Evolutionary and Structural Analysis of Pathogen Proteins.

Final year UG project 2024-25
2024-11-03 (Week 7)

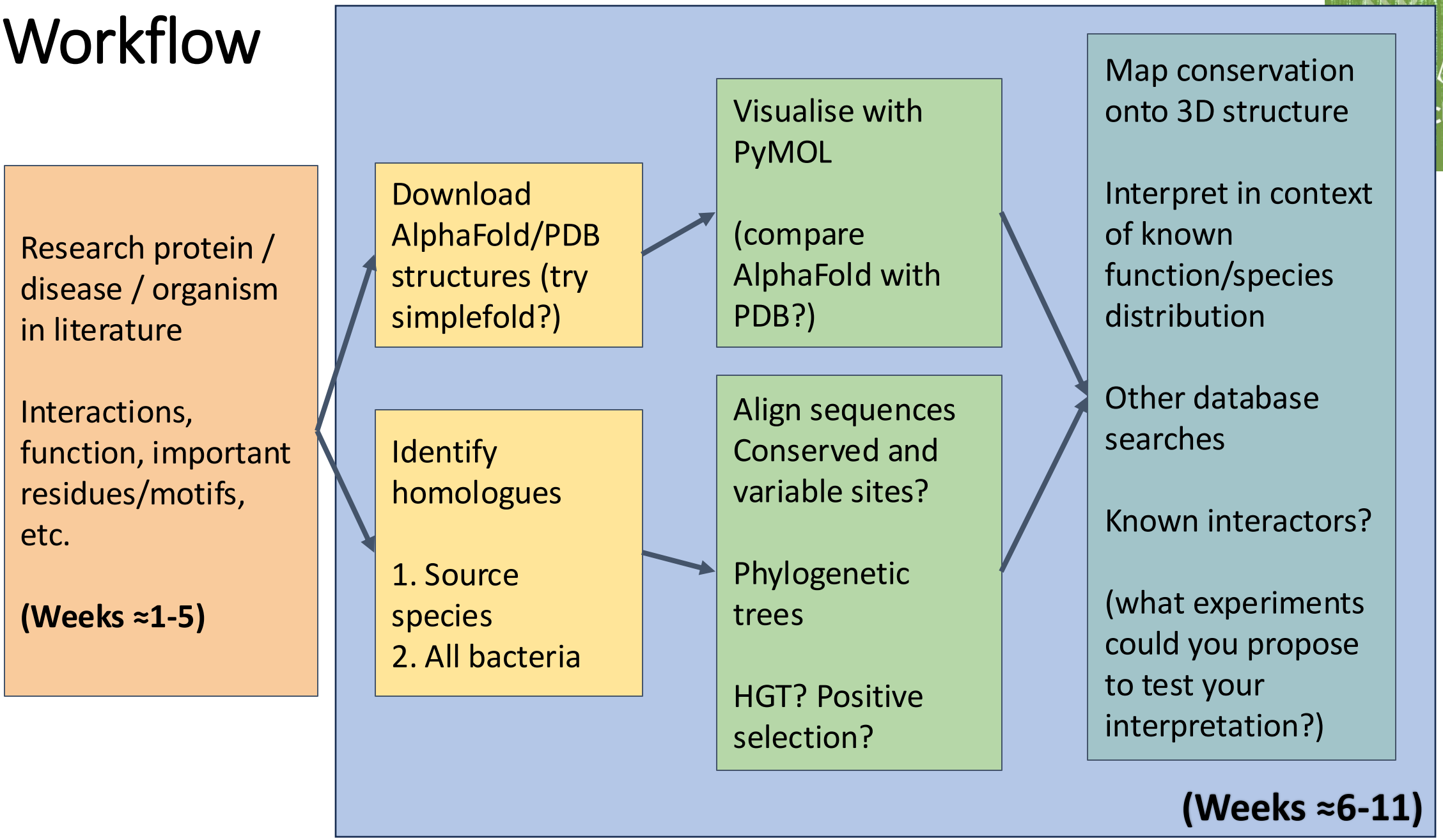
Candidate proteins – start points

Any changes needed?

Organism	Host	Gene/Protein	PHI accession	Student
Escherichia coli	*Homo sapiens*	*espY*	PHI:8647	LB
Shigella flexneri	*Homo sapiens*	*ipaJ*	PHI:9253	LT
Candida albicans	*Mus musculus*	*sap6*	PHI:10193	IM
Pseudomonas aeruginosa	*Homo sapiens*	*tplE*	PHI:6646	AE
Vibrio vulnificus	*Mus musculus*	*vvhA*	PHI:6877	JT

http://www.phi-base.org/

Workflow


Research protein / disease / organism in literature
Interactions, function, important residues/motifs, etc.
(Weeks ≈1-5)
Download AlphaFold/PDB structures (try simplefold?)
Identify homologues
1. Source species
2. All bacteria
Visualise with PyMOL
(compare AlphaFold with PDB?)
Align sequences Conserved and variable sites?
Phylogenetic trees
HGT? Positive selection?
Map conservation onto 3D structure
Interpret in context of known function/species distribution
Other database searches
Known interactors?
(what experiments could you propose to test your interpretation?)
(Weeks ≈6-11)

Your questions/comments/issues

(What would you like to talk about?)

Points for discussion?

- What would you like to cover?
- **Keep investigating your protein**
- Make a tree for your sequences
  - How does the tree look?
  - Are organisms/species grouped together how you'd expect? Any surprises?
  - Is there evidence supporting gene duplication (two clades, with the same organisms repeated)?
  - Is there evidence supporting horizontal gene transfer (organism in the "wrong" clade)?
- Compare to sequence data
  - Do patterns of sequence variation map onto your tree?
- Galaxy workflow or scripting?
  - Have you created a workflow for your analysis: alignment? making a tree?
  - Have you noted/saved the tool versions/parameter settings for writing your thesis?

Thursday Seminar

- Kimberly Kline (University of Geneva)
- “Exploring microbe-host interactions in Enterococcal biofilms”
- 2pm, HW111-114
- (HW613 if you want to talk to her immediately after the seminar)
  - “Bacterial Adhesins in Host-Microbe Interactions” https://www.cell.com/cell-host-microbe/fulltext/S1931-3128(09)00178-4
  - “A tale of two pili: assembly and function of pili in bacteria” https://www.cell.com/trends/microbiology/fulltext/S0966-842X(10)00040-5?code\u003dcell-site=
  - “Gram-Positive Uropathogens, Polymicrobial Urinary Tract Infection, and the Emerging Microbiota of the Urinary Tract” https://doi.org/10.1128/9781555817404.ch19

Next Week's Group Meeting

Monday 10th November 13:30 HW324

Useful Links

Useful tools (many others are available)

GalaxyEU: https://usegalaxy.eu/

- Sequence alignment (e.g. MAFFT), phylogenetics (e.g. RaxML), positive selection (e.g. codeML)

iTOL: https://itol.embl.de/

- Visualisation/annotation of phylogenetic trees

PyMOL: https://pymol.org/2/ and/or ChimeraX: https://www.cgl.ucsf.edu/chimerax/

- Protein structure visualisation/annotation

Jalview: http://www.jalview.org/

- Visualisation of multiple sequence alignments

Useful sites/databases

PHI-base: http://www.phi-base.org/

- Proteins involved in host-pathogen interactions, with linked evidence

EMBL AlphaFold: https://www.alphafold.ebi.ac.uk/

- AlphaFold predictions for proteins from model organisms

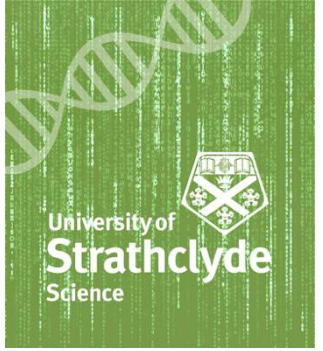
UniProt: https://www.uniprot.org/

- Protein sequence (including homologous sequences) and functional information with evidence

RCSB/PDB: https://www.rcsb.org/

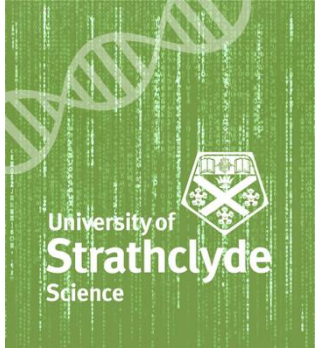
- Repository of record for protein structures

SIPBS CompBiol Sites



- BM432 Project Pages
  - https://sipbs-compbiol.github.io/bm432-project/

- An incomplete little book of bioinformatics
  - https://sipbs-compbiol.github.io/little-bioinformatics-book/



Project Management Tools

You may want tools to…

- Manage your time
  - E.g. Pomodoro technique (e.g. BeFocused, Pomofocus, Forest)
- Schedule work
  - Reminders (macOS, MS Office)
  - Calendar (macOS, MS Office), with email alerts
  - Trello, Asana, etc.
- Manage your project data and information effectively
  - How to name files
  - Project management guidelines (BM432, 2022-23 session; me and Dr Feeney)
  - How to keep a lab notebook
  - Keeping a computational biology lab notebook: https://doi.org/10.1371/journal.pcbi.1004385
  - Organising a lab book