

# Scientific Computing: Challenges and Opportunities

Adina Howe

Assistant Professor

Department of Agricultural and Biosystems Engineering

[adina@iastate.edu](mailto:adina@iastate.edu)

*Slides Modified from NGS Workshop taught in the US*  
<https://github.com/ngs-docs/angus>

# Challenge #1:

**Most biologists still don't know much about computational science.**

- Among many biologists, there is a general fear or skepticism of computers.
- This leads to shallow thinking about computational science.

# Challenge #2:

**Most computational scientists still don't know much about biology.**

- Existing computational solutions may not use appropriate heuristics, or default parameters.
- “It works on my data...”, but their data  $\neq$  yours!

# Challenge #3:

**Both biology and computational science  
are deep, complex fields of study,  
inhabited by extremely smart people!**

- None of this is easy, on any side of things.
- If it were easy, they wouldn't need people as smart as all of us to do it, right??
- A one week course can't possible teach you everything.

# Challenge #4:

**Sequencing technology is changing very fast.**

- We don't understand its limitations or biases very well.
- The software and compute infrastructure lags behind volume of data & type of data.

This is not the #1 problem you will face with bioinformatics.

Here is the #1 problem:

How do you know if your computational answer is right or wrong?

# This is not the #1 problem you will face with bioinformatics.

Here is the #1 problem:

How do you know if your computational answer is right or wrong?

**If you can't answer this question, then what's the point of doing the computation?**

# Controls

- Just as with experiments, you can put negative and positive controls in your bioinformatics.
- e.g. with BLAST,
  - Do you see expected matches with the parameters and database you're using?
- Positive controls are often easier than negative, in “discovery” science...

# Internal controls

- Use molecules and sequences for which you have expectations.
- “I know this gene comes up, based on qPCR. I expect to see it in my mRNAseq.”
- Or, “human? I didn’t expect to see human!”