Identifying function of unknown yeast genes

Shubhi Sharma, Jill Keeney (Ph.D.), Juniata College, 1700 Moore St. Huntingdon, PA 16652

1. ABSTRACT

Saccharomyces cerevisiae (Baker's yeast) has a relatively simple eukaryotic genome that can be easily manipulated to study gene function. Yeast as a model organism can be used to study pathogenic genes, as well as in the invention of new technologies. Although yeast is a simple eukaryote nearly 10% of the genes remain unknown after 20 years sequencing the genome. Through this project, we wanted to identify the function of YER066W, a gene of unknown function. Different bioinformatics tools were performed to compare our gene to genomes in other species. Through the comparison of the alignment of sequences using BLAST, we observed that our gene contained a WD40- repeat protein domain which has functions ranging from signal transduction and transcription regulation to cell cycle control. In order to investigate differences in metabolism, we grew YER066W deletion and wild type strains in different carbon sources: sucrose, raffinose, mannitol, galactose, maltose and glucose. No difference in growth rate was observed. We then compared the morphology by looking at the bud size between YER066W deletion and wild type strains. Fixed cells were compared for the number of unbudded yeast cells and different bud sizes between the YER066W deletion and wild type strain to compare if there was a difference in the frequency and morphology between the strains; no differences were observed. We hope to design further experiments which will enable us to investigate the function of YER066W and the WD40 domain.

2. BACKGROUND INFORMATION

Saccharomyces Genome Database (SGD)

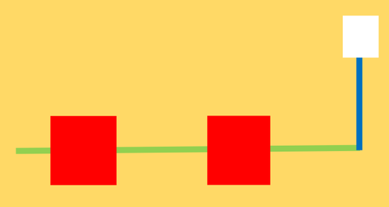
- Community resource with information about the yeast genome.
- *YER066W* is classified as a putative protein with unknown function.
- Non-essential gene identified in a screen for mutants with decreased levels of rDNA transcription.

The gene ontology (GO) is used to compare a gene between species by looking at the molecular function, biological process and cellular component.

- GO for YER066W protein has an unknown biological role and cellular location.

SGD shows that YER066W contains a WD40 protein domain

- The WD40 repeat domain is part of the β-propeller domain-containing protein family.
  - Has a central peptide-binding pocket.
  - Involved in a variety of cell process; protein-protein/DNA interaction.
  - Short ~40 amino acid motifs, often terminating in a Trp-Asp (W-D) dipeptide.
- YER066W has two WD motifs:
  - 2-39
  - 64-99



3. WD40 motif structure

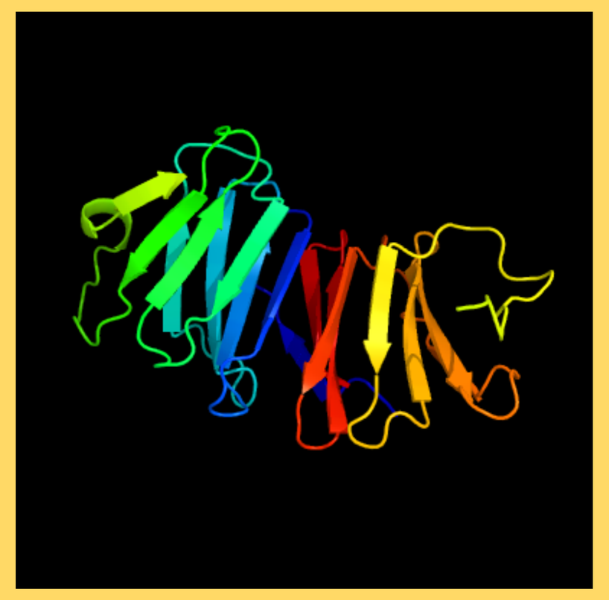
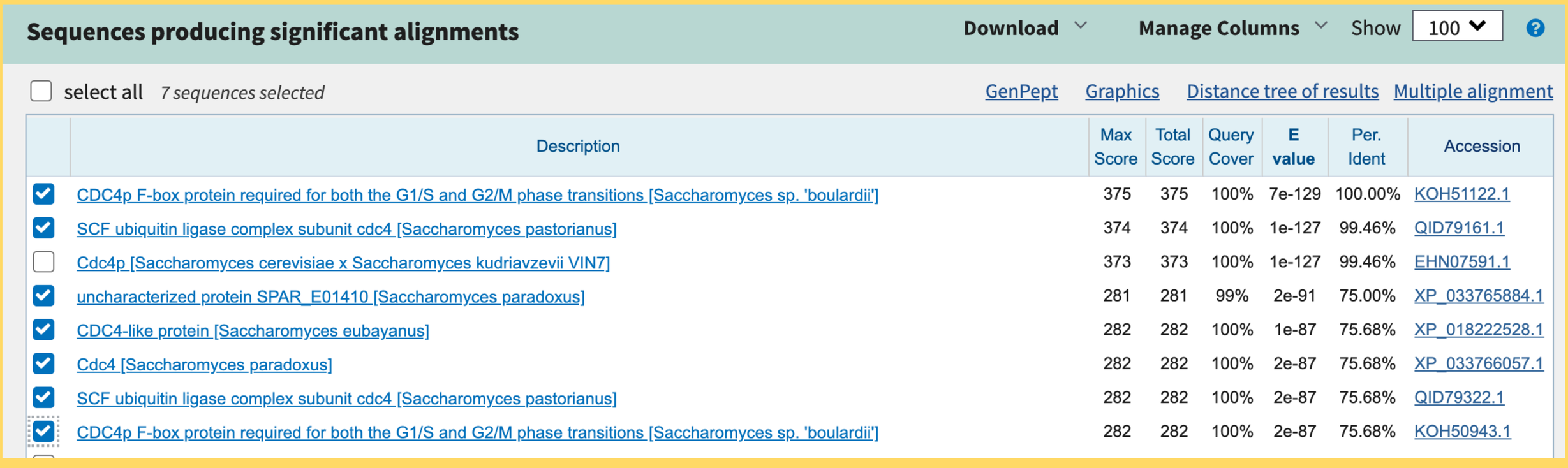


Figure 1: The crystal structure represents the WD40 domain and the β-propeller plates. This motif is in YER066W.

4. Bioinformatic analysis

YER066W blastp of other species involved in cell cycle.



Sequences producing significant alignments — Download — Manage Columns — Show 100

select all — 7 sequences selected — GenPept — Graphics — Distance tree of results — Multiple alignment

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
☑	CDC4p F-box protein required for both the G1/S and G2/M phase transitions [Saccharomyces sp. 'boulardii']	375	375	100%	7e-129	100.00%	KOH51122.1
☑	SCF ubiquitin ligase complex subunit cdc4 [Saccharomyces pastorianus]	374	374	100%	1e-127	99.46%	QID79161.1
☐	Cdc4p [Saccharomyces cerevisiae x Saccharomyces kudriavzevii VIN7]	373	373	100%	1e-127	99.46%	EHN07591.1
☑	uncharacterized protein SPAR_E01410 [Saccharomyces paradoxus]	281	281	99%	2e-91	75.00%	XP_033765884.1
☑	CDC4-like protein [Saccharomyces eubayanus]	282	282	100%	1e-87	75.68%	XP_018222528.1
☑	Cdc4 [Saccharomyces paradoxus]	282	282	100%	2e-87	75.68%	XP_033766057.1
☑	SCF ubiquitin ligase complex subunit cdc4 [Saccharomyces pastorianus]	282	282	100%	2e-87	75.68%	QID79322.1
☑	CDC4p F-box protein required for both the G1/S and G2/M phase transitions [Saccharomyces sp. 'boulardii']	282	282	100%	2e-87	75.68%	KOH50943.1

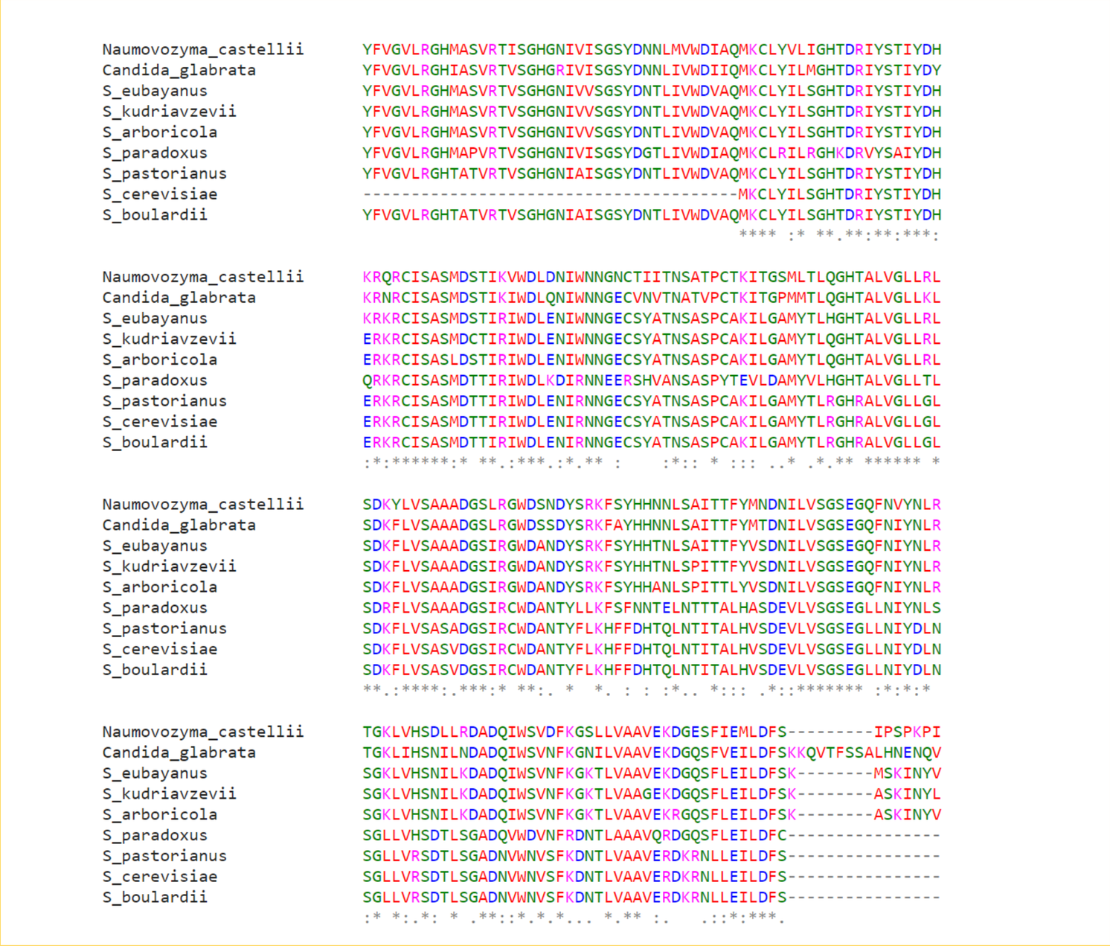


Figure 2: Using T-COFFEE the multiple sequence alignment (right) acquired from BLAST (left) compared YER066W proteins in similar species. Sequences from these non- Saccharomyces cerevisiae genes are thought to be related to CDC4p which is involved in cell cycle phase transitions.

5. YER066W deletion does not affect yeast cell-cycle

Table 1: An OD600 ranging from 0.5 to 1.2 reflects the culture of cells growing in log phase.

Type	OD from overnight (10:20 AM)	Final OD (4:00PM)	Initial OD in conical flasks (11:50AM)	Final OD in conical flasks (6:40pm)
WT	1.6	0.25	1.6	0.516
2199	1.6	1.5	1.7	0.586

Table 2: Bud sizes between Wild Type (WT) and deletion strain

Type	Unbudded	Small	Medium	Large
2199	30	41	20	9
WT	31	51	10	8

- Wild type and deletion strains were grown to log phase to observe budding and morphology. This was done in order to determine if the yeast strains are involved in cell cycle, which is one of the function of WD40 domain. To assess the cell cycle, statistically 100 cells of each strain were counted.
- Budding Index: Small bud (1/3 or less than the mother cell), medium bud (1/3-2/3 of the mother cell) and large bud (2/3 or greater than mother cell). Unbudded cells are not involved in the cell cycle. The results are shown to the right.

6. Growth in various C-sources

SGD data shows that growth in different carbon sources may vary.

- Tested carbon sources: sucrose, glucose, galactose, raffinose, maltose and mannitol.
- Same growth observed for both wild type and mutant in all carbon sources, except for mannitol. Yeast doesn't use mannitol.

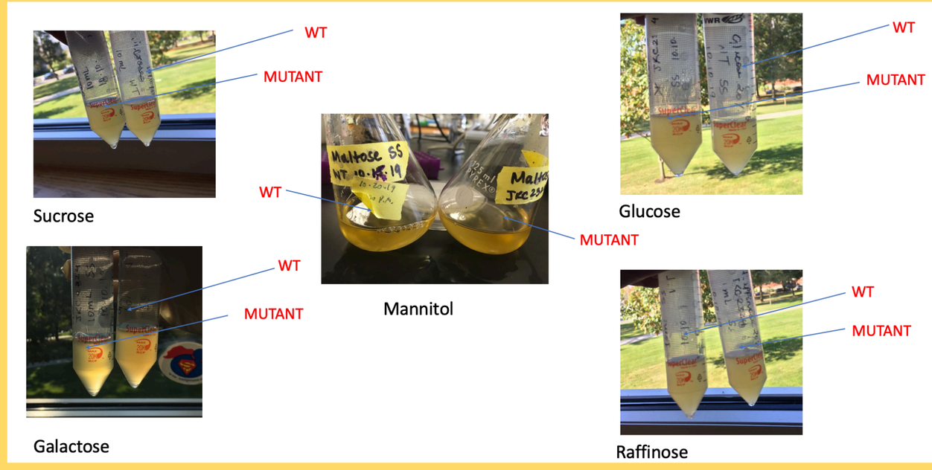
Carbon source (2%)	Concentration (w/v)
Glucose (control)	20%
Sucrose	20%
Mannitol	20%
Galactose	20%
Maltose	20%
Raffinose	10%

7. Results

Molecular Weight 20.7 KDa
Philius predicted protein: Globular

The whole genome expression indicated a difference in growth among different carbon sources.

- No difference in morphology and bud size was observed between the wild type and deletion strain.



8. Summary

- YER066W contains the WD40 protein domain which is involved in a variety of cellular functions such as signal transduction, protein-to-protein interaction, etc.
- The gene expression from the sequence alignment shows that our protein may be involved in cell cycle phase transition.
- In conclusion, no difference in cell cycle growth was observed between the wild type and deletion strain.
- For future studies, we hope to explore the functions of YER066W related to the WD40 protein domain.

9. REFERENCES

1. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1462-2920.13617)
2. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3835864/).
3. https://ojs.library.ubc.ca › index.php › expedition › article › view
4. https://uwaterloo.ca/chem13-news-magazine/april-2015/activities/fermentation-sugars-using-yeast-discovery-experiment
5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3153487
6. https://www.microbiologyresearch.org/docserver/fulltext/micro/133/7/mic-133-7-1675.pdf?expires=1569692668&id=id&accname=guest&checksum=48E8DC2527D03D6EB707AD57796BD872
7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3216861/
8. https://aem.asm.org/content/69/7/4144
9. https://www.molbiolcell.org/doi/pdf/10.1091/mbc.E15-07-0499
10. https://aem.asm.org/content/71/11/6465).
11. https://pubmed.ncbi.nlm.nih.gov/29026209/
12. https://pubmed.ncbi.nlm.nih.gov/21468892/
13. https://www.ncbi.nlm.nih.gov/pubmed/30069656

This research is funded in part by NSF award # 1624174 to JB Keeney