

Motivos, patrones y perfiles

Representación de Alineamientos Múltiples,
Homologías remotas y
Bases de datos secundarias

Esquema

- Representación de alineamientos múltiples
- Bases de datos secundarias o de patrones
- PROSITE

I- Representación de los AMS

- Los AMS se utilizan para representar o caracterizar familias de secuencias relacionadas.
- No resulta práctico trabajar directamente con los AMS por lo que se han desarrollado diversas maneras de representarlos.
- Los distintos métodos de representación de AMS forman una *jerarquía de modelos*: cada método es un caso particular del que le sigue en complejidad.

Una jerarquía de modelos para AMS

- Hay muchos métodos
 - **Secuencia exacta**
 - **Secuencias consenso**
 - **Expresiones regulares o patrones**
 - Perfiles o Matrices de pesos posicionales
 - Modelos ocultos de Markov
- Cada modelo generaliza al anterior a la vez que resulta algo más flexible pero también, posiblemente, más complejo.

Motivos señales o patrones

- Consideremos un alfabeto como el del ADN o las proteínas.
- Un *motivo* (*patrón* o *señal*) es una forma de caracterizar un conjunto de secuencias de este alfabeto.
- Dada una secuencia, S , y un motivo M diremos que M está presente en S si cualquiera de las secuencias descritas por M ocurre en S .
- P.ej. $M="TATA"$, $S1="GATTACA"$ y $S2="PATATA"$
 M está presente en $S2$ pero no en $S1$

Aplicaciones de los Motivos

- Los motivos representan zonas conservadas entre las secuencias que suelen asociarse a *características funcionales del grupo de secuencias*.
- Una vez se ha construido un motivo o patrón de un grupo de secuencias puede utilizarse
 - Para asociar una nueva secuencia con la familia de secuencias que lo ha generado (si presenta el motivo es de la familia y puede que comparta sus funciones)
 - Para buscar secuencias que pertenezcan a aquella familia

Modelos para AMS y motivos

- Una manera natural de representar un AMS es a través de los motivos o patrones que contiene.
- La jerarquía de modelos para AMS a la que hemos hecho referencia es, pues, también una jerarquía de modelos para motivos: El patrón característico del alineamiento es el "motivo" que lo caracteriza.

Descripción de motivos (1)

Palabra exacta

- La manera más simple de describir un motivo contenido en un AMS es a través de la secuencia exacta de letras (la "palabra") que lo forman

– Muy preciso si se presenta pero no admite variaciones

G	A	T	T	A	C	A
G	A	T	T	A	C	T
G	A	T	T	A	C	T
G	A	T	T	A	C	A
G	A	T	T	A	C	C
	A	T	T	A	C	

G	A	T	T	A	C	A
G	A	C	T	A	C	T
T	A	T	T	A	C	T
C	A	T	T	G	C	A
A	A	T	T	A	C	C
	A	?	T	?	C	

Descripción de motivos (2)

La secuencia consenso

- Si en alguna posición aparecen cambios en la palabra exacta se pueden utilizar caracteres diversos para indicar estas variaciones.
- Por ejemplo
 - Si todas las secuencias tienen el mismo residuo en una posición dada se pone la letra mayúscula
 - Si la mayoría tiene la letra se pone minúscula
 - Si hay empate se ponen las letras empatadas

Un ejemplo de secuencia consenso

	1	2	3	4	5	6	7	8	9	10
I	Y	D	G	G	A	V	-	E	A	L
II	Y	D	G	G	-	-	-	E	A	L
III	F	E	G	G	I	L	V	E	A	L
IV	F	D	-	G	I	L	V	Q	A	V
V	Y	E	G	G	A	V	V	Q	A	L
	y	d	G	G	A/I	V/L	V	e	A	I

Descripción de motivos (3)

Expresiones regulares

- La secuencia consenso resulta útil si hay pocas variaciones.
- Buscando una mayor flexibilidad se propuso el uso de expresiones regulares
 - *Una expresión regular, a menudo llamada también **patrón**, es una expresión que describe un conjunto de cadenas sin enumerar sus elementos ([Wikipedia](#))*
 - Son ampliamente utilizadas en informática, en entornos UNIX/Linux especialmente, para manipular cadenas de caracteres de manera muy flexible.

Sintaxis de expresiones regulares

- Una expresión regular o patrón describe un conjunto de caracteres sin enumerarlos.
- Para ello utiliza algunos *constructores* cuya sintaxis puede cambiar entre lenguajes o S.Op.
- En general todo sistema de e.r. dispone de
 - *Comodines: Apto para cualquier caracter*
 - *Ambigüedades: Se presenta/prohíbe varios caracteres*
 - *Factores de repetición: Número de veces que se presenta [o puede presentarse] un caracter*

Sintaxis de expresiones regulares

Caracteres comodín

- Si en una posición dada puede aparecer cualquier carácter se indica con el signo “comodín”
- Aunque en informática éste es a menudo un “*” aquí se utilizará una “x”

G	A	T	T	A	C	A
G	A	C	T	A	C	T
T	A	A	T	A	C	T
A	A	T	T	A	C	C
	A	x	T	A	C	

Patrón: A-x-T-A-C

Sintaxis de expresiones regulares

Ambigüedades

- Si en una posición dada puede aparecer varios caracteres distintos podemos indicarlo de dos formas
 - Aquellos que pueden aparecer: entre “[” y “]”
 - Aquellos que no se encuentran en la posición: entre “{” y “}”
- Una misma secuencia se puede indicar de maneras distintas. *P.ej: [ATC] equivale a {G}*

G	A	T	T	A	C	A
G	A	C	T	T	C	T
T	T	A	T	C	C	T
A	T	T	T	A	C	C
	[AT]	x	T	{G}	C	

Patrón: [AT]-x-T-{G}-C={CG}-x-T-[ATC]-C= ...

Sintaxis de expresiones regulares

Elementos repetidos

- La repetición de un elemento se indica con éste entre paréntesis: "("y")"
 - A(4) indica una "A" repetida 4 veces
 - x(3) indica un caracter cualquiera repetido 3 veces
 - Si el elemento que se repite es uno cualquiera ("x") puede asignarsele un número variable de repeticiones, incluso el cero
 - x(2-4): "x-x", "x-x-x", "x-x-x-x"
 - x(0-2): "", "x", "x-x"

Ejemplos (1|2)

- Las secuencias en negrita verifican el patrón

A. Patrón: [AC]-x-V-x(4)-{ED}	B. Patrón: A-x(0,1)-{V}
1. ACVCCCCE	1. ACA
2. CDVAAAAC	2. AE
3. ACVVVVVD	3. AVE
4. AVVVVVVV	4. AV
5. CEVACCAC	5. EA

Ejemplos (2|2)

- Generar el patrón de las secuencias siguientes

A. Secuencias:	B. Secuencias:
1. ACCM	1. SCCCMM
2. ADVMM	2. SCVMM
3. ACEM	3. TCAM
4. ACMM	4. TCMM
5. ACTM	5. TCSM
A-[CD]-x-M	[ST]-C-x(1-2)-M

Ejercicios (1|2)

- Indicar que secuencias verifican los patrones o explicar porque no lo verifican

A. Patrón: A-x-V(3)-{AL} Secuencias:	B. Patrón: [ST](2)-x(0,1)-V Secuencias :
1. AAVVVA	1. XSSV
2. AVVVAM	2. TSEV
3. ASVVVM	3. SAV
4. ACVVVC	4. TTVV
5. ACVVVE	5. TTTV

Ejercicios (2|2)

- Generar el patrón de las secuencias siguientes

A. Secuencias:	B. Secuencias:
1. MAS	1. CCCSAC
2. MCAT	2. CCVTC
3. MVAT	3. CCASSC
4. MAAS	4. CCMSC
5. MTAS	5. CCSTC

II- Homologías remotas

- Cuando realizamos búsquedas una secuencia con BLAST sucede a menudo que no se encuentran secuencias similares
- Esto no significa que *parte de la* secuencia no pueda estar conservada entre los miembros de una misma familia
- Un alineamiento múltiple puede revelar la existencia de patrones comunes a un grupo de secuencias.
- Esto se suele denominar “homología remota”, en contraposición a l'homología "normal" que se supone que existe entre secuencias con un alto grado de similitud.

Bases de datos de patrones

- Los patrones que se han obtenido de un AMS pueden actuar como "firmas" o características de las familias de secuencias de las que se han obtenido.
- Estos patrones pueden utilizarse para:
 - Decidir si una nueva secuencia pertenece a una familia que presenta el patrón que la caracteriza.
 - Buscar secuencias que contengan el mismo patrón que una dada
- La base de datos secundaria más importante es PROSITE

Predicción de función de una secuencia

- La asociación entre patrones y familias de secuencias nos facilita una aproximación a la predicción de la función de una secuencia
 - Si tenemos una secuencia de función desconocida...
 - Y esta secuencia contiene un cierto patrón ...
 - Que sabemos asociado a cierta familia de proteínas con una función concreta ...
 - Podemos deducir/predecir que ésta es la función que realiza también dicha secuencia

Un protocolo para la predicción de la función de una secuencia

- Deseamos conocer la función de una secuencia
 - Miramos, con BLAST si posee homólogos en una BD
 - Si es así podemos mirar de asociar su función con la de éstos
 - Si no posee homólogos cercanos podemos
 1. Realizar un AMS con las secuencias más similares a ella
 2. Extraer el patrón que las caracteriza
 3. Buscar en BD secundarias (PROSITE) si dicho patrón puede asociarse con una familia de función conocida
 - En la práctica podemos omitir (1) y (2) y someter directamente la secuencia a PROSITE

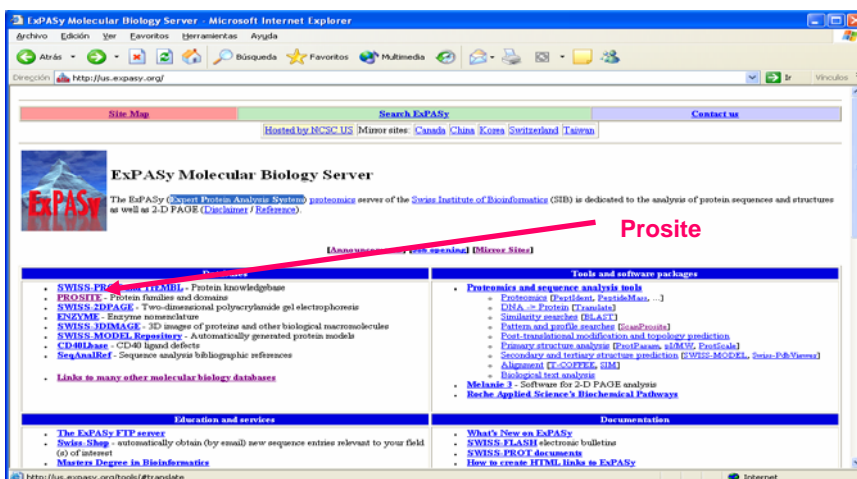
Más información sobre homologías remotas

- El concepto de homología remota en contraposición al de homología proxima o de alto grado de similitud está explicado en este tutorial:
[*Distant homologies: motifs, patterns, profiles*](#)
- Un artículo de revisión donde se desarrolla la relación entre homologías, patrones y bases de datos secundarias es el siguiente
[The role of protein databases in sequence analysis](#)

III- PROSITE

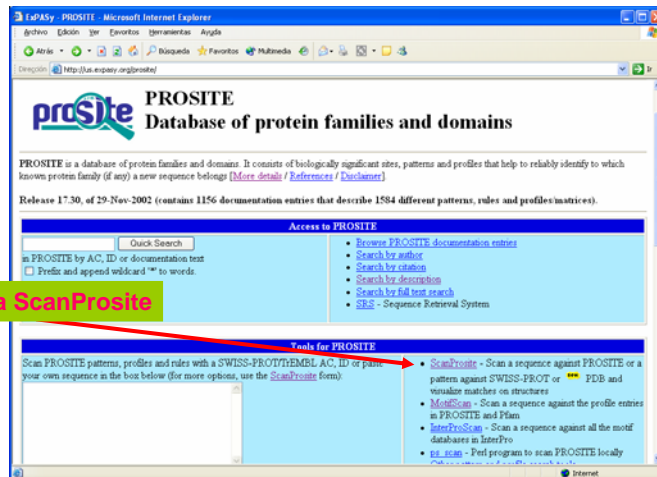
- Es una base de datos con patrones y dominios de proteínas. Desarrollada por ExPaSy ([Expert Protein Analysis System](#))

Una visita a ExPaSy: <http://us.expasy.org/>



Una visita a PROSITE

<http://us.expasy.org/prosite/>



¿ Qué contiene PROSITE?

- Contiene patrones de aminoácidos con significado biológico, dispuestos de manera que los programas informáticos (por ejemplo: *ScanProsite*) puedan determinar la familia de pertenencia (si se encuentra en PROSITE) de la nueva secuencia de manera rápida y eficiente.

Un ejemplo de entrada en PROSITE

NiceSite View of PROSITE: PS00706

General information about the entry

Entry name:	PRION_2
Accession number:	PS00706
Entry type:	PATTERN
Date:	DEC-1992 (CREATED); DEC-1992 (DATA UPDATE); JUL-1998 (INFO UPDATE)
PROSITE documentation:	PD002063

Name and characterization of the entry

Description: Prion protein signature 2.
Pattern: E-x[ED]-x-K-[LIVM]C-x-[KR]-[LIVM]C-x-[Q]-M-C-x(2)-Q-Y.

Numerical results

- SWISS-PROT release number: 40.7, total number of sequence entries in that release: 103373
- Total number of hits in SWISS-PROT: 44 hits in 44 different sequences
- Number of hits on proteins that are known to belong to the set under consideration: 44 hits in 44 different sequences
- Number of hits on proteins that could potentially belong to the set under consideration: 0 hits in 0 different sequences
- Number of false hits (on unannotated proteins): 0 hits in 0 different sequences
- Number of known named hits: 0
- Number of partial sequences which belong to the set under consideration, but which are not hit by the pattern or profile because they are partial (fragment) sequences: 0
- Precision (true hits / (true hits + false positives)): 100.00 %
- Recall (true hits / (true hits + false negatives)): 100.00 %

Comments

- Taxonomic range: Eukaryotes
- Maximum known number of repetitions of the pattern in a single protein: 1
- Interesting site in the pattern: 13,divalide

Cross-references

True positive hits:

PRIO_AOTTR (P40245), PRIO_ATRGE (P40246), PRIO_ATEPA (P21446),
PRIO_BOVIN (E10272), PRIO_CALJA (E40247), PRIO_CALMO (P40248),
PRIO_CARDA (E22444), PRIO_CARFA (E40249), PRIO_CARHE (E24443),
PRIO_CERAP (E40250), PRIO_CERAR (E40251), PRIO_CERAT (Q25142),
PRIO_CEREL (E22142), PRIO_CERHO (Q25143), PRIO_CERPA (Q25144),
PRIO_CERTE (Q25145), PRIO_CERIC (E22146), PRIO_COLOR (E40252),
PRIO_CREGR (Q25148), PRIO_CREMT (Q25149), PRIO_CRLCA (Q48254),
PRIO_GORGO (E40253), PRIO_HERAR (E04158), PRIO_MAFCA (E40254),
PRIO_MANSF (E40255), PRIO_HESAU (E04472), PRIO_MOUSE (P24222),
PRIO_PORPP (E42114), PRIO_PORPP (E40256), PRIO_POROW (E24223),
PRIO_PANTR (E40257), PRIO_PIG (E42227), PRIO_PORPP (E40258),
PRIO_PORPP (E40259), PRIO_PORPP (Q25111), PRIO_PAT (E40260),
PRIO_SARS (E40261), PRIO_SHEK (E22907), PRIO_SIGHT (Q25112),
PRIO_THERG (Q25170), PRIO_TRIUV (E34789), PRIO_TRAST (E40262),
PRIO_BOVIN (G04488), PRIO_TRAST (E40263)

Un ejemplo de entrada en PROSITE (continuación)

Cross-references

True positive hits:

PRIO_AOTTR (P40245), PRIO_ATRGE (P40246), PRIO_ATEPA (P21446),
PRIO_BOVIN (E10272), PRIO_CALJA (E40247), PRIO_CALMO (P40248),
PRIO_CARDA (E22444), PRIO_CARFA (E40249), PRIO_CARHE (E24443),
PRIO_CERAP (E40250), PRIO_CERAR (E40251), PRIO_CERAT (Q25142),
PRIO_CEREL (E22142), PRIO_CERHO (Q25143), PRIO_CERPA (Q25144),
PRIO_CERTE (Q25145), PRIO_CERIC (E22146), PRIO_COLOR (E40252),
PRIO_CREGR (Q25148), PRIO_CREMT (Q25149), PRIO_CRLCA (Q48254),
PRIO_GORGO (E40253), PRIO_HERAR (E04158), PRIO_MAFCA (E40254),
PRIO_MANSF (E40255), PRIO_HESAU (E04472), PRIO_MOUSE (P24222),
PRIO_PORPP (E42114), PRIO_PORPP (E40256), PRIO_POROW (E24223),
PRIO_PANTR (E40257), PRIO_PIG (E42227), PRIO_PORPP (E40258),
PRIO_PORPP (E40259), PRIO_PORPP (Q25111), PRIO_PAT (E40260),
PRIO_SARS (E40261), PRIO_SHEK (E22907), PRIO_SIGHT (Q25112),
PRIO_THERG (Q25170), PRIO_TRIUV (E34789), PRIO_TRAST (E40262),
PRIO_BOVIN (G04488), PRIO_TRAST (E40263)

Retrieve an alignment of SWISS-PROT true positive hits:

[Chatral format, color, condensed view] [Chatral format, color] [Chatral format, plain text] [Fasta format]

PDB
IDWY; IDWZ; IXXD; IXXI; IEIG; IEIJ; IEIP; IEIS; IEIU; IEIW; IOLX; IOLZ; IOMD; IOMI; IOM2; IOM3; IRID; IAOZ.

View entry in original PROSITE format
View entry in plain text format (no links)
Direct SwissProt submission

You would like to retrieve all the SWISS-PROT entries referenced in the DR lines of this entry (with the exception of false positive hits), you can enter a file name. These entries will then be saved to a file under this name in the directory out-going of the EMBL anonymous ftp server, from where you can download it. (Please note that this temporary file will only be kept for 1 week.)

File name:

Objetivo de PROSITE

- Su objetivo principal es determinar la función de nuevas proteínas no caracterizadas, cuyas secuencias se han obtenido por traducción de DNA genómico o cDNA y que son demasiado distantes de las ya conocidas para poder ser alineadas globalmente por los métodos ya vistos.

¿Cuándo utilizar PROSITE ?

- Cuando la nueva secuencia sea demasiado distante de las conocidas hasta el momento como para poder obtener un alineamiento global.
- En estos casos puede suceder que determinadas secuencias de aa se conserven debido a su importancia biológica (patterns, motifs, signatures, fingerprints). Por ejemplo en las zonas de unión de un enzima con un substrato, etc.

¿Cómo utilizar PROSITE ?

Básicamente consideraremos 3 posibilidades:

1. Introducir un código de identificación y que nos dé la información relacionada.
2. Introducir una secuencia de aminoácidos para que busque posibles patrones contenidos en ella.
3. Introducir un patrón y buscar proteínas que cumplan dicho patrón.

Identificación de la función de una secuencia de aa
con PROSITE (1)

- Supongamos que acabamos de obtener la siguiente secuencia humana de aa cuya función deseamos conocer

```
MANLGCWMLVLFVATWSDLGLCKRKPGGWNTGGSRYPGQGS PGGNRYPPQGGGGWGQP  
HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA  
VVGGLGGYMLGSAMSRP I IHFGSDYEDRYRENMHRYPNQVYYRPMDEYSNQN NFVHDCV  
NITIKQHTVTTTTKGENFTE TDVKMMERVVEQMCITQYERESQAYYQRGSSMVLFS SPPV  
ILLISFLIFLIVG
```

Identificación de la función de una secuencia de aa con PROSITE (2)

Entramos la secuencia

Scan a protein for PROSITE matches

Enter a SWISS-PROT/TrEMBL accession number (AC) (for example **P05130**) or a sequence identifier (ID) (for example **NOTC_DROME**), or paste your own **protein** sequence in the box below:

```
GQGSFGNRYPPGGGGGQOP
HGGGWGQPHGGGWGQPHGGGWGQGGGTHSQW
NKPSKPKTNMKHMAGAAAAGA
VVVGLGGYMLGSAMSRP I I HFGSD YEDRYRENMMHRYPN
QVYYRPMDEYSNQNFVHDCV
NITIKQHTVTTTTRKGFNFTDVKMMERVVEQMCITQYE
RESQAYYQGGSMVLFSSPPV
ILLISFLIFLIVG
```

and specify which motifs to use:

Scan patterns profiles rules [\[User Manual\]](#) (You may also specify a PROSITE entry in the box to the left)

Exclude patterns with a high probability of occurrence

Iniciamos la búsqueda

Your e-mail (optional): (will send results by e-mail)

plain text output

Identificación de la función de una secuencia de aa con PROSITE (3)

ScanProsite - Microsoft Internet Explorer

Search a sequence against PROSITE

Sequence:

```
RRLDGLMELV LPAATWDLG LCRDFEFDG WFTGGSEYDQ GSPFGSEVY PGGGGGQOP
HGGGWGQPHG GSGGPNQGG WQPHGGGQGG QGGGTRHQQN KPSKPKTNMK HMAGAAAAGA
VVVGLGGYML GSAMSRP I I HFGSD YEDRYRENMMHRYPN
QVYYRPMDEYS NQNFVHDCV
NITIKQHTVTT TTRKGFNFTDVKMMERVVEQMCITQYE
RESQAYYQGG SMVLFSSPPV
ILLISFLIFL IIVG
```

PROSITE Release 17.25, of 04-Nov-2002 (Excluding patterns with a high probability of occurrence and high threshold profile hits)

> **EP0000063** **P200291** **PRION_1** Prion protein signature 1 [pattern]

113 - 120 AGAAAAGAVVGGGGY

> **EP0000063** **P200706** **PRION_2** Prion protein signature 2 [pattern]

200 - 210 KLDVQDTEKPVQKCLLQY

Graphical summary of hits (Java applet)

Click on item to see a description. Drag the two red cursors to select a zoom region.

2 hits with 2 PROSITE entries

Ha encontrado 2 patterns

Posición de los patterns en la secuencia de aa.

Identificación de la función de una secuencia de aa con PROSITE (3)

- Los dos patrones encontrados están incluidos en la familia de los priones:

>[PDOC00263](#) [PS00291](#) **PRION_1** Prion protein signature 1
113 - 128 AGAAAAGAVVGGGLGGY (totalmente conservado)

>[PDOC00263](#) [PS00706](#) **PRION_2** Prion protein signature 2
200 - 218 EtDvKMMerVVeQMCitQY (no totalmente conservado)

Identificación de la función de una secuencia de aa con PROSITE (4)

The screenshot shows the PROSITE website interface in a Microsoft Internet Explorer browser window. The page title is "NiceSite View of PROSITE: PS00291". The main content area is titled "NiceSite View of PROSITE: PS00291" and contains the following information:

- General information about the entry:**
 - Entry name: **PRION_1**
 - Accession number: **PS00291**
 - Entry type: **PATTERN**
 - Date: **APR-1990 (CREATED); DEC-1992 (DATA UPDATE); JUL-1998 (INFO UPDATE)**
 - PROSITE documentation: [PDOC00263](#)
- Name and characterization of the entry:**
 - Description: **Prion protein signature 1**
 - Pattern: **A-G-A-A-A-G-A-V-V-G-L-G-Q-Y**
- Numerical results:**
 - SWISS-PROT release number: **40.7**, total number of sequence entries in that release: **103373**
 - Total number of hits in SWISS-PROT: **44 hits in 44 different sequences**
 - Number of hits on proteins that are known to belong to the set under consideration: **44 hits in 44 different sequences**
 - Number of hits on proteins that could potentially belong to the set under consideration: **0 hits in 0 different sequences**
 - Number of false hits (on unrelated proteins): **0 hits in 0 different sequences**
 - Number of known missed hits: **0**
 - Number of partial sequences which belong to the set under consideration, but which are not hit by the pattern or profile because they are partial (fragment) sequences: **0**
 - Precision (true hits / (true hits + false positives)): **100.00 %**
 - Recall (true hits / (true hits + false negatives)): **100.00 %**
- Comments:**
 - Taxonomic range: **Eukaryotes**
 - Maximum known number of repetitions of the pattern in a single protein: **1**
- Cross-references:**
 - True positive hits:

Identificación de la función de una secuencia de aa con PROSITE (5)

The screenshot shows the PROSITE database entry for PS00706, titled "PRION_2". The entry is displayed in a browser window with the following sections:

- General information about the entry:**
 - Entry name: PRION_2
 - Accession number: PS00706
 - Entry type: PATTERN
 - Date: DEC-1992 (CREATED), DEC-1992 (DATA UPDATE), JUL-1998 (INFO UPDATE)
- PROSITE documentation:** [\[1-3\]](#)
- Name and characteristics of the entry:**
 - Description: Prion protein signature 2
 - Pattern: E-x-[ED]-x-K-[LIVM](2)-x-[CR]-[LIVM](2)-x-[QR]-M-C-x(2)-Q-Y
- Numerical results:**
 - SWISS-PROT release number: 40.7, total number of sequence entries in that release: 103373
 - Total number of hits in SWISS-PROT: 44 hits in 44 different sequences
 - Number of hits on proteins that are known to belong to the set under consideration: 44 hits in 44 different sequences
 - Number of hits on proteins that could potentially belong to the set under consideration: 0 hits in 0 different sequences
 - Number of false hits (on unrelated proteins): 0 hits in 0 different sequences
 - Number of known missed hits: 0
 - Number of partial sequences which belong to the set under consideration, but which are not hit by the pattern or profile because they are partial (fragment) sequences: 0
 - Precision (true hits / (true hits + false positives)): 100.00 %
 - Recall (true hits / (true hits + false negatives)): 100.00 %
- Comments:**
 - Taxonomic range: Eukaryotes
 - Maximum known number of repetitions of the pattern in a single protein: 1
 - Interesting site in the pattern: 11444444
- Cross-references:**
 - True positive hits:
 - PRIO_AOTTE (E40245), PRIO_ATRGH (E40246), PRIO_ATEPA (E11445),
 - PRIO_BOVIN (E10223), PRIO_CALJA (E40247), PRIO_CALMO (E40248),
 - PRIO_CARDO (E23444), PRIO_CAFPA (Q16501), PRIO_CAFMI (E24443),
 - PRIO_CRRAP (E40249), PRIO_CRRAR (E40250), PRIO_CRRAT (Q25445),
 - PRIO_CRRSL (E23445), PRIO_CRRTO (Q25446), PRIO_CRRSA (Q25447),
 - PRIO_CRRTO (Q25448), PRIO_CRRIC (E27447), PRIO_COLOU (E40251),
 - PRIO_CRROR (Q60505), PRIO_CRRMI (Q60468), PRIO_PRLCA (Q18754).

Conclusión del ejemplo

- Parece bastante claro que se trata de la secuencia de aa correspondiente a un prión ya la búsqueda ha identificado 2 “patrones” ambos correspondiente a la familia de los priones.