

EnsemblBacteria

<http://bacteria.ensembl.org>

1. Go to the EnsemblBacteria homepage.
 - a) How many bacterial genomes are accessible through Ensembl?
 - b) How many *Haemophilus influenzae* genomes are accessible?

2. Go to the genome information of *Haemophilus influenzae* Rd KW20.
 - a) Exactly how many base pairs are in this genome?
 - b) How many annotated coding genes are there?

3. Search this genome for the *mdh* gene.
 - a) What is its gene ID number?
 - b) What type of enzyme does it encode?
 - c) What are its chromosomal coordinates?
 - d) How long is the transcript produced for this gene?
 - e) How many amino acids are in the protein encoded by this gene?

4. Use the pan-taxonomic comparison tool (at the left) to make a gene tree image.
 - a) Which bacterial species have *mdh* genes most similar to *H. influenzae*?
 - c) How many homologs (orthologues) of *mdh* are in *Drosophila melanogaster*?
 - d) Which *D. melanogaster* orthologue is most similar to the *H. influenzae* protein?
 - e) What is the % identity between the *D. melanogaster* and *H. influenzae* proteins?
 - f) Generate an alignment of these two protein sequences. How many residues are identical? How many are similar? How many "gaps" are there?

5. Return to the homepage and search for the *Bacillus subtilis* QB928 genome.
 - a) Exactly how many base pairs are in this genome?
 - b) How many annotated coding genes are there?
 - c) Does *B. subtilis* have the *mdh* gene?

6. Return to the homepage and search for the *Mycoplasma genitalium* G37 genome.
 - a) Exactly how many base pairs are in this genome?
 - b) How many annotated coding genes are there?
 - c) Does *M. genitalium* have the *mdh* gene?