ASSIGNMENT (an ~easy one related to BLAST analysis!) 20% of final grade

 - you will need the files in this folder.   Note that you may want to change the
filenames, since the blast database name is equivalent to the filename...

Assignment: use of BLAST to Discover putative Orthologues

  Orthologue DEFINITION: Genes that are found in different species that evolved from a
common ancestral gene by speciation.

A common first-step in discovery of Orthologues is to do a "reciprocal best
BLAST". That is, you take protein X in Species A, and BLAST it against all proteins in
Species B. The top (significant!!) hit in Species B is then BLASTed against all proteins in
Species A. If it's top (significant) hit is the Protein X, then those two proteins are
considered to be Orthologue candidates. (there is more work to do after this, but this is
a good start)

Using BioRuby to blast and parse the blast reports, find the orthologue pairs between
species Arabidopsis and S. pombe. I have uploaded their complete proteomes to the
Moodle for you. You do not need to create objects for this task (the existing BioRuby
objects are sufficient)

To decide on "sensible" BLAST parameters, do a bit of online reading - when you have
decided what parameters to use, please cite the paper or website that provided the
information.

Bonus: 1%

Reciprocal-best-BLAST is only the first step in demonstrating that two genes are
orthologous. Write a few sentences describing how you would continue to analyze the
putative orthologues you just discovered, to prove that they really are orthologues. You
DO NOT need to write the code - just describe in words what that code would do.

(You can learn about orthology analysis by reading online, and that will give you ideas
of how to take the next step)