Bioinformatics Commandments

Yazdan Asgari

2019

The Web Is Not a Secure Place

- Sending your data to a Web server is like making it public

- If you work for a pharmaceutical company, this is a bad idea

- Likewise, don't submit patients' private data over the Web

Keep Track of Database and Program Versions

- Reproducibility is important in science

- The only way to reproduce a bioinformatics experiment is to use the same tools and resources

- Remember (and write down)
  - Which program was used (and which version)
  - Which database was used (and which version)

3

Keep Track of Your Sequences

- Always write down your sequences' accession numbers

- Keep these for each sequence:
  - The ID : Identification number
  - The AC: Accession number

- The AC should never change

Remember Your Program Parameters

- If you change default parameters, write down your changes!

- This will be the only way to reproduce your experiment.

Save the Right Results

- Save the flashy graphs for publication

- Save the boring ASCII files for further work!

Know Thine E-Values

- Alignments are evaluated with E-values

- People talk about percentages of similarity by the coffee machine but they **write down** E-values in publications!

- The lower the E-value, the better the alignment

Know Thine Alignments

- Evaluating alignments is difficult

- Develop a sense of what is good and bad in a multiple-sequence alignment:
  - Good ⇔ Nice, ungapped blocks
  - Bad ⇔ Messy, gapped blocks

Check Borderline Results with Different Programs

- Three short-sighted witnesses are more Informative than a single eagle-eye witness

- Better to have three different programs giving the same borderline result than a single good result

Bioinformatics: Life Sciences, Robert Lessick

Be Careful with Unpublished Methods

- The Internet is **great** because you can find everything on it

- The Internet is **a pain** because you can find everything on it

- Good methods are published in peer-reviewed journals **before** being put online

Data and Wine

- Databases are not like good wine . . .
  - They don't improve with age.

- Databases are more like vegetables, salads, or fruit juice . . .
  - The fresher, the better!

Everything Is Not For Free on the Internet

- Everything is **mostly** free for academics

- If you are a company, you may have to pay a royalty, even if the information looks free . . .

- It's all about trust . ☺

Biting the Right Bullet

- You **can** do everything online, but over time, it will be
  - Less reliable
  - More time-intensive
  - Less flexible

- If you find yourself doing Bioinformatics nonstop for a few months . . .
  - Time to learn a bit of Mathematics, Statistics, Physics, Programming, and Linux!

Bioinformatics: Concept

bioinformatics involves more than just algorithms and methods:

it is emerging as a discipline that requires a way of thinking to help the user to find the right tools, to learn how to use them, and to know how to interpret the results.