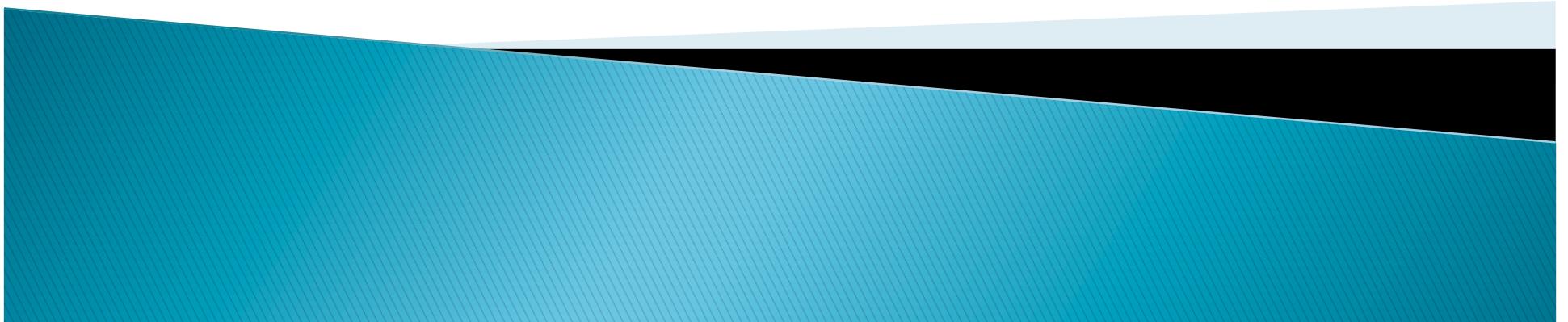


# This Is the Welcome Talk

Titus Brown

6/6/11



# About this course

- ▶ Modern biology faces a number of challenges.
- ▶ One of them is sane and appropriate integration of sequence data into biology.
- ▶ We would like to see if we can help you from that.
- ▶ Many “sub”challenges



# Challenge #1:

**Most biologists 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 don't know much about biology.**

- ▶ Extant computational solutions may not use appropriate heuristics, or default parameters.
- ▶ “It works on my data...”, but their data != yours!
- ▶ Solutions/programs may not be couched in the right terms for the biology, or with proper appreciation for biological complexity.



# 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 two 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.



# The opportunity:

- ▶ The sequence is coming, if not already here.
- ▶ “In the land of the blind, the one eyed is king.” -- those prepared to *think* about how to use sequencing technology to answer their question will have a substantial leg up.
- ▶ Who knows? Some of you might even like this mix!



# Our goals

- ▶ Provide a “safe place” to experiment.
- ▶ Lots and lots of help (in the form of TAs)
- ▶ Provide lots of data sets, tools, scripts.
- ▶ Research specific help?



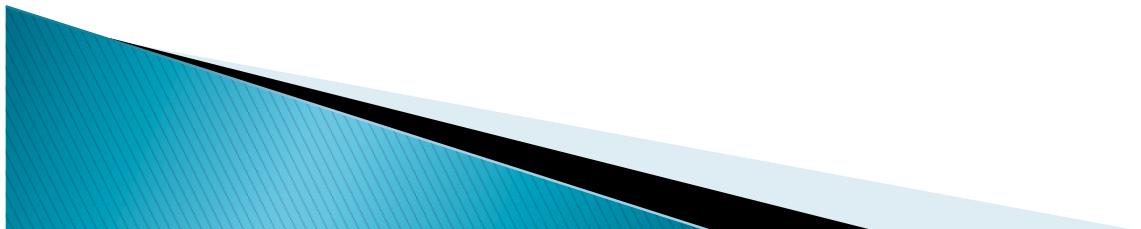
# Our requirements of you

- ▶ Nothing.
- ▶ This is a requirements free zone.
- ▶ You can safely skip the entire course...



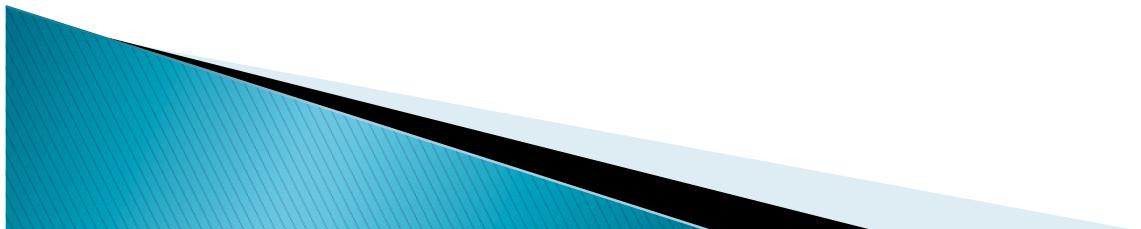
# Our expectations

- ▶ Questions!
- ▶ Ask for help when you need it!
- ▶ A certain amount of tolerance may be needed, by you of us...



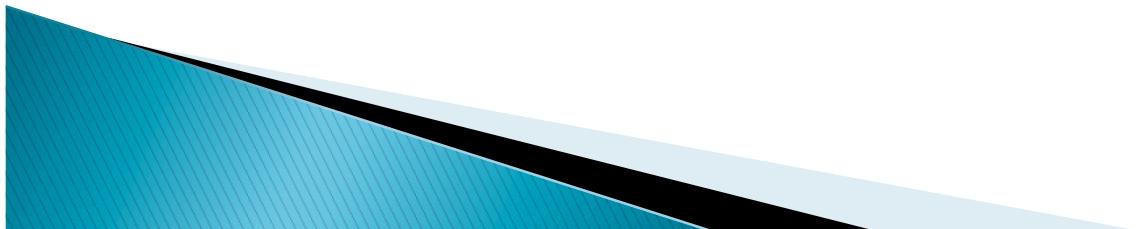
# Our hopes

- ▶ Enthusiasm!
- ▶ Engagement!



# Daily schedule (tentative)

- ▶ 9am – lecture
- ▶ 10:30am – tutorial 1
- ▶ 12–1pm – lunch
- ▶ 1:15pm – tutorial 2
- ▶ 3pm – free time!
- ▶ 5–6:30 – dinner
- ▶ 7pm – tutorial/lecture



# Weekly schedule – tentative wk1

- ▶ Tuesday – BLAST, cloud computing, scripting
- ▶ Wed – mapping
- ▶ Thursday – assembly
- ▶ Friday – pipelining
- ▶ Saturday – statistics & graphing



# Weekly schedule – tentative wk2

- ▶ Monday – mRNAseq
- ▶ Tuesday – ChIP-seq
- ▶ Wednesday – resequencing
- ▶ Thursday – genome assembly
- ▶ Friday (until noon) – postmortem!

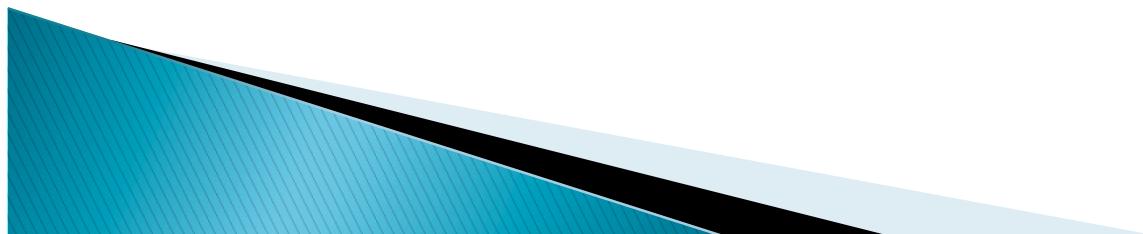


# Dramatis personae

- ▶ Titus Brown (that's me)
- ▶ Ian Dworkin -- co-instructor
- ▶ Rosangela Canino-Koning (“Rose”) – cruise director
- ▶ Chris Chandler – TA
- ▶ Jason Pell – TA
- ▶ Likit Preeyanon -TA
- ▶ Chris Radek: go-fer.



# “Cruise director”?



# For this evening...

- ▶ Get connected to the network!
- ▶ Make sure you have an EC2 account. Hint: if you didn't get a phone call, you don't.
- ▶ Enter your dropbox information.
- ▶ Once I set up dropbox access, put your slide in the “welcome” folder...

