

WELCOME TO PUBS!

Physical Underpinnings of Biological Systems - 2014

<http://fraserlab.com/pubs/>

Introductions



James/Jaime
instructor



David/Iggy
course coordinator



Joe
instructor emeritus



Zairan



Garrett

Alain



Ben



Laura
microscopy coordinator



Kyle



Samuel



Tanja



Clint

Interdisciplinary Graduate Training in Teaching Labs

Ronald D. Vale,^{1,2,3*} Joseph DeRisi,^{2,3} Rob Phillips,⁴ R. Dyche Mullins,^{1,2} Clare Waterman,^{1,5}
Timothy J. Mitchison^{1,6}

Intensive, short-term courses meld students and faculty and new techniques in pursuit of genuine research questions.

Science 21 December 2012:
Vol. 338 no. 6114 pp. 1542-1543
DOI: 10.1126/science.1216570

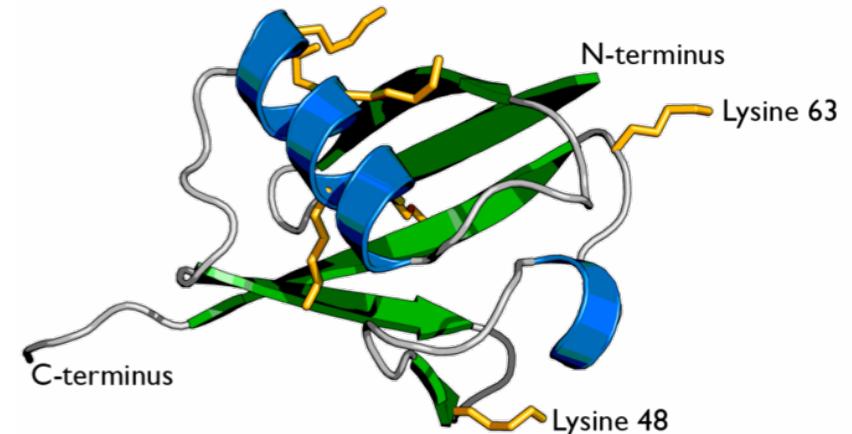
- We have three graduate programs (BMI, BP, CCB) represented - and many diverse scientific backgrounds - this is a huge advantage
- David/Iggy and Dan Bolon worked really hard this summer to get everything in place
- This course is an experiment in hands-on **team**-based learning. You will be exposed to: deep sequencing, genetics, chemical biology, systems biology, protein biophysics, evolutionary biology, statistical mechanics, computational biology... etc...
- Lecturers (and we have a great line up of faculty!) will reinforce broad themes, but you will drive the research questions, day-to-day experiments, and code forward!
- So... why is it called PUBS?

Ubiquitin is the central protein in “proteostasis”

- Ubiquitin (Ub) is a PTM that targets proteins for degradation
 - proteins marked with a tetra-K48 Ub chain targeted for proteasomal degradation
- but... Ub contains multiple other lysine residues
 - these lysine residues can direct other functions (e.g. DNA damage response, membrane trafficking, transcription - discussed in assigned *Finley review*)

**A major question in the Ub field:
What are the roles of non-K48-linked Ub chains?**

- Almost every part of Ub is used in some protein-protein interaction surface
 - most important is the “hydrophobic patch”
 - Ub-Ub, Ub-E2/E3 Writer, Ub-Dub Eraser, Ub-Reader interactions
- Biophysicists love Ub too!



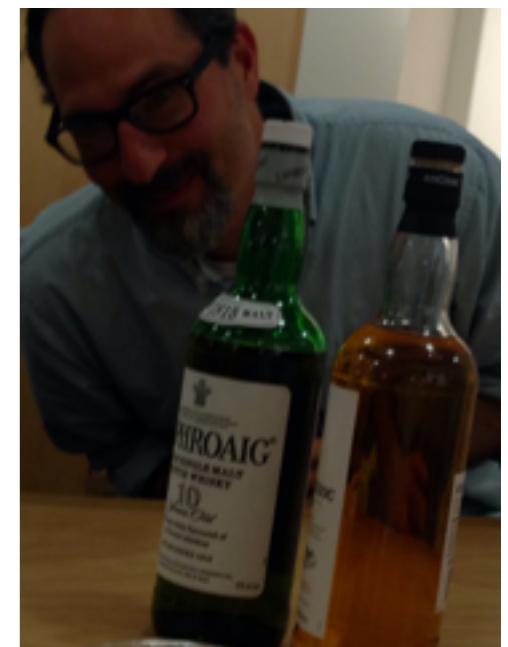
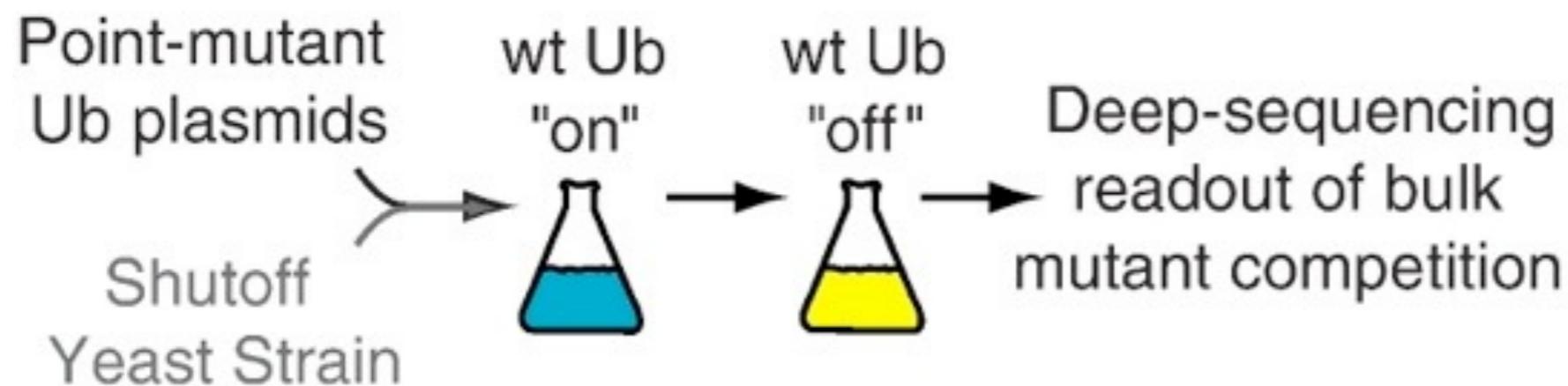
Ubiquitin is highly (ULTRA) conserved in evolution

Organism	Sequence Alignment	Swiss-P
Amoeba	MQIFVKTLTGKTITLEVESSDTIENV KQK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P49634
Green alga	MQIFVKTLTGKTITLEVESSDT V ENV KSK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P42739
Chlamyd. reinhardtii	MQIFVKTLTGKTITLEVESSDTIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P14624
Mouse	MQIFVKTLTGKTITLEVEPSDTIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P62991
Human (*)	MQIFVKTLTGKTITLEVEPSDTIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P62988
Slime mold	MQIFVKTLTGKTITLEVE GSDN IENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P08618
Purple sea urchin	MQIFVKTLTGKTITLEVE PSD IENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P23398
Eimeria bovis	MQIFVKTLTGKTITLE DVEPSD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P46574
T. pyriformis	MQIFVKTLTGKTITLE DVEASD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P20685
C. elegans	MQIFVKTLTGKTITLE EASD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P14792
Red alga	MQIFVKTLTGKTITLEVE ASD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P42740
Neurospora crassa	MQIFVKTLTGKTITLEVESSDT IDNV KQK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P13117
Baker's yeast	MQIFVKTLTGKTITLEVESSDT IDNV KSK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P61864
Inky cap fungus	MQIFVKTLTGKTITLEVESSDT IDNV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P19848
Garden pea (**)	MQIFVKTLTGKTITLEVESSDT IDNV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P03993
Euplotes eurystomus	MQIFVKTLTGKTITLE DVEQSD T IDNV KTK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P23324
Potato late blight fungus	MQIFVKTLTGKTITLE DVEPSD S IDNV KQK I Q DKEGIPPDQQQLI FAG KQLE D GRTL S DYN I Q E STL L HLVLRLRGG	P22589
Leishmania major	MQIFVKTLTGKTIALE EVEPSD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE E GRTL S DYN I Q E STL L HLVLRLRGG	Q05550
Sauroleish. parentolae	MQIFVKTLTGTTIALE EVEPSD TIENV KAK I Q DKEGIPPDQQQLI FAD KQLE E GRTL S DYN I Q E STL L HLVLRLRGG	P49635
T. brucei brucei	MQIFVKTLTGKTIALE EVEASD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE E GRTL A DYN I Q E STL L HLVLRLRGG	P15174
Trypanosoma cruzi	MQIFVKTLTGKTIALE VESSD TIENV KAK I Q DKEGIPPDQQQLI FAG KQLE D GRTL A DYN I Q E STL L HLVLRLRGG	P08565

1 10 20 30 40 50 60 70 76

...only 3 substitutions from yeast to human

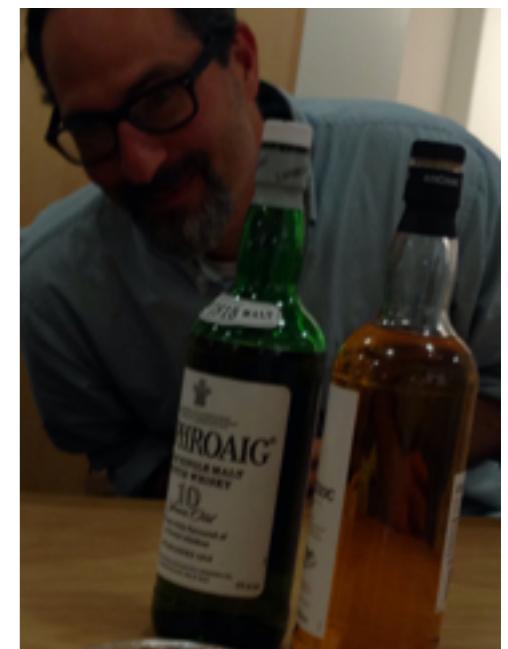
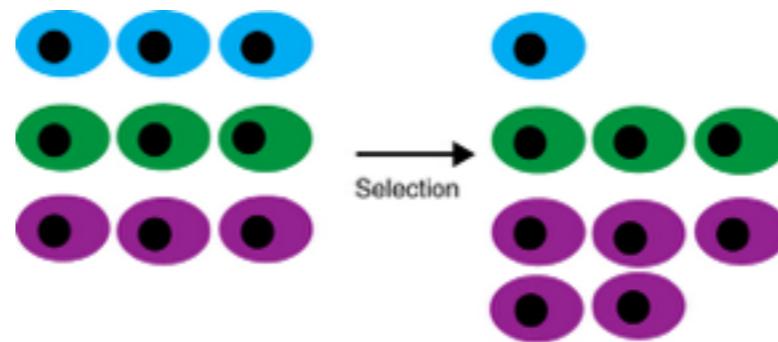
Dan Bolon asked how this compares to growth in rich media...



Dan Bolon

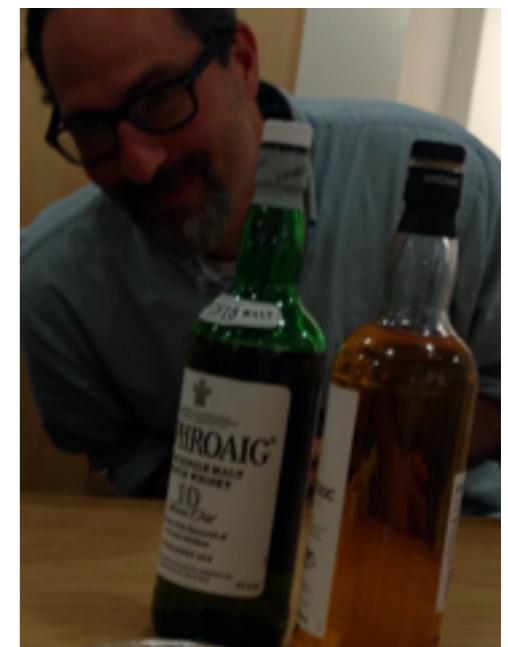
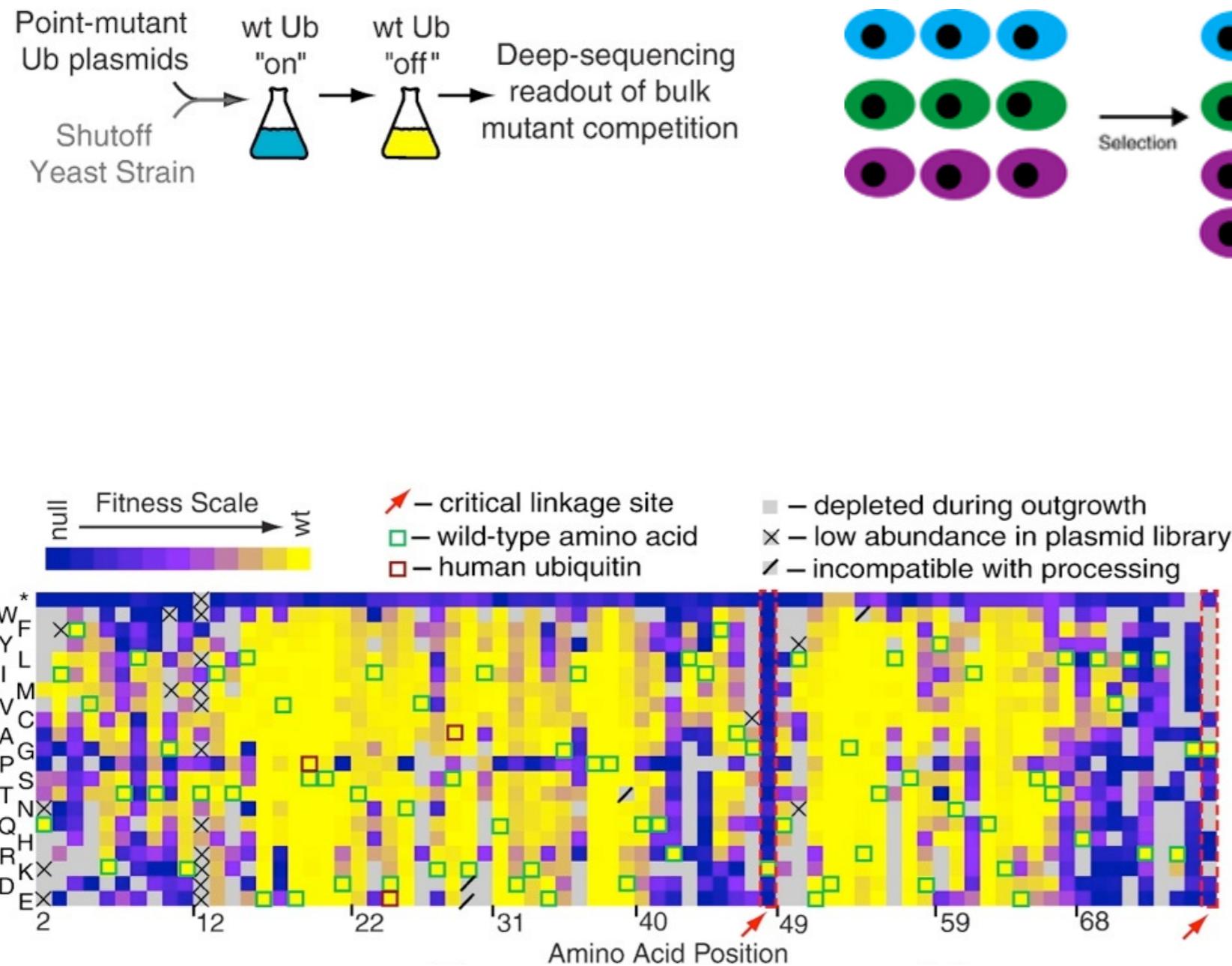
Over time the “**bad**” Ub variants die out and the “**good**” dominate

Point-mutant
Ub plasmids → wt Ub
"on" → wt Ub
"off" → Deep-sequencing
readout of bulk
mutant competition
Shutoff
Yeast Strain



Dan Bolon

Dan assembled this in a matrix of fitness values by amino acid type by position



Dan Bolon

Ub is highly conserved in nature... yet highly mutable in these selections

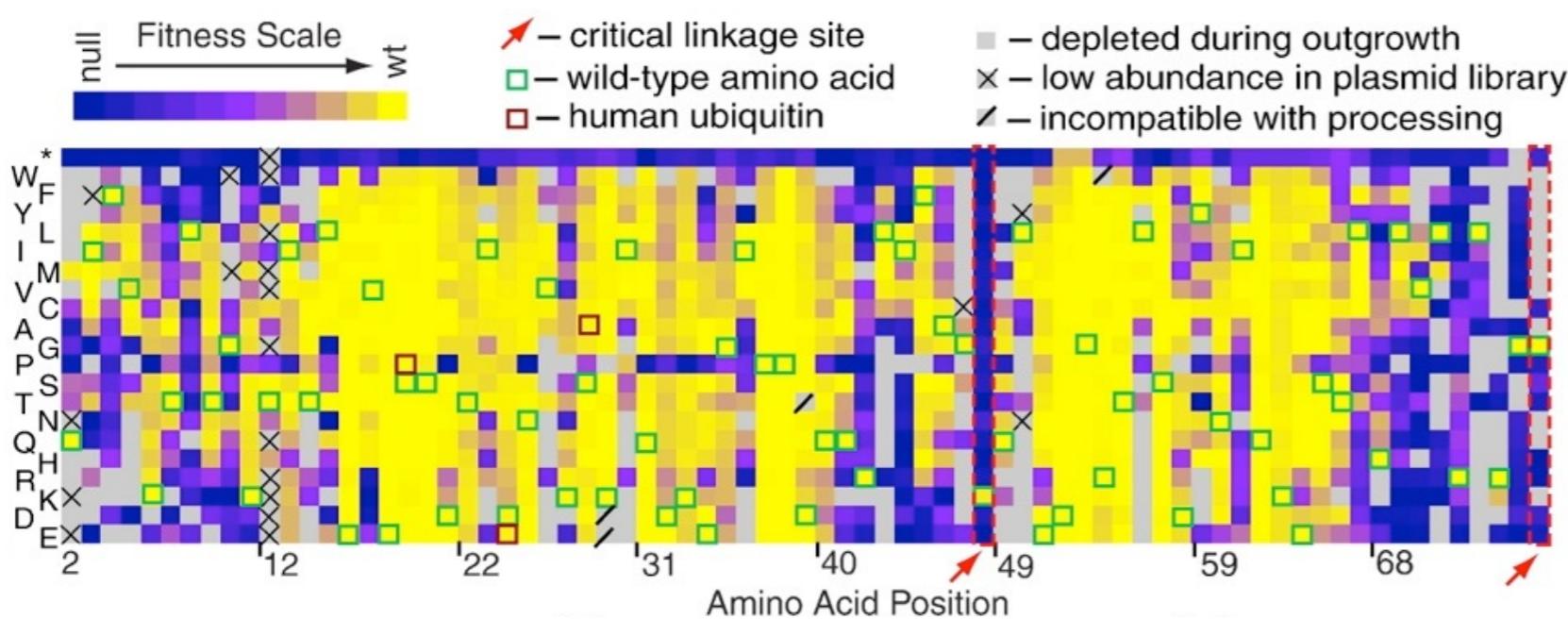
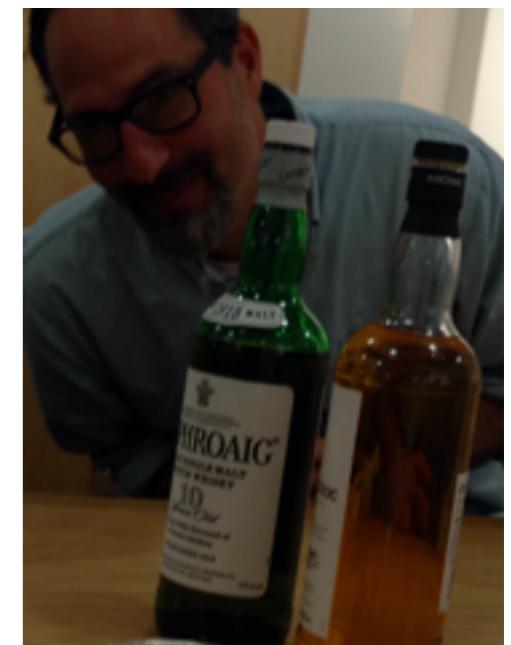
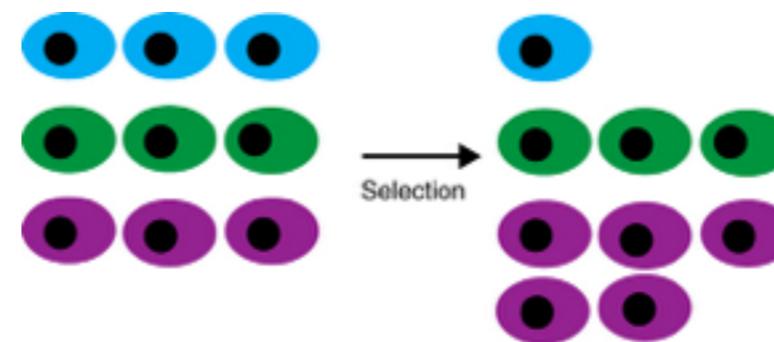
Point-mutant
Ub plasmids

Shutoff
Yeast Strain

wt Ub "on"

wt Ub "off"

Deep-sequencing
readout of bulk
mutant competition



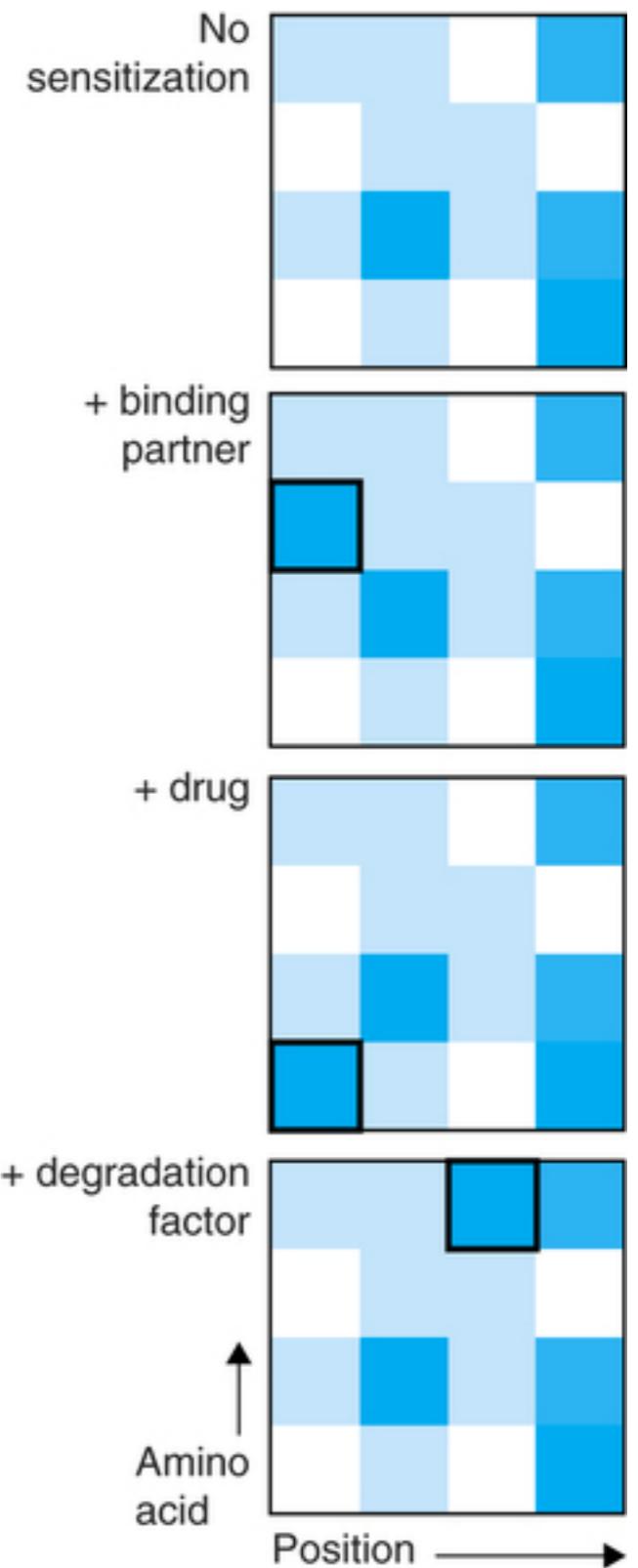
Dan Bolon

Why is the evolutionary history so different from the selection experiment?

hint... assigned reading: Fowler and Fields, Nature Methods, 2014

Why is the evolutionary history so different from the selection experiment?

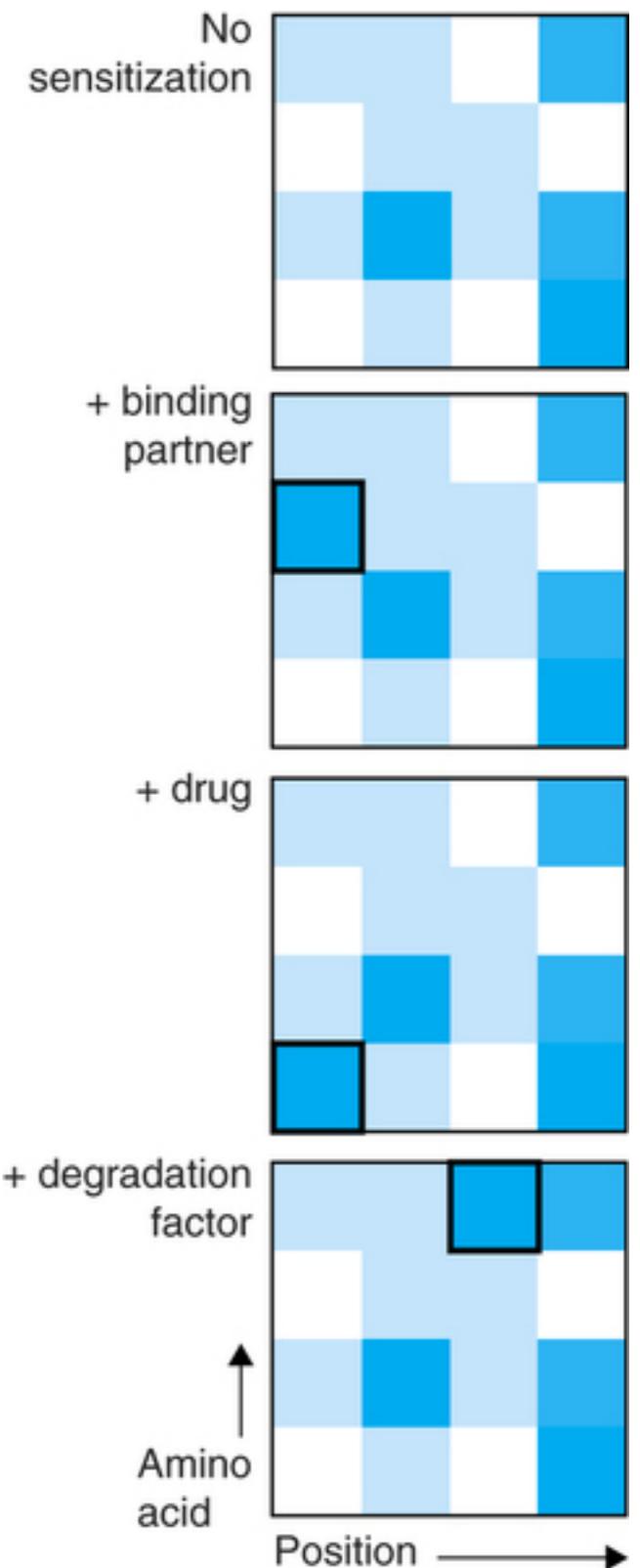
How do different environments
(chemical perturbations) alter the Ub
fitness landscape?



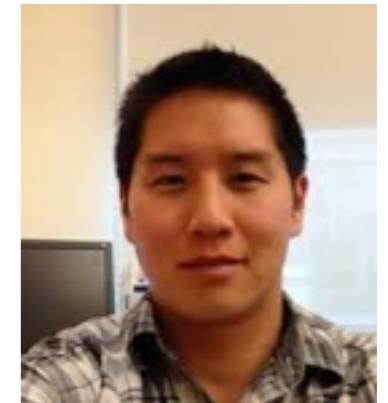
Why is the evolutionary history so different from the selection experiment?

How do different environments
(chemical perturbations) alter the Ub
fitness landscape?

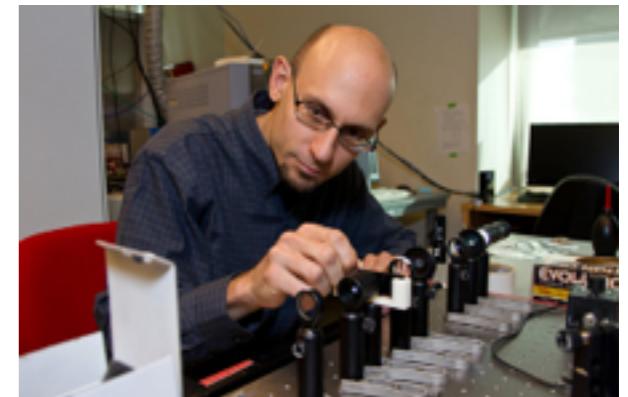
**EACH TEAM WILL EXAMINE
A DIFFERENT PERTURBATION**



- Do the perturbations have different effects the landscape of all possible mutations?
(sequencing - CAT)



- Is the effect equal for all cells or does it create populations with different growth rates?
(microscopy - NIC)

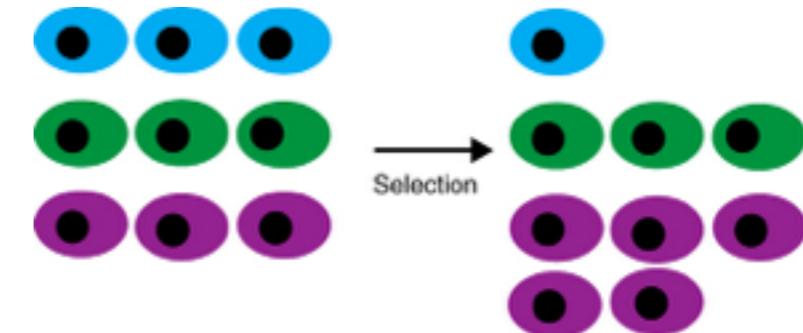
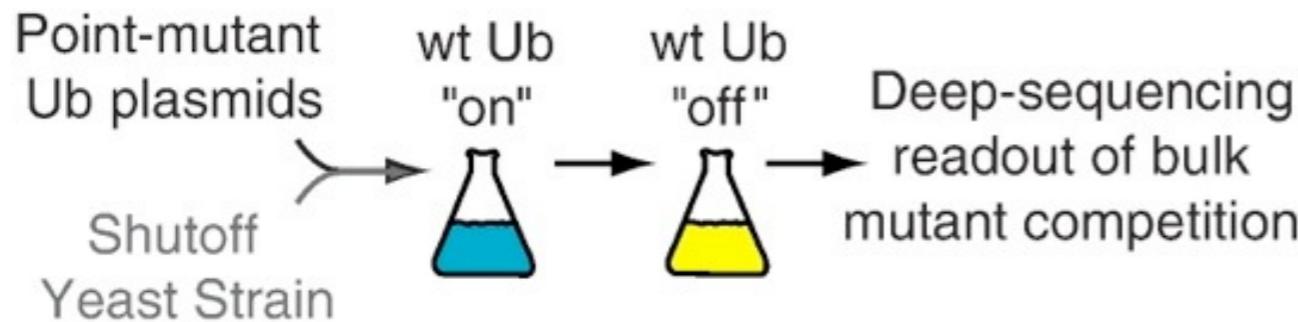


- Can we model the selective pressure biophysically?
(protein design - QB3 Cluster)



- Week 1: transformations and determining the optimal concentrations of chemicals - one late day
 - Week 2: sample the transformed library (2x) - two very early and very late days
 - Week 3: prepare the library for sequencing
 - Week 4: analyze the effect of the chemical perturbation on all possible Ub mutants
 - Week 5: **Presentations** and compare datasets between teams
-
- Week 6: compare bulk and single cell growth rates
-
- Week 7: computational protein design to explain sequencing results
 - Week 8: comparisons between design and selections
-
- November 25th: **Final Presentations** and Party!

How does the sequencing experiment work?



Ubiquitin sequence divided into eight regions for accurate and efficient analyses

MQIFVKTLTGKTITLEVESSDTIDNVKSQIQDKEGIPPDQQQLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGG

1 2 3 4 5 6 7 8
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

Generated libraries of saturation mutants at each amino acid position

Transformed libraries into yeast, outgrowth, shutoff and sampling over time in competition

but wait - the Bolon approach took weeks for library preparation for sequencing!

- The Bolon/Mavor Barcoded library will help us out here

~5000 alleles

50,000 BCs

~70 Billion possible BCs



PCR product



261/31 bp paired end read ↓

Association of nucleotide sequence with barcode (done)
Association of amino acid changes with barcode (you!)

- Pickles are a way to dump out python data structures as files, allowing easy transfer of data between scripts
- ```
import cPickle as pic
data = pic.load(open("filename.pkl","rb"))
print data
```
- We are giving you 3 pickles (<http://fraserlab.com/pubs/>):
  - allele\_dict.pkl - contains a dictionary where:  
key = barcode nucleotide sequence  
value = residuenumber\_codon  
(residuenumber is in protein space, codon is in nucleotides!)
  - translate.pkl - contains a dictionary where:  
key = codon  
value = amino acid
  - aminotonumber.pkl - contains a dictionary where:  
key = amino acid  
value = number  
(useful for plotting)
- Many barcodes can map to the same codon, and (for some amino acids) many codons can map to the same amino acid
- We want to know how many barcodes there are for each possible amino acid mutation of Ub - 3min presentations from each team tomorrow!

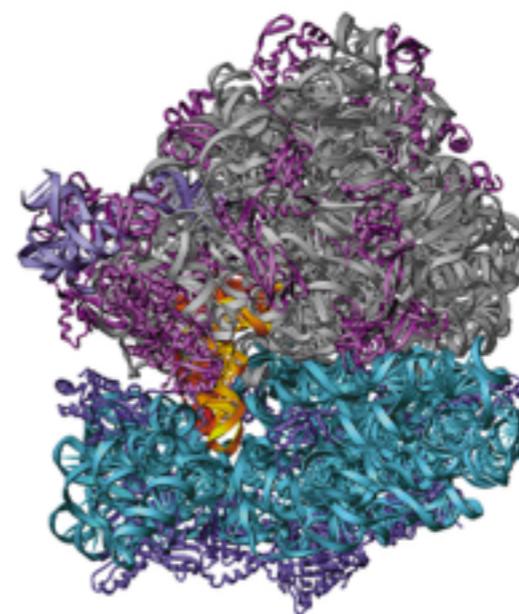


# This week

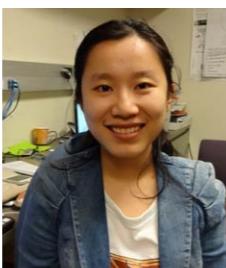
- Tuesday: brief presentations of `ribosome_barcodes.py`; transformation (competition between teams for highest efficiency)
- Wednesday: growth curves to determine the optimal chemical concentration (team organization will be key for taking multiple time points)

# Today, we have to accomplish 3 tasks

- The teams need **names!**  
Each team will get a different  
chemical perturbation
- Joe needs to give each team  
an **account** on the server  
<http://fraserlab.com/pubs/server/>
- We need you to convert the  
barcodes from nucleotide  
space to amino acid space  
**(ribosome\_barcode.py)**



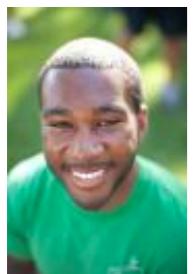
Teams?



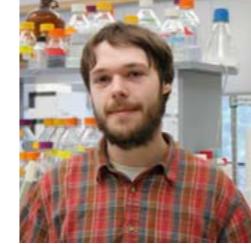
???



???



???



???



???