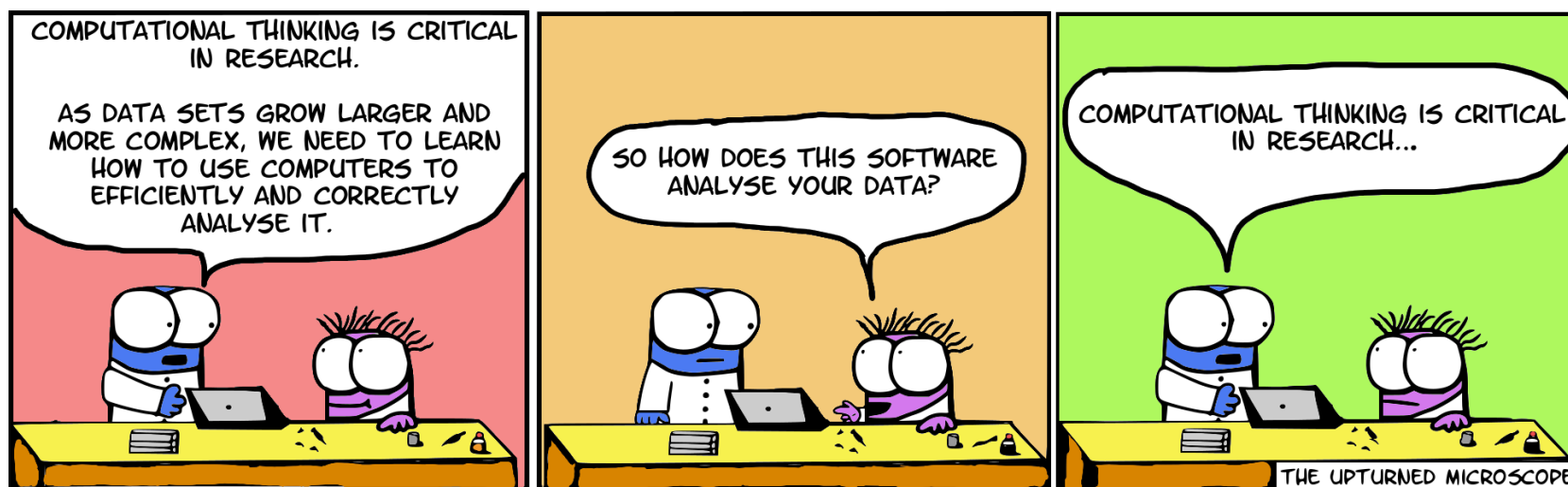


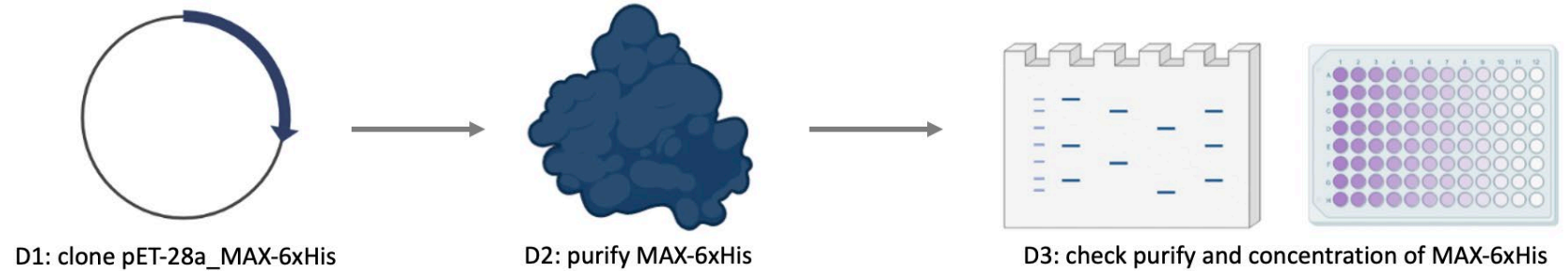
M1D7: Analyze SMM data to identify putative small molecule binders

1. Prelab
2. SMM analysis
3. Examine chemical structure of hits



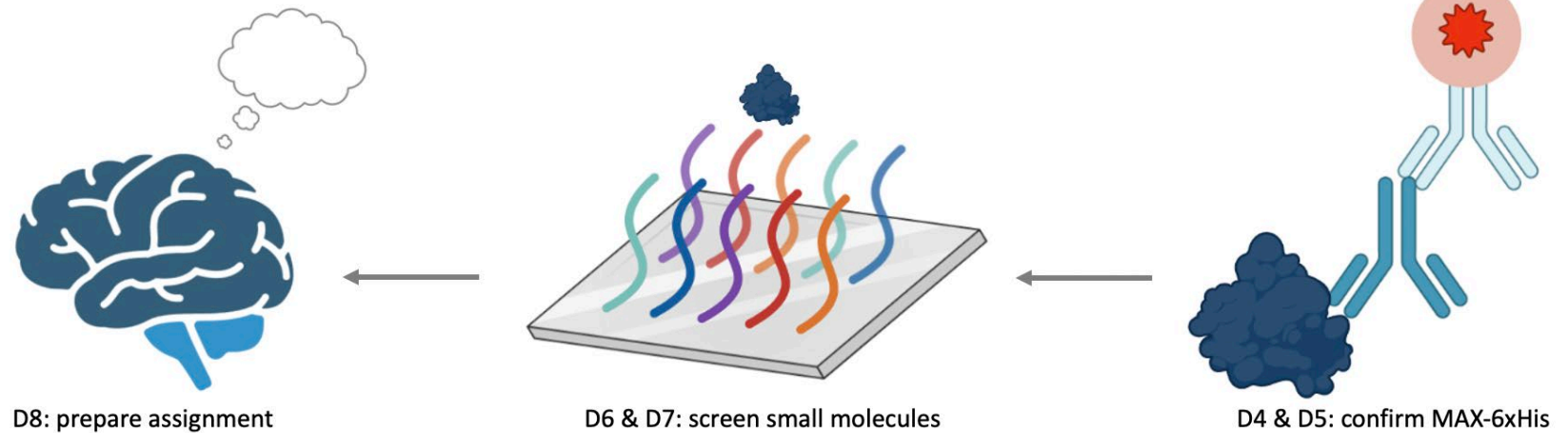
Mod 1 Overview

Last lab:
Prepared SMM slides



Instructors scanned slides and prepared computational analysis

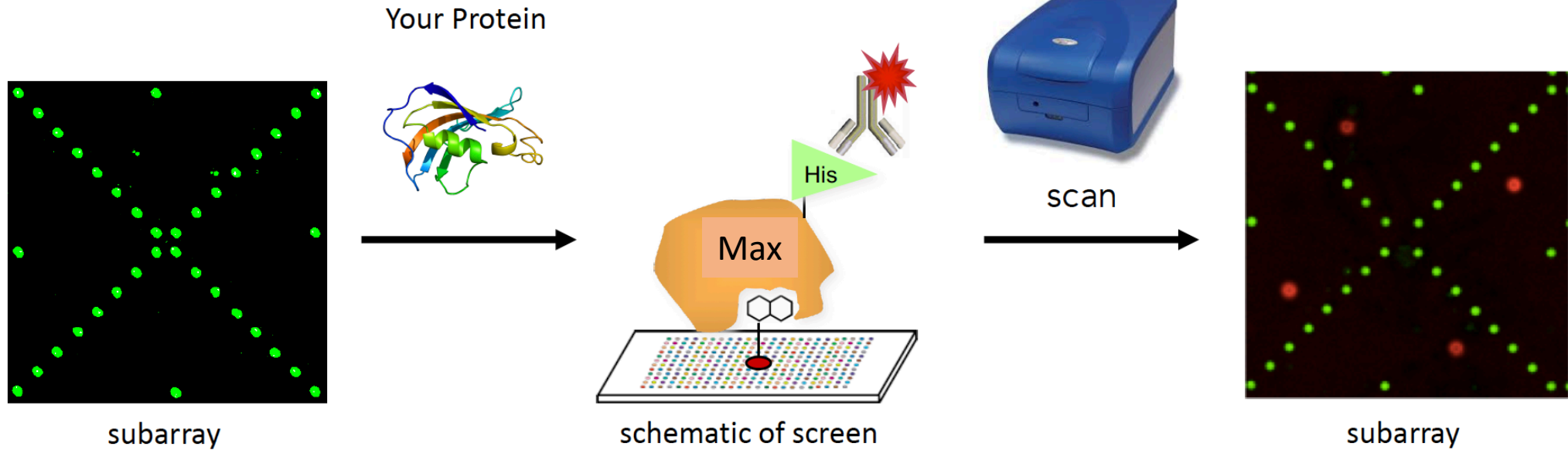
This lab:
SMM analysis to identify putative binders



SMM workflow

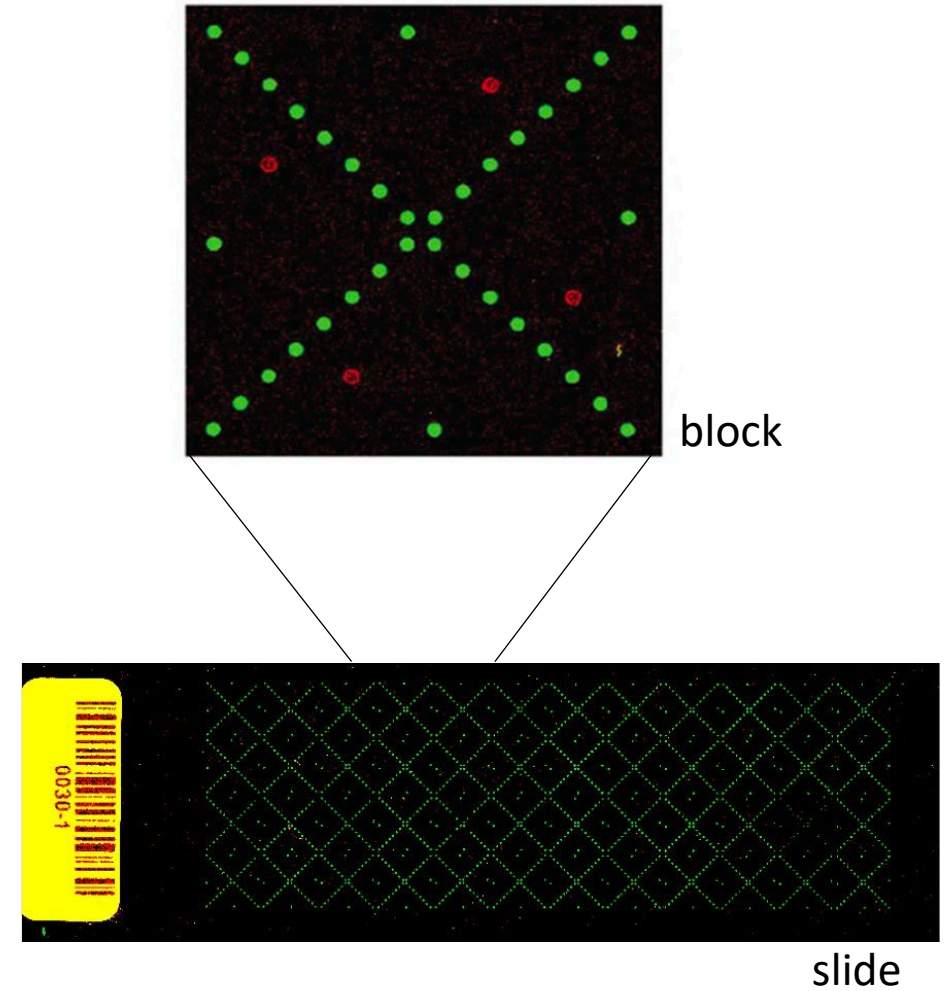
SMM Screen

Data Acquisition



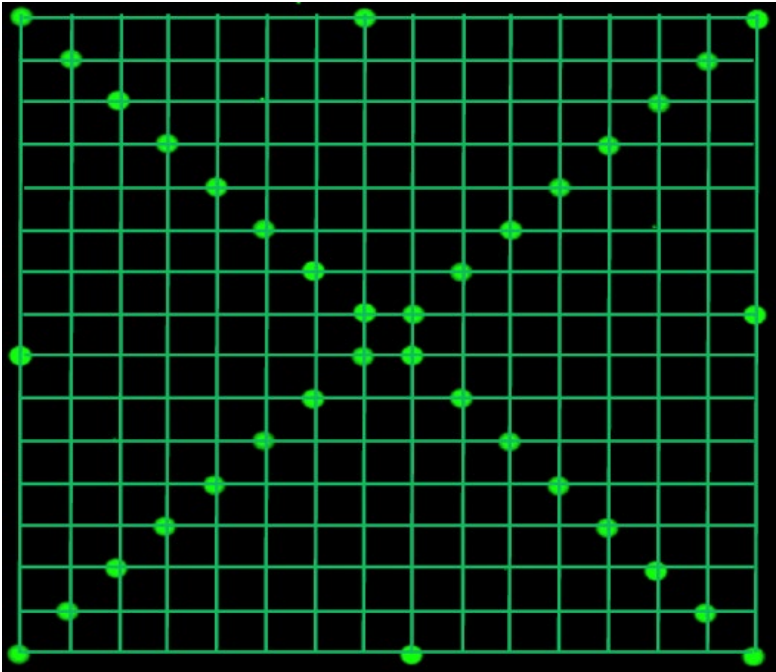
Workflow for SMM data analysis

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

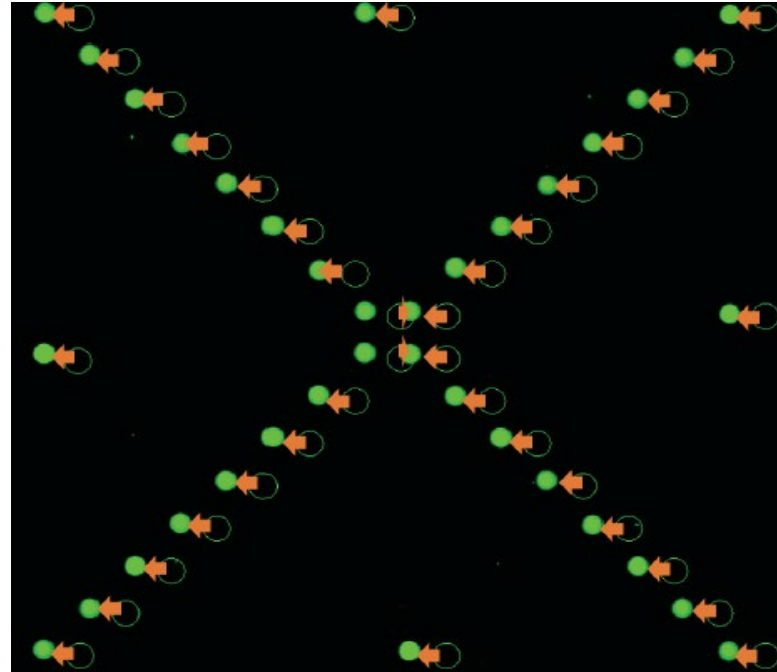


Align SMM using sentinel spots

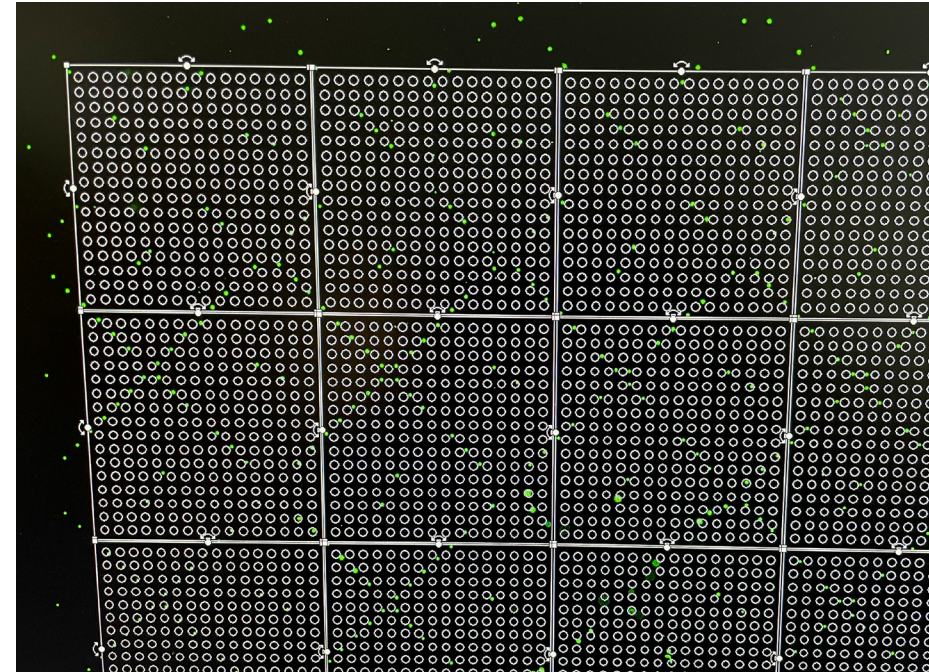
- Slides are printed in block patterns (16 rows x 16 columns)
- Each ligand spot is identifiable via intersecting lines between sentinels



Concept



Process

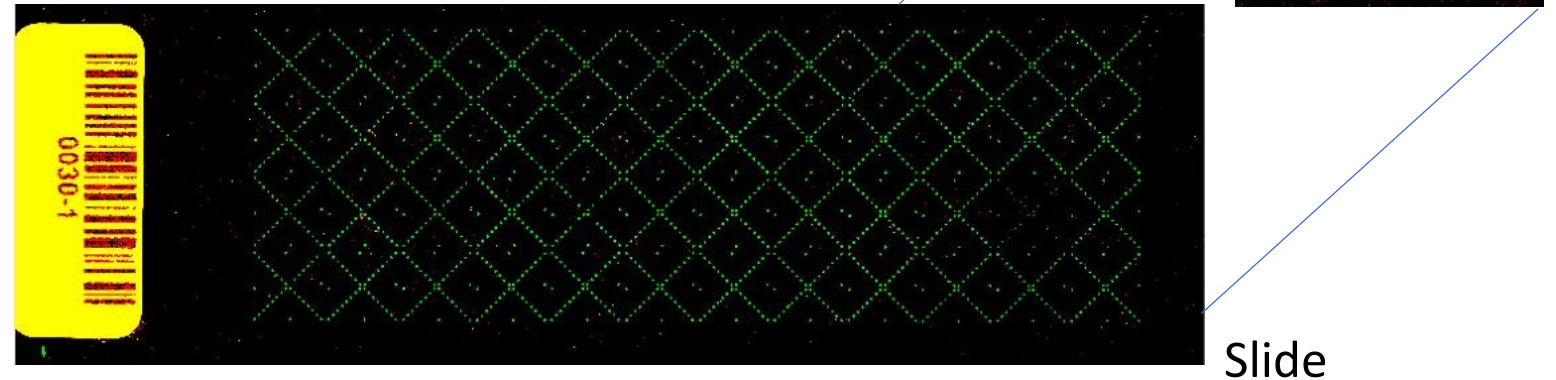


Real life

How do you identify hits from the SMM data?

First, consider bias that exists in the data set

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



Then, identify hits with significantly higher fluorescence over background

Lastly, manually confirm hits to eliminate false positives

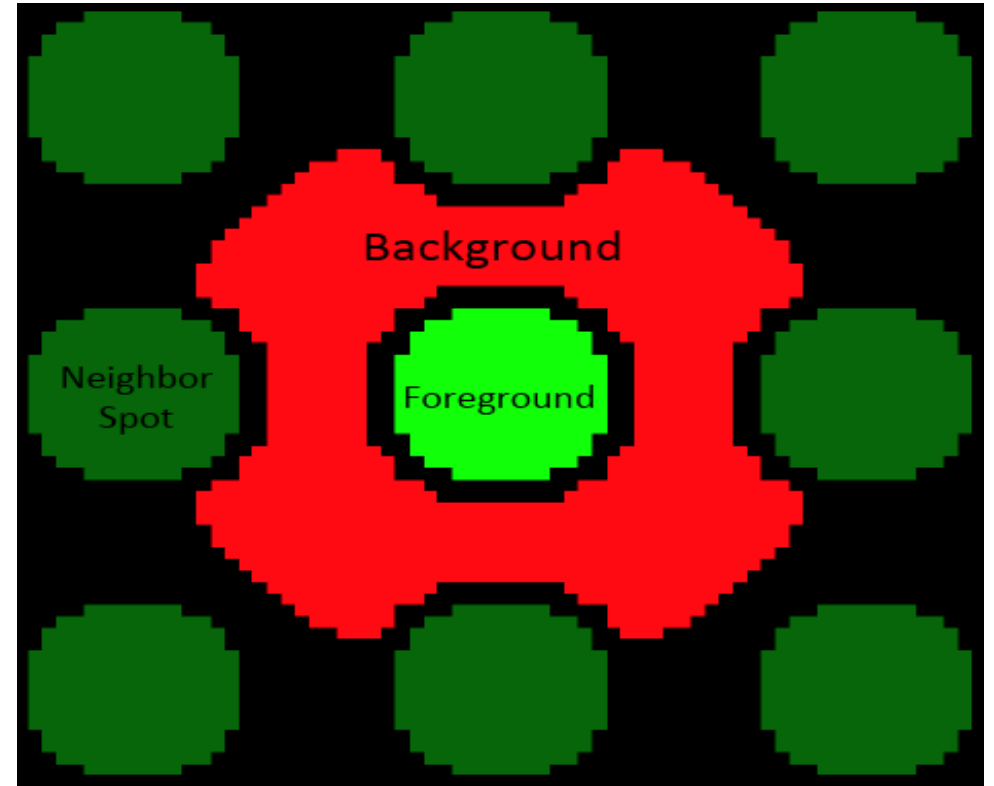
Fluorescence represented by an array of numerical values which are used to calculate signal to noise ratio

```

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 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

```

- Each pixel is represented by a number that indicates intensity of the signal



Signal-to-noise ratio

$$(SNR) = \frac{\mu_{\text{foreground}} - \mu_{\text{background}}}{\sigma_{\text{background}}}$$

Identifying hits with significant fluorescence

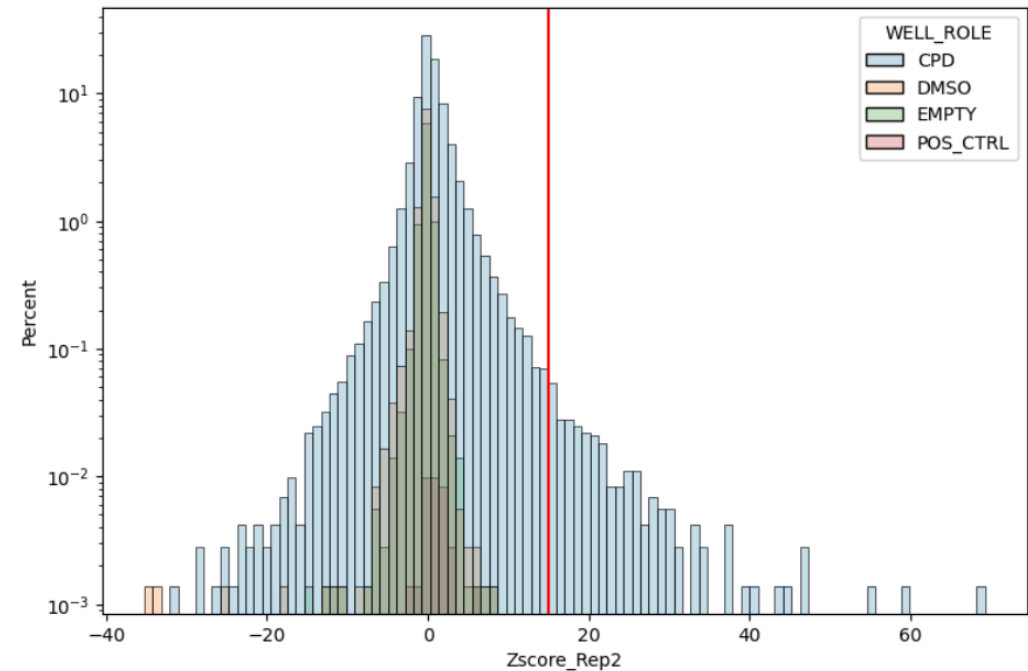
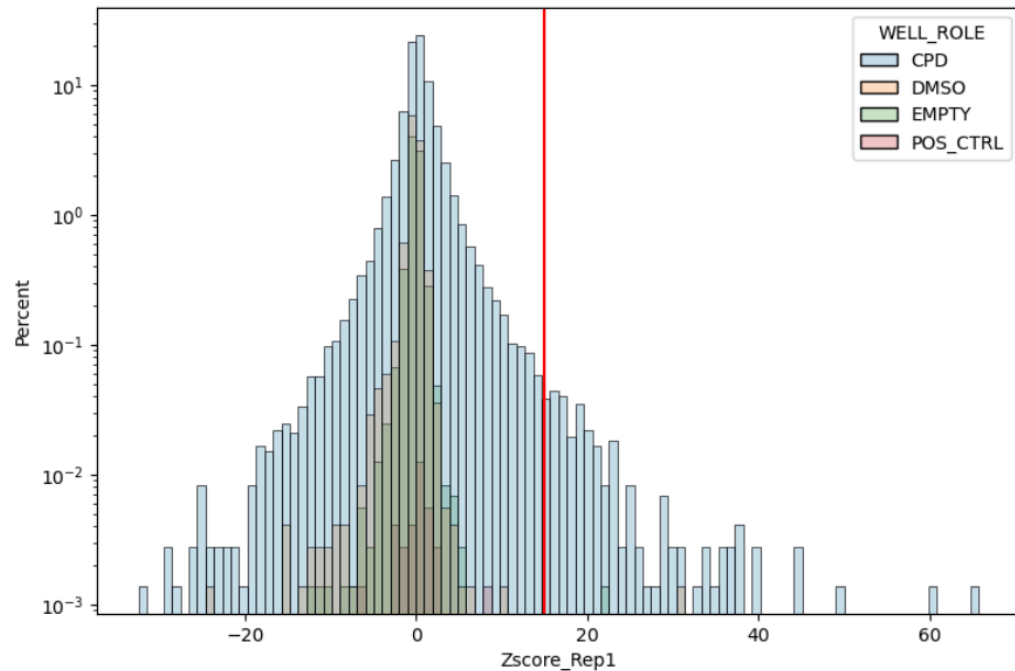
$$\text{Robust Z-score} = \frac{\text{SNR}_i - \text{median}(\text{SNR})}{\text{median}(|\text{SNR}_i - \text{median}(\text{SNR})|) * 1.48}$$

median absolute deviation (MAD)

scale factor for the normal distribution

Robust Z-scores help eliminate the influence of outliers

Robust Z-score calculated for all compounds

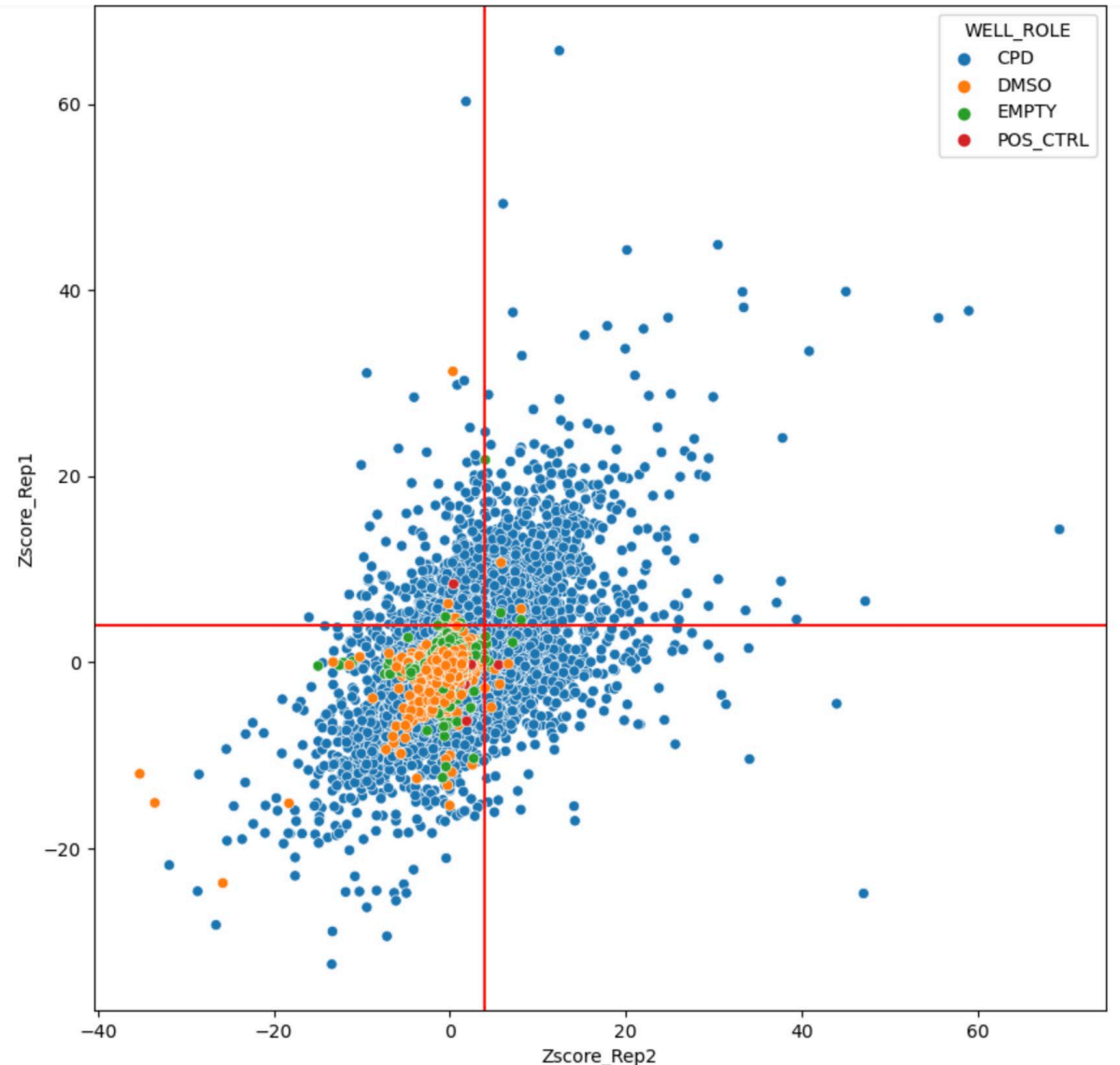


Histogram of robust z score of wells in both replicates

- Why are the empty wells clustered around zero?
- Where do we expect to see the putative binders represented?

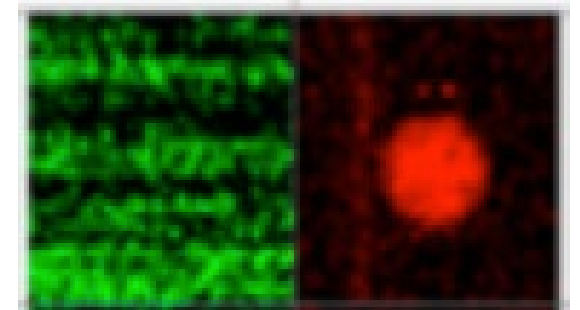
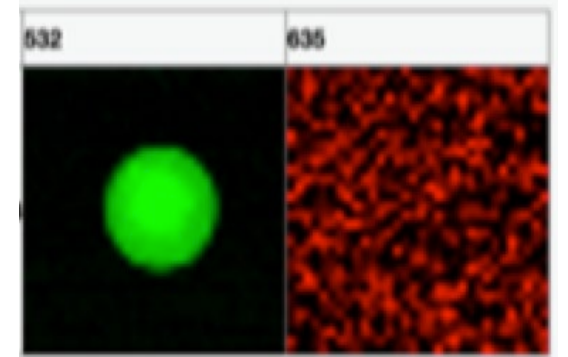
Use a scatterplot to compare consistency of replicates

- Expect to see a linear relationship between replicate z scores
- What does it mean if there are replicates that do not have a linear relationship?
- Where do we expect to see our putative binders in this graph?

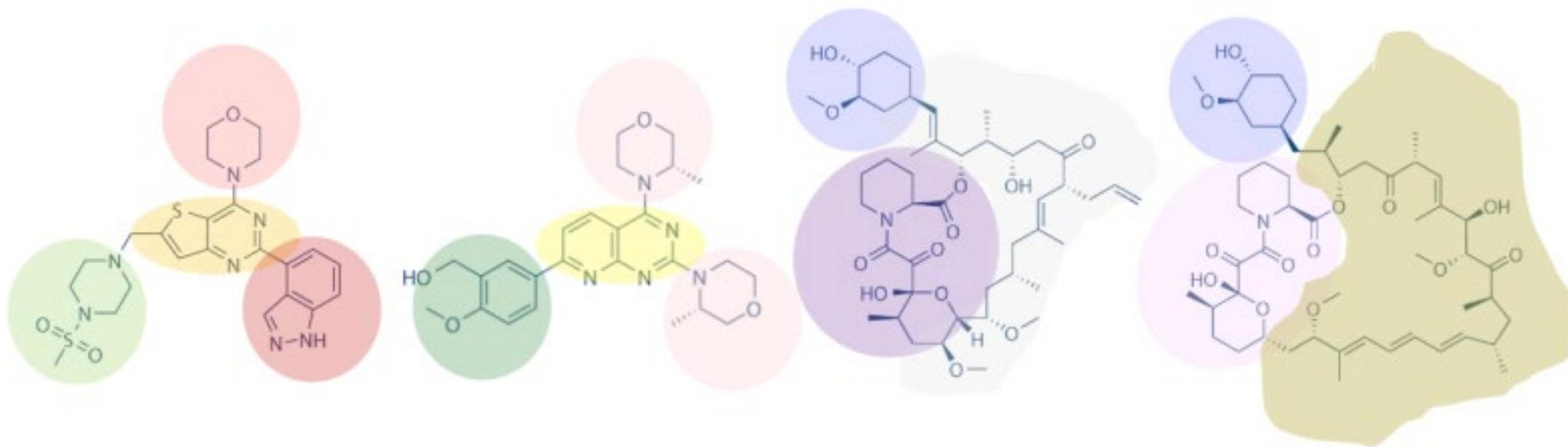


How do you validate hits manually?

Internal_ID	Zscore_Rep1	Zscore_Rep2	Slide_Rep1	Slide_Rep2	Loc1(block, row, col)	Loc2(block, row, col)
KI10167	37.029230	24.797054	50033718.0	50033720.0	2,2,3	2,7,11
KI10451	38.126523	33.380424	50033718.0	50033720.0	42,12,8	42,15,14
KI10796	16.684962	18.340951	50033718.0	50033720.0	45,10,16	45,12,4
KI10776	19.640085	18.046225	50033718.0	50033720.0	41,1,6	41,8,10
KI11103	16.929258	19.408047	50033718.0	50033720.0	42,1,12	42,6,8
KI11145	18.763623	15.206143	50033718.0	50033720.0	8,10,14	8,14,16
KI12064	25.653114	15.651485	50033718.0	50033720.0	41,11,4	41,15,5
KI20071	22.544108	24.043693	50033689.0	50033693.0	26,12,11	26,5,4
KI20165	25.236351	23.586604	50033689.0	50033693.0	4,2,3	4,7,11
KI20173	36.988660	55.511073	50033689.0	50033693.0	4,16,2	4,6,9



How will you identify common structures?



For Today

- Work through SMM analysis procedure
- Evaluate chemical structures of identified hits
- Work on homework and begin thinking about Data Summary

For M1D8

- Answer the wiki questions in the Homework section to begin work on the Data Summary Implications and Future Works section
- With your Lab Partner, incorporate the feedback on your methods homework and revise to include experiments from M1D4-M1D5