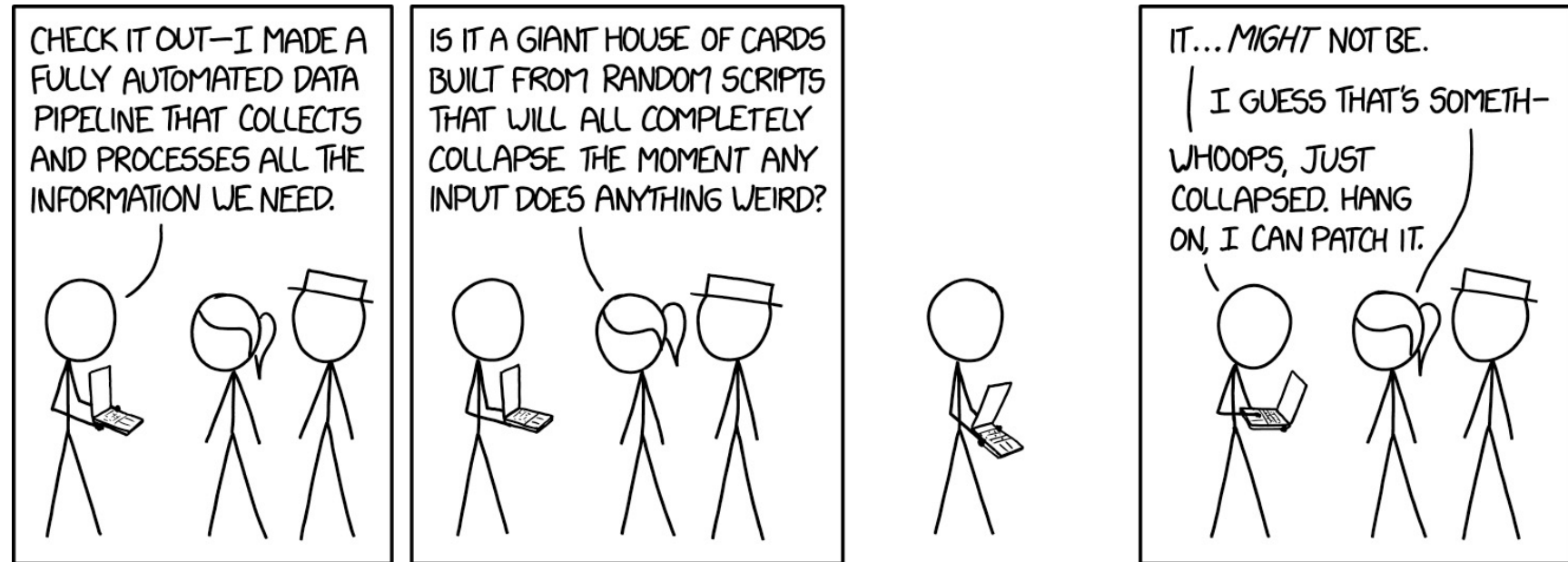
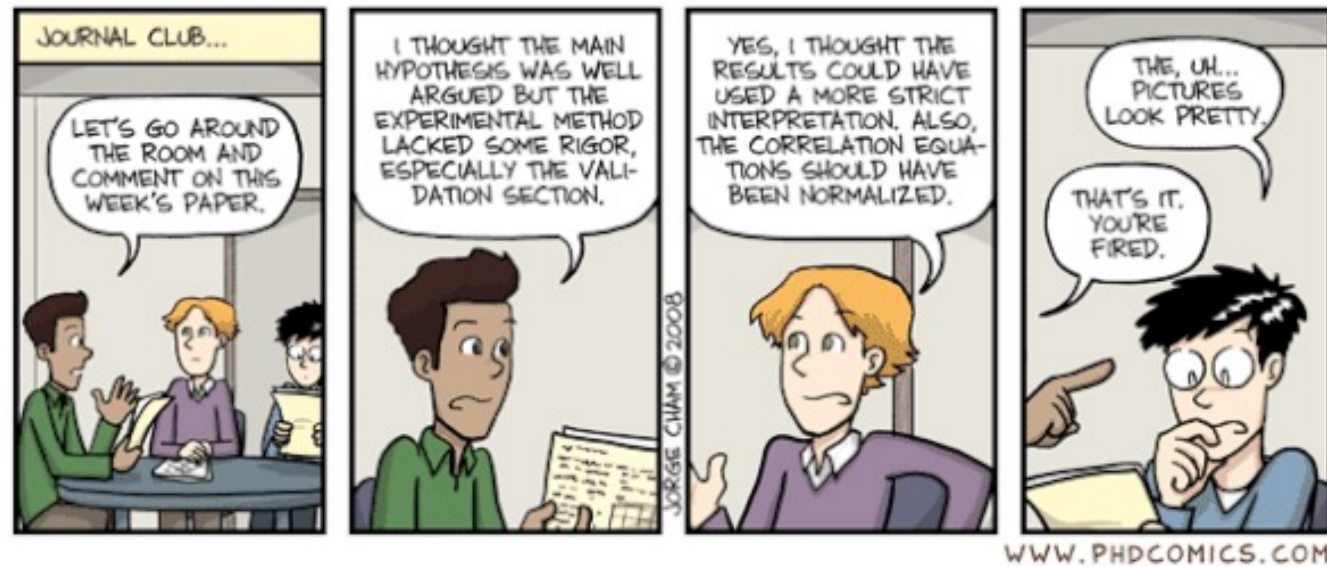


M2D4: Analyze SMM data to identify putative small molecule binders

- Prelab discussion
- Analyze SMM data
- Quiz



Notes on Journal Club...



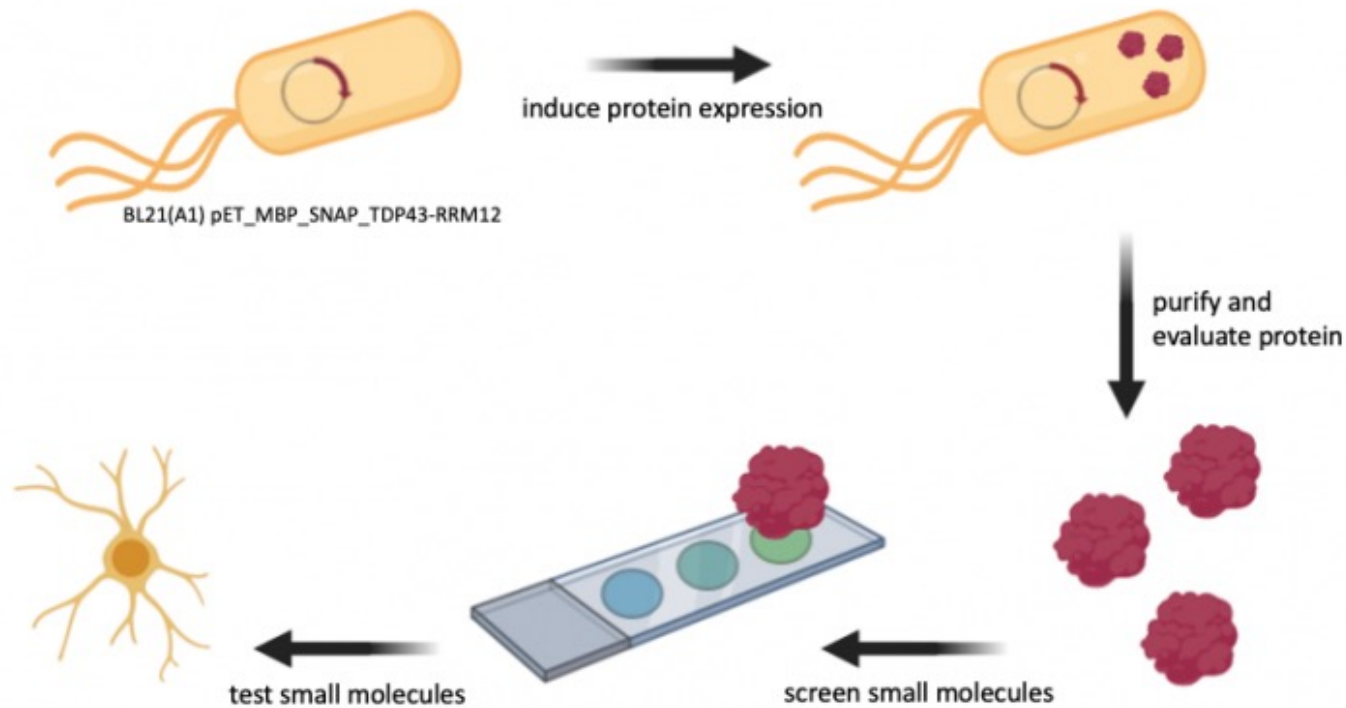
- Be sure to practice
- Time yourself– it's very easy to run too long on Zoom!
- Journal Clubs will be recorded, but not shared
 - You will watch them with Noreen as part of feedback
- Participation points are awarded for asking good questions of your classmates

submit
slides
1 hr prior
to class

2 questions

Overview of Mod2 experiments

Research goal: Identify and characterize small molecule binders to a protein drug target



Use jupyter notebook to:

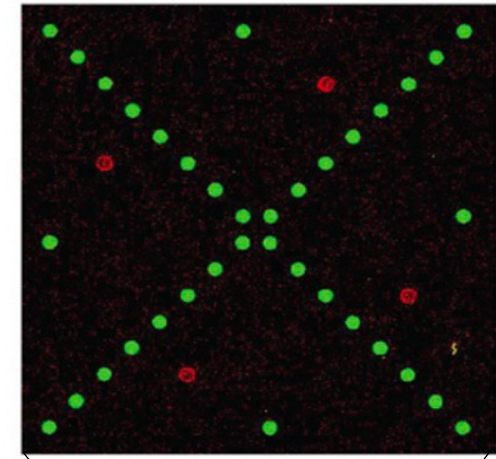
- analyze hits
- evaluate controls

Evaluate binders for common features

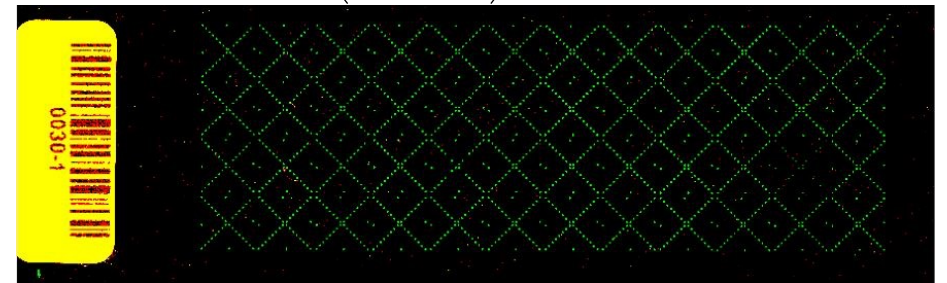
Workflow for SMM data analysis

1. Align spots using fluorescence on 532 nm channel (sentinel spots)
↓
2. Quantify fluorescence on 635 nm channel
↓
3. Identify 'hits' with improbably high fluorescence
↓
4. Complete 'by eye' analysis of putative hits to manually remove false positives

*le47-SMAP
- protein*

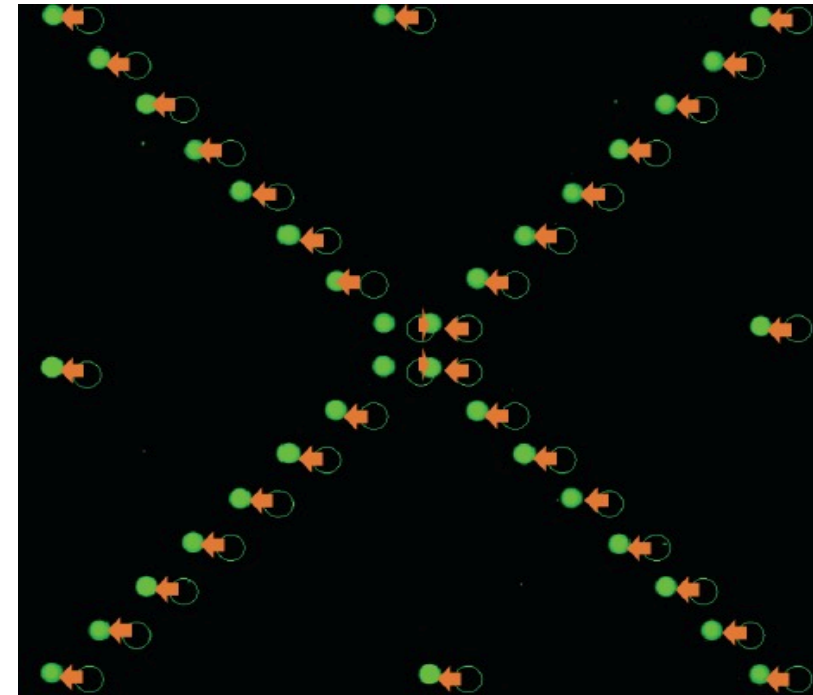
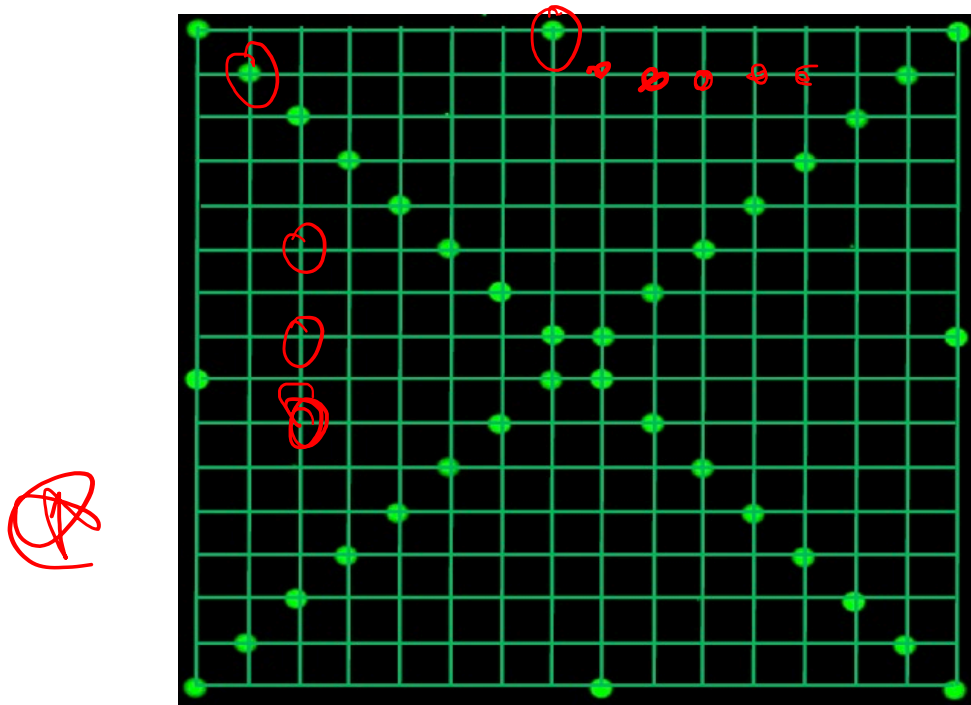


*Genepix
Software
SMM
Scanner*



Align SMM using sentinel spots

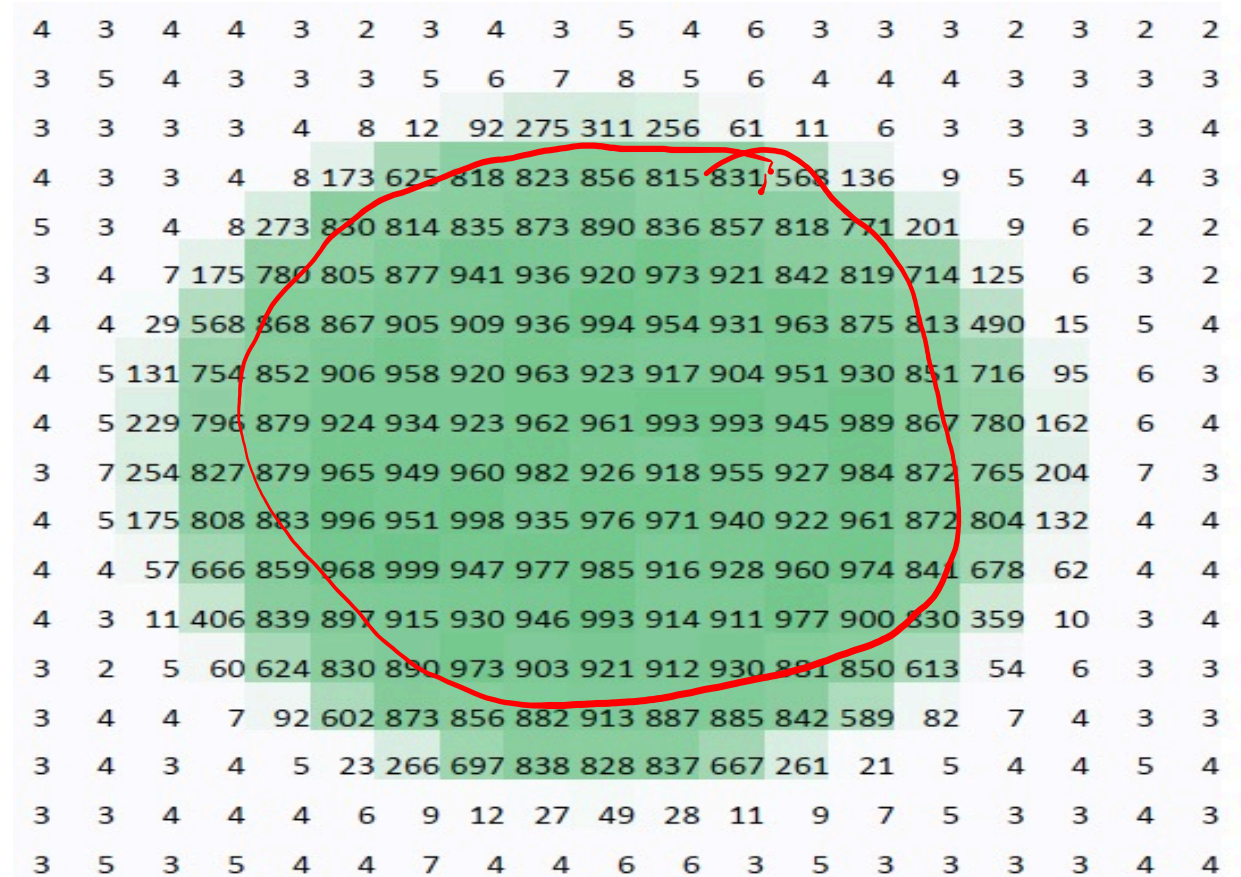
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels



alignment

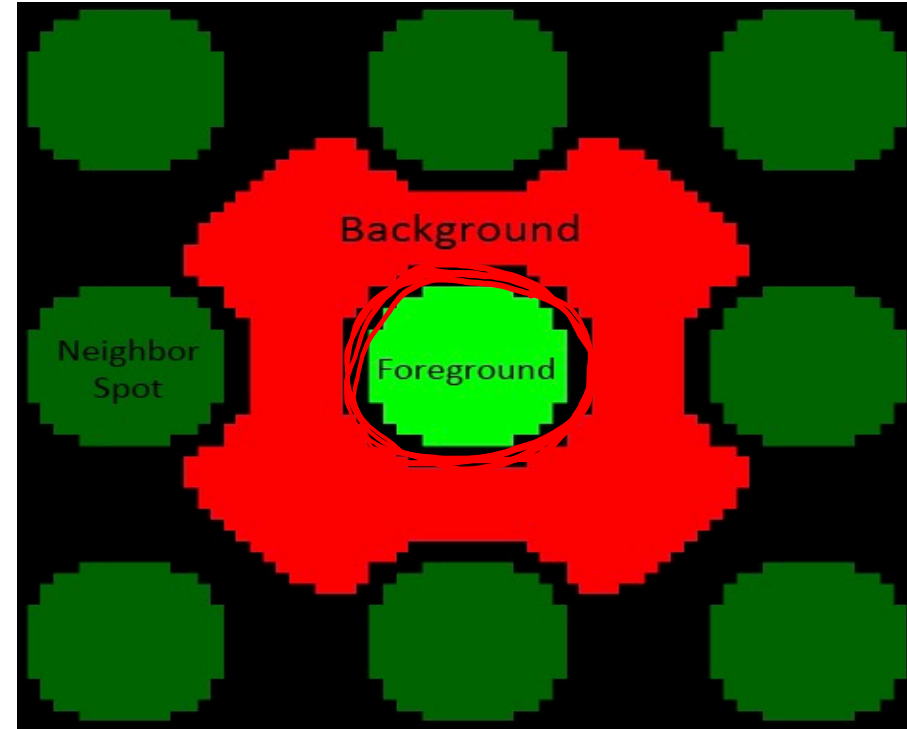
Spots are represented by an array of numerical values

- Each pixel is represented by a number that indicates intensity of the signal
- Computational analysis used to define 'hits'



Fluorescence is quantified to identify hits

- Foreground: *- location of SM spot*
- most signal
- Background:
- nonspecific
- residual ligand

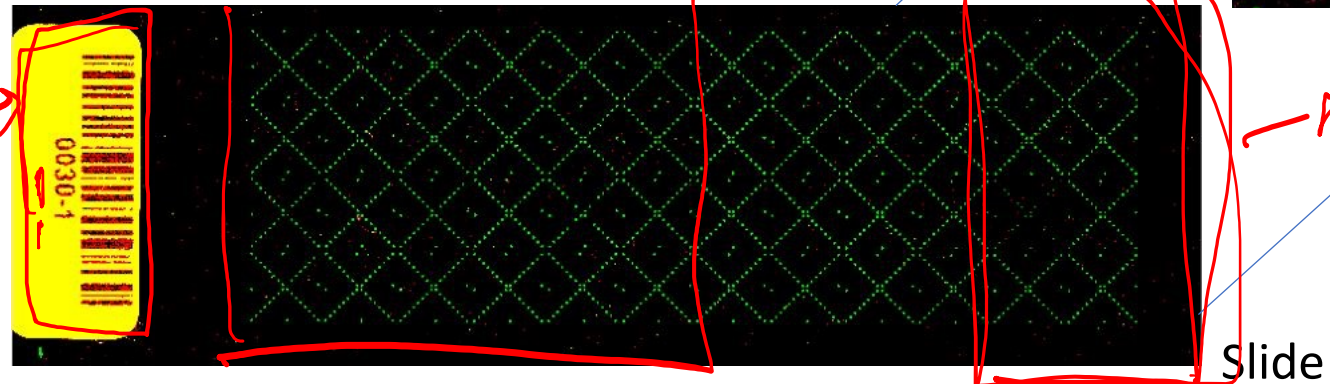


$$\text{Signal-to-noise ratio (SNR)} = \frac{\mu_{\text{foreground}} - \mu_{\text{background}}}{\sigma_{\text{background}}}$$

How will you identify hits from the SMM data?

First, consider bias that exists in the data set

- Across all slides
- Within each slide
- Within each block



high signal

Then, identify hits with significantly higher fluorescence over background

Z-scores

Lastly, manually confirm hits to eliminate false positives

by eye

Identifying hits with significant fluorescence

Robust Z-score =

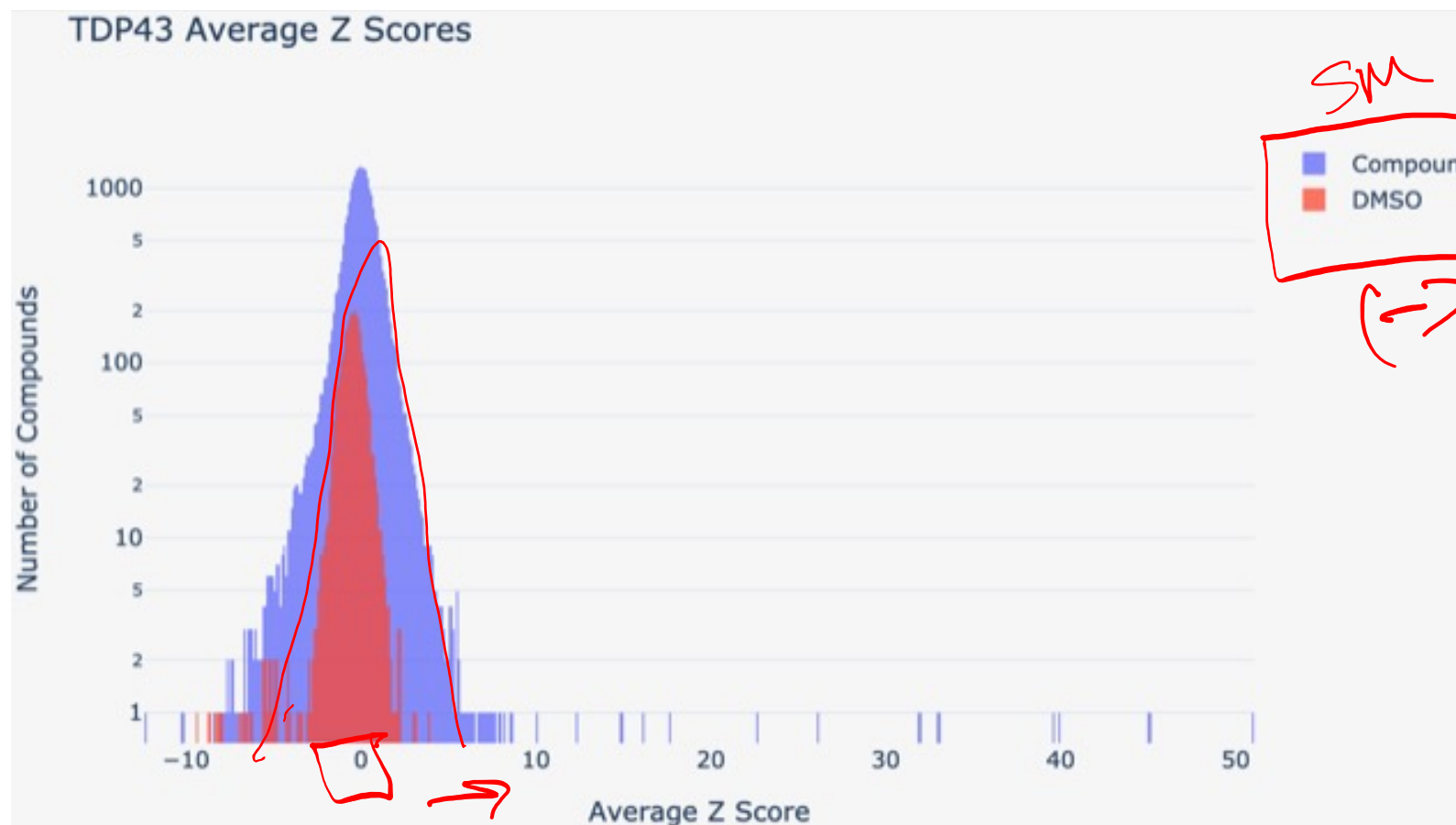
$$\frac{\text{SNR}_i - \text{median}(\text{SNR})}{\text{median}(|\text{SNR}_i - \text{median}(\text{SNR})|) * 1.48}$$

median absolute deviation (MAD)

scale factor for the normal distribution

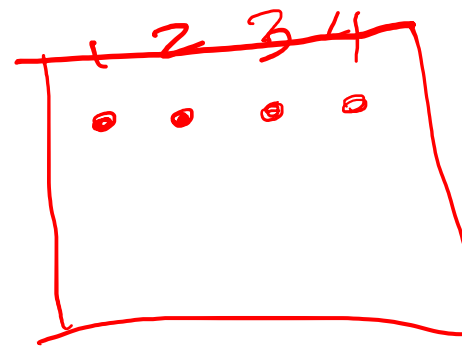
Robust Z-scores eliminate the influence of outliers

Average Z-score calculated for all compounds



-- Replicate spots averaged
-- Each count = unique compound

How will you determine a threshold Z-score?



Height at x (average z-score) =
number of compounds (y) with
that z-score or higher

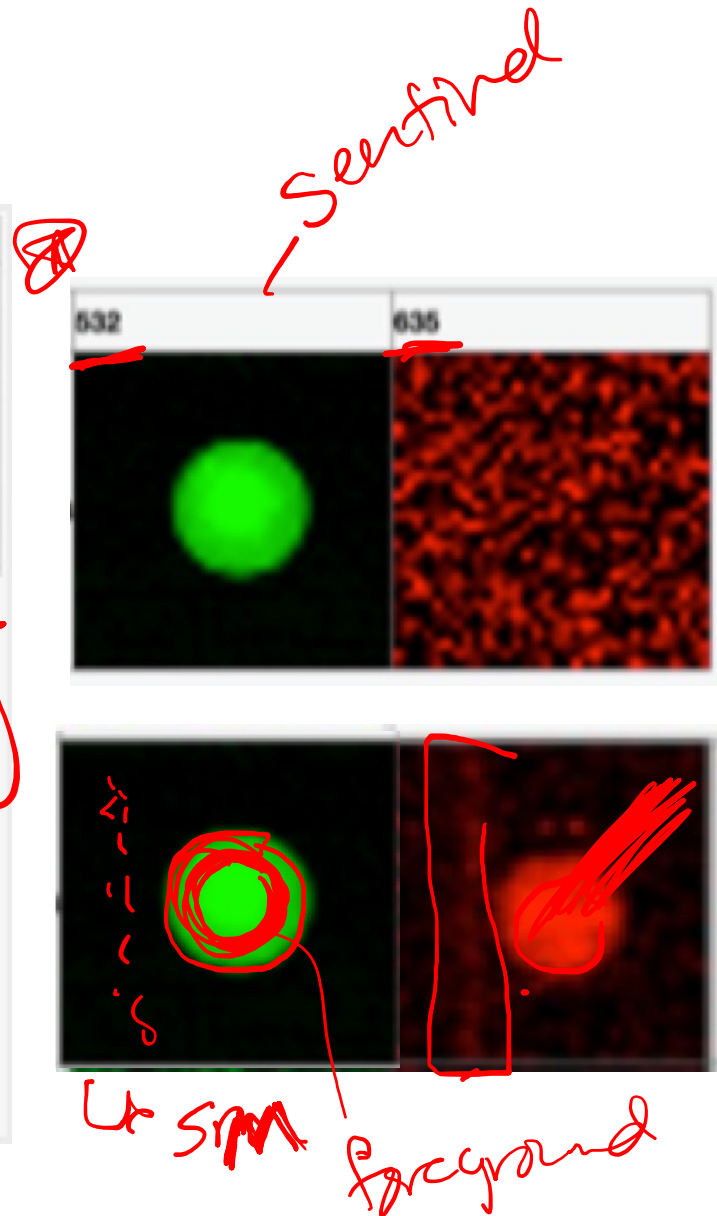
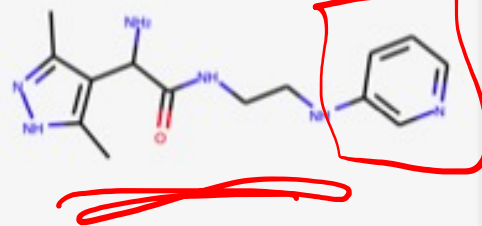
set threshold too \uparrow =
miss hits!

set threshold too \downarrow =
false positives

How will you validate hits?

T	ID	Robust Z	SMILES	Validated
49592	13:KI0001...	51.03151	C[C@H](C...	-1
42089	11:KI0001...	45.09263	CC1=C(C(...	example
6782	02:KI0001...	39.91118	CCNC(=O...	-1
29108	08:KI0001...	39.59436	C1C(C2=...	-1
44736	12:KI0001...	33.03555	C1CN(C2...	-1
29660	08:KI0001...	31.94118	CC1=NC2...	-1
11360	03:KI0001...	26.13059	C1CN(CC...	-1

532	635
0011-08	
0012-08	
0014-08	



M2D5HW (Due 4/13)

- show & discuss
- did as data
→ can show

- Create a figure of the purity and concentration data
- Figure must include a title and caption
- Write associated results and discussion paragraphs
 - Mod2 results text will not include interpretation of the data shown in the figure
 - Separate discussion section associated with figure with interpretation
- Review guidelines on the wiki homework tab!!

wiki

RESULTS

1. What was the overall goal of these data?
 - State concisely as an introductory sentence.
2. If applicable, what was the result of your control?
 - Was it expected?
3. What was your result?
 - Was it expected?
4. What does this motivate you to do next?
 - Specifically, what experiment follows?

DISCUSSION

1. What evidence do you have that your result is correct or incorrect?
 - How do your controls support your data?
2. In sum, what do your data suggest or indicate?
 - Do your data support your hypothesis? Why?
3. What does this motivate you to do next?
 - Specifically, what is the next research question?

For today...

- Work on SMM analysis
- Work on Journal Club
- Get a start on the homework due M2D6

Next week...

- JOURNAL CLUB!

For M2D6 (4/13)...

- Draft a figure, results, and discussion section of the purity and concentration data