Wednesday's Challenge - what yeast is in the brewer's oak barrel?

Jacob L. Steenwyk & Antonis Rokas

In this workshop challenge, we will reenact a plausible research scenario. Consider the following. The Brewmaster of a major brewery has sent you a yeast species isolated from one of the Brewmaster's oak barrels used for beer storage. To determine the identity of the species and set the stage for further analyses, you have sequenced and assembled the yeast isolate's genome and now wish to create a phylogeny to understand who it is most related to. Thus, we will build a species phylogeny by

(1) Identifying orthologs, (2) Aligning orthologs, (3) Trimming orthologs,
(4) Creating a supermatrix, and (5) Inferring a putative species tree.

After you have inferred a species tree, compare it to the phylogeny in Shen *et al.* 2018, Cell, which can be viewed as a pdf here:
https://jlsteenwyk.github.io/publication_pdfs/Shen_et_al_2018_Cell.pdf
Hint: The easiest way to find taxa on the phylogeny you created and the Shen et al. 2018 phylogeny is to use *control+f* to search for taxon names in Figure 2.

Useful information:
1) How to extract single FASTA entries from multi-FASTA files
To extract a single FASTA entry from a multi-FASTA file, use *samtools faidx*.
For example, execute the following command to extract *Brewery_genome_1* from the multi-FASTA file *Brewery_genome.fa*

```
samtools faidx Brewery_genome.fa Brewery_genome_1
```

2) outgroup taxa are
Starmerella apicola
Starmerella bombicola
Wickerhamiella versatilis