



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

Trabajo de Fin de Máster en Biotecnología Molecular y Celular de Plantas

# **Estudio bioinformático de la familia de genes Patelina**

Trabajo Fin de Máster realizado por

**Estela Alemany Campos**

Director

**Dr. José Miguel Mulet Salort**

Codirector

**Antonio Bustamante González**

**Valencia, enero 2017**



## ÍNDICE GENERAL

1. INTRODUCCIÓN.....	1
1.1. Estrés por frío .....	1
1.1.1. Concepto de estrés.....	1
1.1.2. Tipos de estrés. ....	2
1.1.3. Relevancia del estudio del estrés en plantas.....	2
1.1.4. Mecanismos generales de respuesta a estrés. ....	3
1.1.5. Estrés por bajas temperaturas.....	5
1.1.6 Sec14.....	6
1.1.7. Patelinas, familia multigénica.....	6
2. OBJETIVOS .....	11
3. MATERIALES Y MÉTODOS .....	15
3.1. Bases de datos.....	15
3.1.1. UniProt.....	15
3.1.2- GeneBank.....	16
3.1.3. Protein.....	16
3.1.4. TAIR.....	16
3.1.5. Prosite y ProRule .....	17
3.2. Software bioinformático .....	17
3.2.1. BLAST.....	17
3.2.2. Clustal Omega .....	18
3.2.3. TreeView.....	19
3.2.4. Coils .....	19
4. RESULTADOS .....	23
4.1. Elección del gen. ....	23
4.2. Características generales del gen <i>BVCRIO4</i> .....	24
4.3. Búsqueda de secuencias homólogas.....	28
4.3.1. Estudio filogenético de CRIO4.....	28
4.3.2. Patelinas conservadas en otras especies.....	30
4.4. Clasificación patelinas en función de la estructura de dominios .....	32

5. DISCUSIÓN.....	41
6. CONCLUSIONES.....	47
7. BIBLIOGRAFÍA .....	51

## ÍNDICE DE FIGURAS

Figura 1: Secuencia de nucleótidos del gen BvCRIO4 .....	25
Figura 2: Estructura de dominios de la proteína CRIO4 .....	26
Figura 3: Predicción de <i>Coiled Coil</i> en la proteína CRIO4 .....	27
Figura 4: Diagrama de la estructura de dominios de CRIO4 .....	27
Figura 5: Cladograma de CRIO4 y la familia de PATLs de <i>Arabidopsis thaliana</i> ..	28
Figura 6: Cladograma de CRIO4 y las PATLs de las especies más cercanas. ....	29
Figura 7: Cladograma de CRIO4 y las PATLs de las especies que presentan motivos ricos en prolinas .....	30
Figura 8: Relación de patelinas conservadas en otras especies (%).....	31
Figura 9: Relación del número de patelinas conservadas por especie (%) .....	31
Figura 10: Predicción del dominio <i>Coiled Coil</i> en las patelinas de <i>A.thaliana</i> .....	33
Figura 11: Representación gráfica de la estructura de dominios de las PATL1 y PATL2 de <i>Arabidopsis thaliana</i> .....	33
Figura 12: Representación gráfica de la proteína CRIO4 .....	34
Figura 13: Representación gráfica de PATL3 de <i>Beta vulgaris</i> .....	34
Figura 14: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRIO4 con a la estructura de PATL1 y PAT2 de <i>Arabidopsis thaliana</i> .....	34
Figura 15: Representación gráfica de la estructura de dominios de la PATL5 de <i>Arabidopsis thaliana</i> .....	35
Figura 16: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRIO4 con a la estructura de PATL5 .....	35
Figura 17: Representación gráfica de la estructura de dominios de la PATL4 de <i>Arabidopsis thaliana</i> .....	35
Figura 18: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRIO4 con la estructura de PATL4 .....	36
Figura 19: Patelinas conservadas en otras especies que rigen la estructura de la PATL4 de <i>Arabidopsis thaliana</i> .....	36
Figura 20: Representación gráfica de la estructura de dominios de las PATL3 y PATL4 de <i>Arabidopsis thaliana</i> .....	37

Figura 21: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRIO4 con la estructura dominio Sec14 y dominio GOLD.....37

Figura 22: Patelinas conservadas en otras especies que solamente han conservado los dominios Sec14 y GOLD.....37

# **INTRODUCCIÓN**





## 1. INTRODUCCIÓN

### 1.1. Estrés por frío

#### 1.1.1. Concepto de estrés.

*Lambers et al.* (1998) define el estrés como cualquier factor ambiental, biótico o abiótico, que reduce la tasa de algún proceso fisiológico, por debajo de la tasa máxima que podría alcanzar. Y especifica tres escalas temporales en la respuesta de las plantas al estrés:

*Fase de alarma:* es el efecto inmediato cuando se presenta el estrés. Ocurre en una escala de segundos a días. Las plantas reaccionan ralentizando o deteniendo sus funciones fisiológicas básicas, reduciendo su vigor, y activando mecanismos específicos para hacer frente al estrés. Las plantas que no poseen mecanismos adecuados de defensa o respuesta frente al estrés experimentan daños irreversibles y mueren. El desenlace es el mismo cuando la situación de estrés es muy intensa y supera la capacidad de respuesta de la planta.

*Aclimatación:* es el ajuste morfológico y fisiológico que realiza la planta, como individuo, para compensar su peor funcionamiento después de la exposición al estrés. Ocurre en una escala de días a semanas. La activación de los mecanismos defensivos o de respuesta da lugar a reajustes del metabolismo celular para adaptarse a las nuevas condiciones, a la activación de los procesos de reparación de la maquinaria celular dañada y a la expresión de las adaptaciones morfológicas.

*Adaptación:* es la respuesta evolutiva que resulta de cambios genéticos en las poblaciones, conduciendo a una compensación morfológica y fisiológica. Ocurre en una escala temporal mucho mayor que la aclimatación, y tras muchas generaciones.

### **1.1.2. Tipos de estrés.**

Entre las posibles clasificaciones de los factores de estrés, en general, podemos clasificar en dos grandes grupos, estreses bióticos y abióticos (Azcón Bieto y Talón, 2008). Siendo los estrés bióticos los originados por la acción de otros seres vivos: animales, plantas (competencia, alelopatía, etc.), microorganismos (bacterias y hongos), y otros agentes fitopatógenos (virus y viroides).

Y por otra parte, los estreses abióticos, dependiendo del agente causal, podemos dividirlos en físicos y químicos. Entre los estrés físicos encontraremos el estrés por déficit o exceso de agua, temperaturas extremas (calor, frío y congelación), salinidad (en su componente osmótico) y radiación UV. Y entre los factores químicos destacan la contaminación atmosférica por metales pesados, compuestos tóxicos de diferente origen, salinidad (en su componente iónico o tóxico) y carencia de elementos minerales esenciales para la planta.

### **1.1.3. Relevancia del estudio del estrés en plantas.**

La importancia del conocimiento de los factores de estrés en las plantas de se deben a los siguientes motivos:

Elemento crucial para la elaboración de modelos predictivos, como por ejemplo en el estudio de los efectos en el tan inminente y problemático cambio climático que estamos experimentando.

Para el análisis de la interacción de las plantas con los factores ambientales con el fin de comprender la distribución de las especies en los diferentes ecosistema.

El rendimiento y capacidad de supervivencia de los cultivos limitado a los impactos de estrés ambientales, siendo estos los factores que más limitan la productividad agrícola a nivel mundial. Estos estrés no solamente tienen un impacto sobre los cultivares actuales, sino que también son un obstáculo importante en la introducción de cultivos en zonas que actualmente no están siendo utilizadas para la agricultura (Gaspar et al., 2002).

Siendo crucial comprender los procesos fisiológicos subyacentes en la tolerancia de los cultivos al estrés a la hora de establecer programas de mejora genética, tanto desde abordajes tradicionales como biotecnológicos (*Ali Dib et al., 1990*).

#### **1.1.4. Mecanismos generales de respuesta a estrés.**

Las plantas están ligadas a su hábitat, no pueden huir de la amenaza de muchos agentes estresantes ambientales y antrópicos, y por lo tanto necesitan mecanismos especiales para evitar el estrés o adaptarse a él (*Lichtenthaler, 1998*). La flexibilidad de su metabolismo les permite hacer frente a su entorno en constante fluctuación (*Gaspar et al., 2002*).

La manifestación de las respuestas de las plantas frente a unas condiciones ambientales adversas implica la puesta en marcha de una secuencia compleja de acciones. En primer lugar, se produce la percepción por la planta del estímulo estresante, siendo este proceso el menos conocido. En general, a señal de estrés es detectada por una serie de receptores y seguidamente es amplificada a través de cascadas o rutas de transmisión de señal hasta llegar al núcleo de la célula.

La transducción de señal produce cambios transitorios en los niveles de determinados iones, como el calcio y de moléculas como lípidos, especies reactivas de oxígeno, especies antioxidantes, óxido nítrico, entre otras que actúan como segundos mensajeros, amplificando la señal de estrés a través de la célula hasta inducir un cambio en una molécula efectora, como una quinasa o un factor de transcripción. Las hormonas también tiene un papel importante en las rutas de transmisión intracelular de la señal de estrés, donde se ha descrito como el ácido abscísico participa de forma activa en la señalización de muchas respuestas al estrés abiótico (*Toumi et al., 2010*), y se ha detallado también la implicación del etileno, el ácido salicílico y el ácido jasmónico, en la transmisión de la señal de infección por patógenos (*Dempsey et al., 1999; Dong, 1998; Jameson y Clarke, 2002; Vlot et al., 2009*).

## Introducción

Finalmente tiene lugar la regulación de la expresión génica en el núcleo donde se induce o reprime la expresión de diversos genes, o los factores de transcripción que los regulan, lo que a su vez provoca cambios en el metabolismo celular, que se traducen en alteraciones en el crecimiento y desarrollo de la planta, cambios que según *Azcón-Bieto y Talón (2008)* pueden ser entre otros:

- Alteraciones en el desarrollo de la planta. Normalmente se aprecia un menor crecimiento vegetativo, así como una reducción del número de estructuras reproductivas que aceleran su desarrollo para poder asegurar la siguiente generación.
- Muerte celular y abscisión de los tejidos dañados, lo que elimina el foco de la infección en estrés biótico, disminuye la superficie de transpiración y permite reciclar nutrientes.
- Aumento o disminución en la actividad de las rutas alternativas de disipación y obtención de energía, como la fermentativa.
- Cambios en la actividad hormonal. Además de participar en la percepción y transmisión de la señal. La modificación de los niveles hormonales puede incrementar la resistencia a estrés.
- Síntesis de nuevas proteínas, como la ubiquitina, las proteasas implicadas en la degradación de las proteínas dañadas o marcadas con la propia ubiquitina , y las proteínas de choque térmico. Estas últimas no solo inducidas frente a temperaturas extremas, actuando como chaperonas plegando proteínas desnaturalizadas por el estrés y previniendo la formación de agregados proteicos irreversibles.
- Síntesis y acumulación de compuestos osmoprotectores que actúan restaurando el potencial hídrico o bien como protectores de la estructura de membranas y macromoléculas.
- Síntesis de metabolitos secundarios protectores, como los fenilpropanoides.

### 1.1.5. Estrés por bajas temperaturas.

El frío está considerado como uno de los estreses abióticos más importantes, afecta severamente al crecimiento y desarrollo de las plantas, y limita su distribución geográfica. Solamente un tercio del área total de la Tierra está libre de heladas y el 42% de la superficie terrestre soporta temperaturas por debajo de los  $-20^{\circ}\text{C}$  (Larcher, 1995).

Para entender el efecto del frío sobre las plantas, conviene distinguir entre aquellas temperaturas que se encuentran por encima y por debajo de los cero grados Celsius. Puesto que el término “estrés por bajas temperaturas” puede resultar arbitrario. Se puede clasificar en función de la temperatura en: “estrés por frío” para referirse al estrés causado por temperaturas comprendidas en el intervalo entre  $0$  y  $15^{\circ}\text{C}$ , también denominado *chilling stress*. Y por otro lado, “estrés por congelación” para designar al *freezing stress*, estrés causado por temperaturas por debajo de los  $0^{\circ}\text{C}$ . Siendo el enfoque de este trabajo el estudio del estrés por frío.

Otro enfoque es la clasificación en función de la sensibilidad al frío de las plantas (Goñi, 2010):

- Plantas sensibles a las bajas temperaturas (*chilling-sensitive*): especies que presentan síntomas de daño tras tiempos prolongados de exposición a temperaturas por debajo de los  $2 - 15^{\circ}\text{C}$ . Se entiende por planta moderadamente sensible aquella que se altera en el intervalo entre los  $2 - 7^{\circ}\text{C}$ , y muy sensible cuando no soportan temperaturas por debajo de  $15^{\circ}\text{C}$ .
- Plantas resistentes a bajas temperaturas (*chilling-resistant*): son capaces de tolerar bajas temperaturas por encima de  $0^{\circ}\text{C}$  pero son seriamente dañadas cuando se forma hielo en su interior.
- Plantas resistentes a las heladas (*frost-resistant*): especies capaces de soportar temperaturas por debajo de los  $0^{\circ}\text{C}$ .

La respuesta al estrés por frío es un proceso complejo que implica muchos cambios a nivel celular. Se produce una reprogramación de la expresión génica,

modificaciones en el metabolismo celular (*Chinnusamy et al., 2010*), y cambios en compartimientos celulares (*Miura y Furumoto, 2013*).

### 1.1.6 Sec14

El gen de remolacha *BvCRIO4*, objeto de estudio en este trabajo, codifica una proteína con un dominio Sec14. El dominio Sec14, también denominado CRAL-TRIO, representa una unidad estructural, ancestral y versátil, que está restringida a eucariotas (Sha et al., 1998). El dominio Sec14 se encuentra en 1550 proteínas anotadas en la base de datos del NCBI (*Nacional Center for Biotechnology Information*; <http://www.ncbi.nlm.nih.gov>). La diversidad de esta gran familia de proteínas que contienen el dominio Sec14, propicia que sus miembros participen en multitud de procesos celulares. Se han descrito múltiples roles para estas proteínas de transferencia (PITPs tipo Sec14,) inicialmente como proteínas de transferencia de fosfatidilinositol (PI) y fosfatidilcolina (PC) entre membranas, como sensores de la composición de la membrana, reguladoras de rutas de tráfico de membranas, siendo esencial en levadura; y en humanos, disfunciones de las proteínas Sec14, pueden dar lugar a enfermedades neurodegenerativas, ceguera y algunos cánceres (*Bankaitis et al., 1990; Sha et al., 1998; Cichowski y Jacks 2001; Fishman et al., 2004; Curwin y McMaster, 2008*).

### 1.1.7. Patelinas, familia multigénica

El gen *CRIO4* de remolacha, objeto de estudio en este trabajo, es ortólogo a los genes PATLs de *Arabidopsis thaliana*. Las PATLs o patelinas, son una familia de proteínas tipo Sec14 de *Arabidopsis thaliana* compuesta por seis miembros. Poseen una región N-terminal variable en longitud y composición, donde también pueden encontrarse los motivos *Coiled Coil*, un motivo de plegamiento y oligomerización de proteínas (Burkhard et al., 2001), secuencias PXXP de unión a dominios SH3 y secuencias poliprolina (Feng et al., 1994). Los dominios SH3 participan en interacciones proteína- proteína implicadas en dinámicas de

## Introducción

citoesqueleto, rutas de señalización tirosina quinasas y tráfico vesicular (Kay et al., 2000). El extremo C- terminal de las PATLs está muy conservado y es el que define a la familia por contener los dominios Sec14 y GOLD, que se encuentran en otras proteínas eucariotas implicadas en el tráfico de membranas. El dominio Sec14, tiene actividad en la unión y transferencia de fofolípidos (Peterman et al. 2004) y el dominio GOLD esta relacionado con la función de Golgi y en el trafico lipídico (Anantharaman y Aravind, 2002).

El análisis de las seis patelinas de levadura nos indica que además del dominio SEC14 y GOLD que es el distintivo de la proteína, en el N terminar pueden aparecer dominios PXXP, poliprolina o *Coiled coil*. No hay información en la literatura sobre la presencia de esta familia en diferentes plantas, así como si los diferentes dominios están conservados o se presenta un proceso de *domain shuffle*. En el presente trabajo trataremos de conseguir toda la información posible sobre esta familia en todas las bases de datos así como su representación en los genomas de plantas.





## **OBJETIVOS**



## **2. OBJETIVOS**

Los objetivos del trabajo propuesto consisten en el análisis bioinformático de la familia de proteínas PATL en plantas, a la cual pertenece BvCRIO4. Para investigar la conservación de esta familia y la dispersión de sus diferentes miembros en diferentes genomas de plantas.

Para ello, se realizará un primer estudio en bases de datos moleculares y una búsqueda de similitud con otras especies o proteínas relacionadas. A continuación, una búsqueda en bases de datos de motivos y dominios proteicos verificará los resultados obtenidos anteriormente y permitirá determinar la presencia o ausencia de esta familia en diferentes especies de plantas., permitiendo enmarcar la proteína en una/s familia/s proteica concreta. Por último, la estructura de dominios de CRIO4 en comparación con los resultados obtenidos en otras especies proporcionará información sobre la localización de los dominios catalíticos y motivos proteicos conservados.



## **MATERIALES Y MÉTODOS**



### **3. MATERIALES Y MÉTODOS**

#### **3.1. Bases de datos**

##### **3.1.1. UniProt**

El nombre de esta base de datos proviene del nombre en inglés *Universal Protein*. Es una de las bases de datos más importantes que existen sobre fichas proteicas, creada a partir de Swiss-Prot, TrEMBL y PIR. Para acceder a cada una de las fichas rellenamos el campo “*Query*” con el “*GeneName*” o nombre de la proteína de estudio y, pinchando en el botón “*Fields*” seleccionamos el campo “*Organism [OS]*” y ponemos el nombre del organismo que queramos en el campo “*Term*”. En cada entrada de UniProt aparece gran cantidad de información minuciosa sobre proteínas. Esta base de datos ha sido usada durante todo el trabajo. Dentro de toda esta información contenida en las fichas se pueden destacar varios apartados importantes como “*general annotation*” para la búsqueda de características moleculares y funciones biológicas y biotecnológicas; “*sequence annotation*” para la búsqueda de información de secuencia; “*Sequences*” para visualizar la secuencia aminoacídica de la proteína y descargarla en formato FASTA (aparece un botón con este nombre), que permite usarla en herramientas bioinformáticas; “*References*” para obtener bibliografía de ampliación; y “*cross-reference*”, que provee de referencias cruzadas hacia otras bases de datos como GeneID, Ensembl o GenBank. Las secuencias de cada proteína a estudiar han sido descargadas en formato FASTA y puestas en formato multiFASTA para facilitar su uso en los distintos programas bioinformáticos utilizados. Esta base de datos también puede usarse para obtener homólogos de nuestra proteína de estudio, utilizando la herramienta Blast, que compara las secuencias de aminoácidos en formato FASTA y muestra todos los posibles organismos con secuencias homólogas para esta proteína.

### **3.1.2- GeneBank**

Esta base de datos se encuentra en el sitio web del NCBI (*Nacional Center for Biotechnology Information*) y es la base de datos de secuencias genéticas del NIH (*Nacional Institute of Health*), que proporciona una completa información sobre secuencias nucleotídicas. También contiene estructuras 3D, secuencias aminoacídicas y muchas referencias cruzadas hacia UniProt. Además, permite visualizar información sobre la estructura génica del gen, con la región promotora, la región poliA, exones e intrones. La hemos usado gracias a las referencias cruzadas de UniProt para obtener las secuencias nucleotídicas de las proteínas estudiadas (CDSs), dentro del campo “*Features*” de la ficha, pinchando en el botón CDS de cada una de las fichas, permitiendo obtener el ORF del gen en formato FASTA.

### **3.1.3. Protein**

Una colección de secuencias de varias fuentes, incluyendo traducciones de regiones de codificación anotadas en GenBank, RefSeq y TPA, así como registros de SwissProt, PIR, PRF y PDB.

### **3.1.4. TAIR**

El recurso de información de Arabidopsis (TAIR) mantiene una base de datos de biología molecular y genética para la planta modelo *Arabidopsis thaliana*. Los datos disponibles de TAIR incluyen la secuencia completa del genoma junto con la estructura génica, la información sobre los productos génicos, la expresión génica, el ADN y las reservas de semillas, los mapas del genoma, los marcadores genéticos y físicos, las publicaciones y la información sobre la comunidad de investigación de Arabidopsis. Los datos de la función de los productos genéticos se actualizan cada semana a partir de la última literatura de investigación



publicada y de las presentaciones de datos de la comunidad. TAIR también proporciona enlaces a otros recursos de Arabidopsis.

### **3.1.5. Prosite y ProRule**

Estas bases de datos también proporcionan información sobre dominios y motivos, siendo muy completas. Se complementa con ProRule, una colección de reglas basadas en perfiles y patrones, que aumenta el poder discriminatorio de los perfiles y patrones, proporcionando información adicional sobre los aminoácidos funcionalmente y estructuralmente críticos.

## **3.2. Software bioinformático**

### **3.2.1. BLAST**

El algoritmo BLAST (*Basic Local Alignment Search Tool*) es una herramienta de búsqueda básica de alineamiento local entre secuencias que determina el porcentaje de similitud entre ellas. Se encuentra alojada en el NCBI y en UniProt. Usamos esta herramienta porque sus implementaciones son muy eficaces y permiten trabajar con resultados rápidos en las búsquedas de secuencias en los servidores gratuitos. En nuestro caso usamos el blastp, basada en una matriz de intercambio de aminoácidos para puntuar los alineamientos locales que se realizan entre la proteína de referencia introducida y todas las entradas presentes en la base de datos UniProt, introduciendo la secuencia aminoacídica de la proteína betagalactosidasa. Así, podemos buscar organismos homólogos que contengan una secuencia similar a nuestra proteína, basándonos en una medida de e-value (indica la probabilidad de que la similitud que se encuentre sea debida al azar), p-value significativos (entendiendo por significativos los alineamientos que arrojen un e-value menor o igual a 0,02) y % de similitud o identidad. Además, hemos usado también un algoritmo tipo BLAST llamado bl2seq para la búsqueda de similitud entre dos secuencias de entrada. Las secuencias presumiblemente

homólogas a nuestra proteína se recogen en formato FASTA, creando un archivo multiFASTA.

Hemos usado la matriz de intercambio BLOSUM 62 (*BLOcks of Amino Acid Substitution Matrix*) con porcentaje de identidad mínimo de 62%, que viene por defecto en el BLAST del NCBI. De esta forma, determinamos un grado de exigencia tal que todas las familias de proteínas que tengan un 62% de identidad se agrupan como un único representante en el análisis. Si usáramos un porcentaje de identidad mayor, obtendríamos proteínas con una relación evolutiva más estrecha.

Con este programa podemos realizar una rápida búsqueda de similitud contra una base de datos completa, permitiéndonos variar ciertos parámetros de búsqueda y obtener gran fiabilidad en los resultados obtenidos.

### **3.2.2. Clustal Omega**

Es una aplicación que realiza alineamientos múltiples de proteínas y secuencias nucleotídicas desde secuencias que se encuentran en los archivos multiFASTA creados anteriormente. Se basa en la utilización del algoritmo Clustal de alineamientos progresivos, que agrupa secuencias y realiza alineamientos por pares para generar matrices de distancia en los que se puede estudiar el % de identidad entre cada par de secuencias. A partir de éstas, se elaboran árboles filogenéticos que agrupan los grupos creados según su relación evolutiva.

En nuestro trabajo, damos al programa como información de entrada las secuencias aminoacídicas y nucleotídicas en formato multiFASTA, para obtener verificación de que los homólogos de *BvCRIO4* obtenidos con BLAST son correctos y para determinar la presencia de dominios en nuestra familia proteica de estudio. Hemos tenido que realizar un alineamiento general para luego ir eliminando candidatos no válidos, a fin de encontrar un buen alineamiento múltiple.

### **3.2.3. TreeView**

Éste es un programa que permite abrir los archivos con la extensión .dnd para el estudio de la filogenia de la familia proteica. Nos permite visualizar y comparar los árboles filogenéticos obtenidos desde los alineamientos de las secuencias de aminoácidos y nucleótidos. Permite ver el árbol de tres formas distintas: árbol sin raíz, cladograma o filograma. Además permite enraizar el árbol y ordenar las secuencias, ya que podemos definir un grupo externo, que se encuentre más alejado evolutivamente del resto de proteínas.

### **3.2.4. Coils**

Es un programa que compara una secuencia con una base de datos de bobinas en espiral de dos hilos paralelos conocidos y obtiene una puntuación de similitud. Al comparar esta puntuación con la distribución de las puntuaciones en las proteínas globulares y en espiral, el programa calcula la probabilidad de que la secuencia adopte una conformación de la bobina en espiral.



## **RESULTADOS**



## 4. RESULTADOS

### 4.1. Elección del gen.

El interés de nuestro laboratorio por BvCRIO4 surge a raíz de ser identificado en un escrutinio de genes de remolacha capaces de conferir tolerancia a frío al ser sobreexpresados en levadura (Izquierdo AC, 2016). El frío es uno de los estreses abióticos más importantes, afectando severamente al crecimiento y desarrollo de las plantas y limitando su distribución geográfica. En concreto, se sabe que la sobreexpresión de CRIO4 confiere a las células de levadura la capacidad de crecer a temperaturas bajas y tolerancia lumínica (inhibidor de la glicosilación de proteínas). Puesto secuencia de CRIO4 se han confirmado diversos motivos y dominios conservados en la proteína que parecen dirigir su función hacia el tráfico vesicular, y revela múltiples potenciales sitios de fosforilación en CRIO4 sugiriendo una posible función de señalización. El dominio Sec14 es un dominio de unión y transferencia de fosfolípidos, y los resultados obtenidos indican que CRIO4 establece una unión con fosfatidilinositol y fosfatidiletanolamina. En un estudio de localización subcelular se demostró que CRIO4 está asociada a la membrana plasmática y en membranas intracelulares, formando parte de vesículas, lo que parece confirmar el papel de CRIO4 en el tráfico intracelular, función que compartirá con sus ortólogos. El patrón de expresión de *CRIO4* en remolacha mostró que *CRIO4* se expresaba en los tejidos analizados, concentrándose su mayor expresión en hoja y raíz. También se demostró que la expresión de CRIO4 aumenta tras tiempos cortos de exposición a estrés por frío y estrés por calor. También se han realizado estudios fenotípicos de diferentes líneas mutantes en genes PATLs de *Arabidopsis thaliana*, obteniéndose una gran variabilidad en los resultados que podría ser la redundancia funcional entre los miembros de esta familia multigénica. Estudios fenotípicos de las líneas transgénicas de *de Arabidopsis thaliana* que sobreexpresan el gen CRIO4 muestran una mejoría en su crecimiento y desarrollo a bajas temperaturas, aumentando por tanto su tolerancia a estrés por frío.

La gran variabilidad de esta familia en *Arabidopsis* nos plantea la pregunta de si esta variabilidad se da en otras plantas o es una particularidad de *Arabidopsis*.

#### 4.2. Características generales del gen *BVCRIO4*

El ADNc correspondiente a *CRIO4* tiene una longitud de 2150 nucleótidos, contenido en un marco abierto de lectura (*Open Reading frame*, ORF) de 1751 nucleótidos (Figura 1). El codón de inicio de la traducción se encuentra en la posición 70, y el codón de parada en la posición 1821.

*CRIO4* codifica una proteína tipo Sec14 que conserva la misma estructura de dominios que las PATLs (*Patellins*) de *Arabidopsis thaliana*. La proteína CRIO4 está compuesta por 583 aminoácidos y su peso molecular estimado es de 65,33 kDa (Figura 2).

La proteína CRIO4, presenta una estructura multidominio muy semejante a las PATLs de *Arabidopsis thaliana*, con los dominios SEC14 Y GOLD en el extremo C-terminal y motivos prolina y *Coiled coil* en su extremo N-terminal.(Figura 3 y 4) El conjunto de motivos y dominios que componen la estructura de la proteína CRIO4 señalan hacia una función en el tráfico vesicular, y el elevado numero de potenciales sitios de fosforilación indican un posible papel en señalización.

En plantas, la localización subcelular de CRIO4 está asociada preferentemente a la membrana plasmática y a membranas intracelulares, aunque también puede establecer un patrón reticular.

Estudios fenotípicos de sobreexpresión del gen CRIO4 en líneas transgénicas de *Arabidopsis thaliana*, demuestran que la sobreexpresión de CRIO4 parece mejorar el crecimiento de las plantas a temperaturas bajas, aumentando así su tolerancia a estrés por frío. (Izquierdo. AC, 2016).



## Resultados

<b>BvCRI04</b>						
TNNCCGGGC	TGCAGGAATT	CGGCACGAGC	TCATTTCTCT	ACATCAAAAA	CACAACAAAG	AGATCACCCA
TGGCGGAAGA	AACCCATAAG	CCAGAATCAA	CGGTGGCTGA	AGTGGTGGTT	CCAGTAGCCG	AGAAACCAGCT
GAGAAGCCAG	CTGAGAAGGC	AGTTCTACCA	CCTGAAGCTG	AGAAACTAGC	TGCAGCTGAA	TCAGCTGAAG
CCGAGAAGCC	AGCTGATTCA	GCCGAGGCTA	AGATAGCTCA	ACAAGTCTCA	TTCAAAGAGG	AGACTAATGT
TGCAAGTGAG	CTACCTGAGC	TACATAGAAA	GGCTCTCGAG	GACTTGAAGA	AACTTATTCA	AGAAGCCCTC
GAGAAGCACG	AGTTCTCTTC	TCCTCCTCCT	CCGCCTCCGC	CTGCTCCAGC	TAAAGTTGAG	GAGAAGGCCG
AAGAGAAGAA	AGAGGAACAA	CCTCCATCCA	CCACCTCCAC	CACCACCACC	ACCACCACCG	CGGTTTCAGA
TGAGGTTGCT	GTTGCTCCTC	CATCCGAAGA	GGCCCCGAAA	ACTGACGAGG	CCTCTCCGAA	AGTGGAGGAG
GAGCCTGCAA	AAATAGTTGA	GCAACCACCT	ACAACACCGG	CAGAAGAACC	TGAACCAGCA	AAAACACCTG
AGGTTGTTGT	TGCTGAAGAG	GAGAAAACCTG	GTGAGGATAT	TAAAGAAACT	ATAGTAGTCG	AGGTTGCGAC
AACTACAGCA	GCACCAGTAC	TAACAGAACC	AGAATCTGTT	GAGGAGACAC	TGAAGTTGTA	GTGGAAGAAT
CACCAAAGGA	GCCAGAAGAA	GTATCAATAT	GGGGAATTCC	ACTTCTTGCT	GATGAAAGAA	GTGATGTAAT
TCTATTGAAA	TTCTTAAGAG	CAAGAGATTA	TAGAGTGAAA	GATGCTTTCA	CTATGATTAG	AAATACTGCT
CGTTGGAGAA	AAGAATTTGA	GGTTGATTCT	TTACTTGATG	AAGATCTTGG	AAATGATTAT	GAGAAAGTTG
TTTTTACACA	TGGAGTTGAT	AAACAAGGTC	GTCCTGTTTG	TTATAATGTG	TTTGGAGAGT	TTCAAATAAA
GGAACTTTAT	CAGAATACTT	TCTCTGATGC	AGAAAAAAGG	AAAAAGTTCT	TGAGATGGTT	GATTCAATTC
CTTGAAAAAA	CTATTAGAAC	TCTTGATTTT	AGTCCTGAAG	GAATTAATTC	TTTTGTTCTT	GTTAATGATT
TGAAGAATTC	TCCTGGGTAT	GGTAAGAGAG	ATCTTTACAA	AGTTATTGAC	AAGTTTCTTG	AGATTCTCCA
GGATAATTAC	CCAGAATTTG	CTGCTAAACA	GTTGTGCATC	AATGTTTCAT	GGTGGTACTT	GGCATAACAAC
TGGATCTATT	TGACTGTATT	TACACGCAGG	AGCAAGAGCA	AGTTTGTGTT	TGCAAGCCCA	TCTAAAACCTG
CTGAGACCCT	TTTCAAGTAC	ATAGCTCCTG	AGCAGGTGCC	TGTTCAATTT	GGTGGGCACA	GTAAGTTTGG
TGAGCATGAG	TTTTCCCCTG	CTGATACTGT	CACTGAAGTT	ACCATCAAGC	CTGGCAGTAA	ACATCCCCTT
GAATTTGCAT	TTTCTGAGGA	AACTGAGTTG	GTATGGGAGC	TTAGAGTAAT	TGGATGGGAT	GTAAGTTACG
GAGCCGAATT	CTTACCAACA	AAGGAAGGAG	GATACACTCT	AAACATTGCC	AAGCCAAAGA	AGGTAACACC
AGCTGATGAA	CCAGTCATTT	GTGACACTTT	CAAGGTTACT	GAGCCAGGAA	AGGTTGTAAT	TACAATTGAC
AACCAAAGTT	CAAAGAAGAA	GAAGCTCCTT	TACAGGTCTA	AGGTCAAGAC	TTCTCAGTGA	TTTTATCATA
ATCAATTTCA	CAACATTTCA	ATTCATTCAA	AAAATCATAG	TGACTCTTAA	TAAGTTACTT	TTTTATTTTT
ATTTTATAAG	GAGTTTTGGG	AATTATTTAT	TTATTATTTG	TTTTATTTTT	GTATTTTGTT	GGATTTTTTT
TTTCTTTTTT	TGGGGTGTGT	CCTGGTCCTT	GTGTTTCTCA	AGTTTGGAGG	ACATATACAG	AGGAAGAAAG
GGGTGCAAAA	ACTGTGTGAA	TTGATACTCT	TTTTGTCTGT	ATATTCTTAT	TCTATGAAA	TTTTAATTAT
GGTTGTTGAG	TTTGTGCTAA	AATTATAAAT	TTTCATTCT			

*Figura 1: Secuencia de nucleótidos del gen BvCRI04.*

## Resultados

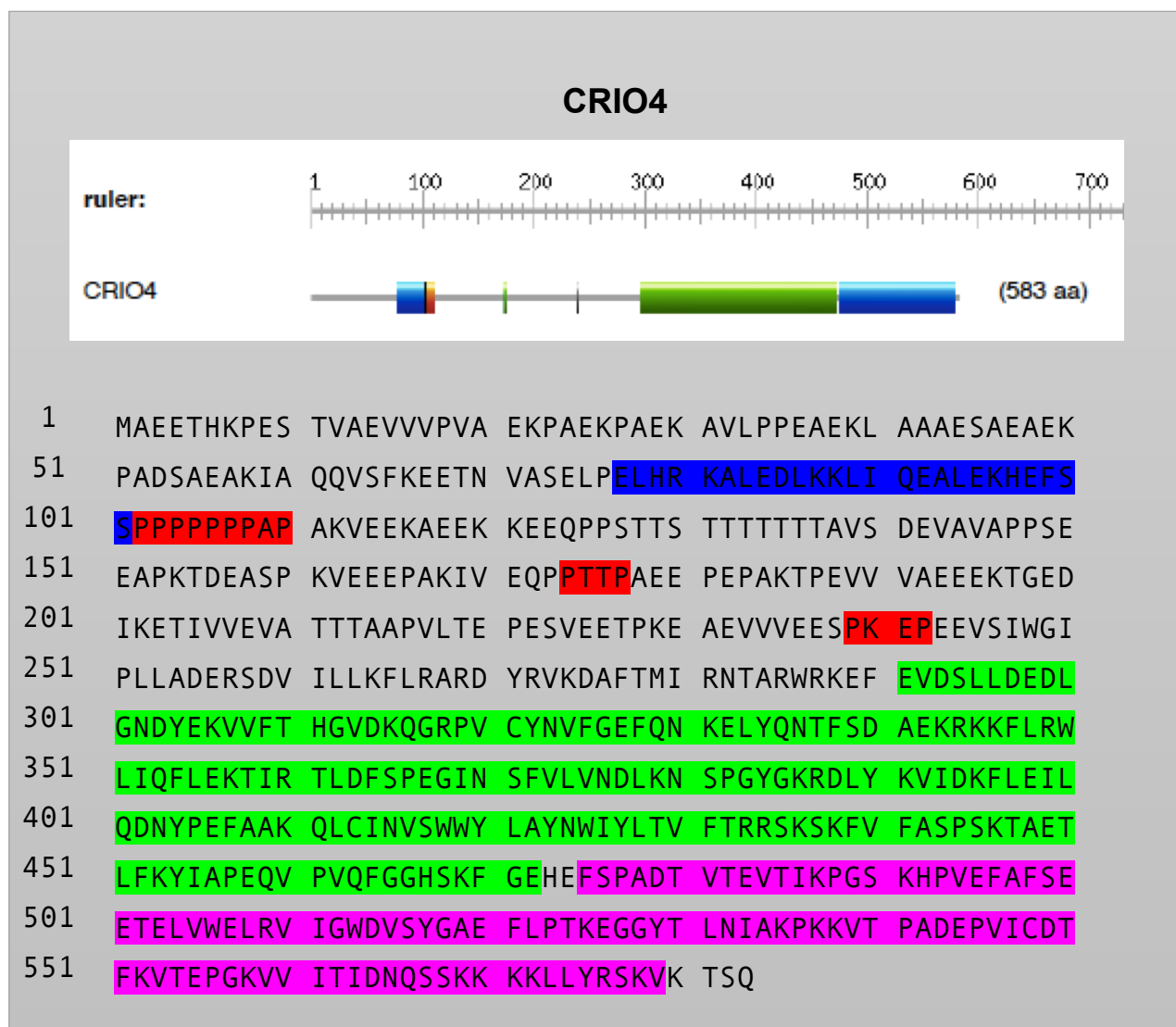


Figura 2: Estructura de dominios de la proteína CRI04. Subrayado en azul el Coiled coil, en rojo los motivos PXXP, el dominio Sec14 en verde y el dominio GOLD en violeta.

## Resultados

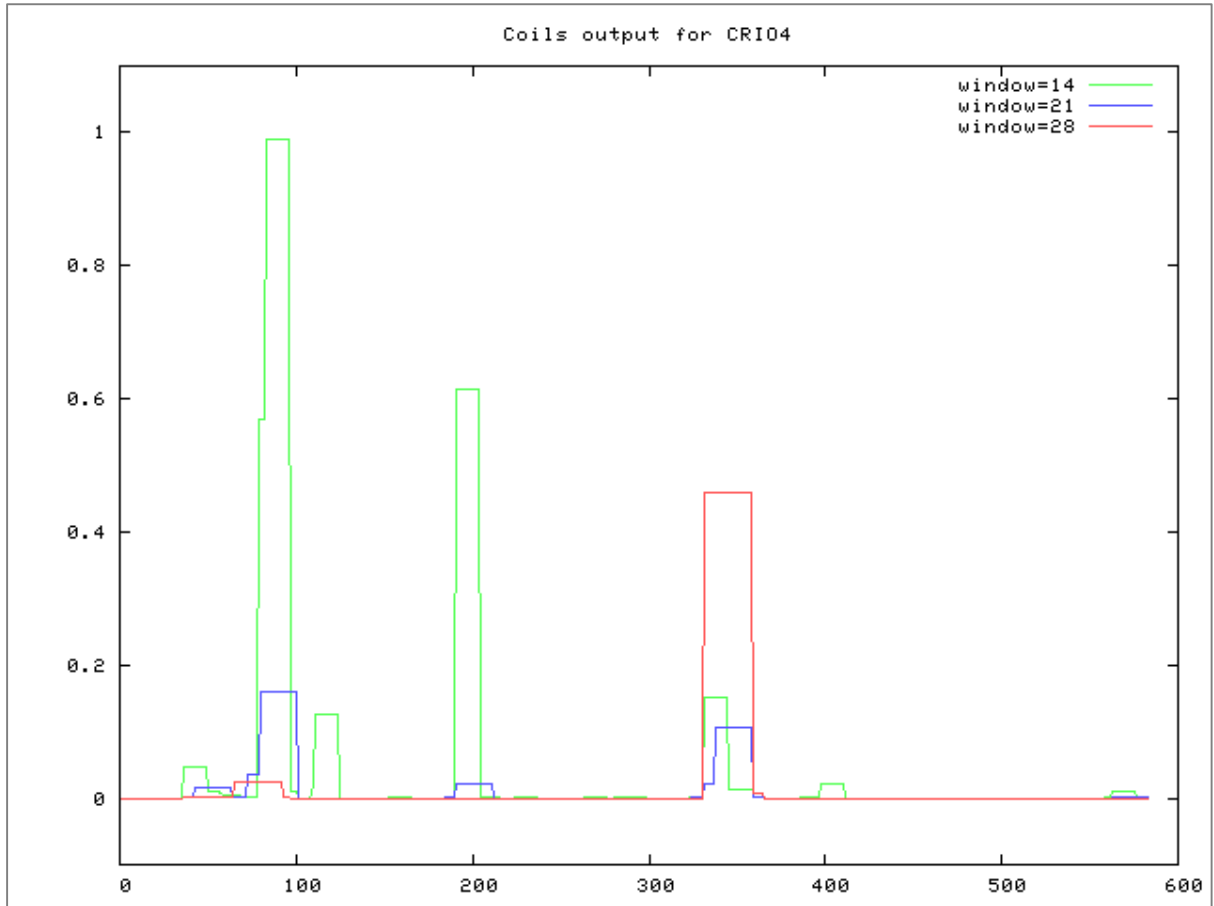


Figura 3: Predicción de Coiled Coil en la proteína CRIO4

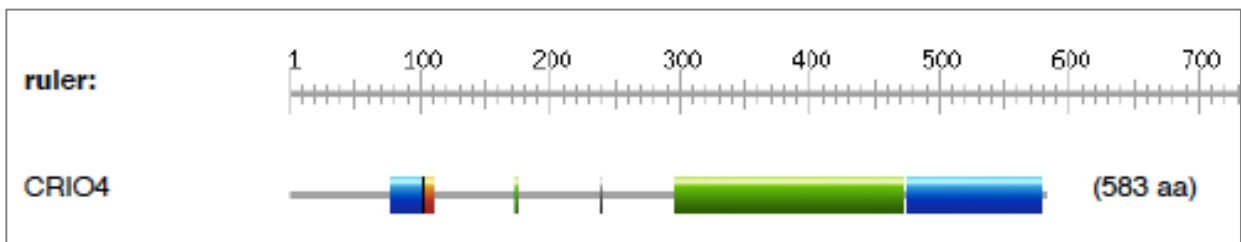


Figura 4: Diagrama de la estructura de dominios de CRIO4.

### 4.3. Búsqueda de secuencias homólogas

Con la intención de identificar los genes ortólogos a *CRIO4*, se realizó una búsqueda exhaustiva en la búsqueda de patelinas en otras especies en las bases de datos de GeneBank, Protein, UniProt y TAIR. Con todas las secuencias recopiladas, se realizó un análisis de similitud de secuencias de aminoácidos con el programa BLASTp del NCBI.

De este modo se, el programa BLASTp permitió identificar como ortólogos de *CRIO4* en *Arabidopsis thaliana*, a la familia de las patelinas cuyos genes aparecen en el siguiente orden de homología: *PATL3* (At1g72160), *PATL2* (At1g22530), *PATL5* (At4g09160), *PATL1* (At1g72150), *PATL4* (At1g30690) y *PATL6* (At3g51670). Estos mismos datos fueron corroborados mediante un alineamiento con la herramienta WU-BLAST de la base de datos de *Arabidopsis thaliana* (TAIR).

#### 4.3.1. Estudio filogenético de *CRIO4*

La herramienta Clustal Omega permite obtener árboles filogenéticos basados en homología de secuencia. Se sometieron a análisis con Clustal Omega el conjunto de las secuencias de todas las PATLs de *Arabidopsis thaliana* y de *CRIO4* y el árbol filogenético generado se muestra en la Figura 5.

En cuanto a homología, *CRIO4* parece ser más cercano a *PATL3*. Sin embargo, el análisis filogenético enmarca a *CRIO4* más próximo a la rama de *PATL1* y *PATL2*.

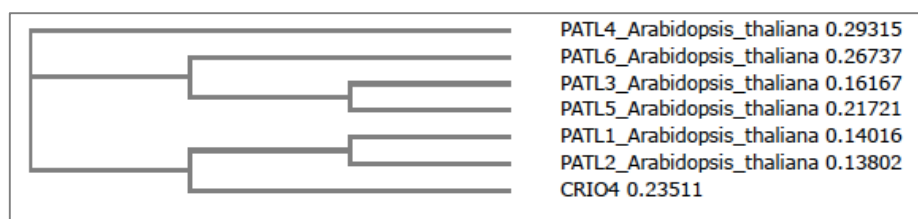


Figura 5: Cladograma de *CRIO4* y la familia de *PATLs* de *Arabidopsis thaliana*, a partir de secuencias de aminoácidos.

## Resultados

Del mismo modo, se realizó un segundo árbol filogenético utilizando todas las patelinas encontradas con similitud a *CRIO4*, obteniendo el cladograma que se presenta en el Anexo I dada su extensión.

Analizando las especies más cercanas a nuestra proteína, obtenemos el siguiente árbol filogenético (Figura 6), donde aparecen como era de esperar, la PATL3 de *Beta vulgaris* en la misma rama que *CRIO4*. En la rama derivada más cercana, aparecen *Erythranthe guttata* y *Nicotiana tabacum*, ambas descritas con resistencia a estrés abióticos.

*Erythranthe guttata* es una dicotiledónea de flores silvestres, cultivada en Australia y Norte América, incluyendo hábitats como el desierto, alta montaña, e incluso se han llegado a encontrar en minas de cobre, por lo que les confiere cierta predisposición al frío y a estreses abióticos.

*Nicotiana tabacum*, descrita con resistencia a estrés por bajas temperaturas y a salinidad.

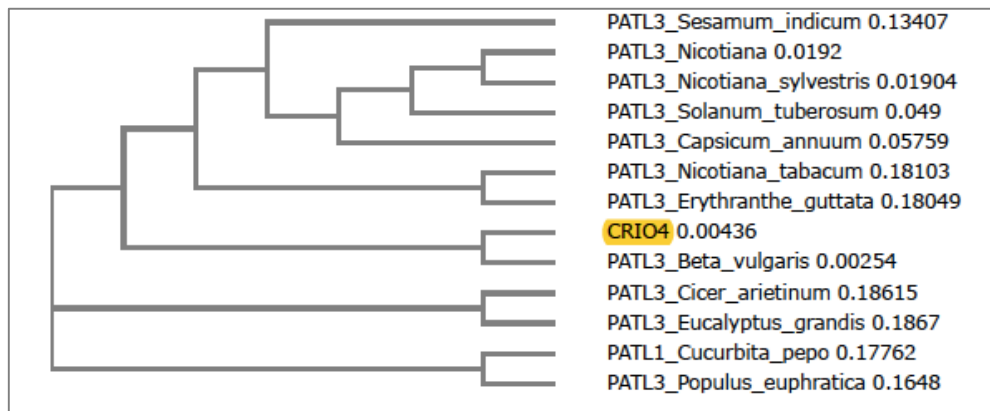


Figura 6: Cladograma de *CRIO4* y las PATLs de las especies más cercanas, a partir de secuencias de aminoácidos.

Por otro lado, se realizó un análisis de presencia de motivos poliprolina conservados, dando lugar a una nueva relación de patelinas con esta característica y el árbol correspondiente, en las que destacaríamos la cercanía de la PATL2 de *Arabidopsis thaliana* y las patelinas 3 de varias especies de interés agronómico como: arroz, judía, maíz, cacao, vid, tabaco, entre otros (Figura 7).

## Resultados

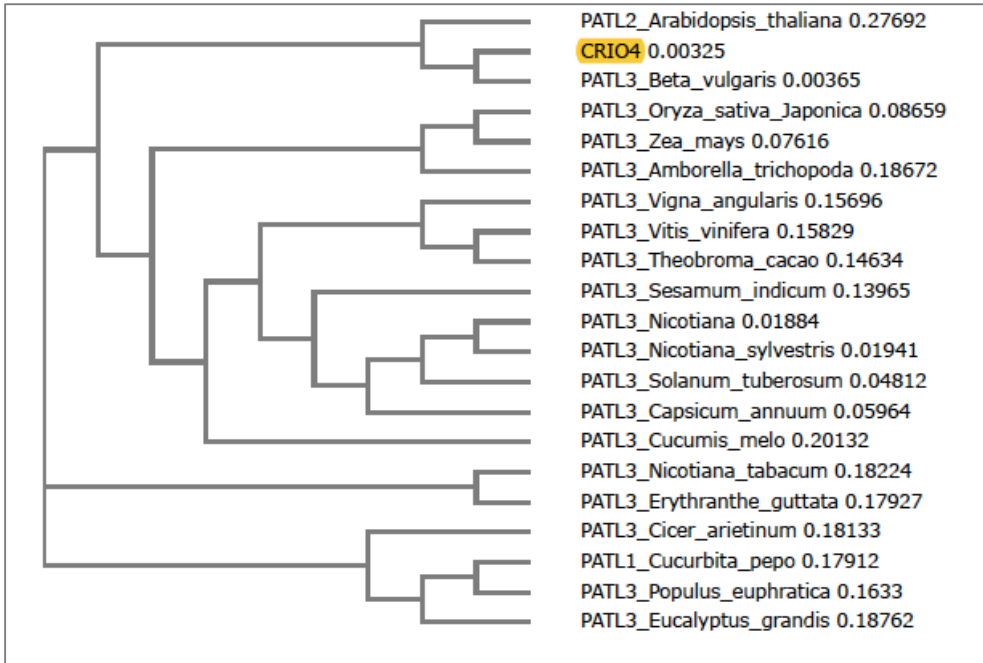


Figura 7: Cladograma de CRIO4 y las PATLs de las especies que presentan motivos ricos en prolina

### 4.3.2. Patelinas conservadas en otras especies

Las patelinas más conservadas son las PATL3 y PATL6, conservadas en el 80% de las especies que presentan homología con BvCRIO4 (Figura 8). La peculiaridad de la estructura de dominios que presentan estas dos patelinas es la más sencilla, ya que solamente conservan los dominios Sec14 y Gold en la parte C-terminal y una parte N-terminal variable en longitud y composición.

La remolacha (*Beta vulgaris*) conserva las PATL3, PATL4 y PATL6, además de BvCRIO4.

## Resultados

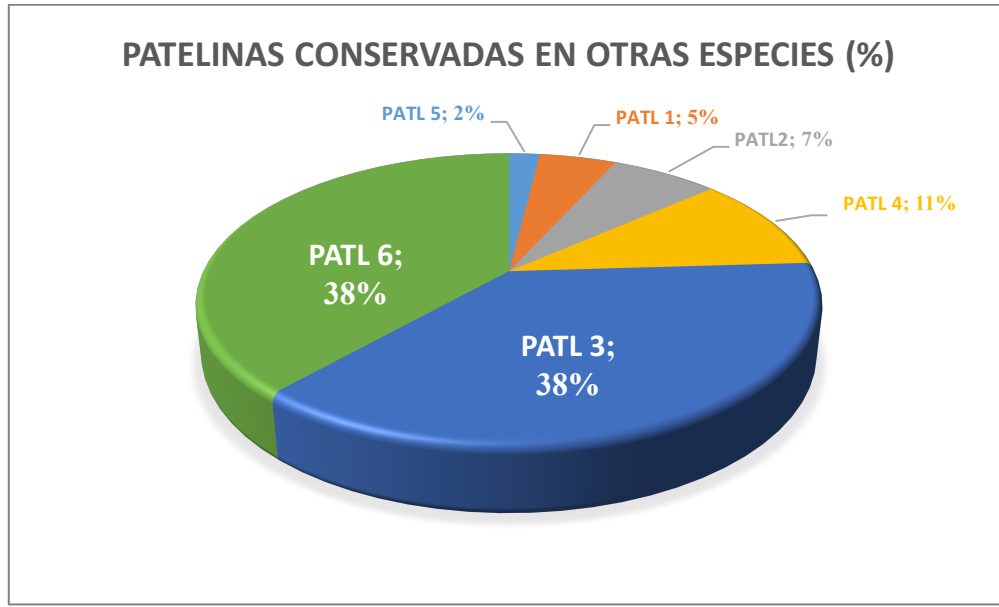


Figura 8: Relación de patelinas conservadas en otras especies (%)

La mayoría de especies conservan una o dos patelinas, muy pocas son las que conservan cuatro o más patelinas, como *Camelina sativa*, *Theobroma cacao* y *Arabidopsis thaliana* que llega a conservar las seis, aunque se predice que se esté dando solapamiento de funciones entre los miembros de esta familia (Peterman et al., 2006) (Figura 9).

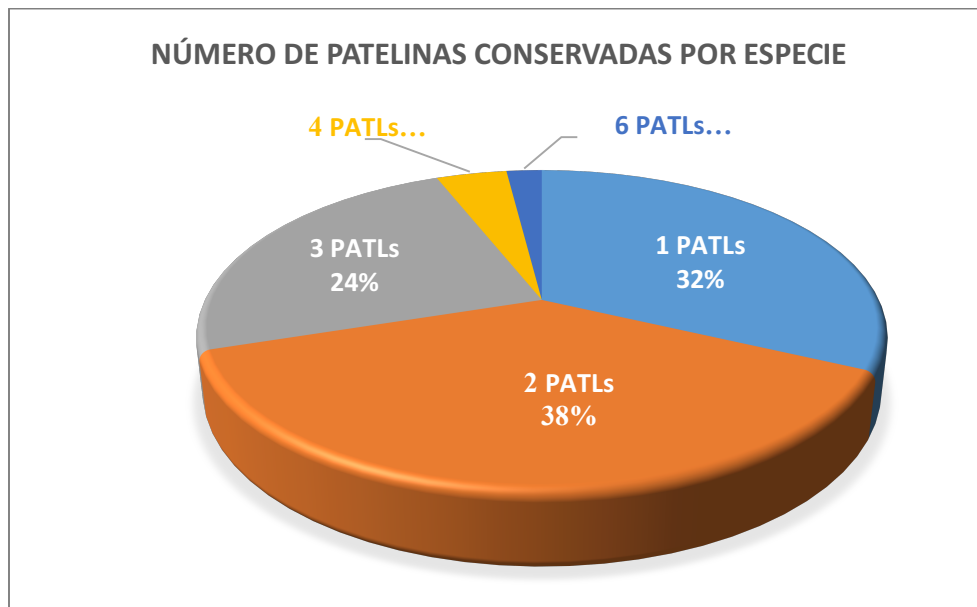
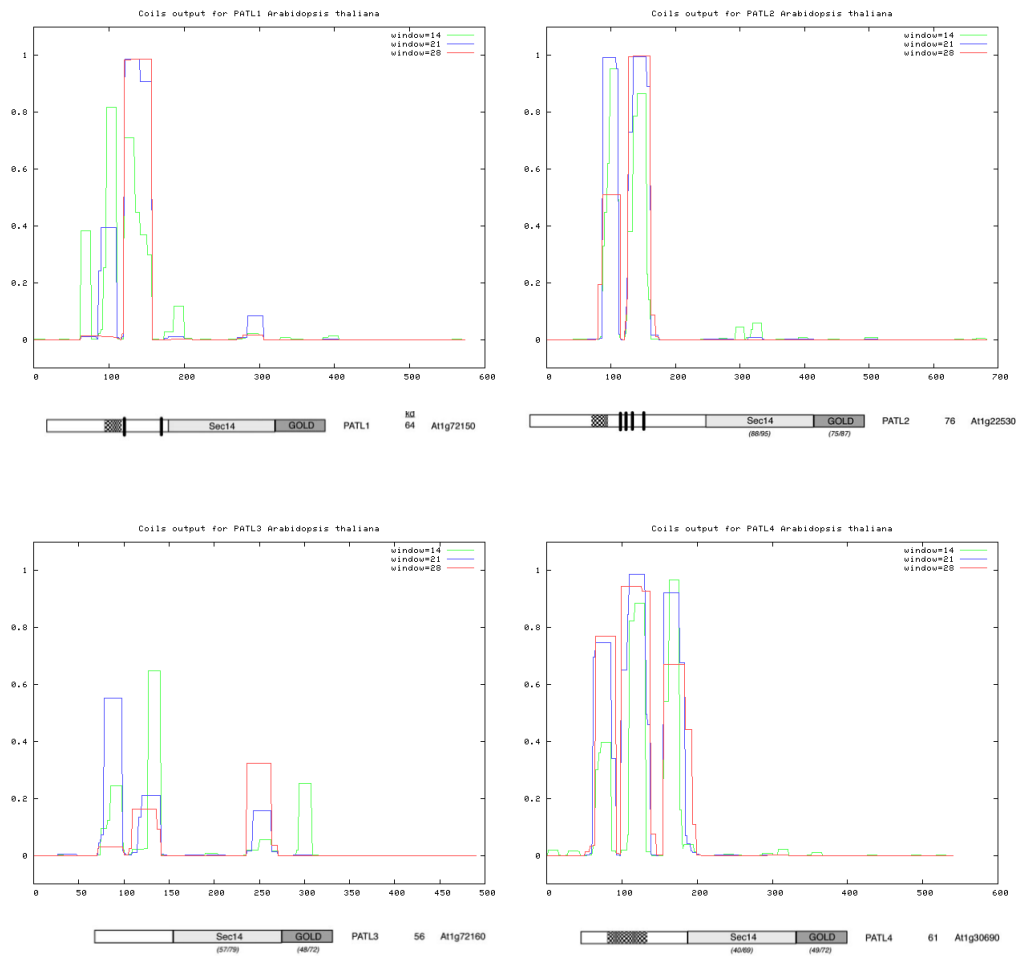


Figura 9: Relación del número de patelinas conservadas por especie (%)

#### 4.4. Clasificación patelinas en función de la estructura de dominios

Podemos clasificar las patelinas en función de su semejanza a las estructuras de las patelinas de *Arabidopsis thaliana*. Para ello, en primer lugar hemos delimitado mediante la herramienta bioinformática *ExPASy-Prosite*, la presencia de motivos ricos en prolina y mediante el software *Coils* predecir las regiones *Coiled coil* en las patelinas (Figura 10 y Anexo II y III) para así poder delimitar en los diagramas de estructuras de dominios de las patelinas los motivos de interés para su clasificación.







## Resultados

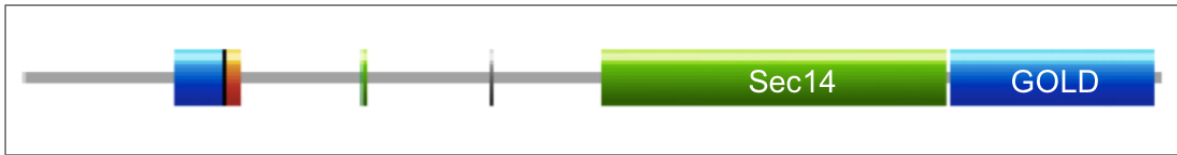


Figura 12: Representación gráfica de la proteína CRI04



Figura 13: Representación gráfica de PATL3 de Beta vulgaris con un 99% de homología con CRI04

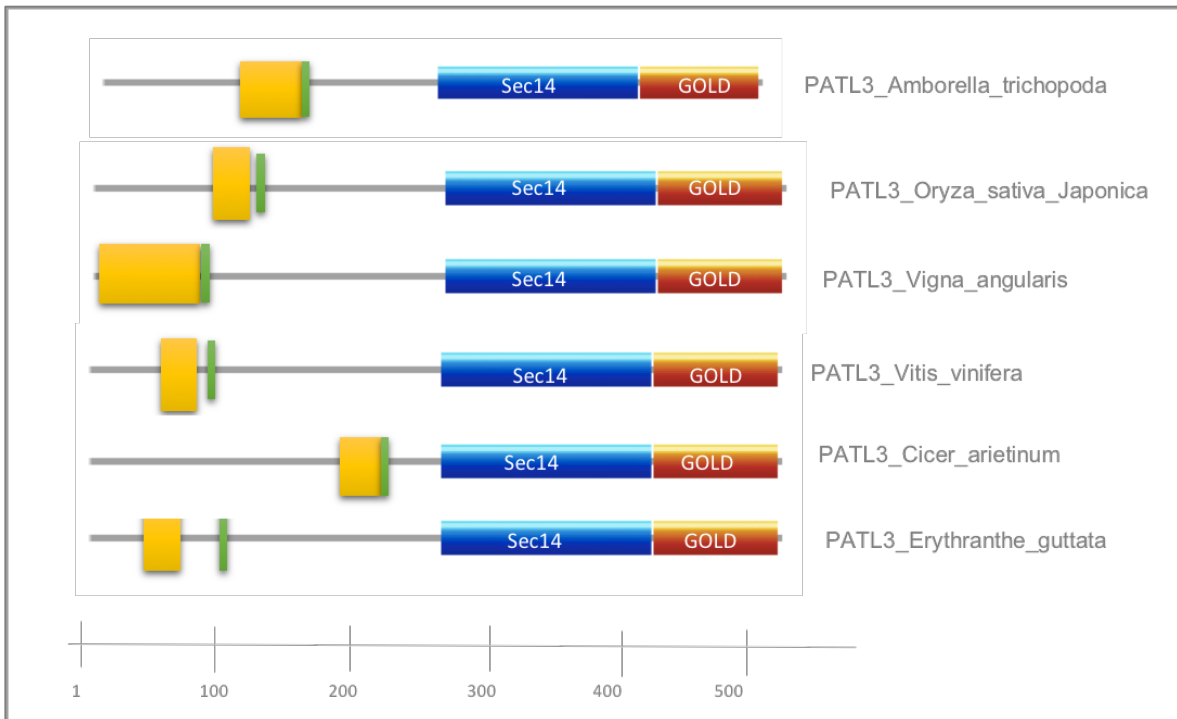


Figura 14: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRI04 con a la estructura de PATL1 y PAT2 de Arabidopsis thaliana

En un segundo grupo incluiremos aquellas patelinas que conservan los dominio Sec14 y GOLD y secuencias PXXP, como la PATL5 de la planta modelo, *Arabidopsis thaliana* (Figura 15), en las que encontraremos especies de

## Resultados

importancia agronómica como maíz (*Zea mays*) y el cacao (*Theobroma cacao*) en la Figura 16.

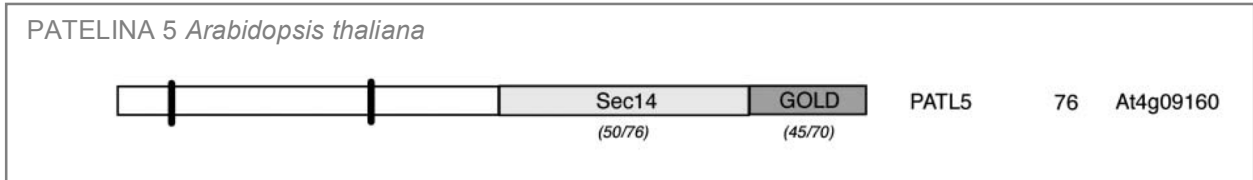


Figura 15: Representación gráfica de la estructura de dominios de la PATL5 de *Arabidopsis thaliana*

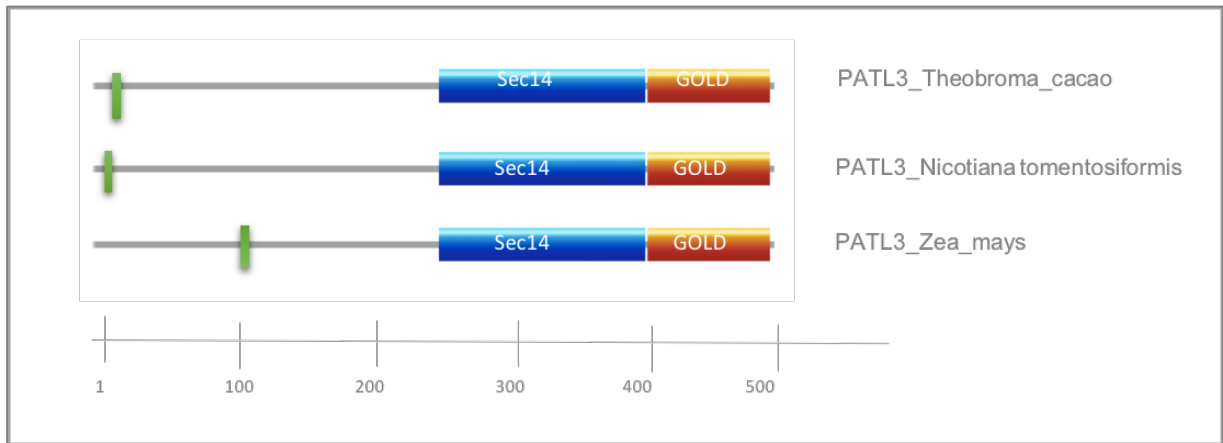


Figura 16: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRIO4 con a la estructura de PATL5, con motivos PXXP, dominio Sec14 y dominio GOLD.

En un tercer grupo incluiremos aquellas patelinas que si que conservan el motivo *Coiled coil* pero no mantienen los motivos ricos en prolina (Figura 17). Aquí encontraremos una serie de patelinas de especies con elevado interés agronómico como es el caso del tomate, melón, los cítricos, que siendo cultivos importantes en las líneas de mejora genética (Figura 18 y 19).

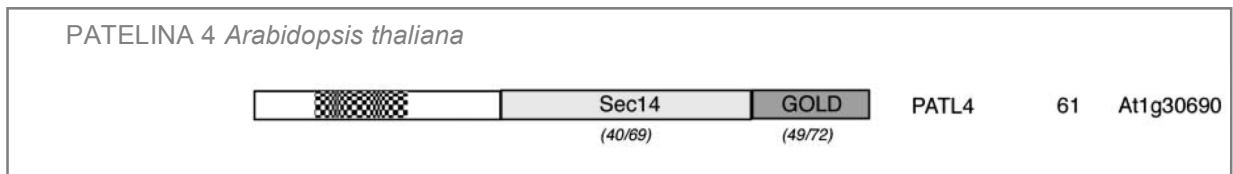


Figura 17: Representación gráfica de la estructura de dominios de la PATL4 de *Arabidopsis thaliana*.

## Resultados



Figura 18: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRI04 con la estructura de PATL4 con un dominio Coiled coil, dominio Sec14 y dominio GOLD conservados.

PATL1_Camelina_sativa	PATL3_Medicago_truncatula	PATL4_Theobroma_cacao
PATL2_Capsicum_annuum	PATL3_Populus_euphratica	PATL4_Vitis_vinifera
PATL2_Camelina_sativa	PATL3_Prunus_mume	PATL4_Solanum_lycopersicum
PATL2_Nicotiana_sylvestris	PATL3_Ricinus_communis	
PATL2_Raphanus_sativus	PAT3_Vigna_radiata	
PATL3_Brassica_oleracea	PATL4_Citrus_sinensis	
PATL3_Camelina_sativa	PATL4_Cucumis_melo	
PATL3_Citrus_sinensis	PATL4_Glycine_max	
PATL3_Cucumis_sativus	PATL4_Musa_acuminata	
PATL3_Fragaria_vesca	PATL4_Nicotiana_tabacum	
PATL3_Glycine_max	PATL4_Oryza_sativa_Japonica	
PATL3_Jatropha_curcas	PATL4_Solanum_lycopersicum	

Figura 19: Patelinas conservadas en otras especies que rigen la estructura de la PATL4 de *Arabidopsis thaliana*

Y por último, un cuarto grupo en el que incluiremos aquellas patelinas que solamente conservan los dos dominios que definen a la familia, el Sec14 y el GOLD en la parte C-terminal (Figura 20). Este es el grupo más numeroso en cuanto especies que conservan esta estructura de dominios, encontraremos especie como el tabaco, maíz, cacao, entre otras de interés agronómico (Figura 21 y 22).

## Resultados

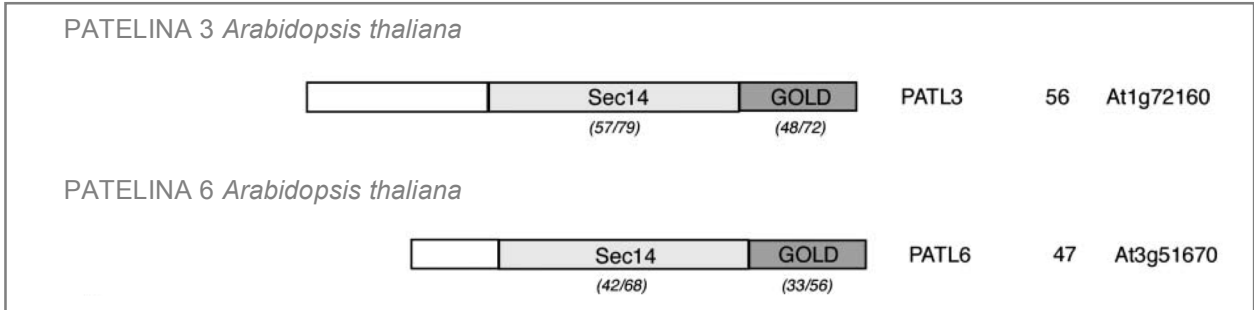


Figura 20: Representación gráfica de la estructura de dominios de las PATL3 y PATL4 de *Arabidopsis thaliana*.

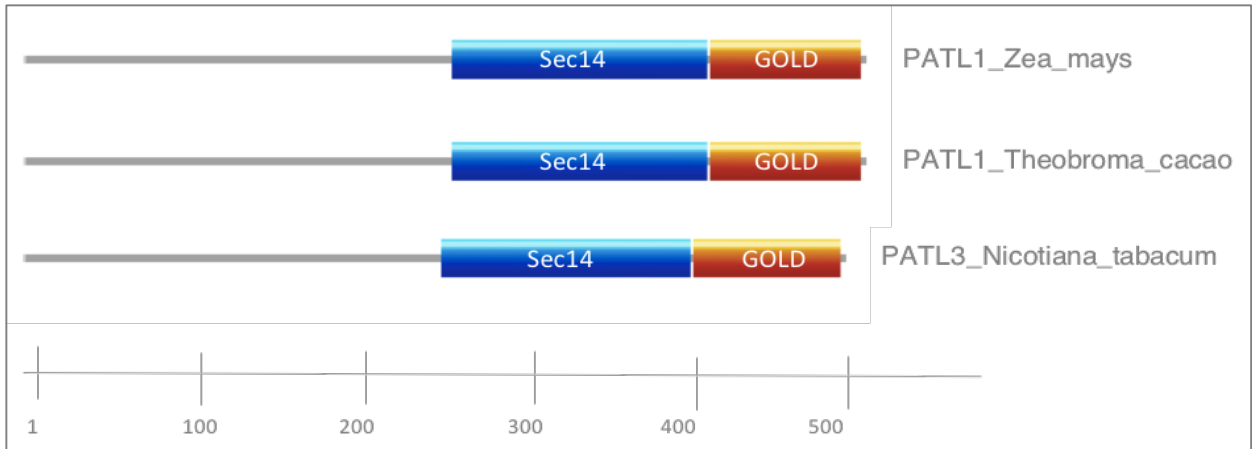


Figura 21: Representación gráfica de la estructura de dominios de las patelinas homólogas a CRI04 con la estructura dominio Sec14 y dominio GOLD

PATL2_Brassica_rapa	PATL6_Brassica_oleracea	PATL6_Nicotiana_attenuata
PATL2_Brassica_oleracea	PATL6_Brassica_rapa	PATL6_Nicotiana_sylvestris
PATL3_Arachis_ipaensis	PATL6_Brassica_napus	PATL6_Oryza_brachyantha
PATL3_Arachis_duranensis	PATL6_Camelina_sativa	PATL6_Perkinsus_marinus
PATL3_Brassica_napus	PATL6_Capsicum_annuum	PATL6_Populus_euphratica
PATL3_Capsicum_annuum	PATL6_Citrus_sinensis	PATL6_Prunus_mume
PATL3_Gossypium_arboreum	PATL6_Cucumis_sativus	PATL6_Raphanus_sativus
PATL3_Gossypium_raimondii	PATL6_Eucalyptus_grandis	PATL6_Sesamum_indicum PATL6_Ricinus_communis
PATL3_Juglans_regia	PATL6_Fragaria_vesca	PATL6_Setaria_italica
PATL3_Musa_acuminata	PATL6_Glycine_max	PATL6_Solanum_lycopersicum
PATL3_Nicotiana_sylvestris	PATL6_Gossypium_arboreum	PATL6_Solanum_pennellii
PATL3_Sesamum_indicum	PATL6_Gossypium_raimondii	PATL6_Solanum_tuberosum PATL6_Tarenaya_hassleriana
PATL3_Solanum_tuberosum	PATL6_Ipomoea_nil	
PATL3_Tarenaya_hassleriana	PATL6_Jatropha_curcas	
PATL5_Zea_mays	PATL6_Musa_acuminata	
PATL6_Brachypodium_distachyon	PATL6_Nicotiana_tabacum	

Figura 22: Patelinas conservadas en otras especies que solamente han conservado los dominios Sec14 y GOLD.



## **DISCUSIÓN**





## 5. DISCUSIÓN

El interés de nuestro laboratorio en el estudio de las patelinas surge a partir de haber sido encontradas en un escrutinio para buscar genes de remolacha capaces de conferir tolerancia a frío al ser sobreexpresados en levadura. *BvCRIO4* es un gen de la familia de las patelinas. En *Arabidopsis* esta familia está caracterizada a nivel de secuencia y se conoce que tiene seis miembros, a diferencia de remolacha donde solo existen 3 miembros. Estos miembros nos sirven como arquetipos para diferenciar y clasificar a los miembros de la familia en función de su estructura de dominios. Para empezar PATL3 y PATL6 solo tienen los dominios SEC14 y el dominio GOLD que son los que determinan a la familia. Estas proteínas son las más representadas en todos los genomas. No hemos encontrado ningún genoma de planta en el que no estén representadas las patelinas lo que apunta a su papel esencial. Hasta ahora se ha caracterizado su papel en la formación de la patelina (Peterman, 2006), y por ser capaces de interaccionar con proteínas de movimiento víricas (Peiro, 2014) así como su capacidad de conferir tolerancia a frío (Izquierdo AC, 2016). No obstante su función está todavía inexplorada en su mayor parte.

El hecho que 3 y 6 sean las más frecuentes, y a su vez las que tienen la estructura más sencilla indica que posiblemente evolutivamente son anteriores y que el resto de patelinas más complejas aparecieron por la incorporación progresiva de dominios. Patelina 4, la que incorpora un dominio *Coiled coil* es la siguiente más representada, con un 11% de presencia. Esto apunta a que posiblemente este dominio fue el primero en incorporarse, la siguiente es Patelina 2, que incorpora el mencionado dominio más los dominios PXXP o poliprolina. Curiosamente las menos representadas en los genomas de plantas, según nuestros datos, serían las del tipo Patelina 5, que son las que no tienen dominio *Coiled coil*, pero si que tienen dominio PXXP, posiblemente esto sugiera que este dominio puede perderse en etapas evolutivas posteriores a partir de Patelina 2 o Patelina 1, aunque sin un análisis filogenético más profundo no podemos

descartar que se haya originado independientemente a partir de Patelinas del tipo 3 o 6 que no tienen ningún dominio específico en el N terminal.

Otro resultado de este estudio es que hemos identificado plantas con 1, 2 o 3 patelinas, así como plantas con 4 o 6, pero no plantas con 5 patelinas. Esto nos sugiere que en algún momento pudieron producirse procesos de duplicación génica que aumentaron el número de copias de 1 a 2 o de 2 a 3, pero el hecho que no hayan plantas de 5 copias indica que en los organismos donde aparecen 4 o 6 copias se debe a que duplicaciones del genoma de plantas que originariamente tenían 2 o 3 copias. La posición de las patelinas dentro del genoma de *Arabidopsis* nos puede dar pistas sobre este suceso. Las patelinas 1-4 se sitúan dentro del cromosoma 1, y de estas solo Patelina 1 y 3 están situadas en tándem, lo que sugiere un origen por duplicación génica. No obstante las dos copias en tándem no pertenecen al mismo grupo. Patelina 3 está en el grupo más numeroso (sin dominios en el N terminal) mientras que Patelina 1 pertenece al grupo que incorpora dominios PXXP o prolina y *Coiled coil*. Patelina 1 y 2 están muy cerca filogenéticamente, o que sugiere una duplicación y una translocación reciente. Patelina 5 se sitúa en el cromosoma 3 y Patelina 6 en el cromosoma 4. Esto sugiere eventos de duplicación génica, pero también que los conocidos eventos de duplicación del genoma de *Arabidopsis* influyeron en aumentar el número de miembros de esta familia. De hecho según el análisis filogenético de las patelinas de *Arabidopsis* (Fig. 5) Patelina 3 y 5 aparecen muy cercanas, lo que sugeriría que es fruto de la duplicación del genoma o de una translocación. La función de las patelinas es esencial. Si existiera una única copia una mutación que implicara una pérdida de función, o una nueva función, posiblemente sería letal para la planta. Sin embargo al producirse una duplicación, ya sea a nivel de gen o de genoma entero, en las primeras etapas los dos genes tienen una función redundante, pero una mutación nueva (por ejemplo, la incorporación de un dominio *Coiled coil* o una mutación que implique la aparición de una prolina) puede implicar la aparición de un nuevo miembro de una familia y

## Discusión

funcionalidades diferentes, puesto que la función esencial se mantendría gracias a la duplicidad de genes.

A la vista de los resultados lo más probable es que Patelina 5 y Patelina 6 provengan de eventos de duplicación de genomas ya que se encuentran en una única copia en su cromosoma. Si tenemos esto en cuenta, y además consideramos el hecho que Patelina 5 (con dominios poliprolina o PXXP, pero sin *Coiled coil*) es la Patelina menos representada en genomas de plantas pero que es muy similar en secuencia a Patelina 3, esto nos sugiere que posiblemente esta provenga de una patelina del tipo 3 que ha obtenido dominios PXXP y no una que proviene de Patelina 1,2 que ha perdido el dominio.



## **CONCLUSIONES**



## 6. CONCLUSIONES

1. Las proteínas Patelina están presentes en todos los genomas de plantas investigados en al menos una copia, lo que implica a una función esencial en plantas.
2. Hemos encontrado plantas con 1, 2, 3, 4 ó 6 miembros de la familia patelina, lo que implica que durante la evolución se han producido numerosos procesos de duplicación génica o de duplicación genómica.
3. Las más representadas son las del tipo 3 y 6, lo que sugiere su función esencial y ser las originarias.





## **BIBLIOGRAFÍA**



## 7. BIBLIOGRAFÍA

- **Ali Dib T., Monneveux P., Araus J.L.** (1990) Breeding durum wheat for drought tolerance. Analytical, synthetical approaches, and their connections. In: Symposium on Wheatbreeding. Prospects and future approaches. Varna, Bulgaria, pp. 1-33.
- **Anantharaman V., Aravind L.** (2002) The GOLD domain, a novel protein module involved in Golgi function and secretion. *Genome Biol.*; 3: 0023.0021–0023.0027.
- **Azcón-Bieto J., Talón M.** (2008) Fundamentos de Fisiología Vegetal. Capítulo 29: Fisiología de las plantas y el estrés (2nd ed.) Interamericana-McGraw-Hill, Madrid, pp.577-597.
- **Bankaitis V.A., Aitken J.R., Cleves A.E., Dowhan W.** (1990) An essential role for a phospholipid transfer protein in yeast Golgi function. *Nature*, 347: 561-562.
- **Burkhard P., Stetefeld J., Strelkov S.V.** (2001) Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.*; 11: 82–88.
- **Chinnusamy V., Zhu J.K., Sunkar R.** (2010) Gene regulation during cold stress acclimation in plants. *Methods Mol. Biol.*, 639:39-55.
- **Cichowski K., Jacks T.** (2001) NF1 tumor suppressor gene function: narrowing the GAP. *Cell*; 104 (4): 593-604.
- **Curwin A.J., McMaster C.R.** (2008) Structure and function of the enigmatic Sec14 domain-containing proteins and the etiology of human disease. *Future Lipidol.*; 3(4): 399-410.
- **Dempsey D.A., Shah J., Klessig D.F.** (1999) Salicylic acid and disease resistance in plants. *Critical Reviews in Plant Sciences*; 18: 547-575.
- **Dong X.** (1998) SA, JA, Ethylene. And disease resistance in plants. *Current Opinion in Plant Biology*; 1: 316-329.
- **Feng S., Chen J.K., Yu H., Simon J.A., Schreiber S.L.** (1994) Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science*; 266: 1241–1247.
- **Fishman G.A., Roberts M.F., Derlacki D.J., Grimsby J.L., Yamamoto H., Sharon D., Nishiguchi K.M., Dryja T.P.** (2004) Vovel mutations in the cellular retinaldehyde-binding protein gene (RLBP1) associated with retinitis punctata albescens: evidence of interfamilial genetic heterogeneity and fundus changes in heterozygotes. *Arch. Ophthalmol.*; 122(1): 70-75.
- **Gaspar T., Franck T., Bisbis B., Kevers C., Jouve L., Hausman J.F. and Dommes J.** (2002) Concepts in plant stress physiology. Application to plant tissue cultures. *Plant Growth Regulation*; 37: 263-285.
- **Goñi, O.** (2010) Aislamiento, caracterización y funcionalidad de quitinasas y 1,3-β-glucanasas inducidas diferencialmente en frutos de *Annona cherimola Mill.* por bajas temperaturas y elevadas concentraciones de CO<sub>2</sub>. Tesis doctoral. Universidad Complutense de Madrid.

## Bibliografía

- **Izquierdo-García AC.** (2016). Caracterización del gen CRIO4 y su implicación en tolerancia a estrés por frío. Tesis doctoral. Universitat Politècnica de València.
- **Jameson P.E., Clarke S.F.** (2002) Hormone-Virus Interactions in Plants. *Critical Reviews in Plant Sciences*; 21(3): 205-228.
- **Lambers H., Stuart-Chapin III F., Pons T.L.** (1998) *Plant Physiological Ecology*. Springer-Verlag, New York.
- **Larcher W.** (1995) Plants under stress, pp 321-375. *Physiological Plant Ecology*, 3rd ed. Springer. Berlín.
- **Lichtenthaler H. K.** (1998) The Stress Concept in Plants: An Introduction. *Annals of the New York Academy of Sciences*, 851:187-98.
- **Miura K., Furumoto T.** (2013) Cold Signaling and Cold Response in Plants. *International Journal of Molecular Sciences*, 14: 5312-5337.
- **Peiro A, Izquierdo-Garcia AC, Sanchez-Navarro JA, Pallas V, Mulet JM, Aparicio F.** (2014) Patellins 3 and 6, two members of the Plant Patellin family, interact with the movement protein of Alfalfa mosaic virus and interfere with viral movement. *Mol Plant Pathol.*; 15(9): 881-91. doi: 10.1111/mpp.12146. Epub 2014 Jun 5.
- **Peterman T.K., Sequeira A., Leavitt H., Oddone A., Kremer M.** (2006b) Genome-wide analysis of the patellin family of Arabidopsis. Book of abstracts, American Society of Plant Biologists (ASPB) Meeting 2006-Boston, Massachusetts, USA.
- **Peterman T.K., Sequeira A.S., Samia J.A., Lunde E.E.** (2006a) Molecular cloning and characterization of patellin1, a novel sec14-related protein, from zucchini (*Cucurbita pepo*). *J Plant Physiol.*; 163(11):1150-8. Epub 2006 Mar 20.
- **Toumi I., Moschou P.N., Paschalidis K.A., Bouamama B., Ben Salem-Fnayou A., Ghorbel A.W., Mliki A., Roubelakis-Angelakis K.A.** (2010) Abscisic acid signals reorientation of polyamine metabolism to orchestrate stress responses via the polyamine exodus pathway in grapevine. *Journal of Plant Physiology*, 167: 519-525.
- **Sha B., Phillips S.E., Bankaitis V.A., Luo M.** (1998) Crystal structure of the *Saccharomyces cerevisiae* phosphatidylinositol transfer protein Sec14. *Nature*, 391: 506-510.
- **Vlot A.C., Dempsey M.A., Klessig D.F.** (2009) Salicylic Acid, a Multifaceted Hormone to Combat Disease. *Annual Review of Phytopathology*, 47: 177-206.

**Nota:**

La normativa establece un número máximo de páginas, por lo que la copia impresa concluye aquí. No obstante en el PDF que se ha subido a la aplicación que tienen disponible en su intranet se han incluido los anexos con los datos de secuencias y estructura de dominios por si fueran de interés.



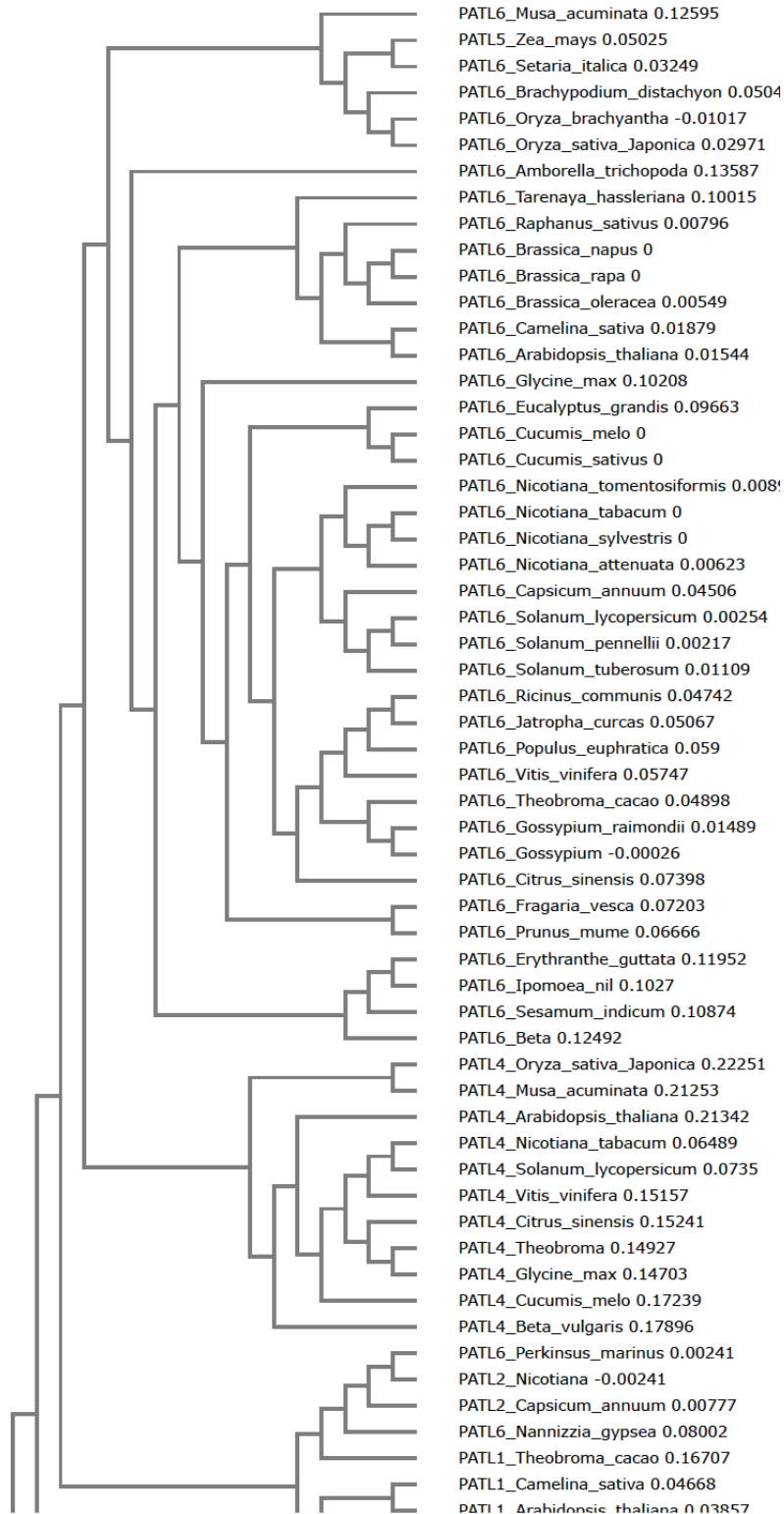
**ANEXOS**

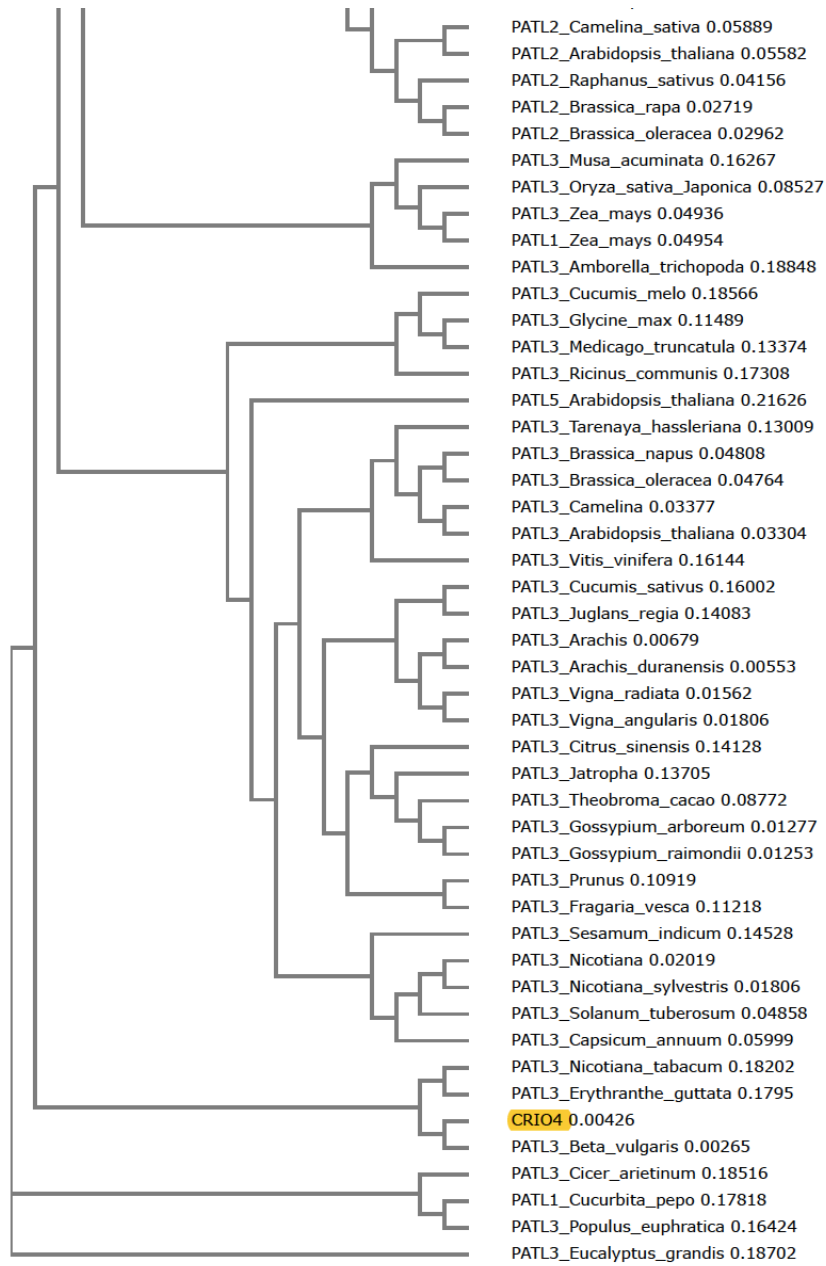




## 8. ANEXOS

### ANEXO I. Cladiograma especies homólogas a BvCRIO4





## ANEXO II. Secuencias de aminoácidos de especies homólogas a BvCRIO4 con motivos poliprolina.

>CRIO4

MAEETHKPESTVAEVVVPVAEKPAEKPAEKAVLPPEAEKLAASAEAEKPADSAEAKIAQQVSFKEETNV  
ASELPELHRKALEDLKKLIQEALEKHEFSSPPPPPPAPAKVEEKAEKKEEQPPSTTSTTTTTTAVSDEVA  
VAPPSEEAPKTDEASPKVEEPAKIVEQPPTTPAEEPEPAKTPEVVVAEEKTVEDIKETIVVEVATTTAAPV  
LTEPESVEETPKAEVVVEESPKPEEVS IWGIPLLADERSDVILLKFLRARDYRVKDAFTMIRNTARWRKEF  
EVDSSLDEDLGNDYEKVVFTHGVDKQGRPVCYNVFGFQNKELYQNTFSDAEKRKKFLRWLIQFLEKTIRTL  
FSPEGINSFVLVNDLKNSPGYGKRDLYKVIDKFLEILQDNYPFAAQQLCINVSWWYLAYNWIYLTVFTRRSK  
SKFVVFASPSKTAETLFKYIAPEQVPVQFGHSGKFGHEFSPADTVTEVTIKPGSKHPVEFAFSEETELVWELR  
VIGWDVSYGAEFLPTKEGGYTLNIAKPKKVTPEDEPVICDTFKVTEPGKVITIDNQSSKKKKLLYRSKVKTS  
Q

>PATL3\_Beta\_vulgaris

MAEETHKPESTVAEVVVPVAEKPAEKAVLQPEAEKPAASAEAEKPADSAEAKIAQQVSFKEETNVASEL  
PELHRKALEDLKKLIQEALEKHEFSSPPPPPPAPAKVEEKAEKKEEQPPSTTSTTTTTTAVSDEVAVAPP  
SEEAPKTDEASPKVEEPAKIVEQPPTTPAEEPEPAKTPEVVVAEEKTVEDIKETIVVEVATTTAAPVLT  
ESVEETPKAEVVVEESPKPEEVS IWGIPLLADERSDVILLKFLRARDYRVKDAFTMIRNTARWRKEFEVDS  
LLDEDLGNDYEKVVFTHGVDKQGRPVCYNVFGFQNKELYQNTFSDAEKRKKFLRWLIQFLEKTIRTLDF  
SPEGINSFVLVNDLKNSPGYGKRDLYKVIDKFLEILQDNYPFAAQQLCINVSWWYLAYNWIYLTVFTR  
RSKSKFVVFASPSKTAETLFKYIAPEQVPVQFGHSGKFGHEFSPADTVTEVTIKPGSKHPVEFAFSEET  
ELVWELRVIGWDVSYGAEFLPTKEGGYTLNIAKPKKVTPEDEPVICDTFKVTEPGKVITIDNQSSKKK  
KLLYRSKVKTSQ

>PATL3\_Amborella\_trichopoda

MAEEAQPPAAPATQALTAPPPAEEAPPATKEVKVEELEPASTKEVEVPKAEVAEDPKPEAETPPEASPE  
SKEAVEESIPQSVSFKEESTRVADLEDPEKKALEELKTLIDQALKENGLSTPPPPPPPPKEQEKPAPEK  
GEEKPAPEKEAEKKEEKPAEKKEEQKPAASSAVEAPAEPEVEKPKPAEPTAAEITPEAPAAPETTPPS  
KAEDVPAPASEA PPAVSVETAAAPPAAPAAEDDGAKTVEAIEETVVPASAPAAEAPETEAPVVTEE  
PAQEAQVTPPEVFIWGIP LLGDERSDVLLKFLRARDFKVKDAFTMIKNTVQWRKSFIEGLLEEDLGS  
DLESVVMNGVDREGHPVCYNV YGEFQNKIYQKTFADDEKRQRFLRWRIQFLEKGIRQLDFTPGGIST  
LVQVTDLKNSPGPGKRELQATNQAL SLLQDNYPEFVAKQVFINVPWWYLA FNRMISPFLTQRTKSK  
FVVFAGPSKSTETLFKYIAPEQVPVQYGGLSKENDT EFSTADAATEFTIKPASKNTVEIPVTEACIL  
VWELRILGWDVSYGAEFVPSAEDGYTVIVQKARKIAITD EPVIRNSFKIGEPGKIVLTVDNTSSK  
KKKLIYRYKTKKTVDISV

>PATL3\_Cicer\_arietinum

MAEEPQKPAEQVTTTSETAVEKPPQAEEVVAAAASAAAVPPAAAEAEENPAESKEVVPEAENVKSDAAA  
ADETAAVDNKISQSVSFKEETNVVSELPEPQKKALDELKLLIQEALSKHEFAVPPPPPPKAAEATPQDDK  
KPQE

EKKKEEVAAEKKKEEVEEEEKKKEEVEEEEKKKEEVVVEKKVVEEVEKKEEKQGSSEEPKTEAKPETEAVEKKVD  
VTVVELVEKIATSAAEDGAKTVEAIQESIVSVTVTNGDGEQPAADVLPSTPEEVEIWGIPLLADERSDVIL  
LKFLRARDFKVKEAFTMIKQTVIWRKEFGIEGLLQEDLGTDWDKVVFTDGYDKEGHPVCYNVFGFENKDLYQ  
KTFSDEEKRNFIRWRIQFLEKSVRKLNFAPSAISTFVQVNDLKNSPGLGKRELQATNQALQLLQDNYPEFV  
AKQIFINVPWWYLAFSRMISAFLTPTKSKFFAGPSKSADTLFKYIAPEQVPVQYGGLSREGDQEFTTADPA  
TEVTIKPATKHAVEFPIPEKSTLVWEVRVVGWVDVSYGAEFVPSAEDGYTVIVQKNRKIAPADETVINNTFKIG  
EPGKVVLTIQTSKSKKLLYRSKTIPISE

>PATL3\_Cucumis\_melo

MTDQEVVITHVPLPDKPTNDLPLPEPPVKDSFKPEGPLADPAESEVLKPAAGDDKVSADVDSFKEESTKVA  
DLSDSEKKALEEFKQLIQEALNKHEFTS P P P P P P P P P P S T L P A K V E E A P V Q S E V V V D K T D E L I D D V A K H S D E K E  
E P P K S E D K T A E T N E E G E K V K K S N E T T V P A E E K E V V A V K T E S S V D D G A K T V E A I E E T I V A V T V S A T A S T E E A V  
D A A A N P P P A A V E P E E V S I W G I P L L A D E R T D V I L L K F L R A R D F K V K E S L T M L K N T I Q W R K D F K I E E L L E E D L G S  
D L E K V V F M H G S D K E G H P V C Y N V Y G E F Q S R E L Y Q K T F S D E E K R E K F L R W R I Q F L E K S I R K L D F N P G G I C T I V Q V  
N D L K N S P G L G K W E L R Q A T K H A L Q I F Q D N Y P E F V A K Q V F I N V P W W Y L A V N K M I S P F L T H R T K S K F V F A G P S K S A  
D T L L R Y I T A E E L P V K Y G G M S K D G E F E T C D S V T E I T V K P S A K H S V E Y P V A Q G C A V T W E V R V V G W D V N Y G A E F V P  
S G E G S Y T V I I D K A R R V G S S S Q D H P V I S N T F K I S E P G K V V L S V D N P T S K K K K L L Y R F K T K S L

>PATL3\_Erythranthe\_guttata

MAEETKKTAEAAACEEVVVADVPVAEKPTTVVEKEAPPQPEPEPEPEAEKVEKAAAAGEKEEASEEKK  
VAESASFKEESNKVDDLIDPEKKALDELKKLIHEALCKREFTA P P P P P P P P A V K E E E K V E E E P K P E E K K E E E E  
T K T E E K K E E E P K T E A C A E V P A A E P V K E V P A A V V E E K A P E P P C E T V V E K I E E K V E S L V E E I K E T I V H E V T A P A P  
P P P C E E E P I A A A A A A A E E T T P E E V S I W G I P L L A D E K S D V I L L K F L R A R D F K V K D A F A M L K S V V T W R K E F K I D E  
L I E E E G I I N G L D K V V Y I H G C D K E G H P V S Y N A F G E F Q D K E L Y N N T F A D A E K R T K F L K F Y I Q F L E K N I R K L D F K P  
D G T C T I V Q I T D L K N S P G L L L Y K K E L R Q A T N Q A L Q L L Q D N Y P E F V A K Q V F V N V P W W Y V A Y N R V I S P F L T Q R T K S  
K F V F A G P T K T A E T L L K Y I A P E Q I P V Q Y G G L S K Q D A E Q E F T T A D P A T E E I I K P A C K H I V E L P I T E A G T F V W E A R  
V I G W D V S Y G A E F V P S A E G G Y T W I V Q K S R K I G P V D E T V V S C S F K V G E T G K V V L T F D N Q T S K K K K L L Y R S K T K A S  
E

>PATL3\_Nicotiana tomentosiformis

MADTAPT VHTLQPPSTAEEDT P P P P P P P A V E P A A T E P P L A S V E V E P V K Q E S V T E A E E H H P A V T V V E K D T P L  
T E P P P E E P T L V A A E Q V S V T V P E K E A E G V S V T E S D K S K P A D V K K I P E S L V S F K E E S N I V S D L S D S E R K S L E E L K  
F L V Q E S I K N Q T F T T G T P A K T Q E I Q T P S T E I T D L P Q E V S I W G V P L L K D D R T D V I L L K F L R A R D F K I K E S F A M L K  
K T I L W R K E F N I E E L V D E D L G D D L D K V V F M H G H D K E G H P V C Y N V Y G E F Q N K E L Y N K T F G D E E K R N K F L R W R I Q F  
L E R S I R K L D F N P G G I N T I F Q V S D L K N S P G P G K R E L R L A T R Q A L H L L Q D N Y P E F V A K Q V F I N V P W W Y L A F Y T M I  
S P F M T Q R T K S K F V F A G P S K T A E T L Y K Y V S P E Q V P V Q Y G G L S V D Y C E C N P E F T V N D P A T E I I V K P A T K Q T V E I I  
V N E K C I I V W E L R V L G W E V T Y S A E Y V P N T D S G Y T V N I Q K P R K M T P T D E P V V S S S F K I V E L G K I L L T I D N P T S K K  
K K L L Y R Y K D K P Y S D

>PATL3\_Oryza\_sativa\_Japonica

MAEEPQPEAAPAAVAATTEVAVAEKAPVEAEKEKKVEEETPAVEAEAKEEKKDEAAAAAAGGDEAGAIEG  
TGSFKEESNLVADLPDPEKKALDEFKQLIAAALAACEFNL P P P P P P P P KAKVEAAVEETKAEETKAE E P K AEE  
PAKEE E P K AEEVAAAAAAPPEAGTEEPKAEASSEAKTEEPKAEAAADEPAKEESKAEAAPAEAKP  
AEPEPEEKT V V V T E E E A A T K T V E A I E E T V V P A A A A P A A A A T E E A A A P E P E V Q A A A A P E P V L I W G V P L V G D D  
ERTDTVLLKFLRAREFKVKEAMAMLRSAVLWRKRFGIESLLDADLALPELDSVVFYRGADREGHPVCYNVYGE  
FQDKDLYEKAFGDEEKRRERFLKWRIQLLERGILSQLDFSPSGICSMVQVTDLKNSPMLGKHRAVT  
RQAVALLQDNYPEFIAKKVFINVPWWYLAANKMMSPF LT Q R T K S K F I F A S P A K S A E T L F R Y I A P E Q V P V Q F  
GGLFKEDDPEFTTSDAVTELTIKPSKETVEIPVTENSTIGWELRVLGWEVSYGAEFTPDAEGGYTVIVQKTR  
KVPANEPIPIMKGSFKVGEVPGKIVLTINNPASKKKKLLYRSKVKSTSESV

>PATL3\_Theobroma\_cacao

MADETPAPAPPVAA P P P P P P P P AVTESDSPAVVEKEDPSPSPVDVVSVTVTESVSTAIEEKEEQTPAAEP  
VEVVPDSTDKGEEEPPL P P P P P Q P V K S E S E S P A V T E K A K E E E P P A V T E K A K E E E P P A A T E K A K E E P P A A T V A V L  
ETELQQPPELPPQEVVTESSLAAMMEKEEGGAPEPTVATTTTTTSAEEGVAVVVEEKKIPQNLGSF  
KEESNKVADLSDFERKALEELKHFVQEAIDTHLFTSETKSEENPEKEKKEE P K E V S I W G I P L L K D D R S D V I  
LLKFLRARDFKVKDAFVMIKNTIRWRKEFGIDELLEDLGDMEKVVFMHGQDREGHPVCYNVYGEFQNKDLY  
QKAFSDEEKRMKFLRWRIQFLEKSIRKLDSPDGINTIFQVSDLKNSPGPGKRELRLATKQALQLLQDNYPEF  
VAKQVFINVPWWYLVFYTMISPFMTQRTKSKFVVFAGPAKSAETL F K Y M A P E Q V P I Q Y G G L S V D Y C D C N P E F G D  
ADPATEITVKGTKQTVEITIEYKCDLVWEIRVVGWEVSYGAEFMPNAKDSYTVIIQKPTKMSPKDEPVVSQC  
FRVGELGKVL L T V D N P T S K K K K L V Y R F K V K P H C D

>PATL3\_Vigna\_angularis

MAENDSN P P P P P P P P A A S P P Q E P P V P P S L S S E N D A K A D D V I E E K V K A E E V L V A E D K D E V K G G E E A V V V E E K E  
KEEKEEKEEKEKKEEKEEVTGETESESLKEESNRVSDSERKGIEELKVVREELERKEEKEEVS IWGVPLFKD  
QRTDVILLKFLRARDLRVNDLVMIQNTLRWRKEFGIDALLEEDLGEELKVVFMHGQDREGHPVCYNVYGEF  
QNKELYQKAFSTEENRTKFLRWRIQLLERSIRNLDFTPGGINTIFQVNDLKNSPGPAKRELRIATKQALQLLQ  
DNYPEFVAKQVFINAPWWYLAFNLTISPFLTPTRTKSKFIFAGPSKSPDTL F K Y I S P E Q V P V Q Y G G L S V D F C D C  
NPDFMTSDPVTEIPIKPTTKQTVEIAIYEKCIIVWELRVVGWEVTVYNAEFKPKDAKDAYTVVIQKATKMSPTDE  
PVVSNSFKVGELGKLLLTIDNPTLKNKRLLYRFNIKPYCD

>PATL3\_Vitis\_vinifera

MADETDAAAPPPTLLAAASVVPETAAPPETAASEEVLTVLDVAPPVTEKEVSPQK P P P P P P E E V V A V D S E K  
KVPQNLVSFKEESNRLADLSEERRALEELKQSVQEALRNGIFTSQPQ P P P P P P P P Q S A E K P P E K I E E A S E K R  
EPNPVAESEISTQEE S A K D E N V K P T P N P T I E S I L K H E S P T Q E D V S I W G I P L L K D E R S D M I L L K F L R A R E F K V K  
EAFAMLKNTIFWRKEFGIDALVDDDLGEHLEKVVFMHGFD RD G H P V C Y N V Y G E F Q N K E L Y Q K T F S D E E K R M K F  
LRWRIQFLERSIRKLDFTPGGVNTIFQVNDLKNSPGPGK W E L R Q A T K Q A L Q L L Q D N Y P E F V A K Q V F I N V P W W Y

LAFYMMISPFLLTQRTKSKFVFASPAKSAKTLFKYIISPEQVPIQYGGLSVDYCDCNPDFGIADPVTEITVKPST  
KQTV EILVSEQCIVWEVRVVGWEVAYGAEFI PDAEDEYTVVVQKATKMAPTDDPVMCNSFKIKELGKIVITI  
DNPTS KKKKLLYRFKVKPYSN

>PATL3\_Zea\_mays

MAEETQPEAAAEAEV VVTEPAPAPAEAEV PAAADAEAVTE TEKKADEPAVTADDAGKGTGSFK EESNLVED  
LPDPEKKVLDEFKHLIAAALAAGEFNL **PPPPPP** KAKEEPKAEETKTEESKTEDPAEE **EPK** AESAAE EPKAEV  
AANAPDEEVKTEVPPVVEEAKAETVAVEAKPAEPEPQEKTVVVAEEEEPATKTVEAIEESVVSAD EAAAPEPVLV  
WGVPLVGDDERTDVTLLKFLRAREFKVKEAMAMLKSAVLWRKRFGITSLLDADLGLPELENNVVFYRGADREGH  
PVCYNVYGEFQDKDLYEKAFGDDEKRERFLKWRIQLLERGILSKLDFSPNGICSMVQVTDLKN SPPMLGKHRA  
VTRQAVTLLQDNYPEFIAKKVFINVPWWYLAANKMMS PFFTQRTKSKFVFASPAKSAETLFRYIAPEQVPVQF  
GGLFKEDDPEFTTLDVTVELTIKPSKETIEIPVTENSAI VWELRVLGWEVSYSAEFTPDTEGGYTVIIQKTR  
KVPANE EPI MKGSFKVGE PGKLV LTVNNPASKKKKLLYRSKVKSISE

>PATL1\_Arabidopsis\_thaliana

MAQEEVQKSADVAAAPV VKEKPI TDKEVTI PTPVAEKEEVAAPVSDEKAVPEKEVTPEKEAPAAEAEKSVS  
VKEEETVVVAEKVVVLTAAEEVQKKALEEFKELVREALNKREFTAPVTPVKEEKTEEKKTEEETKEEEKTEEKK  
EETTTEVKVEEEKPAVPAEEEEKSSEAAPVETKSEEKPEEKAEVTTTEKASSAEEDGKTVEAIEESIVSVSPP  
ESAVAPVVVETVAVAAEAPVEPEEVS IWGVPLLDQERSDVILTKFLRARDFKVKEAL TMLKNTVQWRKENKID  
ELVESGEEVSEFEKMFVAHGVDKEGHVVIYSSYGEFQNKELFSDKEKLNKFLSWRIQLQEKCVRAIDFSNPEA  
KSSFV FVSDFRNAPGLGKRALWQFIRRAVKQFEDNYPEFAAKELFINVPWWYIPYYKTFGSIITSPRTRSKMV  
LAGPSKSADTIFKYIAPEQVPVKYGGLSKDTPLTEETITEAIVKPAANYTIELPASEACTLSWELRVLGADVS  
YGAQFEPTTEGSYAVIVSKTRKIGSTDEPVITDSFKVGE PGKIVITIDNQTS KKKKVL YRFKTQ

>PATL2\_Arabidopsis\_thaliana

MAQEEIQKPTASVPV VKEETPAPVKEVEVPV TTEKAVAAPAPEATEEKV VSEVAVPETEVTAVKEEEVATG  
KEILQSESFKEEGYLASELQEA EKNALAELKELVREALNKREFTA **PPPPPP** APVKEEKV EKKTEEETEEKKEEV  
KTEEKSLEAETKEEEKS AAPATVETKKEE ILAAPAPIVAETKKEETPVAPAPVETKPAAPVVAETKKEEILPA  
APVTTETKVEEKVVPVETTPAAPVTTETK EEEKAAPVTTETK EEEKAAPGETKKEEKATASTQVKRASKFIKD  
IFVSVTTSEKKKKEEKPAVVTIEKAF AADQEEETKTVEAVEESIVSITLPETAAYVEPEEVS IWGIPLLEDER  
SDVILLKFLRARDFKVKEAFTMLKNTVQWRKENKIDDLVSEDLEGSEFEKLVFTHGVDKQGHVVIYSSYGEFQ  
NKEIFSDKEKLSKFLKWRIQFQEKCVRS LDFSP EAKSSFV FVSDFRNAPGLGQRALWQFIKRAVKQFEDNYPE  
FVAKELFINVPWWYIPYYKTFGSIITSPRTRSKMVL SGP SKSAETIFKYVAPEVVPVKYGGLSKDS PFTVEDG  
VTEAVVKSTSKYTIDLPA TE GSTLSWELRVLGADVS YGAQFEP SNEASYTVIVSKNRKVG L TDEPVITDSFKA  
SEAGKV VITIDNQTF KKKKVL YRSKTQA

>PATL3\_Arabidopsis\_thaliana

MAE EPTTTTLVTPEKLPSPSLTPSEVSESTQDALPTETETLEKVTETNPPETADTTTKPEEETA AEHHPPT  
VTETETASTEKQEVKDEASQKEVAEEKSMIPQNLG SFKEESSKLSDL SNSEKKS LDELKHLVREALDNHQFT

NTPEEVKIWGIPLLEDDRSDVVLLKFLRAREFKVKDSFAMLKNTIKWRKEFKIDELVEEDLVDDLDKVVFMHG  
HDREGHPVCYNVYGEFQNKELYNKTFSDDEEKRKHFRLRTRIQFLERSIRKLDFFSSGGVSTIFQVNDMKNSPGLG  
KKELRSATKQAVELLQDNYPEFVFKQAFINVPWWYLVFYTVIGPFMTPRSRSKLVFAGPSRSAETLTKYISPE  
QVPVQYGGLSVDPDCNPDFSLEDSASEITVKPGTKQTVEIIIEKCELVWEIRVTGWEVSYKAEFVPEEKDA  
YTVVIQKPRKMRPSDEPVLTHSFKVNELGKVLLTVDNPTSCKKKLVYRFNVKPL

>PATL4\_Arabidopsis\_thaliana

MTAEVKVEEKQVESEVVIAPAVVPEETTVMKAVVEETKVEEDESKEGVEKSASFKEESDFFADLKESEKKA  
LSDLKSKLEEAIVDNTLLKTKKKESSPMKEKKEEVVPEAEVEKKEEAAEKKVEEKKSEAVVTEEAPKAET  
VEAVVTEEIIIPKEEVTTVVEKVEEETKEEKKTEDVVTTEEVAETIEVEDEDESVDKDIELWGVPLLPKSGAE  
STDVILLKFLRARDFKVNEAFEMLKTKLWRKQNKIDSILGEEFGEDLATAAYMNGVDRESHPVCYNVHSEEL  
YQTIGSEKNREKFLRWRFLMEKGIQKLNLPKGGVTSLLQIHDLNAPGVSRTEIWVGIKKVIETLQDNYPEF  
VSRNIFINVPWFYAMRAVLSFPLTQRTKSKFVVARPAKVRETLKYIIPADELPVQYGGFKTVDDTEFSNETV  
SEVVVKPGSSETIEIPAPETEGTLVWDIAVLGWEVNYKEEFVPTTEEGAYTVIVQKVKMGANEGPIRNSFKNS  
QAGKIVLTVDNVSGKKKKVLYRYRTKTESSS

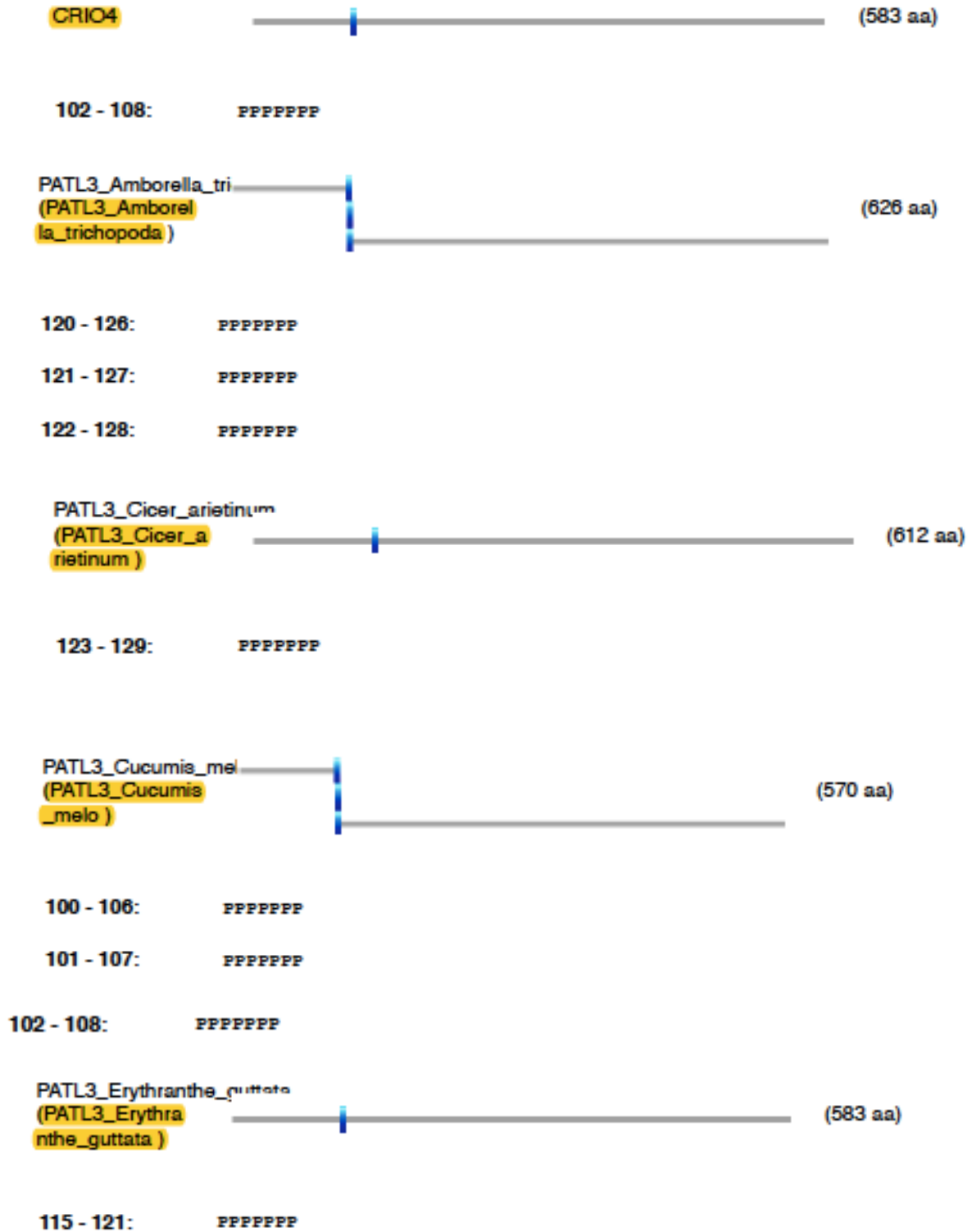
>PATL5\_Arabidopsis\_thaliana

MSQDSATTTPPPPLTSDVSMPSGEEDEPKHVTSEEEAPVTSETNLKLPMPLEESNHTAEVSEKVTPET  
MTLESEGLNHAADSEQTHEVTPETETAKLEVLNHTAEDSEQTHEVTPKETVKSEFLNHVAEDSEQTHEVTP  
ETETVKSEVLNHAADSEQPRGVTPPETETSEADTSLLVTSETEENHAAEDYSETEPSQKLMLEQRRKYME  
VEDWTEPELPDEAVLEAAASVPEPKQPEPQT **PPPP**STTTSTVASRSLAEMNREEAEVEEKQKIQIPRSLGS  
FKEETNKISDLSETELNALQELRHLLQVSDSSKTSIWGVPLLKDDRTDVVLLKFLRARDFKPQEAYSMLNKT  
LQWRIDFNIEELLDENLGDDLDKVVFMOGDKENHPVCYNVYGEFQNKDLYQKTFSDDEEKRRERFLRWRIQFLE  
KSIRNLDFVAGGVSTICQVNDLKNSPGPKTELRLATKQALHLLQDNYPEFVSKQIFINVPWWYLAFYRIISP  
FMSQRSKSKLVFAGPSRSAETLLKYISPEHVPVQYGGLSVDNCECNSDFTHDDIATEITVKPTTKQTVEIIIVY  
EKCTIVWEIRVVGWEVSYGAEFVPEKGYTVIIQKPRKMTAKNELVVSFSFKVGEVGRILLTVDNPTSTKMK  
LIYRFKVKPLACE

>PATL6\_Arabidopsis\_thaliana

MDASLSPFDHQKTQNTPEPKSFITSLITLRSNNIKEDTYFVSELKPTEQKSLQELKEKLSASSSKASSMWG  
VSLGDDKADVILLKFLRARDFKVADSLRMLEKCLEWREEFKAEKLTEEDLGFKDLEGKVAYMRGYDKEGHP  
VCYNAYGVFKEKEMYERVFGDEEKLKFLRWRVQVLERGVKMLHFKPGGVNSIIQVTDLKDMPKRELRVASNQ  
ILSLFQDNYPELVATKIFINVPWYFVSVIYSMFSPFLTQRTKSKFVMSKEGNAETLYKFIKIRPEDIPVQYGGLS  
RPTDSQNGPPKPASEFSIKGGEKVNIQIEGIEGGATITWDIVVGGWDLEYSAEFVPAEESYAIVVEKPKMKK  
ATDEAVCNSTTVEAGKLILSVDNTLSRKKKVAAYRYTVRKSTTTV

**ANEXO III. Localización de los dominios Sec14, GOLD, motivos poliprolina y Coiled Coil.**





PATL3\_Nicotiana  
(PATL3\_Nicotiana)  (523 aa)

22 - 28: P P P P P P P

PATL3\_Oryza\_sativa Japonica  
(PATL3\_Oryza\_sativa\_Japonica)  (613 aa)

112 - 118: P P P P P P P

PATL3\_Theobroma\_cacao  
(PATL3\_Theobroma\_cacao)  (607 aa)

15 - 21: P P P P P P P

16 - 22: P P P P P P P

PATL3\_Vigna\_angularis  
(PATL3\_Vigna\_angularis)  (476 aa)

8 - 14: P P P P P P P

PATL3\_Vitis\_vinifera  
(PATL3\_Vitis\_vinifera)  (530 aa)

120 - 126: P P P P P P P

121 - 127: P P P P P P P

(PATL3\_Zea\_mays)  (556 aa)

99 - 105: P P P P P P P

PATL2\_Arabidopsis\_thaliana

(PATL2\_Arabido  
psis\_thaliana)



(683 aa)

117 - 121:

PPPPP

PATL5\_Arabidopsis\_thaliana

(PATL5\_Arabido  
psis\_thaliana)



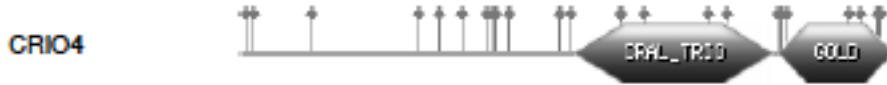
(668 aa)

249 - 253:

PPPPP

CRIO4 (583 aa)

MAEETHKPESTVAEVVVVPAEKPAEKPAEKAVLPPEAEKLAASAEAEKPADGAEAKIAQQVSEPK  
 EETNVASGLPELHRHALEDLKKLIQEALKEHEFGSPPPPPPPAPAKVEEKAEEKKEEQPPSTTSTT  
 TTTTTAVSGDEVAVAPPSEEAAPTDEASPKVEEPAKIVEQPFTTPAEPEPAKTPFEVVVAEEKTV  
 EDIKETIVVEVATTAAPVLTPEPESVEETPKAEVVEEGSPKEPEVSGINGIPLLADERSDVILLEK  
 FLRARDYRVDAFTMIKNTARWRKEFEVDGLLDEDLGNDYERVVFTBQVDEQGRPVVCYNVFGCFQK  
 KELYQNTFSDAEKRRKFLRWLIQFLEETIRTLDPSPEGINSFVLVNDLKNSPCYGKRDLYKVIDKF  
 LEILQDNYPEFAARQLCINVGWNYLAINWIIYLVFTTRRSGSKFVFASPSRTAETLFRYIAPEQVPV  
 QFGGSKFGEHEFPADTVTEVTIKPGSKHPVEFAPSEETELVWELRVIGWVDVSYGAEFLPTEKGG  
 YTLNIAKPKKVTPADEPVICDTFRVTEPGKRVVITIDNQSSKKKLLYRSKVTGQ



**PS50191 CRAL\_TRIO** *CRAL-TRIO lipid binding domain profile :*

297 - 472: score = 16.700

DEDLGN DYERVVFTBQVDEQGRPVVCYNVFGCFQKELyqntfSDAEKRRKFLRWLIQFLE  
 RTIRTLDPSPEGINS-FVLVNDLKNSPCYGKRDL---YKVIDKFLEILQDNYPEFAARQL  
 CINVGWNYLAINWIIYLVFTTRRSGSKFVFASPSRTAETLFRYIAPEQVPVQFGGSKFGE

Predicted feature:

DOMAIN	297	472	CRAL-TRIO
--------	-----	-----	-----------

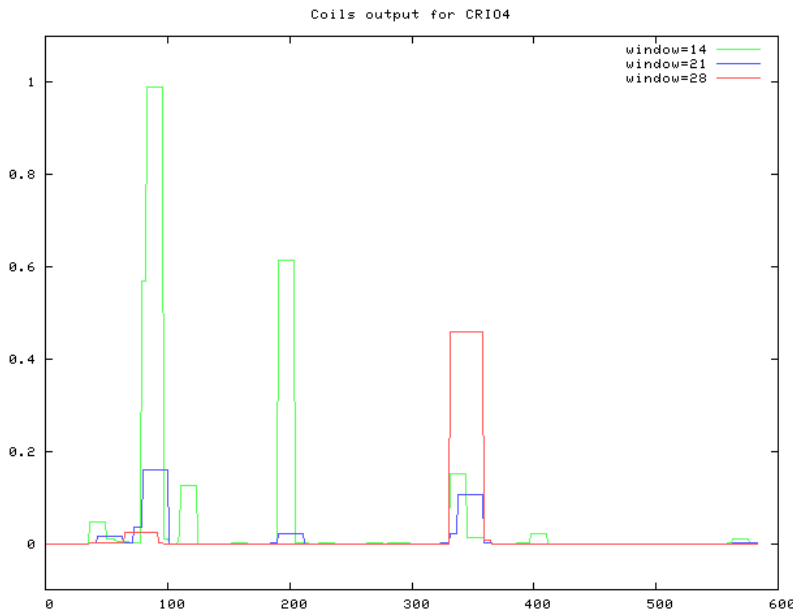
**PS50866 GOLD** *GOLD domain profile :*

478 - 579: score = 9.645

ADTVTEVTIKPGSKHPVEFAPSEETELVWELRVIGWVDVSYGAEFLPTEKGGYTLNIAkpk  
 kvtpADEPVICDTFRVTEPGKRVVITIDNQSSKKK--LLYRSKV

Predicted feature:

DOMAIN	478	579	GOLD
--------	-----	-----	------



PATL3\_Beta\_vulgaris (579 aa)

MAEETHKEPSTVAEVVVVPAEKPAEKAVLQPEAEKPAEAESAEAEKPADGAEAKIAQQVSPKEETN  
 VAGELPELHRALEDLKKLIQEALKEHEFGSPPPPPPAPAKVEEKAEKKKEEQPPSTTSTTTTT  
 TAVGDEVAVAPPSEEAAPTDEAGPFVEEPAKIVEQPPTTAEPEPAKTFEVVVAEEERTVEDIK  
 ETIVVEVATTTAAFPVLTEPESVEETPKAEVVEEGPKEPEVEVGINGIPLLADERSDVILLKFLRA  
 RDYRVKDAFTMIRNTARWRKEFEVDSLLDEDLGNDEKVVFTHGVDQGRPVVCYNVFGEPQNKELY  
 QNTFSDAEKRRKFLRWLIQFLEKTIRTLDPSPEGINSFVLVNDLKNSPCYGGRDLYRVIDEFLIIL  
 QDNYPEFAAKQLCINVSWWYLAYNWIYLTVFTPRSEKSFVFPASPKTAEATLFRYIAPEQVPVQFGG  
 HSKPGEHEFPADTVTEVTIKPGSKPIEFAPSEETELVWELRVIGNDVSYGAEFLPTKEGGYTLN  
 IAKPKVTPADEPVICDTFRVTEPGKVVITIDNQSSKSKLLYRSKVTGQ



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

293 - 468: score = 17.105

DEDLGNDEKVVFTHGVDQGRPVVCYNVFGEPQNKELYqntfSDAEKRRKFLRWLIQFLE  
 KTIRTLDPSPEGINS-FVLVNDLKNSPCYGGRDLYRVIDEFLIILQDNYPEFAAKQL  
 CINVSWWYLAYNWIYLTVFTPRSEKSFVFPASPKTAEATLFRYIAPEQVPVQFGGHSKPFCE

Predicted feature:

DOMAIN	293	468	CRAL-TRIO

**PS50866 GOLD** GOLD domain profile :

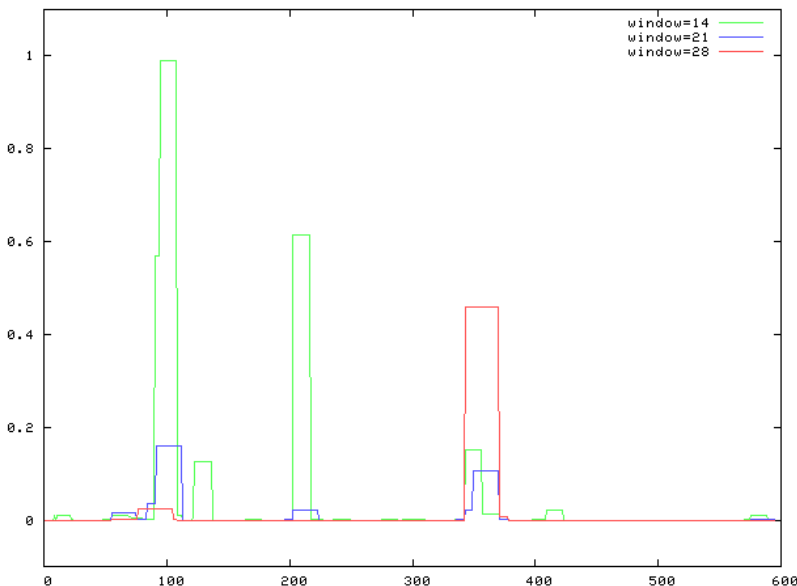
474 - 575: score = 9.576

ADTVTEVTIKPGSKPIEFAPSEETELVWELRVIGNDVSYGAEFLPTKEGGYTLNIakpk  
 kvTPADEPVICDTFRVTEPGKVVITIDNQSSKSKLLYRSKVTGQ

Predicted feature:

DOMAIN	474	575	GOLD

Coils output for PATL3 Beta vulgaris



PATL3\_Amborella\_trichopoda (626 aa)

MAEEAQPPAATQALTAFFPAEEAPPATKEVVKVELEPASTEEVEVPTAEVAEDFPPEAETPPEA  
 SPESKAVEESIPQSVGFKEESTRVADLEDPERKALEELKTLIDQALKENGLSTPPPPPPPPKBEQ  
 KPAPFKGEEKPAPEREAKEEKEEKPAAEKEEQKPAASSAVEAPAEPEVEKPKPAEPTAAEITPEAPAA  
 PETPPSKAEDVPAPAGEAPPVSVETAAPPAAPAAEDDGCARTVEAIEETVVVPGASAPAAEAPET  
 EAPVVTEEPAQEAQVTPPEEVF IWCIPLLGDERSDVLLKFLRARDFKVEDAFTHIKNTVQWRKSF  
 IEGLLEEDLGSDLSEVVMNCGVDRECHPVCYNVYGEFQNKIYQKTFADDEKQRFLRWRIQFLEK  
 CIRQLDFTPGGISTLVQVTDLEKSPGCGKRELRQATNQALSLLDQNYPEFVAEQVF INVFWWYLAF  
 NRMISFFLTQRTESEKPFVAGPKSKSTETLFRYIAPBQVVPVQYCGLGKENDTEFSTADAATEPTIKPA  
 SKNTVEIPVTEACILVWELRILGWDVSYGAEFVPSAEDGTYTVIVQKARKIAITDEPVIRNSFKICE  
 PGIIVLTVDNVTSSEKKKLIYRYTKRKTVDISV



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

336 - 511: score = 19.496

EEGLGSDLESVVMNCGVDRECHPVCYNVYGEFQNKIYQKTFADDEKQRFLRWRIQFLE  
 KIRQLDFTPGGISTLVQVTDLEKSPGCGKRELRQATNQALSLLDQNYPEFVAEQV  
 FINVFWWYLAFNRMISFFLTQRTESEKPFVAGPKSKSTETLFRYIAPBQVVPVQYCGLGKEND

Predicted feature:

DOMAIN	336	511	CRAL-TRIO
--------	-----	-----	-----------

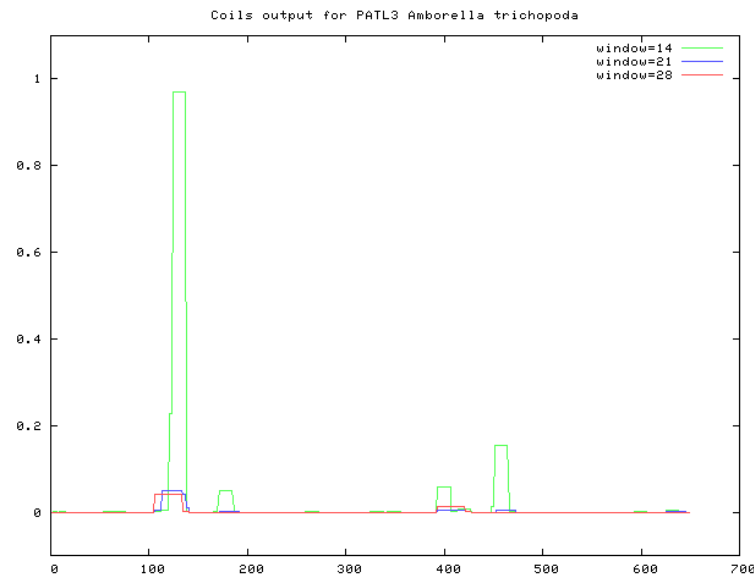
**PS50868 GOLD** GOLD domain profile :

487 - 618: score = 11.191

TETLFRYIAPBQVVPVQYCGLGKENDTEFSTADAATEPTIKPASKNTVEIPVTEACILVWE  
 LRLGWDVSYGAEFVPSAEDGTYTVIVQKARKIAITDEPVIRNSFKICEPGIIVLTVDNVTS  
 SEKKK--LIYRYTK

Predicted feature:

DOMAIN	487	618	GOLD
--------	-----	-----	------



PATL3\_Cicer\_arietinum (612 aa)

MAEFPQKPABQVVTTTGETAVEKFPQAEVVAAAGAAAAPPAAAAEAENPAESKEVVPEAENVKSD  
 AAAADETAAVDNKIGQSVSFKEETNVVSELPEPQKALDELKLLIQEALSGHEFAVPPPPPPFAA  
 EATPQDDKPKPQEKKKEEVAEKKKEEVEEKKKKEEVEEKKKKEEVEVVVEKKEVEEVEKKEEKQGGG  
 EEPRTAKPSTEAVEKRVVDTVVVELVERIATSAEEDGAKTVEAIQESIVSVTVTNGDCEQPAADVE  
 LPPSTPEEVEIWGIPLLADERSDVILLKFLRARDPKVFEAFTHIRQTVINRKEFGIEGLLQEDLGT  
 DWDEVVFTDCYDKEGHPVCYNVFGFENEDLYQKTFSDDEEKRNKFIWRRIQFLERSVRELNFAPSA  
 ISTFVQVNDLKNPGLGKRELKQATNQAQLQLQDNYPEFVAEQIFINVPWNYLAFSRMISAFLTFR  
 TSKFFPAGPSKADTLFKYIAPQVQVYGGLSREGDQEFTTADPATEVTIKPATKHAFFPIPE  
 KSTLVWEVRVVGWVDVSYGAEFVPSAEDCYTIVQKRNKIAPADETVINNTFKICEPGKVVLTDNQ  
 TSKKKLLYRSKTIPIGE



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

325 - 500: score = 18.965

QEDLGTDWDEKVVFTDCYDKEGHPVCYNVFGFENEDLYqkTfSDDEEKRNKFIWRRIQFLE  
 KSVRELNFAPSAIST-FVQVNDLknPGLGKREL---RQATNQAQLQLQDNYPEFVAEQI  
 FINVPWNYLAFSRMISAFLTFRTEKFFPAGPSKADTLFKYIAPQVQVYGGLSREGD

**Predicted feature:**

DOMAIN	325	500	CRAL-TRIO

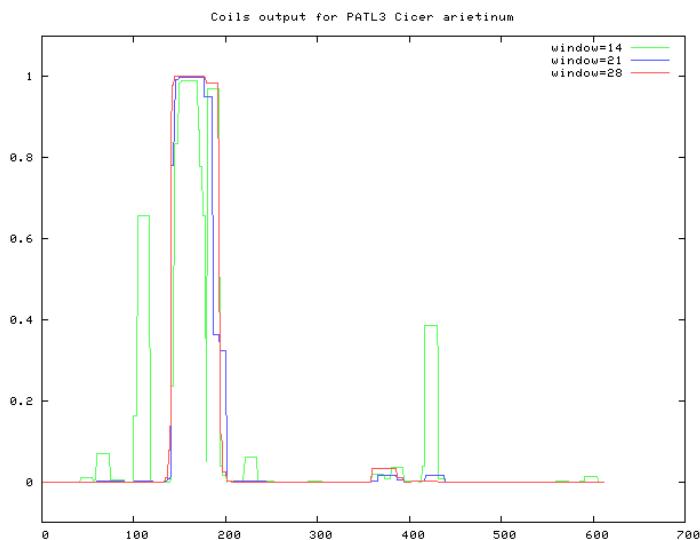
**PS50866 GOLD** GOLD domain profile :

501 - 607: score = 10.937

QEFTTADPATEvtikPATKHAFFPIPEKSTLVWEVRVVGWVDVSYGAEFVPSaDGYTVI  
 VQknrkiAPADETVINNTFKICEPGKVVLTDNQTSKKE--LLYRSKTI

**Predicted feature:**

DOMAIN	501	607	GOLD



PATL3\_Cucumis\_melo (570 aa)

MTDQEVVITHVPLPDKPTNDLPPLPEPPVVDSDFKPEGLADPAESEVLEKPAQDDKVSADVDGFKKE  
 STHVADLSDGSEKKALEEFQQLIQEALNKEHFTGPPPPPPPPSTLPAEVEEAPVQSEVVVDKTDDEL  
 IDDVAKHSDKEEPPKSEDEKTAETNEEGEERVKESNETTVPAEKEVVAVKTESSVDDGAKTVEAI  
 EETIVAVTVSATASTEBAVDAANPPPAAVEPEEVSINGIPLLADERTDVIILLRFLRARDPKVKES  
 LTMLENTIQWRKDFRIEELLEEDLGSDLERVVFMHGSDEKGFVPCYINVYGEFQGRELYQKTFSDDE  
 KREKFLRWRIQFLERSIRKLDPNPCCICTIVQVNDLKNSPGLGKWELEKQATKHALQIFQDNYPEFV  
 AKQVFINVPWYLAVNEMISPFLLTHRTSEKPFVFCGPKSADTLLRYITABELPVRYGGMSEKDFE  
 TCDGVTEITVRFSAKHSVEYFVAQGCVAWTWEVRVVGWVDVNYGAEFVPSGEGSYTVIIDKARRVGGG  
 SQDHPVISNTFRKISEPGKVVLSVDNPTSKKKKLLYRPFKTESL



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

285 - 460: score = 19.202

EEDLGSDLERVVFMHGSDKEGHPVCYNVYGEFQGRelyqktfDDEEKREKFLRWRIQFLE  
 ESIRKLDPNPCCICT-IVQVNDLkncPGLGKWELE--RQATKHALQIFQDNYPEFVAKQV  
 FINVPWYLAVNEMISPFLLTHRTSEKPFVFCGPKSADTLLRYITABELPVRYGGMSEKDFE

Predicted feature:

DOMAIN	285	460	CRAL-TRIO
--------	-----	-----	-----------

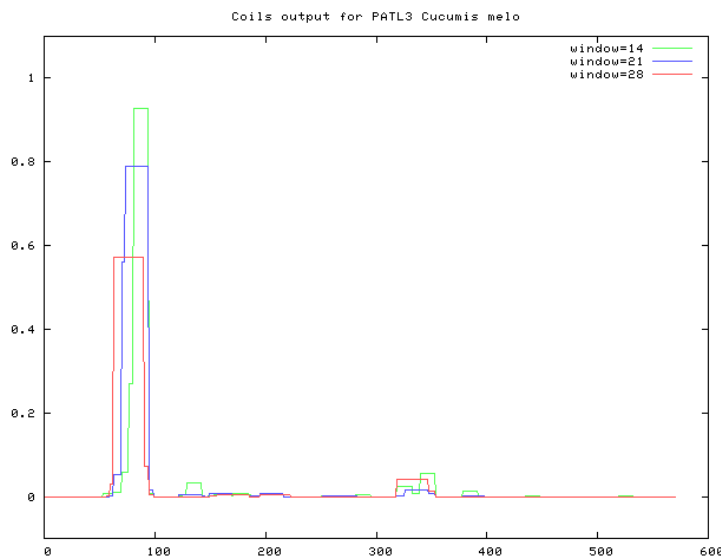
**PS50866 GOLD** GOLD domain profile :

465 - 567: score = 10.799

DSVTEITVRFSAKHSVEYFVAQGCVAWTWEVRVVGWVDVNYGAEFVPSgEGSYTVIIDkarr  
 vgcSQDHPVISNTFRKISEPGKVVLSVDNPTSKKKK--LLYRPFK

Predicted feature:

DOMAIN	465	567	GOLD
--------	-----	-----	------



PATL3\_Erythranthe\_guttata (583 aa)

MAEETKKTAEAAACEEIVVVVADVPVAEKPTTTVVVEKEAPPQPEPEPEPEAEKVEKAAAAEGERKEEA  
 SEERKVAEGASFRREGNVDLDLIDPEKRALDELKELIHEALCREFTAPPPPPPPAVKEEERKVEE  
 EPKPEEKKEEETKTEEKKEEPTTEACAEVPAAEFVKEVPAAVVEEKAPFPCCETVVEKIEEKVE  
 SLVEEIKETIVHEVTAPAPPPPCEEEPIAAAAAAAEETTPEEVSIGCIPLLADERSDVILLKFLRA  
 RDPKVEDAFANLKSQVTVNRKEFRIDELIEECCIINGLDEKVVYIHGCCDEKGPVSYNAFGEFQDEKEL  
 YNNTFADAERKTRKFLKFIQFLEKNIRKLDKFKPDGCTCTIVQITDLEKSPGLLLYKRELKQATNQAL  
 QLLQDNYPEFVARQVFNVPWVYVAYNRVISPFLLTQRTEKRFVVFACPTTAETLLKYIAPEQIPVQ  
 YGCLSEQDAEQEFTTADPATEEIIKPAKHIIVELPI TEACTFVWEARVIGWDVSYGAEFVPSABCG  
 YTWIVQKSRKICPVDETVVSCSFRVGETGRVVLTFDNQTSKRRKLLYRSKTKASE



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

294 - 471: score = 20.167

EEGIINGLDEKVVYIHGCCDEKGPVSYNAFGEFQDEKELYNNTFADAERKTRKFLKFIQFLE  
 KNIRKLDKFKPDGCTCT-IVQITDLEKSPGLLLYKRELKQATNQALQLLQDNYPEFVARQV  
 FVNVPWVYVAYNRVISPFLLTQRTEKRFVVFACPTTAETLLKYIAPEQIPVQYGCLSEQDA

Predicted feature:

DOMAIN	294	471	CRAL-TRIO
--------	-----	-----	-----------

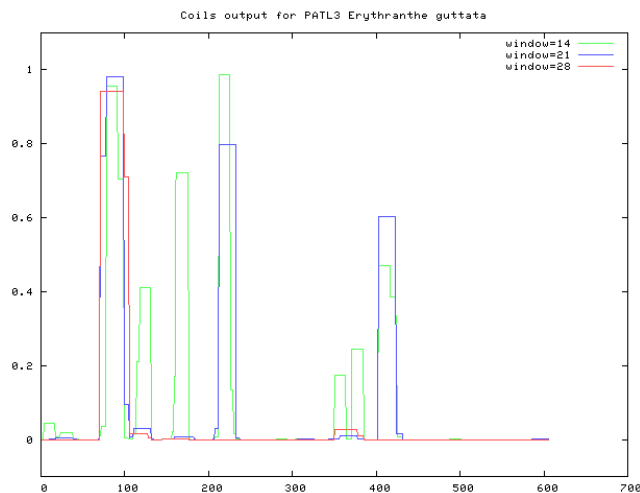
**PS50866 GOLD** GOLD domain profile :

479 - 579: score = 8.861

DPATEEIIKPAKHIIVELPI TEACTFVWEARVIGWDVSYGAEFVPSAEGCYTWIVQKSRK  
 IGPVDETVVSCSFRVGETGRVVLTFDNQTSKRRK--LLYRSKT

Predicted feature:

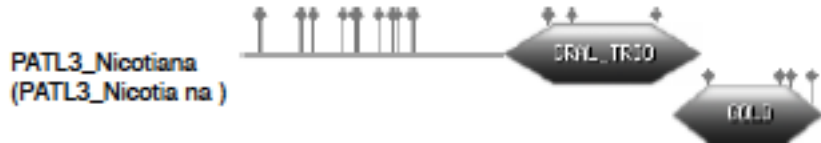
DOMAIN	479	579	GOLD
--------	-----	-----	------





PATL3\_Nicotiana (523 aa)

MADTAPTVHTLQPPGTAEDTTPPPPPPPAVEPAATEPPLASVEVEPVKQEGVTEAEHHHPAVTVVE  
 EDTPLTEPPPEPTLVAAEQVSVTVPEREAEQVSVTEGDSKPADVVKIPEGLVSPREESNIVGDL  
 SDGERKSLLEELKFLVQESIRKQTFPTTCTPARTQEIQTPTSTEITDLPQEVSIWCVPLLDRTDVI  
 LKFLRARDFKIKESFAMLEKTIILWREKFNIEELVDEDLGDDLDKVVVFMHGHDKEGHPVCYNVYGEF  
 QNKELYNKTFGDEEKRNKFLRWRIQFLERSIRKLDNFPGGINTIFQVSDLENSPGCGKRELRLATR  
 QALHLLQDNYPEFVARQVFINVFWYLAFYTMISPFMTQRTESKRVFAGPSKTAETLYRYVSPQV  
 PVQYGGLSVDYCECNPEFTVNDPATEIIVKPAKQTVETIIVNEECIIVWELRVLGWEVTVYSAEYVP  
 NTDGCTTVNIQKPRKMTPTDEPVVSSGFRIVELGRIILLTIDNPTGKRRKLLRYKDEPYSD



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

233 - 408: score = 18.783

DEDLGDLDKVVVFMHGHDKEGHPVCYNVYGEFQNKELYNKTFGDEEKRNKFLRWRIQFLERSIRKLDNFPGGINT-IFQVSDLENSPGCGKREL---RLATRQALHLLQDNYPEFVARQVFINVFWYLAFYTMISPFMTQRTESKRVFAGPSKTAETLYRYVSPQVPVQYGGLSVDYCECNPEFTVNDPATEIIVKPAKQTVETIIVNEECIIVWELRVLGWEVTVYSAEYVPNTDSCYTTVNIQKPRKMTPTDEPVVSSGFRIVELGRIILLTIDNPTGKRRKLLRYKDEPYSD

Predicted feature:

DOMAIN	233	408	CRAL-TRIO
--------	-----	-----	-----------

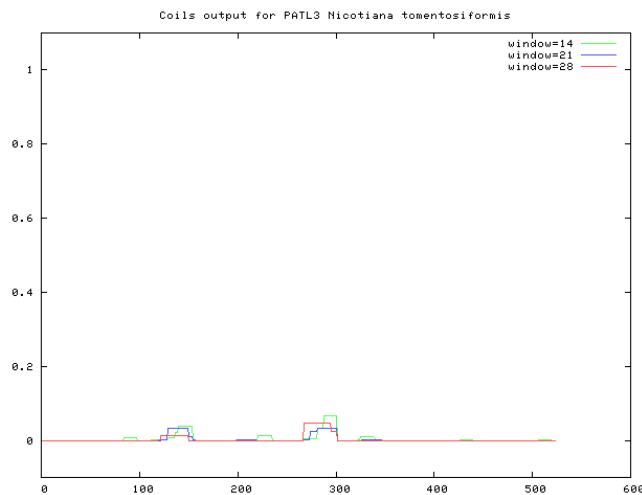
**PS50866 GOLD** GOLD domain profile :

384 - 518: score = 8.837

AETLYRYVSPQVQVYGGLSVDYCECNPEFTVNDPATEIIVKPAKQTVETIIVNEECIIVWELRVLGWEVTVYSAEYVPNTDSCYTTVNIQKPRKMTPTDEPVVSSGFRIVELGRIILLTIDNPTGKRRKLLRYKDEPYSD

Predicted feature:

DOMAIN	384	518	GOLD
--------	-----	-----	------



PATL3\_Oryza\_sativa\_Japonica (613 aa)

MAEFPQPEAAPAAVAATTEVAVAEKAPVEAEKEREVEESTPAVEAEAKEEKKDEAAAAAAGGDEA  
 GAIEGTGSGFKKEENLVADLPDPEKALDEFKQLIAAALAACEFNLPFFFFFFAKVEAAVEETKAE  
 ETKAEETPKAEPEAKEEPEKAEVAAAAAAPPKAGTEEPKAEASSEEAKTEEPKAEAAADEPAKEES  
 KAEEAPAEKAPAEPEPEENTVVVTEEEAATKTVEAIEETVVVPAAAAAPAAAAATEEAAPEPEVQAA  
 AAPEPVLINGVPLVGGDDERTDTVLLKFLRAREPFVKEAMAMLRGAVLWRRKRFGIESLLDADLALPE  
 LDSVVFYRGADREGHFVVCYNVYGEFQDEDLYEKAFCDDEEKRRERFLWRIQLLERGILGQLDFSPSG  
 ICSMVQVTDLKNSPFMLGKRAVTRQAVALLQDNYPEFIAKRVFINVPWYLAANKMMSPFLLTQRT  
 KSKFIFASPAKSAETLFRYIAPEQVFPVQGGFLKREDDPEFTTSDAVTELTIKPSKETVEIPVTEN  
 STIGWELRVLGWEVSYGAEFTPDAGCCTTVIVQTRKVPANEKPIKMGSPKVGEPGRIVLTINNPAS  
 SKKKLLYRSKVESTSEV



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

324 - 499: score = 17.735

ADLALPELDSVVFYRGaDREGHFVVCYNVYGEFQDEdlyekafgDEEKRRERFLWRIQLLE  
 RGILGQLDFSPSGICSMVQVTDLknsPFMLGKH---RAVTRQAVALLQDNYPEFIAKRV  
 FINVPWYLAANKMMSPFLLTQRTKSKFIFASPAKSAETLFRYIAPEQVFPVQGGFLKREDD

Predicted feature:

DOMAIN	324	499	CRAL-TRIO
--------	-----	-----	-----------

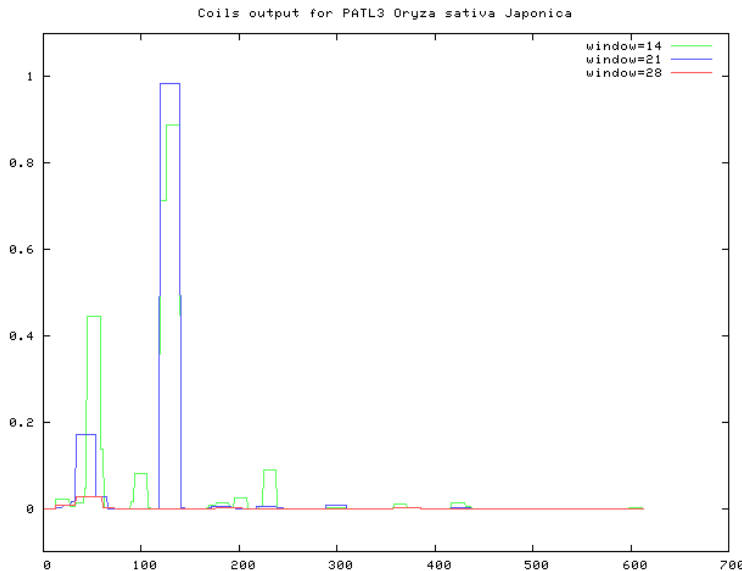
**PS50866 GOLD** GOLD domain profile :

505 - 606: score = 10.430

SDAVTELTIKPSKETVEIPVTENSTIGWELRVLGWEVSYGAEFTPDaeGCCTTVIVQtr  
 kvpANEKPIKMGSPKVGEPGRIVLTINNPASKKK---LLYRSKV

Predicted feature:

DOMAIN	505	606	GOLD
--------	-----	-----	------



PATL3\_Theobroma\_cacao (607 aa)

MADETPAPAPPVAAPFPFPPPPAVTESDGPVVEKEDPSPSPVDVVGVTVTESVSTAIEEKEBQT  
 PAAEFVEVVDGTDKGEHEEPLFPFPQPVKSESEGPVTEKAKHEEPPAVTEKAKHEEPPAATEKA  
 KEPPAATVAVLETELQPPPELPPQEVVTESEKSLAAMHEKEEGCAPEPTVATTTTGAEEGVAVVV  
 EEKIPQNLGSPFEESNEVADLSDFERKALEELKHFVQEAIDTHLFTGETRSEENPEREKKEEPE  
 VSWNGIPLLEDSDRSVDVILLKFLRARDPFRVDAFVMIKNTIRWRKEFCIDELLEDLGDDEMERVVFH  
 HGQDRECHPVCYNVYGEFQNEKLYQKAPSDDEEKRMKFLRWRIQFLEESIRKLDPSPDGINTIPQVS  
 DLKNSPGKRELRLATRQALQLLQDNYPFVAEQVFINVFWYLVFYTMISPFMTQRTESEKVFVA  
 GPASAEITLFRYMAPEQVPIQYGGLSVDYDCNPEFGDADPATEITVKEPGTQVVEITIIYERCDLV  
 WEIRVVGWVSVYGAEFMPNARDGYTVIIQKPTKMSPFDEPVVSGCFRVLGELGKVVLLTVDNPTGKKK  
 KLVYRFEVRFHCD



PS50191 CRAL\_TRIO CRAL-TRIO lipid binding domain profile :

317 - 492: score = 18.252

DEDLGDDEMERVVFHMGqDRECHPVCYNVYGEFQNEKdlyqkafeDEEKRMKFLRWRIQFLE  
 ESIRKLDPSPDGINT-IPQVSDLKnsPGKREL---RLATRQALQLLQDNYPFVAEQV  
 FINVFWYLVFYTMISPFMTQRTESEKVFVAGPAKSAEITLFRYMAPEQVPIQYGGLSVDYC

Predicted feature:

DOMAIN	317	492	CRAL-TRIO
--------	-----	-----	-----------

PS50866 GOLD GOLD domain profile :

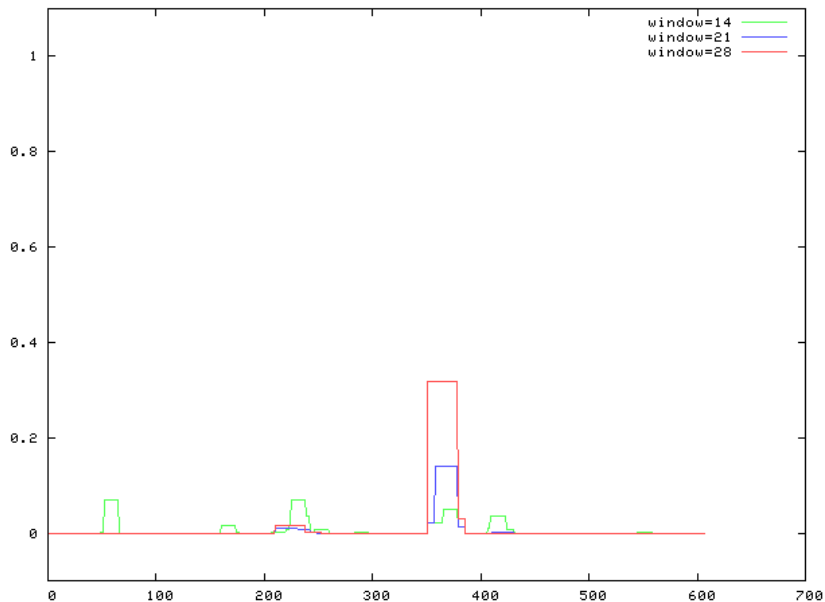
492 - 602: score = 9.899

DCNPEFGDadpateitvkePGTQVVEITIIYERCDLVWEIRVVGWVSVYGAEFMPNARDG  
 YTVIIQkptkmsPFDEPVVSGCFRVLGELGKVVLLTVDNPTGKKK--LVYRFEV

Predicted feature:

DOMAIN	492	602	GOLD
--------	-----	-----	------

Coils output for PATL3 Theobroma cacao



PATL3\_Vigna\_angularis (476 aa)

MAENDGNPPPPPPFAAGSPQEPFVPPSLGGENDAKADDVIEEKVFAEEVLVAEDKDEVKGGEEAVV  
 VEEKEKEEEKEEEKEKEKEKEEVEVTVGETESBSLKEESNRVSDGERKCIIEELKRVVREELERKEEG  
 EVSINCVPLPKDQRTDVIILKFLRARDLRVNDGLVMIQNTLRWRKEFGIDALLEEDLGEELERVVV  
 MHGQGREGHPVCYNVYGEFQNKELYQKAFSTEENRTKFLRWRIQLLERSIRNLDFTPGGINTIPQV  
 NDLEKSGPAPAKRELRIATKQALQLLQDNYPEFVARQVFINAPWYLAFNTLISPFLLTPRTEKRFIF  
 AGPSKSPDTLFRYISPEQVQVYGGLSVDFCDNPDFTMSDPVTEIPIKPTTEQTVETIAIYERCI  
 VWELRVVGVHEVTYNAEFKPAKDAYTVVIQKATKMSPTDEPVVSNKFKVGEKGLLLTIDNPTLKN  
 RLLYRFNIRPYCD



PS50191 CRAL\_TRIO CRAL-TRIO lipid binding domain profile :

186 - 355: score = 18.433

EEDLGEELERVVVFMHGQGREGHPVCYNVYGEFQNKELYQKAFSTEENRTKFLRWRIQLLE  
 RSIRNLDFTPGGINT-IFQVNDLkncPGPAKREL---RIATKQALQLLQDNYPEFVARQV  
 FINAPWYLAFNTLISPFLLTPRTEKRFIFAGPSKSPDTLFRYISPEQVQVYGG-----

Predicted feature:

DOMAIN	186	355	CRAL-TRIO
--------	-----	-----	-----------

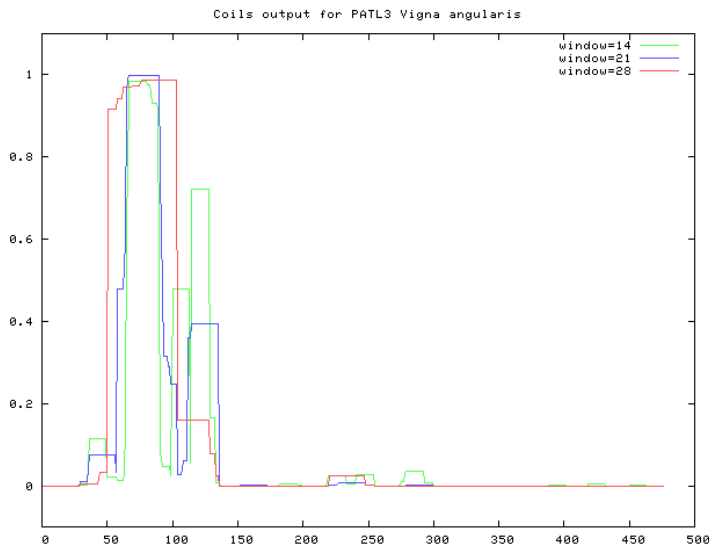
PS50866 GOLD GOLD domain profile :

366 - 471: score = 8.607

DFTMSDPVTEIPIKPTTEQTVETIAIYERCIIVWELRVVGVHEVTYNAEFKPAKDAYTVVI  
 QkatkmsPTDEPVVSNKFKVGEKGLLLTIDNPTLKNR--LLYRFNI

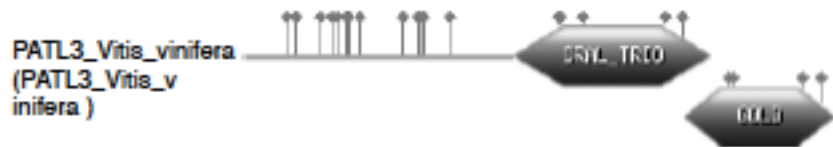
Predicted feature:

DOMAIN	366	471	GOLD
--------	-----	-----	------



PATL3\_Vitis\_vinifera (530 aa)

MADETDAAPPTLLAAASVVPETAAPPETAASEEVLTVLDVAPPVTEKEVSPQKPPPPPEVVAV  
 VDGEKVPQNLVSPKEESNRLADLSEERRALEELEKQSVQAEALRNGIPTGQPQPPPPPPQSAEK  
 PPERIEEAGEKREPFPVAESEISTQEEGANDENVKPTPNPTIESILKHEGPTQEDVSIWGIPLLED  
 ERSDMILLKFLRAREFPVKEAFAMLENTIFWRKEFGIDALVDDDLGHELEKVVFMHGFDRDGHFVC  
 YNVYGEFQNKELYQKTPSDEEKRMFLRWRIQFLERSIRKLDPTPGCVNTIFQVNDLENSPGPGKW  
 ELRQATEQALQLLQDNYPFVARQVFINVFWYLAFYMMISPFLLTQRTESKFPVFPASPAESARTLFFK  
 YISPEQVPIQYGGLSVDYCDCHPDFGIADPVTEITVKPSTEQTVEILVSEQCVIVNEVRRVVGNEVA  
 YCAEFIPDAEDEYTVVVQKATKMAPTDDPVMCNSFKIRELGRIVITIDNPTSEKKLLLYRFRVFPY  
 SN



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

240 - 415: score = 18.825

DDDLGHELEKVVFMHGFDRDGHFVCYNVYGEFQNKelyqktfSDEEKRMFLRWRIQFLE  
 RSIRKLDPTPGCVNT-IFQVNDLENSPGPGKWEELRQATEQALQLLQDNYPFVARQVFIN  
 VFWYLAFYMMISPFLLTQRTESKFPVFPASPAESARTLFFKYISPEQVPIQYGGLSVDYC

Predicted feature:

DOMAIN 240 415 CRAL-TRIO

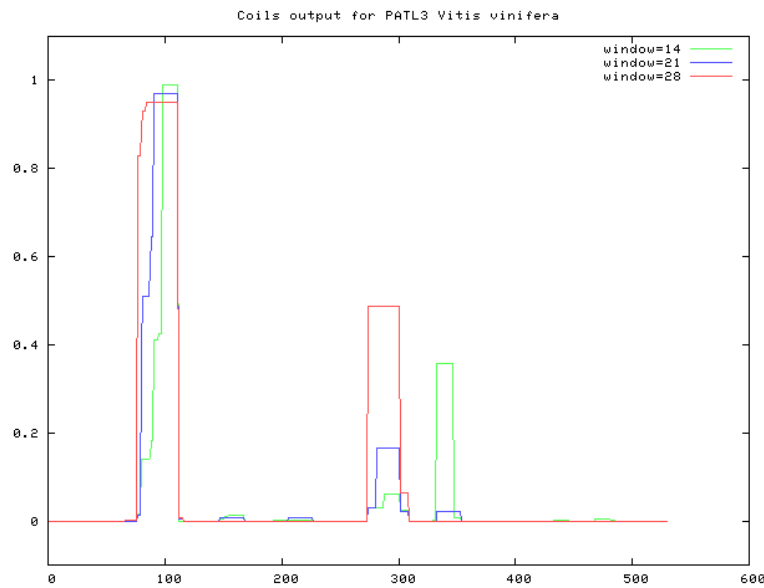
**PS50868 GOLD** GOLD domain profile :

391 - 525: score = 9.114

ARTLFFKYISpeqvpIqygglsvdyedonpdfgiadpvtEITVKPSTEQTVEILVSEQCVI  
 VNEVRRVVGNEVAYCAEFIPDAeDEYTVVVQkatkmaPTDDPVMCNSFKIRELGRIVITID  
 NPTSEKKK--LLYRFRV

Predicted feature:

DOMAIN 391 525 GOLD



PATL3\_Zea\_mays (556 aa)

```

MAESTQPEAAAEEAEVVVTEPAPAPAEAEVFAAADAEAVTETETEKADPAVTTADDAGKGTGSGFKEES
NLVEDLDPDEKRVLDPEFKHLIAAALAAGEFNLPPPPPPFKAKEEPKAEETRTEESKTEDPAEEEPK
AEGAAEPPFAEVAANAPDEEVVTEVPPVEEAARAEVAVEAKPAEPEPQERTVVVVAEEEPATKTVEA
IEEGVVGADEAAAPEPVLVNGVPLVGGDDERTDTVLLKFLRAREFVKEAMAMLESAVLWRRKPGIT
SLLDADLGLPELENNVVFYRGADREGHPVCYNVYGEFQDKDLYEKAFGDDEKRRERFLEWRIQLLERG
ILSKLDFSPNGICSMVQVTDLENSPPMLGRHRAVTRQAVTLLQDNYPEPIAKKVFINVPPWYLAAN
KMMSPFFTQRTSKFVFPASPAKSAETLFRYIAPBQVPVQFGGLFKEDDPEPTTLDTVTELTIKPSS
KETIEIPVTENSAIVWELRVLGWEVSYSAEFTPDTEGGYTVIIQKTRKVPANEPIPIMKGSFVVGEP
GKLVLTVNNPASKEKLLYRSEKRSISE
    
```



**PS50191 CRAL\_TRIO** *CRAL-TRIO lipid binding domain profile :*

269 - 444: score = 17.357

```

ADLGLPELENNVVFYRGaDREGHPVCYNVYGEFQDKDlyekafgDDEKRRERFLEWRIQLLE
RGIL-----SKLDFspNGICSMVQVtdlknspplgKRAVTRQAVTLLQD
NYPEPIAKKVFINVPPWYLAANKMMSPFFTQRTSKFVFPASPAKSAETLFRYIAPBQVPV
QFGGLFKEDD
    
```

Predicted feature:

DOMAIN	269	444	CRAL-TRIO
--------	-----	-----	-----------

**PS50866 GOLD** *GOLD domain profile :*

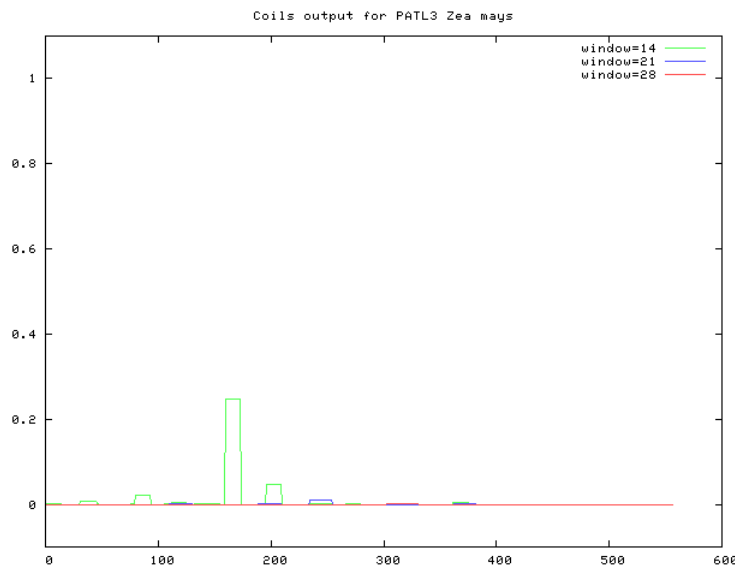
446 - 551: score = 10.683

```

EPTTLDTVTeLTIKPSKETIEIPVTENSAIVWELRVLGWEVSYSAEFTPDTeGGYTVII
QktrkvpANEPIPIMKGSFVVGEPGKLVLTVNNPASKEK--LLYRSEV
    
```

Predicted feature:

DOMAIN	446	551	GOLD
--------	-----	-----	------





PATL1\_Arabidopsis\_thaliana (573 aa)

MAQEEVQKSGADVAAAPVVKPKITDKVETIPTVVAEKKEEVAAPVSGDEKAVPEKEVTPKEKAPAAEA  
 ERGVSVKKEESTVVVVAERVVVLTAEVQKHALKEEPKELVREALNKREPTAPVTPVKEEKTEEKTEE  
 ETKEEKTEEKKEETTTEVKVEEEKPAVPAEEEEKSSEAAPVETKSEKPEEKAEVTTTERAGSAEE  
 DGTETVEAIEESIVSGPPEGAVAPVVVETVAVAEAEPEVEEVSIVGVPLLQDERSDVILTRFLR  
 ARDFVKREALTMLNNTVQNRKENKIDELVESGEEVSEFERMVFABGVDRKGVVVIYSSYGEFQNK  
 LPSDKEKLNKFLSWRIQLQEKCVRAIDFSNPEAKSSPVFVSDFRNAPGLGKRALWQFIRRAVEQFE  
 DNYPEFAAKELFINVPWWYIPYYTFCGSIITSPRTRSKMVLGAPSKGADTIFFKXIAPEQVPVRYGG  
 LSKDTPLTEETITEAIVKPAANYTIELPASEACTLSWELRVLGADVSYGAQFEPTEGSSYAVIVSK  
 TRKIGSTDEPVIITDGFVVGEPGRIVITIDNQTSKKEKVLRYRFTQ



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

295 - 468: score = 15.484

SGEEVSEFERMVFABGVDRKGVVVIYSSYGEFQNKelfsDKEKLNKFLSWRIQLQEKCVR  
 AIDFSNPEAKSSPVFVSDFRnaPGLGKRAL---WQFIRRAVEQFEDNYPEFAAKELFINV  
 PWWYipYYTFCGSIITSPRTRSKMVLGAPSKGADTIFFKXIAPEQVPVRYGGLSKDTP

Predicted feature:

DOMAIN	295	468	CRAL-TRIO
--------	-----	-----	-----------

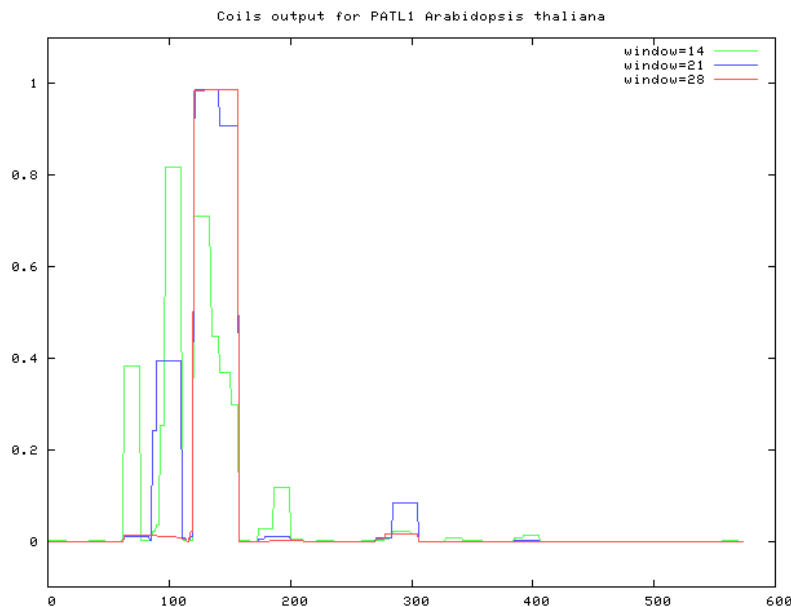
**PS50866 GOLD** GOLD domain profile :

471 - 572: score = 13.198

EETITEAIVKPAANYTIELPASEACTLSWELRVLGADVSYGAQFEPTEGSSYAVIVsktr  
 kIGSTDEPVIITDGFVVGEPGRIVITIDNQTSKKEK---VLYRFTQ

Predicted feature:

DOMAIN	471	572	GOLD
--------	-----	-----	------



PATL2\_Arabidopsis\_thaliana (683 aa)

MAQEELIQPTASVPVVKKEETPAPVKEVEVPVTTTERAVAAPAPEATEEERVVSEVAVPTEVTAVKEE  
 EVATGKEILQSEGFKEEGYLASELQBAEKNALAELKELVREALNREPTAPPPPPAPVVEEKVVEE  
 RTEETEEKKEEVTEEEKSLEAETKEEEKSAAPATVETKKEEILAAPAPIVAETKKEETPVAPAPVE  
 TKPAAPVVAETKKEEILPAAPVTTTETVVEEKVVPVETTPAAPVTTTKEEERAAAPVTTTKEEERAA  
 APGETKKEEKATAGTQVKRASFKIDIFVGVVTPGKKKKEEKPAVVTIKAFPAADQEEETRTVEAV  
 EESIVSITLPEATAAYVEPEEVSINGIFLLEDERSDVILLKFLRARDFFVKEAFTMLKNTVQWRKEN  
 RIDDLVGEDLEGSEFERLVPTHGVDKQGHVVIYSSYGEFQNKKEIFSDKEKLSKFLKWRIQPQEKCV  
 RSLDFSPKAKSSFPVVSDFRNAPGLGQRALWQFIKRAVQFEDNYPEFVAKELFINVPWWYIPTYK  
 TFGSIIISPRTRSEMVLGSPKSAETIFRYVAPEVVPVRYGCLSEDSPPTVEDCVTEAVVVKSTKY  
 TIDLPAEGSTLSWELRVLGADVSYGAQFEPSEASYYTVIVSENRKRVGLTDEPVIIDGFRASEAGK  
 VVITIDNQTFKRRKRVLYRSKTQA



**PS50191 CRAL\_TRIO** *CRAL-TRIO lipid binding domain profile :*

404 - 576: score = 14.463

EDLEGSEFERLVPTHGVDKQGHVVIYSSYGEFQNKKEIFSDKEKLSKFLKWRIQPQEKCV  
 SLDSPKAKSSFPVVSDFRNAPGLGQRAL---WQFIKRAVQFEDNYPEFVAKELFINV  
 PWWYIPTYKTFGSIISPRTRSEMVLGSPKSAETIFRYVAPEVVPVRYGCLSEDSP

Predicted feature:

DOMAIN	404	576	CRAL-TRIO
--------	-----	-----	-----------

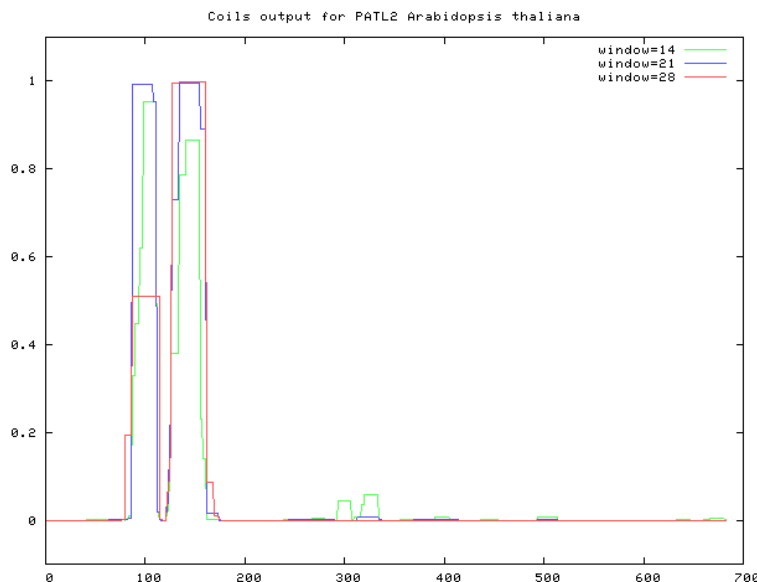
**PS50866 GOLD** *GOLD domain profile :*

580 - 681: score = 10.637

EDCVTEAVVVKSTKSYTIDLPAEGSTLSWELRVLGADVSYGAQFEPSEASYYTVIVSENR  
 kvGLTDEPVIIDGFRASEAGKVVITIDNQTFKRRK--VLYRSKT

Predicted feature:

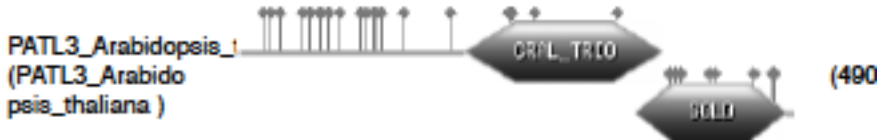
DOMAIN	580	681	GOLD
--------	-----	-----	------





PATL3\_Arabidopsis\_thaliana (490 aa)

MAEEPTTTTLVTPKLPSPGLTPSEVSESTQDALPTETETLEKVTETNPPETADTTTRPEEETAAB  
 HHPPTVTETETASTERQEVKDEASQKEVAEERKSMIPQNLGSPFEESSKLSDLNSERKSLDELKH  
 LVREALDNBQFTNTPEEVKIWGIPLLEDDRS DVVLLKFLRAREPFVKDGFAMLENTIEWRKEFKID  
 ELVEEDLVDDLDKVVFMHGHDRECHPVCYNVYGEFQNKELYNKTFSDDEERKHFRLRTRIQFLERSI  
 RRLDFSSGGVSTIFQVNDMKNSPGLGKRELRSATEQAVELLQDNYPEPVFRQAFINVPWWYLVFYT  
 VIGPFMTFRSKSLVFAGPFRSABTLFRYISPEQVPVQYGGLSVDPDCDCNPDFSLEDGASEITVKP  
 GTRQTVETIIIEKCELVWEIRVTGWEVSYKAEFVPEEKDAYTVVIQKPRKMRPSEDPVLTBSFKVN  
 ELGKVVLLTVDNPTGKREKLVYRFNVEPL



PS50191 CRAL\_TRIO CRAL-TRIO lipid binding domain profile :

202 - 377: score = 19.650

EEDLVDDLDKVVFMHGHDRECHPVCYNVYGEFQNKELYNKTFSDDEERKHFRLRTRIQFLERSI  
 RRLDFSSGGVSTIFQVNDMKNSPGLGKRELRSATEQAVELLQDNYPEPVFRQAFINVPWWYLVFYT  
 VIGPFMTFRSKSLVFAGPFRSABTLFRYISPEQVPVQYGGLSVDPDCDCNPDFSLEDGASEITVKP

Predicted feature:

DOMAIN	202	377	CRAL-TRIO
--------	-----	-----	-----------

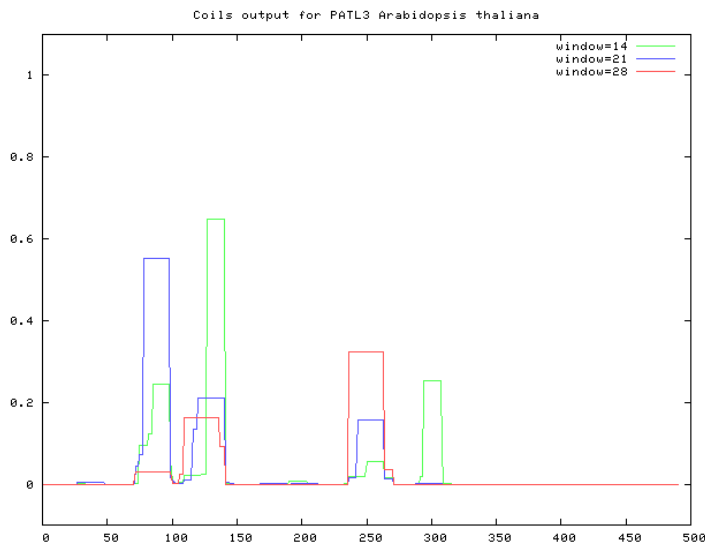
PS50866 GOLD GOLD domain profile :

353 - 487: score = 10.960

AETLFRYISPEQVPVQYGGLSVDPDCDCNPDFSLEDGASEITVKPGTRQTVETIIIEKCEL  
 VWEIRVTGWEVSYKAEFVPEEKDAYTVVIQKPRKMRPSEDPVLTBSFKVNELGKVVLLTVD  
 NPTGKREKLVYRFNVEPL

Predicted feature:

DOMAIN	353	487	GOLD
--------	-----	-----	------



PATL4\_Arabidopsis\_thaliana (540 aa)

MTAEVKEEIQVESEVVIAPAVVPEETTVEAVVEETKVEEDESQKPEGVERGASPKESGDFPADLKE  
 SEKRALSDLKSLKEEAIVDNTLLKTRKKEGSPMKERKEEVVPEAEVEKKEEAABEKVEEKKSE  
 AVVTEEAPKAEVTEAVVTEEIIPKEEVTTVVEKVEEETKEEERKTEDVVTEEVKAEITIEVEDEDES  
 VDNDIELWCVPLLPKRGAESTDVILLKFLHARDPEVNEAFEMLRKTLKWRKQNKIDGILGEEFGED  
 LATAAYMNGVDRESHPVVCYNVHSEELYQTIGSEKNREKFLRWRPQLMEEGIQKLNLEKPGCVTGLLQ  
 IHDLEKAPCVSRTEIIVWGIKKEVIETLQDNYPEFVGRNIFINVPFWFYAMRAVLGPFLLTQRTESKPV  
 VARPAKVRRETLLEYIPADELPVQYGGPRTVDDTEPQNETVSEVVVVRPGGSETIEIPAPETEGLVW  
 DIAVLGWEVNYKEEFVPTTEGAYTVIVQKVRKMGANECPIRNSPFNSQAGKIVLTVDNVSGKKEKV  
 LYRYRTETEGSS



PS50191 CRAL\_TRIO CRAL-TRIO lipid binding domain profile :

258 - 428: score = 18.727

GEEFGEDLATAAYMNGVDRESHPVVCYNVHSEELYQTIGSEKNREKFLRWRPQLMEEGIQK  
 LNLEKPGCVTGLLQIHDLEKAPCVSRTEIIVWGIKKEVIETLQDNYPEFVGRNIFINVP  
 FWFYAMRAVLGPFLLTQRTESKPVVARPAKVRRETLLEYIPADELPVQYGGPRTVDD

Predicted feature:

DOMAIN	258	428	CRAL-TRIO
--------	-----	-----	-----------

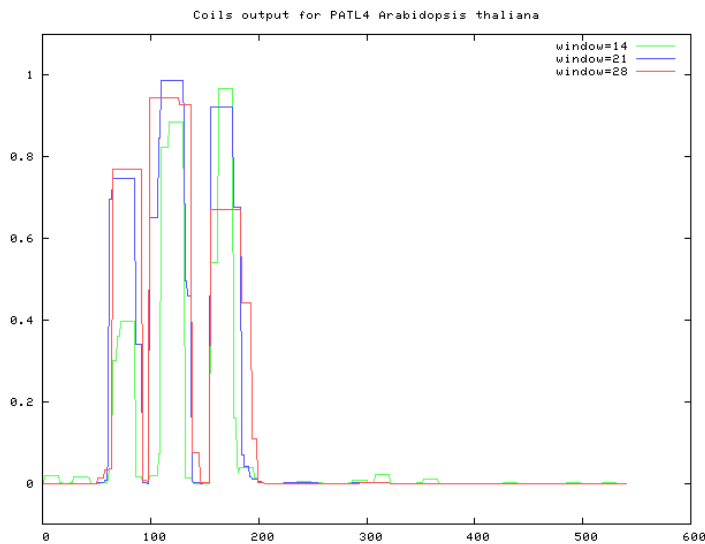
PS50866 GOLD GOLD domain profile :

433 - 534: score = 10.430

NETVSEVVVVRPGGSETIEIPAPETEGLVWDIAVLGWEVNYKEEFVPTTEGAYTVIVQKVR  
 KMGANECPIRNSPFNSQAGKIVLTVDNVSGKKEK--VLYRYRT

Predicted feature:

DOMAIN	433	534	GOLD
--------	-----	-----	------



PATL5\_Arabidopsis\_thaliana (668 aa)

```

MSQDGATTTFFPPLTGDVSMPSGGEDEPRHVTGEEBAPVTGETNLKLPMPLEESNHTAEVVSSEK
VTPETMTLESEGLNHAADGSEQTHEVTPETETAKLEVLNHTAEDGSEQTHEVTPERETVKEEFLNHV
AEDGSEQTHEVTPETETVKEVLNHAADGSEQPRCVTPTPETETSEADTGLLVTSGETEFPNHAADY
SETEPSQKLMLEQRRKRYMEVEDWTEPELFPDEAVLEAAASVPEPKQPEPQTTPPPPTTTSTVAASR
LAEMNREEAEVEERQKIQIPRSLGSPREETNKIGDLSETELNALQELRHLLQVSDSSKTSINGV
PLLEDDRDTDVVLLKFLRARDPKQEAYSMLNKTQWRIDFNIEELLDENLGGDLDKRVVFMQGDKE
NHPVCYNVYGFQNKEDLYQKTFSEDEKRERFLRWRIQFLERSIRNLDFVAGGVSTICQVNDLKNSP
GPGKTELRLATKQALHLLQDNYPEFVSEQIFINVPWYLAFYRIISPFMSQRSKSLVFAKPSRGA
ETLLRYISPEHVFPVQYGGLSVDNCECNSDFTDDIATEITVKEPTTKQTVEIIVYERCTIVWEIRVV
GWEVSYGAEFVPEKEGYTVIIQKPRKMTAKNELVVSHSFPVGEVGRILLTVDNPTSTKRMLIYRF
KVEPLACE
    
```



PS50191 CRAL\_TRIO CRAL-TRIO lipid binding domain profile :

377 - 552: score = 17.958

```

DENLGGDLDKRVVFMQGDqDENHPVCYNVYGFQNKEdlyqktfSEDEKRERFLRWRIQFLER
SIRNLDFVAGGVST-ICQVNDLknsPGPGKTEL---RLATKQALHLLQDNYPEFVSEQI
FINVPWYLAFYRIISPFMSQRSKSLVFAKPSrSAETLLRYISPEHVFPVQYGGLSVDNC
    
```

Predicted feature:

DOMAIN	377	552	CRAL-TRIO

PS50866 GOLD GOLD domain profile :

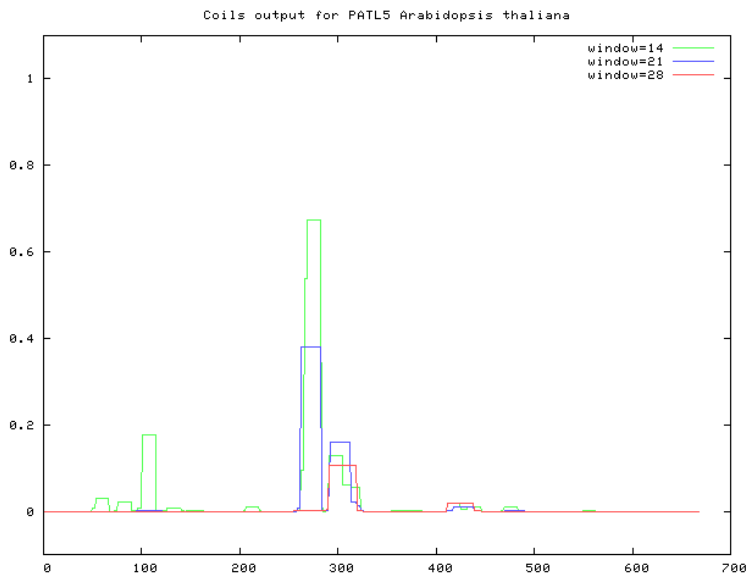
556 - 662: score = 10.937

```

SDFTDDIATEitvkPTTEQTVEIIVYERCTIVWEIRVVGWEVSYGAEFVPEKEGYTVI
IQkprkmtAKNELVVSHSFPVGEVGRILLTVDNPTSTKRM--LIYRFV
    
```

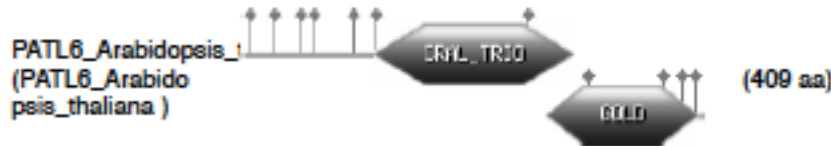
Predicted feature:

DOMAIN	556	662	GOLD



PATL6\_Arabidopsis\_thaliana (409 aa)

MDASLSPFDHQKTKQTEPKKSPITSLITLRSNNIKEDTYFVSELEKPTBQKSLQELKELGASSSKA  
 SSMNGVSLGGDDADVILLKFLRRADFVADSLRMLKCLEWREEFPAEKLTEDLGFRDLEGEV  
 AYMRGYDREGHFVVCYNAYGVFKEKEMHYERVFGDEEKLNRFLWRVQVLERGVNMLHFFKPGGVNSII  
 QVTDLEKDMFKRELRVASNQILSLFQDNYPELVATRIFINVWPYFVSVIYSMPSPFLTQRTKSKFVMS  
 EGNAAETLYKFIKPEDIPVQYGGLSRPTDSQNGPPKPAEFKSGGKRVNIQIEGIEGGATITWD  
 IIVVGGWDLEYSAEFVFNAAEESYAIIVVEKPKKMKATDEAVCNSTTVEAGKLILSDNTLSRKKKVA  
 AYRYTVRSTTV



**PS50191 CRAL\_TRIO** CRAL-TRIO lipid binding domain profile :

116 - 294: score = 17.511

EKLTEEDLGFKDLEGEVaymrgyDEEGHPVCYNAYGVFKEKemyervfgDEEKLNRFLRW  
 RVQVLERGVNMLHFFKPGGVNS-IIQVTDLEKDMFKREL---RVASNQILSLFQDNYPELVA  
 TRIFINVWPYFVSVIYSMPSPFLTQRTKSKFVMSKGNAAETLYKFIKPEDIPVQYGGLSR  
 PTD

Predicted feature:

DOMAIN	116	294	CRAL-TRIO
--------	-----	-----	-----------

**PS50866 GOLD** GOLD domain profile :

270 - 404: score = 11.560

AETLYKFIKpedipvqygglsrptdsqngppkpaefskGGKRVNIQIEGIEGGATITW  
 DIVVGGWDLEYSAEFVFNAAEESYAIIVVEKPKKMKATDEAVCNSTTVEAGKLILSDNTLS  
 RKKKVAAYRYTVRSTTV

Predicted feature:

DOMAIN	270	404	GOLD
--------	-----	-----	------