

FungiDB: Genomic Segments search: DNA motifs

1. Find fungal genes downstream of a regulatory DNA motif.

Transcriptional start sites are often located within a certain distance upstream of the genes or gene clusters that they regulate. In fungi, DNA motifs are also important for regulation of processes linked to host cell invasion or production of secondary metabolites. Readily available genomic data facilitate the discovery of regulatory motifs via examination of orthologous sequences.

The basic-helix-loop-helix (b-HLH) motif is involved in transcriptional regulation and cell-type determination. *S. cerevisiae* transcription factor PHO4 is one of the b-HLH motif proteins and it positively regulates the acid phosphatase gene PHO5 and binds to the CACGTG DNA motif. Identify genes with upstream CACGTG DNA motif in *Saccharomyces cerevisiae*.

- a) Find the CACGTG DNA motif in the *Saccharomyces cerevisiae* genome.
- Navigate to the *Search for Other Data Types* panel
 - Select *DNA motif pattern* from the *Genomic Segments* menu and look for CACGTG in *S. cerevisiae*.

The image shows two panels from the FungiDB search interface. The left panel, titled "Search for Other Data Types", has a search bar and a list of categories. "Genomic Segments" is expanded, and "DNA Motif Pattern" is selected. The right panel, titled "Identify Genomic Segments based on DNA Motif Pattern", shows the search configuration. Under "Organism", "cere" is selected. Under "Pattern", "CACGTG" is entered. A "Get Answer" button is at the bottom right.

The image shows a summary of the search results. It displays "(Segments)" and "Strategy: DNA Motif *". There is a yellow button labeled "DNA Motif" with "1906 Segments" below it, and a red "Add Step" button. On the right, there are links for "Rename", "Duplicate", "Save As", "Share", and "Delete".

The image shows a table of search results. The table has columns for Segment ID, Organism, Genomic Location, and Motif. The first few rows are as follows:

Segment ID	Organism	Genomic Location	Motif
BK006934:1958-1964.f	Saccharomyces cerevisiae S288c	BK006934:1,958..1,964 (+)	...CATTAGCACTACCATGAATGCACGTGTGCGTGTCTCATCACTGCT...
BK006934:1958-1964.r	Saccharomyces cerevisiae S288c	BK006934:1,958..1,964 (-)	...AGCAGTGATGAGGACAGCGACACGTGCATTCATGGTAGTGCTAATG...
BK006934:107613-107619.f	Saccharomyces cerevisiae S288c	BK006934:107,613..107,619 (+)	...ATTGCAAAGTAGTATTTTGTACGTGATTTTATCCAATTAATAA...
BK006934:107613-107619.r	Saccharomyces cerevisiae S288c	BK006934:107,613..107,619 (-)	...AGTTTTAATTGGATCAAAATCACGTGACAAAATACTACTTTGCAAT...
BK006934:120032-120038.f	Saccharomyces cerevisiae S288c	BK006934:120,032..120,038 (+)	...ATGAGTGTGGCGATAAGGCCACGTGAATGTACAACCTTCTCTCTT...
BK006934:120032-120038.r	Saccharomyces cerevisiae S288c	BK006934:120,032..120,038 (-)	...AAAGAGAGAAGTTGTACATTACGTGGGCCCTATCGCCACACTCAT...
BK006934:123883-123889.f	Saccharomyces cerevisiae S288c	BK006934:123,883..123,889 (+)	...CAACCTCAAGTAGAATACTACAGTGCATCGGTAGCATGGCAAG...
BK006934:123883-123889.r	Saccharomyces cerevisiae S288c	BK006934:123,883..123,889 (-)	...CTTGGCGATGCTACCGATTGACGTGTAGTATTCTACTTGAGGTTG...
BK006934:144759-144765.f	Saccharomyces cerevisiae S288c	BK006934:144,759..144,765 (+)	...TTTCCAATGTTTCTTCTCAGTGTGTTGGAAAAGAAAAGGAC...

- Your search should return multiple DNA segments containing GACGTG motif. Next, let's look for putative regulatory targets of this motif by searching for genes that are located 600bp downstream of this sequence.

b) Identify genes with the CACGTG motif located 600bp upstream of an open reading frame.

Use the genomic collocation function to set relevant distance conditions for genes in relation to the CACGTG motif.

- Click *Add Step*. Choose *Run a new search for Genes > Taxonomy > Organism* and select *Relative to genomic location*.

- Set up the collocation using the following guidelines:

Return each Gene from step 2 whose upstream region (600bp) overlaps the exact region of a Genomic Segment in Step1 (CACGTG) and is on either strand.

(Genes) Strategy: DNA Motif*

Organism: 6350 Genes

DNA Motif: 1906 Segments (Step 1) → 449 Genes (Step 2) → Add Step

Rename
Duplicate
Save As
Share
Delete

449 Genes from Step 2 [Revise](#)

Strategy: DNA Motif

Click on a number in this table to limit/filter your results

Gene Results | Genome View | **Analyze Results**

First 1 2 3 4 5 Next Last Download Add to Basket Add Columns

Gene ID	Transcript ID	Organism	Genomic Location (Transcript)	Product Description	Match Count	Region	Matched Regions
YHL038C	YHL038C-126_1	<i>S. cerevisiae</i> S288c	BK006934:23617..25509(-)	Cbp2p	2	25510 - 26109 (-)	BK006934:25779-25785:r: 25,779 - 25,785 (-); BK006934:25779-25785:f: 25,779 - 25,785 (+)
YHL036W	YHL036W-126_1	<i>S. cerevisiae</i> S288c	BK006934:26241..27881(+)	Mup3p	2	25641 - 26240 (+)	BK006934:25779-25785:r: 25,779 - 25,785 (-); BK006934:25779-25785:f: 25,779 - 25,785 (+)
YHL004W	YHL004W-126_1	<i>S. cerevisiae</i> S288c	BK006934:99219..100403(+)	mitochondrial 37S ribosomal protein MRP4	4	98619 - 99218 (+)	BK006934:98631-98637:f: 98,631 - 98,637 (+); BK006934:98658-98664:f: 98,658 - 98,664 (+); BK006934:98631-98637:r: 98,631 - 98,637 (-); BK006934:98658-98664:r: 98,658 - 98,664 (-)
YHR001W-A	YHR001W-A-126_1	<i>S. cerevisiae</i> S288c	BK006934:107826..108122(+)	ubiquinol-cytochrome-c reductase subunit 10	2	107226 - 107825 (+)	BK006934:107613-107619:f: 107,613 - 107,619 (+); BK006934:107613-107619:r: 107,613 - 107,619 (-)
YHR008C	YHR008C-126_1	<i>S. cerevisiae</i> S288c	BK006934:122889..123590(-)	superoxide dismutase SOD2	2	123591 - 124190 (-)	BK006934:123883-123889:f: 123,883 - 123,889 (+); BK006934:123883-123889:r: 123,883 - 123,889 (-)

References:

1. Fisher et al., 1991. C-myc and the yeast transcription factor PHO4 share a common CACGTG-binding motif. *Oncogene*. Jul. 6(7), p. 1099.
2. Wang et al., 2002. A systematic approach to reconstructing transcription networks in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*. Dec. 99(26) p. 16893