Using the PomBase motif search

Proteins that contain a C-terminal CaaX motif are prenylated which facilitates their association with membranes. The CaaX motif contains a cysteine residue, followed by two residues that are typically aliphatic amino acids (AGILV), followed by an amino acid that confers substrate specificity (for example FTase recognize serine, methionine, glutamine or alanine at this poisition and GGTase-I recognize leucine at the fourth position).

In this exercise you will
- Identify putative prenylated proteins in *S. pombe*
- Look at features of orthologues of these proteins in SGD and FungiDB

Let's start by finding all known (experimentally characterized) isoprenylated *S. pombe* proteins. Go to the advanced search (https://www.pombase.org/query) and click on the "protein modification" filter option. Search for "isoprenylated residue", select it from the list and click submit.

This search only finds 17 genes (these are experimentally characterized isoprenylated proteins) - there are probably more uncharacterized ones.

Use the motif search (https://www.pombase.org/motif_search) to find all proteins containing the strict CaaX box motif
C[AGILV][AGILV][MSQACLE]$
(i.e. proteins with C-terminal cysteine-aliphatic-aliphatic-M/S/Q/A/C/L/E amino acids). This should match 21 genes.

We can find more isoprenylation candidates by relaxing the CaaX motif. Remove the requirement for aliphatic amino acids at the second and third positions of the motif "[AGILV]" and instead allow any amino acid at those positions "..".
This changes the motif to C..[MSQACLE]$

Go back to the motif search and find the 46 proteins matching this relaxed motif. Click the button to send this list to the advanced search.

We now want to find proteins that match the relaxed motif, but are not in the published list of isoprenylated proteins. Tick the two queries, and then press the 'subtract / not' button - this subtracts the bottom query (older - isoprenylated proteins) from the top query (newer - proteins containing the CaaX box motif). This should leave a list of 32 genes.

	Genome location	
	Number of exons	
	Taxonomic conservation	
	Characterisation status	

Combine queries:

Union / or	Intersect / and	Subtract / not		Delete

	Results	Query history
☑	46	gene list: [SPAC110.03 SPBC609.03 SPAC16A10.04 SPAC30D11.11
☑	17	isoprenylated residue (MOD:00703)

Click on the 'results' hyperlink to see the list of 32 genes matching this query.

SGD also annotate protein modifications. In some cases, a prenylated protein may have been published in *S. cerevisiae* (and annotated as such in SGD).

In the list of 32 genes that match the relaxed motif, click on the mas5 result. Go to the 'orthologs' section on the mas5 gene page, and click on the *S. cerevisiae* YDJ1 hyperlink, this will take you to SGD. On the YDJ1 page, click the 'protein' menu tab, and scroll down to the 'Post-translational modifications' section. Which residue is farnesylated in the S. cerevisiae ortholog? (not all modifications are shown by default)

You can use the advanced search to mix and match different queries.

Now let's find putative isoprenylated proteins that
1. are inviable when mutated
    and
2. are conserved to fungi but not to vertebrates.

Go back to the advanced search https://www.pombase.org/query
Click on the 'phenotype' filter, type and select 'inviable vegetative cell population', then 'submit'

Then, use the taxonomic filter to find:
● genes "conserved in fungi"
● genes "conserved in vertebrates"

- Combine the putative CaaX box gene list (32 genes) to "inviable vegetative cell population" using the intersect/AND operator
- Combine the resulting list with the "conserved in fungi" gene list using the intersect/AND operator

- Combine the resulting list with the "conserved in vertebrates" gene list using the NOT operator

This query should look like this:
(((conserved in fungi (PBO:0011064) AND gene list: [SPBC609.03 SPCC18.04 SPAC3C7.05c SPCC417.05c SPBC530.04 SPAC959.04c SPAC806.04c SPBC1734.11 SPAC9E9.15 ...]) AND inviable vegetative cell population (FYPO:0002061) [single allele genotypes]) NOT conserved in vertebrates (PBO:0011070))

And should identify one gene - Pof6 SPCC18.04
Open up the pof6 gene page in PomBase

In a new browser tab, go to fungiDB and search for the *S. pombe* pof6 page using the search box in the top right corner. Find the *A. nidulans* FGSC A4 ortholog of Pof6, and open the gene page for this gene.

Compare the phenotypes of pof6 (PomBase gene page) to the phenotypes annotated for RcyA (FungiDB). What is the most striking difference in deletion phenotypes of this gene in the two different yeasts?

Go to the *S. cerevisiae* ortholog in SGD (linked to from the PomBase pof6 gene page in the ortholog section), and click the phenotype tab. Is the S. cerevisiae deletion mutant viable or inviable (you can use the filter search to narrow down to the phenotype of interest).