILIFSMSSPVAISIIGEMIKSESEKIFAAAYDN
VWYEEKEEKTKKELLVLKASLRPHSLHYRRLTSFNMETF GELLQASYSYLTMLLQFSEIN

>RPRC000526_OR58_Rprolixus

MVYSFLKLTGMEEEGQLRFLNLGTQLFICAYVTMVTGIELIWHGHGNLYEIVNALSTFTVAINYLT KVVIILMYRRE
IRQLFARLEHLHME LLKDEDQRHIIIDMERFSKMFLRFYKISLNFYPILSLITNFINDYKTD FRKIYLSLKLTPWTI
SDIWSY SAGAMLILWISVMMLSSFFSFIAMELAFTFEITAF LKVLQGR LNNMNEKDENIYGLHRDI IKLVTD FNALFS
GQMYWDILVSSIQPCGSGFILIKALKRKDPEATELFYKMILVVIGPFILCAGCQQISTESEKLHEASYMIPWYEQTPR
KRKNLIQMLTVTTK PSTINFKGIIVFNYS CFAAVAQGIYSYLMVMNKFATDD

>RPRC000733_OR78_Rprolixus

NFKRLSPIQLTNYILTRCGLLNQEGRPWTVIIATYHAIMLILVGFLT VANVYKKLDDFNEFILPLLLLPIITHVSGKF
FHLLLNQRKLRKLLDGLEHLRQDELGDSVYAEHF IKADKMNKSIVYYLDLNMHLAPVISTGLNVLDYLSASETPTLM
IPIWIPWEHQKSWPYVA AVIASTVISFTFATVYASLYSLVTIALQMSALLQVLQAKMEEEEKANDKSI FRQHFNILOV
IQQLNNLLSGQYCLEILVSTLQPCGACYMLIKY LKSGDSGWLDCLYKVMVALGASTILCAGGEEINSQVENLHQ SAYR
SSWYEEQPAARKDLLILMTVTARPLQFQYKGLVT FNFHRFAMVIQGVYSYTTMLSNLETA E

>RPRC000476_ORco_Rprolixus

MQKVKMHGLVGDLPNIRLMQMTGHWLLEYHEETGGMVRLMRMAFCWFTTLIVFTQFAFITCFLIILETYDADQMAAAT
ITTLFFLHSVTKFMYFALRSKYFYRTL SAWNQMN SHPLFAESNARHRATALARMRKL LMIIGSGTIFSVLAWTTVTFL
DDPHRQITDPENANETITVEKPM LMVDAWYPWD AKYGMTYFMSFIYQLYWLFISLSHANLLDILFCSFVLFACEQLKH
LKEILQPLMELSAALDSVVPNSGDLFRASSTASNIPLIGNDGN EFDVRGIYSNQRDFS GFGP GIIAALPSNGGGVGPNG
LTKKQELMVRS AIKYWVERHKKHVVKFVTSIGDCYGSALLHMLTSTVTLTLLAYQATKIEGVDVYASTVIGYLVYTLG
QVFVFCIHGNELIEESSVM EAAYSCHWYD GSEEAKTFVQIVCQQCQKSLTVSGAKFFT VSLDLFASVLGAVVTFYFMV
LVQLK

>RPRC001628_OR17_Rprolixus

MERFLQNHFS EDNENLQQFVRKDYGILFQISGFYRRMDTVKEKFLATS YVFLIITMLTYQAILFLITMFHTDNDLIIM
FQS FHFIMLYIFSIIWQFFIFVNREKVARIFKAIHHGLTYDEEREIYREQVKANQFYRKLSIKLP SMCFLVAVIV
ELGPLIDRYVRNIQYEMYSPGGLSQNIPIPIWIPYGTDTNFKLIITMLLKL PVDIWTTVTCGTCLMMVLESANVLHLE

MELLIHSLNHLPSRTENLLNKTTSNKRKINITKLREDKEMASCYKRNLRDNAIHYQRIIRLFLDVQDILSVPMMLVTFV
SESIFLALSASAIIVFGKGQVDLVMSVLGIFVSESMALLVSCWYAERIRQSSSENLRDTLYATDWIYAPKDAVIYIQIMS
TRALKSLTLMFGLFPITFETFANIMNTGYSYFNLIYATNSE

>RPRC000337_OR36_Rprolixus

MAPPENDRSYLKGYDKVGGWMLRSATVYSLDIISSTKEKFLFCIQQAVHIIISLVIMWTLAIKTIYIYNDTSAILEFL
HLLFGATVLWVQTLNIVKRDSMENMMLYIGKDFYDYPNEDMSEEREIKKNTLDWYEKFWKFCLIIITGSICFLFIIRI
ALMFFFQHQEYDLTKEAHPYLLLKVYYPMDLNNPMYLI FAVAHQLV IASAVAVLNAQATVFIRFIGCMVAEFRVLAH
SIERMESRIHFLCKPNIETEKIGNKELYVKKFHECSIVCLKNNIKHHQQQLVDFFISLQDVHGFDSLAAFLVAVIVLCF
GSFMVIVGLKSNLADCLVTIALLTMEMTNIFFLCYLSEELERLKFCESLQFMNWRDMSLKFKKMLIIINECNKKKF
LLKTNLTGVTASRETFSILTTAYSYVNVLREFFG

>RPRC000041_OR39_Rprolixus

MKDDKEVERFLIKEYFYIHKIGLAYVPWDKESAIISIFQYAIYLIVLSISTALLGYTAFLTIDDIALLFSQDLHITLIC
VVAIINIINNSMIYRNDMVEIHKNLMNGLCNYETGEFYEGEKKNIFKQKKFLLLWPLYYGFCASMAIVIGPLMDSYYG
VDAVHTSYQPQVYLKLPFAMWTPFDMGSQVPLHLILSVFQAQFMVIINNLTISAGVAMYVVVLNLYLCLHLNLLTNSIDNL
DVRATELYKVRQGNRASSKLDRVLYNKCYNDCLIENIKHHQSILRLHKMYVNLSSVPLFVFPFIAGAVLMSMAVVFHMS
DDVRIGPKLTSGTVVFGEMLNMLMLCVYGEKLLKLLSEGARDTIYSIRWYEMDKCHQTVSIFQQATLKPMNVKCVYLL
DINMETFATLINSTYSYINLVYAFK

>RPRC001291_OR44_Rprolixus

MFTKEIRNTIKLLEWTGHWFNYDNKTNEKYLKIFQIFRLTYLGIALLSNAYTIYKGLKKSLESFGVFFPMTIMCLLC
AMLIAINNVRMKDCLVLKNSMIENHKENWEHEIFQEDSEFIWNIVKKYLKTIIMYLLYTPIPLFIDFIAYNCNLYDG
PPLNLPQTLSGFMDDQPV RDGKYTYIILFSQMLTGIGIPNYSGNVFGFCIITVYIRVEMKIIQKKILLNDIDNNLKK
VYYSKLVEI IERYLQIKSILKDLEAVYGAQFALQNALAPISTSFCNLIVMLEDKSVAPAYLICSLVLTGPFICCLA
GESLQEENSNINKLIYNLKWYEWKSRDGRMVLMMLEQTSLSMTITCKGLSPMGHMTFQEIMNKTYSYCMLLRSL

>RPRC001292_OR46_Rprolixus

MTPYEQLLFLCLSFGGYWFRNLNDKNKLTFLHILQNIILYANVIYEIIIGLYIEGDLQELLAGKFLYLPVSI FELTNHLI
LFCLRDDYDRMLLTQE IFAKDLEERRKFQYYKYLRWGIIVYGT AQFMLWFVSFFLVLFKEIIAVIFRLEEF SFVPP
MPIENIIGLEEGRHLHNWLKLI FGFLYGTAWILSYTSIMMGTFHFMAFTIIEFRIFRYQLEVYHKNMLIENFEHTMKT
IEKQCLLARLFENLKNTSSLFLGLQNSMMALNICFLLHTFTKITNAIAWYIGLYGFIGFTMIITATCYFGDLLEDEGD
AVNNILYDLWPYSLPPKQRRYFALALLQGRSLKISYKGILPTNLETY LQFMQAVYSYIMLLNSV

>RPRC000208_OR61_Rprolixus

MVYSFLKLT DGMEEEGQLRFLNLGVQLFICAYVTMTGIELI WGHGNLYQMVNALSTFTVALNFLT KVVSILFYRRE
IRQLFARVEHLHMKLRKDEDHHRILDLERFSKICLWSYKISLMMYPFYSFITNFIIDFKNDFRKIFFTVQLLPWNI
SDIWSYTTGALLVLVILSVMIYSFYSFMAELAFTEITAF LKVLQGR LKNMNEKDENIYGLHRDI IKLVTDFNALFS
GQMYWDILISSVQPCGFGFSLIKALKRSDPGATEIFYKIILVIMGPFILCACGQQISTESEKLHEASFMIPWYEQTLR
MKKNLIQMLTVTTKPTVNF RGII VFNYSCFAAVAQGIYSYLMMINQFATDE

>RPRC000200_OR74_Rprolixus

NFKRLSPIQLIYIYLLTRCGLLNEDGRPWTVIIATYHSIMLILVGLLTVAKVYKKLDDINEFILPLMLWAMIIHISGKF
FYLLKNQRKLRLSLDRLEQLRQEELGDSVYAEHF IKADKINKTILYLLNLNIHLVPVISTGFNVLHDCLSTSETPSLM
MPIWIWPEHQKSWPYAAAVIATTIMAFSGVIVFATLYSLVTITLQLSALVQVLQAKMEKKQANDKSI FRLHINVLQV
LQQNDLLSGQYCLEILLSSLQPCGFCYMLIKYLKSGDSRWVDCLYKVFVSLFATSFLCACGEEINSQVEYLHQ SAYR
GSWYEEQPAARKDLIILITITSRPLQFYKGLVTFNLRRLATVIQGFYSYTTMLSNLETAD

>RPRC000468_OR77_Rprolixus

NFKRLSSIQLIFDLLTRCGLLNQEGRPWTVIIATHHSIMLILMGLLTVAKVYEKLDDFNEFILPLMLWAMIIHISGKF
FYLLHQRKLHLKLLDRLEQLRQEV LGDSVNAEHFTKADKMNKTILYLLNLNMHLAPVISTMLNVLHDYLSASSETASLT
LPIWIWPEHQKSWPYVAACMVIAFTGAIVFSALSSLVTITLQISSLLKVLRAKMEEK RANYKSIIRQHINILQV
LQQFNDLLSGQYCLEILLSALQPCGFSYMLIKYLKSGDSRWVDCLYKAFVALFATSFLCACGEEINSQVENLHQ SAYR
SSWYEEQPAAKDLIILMTVTARPLQFYKGLVTFNLRRLAMVIQGFYSYTTMLSNLETAE

>RPRC006981_OR83_Rprolixus

KTNKLSPIKQAFRAIRFSGLIVEKENQWAI SLTVYHILMISYGAIAVGESGFQHLVLSLNS SAYRAFSVPFINSFYRI
ELCLVEENLLQFVEVLSTSLYINDSTKALNILLHQKKIKDLFNRLDKFSLKILQDQEAGEHVKSIDKENLFAKSFRV
KFI SVFIPILLWLPPISLITGIIGDYLTDFVKPHLPYQLWLWPWSMDEFLPYLAGMVVFTMLELTTCIFYTSFTIMFLT
FTNELSACLRLILQARLETNGPTDKNIYKYHNVIIDILQTYNGIFCGPLYIEILISTLQPCGFLYAF TKVFKNYNPAAF
EQILKALLCVAAPYIVCSGQEI SNQKRLNKLKILNEDKYGDSMSSVCLGTVH SVLCPQLVQEFSGRCWEVGFFVLE

>RPRC000431_OR84_Rprolixus

TPEMMPVQRAYHQIRFSGLILEKDNPKTIFFAVYHMFVNYVGLASGVDLIRGNQSLIDFVQSLGLAMIYINNIVKA
TNILLHQEIKELFARLDKASLEITRNQAEKHIELRPENLQTKFIVTYSKLLMWFPFSSFTTGLLSDYLSDFVKPHLP
FPLWIPWRMTTFWSYLATMVFDTMLSVTTSYYYYVSFTALLFTTTELTACLRVLQHRLETNGPADKNVYIYHHTILEL
LQDYNKIFSGPVYSEILISALQPCGFLYAFIKLLKRQDPASVDLATKGVLTVLTPFVPCSCGQTISSQIERLHASAYM
GKWYEEKPKVRRDLLTMLTVTTTRQADMNYRKFISSYFVVCYATVIQGIYSYLMMI IQIGGD

>RPRC002158_OR93_Rprolixus

IIKRFNNHNQGDKEIIITGYWVSKWCGFLYDQSWTSVIHITRLILFLILTILHILIAVLLYEDMSLLEKVLITNYAI
VYINCLFMVPLPIYNSKKNQLLCSIIKTEFQEAKIQFSPKQLEVKKYTKRFVRNFAYNVAFYLYGSLIINFIRRYLFE
GVSLKILPVNGWVPFELNSLIRYTFVFIFQIVASISGMSAHVGFLIIFVTHSVSLCEQFEILSIHIHETFKHVYFGET
MECFEVNNNREKYIQFRLKYAVYQHKGKLLQYFNLYQEIYSMFLFIFALGTGAMICTAVYVITDPMSSLSLVIVSFACL
LIPEIAIIGIYCWFQGOVTKNSYIQESVYNTPWYTQSVTTQKLLQNMLTACQRRERIVKGGGLQQFSLGLSELLQVS
FSWFNVLNAIHQ

>RPRC000188_OR101_Rprolixus

MANSNSIPSNIKPKVLRVPLHTLQIAAMWPPFTINRVIKTLMRIYLVSSAAILLALCTVGLLVKTIITTKDLVDRSEAID
IITLTSSALYKQLFFVKHYNELNDLVMTADILEVPVGVWLKYPYLSAAHCTCGFLAIGFWMVPIILKVSFGITTLPEM
RLPMNVYDWGPTGWQFCAIYVCCCFGLAFSTHVYMAVDGFLFTAVYTANGALKLLADRIMKIRSIIKTEGKETLNEIIL
SELKKCAKTHALILQFIRKLEVIFRSLIVADVLAHAIISLSFAMLQMSSESKGILEWTKMVFFIFYCFLHQYLNSSFQGD
LINKQENLRYVLEEVPEECSLPVRSTTLIMITATTNPIKLSAARMYFLQYATFSEFVKNMISTYMLRQLQDNSV

>RPRC002141_OR110_Rprolixus

MRKGFLDNYGFLMSMAGIYLTFFGNWIIISMLHLGVGSLQLICYFYTIYLTNDMGNVIELLHSAVIVMTITGPAIINGL
LQKRRLERIMRALGTGVYDYGVTLDLDDSLKEIREFSIQTRARKNLCRIFLAIVVANLFAIGLIKPIAMNLLTTIEYD
DDFDGINKRVPVHCWYPLDSRVPLNNLLCFIFQYVIVVTIPLGVCGLDLIYLCMAEEIYIQLKIVGTSVKNIKKRAMK
VMKLNGTNDINKAMSRCLGHSIDHHNALSSLFEDFRNYHFYIILSVMTGATLMLCLAGVIFTTDSVAPVFRLTYVVF
NTELHVHLVFCWYGEQIKEMSIKIGDDMYDADWFSSESGHIKIHLMI IQKIVRPFKVTAGGFMDASMDTYANVLSSAY
SYFNLVKAAREN

>RPRC013097_Ir41c_Rprolixus

TLTLDGTEGRMLTTFCEIYNCSFEVYDDGNFWGTLNSTNGTAEGVIGSVFTRKVEVGYVGIYQWPEVNFHIDYSASWL
YSTLNLVLPKVPFISAWKTPILPFDWQIWLATFVSWIFGTIAFFTSTWAIQMHGQILSIKIGKLRFSMTTMAVLG
VLVLQSPPRYLLPKISSRLRIVFICTELLALILTSCYAAELASYLTVPQFSKPVDTRETFVDSNLIWIAPHEVYTLCLK
DTGDPDIADKLVASCKVYKIDELTEKAETGKFGVAIEVLPGGAVADVPLTEKVIKSHFMKESLYGSPCALMFNKAS
PYVHHFNNMLHRLLDAGIITFWEEDVARKYMSVRKQMAIKESKVKRSENLVKKLTIDHVQGTFLIYTSGLICALITFL
FEIYYFKKKQNMHEWVF

>RPRC000047_Ir75e_Rprolixus

MKVKDFLFAIISVYWLDSFGIGLPSISISHIISKHFLEIQGFQLYVCDKRDIAIKIMKAFYNEPRLLSSKLLTTANSTR
DTPLHFIQPSVLGIFVDLSCEKGIRFLNESSTKFNASYNWLFWTTNYYNLLTLLRPTRVRTDSNVTIALSGAGDTIRM
LDLFRMHISLPFTVTEVELWYSNGTYIHLKDPPERKNFQRAVIPATLLLTSMYDLNNSIDPLLDQAYMTGIDGAQRF
GLSTFLTADIFNITYSYTLAKSWGYPKKDGSWDGMMKHFQSGISEVGFGPAQLSFGREPLIDFLMSLYLFFKCTFTFK
QPKLFGTSKALIIPINYKTWICIFIVVILGAILRAMTVHDKSEYSNDNWSGSTFLVLAALSQQGMPDNPDKNPTRLF
YFFFLFFTYFIFLYYGISILNGLLLPAPNAIQNMDQLLES DIRIGIQDLPHYELENWNSWTTGVNKKLQALKPPH
KKYYGTVEGLQEIKKGHFALYNFVGQFYDSVPEVLTNAEIASLTEIEKFRPRKNGALVVENSYPKEIFKRKLFFMREV
GIIDRERIYWMRPKPYASWELDALSVGMKPLSIAFIIILGAFIAIIFCMTEVLIRERLKKRNKMLKSKKIAFTRKFR
N

>RPRC000847_Ir102_Rprolixus

MKLLILCIITNVYPLPASSVETNNNLYDCIHLIANKSVLVKTIILREEIENYKVSVDLLKLKNWPKMIMNSNQSSLK
HLQGVYLLFPDSNCTDYLNLFEKSSKMINSKYIIYTESMFAMFTLFDISWKFRVIDIIALYGSKEDTILVFTYFPYGT
HGCMTQPVLLIDSCFRKYFYDYSLFSRKNKLYNYHGCRIKVTLKDKTVESVDTNADDLAFKLLKFLSIKMNFTMLIN
VADAGENAYAAAYQNFTSASVISVARGDSELGVGRFSQILDYHPMVITSKEISMDCFTWAVPIKAGKVPSIWVYVNE
FDYLTWILIVLVYFLIALLFNYIFSIFIRPKNDRNLNLVHIFGTFINVSVKSHIFHCTERVMYASWLLYSFVTISAYQ
ASGLSLVTVPPEIPEIDTTQYLLQTLGALKSEPKIYHVLASARETSEQIKLVDSFELILPDENFSSAVHNIYKKRNL
ALFHTSDKLAKTEYEIKLKLNTNKIVHIINDCLISSPTSPFILRKDNFLIEPINDFVTRLFESGILKYWDRNTLEFST
HNSTNLHSASMNIIINLFGGLAVLIIGYVMAIFIFLAEILSKTWNL

>RPRC005659_Ir101_Rprolixus

MVDGVVFLLENYFDGDCLTLFTLRLQATELELLLLLQHLKMKAFYLLRPVNVIVTKVPFPCFYIILLGEKDYLLDNDP
AKINSKFFLNKETELSKYIVIAQNKLNFQQKSKALITLFRRGHPDALLNPEDNLFAYTLFPFGEKMKNCPENGIEIK

LLDKWVNNRFHKNLYPHKVPQLFNGCNLDVSFIEYVPYTIYHGRNEVKGIDGVMVKLIGESLNMTVRWLLRESWFSYL
KNGSHYGILPDILYGRTHLACGGLIRSPELAGYVHFTTTYAISYLTWYLPQPPYEQTWKTMYKAFQKKLLLLLSIAVIF
TIPVFIAIIAKLADNELSLFKCITRCYVAAWALAVGNSTGKKPRTTSVRIIYAVWMIYAMHLALAYTASLAGLLTKSA
HRPRVKNFNDLINTGLPTAMLDVHQSATTNISDPEVRQALDKFLLIEDTDEAISLMALHGNMSILDNSAYVDYTICSK
FSKKPIYNSGLSTFPMKTGLMMRKSHFLYKIINKMIIRLVESGINDKIIKSTTKCRKTSRRLKSGQFKASTLRGLSGA
FLVLGVGLLVAIAVFFLEQIIYQRNLYRNDKEAII PYLN

>RPRC002996_Ir104_Rprolixus

MKLKWFVLNLLQVISADLINSSKENFFLNIVANLPREKGGNLI FLQENFTNLLPKLAQIIAYSLLAVKSQYLLATPSN
CIIILSTNTADLNKMHFSYGTCIVTVFYGITYNASNTKFLKNQTNSAIIIPIDSSSVIHLKTPYTQYEINEYNRLLLLK
KILEKPPTNLYKKSVRVTMFPKRFTSVKIDGEKFVGFDDILSILEERLNFSTVMIKNNGRYGYQHPNGNFTGSLGDV
IHGKADLATNSIFIKYNSNDIQFTTVVMLEELCIAVPASEMIPQWMVVFVLFWEPEVWFGILLTYTIGGILWQKLVNN
FGKPRKLHINSSTTWLTMAKIILQMSVNGFSTRFWQRIYLLSLILGSMVLTKCLEGSLLDKLTRNSYWKEISSLEELG
KSPLKLYTLSENLRDFTFLKSNKLQSKFYVYKNVSEDLVQLVYRRKDVAYLGRKIDLIIRRIGEIRTNPRAVHIIIECP
RSYHLGYLLRKDSHLYQPINTLLGRLLAVGIIIDLQNNYINATMAKQLAKKFRKNQRKFQPVKLRQLNLTIGLLVLVVF
GYTISSLVFIGELLVKRFLVKRNEG FVYVN

>RPRC002112_Ir75a_Rprolixus

MMLYKSYLLSTLFFYSVAAEFPLNLI SNYFNQLEINVVTILSCTSLEDKVTILKYLNSRGFQVSYSWYSIERIPLHRR
GAVLDLDCPGRNKILWKINKEKAFGMHTEWLLLRVQKESDQHLNGQIFMSEAVISVAESYALPGSCVKLLQLIDDVE
KLYYFDMYRVTIREFMQFQLDWSGSKKLPLPEIVPRCEYRKRNSNFGGITMKSASIIKFPDFFKGFESLDYREVDTWAK
MHFPMVLLTEQLNFIQNIITITDLYGWETNGSFDGLMGLLQREEIDFGATGFFMRKDRMKVSDFS AETFYVRTAITFK
QPTLSSVSNI FVLFPFSKLVWIAGASLTLLVAVILASEYFLTCKISIIYPYKEKHAATADLITVMSTVCQQGTENPIS
LPSRITVFLFSLFAFFLYTSYSANIVALLQSSAPVIKTINDLTKSHLKFKVQNDIYNEIYFKEAVDKAIIIDLYKVKVK
PQGKEAYCSPDIGVKYMRGTGYAYDGETNLIYKFISDTFEEDEKCSLSEMDLFYLPALAVPVVKRSGHREHLTRMMSW
QREVGMFSRIKSVWLAPKPQCGSNGSGFVNVGLADFLPALLVFVYGAAVTTGLFILEILFHYNKLRWRNSICGGKIV
VTDLRSI

>RPRC017346_Ir75i_Rprolixus

MWKIVVLTFLNLLLPAAIKTESFLPSITEIIKMHFRSAEGTNLFFCQKEDALLTLKKS LNERQSYFINHLGEVKKLQL
TFLMPQVLGVFLDLTCEGGHVFLKQNSGAFNATFRWLFWCNDPTATTKLLHPTRLMTDSDVTLAYPNGTDIILQDVYR
MHDTLPIVATNIQRWSKEGTVTYLAEPPPRVDFKGIALTGSLLLVDTVINKTEEIPLLFDPAYLPKVDANRRYGF TLF
KHLSEYYNFKYKIIYSDVWGYPQKNGSWSGMVGVQVKREADMGLSGVHFSVPRTVIDFVGILDTTVVCF AFLQPQLF
GSYKALILPMNATVWACLMMIAVVGVVVFRIVAIYDDTSQSNDSWGSFLLVLA AISQQGIPDSTDKLPTRIIYLSLL
GVSF LAAVYNTAILNGLLLPAPNAIQNMDQLLASDMKIATSDTKYLLYDMTHNDSLTMKLRKKMEKSKLPNGGYFTI
SGGV DMLKHGGLALYIMTREAYPDINSKLSAAEICS LTEIEKYRPFKTSGLVQKNSPFKEIFNRKFTLMRETGIMSRG
KLLWNPRKQECNWKQDALVVGVEPLTVAYAVLSLGTILSFVIFLIEISAEMKQKKFTRKNSSSRLFKKNQF

>RPRC017346_Ir75p_Rprolixus

MIFPKEYFLIIFLNYANLGMASSNAPVKFVRNYFLFKNINSVNLICCSKEESFSLIDYLSKNSSFLT MHKYTSVVNST
NYWFYNSMKNTAIQRIGIYVNLNCEFS AEMLEKHFFTYYSAGNYSWIFWSENYYEGFKKIENIKIQWNSDIFLAIPGI
RMYKIHRIHSTAISLPIYASWNESSEFNFYQVNNIHNKGLLVRAFITIKNWQNETNVSESKLLNRYVDDPSEILPK
LSYTVHNLLAEIYNFSWNLSLSFNKLSNLYLNQND AEGQFLALQSGLDVGLNQLTSKINAVDVIPSICKFRQGFV
FRHPIIFTEATDLLRPFSTKVWLCLTITIFLCAIGLRIIFLHRSNEEICESLSISILDSVGVLCQQGISSPSMKTSA
RIAHVSLLLLGSWVFAFYNA AIVILLSPYQRYINSVDLLMKKQFVLGAEDLPYNRYYLGMYNDTRAQM VYDLKLKGE
TPEKDKFYKLSDGISMLENGRFAFFSHITNVYSKVGHFKTADFCSLTEIKVYEPFQVSGFVRKSLPYKELFIQGLILL
KETGIIITRENNYWLNNKPFCTGAVKSI AFNMHSLFLAYAVYAFGFLISLLILLLELLCKYKQMY

>RPRC000007_Ir75n_Rprolixus

MSLMQYCTKLLLLVSYIFASFGLEFKLISNYFKTQHLSVITACVCNRKDAIFLARNLNAESFRSKVHTPQSNDNLEFK
ARQYEQGIYIDLTC EYNQKLMETEYNSMNTSFKWLAWLGKASFNC LDGSFIGADTDMVVATEAGMLYEAYRISLAHPA
TVNLIGNWTSSGEFTIGRPFPRDLSGATLRGSVYVYELKVGNFNSNIILNKMIEPDLDIFHRYNFALEEEELRFIYKY
SLQPVFANATGHLLPNGRFIGMVGDIAENRSDIGLSSAQPMATRQNVVRFLTTLQOMKMI FVMNPKTVGTYKALLP
LSEG V WATLAGITVATVLCYHF IQRYSMNRQDDYEISAGVVYAIGLVAQQGIEGKT VTFSSRLMSVMLLFLAII LSAC
YNASIMNGLLTPAPIKIKNVVQLVD SHFDLGMVDLPYLTMDKNYELYFSKKALKKIFSTKDSIVDIQSGLERVKNSKF
AFAAEDVLIRTE TNRLFN NIEKCDIKELEITRFPLAMPIQKNCTYRELFTRGLLWIREVGLSKRLKDHWYEPAPFCLG
NPQFINVTLEAVSVALLIFLFGLLLSMFI FLGENVVMKRKMKRKGIIKVTRIKQKFNL

>RPRC000589_Ir25a_Rprolixus

MDYIRRNRKLGVIDGNVNVVAGDDAKAILEQLCEVLNSSV TAGKAPYLVLDTSLTGVPSESVKTFSAALALPTVSA

SFGQEGDIKQWRNIDGEKLFILQISPPADI IPEIIRNIVTMQNITGAGILFDDSFVMDHXYKSLLLNVAAARHIIAHV
GNDREIRTQLLRFRELDIVNFFVLGRLTMIKAVLDLANNKYFGRKYAWYAITQDKGQLKCCTNATILHVKPEPEAAS
RERLDNLRATFNLVEEPEITSIFYFDFFLHAILAIKGLTENELLPKDFNRRTTCDYKEYQEIIRKDLDLRKYLKAVTE
PFSYAPFIIIGANGHSYFIFIMKLEKVSIVNSQSESVESLGTWKAGLSSHISCKDPVAMSNMSAVTIYRVVTVKQHPFV
IEYQDGGGETKFKGYCIDLIEEIRS FVNFEYEIYVAPDNSFGMDENGWNGMIKELIEKRAEIALGSLSVMAERENVV
DFTVPYYDLVGITILMKKQITPTSLFKFLT VLENDVWLCILAAYFFTSFLMWIFDRWSPYSYQNNREKYKDDEEKREF
NLKECLWFECMTSLTPQGGGEAPKNLSGRLVAATWWLFGFII IASYTANLAAFLTVSRLDTPIESLDDLAKQYKI QYAP
INGSSSMTYFKRMADIEERFYEIWKDMSLNDSLNPVERAKLAVWDYPVSDKYTKMWQAMKEAQPPATLEQALKRVMDS
KSSTEGFAFLGDATDIKYQVLTNCEFQMVGDEF SRKPYAIAVQQGSPLKDQFNNAI LQLLNKRKLEKLKEKWWTENPE
RKKCNKQEEQSDGISIHNIGGVFIVIFVVGIGLACITLGLEYYWYKFKKPTTNQPPPAAGIGLKALTNEQHRQAALGKL
SVTGLSDYTNQQFRARNHIR

>RPRC017353_Ir75j_Rprolixus

MYIPHFIKMLIVIVLLFFNNQADCRMIVESGSKNAIMYKVIGEFIVGSSAGTHLYFC SRKDATNILKHFTDKHTINSR
VLNFKLSSNEIIASFLQPQITEIFLDLTCRSGIQFLVQS FQHFNSSYRWLLWTNNYNQ TIDHLNPARILVDSAVTVAV
ASEDETIVYDLYKMHYSTPFTVTKLKLWPATGTSINFGMPKPRNNFMGTVIKASVVARFLNYSQDNI IELMLDPSYMP
GMDAKRRFHLRMFIYLAEHYNFTPSYKFARTWGIPTKNGSFIMGIGQLQRNETEISAVYTIFSVPRSAVIDYVLAVTY
AKGSILFLQP KLLGTYKALVVPLQPIVWVAILVISIVTII SFKVVGRCHELWSDSTLMVVATIAQQGGPD MENLSKRL
WHLAALNFAFFITLYYNI AILSGLLLPAPNAIQNM DQLLASNLKFGVETEPYIRNDAESNDELSVRLRAKMAKSKLSN
KVYFDSDEGLELVRRGGFAFYNLVTD SYAKIRDKYTEAEVCSLTEIDKFSKTHIGPIVQNNGPYKELFRRSILLMKEN
GMIKNQFKFWVASKPECIWKQEKFTIGTRALTIAFSVLLIGL FVSGIVFILEVRNFRKKRKQVVFYSKELSLLSQKIL
P

>RPRC0000383_Ir40a_Rprolixus

MECSALSKILSTLPTSNAALIVDDEQDEQWLN FVSDNLFKSGILTVLYRLHSDSTNNSYLQAFTDDIKLRTSGFLDTV
YVFNVDTNAAETILQLVMTENVARRNILLVFLWQDLFVSRRFRDFLQEGMRVCV VINNRYGVFQVLYNQARPDGGKGL
DLVNWWSHTTESMFHSPLLPPAIKTYQNFNGRI FKPVLHKPPWNFVIYSNITNKILVQGGRRDDKLLRLLSNKLNFKY
EYFDPDRSQGSAIVNGTMQGV LGLIWQRDVE MFIGDLTVTYERSLAVEFSFLTLTDSEVFLTHSPGR LNEALALIRP
FQWQVWPAVVTTVLLAGPLLYALVWSANGWTSTRSNLCDCIWITTTIFLRQSI PQAYTPSRVRLLMVTLYLVSTYVIG
DMYSANLTSMLARPAREKPI SNLEQLYQVMKYSD FQLLVEKHSASHAVLENGTGLYKSIWEKMKSQPIYLIESTE EGM
LLVREQKNVALIGGRETFLYETRFRGAHHFHLSEKLYTRYSAIALQIGCPFVHNFDILMRLFEAGVLTKMTQEEYRK
LEDTRPSNEEKKDVGEGRGQAESE DTHRMAASMEMLQGAFFLLLAGYL VAGLILFAEWCSIGKFKPVKLSIEKKPDL
YKKSCM

>RPRC17348_Ir75b_Rprolixus

MSIIATSTKHGWI IWFYFVGIVKCANVDLVDQQRQI IFNYFELTN IETINLITCWNIREKYAMLQESSEKGLRVSF SF
ESMKPWPEFRRGVLDLGC PFTEYYLTRISA EKMFGIHTEWLLIDVTDNTDDNHVEGQIYESNGVLRASNSYSLPGSC
VRLVQIMTNSKMAFYFNMYRVMTREPMQFQLDWSGS IDDLPSLLTPCSNYNRSNFGGATLNAAVVVEFPHLFEGFDSM
LHMDVDIWPKSHYPITILFAEQ LNFRINF TVVPSFGWL VNGTDFDGMIGMLQREEAEIGATGSFMRGDRMNAAQYTTD
TFEPRSAVLFKQPPLSSAYNIFLLPFSRLVWIACAILSLLVSLILAIEFVLTKYKSVRAYEQQNV ELADLITMVL SAA
CQQGT ELIPISMPARITVFLFSLAAFFLYTSYSANIVALLQSTAPVLKSLEDITNSHLGIKIQDLRYNKKFFKETADT
NIINLFQRKVEPFGDEVYCLPEEGMRHMRSGKFAFLVEANVGYKII SNTFEEMEKCGLGEMKMFFNPKLSIPLVKRSG
HREHITRMITKQRETGLLERIRLYWLP HKPLCESKSSVFISVGIADFLPALHVFLYGAAITVGVFIFESAF CYGKKIY
VLKKSQ LKNVRKVI PSYKNVH

>RPRC002763_Ir8a_Rprolixus

MEEKRRIIVQQFQHLERRIQQNWEENKV FELDAPKIKEPTENKLFVTFPFPYMNGRLHLGH SFSLSKCEFAVRYNRLK
GKRCLFPFAFHCSG IPIKACADKLKELKEYGFPPQFPTDIPLQADGYTDEGEITKDKSKGKKVDWRRTFITTDVNPY
FSSFVQWQFIRLRDAQVKYGTRYTIFCERD GQPCMDHDRASGEGVGPQEYTVILLKLLGPYPSKLS ELRGRPIYLAA
TSHRLETLYGQTNCWIHPNLSYICFESKFGMC ICTKRAARNMAFQGLAMEYKGIGHYIELKGEIALLCICV

>RPRC002763_Ir107_Rprolixus

MKFTAISVGAGIDGSFMLELSRHLNFTPVII KAKNYDEAMSKMLAKAVGMSVNAWSMRFLNEHVSMT HAMYSDQK CVA
LRKGSVLRGWHVFLQTFRWDVWLAI ICTAIITNWIVALVTRRRS WYEAMPTVLRAMFTVPLRRPPKSTKERII IASCL
LFGIVIMTTYQGNLYYFIKTKVKHKPPTTLSEIRQEIRGRHLPNDTTLTNRVFEKYAIRIRRKIFD LLMQTGQLSNLY
LVPECLYTANFAYALRKGSVWLAPLNRFLLSMF EAGFPEAWYRRTVFTKKRISYKGKRKT VKSARVFTLRDLEVAIFV
LLGLTL SFVFLLEILSATVASRNLF LRWKLKIRHDYVN

>RPRC017352_Ir75l_Rprolixus

MDSLLQMMNMVNISLFLTIYLVHLPFCSASSMFVRDFFISRSINTIEIHACTTND AIKLVKDLTNDG FQVILARHNR

HEENYYRRNHYYKHGIYIDTNCKAGVDLIYMESSLNSFNSSFCWLLYHQNYKSIISTLNQLQLGMTSDVALAVPFDSI
YILLYDLYKVLAQQQLISQLIGRWSQSEGLINFALPRRTYNMGRTEIKAGIYFLEKVYHLENVVNELQDDRYNAEHD
ANRFTYAIISLHIADLYNFSINLVIEDKAWGTKQEDGHFDGLMGLLEREEVNLITITAMQFMSARLDVSDFIFFSHYFNF
PFIFKQPRALGTYRALSLLQTPAWICVGCMLLISIFIYIWIVQNVYLRANDPDGEATWSGAVIQVIGVVTLOGLSQEP
YRIAGRIACLFILLALMISVYYNAATMNALLSPAPVSIRNLQQLISSNMPLFLMDIPYFNTRTTTTRERLLSKEVMHRI
LSQRQPDFIASLDEGLEIIRTRGAFLGEDFTLYTAIENRFQDSDKCDLSAISALPNVPAGSYIKKYSHLKEILFRGNL
RMKEHGLINRQGLHWFPPKPTCGSQQMLLLVLDLVVTVIAVSIMICGIVVSITLIFLECFEFARINNFYKNDGKIEVKQ

>RPRC000846_Ir106_Rhodniusprolixus

MSNLCTLVMIYLSLSVSKGSDDLYKAMDDVADVHEDRIASLEKIFSITLSGFLCMQVIIDGAPFSAALLKDIHQNTNL
RPIIVSDTGTMEMRCKAFVVISQNIYKVPOLIDGLPLYAENRLVIFVKSHRYLVTVQFAMKKAHIFGNIDVFIPTPDNY
TVGFRMGQTLRKFNLIEDLSGEWTDQENQMDFHGRTIKVLTYNCSMFASKIGPPDSEGKPSFEAGAEMIIFKEIAARLN
LTYKIYTPKSYVNKLNVNPLFPVDKWEKLTLEDPGDNGTDLAFCCLLNSFKRRKSELTMSESLMLLCFRFLVPRPRPIQ
TEWYGIFKFPFANDLWFTIAISIGLTTFFVLQRVTLSSRILETQGISRYYGLGGSFLLQVLSILVLSSIPSSTEHQGPSR
YLLTCAALFTVITTSFSSSLSSNLATQLYSPKLDNIDDFIREQYAWIYVKKPDTEEMFNVANKDHARWSVRAKYAKN
FDSYNALLRKGYILWGAQVENGFLVPLEQLIDKKILRKFEASSECFIEYFSGFAFKVGSFLGPNNLIQKYKESGL
ISKVLIDLNKKRYKKYRYVRRVFHFTKAKLSGPQGIPLNQLIGIFEIWAAGTFLAIVVFILEYFQKFKL

>RPRC17347_Ir75c_Rhodniusprolixus

MNYLFLFLFTALLSLHLVHSLPGSLIINFFDQTHIDSITIISCVSLGRKLSALREITAGGLQVSFSWKTIDPTPKQRR
GAVLDLDCPKINTIIAEINKEKIFNTHTEWLVINRINSKSSYSKKGNYLSEGVKAISASYALPGSCVRVQLILDGTG
EAYYFDMYRVSMRPMKFQLDWAGPYESSLPNEMTQCAERNRSNLEGIFLKGSSSVVHKELIENFDKIHRDDVPEVDF
WAMVHYPIAKLIAEQNSVLEMVLMDDYGYLTNGSFDGLMGLLQORDEVDAATGLIIRHSRFIAVQFTAETFELKTAV
ISKQPPSSASNIFLLPFSQLVWIIACCLLSLMVTAILILEYTL SKHMPQPYSPQYTTMADLITMVLSTICQQGTALA
PVSLSARITVFLFSLFAFFCFTSYSANIVALLQSSAPILKNLNDITNSPLGVKVQNLTYNRNFNRSABAEDVINLYKK
KVVPQGEKAYCEPIDGMKLMRTGKFAFSVETNHAYKIIISDTFEEHEKCGLSEIRIFFIPKLAVGVVRRSGYREQITRK
VTWQREIGLLSRIKSVWLPQKPHCESGTSTFISVTIEDFIPALLVFAYGAVAAFSILAMENVHYFRIKYFSNAIATDI
KKIQLNPEVMIIRGKHKGK

>RPRC000314_Ir75k_Rprolixus

MNIVWSILILCTIFMVSFTISSPISALITDFFKLKHVGILYLQTCDTYNATQFMKFLSDEFIFSMNINSPRSLANIVQ
KSDQRIGIFLDLFCEESIQLFLKATSLHLFTITYEWLILLKSESDFSLLENTGIGMSSEVTTVISSQSILYDLHRPDV
SSLLIKTPVGRWSDLYGLDYNRTPRVLDLEGETIIGSMFFIQDVYDHKNISQQMQDWSFNKNDDVNRFSYVLHIHIA
DIYNFKLELLVTDGWAQKPDGHYSIGIIGHLETRASNVSVAPLLILPHRAAIADVLTPMSDYWKPFVFKQPRALGTYR
ALFLPFEPATWTVAFILFFAALICYKFSINENHRPTFSENIILFIGIIAQQGLTRDPKTIGERLMCFMTLIMGAMVV
LYYNAATMSALLSSTPPKINKLDQLVDSNVPLTLMADAPYNNKEVRNRLLDKNVRDKVDKYSKTTYGLEEGLQIILKG
GAFSAEYFAIFEGIDKKFKDSQKCDLTYIPAYKPLYVSSYFPPKSHIKELYNRGFLRLMEHGLLRREKEKWIWNKPKC
AGKEQFLFVSLEAITIALATLGLGVTTISIVVLIFEIYAYSKSKKLLNVKRRKNLHTESSLSLILPVIN

>RPRC017345_Ir75o_Rprolixus

MKFIFSDFILIFTYFLFKFNHFGNCHGIPVETIKSFFSFKNVGAANIFPCTNRDGVDFLLQMSIRSDKNSLLTGVLK
AQYQKLDFYNIFKSEKHPLGTGIDLRCESTKYVFNKASMELMINMSHVFLLYSQNYDEIMDYLNQNTSIAFSDSVSMT
RGRIYDLYRIQNDMALIVNEIGSWARSSTTLRTPPIRDNLQGSFIKASMLVTRLEVNGTCAEQQLLDNTFKPQLDVLS
RNIYANNKHIESIYNFSYLINITNDWGYLQKDGHFGKGSMDALEEKTVEVGLTSLRFFKERYIVSFFLPASSKYRTC
FIFKHPSTLASYSAFAMPFTLPSWMCTFATIMLSGLCLRMIRNFNIKDKPPNEKSFSASVLVVLGIFCQQGFYPPDSFR
APVRILCVIMEVTSVMVYSYGAEVVGFLLSPSPKFLNTVDKLIESPLKMYSENISYHRSYFTGNVSEKITRAYERAI
KGPTIEQDKFINLEKGLEEVVRKGTMALYQDNTMYTVIGKTFNSPEKCSLFEIEMIRLLGSPVIRKKS PFKELMFRGT
MRVLESGLQYRETRKWYAQRPECIGQQESNAVTLEATLLAQVLYASGLIIAIVLFFMEKQHYFLNKNKAKLKRPRIMV
TPSVVGIGNIH

>RPRC017355_Ir75h_Rprolixus

MKVKDFLFAIISVVYWLDSFGTGLPSSIPHIISKHFLEIQGFQLYVCEKRDATKIMRAFYNKPRLLSSKLLTTANSTR
DTQLHFMQPPVLGIFVDLSCEKGIRFLNESSTKFNASYNWLFWTTNYNLLTLLRPTRVRTDSNVTIALSGAGDTIRL
LDLFRMHISLPFTVTEVELWYSNGTYIHLKDPPEKKNFQRAVIPATLLTSMDDYDLNNSIDPLLDQAYMTGVDGAQRY
GMSTFLTIADEVNITYSYTMAKSWGYPQKDGSWNGMMKHLDSGISSEVGFPAQLNFGRAPIIDYLMTFYLFKCTFTFK
QPKLFGSYKALIVPINNNAWICIFIVVILGAILRTMTVHDKSEYTNNDNWSGSTFLVLAALSIOGMPDNPDKNPTRLF
YFFFLFFTYFIFLYYGISILNGLLLPAPNAIQNMDQLLESDIRIGIQDLPHYHFELENWNTWTVGVRKKLQALKPPH
KTFYGTVEGLQEIKKGHFALYNFAGQFYDRVPEVLTNAEIASLTEIEKFRPRHTGALVMDNCPYKEIFKRKLLFMREV
GIIDRERTYWMRKPYPANWELDALSIGIKPLSLAYGIIIMGAFIAIVFWLAEMLIKRRKSRNNLLTRPRIAFT

>RPRC002124_Ir75f_Rprolixus

MWKLVLFTLLLNQMLPLAKSENILPSMVKIIKMHFREREGTNLFFCQKIDALMMLKNIVNAARSYNVNHLRKDNMTKR
LQITFMTPQIMGIFLDLTCSEGQYFLKENSDFFNASFRWLLYTIKPAETAELLQPTRFMTDSDVTLASPNGSDIILQD
VYRMFKTLPIISTDIQRWSEEGTVTQLADPPPRVDFQGVTLTASLLMVDTVINKSEEIPLLYDPSLSDDGKSWYGLTL
FTHMAEYYNYTYRFIFSDVWGYPLKNGSWSGMVGQVKRHEADMGLSAAHFSVARTKVIDFVGIIDITRVCFAFLQPKL
FGSYKALILPLDVTWVACLVTAVAGVVVFRIVANYDDTSQQNDSWSGSFLLVLAASQQGIPDSTDKLPTRIIYLSL
LGVSFLAAVYYNTAILNGLLLPAPNAIQNMDQLLASDMKIATSDTKYLLYDMTHNDTLTVKLRKKMEKSKLPNGGILT
VSGGVDLVKKGGLALYILTRESYQEINYKFTAAEICSLEIEKYRPLPTGGIVQNNSPFKEIFNRKFTLMRETGIMSR
GKLLWNPRKPECDWKQDALVIGAEPLSMAYAVLSLGTIVAIVLVFIEVHTFRKQKKLLRRNHSSKFFMHRQPH

>RPRC0000247_Ir75m_Rprolixus

MLVVQYFLILLLLICSTPLTFCVEVKIIRNYFKFQSIYVVTAYVCNRKDALFLFQNLISIEKFRCLKVHATNWNNNLDFK
ARFPFQGFIDLTCDVNQKFIINEYKPMSTGFKWFAWLGNESFNCLGGSFIGADTMVVATEAGMLYEAYRISLAHPA
TVNLIGNWTSSGEFTIGRPFPRDLSGATLRGAVFFYHLKVDKTNYSVDILNKMIEPDLDFVQRFHYALEEDLRFIYKY
NLQPVFTNATGIHLNPKFTGEIGDIAYNKSDIGLSPVRYLAMRAKYLSYLPDLQVFRIIFGFMNPKTVGTYKALLP
LSEGVMATLAGITVATVLCYHFIQRYSMNRRDDYDISAGVVYAIGLVAQQGLVGHSVTISFRIMSVLSLLLALLLSNY
YNASIMYGLLSEAPTIIKTLDQLIDSHLDLGLINLPYLTLKENKDIYLSKKLQKKIYNTKNSVLDLQPLQVRVTSS
AFAVEDGLFRIEANNMFTNREKCDLKELELTRFPIISLALQKNSSYVELFTRGLLWIRETGISKRLTDHWYKPAFCLA
TTEFVHVPLEAVAVALLIFIAGVSLISIFILLGELILIKPQQNNKVSDDVIVN

>RPRC000469_Ir76b_Rprolixus

MARSSLINFLILGLCSNYKSDYLKNTTEVDFSCCLKESHEIDKDVLRGKTLNVLTLDAMPLNGAKIQAKNVYEGFGIT
FEILRTLQOKYGFTYNLTKLIPPVIGDSKNGALGKLVSRVDMVA AFLPITRNSHEFVKWGIDLWHTSYFVLMQRPLD
SSNSGGLAPFDMHVWILILVSLISVGPIMYWIMWLVRVKLCQDDDAQLYLSLACVWFVYGALMKQGSTLNPMSDSTRM
LFATWWLFILILTAFYTANLTAFLTLSLAKLPINDISDVAKSSNKWFSSDGGAIYYAVKDNDGDLNILT KSIDNGRGL
FVDVSREATVTSRLSDGWLYLDDWYTLTRLMFN DYKKRKYLDNPNMESESCTYVLT EKHVLERPLAFAYQMSVLP
FDKVVRRFIESGIIKHLITLDPDAVICPLNLGSKDRRLRNSDLFTTYMVVMTGFVAGLSVFLGEILWRLTRGPLIAD
ALQPELLTAMQQNGQNYFTIDKKFMPLRTPSSMILHYG

>RPRC000105_Ir75d_Rprolixus

MSEIKFNKSNLLEPLLDKSYKYGIDGICRYGLAVFLHLSDMCNFTYSFTLNNKWGVQLNNGTWNMGIGQLQLRQTEVG
LAPYKYDVQATKAIDYVTTINTFRSCFTFLQTKMFGTNKALIRPLDYSVWACLLIITIAAVLMFRLVASYEGDTMYND
NLGGSVLIVIAAFTQQGLPDSDEKASTRLIYLSLLIISFFTVTYYNTAILNGLLLKAPNAIHNIEQ LLESELEMGLLD
RPLLREALRKNNTITKEVRKKMLMGEPESEHYSSIVESVHRIKNGQFALFTKDEEVYAE LLNQLSDAEVCTLSEVQVS
NPFHVGALAVDDSPYKEIFSRKFTLMKERGVLD RQQQYWREQKPECHWRQDPLIVSTGPLALAF LILIVGMAATPFVL
MLEMSIHKINRIKSIAPVKPF GKTD

>RPRC008446_Ir93a_Rprolixus

MAVPYSIALYGIRNEFVSERKANDTLLILVDPFFMYPKFRYFENKLKEIVNSRIHKFLDDGSLSVIYNGRDLKSKEDL
TAIFSITSCEEMWNLYSNFTGNGIIFITITEPDCPRLPQHVGTTLPVYERGSEISQLILDLSKEKLDWQSVTIVHDN
SISDKLVEKITLAVTKSLPITNSTCAISLYKIESSKNDVDVKRNKEIFNTISSLPSLEINRNFLILAEVDIIPVYYES
AKSVGLVDPTSKWLFIGMKTD FSNHNNINKFIHIVGDGENVAFIYNSTDDTGLCLNLLCHAEELVGNLAVALDYSIE
EEIRLSEQVSDEEWEVIKPTKQERREAILNFMKNKQDDIGTCDNCTLWYFKSSESWSGMDYFHKGNASLLEVGYWAPKP
GPVLVDELFPNIVHGFRGRSIP IATFHYPWPQVIKYDDVGKPT EYKGLVFEIINELSNLNF TYDVIIISNRTVLKSI
TNSLKIDEKLGEVSLDGRIETSAWKQALKLLENKRVLIAAAAFVTVEDRKKEVNFTYSSISIEAYAF LVSRPKELSRAL
LFILPFSSDTWLCIIGAILLMTPLL CFVHRISPFDYDHYSHRGKG GFTKMMNCFWYLYGALLQQGGGIMPEADSGRLVI
GTWWLVVLVVVTTYSGNLVAF LTFPKMDKII SNVDQLMERRESLSWGMPEISTLHSILKSTDNSKLNALSDGAKLH
LTPEIISDIQNGKHIYIDRKTI LAFVMKQEFIRTNRCDFSLGEEEFLEEHLAMALPVHTPYLKI FNSRIYEMHKVGLI
QKWLVDYLPKRDKCWDAKLSGESNTHTVNMDDMQGSFLLFLGVSLGIMLIIAEYLYHKWKTTQEKHVIQPYVS

>RPRC000379_Ir103_Rprolixus

MIACLAEP LTLTSLNANYSLINQDNLANIFAYA FERHYCAQISFEQSEYASFLMKLIHETRTSPTIAIGDQNEKLSCP
YHIIFTLTLKALPYLLSKTKIADRALLCLVATASF KDFQLFIKEEAVKYFGQSSGSLVLTNPDIVRYSLISNRFI
KENIGNPKWMSDVVR FEDLNGLELTVSTFNCSIFSQIGPLDSNGHPKWIYGV EIDFLHLLSNRLNF TYRIVIPTDEK
RGFINEQGMWSGGLLKL VGEKQVDLAFCGILLTDY LKMSSIEFAPPITKICHKFLVPRPVLLTNQVFGIFKPF RTHLW
AFITATTLFMALCFRWIADISREIPTIVSRQNSVYRKYPKSLFIVWAILFGNF PKSLVTKGPLRHIVTWWALFCMLMS
ACFASSLISHLTTPTYSEKMDTLKELVEQEYFWTTNTMLDFRGIFNLEDQWQREWIKKFFLSADEENQLIKQAKMA
VFGGQLDDAFYVYDDICPSC LRAYEVSSQCISSHTTGFILPNGSPYKNIFTHFVRIFQERGLFIKKFRDISDKEALKK
QYIREVVVPK KRTIVWRQQLELDNII GIFYAFALGILVSVLVFLEWTI

>RPRC001826_Ir21a_Rprolixus

MMNIKIFITIFFLSTSLVMSEARGGRENVTIASAAQYLFNRNVIVRRTCTSEYGVYPYLVPLLSEIVQTYMSTNVVVLL
YDEYFYFSSSQFQLILDNLLKMSISLRHGKLDTSLATPRAPPGFMQVNNNERLAFIVFSKIDIDVSGEAVVGNVLSDTKV
VLISQSSSYQVNQYLSKKSARRIVNLLVIVDMTLKVDLYAQKRKNKIGECDIMLYTHKMYADALGN SHPMIITAWRRD
SFTRTNITLFPDKFKGGLNGLHLIVSASNKPPFVFRKRCPDVSEPSQLTTETGLEVRLIKLLGKIMNFTVEFKEPSFE
ANDGEDTADVVLNELLEGRVQVALGGMYSRYQLFSFIFPHIQDCASFISLTSTALPKYRAIMGPFLLWDVWLALTI
VYLLAIVPISFSVWHTFTPTITNDFKELENMFYVFGTFTNCFSSFSGTNSWSRGDKTATRFFIGTYWVFTIIITACYTG
SIIAFITLPVFPEVIDTTRQFVEEKYKLSTIRNGIWHSLVNESSDPVALQLLPDIDLVS DYIEGLRNVSESVKRHRKS
VFLGSKHLLLEYVVRTNYTPEDTSKRLLFHISKECFVPLMVSLAFPPGSYFIEHMNVALENLIQAGFLAKIIREVDWII
YRSSTGNLLQERSTGLRTPPEDRELTLDDTQGMFLLLGAGFAIAVAALTGEIVACVWKETRAPPEFGKPTLGERILER
IYKWAYCKGCLLSPIEVQERLSRLFLDRRESTENRNLSELFGEVDREMNDRIYNNRIEIMSTIHLQAPRTEQPIR
LRSF

>RPRC000489_Ir41b_Rprolixus

MGQVIYVFI MLTNALISIKVAAYFILPNKEEDNCKNMEILTKQVITTYL SHKCIALITTKEMFQDFNTDLGTSFMQI
TPQEKCDQGLVDIIAEGFKQNC DGIYQVEDPLCFLKIWH TSLQMTFHRLNPKFFFLPDCKNKRNHAMEILSAQESDI
TVNIVVAEIDFNDTNKSKFGLNH HEDWPITLWTNDFSKLASSPRRTKIFVDSWHPKIGFRFNNNIFYDKILNLEGRAM
RVATFTYSPYTIVGFIEGIEMRVLLFCRLYNCTIQIVYDDALWGTIQGNGSGDGILGMVYKDEADVGFAGIYLWIDP
ASYLDYSATWLYAAVTLVPKVPVPLSGWRVPFLPFDLFMWSAIFLSLAISSIFIYSASWISSRRNHNANGRELEGIKD
ITLMVLGMVYQPPSKNLEDNHVSINQIVTSVAISC VLLTSCYSAGLSSH LTVPQFTKPIDTFDELSTSGLYWLAIHE
AFVYSLLPVEDEKVQTVAKHFKKLPEKRLLAEAPKQKYGLALERLPAGHFAEQSHFTDEVISNSHVMKENLFGSLLIF
ALRKGSPTTYRLNKLNVNILLSSGIMYYWEGAVSRKFLNTRKQMALMEARNAHSETGP IILKLSHVQGAFI IYGIGLLL
AICTIAVEIFLFKKYNTPTITIRIYPYTY

>RPRC000538_Ir105_Rprolixus

MLRFISCFYFLLIFVCVNSSELFLNLSLADITHNTNIKLILLNKELWRPFVYSFHGSNCIQVTDDNSIFANIILKTLHN
VHQSPALFTANIQEVQNRCHSHLVLAKKLDKLP SLLVQIPRWAADRILCL IYKTTDNEFHRLKELLVAHLTVRGCVY
CLLPLVFAKSVGTYLSKGHILLSSYLPKYDGIYKFSGREIVVSTTNFSIYTRIWPSNDSSRPAELGGIELNLFQDMTK
RMNLTYKLRLSMEGIGVKENGRWKTGILGDIAEQRADVAIRGLWVSM DKNVSGIVMSSSFRAIKMFLVPRPKRLMGH
WYGVFELFQKYLWSVIAITLIVYTLILYATWKLTKSNIRGKDMKSLERAFFQLYGMLVLGSWPKVNPLQTSQRHLFMW
WAIFALLISTSFSSSLVSKLTAPTYSRSVNTIEDFLQEGLTWTLRKPF SFYIVFNNND DTHKRWLSQIEYINCTEEFN
QILQKDDKQVILGYEINDSFHLFDDAIQEDLLEKYVTTNERLIEFFIAFALREGSIFNQAINRYIRYFVESGLVQHIT
DLEIRHKAIFKQSVYKLRKPERIIEQDFKSLTLEELKGIFTILIIIGYSVAFLIVLFERLHKS KVIHTSI

>RPRC017356_Ir75g_Rprolixus

MSEIKFNKSNLLEPLLDKSYKYGIDGICRYGLAVFLHLSDMCNFTYSFALNNKWGVQLNNGTWNMGIGQLQHRQTEMG
LAPYKYDVQATKAIDYVTSINTFRSCFTFLQTKMFGTNKALIRPLDYSVWACLLIITIAAVLMFRLVASYEGDTMYND
NLGGSVLIVIAAFTQQGLPDSDEKASTRLIYLSLLIISFFT VTYNTAILNGLLLQTPNAINNIEQLLESELEFGLLD
RPLLREALRKNDTVTKEVRKKLMLGEPSEHVYSSIVESVHRIKNGQFALFTKDEEVYAELLNQLSDAEVCTLSEVQVS
NPYHVGALAVDDSPYKEIFSRKFTLMKERGVLD RQHQQYWREQKPECHWRQDPLIVSTGPLALAF LILIVGMAAAPFVL
MLEMSIHKINRNKSIVPVKLF EKID

>RPRC000191_Ir41a_Rprolixus

MNNMPKLIFSFTILSYVIGSISVKVATCFILPNEEDNNCKNIEILTKQVLSTYYLSHKCTVLITMMEMFHEFNTDLGT
SFIQIAPQEKCDYGLANKIAEGFKQNCDSYIYQVDDPLCFLKIWHISLQMTFHRHNPIFFFLPDCKNRRNHAMEILSA
PESDITVNIIVVELEFNDNNKFKLNHSEDWPVTLWTNDFSKLAGSPERTKIFVDSWHPKIGFRFNNNLFYDKILNLKG
RVLRVAAFTYPPYTIVEKNTIDGIEMRIAF LFCRLYNCTIETVYDDGLWGTILGNGSGDGILGMVYKNKADIGFGGIY
LWLEPAYYLDYSDTWLYAAATLLVPKVPVPLGGWRVPFLPFDLAMWTA VFLSLAISSFCISFASWLINRKRYNV DVYKL
LGINDTILMILGMSVLQPPSRQLSDNHVSINQMVTSVAIMCLLLTSCYSAGLSSYLTVPQFTNPIDTFDELSASGLYW
LAIHEAWVYSLQQVEDEKVQAVVKHFRKLSEERLLSEAPKGYGLAVERLPAGHFAEQSHLTDDVISNSHIMKENLFG
SLLTTALRKGSPTTYHLNKL MNVLLNSGIMYYWEGAVSRDFL NTRKQMALMEARNARSKTEPTNLQLSHVQGAFI IYG
IGVLLASCTIVIELYRSKKYKTSTKIRIHPYTF

>RPRC000328_Ir68a_Rprolixus

MLTHLARKLSKKVNCLTLITDSYNNFVFSNEMFHTFHGIPYFKVLVKDYEDLLAPNFQTLTVIQQS QLKGCNAYIILI
ANADQTRAFRLRFGDRHRVLDTRAMYIVIYDPRLFHRELYYLWKKVNVNIFIRKFQEVPRYEFSTVPFPASINSKFVPQ
RIDTWRGGKFQDKELFTDKTYNLKYNTLRIAVFDHMP SVVRIATPHSDENSQLANS PAKFGGLEVQILQSLSKAMNF
QTVLYEPMNADVEAWGRKQLNGTFSGILGEITTGKCDMALGNFLYTPYNLKLIDL SRPYITQCFTFLTPESTTDNSWK
TLILPFKLYMWIAIILTLITCSFIFYGLAKFHTSLNQYPTQKQKADRAFWKDLLRLWKDGKSKNTMQIMSVNNLTNHG

EENGPKGLYL FQKMEDSLMNAYSMMLLVSLPKMPTGWSLRTL TGWWWIYCLLVTVAYRASMTAILAHPAPKITIDTLE
ELATSSIGCGWGEENKFFTTSLDVAGQVRGKKFDVVDADA EAIMKVAEGGYAYYENIFYLLHAKVKQKLIETIV
KDLHIMGDCVINMPISLGLQKNSPLKPRVDVFLQRVIEAGLVKKWLGDMVKTTAAEAPLNKAEIKALMSLKKFSGAL
VALLSGYVIGILVLIGEKIYWYFIMRDPMYDKYSKTLNFGKIT

>XP_026683431_1_GR1_Dcitri

MPHHTSSAPRRHSPTDDVILDDFQILSLEKDQKPMWSVTTNNKTGKLEGHGIAFEILNTLSAKYEFSEYIILPSRKT
LLDQDGSIMNILKTGEVDMVAAFVPVLNIKKNKMAWGIELGQSQYVVMKRPHE SATGSGLLAPFEIEVWFLILISLV
LIGPIMHAIMRVRRWLSGTQSELYPISACTWFVYGALMKQGSTLNPRTDSTRIMFASWWIFIMILTAFTYTANLTAFLT
LSQFTLPIHSVEDIANPTNTWCAEAGGAVEYSIKVGRKGP

>XP_026680610_1_GRionotropic_kainate3-like_Dcitri

MFRPWFYFLLLSLFAQSALSPLPVIKIGGLFDSSPIDAVNEKVFMKAIERNVDEELLGETKLVGLTMRVPELMSFQVE
RSVCELIKKGCIAIFGPRSAMTTPIVESVCDTKEIPHIQTRWDHATPRGSCSVNLYPDANVLAHALVDVVRSA GWKHF
TIVYYDNDGLFRVKLLLEMYNNKGHTIIVRKLSDDG NFRPTLRRIKQQKQTFILIDVPTDILHEVLKQAQQIGLMSSE
QNFIINNLDMHTVDLEPFMYSGTNISGIRMVNNTTSPNLLKAFDGFEDATPDTIALETVLIYDGVFLLAQSLNIPMEP
EPVNVRRERADKDEEEAKEDKGEGEPAPDGEKSDASPAEETSSGDEEPPTCENDIAWPHGYTIINDIKLNTLPRGL
SGPIRFDHQGFRTSFQLELTRLTEEGIQGIAKWSPKERIKFYNEGPTGDFSNTTFRIVIALTPPYGMEKESAARLTGN
DRYEGYAVDLIHELSSLCKFNYTLEVQADSKGGSWDNKTQRWNGMIGKVL YGEADLAITDLTITRDRESVVDFTMPFM
NLGKTFVTTDALYESG

>XP_026679576_11_GRionotropic_kainate2-like_Dcitri

MHSAMESSRPSVFVKSNEGVERVVKKEGKYAFFMESTGIEYEVEKNCDLMQVGGLLDSKGYGIAMPTNSPHRSTVSG
AVLRLQESGKLRDLKNKHWEHNQTVKCEKDEDETDNPELGIDNVGGVFVVL IAGCGVSFLFSLLEFLWNVRKVAVEE
KLSPYEA FMLELKFAIRCYGTSKPVRRAKASTPSLKSNSLANLDLFNNKMNMINYSTR

>XP_026688648_1_GRionotropic_NMDA2B-like_Dcitri

MTTNQFAAMLFQWTVLLLLLIEPNAICLQLAERNRSHITLGRKGSSIKLGSTDNKIYNTSSRYIRTSTTETPTPATR
PPRVTTSKQTLTSTIVKLFFPPVKPVTTMFPVHSHRLHEKLHLGLVVPHTSFGKREYTKAYKSALDELKRSNKVKFN
FFKSHEISVHMAMNSLTPSPTAILKSICQEFLSVNVSAIYLMMNYEKYGRSTASAQYFLQLAGYLGIPVIAWNADNSG
LERRASQSSLQLQLAPSLEHQTAAMLSILERYKWHQFSIVTSQIAGHDDFIQAVRERISYMTDRFKFTLINAVLFTKA
SDLEDLVHSESERVVLLYATREEAVRILSIANDLKITGENYIWVVTQSVIENLQTPSQFPVGMGLGVHFDTSASLVQGI
STAIKVYAYGVEDYMSDSSNANLSLNTQLSCDGVGESRWKTGDRFYRLRNVSVEGDTGKPNVEFTADGVLKAAELKI
MNLRPNLGSQLKWEIEIGVWKSWEKEGLDIKDIVWPGNSHTPPQGVPEKFHLKITFLEEAPYIQMSPPDPVTGKCNMNR
GVICRVAGDADMDKIDVAMAHHNESFYQCCSGFCIDLLGKFAEELGFTYELVRVEDGKWTQENGKWNGLIADLVNRK
TDMALTSLVINSRRESVVDFTVPIMETGIAIVVAKRTGII SPTAFLEPFDTASWMLVGVAIQASAFITFFFEWLSPS
GFDMKTTALQPSHRFSLFRTYWLWAVL FQA AVHIDTPKGFTAKFMTNMWAMFAVVFLAIYTANLAAFMITREEFHEF
TGVD DVRI

>XP_017305182_1_GR1-like_Dcitri

MLTHNQNTTTRRFELQAYVDVINTADAFKLSRLICNQFSRGVFSMLGAVSPDSFDTFHSYSNTFQMPFITPWFPEKVL
APSSGFLDYAISM RP DYHQAIIDTVKYYGWKNIIMYDSDHGLLRLQQIYQSLKPGSDSFQVVTVRRIQNVSE AIDFL
HDLELLDRWGFKHIVLDCATDMAKAI VVSHVRKVTLGKRTYHYLLSGLAQ SALMYDAVLVIVETFNKLLKRKPD L FKA
NVNRRGQMMNNGSRMDCNTSKGWVTPWEHGDKISKNLRKVEIEGLTGDIRFNND SKRMNYTLHV VEMTVNSAMVKVEI
EGLTGDIRFNND SKRMNYTLHV VEMTVNSAMVKGNDKYEGFCKDLADLLAQRLKISSENQ GIEPILLFAEGSDYVSDQ
VYRAYLEEMICGSILDGSVQHIALVRQILDGVDRDLADQSDVLYGT VKDASTYHFFETATFD TYRKMYFYFMQMHKDKV
FVSSETEGV NKVRSSHGKYAFLIESPRNDYENARQPCDTMKVGNNLDVKGFGVATPIGSPLKNMKFGGT VYF GNPNN
AGELSRLQKKWIDTSQCKSSKSQDTHSELALSNVAGVFYILGAGLLLAMAVALVEFCYNTHMEASNNKVPVSDVMKN
KARMTIGAREFDNGRVSQMYASGNTLNPLEDQVHSNTHTQV

>XP_008486318_1_GRU1-like_Dcitri

MNFTPIITEPPDGNRYGWRSPDGNFTGMLGELINSRAEMGFNSIFIKNYHTTDITFTTAFD VDDVCLVVPKSDLKPEW
QGIFLAFDETTWLLIVFTYFVCVIVWYYIKSVIYDSFDHALLVLNVFQAFILAPLHQVPKY YSEKMIFLSILLFTLVI
SNSFQGS LVRLLTHPIEPDINSLADLDESGLPIRSTANNLRETFTYDNGNAIYRNLFNKVFVTPDKSENILTRISY
RRDAAGIVRKNIATQAMKNYIKDDKVLLHIVPECPISYNLALMFKKQSLYIHVINDIILLVQSGILNSYYITNTSSI
IHVYNQHKHHQIFTINNVYIAFII LMLGYTLALMVFLFEYCCVHISNRRCPMYCTTCNALETGDTKATNKDSLKSKDD
NLYLPFIL

>XP_022200717_1_GRI_kainate1-like_Nlugens

MSAKVAFVRPKNGRAPVSSGLFHPADDKQEVALRYAVERINADRSILPRSRLLAQIEKIP PQDSFHASKRVCHLLRS
GVAVRSDSQSRANVL

>XP_022200714_1_GRI_kainate2isoformX2_Nlucens

MEIPHLETRWDYRLRRESCLVNLYPHPTTLSKAYVDLVQAWGWSFTIIYENNEGLVRLQELLKAHGPSEYPITVRQL
GEGSDYRPLLKQIKNSAESHVLDLDCSTDRIYNVLKQAQQIGMMSDYHSYLITSLDLHSVDLEEFKYGGTNITAFRLVD
PEKPELQVRVVDWINGEQKLGRKLDLITHSQYGRDNITFIKTETALMYDAVHLFAKALHDLDTSQRIDIKPLSCDALD
TWPBGYSILINYMKIVEMRGLTGVIKFDNQGFNSFVLDIVELNKEGLSKIGTWNSSSEGVNFTRTYGEVYTQIVESLQN
KTFIVTTILSSPYCMRKDSSEKLTGNAQFEGYSIDLIHEISRILGFNYTFRLVPDGRYGSFNRETKEWDGMIKELLEQ
RADLAIADLTITYDREQAVDFTMPFPMNLGISILYRKPIKQPPNLFSSFLSPLSLDVWIYMATAYLGVSVLLFILARFTP
YEWNNPHPCNPNPDHLENKFSLFNCLWFAIGSLMQQSCDFLPKFSPEYWDNPHPCNDEPDVLENQFSLNLSLWFTIGS
LMQQGSDIAPKAVSTRMVAGMWWFFTLIMISSYTANLAAFLTVERMDSPIESAEDLAKQTKIKYGALRGGSTAFFRD
SNFSTYQRMWSFMESARPSVFTSSNVEGVERVTGKGGSYAFLMESTSIEYVIERNCDLTQVGGLLDSKGYGIAMPPNS
PYRTAISGAVLKLQEEGKLHLKTRWWKEKRGGSRCDDTSKSSSAANELGLANVGGVFVVLMMGMGVACVIAVCEFEV
WKSARKVAVEERESSLCSEMVTELKFAFCCGGTTRPAQKKPPPKSAMTGADQRFMPLGSYPQYGFITKEPLA

>XP_022200713_1_GRI_kainate2isoformX1_Nlucens

MEIPHLETRWDYRLRRESCLVNLYPHPTTLSKAYVDLVQAWGWSFTIIYENNEGLVRLQELLKAHGPSEYPITVRQL
GEGSDYRPLLKQIKNSAESHVLDLDCSTDRIYNVLKQAQQIGMMSDYHSYLITSLDLHSVDLEEFKYGGTNITAFRLVD
PEKPELQVRVVDWINGEQKLGRKLDLITHSQYGRDNITFIKTETALMYDAVHLFAKALHDLDTSQRIDIKPLSCDALD
TWPBGYSILINYMKIVEMRGLTGVIKFDNQGFNSFVLDIVELNKEGLSKIGTWNSSSEGVNFTRTYGEVYTQIVESLQN
KTFIVTTILSSPYCMRKDSSEKLTGNAQFEGYSIDLIHEISRILGFNYTFRLVPDGRYGSFNRETKEWDGMIKELLEQ
RADLAIADLTITYDREQAVDFTMPFPMNLGISILYRKPIKQPPNLFSSFLSPLSLDVWIYMATAYLGVSVLLFILARFSP
YEWNNPHPCNDEPDVLENQFSLNLSLWFTIGSLMQQGSDIAPKAVSTRMVAGMWWFFTLIMISSYTANLAAFLTVERM
DSPIESAEDLAKQTKIKYGALRGGSTAFFRDSNFSTYQRMWSFMESARPSVFTSSNVEGVERVTGKGGSYAFLMEST
SIEYVIERNCDLTQVGGLLDSKGYGIAMPPNSPYRTAISGAVLKLQEEGKLHLKTRWWKEKRGGSRCDDTSKSSSA
ANELGLANVGGVFVVLMMGMGVACVIAVCEFEVWKSARKVAVEERVSSPLT

>XP_022200313_1_GRI_kainate2-like_Nlucens

MGEGLRLAVLDLPSAEELQRPGLWLEAALRVHTYLAVAVDLDCPRAHALLSESSSLGLFNSTSFYWLLLRDDASLHPEQV
IDSNLSLSLGLDTELTAVNASSGRPRLYDVYQTAPPPWGRTFAPAMNLPKGPVNWNPRTWKLDGRTNLHGLVLNATI
VIQKYGNNGWRPSESEIEIYLTNYEHNYDTMTKFCYAVFLYATQLFNFRARMFPTPSWGKYKNGTDFDGMVSLLMNGT
ANVGITTAGIRTVRLPVIDFISSQTWLNAPKFI LRHPRTARSTLNI FLHPLSPGVWFSIASIVVVTVAGTHFIGWVQV
KLSQLNVTDNWSSCLILVVGAI CQQGSSEVPKQ MAGRIFFLSLLVFTFVVYQFYTCISIVSTLLTEPPKTIHQLSDIM
RTHMTIGVEDILTARDTLYRFQADPEIRKMYLTKIKPKKEDAFHTVEEGVALMKRGGFAFYVDPAAAYKLIEDTFSEQ
EKCDLTEVWMHEPDVHTLPIQKRSPYKKLFNFGMRLALEYGVVKRERRKWHSSRPECVLAKNHRSHVSI GLIEFRPAL
LLLAAGMMLAFFVLVHEIMHHRGAMRHARQRPRYRGHDERMDKAGRGFRAGGYTRSTRDDAAGIVISGDYADERGRR
GFEGERGAMDTYRGARRADADVLEARLGLAKMRRSSLLPAVDPNP

>XP_022187223_1_GRI_kainate4-like_Nlucens

MNDITHDEGRFGIKLGNGSWTGAIKELINGKSDVFYCGLWINLQRAEAIGFTSPLLQOEITFFVPLERGLQWSWISLF
LPFTATVWVIAFTVFLLTFLVLYVAAHITKYCPSLQRISFLPTPHNSKGLVRFNLRSI

>XP_022187829_1_IR25a_Nlucens

MFNKPMSCKQLSIGASLVLDFTWTGWRRGREL CENIGVPYLKAAKPVFPLRLVDDTLTSLRNATDSILIFENDLQLQ
QSLYYLVEFSRLRVYAVNGLNASWVKTISSMLPIPSNLVVIAQTNHINSIFKQAMRAKLIQMDDRWNLVFTDFNYERF
HRGILSVATTLYTMKKKACCPALQLGENCVCPTTEVESAFNLNLLPKLIVDIFEKLTNSLIYTGPIVADCNIKTNAT
NPLWTAIFYKYISKVSSSENGVGYARKMMELYPRLEFQVVMTTSRQGYLPEVPTADSKLATWDETFGLQTAPGLIVNPVK
RFFRIGVAESIPWMFPVINPKTGSKLLDDEGKEVWEGYCLDFIQRLAKEMNFDL FELVKADSFGERAPDGTWDGLVGD
SRGKTDIVVASLTMTSREEVVDVAPYFEQAGFTIVIRKPVKRTSLFKFMTVLRLEVWLSILGSLILTGIMIWL LDK
YSPYSAQNKKLYPYPCRHF TLKESFWFALTSFTPQGGGEAPKALSGRTLVAAYWLFVVLMLATFTANLAAFLTVERM
QSPVQNL DQLARQSRIN YTVTLSQDAHKYFQNMKNAEDTLYRMWKEITLDSSTVQTQYRVWDYPIKEQYGHILQAIER
TGPVPNSTVG IQK VLESEQGEFALIHDSSELRYRITQNCNLTEIGDIFAEQPF AIAVQQGSQ LNEEISRRILDLQKDR
YFESLSAKYWNSSNRGACHNIDNEGITLES LGGVF IATLCGLALAMVTL CGEVIYHRRKQGGQQTNSSINPENLKI
HFGILT KKEFGTTAIMREKSQIINPSVTYISVFP RSQLY

>XP_022203339_1_IR21a-like_Nlucens

MLVEEGAILDHCNSTSGTLPVVS LNMNHIADKYLND CFTTVLYDDYFYKSHRFNVIIEQLLKTYPPYSVHGRLLITGRS
MKVKGDI IKKRREQKCN NFVLLTKELEMVSDVLGKQPESRIIIVPLTSLWQLSSFLTSPISQNLVNMLILFDP SLTVE
GLEARNNSKDITIYTHLLFTDSMGATVPKVMVTVWRCDRLTQPKLNLFPNKFQSGFNGYRFRLSGAQQPPFVFRKKNQQ
SKSNADERITTWDGYEVRLQLVSEALNFTFKIEEPADSGNASDRRVIKEVRNLRADIAIAGIPATADRMNEFSFSTV
HSYDCANFITLTSTALPKYRAILGPFRLEVWICLILILIYLF AIIPI SFSARHTLKHLLLED PWEIENMFYVVF GTFTNCF

TFIGRGSWTSKGHNATRLFVGTYWVFTIIITAAYTGCIIAFITLPAYPPVLDTAQQLLKANYRIGILNGSGWQAAAFEN
SSDKVTDRLKKADFVPNLALGLKNVSKAVYWKPYAFLGSKEEMKFMINNTNYTPSSRHKKSFLHLSKECFVPLMVSAV
FSKDAIYKHPLDVTIMRVIQSGLFMKLSRDVEWDQYRASTGKLMQVSSSNLKLGAPEDRQLTLDDTQGMFLLLGAGVL
LAFVSFFTECAVAVRRRCKAWLHQALDDVSTSRNVSVVLDAQTESISDYLGLHSENRCSSQIDDSAPLSKHPIISLSL
MSLEESISPHGSARRFSL

>XP_022197172_1_IR40a_Nlucens

MTTIHKLEDDDDKSKLQVDIKKHMTSHKRVVYLIIGRKKTLENVMMIKKEDLIQRNIIYIFLWNKFPPSRNFSAAII
EAVQVCVFKTRAGNFNFYNNRKFPPNGKGKLELVNWWWSQRDKLFVHPLLPQTEDVYRDFHKRRFKIPVVNKPWFVVS
YSDNNVIQSVRGRDKKLLDLISQKLNFKLVHFDPPDHSQGSIAVNGSMQGVLTGIWKREVEMFIGDLTVTYERSQAVD
FTFPTLVDNEVFLTHAPGYLSEAVAILRPFHWKVWPPVIVTFLVSGPILFFFMNRNSHARKRKERIVKKDKGKLLCDCI
WLTSAILFKQSCSLVRTNSTQVRVVLTLYLGLATYVIGDMYSANLTSMLARPARETPISSLEQLAGAMEKKGIQLLAE
HSASHGNLENGTGLYRRIWNTMKRQPKYLITTAERGIELVKDNKHAFALIGGRETFYNNTRYAGILAKITEDEYQKLI
EDKNANDNKKVETVAEGRKAKEEQDGSQPRAMSDTLSGAFLVLFIFYCASALLLMEIAYFKSKRKKIQSIVERGR
NGRANSRAMIHTKYNICD

>XP_022190951_1_IR21a-like_Nlucens

MKVKGDIKKRREQKCNFVLLTKELEMVSDVLGKQPESRIIIVPLTSLWQLSSFLTSPISQNLVNMLILFDPSTLVE
GLEARNNSKDITIYTHLLFTDSMGATVPKVMTVWRCDRLTQPKLNLFPNKFQSGFNGYRFLSGAQPPFVFRKKNQQ
SKSNADERITTWDGYEVRLQLVSEALNFTFKIEEPADSGNASDRRVIKEVRNLADIAIAGIPATADRMNEFSFSTV
HSYDCANFITLTSTALPKYRAILGPFRLEVWICLILYLFALIPISFSARHTLKHLLLEDPWEIENMFYVFGTFTNCF
TFIGRGSWTSKGHNATRLFVGTYWVFTIIITAAYTGCIIAFITLPAYPPVLDTAQQLLKANYRIGILNGSGWQAAAFEN
SSDKVTDRLKKADFVPNLALGLKNVSKAVYWKPYAFLGSKEEMKFMINNTNYTPSSRHKKSFLHLSKECFVPLMVSAV
FSKDAIYKHPLDVTIMRVIQSGLFMKLSRDVEWDQYRASTGKLMQVSSSNLKLGAPEDRQLTLDDTQGMFLLLGAGVL
LAFVSFFTECAVAVRRRCKAWLHQALDDVSTSRNVSVVLDAQTESISDYLGLHSENRSSQIDDSAPLSKHPIISLSL
MSLEESVTPHGSARRFSL

>XP_008180387_1_GRIkainate2_Apisum

MQSAATAVVVFLFLLSGFADCAAEDVERITIVGLFPSEDSVEQMAFELAIHKVNLDPTLSNDVKLEGRVEIVDIDDGYQ
TSKIVCESLESVGGAIFGPAGYESSVIVQSI CDSMEIPHIETHWKMNLRQQPNYYINVPDPVVLVSRGYTAIVRDMDW
TSFTLLYQRDEGLARLQHLIQDYSGLTKLSDTELSAIIKLSENYDYRPMLEKVKKSLESHIVLDCDTDIILTVLGQ
AEDVGLMDDYHSFIITSLDAHTIDYGHLOQFKRNTITAVKLIDPSSPTVTNIMADLEFVQQRMNLMNEVFRAETITVNA
LLMFDAVNVAKALRGIGGTAKAIKAEPNSCANRSITGWSSGFSLINFMRVVETDGLTGKLRFDQDSGYRNYFTLEMVE
LTNTGFKKIGVWDPQKEMSYTRTNQMLDDLNVANMNKTFIVASKITEPYMMLKEDHKNRVGNDDKYEGYVVDLIHMS
EENITYEFKLRNDGNGKKDKKTGKWDGLIGEVHELRLADLAVCDLTIHDDRRTAVDFTTFFMNLGISILFSKPKEPET
NLFSFTQPLSFHVWVYTATAYLGLSIIILYLARITPNEWQNPHPCASEPEELENLSLNLINCLWFSILCQGSVLP
RAFSTRVCASMWFFALIVTQSYTANWTAFLTSSRKESAIKRVEDLDKQSTIKYGCVRGQSTAGFFENSVDNLYQKMY
SVMETYGDTVMMYDNKQGVDRVKKEREAYAFFMESSTIEYEVQRNCDLTEVGTWLDNKAYGIAMPFNAPHRRTAVSMAL
LKLSESGKLMELKDKWWSVSEDKMCPKPKDSAEILDVNEVGGMFVILILGCMGLGFLFSLLEFLWNIRKVAVAEKLSPW
EAFKLELKFVLKCHHSTKPVRRHTTYSDEPSSD

>XP_029346616_1_GRIkainate5-like_Apisum

MRELKDKWWKVQGDNKCEEEDESDELGFTKIGGVFVVLVLGCLIAFMFSILEFLWNIRKVAIEEEITPKEALILEWKF
AMKCDGGVKPLRRRHIDTGTNNSDTS

>XP_029345828_1_GRIkainate1_Apisum

MFVSYDFVVIICLCNVVAQENDNIPLSIVGFFDKENSIEQQTFELSIQKVNIDPKFANVYLKAEIKIIDSTDTYMTGL
KVCETMESGIGAVFGPSHLESSNIVQSLCETMEVPRIDTSWENTPVENFNFYFNPHPEPTLLAKGYTAIVHMDWKSF
TLLYQRPECLQRLQDLIQDYSGRTKPNDKQAAAIISIIQLPEGNNFRPIKDIKKSLEGHIVLDCDADLILTVFKQAKE
VNLLDDYHSFIITSLDAHTVDFSSIVQNLRTNITTVRLIDPLSPFVENIVRDLNLFVQQRMNLMNEPLKADKLTVNAIL
IYDAVNVAKALKGLGMTNKIITEPLQCTDSPFIPWSNGFKLINFMRIETEGLTGLLRFDNKTGHRSYFTLEMVELV
DTGFKKIGLWDPERGMYTTRTSLEMLRDLYAGSKNKTIFIVSTKIVSKNDCLTYSQRADLGVCDLTIHERRSAVDFTM
PFMNLGISILFSKPEEPQTNLFSFTQPLSFQVWIFTATAYLGLSLVLFLLARITPNEWQNPHPCNHPVELENLSL
NCLWFSMGSLCQGSIDILPRAFPTRLCAAMWFFALIMTQSYTANWTAFLTSNRMETIKNVDDLKKGGTGDIKYGC
VTDQSTASFFQNSVDNLYQKMWSVMELNGDSVMVSDNKQGVDRVKKERNHYAFFMESSIEYEVQRNCDLTEVGYWLD
NKAYGIAMPFNAPHRITLVNMAVLKLSSESGALMNLKNRWWSVSDDKRCKDLKKDSAEILDVNEVGGMFVILILGCLIAFL
FSILEFLWNIRKVAVEEKLTLWEAFMVELKFVLKCHGTSKPVRRHVEDSSSDVTTK

>XP_029346343_1_GRIkainate2_Apisum

MCDALDLPHEARLDLEMDPKRFSINLYPKQSLNKAYSDVISFLNWTKMAIIYEHDYDLFRMKDFMKTVPDRKTVEV

YLRHATPETYRSVLNEIKNMEIYNMLVDTKSEHVDHFLRGILQLQMNDYKFHYLFTSFDMEMFDLEDKFYNFVNLTTF
RMVDIEDVGVREVLKHMEQYSTEKNQSINKFMLTQCEPALMYDSVQVLAVGLGALQQSHTLHYPNISCKDEVWPWSNGS
SLRNFINSVEIKGLTGPIQFKNGQRTDFKLDLMLKLQHALAKVGWETPGLGVNITDMSAFHDTAPYNITLVVVTTIVEF
PYVMHHLQKNYTGARFYGFCDLLEIIAKQVGFYILDLPDNKYGAQDPVTLEWNGIVEQLIKHKADLAVGSMTIN
YVRESVIDFTKPFMNLGISILYKVPSNQPAQFFSFMNPLALEVWIYWLVAIVVVSVTLYIVAKFSPYEWEDQTSGLKF
SSPNIATNPQLNRFTLSNSFWFTIGSLMQQGSDISPKAFSTKIIGCVWWSFTLIIVSSYTANLAAFLTVERMVSPIE
NAEDLAAQTNIAYGTLDSTMTTFFRDSMIETYKKMWAYMQKKKPSVFVTSYDDGIERVLQGNFAFLMESTMLDYAVQ
RNCNLTQINGLLDSKGYGIATPMGSPWKDQISLSILDLQEKGEIQMLYNKWWKPPNDMCMVPEESGAGAKTNTLDFKNI
CGVFAVLLVGLIISIMVSMFEFYHYRKKIKKIEQRIGPQRSTWRDLVDELWFLVKCQGSQRQRTVLTHRCEQCHPKVRT
AERRSRDFMEGDRL

>XP_029345195_1_IR25a_Apium

MEDVWKGYCIDLIEKLSKEMNFKYELVVKDKFGSLDPVTNQWNLIGGLVEGELDIVIAALTMTSREEVIDFIAPYF
EQTGISIVIRKPSRKTSLFKFMVTLKPEVWLSIVAALAMTAVMIWVLDKYSYPYSAQNNKTKEYEQFRHFTLVESFWFAL
TSFTPQGGGETPKAISGRVLVAAYWVFWVLMATFTANLAAFLTVERMQTPVQSLQQLARQSRINYSVIDGSDAHFFF
RNMKMAEDILYNVWKEIALNQTNNRKDFRVWDYPIKEEYQILAAIERTGTVPNRSVGYQMVLADNEQGEFALIHDS
IEYEVYNNCNLTEVGEIFAERPYSIAVQQGSLIQEEISRKILDLQKDRFFELLNAKYWNASKVSMCPNADDSEGITLE
SLGGVFIATLVGLLIALITLAFEVVFYFKHKRAKVAEVSVVNNTVHKDKLMYGHELFMTLGRNSNLDQTRWANKIKLD
STTGRLNNALFFRNKFQN

>XP_029345137_1_GRIkainate 3_Apium

MCASVVTLCWSFVMLLLLVGASEDDIVKPSASNKVVKIVALFNENDDSANQLAYNVSAKIVNIMDTVLPKNKIVPRSE
YVNDMYNVTRVVCDSLKLGIAAAFIDGTDDEDIANVRSISTRTQIPFIETHWKTFRHPPDPYAINIYDPDLLAKAI
RDIILDMDWNSFTAIYESPESLIRLKALLMHFDYGHKAPGKSIKIVQIPPTDDFRPLFKELKLSGEKHIILDYEV
MHILGQAEVHFMDYQSFVITNLDAHTLSFSEFHKSANITLIRLINPESQHVRNAVEEIVFNEQKSGRISVLSPTD
MKTKTALLYDAVNFFATSLHGLVATQTMGPMRIACDDIKPVWHGYSLINYMVRMEVNGLTGKMRFDAETGNRNYFKLD
VVRVQESKKHRLGSDPELNMTLTRSVSEINSEFAQSITNKTFIIVVGKLVQPYLMRCNATEKGMDKDEECFEGFAYDL
VEEMAKYNGFKFKFTTNQDYGIMNHKTGKWNGMIGELQSMRADLAICDLTITFDRRNAVDFTTFFMTLGISILYAKPE
KKKPQLFWFLNPLSFSVWMTATAYLGVSLFLFMLARMTPEFELPHPVKPGDDTLENSMTLLNCLWFSIGSVLCAGC
EVLPAKAVSTRLVAGMWWFFALIMTSSYTANLTASITSGRLDTPIKNVDDLKSDSNIEYGCYEEGSTASFFQKSNLSLY
QRMWSVMEASNPTVFTKSNQEGVDRVLKKGKRYAFLMESSIEYQTERNCNLMEIGNTLDSKGYGIAMPNNSPYRTLI
SESVLRLQESGFMRELKDKWVKVQGDNKCEEEDESELGFTKIGGVFVVLVLGCLIAFMFSILEFLWNIRKVAIEEEI
TPKEALILEWKFAMKCDGGVKPLRRRHIDTGTNNSDTS

>XP_016662281_2_IR40a_Apium

MATAERCVPALMLFTVLGTCDVVLVNNVMSVSFTSSEINVAVDRYPPCSGRRTSHVLHSAIDVYAHGKLAVGSNVTI
YINIRLLGLHRDGDFFLTQIRRSNLMsrNVVYVFLWLRSSVSRTFKADILEAMRVCVITSPPRGFYQIYYSQASAR
PGYGSSLKMNWWSAMDGLVRFPPLPPPKQVYKNFEGRYFNPVVLHKPPWTFVEYLNDSTRVEGGRDDKLINLLADKL
HFQFKYIDPPDRTOGSGGLDRGSSMQGVLGLIWQREADWFGDLSITYERNLVVDFSFLTLDNEAFLTHAPGRLNEAF
SLIRPFHWSVWPLLLITVIFAGPILYILVDTTDGHPQGKSMLYWKCVWWSVTVFLQQAIIIPSENNKIRFVAGLLMLS
VTYVIGDMYSASLTSILARPPKEPPINTLKELSEAMRDSGLQLLVEVQSASQAMLENGTGVEELSQLVTRQREYLIG
STEKGMQLVRDNKNYAVIGGRETFYDIKRFGAQHFHLSKLNTRYSAIAFQACPYRDNFDDVLMRLFEGGILSKIT
EEYQKLNDKLMGSEEFDDSTSVVIEPVLEGSEPRQEDDDKQLTIAMSMKTLQAFYVLAIGSILAGLLLLLIEMRSHDK
LEKDKVIKLVLEAPFVYKRKVPNKFNRLYDLKE

>XP_029346904_1_IR93a_Apium

MLLTASGCPGMMILTTVYNIIAHNILITICNGLNIYNIDTGPRNNDTVAAIINPKFTASKNNEIWNTINGHIYKLKQ
DVFKQAGVSVEVYKTTNIKLSKDIVSIFIVSYCEDIWDLYYKAEFEGLNGLLYISITDTCPRLPPTDEAITIPLTVHH
SELSQMILDLRMSNAFSWKSAVLMHDNSIGDSVLQHIIVTSLTKYYPNIMSPSITIFEIYTQGEWKRRKLFMEDLQH
FLKMSEINSNYICIVSILYVPLILDVAKSLNLMATAENSWLIIIPDIDSSRNNTSSFTNLLSEGENISFIYNSTKTGSK
CIGGILCLVDELMSVFIAMAFSALIQOEIELSQRVSEEEWDEIRPSKIDRRQSMVSFIKFRLESQVCETCPLWQIDSG
VTWGQEHFGQGCYILPVGNWNTKTGLKLTEPLFLHLANGFRGIALPIATFNFPWQIVNFNRSGLHIGYSGLVFDIIN
QLAKTLNFTYNVIVISNTEQMNTTRTLFMQNNVLGEHDAVVSKPLWDKMLDLVRSEKVFIAAAFAVKEANQILVNYT
THISLEPHQGILVARPKELSRALLFTAPFTLLTWLCIAIVVGLMGPLLVNFHVLSPYIEYHNIIPKGGLSNPLNCFWYV
YGALLQQGGGHLPDGSGRLVVGTTWVLFVIVITTYSGNLVAYLTFPQMDSMVSNVADLMARKPQGYSWGIPKTSNLH
SLLTTLPPDDTMVKELIKNAEHHEELSRSIIEVRVSGKHAFIHRRTNLMIKNDFLKTNRCDFAIAGNEDFAEEKLAMM
LSKESPYLSRINREIEKMHKVGGLINKWLVDTLPKKDQCWTNTQLEVTNHKVNLDMDQGSFIVLLLVGLSSLSVSVFVEY
ILHKYINRRQIVITPFIN

>XP_029347313_1_IR93a_Apisum

MLVNGEVDMAAAFI PVL PGLDDIVKWGIDLTQFQYVVL MKRPKESATGSGLLAPFEMEVL LILMSLIAVGPIIYGIM
MLRHKL CGHDSGIIYSIPTCVWFVYGALMKQGSSLNPD TDSARLIFATWWIFIMILTSFYTANLTAFLTL SKFTLP IK
RIDDIASNEYRWISSEGS AVEYIVKVDNDLKPLKQSMFEGNGQFMVINIDEIDTILKHISSG TLLLHDKNWNLFMLR
IERSITDDKERCRYVLT TDPYLTRSMSFVYPKHSILPPLFNPIMLSYMESGIVRHLQTKDLPEAVICPLNLGSKERQL
RNSDLYTTYAVVVC GFSMAAAVFTLELLSTRTGWFVAADLARRSHDPLRSGAVRCEISVFGQPSKAMYPF SHRDNMVN
AASATHFPFRGSSSSSTSLPSSSSSPFLLLDGKSSLRRFQY

>XP_029347065_1_IR75a-like_Apisum

MTPYFPNGSAVSSAVYCKALKTLGLCHATGKGLSVFSRTVAKNTFLCTCASYFGVSNNRFLNLQNDWLLFNQKSI SNT
SDTLTIASVVFETYLGKSYVLPDSGVFLFQKIGNNVWEIWSGFRASKLDTIRVYKTGLASKYHIEINDSTEIRMDFRG
ATLKATTVIFNPSEFIGFNQTITSGLDVYTYMHYPMIQILGYQLNFNYTISITDDHGWSNGNGSFSGLIGILQREESD
FGAAGSLMRLDRMTAVDFTVGTVSLESNILFKQPM LSSITNIHIKPFKYEVWQITLYMLIGFVLILLFLNKFKANRSP
SITVLDII GLVHGAICQQGYTIVLILNSIKIVVFLVLTTFFLFNAYSASIVSLFQSTSSSINNVDKDLHEKSMTMSI
QIASYAKPYFNETTNEILRKLYDMKIKPYDENNSFTDAQKGIERIRTEFHGFMVEKTSAYQI INKKWREEEKCGLYEI
QLFKLPVLAIPVVKKSGHKDVFKQKYVFVYVLR

>XP_029343909_1_IR21a_Apisum

MLPSLDTISYKLFNVY EYSSKRTRVSESRVCVNYIIFCQDPEITATKFGNAIDSRVFIVTIGSKWQIQEFFKSPASQNI
MNLIIASAGPALQEKKNIKRSYHVLTVGVKANRQNSIGNTKEILPEVADIKLYTHEMYIDGLGSSAQ TILTTWKLNK
FTRPEVDLYPVKLVNGFRGHRFILSAIENPPLVFRSVDKVL TQEQSTSWDGVEIRLLQIISRVLNFTLGIHDATFSKS
RMLIFFFIRGEGDDRIIGDLVASKADIGISGLYMTNARYSLVDFSPVILQDCGTFVSLG SFALS KYRAIFGPFHWSIW
VMVVITYMVAIFPIAFTNNRNVGTLCSPKQLESMCCYMFGTYNLFTFKDVKSWTNTKMGSTRLFIGTYWIFTIIT
TAYTSSIIAFITLPEQPVIVDKSYQLVDQGYRVM TLDKGGWQHYLNITNDTMSRRLLSNIKLMKNLDDAIDYIVRTRF
ILDYAFLGSKISLTYLSQNNFIQKYVPIKIFLHVASECYVPFNIGIAYKKNFVFRNIFSNFILRAQQSGLVTKI IKDI
EWEITQKSGPNLIIAPEDRQLALDDVQGMFVLLGGGIFLAAFTLSIEFVKCKRENKRKIAQIKIEKHLKKKKRSKSLTV
RGNTTAEPFRPLTAF

>XP_008183092_2_IR25a_ApisumMLTMYKPIRGITILLWINTLFNIGTSQNVQTVNILFINDRTNEVAEDTLNVA
LNYIRRNPVRVGLMIDGLYSVKIGGDDASAILETLCVNYNASIRNNKPPHLVIDTTINGVASEAVKSFTAAALAPT VSA
SYGQTGDIRQWRNL DGEQQKYLIQISPPADLIPEIVRSIVVAQNITNAGIMFDDTFVMDHKYKSL LQNIPTRHIIAAI
DDTTSIKLHLTRFRDVIDVNFVVLGKLSIIKSVLDHANSNKLFG RKYAWHVITQDKGSLKCGCSNATILFVKPEPDAG
SRERLSNLRTTYGLTSTPELKAAFYFD FYRSLLAIRSMMSNGEWPTNVTYTTTCDEYNEENPLPRRNVDLRRYLKDMT
EPPSYAPFLIDTNGHSYEEFTMRLEKVTVLNSQSVSAENVGSWKASLNSPIIVKDAANMTHFSAVTVYRVVTVLQNP F
MIQIDDEDGKGVKFKGYCIDLIEEIRKLIGFEYEIYIAPDNNFGNMDENGQWNGMVKELVEKRADIALGSLSVMAERE
NVVDFTVPYYDLVGITILMKKPQTPTSLFKFLT VLENDVMMCILGAYFFTSFLMWVFD RWSPPSYQNNRIKYKDDEEK
REFNLKECLWFCMTSLTPQGGGEAPKNLSGRLVAATWWLFGFIIIASYTANLAAFLTVSRLDTPVESLDDLSKQYKIQ
YAPLNGSSAMTYFQRMADIETRFYEIWKDMSLNDLSSEVERAKLAVWDYPVSDKYTKMWQAMKEAKLPNTLEEA IERV
QSSKSSSEGFAYLG DATDIRYQVMIDCHLQMVGDEF SRKPYAIAVQQGSPLKDQFNNAI LLLL NKRKLEKLKETWWNL
NPERIQCEKQDNQSDGISIHNIGGVFIVIFVIGIGLACFTLAF EYWWYKYKKSSRVANTMNP KQITMNRGGEFTY PVVP
TFNANS GMRSRSIIHGFRQSIGQSSPKQK

>XP_016658040_1_GRIkainate 2_Apisum

MCDALDLPHEARLDLEMDPKRFSINLYPKQSLLNKAYS DVISFLNWTKMAIIYEHDYDLFRMKDFMKTVPDRKTVEV
YLRHATPETYRSVLNEIKNMEIYNMLVDTKSEHVDHFLRGILQLQMNDYKFHYLFTSFDMEMFDLED FKYNFVNLTTF
RMVDIEDVGVREVLKHMEQYSTEKNQSINKFMLTQCEPALMYDSVQVLAVGLGALQQSHTLHYPNISCKDEVPWSNGS
SLRNFINSVEIKGLTGPIQFKNGQRTDFKLDLMKLKQHALAKVGEWTPGLGVNITDMSAFHDTAPYNITLVVVVTIVEF
PYVMHHLQKNYTGNARFYGFCDLLEIIAKQVGFDYILDLVPDNKYGAQDPVTLEWNGIVEQLIKHKADLAVGSMTIN
YVRESVIDFTKPFMNLGISILYKVPSNQPAQFFSFMNPLALEVWIYWLVA YVVVSVTLYIVAKFSPYEWEDQTS GKLF
SSPNIATNVPQLNRFTLSNSFWFTIGSLMQQGS DISPKAFSTKIIIGCVWWSFTLIIVSSYTANLAAFLTVERMVSPIE
NAEDLAAQTNIAYGTLDSGTTMTFFRDSMIETYKKMWAYMQKKKPSVFVTSYDDGIERVLQGNFAFLMESTMLDYAVQ
RNCNLTQINGLLDSKGYGIATPMGSPWKDQISLSILDLQEKGEIQMLYNKWWKPPNDMCMVPEESGAGAKTNTLDFKNI
CGVFAVLLVGLIISIMVSMFEFYHYRKKIKKIEQRIGPQRSTWRDLVDELWFLVKCQGSRQRTVLTHRCEQCHPKVRT
AERRSRDFMEGDRL

>APZ81419_1_IR25a_Alineolatus

MPSFTARLPGTTSVAKVFTVFMFYVSL LQKVHSQSATSINVMFVTEDRNDIARLAFDVVSDYVKRNSKLGIEMEVFRV
TESGSDAKFLLNLCETFNASAKAGKPPHII LDTSVVGVTSEAVKTF SRALGIPTLSASYGQEGDLRQWRAL EVEIAK
YLLQINPPADIIPEVVR SIVILQNISSAGIVFDDSFVMDHKYKSLLLNVPARHIMGRVRNIQEIRNQLTRFKELDIVN

FFILGSLSTIRNVLNEANGMKFFDRKYAWHAITQDKGQLKCDCSNATILHIKPEPDGSKERLDNLRTSYNLVEEPEI
TSVFYFDFFLRGLLAAKALIEKAPWPKDYNTSCDNYDENHDFIRKDLRLSSLDVKEAYSAPFLISTNGKSFMEF
NMKIEKVIVNSIAESAEIGTWKAGLSNQIQTKDIASMRNFSAVTVYRVVTVKQKPFVIETFENGKPKYSGYCIDLL
EDIRSFVHFEYDIYVAPDNAYGNMDPSGNWNGMIKELIEKRAEIGLGALSMAERENVVDFTVPYYDLVGITIMMKKQ
TTQTSLFKFLTIVLENEVWLCILASYFFTSLLMWVFDWSPYSYQNNREKYKNDEEKREFNLKECLWFCMTSLTPQGGG
EAPKNLSGRLVAATWWLFGFIIIASYTANLAAFLTVSRLDTPIESLDDLAQYKIRYAPINGSEAMTYFQRMADIEER
FYEIWKDMSLNDLSSEVERAKLAVWDYPVSDKYTKMWQAMKEAGLPATLEEALDRVRKSQTTSEGFAFLGDATDIKYL
VLSSCDFQIVGDEFSRKPYAIAVQQGSPLKDQFNAILQLLNKRKLEKLKEKWWTENEDRMQCEKQEEQSDGISIHNI
GGVFIVIFVGIGLACITLGLYWWYKYKKPASPKQVGPMAQIIISTNATNKQLSVTGLMDYNTREPRARYPIRRTAVNA
TQDYSRPTFFAQQERLSHW

>APZ81411_1_IR8a_Alineolatus

MRTCLWFHVVLFAHEVAGQGVKLLVVKDNNAGIWDVSVTSFFEQLPVTVDKEDTNSTINDLCEVLKEGVWGVLDLTW
SGLDEIKAVCNTWGLPYVRLEYGITQYLRGADKSLATIRKAPDAALIFQTEEQDQSLFYLIRESSMRVILFKGLSDK
EAETLTMTMRPTNPNFIIFADTPSMNVMLKAVERNLVRYDDRWILVFLDNEHNSFDRKTLVKRVTLATPTIDANAAAN
FAENVAETLEEVAKTSGIDLSPVPAQCEGSSTAAKDLTVFQEKLEIVEKKPWLDRQQUESTMALHLDMMDWTAESSKG
EKLFIGSWNSKKGLTIAGNVTKIPRFFRVATGYMVPFAYPVIDPSTGAPKLDDKGNEVWEGYCIDLINRLAEDMDFDY
ELTTSYNFGRKLPNGSWDGLIGDLASGRVDIIVAALTMTSREEVIDFVAPYFEQTGFSIVIRKPLRKTSLFKFMTVL
RVEVWFSILAALCLTAFMIWFLDKYSPYSARNNKDKYPYPTREFTLRESFWFAVTSFTPQGGGEAPKSLSARTLVAAAY
WLFVVLMLATFTANLAAFLTVERMQSPVQSLKQLARQSRINYTVVQDSDAHSYFRNMKFAEETLYRVWKEITLNASAN
QSQYRVWDYPIKEQYGHILISMEKTGTVNSTEEGFQKVRENEDAEFALIHDALIEIKYEVYRDCNLTEIGEPFAEQPYS
IAVQQGSHLNEEISRRILDQKDRYFESLSGKYWNSTMKGKCDSSDEDEGITLES LGGVFIATLFLGLVLAMLTGLGIEI
VYERKAKKNVIVKVSAPKEKSEKSEKKEKMMNNPFFNDKLF SREFGSFPKKPSKLLAPKPKVSFITVFPDQLY

>APZ81415_1_IR75d_Alineolatus

MGALLPYSLITQYFINIHVSSIIVVSCCTTSQTAQLLRHLSQRGITASWAVDNLSPLEVRRSGIVLDLSCNQSKIEILH
DMSSRKMFGLMEWLLMSEGSAPEEAELPDLYILPGSSVTL SVTSPSSISFYDYTRITRRLPYKFTLLGAVARDEDVL
PQWKRPSRVNYEQNLLTTVSVIHSLDIRKLTDPDVAEEDRWPAIHFPVVVNVAYQLNFKFDLRLESVHGWKFPNGSFE
GMIGVMEREEDVDFGASGVIMREDRKHVDYTVDYFEFKTGIIFKQPSLSSVSNIYLLPFSRHVWAACGGLLLFVLIIL
CIAVSSGDAQTFTTPATFLDMVNIVLGFVCQQGSYLAPVTISGRIVVFVSSLAALFLYTSYSANIVALLQSTSSVLKT
LKDLTNSHLGLKVQINEYHLGYFLEAVDEDVITLYNKVKVNQPETFVNGTRGVEFMRTGDFAFCEFDLAYKQISKTF
QEEKCGLGEMHLFFVPRLSIPVIKRSRGHREHFTQTIIWQWESGMLDRISRIWLARRPRCESTGGGYLRVGLKDFNPA
LKVILVGIIISIWFFLCELITDRGFKAYYRKIKHNQEKIMGDDGHLIADLCFSPWKMLFKNKR

>APZ81412_1_IR21a_Alineolatus

MKLEVLVFTTLVHLTHSVKITKLLNSMAFADLETISCMAPDGIPYMPVPLINSIAKRYLKDHATVILYDDYFYHPRLK
NMIDILISNYAYPLRHGLVNTTMAKPTVPAGILEARENEQMAFIVFTKESEIGAEAIREFTGHTMTLLIAQTSVYHV
KLFLQTKLAADITNLLVFVDPMIKIDHFVQKTARVLKECDILIFSHKVITDSLGISMPVIVTAWRRNHLTRQVQLFPP
KYKRGLGGLHLVASASEIPPFVFRKHGHDSGAGYTITKWDGIEVHLLYMLSQMLNFTVEYKEPEFNEEDVAQTVIKD
LHTKKTTLAIGGVYLTPERIGGLTFSFPHTQDCASFISLASTALPKYRAIMGPFLLWDVWLALTAVYLLAMFPIAFSVW
HSIKPLNDIREVENMFVYVFGTFTNCFTFTGKNSWSKADKTATKFFIGTYWIFTIIITACYTGSIVAFITLPTYPET
IDSSKQILLEEDYKISLLGSGGWEGLFNDTEDPVASKLYESVERVPNLYSGLRNVTRNVHSWRQSAFLGSRRLLEYTVK
TNFTPDEDSKRLMFHLSDEC FVPLFVSIVMDKRTNYLEEFNNALERIIQSGFMTKIVREVIEWQEYRSASGKLLTMHKG
LKGAPEDRELNLDDTQGMFLLL GAGFGIGLLVLIIEISVWSSEQRKNRQFGELTLKQRAINKLKEHWETLYACLLAPA
NSGIIYFRERRVSSAFGEYVTQPYASWSITSIPSTVEPPSPSPNGEISSAPTVGNAASLSMDQLSFPPEKPIRMMS
F

>APZ81410_1_IR76b_Alineolatus

MSPFVHMLLVAMCANYQTNLLATDNQNFTCILKSEEQIKKEVYKGVKIKILTFDEMPLSGARKDGKGMIGEGVAFEL
VETLKEKFGFDYTVRMAPIVGDESHGALGKLVSRIDMVAAFIPVLPDAHEFVKFGKDLSQAAYYVMLKRPADSNSG
SGLLAPFDTVVWLLILVSLAVVGPVFYGVMMWRDLRCPGDIDQVYPLSTCVWFVYGALMKQGSTLNPLADSARMVFAT
WWIFILILTAFYTANLTAFLTLSISTLPIKEIDDVAKDNRHWFALQGGPIEHAIKDKEDEKLRLRDSAASGRATFLE
TKQESIILQKITNDWYYLDDSYSLTRMMYDDYNRKSMDNAESSLRCAAYVLTEKPFLVRSALAFAYGKDSPLPDLFNPIL
ERFIESGILQHKLNLDLPDAVICPKDLGNKERKLRNADLWTTYLVVFSGVSVAFMIFLIEIIWRFYRKVKGSNQGVFN
KQQMFQSRNLNTDKLLAMRDQVQTKINGRDYYMVTNKGGSNHFIPLRTPSALLFQYG

>APZ81409_1_GIR1_Alineolatus

MIREWMNMFTAALFLACAHVASALPEVIKIGAVFEEGDEWSRYAFEAIAIKIINEDESTLPGIKLEAAIAEPIQYD
IVGVENHVCDMMKAGVVGIVGPHSTEMSNHVQSLCDTMEIPHISARWDGQQRSSCLVNLYPHPSVLAQVAADMVRTW

DWKGFTVLYDDFNALRKIGELVKVADDKGLIVSVRQIKGTEEEESNYRYVLQQVKHSGETNIVVEVSRERLFDVMMQA
QQVGLVGGEYSYIITSLDFQSIDLEPFKWAGTNITGIRIVNPDQPHFKEVMKLIEMKQTQEGEGEGAGEEGGEGEGEK
EEVGENNGDEAEAGADADELGEGEEEGGGKEENKRRKRRTQDDEENEVIGEEEEGAEGEEEEGGGEEVGENGEG
GGEGPDDGAEGNQEEENEGEGEEEEAGGGAPEEAPAEIEEEEYEIPPIEALLIYDAVNLVAEALHNLDIVEPREIDCRQN
NAWESGYSVINFMKSEQAGLTGLVKFDNEGFRSEVVLEIVELVQEGRLVKGNWTKQDGVTHYIGGEAGPTDIGDDL
RNTTFVVLIALTHPYGMLKEDSRQLMGNDRFEGFGIDLIHELMSMSGFNYTFRVQEDKSSGSPKTLENGTRVWSGMIG
EVLAGRADLAIADMTITRERERDVDFTMPFMNLGIIILYRKPMAMPPSLFSFLSPFSYEVWGYILSAYLGVSFLLFIMA
RISPYEWTNPYPCIEEPTLETQFSLSNSLWFTTGSILQQGSDVAPISVSTRMVAAIWWFFTLIMVSSYTANLAAFLT
IEQKIEPFTDVEGLANSEGIKYGAKKGGATANFFRDSNEPIFQKMWTFMEANPDVMPTSNEAGVARVQENTDYAFLME
SASIEYEQERKCELTMGDLLDSKGYGIAMRQNSSYRNVLNRNVIKLQEKGKLTQLRDKWWKEKRGGGACNAEEEGGE
ASELGLDNVGGVFVVLGGCILSVFLAFGELLCDIYGREDKVSFKDELIEEIKFIARCHGTVKPVKKEPANSSSSSS
KSGSRSGSKSGSKTRSSRTRTRSGSGSPSVERHRFSSRLYSDDL

>APZ81408.1_IR93a_Alineolatus

MSTGLFLVQFVLLMIMLKDHENNAIKLHNDKRANDTLIIIDESFVDSLDRNIEQQVRQIVSEISSRILKKGAVDIL
YHSNANIYLEPDVTAVFMSCLDLWTLFNRGKKFDILFISLTEANCPRLPPDSGITVPLYRRGWEIPQIILDLREDG
SLTWETTAIIFDDNLEEDMLKSIIEVLNRPKKSSMACSVVLYRLHSFPKDIQNKKKKNMDNLLNLPIDMVNNFLV
LIDQRKITPTLELVKKMGLVIPTAQWMFVVRNLNVMKRRTSPEKRYVDLIGEGENVAFLINSTRADSRCDMGLLCNAR
QLVEKLIIAIEKSIEQEIILADSLDEEWDVLKPKKLERREHILDFIKKKTRDEASDCDCTEWIIRSSDSWGMDFVQ
SKGSYGNKESNSTTGALLEVGHWPQPSGLVLVDHLFPNIVGGFRGRTPIASVHFPPWQFVKYDEFGQPSEYGGVFN
VLNELADKLNFTYIEIVLLPNGTSAANKFTLHKELGEVVIDSSVEFAAWDQVVLDLKSNKVFLGAVAFVETEERKADVN
FTHPVATDAYAFLVSRPKELSRALLFIQPFGETWLCIIATILLAGPLLWFVHRVTPFYDHYSHRGKGGYTRLYNCFW
YLYGALLQQGGGVMPEADSGRIVIGTWLVLVVTYSGSLVAFLTTPKMDKVISNVDQLLERSAVSGDGMITWSFP
KISTIHRLLKDTDNKKFNMFYEASEKLEQLTPEIIAKIQNGEHVYIQRKTMLLYIMKQEFRLTQRCDYSIGSEEVLEKE
RLGLVVKSNSPYLKIINQHIHDMHKVGLINKWLEDKSLPKDKCWMSTLGSSSSSTHTVNMSDMQGCFFLLFIGVFTSIL
LIGGECFMKWWKLKQKSIIPFIS

>XP_024219750_1_IR25a_Hhalys

MCRLLLLLLLLLEHALSQTLOTINIMFINEAKNQLAEMSFDVVLNLYLKKNPKLGVKVEAAVRVSIISGTDAKAILESICEA
YNGTVSDGKPPHLVLDSTMNNVPSEAVKTFDALALPTISSSFGQEGDLRQWKTLDNEKQKYLIQINPPADIPEIVK
SIVQLQONITNAGILYDDSFEMEYKYKSLINMATRHIIHVHNADSIENQLMRFRNLDIVNFFILGGLSTIKMALDTA
SKKQYFGKKFAWHVITQDKGQLSCSCSNATILYVKPEPEPGMKERLDGLRNSFNLVEEPEITSIFYDFFLHSILAIAIK
NLLDEDDWPKDFNYTLCDDYRADREIVRKEIDLMKNLRVSEPPSYAPFLLRNGNSFPEFVMKLEKVTIVNSQSESA
ESMGTWKAGLISPIILKDATAMNNFSAVTVYRVVTVKQKPFVIEYEEDGRKKYKGYCIDLIDEIKTLVGFDYIEIYVAP
DNQFGNMDENGNNWNGMIKELVDKRAEIALGSLAVMAERENVIDFTVPYYDLVGITILMKKQTTATSLFKFLTVLNDV
WLCILAAAYFFTSFLMWVFDWRSPYSYQNNREKYKDDEEKREFNLKECLWFCMSTLTPQGGGEAPKNLSGRLVAATWWL
FGFIIIASYTANLAAFLTVSRLDQPIESLDDLAKQYKIQYAPLNGSVAMTYFERMANIEKRFYEWKDMSLNDSLSEV
ERAKLAVWDYPVSDKYTKMWQAMKEAKLPNTLEEALDRVRQSKSSSEGFAFLGDATDIRYQQLVNCEFQMGVDEFGRK
PYAIAVQQGSPLKDQFNAILQLLNKRKLEKLKEKWWSENPERQKCEKQDDQSDGISIHNIGGVFIVIFVGIGLACIT
LGFEYWWYKYKRPADGGGGPMVVKPTIGGGRNVEKLSVTGLADFGHHTTFRSRNTHNSNMRRGNLSHIPTSQW

>XP_014272502_1_GRI_kainate2-like_isoformX2_Hhalys

MHCFYCLLLLLIFNGAFSTTIPIAGFFNADKTEQEKVFKLAVENNVNNVQDKLFPEDSLVLDPKVEHVLSFDSYNTSR
RVCDLASEGIGAIIFGLHSPEAVGITESLCDAMDIPLIKTFWDVSAMPMKSAINLYPDPKILSKALVDVIKMDWTSYA
ILYDSDFGLLRLQELLKSHKRAYTTEGHVPFTVWQLEPEGDVRPLLKQIQKTSETRIVLDCATERIEVLKQAKQVKM
LGDYHSYVLTSLDAHTLDYSELQLDTTNITTLRLVNPQRHYVQSTVQDWVFFHKTLLEDKVPVTADTLKTDSALTYDA
VKVYSEAVQKIRRDSQEEVEVTPLLCSQPTKWLNGSLIIDEIKKLEIEGMSGSIKFDESGQRIFFTIDITELTKNGFR
KIGNWDPERGISYTRTRSQMRKEMYQSIENKTFVVSVRGEPYLKEVDSNLEGNARYAGYSMDLIDEIAKDLKFEYKF
YLAPDKAYGSFNKDTKQWNGLIKELRERKADLAICDLTITYERRSAVDFTMPFMNLGISILFSKPTKQPPDLFSFLSP
FSIDVWIYMATAYLGVSLLLYFLARISPMWKNPHPCNKDPEELENTLAIYNAIWHNIGSLMQQGSDIAPQALSTRVV
AGMWWFFVLIMISSYTANLAAFLTMDRMDATIESVEDLANQNKIKYGVKGGSSANFFRDSNVSLYQKIWSQMESARP
SVFTKSNDEGVERVLRGKRAYAFFMESTTIEYQKEKHGCSLMQVGGLLDSKGYGIAMPFNSPYRIASGSVLKMQESGK
LQQLKDKWWKNSDEGENKCPVEEDASSSELGANVGGFVLVLLVGCIAAFFVAILEFLWNVRKVAVEEKISPGDAFFL
ELKFAIQCYGTTKPVKQEEESVAEEDLVECEIEGVVEEEEDRGFFGDENMEEYMLRNGFSNNIRAKSTQSYS

>XP_014272501_1_GRI_kainate2-like_isoformX1_Hhalys

MHCFYCLLLLLIFNGAFSTTIPIAGFFNADKTEQEKVFKLAVENNVNNVQDKLFPEDSLVLDPKVEHVLSFDSYNTSR
RVCDLASEGIGAIIFGLHSPEAVGITESLCDAMDIPLIKTFWDVSAMPMKSAINLYPDPKILSKALVDVIKMDWTSYA

ILYDSDFGLRLQELLKSHKRAYTTEGHVPFTVWQLEPEGDVRPLLKQIQKTSETRIVLDCATERIIEVLKQAKQVKM
LGDYHSYVLTSLDAHTLDYSELQDTTNITTLRLVNPQRHYVQSTVQDWVFFHKTLLEDKVPVTADTLKTDSALTYDA
VKVYSEAVQKIRRDSQEEVEVTPLLCSQPTKWLNGLIIDEIKKLEIEGMSGSIKFDESGQRIFFTIDITELTKNGFR
KIGNWDPERGISYTRTRSQMRKEMYQSI SNKTFVVS RVGEPYLKEVDSNLEGNARYAGYSMDLIDEIAKDLKFHEYKF
YLAPDKAYGSFNKDTKQWNGLIKELRERKADLAICDLTITYERRSAVDFTMPFPMNLGISILFSKPTKQPPDLFSFLSP
FSIDVWIYMATAYLGVSLLLYFLARCTPDEWDNPHPCNPDPTELECIFSLHNCWLWFSIGSLMAQGCDDL PKALSTRVV
AGMWWFFVLIMISSYTANLAAFLTMDRMDATIESVEDLANQNKI KYGV LKGGSSANFFRDSNVSLYQKIWSQMESARP
SVFTKSNDEGVERVLRGKRAYAFFMESTTIEYQKEKHCSLMQVGGLLDSKGYGIAMPFNSPYRI AISGSVLKMQESGK
LQQLKDKWKNSEDGENKCPVEEDASSSELGIANVGGVFLVLLVGCIAAFFVAILEFLWNVRKVAVEEKISPGDAFFL
ELKFAIQCYGTTKPVKRPQEESSVAEEDLVECEPIEGVEEEEEEDRGFFGDENMEEYMLNNGFSNNIRAKSTQSYS

>XP_014289152_1_GRI_kainate2_Hhalys

MLIKIVLITFLFTNSYTSSEDEKNEEMVKTEVRIGALFDKEDETLIKTFTAAVELVNSDEELLPNMTLVPIFLTGIPE
YDSLEVGGKVCELMSYGVAAIFGPQSPFTSYHVQSLCDTMEPHLSTKWDL SQRRSSCLLN IYHPSSLTQAITDIVA
AWNWKGFVTLYDDFDALRKIQGI IKLADDKGYLVTVRQLESDDDNYPVLKEIKHSDETNIVIDCAVEKLENLLIQAQ
QVGIMGSRYNYIIATLDFQTINYEPFMWGGTNI TGIRLVDPDDPYVRNVTQTEERGEEDGMESTTTGLGDISTTEPPI
EEEEEQSSIPTVEAALLHDAVLLFVKALQHLSPLNIKPLECGPHSSFDGYTVINYMRLSEIKGLTGVIKFNHEGYRT
DIQLDVISLTEEGLQKTGTWNTTTGLAMDPPDTNEG FVVDAGEDLRNKS FVVI IALTHYPYNMLKEDSKTLTGNERYEG
FGVDLIHEL SLMSGFN YTFVEQYDKNSGSPTTLKNGSRVWNGMIGEVQAGRADLAIADITITRERERDVDFTHPFMNL
GISILHRKPSKAPPSLFSFLSPFSNDVWGFMLAAYFGVSLLLFVMARLSPNEWTPYPCIEEPEHLENQFSLNLSLWF
TLGSVMQOGSDVAPISVSTRMVASIWWFFTLIMVSSYTANLAAFLTVENNVSPFSDVKELAEQSDIKYGAKKGGATAN
FFRDSKEELYQKIYKFMGDNPD CMMPNNDVGVN RVVSEDKYAFFMESASIEYEVQRKCQLAMVGDLLDSKGYGIVMRQ
NSTYRNALNKNVRLQENGKLSQLKDKWWEKRGGGACTDGAEEGEASELNLNNVGGVFVVLVAGCFIGVFLSFCEVL
WDISQRDEKISFKHELMEEIKFIMKCKGTVPVRKNSLTLCNNLSKSSSKSSSTSKRSSKNSFSFRLP

>XP_014274775_1_IR25a_Hhalys

MWLQDVGHKPRNCKQDNDSSQIDG EIHVVILKEENGTIWENVKTEEVWEEQHVIILNRENEEESFNKVCEELSSGAWMV
LDLTWSGWDSVAEIIAGIRYL RADLGISP FMRATELT LIKLRNSTDAALIFQHSHHFEQSLWYLVRESSLRVTVHGLD
DESAESLLDMRPS PSSYVIFADHDKANTILKKA VD KKLVLVDDR WAMV FMDIGTAPINKTMLKKRIMQVYMPKSI CCT
DPGQMLPCHCTEPFNFEIEGAKQLKIALSKAISSSLTEGLSAEPLNYS CSSSALS PKNDSIFYDNLDKALSDSWLLHK
TYGEVRLNLRFEIRNIGENRAQQ LGNWDPIAGIRTAAMPRVKRFFRIGTGFTSTPFAYPTNQISP DGSARWEGYALDLI
ERLAESMKFEYQLITVENHMFQGRFENG SWNGLVGM LATGRVDMVVGSLTMTS EREEVIDFVAPYFEQTGFSIVLRKP
LKKTSLFKFMTVLRVEVWLSILGALCLTALMIWFLDRYSPYSARNNKEKYPYPCRDFTLKESFWFAVTSFTPQGGGEA
PKLSARTLVAAYWLVVLM LATFTTANLAAFLTVERMQSPVQSLKHLARQSRINYTVVKDSDAHNYFRNMKFAEETLY
KYWKEITLNSSSDQSQRVWDYPIKEQYGHILNAIEKAGPVANI QEGLRKVIESEQA EFALIHDLLELKYHVFKNCNL
TLIGEPFAEQPYAVAVQQGSHLNEEISRRILDQRERFFEATASKYWNSTIKSKCDNVEDEGEITLES LGGVFIATLV
GLFIALLTGAEVWYHKKKSKNQVTMKGQQQLQSIIRDEIFTTKEFGHHNFANRKASLLATKPKVKQITVYPRGQLY

>XP_024214423_1_IR93a_Hhalys

MVKFILLQHLFISIVAYEYYTTPSDRKTTDNL LVIIDPNFLHPQYKGLNNNIRH MVSSIANKYLKGESLLITYTNNLN
YRIKDDITAVFSVIHCEDIWDLFFEQQETSILFIAITEPNC PRLPKDFGITIPLYEPSLEFSQILIDLRTEDFLAWEA
VNLI FDDSIDSKLIDNILEAQTKSIPLTRKSASVSLYRINGTVNKLIRKKEVLNVMQLIPRTEMFFNYLVVNFELVP
TILEIAKLI GLMSPRNRWLFVRLRSDDSKANNVKNYIHLLEEGENVAFLYNFTNINYS CNYSLYCQANELLDHLIVSL
DRSMREEKLLSEQVSEEEWDLIRPTKSQRRNSI INFIKARQKDISWCDNCTWWRIISSETWGSDFLEKKELLES GDWS
PRSGPVLVDQIFPNVAHGFRRKI IPLFTFHNPPWQIVKYDQNGKPSQVKGVIFEVVDHLAKSLNFTY EIIILMSNTSLP
ANQTKFYRFNESVGDVVL DQSTEF LAWEQVVR LIQNKKVLIGAAFTITEKRKKYLNFTLTIRTENYAF LVARPKELS
RALLFIQPFSTDTWQCIVA AVLVM TPLLNFVHRVSPFYEYHSQREKGGYMKMMNCFWYLYGALLQQGGGVMP EANSGR
LVIGTWLVVLVLT TYSGNLVAFLTFPKMDKVISNVDQLMEQRGEVTWGMPE DSTLHIILKSTDNDKLNELSDSAQL
HRMVTQDIVSQIRKGEHVYIDRKSILLYLMKQELLTTNRCSLSIGEEEF LAEHLAMVISPSSPYLELINKQIYKMHQV
GLIDKWLTDYLP TKDRCSNTLSSESQTHTVNLDDMQGSFLLFLGVTLGFILII GELLFKKWKKTQEKQVIHPFVT

>XP_024214586_1_GRIkainate 4-like isoform X2_Hhalys

MKTA IRLQLSAEGFLVSFSLDTVDDLPI TRSGGVLDLTCNFSLSVINQFTRDKTMGVHTEWLLLDGSSLDVIKDAYI
LPGSFVTTIAQILGEFVYLDVYRTSPYRPLKYTVLNDQTELEQFIQLPDRPSRND FEGISLLGAAVLYYPHMF CGFDCR
DHPEVDTIAKTGFPI SQHFQEQLNFTLT FQILNDYGWKTNTFTSGLMGLLQREEIDMGVIGIFMRPDRIGTVDFTGDT
FEIKSLVIYKQPALS AVSNIFVL PFTRMVWVCCVALSVLTGIFLMADVATSFGNRQT FEESVTTSDVVTLIIGCICQQ
GTTLVPKTLSARI IIFLFSLCCLFFYTSYSANIVALLQSSSATFRSLSDLTNSPLGIGIQDVIYNKIFFGEATDENVR
ELYDKKIAPQGPKAYLNPVEGIKKLRSGMFAFDVEIHWGYKII SDTFHENEKCDLDEMRI FL LPKLSIPVVKKSGYRE

HFTRMNTWQRDVGLHSRIRQRWLPKKPICDNTGRGYVSVGLTDFKPAALLVMVYGISFSIAAFFLELFTRSKLFYRCKF
LNKSKRKRKRFNDWDRVM

>XP_014291746_1_IR21a_Hhalys

MSVINFEIKPKLVKIIFFDDLHKADYKFKLNLDKIIKGIHNSAVVFGQKSSAIEAINDNEQPAWIIFSAQPESLPMKGK
ILLIIHVSFVLKEWLLSPKVKKHQQLVVISDPALKKSKGSEKDCDTGLYTQFINPHSFGRGKALILTSWRNHNFTQ
SVKIFPSEEQTLYNGNHLVVYARDQPPFAFKRILENGEVIWDGVEVRLLLLLMKEILNFTMEFQDHIGKQSSDTLQILR
QGNTDLVVGGFVMTKEIYGKTSMIYPHFMDCAAFISLTSIALPKYKAVMGPFLLWDVWISITLCYILAIPIAFSAWHT
LGPLIQHPSEIENMFYVFGTFTNCFTFRGQFSWTKSVKNSTKLFITGYWMFTIIITACYTGSIIAFITLPVYPEVMN
TMHQLLHKNYRVTTMADEGWWTLLIGSDDDVASGLAGTAETVNNVLEGLSTVIKSSKEDKPVTFGLGSSEHLKHILKSN
CSASDVSKRQLFHISRQCFFVPMISMIMPHDSIYIDFSRSLIKAIEAGFINKIQDDLEWNLYKSSARQTLLOGNLKL
EALERQLTLEDTOGMFLLGCGFAFAVCAFGFELGTWIKNRKGEKLTIVVKERILAAASRRVSA AFLTPYKSHIRDHYP
LYFRNVSRKSSLFGSMIILEPSDRQKPPTAVSEIHLMPKVEKPIRLLSF

>XP_024218345_1_GR1-likeisoformX3_Hhalys

MHHKEEPIMLRKPDGSLTGNDRFEGYCKDLADLIAKRLGINYELRIVKDGQYGAENS DVKGGWDGMVGELIRKEA
DMAIASMTITSERERVIDFSKPFMYLGISIMIKKPVKQKPGVFSFLNPLSKEIWVCVIFSYIGVSIVLFIVSRFSPYEW
WRLLSYDDPHHSLHATSSSSNVLANDF SVLNSLW FALGAFMQQGC DISPRISGRIVGSVWWFFTLILISSYTANLA
AFLTVERMVAPINSPEDLASQSEVQYGTLYHGSTWDFFRKSQITLYSKMWEFMSSSRSHVFRVITYDEGIRRVQRSGKY
ALLIESPKNDYINEREPCDTMKVGRNLDSKGFGVATPLGSPLREVVNLAVLSLIETGELAKLKNKWWYDRTECKHGDK
DTSRNELSLSNVAGIFYILIGGLILAMAVALIEFCYKSHSEATRAKIPLSDAMKAKARLTIGGREFDTRKRETLFQYY
SSGNAINQPESDQIHSNTHTQV

>XP_024218340_1_GR1-likeisoformX2_Hhalys

MTVNSAMVKVAEWTDESGFQPVAAKYIRIKQPSDIERNRTYIVTTIVEEPIMLRKPDGSLTGNDRFEGYCKDLAD
LIAKRLGINYELRIVKDGQYGAENS DVKGGWDGMVGELIRKEADMAIASMTITSERERVIDFSKPFMYLGISIMIKK
VKQKPGVFSFLNPLSKEIWVCVIFSYIGVSIVLFIVSRFSPYEWRLLSYDDPHHSLHATSSSSNVLANDF SVLNSLW
FALGAFMQQGC DISPRISGRIVGSVWWFFTLILISSYTANLAAFLTVERMVAPINSPEDLASQSEVQYGTLYHGSTW
DFFRKSQITLYSKMWEFMSSSRSHVFRVITYDEGIRRVQRSGKYALLIESPKNDYINEREPCDTMKVGRNLDSKGFGVA
TPLGSPLREVVNLAVLSLIETGELAKLKNKWWYDRTECKHGDKDTSRNELSLSNVAGIFYILIGGLILAMAVALIEFC
YKSHSEATRAKIPLSDAMKAKARLTIGGREFDTRILDSTAETIK

>XP_024218337_1_GR1-likeisoformX1_Hhalys

MTVNSAMVKVAEWTDESGFQPVAAKYIRIKQPSDIERNRTYIVTTIVEEPIMLRKPDGSLTGNDRFEGYCKDLAD
LIAKRLGINYELRIVKDGQYGAENS DVKGGWDGMVGELIRKEADMAIASMTITSERERVIDFSKPFMYLGISIMIKK
VKQKPGVFSFLNPLSKEIWVCVIFSYIGVSIVLFIVSRFSPYEWRLLSYDDPHHSLHATSSSSNVLANDF SVLNSLW
FALGAFMQQGC DISPRISGRIVGSVWWFFTLILISSYTANLAAFLTVERMVAPINSPEDLASQSEVQYGTLYHGSTW
DFFRKSQITLYSKMWEFMSSSRSHVFRVITYDEGIRRVQRSGKYALLIESPKNDYINEREPCDTMKVGRNLDSKGFGVA
TPLGSPLREVVNLAVLSLIETGELAKLKNKWWYDRTECKHGDKDTSRNELSLSNVAGIFYILIGGLILAMAVALIEFC
YKSHSEATRAKIPLSDAMKAKARLTIGGREFDTRYSSGNAINQPESDQIHSNTHTQV

>AXU25112_1_IR6_Clividipennis

MSELFQSNIEEVSRCVITNPRHYSYHMYNQASADGQSELTRVNWWTRRKGFYTNLLLPKSSELYKDMKGRMFVIPV
LHKPPWHFVSFENETVHVEGGRDEKLLSILAELKNFRYVYFDPDRSEGSSIINNMPGVLGLIWQGEVDIFIGDLTIT
YERSLAVEFTFFTLSDSEIFLTHAPGKLNEALALVRPFRWEVWPV IIVTVALSGILLFVLQKVR LGWHSVKLSELCD
VWIIITIFLRQSLVTKVHGDRLRTVILLYLVATYVIGDMYSANLTSLLTRPAREKPISTLEQLNDAMTNKNFKLLVA
RRSSSHGVLENGTGLYQIVWNKMKNQNYLLNTVEEGIEQVKHFKTVALLGRETFFYN AKRYGVRHFH LSEKLFTRY
SAIALRIGCPFIDQFNKQILNMFECGILSKITEDEYQKLRKKQLSKTGETPISP AESFIGIGEGGSQSENDET LRAMS
MKTLQGAFFVLIVGCCFGALVLIFEVYTYINIRKR

>AXU25111_1_IR5_Clividipennis

MNNINRERRQILATLFENVRLENLATYCESASDSPFHFTARCLWVAGTRRACLVDINAGYKWISDTFEEAEKCSLSQ
IKLFKTEMLAMPVAKRSGYREIIAQKLCWQQENGLYDKIAATWIPPKPRCVGDSGGFTSVGTTDFLPAILALIYGCSI
ATFILILELLYKHRTHFRSMISGPGKDSDVSTNKRDEMKRVS

>AXU25110_1_IR4_Clividipennis

MEYANRQDTLSERRVAVRVMKSWSEVDRVLYSKLGIFVDLSCLKNAHFLKEADQSRFRLNHTWLWVDPNQYAMANT
PQSPYSEVVVASKDTLHQITKFVDELPELHEVGTFSSNITTLSLTIKPNLNNKVFVSYVMEIYD PANLTNQML
DPDFQPGFDPGNRFYATMLLLGDMYNSMEFFQEYDWGEHLPNGTWTGVIGSIQSGRTDMSISPLFPAERLDISYA
PPVIHNYEFVIAFQQHRKLGTYKAQTMELTPTVWALVGLFALAGGVFTFTFMQGSFIRALPAGMIQVVGSLASQGLC
IDPESASSRIAGLSLMLIGLLLSNYNAATMTALLSEAPPEIKSFDEFLKTDIPINMVDAAYSTSKLGRIHVFPQEAW

PKTGLTEGSPPPKIMTLEDGVKSMKKGDFAFFAESYLIAKGVNEAFTDDDKCSLTIFPTMATVNRAFRFLKKDGHLL
FSGRIIRSMEVGIIISKVRDDWIAKPPSCFRNVFFYQVQLEAVTVAFVVLLGGMILSLLIYFCELI
PCLKTANMKQPST
LTTSVD

>AXU25109_1_IR3_Clividipennis

MPLAFLGWILMILDLFGTSTFGKLDRIRIDDDYFNRKMVYLAREVMNEYMADRRACIIVVSDEGLLDDLKDFRNASFIR
VLFNGTKTKCDPNMHNHLIQAFYHKCPRYVVQISKPTCFFPAWFLARNGSTYEKHNPRVLFPLPIKPHATVYGEEVLAM
NQTNISHDILIAETSDSKEWPITLYTNNFWQYVGEPEGRIGRILYLDEWSWETGFKRNVLDLYPDKVRDLRGKVLRLSAFP
YLPYGNNEPMDGSEARILLEFCVVYNCTVVDVDDGHLWGSIIYPNNGTGVGEAGTIYMELSDFGVGANYLWLEFWPYLE
FSNAYLYGALTVMVPRPELLSGLLTPFLPFPLSLWLILVFTVFMSSAVGLHVVTDATRRFAPHFREEIYHNHKKFITVTD
SLIRSIGMFVLQQPQRLVTGSPVRHLFTAFELIYLVITSAYAAELYDYLTIPRTTKPIDTVEDLADSNLIWMTDHEVW
VFGILHAEDPKIKKAASNFRALPTPELIKLGGETSIGLGIERMAGGHYTELPYITDKFVSI
SRVMRDNYWWSPLVVM
QKGSPIANRLNDIIGRMESGGLYYYWEADCVRKYLN
YTKQLDMIWSTRPIKYPKVLNVADLEGAFILYFIGTAMAVG
LFFLELYFKRVLKLERSFLDPTPKWFD
EYLEEKNKILK

>AXU25108_1_IR2_Clividipennis

MLSPALDGVEIRVMDAFCVAANCQLSPVTDDDLWGELYENGTSIGIVGNVLQDNADVGLGAVYLLWYYEHIEFVFPYMP
SKVTVLLPKPSPLPEWKVPIAPFDWVLWLSLLGSIVLAAIVLFLMNKCLLKYSRSHIPREFESVSGIILRAVGMAVL
QSPENIFTNGSTLRIALTTFQVFFLLYTTIYSSALSSVLTI
PRYYPPI
DNMRDLYLSGLPWAADHIAWVTTLKEADDP
IIKGLLGNFVFDQETLTVLSKKGHGFTIELTNGGHVSEAAFLAADMINNLHTMKQPMYHTYSTMIIRKGS
PYADEL
RKIVHRCIDTGLVQMWESMMLTSYGSSSIQTAFQLSKKSRSNEHTGLVKLKLHTQG
AFFILILGTVTGTIIFFMECF
FHPRRGGFNGYRQDDVGGLGILVGLGYA

>AXU25107_1_IR1_Clividipennis

MPGVLGLIWQGEVDIFIGDLTITYERSLAVEFTFFFTLSDSEIFLTHAPGKLNEALALVRPFRWEVWPV
IIVTVALSGI
LLFVLQKVRLGWHSVKLSELCDSVWII
TTIFLRQSLVTKVHGDRARLTVILLYLVATYVIGDMYSANLTSLLTRPARE
KPISTLEQLNDAMTNKNFKLLVARRSSSHGVLENGTG
LYQIVWNKMKNQNQYLLNTVEEGIEQVKHFKT
VALLGGRET
FFYNAKRYGL

Anexo IV: Secuencias analizadas

Ecdysis

Hormona disparadora de la Ecdysis (ETH)

>TRINITY_DN1204_c0_g1_i1.p1

ATGAGAACAATGTATACAGTTCACAATCGATTTCGTAACAGAAGGGGTGTTCTTGATACTATCAGCCACATTTCCTAAGC
GTGTTTCTCACTGCACAGCGAAGAAAGTATCGACTTGCATTCCAAGTACAACCGACCAGGAGGTCTTCTAGCGACTGCT
GCAGAGGAGCCATCACTACACACAATGCAACGCAGAAACGATTTCTTCTGAAGGTGAACGGAAAAACGGAAACAAA
ATACCAAGAATGGGCAGACGCAACATTGTCTATCCAGAGAAGAAAGAAGAGCTCGAGAGTCCGTTTGTGGTAAGCAGG
AGGAACGATTTCTTCTGAAAGCACACAAATCAATACCGAGAATTGGACGCAGGAACAGCAATAGTGAGCCACAATCA
TCGCTGGCTGATGATAATCAAGCAGCCTACAATCAACGACTTTCAAAGAAATTTGACTCTGAGGGCATGGAGAAAGCT
GAAGCATCAGCGTGGCCTTGGTTTAGAGCGCCTGGCTTGTATGTTCCACAAAAGAGAGCCATCTATTTACAGAAAC
CCAGCTTACTATGGACCAGGTGTAATCATGTGGGATACAGACGAACATCCGTCTCTACTTCTGGAAGATAGTGAAGCA
AATTTCCGAAGTCTTGACTAA

Corazonina (Crz)

>TRINITY_DN69828_c0_g1_i4.p1

CACAGCTCAGTACCACAACAAGGACGATATCAGCTCAAGCATCAAGTACTGTACCTACCTACTAGTTGCACCACAATG
CTACTGACCCGGCTGCCCCGTGCTCTGCTGCTTGTGCTGGGCTGCATGGTGTGCGCCCTGTTTCGCACAGACATTCCAG
TACTCGCGCGGCTGGACCAACGGCCGAAAGCGGGCACTGCTGCAGCCCCAACCCACTGCAGTTGCAGCCCAACCCCTCTG
CAGCTGCAGTCGTCTGCAGCCAGGGCACGCGACATCGACTGCCAGCTGCAGCGGTTGCGAGCGATGCTCGAGGGCCGC
ACCGACCCCGAGCCACCCTGCACTCTCTCCCCCTGTCTGCACCATGTGAACTGTGGGGCACAGGCTGCCGCTCAACAA
CTATAA

Hormona de Eclosión (EH)

>TRINITY_DN10952_c0_g1_i1.p1

GACGGTCTTCAATCGACGTCAAGTGCACCGACGCGTCACAGCGCCACGACGCTTTGCGTCAGGTAGCGTCTCACTTCTT
CTTGGGTCAACATTCTGCGCACACCTCTGGTCTGCTCGACACTTCAAAATGATGTACTCTGGATTGATTTTGACGTTT
TTATTAGTTTCTCGGTCAAGTGCAGCTCAGCACGCATTGGAGTTTGCATAAGGAACTGTGGCCAGTGCAAGCAGATGTTTGGG
CATTACTTTACGGGCAAGCCTGTGCAGAGTCATGCCTTGCTACAAAAGGCTTGAAATCTCCCGATTGCAACAACCCC
AAAACGTTGAGAGGGTTAATAGTAAGACTCTACAGATAA

Péptido Cardioactivo de Crustáceos (CCAP)

>TRINITY_DN7706_c0_g1_i14.p1

CCAGCTTTAAACAGTTGACCGAAGTTGGTAGCGGCTCATTTCTCATCCAGAAGCCGGTGACGCCAAAGATGCAGCTG
CATGTGCGACTGTGTTGTGCAATGGCTGCCCTACTCTGTATTTTGACACAACCTACCAACGGCGATGACGTCATCATA
CAGAAAAGGGTAGACGACTCTGACTTGGCTGCTATATTCGATCCAAAAAGAAAGAAACCTTTCTGCAATGCATTACA
GGTTGCGGAAGGAAACGTTCCGATGAATCGATGGCAACTCTTGTGGACCTGAACTCCGAGCCCGCGTTGAGGAGCTG
AGCAGACAAATACTCTCCGAAGCTAAACTGTGGGAGGCAATCCAGGAGGCGCGCCAGGAGCTCATGAGGAGACGCCAG
CAGGAGGGGAAGCTGCCACCAATTCCGTACTCTTTAACAGGCATGAGAAGAAAGAAGCGATCTCCAGAGAAACAGCTC
ACAAAATAA

Bursicon-alpha (Burs)

>TRINITY_DN5062_c0_g1_i2.p1

ATGACCTGCACGATTCAAATATTTATGCATGGAGCAGTCGTTCTTGCCATGATGCTGGCTGAGAGCTTCCAGAATGAC
GCCAAACCGGCGCGGCGCAGTGTCTTCTGATGAGTGTGAGGTCACACCCGTCATTTCATGTGCTTCAATACCCCGGCTGT
GTTCCAAAACCAATACCCTCATTCGCATGTACCGGAAAATGTAGCAGCTACTTGCAGGTTTCAGGATCGAAAATATGG
CAAATGGAAAGATCATGCATGTGCTGCCAAGAGAGTGGAGAAAGGGAGGCCAGTGTCTCTATTCTGTCCGAAAGCC
AAACCAGGAGAGAAAAAATTTTCGCAAGGTGAATACAAAAGCGCCTCTTGACTGCATGTGTGCTGCCCTGCACTGGGGTA
GAGGAGACTTCAGTAATACCACAAGAAATAGCTGGATATGCTGATGAGGGTCCCATGAACAACCATTTTCAGATTACAC
AATTCTCGATAG

Partner of Bursicon-beta (Pburs)

>TRINITY_DN9187_c0_g1_i1.p1

ATGGAAACACATACCTTAATCCTCCTGCTTTTTCACATCAACAGAAGTATGGTTTTATAACGCTTATGCCGGTGAAGAA
GAAGTGTGTGAGACACTACCTTCAGAAATACATATACCAAAGAGGAATTCGATGAGCTCGGGGAGACTGCAGAGAACT

TGCAATGGTGATGTAGCAGTGAACAAGTGTGAAGGAGCCTGTAATTCCCAAGTGCAGCCCAGTGTTGTACCCCCTACC
GGATTTCTCAAGGAATGTTTCTGTTGCCGTGAGAGCTTTTTGCGCGAGAGAACGATCACGCTGACCCACTGCTACGAC
CCGGATGGCATGAGGCTGACTGGCGCGCCCCAGTCGCAGATGGACGTCAGGCTGCGAGAGCCGGCCGAATGCAAATGC
TACAAGTGCAGTGATTATTCTCGATAG

Ovogenesis

β Spectrin (β Spec)

>TRINITY_DN572_c0_g1_i18.p1

AACATAAACGACCGAAGAAACGTGAACCGCGAACGAGCTCTCCAGCTCATGGAGAAGCTGAGAGATCAACTCCAATC
CACCAGTTCCTTCAGGACTGCGAAGAGCTGGGCGAATGGGTGCAGGAGAAGAATGACGGCGCAGGACGAGACCTAC
CGCAGCGCCAAGACGGTGCACAGCAAGTGGACGCGCCACAAGGCCTTCGAGGCCGAAATCGCCTCCAACAAGGATCGC
CTTGTCGCGGTGCAACAGAACGGCGAAGAGCTGATGAAAGAGAAGCCCCGAAGTGAAGGAGGTGATTGCGCCGAAGGTG
AGCGAGCTGCAGCAGCAGTTCGACGACCTGGAGACGGTGCAGCGCGACAAGGGCGAAGCCTGTTTCGACGCCAACCGC
GAGGTGCTGCTGCACCAGACCTGCGACGACATCGACACCTGGATGAACGATCTCGAGAAACAGTTCGAGTCGGCCGAC
ACCGGCAACGACTTGGCCTCGGTCAACATACTCATGCAGAAGCAGCAGATGATCGAAACCCAAATGGCCGTCAAGGCG
AAACAGGTGACCGAACTGGAGACGCAGACCGAGTACCTGGAGAAAAACCGTGCCCCGAAAAGATCGACGAGATCAAGGTG
AAGAAGTCGCAGGTGGAGCAGCGCTTCGAGCAGCTGAAGGCGCCGCTGCTGCAGCGCCAGCGCGCTCGAGAAGAAG
AAGGAGGCGTTCCAGTTCAGACGCGACGTCGAGGATGAGAAGCTGTGGATCGGCGAGAAGATGCCGCTGGCCACCAAC
ACCGACTACGGCAACAGCCTGTTCAACGTGCACACGCTCAAGAAGAAGAATCAGTCACTGGCGACAGAGATCGACAAC
CACGAGCCGCGCATCCACACAGTGTGCCACAACGGACAGAAGCTGGTGGACGAAGGCCACGAGGACAGTCACGAGTTC
AACGAGCGGATCGCCGACCTGCTGCAGCGCTGGCAGCTGCTCAACGCGGCCGTCAAGCAGCGCAACGCCCAACTGGAC
CAGTCCGAGAAGGCGCAACAGTACTTCTTCGATGCGTGCGAGGCGGAGTCGTGGATGAGCGAACAGGAGCTGTACATG
ATGGTAGAGGACCGCGCAAGGACGAGATCTCGGCGCAGAACCTGATGAAGAAGCACGAGGCGCTGGAGCGGGCCGTC
GACGACTACGCGGACACCATTTCGCCAGCTGGGCGAGACGGCGCGCCACCTCACCCAACAACAGCATCCGCACAGCGAT
CTAATTGCTGTCAAACAATCGCAGGTTGACAAGCTGTACGCCGGTCTCAAGGACCTGTCCGGCGAGCGTCGCGCCAAG
CTCGACGAGGCACTGCAGCTGTTTCATGCTGAACCGCGAGGTGAGCGACCTCGAGCAGTGGATAGCCGAGCGCGAGCTG
GTGGCAGGCTCGCACGAGCTCGGCCAGGACTACGATCACGTACGCTGCTGTGGGAACGCTTCAAGGAGTTCAGCGC
GACACGGAGGCGACGGGCGGCGAGAGGGTGACGGGAGTCAATGCTATTGCCGATCAGCTGATTGCTGTGCGACACAGC
GATTCGGCGCTGATTGCTGAGTGGAAGGATTTCGTGAACGAGTTCGTGGCAAGACCTGCTGGAGTTGATTGAGACGCGC
ACCCAGATGCTGGTTCGCTCCAGGGAGCTGCACAAGTTCCTCCACGACTGCAAAGACGTGTTGGGCCGATTTTCGGAG
AAACAGCACGCCATGTTCGGATGAGCTGGGTTCGCGACGCCGGCTCAGTCAGCACGCTGCAGCGCAAACATCAGAACTTC
TTGCAGGATTTGCTCACCTTACAGTCGCAGGTGCAGCAGATCGAGGAGGAGTCGGGCAAGCTGCAGGCGAGCTACGCG
GGCGAGAAGGCGAAGGAGATCACGAACAGGGAGGGCGAGGTGGTGGCGGCGTGGGCCGGGCTGCAGGCTGTGTGCGAC
GCGCGCAAACAGAAGCTGGCCGACACCAGCGACCTCTTCAAGTTCCTTCAACATGGTGCGCACACTCATGCTCTGGATG
GACGACGTCGTTAGGCAGATGAACACTTCTGAGAAGCCCAGAGACGTGAGCGGAGTGGAATTGCTGATGAACAACCAC
CAGAGCCTGAAGGCCGAAATCGACGCCCCGCGAGGACAACCTTACCGCCTGCATATCGCTCGGCAAGGAGCTGCTGTCC
AGGAATCACTATGCTTCCAACGAAATCAAAGAGAAGCTTGTAAAGCTTAAGCAATCAAAACAACGCCCTTCTACACAGA
TGGGAAGAACGCTGGGAGAACCTGCAATTGATTTTGGAAGTGTATCAGTTTGCCCGTGACGCAGCCGTGGCTGAAGCC
TGTTTGGTGGCACAAGAACCTTACCTGATGAGCTCTGAACCTTGGGCACACAATCGACGAAGTTGAAAACCTGATCAAG
AAGCACGAGGCTTTTGAAAAATCGGCCACCGCCAGGAAGAGAGGTTCTGTGCTCTGGAACGATTAACCACCTTTGAA
TTGAAAGACCTGAAACGAAGAAAAAGAAGAGGAGGAGCGAAAGAGACAGGAAGAGTTGGCCGCCAAGCAGGCCCTCGCCA
CCACCTCGCCACACGATGGAGGCGGTGACAGGCCTGACGGTGATACTTCAGCTGAAGGTGAAGCAGCCCTGAGCGAG
AAGGAAGACAGCAGAAAAAGAACGTCTAGCCCGGCAGCTGAAGATAACCAATTGGAAGGCACTTTGGTGAGGAAACAT
GAATGGGAGGCCACCACTAAGAAGGCTTCAAACAGATCGTGGGACCGCGTGTGGGTGGTGATAAGGGGCGGGGAGGTG
GCCTCGTACAAGGACCAGAAGGCGGCCAAGGCGACGCCGAGGTGTTCTTCAAGGGCGAGGCGCCGCTCGATCTCAAC
GGCGGCAGCGCCGAGGTGGCCACCAACTACACCAAGCGCAAACACGTGTTTCAGGCTGAAGCTGGCCGGCGGCGCTGAG
TACCTCTTCCAGGCGAAAGACGACGAAGAGATGCGTGTGTGGGTGGGGGCGCTGCAGTCGGCGTGCCAGGGCTCCTCG
GCGGCGGCCGCTCTCGCACCTCGCAGACGCTGCCCCGCCCCGCGCGACGCGCCCAAGGACGAGCCCAAGCGGCGCTCC
TTCTTCACACTCAAGAAAACTAA

Armadillo (arm)

>TRINITY_DN6159_c0_g1_i19.p1

ATGAGTTACCAAATGCAGTCTGCTCAAGCCAGAAACATGAACTACCAGGGTGGCGGCCAAGGGGGCGGCGACATGCCG

ATGGGCGGCAAAGAGCAGCAGACGCTGTTGTGGCAGCAGGGCAACTACCTGGGTGACTCAGGCATTCACTCGGGCGCC
ACCACCCAGGCGCCGTCGCTCACCGGCAAGGAGGACGACATGGACGGCGATCAGCTCATCTTCGACCTCGACCAGGGC
TTTGCGCAGGGATTACGCAGGACCAGGTTGATGAAATGAACCAGCAGCTGAGTCAGACGCGTTCGCAGCGCGTGCGA
GCGGCCATGTTCCCCGAAACTCTCGAGGAAGGCATTGAGATTCCGTCGACACAGTTCGACCAAGGCCAGCCGACCGCC
GTGCAGCGCCTGTCCGAACCCAGCCAGATGCTCAAGCAGCGCGTCGTCATCTCATCAACTATCAGGACGATGCAGAT
CTAGCGACGCGCGCCATTCCCGAGCTGATCAAGCTGCTGAACGACGAGGACCAAGTGGTGGTGTGCGAGGCGGCCATG
ATGGTGCACCAGCTGTCCAAGAAGGAGGCCTCGCGCCACGCCATCATGAACAGTCCACAGATGGTAGCTGCCTTGTTA
CGAGCGATCTCGAACAGTAACGATTTGGAACGACGAAAGGCGCAGTGGGTACCCTGCACAATTTGTCTCACCATCGC
CAGGGTCTGCTCGCCATTTTCAAGAGTGGAGGCATACCGGCTCTTGTCAAGCTTTTGAGTTCTCCAGTGGAGTCAGTG
CTGTTCTACGCGATCACCACGCTGCACAACCTGCTGCTGCACCAGGAAGGCAGCAAGATGGCCGTGCGCCTGGCAGGT
GGGCTGCAGAAGATGGTCGCCCTGCTGCAGCGCAACAACGTCAAGTTCCTGGCGATTGTCAACCGACTGCCTGCAGATC
CTCGCCTACGGCAACCAGGAGAGCAAGCTGATCATCTGGCGTCGACGGGCCCCGTGAGCTGGTGCGCATCATGCGC
AGCTACGACTACGAGAAGCTGTTGTGGACTACGTCGCGCGTTCCTAAAGTGTGTGTCGGTGTGCTCAAGCAACAAGCCA
GCCATAGTGGAGGCGCGCATGCAGGCTCTGGCCATGCACCTCGGGCATCCAGCCAGCGTCTGGTGCAGAACTGT
CTCTGGACCCTCAGGAATCTGTCCGATGCTGGAACCTAAGGTGGACGGCCTCGAGGGCCTGCTGCAGTCGCTGGTGCCT
CTGCTGGCGTCGACCGACGTCAATGTGGTGACGTGCGCCGCGGGCATTCTGTCCAACCTGACGTGCAACAACCCAGCGC
AACAAGGTGACCGTCTGCCAGGTGGGCGGAGTTGACGCCCTGGTGCATACCATTGTCAACGCAGGGGATCGCGAGGAG
ATCACTGAGCCCCGTGTTTGCGCCCTCCGTCACCTGACCTCGCGTCACGTGGACTGCGAGATGGCGCAGAACGCGGTG
CGCATGCAGCTGTACGGCATGCAGGTGATAGTGAAGCTGCTGCACCCGCCCAGCCGCTGGCCGCTCATCAAGGCCGTC
ATCGGACTCATCCGCAACCTGGCGCTGTGTGTCAGGCCAACACGCCCCGCTGCGCGAGCACGGCGCCATACATCATCTG
GTGCGGCTGCTCATGCGCGCCTTCCAGGACTCGCAGCGGCAGCGTTCGTCGGTGGCGAGCAGCGGCGGGGGCGGCGGC
GGCAGCCAGACCGGGGGCGGCAGCAGCGTGTACGCGGACGGCGTGCGCATGGAGGAGATCGTCGAGGGGCACAGTGGGC
GCGCTCCACATTTTGGCGCACGAGTCGCACAACCGCAGCATCATACGCAGCCAGCACGTGATTCCGATATTTGTGCAG
CTGCTGTTCAACGAGATTGAGAACATACAGCGGGTGGCGGCCGCGCTGCTCTGCGAACTGGCCACTGATAAGGAGGGC
GCCGAGATGATTGAGCAGGAGGGGCCACCGCCCCGCTCACCGAGCTGTTGCATTCAAGGAATGAGGGTGTGCTACT
TATGCAGCCCGCGTGTATTCCGCATGTCCGAGGACAAGCCACAAGACTACAAGAAGCGTTTGTCAATGGAGTTGACC
AACTCTCTGTTTCAGGGAAGACCAGAACCCGTGGAACGGTGGTGTATCTAGGCATGGGTCTGATTTACAGGACATGCTT
GGCCCTGATCAAGGCTATGATGCTATGTATACGCAAGGCCCCACCCAGTGTACATAGTAGCCACGGAGGCAGGCAGTAT
CAGCAAGGTTATGACCAGATACCAGTAGACTCGATGCAGGGTCTAGACCTGGGCTCTCAGCACGACTCCAACCTACAGC
CCAATGGAGGTGGGAGCCGATTTGGATTTTCGATCATTGACGAGCTGCCCCAGCCGCCTCAAGACAATAACCAGGTG
GCCGCTGGTATGACACTGATCTGTAG

vas (vasa)

>TRINITY_DN3339_c0_g1_i1.p1

ATGCATCATCGTAAGAGATCTCGAAGCCGATCCAGGTCATCAAGTCCAGAACACTATAAGAGAACTAAAGGTTCTCAT
GGACGTGAGAGGAGCAGTAAGGAATATAGACGAAGGAGCAGATCAAGAGAGCGATCTCGTAGGTCCCACAGTAGGGAT
CGAGACAGAAAACCTTTTCAAGAAGGACAAAAAGTCATCTGATTACCGGTGCAAGCGAAGTTATGGAAAAGGAAGGCGA
ACACGTTCTAGGTCTGATTTCGAGATCGCGCTCCAGATCGAGATCGCATTCCAGAGAAAGGGGTAAACTCAAAGACGAG
AAGAAAGATTGAAAAGCTGACGATGAGCCTTTTCGATCCCAACCTGAGACAAAGAAGAAGAGCAAAAAAGGCTTGAA
TTGGAAATGCAAAAGAGAAGAGAGCGGATTGAAAGATGGAGAGCTGAAAGGAAAAAGAAGGAATTGGAAGTTGTAAAA
AAGGATCTGAAGTCATCTTTAGCCAATCTTCAACTTCCGACCAAGAAATGGAATCTTGAGGATGACAGTGACGAAGAA
GAAGATGGGAAAGCTGATAACAACAAGAATGATGGCGAGGAAGAAGTGGATTCTTTGGATGCTTTCATGCAGGGTGTA
CAAGAAGAAGTGAGGAAAGTGAATAAGATAGAAGGAGCCAAGAGAACGGCCGATGGCCAGGGCAAGGTTCCGCCGGCC
GGATCCGGAGGCCTGGTCATAGTGACCGGTGTTGCCAAGAAAAAAGTCGACAACAAAAAAGGAGAACTCATCGAACAG
AATCAAGATGGACTCGAGTATTCATCAGAAGAAGAGGAAGAAGATTTGTCTACGACTGCAGCTAGCATAGCCAATAAG
CAGAAGAAAGAGTTGGCTAAAGTCGATCACACAAGCATCGATTATACTCCTTTCAGAAAGGACTTTTACACGGAGGTT
CCCGAAATTGCTCGGATGACACCAGAGGAGGTGGAAGCCTACAGAGAAGAAGTGAAGGTATAAGAGTGAAAGGCAAA
GGCTGTCCAAAACCTATCAAGTCCTGGGCTCAGTGTGGTGTGAGCAAAAAGGTTCTTGATGCATTGAAAAAGCAAGAT
TACGAGAAACCTACACCAATTCAGGCTCAAGCAATACCTGCTATCATGTGAGGACGTGATCTAATAGGAATAGCAAAA
ACTGGAAGTGGGAAAACACTAGCATTTGTTCTTCCAATGCTTCGCCACATCATGGACCAACCACCTCTAGAGGCGAAT
GATGGCCCACTTTCATTGATAATGGCCCCCTACAAGAGAATTGTGCATGCAGATTGGCAAAGAACTAGGAAGTTCACA
AAATCGCTCAATACTCGTGTGCTTTGCGTTTATGGTGGCACTGGCATCTCCGAACAGATCGCTGAACTGAAACGAGGA
GCTGAAATAATAGTTTGCACACCTGGCAGAATGATCGACATGCTTGACGTAATAGTGGTCGAGTAACAAATTTGAAG
AGAGTTACATACATCGTTCTTGATGAAGCAGACAGAATGTTTGACATGGGCTTTGAACCACAGGTTATGAGGATAATT
GAAAATATTAGACCGGATAGGCAGACTGTGATGTTTCAGCGCTACTTTCCCCAGACAAATGGAGGCATTGGCAAGAAGG

ATACTCACAAAACCTATTGAAATTCAGGTTGGAGGAAGGAGTGTGGTATGTAAGGAGGTGGAACAGCATGTTATAGTT
TTGGAGGAAGACCAGAAGTTTTTGAAGTTGCTGGAAGTCTGCTGGGACATTATCAGAATCAGGGCAGCGTAATAGTGT
GTTGACAAACAGGAGAATGCTGACGCACTGTTGAAAGATCTCATGAAGGCTTCATATCAGTGTATGTCGCTACATGGT
GGAATCGATCAGTATGACAGAGATAGTACAATAGTTGACTTCAAAGCTGGAAGAGTCAAGCTTTTGGTTGCAACATCA
GTGGCAGCTAGAGGCTTGGACGTGAAACAGTTGATTTTTGGTAGTGAAGTACGACTGTCCGAACCATTACGAAGATTAC
GTGCATCGGTGTGGAAGGACAGGTAGGGCAGGAAACAAGGGCCACTCGTACACATTATTACGCCGGAACAGGAGCGG
TACACGGGCGACATAATCAGGGCGCTGGAGTTGTCTGTCGGTGAGTGTTCGGGACGCGGTGAGGGCGCTCTGGGAGAAC
TACAAGGCTAAGCAGGCGGCCGAGGGCAAGAAGGTGCACACGGGCGGCGCTTCAGTGGCAAGGGATTCAAGTTCGAC
GAAACAGAGGCCATGATGGTCAACGAAAAGAAGTTCAGAAGGCGGCGCTCGGTCTGCAGGACTCTGACGACGAG
GACCTGGAGAACGACATCGATCAGCAGATCGAGAGCATGTTGGCGCCGAAGCGCATCGTGAAGGAGATCAAGGCGCCG
CTGGGATCGGCCAATGCGGCTGCGGCGGCGGCTGCGGGCGGGCGGCGGCGGCGGAATGCCGTCGACCGACAAGTTG
GAGCTGGCCAGGCGCCTGGCGTTCGCGCATCAACATGGCCAAGAACCTGGGCGCCGAGGCGAAGGGCGCCACTCAGCAG
GCGGCGGAGGCCATTCTCAAGGGCGGCGCATCGCAGCCATCATCACCGCCAAAACCTGTTGCCGAGCCTCGCCGCC
AAACTCAACACGCGGCTCAACTATCAGCCTCGCGAGGAGGAGGACAAGCAGGAGGAGGAGGCCGTCGAAACGTTTCGC
AAGTACGAGGAGGAGCTCGAGATCAACGACTTTCCGAGCAGGCTCGCTGGCGGGTCACTCGAAGGAGGCGCTCGCC
CAGATCTCCGAGTACTCGGAGGCGGCATCACAGTCAGGGGCACCTACTACCCGACTGGCAAGGCGCGCCCGCAAGGC
GAGCGCAAACTCTTCTCGCCATTGAGAGCACCAACGAGTTGGCAGTCTCTAAGGCCAAAATCGAAATCACTCGTCTC
ATCAAGGAAGAATTCTCAAGCTGCAGACTTCTGCTCATCACAGTATCAACAAATCTAGATACAAAGTTGTGTAA

P-element induced wimpy testis (piwi)

>TRINITY_DN185_c0_g1_i9.p1

ATGAGTAAATGGCAGTTAGCCAATGGAAAATACCTGAACGAAAGAAAAATGGCACCAAGGGCCGCAAAATTTGAACTG
GAACTCAACCACTTCGAATTGTTTTTCAAGAAACAAAATTTTATCGCTATCCACTACGATGTATCATTCAAACAGAC
AAACCTCGACGTATGTTCCGGCAAATAATGGAAACATTCGACGAAAAAAGTTTCCCAACAATTACCCAGCTTTTCGAT
GGGCGTAAAAATCTATACAGCGCAAAGGAACCTCCATTTCGGAATGGAGGTGACAGATACAGTGAAAGTTTTCAACGAC
GAGCGTTATATGGATGTAGAGTATGAAGTGACGGTTAAGTTTCGCGTCAAGAGTCGATATGAGTCAACTGAGCCAGTAC
TTGAGCGGCAAAGGCCAGTCTTATCAAACACCTCAGGAGGCGTTGACAGGAGTTCGATATAGTACTGCGAAATCCAGT
GCATTGTCATTCTAGTGTGCGGAAGATCGTTTTTCTCGAAACCAAGATCCATCATTGAAGTGGGAGAGTCTGGAG
CTTTGGTATGGATTCTACCAGAGTGCCATTCTTGGATGGAAGCCTTTTCTGAACGTTGATGTTGCTCATAAAGGATTT
CCGATGGGCGAAAGGTTGTTGGAGACCTTGTGTCTGTTATCAAAGGTGTGAGATCAACGACTTGAGGAATTTGAAACAT
TTGGACAGTTACGTACAAAATGATTTTCGAGAAGTATATTAAGGATTGAAAGTGGAGTATCAAATCCCACAAAGACCT
GACACAAAGCGTGTTTACAAAGTCAACAAATTTGGTGAAGAATGCTGTAGAGCAAAGATTCATTTTCGAAAAAGACAAC
AAAAAGTGGAAATGACCGTTGGTGAATACTTCCAACGCGAGAAGAAGTGCAGTGTGCAGTCTCCCTACCTGCCTCTG
GTACACATTGGCCCTCTAAATAAAGAGTTCTTCGTTCCAATGGAGATGTGCACCATTGTTTCGAGGTGAGTCAATAAAC
CGTAAGCTGACCCCGGTACAAACGGCTGAAATGGTCAGAAACGCTGCCAAGCCTCCCGACGAGAGGAAAAGAAGATC
GCCTCGGCGTTGAAAAGGGCAAACCTTCAACAATGACAAATGCGTACAGGAGTTTCGGAATTCAGGTATCGGACAGATTT
GCCCAAGTAAGTGGCAGAGTTCTTGATCCACCTACTTTGGAGTATAACAGGCAAATGATCACACCGCAGAAAGGAGTA
TGGCGATCTGGTAAGTTTCTCAACCCGGTTCAAATCCAAAATTTGGGCGATCATCAACTGTGATCGTCTGATACCAATGAC
GGACAGTTACAGCGATTTGCATCGGGAATGTGAGAACAAGGCAGAACACTTGGTGTGAATATTGGGGCAATATCCAAG
ATAATTCCGTTCTCTCAAATCCAACCCAACCGACCGAATTGGAGACAGGAGTTCAACAAGCTGCTCTGCTCCCTGCGA
GACAGCAAAACTGAAATTTGTGATTGTTGTTATCCCTGAGATTGGAGACATTTATGCCAAGGTGAAACAAACGGCTGAG
CTAAGTGTGGAAATTCTAACACAGTGCATAAAATCGAAAACAATGTTCAAAATGAATCCGGCCACAATTGGAAATATC
CTCTTGAAAGTCAACTCGAAGTTGAATGGGCTCAACCATAAAATCGGCGGACGACCAAAAGCTGCTGGCAAGACCTGCT
ATGATCATGGGTGCTGACGTCACTCACCCGTCTCCTGATCAAACATAATTTCCCTCCGTTGCTGCTGTGTGTCAGCAAGT
CATGATGCAAACGGGTTTCATGTACAATATGATGTGGAGATTGCAGCCTGCCAAAATGGAAATCATTGAAGATCTGCAG
AATATTGTGGTGGCCAGTTGAAATTTTTCTTTTCAGAGGACAAGATGCAAGCCGGAACCATCTATTTCTTCAGAGAT
GGTGTGAGTGAAGGCCAGTTCAATCAGGTTCTGAGCGCCGAGCTAACAGCCATCAGACAAGCGTGCCGGACGCTGAAC
GAAAACACAAACCGGAATCACATTCCTGGTGGTGCAGAAGCGTCATCACACGCGCTTCTTTCCGAAGAACGATCGC
GACAAAGATGGCAAGTGGGGAACGTCGCCGCGGCACCATTTGTGACACGAGATCACCCACAAGTCGGAGACCGAC
TTCTACCTGGTTAGCCACGCTAGTATCCAGGGTACCGCACGACCCACCAAGTACCACTTGCTGTGGGATGACAACAAT
ATCGACGAGGACGACCTCGAGGAGCTGACCTACAGTTTGTGTACCTGTTACCAGGTGCACCCGATCTGTCTCCTAC
CCTGCTCCACCTACTATGCGCACTTGGCCGATTCAGGGCACGCGCTACCTTGAAGGGCAACAGACTTCACCGAAT
TTGGCCCGTGAAATGGAGAGCCGTATGATCGAAGTCAATTTCCACAAAGCAAATCCAATGTTCTTCGTCTAA

Twin

>TRINITY_DN11696_c0_g1_i1.p1

ATGACAGTATTTGTGTGGCGTTGCAAAGTATTTCAAACATTGATTTCGTTGTTGTAGTTTATCTTACAAGACAAATTTA
AAAACTAACCCCTATGGAAAAACCAACAGCTTACTTCCGTCTAACTGATGGATCTGATACATTCCAGGTGAGTTTTTTT
TTCAAGAATGAAACATTGAATATAGAACGTGAGTGTAATTTCAACAGAAAATTGGATGAAAAAATTGGAGACTTTTTT
AATAGAGTTAGAGCGAACATAGATAAATTGAAATCATCCAAACGCAAGAATAAGAAGCATAAAGGACTGACTGATTTA
GAACTTGATATCAGTGAATCTTCGGCAGGTTTATTGCAAGGTGGTGTCTTGATAGATAGTGAGACCATTTCGCCGGGAT
GCAATATTCTATAACAAGTCCACGGAAACCCCTATCCCTCATTATCAATAAAGTAGAGTACACTGTATCAATTAATTCC
CCATGGGTGAATGAACTATGTCTTCCAACAAGCATGCTGGCTGGATTTTTAATCTACCCCTCAAAGTTCAATATTTT
TTCTCTGAACAAAAATATTGTTCTTTTCGAGTGGTTCAGGAAGAAATGTAAGCATGCACCTAGAAGTGATGATCTAGGA
TGGGAAAGAGTTGGAGAAGGATTTACTTACTCTACCCTAATGAAGATATTGGATATTATTTAAAGTTGACATGTGAT
CCGAGTAATGAGAAATTCAAGGGTCTGAAGCTTCAACAGTCTCGTCAGGTCTTATTGAAGCTGGTCCAGGAATGTGC
CCATTTGAAGATCGCCATCCATTTACCAGAGAACGATTATCCGGGAATGAATTTTCGTGTGATGACTTACAATATTCTC
TCTGATATTTACACCGAAACAGCGGAAGCTCTCAGTCATCTATTTTCACTACTGCCCCCGTTACGCTTTGAAAATTGAT
TACAGAAAGCAACTTCTAATCAGAGAAATATTAGGATATAATGCCGATTTGATTTTCTTGCAAGAGGTGGATGATTGG
TTTTTCAATAGAGACTTGAAATATGTGTTGGGTCTTAATGAGGGTGCTATGCACAGAAAGGTAGCACTGTGAGA
GAAGGTGTTGCTTGCTTCTACTCCAAAGATAGATTGAGACTCCTACACTCCGAATCGATTATACTGGGAGAAGAAGTT
CAGAACAAACCCCTTGCTTGAGCGAGATATGGGACGTTATTAGCAGAAAAGAACCCTTAAACTTCGTGTGATAGAAAGG
CATACAACCTTTACAGATTTTGACTCTGGAAAGTTTGGACAAATCAGATGAGTTATTAGTTGTTGCAAAACACTCATCTA
TATTTCCACCCAGATGCTGATCATGTTTCGTTTAAATTCGAAGGAGGAATATCCATGAAGTATATCACTGACTACATTGGA
AAACTAAAAACAGAAAATCCAAATCGAAATGTATCACTCATTTTTCTGCGGTGACTTCAACAGTGTTCCAGAATGTGGA
ATTTACAACTTATGACAACTGGAATTGTACCAGAGGATTACGTTGACTGGACAAGTAATAAAGAAGAAGCTGTAGAA
GGACTTAGTGAGTCAGCCGTGGAGAATTGCCAGTGCGTGTGGCACCCCGCAATACACAAATCTCATCCAGGAGTTT
TCTGGTTGTCTTGACTACATATTCTATCAAACCTGATCGACTAGCTGTGACTCAGGTGGTGCCATTGCCTACAGAAGAA
GAATTGAGACAGCATACAGCTCTCCCCAGTGTTGTATTTCTCTGATCACCTAGCTCTGATAGCAGATCTACGATGG
ACTAGTTAA

aubergine (aub)

>TRINITY_DN185_c0_g1_i9.p1

ATGAGTAAATGGCAGTTAGCCAATGGAAAATACCTGAACGAAAGAAAAATGGCACCAAGGGCCGCAAAATTGAACTG
GAACTCAACCACTTCGAATTGTTTTTCAAGAAACAAAATTTTATCGCTATCCACTACGATGTATCATTCAAACCAGAC
AAACCTCGACGTATGTTCCGGCAAATAATGGAAACATTCCGACGAAAAAAGTTTCCCAACAATTACCCAGCTTTTCGAT
GGGCGTAAAAATCTATACAGCGCAAAGGAACCTCCATTCCGAATGAGAGGTGACAGATACAGTGAAAGTTTTCAACGAC
GAGCGTTATATGGATGTAGAGTATGAAGTGACGGTTAAGTTTCGCGTCAAGAGTCGATATGAGTCAACTGAGCCAGTAC
TTGAGCGGCAAAGGCCAGTCTTATCAAACACCTCAGGAGGCGTTGCAGGCAGTTGATATAGTACTGCGAAATCCAGCT
GCATTGTCAATTCATTAGTGTCGGAAGATCGTTTTTCTCGAAACCAGAATCCATCATTGAACTGGGAGAAGGTCTGGAG
CTTTGGTATGGATTCTACCAGAGTGCCATTCTTGGATGGAAGCCTTTTTCTGAACGTTGATGTTGCTCATAAAGGATTT
CCGATGGGCGAAAGGTTGTTGGAGACCTTGTGTGCTTATCAAAGGTGTCAGATCAACGACTTGAGGAATTTGAAACAT
TTGGACAGTTACGTACAAAATGATTTTCGAGAAGTATATTAAAGGATTGAAAGTGAGATCAAAATCCCACAAAGACCT
GACACAAAGCGTGTTTACAAAGTCAACAAATTGGTGAAGAATGCTGTAGAGCAAAGATTCAATTTTCGAAAAAGACAAC
AAAAAAGTGGAATGACCGTTGGTGAATACTTCCAACGCGAGAAGAAGTGCAGTGTGCAGCTCCCTACCTGCCTCTG
GTACACATTGGCCCTCTAAATAAAGAGTTCTTCGTTCCAATGGAGATGTGCACCATTGTTTCGAGGTGAGTCAATAAAC
CGTAAGCTGACCCCGGTACAAACGGCTGAAATGGTCAAGAACGCTGCCAAGCCTCCCGACGAGAGGAAAAGAAGAATC
GCCTCGGCGTTGAAAAGGGCAAACCTCAACAATGACAAATGCGTACAGGAGTTCCGAATTCAGGTATCGGACAGATTT
GCCCAAGTAACCTGGCAGAGTTCTTGATCCACCTACTTTGGAGTATAACAGGCAAATGATCACACCCGAGAAAGGAGTA
TGGCGATCTGGTAAGTTTCTCAACCCGGTTCAAATCCAAAATTTGGGCGATCATCAACTGTGATCGTTCGTACCAATGAC
GGACAGTTACAGCGATTTGCATCGGGAATGTCAGAACAAGGCAGAACACTTGGTGTGAATATTGGGGCAATATCCAAG
ATAATTCCGTTCTCTCAAATCCAACCCAACCGACCGAATTGGAGACAGGAGTTCAACAAGCTGCTCTGCTCCCTGCGA
GACAGCAAACTGAAATTGTGATTGTTGTTATCCCTGAGATTGGAGACATTTATGCCAAGGTGAAACAAACGGCTGAG
CTAAGTGTTGGAATTCTAACACAGTGCATAAAATCGAAAACAATGTTCAAATGAATCCGGCCACAATTGGAAATATC
CTCTTGAAAGTCAACTCGAAGTTGAATGGGCTCAACATAAAATCGGCGGACGACCAAAAGCTGCTGGCAAGACCTGCT
ATGATATGGGTGAGTCACTCACTCACCCTGCTCCTGATCAAACTAATATTTCCCTCCGTTGCTGTGTGTCAGAGT
CATGATGCAACGGGTTTCATGTACAATATGATGTGGAGATTGAGCCTGCCAAAATGGAATCATTGAAGATCTGCAG
AATATTGTGGTGGCCAGTTGAAATTTTTCTTTTCAGAGGACAAGATGCAAGCCGGAACCATCTATTTCTTCAGAGAT
GGTGTGAGTGAAGGCCAGTTCAATCAGGTTCTGAGCGCCGAGCTAACAGCCATCAGACAAGCGTGCCGGACGCTGAAC

GAAAACTACAAACCGGGAATCACATTCTTGGTGGTGCAGAAGCGTCATCACACGCGCTTCTTTCCGAAGAACGATCGC
GACAAAGATGGCAAGTGGGGAAACGTCCCGGCCGCCACCATTTGTCGACACGCAGATCACCCACAAGTCGGAGACCGAC
TTCTACCTGGTTAGCCACGCTAGTATCCAGGGTACCGCACGACCCACCAAGTACCACTTGCTGTGGGATGACAACAAT
ATCGACGAGGACGACCTCGAGGAGCTGACCTACAGTTTGTGTACCTGTTACCAGGTGCACCCGATCTGTCTCCTAC
CCTGCTCCACCTACTATGCGCACTTGGCCGCATTTCAGGGCACGCGCTACCTTGAAGGGCAACAGACTTCACCGAAT
TTGGCCCGTGAAATGGAGAGCCGTATGATCGAAGTCAATTTCCACAAAGCAAATCCAATGTTCTTCGTCTAA

Bicaudal (Bic)

>TRINITY_DN4949_c0_g1_i1.p1

ATGAACCAAGAAAACTAAAGAACTGCAAGCCCAGGTGCGCATTGGAGGGAAAGGAACACCCCGAAGAAAGAAGAAG
GTTGTTTCATGCGACTGCTGCTACCGACGACAAGAACTACAAAGTTTATTGAAGAACTATCGGTAAATACTATTCCG
GGCATTGAAGAAGTCAACATGATAAAGGATGACGGAACGGTCATCCATTTCAACAACCCCAAGCTCAGGCTTCATTG
GCCGCTAATACGTTTCGCAATTACTGGTCACGGTGAAAATAAGCCTATCACAGAAATGATTCCCGGAATCCTCAATCAG
TTGGGCCCCGAAGGATTGACTCAGCTGAAGCGTTTGGCTAAAACAGTAGCCAGCAACGCCAGCGCGGTAAGGCTACC
ATCGAAGAAGACGACGAAGTCCCTGAGTTGGTCGAGAAGTTTCGACGAGCCAGCAAGGAAGAAGTACCTGTCAGTATT
GCCACTGAAAAGGATGCTTCTGGCGAAGGCAAGGAGAACACCGCTCCAGTTCTTGAAGCGGAAACAAAAGAAAAGCCA
AAGTAG

Bicaudal C (BicC)

>TRINITY_DN11822_c0_g1_i1.p1

ATGACTGAAACAAACACTCATGTCTTATGGCCATCCAGATTAAAAATTGGATCAAAATCCCACAAAGACCCCTATATA
AGAGTACTCGGACGTTTAGAAGACATAAAGATTGCAAAAAATAAGATTATGGCAGTTCTTGACACAAGGGGCAGTAGA
GTGATGATGAAACTGGAGGTGAGCTATACAGATCACTCGCACATAATCGGCAAGGGAGGCCTAGTCATAAAGGAGGTG
ATGGAGGAGACCGGCTGTACATACATTTTCCAGACTCGAATAGAAAACAACCCCAACGAGAAGAGCAATCAGGTGTCA
ATAGCCCGGCGAAATTAGTAAGGTGGAAAAAGCTAGATCTAGAATAAGGGAACCTTCGTCTTTGGCGTTTACATTCCAA
CTACCGGTGATGGGCACAGTGCATTTCGGCTTCGGAGCTGGCGTCGTCTGATACATCACCGGCTTACAAGAGAGATACAAT
GTGCACGTCAAATTTCAAACACAATCGAAACTCTACTCGACTCTGGTCTGGTCAAAGGCTGTGAACAGGAAGTGAAG
CAAGTGAGGAGGCCACCCCTCCAGCTGGTCCATCGTTTTATGTGGCCCAATGGCGGTTGGCGTTCAAATCAAATGTTT
ATGGAGATATCACCGCATCATCATCAAATTGTTCTGGTGAATGGTGATAGTAATTTGTTGAAGATTATGGGATTGACA
AAAACAAAAATCTTATTCCCAGATGTGGATAATCCTTGTGTACCAACTCTTAAAAAGAGCAATGTGTCAATCACTGGT
GCTATTCATGATGTGTACTTAGCTCGCCAGATGCTAATGGGATCATTGCCGTTATTATTAATGTTTGAAGTAGAGGAC
AAATTTCCAAGCACCGACCAAATAAACACAATCATGTACAACCTTACAACCTATCCATTATTGTGAAACAGAAACCTCGA
CTCAACTCGATCTCAATAATCATTAAGGCATTGAAAGACAAGCAGACAATATTTATGAAGCTTACAGACAATTGAAA
GGACTGAAAGAACAAGAGTAATTGCAAGTATTCTTTCGACGTACAATATTCCCAACTTCCCAACTCTACTCCAATG
AAACCTGTGAAAGTGAATGAGCCAATGGGCTGCCAGAATATACCAGGACTACCAGGAAATCTGCCGGATCCAGAAGT
ACTCAACTATTCTGTCAATTGTGCGAGTTATTTTCTCAACATCAACATCAGAGAAATCTCACCGAAAATGCTCTTGGA
CAAGCTGTATCTTCATTGTCTCTTCTGATATTTTCAAGGAATCTCATTTCATCGCAAAACAGTAGATGCAGCAGCCCA
ATGCTGAGTCCCAATTCCAAATCGGCTCATGGCAGTGGTCAAATACCCAACACAGTAGCGATTTGAGCATGAGTACC
GGTAGTAATAATTGGAGAGCTCCTGGCTTTGGTAGGAATTCGCATATTGAATACCAAGAGATGAAGCGCCAGGCTTTA
ATTGCTCTTCAGACGAAACCTAATGCAGATGATCCTCGTGTTCACACGTGACATGGGCAGGCTATGGCTTGAGTCGA
TCTGGTCCAGCAAATGTACAAATAATGAACCACAGACCTCGTCAGACGTTGATGGAGCATGGGGAGGCGCGAGCTTC
GAGAACCAATACGGTACCAGCAATATTTTAGAATCTGTATCTAGTCAACACATGGAGGTGATCACTAGTTCGGAATCA
TATACGATTCTCTCTCCTGGCTTCTCTCAATCTGGAATGTACATTCTCTTATCCACATGCATCACATTGATTG
CCTACATTTTCGACTTTAACCAGAGCAGGACTTGAAAGAAATCGGGGTTTCCCTCGTAGGAGCTCGACGTAGGATTCTG
GCATCAATTGCAGAACTGAATCGGTCTCAGTCGAATTTGGCACCTGGCCTTTGGGACTATATGGAGAGATCCAACCTCT
TCCAATGCTTCTTCAGGACCAGAATGA

Dynein heavy chain 64C (Dhc64C)

>TRINITY_DN50570_c0_g1_i1.p1

GCTGTCAAATCACAGGAATTACAAGCTAAAAATGAAGCTGCCAACGCCAAGCTGCGTCAGATGTTGCAAGACCAGCAA
GAAGCCGAAAAGAAAAAGGTGCAGAGCCAAGAAATTCATCTGAAATTGAGAAGCAAACCATTCAAATATCTCAGAAA
CGAGAGGATGTAATGGCCGACCTCGACAAGGTGGAACCGGCCGTCATTGAAGCTCAGCAAGCTGTGAAAGAGATAAAG
AAACAGCAGTTGGTTCGAAGTGAGATCGATGGCCAACCCGCCTCCCATCGTCAAACCTGACCTCGAGTCCATTTGTCTG
CTGCTTGGTGAGAGTGCCAGCGACTGGAAATCCATCAGGACAGTAGTGATGCGAGAAAATTTTCATCAATTCTATCGTG
TCAAATTTTCAGTACGGAAAACATAACCGATGAAGTTCGAGAGAAAATGCATTTCAGTTACCTGAACAATCCAGAGTAT

ACTTTTAAAAAGTTAGCCGAGCCAGTATGGCTTGTGGTCCTATGGTCAAATGGGCGGTTGCTCAGGTGAGCTACGCG
 GACATGCTGAAGAAAGTGGAGCCGCTGCGCGAAGAGCTGCACAGCCTGGAGGTGCAGGCGGCGGAGAACAAGGAGAAG
 GGCGAGGAGATGAAGGCCTTAATTGCACAGCTGGAGAAGAGCATAGCCGCCTACAAGGACGAGTACGCGCAGCTGATT
 TCGCAGGCGCAGGCCATCAAGGCCGACCTGGAGAACGTGCAGGCCAAGGTGGATCGCTCCATTGCGCTGCTCAAGTCG
 CTCGGCATCGAGCGCGAGCGCTGGGAGGCCACCAGCGAAACCTTCAAGTCGCACATGTCAACCATCGTTGGAGACGTA
 CTCCTGTGCGCTGCCTACCTCGCTACGCTGGCTACTTTGATCAGCATTACCGTACCAACTTGTTCGCCACCTGGTGC
 AATCATCTTACCCAAGCTTGCCCTCAACTTCCGTTTCGGATATCGCTCTCACCGAATATCTGTCCAATCCAGACGAAAGG
 CTGCGTTGGCAGGCGAATGCCCTGCCATCTGACGATCTGTGTACGGAGAATGCTATCATGTTGAAGAGATTCAACAGA
 TACCCTCTGATCATTGACCCGTCTGGACAAGCAACTGAGTTCATTTTGAATGAATTCAGAGATAGAAAGATCACCAAG
 ACCAGCTTTTTGGATGACTCGTTTCAAGAAAGTCTTGAATCTGCCCTGAGATTTGGTAACCCACTTCTAGTTCAAGAC
 GTTGAGAACTACGACCCGATCCTGAACCCGGTGTGAACCCGCGAGCTGCGCCGCACGGGCGGCGGGGTGCTGATCACG
 CTGGGCGACCAAGGACATCGACCTGTGCGCCTCCTTCGTTCATCTTCTGTGACGCGCGACCCACCGTCGAGTTCCCG
 CCCGACATCTGTTGCGCGCTCACCTTCGTCAAACTTACCGTCACGCGCAGCTCGCTGCAGTCGCAATGCCCTCAACCAG
 GTGCTCAAGTCCGAGCGGCCGACATCGACGAAAAACGCTCCGATCTGCTCAAAGTGCAGGGGAATTCCACCTTCGA
 CTACGACAGCTTGAGAAGTCTCTGCTTCAGGCTCTCAATGAAGCCAAAGGAAAGATCCTGGATGATGATTCAGTCATC
 ACCACTTTGGAAACTCTGAAACAAGAAGCTGCAGATATCAGCAAGAAAGTTGAAGAAACTGATAAGGTTATCGCTGAA
 ATCGACACTGTTTACAGCAGTATAAACCATTGTCTCAGGCATGCAGTAATATGTACTTCACAATGGACAGCTTGAAT
 CAAATACATTTTCTGTACCAATACTCACTGCAGTTCTTCTTGGATATATTCAGCTCTGTATTGCTGAATAACCCTAAG
 CTTGATTCTATCACCAGCTACAACATCCGTCTGGGCATAATAACTAAAGATCTGTTTGCCGTTTGCTACGAGCGTGTT
 GCCAGAGGTATGTTGCACACTGATCGCCTGACTCTGAGCATTCTGTTGTGTCGCATCCACCTGAAGGGCGTGCCAGT
 GAGCCCTCTCTAGACCAGGAGTTCCAGTTCTTCTGCGCGGCAAAGAAGGCGTCTTCAACGCTCGAGCGCCTGCCATT
 GAAGGACTGTCTCAAGAACATCATGATGCTATGTTGAGACTGACATCTAGATTGCCAACGTTCAAAAATCTGGCCAAG
 AAGATTGAGGATATGCCGGAATTTAGTGCTTGGCTGCAACAGAGCACTCCTGAGAATTGTGTTTCTCGATTGTGGGAA
 GAGGAGAAAGCTTTGAGTCCGATTGGCACTGCCATGTATCAACTTCTACTCATAACAGGCGTTCCGACCGGACCGTGTG
 ATTGCGGCTAGCTACTTGGTGTGCTGAGCTGTGCTGGGCGAGGGCTTACGCGCTCGGCGGAGCGCGAACTCGACCTG
 GGCACCATAGTGGAGACCGAGCTGAACGCCAACGTGCCGGCGCTGCTCTGCTCGGTGCCCGGTTATGATGCCAGCTCG
 CGAGTTGACGATCTGGCCGCCGAGAAGGGCAAGCAGATCGCCAGCATCGCCATCGGCTCGGCGGAGGGCTTCAACCAG
 GCCAGCAACATGATCAACGTGCGCGTCCGGTCCGGGCGCTGGGTGCTGCTCAAGAACGTGCATCTCGCGCCCCAATGG
 CTCGTGCAGCTCGAAAAGAACTGCACAGCCTACAGCCGCATGCCAACTTCCGACTGTTTCTCACCATGGAAATCAAC
 CCTAAAGTACCAGTCAACCTTCTCCGAGCCGACGACGATATTCGTGTTTCGAGCCACCGCCTGGCATCAAGGCCAACCTG
 CTGCGTACATTACGACGGTACCCGCTCGCGCATGATGAAGGTGCCAACGAGCGCGCTGCTCTCTACTTCTCTCCTC
 GCCTGGTTCCACGCCATAGTTCAAGAACGTCTGCGCTACGTGCCACTCGGCTGGGCCAAGTACTACGAGTTCAACGAG
 TCCGATCTGCGAGTGGCCTGCGACACTCTCGACCTCTGGATAGAAGCCACTGCTATGGGAAGAACCAACCTCAGTCCC
 GAGAAAGTGCCCTGGGATGCTCTGGTCAACTGCTCTCGCAGTGCATCTACGGTGGAAAGATCGACAACGATTTTCGAC
 CAGAGACTGCTGGCGTCTTCTGGCCAAGCTGTTACGCGCGCGAGCTTCGAGTCGGACTTTGCGCTGGTGGCGAAC
 GTGGACGGCGTGGCCGGTGGTCCGGGAGGCCACAGACACATCACCATGCCCGACGGCACCCGGCGCGATCACTTCTTG
 CACTGGATTGAACTCTCTCCGATCGGCAAACCTCCATCTTGGCTCGGTCTGCCCAATAACGCCGAAAAGGTTCTTCTC
 ACGACGAGAGGAACGGATCTGGTGAGCAAGCTGCTGAAAATGCAGCAGTTGGAGGATGAGGATGAGCTGGCCTACAGC
 CGGGAGGACTCGGCCGGCCTGGATGCCTCGGCCAGAGAGAGCCAGATCGACGGCCGCCCCGCTGGATGCGCACCCCTG
 CACACGTCCGCTCCACCTGGCTACAGCTTCTGCCCAAATCACTTCAGACTTTGAAGCGAACAGTTGAAAACATCAAG
 GATCCACTGTACAGATATTTTCGAGCGAGAAGTGAACCTCTGGATCTCGTCTACTCCGCGATGTGATTCATGATCTAGAG
 GACGTTGTGGCCATTTGCCAGGGCGAGAAGAAGCAGACCAACTACCACCGCGCCATGCTGAGCGACCTGGTGAAGGGC
 ATTCTGCCGGGCGAGCTGGCGCCGCTACACGGTGCCGCGCGGCTGCACCGTCATCCAGTGGATCACCGACTTCCGCAAC
 CGCGTCTGCAACTGCAAAAAGTCTCCACGCTCGTGTACAGGGTGGAGCCAAGGAGATCAAGACCTTCCCAGTGTGG
 TTGGGAGGACTTCTCAATCCAGAAGCTTACGTGACTGCAACTAGACAGTGCATTGCTCAGGCGAACAGTTGGTCACTG
 GAAGAGCTAGTGCTGGATGTGAACATCACTGATTCCAATGATCCGGCCGTTGACGAGTGCAGTTTGGAGTTGTGCGT
 TTGAAATTGCAAGGAGCGCAGTGTGTAACAATCAATTACTCTTGACGTCAACTATCATGATGGATCTACCGTTGACA
 CTTCTTCGATGGGTTTCGAGGTACCCCGGCACTGAAGTTCCAGGCAAATAACCTTCCCGTGTATTTGAATTCTACC
 AGAACCGAGCTTCTGTTCACTGTGGATTTGAACATCGCTCCTGGCCAAGACCACACAGTTTCTATGAGAGAGGAGTG
 GCTGTTCTCACCTCTACAGCTCTCAACTGA

Bicaudal D (BicD)

>TRINITY_DN45988_c0_g1_i1.p1

ATGGAGACAACGGGACATGAAACCGTATCAGATTTGAAACAAGAAATCGAAAGATTGAGTCGTGAATTAGATCAGGCA
 AACAGTGAAAGAACTCAGTCCGCGCAATATGGTCTTGTGCTTCTGGAAGAGCAAGAGCATTTACGACAAAGATTTGAA

GAGTTGGAGAGCTTGTACGAAAATACTAAGCATGAACTTGACATTACTCAGGAGGCCTTGGCGAAATTCCAGACTACT
CACAAAGTAACAACAAAGTCCGGCATTGAACAAGAAGAGTCATTGCTGTGCGAATCTGCGGCTCGAGAAACGTCATTG
AGCTCTATAATTATTGACCTAGAAAATGACCAAAAATCTGTCCGTCAAGAATTGGAACGTGTTCAAGAGGAACGAGAT
CGGATGATCCAGGAGAATGCAGAACTTGTGAAGGATGTACAAGAAGCTGAAGCCGAGAGAAAGAGTCTACGAACAGAG
ATAAAAGAAATGAAGTTCAGGGAAACCAGGCTACAAACCGACTACGCAGAATTGGAAGAAGAAAACATTGCCCTCCTC
AAACAAGTTTCTGGACTCAAATCATCTCAGGTGGAATTTGAAGGTGCTAAGCATGAAATTAGAAGGCTTCAAGAGGAC
TTGGAAGATCTTCAACAACAAGTGAACGAAGCTAAGAAATTGAAAGAAATAGCCGAGAAACAAATGGAAGAAGCATT
GAAACTTATCAGATTGAAAGAGAAGCAAAAAATGCATTGAAGAAAGAACTTGACAGCAGGATCAGCAGCGAATCCATG
TATAATTTAAGCAACCTCGCCTACAGTATACGTGGAAAAAGAAGAAATGCAGCCGAAATATTTGGGTTCATCCAATAA

egghead (egh)

>TRINITY_DN19137_c0_g1_i1.p1

ACCATGTTGAGCAGCAGAGCGAAGCACTGTCTCCACTGCCTCCTACTGTTGCGGTGTGATTGTGCTTTTTCGAACTGTTT
ACCGGGCGGCTACGGCTTAAAGAAGACAATACCTTGGACTTTGATCCCTGGGAACGCTACGGTATTTTGGGGGCCTCC
TTACTCTACTTTCTACGACTCCTAACCTTCCTTCCACTGCCTCAAGTGCTCCTTAACCTTCGCCGGCCTGACACTCTAC
AACGCGTTCCCCGATAAAGTCACATTAAGGATCACCCTGTTGGCTCCGTTTATTTGTATACGGATTGTGACTCGG
GGTGATTTTCCCCAGCTGGTTAAGGCTAATGTTGAACGCAACATGGCTAAATGTACAGAGGCGGGGTAGAGAATTTTC
CTATTGGAAGTTGTGACAGATAAACGAGTGGGACTTCCCGTTACCCGAAGAATACGCGAGGTGGTTGTGCCAACGACC
TATAAAACCAATCGGGAGCGTTATTCAAAGCTCGTGCCTCCAGTACTGTCTCGAAGACGATGTTAACATGCTGGCA
GACACTGATTGGGTAGTGCATCTCGATGAGGAGACCCTCCTGACAGAAAATTCTGTGCGCGGTATTCTGAATTTTGTG
CTGGATGGGAAGCATCATTTTCGGACAGGGTTTGATAACGTATGCCAATGAAGAGGTGGTGAACCTGGGTGACAACACTG
GCGGATAGTTTTTCGTGTGGCG

exuperantia (exu)

>TRINITY_DN8816_c0_g1_i1.p1

ATGGTATCATCAAAGGTGTCAACATCAGCATCGGTGCCGGAAGAGTGTCTGCCGGAAGGCAGGTTCCGTCTGGTG
GGTTGGGACGTGGACGCCACCGACAAATGGTGACTGACGAAATTTGCCAGCTGGCAGCTTACACGCCCCGCGAAGAC
TTTTTCGCAGTACATAATGCCCTACAAGGATCTGACGCTGTTTGCAGCAGCGCCGTACATGATGTACACGAAAAACAATC
GGCAGGTTTTCGCGCGCTCAAGGACAAAACCAACAATAAGATGCTTGAACCAATCAGAAATTTAGCTCTGGTTGAC
TTTCTAATGTGGCTGGAAGAGATGAAAAATCAAGAGCCAGGCTGTACTGGTCTTATTCTGCTCTGCCATGACGCCCTC
GAATTAGCACCTCTCCTTTTGGTGCAGGCTCTCAAGAGGTATGACCTGTTGGACCGTTTCTCGAGCATTGTGAAGGGC
TTCAGCAACTGCTATGCGTTTATCAAGTTCAAGTGCGAGTCAAGTGTCACCTCGTTACGCTGAAAACCTGTGCGAAA
ATTCTGCTCGACAAGGAAGACGAACAACTCATAGCGCAAAAAATCGAGCGAGAATCGCTTATCAGTTGGTGCAGCAC
TTGCGCGCCGGCCTAGCGGACAATAGCCAGAAGGGCGACCTGATCAGCGATCTTGGCGAGCTCGGGCAGTTACAGTGG
ACGGTGAGCGACAAGGAGGCGTCAGTCGCGCCGCGACTCGAGCAGCTGCAGCGACAGATGTGCTCAGCCGCTGTTT
AAGTCGTGCTCGCCGAACGCAGCGACCGAGTCGCAGTCGCAGGACAAGCCGCGGCGACCCAGCATGCAGGAGAGACGT
CGCGCTGTGCGAGCTGAGGACTCAACTGCTTGAAGCGATGATTGACTATGACACGTTGAGCAAGACTTGGTCTGAGGGT
CAGAAGGAGAAGCTGGTTGGCCTGCTGGAGGCGGCGTTGAGCGAGTGGAACGAGGCTGACCGGCGCGAAGTGGTGGCA
ATAGCCGTGCGTCACTTCGAGTGTCCGCCGAGGGCCCCGACGCCAGGGGCGCCGCAAGTCAGGGACGGCGGCTAGCGAT
GGCGGCAACTCGACGCCGCCCAACGCCACCTCCTCCGCCTCTGCCTCCTCCGCCTCCTCCGCCTCTCCGTCCAGC
CTCACGCCCCACCCACCATCAACAGCACCACTACCGAATAG

Integrin alphaPS5 subunit (ItgaPS)

>TRINITY_DN39338_c0_g1_i2.p1

TTTTACCTTGAGAACAGACAGGAGAAGTGTGTTGGTGCTTCACTTTGTGCTACAGACTTCAATAACGATGGAATTGAA
GACATTATTGTTGGAGCACCATTTTTTGTCTACAGCGAAGAATCGTTCAATGAAGGTGGAATAAGAATAAATTTTCG
AATGAGAATGATTGGCGCAATCCGTTACGACAAGTTATCGTTATGGTCAACAGAAGAAGTTAGGAGCTCAGTTTGG
CTTTTCAATTGTCTGCTATTGGGGACTTGAATAAAGATGGATTGCCAGATATTGCGGTTGGAGCACCCTATGAGGATGGT
GCAGGTGCCATTTACATATTTCTTGGTGATAAATTGAAAGGCATTTTCAAGCGTTTTAGTCAGCACATTGCAGCATCT
GAAATTCATCATTCACTACTCGGCTTTGGAATAAGTATGACTGGTGATCAAGACATTGATTCAAATCTTTTCAATGAT
ATTGGTGTGGGAGTTATTTATCTAGCACAGTTGTTCTAATAAAATCTAAGCCAGTTATCAAATTAAGAGGAGATCTC
AAAATATTATCTGAATTTGAGCAAGGGGTAATCAGTGTGAACGATAAGCCTCTTCTTATTTCTGCTTGTGTGCATTAT
GAAGCAGGACAGTTGAAAAGTGATACGGCATATGTGCAATGAAAATTGATGTGGATAAAAAAGAGAAGAGAACAAAA
TCAGTTTCAGGTCAATTCCACCTTGCCATTTTGAAAGGAACAACAATTGAAGCTGTGCGAGAAATCCCCTACAACACC
ACGTTCTGTTTGAATTTACAGTTTCAGTTGCAAGAAGAAATAAAAAACCGAGATGACGCATTGTTGATTGAAATGCAG

AATTATACTTTGAATAACGATGAACGATGCGAAAAAAGGAATGCCCAATGACAGATCCGAAATTGGAAGGAGATCTG
ACAGAACTGAGACTGCCTTTTCGAACTCATATTGCTGATCTACAGCTGTCTTTGGAGTTCATCAATATATCTGGATCT
ACATACAGGCATACCTCAGCGAAAAAGGATCTAGCAATTAGAGTGAAATTAATCAACAAAAGAGATACGGCAATTGGG
ACCAAGCTGGATATTGAGATGACTTCAATTCTGAGTGTTAAATATATTGAGGGAGATGAAAGATGTAAATCATTTACT
CAATTGGAAACAGAGGCTGGGCGAACTCATTATTATTGCAACTTCAGGAATGTAAAAAGGAACGAAGAGCGGACGATT
ATAATGCATTTATCCACCAGTGATTTAACGCCAGAAGTGGATGAATTAATTTTCAATTTCACTGTGTCTAGCTCCAGC
AAAGAATCAGAAGGTGATTTGAAGAATAATAATGCATCCCTGGCATTACTTGTGAGGTGGATACTCCGTTGTTTATC
TCTGGTTATCCAAATCCTACAGAGATTATAATTTCAAAAAGATGAGAAAAAGTGACATTTTCAAGCATGTTTATGATGTG
TCTAACAAAAGTCTATGAAGTTGGCTGTGCGGGTCAACATTACAGTTCCTATTCTGTGATAATCAGCGATGCTGGA
AGAGTACCGATGTTTCTTCTTCAGGGTAACTGCAGTTACCAAATGATGAAGAAATTACCTGCTCAAATGACGCAGAT
GAAAATTTGAAGACATTTCCTCACCGAGTGGATAGATATACCTGATCAGAATTTTATAACAAGTCTTGGCGAGAACTG
AAATTGAGAGCTGAGAAAAATAATGAGAAATCATCTCGACTGCAACTATCAATCACAAGTATTGAGCACAGAACTCTTCC
ATTGCACTGGCTTGTGTCAACATCTCCTGTTCTAGTTCCTATTGCAAGGAAATGCAACCATTGAACTTGAAATGAGA
TCTTCATATAAAGTATTTGAATATTTGATTGACACTGATGGCATGAGGAAATTCGTAACCTGAAGCGAGCGTATTTCTC
TCTCACTATCGGAAAAAATTCCATTTAACCGGCATTGCGGAGGCGGTTGTTGAAACGAAAAATAACAGTAGACATCCAT
AAGAATCCAATTTGGCCATTGGTTCTTGCCATTGTTCTCACTGTCTATCCTGCTCATAATAATTACTTTTGGTTTTATAC
AAGGCGAATTTTTTCAAGCGGAACAGAGTAAACAAAGTTATTTCAACTGACGCAGCAGCGACAAGCACCGACAATTTG
TTGACTGAATAA

cappuccino (capu)

>TRINITY_DN5275_c0_g3_i3.p1

ATGTCTCGAAAGGACAAGAACATTGGAAGCGGAGGCTTTCTGGAAACCCTGTTCCGGCCGAAGAAGACCGGATTCAGA
GGCGGCTCCAGTCAGCTTCCGAGGCCTCGCTCCGAAACAGACTTCAATGAAACCGATGAACAAGTGGATGTCATCGAG
CACCTCGACGAACAGTCACTCAATAGTAAATTCGAGGAGATGTTGGATGACATGAACCTTATCAGAAGAAAGAAAGGAA
CCCCCTCCGAGCTCAACCTACTATGAATAAGAAGAAAATGCTCATTAAATCATTATAAGGGTCTAGTGCAAGAAAATAGA
AACAAATTCGATAAACCGGCTGAGTACATCCAGTATCTTTCAACGCCAGATCTTAGTGTGAATAAAATGTACAACCTGT
ATAGAATCTCTAAGGATAGCCCTAACTAATAATCCAGTTAGTTGGGTCCATGAATTCGGCACCACCGGATTAACAA
GTCTTATCAATATTGAACGAGTGTTATCGAAACGACAGCGCTACGCAATCCAGTCCAGTGCTGCGCATCTGCGATCTG
CGCGCATCATGAACAACACAGTCCGGCTGAAACAGATGTTTCGGCCAGAAGGAGGCGTTGACGGTAGTTGCGCGGTGCG
CTGGACCTTCCAAAGCCGTCCGTAATGCTGGAGGCGGTCAAAATGCTGGGCGCCGTGTGTCTGGTGCATGACTGGGAG
GACGACGGGTTCTCCGGTCACAAGAAGGCGATCGAGGCGATCACAATGAGTGGAGAGCTGAAGGGGAGAGAGCGGTTT
ACGCCCCGTGTCGAAGGACTGATGATCAACGGCAACGAGCCGCTGAGAGTCGCTGTTTACAACCTGATCAATGCCATT
GTGTCCACGCCTGATGATATGGACTTCAGGCTGCACCTCAGAAATGAGATTATGCGCTGTGGACTCATGGATATTTTG
GAGACGTTGGAGAAAGATGCCAGTGAAGATCTTGAAAGCATTGAAAATATTCTGGACCACAAAGAAGAAGACTAC
TACGAGTTCATTTCAGAGGTTTCGACAATGTGAGACTGGAGTTGGATGACGTCAATGACTGTTTTGAAACAGTCAAAAAC
CTTGTCATGGATACACCGGCCGAACCTTATCTACTTTTCAGTATTACAGCATTTATTGTTTCATCAGAGATGACTCTGTG
ATAAGAAGTGCATATTATAAGTTGATAGAAGAATGCGTCTCGCAAATAGTTCTCCATAGAGGCGGGTGTGATCCTGAT
TTCCGAGCCACGAAAAGGTTTCAAATAGATGTCCAACCGTTAATTGATACTCTAGTCGAAAAATCAAAAGTTGAAGAA
GAAAGACGATTAGAAGAGGTTCTTGGAAGCTTGAAGAAGCGTTAGCAATGAGACAGGAAGCTGAAGCCAAATTAGTT
CAAGCCGAAAAGACAATCTCCGAACCTTCAAGCTGGTGGGGCTACCAAGGCACCGTCATCGGGCGGCGGCCACCAACC
CCGGCTCCACCGCCACCTCCGCCCATGCCCATGCTGCGTCCGCCGGCTCCGCCAATGGGAGCGTCCAATGCTCCGCCG
CCCCCGCTCTGCCCGGCATGATGCCTCCCGGGGGCGGACCTCCGCCGCCGCTCCCATGCCAGGCATGGGGGGACCG
AGGCCTCCCCCACCCCCCGTGGGCGGATTTGCCGTGCCCATGCCGCCCAAGCCCCCGATGTACTGCCGCACGGCATG
AAGCCCAAGAAGAAGTGGGAGGTTGAGGGACTCAAACGCGCCAACCTGGAACAGATCGTACCTCAAAAGCTATCCGAG
AAATCGTCTGGGTGAAAGTACAAGAAGAAAAGTTAGCGTCACCCGACATTCTCAACGGACTAGTAGCCAAGTTCTCG
TCGAAGCCATCCTCCAAACGGTCGGACGACGTGATAGACAAGGAGGCACCTTGAAAAAGTCAAAGCTCTACGGGTA
TTAGATGGAAGTCTGCTCAAAACCTGTCCATACTGCTAGGAGGTAGTCTGAAGCATTATAAGTTATAAGGATGTAAAA
AGATGTATATTGAGATGCGACTTCAATGTACTGTCTGCCAGTATTCTAGAACAGTTGATCAACTACCTTCCTCCGCC
GATCAACTTAAACGACTAATTGAGCTAAACGTTGACTACAACCAGTTGACTGAGGCTGAACAGTTTGTGTAAACGATT
GCAGAAATTAAGGCTTCTTCCAAGGCTGAAATCAATGAGATTCCGGCAACAATATCCCGAGTTGGTGCAGGATATA
AAACCGGATATAGTTGCTGGGGCTGCTGCTTGCAGAAGAAGTGAAGCATAGTAGAAAGTTCACTCAAATATTGGAACG
ATTTTACTAATGGGCAATTACATGAACAGTGGTTCAAGAAACGGGCAAGCGTTTGGCTTTGAAATCAACTTTCTCACC
AAGCTGACAAGTACAAAAGATGTTGAGAACAAAACAACGTTGCTGCATTACCTAGTCGACACGATAGAGCGCAAATTC
CCCGATCTGTTGACGTTCCGCCGACGAGCTGACGCACGTCGACAAGGCAGCCCGAGTGTGACTGAGACCATCGAAAAG
GCATTGCGCGGCATGGAGGCACAGGTGCGCAACCTGGAGACAGATCTGAGCAACAGTCGACAGCCGAGTGCATGAC

GATTTGTTTCGCCGACACAATGTCTGCGTTTCGCTTCGGACGCCAGGAAACAGTGTGGGTTGCTGCAGAACATGTTCCGT
AACATGGAGACGCTGTATGCGGAGCTAGCCGAGTACTACGCGTTTCGACAAACAGAAGTACACTCTGGAGGAGTTCTTC
ACCGATATTAAAGGCTTTAAGGATGCGTTCCGTTCAGGCGTACAAGGACAATCTCCTGGTGAGGGAGACGGAGGAGAAG
TCTCGAAGGGCAAGGGAGGCGAGAGAGAAAGCGGAGCAGGAGAGAGCCGAGAGGGCCGCTAGGAAGAAGCTGGTCGAC
ATGACTAACGATCAGACTCAGGAAGGTGTTATGGACAGTCTGCTGGAGGCTCTTCAAACGGGCTCGGCCTTCACGAGA
GACCAGAGACGCAAACGGCAACGTCTCCGAGGGTCGATGGTGGACCCGACAGGGTCTTTTTGCGATCGTTCTCGTGT
CCGTCCATCTATCAGCACGATGTCTATAGAACGGAAGATCGTACACTTTTTGAAATTTAGAATTTAA

Dynactin 1 (DCTN1)

>TRINITY_DN22476_c0_g1_i1.p1

GAATCAAAAGATGATATGGCTGATTTGGCCGATACTCTGGAGATGGCAACTCTGGATAAGGAAATGGCCGAAGAGAAG
GTGGATACGCTCCAAGTGAATTGGAAAGCGCCAAAGAAAAAATAGAAGAGTTGACTCTAGATTTAGAAGTCTCAAA
GAAGAACTGTTCGGACAAAGGATCTGGAGAAGCTTTCGCTCTCATCTTATCAAACAAAGCAATTGGAACAACAGAACAAT
AGACTGAGAGAGACTCTAGTCAGGATGCGTGATCTGTTCAGCTCATGAGAAACACGAGACTCAAAAACCTTCAGAGAGAT
CTAGAAACTAAAAGAAGTGAATTTGCGGACTTGGCTAAAGCAAATGAGAAGCTTCAAGCAAGACTAACAGATATGGAA
AGTCAAGTTGCCGATCTCCATGAACAGGTGGATGCTGCTTTGGGTGCAGAGGAAATGGTAGAGCAACTCGGCCAGCAG
AAGCTCAGCTTAGAAGAACGTGTCAAGGAAGTGGAGGAAGCAGTCAATGACTTGAAGCCATTCAAGAAGTGAACGAT
CAACTCCAGGAAGACAGTAAAGAAGTGAATCCGAAGTGAAGAGAAGAGGTTGATTTGGCCAATGCTGCCACTCGTGAG
GCATTGAAGCAGAAAGAAGCTGTTCTTGAAGTCTTTCGCTGATAGGGAATTGACAATTTTGAAGTTCAGGGAAGTGT
CACAAATTGCAAGACCAATGTATGGATTTGCGTGCTCAACTTGAACGTGAATCATCTTCTAAGGCAACAATTGCTGCT
GCAATTCCTGAAATTCAAGACTTCAAGAAAATGTTTTCGCTGAAACGAAAGCTCACGCTCGTGCCATAGACTTGGAGTTG
AGGCGTATTGAGGTTTACCAATCACAACAGCATATACAATATTTGGTGGCCTACATGCCTGACTCTTTTCATGAATAGA
GGAGGTGATAACGATGCTCTGCTGGTCTCTTGTGATTCTCGGCTGGTGTGGAAGTGCAGCGTGCTCATTTCCCAG
GCGAAGGACAAGTTCAGGCGCTCGAGCTGGTTCGATAAGAAGAGTCTAACCAGGCGCATTCGGTCGAGCAGTTCGCG
TCGAGAGCTCGCCTGGTTCATGCACTTGCACAACCTGCAGGGCATTCTCAGGCAGTTCATGTTTCGGCCTCAACACGTGC
AGTCCCAGACACTGCTGAAGGTGGGAGGAGCTCTGCCTGACATGATCCAACAGGAGAAGGTGGTCGACAGCTATATA
GATATGATGAAACGAGATCAGTTGGATGAGAACGTGAACACAGAAGCTTTGGAGCGTTGTGTGACGCTACTTCAATACG
ACACATCCAGTGCTTCTTTTGGCCAATGGAGAGACACAAGTCCACCAGAGCTATCTCGTTCAAGATGTGGGCAAGGCT
CTTCTGGCTGCTACAGATTGCATCAGAACTGATTCGACTATCATTTCAAGTATTGGCTCAGACGGGTCCCGATCCGAGC
AATATAATGCTGCTCTGCCAGCACCTGTGCACGGTGTGCGAGGTGATCCAGCAGCAGCTGAAGCAGATCCGGCGCAAG
AGCGTGGACTGGCCGAGCTGCCGCTGCCCGCGCTCGACCTGCACCAGTGCACCCTGGCAGCGGCGCGCTCGCCAAG
GTCATGCGTGACGTGCTTAAAGCTGCTCTCGTACAGATACCTGTTTCAGGAGAGATAGAAGGTTCAATAGACGGTGCC
AAGCTGTTGGAAGCACTGAATGCGTCTTGGGAG

oo18 RNA-binding protein (orb)

>TRINITY_DN2621_c0_g1_i1.p1

CATGTGAATGTGGAGTGGCCCAAGGGTGTGAACAAGGGCGCTGGACTGGACAACCAGAACCACAAGAAGGGATACGTC
TATGTGATTTTTTGAATGATAAGCAAGTTAAAGCATTACTAGCAGCCTGCACACAAGATTTAAGCAAAGGGCCAGAT
ACGCTGTTCTACCCGGTATCATCGAAGAAATTCAAATTGAAATTGGTACAAGTGATACCGTGGGCGACCTGCGACAGC
AACTACATGAAGTCGTCGCACACCAATCCTAGCGCCAAATGCACTGTGTTTGTTCGGCGCATTACACGGCATCATGCAC
GCCGAATTCCTCTTCAACGTCTTTGACAGTCTCTTTCAAGGCGTCGTCTTTGCAGGTATTGATACGGATAAATATAAA
TATCCGATCGGCTCAGCCCGTGTGATGTTTCGACAACATGAAATCTTATATGCGAGCTGTGAATGCTCAGTTTGTGGAC
ATCAAAACTCAAAAATTTCAAAAGACGGTCCAAGTTGACCCCTACCTGGAGGACTCTGAGTGCAGCTACTGTCCACC
CAGCAGGGGCCCTACTTTTGAAGGCGTCCGCTTGCTTCAGCTACTTTTGCCACAACCTGCTGGCAGGTGCGCCACCCG
GCCAACACACCCCTAGGCGCTCACCAGCCGCTGATGCGCAACTCGAAAACGCCATCGGGCAGCAACCTGGCAACACAG
CAGCTGGCGCAGCACCGTCGCTCAGCTCAAGGCCTGAGCCTGACGTACAGCTAGGTGGCCGAGCAGCAGCATGACG
TCACAGCTGGGAGGCCGTAGCAACAGCATGACGTACAGCTAGGAGGCGTACCGGCCTATATTGGGGCAGGTCCCGTT
GTCGGCGATGTAGCCGCTGCCCTGCTGAACAATAATACTATTACTGGAAAAACGACAATGCTCCCGCTTCAGCTGTG
GGCCAGCTAATGTTAATGCCGGTGTGGTGGTATATGATCAGGGGCAACTTGGATTAA

18 Wheeler (18w)

>TRINITY_DN7562_c0_g4_i1.p1

GTTCTGAATTTATCGCACAAACGCCCTCACTCGTATTGATGCAAAAACATTCAAAGACTTATTCTTTTTGCAGATCTTA
GACCTTAGGAACAATTCAATCGGTTTCATCGAAAACAACGCGTTTTTACCCTCTATAATTTGCACACTCTCAATCTT
GCCGAGAATCGTCTTCATCATATCACAGCTACTCTTTTCAACGGTTTATTTGTGCTGACCAAATTAACACTGAGTAAT

AATTTGATTGTGAACATCGACCCTCAAGCATTCCGTAATTGTTTCAGATTTGAAAGATCTTGATCTCAGCTCGAACGCA
ATTCAAGAAGTACCGGAAGCCCTAAGTGAACATAAATTTCTCAAAACATTGGACCTGGGTGAAAATCAAATCAGTACT
TTCCGTAATGGTAGTTTCAAAAATTTACAACAGTTGACTGGATTACGTTTGATCGACAATGACATTGGTAACCTCACA
AAGGATATGTTTTGGGATTTGACAGGGCTTCAAGTTCTGAATCTAGCCAAAAATAAAGTGCAACAAATCGAAAGAGGC
ACATTTCGAGAGAAATACCCAATTAGAAGCGATCAGATTGGATGCAAATTTTCTGTGCGATATAAATGGTGTGTTTTCA
GCATTGGCTAGTTTATTGTGGCTGAATCTCTCCGAAAATCATCTAGTTTGGTTTGATTACGCTTTTGTTCCTACAAAT
TTGAAATGGTTAGATGTTTCATGGTAATTACATAGAAAACTGGCCAACTATTATAAACTGCAGGATGAGTTGCACATT
AAAACCCCTTGATGCTAGTCATAACAGAATAACTGAACTGAATGAATTGTCTATACCAAACAGTGTTGAAGTTGTTTTTC
ATAAATAATAAATTTTCATATCTCAGGTGAAAGTGAACACGTTTTTGGAGAAGAGAAATTTGGCACGAGTTGATATGTAT
GCTAATGAGTTGACTAGATTGGATTTGAATGCTCTTCGATTGGTCCCGGTACCTGAGAATAGAACTTTGCCAGAATTT
TACTTGGGTGGAAACCCATTTCACTGCGACTGTTCCATGGACTGGTTGCCGGTAATTAACAACATGAGTCTGAGACAG
CATCCCCGAGTGATGGACTTGGAGAATGTCATGTGTAGGATGAGCTACTCAAGGGGCTGGTTACCTGTGATTTGATGCT
GACTCCCGTCTCTCCGAGTTCTCTGCAACTATGAAACTCATTGCTTTGCACTGTTGATTCATGTTGATTTGATGCT
TGCGACTGTGAGATTGCTTGGCCGACCAACTGTTTCTGCTATCATGACCAGACGTGGAGCACTAATGTGGTTCGATGCT
TCAAGACAAAACAACGATCAGATCCCAAGACGAATTTCCCATGGATGCTACTCATGTCTATCTGGATGGAAATAATTTA
AAAGAACTACAAAATCATGTATTTATCGGAAGAAAGAATATGCGCACACTGTTTGTAAATGCTAGTAGAATTCAAACT
ATACAAAATAGAACTTTCAATGGACTGAATGCTCTACGTGTTTTACATTTGGAGGACAATTTAATCAAAGAATTGAAG
GGTTATGAATTTGATCATCTTTCTCAACTGCGTGAATGTTTGTGAAAATAATTTGATAGCTCATATTGGTAACTCA
ACACTATTCCCATTTGAAATCGTTGGAAATTTCTGCGTTTGGATGGAAACCGTCTTGTGAGCTTCCCCGATGGCAGCTG
ACCTTGAATCTTATCTAGTCGAAGTAACGTTGGGCAACAACCGATGGAGTTGCCGCTGTAAGTTTCTGAAAGAGCTG
CGATTATGGGTGTGCGATAACTCTCAGAAAGTGGCTGATAGCGGAGACATCTGGTGCTACGCGAACGAAACGAGAACA
CCCTACCGGCGCGAGGTCGACCTCAATAACACCGCCTGCAGTGATTACTACTCGGACAGTTCCGGTGATTTCAGAGTCTG
ATCGTTTCAGATTACTTGCCGATGGTGATCGTTACACTATCGGCACCTGTGATAGTTGTTGTTTTGAGTGTGGTAATG
TTCATTTTCCGCGATTCCGGTGAGGGTGTGGTTGTTCTCGACGTACGGCGTCAGACTGTTCCAATTCAAGGCGCCGTC
GCCAAGCACTACGGGGAGGATAGGGAGAACTGTACGATGGGTATGTTGTATATAGTCCGAAGGACGAGGAGTTTGTG
CTGCAAAGCATAGTGGCCGAGCTGGAGCACGGCAGTCCGTCGTTCCACTTGTGCCTGCACTACAGGGACCTGCCGCAC
TCGGCAGCGCCCTACCTGCAGCACACGTCGCCGGTGGTGGTGGAGGCGGCCGAGGCCAGTCGGCGTGTCTACTTGTG
CTTTCAAGAACTTTTTACAGACCGAATGGTCGAGGTTTCGAGTTCCGTTTCAGCGCTCCACGAAGCGTTGAAAGGCCGA
GTGTTCAAATTTGGTGCTGGTCGAGGAGGGCAGTATTCTGCCCGAGGCTGAGCTCGACCCCGATCTGAGGCCTTACCTG
AAAACGGGCACAAGGATCAGGTGGGGCGAGAAACGTTTTGGGAGCGACTGCGCTATGTGATGCCCAGTGGCGGCGAGT
GATAAAGAGGAAGTCGTGCAATTACAGGAGGAACATCAACAACATATAGTCTAGACGGAGGCGGGCCAGGCGACAAT
CACGCTTATCTGCAGGAGAAGACGCTGCAGGCGCCGTACCCGCTGAGAATCTGCCGCGGCGGTACAAGGAGGACGCC
AACTACTCGTCAGCGACAACGCGACGCCGTGCCGCGACCCGTGCACCGACACCACGGCTACGCCAATCCCCGAGAGA
CCCTCGTCCGAGCACATCTACTCGTCGATAGACTCGCCTGACTACACGGGCTTCGAGAGGAGCGACTCGCGGACGCAG
CAGCGGTGCCGCGCTCAACAACCACAGCCGCCGACCCGCACTCGCAGCCGCCGCCCGCTGGAGGACGGCAGGC
GCCGCCGTAGTCAACAACGGAGGTGTCCAGGCCTACCTAGTCTGA

brainiac (brn)

>TRINITY_DN915_c0_g2_i2.p1

ATGACTGCACAAACAAAAGAGGAATTTCAACCGAAAGAAGATCCTCCAACCTACGAATGCAGCTGCGTCGGATAAAAAA
GTGAAACAGGAAGCAGTATGCTGTTTTTGTCCAATACGACCATGTCCTTTGATTCTGCTCATGCTGATAACGCCAATC
TTTGCTTGGGAATGTACTCGGCTTTCTATGATCCACACTACGTCTATGTGATTGTTGCCTACGGCATAACCATCATCC
AATAACACCGGTGCGCAAAACAAGACAGATTTACCACAAGCCGACGCATCCACTCAAAAAGAAGGTGAGAAGATTTTC
TATCTGCCATTTGAAAATCATTTTCCACAATCATCAGAAGAAAAATCGGGTGAACATCTAGCTGACAGCGATCAACAG
AACGACACACTGCAATCAAAAATCTCCTTTTTCTGCTAGATTTTCCCTACTACTATCGTTTGATATTCTATGGAGAC
ACTGATGAGGGAAAATCTTCAGGGAAAGCTGCTACAAGTGGGGGAAAAGTGTGGAAGGATTACAAAGTGTGGTGAAC
AATCTACTAAAGGAACAGAAAATTAAGGAGGACACATCCCCAGTTATCGTTAAACTTCAGCAGAAATGAAGCCAAACT
GATAAACCAGCTCTGATAAAAATGAGATCCCTGTTTCTGAAACAACCGCTTTGGCTATCATCGACCCCTACCAAGAG
AGACATTTCCATTACGAGGAGGAATTTCTTCAACATTTCCCAACCCCAATCTCTGTGAAAACCTACGGTGAAAACGTCACA
ATAAATCGGTATACACTCGTTTTTGTGTGGATTTTGTAGAGGCGGAAAGCCATCAGGGAGTCTGGGCCCAGCTGGCT
GATAACCAATTCTACAAATACGTATTCTTCTCGGGTTCCCAATGTCTAATGACGTCATGAACAAAGTAATGCACGAA
AGTCACAAGCATCACGACATCATACAAACATCGTTTCATTGAGGGCGCAAACACGGAGTCGGCACAGAGTGTGATGTTT
TTGAAATGGGTTTTCGAAGGAATGTCCAGAACACAGTGGGTATGAAGACTGAGGATGATGTTTTTGTGAACCTCCCC
TATGCTTTGTACTCGATGAGTGAAAAACATTCTGCAGAAAGGGCCATCTATGGTTTTGTGGATACTGTAGAGGAAGGA
GCAGAGTTACATACCCATCATGGTTACCATGATGGGGTATCAACTACTGCAGTTACAAGCAAAGTAGATAAAAAAGGA

CAAAAATTCCCTGCTTTTGTAGGTGGTGCTGGTTACATCCTTCCTGGGTTCTTAGCTGAAGAGCTGTATCATGCCGCA
AGTGATGTACCAACCATAACAACATGAAAGTCTTTTTGTAACAGGAATCTTAGCAGAAAACTGAACATATCCCGAGTG
GATGTTCCATGGTTTGATGCCTCCTCTGTCACCGACATCCCCAGAGATGACCCCTGTTCTATAATGTCGGTCTCCTTC
ATAATTCATCCGTTCTCGGCTGACGATATGACGTCATTCTGGAACATGTTGAGGCAGAAGAGTTTCAACTGTGGTAAT
TTCAAGTATTTTGC GGC GAAGATGAAGTCGGATTACACCAGTCTTATGTATTTCGTCGAGTTTGGATAATCCGAAGGTG
AATGTGGTTCAATCGTTTCAGTGTGACTCCGTATCAGGGGAATGAGGAATCCAAGCGATTTGAGTGA

squid (sqd)

>TRINITY_DN4198_c0_g1_i7.p1

ATGTCTAGGCTAATAGTGAAAAACCTACCAACAATGTCACCAGTGAGAGTTTGAAAGAGCATTTTCAGTGAGAAAGGA
GTGGTGACCGATGTACAACCTGAAGTACACGAAGAAAGGCGAGTTCAGACATTTTCGCATTTGTTGGCTACCAGTCCGAA
GAGCAGGCCAAAGCTGCAATCGAATACTTCAACAACACCTACATACGTTTCAGTACGTATTGAAGTGCAGCCGTGCTCA
GAACTTGGTGATGCAAGCAAAACCGAAAATCGTGAGCAAAATACGCTGCAGACAGTTCTGCCTACAAGAAAACCTGCATCAG
GATGAGATCAAAGCAAAAGCCAAGGCGGAGAAGGAGGCCAGAAAAGAAGAGCAGGAAGGAGGCCAAAGTTCAAGAACAAA
ACTGAAGAGATGATATCGGAGAGAATGAAACAGCACAAAAACGATCCTCTATTCTTGAATACTTGGAGGCTCATGGC
AAAGATGCGGCGAATATTTTGGAAAAGATGAAAGTGTGAGAAGACTCCGGAGTAGAAGACGAAGATGATGACAATAAG
GAAGAAGAGGCCAAAAGAAGTCACTGCAGCCACAGCTGAAATTTTCGGATTTGGATTATCTCAAGTTGAAAATGAAGCCG
GGCGCCGACGAAAAGCAAAAACCTGCAAGAAGTGGTGCAAGAGAAAAGAAAAGGAAGAAGCGGCGAAAAGCAAAGTATTTTC
ACGGTGAAAATCAGCGAACTGCCGTGCACAGTGAAGAAGTGCGAATTGAAGGAATTGTTTCAGAGAGCTGCATCCCACT
TCCATCCGCTGCGCGCGCCACAAGGGATTTCGCTATGTGCGGCTTCAAACCGAACAAAACATGCGGCAGGCGCTC
ACCAAGAATAAAAGCTTTGTTGGGAATAAACAGATAATAGTGAAAGAATTCAAAGGAGAAGGTGATAAAGAAGAAACA
AAGGAGGAGAAAAGAGAAAAGATGGAAACGTCAAGAGGAAGCCTTGATGCAGGAGGAGAATATTGGGGAGTCAGGCCGG
ATATTTTGCTAGAAATCTAGCCTACACAGTCACCGAGGAAGAACTGGAACAGTTATTTCAGTAAATTTCGGGCCAGTGACA
GAGGTTAGCGTGCCGGTTGACCGACACACACGTCAGTGAAGGGCTTTGCGCACATCACCTTTCTCATGCCAGAGCAC
GCGGTGCGCGCGCACACACAGCTGGACGGCACAGTGTGTCACGGGCGCATGCTGCATCTGCTGCCCCGGAAGGCCAAG
ACCGAACACAAGATGAGGACGAAGAAGAAGGAGGCCAGTTCAAACAGAAAAAAGCGAAAAAACAGAAAGAAACGGCG
ACGCAGGCCTATAACTGGAATACACTGTTTCATGGGGCAGAACGCTGTGGCAGATCTCATAGCTCACAACCTACAACACC
ACTAAGAAGAAAAGTTCTAATGGGCGATGACACAGCTGTTTCGCTGGCCCTTGGCGAAACTCAGATTGTGGCCGAGACC
AGAGAATTTCTCGAAGACAACGGCGTCAAATTGGAGGCCTTCAATCAGGCACCCAAGCATCGTTTGAAGACTGTTCTG
CTGGTGAAAAATCTGCCAGTGAATACGACAGCAGCTGACATTTCGTGACATGTTTGCTAAGCATGGTGAAGTGGGCGAGG
GTTGTTCTGCCTCCCTCTGGAGTCACAGGAATCGTTGAGTTTCTGGAGCCGTCTGAGGCAAGAGCCGCTTTTTGTGAAG
TTGGCCTACACCAAGTACAAGAACCTGCCGCTGTACCTGGAATGGGCACCCGATGATGCTTTTACATCACCATCGACT
AAACAAGGCGCTGCTGGTAGTGGAGAGATTGATAAGAAAGAAGATAAAATGAAAGAGATTGAGGAAGAAGTGAAGAA
GAACCGGAACCGGACACTACTCTTTTCATTAAGAATCTGAGTTTTTCAACCAACGAAGAACTATCAAAGATCATTTT
TCGAAATGCGGGAAAGTAGCCAGCGTAACCATTGCTAGAAAAAAGATCCAGCTAAAGCCGGGGCCCTACTGTCAATG
GGATACGGCTTTGTGCAATTTTATAAAAAGAAATCGGTGGACCAGGCTCTGAAACAGTTGCAGCAGAGCACTCTGCAA
GGACACAGTCTCGAGTTGAAACGTTCCAACAAGACGCTGGCTGTGAATGTGGCTGCCGAACGCAAGAAGACAAATGTC
GAAAAACAGACTGGAGTGAAAATACTCGTGAGAAATATTCCCTTTCAAGCCACACACAAGGAAATCGATCAGCTTTTC
AGGGTCTACGGAGAGATAAAATCGTTGAGGTTGCCAAAGAAAATGGTTGGAACAGGACCTCACAGAGGATTCTGTTTT
GTCGAGTATTTGTGCAAGAATGATGCCAAGCGAGCAATGTTGGCGCTGTGCCAGAGCACGCATCTGTACGGCAGGCGG
CTGGTGCTGGAGTGGGCACGCGCAGAGGACTCGCTCGAAGACATCAGAAAGAGAACGGCCGACAACCTTCTGGCCAAA
GAACCATCTCATCAAGAGGCAAAAGGCTACAGCAGACCTAGAAATCGGAGAGAACGACCCTGATGATTGA

chickadee (chic)

>TRINITY_DN499_c0_g1_i1.p1

ATGAGCTGGCAGGATTATGTTGATACACAATTAGTAGCAACAAAAGTGTGTTACAAAAGCAGCCATCGCTGGATTAGAT
GGAAATGTTTGGGCGAAAACAGACGGATTTGAAGTCTCAAAAGATGAACTATTAAAATTGGTGAATGGTTTTCGAAACG
GAAGACCTCCTAACGTCATCAGGAGTAACGCTCGAGGGAAACCGATACATATACCTTTCCGGAACAGACAAAGTAATT
AGAGCGAAACTCGGCAAACTGGAGTACATTGTGTAAAAACGCAAAAAGCTATTGTTGTATCTTTGTATGAAGATCCT
ATCCAACCACAGCAAGCTGCATCTGTAGTCGAAAAGCTCGGAGACTACTTAGTGGGATGCGGTTTTTAG

Egg-derived tyrosine phosphatase (EDTP)

>TRINITY_DN2829_c0_g1_i1.p1

ATGGAGGAAGAAATAGACGGGTGCAATGAAAGTGAAATCCAAGATCTCCTCATGTATTTCTCTAAAAATACGTACCGA
GCAAAAGAATCTGACCATGAGAGTCAACGAATATTAGATAAGTGCATTGAACTTATATCTTTGGACTACAATCATGTG

TTGGTGAACAACCTCGAATGGAGAGCTGAGTGCCCACTACCCGGCTCAAATAGTTATAATGGAGCACGAGAGAGCGCAG
CAGCCCCGGTTCGCCGTCACCGGAAGTCCGCTCACAGGACACAATCTATGAGAGCATGTGTGATGTGTCCAAGTTGAAA
GAACTGTTTCGGAAAAGCGAAGCCGGCCAGGTGTGAGGACGATTTCTGTTCAGTCATTTACTTCAAAGGAAAATAT
GTTTGCAGATCAGCCACTCTGTCTGGTGAATTGAAATCTATGGCAGATCAGGCCTGGATTTTCTCTCAACAACCTGAA
ACCGATTGAGATTGAGTGCAATTAGAAGAGGAGATGTGCGGGCAACTGGCAGTTATTCAACCGAGTGCGCAGTCATGAC
ATTGCGCTGCTGCAAACTTTGAATATTGGCACCATTGTTGACTTCATGGTGGAAAAGAAAAAAGTAAAAATATGGAATA
AATGTTTCTCTTCAGAGAAAAGTGGACAAAAGAAAAATCGATATGCTGACTTCAAGATTGTCTCCCTGCCATATCCGGGC
TGTGAATTTTTTCAAGAACTATCGAGACAATGATTACATCGCCAATGGTTTTGGTGTGTTGACTGGGAGCAGGCATATGCA
GATGCAGTTCTGGAGGTGCCTGATGAACCCATCGCCAATCAACTCAAGATTAAGTGAATGATTATAAGATTTGGGAC
CTGAAGGAGATGACGGAGAATACTACGTGCGCCTGCTGGTGCCTACCTTGTGGAGGGCAGCTGCGGCATGCTGGTGCAT
TGCATCAGCGGCTGGGATCGCACGCCGCTCTTCATCTCGCTGCTGCGACTGTGCTGTGGGCCGACGGCGTCAGCCAC
CGCTCGCTCACCACGCGTCAGATCCTCTACCTGACGCTCGCCTACGACTGGATGCTGTTTCGGCCACGACCTCGTCGAC
CGCTCAGCAAGGGCGAGGAGATTCTCTTTTGTCTTCTACATGCTCAAGGCCATATCGGGCGAGGAGTTGAGTGTC
AACAGGAAAGAGAAGCAATCTGTAAACAGCCGACGCCAGTTATAAGAACCGACTCTGAAACGCACCTGGAAGGCGTT
CTATTTGAGGCTGGCTCCCAAGGCTCACTGAATTCGTTCATGGAGCTCCCTCAGCAGCAGAAGCCAAGAACTGCTAAC
TTCTATTGTACTGATGATACTAATGGCAATGCCGTCCAGGTGCCAGTTCGTACCGCTCAGCCTCAGTACTCGTGTTCG
TCAGCCAGCACGTGCTCATCAGTGGCCAGCTGCACCGATCAGCTGCGATCAGCTGATCAGCTGCGATCGACCGACGAC
GAAGCCCACAACGGCAACGCCGCCACCTCAACCATTCCAAACCCACACGAGTCTGTGGCTGTGCCTTCCAGACCT
GTGGTCAGACATCGTAACGAATCCACCGGATCTCTTAATGTGCGGAGCTGGCAATTTATTACCGGTTCTGGCAGCCTA
AGAGGATCATCGTCTTCCAGCTCAAATGTAACGACCAGATTGAGTGACTCTGCATTTTCATGAGGATGAACCTACGAAA
TTGGCTGGTGACACACCGACCGGTCCCACTCAATTAGAAGATTGTTTCATCTACAGTAATGGAACACTTGACGCTGTG
CCTTTACGGAGCGAGAGATTGAGAAAAGTACGGGAATTGTTTCTCTACTGTTTTTGTCTCAACCATAGGACTTAGAACG
AAGGATGGCTCCGAGAAAAGCTTAACTAGCCAAGCGTTGAATAGTATAGTTGAGCAATGCAACAAATATGTGTACAAT
AAATATTAG

Dynein light intermediate chain (Dlic)

>TRINITY_DN1812_c0_g1_i8.p1

ATGTCAAATGTACTATAGAAAAATCGGATTACCTGGGTTTTCAATAAATGATAAGGAAGAACAAGATGAAAAGGAT
GATAATCTATGGACCTCTATCCTTAGTGAAGTTCAAACAAGAGGTAGCACCAAATTGCCATCATGTAAATCGGTACTT
GTATTGGGTGACAACGAATCTGGTAAAACAACATTGATCGCAAACTCCAAGGAGTTGAAGATCCAAAAAAGGATCG
GGCTTGGAGTATGCATACATAGATGTTTCGAGACGAATATCGAGAAGATCACACAAGACTCAGCGTGTGGGTGCTGGAC
GGCGAGCCGTCGCACGCGCACCTGCTGAGGTTACGCTGACGCCCCGCTCGTTTTCCGCACACCCTCGTCATGTTTCGTG
GTTGCCATGACCACGCCCTGGAGTCTCCTAGAACAATTAGAAAATTGGGCCACAGTCTACAGGATCACATCGATCGG
CTCAAATTTCCACCCGACCAGCTGCAGGAGTACAGGCATAAGTGTACGAGGAGATGGCAGGAGTATGTGGACGAAGGT
GAAGATCTAACTTGGATTCTCATCTAAGGAGAACTTCAAGAACTTGGATGGCGAAGACGGTGTGGATGGGGATGTT
TTCCCTCTAGGCGAGGGAGTGTCTACAAGAAATTTGGGCCTAGATATAGTTGTAGTTGTAACCTAAGACGGATTATATG
TCAACATTAGAAAAAGAATCCGACTATAGGGACGAATATTTTCGATTTTCATCCAGCAATGGATACGAAGGTTTTGTTG
CAATATGGAGCTGGATTATTCTACACATCAGCCAAAGAAGATAAAAAATTGTGATTTACTCTATAAATATCTCACTCAT
AGAATTTATAGTTTTTCTTTTTCGAACACCTGCACTTGTGCTGAGAAAGATGCTGTTCTTATACCTTCTGGATGGGAT
AGTATGAAAAAATTAGTATATTGTTTGAATAATATGCAAAGCATGAAGCCAGATGATTACTACAGGGATGTAATAGCT
CCTCCAGTAACAAGAAAGACTGTGACGCGGGAAACGGAAATCCAGGCGGAGGAGGAGCAGTCGTTCTTGTGCAAGCAG
CAGGCGCTGCTGCAGCAGGGCGCGCCGCCATGCGCCAGGAGTCGCCGATGCGGTGCGCCGGCCGGCGTCCAGAAGACG
GGCGACCGCCGCATACAAGGATCACCTTCGGTACAGGGATCCCCAAGAAGATGGATCCTACGAAGCCAGGAGCTGCC
GGAAGTGGCGAGGGTGTGTTTAGCCAATTTCTTCAACTCGTTGTTGCATAAAAAGACTGGAGCGGCGGGAAGCCCATCT
TCATTAAAATCTCCAAACGACCTGTGACAAGTCCAGCGCCGCACTCGCCGGTGGACAAGGCGGCGATGCGTTTCGGAC
GCGGCTGCTGAGCTGGACCGACTGACGCGCAAGAAGCCCCCCCCACCCGCTCATCACGCTCAACAACCTCAACTCGAGC
TCCTCCAACCTCCAGCCATCGACTCCCGACCACAACCTCCTCTGAATGCTAG

cornichon (cni)

>TRINITY_DN23649_c0_g1_i1.p1

ATGGCTTTCAATTTCCCCGCATTTTCTTACATAAATTGCTTTAATAAATTGATGCTTTCTCTCATTTTCTTCGCAATCTTT
CACGTGATTGGCTTTGATGAATTGAAAACCGATTACAAGAACCCTATCGATCAGTGTAACAGTCTCAACCCGCTTGTT
TTACCGGAGTACCTCCTGCATATTTTGTTCACATTTCTCTTCATATTCTCTGGACAATGGTTACCTTGTGCCTGAAC
CTACCGCTGATTGCCTATCACATCAACAGGTACCGCACCCGACCAGTTATGTCTGGACCTGGTCTGTATGATCCTACA
AGCATCATGAACGCCGATGTCCTGACGAAATGCCAAAGAGAGGGCTGGATAAAGCTAGCATTCTATCTTCTCTCCTTC

TTCTATTATCTTTATGGGATGATTACGCTCTAATCACTACATAA

Heat shock protein 60C (Hsp60C)

>TRINITY_DN10510_c0_g1_i1.p1

ATGTTTCAGGTTACCTCTTGCACTTCGTACTACAGCCATAAGGCAGCTGAACTCTCAGCTCAACCGCAACTATGCTAAG
GACATACGCTTTGGCCCAAGAGTGCCTGCGCTCATGCTTCAAGGTGTCGACATTCTTGCCGACGCTGTTGCGGTTACC
ATGGGACCAAAGGGAAGGAACGTCATTCTCGAACAATCGTGGGGAAGCCCCAAAAATAACAAAGGACGGCGTGACAGTA
GCTAAAGGCGTTGAGCTGAAAGATAAATTCCAGAATATTGGAGCCAAATTAGTTCAGGATGTAGCCAATAACACAAAT
GAGGAAGCCGGTGACGGAACCACAACCTGCCACAGTGCTAGCTCGCGCCATTGCCAAAGAAGGATTTCGATAAAAATTAGT
AAAGGAGCTAATCCCATTGAAATCCGAAGAGGTGTAATGCTTGTGTTGACGCCATAAAGGAACACCTTAAGGCACTA
TCGAAGCCCGTCACAACCTCAGAAGAGATTGCACAGGTGGCTACTATTTCGCCCAATGGTGATAAGGCTATTGGTGAA
CTTATCTCTGCTGCAATGAAGAAGGTTGGAACGGAGGGTGTAATTACTGTCAAAGACGGCAAAACTCTTGAAGATGAA
CTGGATGTCATTGAAGGAATGAAATTCGATAGAGGATACATATCTCCGTACTTCATCAACACATCTAAAGGAGCTAAA
GTGGAATTCCAAGATGCTCTGGTTCTACTCAGTGAGAAGAAAATTAGCAGCATCCAGTCCATTATTCCAGCCCTGGAG
CTAGCCAACCTCTCAGAGGAAGCCTCTGGTGATTATTGCCGAGGATGTTGATGGAGAAGCTCTCAGCACATTGGTCGTC
AACAGGCTCAAAATAGGATTGCAGGTAGCCGCTGTCAAAGCTCCTGGATTTCGGCGATAATCGCAAGAGTACTTTGCAA
GATATGGCGATCGCCACTGGCGGCATTGTTTTTCGGCGACGAGGGTGCCAATCTGAAACTGGAGGACCTCACCCCGACC
GATCTGGGCCAAGTTGGAGAAGTGGTCATCACCAAGGACGATACTCTCCTGCTCAAGGGAAAAGGCAAAAAGTCTGAT
ATCGATAGAAGAGCCGAGCAGCTGAGGGATCAAATTCAGGACACAACATCTGAGTATGAACGTGAAAAGTTGCAAGAA
AGACTAGCTAGGCTTGATCTGGTGTGGCTGTTGTGAAAGTTGGAGGCTCGAGTGAGGTGGAAGTGAATGAGAAGAAG
GACCGCGTGACGGATGCTCTGAACGCGACGCGGGCAGCCGTCGAGGAAGGCATCGTTCCCGGGGGCGGCACCGCTCTC
CTGCGCTGCGCCGCCAAACTGGCAGAAGTTGTACCGCCAACTCCGACCAGAAAACGGGCATTGACATTGTGCAGAAG
GCGCTGCGCATGCCCTGCATGACCATTGCGCAGAACGCCGGTGTGGATGCGTCGGTAGTGGTCAGTAAAGTGCTGGAG
GGCAAAGGCGACATGGGCTACGATGCTCTCAACAACGAGTACGTGAATATGATTGAGAAGGGAATCATCGATCCCACT
AAGGTGGTAAGAACAGCACTAACAGATGCAGCTGGTGTGGCATCACTGTTAACGACAGCTGAGGCGGTGGTCACCGAA
CTCCCGAAGGAAGAGCCGGCCATGCCAGGAATGGGCGGCATGGGTGGTATGGGAGGAATGGGTGGTATGGGCGGTATG
GGAATGTAG

maelstrom (mael)

>TRINITY_DN1714_c0_g1_i4.p1

ATGAGTTCTAAGAAGAAGAAGCAGGCAGCCCCGAACGCTTACCTCATCTTTATGCGGGAATTTCAAGAAATGCAGAGA
AATCAGGGACGTAACCTATACTATGGAAGAATGCGCCGTTTATTGCTCAGATTTGTGGGCGAAATTGCCGCTCGAAGAG
AGAGCCTACTATCAACACCAGGCTGACGATGCAAAGAATAGACTCCGAAACGGAAAGCTGAATAGTAAAGGGGAAGCT
TTACAAAAAATACACCAGTTCCACGTTAGCGAAGAAATGAAACTGCGTGAAATGAATGAAGAAATTGAATTTTCAGTG
CAATGCCTCGATACCAAAACACTTCTCTCGAAAGTCTTTCATGTGATTTCATATTAATACTACTTCTGCAAAACTGACAGA
GGCGTTTTTTTATCCAGCTGAAATTGCACTGGCGGAGTTCACTCTAAACGACGGAATAATTCGTGTTGTACACAGCTTT
ATCGATTCTGGTGAATTACCTAAAGGATATAGGTCAACGGCAATGGAGCATTGTAACGCCACGCACAGAACGCCGTTG
GACCTAAAGGTTGGAAATAAAATTTTGAAGTATTAGAGTTGAGGACACCGTAGAACTAATTATGGAAACTCTTCAG
GCGGGTCCAGGAATGGGAGAACTGCCTCGCCTATATACGATGCCTGATGAGGTAAACACAATTCTCAATGAGTCCTGC
GTCTTAAGCGTCTTGGGAGAACTTTGCAGAATCACTAAAAAACCCGAAAACACATTTAAAGTGTATAATTTACCGAAG
CTGTTTTTACGAATTGCGCAATAAAGCTCGTAGTGACACATTCCCCTCTGTGCTTTTAGCATTGAGTGAGCTGGAAAAA
GACGTTTTTCAAGTATACAAACGGAATCGCTTGCCCGTACCACGAAGAGGAGGACGTTACTTCTTATTGTTCAAAATCA
ATTGTTTACGCGATGGGTGTATATGATTTGTGACCAGTTTGCAAGGATCTCGGCGTGCGAATGCTGCCTGGCTTTTCAT
TGTCCCCGCAACGCTGATCTGCACGGATCGTCAAGATCTAGGGTCTTTTCTACAAGTGACAAAGCAAACGTTAAGACT
TCTGCCTTGAAAGAAAACCGTTCTACTAATTATGTACCACAAAAAGAGCCGCAAATTCGTGAGCCGCATTTTTTCGAGG
GCCAGTAGCTATTATGACTCCGATGCAACGAGTGTAATTCAGATATGATGGGAGCGAAGGATAGTGACTCGCCGTCG
CAGTGGAGAATGCCGCGAGAAGAGGAGGGGAGGTCCAGTGGAGGGGAGCCGCCGCCCTGCTCCCGCCCCCTGCCCCC
CAGCAACGCGGCTGGAGCAACGCAGCCAGCACTAAAGGCGCAGCTGTTCCGCAGTTCGACAACCTGGACGACTTTCCG
TCCATCGGCGGCTCTGCAGCCCCCTAAACCAAAGCCAAGTGCCAATGAAAATCGTTGGGGTGCCCGTGGACGAGGTGCG
CATTTGAGTGATGATTTGAAGGACATGCGATTGGAATGA

Peroxinectin-like (Pxt)

>TRINITY_DN1715_c0_g1_i5.p1

ATGAAAATTGCCATTGCATGTGTTGAGGCTTTTCGTGTACAAGAACCCTGGCGGTGTAAAAATGGTATGCAGCAGTGTG
GCAGTGGCCATTGTGATGTGGTCAGTGGTGCAGTTTGGCAGTCAGTCGGTTCTGATGGCTGCCATGAGTGATGCCTCT

CAAATGCTCAAGGCACCCGCCCTCACCTTGCAAGCCTATTTGGAGTCTGACCAACCCCCCTCCCCCACGCACCCGT
CGTGCGTTGAAGCAGCGTCTGCCCTCCATCTCGAGCAGGGAAGTGGAGACGGCGCTGGACTTCGGCCGCACGTTGATT
GACAAGCAGTCGCGTCACGAGGCCAACATCGCTGGCTCGGGCGTGAAGTGTGAGGAAGGGCACACCCTCGCAAGGCATG
CTGGCCGAGTCGCGAGCCGACCCGCGACGCCCTAGGCCACGCCCTAGACGCACATGTGGCGCTGAAGGGCGCGCAGTAT
ATCGGCAGCAAGTATTGTCTCAGAGCTGGCATAGACGACGAGTCCTGTGCCAACCAGATCTCATTCTTGTCTCTGGGC
TCCACTTCGCTGGGCCAGAATTGTTCTCGCATTACAAGACTGAACTCTACTGTTTCGGAGTTGAACGTGTACCGCTCC
TATGACGGCTCTTGCAATCATAGAACACACTCTGCTTGGGGAAGATCGCTCACTCCCTACAGAAGAGTGCTTGCTGCT
GATTATGAGGATGGTCTACATGTGATGAAGGGCTCCCGCAAAGGAAAGCCGCTAGTGAGTCCCGCCAAAGTGAGTGTC
ACACTGACAAGCAACTTGATCTCCCGGACACTTCGAAGACACTGGCCGTCATGCAGTGGGGACAGTTTATCGCACAT
GATATGGCTCACACCGCCTCCTCAAGAATGTTGACAACCAGTAGCACGATCTCATGCTGTAGTAGCAGCGGCAGAAAA
CTATCTCCGAGACACGAGCACTCCTCTTGTGCCCCCATCGACATTCCCATCACTGATCCGTTCTACTCGAAGCTCGGT
CACACTTGTATGCCCTACGTTTCGCTCACTGATGGCTATCAGATCTGACTGTACTCTGGGGCCTATTCAACAGATGAAC
CAAGTAACCTACTACCTGGACGGATCCGTTATCTACGGTTCAACAGCTGAAGTGGCCACGTCCTCTCCGCCAATTCCAG
AACGGACGTTTGCGCACAAAAATACACAACGAACGCAATTTCTTCCCGCTCCAAGGCTTCCGAATACAAGTGCTTA
CATAACGACACTTGCTTCCAATCTGGTGATGTGCGTGTGAACAGCCAACCAGGTCTAGCAGCTTTGCACACCCTGTTT
CTGAGGGAGCACAAACCGCATGGCCAGTCGTCTGTCCATTATGAACCCAGCTGGGACGACGAACGCCTCTATCAGGAG
GCCAGGCGCATTGTCTATCGCTGAAATACAACACATCACCTACAACGAATGGCTGCCCGTTGTACTAGGTGTGAAATAC
ATGAAGCGCATCGGTTTAAACGATCCAGCCGAGAAGTTTCAGTGCCAAGTTCAACGATCAGATTGATGCCTCGACCAGC
AACAGCTTCGCCACGGCCGCAATGGCCTTCATCAACTCAATGATTGAGGGACGAATCGGGCTGTACGAGGAGGATCGC
AATCCCGAGTTTACCGTGCAGCTGAAAGACGTCTTCAACAAGCCTGCACTGGTGCAGAAGAGCTCGAATCTGGATGCT
TTGATCCGTGGTCTAGCCACACAAAGCAGCCAGAGTATGGATCTTCAGTTCTCAGAAGCTGTAACCAATTACCTCTAC
AGAGAGAACCAATCTTATGGTTTAGACAGACTGAGTATCGATATCCAGAGAGGTAGAGACCATGGGTTACCCAGCTAC
AACAAGTTCAGAGAATTCTGTGGTTTGGCCAAAGTTAAAGAATTCGCACAACCTGACTGACAAGATTCCAATTGAGGTG
GTTAGCAAACCTGAAGAAGCTGTACAAGAGTGTGACGAGGTGGATCTAGTGGTCGGTGGACTGGCAGAGTCGCGCGGG
GAGGAAGGCCTCGTCGGTCTACGTTCAAGTGTATTTTGTCCGAACAATTCACAAGAACAAGGATTGGTGACCGTTTC
TTCTATGACATTGCTGGACAGCCGGGGTCATTCAATGAAGATCAATTGAGAGAAATCAAAAAAGCCAATCTAGCAAGG
ATTTTCTGTGATAACAGTGATAATATCAGGAACATGCAAGCTTTCGTCTTCCACAAGCCATCAACACCAAATAACGAC
TTGGTGCCTTGCGGCGAAGTTTCTTTGATTGAGCAAATGGACTTGGAACCATGGCGGGAAGCGCAATCTCCATCTCTG
TTGCCATTGTAA

quail (qua)

>TRINITY_DN2076_c0_g1_i19.p1

ATGGGCATGCGGGATCTGAAGAGAGCAGTTAGCCGCGAAATGAACGACTCCATCAACAATGTGGATCCAGCCTTCAAA
CTCGTCCCCAGAAATTCTACTGCTTTTCTCATATGGAGGGTTGAGAATCTGACAATCGTTCCAGTTCCCAAAGACCAA
TATGGAACATTTTTCGAAGGTGACTCTTACATAATATACGCCGCCTCGGCTCCAGGGCAATTCTGTTGGGCCAGGAATA
AAAAATCAAGACGTGCGTGGACCTCTAGAAACGCACATTCACTTCTGGCTGGGCGAAGAGACAAGCCAGGACGAGTCA
GCGGTGGCTGCTTATAAACTGTTGAGCTCGATGATTACTTATCGGGATCACCGGTGCAGCACAGGGAGGTGCAGGGC
ACTGAGTCACCGCGATTCAAGTCTTACTTTGAAATGGAATCAGGCTGTTGAAGGGCGGAGTGAGCTCAGGCCTGCAC
CACGTATGGACACATTTCGAGCCGCGCCTGTTCCGCGTGAAGGGCCGGCGCACACCGTTAGTTTCGCCAGATGCCCGCC
ATCGCTTGGGAGCACTTCAATCACGGCGACGTGTTTCATCATCGACACCAGTGACGTCTATCTTTGTGTGGATTGGTAGT
GGGGCCAATAGCATGGAAGGCTTTTGGCCACCAGGGTAGCCGTCCAGTTGCGCGACGAACACAACGCGGTGTCAATA
GTGTTTCGTGGACGACGGCAAGGAAACAGAGTTGCCCGAGTCCGAGCGACTGTTGCTCGGCGTCTACGTGGATCTGTGCG
CCGACGGCGCGTCAAGCCGTCAAGAAATCGGCGGGCGACGACCAACGCGTGGAGCGCTTCGAACGGGCCACGCTCAAA
CTGTACCAGTGCTCGGACGAGGATGGCACCTACAAAGTGGTCGAGGTCAAGTCGGGTCCGCTCTTCCAAACCGACCTC
AACTCCAATGACTCGTTTATAATCGACAACGGGGCGCTTGGCATTGTTGGTTTGGGTGGGGAAGAAGGCGTCTCCGAAA
GAACGCGTGAAGCGATGCGCAACGCTTACGGATTGTTGAAGAAGAAACAGTATCCAAACAACACCCCGTTCAGCCGC
GTCGTGCAACATGGAGAGCCTGTGAGTTCAAAGCGCTCTTTCTCAACTGGACAGACAAGGACATCGCTGCAGCCAAG
CCGACGTTCAACCGCACTAATTCTGTGAACAACAACAAACGAAACCGGTTGCAACGAAACTGGACGCCGCTCGCTG
CACGCCACGCCAGCTTGGCCGCTGAAACGCAATTGGTCGACGACGGCAGTGGTGAAACTAACATATGGCGAGTGCAC
CAATCACAGCTCATCCAGTGCCCAATTACAGCACGGAGTGTCTACGGCGGCGACTGTTATCTCATACATTACAAG
TATTCTGCCATCGGACGGCAATCTCATGTGCTTTATTATTGGCTGGGACTTCACTCGTCAGTGACAGAGCAGACGGCG
CTGGCGTTTTCACACAGTGGCCGAGGACGAGAAGCTAGACGGCGTGGCCACCCAAGTGCGCGTCTGTGAGGGCAAGGAG
CCGCCCCACCTGCTCGCCATATTCAAGGGACGCATGATCACACTGCTCGGCGATCACAAGGAAGCCTTTCCCGCCAAA
TACCTGCTGCAAGTGCAGGGAAATCGCATGCACAATGCTAGAGCTTCTCAGGTGCCTTTTAGGATGTGCTCATTGAAC
TCCAACGATGTGTTTGTGATGGTGAACGGATCTAGTTGCTATATATGGTGCGGCAAAGGTTCCACAGGCGATGAAAGG

GAAATGGCGAAAAGTATAGCTTTAAAGATTAGTGGTCCGGAGTTTGTCTGCTATATGAAGGACAAGAGAAAGAAGAT
TTTTGGAGTACAATTGGCAACAAGGAACCATACGCAATGACAAACGATTGGCAGAGCCCAGTAATCCTGTTCCAGCA
CGACTATTTCACTGTTCCAATGCTAGTGGCTCTTTCAAAGTGGAGGAAGTTTTAAATTTTCATCAAATCGATCTGGTG
CCTGAAGATGTGATGCTACTTGATGCATGGACATCCATTTTCATCTGGGTTGGCTACAAATCTAACAAAGAAGAACGT
AATAGTGTCTATTCAGTTGGCATTTCGATTATCTTCGAACAGATCCCAATGGAAGAGACGAAAAAACTCCAATTATTCAA
TTGAAACAAGGATATGAACCGCCAAATTTCACTGGATTTTTTGGATCGTGGGATTCTCAACTTTGGGAAAGTGACAAA
TCTTTTGAAGAGGTGAGACACGAACTAGAAGGCCAGAAGCCAGTGCTTCAAGTGGAAGTGAAGCTGACGAACGGAGTA
AATGATTTTCGATGAATGTAAAAAATATCCCACTCAAGATCTGCTACCCAAAGATCCCGAAAACTGCCAGATGGTGTG
GATCCTCTCCACAAAGAGTTACATCTCTCTCATCCAGATTTTCATAACAGCGTTTGGTATGAAATACGAGGACTTTGCT
CAGCTACCGAAATGGAGGCAAGAGAACCTCAAAAAAGCAGCAGGACTTTTCTAA

Rac

>TRINITY_DN786_c0_g1_i4.p1

ATGTCTGAAAGTCCCGAGCCCCAAAAAACTTTTTGTGAACTGAACGACATTAAATTGAAAACAAATTCGCTCGATAGG
AAAAACAAAACAAGACTATCGTCTTTGATAAGATTTAGTTTGAAGAGGACATTTCAGTATGAGTTGAAAAACAAGGAG
CTTGAGAGAGGAGGGGAGGATGAAGCGGTAACCTAGTACCACAACGACAACGGGAAAGAAAGGAGGGAAGAAGGGGAAG
AAGAATGCGTTGACAAAGACGAAGACGATTGGGAAAACGGAGGGGGGTGAAGGAAGAACAGCCAGCCCGGAGAATAAG
GGATCTTCGTGCAAGGCGTCGTGATGGTGCGCCGACTGAGTCTGGACAAGATAGTGCCGAAGCTGAAGGGAAGCTCC
AAGTCGAGAGAGGAGAACGGCGACGGCCGACATCGACAACGCCCCGACATCAAGGAGCAGGCGGCGCCGGTCGATCTG
CGCCAGGAGCTGGAGGAGGCGGAGAAGAAGCGGCGCGACGACCTGCCCGTGCCCAAGATGTGCGGAGGAGTCGGGCGAC
CTGAGCGGAGACGACTACAGGACTCCGGTCTGCCAGTCGCCCGAGGGCTCCGTGGCCGCCCCCTGACAGTGTGCGGCTCG
CTCACCGTCTCCAGCTCGGAAGCGTCGTTTCGAGTCGGCCTTCCAGACCTCCAAGTCCAGTCCCGACTCGGCCAAGATG
GAGCGATCCGTCAGCGACACAGAAACCATCACGCCGACCGACAACGTGCAAGATCCGCTCACTGGTCCACATCCTAT
CCCGTCGACGCGAAAGAAAATCAAACTCTGTTTCCGAAAAGCCAGTAGTAGCAGCACAGGCATCAAAGCCCTCCTCG
ACCCAGTCGAAGGAGAGACCGAAATTAGAGAAATCGCCTTCCATTAGAAATTCAGAGACACAACAAGCACTGAACGAT
AGAAGAGAGTCGTTGTTCAAAATTCCTGTGATCGGCGAGCTAGGAGCTGGAAAAACATCCATCATAAAACGCTATGTT
CACGAATCTTCTCACAAAACCTACCGAGCAACAATCGGCGTTGATTTTGTCTCTCAAAGTACTCAACTGGGATCAAAAT
ACAGTCGTTAGGTTGACGCTATGGGATATTGCAGGTCAGGAGAGATTCGGTAACATGACAAGAGTCTATTACAAAGAG
GCTGTTGGTGCTTTAATTGTGTTTCGATGTTACAAGAGCTGCCACGTTTCGATGCTGTTCTCAAGTGGAAGCAAGATCTC
GACTCGAAAGTTTCAGCTTCCAGACGGGAGTCTTATACCTTGTATATTATTAGCGAATAAATGTGACCAGCCGAAAGAA
GGAATAGTGAATGACAAAGAAAAGATGAACGAGTACTGCAAGCAAAAACAATTTCTCAGCTTGGTTTTGAAACATCAGCT
AAAGAGAATATCAACATTGATAATGCAGCCAAGTGTCTAGTTGAAAAAATCCTTCAAAACGAAAAGTTCAACCAGAAA
GGTGATTACAAAAGGATGGGAACAGAGTGTGCTTGGAGGCCGGGGTCAAGGTGAACTACAAGATGGCAAGAAGAAG
TTCTGTTGTTGA

Ser/Thr-rich caspase (Strica)

>TRINITY_DN3046_c0_g1_i2.p1

ATGGATTTCGACAATCTGATGGAAGAAAGAACTACCGGTGATTCAAAGGATTTGTGTTTAAAAATATCGCATTTCGGC
CAACTTGGAGAAGCAGGAAGTTATGGTTTCAGTGCTGTTTAGTCAGGAAAAGCTAGACGCTAATATAAACTTGGATGTA
ATAGAACTATTAGAGAAAGATATGGATTTCAACGATATAGTGTCAATTATTATTTCTTGTAAAGTGAAGAGAGGCATTCC
CAGTATATTTTCCAAAGAATAACCCAGTTCTTGAAAATCAAATCAGAAAATATTGAAATATTAGAAAGTGTCAACTAC
AGTCTCATCTCTGACTGGTGTGTTATCAAATCCAAAAAACTGGCGGTTGAAAATAATTGAAGGTTTAGCAATTGTTGAA
TGTTTCAATATACTTTCTAACTAGGCTACAAGAAAATAGATGTACGGATTACTTTCAACCAGATAACTCTGAGTTT
ACTGTATGCATTTCACTATTGAAGAAGAAACTTTATATTGAGTTATCAAACCTCATCAATACTAAAAAATTGAGTCTCTC
TATGATTGTTTGAAGAAGTTGATAACTGCTCAACTGATTTATTATTACAACATTCATTAATAAATAATAGTCGATAC
TTTGAGTTTCTACTCATCTACCTATCGGCAGAATCATTAATAAGATATGAAGAATATCCTGAATTGAAATTCCCTAATA
AATCTACTTGAGAAAGCAAACTTACTAGACTTGAAACAAATTTTAGAAAAGTTATTTATTGAAATTCAACGAAAAAAGT
TCTCTAAGCAAGAATCGGAAAATGTAACATCGAAATCACATCGATCGAAAGTGCCTCCCGATCTTGATAGACCATAT
TCTAAGGATTCTTCAAGAATTAATCAAGTGAAAGATGAAGACTTTGAATTGGAGAAAACAAAAAAGTAGCTCTTGCT
AATGATGATGATGAATTGAACTACTACAAAATTGGAAATACACAAGATCTTGGTATTTGCCTGATAATCAACCAGAAA
AATTTTCATAAAATGTGCGAGAAAGGAATCCTAACTTCAGAGAATTTTGGAGCCATAATTTACTAGTGGATCGTCATGGA
ACTGAAAGAGACGTCGAAAGGTTGCAAGAAACATTCAATCACTTCAAAATTTTCAGTTATTGTTGAAAATGACGTGTCA
CATTCATGATTGGAGAGATTATAAAAGACACAATTGACCGCGAGTTCAAACAACACCATTTCAGTATTCTTCTTAGTC
ATTTTGTCTCATGGAGAAGAAGGAATTATCTACGGAGTTAACAGTATACCAGTCACTATAAACTCGCTGGTTGATGCA
GTCCTGAGATCGTATTCTAAGCTCAGGAATATACCCAAAGTCATAATAGTTCAAGCGTGTCAAGGCAACTTGCCAAAT

CCAGTATTAGAGACAGATGGAGGTAGCTGTGGAGTAGCGATTCTACTCAGCCCCAAATACTCTGATATTGATTCCAGA
AAGAATGAAATTAACAGTGAGAAAATAAGGACTGCCTACCATGATGACCTTTTAATGTGCTTCTCGACAGTTAGTGGA
TACGAATCTTACCGGCATACCCGAGAAGGATCAAAATATATTCAAATTCTCTGTGATAATTTGATGAAATATGGACAC
AACGATGATTTTTTTGAGCATATGTACTCGAGTGAATAATGATCTGAAAAAGCTGTACGTTCTATAAAAAATCAATGAA
GGTGTGCTAATACCATCAACTCAAGTATCTGAGACAAGCTCTTGCCACACAAAAAATTGTTTTTAGTGGAACCCCTTG
AGAGCTTTTCCAACCTGTACTGATGAGCCAAAACCTTTCAAATAGACGTCTCAAGTTGTATTGA

Endophilin B (EndoB)

>TRINITY_DN37749_c0_g1_i2.p1

ATGGATTTCAATGTGAAGAAGTTGATTAACGATGCAGGAAACGCATTAAGTAGAGTAGTTCAGATGACCGAAGAAAAA
CTCGGAACGTCGGAGAAAACCGAACTGGACGCACATTTTCGAAAATTTGGCCGAACGATCTGACCAGACCAAACTGTGG
ACCGAGAAATTGCTCAGTGATACTGAGGCTGTGTTGACTCCGAATCCAGGAAACCGCGTTGAAGACTATTTGTTTGAG
AAGATTGAAAAGAAGAAACCGACACGTCTCACTAATCTCGAGTATCTAGGAAATGACATGATCGAAGCTGGCAACGAC
TTTGGAACCGGAACGTCTTACGGAAGTGCTTTAATAAAGGTAGGACAATGCGAGCAGCGTATCGGCTCAATAGAGAAG
GAATTCATAGCATCGGCCAATCACTGCTTTTTTAACGCCTCTCAGAAAGTTCCTCGATGGCGAAATGAAAATATTTTG
AAAGAACGAGGTCTGTTGGAGAGTAAAGATTGGATCTAGATGCGTGTAAGAGTCGAGTGCCTAAAGCCAGATCTATG
TTGGGACAACCAACCAATGCGGAGCGCGACTTGCGCGTGGCACAGTCCGAGTACGACCGACAGGCCGAAATCACGAAG
CTGCTGCTGGAGGGCGTGAATAGTTTCGCACGCCAGCCACCTGCGCTGCCTGCACCGATTTCGTCGACGCGCAGGTCAAC
TTCTACGAGCGCTGCCATCACACCATGCAGCAGCTCAAGTCCGAAATGCAGAGTGATATGAAACCAAGTCCAGTTCTT
CGTCTCAACAGTGAAGATCTGGACCAATCCTCTCCAGCCCATATTCTGTCACTGAAATAGTGCACCTCTCAT

G protein-coupled receptor kinase 2 (Gprk2) / Darkener of apricot (Doa)

>TRINITY_DN207_c0_g1_i2.p1

ATGGAGGACGCCGTTGAGTACACGGCCGAGGCCCGCAACGGACATGGAACGGTGGCTTGCACGTGTGCTCTGGTGGTA
GACAAGGGCATCAGAGACTACATTGCGCCCGCATTTCTGGCCGAAATCGAAAAGCATGCCACTGTGAGAGAGGGAGGC
ACGCTCAGGCTCAGAGCCACCGTTGAGGCCTATCCTTCCGTTGATATTATGTGGCACAGAAACGGCACCAGATTGCGT
CAGTCACGCAGACTCACGTGCACTCTCGACCTCGAGGGTCACGTGAGTTTACGCTGGCCGACGTGGCGGCGCGTGAC
GCGGGCGTCTACACGTGCACCGCCACCAACGAGGTGGGAAGGTGTGAGACGGTGTGCGAGGTGCGGGTTCACGCCGAGG
CTTGTTGAGGCTCCATTGGAAGACGGCGGTGTTTACGCGGGGTGAAGACGGGTTACCATACTCGCAAGCTCCAATGTTC
ACGGTGAAGCCTCGTTCAACCGATGCGGTGGAAGGCGACACGGTTATCCTGATGTGCGAAGTCGTTGGAGATCCAGAA
CCGGAATCTTCTGCCTCAGGGACGGGCTGAAGCCGGACTACTACCACCACGGGCCGGGGTTCGTGGCGGAGCGGGAG
GGCTGCCGCTACCGGCTGCACATCGAGCAGGCGCGGCTCGACTACACGGGCGCCTACGCCATCATCGCCCGCAACCGC
CATGGTGAGGCGCGCTCTAGTCTCGCTACAGATCTACGCCAGAGGTCAAGGAAAGGAGGGAGTCATGAGCTCAGGC
GAAGTCAAGCAAGGAGAAGTGCAGACGTTGCCCGAAGTGAGGAAAGAGCTGAAGGACATGCGATGCTGCGACGGCGAC
AGCGTGACTCTGGAGTGCCACGTGTACGCCACGCCTCCGCCCGACATTGCTGGCTCAAAGATGGAAGATAGTAGCA
CTGGGCGGCGACATCTCGTCGGATGTTGAGGGCGAGGTGGCTCGCCTGACCATACGCAAGGTGTTCCCGGAGGACGAG
GGCGAGTACACGTGCATGGCCTACAACGAGCTCGGCCGAGCACTCACCTCGGCTTGTCTCATCGTCGATGTTCTCTGAA
GAGAAGGAGACAATGGTTACCAAACAGCTTTCAAACCTCCAGGATTCTCTCAGCTCCTACCACCCCTCGCACCACG
CCTTCCAGGAGTATATCACCTGGAGTACAAGTCACTCGAGACTTCTCGACGCCGCTCAACCGTCCGCGAGCGCAAGCGC
ACAAAGCTGTCGTCGCCAAAGTTTTACGCGGTGCCGCACAACCGAGTGGCCCAGGAGGGCGAAACGGTCAAGTTCCAG
TGCTCGGCGGCAGGACACCCGAGCCCCCTGGTTCGTGCTGGGACAAGGACGGACACATTGTACGCCGTCGGCGCGCCTC
ACTCTGCTCGAGCAGGACGATCTGCGCATTCTGCAGATCAAGGAGGTGACGCCCCGAGGATGCAGGACTGTATAGGGTC
ACGCTGGAGAACGATTATGGCAGTGTGGAGGCCACGGCTAGGCTGGAGGTTATAGCTCACAAAGGTACAAAGACAAAA
GGCGTGAGGGCATGGAGTGCTTCGCGGGCGTCACCGACTTTCGGTTCGTCGCTGGTGGGGTTCGGCAGCCCGAGTGGGG
GGCACTTTTCGCCCTCGCCTGCGACGTGCAGGCGTCGCCATCACCCAACATCACCTGGTACAGGGACACAGAAGCAATC
GTGAAAACGGACCGGATCCAGCCGTTATGGGACGGACGCACGGCCCCGTTTGAAGTGAGCCGCGTCGAGCCTGACGAT
GCCGGCGTGTACACGTGCGTGGCTGAGAACGAGCTGGGCACGACGAGGTGCTCGGCCGAGCTGCTCGTGTTGGCCGCA
GATGACCCGAGCGACGCTGACCTCAGGCGCGCCGTTTCGTGCGGGGACTGCCGGAGACCACCAACGCAACCGAGGGA
CAGCCGATGGAGCTGCAGGTGCAACTGCAAGGAACACAGCCGTTAGAAGTGGTGTGGTCAAAGGACAAAACAGAAGT
CCCGATTGCATGGATTTCCGGTATGTGGATTATGGAGATGGAAGGTTTGGACTCCGCATCGCCGATGCGTTTCCAGCA
GATGCCGGCCTCTACTCGTGCGAAGCCTTCAACAAGCATGGCGATGCAATCACCTACGGACACTTGCAAGTGACAGAA
CAATGCGAGTCAACACCTGATAGTACAAAGAAACAACCTCTACTTCAACAAAACGCCGCTACCTCTGGAAGCCGAGCTT
GGCCAGCCGGCCTGTTTCTGCGCCCGGGTGACACTCCGCCCGACTTGCAAGTTCACCTACAGTGGTGGTTCGCTGGT
CGCCCGGCTTCCCAACTGCCCAACAGGTTCAAGGTGGAACGCGACGGTGACACAACAGTGCTGCGAATACCTTGTGTC
CGCGCGGACGATATCGGCGCGGTTAGTTGCCGCGCGCACGTGTGCGCCTTGGGCGACCGCGAGCCGGTGGCCGGCTGC

CGCGCGTGCTACGTCACTTCGTGCACGTGTGCGGCTGTGCGCCGCACCCGGCATGCACGGCTCCTGCCCCGCCCAGCTG
 CTGCACGGACCCCTTGATACGACGGCGCTGCGCGGACAACGGGTCACTACTGCAGGCGGCTACGCGGGCAGCCCCGAG
 CCGGTTGTGCGCTGGCTCAAAGCGGGAAGAGAGCTGTCAAATAACAGTTGCATGACGATAATCAATGAAGACGGTCTG
 TCGTGTGCTGATTCTGGATCCAGTGTGCGCCGATGACAGCGGCAAGTACGTTGTTAGTGTGAAAAACAGCCATGGAGCT
 GATTGTCACTTTGCTTCGGTGGCGGTGGAAGGGCCCCCAGATCCTCCGGCAGGTCGACCCAGCGTTAGAACGTTCAAT
 AGCACAGCCACAATAGCCTGGTCGAGTCCACCCTACGATGGAGGGTGCATGGTCACCGGCTACAGAGTCGAGATGAAG
 AAGTACCTCAAACGTCGTGGACGGTTGTTACAGACAGATGCCACTCATTGTGCGCACACAGTGCACGACCTGGTACCT
 GGTATTGAGTACGTGTTTACGGGTGAGGGCTGAGAATGTGCATGGAGCCAGTGAAGCCAGCGCCGAGACTGTGACGCTT
 TCTATTGGCTCCAAGTCGTCTTCAGGTTTCGGAGGAAGACGACACAGTGGACCAACAACCAGCTGTGACGTGGGACCGC
 GTCACAGTGGAGCCGGGCGAAGTGTTCACCGGCCGCTACGACATCATGGAGGAGCTGGGCAAGGGCCGCTACGGCGTC
 GTGCACAAAGTGCAGGACAAGGCCACGGGCGAGCGCATGGCCGCCAAGTTTGTGCGCTGCATCAAGAAACAGGACAAG
 CAGAAGGTGTTTGAGGAGATTGAGATTATGAACTGCCTCAGGCATCCCAAGCTGCTGCAGCTGGTGTCTGCCTATGAT
 AACACTAGGGAGATCGTCATGGTTACTGAGTACATAAATGGTGGAGAACTGTTTCGAGAGAGTTGTAGCCGATGATTTT
 ACATTGACTGAGAGAGATGCATGCTTTTCATGAGGCAGATATGTGAGGGTGTGACTACATGCATCAGAATAGCGTT
 GTACATTTAGATCTCAAGCCAGAGAACATAATGTGCTACACACGAACCTCTCACCAGATCAAAATCATTGATTTTGGT
 CTGGCGCAAAAGTTGAGTCCTAGCATAATTCTAAGGGTGTGTTTGGTACGCCTGAGTTTATTCCGCCAGAAATCATC
 AACTACGAGCCCATTGGAGTCGAATCAGACATGTGGTCCGTTGGAGTCATCTGCTATGTTCTATTATCGGGACTTTTCG
 CCGTTTCATGGGTGACAACGATGCTGAGACGTTTTCGAATATCACCAGAGCCGATTTTCGACTTTGATGACGAAGCTTTTC
 GACGCTATTTCTCAGGACGTCAAAGATCTGATCAGCAACTTACTAGTTAAACGAAAAGAGAAAAGATTTAGCGCTAGA
 CAATGTTTGTAGAACAATAATGGATGTGACAAATGTGATAAACTATGAGCTGTGTAATACTTAGTACCGATAAGCTGAAA
 AAATTCATAATTCGTAGAAAATGGCAGAAAACCGGAAACGCAATCCGAGCTCTGGGACGTATGGCCACCTTGTGCGCC
 AGTAGGAGAAATTCGGCCTCACCCTGCGAGTCGAGCAGTCCGCGACAGTCTGCTGGCAGGAGGGGGTGTGTGCGCGCATG
 TCCAGTCTGAACGAGGAGGAACAACAATCATGTGACGACCAATGAGGAAATCCATCAGCAGCACAGTATCAGCGCC
 GAATGCAATGGAAAGAACGGTCGGGTGCGGGTTTGCAGCGAGAGGAGTGACAGTGGAATCAGCGACTGTTCTCTATT
 AACGCGCAACCTCCCAGCAACTCCACGGTTGCCAACCGCTCACCAGCAAAAGCCTTCTCAATCTCGGAAGAAAACGAA
 CTGTGCGATGGTTTCGACAAGCATTGCCGAGTCAAACTAACAGTTTCACAAAAGACCAATGGTGGCATCACTGTGCGAG
 TCAAACTAACAGGTACATGATCAAAACAATCTGATAAATCTAAATCATCTCCAAACGAATCTAGTGATAGAACG
 AGTCTCGGATCGAAGCTGACACGGTTTCGTGATGATGACAAGTCCAGGACACTCAAAGAATTGACGACACCTAACTA
 AGATCTGATACAAGTACAAAACAAGTAGATCAGTTCAAGGAAGCGACAAGAATAAGGTTTGCCGATAATTTAAAGCCG
 AGTTTAGATGAAGCATCTTTGAAAGGTGCGAAAGAGTGAAGATGCGTCTAGAACTAGGTTTGCGGATAATTTGAAGCCG
 GATGAAGCTTCTTTGAAAGGTGCGAAACAGTTGAAAGATGCGTCTAGAACTAGGTTTGCGGATAATTTGAAGCCG
 TTAGATGATGCATCTTTAAAGGGTGCAGAAACAGTTGAAAGAAACAATAAGAACTAGATTTGCTGAAAATTTGAAGCCA
 TGTTTAGATGACGCTCTTTGAAAGGTTCAAGACAGTTGAACGATGCGTCTAGAACTAGGTTTGCCGATAATTTAAAA
 CCAAGTTTGGATGAAGCTTCTTTGAAAGGTTCAAAACAGTTGAAAGATGCGTCTAGAACTAGGTTTGCTGAAAATTTG
 AAGCCCAGTATAGATGAGGCACCTTTCCGTAGTAAAGTTGACGTTGAAAATAATAGAAGGAATGATTCAAAACGTAGT
 GTAGGTGGTTATGATAGTACAAATAAAAGACTGGAAGATAGGAACACAACGAAAGTTAATAGTAGTAAATCCCCTGAA
 AATACTAGTGGGGTAAGTAGTTTGAGAGACGCTAGTAGTGGTTTGCCAACTAAAAAGCCTGCTAATAGTGTGTGAGT
 TCGAGTCCGTTTTTTGAATAACAGCTTTGCAACTGATGACAACAGTTTGAACGAAACAGACTGATGGTAGTAAATTT
 GTGGACAATTTTTTCAAGTCCGACGCCGATAGAAGTAGTGGCAGCTTCAGGCGGGAGGCCTAAAAGTGACGACGGAC
 GCACGAGCTCGGTGCTGAGTCTGCTTTGCAGAAGGACGCCAAGGTTCTCGAGAAGAGTGAAAATTTCCAGAAGGCA
 TTTGCTTTCTGGAACGGTGA

Olfacción

Receptores Ionotrópicos (IR)

>TRINITY_DN11355_c0_g1_i14.p1

CTTTTAGAAGAAGCTTCTAAAGATGAGATTTTTCGATATTTTTGGGTTTTCTGATGCATTTCTATTCTACTAATAGCGTT
 CTTTATCAATTAAGCACCCCGGAATCGGTTAATAGACATTTTTTCAGCAATCGCATTAGAAGCTATAGGCCCATCTAGA
 CCAAAAACCATCGGACAACCTACCAACTTTGAGAGGAAATTGCTTCAAGTGTCCACTTTTCGATTGTCCTGTTTCATGTG
 TTCATGAAGACTATCAACAACAAAACAGTAATCTACGACGGCCTAGAGTACACAATGTTTCGTGGAGATCGCCAGACGC
 CTGAACTTCACCTACCAGTTTTCGGCTGTTGGTGGACGAAAGCAACAAGTGGGGCAAGATGTTGGACAACGGCTCGTGG
 ATGGGCGGCATCATCGGATCCATTCTCGACTCGCAGCGGCAGCGCATCAGCGCGTGCATGTGCGCTTCTGCGGCGTC
 TGGAAGATACATCCGCAGTCTGCTGTTTCGTGATCTGACCACGTCTCTTATACAGGAGAAGAACCCTTCTGCTGCGG
 CAACAGACCGAGGTTTCGCTGGAGCTGGATGTGCGCCATTACCGAGTTCCAGCCCTCTGTGTGGCTCTGCATCGCCGCT

>TRINITY_DN17880_c0_g1_i5.p1

```
>TRINITY_DN19602_c0_g1_i2.p1
```

```
>TRINITY_DN2107_c0_g1_i1.p1
```

CTCACATTCTATGATAAAAAAGGATATCCAGTGACTGTGAGGCAGCTATCATCCGATGATAATCACAGGCCAATTCTT
CGCCGAGTGAAAAATGCTGGAGAAAAATTTATCGTGCTGGACTGTTTCGATTGGAAAGCTGTACTCGGTATTGGAGCAA
GCGATGCAAGTTGGATTGATGGGTGACGACACCAACTACATTATAACTAATTTGGACATGCAGACCATCGATCTGAGC
CCATTCATGCATGGTGGCAGCAATATAACTGGTCTAAGGATGTTCAACCCTGAAGATGAGGATGTCAAGCAAGTTCTT
GACAATTGGAAAGAATGTCTAAACACTCGAAAGATGGAGCGGATTGCACTCCAGTCATTGGCCCTGATTACAAAGTT
CATCTCGAAACAATTCTAGTTATGGATGCTGTGTTCTTATTTCGCTGTTGCTCTTCAACATTTGAGTAATACTATGGAT
TTTGAGTGCAGTACCCCTCTCAATTGCTCTGAAGATCGCAGCTGGGAGCATGGCTACAGTGTCTATAAATTTTATGAAAGTG
AGTGAATTGAAAGGTCTATCAGGCTTGATAAAATTCGACCATCAAGGGTTTCGAACCAACTTCAAACCTAGACATAATT
GAATTGACGTCCACTGGACTGAGGGAGAAAGGCACCTTGAATACAACAGAGGGCGTGAATCTGACTGCCGTTGCTTCA
GGTTCATCTGCCGATCAAGCCGTGGAAGATTTAAGAAACATGACGTTTCATTGTCTTAACCTGCTTTGACCCATCCTTAT
GGTATGCTGAAGGAAACAAGTCACAAGTTGACTGGAAATGATAGATTTGAAGGATTTGGGATCGACTTGATTTCATGAG
TTATCACTGAAATGTGGCTTCAATTACACGTTCCGTTTGAACCCAGATAAGAGCAGTGGAAATCCAAATCCAAAAACA
GGGAAATGGAATGGCATGATAGGAGAAGTATTAGCTGGGAGAGCGGATCTTGCAATAGCGGACATAACTATAACAAGG
GAGAGAGAACAGGATGCTGATTTTACTATGCCTTTCATGAATCTTGGTATCAGCTTGCTGTATAAAAAGCCGATGAAA
GAACCACCAAGTTTGTCTCCTTCTTGTACCATTCAGCTACGAAGTTTGGGGATATATGATTGCCGCTTATATAGGG
GTGTCAATATTGCTTTTCTTCATGGCAAGAATTAGTCCAGCTGAATGGACCAATCCATATCCATGTATAGAAGAGCCT
GATAACTTAGAGAACCAGTTCAGCTTGAACAATTCATTGTGGTTCACCATTTGGCTCACTCATGCAGCAAGGATCGGAA
ATCGCACCAATCGCTATCTCGACGCGTATGACTTCTGCTATCTGGTGGTCTTCACTCTGATCATGGTCTCCTCTTAC
ACCGCCAATCTGGCCGCTTTTCTCACAGTCGAGAACACTATTTCCGCGTTTGAAGACGTTGAAGAGCTAGCTAATCAG
AAGACTATTAAGTATGGGGCTAAGGCTGGCGGCTCTACTGCTAATTTTTTTCGGGATTCAAACCATTTCGACCTATCAG
AAAATGTGGCAATTCATGAAAGAGAATCCGAGTGTCTGACACGATGAAGGTGTGAGACGGGTGATCAAC
CCCGATGAGGACTATGCTTTTCTCATGGAATCCACGCTCTATTGAATATGAAGTCGAGAGGAAATGTGAACCTCAGTCAA
ATTGGAGGACTGCTAGATAATAAAGGATATGGAATAGTCATGCGCAAAAATTCACCTTATCGCAATGCATTGAGCACC
AGTGTATATAAGTTTGAAGAAACGGGAAAGTTGACGAAATTAACCAATGAGTGGAAAGAAAACGTTGGAGGGGGA
GCTTGTGTCAGCAAAAAGATGATTTCAGCTAGTGCAGGCTAGTGAAGTGGGACTTGAGAATGTAGGTGGTGTCTTTCGTAGTA
CTAGTTTGCAGGTGTCTTTGTTGCTTTGTTGCTTCTGGAGAGCTTTTCTACGACGTTAGCACTAAAGAAGAAAGG
GGACCTTTGAGAGAAGAGCTGATGCAAGAGTTGAGGTTCATCACACGATGTGAGGATCAACGAAGCCGGTTCGTAAG
GCTGGTAGTGAAGAGGAGGAAGATGAAGAAGAGGAAACCAATAACGACGTCACCTCTGAAATGCCCTGGGCCCTTTC
TCTACACCCTACCGATCATTCGATACTAATTCAAAACAACCTCTTTGA

>TRINITY_DN40231_c0_g1_i4

GGAAAGGCTCATGTTGGTCTTTCTGCTCTCTATCTTTGGTACAATGAATACCAATATTTGGATTACTCAGTATCGTAC
ATACGCACTGGAATCACCTGTTTGGTCCCGAAAGCCAGTTGATATCTCCTTGGCTTTTGCCGATTCAGCCTTTTGAC
AAAGCATTGTGGGCTATCGTAGTGATGACTCTTGTACATCAACTATCACGCTGTATTATGTGTCAAAAATTACAAAT
ATATTTTCAAGAAATTGGAGACGTAGATTTTGAACGTGTGGATGACAGCGTGCTCAGGGTTATTGGGATGCTTGTTCTC
CAAGTGCCAACAGCAAAAAGGCATAAGACCGCCGCTACTGCATTACGCTATGTTATTACTTTCTAGAAAGTTATGTTT
CTGCTTGTGCGTAACCAGTTATCACAGTGGATTGGCTTCTGTTCTGACTTTACACAGACATGAACCAGCAATAGAAACA
GTGCAGGATTTGGTTGATAGCGGCCTGCAATGGGGTGCCACTCACGATGCTTGGATATTTTCAATTCTGCTCTCCCAA
GACCCAAGAACAAAACTCTAGTTGATAGATTCAAGGTGCTGAATGAGAAAAGTCTCAATGAGCGAACACAACTAGA
AATTTTCGCTTTCAGTTTGGAGCGCTTACAAGGAGGCTATTTTGAATTTGGAGACTATATCACGGAGGAATCAATGAAT
CACCTGCAATTGATGAAAGAGGATATTTATTATGAATATTGTGTTTATGCACTTCCAAAAGGTTCCACATACAAACAA
CCGCTTGATACCTTATTTTGAACATCCAGAGTGCAGGCCTCGTTCTGCATTGGGAAGGCGAGGTTGTCCGCAATTAT
TTGCCAGTAAAAGTGACGATATGTAGTTGCTCAGAGCAGTGCTGCACCTACTGAAAGTGAGCCAATCGTTTTAGAGCTG
AAGCACATCCATGGAGCTCTCACAATTCTTGCCGTTGGATTGGCCTCATCAACTCTGACATTCATCTTGAATTATTT
TGGAAC

>TRINITY_DN45996_c0_g1_i6.p1

ATGACGGCATTGAGAATGGTTGATGCCGAGGATGTGAATGTGAGACGAGTTCTTGAAGACATGCAACGCTTCCAGCCA
ACCGGGCATAGTCTCTAAACAAGACTTCTGTCTACAGGCTGGTCTGCTCTGATGTATGATTTCAGTCTTAGTATTC
GCTGTCCGTCTTCAAACACTGGAGCAGAGCCATTGCTGGTCCCCGCCAATGTCTCGTGTGATGCCGAAATGCCTTGG
GATGGCGGCTTGTCTCTCATCAACTACATCAACTGAGTGAAGTTGAAGGGGCTGAGCGGGCCAATCGCAATTCAAGGAA
GGCAGGCGAATCCAATTCAGCTCGATCTGCTGAAGTTGAAGCAGCAGTCTGCTGGTCAAGTTGGCGAATTGGAGTCCC
GCGACCGGAGTCAACATTTCTGACAGAGCGCTTTCTTCGATTCTGGATCGATGAACGTCACCTCGTTGTGCATACACA
ATATTGGAAACGCCGTACGTGATGATCAAGAGGGCGGAGAACTCGACATCGCAGAAATACGAAGGCTTCTGCATAGAC
CTGCTGAGAGAGATAGCTAATATGGTGGGATTCGAGTACCAAATCCAGTTGGTACCCGATGGGAAGTATGGCGTTTTT

GACCATGAGACTGGAGAATGGAACGGGATTGTTAGGCAGCTCATGGACAAGAAAGCCGATCTGGCAGTTGGATCAATG
ACAATAAACTACGCAAGAGAGAGTGTATAGACTTTACAAAGCCATTTATGAACTTGGGAATCAGTATTCTTTTCAAG
GTACCGACAGACAAAGAATCTCGCTTCTTCACATTCTTGGATCCGCTGGCTTTTGAGATCTGGATGCTAGCGTTTGTG
GCTTTCGTGTGCTAGTTTGCCTTTGCCTGTATGCCATCTCTCGTTTCTCGCCCTACGAATGGTACAACCCGAAGCCGTGG
AAAGGCAAGGACGATCCAAGACTGATGGTGAACACTTTTTCACTCAGTAACAGTGTTTGGTTCATCACC GGTTACTCTA
CTTCGTGAGGGCGCTGGACTCAACCCCAAGGCAACTTCGACAAGAATAGTGGGCGGAATTTGGTGGTTCTTTTACACTC
ATCATAATCTCGTCCTACACGGCAAATCTGGCCGCTTCTTGACGGTGGAGAGGATGATCACTCCCATTGAGAATGCT
GAGGATCTGGCTGGTCAGACTGACATCTCCTACGGCACTTTGGATAGTGGTTCCACCATGACGTTTTTTCAGGGACTCA
ATGATAGAAACGTATAAAAAGATGTGGCGGTTTCATGGAGAATAAAAAGCCTTCTGTATTCTGTCGCGACTTATGAGGAG
GGAATCAGGAGGGTCTTGAAGGGAACCTACGCATTCTTATGGAATCCACCATGCTTGACTACATAGTTCAAAGAGAT
TGTAATCTTACACAGATCGGTGGTCTTTTGGACTCAAAGGTTATGGAATAGCAACGCCTATGGGCTCACCTTGGAGA
GATAAGATTTTATTGGCAATTTTAGAGCTGCAAGAGAAAGGAGAAATTCAAATGCTGTATGACAAGTGGTGGAAAGAAT
CCAGGAGACACGTGTACAAGAAATGAAAAAGGAAAAGAATCTAAAGCCAATGCGCTGGGTGTTGATAACATAGTGGT
GTGTTTGTGTTCTCCTCTGCGGTTTGGCTATTGCAAGTATTGTTGGCAATTTTCGAATTTCTGCTATAAATTCTAAGAAA
ATAACTAAACTTGAAAGGAGAGTAGCCTCGGCACCCCATCAGAGCCTCTGTGCCGAAATGTTGGAAGAGTTGTGCTTT
GCTTTGAGGTGTGCTGGCTCGAGACAAAGGCCGCGCTGAAGAGACAATGCTCAAAGTGCTTACCTGGCACAACCTAC
GTTCCAGCCACATTCGATCTCCCTCCTCCTGTTAATAATGGATTGAGATCGCAGAGAAGCAGCATGGCAGTTTCCCTA
GAAGTGCCACCCGACCTACACGGACATACATCCTCCTGTGATCTTGACCATAGGCCTTGA

>TRINITY_DN7133_c0_g1_i5

TTCAAGTACCATGATGGGCTTCTCAAAGGTGAAGGTCAGACACCGTTTACCTTGCACAACTTGACGAAGGCGATGAC
TATAGACCATTGTTGAAGCAGATCAAGGTGTCAACTGAGACTCACATAGTTTTGGACTGCGCAGCGGGAAAGATTCTG
AGCATTTTGC CGCAGGCCAGGGAAGTGGGAATGCTGGGTGACTATCACAGCTATATTGTGACGTCATTAGATGCGCAC
ACAATTTGATTTTGAAGAAGTAAAGCTGGGTGCGACGAATATCACAGCTCTGCGTATGTTGAACCCGGGTCTAGGCGTT
GTCCAGCAAGTAATAGATGAATGGATATTTGAAGAGCAAAGGAAAAATCGCACACCCGATTTGGTAGCTGATACTTTG
AAGTCGCATTTCGGCATTGATGTACGACGCAGTACAACCTTTTTGCAAGAGGCTGGGCTCAACTGGAATTGAAGGAGCCT
CCGCACCGAGTCGAGTGCAACGCTAAAGACAAGCCAAGCCATAAGAACCCAAACGGATTTGCACTTATCAGCTTCATA
AAAACGATCCGAATGGAAGGTATGACTGGCATGATGGGCTTCGATGAGCTGGGAAGGAGAGATTTCTTCTCCTTGGAA
ATAATTGAACTAGTGAAGACAGGCTTCAAGAAAATAGGATCGTGGGACCCGAAACATGGCATCAACTACACGCGTACA
CAAAGCGAAATGTTGAATGAAGTTATGATGAGCATTTCAAATAAGACGTTTATTGTTACTTCGAGGCTGGGTGACCA
TACCTGGATAAGAAGGAAGGGGAAGGCCTCAAGGGTAACGATCGCTTCGAAGGATATTCCTTGGATCTGATTGACGAA
ATCGCCAAGGAACCTGGGCTTACCTACGAGTTCAAACCTGTGCCGACCAACAGTACGGTTTACTAAATAAAAAGACC
GGCCAATGGAATGGACTTATCAGGGAACCTCAAGAAAGGAGAGCAGATCTTGCAATCTGTGATCTCACCATCACATTT
GACCGGAGGAGTGCGGTTGACTTCACTATGCCATTTATGACTCTGGGTATCAGTATCCTTTACAGTAAGCCAGTCAAA
CAGCCGCTGATCTCTTTTCACTTTCTGCTTCCATTCTCCCTTGACGTCTGGATCTACATGGCCACTGCGTATTTGGGT
ATTTCCCTACTTCTCTATTTCTCCTTGCAAGATGCTCTCCGTATGAGTGGGAGAACCCCCACCCATGCAACCCCGATCCG
GACGAGCTTGAAGTACGTTCAATCTGCCAAACTGTCTCTGGTTCGCAATCGGCTCGCTCATGGCACAGGGATGCGAT
CTTTTGCCAAAGGCAGTGTGCTAGATTGTGCGCCGGCATGTGGTGGTCTTCACTCTGATTATGATCTCGTGTGAC
ACGGCCAACCTGGCTGCCTTCTCACCACGAGCGGATGGAGGATACCATCGAGAGTGTGAGGATCTGGCCAAGCAG
AACAAAATCAAATACGGCGTCTCAACGGCGGATCCACGGCTAGCTTTTTCAGGGACAGCAACGTGAGCATGTACCAG
AAGATGTGGAGCACGATGGAGTCGGCGCAGCGTCCGTGTTGCTGAAGAGCAACGACGAGGCGGTGGAGCGCGTGCAG
AAGGGCAAGCGCACCTACGCTTCAATGAGTCCACCTCCATCGAGTACCAGATGGAGCGCAAGTGCACCTGGCG
CAGGTGCGGCACCTCCTCGACTCCAAGGGATACGGCATCGCTATGCCTTTCAACTCTCCATTCCGAACAGCCATAAGT
GGAGCTGTGTTGAAGATGCAGGAATCTGGAAAATTGCAAGAGCTGAAGGATAAATGGTGGAAAAGTTATGCACGGCGGT
GGAGCATGCGATGGGCAAGATGAAGCACCGCCCTCGAATGAGCTAGGACTGTCCAACGTGGGTGGTGTGTTCTGGTA
CTGATATGTGGCTGTGCAAGTGGCCCTTTTCTGTTGCTATGCTCGAATTCTTATGGAATGTCAGAAAAGTTGCCGTTCAA
GAAAAGATATCACCGAAGGAAGCCTTCATGTTGGAATTGAAGTTTGCCATCAGATGTCATGGAACATCAAAACCAGTT
AGAAGACCAAGAGAGCCTAGCACCAAGGGATCTTTCATGCAGTTGGATGCCTTTGAAAAGAAAGGCTCAATGTGA

>TRINITY_DN95621_c0_g1_i1.p1

CGTCTTCTCGCGCAAATTTGAGAAGATTCTCCTCCGCAAGATAGTTTTCCATGCGTCCAAGCGAGTTTGTACCTACTGCGC
AGCGGAGTGGCGGCCATATTTGGGCCGAGTCGGGGCAGACAGCCAGCCAGTGCAGTCGATCTGCGACACGATGGAG
ATTCCACACCTAGAGACCAGGTGGGACTACCGGCTGCGCAGGGAGAGTTGTCTGGTCAACCTCTACCCTCACCCACC
ACACTGTCCAAGGCTTATGTGATCTGGTCCAAGCATGGGGCTGGAAATCTTTCACCATCATCTATGAGAATAACGAG
GGACTGGTGAGACTTCAAGAGCTTCTCAAGGCTCATGGACCGTCTGAGTACCCTATCACAGTGAGGCAGCTGGGAGAA

GGCTCTGATTACAGACCACTTCTGAAGCAGATCAAGAACTCGGCCGAGTCCCACATTGTACTGGACTGCTCAACTGAC
AGGATCTACAATGTACTGAAGCAAGCCCAGCAGATAGGAATGATGAGCGACTATCACAGCTATTTGATCACATCTTTG
GATCTTCACAGCGTAGATCTGGAAGAGTTCAAGTACGGTGGTACCAACATAACAGCGTTTCAGACTGGTAGACCCCGAG
AAGCCAGAGTTGCAGCGCGTCGTCAGAGACTGGATAAATGGAGAACAGAACTCGGCAGGAACTCGATTTAATAACA
CACTCGCAATATGGCCGAGACAACATAACTTTTATCAAG

>TRINITY_DN11316_c0_g1_i8.p1

ATGCAGAACCAATCCTATGTTAGATTTACTTTTGAGAAAAAGGAAGACCCTCTCTTCGCCGAGATGTATAAAAGAAAG
ATTCTGAAAGGGCCCGGACCAGTGCTGAGTGCAGTAGACGGTATAAAAAATGGTGCAGAACACCCACTTTTGCTTTCTAT
GGCGAGGAGGCAATCATCTATCGGTTAATAGAGAACGGTTTTACCAATGAGGAGATCTGCTCGTTGGCAGAAATTGAT
ATGGACGTACCTCTACCATTGTTTTTCCAGTTAAGCGAAATTCGCCGTTGATAGAGTACTTGAACCTTTGCGTTGATG
AGAATGAAAGAGTCCGGAATCGTAGATGAGCAGAACCGGAAATGGAACCCCGCAAACCAGAATGCCAAGGCAATGCC
AACGTACCAGCAGTGCAGTTACAGTCCGTCCTTGTCGCCTACATCGTCTACTTCGCTGGATTAGTTATGTCACCTAGTT
CTGTTGATGTTAGAAAGACTAGCATATAAACGAAAACCGTCAAAGAACTTTACAAAAGCTAAACAGAGTTAGATTT
ATGCAGGCAGCTAAGACTGCTTGA

>TRINITY_DN11355_c0_g1_i14.p1

CTTTTAGAAGAAGCTTCTAAAGATGAGATTTTTCGATATTTTTGGGTTTTCTGATGCATTTCTATTCACTAATAGCGTT
CTTTATCAATTAAGCACCCCGGAATCGGTTAATAGACATTTTTTCAGCAATCGCATTAGAAGCTATAGGCCCATCTAGA
CCAAAAACCATCGGACAACCTACCCAACCTTTGAGAGGAAATTGCTTCAAGTGTCCACTTTTCGATTGTCCTGTTTCATGTG
TTCATGAAGACTATCAACAACAAAACAGTAATCTACGACGGCCTAGAGTACACAATGTTTCGTGGAGATCGCCAGACGC
CTGAACCTTCACCTACCAGTTTTCGGCTGTTGGTGGACGAAAGCAACAAGTGGGGCAAGATGTTGGACAACGGCTCGTGG
ATGGGCGGCATCATCGGATCCATTCTCGACTCGCAGCGGCAGCGCGATCAGCGCGTGCATGTCGGCTTCTGCGGCGTC
TGGAAGATACATCCGCAGTCGCTGTTTCGTCGATCTGACCACGTCTCTTATACAGGAGAAGAACCCTTTCTGGTGC GG
CAACAGACCGAGGTTTCGCTGGAGCTGGATGTCGCCATTACCCGAGTTCCAGCCCTCTGTGTGGCTCTGCATCGCCGCT
CTATTCTGCTACAAGCTGCCCTTCTCAACACGGCCGCCAAACGGTTACCCGCTTTCACTTGACAAGATATTTTCGC
GTGAGCTCGTGCCTGATGGAGTTGCTGGGCATTACGTCGCTGGCCTCCTCCCCCAGTTGGAAACGAAAATCAGCCTAT
CTGCTCATGATACAGTTCTATTTCGCTGGTCATCACCACCAGTTACACTAGCTACATACTCAATCGGCTTTTGACACCA
TACGTTAGTGAAGGAGTGACATCCATCGAGCAACTGATTGAAGGCGACTATTACTGGGGAAAGCCTTACCCGACTTCA
GGATACTTCATCAACTATGTGTGTAACTTAACTCAATGAGGCTATATTGTTGAAAAATCGGCAATATAAATTTATT
TATATTTCACTATTTACCTGAATAGATGTACAACCTACCTTACAGAAACCAGCGCACGTGCCGTTTCATAGAACGTGCT
CGGCTGTTGAGCGTGCGTCGCATCGACTACCTGCTGCGTCACGACGAGCGGTTTGCGCTGGTGGTGGGCACCATCATG
GACACGCTGATCAAGACCGACATCAGCGTGAACCTGTGCGTGCTGCAGCAGCGCCGCTACCGTGCCGTGCCCGACGAT
CTGCAGGCCCTACAGCATTGCGTTTGCGCTCAGAAAACACTCGCCCTACACGCCCCGTCTTTGATCGGCTCATTGCACGC
ATTGCTAGGTCAAGTCACCAATGGGACACTCCACTATCAAAAGTCATTAA

>TRINITY_DN13434_c0_g1_i2.p1

CAGTGGCAGCTGTCATCCTTCCTGACCAACCCAATATCACAAAGCCTAGTTAACCTGTTGCTGTTGTTTGACCCTACA
CTGACTGTGGAAGGCCAAGAAGCGAGAAATAATACAAAAGAAATAATAATTTACTCATCTGCTCTTCACTGATTTCG
ATGGGAGCCACCGTGCCGAAAGTGATGACTGTGTGGCGATGTAATCGTCTCACACAACCACAAGTCAACCTTTTCCCA
AATAAATTCCAGTCTGGATTCAACGGCTATAGATTTTCGTTTATCGGGAGCTCAACAGCCACCTTTTCGTTTTTCAGAAAG
AAAAACCAGCTTAGCAAACTAATCTGGATGAGAGAATCACAAATGGGATGGATATGAAGTAAGGCTTTTACAACCTG
GTTTCCGAAGCTCTAAATTTACCTTCAAGATAGAAGAGCCGCTTGAATCAGGGAATGCTTCTGATATTCGAGTAATC
AAAGAAGTAAGAAAAAATAGAGCTGATTTTGCAATAGCAGGAATTCAGCTATAGCTGAAAGGATGAATGAGTTTTCC
TTTTTCGACAGTGCATTCCTTCGACTGTGCTAATTTTCATCACATTAACGTCAACCGCCTTACCAAAAATATACAGCTATA
TTGGGACCATTCCGGCTAGAAGTGTGGATCTGCCTGATAGTGATCTACTTATTTGCCATAATTCCCTTATCGTTCTCA
TCAAGACACACGCTCAAGCATCTTTTGGAAGATCCGTGGGAGATTGAAAACATGTTTTGGTATGTTTTCGGCACGTTT
ACCAACTGCTTCACTTTTCGTCGAAAAGGTTCTTGGAACAAGTGGAAAGCATAATGCTACACGTTTGTATTATAGGAACC
TATTGGGTATTACAAATAATAATCACTGCAGCATACACTGGCTGTATTATAGCCTTCATAACTCTACCAGCTTACCCA
CCCATTCTGGACACAGCTAAACAGCTTGTGAAAGCCAATTATCGAATAGGAATCTTGAATGGAAGTGGGTGGCAAGCA
GCATTTGAGAATTCCAGTGATAAAGTGACAGAAAAGCTATTGAAAAAGCAGACTTTGTTCCAAATCTAGAACTTGGC
TTGAAAAATGTCTCCAAAGCCGTTTATTGGAACCGTATGCATTTCTTGGTTCGAAAGAGGAAATGAAGTTTATGATC
AATACCAACTATACCCCTAGTTCAAGGCACAAAAAATCACTGTTCCATTTGAGCGAAGAATGCTTTGTTCCCTCTGATG
GTATCCGCTGTCTTTTCAAAAGATGCCATCTACAAACATCCACTAGATGTGACCATTATGAGGGTGATCCAGAGCGGA
TTGTTTCATGAAGTTGTCAAGGGATGTGAGTGGGACCAGTACCGTTTGAGTACAGGAAAGTTGATGTATCATCAAGCA
ACTTAA

GAATCCAAGAAACTCATCCACCATTACGACCATTACATGAAGAATCTCCATGGACTGAAAATAGTGGTCAGATCATAC
CACGATGTATTCTCAGTTCATATGCGTTGGAAAGGCATCGATCAAATGTTTTTGGAAACTTTAAAAGTAAC TTTGAAC
TGTACAGTTGTGTTTCGAGTTCAGAGATGTGGCTTTCGACGCTTTCTATCTGGGAACAGGGAAGATTAAGAAGAAGATG
ATCGATGTACTGTTGAACAGATATTATATGACGGGCAACTATGGCGAGATTTCCATTGTTAGTCCGCTAAACTTCGAT
GAACTATGTTTTATCACAGGACAGGAGCCGAGAGTGTACCAGGGTGATAATCTGAGAAAAGTTTTCAAACCCACTGTG
TGGTTCATGATATTTATCACCATAGCTATTGTCTATTATGCTTACTAGATTGATGTCGTTTCATGATTGGAGGACCTTTA
AGAACAATTACATAAATTC AATCACTAGTTTCAACCGGCAACGTTTCATGTGCGTAATGCTGAATCAGCGTATTTTTCTGG
CTGAGCTGGATGCTATTCTCCTTCATCATAGCATGCTCCTTCTCCAGCTTCATTGTCCACCATGATGACCATACCGATA
TACGACCATAAACTCTCAACCTTAAACGACATCCTAAAATCGGGTAGGAAGATCTATCTCTACGAATCGTACTCAAAC
TCGGTGTGGACGAACACGTCTCCAATTTTGCAGAAAGTTGTTGAGAAATCAATCCCAAACAGTGATAATGTAGTCAA
CAATTCTGGAATTCTCTGAGATACAGAGACAGTGTATCGCTACTGGCAGACTTACCAGAAAATTCTATGAATTCACT
TTGCACAAGGAAGGTCTCATGTTTTATACAATACCCCAATGCATGGGAAGCTTCTTCACCGGATACGTTATGCGGAAA
AAATCTATACTTCTTACCCTATGAGACTGTGTATAAACAGATTCAAAGAGTCTGGCTTTTATGAGAAATGGTATCAT
ATGGCTGAACATATAGCCAGACTGAAGCCCACTGAACCTGAAGAGGAAAAGTATCAAAATCAACCCCTTAAAGTTGAT
GATTTCAAAGTGATTTTCTTCATTCTGATGGGTGGATTGCTTATCAGTTTCTTGGTTTTTCTTCTAGAAGTACAGAGT
TTCTTTGTGAGGAAGAGGATGAAATTGCGAAGATTGAAGAAAATTGCTGAAAAATAGAGTGGAAGAGCTTTCTTAG

ATGGCAAGCGGACAAGAGCGAAAGGGAACCTTTCAAAGTAGAATATCTACAGAAGATCGAGAAAGAAGTGCAAGAGAGA
TGGGAGAAGGAGAAAATCTACGAGATCGATGCTCCCACTTCGAAGAACAAATCAGGGGCAGATTGTGAGAAATTTTTTC
ACAACCTTTCCATTCCCTTACATGAATGGTAGACTTCATTTGGGTCATACATTTACCCTCTCAAATGTGAGTTTGCT
GTCCGATACCATCGGCTGAAGGGCAAAAAATGTCAATTTCTCTTTGGGTTCCACTGTACCGGCATGCCTATCAAAGCA
TGCGCCGATAAGCTC

GTGGTGGACTTTACAGTTCATACTACGATTTGGTGGGCATAACGATACTGATGAAAAAGCCACAACGCCAACATCA
TTATTCAAATTTCTTACGGTGCTTGAGAAAAGATGTATGGCTTTGCATACTAGCCGCTTACTTTTTTACCAGTTTCCTG
ATGTGGATCTTCGACCGATGGAGTCTTACAGCTACCAAAATAATAGAGACAAGTATAAAGATGACGAAGAGAAAAGA
GAATTC AATCTGAAAGAGTGCTTATGGTTCTGTATGACATCTCTAACGCCACAAGGTGGCGGAGAAGCTCCGAAAAAC
CTTTCGGGCCGGTTAGTGGCCGCCACTTGGTGGCTGTTTCGGCTTCATAATAATCGCCTCGTACACGGCCAACCTGGCC
GCCTTCCTGACCGTGTCAGACTCGACACTCCTGTTGAGAGTCTCGACGATCTGTCCAAACAGTACAAGATTTCAGTAC
GCGCCGCTCAACGGCTCATCGGCCATGACTTACTTTTCAGAGGATGGCCGATATTGAAGCGAAATTTTACGAAATTTGG
AAAGACATGAGCCTGAACGCACAGCCTGAGCGAGGTGGAGAGGGCGAAGCTGGCCGTGTGGGACTACCCGGTGAGTGAC
AAGTACACGAAAATGTGGCAGGCCATGAGAGAGGGCCGGCTTACCCAACACCCTCGATGAGGCGATCGAGCGCGTTTCGT
GCTTCAAAGTCATCCAGCGAAGGATTCGCCTATCTAGGTGATGCAACCGACATAAGGTATCAAGTATTGACTAACTGT
GATCTTCAAATGGTTGGTGATGAATTTTCTCGAAAGCCGTATGCTATTGCTGTTCAGCAAGGCTCTCATCTCAAGGAC
CAATTCAATAATGCAATATTACA ACTATTGAATAAGAGGAGATTGGA AAAACTGAAGGAAAATTGGTGGAACCAGAAT
CCAGAGAAAAAGTATGCGAAAAACAAGATGATCAGTCAGATGGAATAAGCATAACACAATATTGGAGGCGTATTTATA
GTGATATTTGTTGGCATTGGATTGGCGTGTATAACTTTGATTTTCGAGTACTGGTGGTACAAGTATCGGAAGCCTCAG
CGTGTGGCCGAGGCAAAAGGCCAAGGCGAGCGAACATTTCGCACGAGATTTCTTGAAAAATTGGCCACCAGCAAGAGC
GAGTTTTCCCATTTTCAGGGCACGACAAGTTTACCATGGGCAAACTACTACAAAACGAAATCCAGGAATAACGCCAGTA
GCAGAACCCAGTGTAGTTGGTGA

ATGACACTAACTGGGTTCAGCAACTCCACTAAGAACACTGGATTTCCTACTGAGACACTTGAAGGAGACAAAGATGAA
TCGGCCCTCAACCCGTCTGTTTGGAACTTATCAGTGCTGTTCAACAGAAAAAGATATTTCTTGGCGCTTGTGCT
TTTACTGTTACCGATGCTAGGAAGCTGGCAGTGAACCTTTACAACCTGCAATAAGTGTGGAGCCCTATGTTTTCTTTCTT
TCTCGGCCGAAAGAGTTGAGTCGTGCTCTACTATTTATGTCTCCTTTACCATCAGTACTTGGTGTTGTATTACGGCA
GCAGTTATCCTGATGACACCCGTGCTGTGTGCTATCAGCCGAGCCAGTCCGTTCTATCGAGTTTATGGCGCCGATCGG
TCCGGTGGTTTTTCATTGCGCAATGAACTGCTTCTGGTATATCTATGGCGCTCTACTGCAACAAGGCGGAGGAGCGTTA
CCGGCTGCCGATAGTAGCAGGCTGGTTCATAGGCACCTTGGTGGCTGGTGGTCTGGTGGTCTGGTGACCACTTACTCGGGC
ACCCTGGTCGCCCTTCTCACCTTTCCCAAAATGGATCACCTGGTGGCTGATGTCGACGAACTACTCGCCAGCAAGGAC
CTTTTCACTTGGGGCATTCCTGCCGCATCCACACTTGAGACATCACTCAAGTCAAGCGATGAAGATAAGCTGAAAGCA
TTGTACAAGGGAGCACGTCTGCACCCTGCCATATCCAAAGAGTTGATTGAGGATATTGCTCTGGCCACCACGTGTAC

ATTTCAGCGCAAACTAGTCTGCTCTTTCTCTTGCACAGACTCTACCTGAACACTAGTACTTGTGATTTTTTCGCTTGGC
ACAGAAGAATTCTTGTATGAACATCTTGCAATGGCTATTACAAAAGGCAGTCCTTATTTGAAAATAATCAACAAGCAA
ATTGATTGGATGCACCGAGTTGGCTTGATAGGGAAGTGGCTGAAACAGTATCTGCCAGAGAAAGACCAATGCTCTGTT
TCAGTGTCTTCTTCCGAAGTCAATAATCACAAGGTCAATCTTGATGATATGCAAGGATCTTTCTTTGTCTATTTCTA
GGTTTCGCTGTTTCAATTCTACTTATTGGGATTGAATTCTTGATGAAAAAATACATTACAGCTAAAGAACGCAGAATT
ATTCAACCATTTCGTAACGTAA

>TRINITY_DN20443_c0_g1_i1.p1

GTTGAGGACATTCTCACTGCTCGCGACACACTTTATAGATATCAAGCTGATCCCGAAATCCGTCATATGTATTTGACT
AAAATCAAACCAAAAAAAGAGGAAGCATTTTACACTGTCGAGGAAGGCGTTGCTCTGATGAAACGAGGTGGCTTCGCA
TTCTACGTGGACCCAGCTGCTGCTTACAACTCATCGAAGACACATTTTCCGAACAGGAGAAATGCGACCTAACTGAA
GTCTGGATGCATGAGCCTGATGTGCACACTCTGCCAATACAAAAGCGTTCTCCTTACAAAAAACTGTTCAATTTTGA
ATGCGACTGGCCCTGGAGTACGGCGTGGTGAAACGCGAGCGGCGCAAGTGGCACAGCCGGCGGCCGAATGCGTACTA
GCCAAGAACCATCGCAGCCACGTCAGCATCGGGCTGGCCGAGTTTCGTCCCGCTCTGCTGCTGCTGGCAGCCGGCATG
ATGCTCGCTCTGTTCTGCTCGTACACGAGATAATG

>TRINITY_DN2107_c0_g1_i1.p1

ATGTTCAAGAATTTAGCGTTTACATTTTTTATTACTTCAATTACAGTAATTACAGCTCTAATGGAGAAGAACATGAT
GCTGTGGAAATACCAATTGGCGGCTTGTTGCACTCTATGAATGATGAAGATGAACTGAGTGCCATAAAATGATAAGGCG
TTTCATTACGCAGTTGATCTTGTCAATAATGATCCAGACACGTTTCGCTGAATTCAAATTGAAAGCACTCTCCAGAGA
GTGAAGGAGCACAACAGTTTCAATGTAGGTTTTAAAGTGTGTGAGCTTCTATCAAAAGGCGTTGCTGGTATATTGGA
CCTCAATCACAGCTCAGCTCATTGCACGTTTCAGTCGATCTGTGACGCTATGGAAATTCCTCACATTGAGGTGAGATTC
GACGCTCTGCAGCGACGTGGATCTTGTCTTGTCAATCTCTATCCGCATCCTTCTGAACCTTCAAATGTTTATGCTGAC
TTAGTTAGAATGAGCCACTGGAAGAGTTTCACTGTCATTTATGAAGACAATGATGGCTTAATGAGATTGAATAAGCTG
CTCACATTCTATGATAAAAAAGGATATCCAGTGACTGTGAGGCAGCTATCATCCGATGATAATCACAGGCCAATTCTT
CGCCGAGTGAAAAATGCTGGAGAAAAATTTATCGTGCTGGACTGTTTCGATTGGAAAGCTGTACTCGGTATTGGAGCAA
GCGATGCAAGTTGGATTGATGGGTGACGACACCAACTACATTATAACTAATTTGGACATGCAGACCATCGATCTGAGC
CCATTTCATGCATGGTGGCACGAATATAACTGGTCTAAGGATGTTCAACCCTGAAGATGAGGATGTCAAGCAAGTTCTT
GACAATTGGAAAGAATGTCTTAAACACTCGAAAGATGGAGCGGATTGCACTCCAGTCATTGGCCCTGATTCAAGTT
CATCTCGAAACAATTCTAGTTATGGATGCTGTGTTCTTATTTCGCTGTTGCTCTTCAACATTTGAGTAATACTATGGAT
TTTGCAGTGACCCCTCTCAATTGCTCTGAAGATCGCAGCTGGGAGCATGGCTACAGTGTGATAAATTTTATGAAAGTG
AGTGAATTGAAAGGTCTATCAGGCTTGATAAAATTCGACCATCAAGGGTTTCGAACCAACTTCAAACCTAGACATAATT
GAATTGACGTCCACTGGACTGAGGGAGAAAGGCACCTTGAATACAAACAGAGGGCGTGAATCTGACTGCCGTTGCTTCA
GGTTTCATGCCCCGATCAAGCCGTGGAAGATTTAAGAAACATGACGTTTCATTGTCCTAAGTCTTTGACCCATCCTTAT
GGTATGCTGAAGGAAACAAGTCACAAGTTGACTGGAAATGATAGATTTGAAGGATTTGGGATCGACTTGATTGATGAG
TTATCACTGAAATGTGGCTTCAATTACACGTTCCGTTTGAACACAGATAAGAGCAGTGGAATCCAAATCCAAAAACA
GGGAAATGGAATGGCATGATAGGAGAAGTATTAGCTGGGAGAGCGGATCTTGCAATAGCGGACATAACTATAACAAGG
GAGAGAGAACAGGATGCTGATTTTACTATGCCTTTCATGAATCTTGGTATCAGCTTGCTGTATAAAAAGCCGATGAAA
GAACCACCAAGTTTGTCTCCTTCTTGTACCATTTCAGCTACGAAGTTTGGGGATATATGATTGCCGCTTATATAGGG
GTGTCAATATTGCTTTTCTTTCATGGCAAGAATTAGTCCAGCTGAATGGACCAATCCATATCCATGTATAGAAGAGCCT
GATAACTTAGAGAACCAGTTTCAGCTTGAACAATTCATTGTGGTTTACCATTGGCTCACTCATGCAGCAAGGATCGGAA
ATCGCACCAATCGCTATCTCGACGCGTATGACTTCTGCTATCTGGTGGTTCTTCACTCTGATCATGGTCTCCTCTTAC
ACCGCCAATCTGGCCGCTTTTCTCACAGTCGAGAACACTATTTTCGCCGTTTGAAGACGTTGAAGAGCTAGCTAATCAG
AAGACTATTAAGTATGGGGCTAAGGCTGGCGGCTCTACTGCTAATTTTTTTTCGGGATTCAAACCATTTCGACCTATCAG
AAAATGTGGCAATTCATGAAAGAGAATCCGAGTGTGATGACCAGTTTCGAACGATGAAGGTGTGAGACGGGTGATCAAC
CCCGATGAGGACTATGCTTTTCTCATGGAATCCACGCTATTGAATATGAAGTCGAGAGGAAATGTGAACTCAGTCAA
ATTGGAGGACTGCTAGATAATAAAGGATATGGAATAGTCATGCGCAAAAATTCACCTTATCGCAATGCATTGAGCACC
AGTGTATATAAGTTTGAAGAAACGGGAAAGTTGACGAAATTAACCAAAATGGTGGAAAGAAAAACGTGGAGGGGGA
GCTTGTGTCAGCAAAAAGATGATTTCAGCTAGTGCGGCTAGTGAACCTGGGACTTGAGAATGTAGGTGGTGTTCGTA
CTAGTTTTCGGTGTCTTTGTTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT
GGACCTTTGAGAGAAGAGCTGATGCAAGAGTTGAGGTTTCATCACACGATGTGCGAGGATCAACGAAGCCGGTTTCGTAAG
GCTGGTAGTGAAGAGGAGGAAGATGAAGAAGAGGAAACCAATAACGACGTCACCTCTGAAATGCCCTGGGCCCTTTC
TCTACACCCTACCGATCATTCGATACTAATTCAAAACAACCTCTTTGA

>TRINITY_DN25091_c0_g2_i1.p1

TCAGAAAGAATCCTACTGATAACAACCTTATTAGTGGAATGATGATTTGTTTTCATTTTCAGTTCTATGATTGTGACC

GTACTCATACGTCTGTGCCCATTTC AATAGTGACAGCGGAAGATCTAGCCGCCAGTAAGCTGAAGATAGCTGTTTAC
AACATATCTTATTTCCACAAC TATTTCAAAGACACGAAGAACGCTTACATTGCGGAGTTTTACCGCAAGAAAATGAGA
AATCAACATAGATACGTTACAATA CCAAGTGGGAATTGAACTGATGCGGACAGCCGGCTATGCGTTTTTCTACTGAAGAA
TCTTATCAATATCCTTTTGATTGGCAGCACATTTACAGATTCCGAAATTTGTACAATATCCGATGTGGAACCTGTCAAA
CCCTACATGGTGCCAAGCCCTATGAGAAAGGATACACCGTACAGAAAATACGTTAACTGTGGTTTTGCTCATTATGAAG
GAGGTCGGTATATTAGCTAGAGAAACAAATCAATGGAGAACCACACAGCCCAAGTGT CATGGTTCCATTGATTACAGT
AGTGTTCATTTTTGAGAGTTTTAACCGTGGCCTACGTTATTTATCTAACAGGGATGATTTTTGTGCGATTGTGATACTTTTT
GCTGAGAATCACACTTATAAGAAGCAAACGGAGTTATTGAAACAGTGTAAGCGACGTTTGAGATTCAAAAGGTTGCAA
GTGCTCTTCAAAAATGA

>TRINITY_DN3618_c0_g1_i6.p1

GATGAAATCAAGAACTTTGCTGATGAACTTTACTGGCTTGATTATTTTTCCACCGTTAGCCGTCAACGATCTGCGACGG
ATTGGGTTGCGATGTCGACTGGCGAAGAACCTTCATCACAACCGACTCGAACCCGTTCTTCGACAGTTTTGTGCGCTGG
CAGTTTCTCAAGCTGCAGAAGGCGAACAAAGTGATATTCGGCAAACGCTACACGGTGTTCTCGCCGAAGGATGGTCAG
CCCTGCGATGGACCACGACAGGTCCAGCGGGGAGGGGGTTCGGACCCCAGGAGTACACCCCTTATCAAGATGAAGGTGGTG
GCGCCCTATCCAAAAAAGCTGAGTATTTTCAAAGACAAGCCCCTGTTCTTGGTGGCTGCCACACTGCGACCTGAAACC
ATGTACGGCCAAACTAACTGCTGGCTGAGGCCTGATATGAGGTACATCGCTTTTCGAAACTAAGTTTTGGCGACGTGTAT
GTGTGCACTTATCGAGCGGCTAGGAATATGTCTATCAGGGCTTCACGGCTAAGGATGGCGAGTTTTAAACCAGTGGAG
TTGATTGGCCAGGACATAATGGGCATAGCACTGAACGCACCACTGGCCGCTCACAAAGTGATCTACACTCTGCCAATG
CTGTGCGATCAAAGAAGACAAGGGAAC TGGGGTGGTGACCAGTGTGCCCTCAGACTCGCCCCGATGACTACGCCGCCCTC
GTCGATTTGAAAAAGAAGAAACCGCTTAGGGAGAAGTACAACATTGCTGATCATATGGTGCTGCCGTTTGAGCCTATT
CCGATAATCGATGTACCAGAATATGGGAATCTGTGTGCAGTTCATCTGTATGAAAAATTGAAAATTCAAAGTCAGAAC
GACAGAGAACTTCTACAACAGGCTAAAGAGATGGCTTATTTAAAAGGATTCTATGATGGGGTTCTCCTGGTTGGAGAC
TTCAAGGGAGAGAAAGTTCAAGAAGTTAAAAAGAAGCTCCAAAAGCTGCTTCTCGACAAGAACGAAGCTGTAGTGATC
TATGAGCCTGAAAAACAGATTATATCAAGATCTGGCGATGAATGCGTGGTAGCCCTCTGCGATCAGTGGTTCTCTCGAC
TACGGCAATGCAGCTTGGAAGACAAGGCACGAAAATGTCTGGGCAACATGAACCTCTATCATGACGAAGTGC GCAAG
AACTTCTTGGCTACTCTGGATTGGCTCCATGAGTACGCTTGTTTCGAGGACTTACGGTCTAGGTACTAGGCTGCCCTGG
GATAAGAAGTGGCTGATAGAATCACTGTCTGACTCTACTATCTACATGGCCTACTACACGATAGCCCATCTCATACAA
GGGAACCTCGTTCAAGGGTGAAAAGCCCCAACGCCTTGAATATCAAAGCCGACCAAATGACACCCGAAGTCTGGGACTAC
GTTTTTCATAAAGGGTCAACCTTCCCCAAAAACACCACTATCCC GAAAAAGGCTCTGGACAAAATGAAGAACGAATTT
GAATTCTGGTACCCGTTGGATATGCGTGCCTCAGGCAAAGATCTCATTCCCAATCACCTGACTTTCTGCATATTCAGC
CATTGCGCCATCTGGCCTGAGCAGGAGGACAAGTGGCCGAGAAGTATACGAGCCAACGGACATTTGCTGCTCAATTCT
GCTAAGATGTGCGAAATCCGACGGAACTTTCTGACCTTGGATGAAGCTATCAAGAAGTTCAGTGCAGATGGTATGAGG
CTTTGTCTGGCTGATTCTGGAGATACTGTTGAGGATGCCAATTTCTCGAGAGTATGGCTGATGCTGGTGTCTGCGT
TTGTACACTTTTCATTGAATGGGTGAAAGAAACAATAGCGAATAAGTCTGCATTGAGATCTGGTCCAAAAGATTTCATTC
AATGATAAAGTATTCTCTCAATGAAATGAACCTGAAATCGCGCGAAGCCGACGACTTCTACAACCCGATGCTGTTTCGC
GACGCGCTGCGCGTGGGTTTTCTTCGAGATGCAGGCGGCGGGACAGATACCGCGAGCTGAGCCAGGCCGAGGGAGGC
ATGCATGCTGAGCTGGTGATGCGGTTTCATCGAGCTGCAGGCCATCATGCTGGCACCCGCTGTGTCCGATGTGGCCGAG
CATGTGTACGGCCTGTTGGGCAAGACGGAAGCGTGGTAGACGCCCGCTGGCCTGCGGAATCTCCAGTGGACGAAGTG
TTAATCAAATCGTCGGCTTATCTGATGGAGCGGCTCACTCGTTTCGTCTGCAGCAGAAGAACCTGCTGGCGCCCAAA
GGCAAGAAGGGCCAACCGCCACAGACCGCTGAGAAGCCGTCCTTGCCACCGTATGGGTGTCGAAAACGTTTCCGCGAG
TGGCAGACCATAGTGTGACCACCATGAAGGAGTTGTATGATAAAAAACAACGGCCAGTTCCTCCGACAACAAAGTGATA
TCTTCGACACTGAGCGTGATTCCCGAATTGAAGAAGTACATGAAGAAGGTGATGCCTTTTCGTGCAGGACGTCAAGTTG
AAAGTGAACGCCATCGGCCTCAAAGCGCTTAGTCTCACCTAGAATTTGATGAGATGGAAGTATTGAAATTGAACAAG
GCCTATCTGGAGAACACTCTTGACCTGGACGAAATCAAGTTCGAGTACGCGTCGAGCAAGGACGCTCCCGCCACATC
AGCGAGAACTGCACCCCTGGTACTCCATTTCATCACGTTCAACCACTCCGCCGCCAGGTGTGCGCTGACGGTTATAAAC
GCACAGGCGCATAGCGGCCTGTTTACACACACCGGTGCCAGTCAAGGCCGGGCGACACTGTGCACAGTGTGGTGAAGCGA
GTGGCGGCGGCTGAGCGCCACATTAAGGACGTGACCAAAGTGGAACATATGGCGCTGGGATGATCCGATTCTGGGACCG
AGAAAGTTTCCGCTTTTCAACGACCCCAACCGCGGGCCGTATCCGCTGGAACCCAACACTGAACTCATCGTTAATGTT
GATGCAAAATCCATCGCCGTGCAGAACAACGGCAAACAAATCAGCCTGGGAGAATCGATAACTTATTTAATTACTGGT
TTATAG

>TRINITY_DN40231_c0_g1_i4.p1

GGAAAGGCTCATGTTGGTCTTTCTGCTCTCTATCTTTGGTACAATGAATACCAATATTTGGATTACTCAGTATCGTAC
ATACGCACTGGAATCACCTGTTTGGTCCCGAAAGCCAGTTGATATCTCCTTGGCTTTTGCCGATTTCAGCCTTTTGAC

AAAGCATTGTGGGCTATCGTAGTGATGACTCTTGTACATCAACTATCACGCTGTATTATGTGTCAAAAATTACAAAT
ATATTTTCAAGAAATTGGAGACGTAGATTTTCAACTGTGGATGACAGCGTGCTCAGGGTTATTGGGATGCTTGTTCTC
CAAGTGCCAAACAGCAAAAGGCATAAGACCGCCGCTACTGCATTACGCTATGTTATTACTTTCTAGAAAGTTATGTTT
CTGCTTGTGCGTAACAGTTATCACAGTGGATTGGCTTCTGTTCTGACTTTACACAGACATGAACCAGCAATAGAAACA
GTGCAGGATTTGGTTGATAGCGGCCTGCAATGGGGTGCCACTCACGATGCTTGGATATTTTCAATTCTGCTCTCCCAA
GACCCAAGAACAAAACTCTAGTTGATAGATTCAAGGTGCTGAATGAGAAAAGTCTCAATGAGCGAACACAACTAGA
AATTTTCGCTTTTCAGTTTGGAGCGCTTACAAGGAGGCTATTTTGAATTGGAGACTATATCACGGAGGAATCAATGAAT
CACCTGCAATTGATGAAAGAGGATATTTATTATGAATATTGTGTTTATGCACTTCCAAAAGGTTCCACCATACAAACAA
CCGCTTGATACCTTATTTTGAACATCCAGAGTGCAGGCCTCGTTCTGCATTGGGAAGGCGAGGTTGTCCGCAATTAT
TTGCCAGTAAAGTGCAGTATGTAGTTGCTCAGAGCAGTGCCTGACCTACTGAAAGTGAGCCAATCGTTTTAGAGCTG
AAGCACATCCATGGAGCTCTCACAATTCTTGCCGTTGGATTGGCCTCATCAACTCTGACATTCTGGAATTATTT
TGAAC

>TRINITY_DN43282_c0_g1_i1.p1

ATGTCGGATTACATTAATAAAACAATGAAGCAGGTGTTGGAGAAGGACAGATGCACCTACGTGTTGACCTCTGACAGC
TATCTCACTCGCTCCACTGCATTTCGCCTATCCCAAGCACAGTATCCTGCCAACGCTGTTTAAATAGAGTAATGCTGTCC
TACGTAGAATCCGGCATCATAAACCATCTAGCCTCCTACAACACTCTGTGAGACACTCAAATCTGTCCCCTGAATCTG
CACAACAAGGAACGTCAGCTACGAAACACAGATCTGCTGACTACTTACTATGTGGTGATGTCTGGCTTCATAACCGCC
GCTGTGGTTCTTATCTGCGAAATTATCTGGAGAAAATGTGCGGAAATGGGTGCGGGAAAGCGGTGAGCTCGTGAAAA
GTGGAGGAAATTTTCGACATTCAAGGCAAACCGCCGCGGTTTGCATTATGGCGGCGAGCGTTTAAAAGTATGCTGAAA
GATGCGCCCGAGGAAAGAGGTTAATGGTAGGGAGTACTATGTGGTGCAGACTAAGCAGGGAGAGAAACGATTGATACCG
GTCAGG

>TRINITY_DN45996_c0_g1_i6.p1

ATGACGGCATTTCAGAATGGTTGATGCCGAGGATGTGAATGTGAGACGAGTTCTTGAAGACATGCAACGCTTCCAGCCA
ACCGGGCATAGTCTCTAAACAAGACTTCTGTGCATACAGGCTGGTCCTGCTCTGATGTATGATTAGTCTTAGTATTC
GCTGTGCGGTCTTCAAACACTGGAGCAGAGCCATTTCGCTGGTCCCCGCCAATGTCTCGTGTGATGCCGAAATGCCTTGG
GATGGCGGCTTGTCTCTCATCAACTACATCAACTCAGTGGAATTGAAGGGGCTGAGCGGGCCAATCGAATTCAAGGAA
GGCAGGCGAATCCAATTCAAGCTCGATCTGCTGAAGTTGAAGCAGCACTCGCTGGTCAAAGTTGGCGAATGGAGTCCC
GCGACCGGAGTCAACATTTCTGACAGAGCCGCTTTCTTCGATTCTGGATCGATGAACGTCACCCTCGTTGTGCATCACA
ATATTGGAAACGCCGTACGTGATGATCAAAGAGGGCGGAGAACTCGACATCGCAGAAATACGAAGGCTTCTGCATAGAC
CTGCTGAGAGAGATAGCTAATATGGTGGGATTTCGAGTACCAAATCCAGTTGGTACCCGATGGGAAGTATGGCGTTTTT
GACCATGAGACTGGAGAATGGAACGGGATTGTTAGGCAGCTCATGGACAAGAAAGCCGATCTGGCAGTTGGATCAATG
ACAATAAAGTACGCAAGAGAGAGTGTATAGACTTTACAAAGCCATTTATGAACCTGGGAATCAGTATTCTTTTCAAG
GTACCGACAGACAAAGAATCTCGCTTCTTCACATTCTTGGATCCGCTGGCTTTTGGATCTGGATGCTAGCGTTTGTG
GCTTTTCGTGTGAGTTTGCCTTTGCCTGTATGCCATCTCTCGTTTTCTCGCCCTACGAATGGTACAACCCGAAGCCGTGG
AAAGGCAAGGACGATCCAAGACTGATGGTGAACACTTTTTCACTCAGTAACAGTGTTTTGGTTTCATCACCGGTACTCTA
CTTCGTGAGGGCGCTGGACTCAACCCCAAGGCAACTTCGACAAGAATAGTGGGCGGAATTTGGTGGTTCTTTTACACTC
ATCATAATCTCGTCCTACACGGCAAATCTGGCCGCTTCTTGACGGTGGAGAGGATGATCACTCCCATTGAGAATGCT
GAGGATCTGGCTGGTCAGACTGACATCTCCTACGGCACTTTGGATAGTGGTTCCACCATGACGTTTTTTCAGGGACTCA
ATGATAGAAACGTATAAAAAGATGTGGCGGTTTCATGGAGAATAAAAAGCCTTCTGTATTTCGTGCCGACTTATGAGGAG
GGAATCAGGAGGGTTCTTGAAGGGAACACGCATTCTTATGGAATCCACCATGCTTGACTACATAGTTCAAAGAGAT
TGTAATCTTACACAGATCGGTGGTCTTTTGGACTCAAAGGTTATGGAATAGCAACGCCTATGGGCTCACCTTGGAGA
GATAAGATTTTCATTGGCAATTTTAGAGCTGCAAGAGAAAGGAGAAATTCAAATGCTGTATGACAAGTGGTGGAGAAT
CCAGGAGACACGTGTACAAGAAATGAAAAAGGAAAAGAATCTAAAGCCAATGCGCTGGGTGTTGATAACATAGGTGGT
GTGTTTGTGTTCTCCTCTGCGGTTTGGCTATTGCAGTGATTGTGGCAATTTTTCGAATTCTGCTATAATTCTAAGAAA
ATAACTAACTTGAAAGGAGAGTAGCCTCGGCACCCCATCAGAGCCTCTGTGCCGAAATGTTGGAAGAGTTGTGCTTT
GCTTTGAGGTGTGCTGGCTCGAGACAAAGGCCGCGCTGAAGAGACAATGCTCAAAGTGCTTACCTGGCACAACCTAC
GTTCCAGCCACATTCGATCTCCCTCCTCCTGTTAATAATGGATTGAGATCGCAGAGAAGCAGCATGGCAGTTTCCCTA
GAAGTGCCACCCGACCTACACGGACATACATCCTCCTGTGATCTTGACCATAGGCCTTGA

>TRINITY_DN7133_c0_g1_i5.p1

TTCAAGTACCATGATGGGCTTCTCAAAGGTGAAGGTGAGACACCGTTTACCTTGACACAACTTGACGAAGGCGATGAC
TATAGACCATTGTTGAAGCAGATCAAGGTGTCAACTGAGACTCACATAGTTTTGGACTGCGCAGCGGGAAAGATTCTG
AGCATTTTTCGCGCAGGCCAGGGAAGTGGGAATGCTGGGTGACTATCACAGCTATATTGTGACGTCATTAGATGCGCAC
ACAATTGATTTTCGAAGAAGTAAAGCTGGGTGCGCACGAATATCACAGCTCTGCGTATGTTGAACCCGGGTCTAGGCGTT

GTCCAGCAAGTAATAGATGAATGGATATTTGAAGAGCAAAGGAAAAATCGCACACCCGATTTGGTAGCTGATACTTTG
AAGTCGCATTTCGGCATTGATGTACGACGCAGTACAACCTTTTTGCAAGAGGCTGGGCTCAACTGGAATTGAAGGAGCCT
CCGCACCGAGTCGAGTGCAACGCTAAAGACAAGCCAAGCCATAAGAACCCAAACGGATTTGACTTATCAGCTTCATA
AAAACGATCCGAATGGAAGGTATGACTGGCATGATGGGCTTCGATGAGCTGGGAAGGAGAGATTTCTTCTCCTTGGA
ATAATTGAACTAGTGAAGACAGGCTTCAAGAAAATAGGATCGTGGGACCCGAAACATGGCATCAACTACACGCGTACA
CAAAGCGAAATGTTGAATGAAGTTATGATGAGCATTTCAAATAAGACGTTTCATTGTTACTTCGAGGCTGGGTCGACCA
TACCTGGATAAGAAGGAAGGGGAAGGCCTCAAGGGTAACGATCGCTTCGAAGGATATTCCCTGGATCTGATTGACGAA
ATCGCCAAGGAACTGGGCTTCACCTACGAGTTCAAACCTTGTGCCGGACAACAGTACGGTTCACTAAATAAAAAAGACC
GGCCAATGGAATGGAATTATCAGGGAACCTCAAGAAAGGAGAGCAGATCTTGCAATCTGTGATCTCACCATCACATTT
GACCGGAGGAGTGCAGTTGACTTCACTATGCCATTTATGACTCTGGGTATCAGTATCCTTTACAGTAAGCCAGTCAAA
CAGCCGCTGATCTCTTTTCATTTCTGCTTCCATTCTCCCTTGACGTCTGGATCTACATGGCCACTGCGTATTTGGGT
ATTTCCCTACTTCTCTATTTCTTGAAGATGCTCTCCGTATGAGTGGGAGAACCCCCACCCATGCAACCCCGATCCG
GACGAGCTTGAAGTACGTTCAATCTGCCAACTGTCTCTGGTTCGCAATCGGCTCGCTATGGCAGAGGATGCGAT
CTTTTGCCAAAGGCAGTGTGCGACTAGATTGTGCGCCGGCATGTGGTGGTTCTTCACTCTGATTATGATCTCGTCGTAC
ACGGCCAACCTGGCTGCCTTCTCACCACGAGCGGATGGAGGATACCATCGAGAGTGTGAGGATCTGGCCAAGCAG
AACAAAATCAAATACGGCGTCTCAACGGCGGATCCACGGCTAGCTTTTTTCAGGGACAGCAACGTGAGCATGTACCAG
AAGATGTGGAGCACGATGGAGTGGGCGCAGCCGTGGGTGTTCTGTAAGAGCAACGACGAGGGCGTGGAGCGCGTGCAG
AAGGGCAAGCGCACCTACGCCTACTTCATGGAGTCCACCTCCATCGAGTACCAGATGGAGCGCAAGTGCACCTGGCG
CAGGTGGGCACCTCCTCGACTCCAAGGGATACGGCATCGCTATGCCTTTCAACTCTCCATTCCGAACAGCCATAAGT
GGAGCTGTGTTGAAGATGCAGGAATCTGGAAAATTGCAAGAGCTGAAGGATAAATGGTGGAAAGTTATGCACGGCGGT
GGAGCATGCGATGGGCAAGATGAAGCACCGCCCTCGAATGAGCTAGGACTGTCCAACGTGGGTGGTGTGTTCTGGTA
CTGATATGTGGCTGTGCGAGTGGCCCTTTTCGTTGCTATGCTCGAATTCTTATGGAATGTCAGAAAAGTTGCCGTTCAA
GAAAAGATATCACCGAAGGAAGCCTTCATGTTGGAATTGAAGTTTGCCATCAGATGTCATGGAACATCAAAACCAGTT
AGAAGACCAAGAGAGCCTAGCACCAAGGGATCTTTCATGCAGTTGGATGCCTTTGAAAAGAAAGGCTCAATGTGA

>TRINITY_DN7270_c0_g1_i1.p1

ATGGCAGTTGTTCTTGCAAGGAATTGTTTGAATTTACGGGGTATTTGTACTAGCTGCGTTGATTTGAATGGCCAAAA
GAAAAGCCAGCCGATTCCCTGATACATAACAACCTCGTACAATGTACAAAAGACTGGTACAATTGGTGGGAGAAGAAT
GATTACTTCAAACCAAAACCTGGCAGTCAGTCCGACGATAAATTTTGTATAGTTCTGCCACCTCCTAATGTGACGGGA
ACTCTACACTTGGGGCATGCGCTGACCGTTGTTTTACAGGATGTTTTATCCCGATGGCATCGGATGAAAGGCTCCGAA
GTACTATGGGTGCCGGGAACAGATCATGCTGGCATTGCTACTCAGGTGGCTGTTGAGAAACATCTGATGAAGAAGAAG
AACATGAGGAGGCTGGACGTGGGTGCGGAGGAGTTCTGAAGCTGGTGTGGAGTTGGAAGAATGACAAGGACGGCACC
ATCAAGAAGCAGTTGCGTATGCTGGGCGCGTCGGTCGATTGGTCGAGGGATACTTTACCATGGATAAGAAGTACAGT
AAAGCTGTGAACGAAGCCGTTATCAGACTGTGGAATGCTGGCCTATTCTACAGGAAAGAAGCACTCGTGAATTGGTCG
TGTTATTTGCAATCAACCATTTCCGACATTGAAGTGGAGCATAGGCTGATAAATGGTCCACAAACTCTCAATGTACCT
GGGTATTCGTACCTGTACGTTTGAATTATTGATGATTTTCGCATACAAAGTTGTTAATTCAGACAAAGAGATAGTA
GTTTCAACCACTCGCATAGAAACAATGCTGGGAGATGTGGCAGTTGCCGTGAATCCAAACGATGAAAGGTACAAGGAG
TTGCATGGAGCCAGGCTGAGACATCCTCTGCGCAATGCTGACATCCCTGTCTGCGATGAGTCGGTTGATGTCAAG
TTTGGCACAGGCGCAGTGAAGGTCACCCCTAGTCACGATCACAATGATCACAATCTAGCTTTAAGGCATAGTCTTCCA
TTTATCTCTGTCATTAATGACCAAGGGCTGATACACAATTTCCCCAACTCGTCTTATTCAGGTATGAAGAGGTTTCGAG
GCAAGAGAAGCAATAAGAGAGGAGTTGAAAAAGCAAGGACTGTACAGGAATCAGAGGCCCCACAGTACCAGTGCCT
ATGTGCAGTCGGTCTGGTGATGTCATCGAACATCTCCTCAAACCTCAATGGTTTATAAATTGCAAAGACATGGCAGCG
AAGGCTCTCGAGGCAGTCAAATCCGGTCAGCTGGTGATACAGCCTAGTTCTTTTCGAGAGGGTTTGGATGCATTGGTTC
GAGAATATAGAAGACTGGTGCCTGTCAAGACAGCTGTGGTGGGGCCATCGTTTACCCATCTACGAAGCATCCAAAAAA
GACAGTGGAGAGAGCGTATGGGTGGCAGCTCACTCGATCGCCGATGCCACTCAGCAAGCGGCCGCTTTTGAAGTG
TCTGAAGCTGAGGTGGCACTGCGCCAGGATCCCGATGTGCTGGACACGTGGTTCTCATCGGGCATATTTCCCATGGCT
GTGTTCCGGTGGCTCAAAAGACTGATGATTTAGCCAAGTTTTACCCGTTGAGCTTGATGATAACCGGCCATGACATT
TTGTTCTTCTGGGTGGCTCGGATGACAATGCTAGGGCAACAGCTAACTGGACAACCTGCCTTTCAAAAAATATTGCTC
CATGGTATTATATGTGACAATGAGGGCAGAAAAATGTCCAAAAGTCTGGGAAATGTGATCTACCCTGAAAATATTATA
AAAGGCGTTACATTAGAGGACCTGCACAAAGAAGTGGAGAAAAATTACAAGTGTGGCCTAATATCGGGGAAGAAATG
AAGAAAGCGCTCAAGGGACAAAAAGCGACATTTCCGCCAATGTGCCCGACTGTGGCACCGATGCTCTGAGGTTACG
CTTATGTGCGAGAATATCCACAGGCACGCAATCGAAATCGACATCCAAGATTTTCGTCAACAACCGCTCGTTCTGCAAC
AAAATCTGGCAAGCGTTCCGCTATACCTTGAATAAGATCTCCGACCTCAATATAGATATAGAGGATGCGTCAGTGTTG
CCCTACGATTTGCACTCACTGTGCTTATCGATCAATGGATTCTGAGTCGGCTGAGCGAGGCTGTTGACAAAATGAAC
CAAGGCATTGAAAATGCTGACTTGCACATTGCTACCACGGCTTTTCGACGATTCTTCCACTCGTTTCTGTGTGATGTG

TATCTGGAAGCCACCAAGCCAGTATTCCGTAAAAGCGACGACAAAGAAATCAGTAAGGCGATTTCAACGCTGATGTTT
TGTCTAGACAGCTCGCTGAAACTGCTGGCGCCCCCTCATGCCATTCTGTGCGACGAAATCAACTATCAGCTGCCTGGA
AAACGGCCCCACAACGCTCATCAATGCTCAATATCCCTCCCAATTCCAGTTCAGAAACACCCATCTAGAGGATGAAATT
GAGTTCACAATGGATGTGATAAGAACTATTAGAAGAATGAAGAAAACATCATGATGTGAAGAAAATGGCTGCTCATGTT
GAAATTGAAAGCCAACAAGGAGACACCCTGACAAGACACTATCAATTGATCAACAGTTTCATAAATGATGATTACATC
ATAACTGATCTCGTAGAGCATAAGAATGAAGAAAATTTTCGTATGCGAGCCTGTCAGCTTTTCATACGCATGTCTTTTTG
AAATTACCGTTATATGAAATTCCCGATGATGAAATTGAGAGACTGTTAAACCAAGAAAAGAAAATTGAACAAAGCTTTA
GTCAAGTTGTTGGAAACTGTGCAAAATGAAAGTTTTAAAGCAAAAGCTCCACTTGAGGTTCAAGATAGTCATCGAAAG
AAGATTCTGAGTCTCCAAGAGAACTGCTGCAGATTCTGGAGGAACTTGACTCGCAACTATTGCACCAAAACGTAGAA
CCTTGA

>TRINITY_DN85391_c0_g1_i1.p1

ACTGCAGAAAGAGGCATTCAACTTGTCAAAGAAAATAAACATTTTCGCTCTAATTGGTGGCAGAGAAACATTTTATTAC
AACACACGAAGATTTGGAGCACATTTCTTTTCATCTAAGTGAGAACTCCATACTAGATATTCTGCAATAGCATTGCAA
GTTGGATGCCCATTTCATACAGAACTTCAATAGAGTTTTGATGCAGTTATTTCGAAGCTGGTATTCTAACAAAAATAACT
GAAGACGAGTATCAGAAATTAATAGAGGATAAAAAACATAAGTGACAACAAAAAAGCGGACACAGTAACCGAAGGTAAA
AAGGCAGCAAAAAGAAGAACAAGATGGAAGTCAGCCGAGAGCAATGAGTATGGACACTCTGAGTGGAGCCTTTCTC
GTTCTTTTCATCGGCTACTGTGCTTCAGTGCTGCTACTACTTTTTGGAAATAGCTTATTTCAAAAGGAAAATGAAGAAG
ATTCAAACAGTTTTTGAACGCGAAAGGAATGAAAGAGACAACCATAATCATAAATATTATTATCGTGATTGA

Receptores Olfativos (OR)

>TRINITY_DN27871_c0_g2_i2.p1

TGGACTGAGGATGTTGCAACAGAAAAGAACTCCAAAAAGATACGAGGCTCATTGTGCACTTTTCATCAGAACCTAAAT
CGGAACCAAAATGATTGTGCTATTATCTCAGGATACAGCATATTGATGATGAGCTTTCTTGTGGCTTTGGATACCTGC
TTCAATATATTTCCTCATGATGCAGGCAAACAATCCAAAGGTGTTTATGAACTACGCTGCAGCCTGCGTGTGTACTAAT
ATACCTCTACTACTGAATTATTACATAGGTCAACAAATTTTCAATCAGAATGAAATATTTTCGTCACTGTTTAAACGGAG
CTGCCCTGGATAGATAAGCCTCACTGGTTTAAAGAGCGTGCTAATTATGATGACCAGAGCCAATGTTGACACGGAG
TTTAAACCCCTATGGAATTTTCGTTCTGAATCTCAGATCATACAAAGATTTGATGAAAGCTGCCTTTTCGTTTGGAAAC
GTACTCTACAGGAGAAAGCAAATAATGAACTAA

>TRINITY_DN77195_c0_g1_i1.p1

AACCTCAAGAGCTGCTCAAAAAATTCAGCCTATGTACAGTTGGTTGTGAGCTTTGTGATAAGTGTGATACTTGTATA
GACATCTACATGATGCTCAAGAGAGATGATACTAGAACAAAGAGTTACCTTTGGAATCTATGCTCTATTTCATAAATTTG
TTCATATATTTGACATATAGTATTGGACAAAAGAATTAACAACAGAACGACATTTTTTCGGCGAGTACTTGTGACATA
CCATGGACTAACAAACCTCATTTGGTTTAAAAAGTCTCTGCATATTATGATGGGCCAGGCCAGTTATGACATGGAGATA
AAACCTTATGGCATTATGTTACTGAACCATGCATCATTGAAGAATATAATCAAGGCATCATTACCTTTGGAAACTTC
TTATTACAAATGAAAACCTGTGGATTGA

>TRINITY_DN10113_c0_g1_i1.p1

TTTAAAAATATGTTTGGATGGCTCAGGAAAAATGATCTCACTACTCAATTTTGTAACTGCTTTACTCCTCCACTGCG
GCATATCTCTTGTTCACGACCGATGATTTCAAGAAGAAAGTGAACCTTTGCTGTGATGTTTCATCATGGTGAGCATGGCA
ATATTTCTTGTTTACCATAAATGGCCAACGAATATTCAATCAGAACGTCATCTTGAGACGCACACTGGCAGAAATTCCA
TGGACAGACAAGCCTCAGTGGTTCAAACACACTCTATGCATGATGATGACCAGGGCTAACGTGGACATTGAGATGCGA
CCCTACGGTATTTACACGCTAAACTATCTGTGCATACAAAGACTTGATGAAGACAGCTTTTTTCCATTGGAAATATTCTC
TACACAAAGAAGCTGAGTAGCCAAATACAACATTGA

>TRINITY_DN4341_c0_g3_i3.p1

ATGATGAAGGCAACCAATGTGAAAATAATGGTGAATTTTGCAGTTGCTTGTGTATTTGTAACTTGATAATACTATTC
TGCTACCACAACGGACAAAAATAGTGAACCAGAATGATATTTTGCACAGTATCTGAGAGAAGTTCCATGGATCGAC
AAACCACATTGGTTTAAACCTCATTATGCATTATGATGAAACGTGCCAACAAATGACACAGAAATGAAACCTTATGGC
CTCTTCGTGCTAAATCATGAGTCGTTCAAAAATATCATAAAGGCAGCATTTTCTATGGAAATGTAATCTATTCTGTG
AAGTCGAATACTTAA

>TRINITY_DN11791_c0_g1_i1.p1

TGCAAACTAACCAACCTCAATTGATTGTAGCAGCAGACAATCTAACCTCAATCCAGCCCTCATTTTGAAGTATGG
CTGCCTTTTGATTACTCTTGGACTCCATTCTATCAAATTGTTTCATCTGTTGACAAGTTTCATATTATTTGACTGCTGT
TTGATTTATGTGGCACAAGATGTGTTTTGTTTCACCTTGCTGTATGCTGTTTGTGGTCAGTTTCAAATTTCTTGAAGAA

AGTTTGATGAATCTGGGTGATTTGGCTCAAGGGAACTGTTCAATGGATACAAGACCACAGAGAAGACATACTCTACT
GCTGCAGTTCAAACTCTGAACAAGGAATGCAAGAATTAATTTACCTCACAACAATGACAACCTGGATGAGAAATTA
TTGCATAAAATGTTTCAGAGAATGTGTCCAACAACATATTTTATTATTGAAATTGCTTCGAGACTATGAAAAATTAATG
GGGAGTATGATAATTATTGACTTTCTCCATGCTATTTGTTCTTTAAGCTTTGCCTTGCTTGAACCTCTCAGTTGAACAC
ACTTTTATTTCATGCGATAAAATTATTTTCTTTTCATTGCTATCTGTATTGTACATCAGTACATGAATAATTTTTTTGGA
GAAATGGTTATACAAAAGCAATTGAGCATATCACGAGCTGTCTACTCTTTGCCATGGGAAGAATATCCGAAAAAATA
AAAAGCTCAATTTTTGTTTATGCTTCTAAGAACCGAGAGACCAGTTGTCATTAATGGATTCAAGATGTACCTTCTTTGC
TATAAAACATTTGTTGAGTTCTTGAAAGCTATCACCTCCTACTATACAGTTTTTAAGGAACGTTTCATCTAGAAAAGTGA

>TRINITY_DN12315_c0_g2_i2.p1

AGGGGGTTTCCCCTCCTTTGGTCACCTACACCCCTTCGAATATCAATCAAATCTCGTTTTTCATCTACATCTTTGCT
CTCCACAGTTTTCTGATGCTTTTTCTGGCCTTGAGGCATTTAGTGGGTATGCTCTGCTTTGGTTGTGTGTAAGAAG
ATAATGACAGATTTTGAAACGATTTATATACTGTTGAATGAGATTGGGAACGATTTCCCTGGGTTTGCTGTAAAAGAT
TATTTAAAAATGAGCGAGCAAGAAATTCAGGTGGCTATGCAGAGAATAATAATGAAGATATGGAGAATCGACACGTT
GATCAATGGAACAATGGAGTTGATAGAGAAGACATCCTTCGAAAAGATTTGGCACTCATTGTGAAATGTCATCAGAAT
CTTAATCGGAATCAAGACGATTGCTCTAGAGTCTCAGGATACAGTATTTTGGTGATCAGCTTTCTTGTTGCTCTAGAT
ACTTGCTTCAATATCTTCTTAATGTTGCAGGCGAATAATCCCAAGATGTTTGTAACACTACGCTGCGGCTTGCATTTGC
ACCAATATTTCCTCTGCTACTTAATTATTACATTGGTCAACAGATTTCCAATCAGAATGAAATATTTGCGCCAGTGCTTG
GTTGAAGTGCCCTGGATAGACAAGCCCCACTGGTTCAAACGCAGCTTGCTAATTATGATGGAACGAGCCAATGTAGAC
ACTGAGTTCAAACCTTACGGCGTTTTCTGTTCTCAACCTCACATCCTACAAAGATTTGATGAAGGCTGCCTTCTCATTT
GGAAATGTGCTCTACAGAAGAAAGCAGGTAATGAATTAA

>TRINITY_DN12588_c0_g1_i3.p1

ATGGCTTTGAAAATTACAATCTACGATTTGAAATTTTTGTGTCTAACCCCTGAAGAGTTGGACGTGAAAAATAATTTG
AATAATTATGATGCTATGTTCAATTCCAATTCAAATTCAATTGATGCTATTTATAATGATGACGGAGAGTATGGTAAT
ATGCTGAGTTTCGAGAAAGGAGGAGTAATATTCGACTTGAATATGATTTGCCATTGCTGAAAGCTGGAATTGTTAAC
GTGATCAAATTTTCATGATATTATATGCAGTCGAGCAAGCGGCATGAACAAAGATTTTGAGATGATGTACATAATTTTG
AATAACACGATTTGTTTTCAAATTTGCGTGTGTCTCTACTCATCAGCAAAGATGGATGATTTGGTTCTGAAAATAGAG
AATTTTCTCATTATAATACCTCTGGCTGGTATGTTATTCCTCTATTGCTTCTATGCTCAACAGTTGCTGAATGAGGGC
GAAAAATTTGCAACAACCTCTTTGGGAAAGCAATTTTATCGACAAACCAAAATGGTTCAAATCATCCTTGTTGATAATC
ATGATTTCGTATCGCTAAAGAATTGGAAATCAAGCCTTTTGGAATTCTATGTTCTCAATCTGAGAACGTTTTCAATGGTG
ATGAAAGCAGCTTACTCCTACTATAGTATGCTCAACTCTTTGAAGAAACGTATATGA

>TRINITY_DN38863_c0_g1_i1.p1

CTGGAAGAGTTCAATGCATTCTACGGACGGAAAAATGCTGATTTGACCAGTTCTGGGGAGAGCCGCAGAGAAAAATGAT
TACCAGTCGAGCGGCTTCCAGGGAAGTTTCGATGACAAAGCATCGTACCATGCTAGTCTGAGAAATTGCATAAGAAAT
ATGGCCATTTCATCATCAAACCTATCCTTAGGAAAAATTCGAATGCTGAATAAAATGTTTCAGCTTCTCAATGTTTTACGCA
AATACTTTTCATCTGTCTGCAACTCTGCTTTTGCTATATTCTGCTTCCAGATAGGAGACTTTGTATTCAAAAATAAAATAC
GGTCTTTTATTAGTTTCAATGTTGTTGCTTCAATATGCCTATTCTGAGAATGGACAAATGGTCAAGGATGAGGTGGAC
AAAATGAGATTAAGCTTCAACGAATGTGAGTGGGAAGGCCACCCATCTTGTTCTATCCTTATTTTAAATGATGCTT
ATTTCAGAGCTCAATTCAACCGAAATTGCAGGATGTATTCAACATTTTCACACTGGATAGGAAGAATATGGCATCTGTG
GTTCAAGCAGCCTATTCTTACTTCAATCTTCTAAGTCGATTAATGAAGAATTGA

>TRINITY_DN12954_c0_g1_i2.p1

TTTATCTTTAAGGACGTGAGTGACCAGCAAAAAATCCAGATTCCAAGTACTGAGAGACCGCATCTGGAACCGTTCCCGG
TTCACGCCTTTCCCTCAAATCACCGTTATCGAAATACTCATCACAGCTATTGTACCTACTTTGTGACGCAAACTGCC
CTGCTCCTAATACTCCACTGGGACTCCTACGACATCAACTCCCGTCCCGAACTTTTCGCCAACTTAAGCTACAACCTC
TTCGTATTCTTTGTGCACAATGGAGGTGTTTTTCGTCAACCGGACGACTTCGCATTTTCATGCAGATCATCGAGTCCGAG
TTCAAAATCACCGGTCTGGACGCGAAACTCATCGAAAAGAAGCGAATAAGAGTAAGTGAATGGAGGAAACATCGCTT
AAAAACTCGAAAGCCCTCGGCAATATACTTTTATTCTTCTGTATAGGATACTCCTTCAACTGTGTTTCATGGTATTCTA
TTGTGCGCGTATAAACTAAGACCCTTCGACGAACCTTCCAATCACGTTTTCTTTCTTCTTCTACCTGCCGACTTCACAAA
ACAAAGGAAAGCGGATACTTATTCATGTTTATAGCTTTACTACAAATCTGGTACCTGTATTTTCGGCTATATCATCGTT
TCAGTCTCTATTCTCGCGTTTATTCGCTTACAACAGCATTTTCACCGAGATCGAATGTTTCTAATCACACTGGAA
GAGCTGGATCAATTCAAATTCGAAAATGAAGGAGGAATTAATTCCGCAACCGTGTGGTCACACCAGAGAGCTGGAG
CGCTTGAGAAGTATAATTGAAACGCTGGGTATACATCATCAGATACTTTTCAATAAAATATCAATGTTTCGAGGAAGGT
GGAAGATACCATCTTTTCTACGTGAACGCGAGTCTGTGTTCACTCCTTTGTCTCGCCATATTCTGCACATCGAGAGTA

ACATCACACCATTTGAAAGTTAGATACGGTATATTGGCACTAAGTATGCTCGCTGGCTGTATGGTCTATTTCAGAGAAC
GGACAACGACTGATGAACAGGGGTGAGAATGCAAGGAAAGCAATCTACGAATGCTCCTGGATTGATAAACCGATTTGG
GTGCAAAAAGCGTTGCTGATAATGATGATGAGAAACACACAGGATTTGGAAATAAATTTCTACGGCATTTCAAAAC
AATAGAAGCAATGTCAGCAGTTTACTGCAGGCTGCATATTCCTACTTCAGCTTCCTCAACAATACAAACTGA

>TRINITY_DN13549_c0_g1_i2.p1

ATTCTGGTCATGACAGTGCTGGTGCTATGCAGTCTGGAACACGCTTTTGTACCTGAGAAAAATTAATAGACTGATGGAT
GTAATCGAAGAAGAGTTTTTGGTCGGGAGGCACACGAGTTTGAAGAAACAGGGCCCCGAAAAATATAAAATACCAGAAA
TTCTTCAACGACAGCCATGAAGAAATGAAACAGCTTCACAATAGAATATTTCTAACGCTTGTGATAATTTTTTGGTGCA
ACTCTCTTGAAGCGGCCAATAAGAACTGGATGGAATACTGGATAACAAAGAGCCCGATGACTGGCCTACCCCTTAC
CTATTCTACTACCCCGAAGGATACAGTACAGCTCCAGTGTTTGTATTAGCATTGTCTTCTACACTCGATTCAATTATTT
TTCATGCTACTGGCTGGCTATTGCATGCTGGTGTCAACTGCAATTTCTAGTACGAAGATAACAGCAGACTTCAAAGCA
TTCTGTCTCTACCTGGAAGACTTCGGCAAAGAATTCTACTTTTTGGCTGGTAATGAACCTTGACTATGACTATGTTAAC
AAAACGGATTTCGAACGACAAGGAAACCTCTCAAACCTGCCAATTATTATGATTCAAACAATGAACAGATTCCAACCTGAA
AGAAATAATCAGAGAGAATATACGAAAACCGAACAGGACCTGAGAGAAGAATTGAGAGAAATAGTGAAGACTCATCAG
GATCTCTATAGGAATATGAGAACATTGGAGATCAACTCAGGCACAATGTCTCTTATCAAAGATGCTGCTATTAGCTTT
GAAACCTGTATTGGACTTTTCGTTATCATGAAGACAGACGACTCAGTAACAGCAATGAATATGGTTGTGGCTGTGGTT
GCAATTGACTTTCTCCTTCTATTCTCATGCAGTAAGGGACAGTTGATTATAAATCAGGGCAACGACGTGCGTTTTTCAG
CTAGCCAAGTGTCCGTGGATCGACAAGCCTCAATGGTTCAAAAAGCTTTGCAAATTATGATGGCCCGCTCTAATTTG
GAAACACAACCTGAATCCTTTTCGGACTGTACGTCGTCAACTACGAGTCTTTCACAAACATTTTGAAAGCTTCATACTCT
TTCTGCAATCTACTCAACGCCATGCCGTAG

>TRINITY_DN58542_c0_g1_i1.p1

ATGGAAGCCCACTCAGAACACTTTCTGAAACTATTTATCTTCATGAATTTTCGTCTACTCCTACTTGCCTTTCTTGAAC
GCAGTCTACATTCGTGTCAGCTCGGAAAGCTCAACGAATTTCAAAAACGTCGCACAAATATCCTACTTCTACTACCCC
ACGAGCGAGGTCACATTCCTCCATTATTGCTTTGGAACCATATTTCTTTACGTTGGTTTGATTTATGCTTCGGTGTTT
AGCTTCTGTATACTTTACTCCAGCTATTTGGCACTGCATTGTGTAGTGTCTGGCATTAGTCAACTGAGGACGTTTCATT
TTGGAATGGGATAAAATTTCTGACAAAAAGTGGAGCATGAAATGGAATGGAAGACTGCTTGCTGGAGATTGTTGTG
TTTCATCAGGAAATTTGCAGTGAAGCAAGACTTTTAAATAAAGGAATGGAGACGGCGACTATGAATCTGTTACAAGGC
TGTGTGGTACAAATATGCTTGAGTCTATTCAATATTCTTGAGGGAGGTGATAAAATCAAGTTTGGTGGATTACTATT
TTTATGCTCCTCATGCTGGTTGGATTCTCTCTTTTGGACAAATTCTAGAGGATAGGATAGCCCGGATAAAAAATGACA
CTATGGGAAAGCCGTTGGGTGGAAAAGCCGCTTTCCGTGAGAAAAACGCTTTTCATGATGATGATCGGTGCCTCGACT
CAAATGGAGCTCAGGCCTTTTGGGTGCGAGCTGGTTTAAACATGCAGAGTTTGCAAAGGTCTTGAGAACAGTGTAC
TCTCTTTTCAACGTTATAAATCATTCAAGCTCCGTTTCATGATACAAATCTATCATGA

>TRINITY_DN14580_c0_g2_i1.p1

CTCAACGAGTCATCGTTTCCTTGCTGTGCGACAACGTGGTCAGGGAGATGAGGTTGTTCCAAATGAATTTGCAGGAGTTG
AACGCTAACGCTGACGCCATAGAGCAACAGGCGAACAATGCTCCGAATGCCACCGACGATTACAACAGATTGAAAGTT
GTCGTTAGGCAGATTTCAATTCATCACCAGCTTATCTACAAGAAAGTTGGAATGCTGAATAAAGGGTTCGACTTTCAA
GTTATGTACAATAACGCATTCATCTGCTTCCATCTATGCCTGGCTATCTTCTGTATGGTGAAAGTGGATCTGCTTTAC
AAAATCAAGTATGGGGTTTCTATTTGCGCTATCACTATAATTTTATTTCATGTATTCCGAAAATGGACAGCGACTGGAA
AATGAGGGAGAAAACCTACGTTTGGCACTGTACAGTTGCAGTTGGATAGGAAAGCCTTCCTGGTTTTGCCGATCTCTA
TTGATTCTAGTCACTCAGAACAATAGGATTCCGAAATTAGAGACTTTCAAATCTTCACCCCTCAACAGAAACAATCTT
AAAGTGGTGATGCAAGCGCGTATTCATACTTCAGTGTGCTTATAAGATTTTCGAGGTGA

>TRINITY_DN64921_c0_g1_i1.p1

AATATCAACATGCTTATATGTGGTTTTGGCCTATCCTCTAGTGGACAGCGATACTTACAGCTTGATATCGCACGGAGAG
GACCAACGAAACCGTGGAACCAACATAATTCAGACGTCCGGACATCACTCGATAAAAAATCTGGATCAAGAGCTCGAT
TTGATTAATAAAGTTCTCTATTCTATCGTGCTATCTTCAGTGGTAGCTGTCACTTTTGGCCCTTTAATGGCAGCCTTG
ATTAAATATCATCCGCAAATATTTATGGGTGCTCTTCACAAAATCGAAGTCACTGGGCTTACTCTACCTATAAACTTT
AAATACTCGGAAAAATACACAAACCTTCCCACCTACTTGTTTATTTACGGTTTGCAAACCTTTTACATGTATTTGTTT
GCTGGTTACATGTATAGTACCCTTGCATCGGGATTGCTGGCTTTAAAAATGACAATATATGACATGAAAGTATTGTGT
CTCATGGTTAAAGAATGGGATATGATTGCTTTTAAAGGAACTGGTGAGAAATGAATGATTATGAAGTACCAAGCAAT
TGTACAACATTTTCGGATCGGCATACCAAATTTTGAAGGAGTCCCTTGTCAACGTGATCAAACTCATGAATTAATA
TGAGAAGAGCTGGTAATCTGAACAGGAATTTGAAAAATATGTATGCCATCCTCAACAACACAATTTGCTTTCAAATA
TGCGTGTGCTTGTACTCAGCTGCAAAGATGGAAGACTTTGTTCTGAAAATTGAAAATATTCTTCTTGTCTGCCTATT

GCTAGTATTTTGTTCCTTTATTGTTCTAGTGCTCAGCAGCTTTTGAATGAGTGTTCCAATTTCCGTACCGCTCTCTGG
GAAAGTACGTTTTTTTGACAAGCCCAAGTGGTACAAATCATCATTCCCTAATGATCATGATACGCAACACAAAAGATTTG
GAGATGAAGCCGTTTGGTTTCTACGAATTAACCTGAAAACCTTTGCTGTGGTTATGAATGCCGCTACTCTTACTTC
AACGCGCTTAATTCTTTGAATAAAAGAATATGA

>TRINITY_DN15317_c0_g1_i2.p1

GTCTATTTTCTCTACATCTATTCTACTGTGGCATTTTACAATCTATTTCAGCTTTGCAGAACACTCATTCTCATACGGTC
GTTGAAGTAATGGTACTAGTCTACTTTTACTCTTATGACTTTTGACAACTTCTACATTCTTTGTCATTATGGCAATAAA
ATCAAAGAACAGAGTGAAAACACAATTAGGACAGTATTTGCCGCACCCTGGTATGATTGCAGTCATGTTGATAAACT
CTTTTGAAAATAGTAATTTTGAGATCACAAAAGGCTATTGTTGTACCTCGTTCAACTCCACTGTTTTTCGATCTTACC
AATTCAACTTTTATTGTAATAATGAAATCTATTGTTTCATTCTACTTGGGACTTGTCAAGATGAGACAGCAAGCGGAA
ATGTTATCAGAAATTTGA

>TRINITY_DN18132_c0_g1_i3.p1

ATGGCCGGTCATAGCACCAAGTGCTCTCTACATCAGCTACAGTCTCAACTCCATTTCATGGCATCTTGACTTATTTGATC
GGCGCAATTCAGGAAATTGAAGAATTTCCCAAACCTGCATCATTGAAATACCATTCCAATATATAGCTCCACCTCAG
GATACCATCTTGCACTATGTATTGGTGACGGTCTTGCAATTGTTTATACCTATACTTATGCTACGATATAACGGCGACT
ATATACGTGTTTTGTGTCACTGGATATACGGCTGTAGCGGTTGAACTGAAGATGTTTTTGGTGAGATTAAAGGAACTG
GATGAGATGTACCTGGCGAATACTTTGAATTTGGATGATAAGAGTCGTAATGACTTGGATGATAATTTTCAAACCTGTC
ATCAAGAACCCTGTGAAGACTCTGGTTGTGGATCATCAGAGGATTTTCAGAAAAATGAACCAGTTGAGCAAAGGCAAC
TCGTACATGATTTTTCTATGTCAACGCATTTGTCTGCATTCAAATCTGTCTGGCGATAATCAGTTTTTCAAACCGGTGCA
ATTTTCGGACAACTCAAGTGCAGTTTCCGTGGAGCCTCAATTATCGCGTTCTGCTATGTCTATTCTGAGTATGGTCAA
GGCATAACAAGAAGAGTTTGAGAATGTGAGAAATGCGTTGTACGACTTTAGTGGCTCTGGAAAGCCAGTCTGGATCCAG
AAAACCTTCCTCATTCTCATGATTAGAAACACTCATATCCCTCAACTGTCGTTCTTCAACACCTTCACCAATGAGAGA
AGGAACCTTGCAATGTCACTCGTGCCGCATACTCCTATTTCAATCTTCTGAATCAGCGAGATAAGCTACTGAATCAT
AATTGA

>TRINITY_DN20422_c0_g1_i1.p1

TACCTAATTCTGAAGCGATCTATACTGAGTGTGGAGCATCGCGCTTTGGAAAAGTACAGCGGCTACACTCGAGCCAAG
CATCCCCAAATTGAAATAATCAGGAATGATCCGATTTATGCAACATTCGTTCAAGATTGCCTGAAAGAGAATGTACTT
CATCACATTGAAATTCTTAGGCTATTCAATCTCTACCAAGATGTTACGTCAACACTCTATGGTGTTATCATTGGAGTC
AGTATGATGGTCCTAGCATCTTTATCGATTGTGCTATCGAAGACAGAATTGTTATCGTTGGAAGCTGTCAAATTCACA
ATAATCTTCATATTGGAGCTCGTAATTGTTTTTGGATATTGTTTGTATGGGCACTTTTCTTACCCTTGGAAGTCAAGAA
ATACCAAGAGCATTATATAACTGTCCCTGGTTCAATTACAATGAAAGTCAAAAAGAACACTGAGAATCTTCCAGTTC
AATTCGTCAAAAAGTATTGAATTGAGGGGTAGCGGACTCTTTTTGATAGATCTTCAATTGTTTGTACAGATTGTGCGG
ACAACCTTATTCGATATTCAATATATTTTCATCCATGGAAAGCCAGCAATGA

>TRINITY_DN20522_c0_g1_i1.p1

TCGAAACCCACGACTACTGCTGCTCGAACAAGATTGAAAGTTTGGACTCGGAACTTCCGCAAGAAAAATAATTTTAAA
GATTTCTTCAATCGAATCATAAAGGACCATCAATCGGTATGCAGTGCTGCAAAAATGCTAAATGGGAAAGTAATTGGT
ATCACAATGGGGTTCATGAACATCTTCGGGTTTCAATTGTGCATTTTCTTGATGTTTATTGTGGAGGTGAAAGAGATG
ACATTGAAAGTGAGATATGTATTCCGCTATTGCTTAGTACTGGGAATACTGTTTTTCTGCACAAAAAATGTCAGAGA
ATAAAAGATAAGGGACAACAGCTCAGGCATGCAGTTTACGGGAGCCCGTGGACCGACAAGCCGAGATGGCTGCGCAAA
TCAATCAACATCATGATGACGATGGCCAATAAAGATATGCAAATCTCGCCCTACGGACTCTATGTCATTGACATGAGC
TACATGACCAATATTATCAAGGCTTCATTTACATACTTCAATGTATTGAGCGCTCTGAAAGCGTGA

>TRINITY_DN20817_c0_g1_i1.p1

GAGAATAATGCACGGCTCAAGAGGTACCTGCAGAAAATCATCATACACCATCAGGATATTATCAAATCGATGAAAATG
TTGAATAACGAGTTCCAACCTGACAATAATCATATTTCATAAATATCTATAGTTTACAGCTTTGCTTGACATCATTTTTT
ATTTTGAAGATGCCATCAATTGTTTCATCGAATGAAGTACTTTTTTGTCTACTGCGTGGTACTGATTATTCAACTGCAG
TGGGCGACTTTTGGACAGGACCTGATTGATAATGGAGACCGACTTCGACTGGCGCTGTACGAGAGTTTCGTGGATCAAC
AAGCCGCTGTGGATGAAGAAGTCGCTGCTGATCATGATGGCGCGCGCCGAGCAGCCCTCGAGTTCAAGCCCTACGGC
CTCTATGTGCTCGACAAACAGTACATGGCCAATCTCTTCAAAGCGACCTACACCTATGTCAATTTG

>TRINITY_DN24825_c0_g2_i1.p1

ATGCTAATTGTTGTATCTCCACTGCTTTTGGACTATATTTCGGAGAAGGAGATAGACCAAAGATCAAGGAAATCAACTAC
AAATTGCCAATTCGGTTTTTGGTTCCCTTTCAGCACTGATTCTTGGACAGGGTTCACTGTTGCCTTTCTGCTCTTGTCT

GGAGAAATTTTCGATGGTTCTCTTCTACGTTGGGGTAGCTCTACCGTTCATAGTTCATCTTGCACTGGAGGTGGGAGTG
CAGTTTTAAATACTGCACGACTCTATTGTGAACATTGAGCCGAGAGCTATTGAGAGATACAATGATAGGCAGCTGGGT
CATCAAGGTCACAGCGATGCTGAAACTCTGGGAGAGGATCCCTTATTCCAAATGTGCATGAGAGAATGCTTGAAGGAA
AACATACTGCATCACATTCAAATTGTCAGATTTTTCAATCTCTTCCAATATGTTACTTCGGCTATTCTTGGTTTTTCC
ATAGGAGTAAGCTTGATATTTCTAGCATCTCTGTGCGTTGTTCTTACAAAGATTCCAATATTCTCGTTTTGAAGCTGTG
AAGTTCGCATTTTTCTTTCTCTCTGGAACCTTTGCTTTGCGTTTCGGTTTTTGTATGGGAACCTCTCTCACTGAAGCA
AGTAAGCAGATTCCAAACGCTTTATACGAATGTCCTTGGTTTCAGACTGCTCAGGAAAGAGGAAAAGAAAACATTACTG
ATATTTGAAAGGAATACACAACAAAGTTTGGTTCTGAAAGGAAGTGGCCTCTTTGTTATCGATCTACAATTATTTGTA
CAGATTGCACGAACGACTTACTCAATCTTCAACATATTTTCTTCAAGCGAGAAATAA

>TRINITY_DN27120_c0_g1_i1.p1

TATGGTCCGATTCTATTGTTTCTGCTTTTGATTGGCGGTGTGTTTCTCTGCACCAAAGCAATTACACTGCTTGAGATC
AATAAGCTGAATGATGCCCACCTATTTAGGAACCTGTTCCGCATGGTGCAGTCTGTTCTGAGCTTTACAAGCTACATA
TTTGTGTTCTGCTGGTATGGTGAGGAAATAACACAGGCGGGAGCCGAAGTATTGACGCGACTGTACTTCTGTCCGTGG
TTCGATTACCCAGAAAATTGAAAAGGGCCATTACACAATGATGATCTGCAATGTGAAAGCTTCGCAAATGTCTGTA
GCTGGTCTCAAGACTCTCAATCTAAAGACGTTTGGCGAGGTAGTGAGTACAGCCTACTCTTACTTCAACATTGTGAGG
AGTATGAAGTAG

>TRINITY_DN53574_c0_g1_i1.p1

AAACGAGTGATTGTCTATTTTAGACCCATAATTGTTACAAAGACAATATTCTATTTTGAATTTTAGCTGATCTTCTT
TATGCTATCACTGAGCAAACTGCAGGCACTGCCCCAACAAATAAACTGTTGCTTCTGTCTAATGAGTTGTGGGTATGAA
CTGCTTTTAGACTGCTGGTGTGGAGAATATATTGCAAGGAAGAGTCAGCTCGTTGGTTTTGCAGCGTATAAATCAGAA
TGGCACAATATGTCATTCAATGTGAGAAGAAGTTTGATGATTGTCATTAATAGAGCCCAGAAGCCGGAACAGTTGAAT
GGACGTACTGGATTACATCCCATGTCACTGGAGACCTACATACAGATTTTCAAAGCATCCTATTCTACTATACAGTG
CTCCAAAACCTGAATAAGGCCGATCAACAACCTGCAAATATGA

>TRINITY_DN74751_c0_g1_i1.p1

GTTATAGCATGCTATTGCTTCTTCCAAATTTATGAGGGAGTGAAAATGATGTCACTGTACAGAATAACGTCTGGCCTG
CTTTTACTGAATTCAATTGCCATTCATGTTTTTATCAACTTTTACTGTGGAGAAAGACTCACCAGTACGAGCAATGAT
TTTCAAAACGATCTGTACTGCTCACATTGGTATACCAGCAGTTTTAAAGTGCGGAAAATGTTGATCATAGTGGCCGAG
TGTCAAAAGAAGCCCATTTTATTTAAGGCCGGCTCCATAGTTCCCATCACTCTGAGTTCATTTGTTACAGCTGTACAA
GCAGCAGTATCCTACTTCAATCTTCTGCTAGCTCTGGCTAAACAA

>TRINITY_DN34345_c0_g1_i1.p1

TTGTTGAAATTTCTTCGCGATTATGAAAACTGAGCCAACTATGATTATTGTTGACTTTTTTACATGCAATTATTTCA
CTGAGCTTCGCACTGCTGGAGATCACAGTTTACAAAAACATTATTGAGTACACTAAGATGCTACTGTTTCTGGTGATT
TGTATTTGGCATCAATTTCTGAACAACCATTTTCGGAGAGCTTATTATACAGAAGCAAATGAGCGTAGCCGATGCAGCT
TACGATTTGCCCTGGGTTGAAAGATCTCCACAATTCAAACTTCAATGAAAATTATGATAATGCGTTCTCAGCAGCCT
GTGATTTTAAATGGGCTAAAAATGTATTTTCTCTGTTATAAAACCTTCATGGAGTTCACGAAAGCCCTGATATCTTAC
TTCATGGTTCTGAGAACATTACAGGAAGGAAAATAA

>TRINITY_DN43788_c0_g1_i1.p1

ATATATCAGATTTACTCAACCACATCATTTGTTAGCCTGATGAAAAACATTTGCATGTTGGCAGCTCTAATTGTTTCAT
GAATTTGTGTACTTCTTTTTTCGGACAACGTTTAGTTTTCGAAGACCCAGGATCTCCATCAGTCTGTCTACTTTTGTGAA
TGGTATAATGAGACAAAGTCATTCAAGAGCTCAATGAACATTGTTTCATATAAGAATGAATAAATCATTACAAATTCTG
GCAGGAAAGATTGTTCCCGTTACTCTGGCTTCATTATCATATCGTGTGTTGAAAGCAGCCTTCTCGTACTTCAATTTATTG
AGAGCACTTTCACATACATGA

>TRINITY_DN3941_c4_g1_i2.p1

GTCAATTTTGACGGCGCTGTGCGATTTTTGTACTGATTAAAAAATTGATGGCCGGTGAAAGTCTGTTTGAAAACGTGGAA
TATGTGCCGTTACCCTTTGAGTACTCCTACACTTTTGTAAATGCAATATGGCGTCTTTCCTGGTCATCTATGCTGTT
CAGTCGATATGGTTCATGATGTCAGCTCCTGCTTTTGCCTGCTGTGTTTACATGGTTACAGTCCAGGCCATGAGACGA
GTTAAAACTGAGCTGGATCTCCTAGTCATAGCTCTAAATGATTTGGACAAATCTATTGGACATGACACAGACCTAGTC
GATGAATATCATCGGCTAATGATGCAACAAAACTGAAATCACAACCTCAGGAATTTTCGTACAACATCACTACTTTATT
ATACAGAGAGTGAATGAACATGAACTGAACTCACGCCTACTATTACTCTACAAGCATTCTGCTTTCTCATC
TGTTCTCTGCTTGTGGTACATTGTAAAAGGCGAAGAAATAGTACTGAAAGTCAAGTACGTGTCAATTCTTGATACCTCTT
CTAATCTATCTATTCTTATATGCAGCAGCTGGCCAGGATATTATAAATGGGCAAGAGCAGTTAGCCATAACTCTATTT

GAGTGTCTCTTGGATGGATAAACCGAAATGGTTTAAAACCTCGATAGCCACCATGATGATATTTGTGAACAAGCCTTTG
TACGTGAAACAGTTCGGCATCTATAAACTGGACATGCCCTATTTTTTCCAAGATTCTAAATGCTGGCTACTCATACTTC
AATCTTCTAAATACAGTGCAGTAA

>TRINITY_DN53423_c0_g2_i1.p1

CCGGCCGCTGGAGCACCTTCCACGTGCTGGTGGTGTCTACGTGCTGCACGTGGTGTCTACGTGAGCGGCCTGCTC
AACACCTACTGTCTCTTACCAGTGTCAACCTGGCCGCGCACCGCGTCTGCTACGAAGTGAAGATGCTCTGTCTGGCG
GTGAGCAGCGCCGCGTCGGCCGCCCAGCCCCGACCCCGTGTTCCTTGCCAATGTGGTTCGCTCATCACCAGAAAATTTGC
AGAACAGTGGAGTTGATTAATCAAAATTGTAACCTTCTCTGTTACAATAAGCAATCAATGCATTGGAATACAAATTTGC
CTCTATTTATACATGACTGTTGAGGCTGACAGCATTGGCTTACAAATGGTTTCACTCGCCAACCTTCTGCTTCATAGTT
TTCATTATTTACCTATACTGCGATGGAGGACAAAGTATCATAAATGAAAATGTGTTGTTGAGAGACTCGCTAGCTCAA
TGTCCCTGGCCGGGTAAACCGTACTGGTTCCGTCGTACCATTCTCTTCATGCTAACTAGAGCAACACCAGAGCTGAAA
GTCAAGCCGTTTGGTCTCTACATCCTTAATCTCTCCACATTCAAATCAGTTTTAAACGCTTCCTACACATTCTTCAAT
ATGATTAGAATCTCAGCAGACAATAA

>TRINITY_DN5629_c2_g1_i1.p1

AACGAAATGAGTGCCTTTATAAAATGGTTGAGCGCTTTTTCAACTCGCGCCTTCAGGCTCATTTATACGTTTGCGGCC
GTCCTACCGCTCGTTTCATATACTGATCGTGCTGTCCAAGGCACATTTTTTTCAGGAAAAAACTAGATGAAGACCGTATT
CCCTCAGTGATGCACCTTCTATGTTCCATGAAACTAAGAACCATTTCCTGGATACTTCCTAGCCAACCTGGCAGCCTCC
ATCTGGTACGCTTTCTCTCTTTACATCTGGAGCATAGTTTTCAAACCTTTTTGTTCATCGGCATTGGATGTATGTGCACT
GAAATGGAGCTGCTATTTGAGGCGGTCAACGAGATTGATCATATTCCACGAGCTGACAAACAATCCACAACATATCGAT
TGCACCATGCAATTCAATCAAGTTCTACTCAAGAGGAGGTTCAATGGTATTATTGAGCATCATCAAGATATAACCAA
ACAATGAAGCTGATGAACCGGATGTTTCAGACTACAATCGTTCTTTTTTTGAACGTTTACTGTTTGAGCTGTGCTTG
TACATAGTTTTTTGTTATGAAATTGGAAGATGTCGTTCAAAGAATCAAATTTGCAACACTTTATATCTTCGCATTGTAT
CTCCAATACCAATACGCTGCTTGTCTCAAAGGATAAAAGATCGGGGAGAGACATTTCGGAGAACTCTGTATGAATGC
GGTTGGGTGGATAAGCCACAATGGATGAAAAAATGCTTTTGATAATGAAATGCATGGCAGATCAACCCCTGGGTTATG
AAGCCCTACGGCTGGTATGTCGTCGATCGTGTATTTCATGGCAAATATATTCAAAGCAACATACACTTATTTGAATGTT
GTCAACGAATTCATCAAAATAA

>TRINITY_DN7031_c0_g1_i1.p1

GACAAGTATTTTTCAAATTGTAATGATAAGAAGAATAAACAATGCATGAGAAAAATGCAATTAATTGATGCGATGGAG
CAAAAACTGTCAAATTGGTCAACCAGCTTGAAATGATTTTGATTTGCTTCTACACTGGATACATGCTCGCTCCAGCT
TTCGGTATAATTCTCCATGTCATGAAATCACAAAGTTTGAACAAAGTAATACTGCCATTTCATGTTTTTCGTATCGCAA
GAAAGTGCCGCCAGGAGCCCACCAGTATTTTCAAGCCTTCAAACCTACTGCCAGTTTTTCTAGTGGAGTCATTCTAC
ATCTACATACAGTTGATCTGCACAGTCATTCTCCACACTCGCTGCATCCTCACTGTCAGCCATGTTTCCATCGAAATT
GATCTGTTCTGCTTGGATATTGAGCAGCTCTGTTTGGAAAATGTTGAAACGTTTCCAGAAGGTTACTACAATGAGAAA
AATTTAGCACATCGGGCGAGAAAATTGGAAGAGGTTGTGAGGAATTTAGTCATTTCACCACCAAATAATAATGAAGAAG
GTGCAGAAGCTGAATAGTTCTTTGGTGTCTATCCCTGATTTACGGGAATTCATTTCCTGACCATCCAGCTGGCAATATTG
ATATTCTGTTTACAGCAGCATCCAGATTTTCTCCAACGCATTAGATATGCAATTATGATGATACCTTTGCTGTTTCATG
CTTATTGCGTACTCTGAGTATGGACAGAGAATAGAGAACCAGGGAAATAAGATGAGACTTGCTCTTCACAATTTCCCT
TGGCAATCGAAACCGAAATGGGTACAAAGCACTCTTAAATTTCTCATGGTTTCGTTCAAATAAGTCTCCTCAAATCAGC
AGCGTCTACAACCTTTTTCAAACAAAATCGACAGAATATGAGTAAGCTGATGAAGTTCGCCTACTCCTATTTCAACCTT
CTGGAAAGAATTTCTACTTGA

>TRINITY_DN76106_c0_g1_i1.p1

TCACATTCCCGATGCAGGATATCTCAATCCGTGGTTGCACTAATGAGGCCGTTTCATATTTCTAAAAACGTTCTTTAGT
TTTGGGATGCTTAGCATTGGTCTCTTCACTACTCTGACTGTGGGAATTGGCAGTCCCTCTTTCTTCACGGCAGGAGGA
TTTCTGTTTGTAGCCACTCTGGAATTGTTTCTTGTATGCATCTATTCCGAAAAATCTATCAACTTCGAGTTTGGCCGTC
TATAACGCAGCCTACAATGTGAGCTGGCCTGACATGAAGGAGTCTCCAGAAAAAGTCTGGTGTGTTAATGTTACGC
GCGCAACGACCTCTTCAGTTGAATGTGAAAAGCAATTTTGTGCCTCTAGATCTCAACACGTTTCATTTCAGGTGATACAA
ACATGTTATTTCGTTTTTCACTGTTCTACGTGAAACACTTGGCAAACAAAACCTGA

>TRINITY_DN8662_c0_g1_i1.p1

TGCTTCAACCATATGAATAACCAAACCTGTATCGAAGAATATCTTTAAAGAAAAGCATCGAGTTCCCAACATCTCCATC
CTAACATCCAACATCACACCCACACCAGCCTATGGGTTCTTTAACTCATTTGTATTAATTTTGCAAATTGATTTGACC
ATCAACTTATACACAACATGGCATGACGTCATTAGCAGATTGTTGACTTTAAAGGAATTCCTACTGTTATGGAGCGGA
ATGCTAGGATATTCCGCATCAAACGATTACAGTAAGATGCTCCGATACGTTTCACAGTCATTATAAGAAGTATGAAAAAT

AGAGAAAATTTTACAAATGAAGAGAGGGAGATACTGGAAGACCGTGATAAGCAATTGTTGAAATACGACAAAGTTGTG
GTTTTTCATGTTGAAGTATATTTTGCACATTATGTGCGGTCTCCTACCCTACGTTGTTATACTTATATCTATCTACAAG
CGTTACAAAAGAGGTGAAAATTATGAGAAAATACCTGAAATAATTCACACCCACTTTCCCAAGGCGTTCAACACTTCA
ATTCTGTTTCAACTGGTGTACCAAACCGGCATTTATTTATGGTACAGTATGTTGTTGTACAATTGGTTCGCACATATCG
AGAAGCATGCTGTTTCAGTTTTCACCTGTTTGGATACGGAAATCAAACCTGTTGTCGCTGACTTTGAGGAAAATTGAAGAC
TGGAGTGGGAAGGAGAGAGAGAGGGGGGAAAGGAAATTGGATCAACGGTTGAGGAAAATTGTACTTCACCATCAGTCG
ATTTGCAGAAGATTATTGGATTTGGAGAAATATTCCAGTAAAGTTGTTTTTCGGATTTGTGAACCTTGTATGGATCACAA
ATGTGTATTTACTTTGGCTTTTCATCACTCTGATAGATGACCTGGTGACCAAGCTGAGGTACATATCTCGCTATTTGATA
ACTTTTGCCTTCGTTTACTATACCTGCGATAGCGGCCAAACTGTCACTGATCTGGGAAAGAAATTGAGATTTGCATTC
TACGAGTGTCTCGTGGGTGACAAACCCCTATTGGTTTAGAAGAACAATATTTCTTATGAACTTGAGAGCCTCTAAAGAA
ATGGTAGTGAGACCATTTGACTGTATGTTCTCGATATTGGCTGTATTGTATGGATAGTGAAAGGTACTTATACATAC
TTAAACTTCGCCAAGAAGCTACTCGATTAG

Correceptor de Receptores de Odorantes (ORco)

>TRINITY_DN24825_c0_g2_i1.p1

ATGCTAATTGTTGTATCTCCACTGCTTTTGGCTATATTCGGAGAAGGAGATAGACCAAAGATCAAGGAAATCAACTAC
AAATTGCCAATTCGGTTTTTGGTTCCCTTTTCAGCACTGATTCTTGGACAGGGTTCACTGTTGCCTTTCTGCTCTTGTCT
GGAGAAATTTTCGATGGTTCTCTTCTACGTTGGGGTAGCTCTACCGTTTCATAGTTTCATCTTGCCTGAGGTTGGGAGTG
CAGTTTAAATACTGCACGACTCTATTGTGAACATTGAGCCGAGAGCTATTGAGAGATACAATGATAGGCAGCTGGGT
CATCAAGGTCACAGCGATGCTGAAACTCTGGGAGAGGATCCCTTATTCCAAATGTGCATGAGAGAATGCTTGAAGGAA
AACATACTGCATCACATTCAAATTTGTGAGATTTTTCAATCTCTTCCAATATGTTACTTCGGCTATTCTTGGTTTTTCC
ATAGGAGTAAGCTTGATATTTCTAGCATCTCTGTGCGCTTGTCTTACAAAGATTCCAATATTCTCGTTTGAAGCTGTG
AAGTTTCGATTTTTTCTTTCTCTGGAACCTTTGCTTTGCGTTTCGGTTTTTGTGTTTATGGGAACCTCTCTCACTGAAGCA
AGTAAGCAGATTCCAAACGCTTTATACGAATGTCCTTGGTTTCAGACTGCTCAGGAAAGAGGAAAAGAAAACATTACTG
ATATTTGAAAGGAATACACAACAAAGTTTGGTTCTGAAAGGAAGTGGCCTCTTTGTTATCGATCTACAATTATTTGTA
CAGATTGCACGAACGACTTACTCAATCTTCAACATATTTTCTTCAAGCGAGAAATAA

Proteína de Unión a Odorantes (OBP)

>TRINITY_DN11975_c0_g1_i5.p1

ATGTTGTACGTATTATATTTTGTATTGTTACTTTCAGCTCTGTGAGCTGTGATAACTCAAATCATGGCTGCGGACTCA
TCATCAGATATGCTCACAGTTTTTCAACAAATGCAGGGAAGAAACCAGTGCAACAGAAGATGATATCAAGACTTTTCAGA
GCTCAACAAATTCCTACCACAACAACCTGGAAAATGTATGCTTGCCTGCATGTTCAATCACTCAGGGCTAATGAAAGAA
GGCAAATACGACTCAGAGGGTGCCTGAAGCTTGTGGGTGAGTTTTTGCAGATAATCCAGTCAAACCTCGGCAAGGCA
AGAACGCTTATCAATTCATGTTTCAGATGAAGTGAAAAACGAGAATGACAAATGTGAAATTGCTTCCAAAATCGCAGAC
TGTACCGTAAAAATGACCTCTCAGGTTGGTCTTTCTTGA

>TRINITY_DN14064_c0_g1_i1.p1

ATGTTGCTCGAAGTTTGCAGATTTTCTGTGTTTCTGTTTGCCTTTCTGCTACAGTCAATGGACGATTCTCTGAAGAA
GAGAAGCAACTGATGAATCAAGTTCACTCACAATGTGTAAGTGAAGTGGCACAGCAGAGGATGTAGTTTCCAAAGCG
ACCAAAGGTGATTTTGCAGATGATGACAATCTCAAGTGCTATGTCAAATGCATTTGGAGTACCCTGACTGTGATGGAT
GATGATGGCAACTTTGATGTTGGTGTCTTAGAAGTAATGTTGCCAGCTGATATGAAGGATACAGTGAAGCTATG
AGCGCCTGCACAGGAGTAGGTGGGGGATCACCGTGTGAAAAAGCTTTTGCAGTCACGAAGTGTGTTGTATAAGGAAGCA
CCAGCTGATTTCTTCTTACCATAA

>TRINITY_DN1864_c0_g1_i1.p1

ATGAAGGCTTCAGCTGCACCTTACACTTGTCTTTCTAACGTTGGCTGTATTCCATTGCTCAGAGGCCAAAAATTGATAAA
GCAAAGAAGGAAGCTGCTATTAAGAAATGCCAGGCTGAGACATCGGCCTCTGACGAGGATGTGATGAAAGTGAGAAAG
GAGCATGTTGTGCTGAATCTGAGGAAGGAAAGTGTTCATTGCTGTGGATTCAACCATTACGACATGCTGAAGGAC
AATAGAGTCAACCTGGAAGGAGTGAACGCGTTCTTTGAGAAGCTGTATGACGAACAGGAGAAGCGCGATATCGCCATC
AAGGCCGCTGCCAGCTGTGCTGCCACTGAATCTGTAAGCGGACTGAACGATTGCCACGTTGCAGCCAAGTACTTCGCC
TGCTTGCAAGGATGAAGGACGACTTTGATGTTTAA

>TRINITY_DN206_c0_g1_i10.p1

AGAATGTGTAGCCCAACAATGAAGTCAGTTGGTTTGTGCATTCTAGTTGTAGCTGTTATGGTGTCTCAGCCGATCCT
CTATCAATGGCGAAGGCGGCGACAGCAGCGACAGCGGCGACTCCGGCGACAGCCGCGACACCGGCGAGAGCCGTTCTG
CCCCCTCAAAACATAGATGATTCTCTAAACAGAGCTAGAAGGGAAGCAGTGGAGATGAACCTGGTAGAGGAGCCAGTT

AAGAGAGTTGCCAGGGAATCACCTGAAGAGAACAAGGCATGCCAACCTACAAGACCAGAACAAGGAGCTAAGGCATGC
TGCGAGTTGCCATTGGTGTTCGCGGGAACCCCAAGATTGTTCAAGGCATGCAAAGAAGAATTGGGTATTCCTGACCAT
AAGCCACCTCCACCACCAGAAGGGCGTGACCTCACGGACCTCCACCACGTGGAACGTGCGTTGCTGAGTGCCTTTTC
AACAGGACTGGTTTTATTGGAAGGAGGAAAGATAAACAAGGAGGCATTACAAAAAGCTCTAGATGAATACTTGAAACT
GATGGAGCTTGGAAGATGTGGCAACATCTACCCTCAATATATGCTATGATGCTCAAAGCAAAGGTGACTTCAAACCA
GACAATGAGAAGTTTCAGTCTGGTTCATCTGAATTCCTGAGATGCTTCAGCAGAAGCCTGTTTATGAAATGTATTCCA
GAAAAATGGACTGACACGGAAGAGTGCAACAAAATGAAGGAGAAGCTTGACAAATGTCCAAAAATGGTGCCTCCAGCA
TTATTCAACAAACGACCTCACTAA

>TRINITY_DN3903_c0_g1_i1.p1

ATGAAGTGTCAAGTGCTACTGGCTTCCTTTGTTTTGGTTGCAGTTTTTCGATCTCAGCTATGGGGGTCTGACCCCAGAT
AAATTGAAAGAACTCAAACCATTGATTGACACTTGTATCAAGCAATCAAAGTTGAAGAAGAAACACTTGGAACCTG
CACAATGGTCACGAAATTCCATCCTCTCAATCTGGAAAGTGCTTTATTGCCTGTATGGCTGAGCATATGAACTGATG
AAGGATGGAAAATTGCAACCAGAAATGACAATGGAGTTCATTGATAAGATGGTTCAGGACAAAGATAAGGCTGCCGAG
ATCAAAAAGGCCGTGATGAATGCTTCAAGAGTGTTGCCGATACTGGAGACAAGTGTGAAATGGCAGCTGAAGTAGCG
ACTTGATGAAGGATCATCATGCTGAGCTTGCCGGTATGAACTGA

>TRINITY_DN43095_c0_g1_i1.p1

CTAGTCTCGCCACAGGTCGGTCTTCACGCAAAAATGAACTCGTTTATTGTTTGCGTCGTTCTCTCCTGCCTTCTGGTA
GCCAACATCAAGGCTGACGAGGCCACCTCGTCGTCGTCTCCAGATGCGGACACGCTGATCACATCGACGACTCTGTG
CCAGCCTCGAACGAAACGGACGCCGCCGCCGCCATCAAGGAGCAACTAGCCAACTCACTGAGAGCTGCAAAACC
TCCAGCCAAGCAAACCTCTGATGATGTTAAAATTATTGAACTGAATCAGTGCCAAAGACTGAAGGAGAGAAGTGCTTC
CTGCAATGCGTCTATGGTGGATTGGGAATTGTGAAGCATGATCAATTCTCAGTAGAAGGAGCAAAGCTATTAGCACAG
AAACGGTTTGATCTTTCCCTGAGGAAGTGGAAAAGGCCAATCAGCTGATTGAAACGTGTTCAAAGGAAGGGATCAGT
GAGGAAGGAGGATTGAAGATGAAAAGGAGGAGAAGACGGAAGAGAAGGGTGAGGGGGACTTTGCCGAGAGTAACAATT
ACCAGACGAGGAATAAGAAGGAGACGGCCGTGA

>TRINITY_DN5192_c0_g1_i6.p1

ATGTATTTCTACCCAGTTTTTCTGTCTTCCATTGCTTTTTTAGCGATCATCCAAGGCGGGAGGGCTGATGGTGGAGAT
GGTGACGAAGACCAATACAGAGTTCTGATTGGAGATTGGATCCAGCAATCAATGCAATGCAAGAAAGAAATGGAAGAT
TCAGGAGTGGAATGAAATTGATGCCTTGTCTCGCTTTCTATGAGGGAGGTAAGAATATTGATCCAAAATTGAAGGAA
TATGACAGCTGTAAGGCATTGATGGCGTGTGTTATGAAGAAATCTGAAGTTATGTCTTCAGATGGAACATCGCTTATC
AAAGACAAAATGTTGGAGGTAGCCGGAATTTTCGAGCGCAAACATAAAGGAATAAAAGATGAATTCAACGATGTCATT
TGGCCAGCCTGTGAGAAATCTACTGGAAGAGATGGACAGATGGCCTACAATTTTCATGGAATGTGTTTTCAAAAAATCT
GCAATGGCTAAAAAGCTGAAAGATAGGTTCTTTGCCTTCTTCCACGATCAACAGGACGACAGTAGATGA

>TRINITY_DN5517_c0_g2_i1.p1

ATGAGCGGCTTCCATAAGTTTGTAATTTTCGGGCATGGCTGTACTGGCCGGAGCTATGTTGATTAATGCAGAGGACACT
ACTATCAAAAATAAAGAACCCATCTCCACACAAGCAACAGCAAGTTTACTGCCAAGCTCCCCCACTGCTCCAGAAAGG
CTAGAAAGGATTATCGAACAGTGTCAGGATGATATCAAGACTGCTCTTCTACAAGAGGCTCTCAACGTTTTGACTGAC
ACATCTCCAAGAGATCTCGTAAAAAAATCAAGAAGCAAGAGAGAGGTTTTCTCCGGTGAAGAAAAGAGGATAGCAGGG
TGTCTGCTGCAATGCGTTTACAGGAAAGTGAAAGCTGTTGATGACCAAGGAATGCCACAGTTCCAGGCCTTGTGAGA
CTTTACTCGGAAGGCGTCCAAGACAGGAAGTACTATGTGGCCACGGTGCAAGCTGTACAACAGTGTGTGTCCGCCTCC
CAGCACTTCCGCTACTACAACCCTCAAGTGCTCAAAGAAGATGGATACACGTGCGACTTGGCTTATGATATGTTCAAC
TGTGTAAGCGACAAGATTGAAGCATTCTGCGGCAGAACGCCTTGA

>TRINITY_DN5568_c0_g1_i1.p1

ATTTGTTGGAAAAAATCATCTTCAAACATTTTCAAGAGCAGAATTTTCATTTTCATTGTTCTCAGTTGGTGTATCAGTTCA
ACGATTTTCAAAGTCTCTTGTAAGCAAAAATGGAGAGATCGCATGTATCAATAATAGTTTTTGTCTTTCATCCCATTTTTTA
TCTTCAGCTCTGGAAGCTAATTTTGCAATGATGCAGCAATCACTGCAGGGAAGTGGTATACCAATGATTCAATCAATA
GCTGGCGAATTGAAGTATTGCATGGATGTAAATGCAGACCAGGAAAGAGATGGACTTGAGGACTACCTTCCCCTACTA
TTTAATGAAGAGCTGCCATCTTCTCTTGGTCAAAGTGCTTTCTCACCTGTCTGTTCAATCGGTTTGGTTTGTAAAA
GATGGATTTCTCGACGCACAAACAGCGAAAACGTTGGTAGAACTTTTTATAAAGATAAACATGACGAGAAAACGATG
GCAAATATTGCAATAAACGTTTGCCGTGTATCAGCCGTGCCAGATGTACTGAATCCTTGTGAAATTGGTTTCTCATTG
AAGAGCTGCTTCGTTGACAGCAACAAAAGGGTAAAGAATTAAGAGGCAGGAAGTAG

>TRINITY_DN5926_c0_g1_i1.p1

ATGGGAATTTATACGACAAACTTGATTTTCACACTTCTTGGCTCAGCGGTTGTTTCTGGGGTGTTTCATAGATGGAGAG
AATGGTTACAGATTGACCAGGCAGGCGCCACCTGATGACGAGTGTGCGCCACCAAGGCCCGGCTCCAGCGAGGACGGT
CTATGCTGTGAAATGCCTCCTGTTTTCCGAACGGCTCACGATAAATTCGAGAGCTGCTTGGAGGAGTTGAGCAGCATT
TTCCCCCTCCGCTCCTCCGCCCCATGGACATCACGGACCGCTCCTCCTGGGGGTCCAGGGGGCTATGGACCCCCCT
CCCCCTCCTCCCAATGGAGGCAACCGAGGCCACCCCCCGGATTCGGAGGACCTCCTGGACATGAACCACCAATTTTT
TCTTGCGCTCATGAATGTTTGTTCACAAAACCTGGACTATTGGAGAATAACAAGGTGAACGTGGAAGCATTGATGAAG
AAACTAGAAGGCGAACTGGGAGACGATGAAGTGTGGAAAAAGTTGGTGCAATCCATTGTTGACAAATGCATGGAGTCA
AAAGACCTCCATCAAACCAAATGTGCACTAGTGGATCCCATGAGCTTGCGAGATGCGTTCTCAGAGACATGTTTATG
AACTGTCCACAAGAGAAGTGAAGGAATCCGACGATTGCAGTAATATGAAGATGAAATTGGAGAAGTGTCTGAACTT
GTCCACCAATGTCAATGAGACTTTCAAACCTCCATCAATGCCGTGA

>TRINITY_DN6387_c0_g1_i4.p1

CCACGAAGTGTAGTTTCATCCGCAGCTATCGACATGATCACGTCCGGACTAATGCAGGCTGTGACTGCTGCCTGCCTA
TTGCTGGTCAACACTTACGCCTATGATTTTCAGTGATCCCTATTTCAATGAACACCTCCAGAGCGCAATGGAGGAGATA
ATGGAAGAGGAACCTTCTATCCATTGGCAGAGTGCAGAGAGACGCAGATCAGAGCCAAGAAGTTGCAGATGAGTACTTC
AAATGCAAAACACAGGAACCTGAAGACATGTTGCGGTAAAATCAATCTGATGAAAAACTATGGCGACAAAGGAAAGATC
TACGGCAAGCAGTGCTATGAGGAAGTGGTCAGCGCATTCAAAAACCAATTCTTCATCAACAGCAGATGATGATGAAAGT
ATGATGGACATGTTTTCTGTGAGAAAAGTGAAAATGATCAAATTGAAGCATATTTGCGTTTACGAATGTATTGGAAAG
AAAACAAAAATCCTGAAAGAAGATGGTTCTTGAATGCAGAAGAAATCAAACAGTATGCAAGAGAATACATGTTCAAT
GAAGAATGGTCCAAGGAGCTGGGAGAAAAAGCACTCGACAAGTGTCTCACTCAATCTTACAACAGCGTGACCAAGATG
TTAGACGAATATGAGATCAAATGCAACCCAACTTCGGTTTCACTTCCATCACTGTCTCTGGAAGAAATTGAAATGACT
TGCCCTGAATCAAAGGTTGACCTGAGTAAGTAA

>TRINITY_DN6777_c0_g1_i1.p1

ATGAGTGCCTAATCAGAGTTGGTTTTGTTTTTTGGTTTTGTTTGTGTTTCATACAGTGAGGCTAGTCCGGCATTAAACA
GAGGCACAGATAGAGCAAGTTGGGAAAGCAATGGCGAATATGTGTATATCTTCATCAGGCGTTCAAAGGAGTTTAATA
CAAAAGGCAATGAAAGGAGAGATAGAAGATGACCGTAAATTGAAATGTTTCTTTGGTTGCATCATGGAAGCCATTCAA
GTGACAAAAAACGGTAGAATGCAGCCAGAAGTACTGAAAAGGAGAGCTAACGCCATGCTGCCGAAAAACAATGAGAGAA
ATGATTCTGCCAACCATCGATTCTTGACGCCATATTGAAAATGAAGACAAATGCGAATTGGCCCATTCGATTGTCAAG
TGCCACTTCTCTGTCAACGGGAAAAATCCATTCTTCTTCACTTTTAA