

5º passo

- Voltar ao site do NCBI
- Atualizar as posições iniciais e finais nos quadrinhos
- A posição inicial vai no quadrinho 'from' e a posição final no 'to' (independente se é Forward ou Reverse, porque o calculo acima já estabelece o início e fim da sequência)
- clicar em Update View

Bos taurus isolate L1 Dominette 01449 registration number 42190680 bre - Nucleotide - NCBI - Google Chrome

ncbi.nlm.nih.gov/nuccore/NC_037343.1?report=fasta&from=60454698&to=60558562&strand=true

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Advanced Help

COVID-19 Information
Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

FASTA Send to

Bos taurus isolate L1 Dominette 01449 registration number 42190680 breed Hereford chromosome 16, ARS-UCD1.2, whole genome shotgun sequence

NCBI Reference Sequence: NC_037343.1
[GenBank](#) [Graphics](#)

```
>NC_037343.1:c60558562-60454698 Bos taurus isolate L1 Dominette 01449 registration number 42190680 breed Hereford chromosome 16, ARS-UCD1.2, whole genome shotgun sequence
GGTCGGGTGCGGGCAGGAGACGCTCTGTCCACGCTCTCCAGTGCTCGGTGGTTAAAGATGGCGGG
GCGGCAGCAGCGGAGCATTGGGGGCGCGGAGCCGCGGGATGGGAAGGGAAGGAGAACTGTGAGGCG
TCGGGAGGCAGCGCGGCGCTGTGTGAGGAGTGGGGTGGCGAGCGCAGGCTGCCGGACCGCTCAGGGC
CGGGGCTGGGTGGGAAAGGAGAGCCGGAGCAGCGCTGGAAACCCGAGGCCAGAGCCGAGGAGGAATG
TGACCAGGGGTCGTGGAGGCGCGGAGTAAGCTAGCGCAGGGATGGGGCAGCAGGTGGGCGCGCTCGGG
GAGGCTCCGGGGCTTCAAGCAGTCGCAACCCCGCGGATCCGGGGCAGCAGCGCAGCCAGGCTCTGGTC
GCAGGGGGACTTGGTGGGGCGCACCGCAGAGGCCGCTTCAATATCTTCAACCAGCATGGTGGTGTGT
CCGGGAGCCGGCTGAGGACGTGGGGGAGGGAGGAGGTTGGGGCGGGTGGGAGGCAGACAGGCA
GCCCTAGGGGGTCAAGCGGGTGGAGGGGTGCCGAGAGTGGGGGCGGAATCGGCTGCCATCTGGGCCAGT
CACGGGAGCGAGTCACACCCCTCCCGCTCCCATCTGGCTGGCAGCTCCGGGGGCGGAGGA
GGAAAGAGAAGTGGTCTTAGGATCTTGGATGGACTGTCGTCTACTGACACGACTGTGGCGCTCGGG
GGGGACGGTGATGGTATCCTCAGGGGAGAGGAGACTGAGGGAGGTTAAGCTTAATCCAGGTGTAGCA
TAAATAGGTTGTATCGAAGGGACCGTGAATTTGGCTGGAGGGATTGGGAAACAGGAACTGAATCTCAG
GACGTAGTGGTTTGTGAAGTACAGCAAGCAAGCCAGGCAATGAGCGGTATTAGGATAGTGAACAGTCCCT
GAAATACACTTCTATGGGAAAGGTAAGAATAGTAACTCAGCATTGGAATTTTTTCTTTTTCCCAAAG
TGGAACTGAAGGGTTTTGCTAAACCATCCTCTGTCTTCTCTTTTTGAAAGTGAAGTGAAGTGGCTCAGT
CGTGTCCGATCTTTGCGACCCATGGACTGTAGCATACCAAGGCTCTCAGTCCATGGGATTTTCCAGGC
```

Change region shown

Whole sequence
 Selected region

from: 60558262 to: 60561562

Update View

Customize view

Display options

Show reverse complement
 Show gap features

Update View

Analyze this sequence

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

6º passo

- O site vai gerar uma nova sequência com 3.301 bases
- Esta é a sequência do FT
- Copiar a sequência da primeira até última base

- Colar a sequência em um arquivo no word
- Revisão -> contar Palavras
- Verificar se os caracteres dão 3.301

FASTA

Showing 3.301kb region from base 60558262 to 60561562

Bos taurus isolate L1 Dominette 01449 registration number 42190680 breed Hereford chromosome 16, ARS-UCD1.2, whole genome shotgun sequence

NCBI Reference Sequence: NC_037343.1

GenBank Graphics

NC_037343.1:c60561562-60558262 Bos taurus isolate L1 Dominette 01449 registration number 42190680 breed Hereford chromosome 16, ARS-UCD1.2, whole genome shotgun sequence

```
GCACCTGACAGGAAAGGATTTTGTGTTTGTGATTTTGTGCTAAAGGTGAGGAATCTCAAAAGAA
AAGSCTAGTAGCAGGAGAAATTCAGCAGAAAGATGACCTGGATCCATCAAAATGTAGAGAAAGAAAT
GGACTAATATGAGGAGATATGTAAGACTGGCAATACCCCTGGAAATCCAGATGAGTCAAGACAGGGATCAG
TGAAGCACCCTACCTGTTCTTCTGCTTCTGAAAATTTGTGGTATTAATACAGAGGAAAGSTAAAGA
TGAATGATATCTTTGCTTCAATGTGCTCATTGATGACGATGGGTTGATTTTCAATGCACTCCCTGG
CCATGCTCAGGTGAGGCTGCTCCCTACTGACTGACCACTGCACCTCACTTTTTTTTTGATCTGGTT
TCTATACCTGCTGGTCTTCACTGGCTTTTCTGCTGCTGGAGTTTACAGCTGCTCCCAATTCAGCTT
GAGCTAGTACTACTGGGATTTTAAAGTGTTCATAGGTTTCAAGTGGTTCAAGTGTGTAACTCAGGTGAGTCA
TTGGGCTTGAATTTTGAAGATGAAATAGTGTCCACACCTAAGTAAAGTGTCTCCCTGGTCTAATGTC
CCCTCACTCCCAATCTGAAAATTTCCCAATTTCACTTTGAAAAATATCTATATTTGCTTAAAGTTTGTG
CTCCATCGTGTCTGACTTCTGCAATCCGTTGGACTATGGCCGTTAGGCTCCTCTGCTCAGTGGATCTC
CAAGCAAGAAATCTGAGGTGGGTTGCTCACTCTCCACGGGATCTTCCCAACCAAGGATCGAAGCCCA
CATCTCATGCTCTCCTGACTTGACAGGAGATTTTGACCACTGGGAGCCCTGCTCAAGTCTAGGCA
GGATCTAAGAAATCCATGGTGGAGAAATCTCATATGCTAAACAAGAAAGTATTTCAATATAAATAGTT
TGCATGTGACAGGACTGGTATGACCAAGTAGGATGCTCAATAAAGTATGATTCATGATTTTGAATG
TTGGTAGATTTGGTGTCTTTTGTGTTTAAAGTAAAGTGTGCAAGTGTGAGGTTGCTAGGCTCAAGCT
AATAAATCTGAGAAATTTTCCATGATCTAAGCATATTTGCAATTTGAAAAATGAGGCTGGAAATTTG
ATGAGCTCTGAAAATTAACATTTTTTGGTACTAATAGTCTTTTTATATCAAAATACTGTTTAAAG
TGCTTATTTACATGACCAATTTAATTTGTCAACAGCCCTAAGAGTGGTACTATGCTATCACCA
TCACCTCAGAGTAGGAACTGATGCAAAAGGTTAAATATTTGCTAAGAACAAATAGCAGCAATAGT
ATGTAATCAGAGCTCAGTTTTGAACTTAAAGCATGCTGGCTTAAAGTGGTCTTTTATATACAAAATACTGTTTAAAG
CTATACCCATCCATGACTGCTTTAAGAAATAGCAGAGAAACAGATCTGTTTCAAGTGGATCTC
TCAGATAAATAAACAACGAAACCAATCTGGCTCTATGATCACTTTTATCCCTTACAGTGTAT
CTATTTTTTGGCAACACTTTGGGAAATTAAGGAAATTTAAGCAAGGAAACCTTTGATCC
AAATGAAAAATCATATGATCAAGTGTAAAGACAGTGTCTTTGCTTATTTGCTTATGTTGGCAATTTCA
GTTATTTCTGATGATACCTGGATGGCAATGGAATTTTCAATGAGCTTTTATTTGATGAGTAACTAAT
GGTCTCATCTCCTGAGATCTTTGATCTTCCAAAGGAGACTCACTCACTTTTCTCCGGACCTT
ATCTTCCCAAGGTTGTTTGTGGTTTTCATGCTTACTGCTTCAAGCAATTTTCAAT
```

Documento1 - Microsoft Word (Falha na Ativação do Produto)

Inserir Layout da Página Referências Correspondências Revisão Exibição

Contar Palavras

Final: Mostrar Marcação

Mostrar Marcações

Anterior Próximo

Aceitar Rejeitar

Comparar Bloquear Restringir Autores Edição

Controle Alterações Comparar Proteger

Contar palavras

Estáticas:

Páginas	1
Palavras	48
Caracteres (sem espaços)	3.301
Caracteres (com espaços)	3.301
Parágrafos	48
Linhas	48

Incluir caixas de texto, notas de rodapé e notas de fim

Fechar

Português (Brasil)

Colar seqüências no bloco de notas

O bloco de notas ficará assim:

>Nomedogene1 | hg19

ACT.....
..... (toda a seqüência)

>Nomedogene2 | hg19

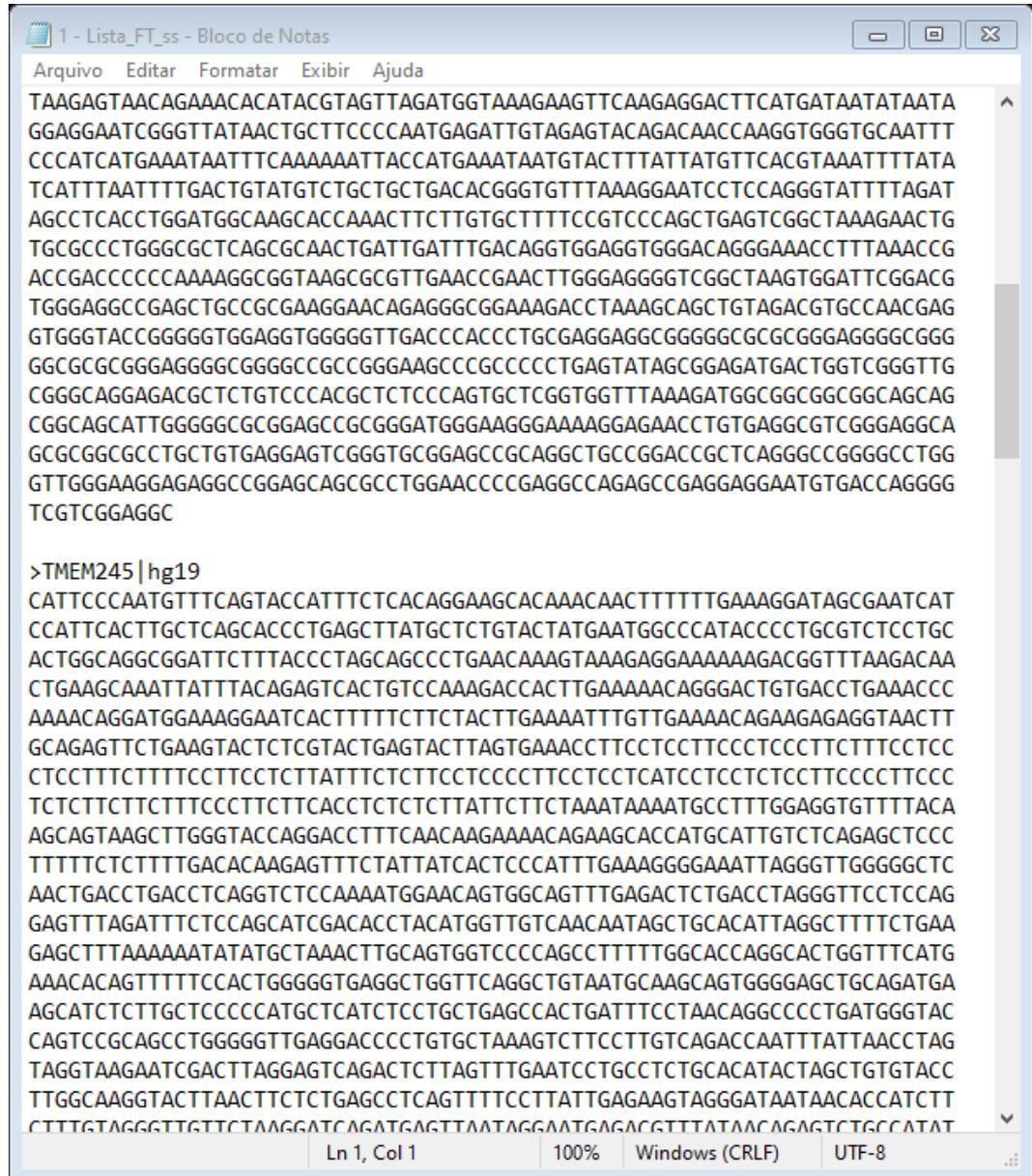
ACT.....
..... (toda a seqüência)

Sempre lembrar de deixar uma linha em branco de intervalo entre o final de uma seqüência e o início do nome do gene seguinte (exemplo imagem ao lado)

- As seqüências são formadas com as letras A, C, T e G.
- Quando tiver um “N” na seqüência é *missing* (a base nitrogenada daquela posição não foi identificada)
- Tem que substituir o “N” por “.”

*Se algum gene tiver o nome com a letra N, tomar cuidado para não substituir/alterar o nome dos genes

7º passo



```
1 - Lista_FT_ss - Bloco de Notas
Arquivo  Editar  Formatar  Exibir  Ajuda

TAAGAGTAACAGAAACACATACGTAGTTAGATGGTAAAGAAGTTCAAGAGGACTTCATGATAATATAATA
GGAGGAATCGGGTTATAACTGCTTCCCAATGAGATTGTAGAGTACAGACAACCAAGGTGGGTGCAATTT
CCCATCATGAAATAATTTCAAAAAATTACCATGAAATAATGTACTTTATTATGTTACAGTAAATTTTATA
TCATTTAATTTTGACTGTATGCTGCTGCTGACACGGGTGTTTAAAGGAATCCTCCAGGGTATTTTAGAT
AGCCTCACCTGGATGGCAAGCACAAACTTCTTGCTTTTCCGTCACAGTGTAGTCCGGCTAAAGAACTG
TGCGCCCTGGGCGCTCAGCGCAACTGATTGATTTGACAGGTGGAGGTGGGACAGGGAAACCTTAAACCG
ACCGACCCCAAAAGGCGGTAAGCGCGTTGAACCGAACTTGGGAGGGGTCCGGCTAAGTGGATTCCGACG
TGGGAGGCCGAGCTGCCGCAAGGAACAGAGGGCGGAAAGACCTAAAGCAGCTGTAGACGTGCCAACGAG
GTGGGTACCGGGGTGGAGGTGGGGTTGACCCACCTGCGAGGAGGCGGGGGCGCCGGGAGGGGCGGG
GGCGCGGGGAGGGGGCGGGGCCCGGGAAAGCCCGCCCCCTGAGTATAGCGGAGATGACTGGTCGGGTTG
CGGCGAGGAGACGCTCTGTCCCAGCTCTCCAGTGTCTCGGTGGTTTAAAGATGGCGGCGGCAGCAG
CGGCAGCATTGGGGGCGCGAGCCGCGGGATGGGAAGGAAAAGGAGAACCCTGTGAGGCGTCGGGAGGCA
GC CGCGCCCTGTGTGAGGAGTCCGGTGC GGAGCCG CAGGCTGCCGGACCGCTCAGGGCCGGGGCTGG
GTTGGGAAGGAGAGGCCGGAGCAGCGCCTGGAACCCCGAGGCCAGAGCCGAGGAGGAATGTGACCAGGGG
TCGTCGGAGGC

>TMEM245 | hg19
CATTCCCAATGTTTCAGTACCATTTCTCACAGGAAGCACAAACAACCTTTTTGAAAGGATAGCGAATCAT
CCATTCACTTGCTCAGCACCTGAGCTTATGCTCTGTACTATGAATGGCCATACCCCTGCGTCTCCTGC
ACTGGCAGGCGGATTTTACCCTAGCAGCCCTGAACAAAGTAAAGAGGAAAAAGACGGTTTAAAGCAA
CTGAAGCAAATTTTACAGAGTCACTGTCAAAGACCACTTAAAAACAGGGACTGTGACCTGAAACCC
AAAAAGGATGAAAAGGAATCACTTTTCTTCTACTTAAAAATTTGTTGAAAACAGAAGAGAGGTAACCT
GCAGAGTTCTGAAGTACTCTCGTACTGAGTACTTAGTGAAACCTTCTCCTTCCCTCCCTTCTTCTCCTC
CTCCTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
TCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
AGCAGTAAGCTTGGGTACCAGGACCTTTCAACAAGAAAAAGAGCACCATGCATTGTCTCAGAGCTCCC
TTTTTCTTTTTGACACAAGAGTTTCTATTATCACTCCCATTTGAAAGGGGAAATTAGGGTTGGGGGCTC
AACTGACCTGACCTCAGGCTCCTCAAAATGGAACAGTGGCAGTTTGTAGACTCTGACCTAGGGTTCTCCAG
GAGTTTAGATTTCTCCAGCATCGACACCTACATGGTTGTCAACAATAGCTGCACATTAGGCTTTTCTGAA
GAGCTTTAAAAATATATGCTAAACTTGCAGTGGTCCCAGCCTTTTTGGCACCAGGCACTGGTTTCATG
AAACACAGTTTTTCCACTGGGGGTGAGGCTGGTTGAGGCTGTAATGCAAGCAGTGGGGAGCTGCAGATGA
AGCATCTCTTGCTCCCCATGCTCATCTCTGCTGAGCCACTGATTTCTAACAGGCCCTGATGGGTAC
CAGTCCGACGCTGGGGTTGAGGACCCCTGTGCTAAAGTCTTCTTGTGAGCAGCAATTTATTAACCTAG
TAGGTAAGAATCGACTTAGGAGTCAAGTCTTGTAGTTGAAATCCCTGCTGCACATACTAGCTGTGTACC
TTGGCAAGGTAAGTAACTTCTGAGCCTCAGTTTTCTTATTGAGAAGTAGGGATAAATAACACCATCTT
CTTGTAGGGTTGTTCTAAGGATCAGATGAGTTAATAGGAATGAGACGTTTATAACAGAGTCTGCCATAT
```

- No fim, salvar a planilha com essas informações, por precaução

Gene	Sentido	Pos. inicial Original	Nova pos. inicial	Nova pos final
ABL2	Reverse	60558562	60558262	60561562
TMEM245	Reverse	98686213	98685913	98689213
LOC101907107	Reverse	61982951	61982651	61985951
PEAK1	Reverse	32554681	32554381	32557681

TFM-Explorer

TFM-Explorer

web server

help

examples

download

[home](#) > [TFM](#) > [tfm-explorer](#)

What is TFM-Explorer ?

TFM-Explorer (**Transcription Factor Matrix Explorer**) is a program for analysing regulatory regions in eukaryotic genomes.

It takes a set of coregulated gene sequences, and searches for locally overrepresented transcription factor binding sites.

The algorithm proceeds in two steps:

- scans sequences for detecting all potential transcription factor binding sites, using weight matrices from **JASPAR** or **TRANSFAC** ([disclaimer](#))
- extracts significant clusters (region of the input sequences associated with a factor) by calculating a score function.

TFM EXPLORER – WEB BROWSER

<https://bioinfo.lifl.fr/tfm-explorer/tfm-explorer.php>

PESQUISAR OS NOMES DOS FATORES DE TRANSCRIÇÃO – 8º PASSO

No site TFM – Explorer, clicar em web server

TFM-Explorer

TFM-Explorer

web server

help

examples

download

[home](#) > [TFM](#) > [tfm-explorer](#)

Enter regulatory sequences [?]

Run TFM-Explorer

Enter the sequences name (optional)

Available organisms: *rat (rn4)* *mouse (mm9)* *human (hg19)* *chicken (galGal3)* *d.melanogaster (dm3)*

Enter a list of RefSeq identifiers (NM_*)

OR

Enter a set of sequences in FASTA format [?]

Paste your data

or upload a file

Paste your data

or upload a file

Enter the sequences name (optional)

Deixar em branco

Available organisms: rat (rn4) mouse (mm9) human (hg19) chicken (galGal3) d.melanogaster (dm3)

Enter a list of RefSeq identifiers (NM_*)

Paste your data

or upload a file

Escolher arquivo

Nenhum arquivo selecionado

Example

OR

Enter a set of sequences in FASTA format [?]

Paste your data

or upload a file

Escolher arquivo

1 - Lista_FT_ss.txt

Anexar arquivo .txt
montado na etapa
anterior

Location [?]

-3000

+300

(boundaries of the sequences in regard of the TSS. authorized values : -10000 to 5000)

Deixar -3000 e +300

Select transcription factor binding profiles [?]

Available weight matrices are public [TRANSFAC \(6.0\)](#) matrices and [JASPAR 2009](#) vertebrate matrices

Use public TRANSFAC matrices

only vertebrates
only insects
all TRANSFAC

Use all JASPAR vertebrate matrices

Use all available TRANSFAC and JASPAR matrices

Select weight matrices in the list below

B\$CRP_C
F\$ABAA_01
F\$ABF_C
F\$ABF1_01
F\$ADR1_01

Upload a file containing a list of weight matrix identifiers [?]

Escolher arquivo Nenhum arquivo selecionado

Não alterar
nada nesta
parte

Adjust parameters [?]

Number of clusters to display [?] (max 25)

Maximum P-value [?]

Ratio (density of clusters) [?]

Diminuir o máximo P- value só se
tiver muitos (?) FT
Se não, deixar como está

Enter your E-mail address (optional):

Reset

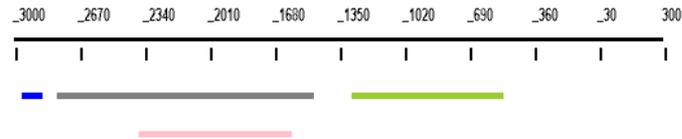
Run TFM-Explorer

Your request has been successfully submitted to **TFM-EXPLORER**
 Your ID is **Nov_02_2021_00_03_57_9363**



Results for job Nov_02_2021_00_03_57_9363 [?]

4 clusters found. The figure below represents the clusters on the input sequences, each of them being described in the table. Click on a cluster for detailed information.



Click on a line to get more detailed information for a cluster [?].

Rank	Factor	Matrix ID	Location	Sequences	P-Value	Correlated with	Legend	Select
1	Esrrb	MA0141.1	+[-2957:-2851]	2	3.39e-05			<input type="checkbox"/>
2	FEV	MA0156.1	+[-2360:-1575]	3	4.67e-05	4		<input type="checkbox"/>
3	FOXO3	MA0157.1	+[-2779:-1467]	4	4.94e-05			<input type="checkbox"/>
4	ELK4	MA0076.1	+[-1283:-850]	3	8.90e-05	2		<input type="checkbox"/>

View transcription factor binding sites associated to selected clusters [?] [view](#)

Summary of pairwise correlations between clusters [?]

correlation coefficient

<0.5

>0.5 and <0.7

>0.7

	1	2	3	4
1				
2				
3				
4				

Na nova página que carrega, tem a opção:

Download results [?]:

Clicar na opção gerada em 'summary in text format'

Download results [?] : complete result files [Nov_02_2021_00_03_57_9363.zip](#)
 summary in text format [Nov_02_2021_00_03_57_9363.txt](#)

Sequences Name	None
Number of sequences	4
Number of matrices	130
Database used	JASPAR
Region	-3000 : 300
Date	Tue Nov 2 00:04:23 2021
Minimum window size	30
Maximum window size	1500
Maximum number of windows to show	25
Minimal P-value	0.0001
Ratio	3.0
URL of results	http://bioinfo.lifl.fr/cgi-bin/tfm-explorer/tfme.py?command=result&run_id=Nov_02_2021_00_03_57_9363

Vai carregar uma nova guia com os resultados (próximo slide)

```
;TFM Explorer
;Date : Tue Nov 2 00:04:23 2021
;Scanned location : -3000:300
;Scanned sequences : None (4 sequences)
;Scanned matrices : (130 matrices)
;Parameters : minsize= 30, maxsize= 1500, ratio=3.0, top=25
;
; headers
; 1 rank window rank
; 2 matrix matrix name
; 3 tf transcription factor name
; 4 info information of content of the matrix
; 5 gc gc content of the matrix (G+C)%
; 6 location location of the window (relatively to TSS)
; 7 pvalue pvalue of the window

1 MA0141.1 Esrrb 12.8061908542 0.523864734014 -2957:-2851 3.38522472321e-05
list of hits
ABL2 -2957 -2945 + 6.40 TGCTAAAGGTGA
ABL2 -2898 -2886 - 5.74 TGACCTGGATAT
LOC101907107 -2952 -2940 - 5.96 TGACCCTGTCCT
LOC101907107 -2918 -2906 - 6.49 TGACCCTGCCCT
LOC101907107 -2866 -2854 - 5.96 TGACCCTGTCCT
LOC101907107 -2851 -2839 - 7.50 TGACCCTGCCCT

2 MA0156.1 FEV 12.1208832256 0.442307695746 -2360:-1575 4.67270375297e-05
list of hits
TMEM245 -2323 -2315 + 5.26 GGGGAAAT
TMEM245 -1977 -1969 - 7.43 ATTTCCCTA
LOC101907107 -2360 -2352 - 7.02 ATTTCCGT
LOC101907107 -2274 -2266 + 7.96 AAGGAAGT
LOC101907107 -2096 -2088 + 6.34 TCGGAAAT
LOC101907107 -2052 -2044 - 7.43 ATTTCCCTC
LOC101907107 -1602 -1594 - 5.11 ACTTCCCA
PEAK1 -1991 -1983 - 5.26 ATTTCCCC
PEAK1 -1798 -1790 - 5.26 ATTTCCCA
PEAK1 -1765 -1757 - 7.43 ATTTCCCTA
PEAK1 -1686 -1678 - 7.28 ACTTCCTC
PEAK1 -1617 -1609 - 5.26 ATTTCCCC
PEAK1 -1575 -1567 + 7.28 GAGGAAGT
```

genes

genes

Estão circulados em destaque os nomes dos fatores de transcrição

Para cada FT tem a lista de genes e as informações associadas à cada um.

O nome de cada FT e dos genes associados à cada um serão necessários para o próximo passo.

9º passo

Montar uma planilha no Excel com duas colunas:

1ª coluna: Nome do fator de transcrição (FT)

2ª coluna: Nomes dos genes associados ao FT

Não importa se tiver que repetir os nomes

- Não colocar cabeçalho na planilha

genes

1	MA0141.1	Esrrb	12.8061
list of hits			
	ABL2	-2957	-2945
	ABL2	-2898	-2886
	LOC101907107		-2952
	LOC101907107		-2918
	LOC101907107		-2866
	LOC101907107		-2851
2	MA0156.1	FEV	12.1208
list of hits			
	TMEM245	-2323	-2315
	TMEM245	-1977	-1969
	LOC101907107		-2360
	LOC101907107		-2274
	LOC101907107		-2096
	LOC101907107		-2052
	LOC101907107		-1602
	PEAK1	-1991	-1983
	PEAK1	-1798	-1790
	PEAK1	-1765	-1757
	PEAK1	-1686	-1678
	PEAK1	-1617	-1609
	PEAK1	-1575	-1567

genes

•
•
•



	A	B	C	D
1	Esrrb	ABL2		
2	Esrrb	ABL2		
3	Esrrb	LOC101907107		
4	Esrrb	LOC101907107		
5	Esrrb	LOC101907107		
6	Esrrb	LOC101907107		
7	FEV	TMEM245		
8	FEV	TMEM245		
9	FEV	LOC101907107		
10	FEV	LOC101907107		
11	FEV	LOC101907107		
12	FEV	LOC101907107		
13	FEV	LOC101907107		
14	FEV	PEAK1		
15	FEV	PEAK1		
16	FEV	PEAK1		
17	FEV	PEAK1		
18	FEV	PEAK1		
19	FEV	PEAK1		
20	FOXO3	ABL2		
21	FOXO3	ABL2		
22	FOXO3	ABL2		
23	FOXO3	ABL2		
24	FOXO3	TMEM245		
25	FOXO3	TMEM245		

Obtendo as vias enriquecidas

BINGO

Para rodar o BiNGO, precisei ocultar a barra de tarefas do computador, ou então não daria para ver a tela toda da análise

Para ocultar a barra de tarefas:

- Clicar com o botão direito na barra de tarefas ou pesquisar
- Configurações da barra de tarefas
- Ativar o modo de ocultar a barra de tarefas da área de trabalho

Barra de Tarefas

Bloquear a barra de tarefas

Ativado

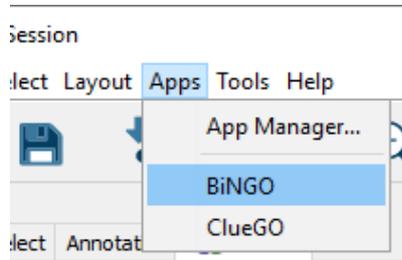
Ocultar automaticamente a barra de tarefas no modo de área de trabalho

Ativado

Ocultar automaticamente a barra de tarefas no modo tablet

Desativado

BiNGO



Clicar em Apps → BiNGO

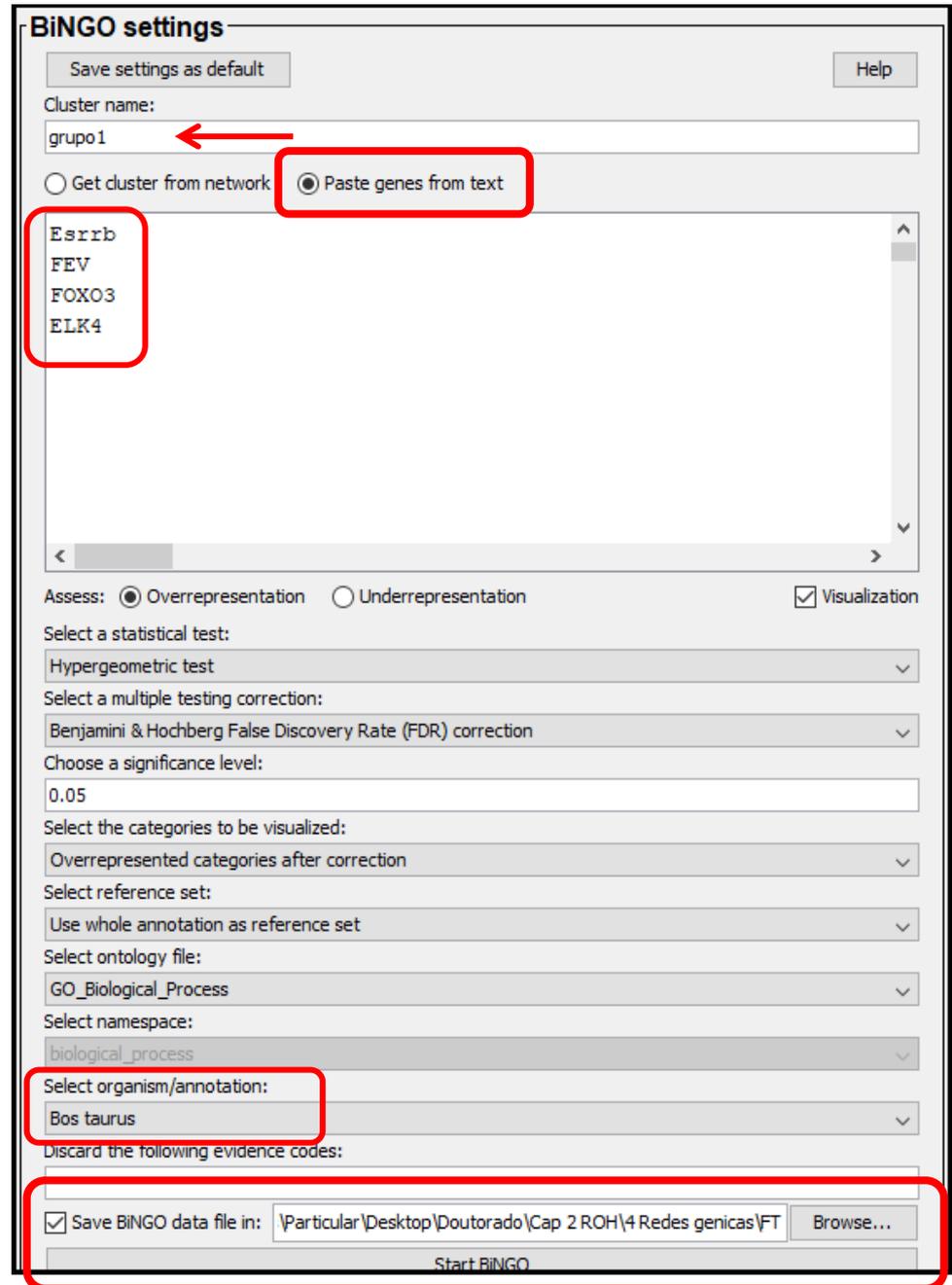
Às vezes tem que ocultar a barra de tarefas da área de trabalho pra poder visualizar toda a tela que aparece

Alterar apenas os parâmetros destacados

- Colocar algum nome em 'Cluster name'
- Selecionar 'Paste genes from text'
- Colocar os nomes dos FT na caixa de texto
- Em 'Select organism/annotation' colocar *Bos taurus* ou a espécie estudada

Nessa parte, se não tiver a espécie de interesse, colocar Homo Sapiens + descrição

- Marcar a opção 'Save BiNGO data file in'
- Browse (pesquisar a pasta para salvar as saídas)
- Clicar em Start BiNGO



Vai aparecer a tela com os resultados dos processos biológicos

The screenshot shows a Windows desktop with several application windows. The primary window is 'BiNGO output', which displays a table of results for a cluster named 'grupo1'. The table includes columns for GO ID, GO Description, p-val, Corrected p-val, Cluster frequency, Total frequency, and Genes. Below the table, there are controls for 'Select All', 'Unselect All', and 'Select nodes'. To the right of the table, there are settings for 'Overrepresented categories after correction', 'Select reference set', 'Select ontology file', 'Select namespace', 'Select organism/annotation', and 'Discard the following evidence codes'. A 'Save BINGO data file in:' field is also present.

GO ID	GO Description	p-val	Corrected p-val	Cluster frequency	Total frequency	Genes
<input type="checkbox"/> 48589	developmental growth	6.6832E-5	8.1977E-3	2/3 66.6%	40/8356 0.4%	ESRRB FOXO3
<input type="checkbox"/> 40007	growth	2.4996E-4	8.1977E-3	2/3 66.6%	77/8356 0.9%	ESRRB FOXO3
<input type="checkbox"/> 1544	initiation of primordial ovarian follicle growth	3.5902E-4	8.1977E-3	1/3 33.3%	1/8356 0.0%	FOXO3
<input type="checkbox"/> 1556	oocyte maturation	3.5902E-4	8.1977E-3	1/3 33.3%	1/8356 0.0%	FOXO3
<input type="checkbox"/> 1831	trophoblast cellular morphogenesis	3.5902E-4	8.1977E-3	1/3 33.3%	1/8356 0.0%	ESRRB
<input type="checkbox"/> 1834	trophoblast cell proliferation	3.5902E-4	8.1977E-3	1/3 33.3%	1/8356 0.0%	ESRRB
<input type="checkbox"/> 6355	regulation of transcription, DNA-dependent	6.2673E-4	8.9419E-3	3/3 100.0%	716/8356 8.5%	FEV ESRRB FOXO3
<input type="checkbox"/> 51252	regulation of RNA metabolic process	6.7803E-4	8.9419E-3	3/3 100.0%	735/8356 8.7%	FEV ESRRB FOXO3
<input type="checkbox"/> 9994	oocyte differentiation	7.1796E-4	8.9419E-3	1/3 33.3%	2/8356 0.0%	FOXO3
<input type="checkbox"/> 1547	antral ovarian follicle growth	7.1796E-4	8.9419E-3	1/3 33.3%	2/8356 0.0%	FOXO3
<input type="checkbox"/> 48599	oocyte development	7.1796E-4	8.9419E-3	1/3 33.3%	2/8356 0.0%	FOXO3
<input type="checkbox"/> 1542	ovulation from ovarian follicle	1.0768E-3	1.1348E-2	1/3 33.3%	3/8356 0.0%	FOXO3
<input type="checkbox"/> 30728	ovulation	1.0768E-3	1.1348E-2	1/3 33.3%	3/8356 0.0%	FOXO3
<input type="checkbox"/> 45595	regulation of cell differentiation	1.1896E-3	1.1641E-2	2/3 66.6%	168/8356 2.0%	ESRRB FOXO3
<input type="checkbox"/> 45449	regulation of transcription	1.5892E-3	1.3105E-2	3/3 100.0%	976/8356 11.6%	FEV ESRRB FOXO3
<input type="checkbox"/> 45648	positive regulation of erythrocyte differentiation	1.7943E-3	1.3105E-2	1/3 33.3%	5/8356 0.0%	FOXO3

Below the table, there are buttons for 'Select All', 'Unselect All', and 'Select nodes'. To the right of the table, there are settings for 'Overrepresented categories after correction', 'Select reference set', 'Select ontology file', 'Select namespace', 'Select organism/annotation', and 'Discard the following evidence codes'. A 'Save BINGO data file in:' field is also present.

The 'BiNGO settings' window is open above the main window, showing 'Cluster name: grupo1' and buttons for 'Save settings as default' and 'Help'. The 'BiNGO output' window has a title bar 'grupo1' and a subtitle 'Annotation: Curator = GO, Species or file = Bos taurus, Type = default Ontology: Curator = bingo, Type = namespace'. A 'Close tab' button is visible in the top right of the main window.

In the bottom right corner, there is a 'grupo1 Color Sc...' window showing a color scale legend with values '5.00E-2' and '< 5.00E-7'.

Os resultados do BiNGO serão salvos na pasta no formato .bgo



- Abrir esse arquivo com o Notepad++
- Selecionar os resultados dos processos biológicos e passar para uma planilha do Excel

```
1 File created with BiNGO v3.0.5 (c) on 01/11/2021 at 23:30:46
2
3 ontology: namespace
4 curator: bingo
5
6 Selected ontology file : bundle://113.0:1/data/GO_Biological_Process
7 Selected annotation file : bundle://113.0:1/data/B_taurus_default
8 Discarded evidence codes :
9 Overrepresentation
10 Selected statistical test : Hypergeometric test
11 Selected correction : Benjamini & Hochberg False Discovery Rate (FDR) correction
12 Selected significance level : 0.05
13 Testing option : Use whole annotation as reference set
14 The selected cluster :
15 ELK4 FEV ESRRB FOXO3
16
17 No annotations were retrieved for the following entities:
18 ELK4
19
20 GO-ID p-value corr p-value x n X N Description Genes in test set
21 48589 6.6832E-5 8.1977E-3 2 40 3 8356 developmental growth ESRRB|FOXO3
22 40007 2.4996E-4 8.1977E-3 2 77 3 8356 growth ESRRB|FOXO3
23 1544 3.5902E-4 8.1977E-3 1 1 3 8356 initiation of primordial ovarian follicle growth FOXO3
24 1556 3.5902E-4 8.1977E-3 1 1 3 8356 oocyte maturation FOXO3
25 1831 3.5902E-4 8.1977E-3 1 1 3 8356 trophectodermal cellular morphogenesis ESRRB
26 1834 3.5902E-4 8.1977E-3 1 1 3 8356 trophectodermal cell proliferation ESRRB
27 6355 6.2673E-4 8.9419E-3 3 716 3 8356 regulation of transcription, DNA-dependent FEV|ESRRB|FOXO3
28 51252 6.7803E-4 8.9419E-3 3 735 3 8356 regulation of RNA metabolic process FEV|ESRRB|FOXO3
29 9994 7.1796E-4 8.9419E-3 1 2 3 8356 oocyte differentiation FOXO3
30 1547 7.1796E-4 8.9419E-3 1 2 3 8356 antral ovarian follicle growth FOXO3
31 48599 7.1796E-4 8.9419E-3 1 2 3 8356 oocyte development FOXO3
32 1542 1.0768E-3 1.1348E-2 1 3 3 8356 ovulation from ovarian follicle FOXO3
33 30728 1.0768E-3 1.1348E-2 1 3 3 8356 ovulation FOXO3
34 45595 1.1896E-3 1.1641E-2 2 168 3 8356 regulation of cell differentiation ESRRB|FOXO3
35 45449 1.5892E-3 1.3105E-2 3 976 3 8356 regulation of transcription FEV|ESRRB|FOXO3
36 45648 1.7943E-3 1.3105E-2 1 5 3 8356 positive regulation of erythrocyte differentiation FOXO3
37 48468 1.8197E-3 1.3105E-2 2 208 3 8356 cell development ESRRB|FOXO3
38 10556 2.0133E-3 1.3105E-2 3 1056 3 8356 regulation of macromolecule biosynthetic process FEV|ESRRB|FOXO3
```

Normal text file length: 6.107 lines: 91 Ln: 91 Col: 1 Sel: 5.495 | 72

Arquivo Excel

GO-ID	GO Description	p-val	Corrected p-val	Cluster frequency	Total frequency	Genes
48589	developmental growth	6,68E-01	8,20E+01	2/3 66.6%	40/8356 0.4%	ESRRB FOXO3
40007	growth	2,50E+00	8,20E+01	2/3 66.6%	77/8356 0.9%	ESRRB FOXO3
1544	initiation of primordial ovarian follicle growth	3,59E+00	8,20E+01	1/3 33.3%	1/8356 0.0%	FOXO3
1556	oocyte maturation	3,59E+00	8,20E+01	1/3 33.3%	1/8356 0.0%	FOXO3
1831	trophectodermal cellular morphogenesis	3,59E+00	8,20E+01	1/3 33.3%	1/8356 0.0%	ESRRB
1834	trophectodermal cell proliferation	3,59E+00	8,20E+01	1/3 33.3%	1/8356 0.0%	ESRRB
6355	regulation of transcription, DNA-dependent	6,27E+00	8,94E+01	3/3 100.0%	716/8356 8.5%	FEV ESRRB FOXO3
51252	regulation of RNA metabolic process	6,78E+00	8,94E+01	3/3 100.0%	735/8356 8.7%	FEV ESRRB FOXO3
9994	oocyte differentiation	7,18E+00	8,94E+01	1/3 33.3%	2/8356 0.0%	FOXO3
1547	antral ovarian follicle growth	7,18E+00	8,94E+01	1/3 33.3%	2/8356 0.0%	FOXO3
48599	oocyte development	7,18E+00	8,94E+01	1/3 33.3%	2/8356 0.0%	FOXO3
1542	ovulation from ovarian follicle	1,08E+01	1,13E+02	1/3 33.3%	3/8356 0.0%	FOXO3
30728	ovulation	1,08E+01	1,13E+02	1/3 33.3%	3/8356 0.0%	FOXO3
45595	regulation of cell differentiation	1,19E+01	1,16E+02	2/3 66.6%	168/8356 2.0%	ESRRB FOXO3
45449	regulation of transcription	1,59E+01	1,31E+02	3/3 100.0%	976/8356 11.6%	FEV ESRRB FOXO3
45648	positive regulation of erythrocyte differentiation	1,79E+01	1,31E+02	1/3 33.3%	5/8356 0.0%	FOXO3
48468	cell development	1,82E+01	1,31E+02	2/3 66.6%	208/8356 2.4%	ESRRB FOXO3
10556	regulation of macromolecule biosynthetic process	2,01E+01	1,31E+02	3/3 100.0%	1056/8356 12.6%	FEV ESRRB FOXO3
48477	oogenesis	2,15E+01	1,31E+02	1/3 33.3%	6/8356 0.0%	FOXO3
31326	regulation of cellular biosynthetic process	2,25E+01	1,31E+02	3/3 100.0%	1096/8356 13.1%	FEV ESRRB FOXO3
10468	regulation of gene expression	2,29E+01	1,31E+02	3/3 100.0%	1103/8356 13.2%	FEV ESRRB FOXO3
9889	regulation of biosynthetic process	2,30E+01	1,31E+02	3/3 100.0%	1104/8356 13.2%	FEV ESRRB FOXO3
19219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	2,40E+01	1,31E+02	3/3 100.0%	1119/8356 13.3%	FEV ESRRB FOXO3
51171	regulation of nitrogen compound metabolic process	2,44E+01	1,31E+02	3/3 100.0%	1126/8356 13.4%	FEV ESRRB FOXO3

- Avaliar os processos biológicos (PB) dos FT
 - Selecionar os FT com os PB mais associados à(s) característica(s) avaliada(s)
 - Salvar os FT selecionados em outra planilha

Dica:

- No BiNGO também podemos fazer essa mesma análise considerando a espécie Homo sapiens ao invés da espécie animal estudada, porque geralmente tem mais resultados para humanos
- Neste caso deu apenas um resultado e foi bem específico para as características estudadas (Número de oócitos e embriões)

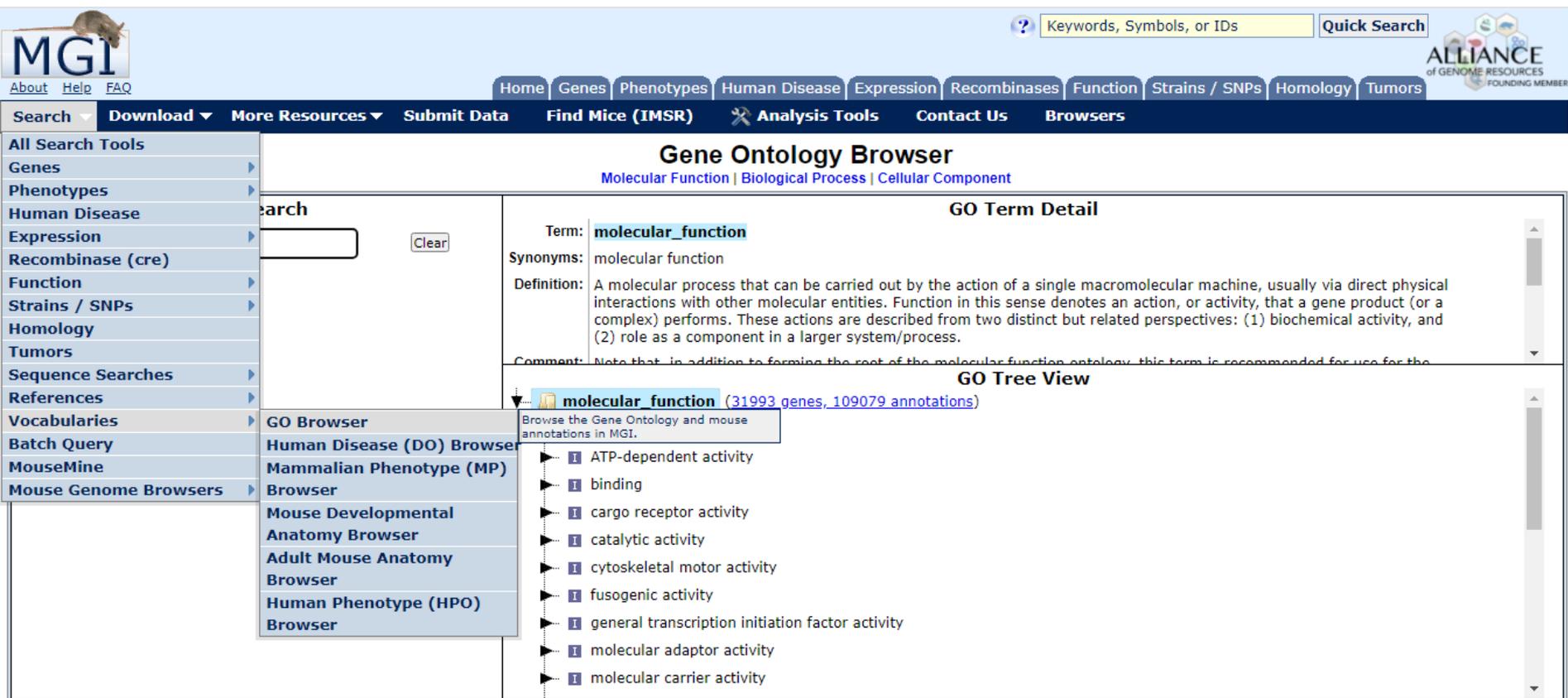
The screenshot displays the BiNGO software interface. The main window shows a network diagram with nodes representing biological processes and edges representing relationships. The nodes include terms like 'gonad development of primary female sexual characteristics', 'reproductive structure development', 'organ development', 'female sex differentiation', 'development of primary sexual characteristics', 'female gonad development', 'sex differentiation', 'reproductive developmental process', 'multicellular organismal development', 'anatomical structure development', 'system development', and 'reproductive system development'. A search bar at the top right contains the text 'Enter search term...'. The left sidebar shows a 'Control Panel' with a search bar and a list of networks: 'Plan1', 'Merged Network', and 'grupo1'. The 'BiNGO output' window is open, showing a table of results for 'grupo1'.

GO ID	GO Description	p-val	Corrected p-val	Cluster frequency	Total frequency	Genes
<input type="checkbox"/> 1544	initiation of primordial ovarian follicle growth	2.0970E-4	3.1036E-2	1/3 33.3%	1/14306 0.0%	FOXO3

At the bottom of the BiNGO output window, there are buttons for 'Select All', 'Unselect All', and 'Select nodes'. The bottom status bar shows 'Node Table', 'Edge Table', 'Network Table', and a 'Memory' indicator.

Dica: se houver dúvida sobre algum processo biológico

- Este site contém algumas definições para alguns dos processos biológicos:
- Mouse Genome Informatics (MGI): <http://www.informatics.jax.org/>
- Para pesquisar o PB: Search -> Vocabularies -> GO Browser



The screenshot shows the MGI website interface. At the top, there is a search bar with the text "Keywords, Symbols, or IDs" and a "Quick Search" button. Below the search bar is a navigation menu with links for Home, Genes, Phenotypes, Human Disease, Expression, Recombinases, Function, Strains / SNPs, Homology, and Tumors. The main content area is titled "Gene Ontology Browser" and includes a search box with a "Clear" button. The "GO Term Detail" section shows the term "molecular_function" with its synonyms, definition, and a comment. The "GO Tree View" section displays a list of related terms, including ATP-dependent activity, binding, cargo receptor activity, catalytic activity, cytoskeletal motor activity, fusogenic activity, general transcription initiation factor activity, molecular adaptor activity, and molecular carrier activity.

MGI
About Help FAQ

Keywords, Symbols, or IDs Quick Search

Home Genes Phenotypes Human Disease Expression Recombinases Function Strains / SNPs Homology Tumors

Search Download More Resources Submit Data Find Mice (IMSR) Analysis Tools Contact Us Browsers

All Search Tools
Genes
Phenotypes
Human Disease
Expression
Recombinase (cre)
Function
Strains / SNPs
Homology
Tumors
Sequence Searches
References
Vocabularies
Batch Query
MouseMine
Mouse Genome Browsers

Gene Ontology Browser
Molecular Function | Biological Process | Cellular Component

Search Clear

GO Term Detail

Term: **molecular_function**

Synonyms: molecular function

Definition: A molecular process that can be carried out by the action of a single macromolecular machine, usually via direct physical interactions with other molecular entities. Function in this sense denotes an action, or activity, that a gene product (or a complex) performs. These actions are described from two distinct but related perspectives: (1) biochemical activity, and (2) role as a component in a larger system/process.

Comment: Note that, in addition to forming the root of the molecular function ontology, this term is recommended for use for the

GO Tree View

molecular_function (31993 genes, 109079 annotations)

Browse the Gene Ontology and mouse annotations in MGI.

- ATP-dependent activity
- binding
- cargo receptor activity
- catalytic activity
- cytoskeletal motor activity
- fusogenic activity
- general transcription initiation factor activity
- molecular adaptor activity
- molecular carrier activity

Ao final desta etapa

- Fazer revisão bibliográfica dos genes e fatores de transcrição:
 - Buscar associação dos genes/FT com as características estudadas
 - Buscar associação dos genes/FT com os processos biológicos das vias enriquecidas
- Na planilha do Excel obtida no 9º passo, selecionar apenas os conjuntos de genes-FT que se destacaram na revisão bibliográfica – descartar os genes/FT não relevantes para a pesquisa

NETWORK ANALYZER