

YEASTRACT+

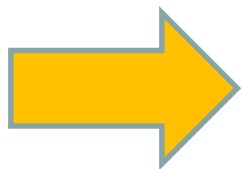
**a platform for the interspecies analysis of
transcriptional regulation
in yeasts of Biotech and clinical relevance**

Miguel Cacho Teixeira

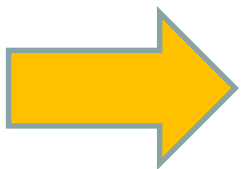
iBB – Institute for Bioengineering and Biosciences, Instituto Superior Técnico,
Universidade de Lisboa

Motivation to study transcription regulation...

... basically it is the first step of gene expression control, thus affecting all biological processes



Stimulate specific biochemical pathways



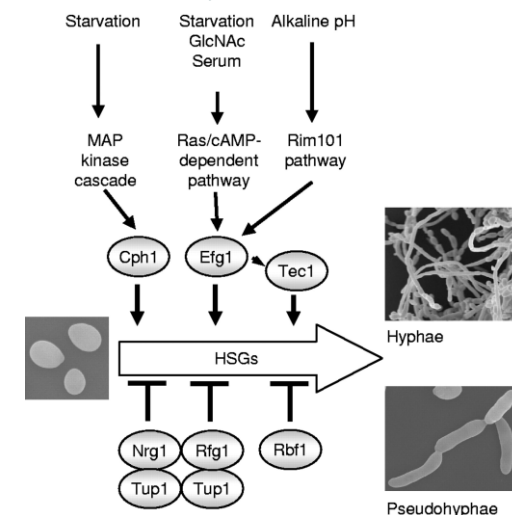
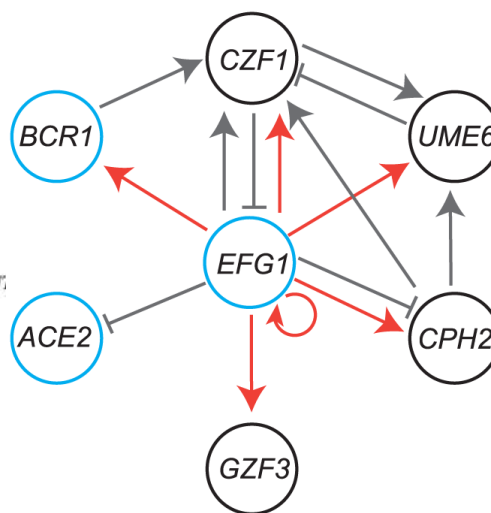
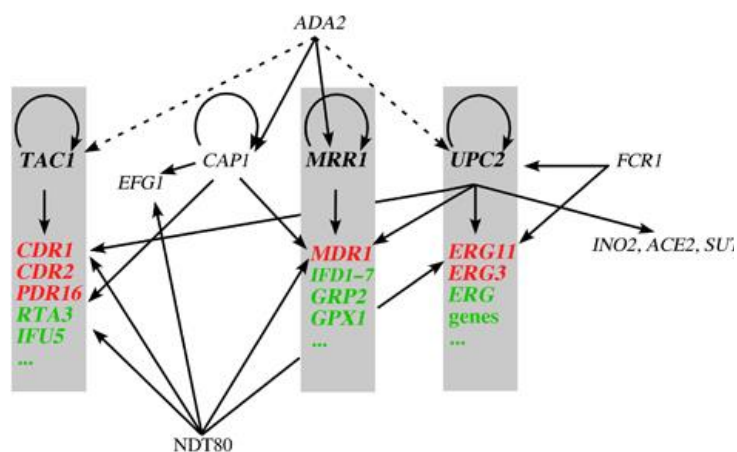
Stimulate fermentation stress resistance regulators

All pathogenesis-related processes are tightly regulated at the transcriptional level

**Azole
resistance**

**Biofilm
formation**

Dimorphism



Sanglard et al., FEMS Yeast Res, 9(7):1029-50, 2009; Holland et al., PLoS Pathogens 10(9):e1004365, 2014; Nickerson et al., Appl Environ Microbiol. 72(6):3805-13, 2006

The YEASTRACT database...since 2006

YeastRACT **TÉCNICO LISBOA** **bSrg iBB** **inesc id lisboa**

Home [Contact Us](#) [Tutorial](#) [Help](#)

Quick search... [GO](#)

[Web Services](#)

[DISCOVERER](#)

Rank genes **NEW**

[Rank by TF](#)

[Rank by GO](#)

Regulatory Associations

[Search for TFs](#)

[Search for Genes](#)

[Search for Associations](#)

Pattern Matching

[Search by DNA Motif](#)

[Find TF Binding Site\(s\)](#)

Utilities

[ORF List ↔ Gene List](#)

[IUPAC Code Generation](#)

[Generate Regulation Matrix](#)

Retrieve

[TF-Consensus List](#)

[Upstream Sequence](#)

[Flat files](#)

About YeastRACT

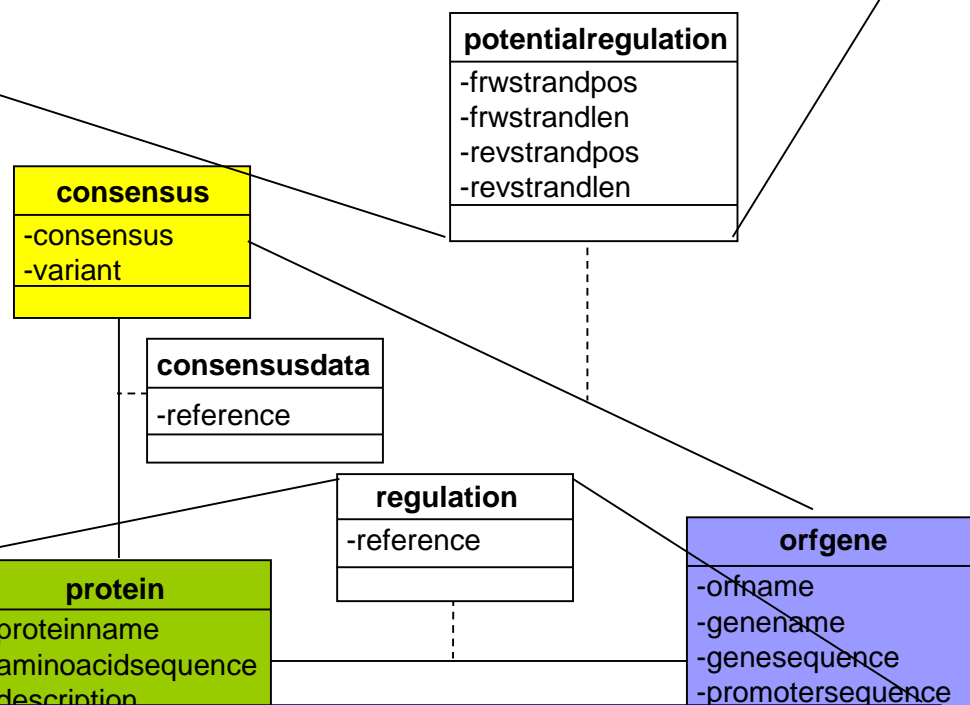
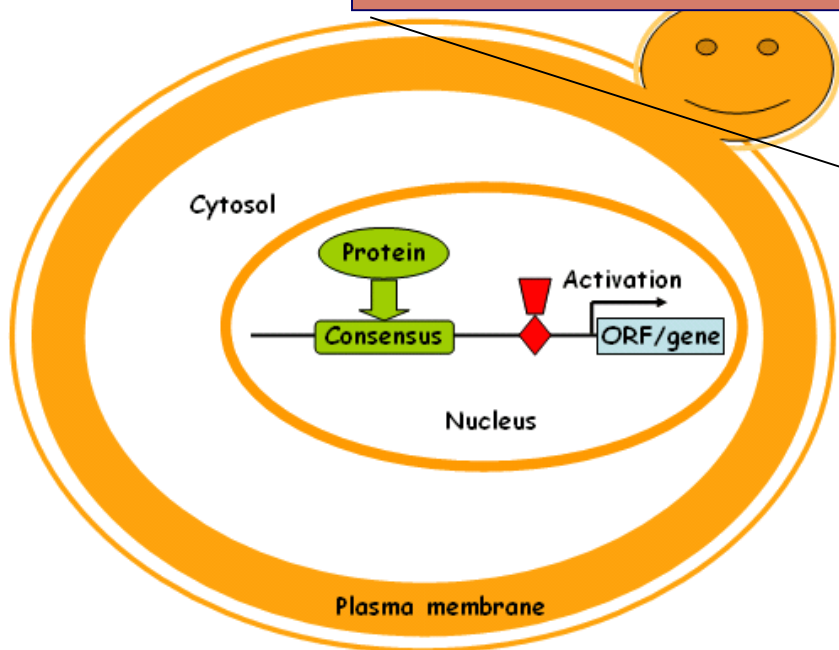
Welcome to YeastRACT

The diagram illustrates the regulatory mechanism of a gene in a yeast cell. It shows a cross-section of the cell with the plasma membrane, cytosol, and nucleus. In the cytosol, a green oval labeled 'Protein' is shown binding to a green box labeled 'Consensus' on a DNA strand. This binding leads to 'Activation', indicated by a red arrow pointing to a red diamond on the DNA. This activation results in the expression of an 'ORF/gene' in the nucleus, shown as a blue box. A smiling face icon is positioned above the nucleus.

Teixeira et al, Nucleic Acids Res, 34: D446-D451, 2006.

YEASTRACT database conceptual model

Potential regulation – a putative transcription factor binding site occurs in the promoter region of the potential target-gene

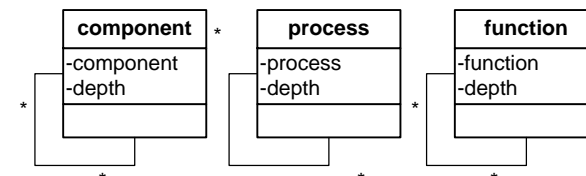


Documented regulation – there is published experimental evidence that proves that the transcription factor binds to the promoter region of its target genes (considered direct evidence) or that it affects the expression of its target genes (considered indirect evidence)

Data collection



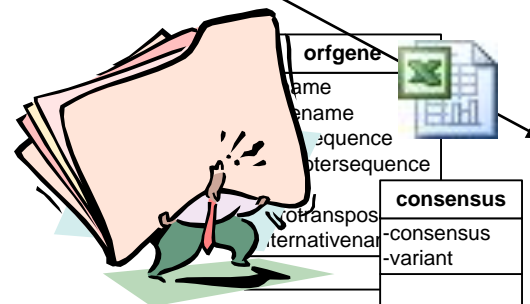
RSA Tool



- >6,000 gene promoter sequences
- >6,000 gene descriptions and associated GO term
- >160,000 documented regulatory associations, collected from over 1300 references
- 268 DNA binding sites
- 118 transcription factors



txt



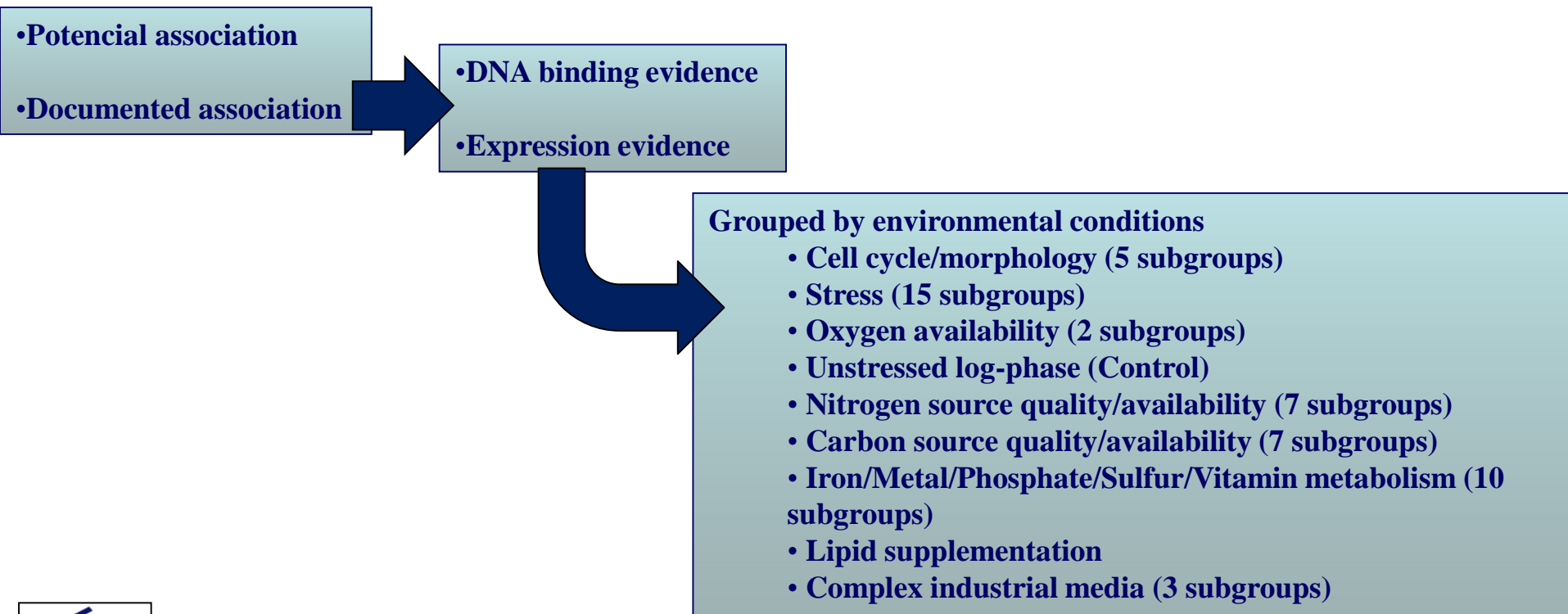
regulation
-regulationmode
-reference
-evidencecode

consensus
-consensus
-variant



Data gathering and structure

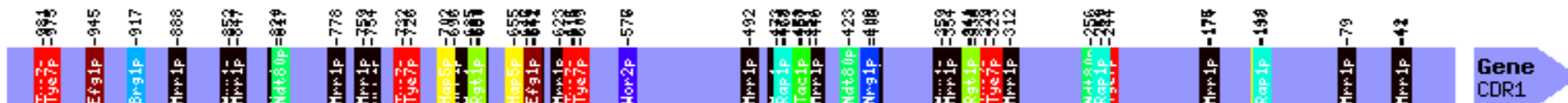
Transcription Factor  **Target Gene**



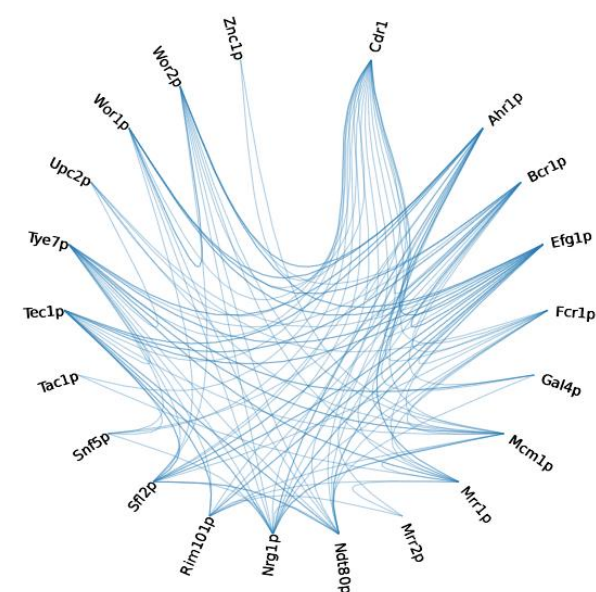
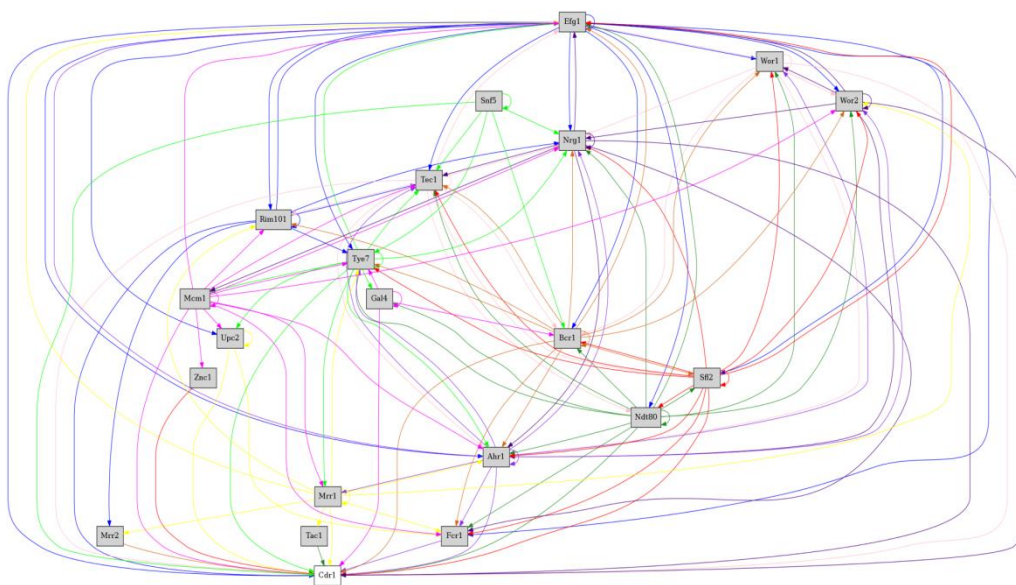
Example 1

Identify the potential or documented regulators of a given gene

Potential regulators of *CaCDR1*



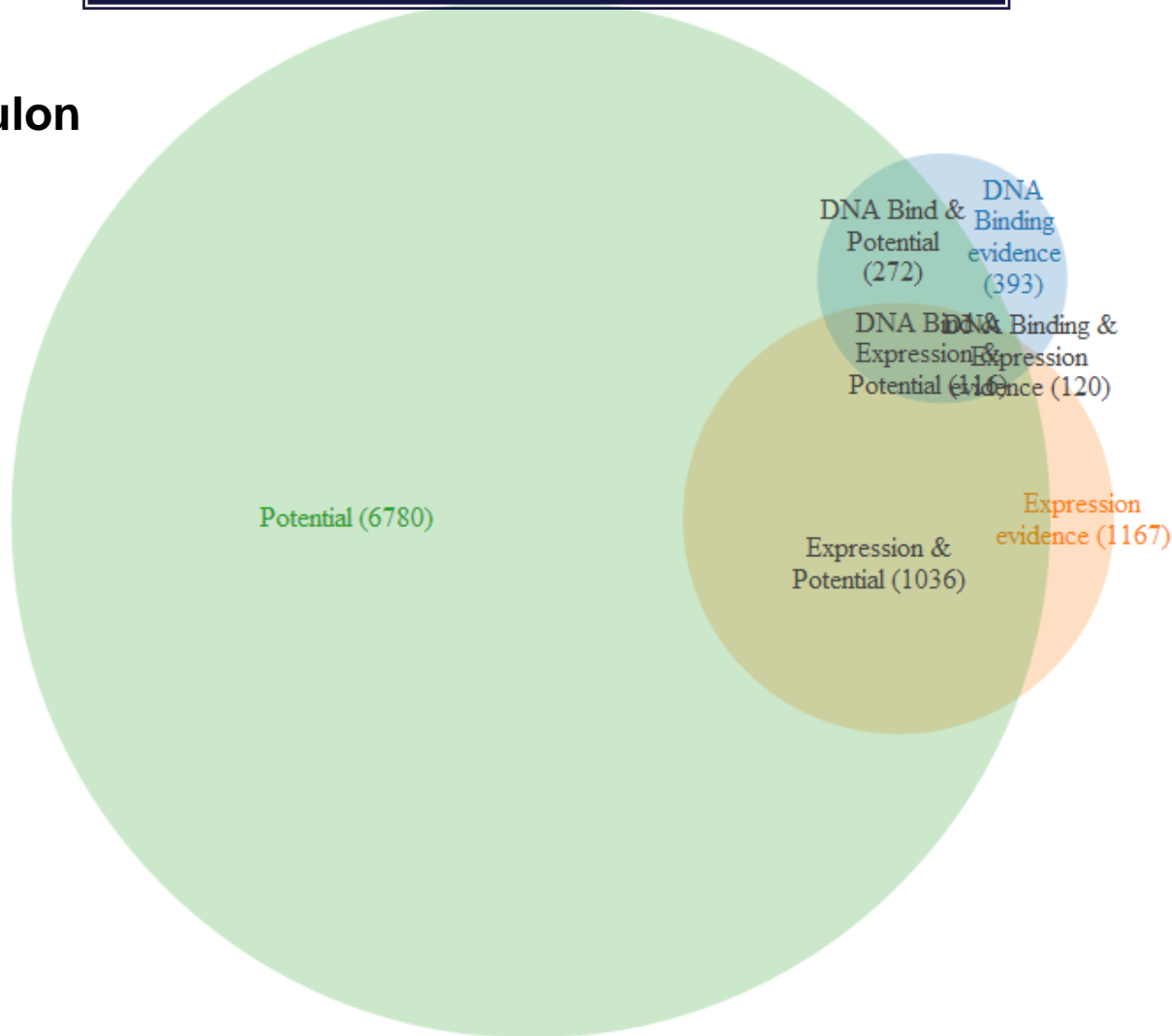
Documented regulators of *CaCDR1*



Example 2

Identify the potential or documented targets of a given TF

The Pdr1 regulon



Example 3

Grouping a list of genes based on

 [GO](#)

Rank genes

[Rank by TF](#)
[Rank by GO](#)

Regulatory Association

[Search for TFs](#)
[Search for Genes](#)
[Search for Association](#)

Pattern Matching

[Search by DNA Motif](#)
[Find TF Binding Sites](#)

Utilities

[ORF List ↔ Gene List](#)
[Generate Regulation](#)

Retrieve

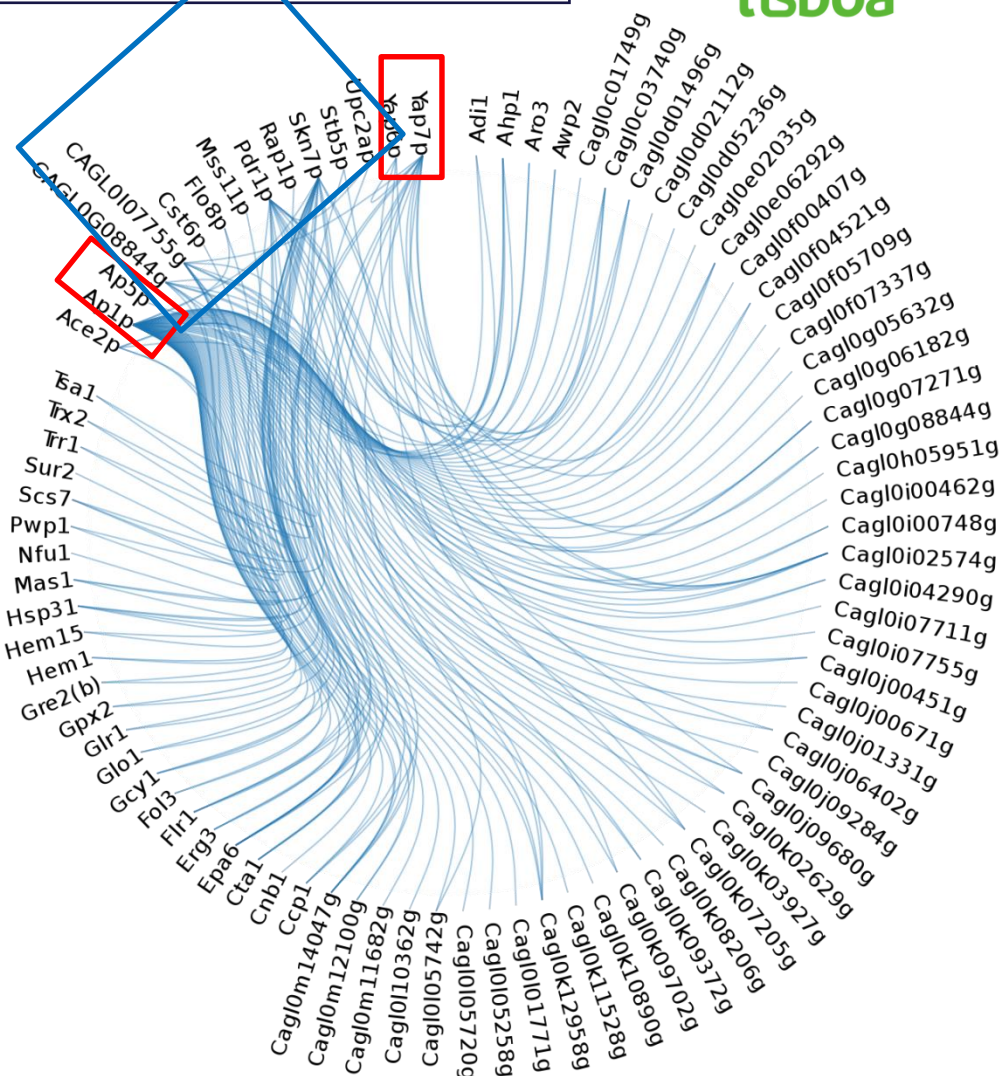
[TF-Consensus List](#)
[Upstream Sequence](#)

About PathoYeastRACT

[Contact Us](#)

Transcription Factor	% in user set	% in PathoYeastRACT	p-value	Target ORF/Genes
<input checked="" type="checkbox"/> Ap1p	100.00%	28.39%	0.0000000000000000	HEM1 EPA8 HSP31 GPX2 CAGL0C01749g CAGL0C03740g CAGL0D01498g CAGL0D02112g CAGL0D05238g CAGL0E02035g GRE2(B) CAGL0E06292g CAGL0F00407g ERG3 SCS7 CAGL0F04521g CAGL0F05709g CAGL0F07337g SKN7 CAGL0G05832g CAGL0G06182g CAGL0G07271g SUR2 MAS1 GLR1 CAGL0H05951g FLR1 ARO3 CAGL0I00462g CAGL0I00748g GCY1 TRR1 CAGL0I02574g CAGL0I04290g CAGL0I07711g PWP1 CAGL0J00451g CAGL0J00671g CAGL0J01331g CAGL0J06402g CAGL0J09284g CAGL0J09680g AWP2 TRX2 CAGL0K02629g FOL3 CAGL0K03927g TSA1 CAGL0K07205g CCP1 CAGL0K08205g ADI1 CAGL0K09372g CAGL0K09702g CTA1 CAGL0K10890g CAGL0K11528g CAGL0K12958g CNB1 CAGL0L01771g NFU1 HEM15 CAGL0L05258g CAGL0L05720g CAGL0L05742g GLO1 CAGL0L10362g CAGL0M11682g AHP1 CAGL0M12100g CAGL0M14047g
<input checked="" type="checkbox"/> Mec1p	1.41%	100.00%	0.0000000000000000	EPA8
<input checked="" type="checkbox"/> Flo8p	1.41%	100.00%	0.0000000000000000	EPA8
<input checked="" type="checkbox"/> Skn7p	28.17%	28.99%	0.0000000000000000	HSP31 GPX2 GRE2(B) SCS7 CAGL0G07271g MAS1 GLR1 GCY1 TRR1 CAGL0I02574g TRX2 TSA1 CCP1 ADI1 CTA1 CAGL0K12958g HEM15 GLO1 AHP1 CAGL0M14047g
<input checked="" type="checkbox"/> Yap7p	15.49%	6.40%	0.000001267781000	EPA8 CAGL0C03740g CAGL0E06292g CAGL0F04521g MAS1 FLR1 CAGL0I02574g PWP1 CAGL0J09680g CCP1 CAGL0L05742g
<input checked="" type="checkbox"/> Pdr1p	19.72%	3.51%	0.000107384426712	HSP31 CAGL0C03740g CAGL0D01498g CAGL0E06292g ERG3 FLR1 CAGL0I02574g CAGL0J06402g CAGL0K02629g CAGL0K07205g CAGL0K09702g CTA1 CAGL0K12958g CAGL0M14047g
<input checked="" type="checkbox"/> Ace2p	7.04%	7.94%	0.000128754993871	CAGL0G07271g ARO3 CTA1 AHP1 CAGL0M14047g
<input checked="" type="checkbox"/> Cst8p	1.41%	50.00%	0.000158453847527	EPA8
<input checked="" type="checkbox"/> Rap1p	1.41%	14.29%	0.003193308826698	EPA8
<input checked="" type="checkbox"/> CAGL0I07755g	12.68%	2.95%	0.004725888851807	EPA8 HSP31 CAGL0I00748g GCY1 CAGL0J00451g AWP2 FOL3 CAGL0M12100g CAGL0M14047g
<input checked="" type="checkbox"/> Upp2ap	1.41%	10.00%	0.008678429751680	ERG3
<input checked="" type="checkbox"/> Stb5p	2.82%	5.88%	0.008816974780704	CAGL0D01498g FLR1
<input checked="" type="checkbox"/> Yap8p	4.23%	4.48%	0.009968041791980	CAGL0C03740g CAGL0I02574g CAGL0K07205g
<input checked="" type="checkbox"/> Ap5p	1.41%	4.76%	0.028484720849128	CAGL0J09680g
<input checked="" type="checkbox"/> CAGL0G08844g	9.88%	2.06%	0.064336261195748	HEM1 SCS7 CAGL0I02574g CAGL0I04290g CAGL0K09372g CAGL0K12958g AHP1

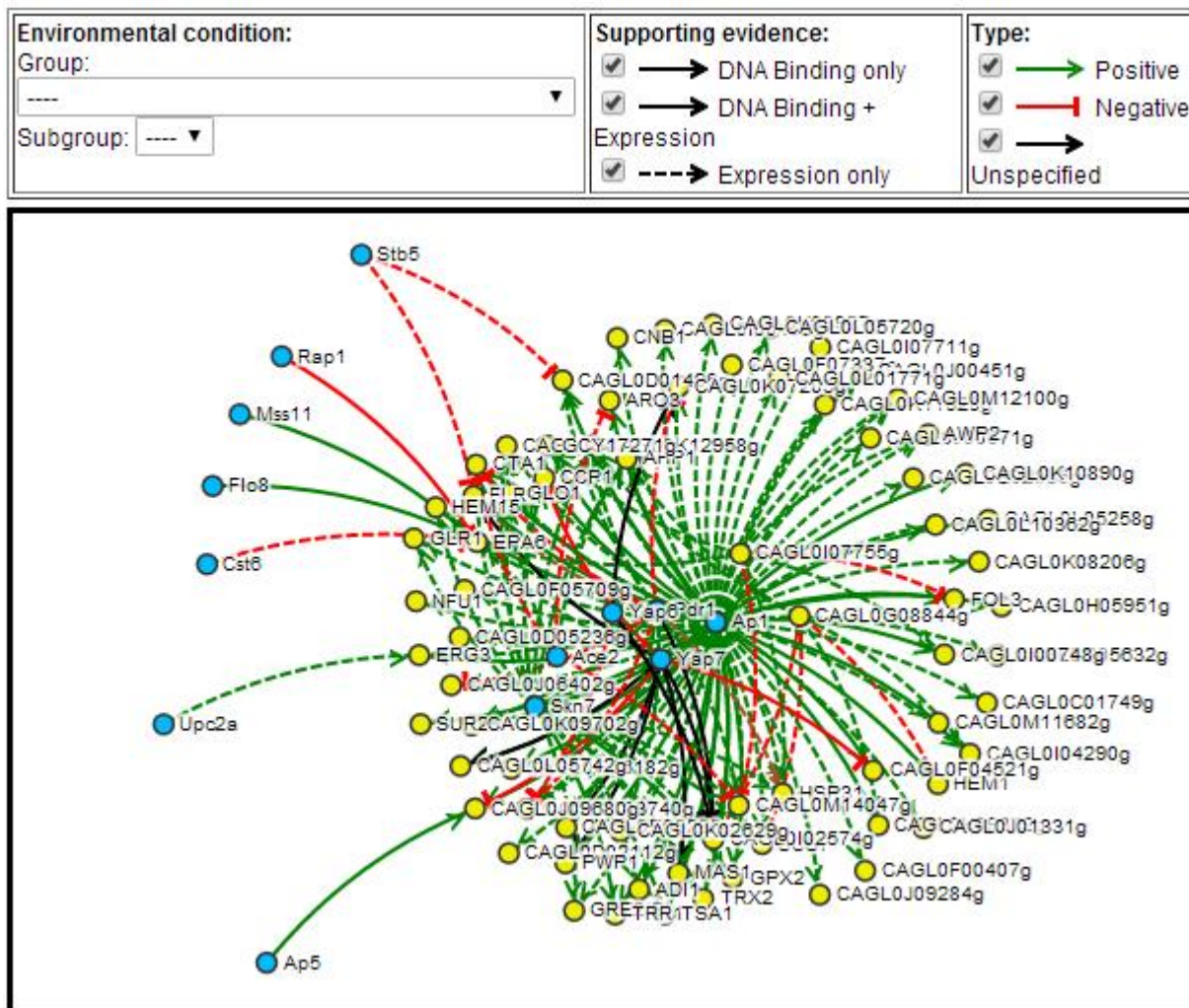
Merhej et al., Front. Microbiol. 2016; 7:1076.



Merhej et al., Front Microbiol. 7:645, 2016.

Example 3

Predicting genomic regulation



Merhej et al., Front Microbiol. 7:645, 2016.

The PathoYeasttract database... since 2016

PathoYeasttract


[Home](#)


[Contact Us](#) [How to cite](#) [Acknowledgments](#)


Welcome to PathoYeasttract


PathoYeasttract (**Pathogenic Yeast Search for Transcriptional Regulators And Consensus Tracking**) is a curated repository of all known regulatory associations between transcription factors (TF) and target genes in pathogenic *Candida* species, based on hundreds of bibliographic references. It currently provides information for the four most prevalent of pathogenic yeasts, *C. albicans*, *C. glabrata*, *C. parapsilosis* and *C. tropicalis*.

Facilities are also provided to enable the exploitation of the gathered data when solving a number of biological questions, as exemplified in the Tutorial. PathoYeasttract allows the identification of documented or potential transcription regulators of a given gene and of documented or potential regulons for each transcription factor. It also renders possible the comparison between DNA motifs and the transcription factor binding sites described in the literature. Finally, the system provides a useful tool for grouping a list of genes (for instance a set of genes with similar expression profiles as revealed by microarray or RNA-seq analysis) based on their regulatory associations with known transcription factors. All analysis can be carried out based on the existing data for the *Candida* species the user is working on or based on transcription associations occurring among orthologous genes and TFs in the remaining *Candida* species or in the model yeast *Saccharomyces cerevisiae*.

Enter
Candida glabrata 

Enter
Candida albicans 

Enter
Candida tropicalis 

Enter
Candida parapsilosis 

Monteiro et al., Nucleic Acids Res, 45(D1):D597-D603, 2017.



Home

[Contact Us](#) [Tutorial](#) [Help](#)

Quick search... [GO](#)

Welcome to PathoYeast

- *Candida albicans* -

Rank genes

[Rank by TF](#)

[Rank by GO](#)

Regulatory Associations

[Search for TFs](#)

[Search for Genes](#)

[Search for Associations](#)

Pattern Matching

[Search by DNA Motif](#)

[Find TF Binding Site\(s\)](#)

Utilities

[ORF List ↔ Gene List](#)

[Generate Regulation Matrix](#)

Retrieve

- >6,000 gene promoter sequences
- >6,000 gene descriptions and associated GO terms
- >38,000 documented regulatory associations
- 74 DNA binding sites
- 117 (of 163 predicted) transcription factors

Welcome to PathoYeastRACT**- *Candida parapsilosis* -****Rank genes**[Rank by TF](#)[Rank by GO](#)**Regulatory Associations**[Search for TFs](#)[Search for Genes](#)[Search for Associations](#)**Pattern Matching**[Search by DNA Motif](#)[Find TF Binding Site\(s\)](#)**Utilities**[Search for orthologs](#) **NEW**[ORF List ↔ Gene List](#)[Generate Regulation Matrix](#)**Retrieve**[TF-Consensus List](#)[Upstream Sequence](#)

- >6,000 gene promoter sequences
- >6,000 gene descriptions and associated GO terms
- >7,000 documented regulatory associations
- 6 DNA binding site
- 11 (of 145 predicted) transcription factors

Welcome to PathoYeast**- *Candida glabrata* -****Rank genes**[Rank by TF](#)[Rank by GO](#)**Regulatory Associations**[Search for TFs](#)[Search for Genes](#)[Search for Associations](#)**Pattern Matching**[Search by DNA Motif](#)[Find TF Binding Site\(s\)](#)**Utilities**[ORF List ↔ Gene List](#)[Generate Regulation Matrix](#)**Retrieve**

- >5,000 gene promoter sequences
- >5,000 gene descriptions and associated GO terms
- >3,500 documented regulatory associations
- 39 DNA binding sites
- 41 (of 123 predicted) transcription factors

related yeast species.



Home

[Contact Us](#) [Tutorial](#) [Help](#)

Quick search... [GO](#)

Welcome to PathoYeast

- *Candida tropicalis* -

Rank genes

[Rank by TF](#)

[Rank by GO](#)

Regulatory Associations

[Search for TFs](#)

[Search for Genes](#)

[Search for Associations](#)

Pattern Matching

[Search by DNA Motif](#)

[Find TF Binding Site\(s\)](#)

Utilities

[Search for orthologs](#) **NEW**

[ORF List ↔ Gene List](#)

[Generate Regulation Matrix](#)

Retrieve

[TF-Consensus List](#)

[Upstream Sequence](#)

- >6,000 gene promoter sequences
- >6,000 gene descriptions and associated GO terms
- >680 documented regulatory associations
- 1 DNA binding sites
- 16 (of 131 predicted) transcription factors

Example 4

Predicting gene/genomic regulation based on regulatory info gathered for other species

[Home](#) > [Rank Genes by TF](#)

[Contact Us](#) [Tutorial](#) [Help](#)

Quick search... [GO](#)

Rank Genes by TF

Rank genes

- [Rank by TF](#)
- [Rank by GO](#)

Regulatory Associations

- [Search for TFs](#)
- [Search for Genes](#)
- [Search for Associations](#)

Pattern Matching

- [Search by DNA Motif](#)
- [Find TF Binding Site\(s\)](#)

Utilities

- [ORF List ↔ Gene List](#)
- [Generate Regulation Matrix](#)

Retrieve

- [TF-Consensus List](#)
- [Upstream Sequence](#)

About PathoYeast

- [Contact Us](#)
- [Cite PathoYeast](#)
- [Acknowledgments](#)

Regulations Filter	Transcription factors	Target ORF/Genes
<input checked="" type="radio"/> Documented <ul style="list-style-type: none"> <input type="radio"/> Only DNA binding evidence <input type="radio"/> Only Expression evidence <ul style="list-style-type: none"> <input checked="" type="checkbox"/> TF acting as activator <input checked="" type="checkbox"/> TF acting as inhibitor <input checked="" type="radio"/> DNA binding plus expression evidence <input type="radio"/> DNA binding and expression evidence <input type="radio"/> Potential		
<input type="checkbox"/> Check for all TFs		
Filter Documented Regulations by environmental condition: Group: ---- ▾ Subgroup: ---- ▾		
Search for Homologous Regulations in: Species: ---- ▾ Consider ---- (use codes)		
Rank genes by: Candida albicans SC5314 The % of Saccharomyces cerevisiae S288c regulated by each TF		
<input type="button" value="Search"/> <input type="button" value="Clear"/>		

Predicted regulation based on orthologous
regulation in *C. albicans*

CAGL0J05060g (Csr1)

CAGL0M07634g (Efg1)

CAGL0I10769g (Mcm1)

Met4 (Met4)

Rgt1 (Rgt1)

Sko1 (Sko1)

CAGL0I07183g (Sfl1)

CAGL0L02123g (Snf5)

CAGL0K04631g (Try5)

CAGL0M12100g (Tye7)

Pdr1

QDR2

Known
regulation in *C.
glabrata*

Predicted regulation based on
orthologous regulation in *S. cerevisiae*

CAGL0M09955g (Sfp1)

CAGL0J01177g (Abf1)

CAGL0E04884g (Adr1)

CAGL0L02585g (Bas1)

Rap1 (Rap1)

Swi6 (Swi6)

Tup1 (Tup1)

CAGL0J04510g (Cdc73)

CAGL0I09614g (Isw2)

CAGL0J07150g (Pip2)

CAGL0C00297g (Set2)

CAGL0J11594g (Sin3)

CAGL0M01540g (Spt3)

CAGL0D04136g (Srb2)

CAGL0H02145g (Ste12)

CAGL0M01716g (Tec1)

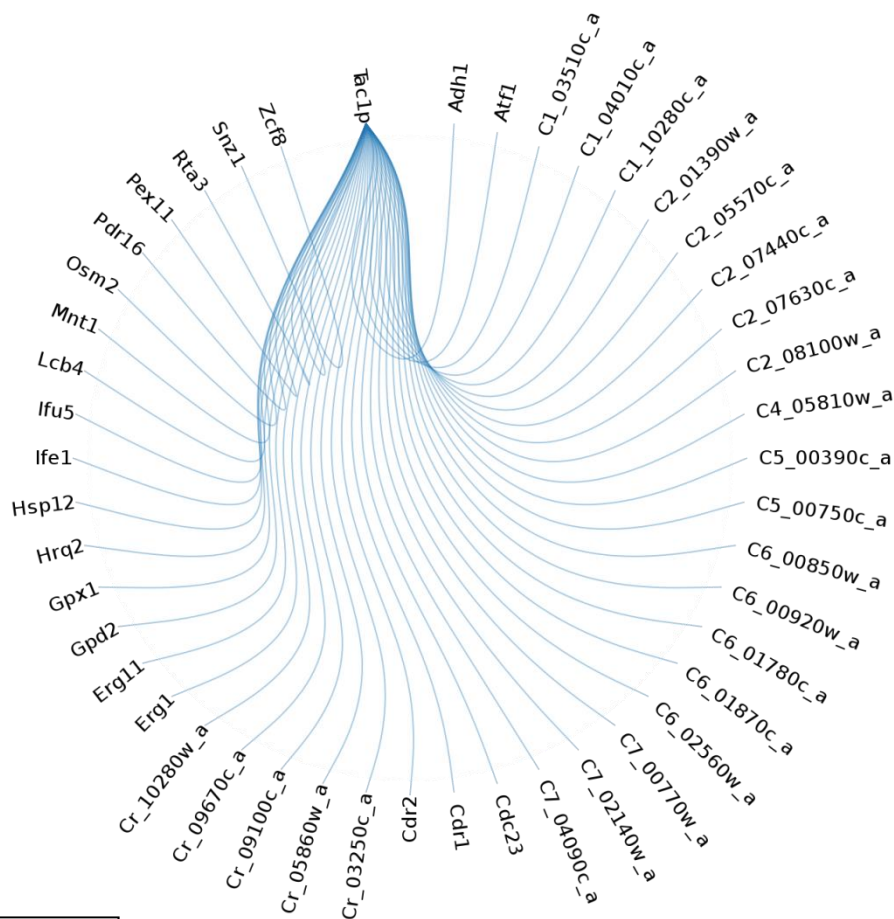
CAGL0F05357g (Ume6)

CAGL0L04576g (Yrm1)

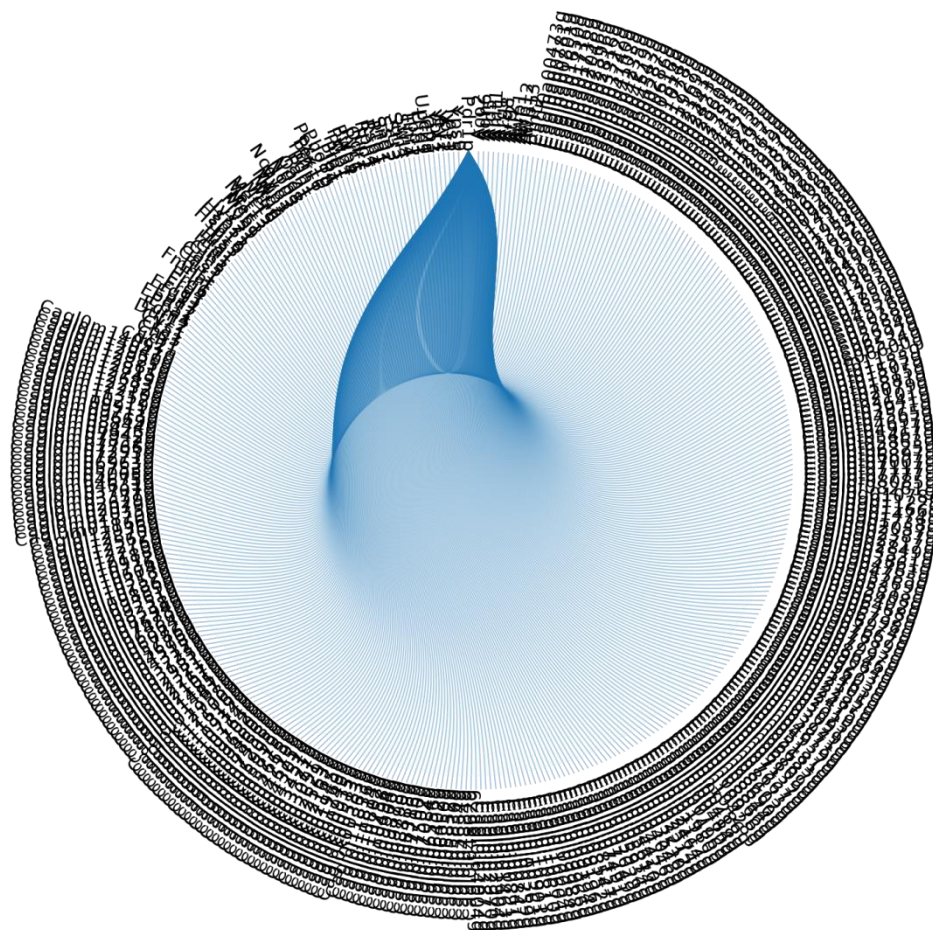
Example 4

Inter-species comparison of genomic regulation

Targets of *C. albicans* Tac1

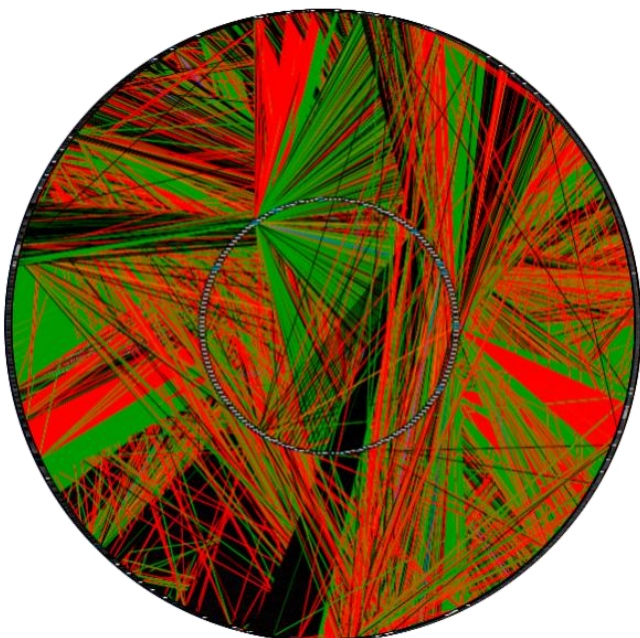


Targets of *C. glabrata* Pdr1

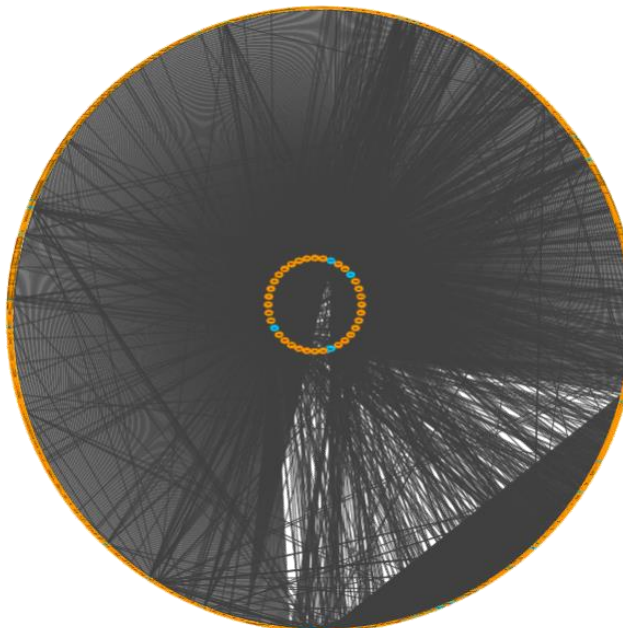


MDR Genes, their Targets and Regulators

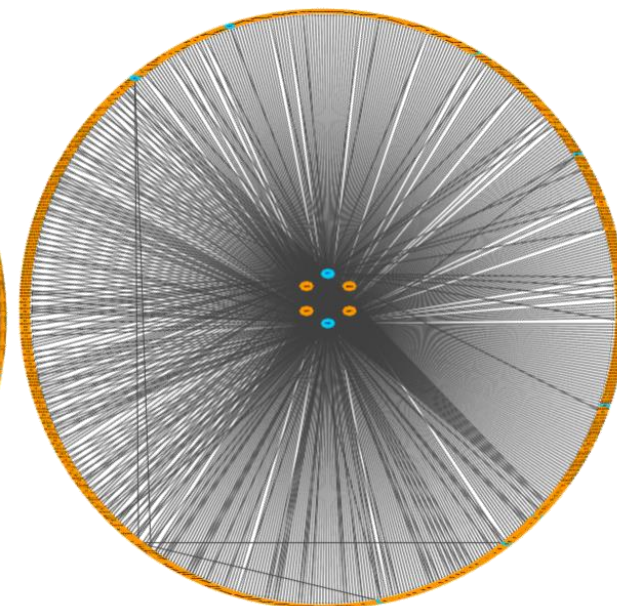
S. cerevisiae



C. albicans



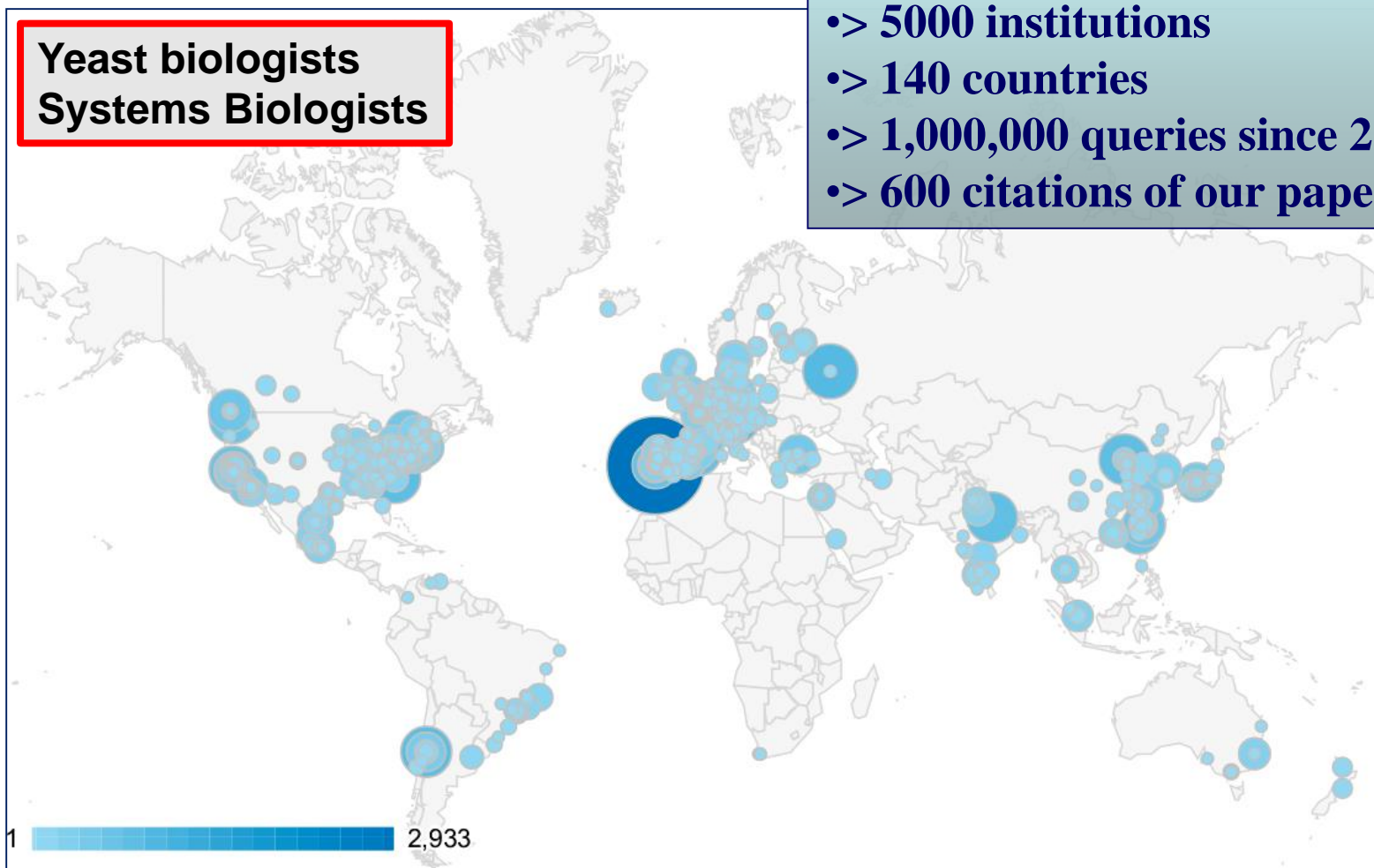
C. glabrata



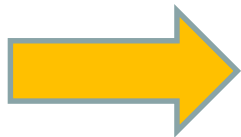
YEASTRACT users worldwide

**Yeast biologists
Systems Biologists**

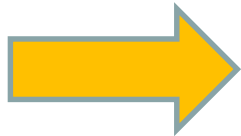
- > 5000 institutions
- > 140 countries
- > 1,000,000 queries since 2006
- > 600 citations of our papers



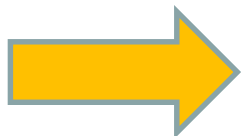
Yeasttract+: perspectives



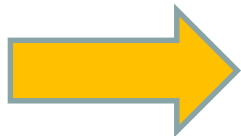
Regularly update the information therein.



**Expand to other species of pathogenic and biotechnologically relevant yeast
species and strains**

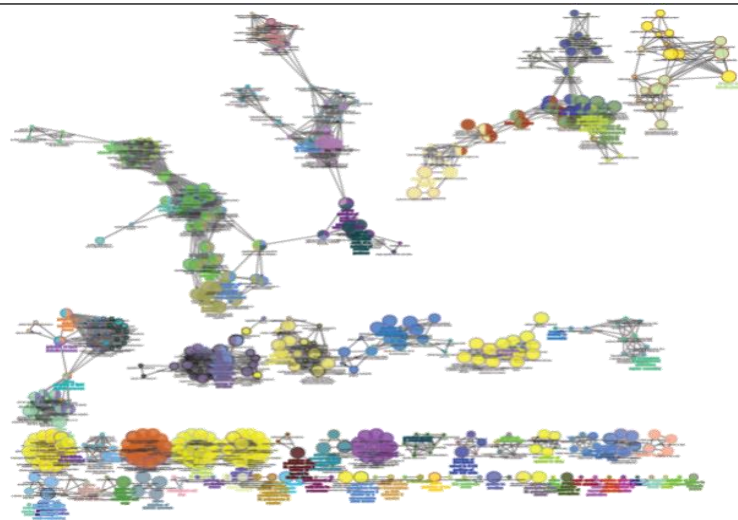


Automate the construction of (predicted) regulatory networks based on genome sequences alone.

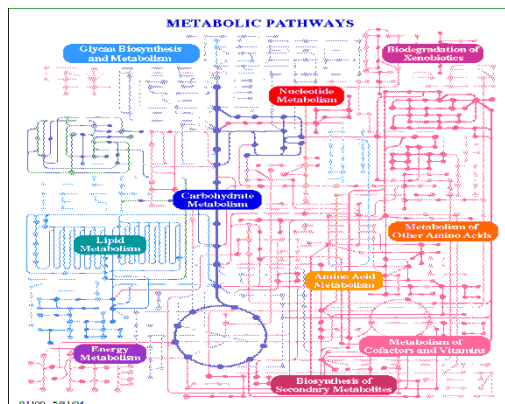


Offer comparative genomics tools for inter-species network comparison

Regulatory pathways



Metabolic pathways



**Predict the global
systems behavior**

**Predict the effect of
perturbations to the
system**

**Optimize the system to
produce a given
metabolite**

Current contributors



**Miguel
Teixeira**



**Isabel
Sá-Correia**



**Pedro
Monteiro**

Pedro Pais

Mónica Galocha

Mafalda Cavalheiro

Margarida Palma

Miguel Antunes

Cláudia Godinho

Arlindo Oliveira

Past contributors

Pooja Jain, Sandra Tenreiro, Alexandra Fernandes, Marta Alenquer

Nuno Mendes, Sofia D'Orey, Sandra Tenreiro, Hélio Pais, Alexandra Carvalho

Tânia Cabrito, Artur Lourenço, Dário Abdulrehman

Nuno Mira, Sandra dos Santos, Catarina Costa, Joana Guerreiro, Alexandre Francisco, Ricardo Aires, Joana Gonçalves, Ana T Freitas

FCT

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

**FEDER, FCT (projects POSI/EIA/57398/2004,
POCTI/BIO/56838/2004, PTDC/EIA/67722/2006, PTDC/BBB-
BIO/4004/2014, PTDC/BII-BIO/28216/2017 and PhD and
post-doctoral grants)**

BioData.pt

