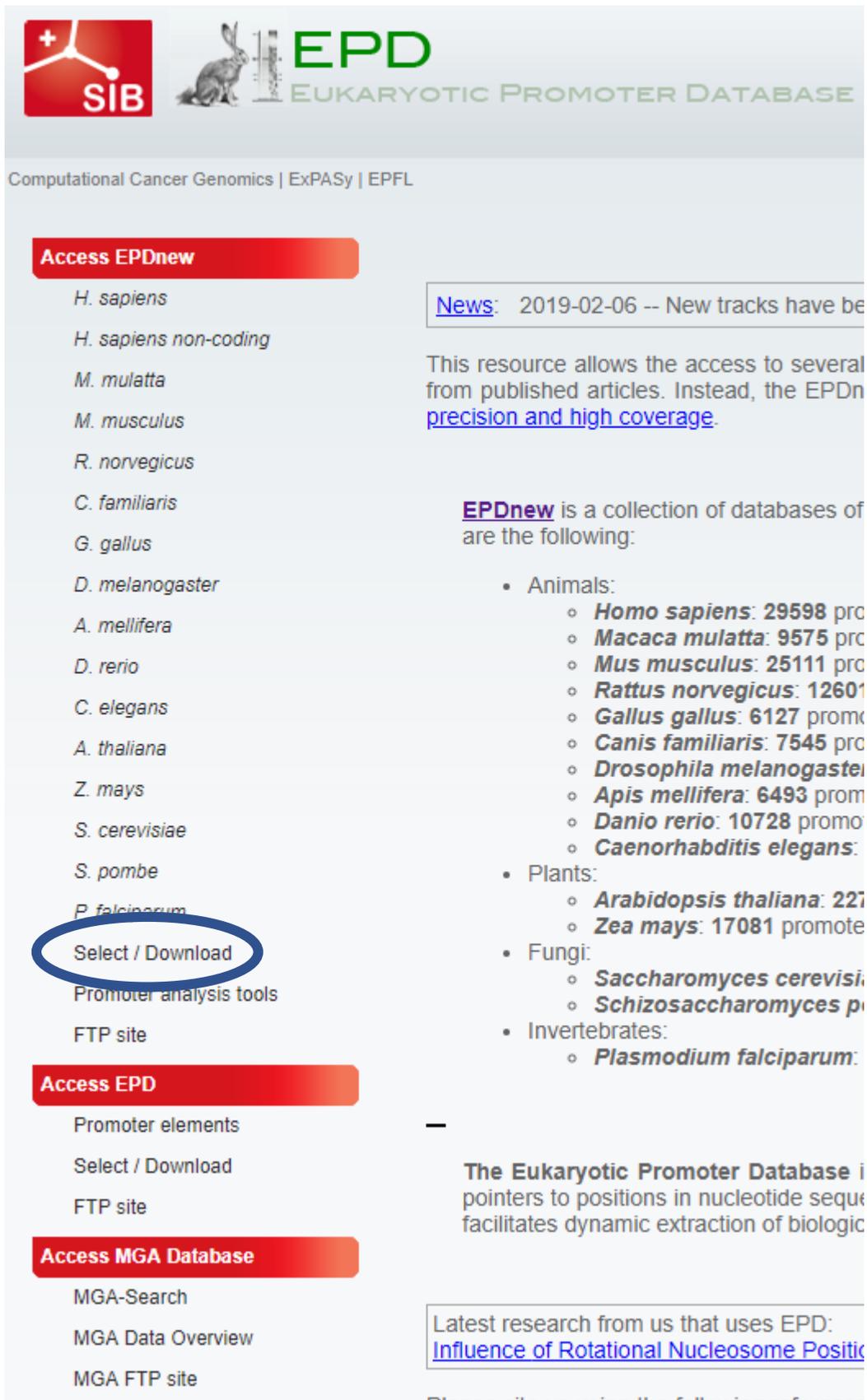


Navigate to <https://epd.epfl.ch//index.php>

Then click on Select / Download



SIB **EPD**
EUKARYOTIC PROMOTER DATABASE

Computational Cancer Genomics | ExPASy | EPFL

Access EPDnew

- [H. sapiens](#)
- [H. sapiens non-coding](#)
- [M. mulatta](#)
- [M. musculus](#)
- [R. norvegicus](#)
- [C. familiaris](#)
- [G. gallus](#)
- [D. melanogaster](#)
- [A. mellifera](#)
- [D. rerio](#)
- [C. elegans](#)
- [A. thaliana](#)
- [Z. mays](#)
- [S. cerevisiae](#)
- [S. pombe](#)
- [P. falciparum](#)
- Select / Download**
- [Promoter analysis tools](#)
- [FTP site](#)

Access EPD

- [Promoter elements](#)
- [Select / Download](#)
- [FTP site](#)

Access MGA Database

- [MGA-Search](#)
- [MGA Data Overview](#)
- [MGA FTP site](#)

News: 2019-02-06 -- New tracks have be

This resource allows the access to several from published articles. Instead, the EPDn [precision and high coverage](#).

EPDnew is a collection of databases of are the following:

- Animals:
 - **Homo sapiens: 29598** pro
 - **Macaca mulatta: 9575** pro
 - **Mus musculus: 25111** pro
 - **Rattus norvegicus: 12601** pro
 - **Gallus gallus: 6127** promo
 - **Canis familiaris: 7545** pro
 - **Drosophila melanogaste**
 - **Apis mellifera: 6493** prom
 - **Danio rerio: 10728** promo
 - **Caenorhabditis elegans:**
- Plants:
 - **Arabidopsis thaliana: 227**
 - **Zea mays: 17081** promote
- Fungi:
 - **Saccharomyces cerevisi**
 - **Schizosaccharomyces p**
- Invertebrates:
 - **Plasmodium falciparum:**

The Eukaryotic Promoter Database i pointers to positions in nucleotide sequ facilitates dynamic extraction of biologic

Latest research from us that uses EPD: [Influence of Rotational Nucleosome Positio](#)

Please cite us using the following referenc

Choose any combination of the four highlighted boxes and click 'select'

Select / Download tool

Use this tool to **select** promoters based on promoter name / ID or **liftOver** them to a different assembly or use them to perform full genome analysis.

Database

Restrict the selection to the following IDs:

Enter one ID per line

Promoters with the following characteristics:

TATA box

Initiator motif

CCAAT box

GC box

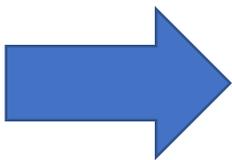
marked as

average expression of at least tags

expressed in at least samples

Additional options:

Select only the most representative promoter for a gene



Note how many promoters are selected (circled)

Choose the start and end base to extract then hit 'submit' to get the fasta

Database:		Selection Parameters	
Database:	human_epdnew	TATA-box:	with
Assembly:	hg38	Initiator:	with
		CCAAT-box:	with
		GC-box:	without
		Marked as:	all
		Average expression:	
		Expressed in:	

Results: 120 promoters selected [SGA file](#) [FPS file](#) [BED file](#)

LiftOver options: hg19 (Dec 2007 GRCh37)

Sequence Extraction Tool (FASTA format)		Downstream Analysis	
From:	<input type="text" value="50"/>	Motif Enrichment	<input type="button" value="OProf"/> ?
To:	<input type="text" value="120"/>	Motif Discovery	<input type="button" value="FindM"/> ?
<input type="button" value="Submit"/>		Chromatin analysis	<input type="button" value="ChIP-Cor"/> ?

Now we need to align the sequences we got. Copy the entire fasta output and go to <https://www.genome.jp/tools-bin/clustalw>

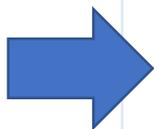
Paste the output in the text box and select the DNA check box, then hit 'Execute Multiple Alignment'.

```
>FP000093 TUBB3_1 :+U EU:NC; range -499 to 100.
GGGCCAGCCTTTACCTACCTCCCCACCCAAAACCGGCAAAAAGCTCAGAGCACCTTGTCT
GCCAAAAGACAGGGAGCTGGGATGGTGCGGGTTGGTCTCTAAACCGGCGTGGGGAAAAAA
GACCTTCCGTACAAAGCCGACGGGTGGGGCTGTCGCAAGGGCGGAACCGAGAGGGTAGCT
GGGGGCGGGGTTCCAGGGCCAAGAGGGGCCATTGTCTCCCTGGAGCCGGCGCCCCCA
CAGCCAGCTCCTCTGGGAGACAGCCCTCTTTTGAATGCGCGGGGCCCTCAGACCGCGC
CCGGCCACGCGCTGGGGATCCTTGGCTGCGGGAGGGGCGCCGATTGCGCGCGGCGGGC
GGGGACGCGCGGTGCGGAGCCTGCGGGCCGGGCGGGGCTCTGCGGCGGCGCCTCCCATT
GGCCACCCGCGGTGACATCAGCCGATGCGAAGGGCGGGGCGCGGCTATAAGAGCGCGG
GCCGCGGTCCCCGACCCTCAGCAGCCAGCCGCGCCGCGCCGCGCCGTCGCGAGCCGCC
GCCAGACGCGCCAGTATGAGGGAGATCGTGCACATCCAGGCCGGCCAGTGCGGCAACCA
>FP000052 STC1_1 :+U EU:NC; range -499 to 100.
ATGTACACACAGAGAAGATAGGGAGTTATTGCATTTGTAGCCTACAAAACAGAACCGAGA
ATGTGCTGTTAAAATTAGAGTAAACTGCTGTAAGCAGGTTAAGTTCTCATCTAAAGAGA
TCACATTTCCCCACCATAACCCCTGCTATCCATTTCCCCCAAGTGGCTCATTAGAAAAAA
GATGGCTAGATTTCAAAAAGCAACTTGGAGAGATTCTATAGGATTTTTCTTAGTTCAA
TCAATACAGAGTTATCTCTTACTTCCACGAAAATAGCTTTTTACACATCTCTGCACACA
CAGTCACACACACATATAAAACATTGGCAGCAGGTACTTTAATTTGCTGGAAAATATTT
CTAAGAAGTCAAAAAGCTCCAGCTGAATTGCATGCCCTCTTATTGGCTCACCAGACCAAT
TGAGGGACCTGATTGGTCTTGCATCTGAGGACCGATAAGAACGGCTATAAAATCCCTGG
GTGCAGCTCTTGGGCCCCAGTTTGCAAAAGCCAGAGGTGCAAGAAGCAGCGACTGCAGC
AGCAGCAGCAGCAGCGGGTGGCAGCAGCAGCAGCAGCGGCGGCAGCAGCAGCAGCAGC
```



Multiple Sequence Alignment by CLUSTALW

ETE3	MAFFT	CLUSTALW	PRRN
Help			
General Setting Parameters:			
Output Format: <input type="text" value="CLUSTAL"/>			
Pairwise Alignment: <input type="radio"/> FAST/APPROXIMATE <input checked="" type="radio"/> SLOW/ACCURATE			
Enter your sequences (with labels) below (copy & paste): <input type="radio"/> PROTEIN <input checked="" type="radio"/> DNA			
Support Formats: FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF			
<pre>>FP000009 MBD3L3_1 :+U EU:NC; range 0 to 100. ACTGCATTTTCCGGCAAGCCAAGGGTTGTCTGCATCTCAAGAGTGGGGTCAGCAAGAGAA ACTCTACGGCTATGGGAGAGCCTGCGTTTACCTCTTTTCCG >FP000007 MYH4_1 :+U EU:NC; range 0 to</pre>			
Or give the file name containing your query			
<input type="button" value="Choose File"/> No file chosen			
<input type="button" value="Execute Multiple Alignment"/> <input type="button" value="Reset"/>			



Finally, we'll take our multiple sequence alignment and create a sequence Logo using WebLogo. Copy the multiple sequence alignment (like the example Below), then go to <https://weblogo.berkeley.edu/logo.cgi>

Paste the multiple sequence alignment into the box at the top, then select DNA/RNA, increase the bitmap resolution and hit create logo

```

clustalw.aln
CLUSTAL 2.1 multiple sequence alignment

FP000009  -----ACTGCATTTTCC----GGCAAGCCAAGGGTTGTCTGCATCTC
FP000008  -----ACTGCATTTTCC----GGCAAGCCAAGGGTTGTCTGCATCTC
FP000005  -----ACTCTGCCTTTG----GACGTGAGAGAGAGCGCACCTTTTAC
FP000003  -----ACTCTGCCTTTG----GACGTGAGAGAGAGCGCACCTTTTAC
FP000001  -----ATCTGCTCTGACTCCCAGGGGACGTGTCTGTCTCTGCGTGTGAC
FP000006  AGCACAGTTGAGTCTCCAGCCTTGACTCTTC--TCAAGAGCCTGTGACTTTCTC----C
FP000002  -----ATCATCTTGGTCATCAACACAACTTGCTTCTCTCCAGACTTGGGCT
FP000004  -----AGCCTCTCCAGCCCCAGCAAGCGACCTGTCAAGCGCCGCTGGACT
FP000007  -----ATCCTTCTCAAATTTCTGAAGGTATGTATATGTG
FP000010  -----AGCAGACAGAGAGAGGAGTTGTCTGGACAGACTGCTCTT

FP000009  AAGAGTGGG-GTCAGCAAGAGAACTCTACGGCTATGGGAGAGC-CTGCGTTCACCTTT
FP000008  AAGAGTGGG-GTCAGCAAGAGAACTCTACGGCTATGGGAGAGC-CTGCGTTCACCTTT
FP000005  TTGAGCT----TCAACATGGGAAAG----GGAAATGAAGACCC-CGATCTCCACTGCTC
FP000003  TTGAGCT----TCAACATGGGAAAG----GGAAATGAAGACTC-CGATCTCCACTGCTC
CAGGGTGAGTGGCAACCTGGGATACCAGAGGGGTATGAGCAAGG-CAGAGGG-ATGG--
FP000001  CTGGACAAA-GGCATCATGAGTTGTC--AGATCTCTTGCAAATC-TCGAGG-----
FP000006  TAAGGTA----CTGCGTTTACATACAGCAAAATGCTATCATT-TACATTATCTAATCT
FP000002  CAGACTC----CGGAGATGAAGCCCTGCTCTGGCCGTGAGCC-TTGGCCTATTGCTG
FP000004  GAAGAACAC--TTACTTTTACATTCTGATGATGATTTTTTCAATTAAGGGATGTCTA
FP000007  GACAGAAGGTGCCAGGCTGGGG-GTGGCAGGCCGGGGGGGCTCTGGCCTGGGATGGAG
FP000010
  
```



[about](#) · [create](#) · [examples](#)

Multiple Sequence Alignment

Upload Sequence Data: No file chosen

Image Format & Size

Image Format: Logo Size per Line: X

Advanced Logo Options

Sequence Type: amino acid DNA / RNA Automatic Detection

First Position Number: Logo Range: -

Small Sample Correction: Frequency Plot:

Multiline Logo (Symbols per Line):

Advanced Image Options

Bitmap Resolution: 96 pixels/inch (dpi) Antialias Bitmaps:

Title: Y-Axis Height: (bits)

Show Y-Axis: Y-Axis Label:

Show X-Axis: X-Axis Label:

Show Error Bars: Label Sequence Ends:

Boxed / Boxed Shrink Factor: / Outline Symbols:

Show fine print: Y-Axis Tic Spacing: (bits)

Colors

Color Scheme: Default Black & White Custom (See Below.)

Symbols	Color	RGB	Symbols	Color	RGB
KRH	green	<input type="text"/>	<input type="text"/>	purple	<input type="text"/>
DE	blue	<input type="text"/>	<input type="text"/>	orange	<input type="text"/>
AVLIPWFM	red	<input type="text"/>	<input type="text"/>	black	<input type="text"/>
<input type="text"/>	black	<input type="text"/>	Other	black	<input type="text"/>

And there's a representative output! There are many parameters you can modify, for example if you have a very long sequence logo but you see something interesting, you can zoom in on that area by changing the Logo Range parameter.

You can also modify the multiple alignment parameters to get slightly

different alignment outputs. Play around and see how different combinations of promoters, alignment parameters and logo output parameters change your final image.

