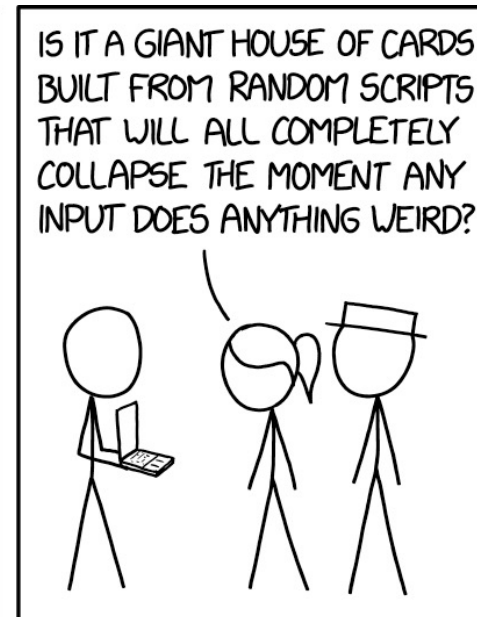
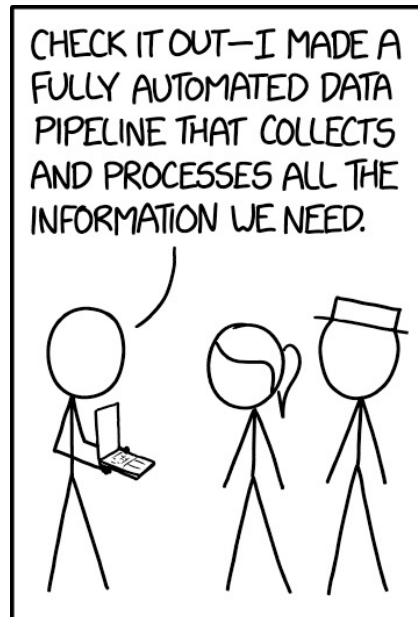
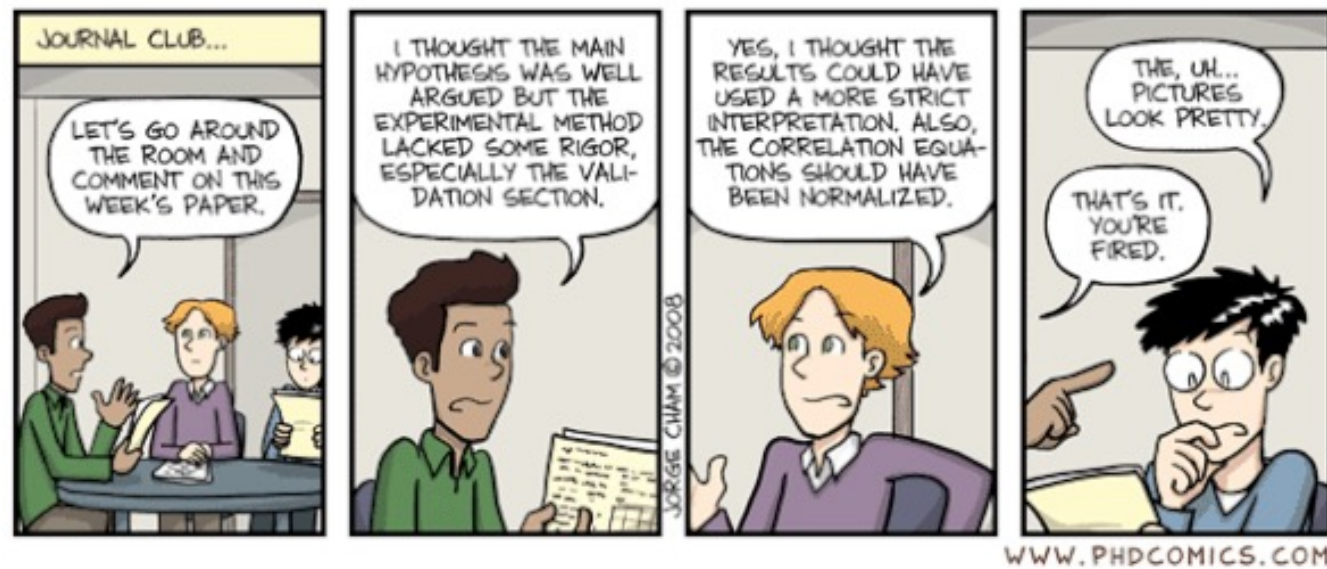


M2D4: Analyze SMM data to identify putative small molecule binders

- Prelab discussion
- Analyze SMM data
- Quiz



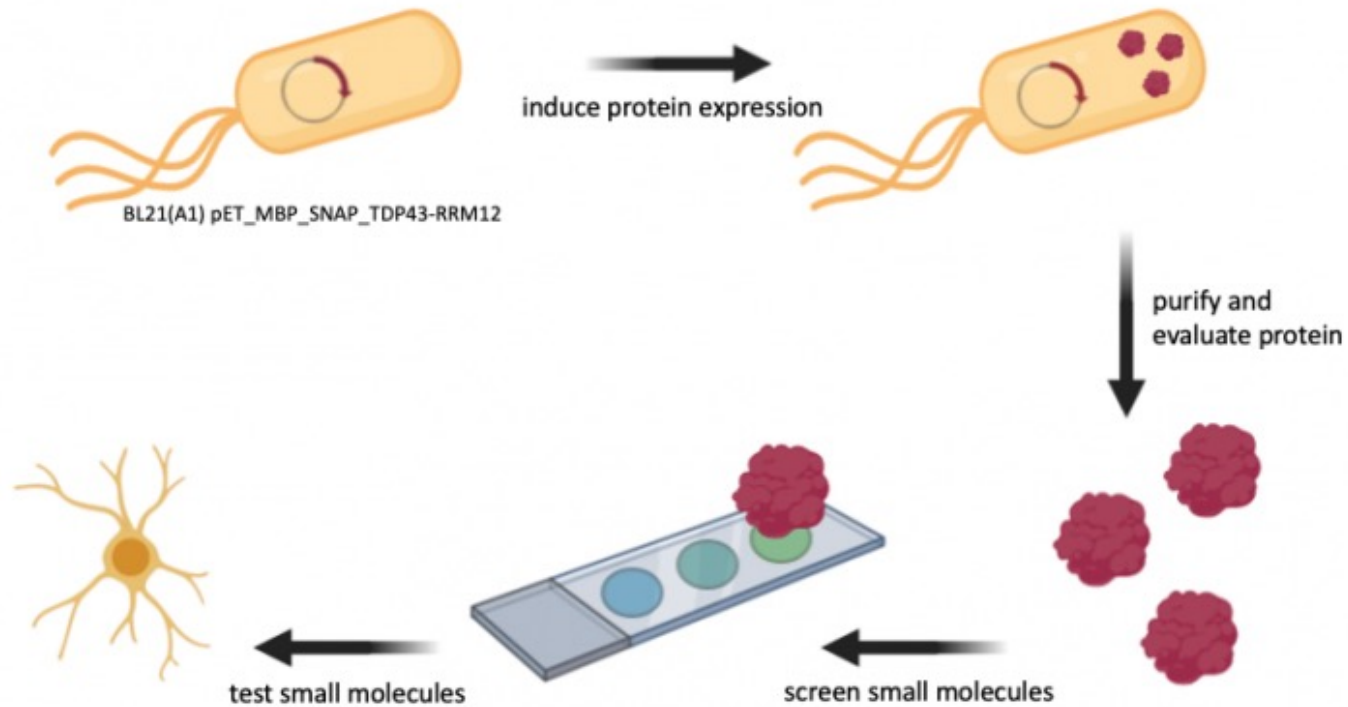
Notes on Journal Club...



- 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

Overview of Mod2 experiments

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



Use jupyter notebook to:

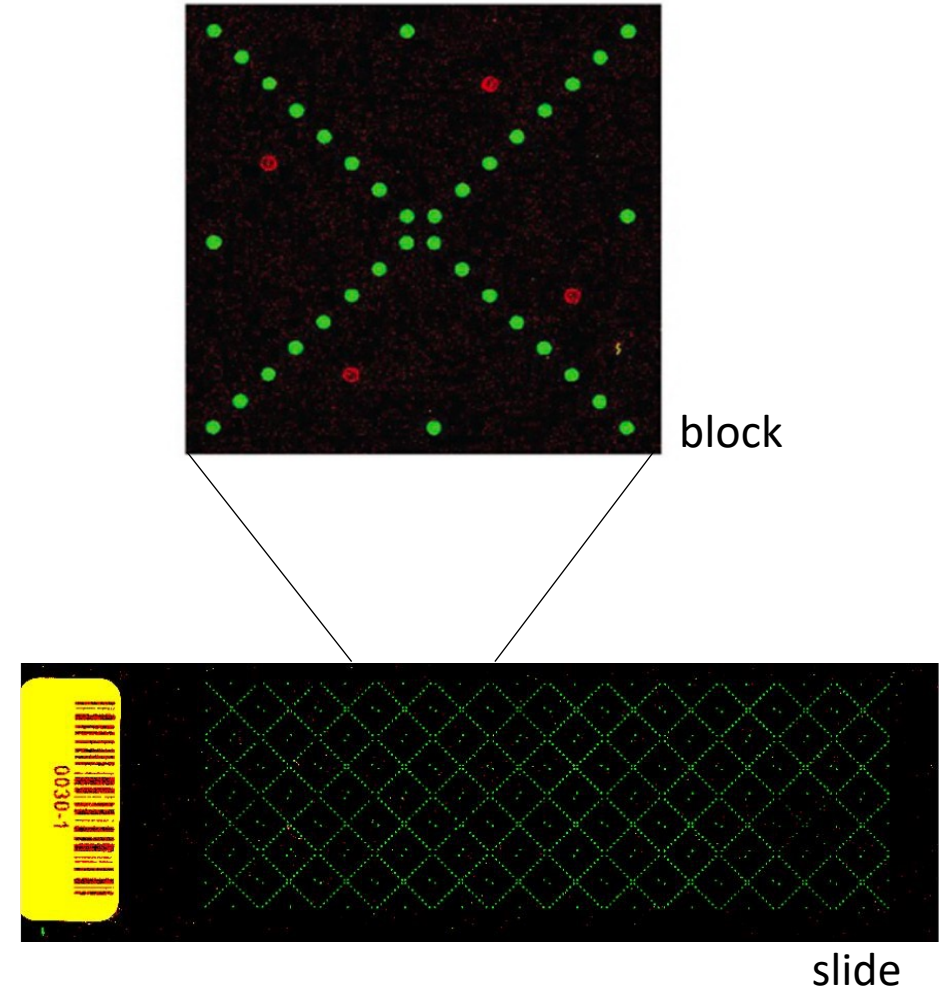
- analyze hits
- evaluate controls

Evaluate binders for common features



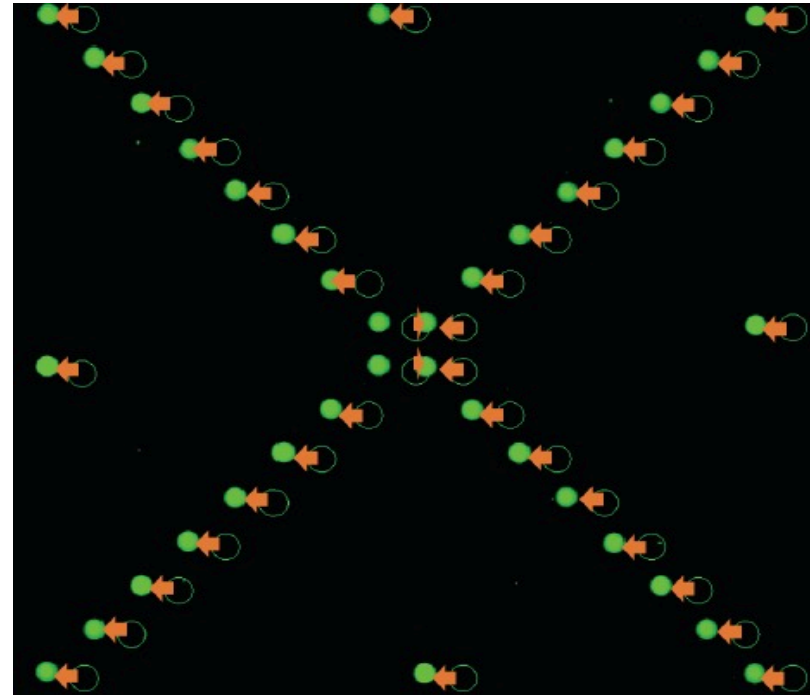
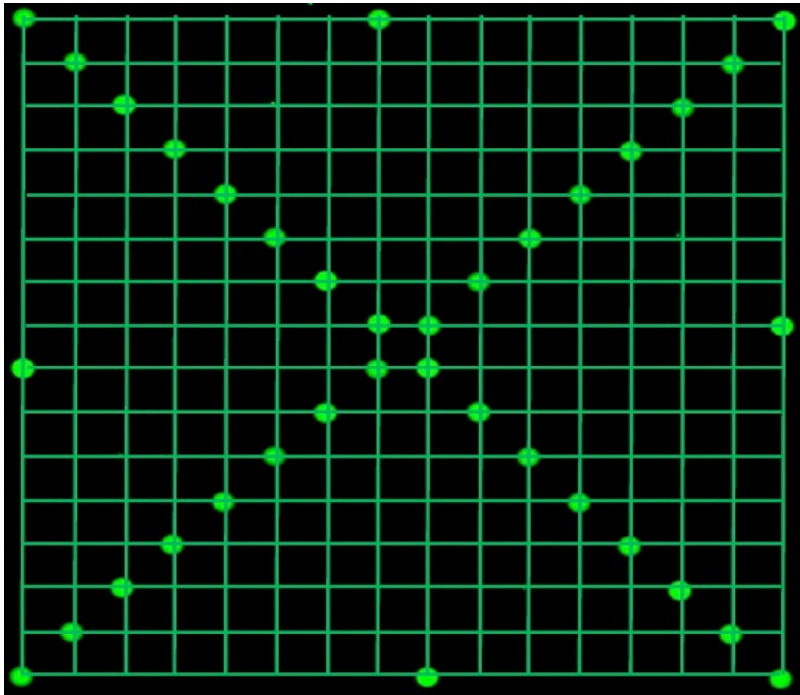
Workflow for SMM data analysis

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



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



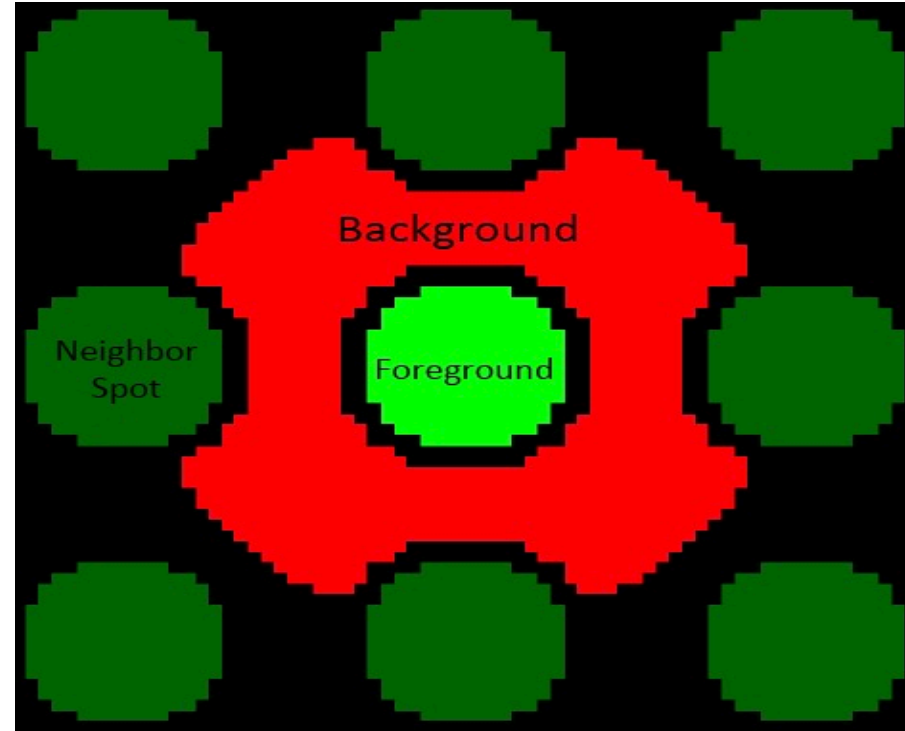
Spots are represented by an array of numerical values

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

4	3	4	4	3	2	3	4	3	5	4	6	3	3	3	2	3	2	2
3	5	4	3	3	3	5	6	7	8	5	6	4	4	4	3	3	3	3
3	3	3	3	4	8	12	92	275	311	256	61	11	6	3	3	3	3	4
4	3	3	4	8	173	625	818	823	856	815	831	568	136	9	5	4	4	3
5	3	4	8	273	830	814	835	873	890	836	857	818	771	201	9	6	2	2
3	4	7	175	780	805	877	941	936	920	973	921	842	819	714	125	6	3	2
4	4	29	568	868	867	905	909	936	994	954	931	963	875	813	490	15	5	4
4	5	131	754	852	906	958	920	963	923	917	904	951	930	851	716	95	6	3
4	5	229	796	879	924	934	923	962	961	993	993	945	989	867	780	162	6	4
3	7	254	827	879	965	949	960	982	926	918	955	927	984	872	765	204	7	3
4	5	175	808	883	996	951	998	935	976	971	940	922	961	872	804	132	4	4
4	4	57	666	859	968	999	947	977	985	916	928	960	974	841	678	62	4	4
4	3	11	406	839	897	915	930	946	993	914	911	977	900	830	359	10	3	4
3	2	5	60	624	830	890	973	903	921	912	930	881	850	613	54	6	3	3
3	4	4	7	92	602	873	856	882	913	887	885	842	589	82	7	4	3	3
3	4	3	4	5	23	266	697	838	828	837	667	261	21	5	4	4	5	4
3	3	4	4	4	6	9	12	27	49	28	11	9	7	5	3	3	4	3
3	5	3	5	4	4	7	4	4	6	6	3	5	3	3	3	3	4	4

Fluorescence is quantified to identify hits

- Foreground:
- Background:

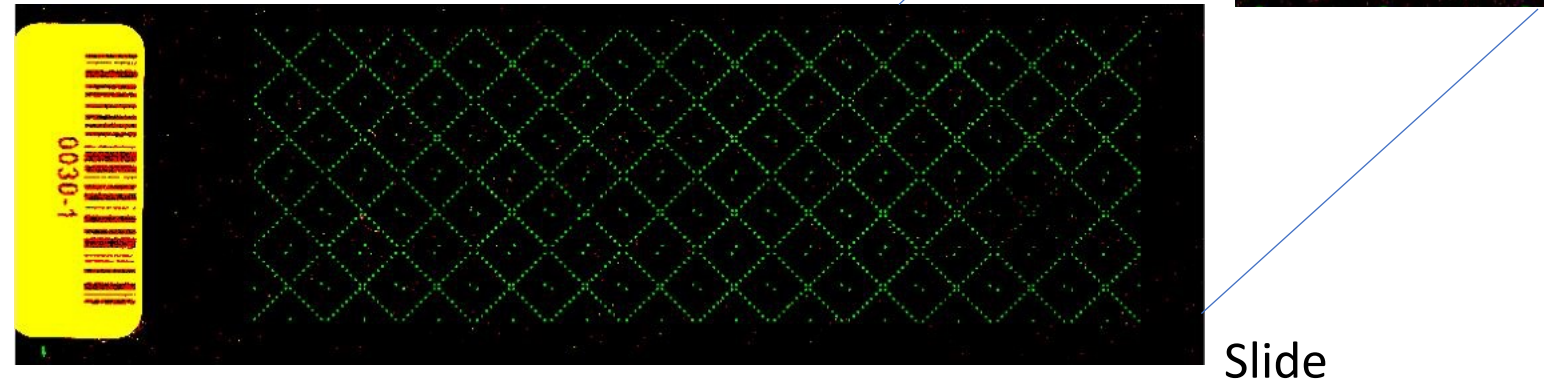


$$\text{Signal-to-noise ratio (SNR)} = \frac{\mu_{\text{foreground}} - \mu_{\text{background}}}{\sigma_{\text{background}}}$$

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



Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

Identifying hits with significant fluorescence

$$\text{Robust Z-score} = \frac{\text{SNR}_i - \text{median}(\text{SNR})}{\text{median}(|\text{SNR}_i - \text{median}(\text{SNR})|) * 1.48}$$

median absolute deviation (MAD)

scale factor for the normal distribution

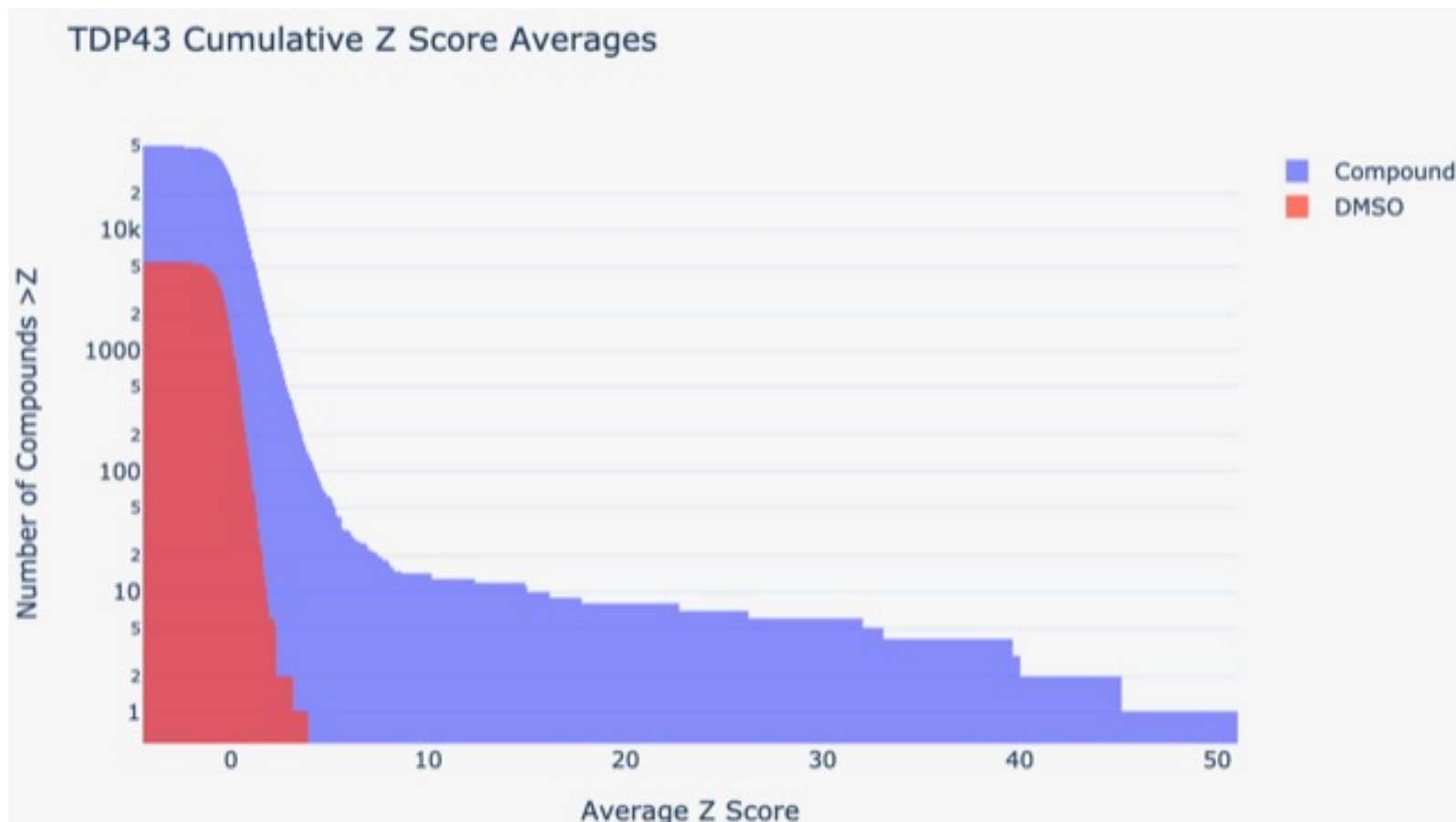
Robust Z-scores eliminate the influence of outliers

Average Z-score calculated for all compounds



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

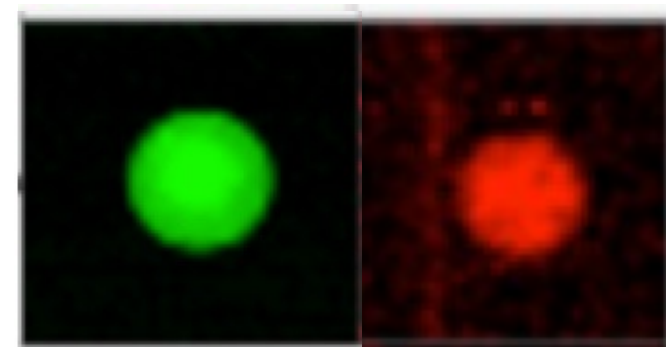
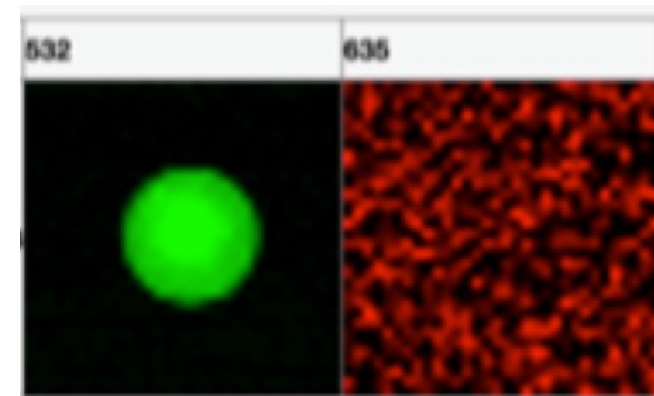
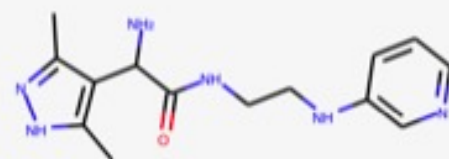
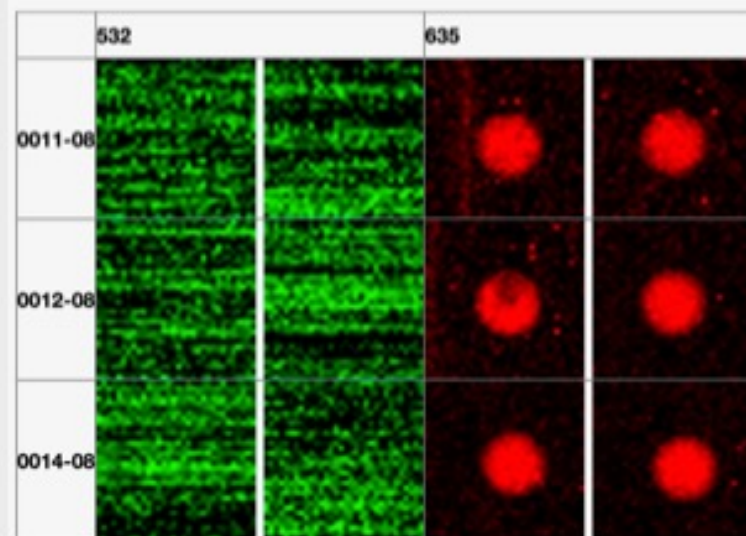
How will you determine a threshold Z-score?



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

How will you validate hits?

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



M2D5HW (Due 4/13)

- 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!!

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