Aplicaciones WEB

INTRODUCCIÓN

DEFINICIÓN DE SOFTWARE:

"Conjunto de requerimentos operacionales, especificaciones, código, guías, manuales y documentación de mantenimiento de un sistema basado en computadora." [Pressman 92]

"El Software es un programa, documento, procedimiento, y de rutinas asociadas con la operación de un sistema de cómputo". [http://coqui.lce.org/rquiles/Software.htm]

SUITE

Conjunto de programas con un objeto en común, que pueden operan entre si, compartiendo e intercambiando datos para resolver un problema específico

OpenOffice.org

Co
wit
Office XP
Office 2000
Office 97

Star Office 5.2

- Sun Microsystems
- Uno de los mayores proyectos de software libre.

OpenOffice cuenta con:

- Procesador de textos (Writer).
- Planilla de cálculos (Calc).
- Software de presentaciones (Impress).
- Software para dibujo (draw).

- Posibilidad de abrir y guardar documentos en formato Microsoft Office (97/2000/XP) y tiene soporte para 27 idiomas.
- Requiere mayor potencia de hardware

SUITE

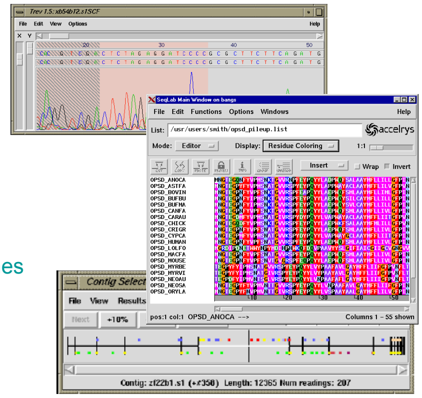
GCG Wisconsin Package

Genetics Computer Group

Seqlab, Línea de Comando, SeqWeb

PROGRAMAS:

Comparación
Búsqueda y recuperación desde Base de datos
estructura secundaria de DNA/RNA
Evolución
Ensamblaje de fragmentos
Búsqueda de Genes y reconocimiento de patrones
Mapeo
Selección de Primer
Análisis de Proteínas
Traducción


SeqLab Main Window on bangs
File Edit Functions Options Windows Help
List: /usr/users/smith/opsd_pileup.list
accelrys
Mode: Editor
Display: Residue Coloring
Insert
Wrap
Invert
pos:1 col:1 OPSD_ANOCA --->
Columns 1 - 55 shown
Contig Select
File View Results
+10%
Contig: zf22b1.s1 (+2350) Length: 12365 Num readings: 207

SUITE

Staden Package Program

ftp.mrc-lmb.cam.ac.uk/pub/staden/staden_package

Medical Research Council

Ensamblaje

Preparar datos de secuenciador para análisis o ensamblaje

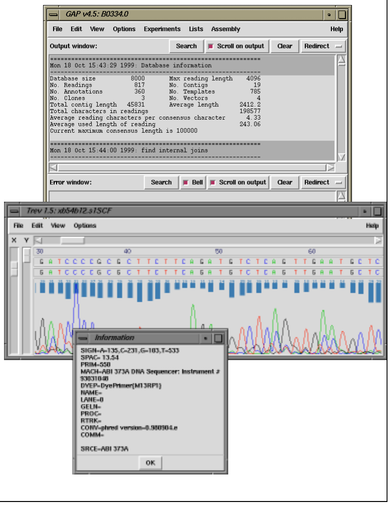
Detección de Mutaciones

Análisis de **Secuencias**

Manipulación de archivos de lectura y **Secuencias de** secuenciador

Any trace file

Archivos ABI, ALF, SCF


GAP v4.5: B0334.0
File
Edit
View
Options
Experiments
Lists
Assembly
Help
Output window:
Search
Scroll on output
Clear
Redirect
Error window:
Search
Bell
Scroll on output
Clear
Redirect
Trev 1.5: xb54b12.s1SCF
File
Edit
View
Options
Help
X
Y
30
40
50
60
G A T C C C C G C G C T T C T T C A G A T G T C T C A G T T G A A T G C T C
G A T C C C C G C G C T T C T T C A G A T G T C T C A G T T G A A T G C T C
Information
OK

```
================================================================
Mon 18 Oct 15:43:29 1999: Database information
----------------------------------------------------------------
Database size          8000       Max reading length    4096
No. Readings            817       No. Contigs             19
No. Annotations         360       No. Templates          785
No. Clones                3       No. Vectors              4
Total contig length   45831       Average length      2412.2
Total characters in readings                          198577
Average reading characters per consensus character      4.33
Average used length of reading                        243.06
Current maximum consensus length is 100000

================================================================
Mon 18 Oct 15:44:00 1999: find internal joins
----------------------------------------------------------------
```

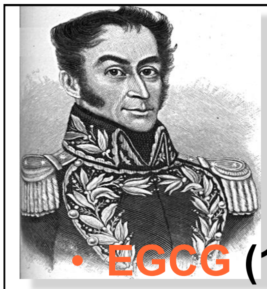
```
SIGN=A=135,C=231,G=103,T=533
SPAC= 13.54
PRIM=550
MACH=ABI 373A DNA Sequencer: Instrument #
93031048
DYEP=DyePrimer(M13RP1)
NAME=
LANE=0
GELN=
PROC=
RTRK=
CONV=phred version=0.980904.e
COMM=

SRCE=ABI 373A
```

SUITE

The **E**uropean **M**olecular **B**iology **O**pen **S**oftware **S**uite

http://bioinf.ibun.unal.edu.co/Pise/5.a/index.php

- 1988, EGCG
- Software libre con Código de Fuente Abierta
- Sus aplicaciones y Librerias están bajo Licencia Pública General (GPL)
  - Edición y visualización de secuencias
  - Alineamiento de secuencias
  - Análisis de Ácidos Nucleicos y proteínas
  - Análisis Filogenéticos

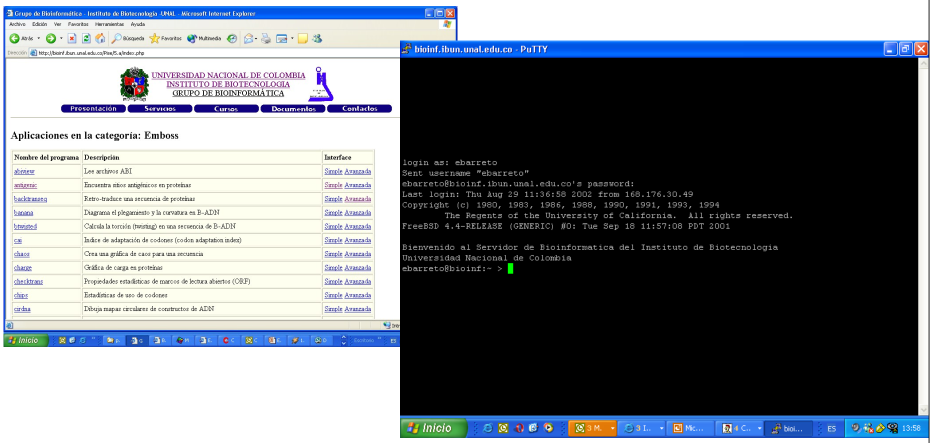


... algo de historia.

- EGCG (1988) surgió como colaboración entre la EMBnet y otros institutos, utilizaba las librerías de GCG.
- Pensado para el Sanger Center.
- Problemas de licenciamiento: imposible distribuir más licencias académicas.
- Se inicia el desarrollo de EMBOSS.

---

emboss


Grupo de Bioinformática - Instituto de Biotecnología -UNAL - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás
Búsqueda
Favoritos
Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/5.a/index.php
UNIVERSIDAD NACIONAL DE COLOMBIA
INSTITUTO DE BIOTECNOLOGIA
GRUPO DE BIOINFORMÁTICA
Presentación
Servicios
Cursos
Documentos
Contactos
Inicio
Escritorio
ES
bioinf.ibun.unal.edu.co - PuTTY
13:58

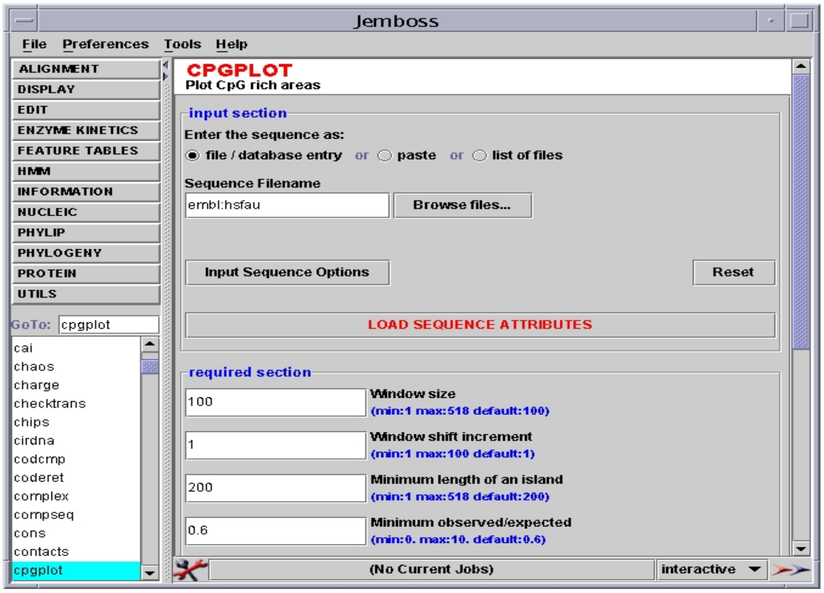
Aplicaciones en la categoría: Emboss

Nombre del programa	Descripción	Interface
abiview	Lee archivos ABI	Simple Avanzada
antigenic	Encuentra sitios antigénicos en proteínas	Simple Avanzada
backtranseq	Retro-traduce una secuencia de proteínas	Simple Avanzada
banana	Diagrama el plegamiento y la curvatura en B-ADN	Simple Avanzada
btwisted	Calcula la torción (twisting) en una secuencia de B-ADN	Simple Avanzada
cai	Indice de adaptación de codones (codon adaptation index)	Simple Avanzada
chaos	Crea una gráfica de caos para una secuencia	Simple Avanzada
charge	Gráfica de carga en proteínas	Simple Avanzada
checktrans	Propiedades estadísticas de marcos de lectura abiertos (ORF)	Simple Avanzada
chips	Estadísticas de uso de codones	Simple Avanzada
cirdna	Dibuja mapas circulares de constructos de ADN	Simple Avanzada

```
login as: ebarreto
Sent username "ebarreto"
ebarreto@bioinf.ibun.unal.edu.co's password:
Last login: Thu Aug 29 11:36:58 2002 from 168.176.30.49
Copyright (c) 1980, 1983, 1986, 1988, 1990, 1991, 1993, 1994
        The Regents of the University of California.  All rights reserved.
FreeBSD 4.4-RELEASE (GENERIC) #0: Tue Sep 18 11:57:08 PDT 2001

Bienvenido al Servidor de Bioinformatica del Instituto de Biotecnologia
Universidad Nacional de Colombia
ebarreto@bioinf:~ >
```

Interfaces

GUIs para EMBOSS

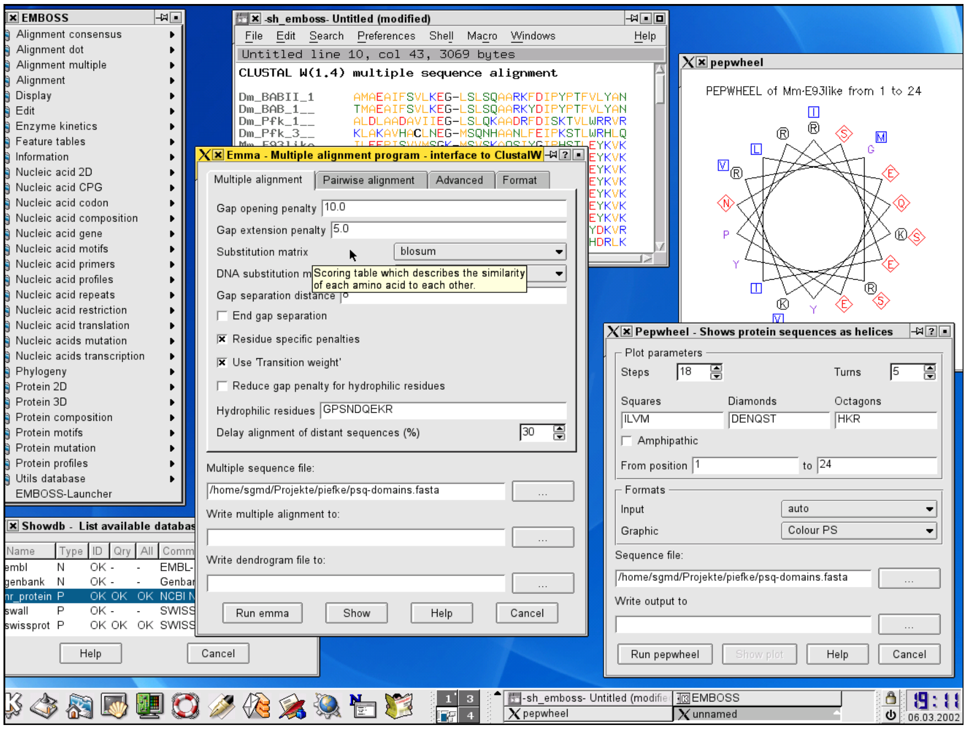

Jemboss
File Preferences Tools Help
ALIGNMENT
DISPLAY
EDIT
ENZYME KINETICS
FEATURE TABLES
HMM
INFORMATION
NUCLEIC
PHYLIP
PHYLOGENY
PROTEIN
UTILS
cai
chaos
charge
checktrans
chips
cirdna
codcmp
coderet
complex
compseq
cons
contacts
cpgplot
Browse files...
Input Sequence Options
Reset
LOAD SEQUENCE ATTRIBUTES
(No Current Jobs)
interactive

CPGPLOT
Plot CpG rich areas

GoTo: cpgplot

input section
Enter the sequence as:
● file / database entry or ○ paste or ○ list of files
Sequence Filename: embl:hsfau

required section
Window size: 100 (min:1 max:518 default:100)
Window shift increment: 1 (min:1 max:100 default:1)
Minimum length of an island: 200 (min:1 max:518 default:200)
Minimum observed/expected: 0.6 (min:0. max:10. default:0.6)


EMBOSS
Alignment consensus
Alignment dot
Alignment multiple
Alignment
Display
Edit
Enzyme kinetics
Feature tables
Information
Nucleic acid 2D
Nucleic acid CPG
Nucleic acid codon
Nucleic acid composition
Nucleic acid gene
Nucleic acid motifs
Nucleic acid primers
Nucleic acid profiles
Nucleic acid repeats
Nucleic acid restriction
Nucleic acid translation
Nucleic acids mutation
Nucleic acids transcription
Phylogeny
Protein 2D
Protein 3D
Protein composition
Protein motifs
Protein mutation
Protein profiles
Utils database
EMBOSS-Launcher
-sh_emboss- Untitled (modified)
File Edit Search Preferences Shell Macro Windows Help
pepwheel
Multiple alignment
Pairwise alignment
Advanced
Format
Run emma
Show
Help
Cancel
Run pepwheel
Show plot
Help
Cancel
-sh_emboss- Untitled (modified)
EMBOSS
pepwheel
unnamed
06.03.2002

```
Untitled line 10, col 43, 3069 bytes
CLUSTAL W(1.4) multiple sequence alignment

Dm_BABII_1      AMAEAIFSVLKEG-LSLSQAARKFDIPYPTFVLYAN
Dm_BAB_1__      TMAEAIFSVLKEG-LSLSQAARKYDIPYPTFVLYAN
Dm_Pfk_1__      ALDLAADAVIIEG-LSLQKAADRFDISKTVLWRRVR
Dm_Pfk_3__      KLAKAVHACLNEG-MSQNHAANLFEIPKSTLWRHLQ
```

PEPWHEEL of Mm-E93like from 1 to 24

Emma - Multiple alignment program - interface to ClustalW

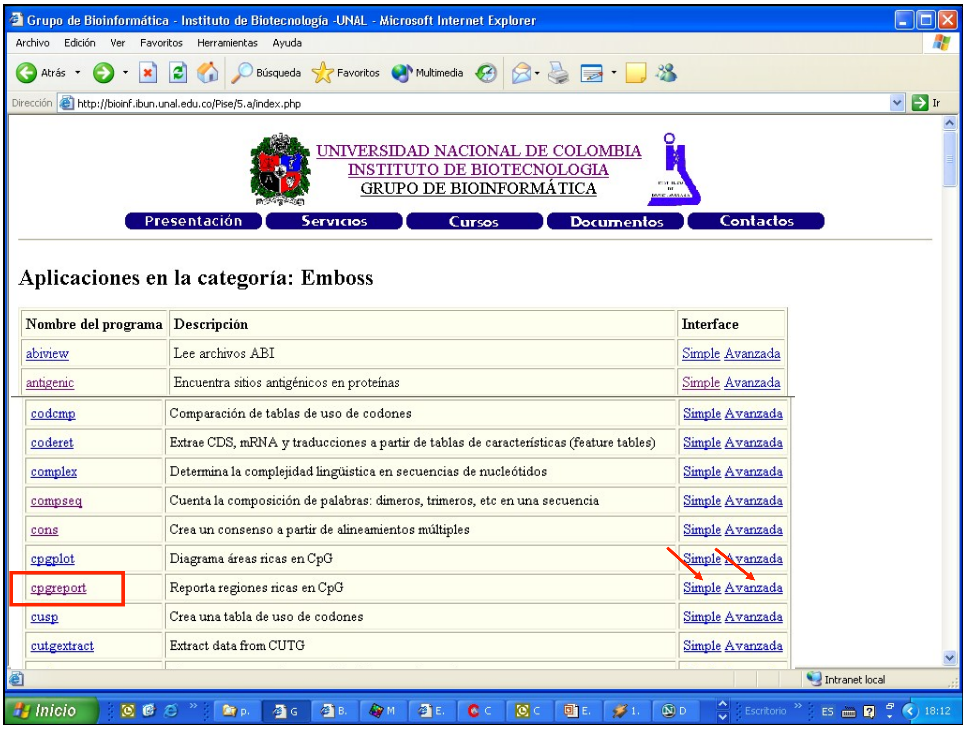
Gap opening penalty: 10.0
Gap extension penalty: 5.0
Substitution matrix: blosum
DNA substitution m… Scoring table which describes the similarity of each amino acid to each other.
Gap separation distance
☐ End gap separation
☒ Residue specific penalties
☒ Use 'Transition weight'
☐ Reduce gap penalty for hydrophilic residues
Hydrophilic residues: GPSNDQEKR
Delay alignment of distant sequences (%): 30
Multiple sequence file: /home/sgmd/Projekte/piefke/psq-domains.fasta
Write multiple alignment to:
Write dendrogram file to:

Showdb - List available databas

Name	Type	ID	Qry	All	Comm
embl	N	OK	-	-	EMBL-
genbank	N	OK	-	-	Genbar
nr_protein	P	OK	OK	OK	NCBI N
swall	P	OK	-	-	SWISS
swissprot	P	OK	OK	OK	SWISS

Pepwheel - Shows protein sequences as helices

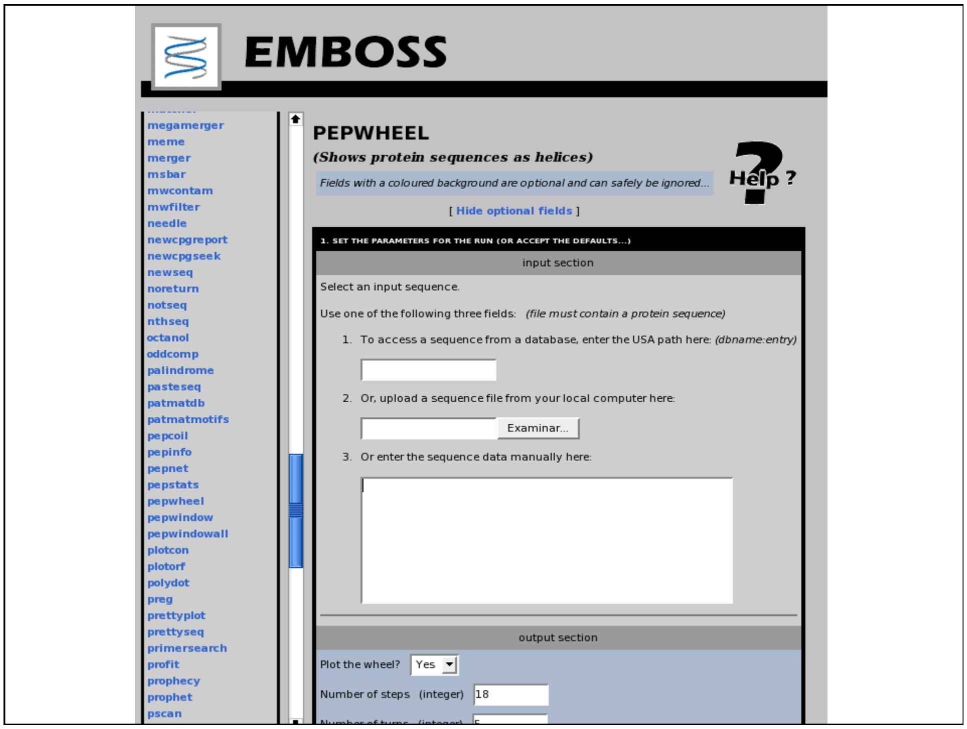
Plot parameters
Steps: 18
Turns: 5
Squares: ILVM
Diamonds: DENQST
Octagons: HKR
☐ Amphipathic
From position 1 to 24
Formats
Input: auto
Graphic: Colour PS
Sequence file: /home/sgmd/Projekte/piefke/psq-domains.fasta
Write output to


Grupo de Bioinformática - Instituto de Biotecnología -UNAL - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás
Búsqueda
Favoritos
Multimedia
Dirección
Ir
UNIVERSIDAD NACIONAL DE COLOMBIA
INSTITUTO DE BIOTECNOLOGIA
GRUPO DE BIOINFORMÁTICA
Presentación
Servicios
Cursos
Documentos
Contactos
Intranet local
Inicio
18:12

Dirección: http://bioinf.ibun.unal.edu.co/Pise/5.a/index.php

Aplicaciones en la categoría: Emboss

Nombre del programa	Descripción	Interface
abiview	Lee archivos ABI	Simple Avanzada
antigenic	Encuentra sitios antigénicos en proteínas	Simple Avanzada
codcmp	Comparación de tablas de uso de codones	Simple Avanzada
coderet	Extrae CDS, mRNA y traducciones a partir de tablas de características (feature tables)	Simple Avanzada
complex	Determina la complejidad lingüística en secuencias de nucleótidos	Simple Avanzada
compseq	Cuenta la composición de palabras: dimeros, trimeros, etc en una secuencia	Simple Avanzada
cons	Crea un consenso a partir de alineamientos múltiples	Simple Avanzada
cpgplot	Diagrama áreas ricas en CpG	Simple Avanzada
cpgreport	Reporta regiones ricas en CpG	Simple Avanzada
cusp	Crea una tabla de uso de codones	Simple Avanzada
cutgextract	Extract data from CUTG	Simple Avanzada


EMBOSS
megamerger
meme
merger
msbar
mwcontam
mwfilter
needle
newcpgreport
newcpgseek
newseq
noreturn
notseq
nthseq
octanol
oddcomp
palindrome
pasteseq
patmatdb
patmatmotifs
pepcoil
pepinfo
pepnet
pepstats
pepwheel
pepwindow
pepwindowall
plotcon
plotorf
polydot
preg
prettyplot
prettyseq
primersearch
profit
prophecy
prophet
pscan
Help ?
[Hide optional fields]
Examinar...

PEPWHEEL

(Shows protein sequences as helices)

Fields with a coloured background are optional and can safely be ignored...

1. SET THE PARAMETERS FOR THE RUN (OR ACCEPT THE DEFAULTS...)

input section

Select an input sequence.

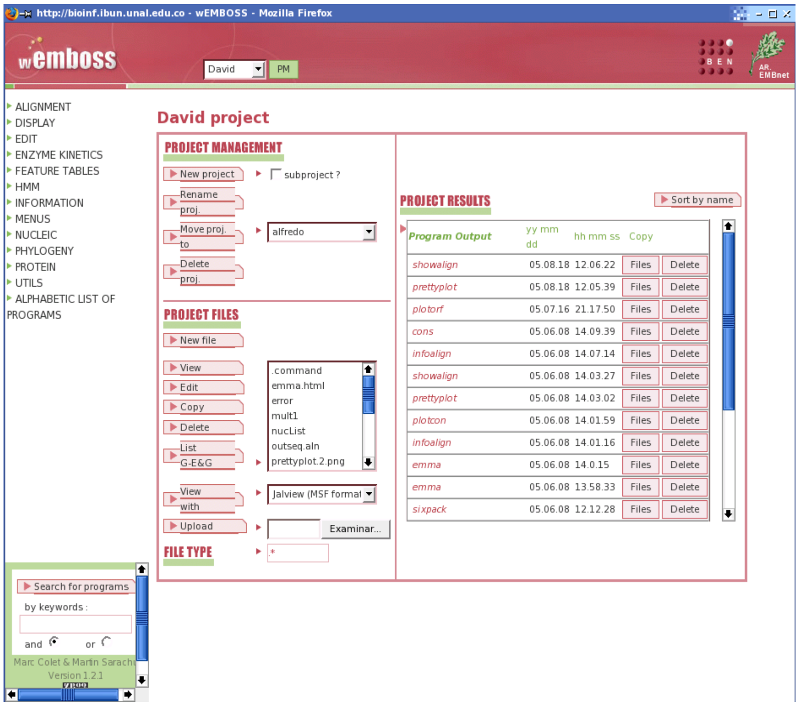
Use one of the following three fields: *(file must contain a protein sequence)*

1. To access a sequence from a database, enter the USA path here: *(dbname:entry)*
2. Or, upload a sequence file from your local computer here:
3. Or enter the sequence data manually here:

output section

Plot the wheel? Yes

Number of steps (integer) 18

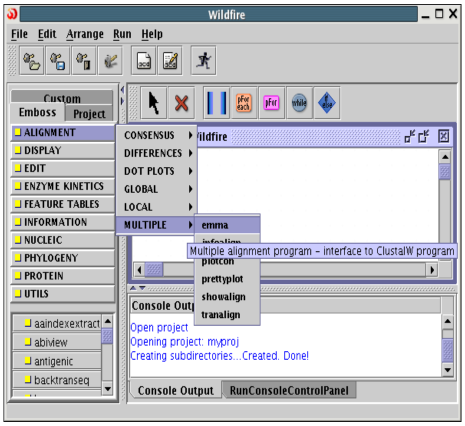

http://bioinf.ibun.unal.edu.co - wEMBOSS - Mozilla Firefox
wemboss
David
PM
B E N
AR.EMBnet
ALIGNMENT
DISPLAY
EDIT
ENZYME KINETICS
FEATURE TABLES
HMM
INFORMATION
MENUS
NUCLEIC
PHYLOGENY
PROTEIN
UTILS
ALPHABETIC LIST OF PROGRAMS
PROJECT MANAGEMENT
New project
subproject ?
Rename proj.
Move proj. to
alfredo
Delete proj.
PROJECT FILES
New file
View
Edit
Copy
Delete
List G-E&G
View with
Jalview (MSF format
Upload
Examinar...
FILE TYPE
PROJECT RESULTS
Sort by name
Search for programs
by keywords :
and
or
Marc Colet & Martin Sarachu
Version 1.2.1

David project

.command
emma.html
error
mult1
nucList
outseq.aln
prettyplot.2.png

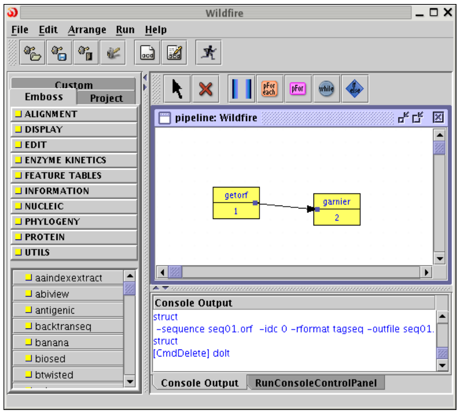
Program Output	yy mm dd	hh mm ss	Copy	
showalign	05.08.18	12.06.22	Files	Delete
prettyplot	05.08.18	12.05.39	Files	Delete
plotorf	05.07.16	21.17.50	Files	Delete
cons	05.06.08	14.09.39	Files	Delete
infoalign	05.06.08	14.07.14	Files	Delete
showalign	05.06.08	14.03.27	Files	Delete
prettyplot	05.06.08	14.03.02	Files	Delete
plotcon	05.06.08	14.01.59	Files	Delete
infoalign	05.06.08	14.01.16	Files	Delete
emma	05.06.08	14.0.15	Files	Delete
emma	05.06.08	13.58.33	Files	Delete
sixpack	05.06.08	12.12.28	Files	Delete

Workflows con EMBOSS

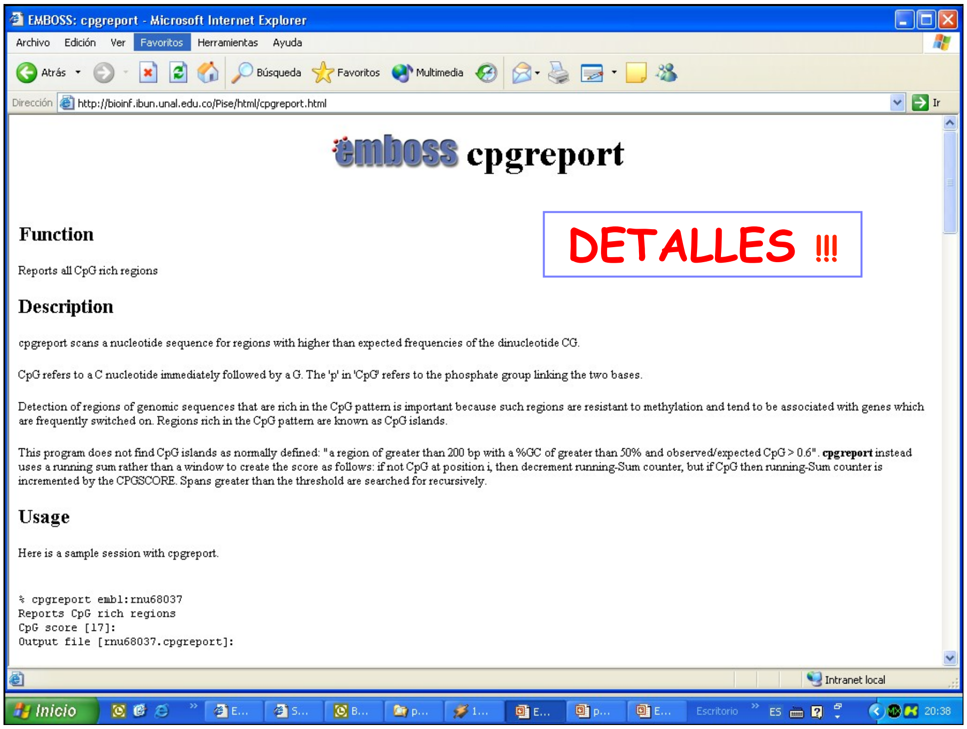

Wildfire
File Edit Arrange Run Help
Custom
Emboss
Project
ALIGNMENT
DISPLAY
EDIT
ENZYME KINETICS
FEATURE TABLES
INFORMATION
NUCLEIC
PHYLOGENY
PROTEIN
UTILS
CONSENSUS
DIFFERENCES
DOT PLOTS
GLOBAL
LOCAL
MULTIPLE
emma
infoalign
plotcon
prettyplot
showalign
tranalign
aaindexextract
abiview
antigenic
backtranseq
Console Output
RunConsoleControlPanel

Multiple alignment program - interface to ClustalW program

```
Open project
Opening project: myproj
Creating subdirectories...Created. Done!
```


Wildfire
File Edit Arrange Run Help
Custom
Emboss
Project
ALIGNMENT
DISPLAY
EDIT
ENZYME KINETICS
FEATURE TABLES
INFORMATION
NUCLEIC
PHYLOGENY
PROTEIN
UTILS
pipeline: Wildfire
getorf
1
garnier
2
aaindexextract
abiview
antigenic
backtranseq
banana
biosed
btwisted
Console Output
RunConsoleControlPanel

```
struct
 -sequence seq01.orf  -idc 0 -rformat tagseq -outfile seq01.
struct
[CmdDelete] doIt
```


EMBOSS: cpgreport - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/html/cpgreport.html
Ir
Intranet local
Inicio
Escritorio
ES
20:38

emboss cpgreport

DETALLES !!!

Function

Reports all CpG rich regions

Description

cpgreport scans a nucleotide sequence for regions with higher than expected frequencies of the dinucleotide CG.

CpG refers to a C nucleotide immediately followed by a G. The 'p' in 'CpG' refers to the phosphate group linking the two bases.

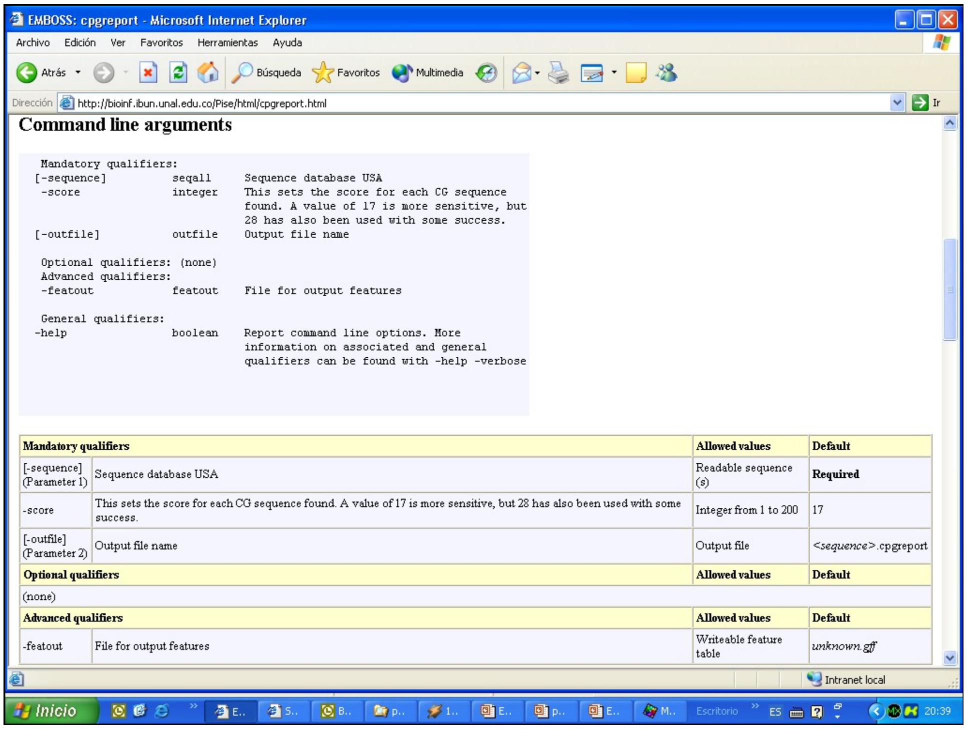
Detection of regions of genomic sequences that are rich in the CpG pattern is important because such regions are resistant to methylation and tend to be associated with genes which are frequently switched on. Regions rich in the CpG pattern are known as CpG islands.

This program does not find CpG islands as normally defined: "a region of greater than 200 bp with a %GC of greater than 50% and observed/expected CpG > 0.6". **cpgreport** instead uses a running sum rather than a window to create the score as follows: if not CpG at position i, then decrement running-Sum counter, but if CpG then running-Sum counter is incremented by the CPGSCORE. Spans greater than the threshold are searched for recursively.

Usage

Here is a sample session with cpgreport.

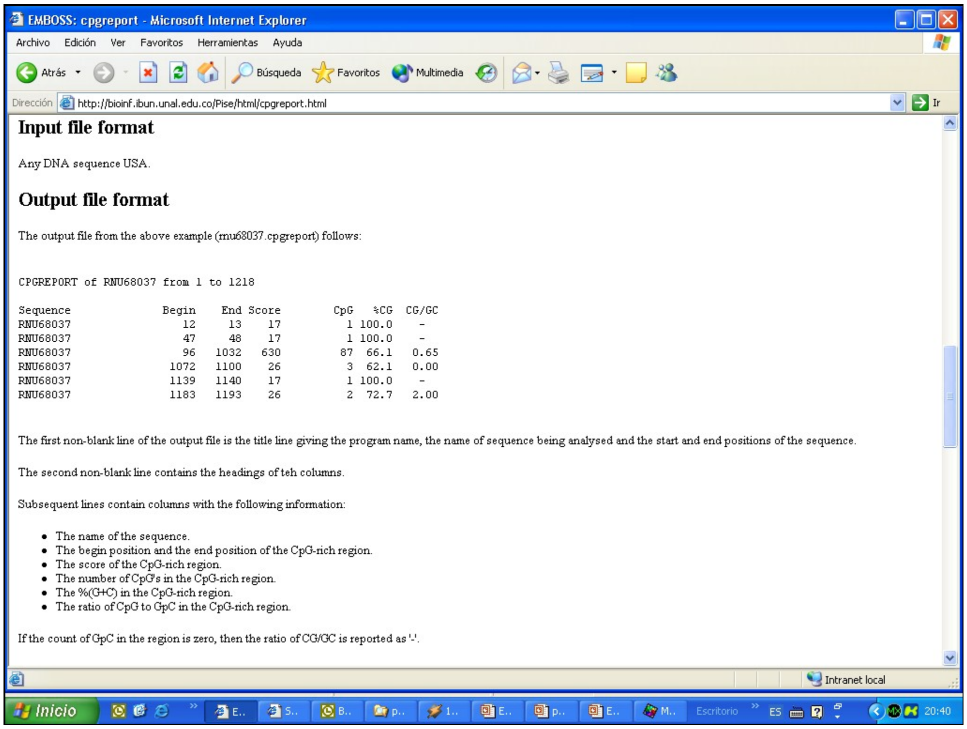
```
% cpgreport embl:rnu68037
Reports CpG rich regions
CpG score [17]:
Output file [rnu68037.cpgreport]:
```


EMBOSS: cpgreport - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/html/cpgreport.html
Ir
Intranet local
Inicio
Escritorio
ES
20:39

Command line arguments

```
   Mandatory qualifiers:
  [-sequence]          seqall     Sequence database USA
   -score              integer    This sets the score for each CG sequence
                                  found. A value of 17 is more sensitive, but
                                  28 has also been used with some success.
  [-outfile]           outfile    Output file name

   Optional qualifiers: (none)
   Advanced qualifiers:
   -featout            featout    File for output features

   General qualifiers:
  -help                boolean    Report command line options. More
                                  information on associated and general
                                  qualifiers can be found with -help -verbose
```

Mandatory qualifiers		**Allowed values**	**Default**
[-sequence] (Parameter 1)	Sequence database USA	Readable sequence (s)	**Required**
-score	This sets the score for each CG sequence found. A value of 17 is more sensitive, but 28 has also been used with some success.	Integer from 1 to 200	17
[-outfile] (Parameter 2)	Output file name	Output file	*<sequence>*.cpgreport
Optional qualifiers		**Allowed values**	**Default**
(none)			
Advanced qualifiers		**Allowed values**	**Default**
-featout	File for output features	Writeable feature table	*unknown.gff*


EMBOSS: cpgreport - Microsoft Internet Explorer
Archivo
Edición
Ver
Favoritos
Herramientas
Ayuda
Atrás
Búsqueda
Favoritos
Multimedia
Dirección
http://bioinf.ibun.unal.edu.co/Pise/html/cpgreport.html
Ir
Intranet local
Inicio
E..
S..
B..
p..
1..
E..
p..
E..
M..
Escritorio
ES
20:40

Input file format

Any DNA sequence USA.

Output file format

The output file from the above example (rnu68037.cpgreport) follows:

```
CPGREPORT of RNU68037 from 1 to 1218

Sequence              Begin    End Score        CpG   %CG  CG/GC
RNU68037                 12     13    17          1 100.0    -
RNU68037                 47     48    17          1 100.0    -
RNU68037                 96   1032   630         87  66.1   0.65
RNU68037               1072   1100    26          3  62.1   0.00
RNU68037               1139   1140    17          1 100.0    -
RNU68037               1183   1193    26          2  72.7   2.00
```

The first non-blank line of the output file is the title line giving the program name, the name of sequence being analysed and the start and end positions of the sequence.

The second non-blank line contains the headings of teh columns.

Subsequent lines contain columns with the following information:

- The name of the sequence.
- The begin position and the end position of the CpG-rich region.
- The score of the CpG-rich region.
- The number of CpG's in the CpG-rich region.
- The %(G+C) in the CpG-rich region.
- The ratio of CpG to GpC in the CpG-rich region.

If the count of GpC in the region is zero, then the ratio of CG/GC is reported as '-'.

FORMATOS

El formato de Microsoft WORD no es un format de secuencia!!!!

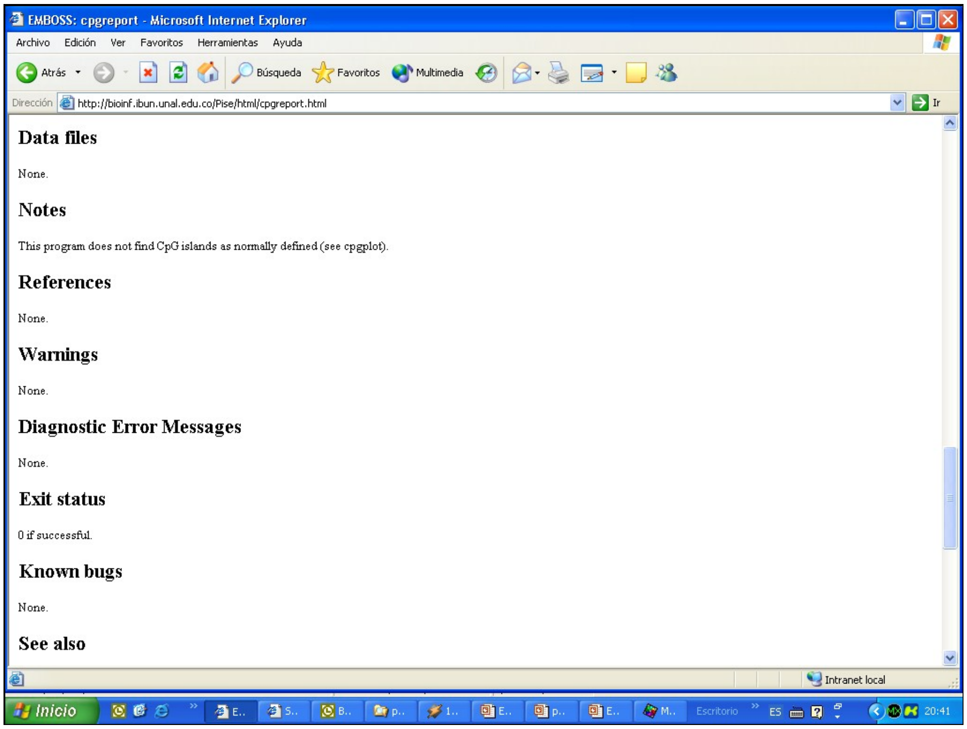
emboss

Una secuencia no requiere ninguna clase de identificacion, pero AYUDA!!!

- Formatos de secuencia son texto ASCII

El formato **fasta** contiene:

>xyz some other comment
ttcctctttctcgactccatcttcgcggtagctgggaccgccgttcagtcgccaatatgc
agctctttgtccgcgcccaggagctacacaccttcgaggtgaccggccaggaaacggtcg
cccagatcaaggctcatgtagcctcactggagggcatt


EMBOSS: cpgreport - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/html/cpgreport.html
Ir
Intranet local
Inicio
Escritorio
ES
20:41

Data files

None.

Notes

This program does not find CpG islands as normally defined (see cpgplot).

References

None.

Warnings

None.

Diagnostic Error Messages

None.

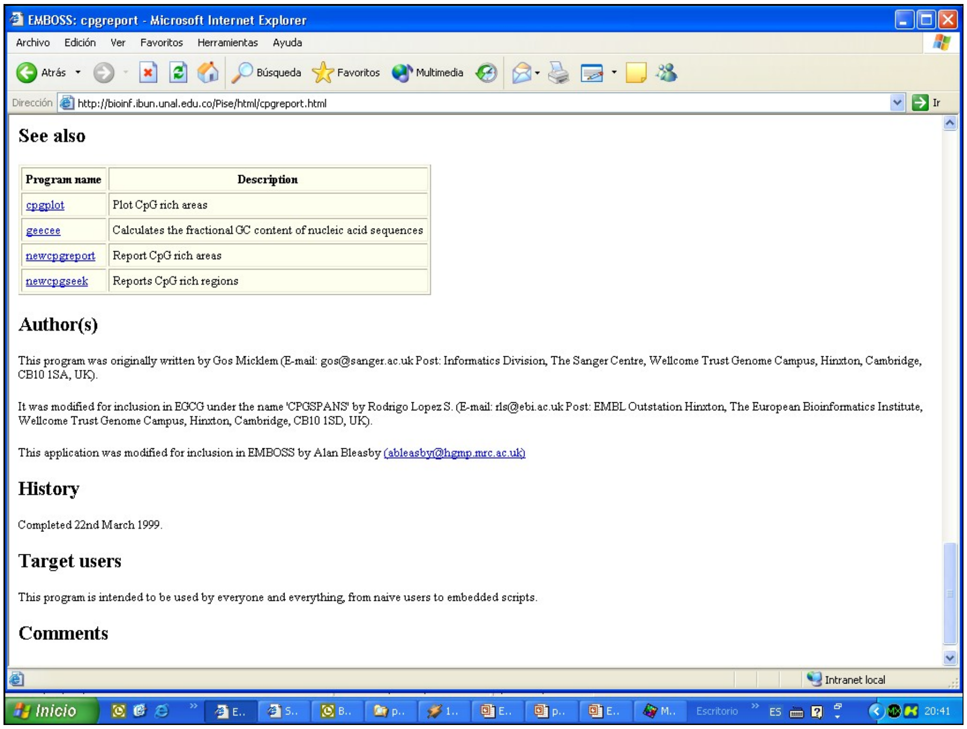
Exit status

0 if successful.

Known bugs

None.

See also


EMBOSS: cpgreport - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/html/cpgreport.html
Ir
Intranet local
Inicio
Escritorio
ES
20:41

See also

Program name	Description
cpgplot	Plot CpG rich areas
geecee	Calculates the fractional GC content of nucleic acid sequences
newcpgreport	Report CpG rich areas
newcpgseek	Reports CpG rich regions

Author(s)

This program was originally written by Gos Micklem (E-mail: gos@sanger.ac.uk Post: Informatics Division, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK).

It was modified for inclusion in EGCG under the name 'CPGSPANS' by Rodrigo Lopez S. (E-mail: rls@ebi.ac.uk Post: EMBL Outstation Hinxton, The European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK).

This application was modified for inclusion in EMBOSS by Alan Bleasby (ableasby@hgmp.mrc.ac.uk)

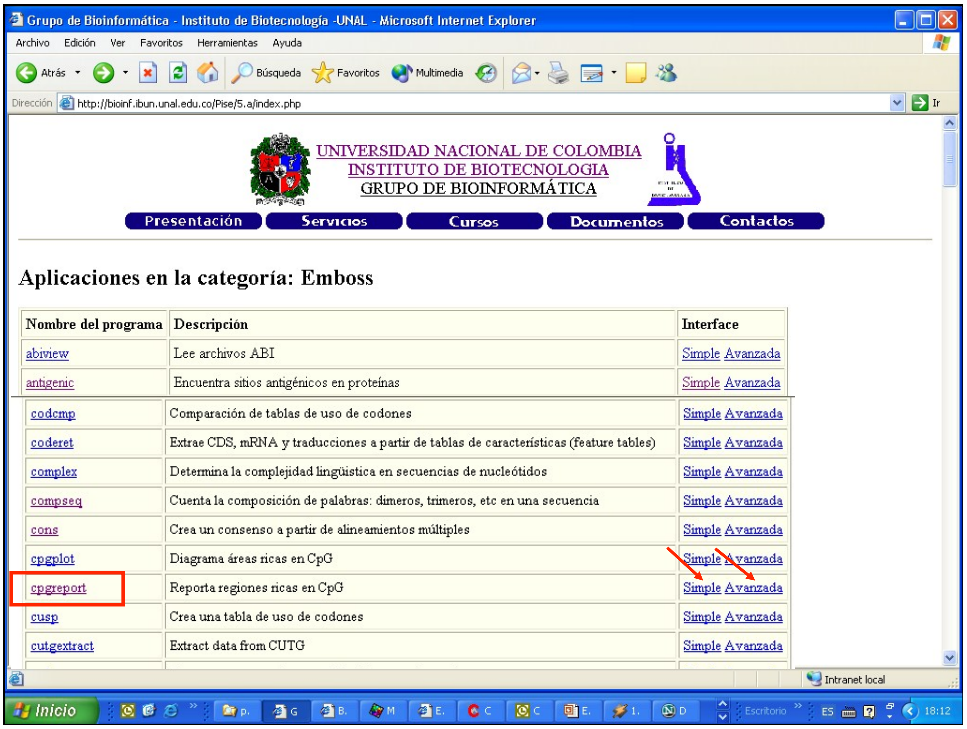
History

Completed 22nd March 1999.

Target users

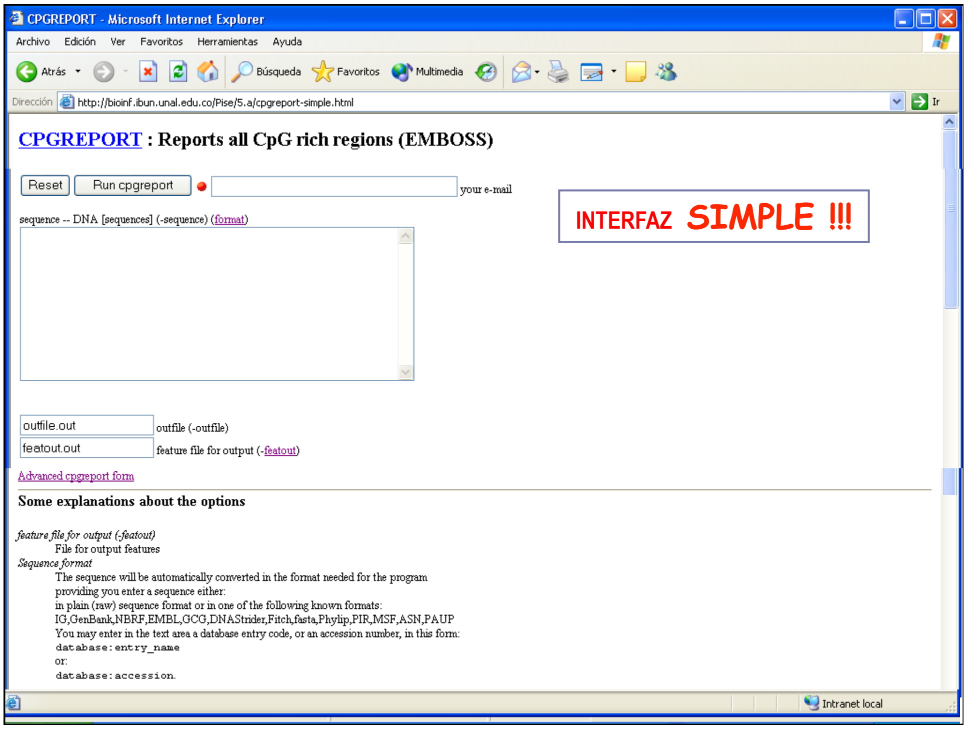
This program is intended to be used by everyone and everything, from naive users to embedded scripts.

Comments


Grupo de Bioinformática - Instituto de Biotecnología -UNAL - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/5.a/index.php
Ir
UNIVERSIDAD NACIONAL DE COLOMBIA
INSTITUTO DE BIOTECNOLOGIA
GRUPO DE BIOINFORMÁTICA
Presentación
Servicios
Cursos
Documentos
Contactos
Intranet local
Inicio
Escritorio
ES
18:12

Aplicaciones en la categoría: Emboss

Nombre del programa	Descripción	Interface
abiview	Lee archivos ABI	Simple Avanzada
antigenic	Encuentra sitios antigénicos en proteínas	Simple Avanzada
codcmp	Comparación de tablas de uso de codones	Simple Avanzada
coderet	Extrae CDS, mRNA y traducciones a partir de tablas de características (feature tables)	Simple Avanzada
complex	Determina la complejidad lingüística en secuencias de nucleótidos	Simple Avanzada
compseq	Cuenta la composición de palabras: dimeros, trimeros, etc en una secuencia	Simple Avanzada
cons	Crea un consenso a partir de alineamientos múltiples	Simple Avanzada
cpgplot	Diagrama áreas ricas en CpG	Simple Avanzada
cpgreport	Reporta regiones ricas en CpG	Simple Avanzada
cusp	Crea una tabla de uso de codones	Simple Avanzada
cutgextract	Extract data from CUTG	Simple Avanzada


CPGREPORT - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/5.a/cpgreport-simple.html
Ir
Reset
Run cpgreport
your e-mail
Intranet local

CPGREPORT : Reports all CpG rich regions (EMBOSS)

INTERFAZ **SIMPLE !!!**

sequence -- DNA [sequences] (-sequence) (format)

outfile.out outfile (-outfile)

featout.out feature file for output (-featout)

Advanced cpgreport form

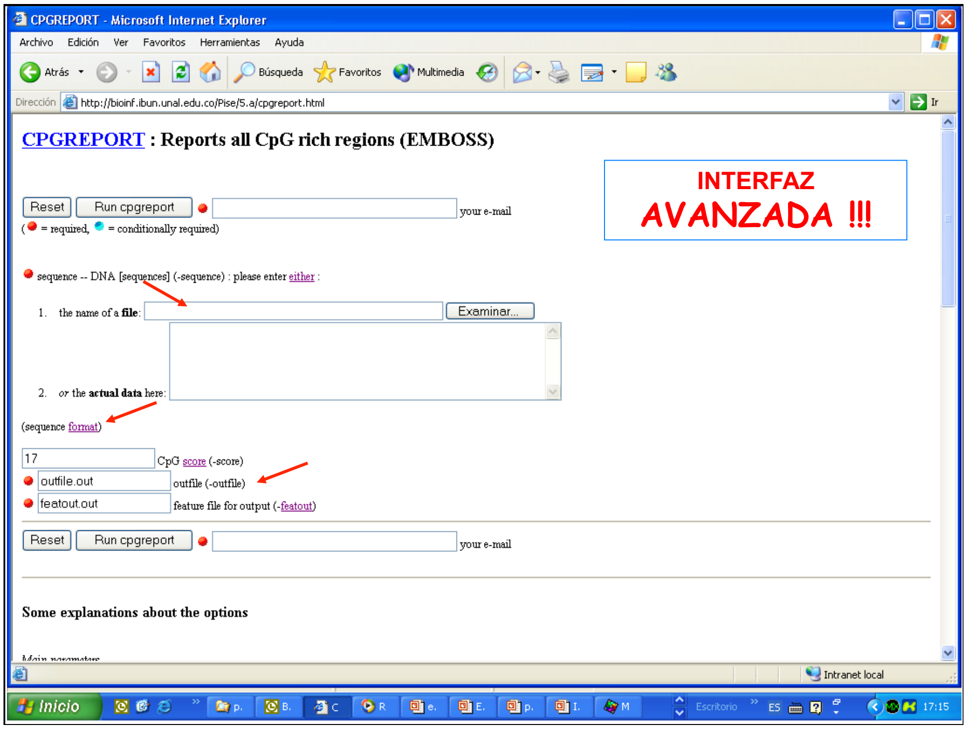
Some explanations about the options

feature file for output (-featout)

File for output features

Sequence format

The sequence will be automatically converted in the format needed for the program providing you enter a sequence either:
in plain (raw) sequence format or in one of the following known formats:
IG,GenBank,NBRF,EMBL,GCG,DNAStrider,Fitch,fasta,Phylip,PIR,MSF,ASN,PAUP
You may enter in the text area a database entry code, or an accession number, in this form:
`database:entry_name`
or:
`database:accession.`


CPGREPORT - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección http://bioinf.ibun.unal.edu.co/Pise/5.a/cpgreport.html
Ir
Intranet local
Inicio
17:15

CPGREPORT : Reports all CpG rich regions (EMBOSS)

INTERFAZ
AVANZADA !!!

Reset | Run cpgreport | ● your e-mail

(● = required, ● = conditionally required)

● sequence -- DNA [sequences] (-sequence) : please enter either :

1. the name of a **file**: Examinar...
2. *or* the **actual data** here:

(sequence format)

17 CpG score (-score)

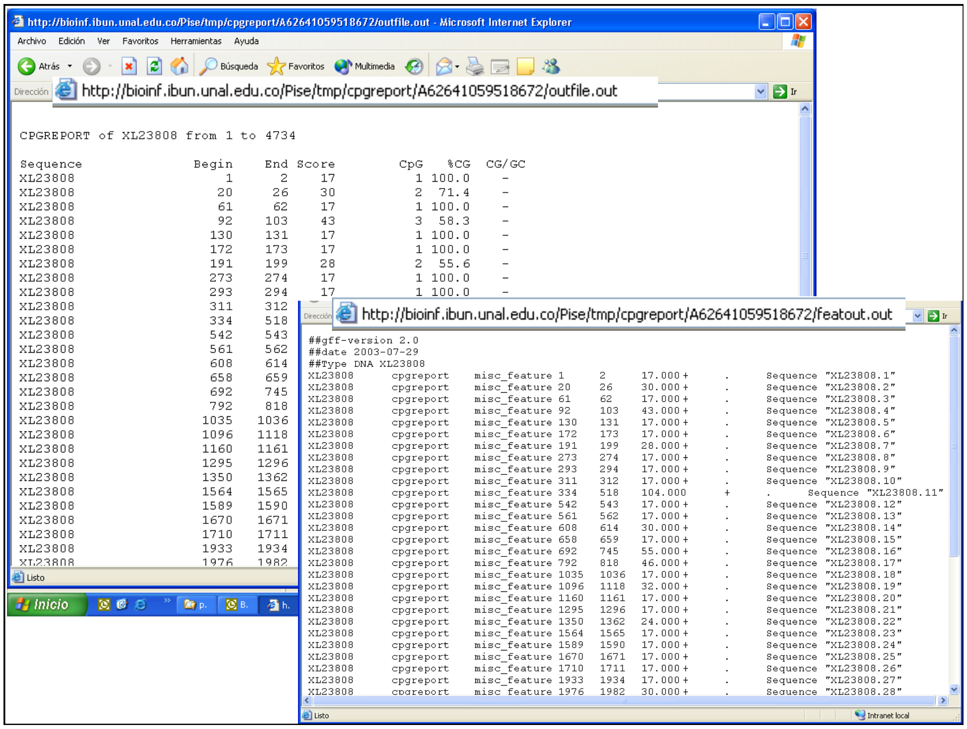
● outfile.out outfile (-outfile)

● featout.out feature file for output (-featout)

Reset | Run cpgreport | ● your e-mail

Some explanations about the options

Main parameters


http://bioinf.ibun.unal.edu.co/Pise/tmp/cpgreport/A62641059518672/outfile.out - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás Búsqueda Favoritos Multimedia
Dirección
Ir
Listo
Inicio
Intranet local

http://bioinf.ibun.unal.edu.co/Pise/tmp/cpgreport/A62641059518672/outfile.out

```
CPGREPORT of XL23808 from 1 to 4734

Sequence                 Begin    End  Score       CpG   %CG   CG/GC
XL23808                      1      2     17         1 100.0     -
XL23808                     20     26     30         2  71.4     -
XL23808                     61     62     17         1 100.0     -
XL23808                     92    103     43         3  58.3     -
XL23808                    130    131     17         1 100.0     -
XL23808                    172    173     17         1 100.0     -
XL23808                    191    199     28         2  55.6     -
XL23808                    273    274     17         1 100.0     -
XL23808                    293    294     17         1 100.0     -
XL23808                    311    312
XL23808                    334    518
XL23808                    542    543
XL23808                    561    562
XL23808                    608    614
XL23808                    658    659
XL23808                    692    745
XL23808                    792    818
XL23808                   1035   1036
XL23808                   1096   1118
XL23808                   1160   1161
XL23808                   1295   1296
XL23808                   1350   1362
XL23808                   1564   1565
XL23808                   1589   1590
XL23808                   1670   1671
XL23808                   1710   1711
XL23808                   1933   1934
XL23808                   1976   1982
```

http://bioinf.ibun.unal.edu.co/Pise/tmp/cpgreport/A62641059518672/featout.out

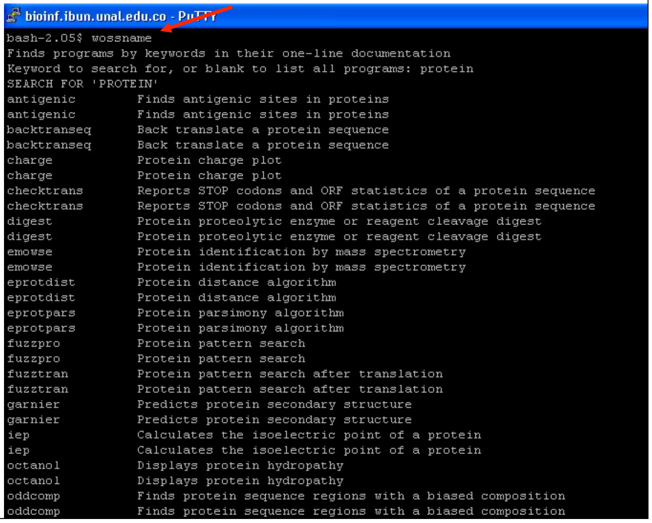
```
##gff-version 2.0
##date 2003-07-29
##Type DNA XL23808
XL23808     cpgreport   misc_feature 1      2       17.000 +     .     Sequence "XL23808.1"
XL23808     cpgreport   misc_feature 20     26      30.000 +     .     Sequence "XL23808.2"
XL23808     cpgreport   misc_feature 61     62      17.000 +     .     Sequence "XL23808.3"
XL23808     cpgreport   misc_feature 92     103     43.000 +     .     Sequence "XL23808.4"
XL23808     cpgreport   misc_feature 130    131     17.000 +     .     Sequence "XL23808.5"
XL23808     cpgreport   misc_feature 172    173     17.000 +     .     Sequence "XL23808.6"
XL23808     cpgreport   misc_feature 191    199     28.000 +     .     Sequence "XL23808.7"
XL23808     cpgreport   misc_feature 273    274     17.000 +     .     Sequence "XL23808.8"
XL23808     cpgreport   misc_feature 293    294     17.000 +     .     Sequence "XL23808.9"
XL23808     cpgreport   misc_feature 311    312     17.000 +     .     Sequence "XL23808.10"
XL23808     cpgreport   misc_feature 334    518     104.000      +     .     Sequence "XL23808.11"
XL23808     cpgreport   misc_feature 542    543     17.000 +     .     Sequence "XL23808.12"
XL23808     cpgreport   misc_feature 561    562     17.000 +     .     Sequence "XL23808.13"
XL23808     cpgreport   misc_feature 608    614     30.000 +     .     Sequence "XL23808.14"
XL23808     cpgreport   misc_feature 658    659     17.000 +     .     Sequence "XL23808.15"
XL23808     cpgreport   misc_feature 692    745     55.000 +     .     Sequence "XL23808.16"
XL23808     cpgreport   misc_feature 792    818     46.000 +     .     Sequence "XL23808.17"
XL23808     cpgreport   misc_feature 1035   1036    17.000 +     .     Sequence "XL23808.18"
XL23808     cpgreport   misc_feature 1096   1118    32.000 +     .     Sequence "XL23808.19"
XL23808     cpgreport   misc_feature 1160   1161    17.000 +     .     Sequence "XL23808.20"
XL23808     cpgreport   misc_feature 1295   1296    17.000 +     .     Sequence "XL23808.21"
XL23808     cpgreport   misc_feature 1350   1362    24.000 +     .     Sequence "XL23808.22"
XL23808     cpgreport   misc_feature 1564   1565    17.000 +     .     Sequence "XL23808.23"
XL23808     cpgreport   misc_feature 1589   1590    17.000 +     .     Sequence "XL23808.24"
XL23808     cpgreport   misc_feature 1670   1671    17.000 +     .     Sequence "XL23808.25"
XL23808     cpgreport   misc_feature 1710   1711    17.000 +     .     Sequence "XL23808.26"
XL23808     cpgreport   misc_feature 1933   1934    17.000 +     .     Sequence "XL23808.27"
XL23808     cpgreport   misc_feature 1976   1982    30.000 +     .     Sequence "XL23808.28"
```

wossname

Reporta una lista de los programas existentes en EMBOSS según la palabra clave utilizada

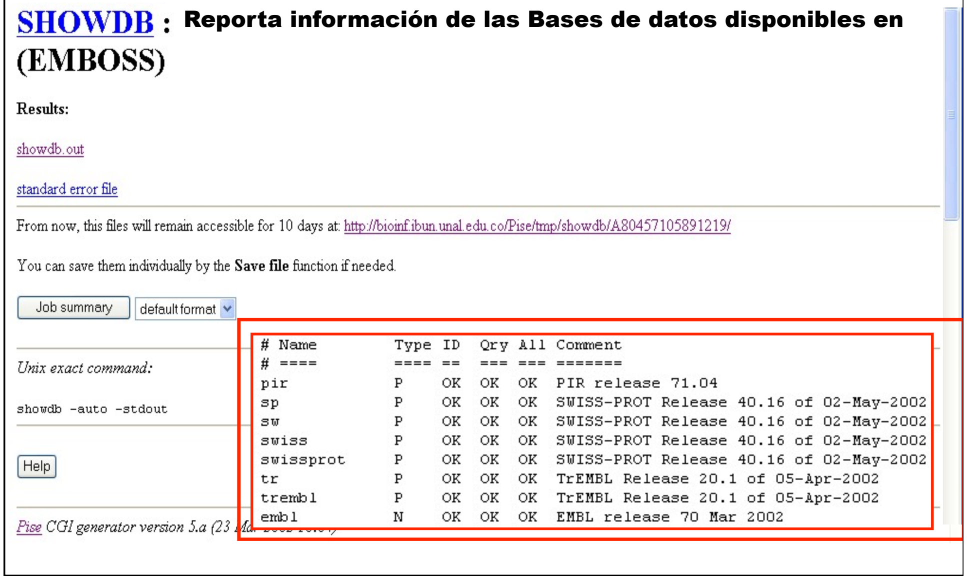
- Ej: protein
  Antigenic
  Backtranseq

unix % wossname
unix % wossname –opt
unix % wossname –help



```
bash-2.05$ wossname
Finds programs by keywords in their one-line documentation
Keyword to search for, or blank to list all programs: protein
SEARCH FOR 'PROTEIN'
antigenic        Finds antigenic sites in proteins
antigenic        Finds antigenic sites in proteins
backtranseq      Back translate a protein sequence
backtranseq      Back translate a protein sequence
charge           Protein charge plot
charge           Protein charge plot
checktrans       Reports STOP codons and ORF statistics of a protein sequence
checktrans       Reports STOP codons and ORF statistics of a protein sequence
digest           Protein proteolytic enzyme or reagent cleavage digest
digest           Protein proteolytic enzyme or reagent cleavage digest
emowse           Protein identification by mass spectrometry
emowse           Protein identification by mass spectrometry
eprotdist        Protein distance algorithm
eprotdist        Protein distance algorithm
eprotpars        Protein parsimony algorithm
eprotpars        Protein parsimony algorithm
fuzzpro          Protein pattern search
fuzzpro          Protein pattern search
fuzztran         Protein pattern search after translation
fuzztran         Protein pattern search after translation
garnier          Predicts protein secondary structure
garnier          Predicts protein secondary structure
iep              Calculates the isoelectric point of a protein
iep              Calculates the isoelectric point of a protein
octanol          Displays protein hydropathy
octanol          Displays protein hydropathy
oddcomp          Finds protein sequence regions with a biased composition
oddcomp          Finds protein sequence regions with a biased composition
```

showdb



SHOWDB : Reporta información de las Bases de datos disponibles en (EMBOSS)

Results:

showdb.out

standard error file

From now, this files will remain accessible for 10 days at: http://bioinf.ibun.unal.edu.co/Pise/tmp/showdb/A80457105891219/

You can save them individually by the **Save file** function if needed.

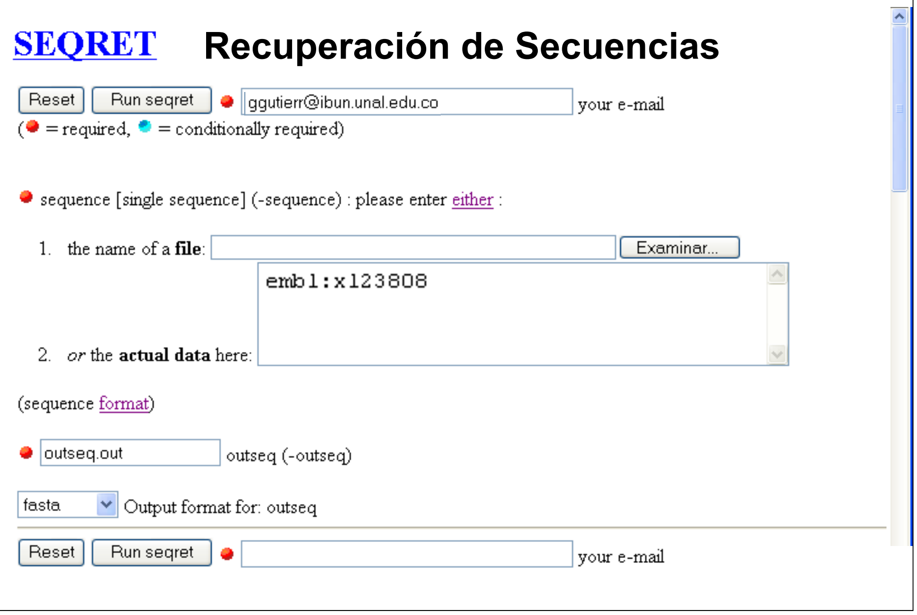
Unix exact command:

```
showdb -auto -stdout
```

Pise CGI generator version 5.a (23 Ma

```
# Name         Type ID  Qry All Comment
# ====         ==== ==  === === =======
pir            P    OK  OK  OK  PIR release 71.04
sp             P    OK  OK  OK  SWISS-PROT Release 40.16 of 02-May-2002
sw             P    OK  OK  OK  SWISS-PROT Release 40.16 of 02-May-2002
swiss          P    OK  OK  OK  SWISS-PROT Release 40.16 of 02-May-2002
swissprot      P    OK  OK  OK  SWISS-PROT Release 40.16 of 02-May-2002
tr             P    OK  OK  OK  TrEMBL Release 20.1 of 05-Apr-2002
trembl         P    OK  OK  OK  TrEMBL Release 20.1 of 05-Apr-2002
embl           N    OK  OK  OK  EMBL release 70 Mar 2002
```

emboss seqret



SEQRET **Recuperación de Secuencias**

● ggutierr@ibun.unal.edu.co your e-mail

(● = required, ● = conditionally required)

● sequence [single sequence] (-sequence) : please enter either :

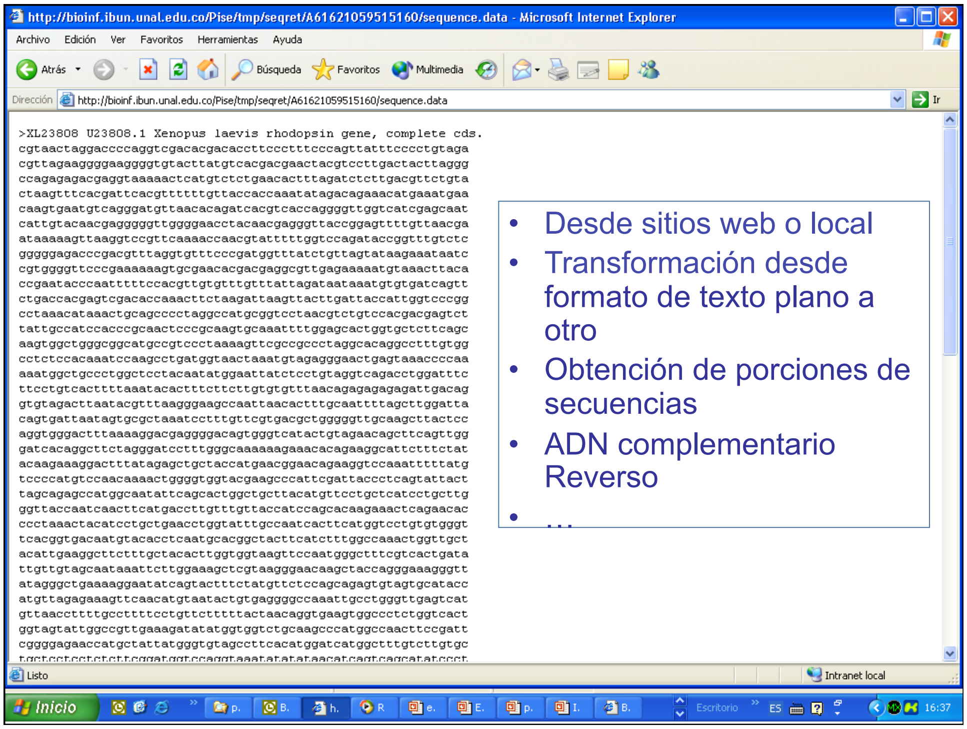
1. the name of a **file**:
2. *or* the **actual data** here: embl:x123808

(sequence format)

● outseq.out outseq (-outseq)

fasta Output format for: outseq

● your e-mail



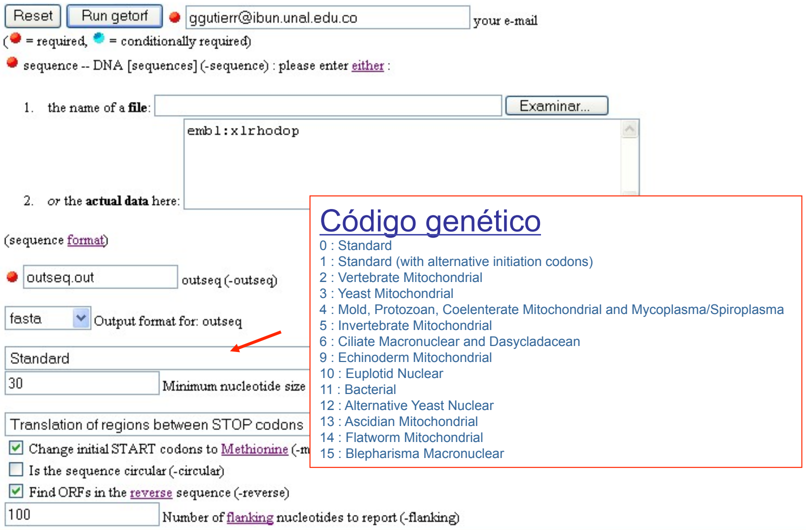
http://bioinf.ibun.unal.edu.co/Pise/tmp/seqret/A61621059515160/sequence.data

```
>XL23808 U23808.1 Xenopus laevis rhodopsin gene, complete cds.
cgtaactaggaccccaggtcgacacgacaccttccctttcccagttatttcccctgtaga
cgttagaaggggaaggggtgtacttatgtcacgacgaactacgtccttgactacttaggg
ccagagagacgaggtaaaaactcatgtctctgaacactttagatctcttgacgttctgta
ctaagtttcacgattcacgtttttttgttaccaccaaatatagacagaaacatgaaatgaa
caagtgaatgtcagggatgttaacacagatcacgtcaccaggggttggtcatcgagcaat
cattgtacaacgaggggggttggggaacctacaacgagggttaccggagtttgttaacga
ataaaaagttaaggtccgttcaaaaccaacgtatttttggtccagataccggtttgtctc
gggggagacccgacgtttaggtgtttcccgatggtttatctgttagtataagaaataatc
cgtggggttcccgaaaaaagtgcgaacacgacgaggcgttgagaaaaatgtaaacttaca
ccgaatacccaatttttccacgttgtgtttgtttattagataataaatgtgtgatcagtt
ctgaccacgagtcgacaccaaacttctaagattaagttacttgattaccattggtcccgg
cctaaacataaactgcagcccctaggccatgcggtcctaacgtctgtccacgacgagtct
tattgccatccacccgcaactcccgcaagtgcaaattttggagcactggtgctcttcagc
aagtggctgggcggcatgccgtccctaaaagttcgccgccctaggcacaggcctttgtgg
cctctccacaaatccaagcctgatggtaactaaatgtagagggaactgagtaaaccccaa
aaatggctgccctggctcctacaatatggaattatctcctgtaggtcagacctggatttc
ttcctgtcacttttaaatacactttcttcttgtgtgtttaacagagagagagattgacag
gtgtagacttaatacgtttaagggaagccaattaacactttgcaattttagcttggatta
cagtgattaatagtgcgctaaatcctttgttcgtgacgctggggggttgcaagcttactcc
aggtgggactttaaaaggacgaggggacagtgggtcatactgtagaacagcttcagttgg
gatcacaggcttctagggatcctttgggcaaaaaagaaacacagaaggcattctttctat
acaagaaaggactttatagagctgctaccatgaacggaacagaaggtccaaatttttatg
tccccatgtccaacaaaactggggtggtacgaagcccattcgattaccctcagtattact
tagcagagccatggcaatattcagcactggctgcttacatgttcctgctcatcctgcttg
ggttaccaatcaacttcatgaccttgtttgttaccatccagcacaagaaactcagaacac
ccctaaactacatcctgctgaacctggtatttgccaatcacttcatggtcctgtgtgggt
tcacggtgacaatgtacacctcaatgcacggctacttcatctttggccaaactggttgct
acattgaaggcttctttgctacacttggtggtaagttccaatgggctttcgtcactgata
ttgttgtagcaataaattcttggaaagctcgtaagggaacaagctaccagggaaagggtt
atagggctgaaaaggaatatcagtactttctatgttctccagcagagtgtagtgcatacc
atgttagagaaagttcaacatgtaatactgtgaggggccaaattgcctgggttgagtcat
gttaaccttttgccttttcctgttctttttactaacaggtgaagtggccctctggtcact
ggtagtattggccgttgaaagatatatggtggtctgcaagcccatggccaacttccgatt
cggggagaaccatgctattatgggtgtagccttcacatggatcatggctttgtcttgtgc
tgctcctcctctcttcggatggtccaggtaaatatatataacatcagtcagcatatccct
```

- Desde sitios web o local
- Transformación desde formato de texto plano a otro
- Obtención de porciones de secuencias
- ADN complementario Reverso
- …

emboss Identificación de ORF

GETORF : Encuentra y extrae Marcos Abiertos de lectura (ORFs)



● ggutierr@ibun.unal.edu.co your e-mail

(● = required, ● = conditionally required)

● sequence -- DNA [sequences] (-sequence) : please enter either :

1. the name of a file:
2. or the actual data here: embl:xlrhodop

(sequence format)

● outseq.out outseq (-outseq)

fasta Output format for: outseq

Standard

30 Minimum nucleotide size

Translation of regions between STOP codons

☑ Change initial START codons to Methionine (-m

☐ Is the sequence circular (-circular)

☑ Find ORFs in the reverse sequence (-reverse)

100 Number of flanking nucleotides to report (-flanking)

Código genético

0 : Standard
1 : Standard (with alternative initiation codons)
2 : Vertebrate Mitochondrial
3 : Yeast Mitochondrial
4 : Mold, Protozoan, Coelenterate Mitochondrial and Mycoplasma/Spiroplasma
5 : Invertebrate Mitochondrial
6 : Ciliate Macronuclear and Dasycladacean
9 : Echinoderm Mitochondrial
10 : Euplotid Nuclear
11 : Bacterial
12 : Alternative Yeast Nuclear
13 : Ascidian Mitochondrial
14 : Flatworm Mitochondrial
15 : Blepharisma Macronuclear

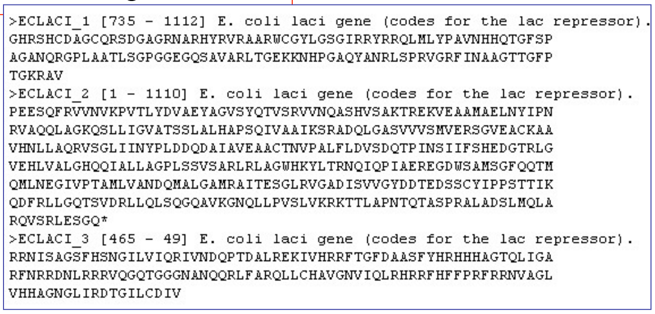
---

emboss Identificación de ORF

Código genético

0 : Standard
1 : Standard (with alternative initiation codons)
2 : Vertebrate Mitochondrial
3 : Yeast Mitochondrial
4 : Mold, Protozoan, Coelenterate Mitochondrial and Mycoplasma/Spiroplasma
5 : Invertebrate Mitochondrial
6 : Ciliate Macronuclear and Dasycladacean
9 : Echinoderm Mitochondrial
10 : Euplotid Nuclear
11 : Bacterial
12 : Alternative Yeast Nuclear
13 : Ascidian Mitochondrial
14 : Flatworm Mitochondrial
15 : Blepharisma Macronuclear

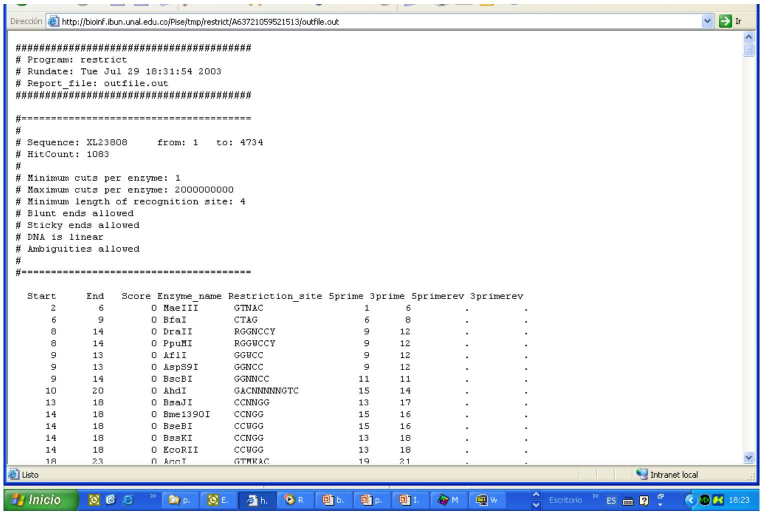
Salida: getorf



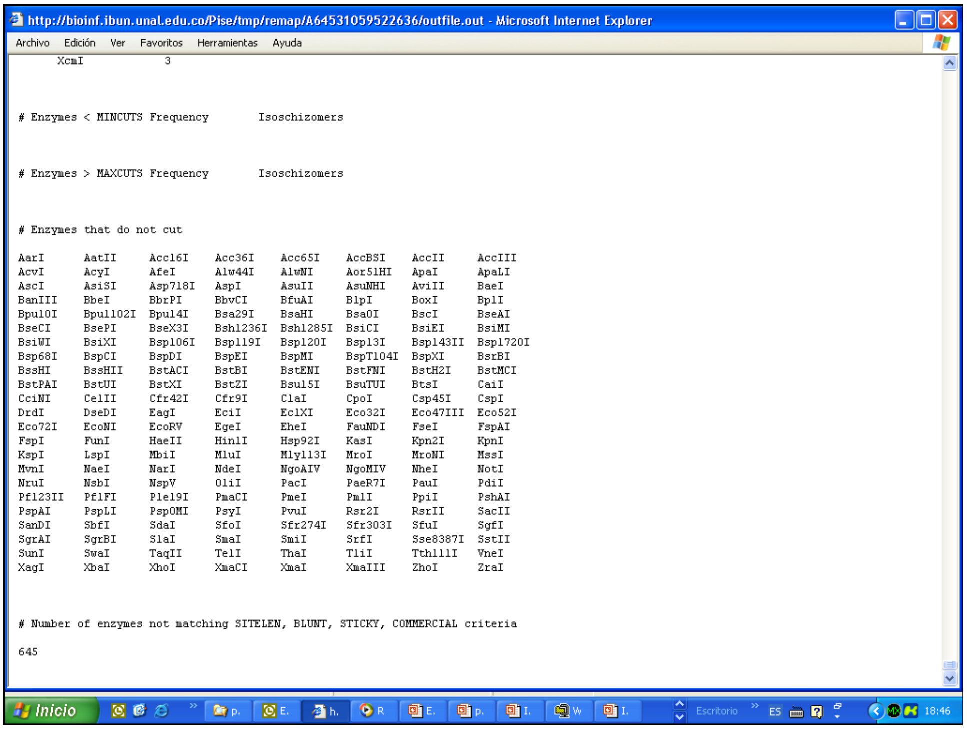
```
>ECLACI_1 [735 - 1112] E. coli laci gene (codes for the lac repressor).
GHRSHCDAGCQRSDGAGRNARHYRVRAARWCGYLGSGIRRYRRQLMLYPAVNHHQTGFSP
AGANQRGPLAATLSGPGGEGQSAVARLTGEKKNHPGAQYANRLSPRVGRFINAAGTTGFP
TGKRAV
>ECLACI_2 [1 - 1110] E. coli laci gene (codes for the lac repressor).
PEESQFRVVNVKPVTLYDVAEYAGVSYQTVSRVVNQASHVSAKTREKVEAAMAELNYIPN
RVAQQLAGKQSLLIGVATSSLALHAPSQIVAAIKSRADQLGASVVVSMVERSGVEACKAA
VHNLLAQRVSGLIINYPLDDQDAIAVEAACTNVPALFLDVSDQTPINSIIFSHEDGTRLG
VEHLVALGHQQIALLAGPLSSVSARLRLAGWHKYLTRNQIQPIAEREGDWSAMSGFQQTM
QMLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADISVVGYDDTEDSSCYIPPSTTIK
QDFRLLGQTSVDRLLQLSQGQAVKGNQLLPVSLVKRKTTLAPNTQTASPRALADSLMQLA
RQVSRLESGQ*
>ECLACI_3 [465 - 49] E. coli laci gene (codes for the lac repressor).
RRNISAGSFHSNGILVIQRIVNDQPTDALREKIVHRRFTGFDAASFYHRHHHAGTQLIGA
RFNRRDNLRRRVQGQTGGGNANQQRLFARQLLCHAVGNVIQLRHRRFHFFPRFRRNVAGL
VHHAGNGLIRDTGILCDIV
```

emboss

RESTRICT Encuentra los sitios de clivaje en una secuencia de ADN para enzimas de restricción


Dirección
http://bioinf.ibun.unal.edu.co/Pise/tmp/restrict/A63721059521513/outfile.out
Ir
Listo
Intranet local
Inicio
Escritorio
ES
18:23

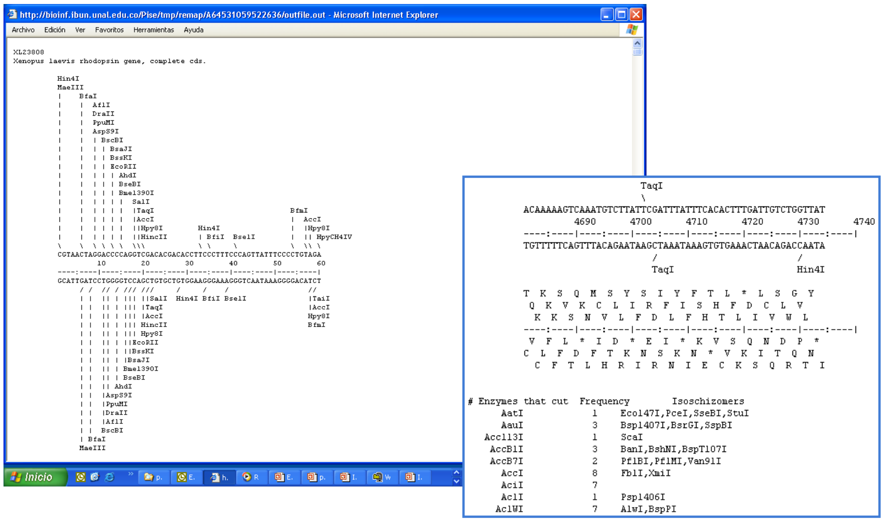
```
########################################
# Program: restrict
# Rundate: Tue Jul 29 18:31:54 2003
# Report_file: outfile.out
########################################

#=======================================
#
# Sequence: XL23808     from: 1   to: 4734
# HitCount: 1083
#
# Minimum cuts per enzyme: 1
# Maximum cuts per enzyme: 2000000000
# Minimum length of recognition site: 4
# Blunt ends allowed
# Sticky ends allowed
# DNA is linear
# Ambiguities allowed
#
#=======================================

  Start     End   Score Enzyme_name Restriction_site 5prime 3prime 5primerev 3primerev
      2       6       0 MaeIII      GTNAC                 1      6         .         .
      6       9       0 BfaI        CTAG                  6      8         .         .
      8      14       0 DraII       RGGNCCY               9     12         .         .
      8      14       0 PpuMI       RGGWCCY               9     12         .         .
      9      13       0 AflI        GGWCC                 9     12         .         .
      9      13       0 AspS9I      GGNCC                 9     12         .         .
      9      14       0 BscBI       GGNNCC               11     11         .         .
     10      20       0 AhdI        GACNNNNNGTC          15     14         .         .
     13      18       0 BsaJI       CCNNGG               13     17         .         .
     14      18       0 Bme1390I    CCNGG                15     16         .         .
     14      18       0 BseBI       CCWGG                15     16         .         .
     14      18       0 BssKI       CCNGG                13     18         .         .
     14      18       0 EcoRII      CCWGG                13     18         .         .
     18      23       0 AccI        GTMKAC               19     21         .         .
```


http://bioinf.ibun.unal.edu.co/Pise/tmp/remap/A64531059522636/outfile.out - Microsoft Internet Explorer
Archivo
Edición
Ver
Favoritos
Herramientas
Ayuda
Inicio
Escritorio
ES
18:46

```
       XcmI             3



# Enzymes < MINCUTS Frequency       Isoschizomers



# Enzymes > MAXCUTS Frequency       Isoschizomers



# Enzymes that do not cut

AarI      AatII     Acc16I    Acc36I    Acc65I    AccBSI    AccII     AccIII
AcvI      AcyI      AfeI      Alw44I    AlwNI     Aor51HI   ApaI      ApaLI
AscI      AsiSI     Asp718I   AspI      AsuII     AsuNHI    AviII     BaeI
BanIII    BbeI      BbrPI     BbvCI     BfuAI     BlpI      BoxI      BplI
Bpu10I    Bpu1102I  Bpu14I    Bsa29I    BsaHI     Bsa0I     BscI      BseAI
BseCI     BsePI     BseX3I    Bsh1236I  Bsh1285I  BsiCI     BsiEI     BsiMI
BsiWI     BsiXI     Bsp106I   Bsp119I   Bsp120I   Bsp13I    Bsp143II  Bsp1720I
Bsp68I    BspCI     BspDI     BspEI     BspMI     BspT104I  BspXI     BsrBI
BssHI     BssHII    BstACI    BstBI     BstENI    BstFNI    BstH2I    BstMCI
BstPAI    BstUI     BstXI     BstZI     Bsu15I    BsuTUI    BtsI      CaiI
CciNI     CelII     Cfr42I    Cfr9I     ClaI      CpoI      Csp45I    CspI
DrdI      DseDI     EagI      EciI      EclXI     Eco32I    Eco47III  Eco52I
Eco72I    EcoNI     EcoRV     EgeI      EheI      FauNDI    FseI      FspAI
FspI      FunI      HaeII     Hin1I     Hsp92I    KasI      Kpn2I     KpnI
KspI      LspI      MbiI      MluI      Mly113I   MroI      MroNI     MssI
MvnI      NaeI      NarI      NdeI      NgoAIV    NgoMIV    NheI      NotI
NruI      NsbI      NspV      OliI      PacI      PaeR7I    PauI      PdiI
Pfl23II   PflFI     Ple19I    PmaCI     PmeI      PmlI      PpiI      PshAI
PspAI     PspLI     Psp0MI    PsyI      PvuI      Rsr2I     RsrII     SacII
SanDI     SbfI      SdaI      SfoI      Sfr274I   Sfr303I   SfuI      SgfI
SgrAI     SgrBI     SlaI      SmaI      SmiI      SrfI      Sse8387I  SstII
SunI      SwaI      TaqII     TelI      ThaI      TliI      Tth111I   VneI
XagI      XbaI      XhoI      XmaCI     XmaI      XmaIII    ZhoI      ZraI



# Number of enzymes not matching SITELEN, BLUNT, STICKY, COMMERCIAL criteria

645
```

emboss

REMAP Muestra una secuencia con los sitios de restricción, traducción etc


http://bioinf.ibun.unal.edu.co/Pise/tmp/remap/A64531059522636/outfile.out - Microsoft Internet Explorer
Archivo
Edición
Ver
Favoritos
Herramientas
Ayuda
Inicio

```
XL23808
Xenopus laevis rhodopsin gene, complete cds.

          Hin4I
          MaeIII
          |    BfaI
          |    |  AflI
          |    |  DraII
          |    |  PpuMI
          |    |  AspS9I
          |    |  | BscBI
          |    |  | | BsaJI
          |    |  | | BssKI
          |    |  | | EcoRII
          |    |  | | | AhdI
          |    |  | | | BseBI
          |    |  | | | Bme1390I
          |    |  | | | |  SalI
          |    |  | | | |  |TaqI                               BfmI
          |    |  | | | |  |AccI                               |  AccI
          |    |  | | | |  ||Hpy8I        Hin4I                |  |Hpy8I
          |    |  | | | |  ||HincII       | BfiI  BselI        |  || HpyCH4IV
          \    \  \ \ \ \  \\\            \ \     \            \  \\ \
          CGTAACTAGGACCCCAGGTCGACACGACACCTTCCCTTTCCCAGTTATTTCCCCTGTAGA
                   10        20        30        40        50        60
          ----:----|----:----|----:----|----:----|----:----|----:----|
          GCATTGATCCTGGGGTCCAGCTGTGCTGTGGAAGGGAAAGGGTCAATAAAGGGGACATCT
               / /  // / /// ///     /     /    /                  //
               | |  || | ||| ||SalI  Hin4I BfiI BselI              |TaiI
               | |  || | ||| |TaqI                                 |AccI
               | |  || | ||| |AccI                                 Hpy8I
               | |  || | ||| HincII                                BfmI
               | |  || | ||| Hpy8I
               | |  || | ||EcoRII
               | |  || | ||BssKI
               | |  || | |BsaJI
               | |  || | Bme1390I
               | |  || | BseBI
               | |  || AhdI
               | |  |AspS9I
               | |  |PpuMI
               | |  |DraII
               | |  |AflI
               | |  BscBI
               | BfaI
               MaeIII
```

```
                               TaqI
                               \
          ACAAAAAGTCAAATGTCTTATTCGATTTATTTCACACTTTGATTGTCTGGTTAT
                   4690      4700      4710      4720      4730      4740
          ----:----|----:----|----:----|----:----|----:----|----:----|
          TGTTTTTCAGTTTACAGAATAAGCTAAATAAAGTGTGAAACTAACAGACCAATA
                                 /                         /
                                 TaqI                      Hin4I

          T  K  S  Q  M  S  Y  S  I  Y  F  T  L  *  L  S  G  Y
           Q  K  V  K  C  L  I  R  F  I  S  H  F  D  C  L  V
            K  K  S  N  V  L  F  D  L  F  H  T  L  I  V  W  L
          ----:----|----:----|----:----|----:----|----:----|----:----|
           V  F  L  *  I  D  *  E  I  *  K  V  S  Q  N  D  P  *
          C  L  F  D  F  T  K  N  S  K  N  *  V  K  I  T  Q  N
            C  F  T  L  H  R  I  R  N  I  E  C  K  S  Q  R  T  I


# Enzymes that cut  Frequency        Isoschizomers
      AatI            1    Eco147I,PceI,SseBI,StuI
      AauI            3    Bsp1407I,BsrGI,SspBI
   Acc113I            1    ScaI
    AccB1I            3    BanI,BshNI,BspT107I
    AccB7I            2    PflBI,PflMI,Van91I
      AccI            8    FblI,XmiI
      AciI            7
      AclI            1    Psp1406I
     AclWI            7    AlwI,BspPI
```

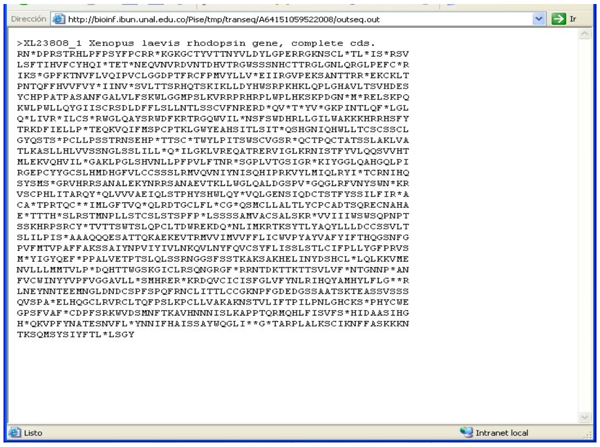
---

emboss

Traducción de secuencias

TRANSEQ

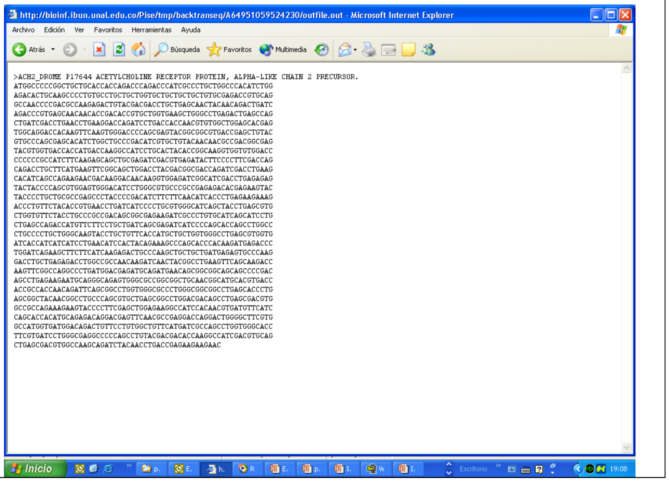
A uno o a todos los tres marcos en sentido delantero o reverso, o todos los seis marcos


Dirección
http://bioinf.ibun.unal.edu.co/Pise/tmp/transeq/A64151059522008/outseq.out
Ir
Listo
Intranet local

```
>XL23808_1 Xenopus laevis rhodopsin gene, complete cds.
RN*DPRSTRHLPFPSYFPCRR*KGKGCTYVTTNYVLDYLGPERRGKNSCL*TL*IS*RSV
LSFTIHVFCYHQI*TET*NEQVNVRDVNTDHVTRGWSSSNHCTTRGLGNLQRGLPEFC*R
IKS*GPFKTNVFLVQIPVCLGGDPTFRCFPMVYLLV*EIIRGVPEKSANTTRR*EKCKLT
PNTQFFHVVFVY*IINV*SVLTTSRHQTSKIKLLDYHWSRPKHKLQPLGHAVLTSVHDES
YCHPPATPASANFGALVLFSKWLGGMPSLKVRRPRHRPLWPLHKSKPDGN*M*RELSKPQ
KWLPWLLQYGIISCRSDLDFFLSLLNTLSSCVFNRERD*QV*T*YV*GKPINTLQF*LGL
Q*LIVR*ILCS*RWGLQAYSRWDFKRTRGQWVIL*NSFSWDHRLLGILWAKKKHRRHSFY
TRKDFIELLP*TEQKVQIFMSPCPTKLGWYEAHSITLSIT*QSHGNIQHWLLTCSCSSCL
GYQSTS*PCLLPSSTRNSEHP*TTSC*TWYLPITSWSCVGSR*QCTPQCTATSSLAKLVA
TLKASLLHLVVSSNGLSSLILL*Q*ILGKLVREQATRERVIGLKRNISTFYVLQQSVVHT
MLEKVQHVIL*GAKLPGLSHVNLLPFPVLFTNR*SGPLVTGSIGR*KIYGGLQAHGQLPI
RGEPCYYGCSLHMDHGFVLCCSSSLRMVQVNIYNISQHIPRKVYLMIQLRYI*TCRNIHQ
SYSMS*GRVHRRSANALEKYNRRSANAEVTKLLWGLQALDGSPV*GQGLRFVNYSWN*KR
VSCPHLITARQY*QLVVVAEIQLSTPHYSHWLQY*VQLGENSIQDCTSTFYSSILFIR*A
CA*TPRTQC**IMLGFTVQ*QLRDTGCLFL*CG*QSMCLLALTLYCPCADTSQRECNAHA
E*TTTH*SLRSTMNPLLSTCSLSTSPFP*LSSSSAMVACSALSKR*VVIIIWSWSQPNPT
SSKHRPSRCY*TVTTSWTSLQPCLTDWREKDQ*NLIMKRTKSYTLYAQYLLLDCCSSVLT
SLILPIS*AAAQQQESATTQKAEKEVTRMVVIMVVFFLICWVPYAYVAFYIFTHQGSNFG
PVFMTVPAFFAKSSAIYNPVIYIVLNKQVLNYFQVCSYFLISSLSTLCIFPLLYGFPRVS
M*YIGYQEF*PPALVETPTSLQLSSRNGGSFSSTKAKSAKHELINYDSHCL*LQLKKVME
NVLLLMMTVLP*DQHTTWGSKGICLRSQNGRGF*RRNTDKTTKTTSVLVF*NTGNNP*AN
FVCWINYYVPFVGGAVLL*SMHRER*KRDQVCICISFGLVFYNLRIHQYAMHYLFLG**R
LNEYNNTEEMNGLDNDCSPFSPQFRNCLITTLCCGKNPFGDEDGSSAATSKTEASSVSSS
QVSPA*ELHQGCLRVRCLTQFPSLKPCLLVAKAKNSTVLIFTPILPNLGHCKS*PHYCWE
GPSFVAF*CDPFSRKWVDSMNFTKAVHNNNISLKAPPTQRMQHLFISVFS*HIDAASIHG
H*QKVPFYNATESNVFL*YNNIFHAISSAYWQGLI**G*TARPLALKSCIKNFFASKKKN
TKSQMSYSIYFTL*LSGY
```

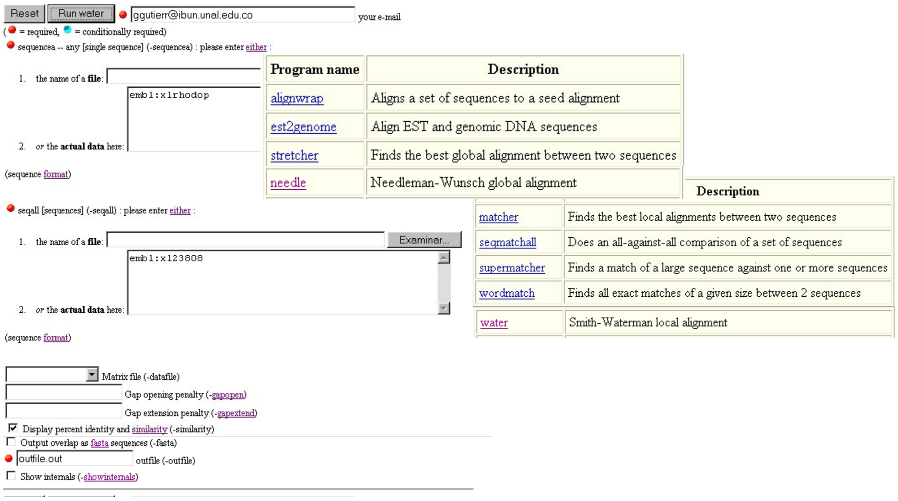
Traducción de secuencias

BACKTRANSEQ Retro-traduce una secuencia de proteínas


http://bioinf.ibun.unal.edu.co/Pise/tmp/backtranseq/A64951059524230/outfile.out - Microsoft Internet Explorer
Archivo Edición Ver Favoritos Herramientas Ayuda
Atrás
Búsqueda
Favoritos
Multimedia
Inicio

```
>ACH2_DROME P17644 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-LIKE CHAIN 2 PRECURSOR.
ATGGCCCCCGGCTGCTGCACCACCAGACCCAGACCCATCGCCCTGCTGGCCCACATCTGG
AGACACTGCAAGCCCCTGTGCCTGCTGCTGGTGCTGCTGCTGCTGTGCGAGACCGTGCAG
GCCAACCCCGACGCCAAGAGACTGTACGACGACCTGCTGAGCAACTACAACAGACTGATC
AGACCCGTGAGCAACAACACCGACACCGTGCTGGTGAAGCTGGGCCTGAGACTGAGCCAG
CTGATCGACCTGAACCTGAAGGACCAGATCCTGACCACCAACGTGTGGCTGGAGCACGAG
TGGCAGGACCACAAGTTCAAGTGGGACCCCAGCGAGTACGGCGGCGTGACCGAGCTGTAC
GTGCCCAGCGAGCACATCTGGCTGCCCGACATCGTGCTGTACAACAACGCCGACGGCGAG
TACGTGGTGACCACCATGACCAAGGCCATCCTGCACTACACCGGCAAGGTGGTGTGGACC
CCCCCCGCCATCTTCAAGAGCAGCTGCGAGATCGACGTGAGATACTTCCCCTTCGACCAG
CAGACCTGCTTCATGAAGTTCGGCAGCTGGACCTACGACGGCGACCAGATCGACCTGAAG
CACATCAGCCAGAAGAACGACAAGGACAACAAGGTGGAGATCGGCATCGACCTGAGAGAG
TACTACCCCAGCGTGGAGTGGGACATCCTGGGCGTGCCCGCCGAGAGACACGAGAAGTAC
TACCCCTGCTGCGCCGAGCCCTACCCCGACATCTTCTTCAACATCACCCTGAGAAGAAAG
ACCCTGTTCTACACCGTGAACCTGATCATCCCCTGCGTGGGCATCAGCTACCTGAGCGTG
CTGGTGTTCTACCTGCCCGCCGACAGCGGCGAGAAGATCGCCCTGTGCATCAGCATCCTG
CTGAGCCAGACCATGTTCTTCCTGCTGATCAGCGAGATCATCCCCAGCACCAGCCTGGCC
CTGCCCCTGCTGGGCAAGTACCTGCTGTTCACCATGCTGCTGGTGGGCCTGAGCGTGGTG
ATCACCATCATCATCCTGAACATCCACTACAGAAAGCCCAGCACCCACAAGATGAGACCC
TGGATCAGAAGCTTCTTCATCAAGAGACTGCCCAAGCTGCTGCTGATGAGAGTGCCCAAG
GACCTGCTGAGAGACCTGGCCGCCAACAAGATCAACTACGGCCTGAAGTTCAGCAAGACC
AAGTTCGGCCAGGCCCTGATGGACGAGATGCAGATGAACAGCGGCGGCAGCAGCCCCGAC
AGCCTGAGAAGAATGCAGGGCAGAGTGGGCGCCGGCGGCTGCAACGGCATGCACGTGACC
ACCGCCACCAACAGATTCAGCGGCCTGGTGGGCGCCCTGGGCGGCGGCCTGAGCACCCTG
AGCGGCTACAACGGCCTGCCCAGCGTGCTGAGCGGCCTGGACGACAGCCTGAGCGACGTG
GCCGCCAGAAAGAAGTACCCCTTCGAGCTGGAGAAGGCCATCCACAACGTGATGTTCATC
CAGCACCACATGCAGAGACAGGACGAGTTCAACGCCGAGGACCAGGACTGGGGCTTCGTG
GCCATGGTGATGGACAGACTGTTCCTGTGGCTGTTCATGATCGCCAGCCTGGTGGGCACC
TTCGTGATCCTGGGCGAGGCCCCCAGCCTGTACGACGACACCAAGGCCATCGACGTGCAG
CTGAGCGACGTGGCCAAGCAGATCTACAACCTGACCGAGAAGAAGAAC
```

Alineamiento de secuencias



● ggutierr@ibun.unal.edu.co your e-mail

(● = required, ● = conditionally required)

● sequencea -- any [single sequence] (-sequencea) : please enter either :

1. the name of a **file**:
2. *or* the **actual data** here: embl:xlrhodop

(sequence format)

● seqall [sequences] (-seqall) : please enter either :

1. the name of a **file**: Examinar...
2. *or* the **actual data** here: embl:x123808

(sequence format)

Matrix file (-datafile)

Gap opening penalty (-gapopen)

Gap extension penalty (-gapextend)

☑ Display percent identity and similarity (-similarity)

☐ Output overlap as fasta sequences (-fasta)

● outfile.out outfile (-outfile)

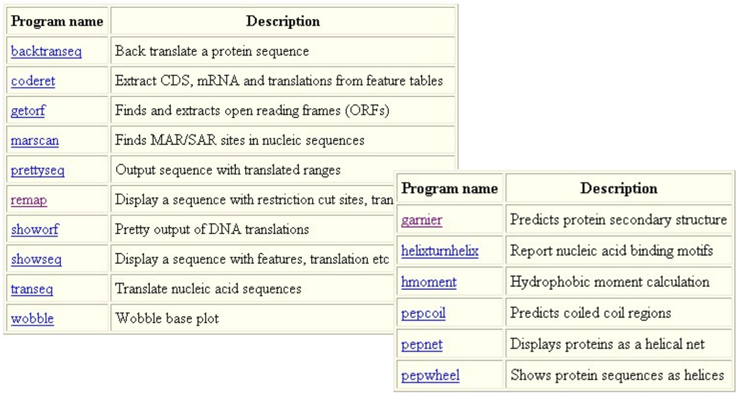
☐ Show internals (-showinternals)

Program name	Description
alignwrap	Aligns a set of sequences to a seed alignment
est2genome	Align EST and genomic DNA sequences
stretcher	Finds the best global alignment between two sequences
needle	Needleman-Wunsch global alignment

	Description
matcher	Finds the best local alignments between two sequences
seqmatchall	Does an all-against-all comparison of a set of sequences
supermatcher	Finds a match of a large sequence against one or more sequences
wordmatch	Finds all exact matches of a given size between 2 sequences
water	Smith-Waterman local alignment

emboss Análisis de Proteínas

Existen varios programas, entre ellos:



Program name	Description
backtranseq	Back translate a protein sequence
coderet	Extract CDS, mRNA and translations from feature tables
getorf	Finds and extracts open reading frames (ORFs)
marscan	Finds MAR/SAR sites in nucleic sequences
prettyseq	Output sequence with translated ranges
remap	Display a sequence with restriction cut sites, tran
showorf	Pretty output of DNA translations
showseq	Display a sequence with features, translation etc
transeq	Translate nucleic acid sequences
wobble	Wobble base plot

Program name	Description
garnier	Predicts protein secondary structure
helixturnhelix	Report nucleic acid binding motifs
hmoment	Hydrophobic moment calculation
pepcoil	Predicts coiled coil regions
pepnet	Displays proteins as a helical net
pepwheel	Shows protein sequences as helices