An automated method identifies dose-responsive genes and quantifies Points Of Departure

David Gerhold, Pei-Hsuan Chu, John Braisted, Yuhong Wang, and Ruili Huang

NIH-NCATS

October 23, 2017

Principles for BMD/POD

- A consensus on BMD/POD method and pathways decisions will facilitate cooperation among Tox21 members, and consistent risk assessment.
- Public BMD Express 2.0 software and visualization tools a winner! Suggest changes to algorithm for identifying “significant genes”.
- Minimize false positives. 21,000 genes >>multiplicity problem!
- The simplest model (most constrained) applicable to transcriptional regulation will minimize overfitting, so minimize false positives.
- Lowest dose is most difficult (no info from lower doses) and most dangerous (worst dose to accept false positives).

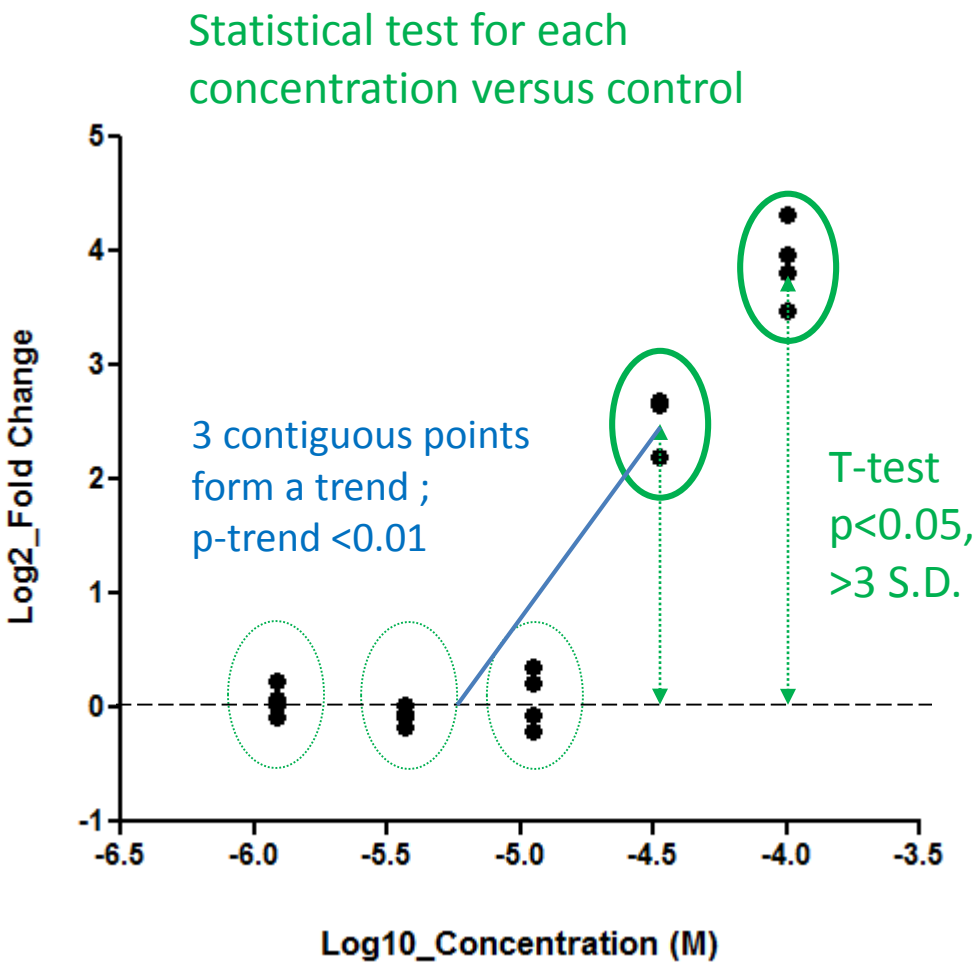
How did we optimize this algorithm without "Truth"?

- Began with BMD Express. Excellent curve fits, but difficult decisions for "significant genes"
- Ruili Huang strategy, Yuhong Wang Hill fit & programming
- Tested on in-house 5 or 6-dose x 4 reps gene expression data, then 6-dose x 3 reps RNAseq data
- In-house data used 3 probe-sets per gene, so adjust for parameters that maximized agreement. Approximation to "Truth"
- For biphasic dose responses, event at lower dose is more important, if significant
- No strong opinion about cutoff for BMD/POD 1.49 S.D? 3 S.D.? 10%? We run 12-20 vehicle controls/microplate, so 1.49 S.D. is often tiny.
- We examined concordance among 3 runs of HepaRG data chemical #1 results. Concordance was not strong by either analysis algorithm, so we focused on run1 data.

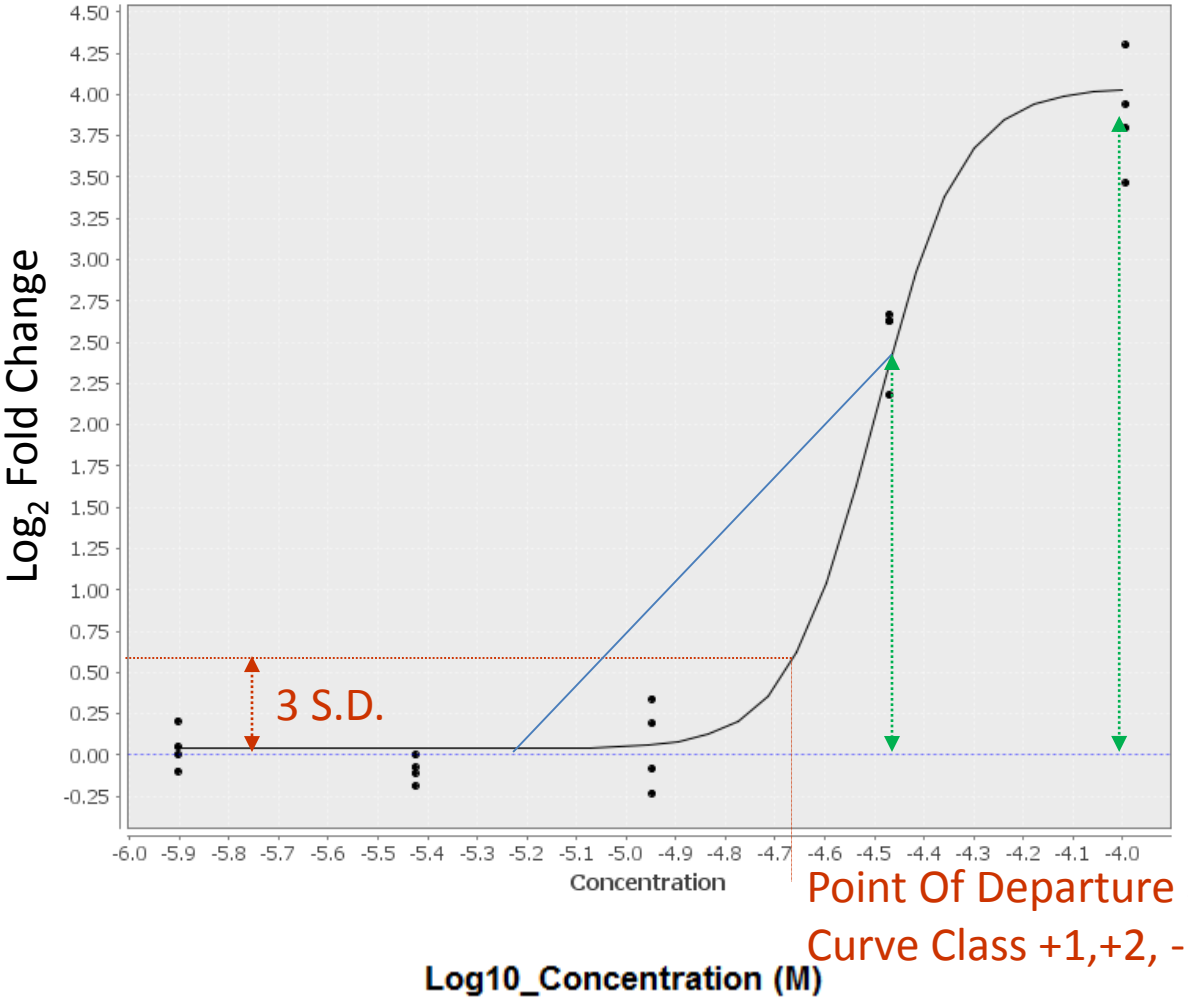
NCATS Point-of-Departure (POD) method

1. For each treatment, genes with low counts are excluded.
2. For each gene x treatment, the Log_{10} dose vs Log_2 ratios are plotted, and fit to the Hill equation (Matrix-fit, 2 iterations, Wang et al. 2010). The Hill equation is a single simple model that avoids illogical curves, and fits nearly all of our known mechanisms for transcriptional regulation.
3. At each treatment concentration, a T-test is applied to determine significant changes from the vehicle controls.
4. Each set of 3 consecutive concentrations are tested for a statistically significant trend upward or downward.
5. Each significant gene x treatment assigned a curve class up: +1, +2, or down: -1, -2. We disregard +0.5, and -0.5.
6. Each significant gene x treatment POD is defined as the lowest concentration where curve deviates significantly (3 SD, $p<0.05$) from controls, minimum change =1.5-fold (or # SDs?). These parameters are flexible to adjust for variance, # reps, # genes).

Test for significant dose-response determination of Point of Departure (POD) and BenchMark Dose (BMD)



Point of Departure (POD) calculation yields the minimum concentration of chemical that causes each gene to change expression. Robust because an additional statistical test requires a significant dose-response trend.



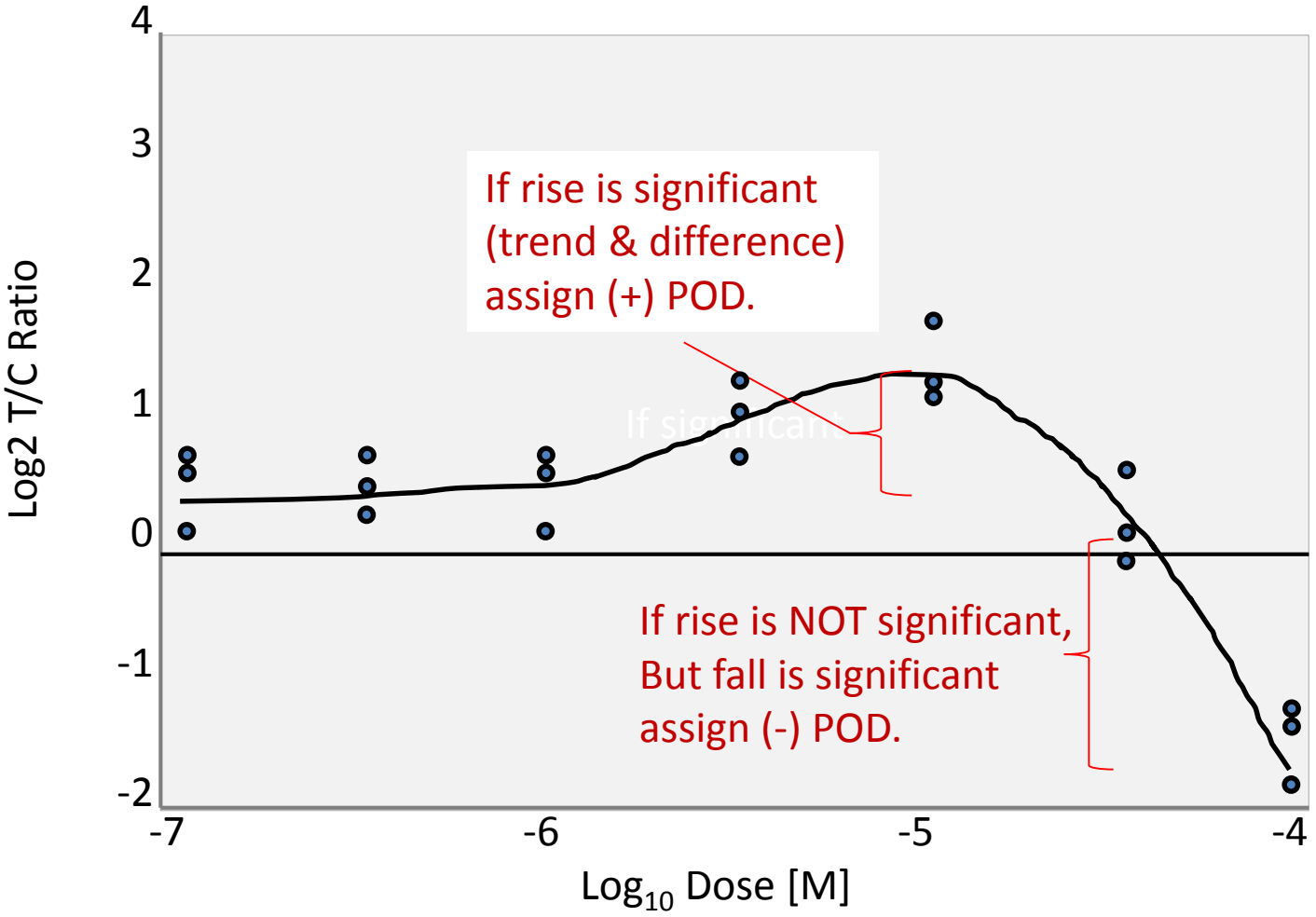
Curve Fit to
Hill equation

1 or 2 points significantly above/below control p<0.05 & >3 S.D., and agreement with p-trend (Curve Class 1 or 2, +/-)

3 contiguous points form a trend ; p-trend <0.01

Point Of Departure and
Curve Class +1,+2, -1, -2

Challenge: Biphasic Response

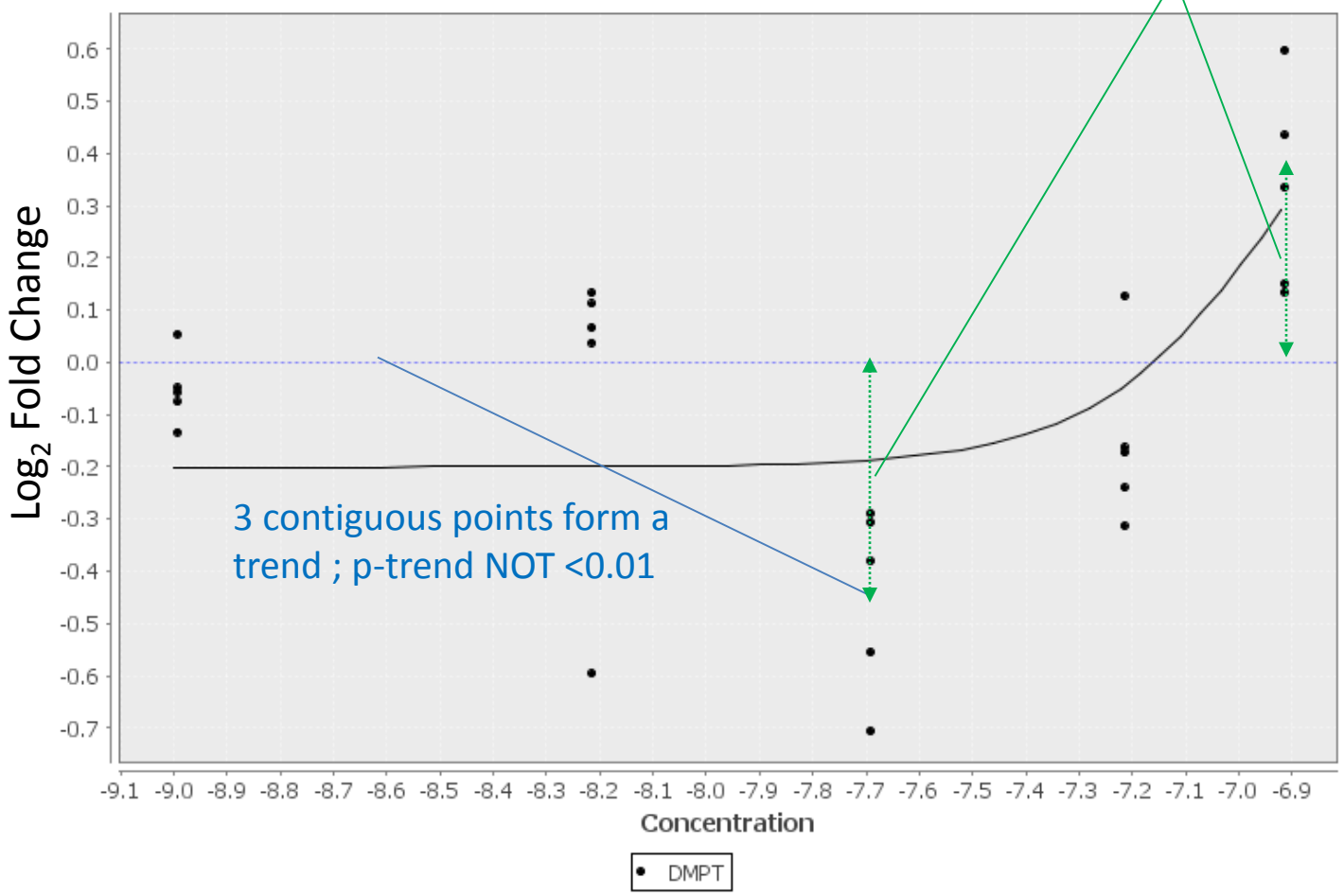


Challenge: Biphasic Response

DMPT_1382511: Curve Class 0, Y at POD_{3SD}=1.2, Control S.D.=0.4

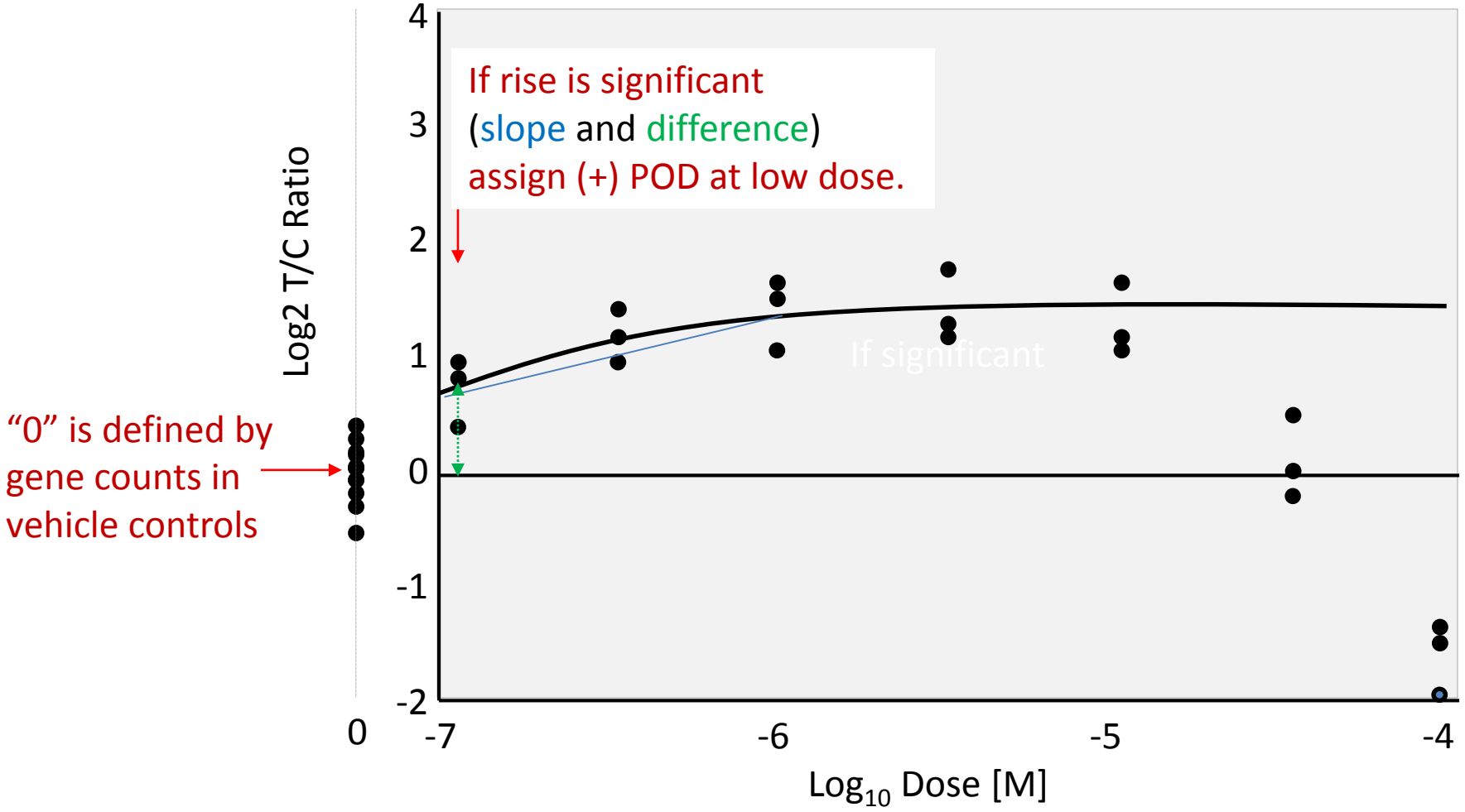
NO significant Point Of Departure

NOT significantly above/below control p<0.05 & >3 S.D.



3 contiguous points form a trend ; p-trend NOT <0.01

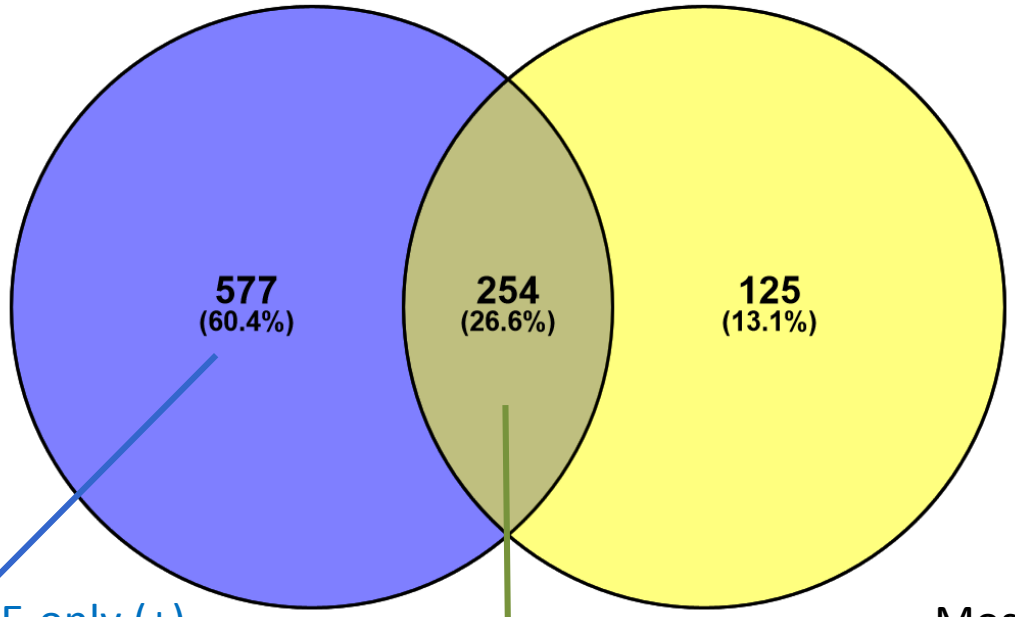
Challenge: Non-zero at Lowest Dose



Comparing BMDE2 and POD Algorithms
HepaRG cpd1 run3

BMDExpress
831 total (+)

NCATS POD
379 total (+), (~1555 including +/-0.5 curve classes)



Curve Class	Count	
-2	63	
-1	71	CC 1/2
-0.5	679	379
0	1422	
0.5	497	CC 1/2/.5
1	133	1555
2	112	
	2977	

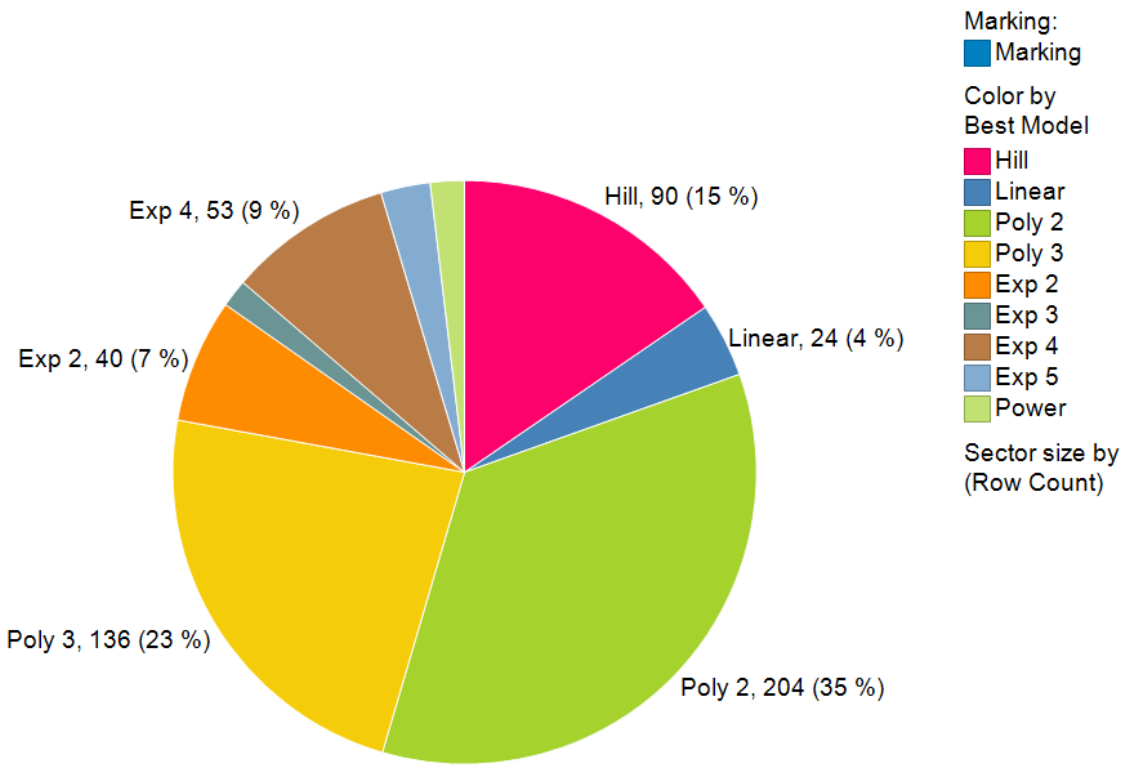
583 BMDE-only (+)
Most are stringency, judgement calls.
A few dozen are questionable polynomial fits

Most agree quantitatively, too.
23 are up/down divergent.

Most are slightly < 2-fold.

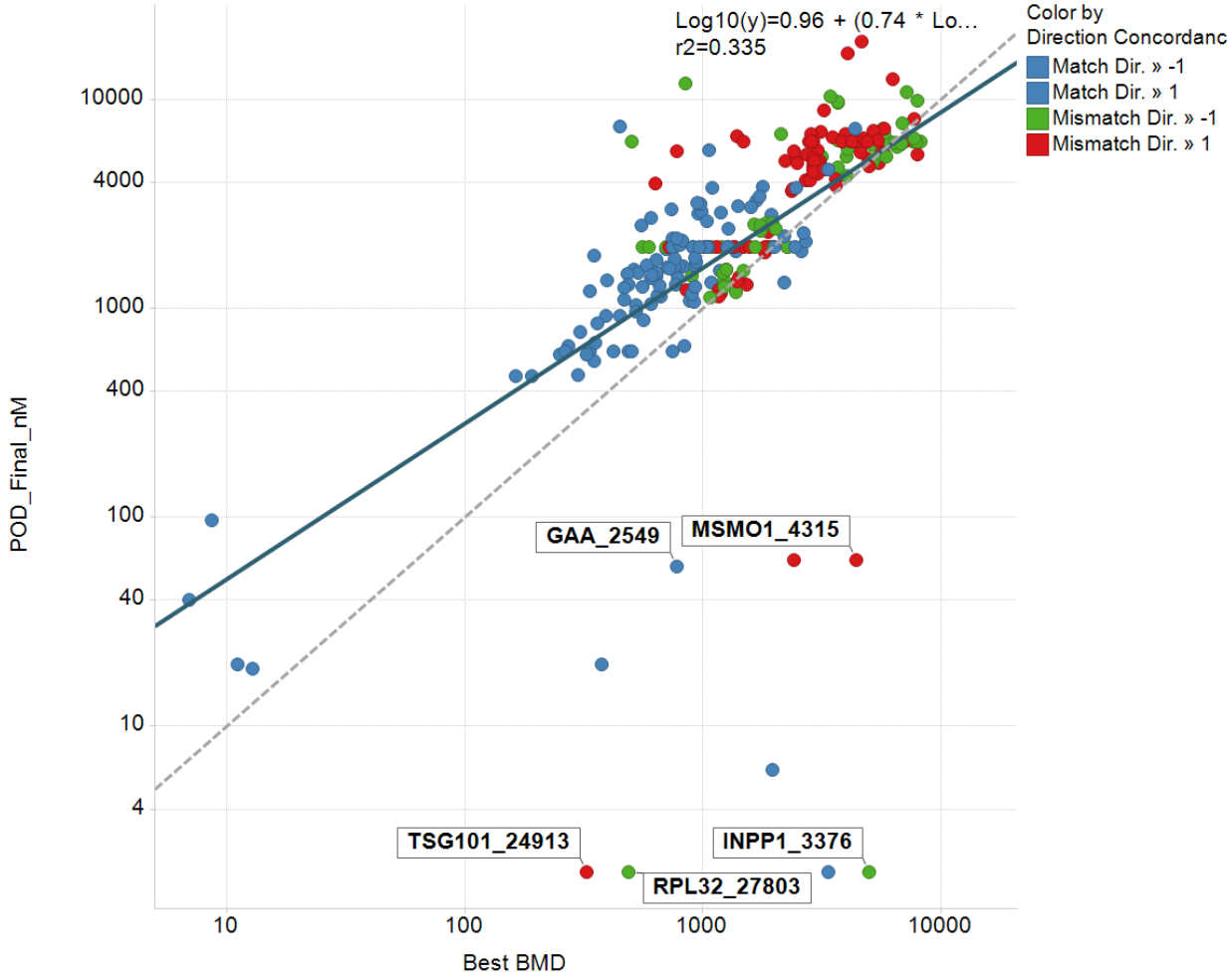
NIH National Center for Advancing Translational Sciences

Best Fit Type for BMDExpress 583 Unique Hits HepaRG Cpd #1

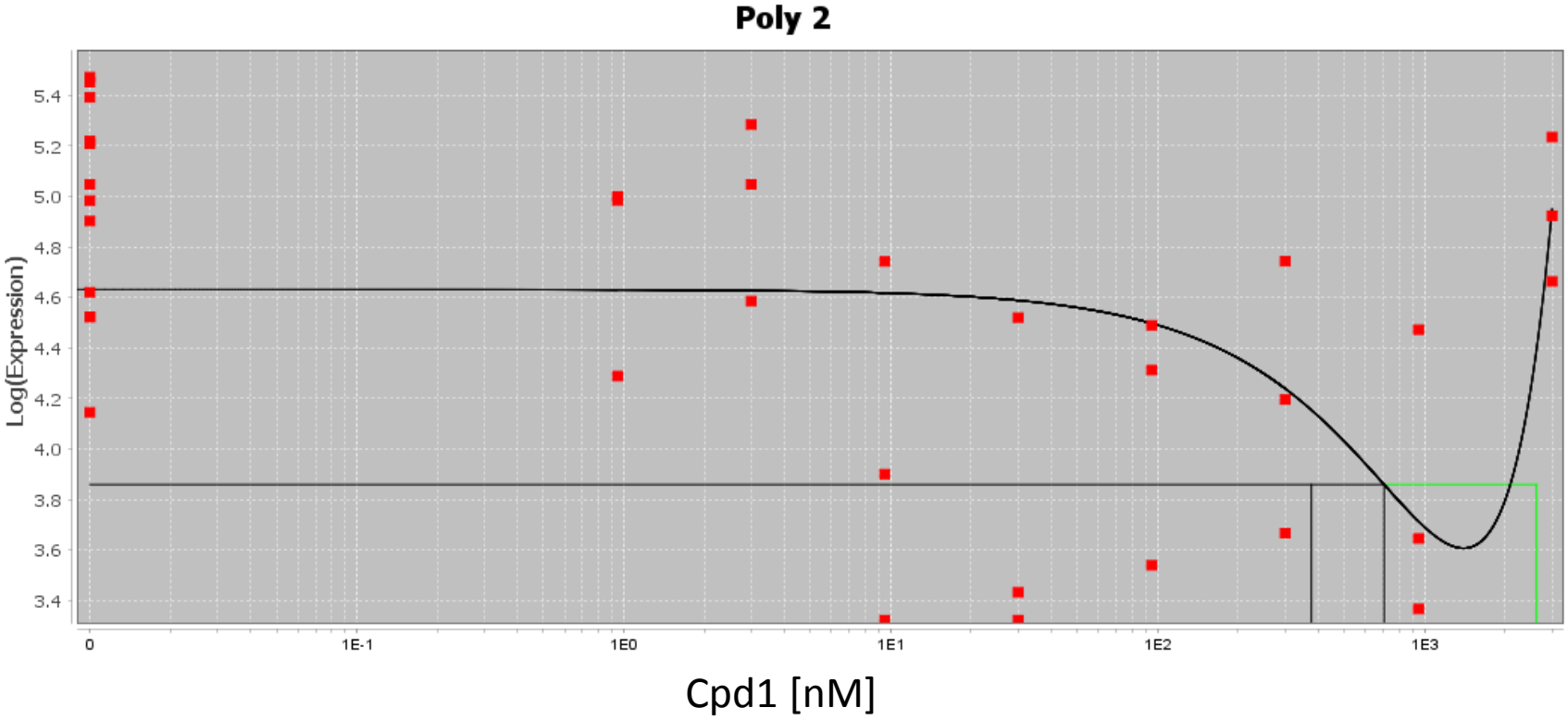


Currently, most are polynomial fits

Quantitative agreement among BMD Express and POD calls



BMD Express 2.0 format notes..
*SLC27A3_6427*_HepaRGcpd1run3

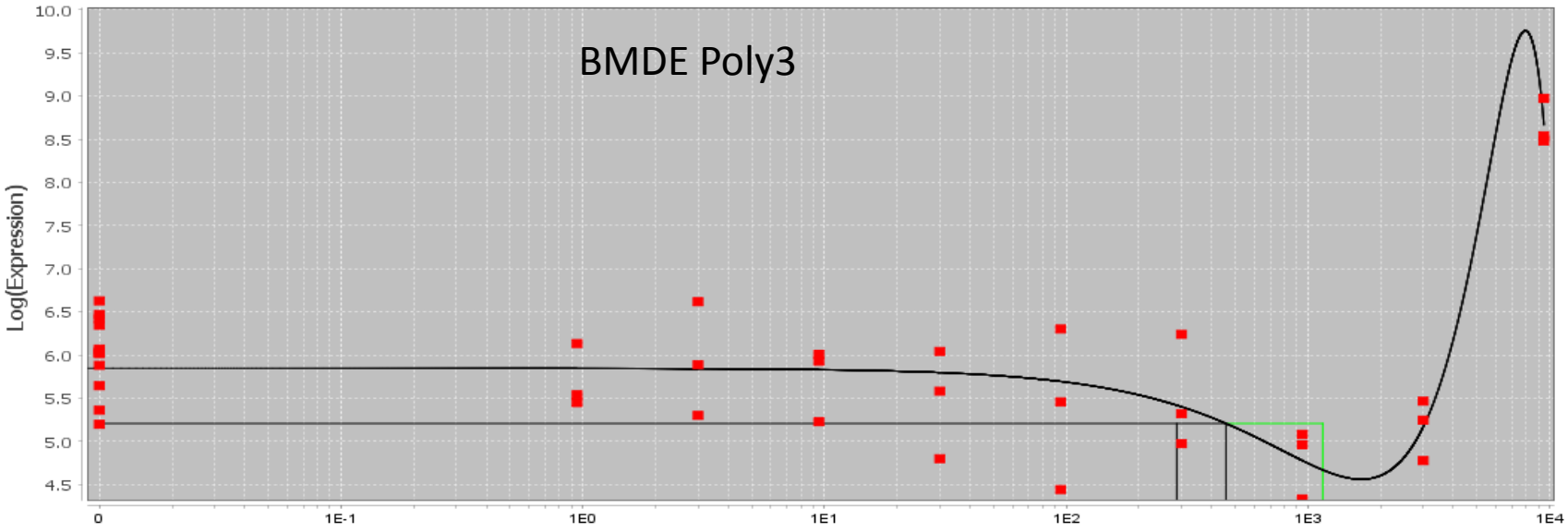


X axis Log_{10} nanoM to accommodate the range.

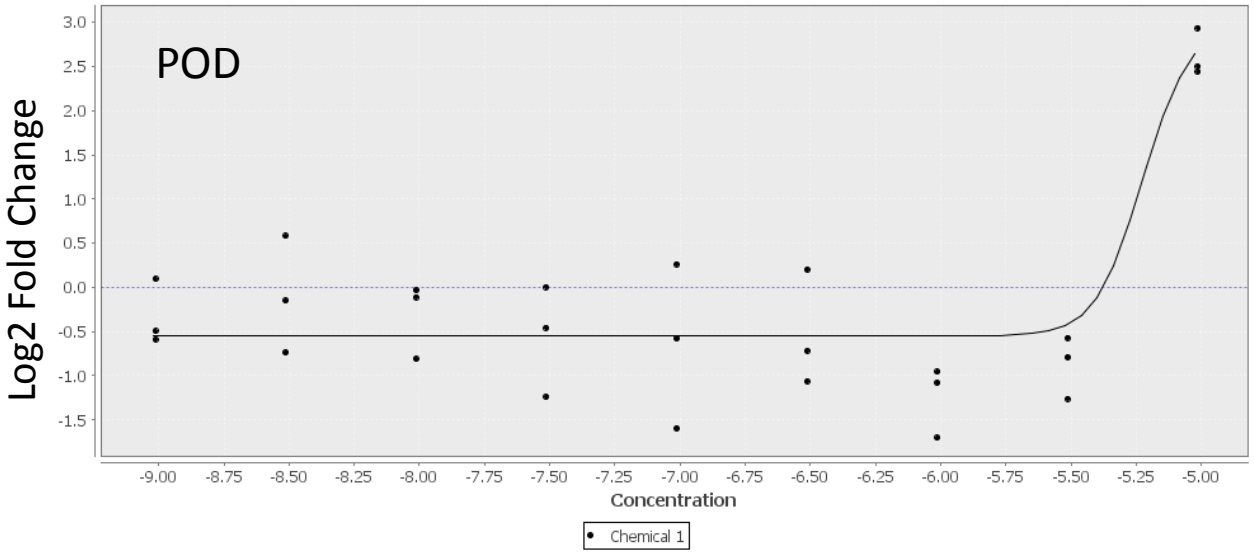
Zero controls are included in BMD fit. Curve fit is affected by scale (nM here), as the zero moves depending on scale. Would curve and BMD change depending on range of concentrations tested?

Does signal exceed noise in this case??

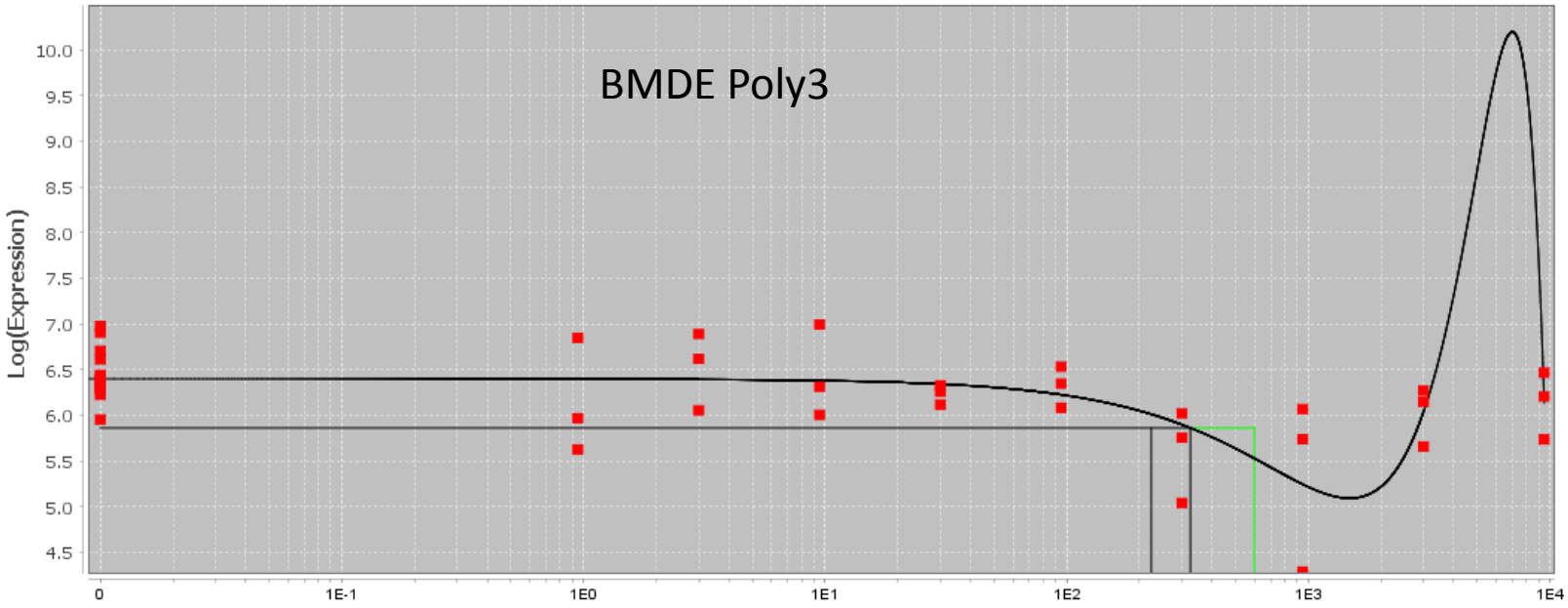
*IER3*_HepaRGcpd1run1



Decrease
or increase?
POD curve class (+1)



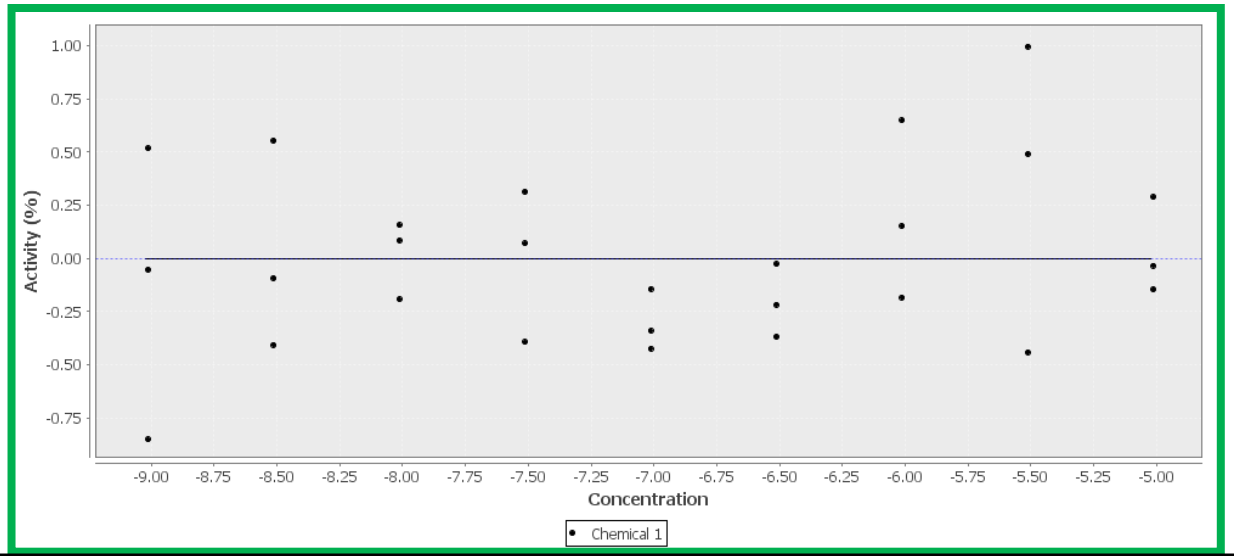
KRT19_10816_HepaRGcpd1run1



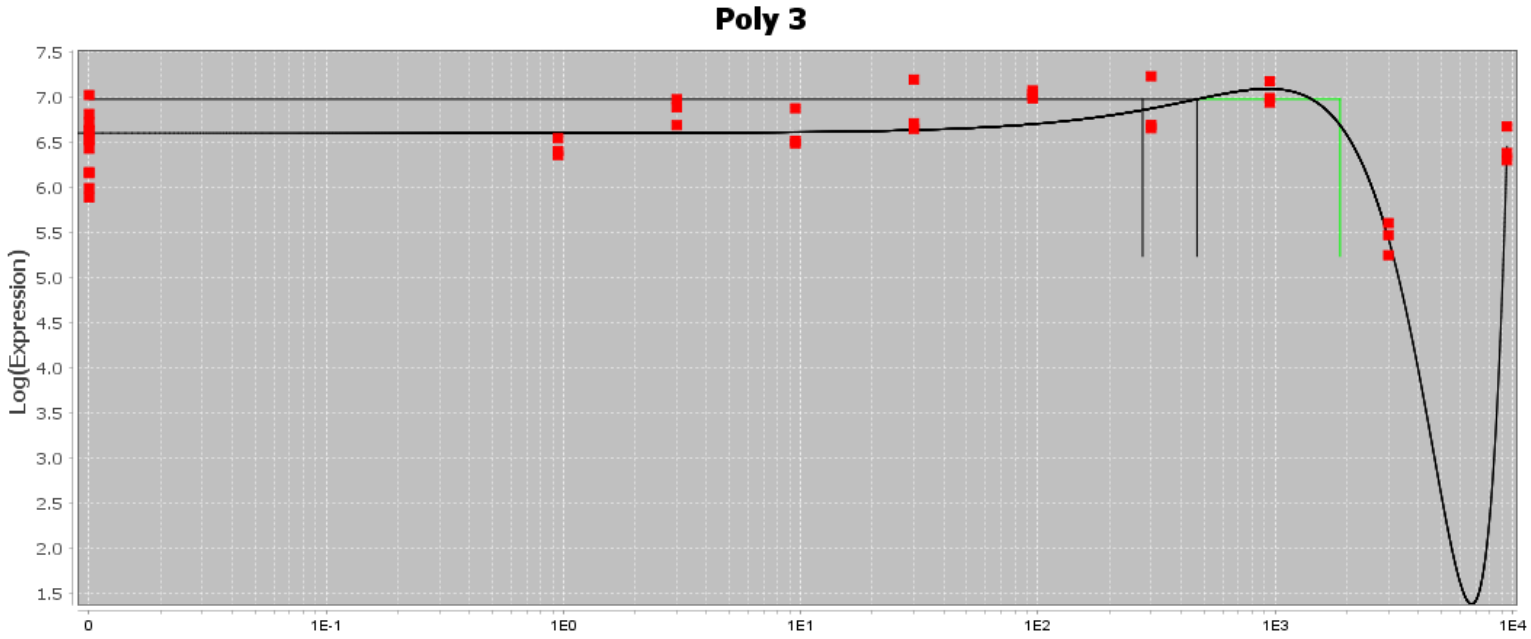
Decrease
or no POD?
(-0.5)

Probe KRT19_3890:
No call BMDE, nor
POD

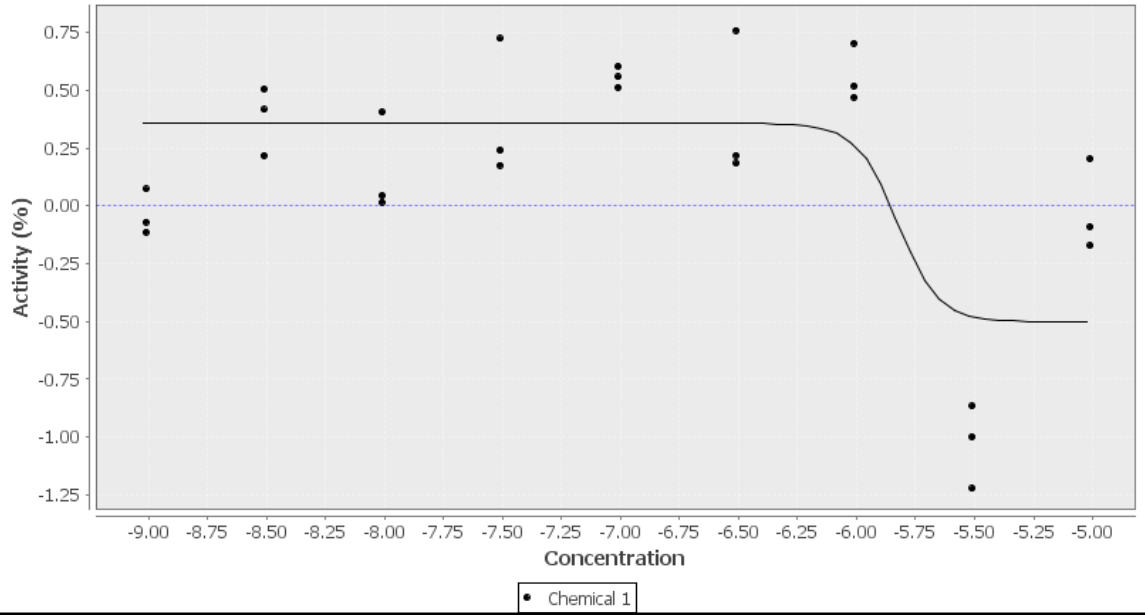
Log2 Fold Change



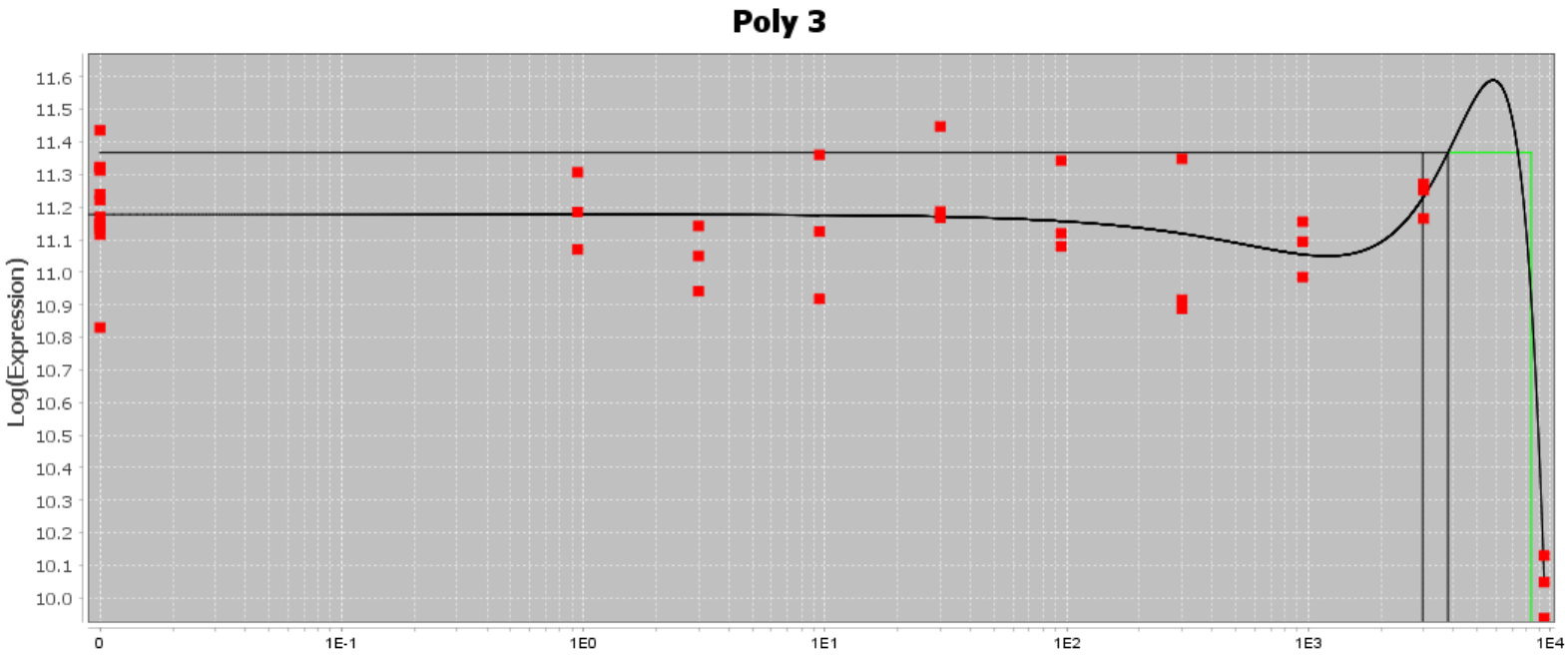
CISD1_HepaRGcpd1run1



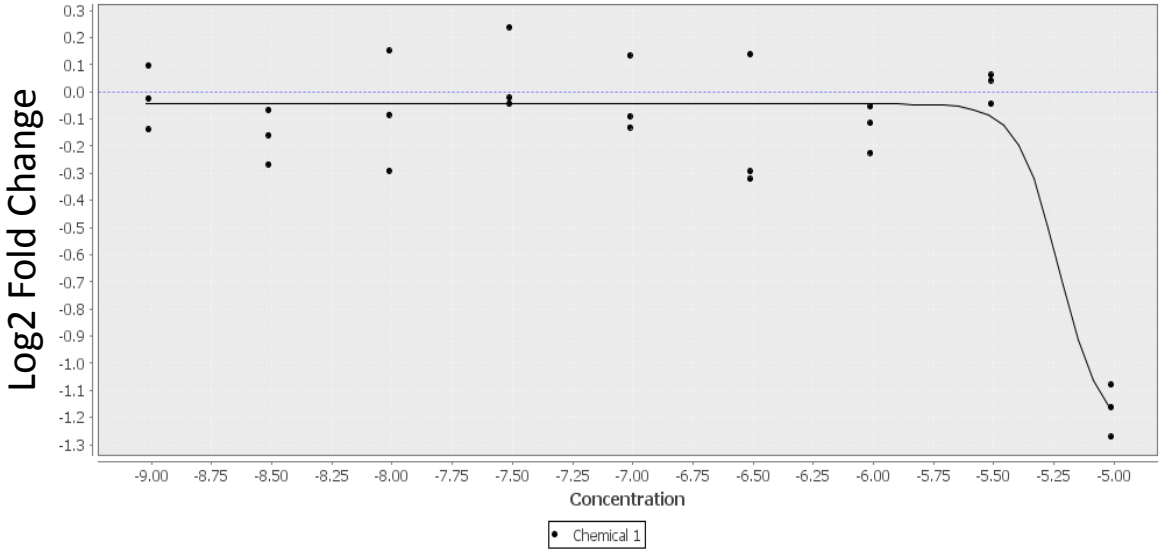
Increase
or no POD?
(-0.5)



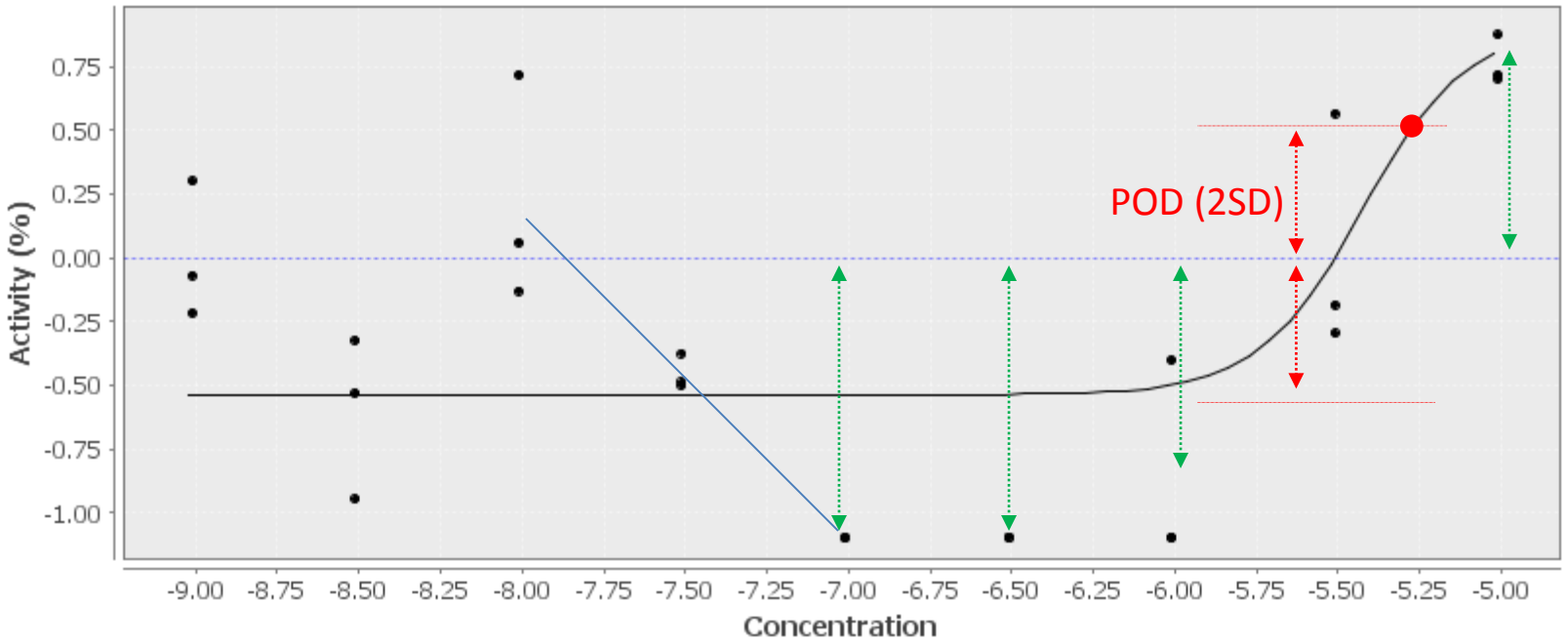
CLIC4_HepaRGcpd1run1



Increase? Or decrease?



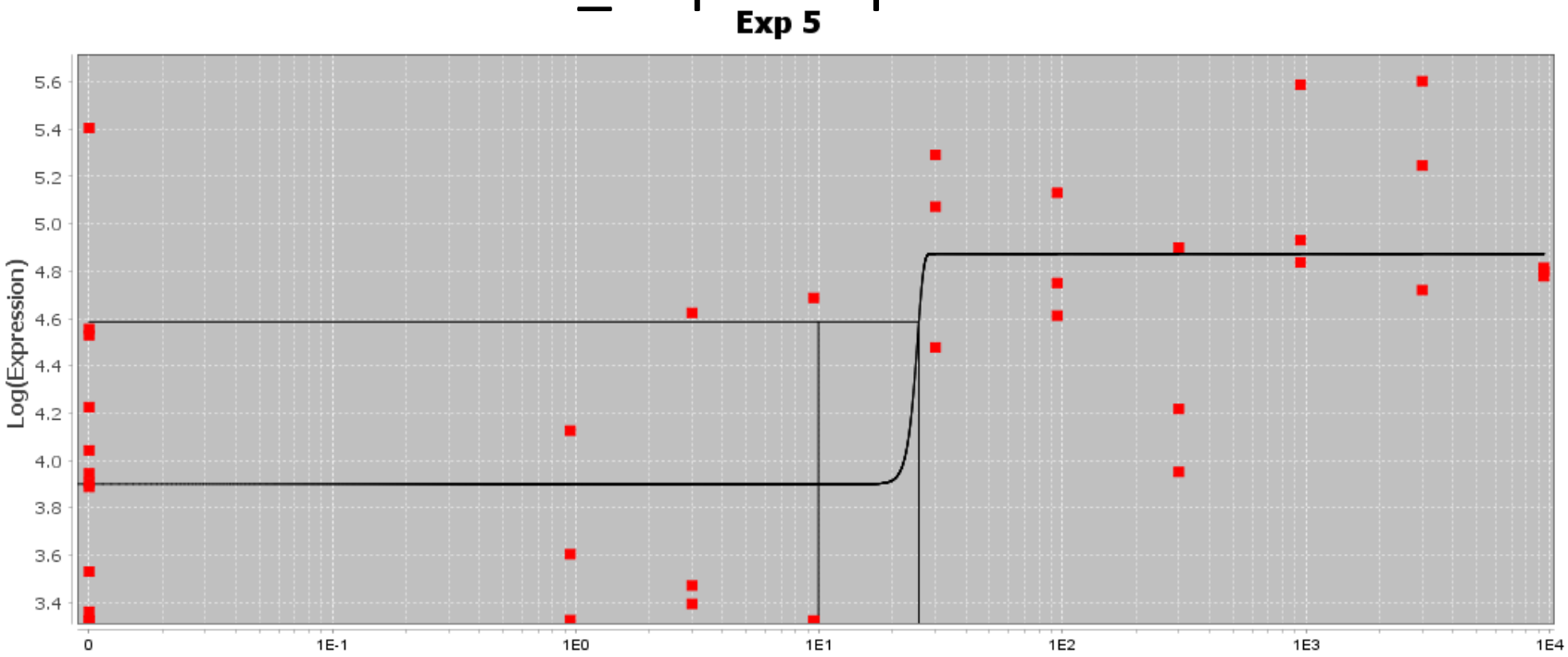
HAVCR1_23_HepaRGcpd1run1



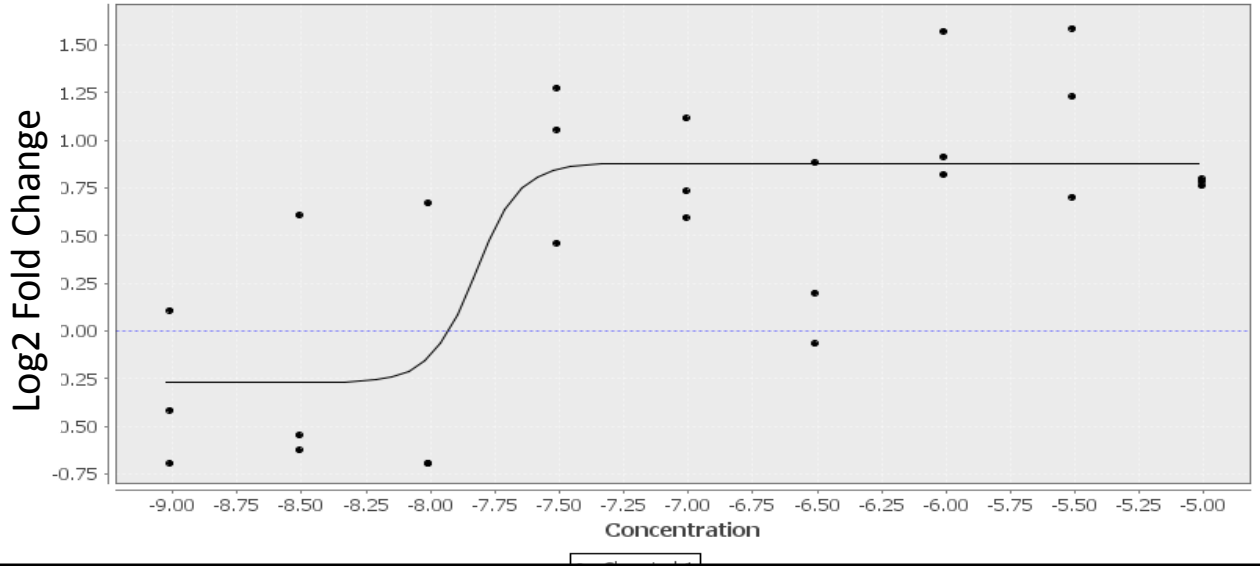
An issue remains to be corrected with NCATS' POD software.
Curve class -2, but POD is called at 5.7 microM.

Since fit line never crosses 2SD, need to discard the -2 curve class, and retest + slope.
Confirms POD at 5.7 microM, but should have curve class +1.

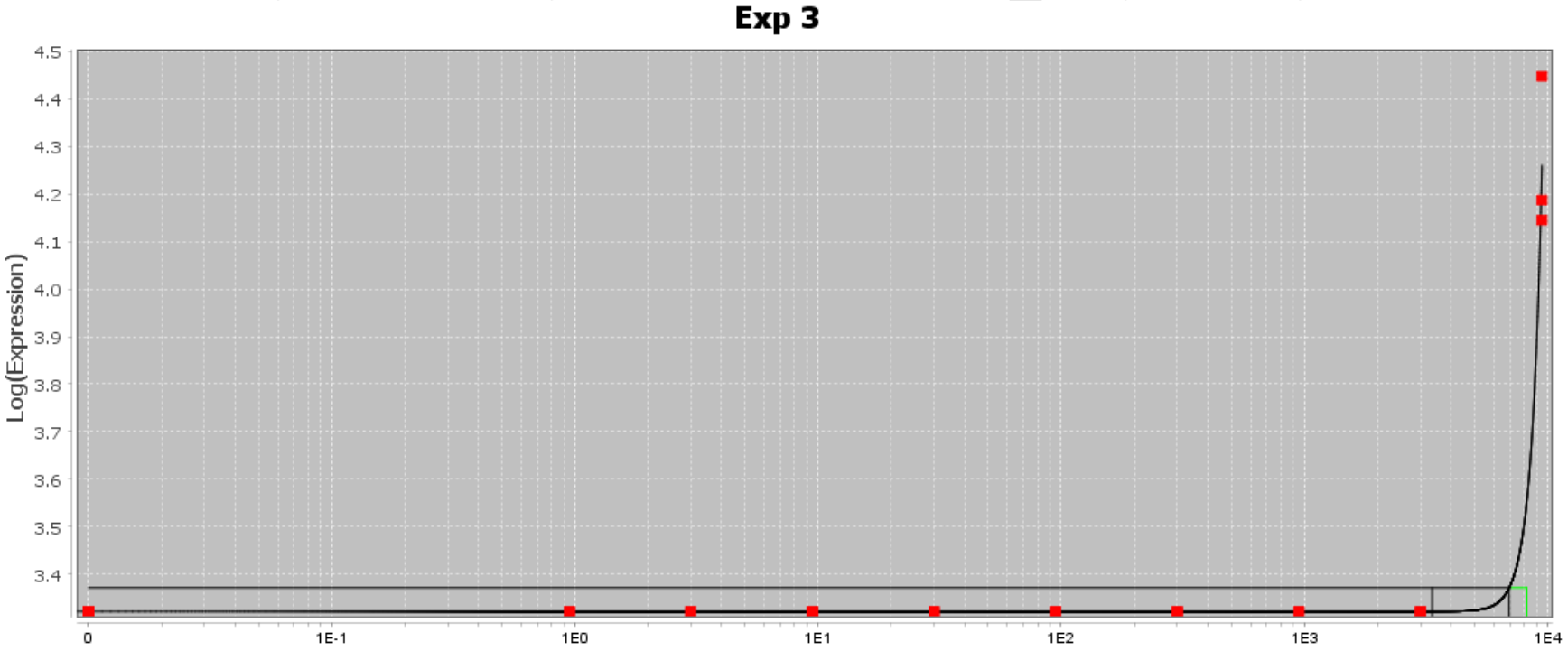
RFC5_HepaRGcpd1run1



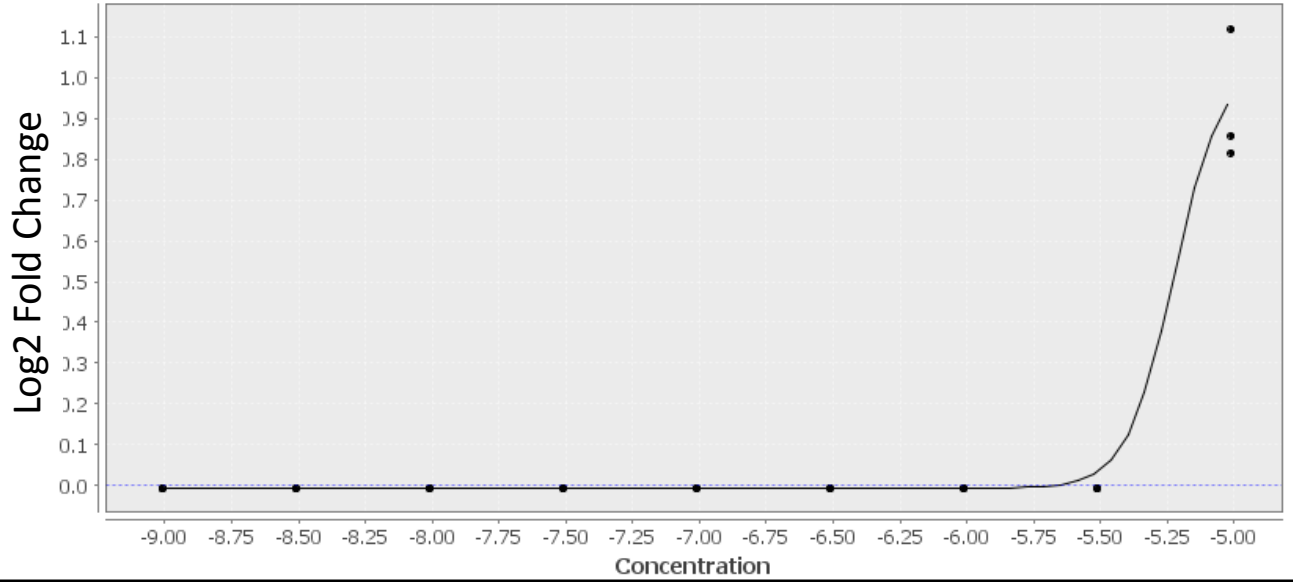
Increase?
or no POD?
(0) fails POD
trend test.
Noisy.



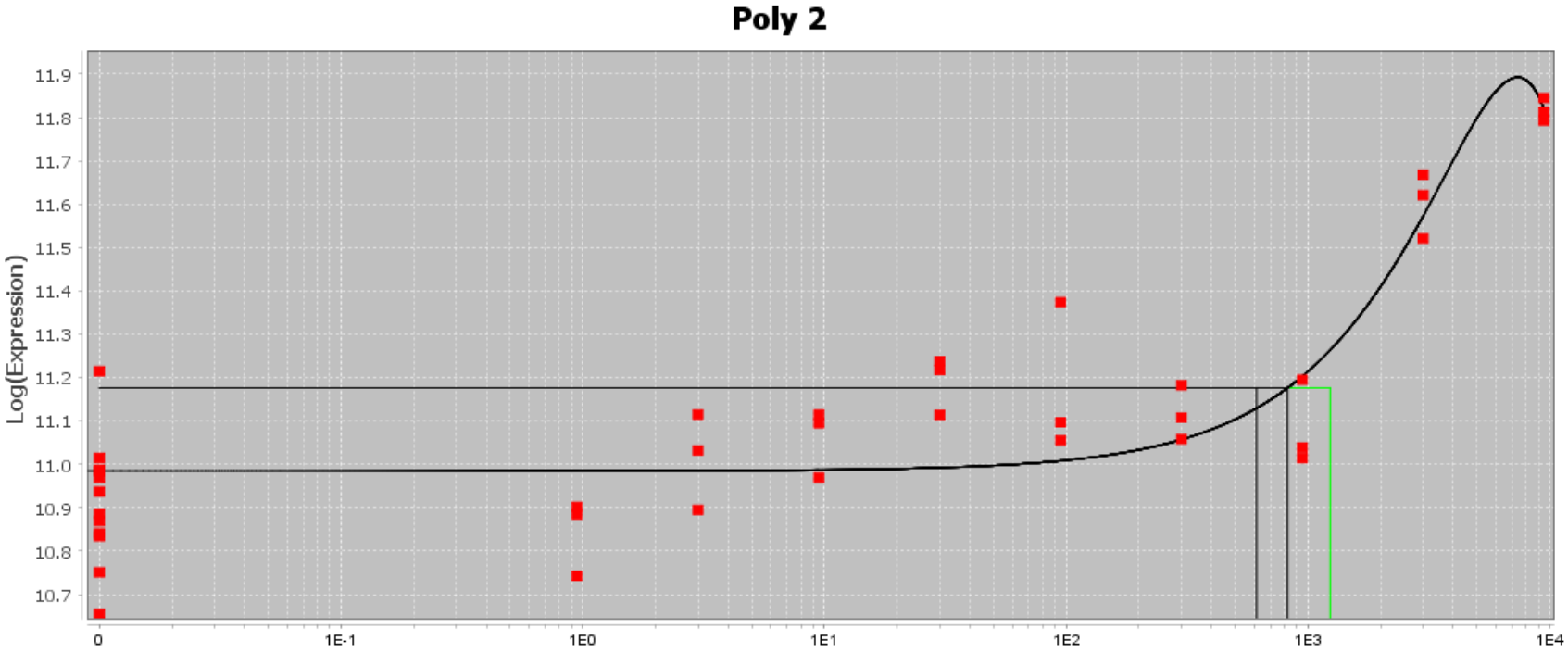
Why no call by BMDE? ATAD5_HepaRGcpd1run1



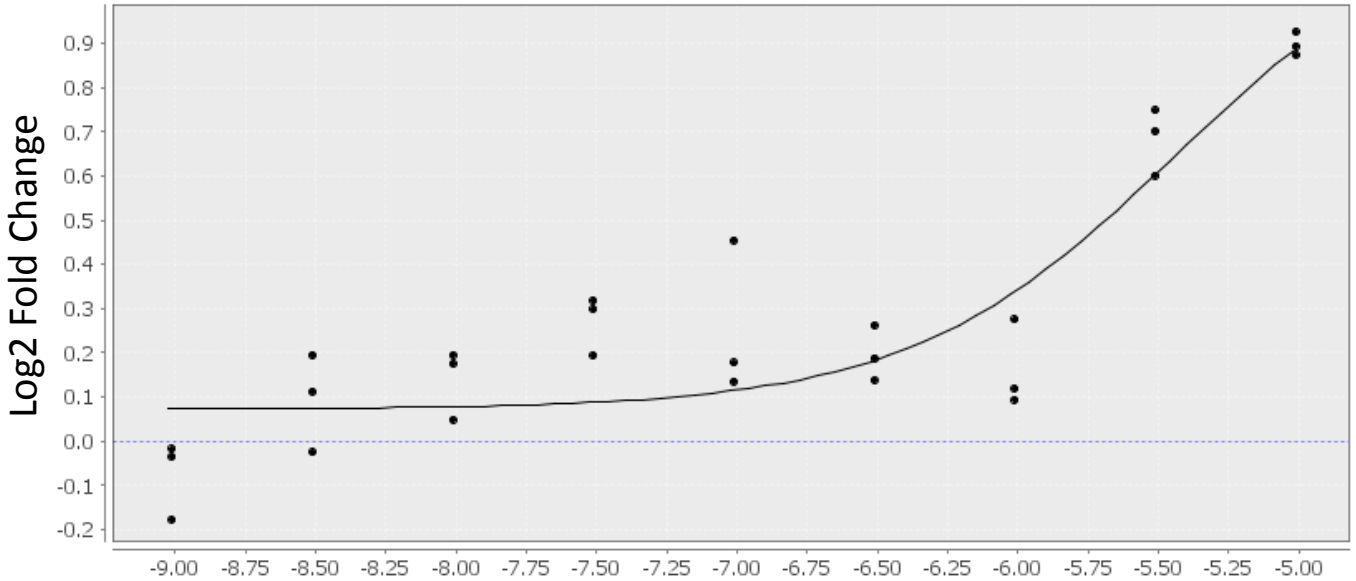
Not significant by BMDE??
Up by POD (+1)



Why no call by BMDE? SOD1_HepaRGcpd1run1



Not significant by BMDE??
Up by POD (+2)



Recommendations

1. BMD Express software – kudos for functionalities and ease of use!
2. A single model (Hill eq.) for consistency, including repeats of same experiment. The simplest most constrained model that is applicable to transcriptional regulation minimizes overfitting, thus false positives. For biphasic dose responses, event at lower dose is more important, if significant.
3. Single (Hill) model also improves performance.
4. Trend Test facilitates true/false positive decisions especially at lowest dose. POD is conservative to minimize False +. Once we make calls for each gene, use BMD/POD # SDs or minimum fold change to adjust stringency?
5. A database is imperative to store data and experiment annotations. NCATS uses enterprise grade database storage. This allows central storage and search for all processed data. With proper Oracle version, supports multi-billions of records. John Braisted can advise.
6. Suggest using BMDE Hill eq, add a trend test, minimum change 3 SD from mean of vehicle controls instead of current 2-fold (range of fit line?)? A SD-basis adjusts for noisy experiments/noisy genes.
7. Suggest retesting these algorithms in cases where there are gene responses at the lowest dose tested (expensive to test every chemical at low [nM]). Trend test helpful here?

Pathways

- Let's judge the BMD/POD calculations based on genes. . . . Then assign to pathways to avoid subjectivity.
- Indexing genes to pathways is useful! Pathway databases e.g. GO, KEGG, Bioplanet, *etc*. are carefully curated, but unavoidable arbitrary:
  - Gene set in a pathway cell type-dependent on. *E.g.* the "PPARg Pathway" or "AHR Pathway" defined by transcriptional changes in hepatocytes.
  - Pathways are hierarchical. E.g. consider "Inflammation" is less actionable than if we separately consider "Neutrophil recruitment" *vs*. "Inflammasome activation" *vs*. "Macrophage activation".
  - Pathway types:
    - Transcription factors, *e.g.* estrogen response or AHR (specify cell type/s?)
    - Metabolic or signal transduction pathways
    - Gene family or enzyme function (disregard these?)
  - For pathways not defined by a transcription factor (*e.g.* metabolic- or signal transduction pathways), only a small subset/# of genes are transcriptionally regulated by a relevant treatment.

Pathways

Can we find rules that mitigate arbitrary nature of pathway databases?

Suggestions:

1. Use only(?) the finest granularity level of pathway hierarchy. Minimizes redundancy.
2. Pre-identify non-redundant pathways of toxicological concern *vs.* non-concern? E.g. “ER stress” but not “Kinase Signalling” or “Transcription”
3. Be conservative: Flag a pathway as transcriptionally activated when several genes in that pathway are induced or repressed. E.g. $\geq$ 10% of pathway, minimum 2 genes.
4. Not TOO conservative: Represent BMD/POD for a pathway as the second lowest BMD/POD for genes in that pathway. (Using mean or median of genes in a pathway would shift BMD/POD upward as we test higher doses).