



Bash scripting

03.Feb 2021



Why?

- Reproducible research!
- Easier to ask for help, input or commands from others



Steps

Might take
more than one
iteration



1. Prototype your analysis on small test dataset
 - Do you get any error codes?
 - Do the results agree with what you would expect?
 - Anything indicating problems in your log files?
2. Debug your code
 - Look at error codes and google them or look at software FAQ's
 - Insert print statements
 - Simplify analysis and see which step is throwing the error
3. Generalize your code
 - Use variables instead of hard coded paths/file names or options
4. Expand your analysis to full dataset
 - Submit analysis manually for each job OR
 - Write wrappers (e.g. automatically process analysis per chromosome)



1. Example: for-loop wrapper

- Usage: Very useful when you need to loop over chromosomes
- Task: extract data for individual chromosomes
- see scripts: 'EX01.for_loop.step*.sh'
 - 'EX01.for_loop.step01.sh': prototyping small test analysis
 - 'EX01.for_loop.step03.sh': generalizing script by introduction of variables
 - 'EX01.for_loop.step04.sh': creating a for loop around analysis



2. Example: while-loop wrapper

- Usage: Very useful when you need to loop over input files
- Task: add a header line to a number of files
- see scripts: 'EX02.while_loop.sh'
 - works on outputs of Example 1



Q&A



Tools that can be helpful

Code editor:

- [Visual Studio Code](#)

Copying files to/from cluster

- [Filezilla](#)
- [MobaXterm](#) (windows only)
- [Shhfs](#) (linux only)