

Concentration-response analysis of image-based high-throughput phenotypic profiling data for chemical bioactivity screening

Johanna Nyffeler, PhD



ORCID 0000-0002-6155-9743

Nyffeler.Johanna@epa.gov

Disclaimer

The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Acknowledgements



Office of Research and Development (ORD) Center for Computational Toxicology and Exposure (CCTE)

Harrill Lab team

- **Joshua Harrill**
- **Clinton Willis**
- Rick Brockway
- Megan Culbreth
- Dan Hallinger
- Terri Fairley

Data analysis

- Ryan Lougee
- Ann Richard
- John Wambaugh
- Derik Haggard
- Woody Setzer
- Logan Everett
- Richard Judson
- Katie Paul-Friedman



National Toxicology Program
U.S. Department of Health and Human Services

- Scott Auerbach

1. EPA's testing vision
2. What is (phenotypic) profiling?
3. Application of HTPP at the Center for Computational Toxicology & Exposure (Nyffeler et al. 2020a)
4. Optimization of concentration-response modeling and potency estimation (Nyffeler et al. 2020b, *accepted*)

Blueprint of Computational Toxicology

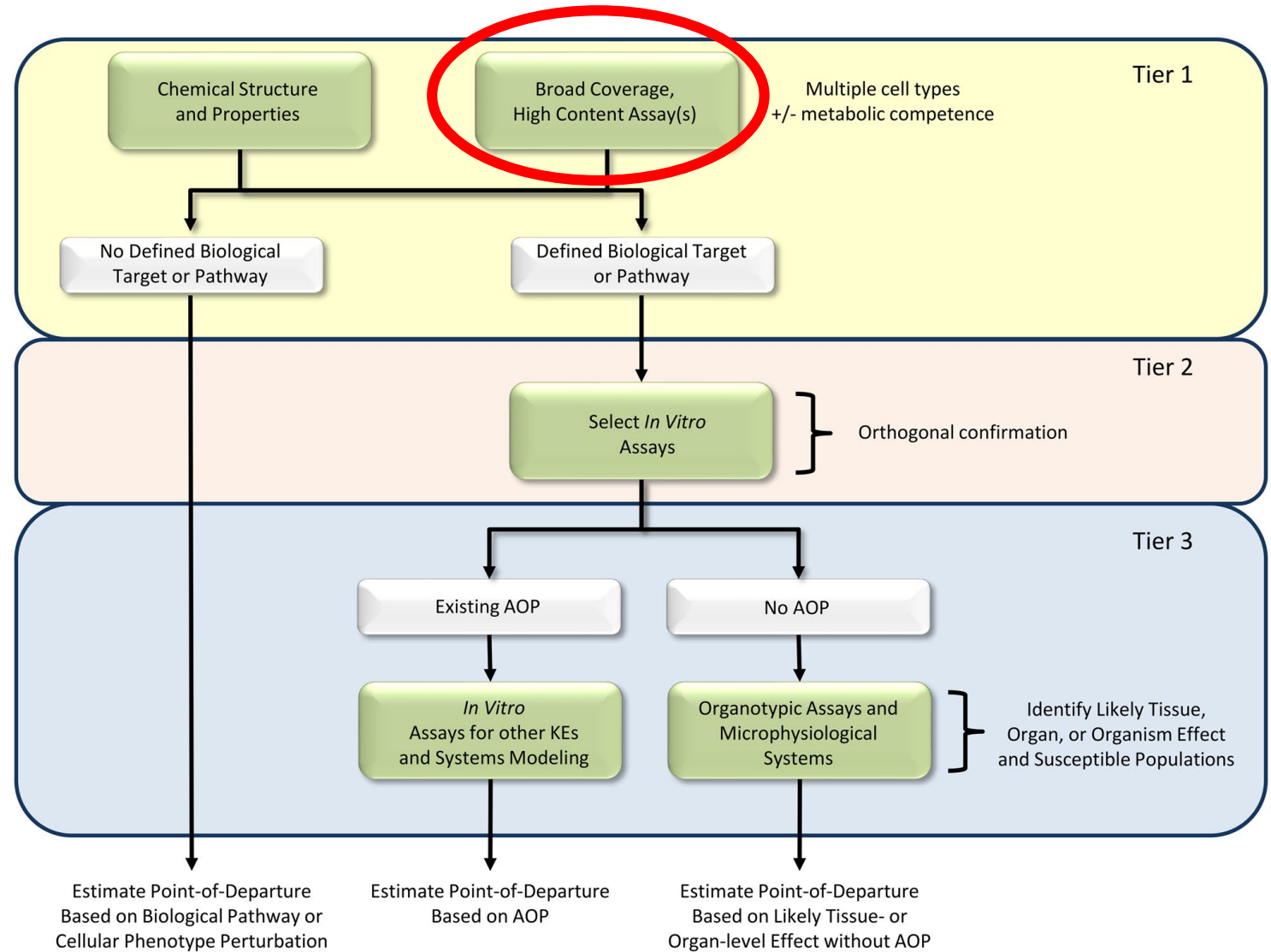
The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency

Thomas et al. 2019

Toxicological Sciences, Volume 169, Issue 2,
June 2019, Pages 317–332

Two profiling assays:

- transcriptomics
- phenotypic profiling



What is (phenotypic) profiling?

What does 'profiling' mean?

Targeted assays

Example: Estrogen receptor agonist assay (NVS_NR_hER)

- Response: decreased radioligand binding
- Positive control: 17b-estradiol
- Number of endpoints: 1

→ For active chemicals, the response is a predictable change in a single endpoint in a known direction

Profiling assays

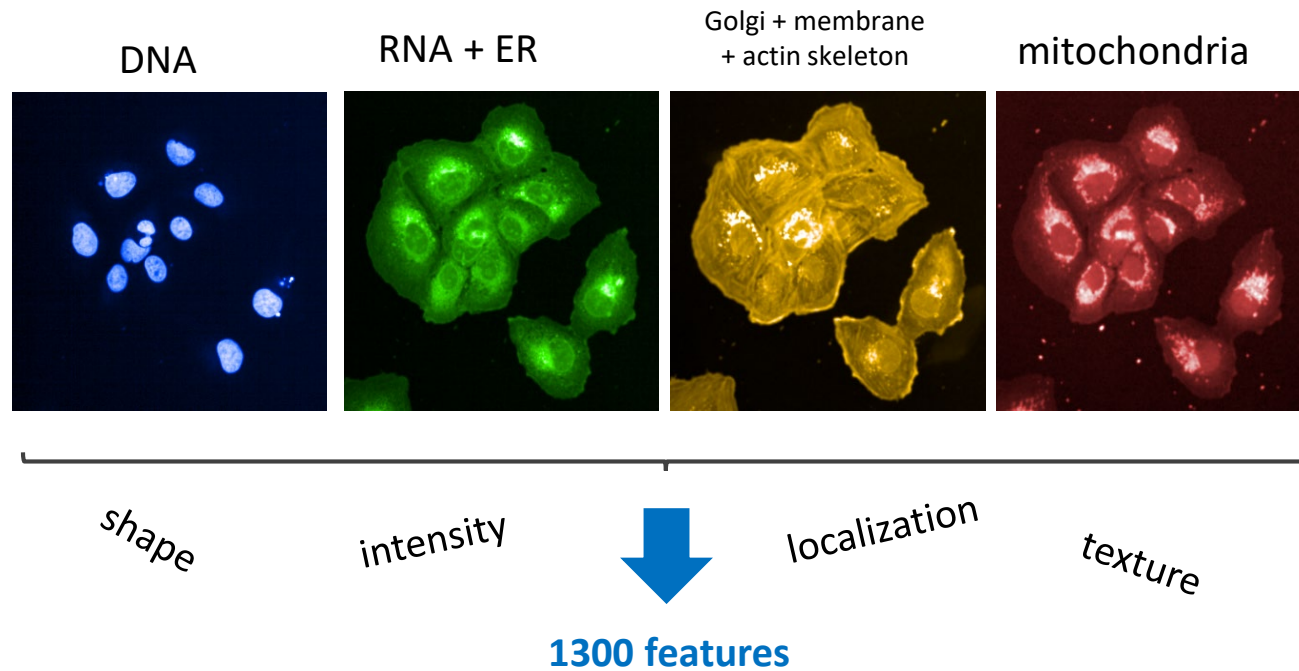
Example: Transcriptomics

- Response: any meaningful change in transcript levels
- Number of 'endpoints': ~ 10'000

→ For active chemicals, responses involve changes in many different endpoints in unknown directions. Vary from chemical-to-chemical.

What is imaging-based phenotypic profiling?

- staining of various cell organelles with fluorescent dyes in *in vitro* cultures
- assessing a large variety of morphological features on individual cells

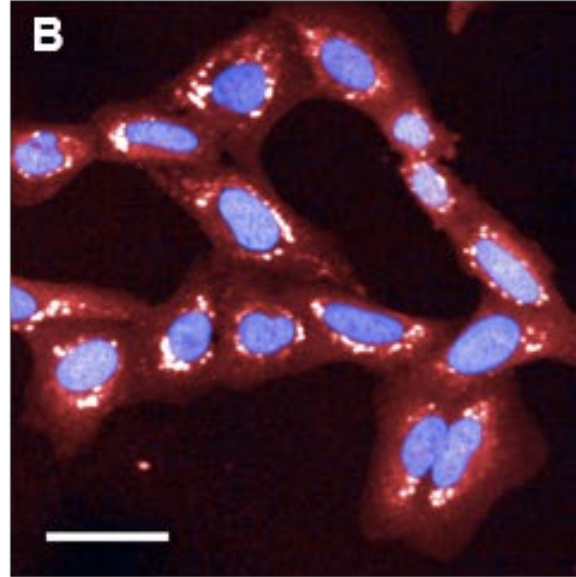
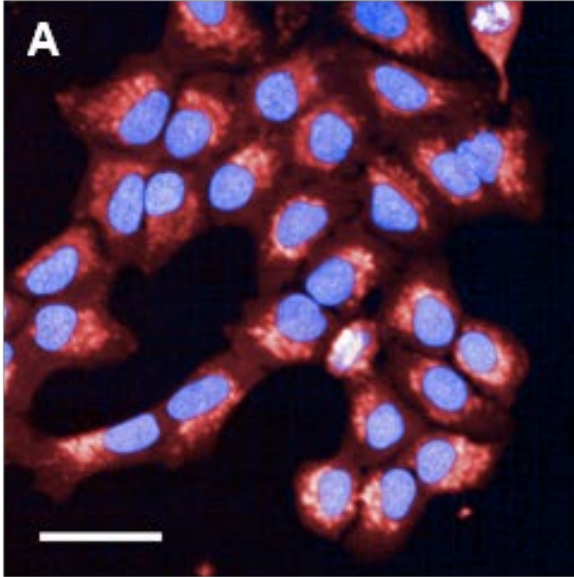


Cell Painting = Cytological Profiling = Phenotypic Profiling = high-throughput Phenotypic Profiling = HTPP

Exemplary chemicals

Solvent control (0.5% DMSO)

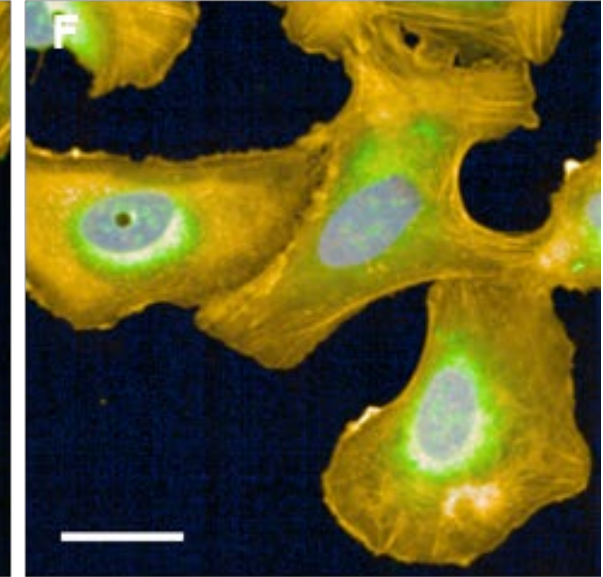
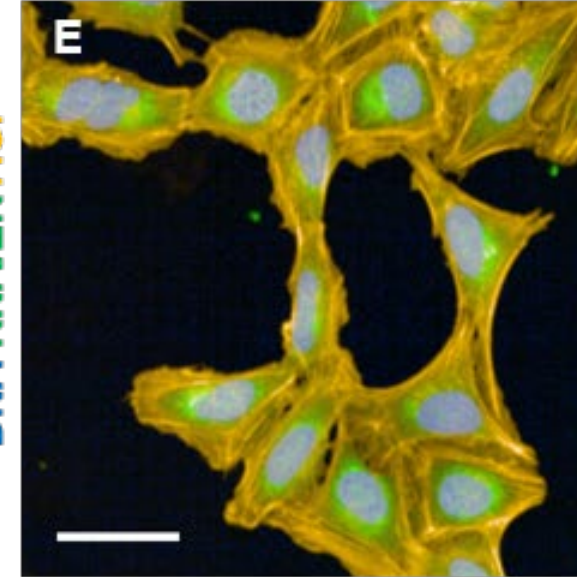
Berberine chloride (10 μ M)



→ Mitochondrial
compactness/texture

Solvent control (0.5% DMSO)

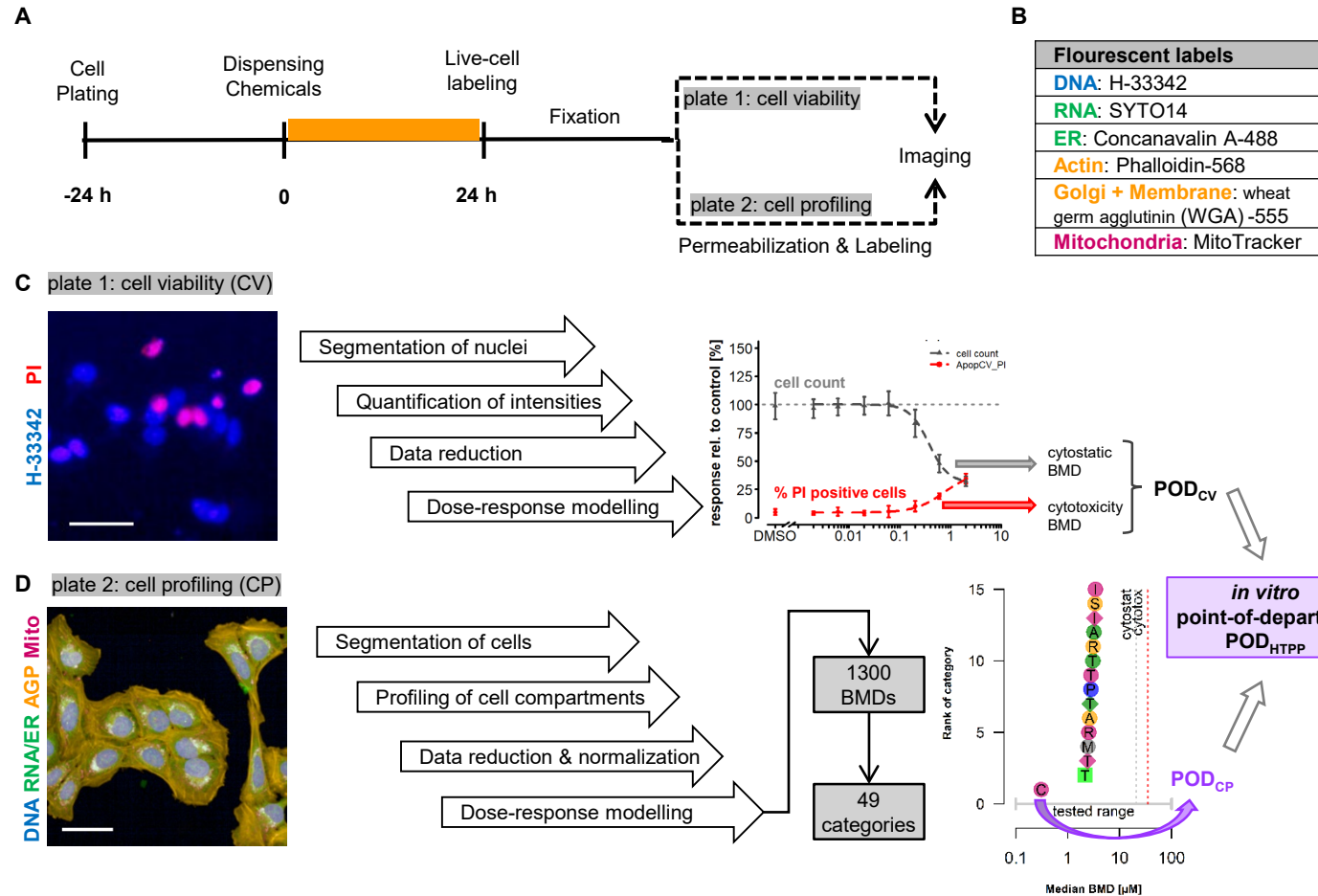
Etoposide (3 μ M)



→ Cells are larger

⇒ **Strong phenotypes are observable qualitatively**

The High-Throughput Phenotypic Profiling (HTPP) assay



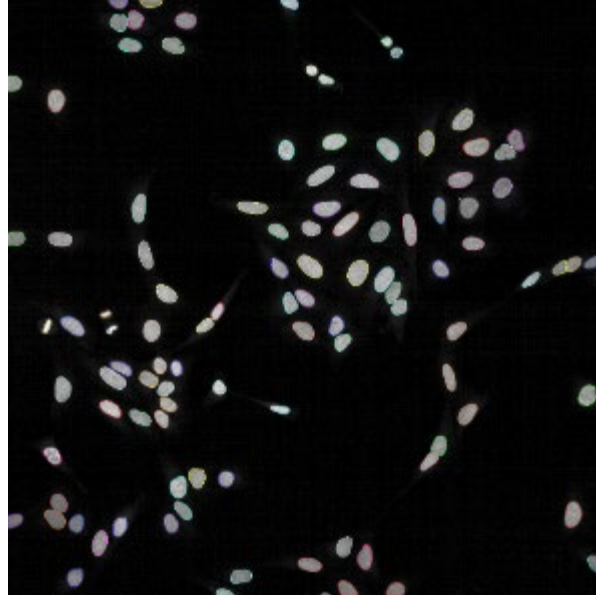
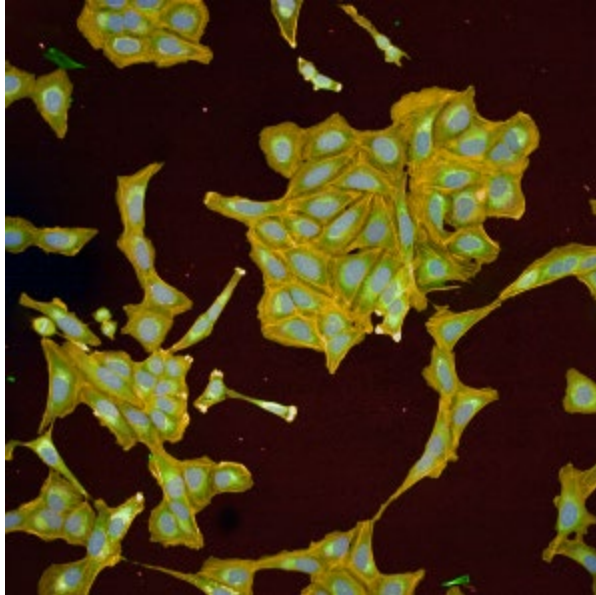
POD: point-of-departure

=

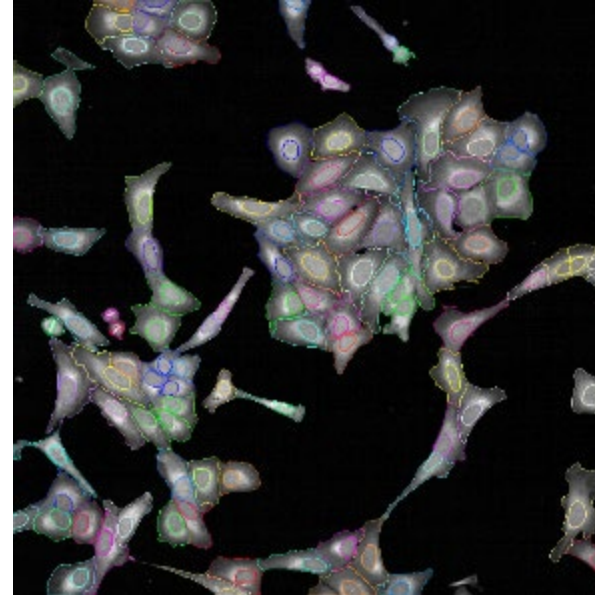
PAC: phenotype altering concentration

Image analysis workflow: image segmentation

1. find nuclei



2. find cell outline



3. reject border objects

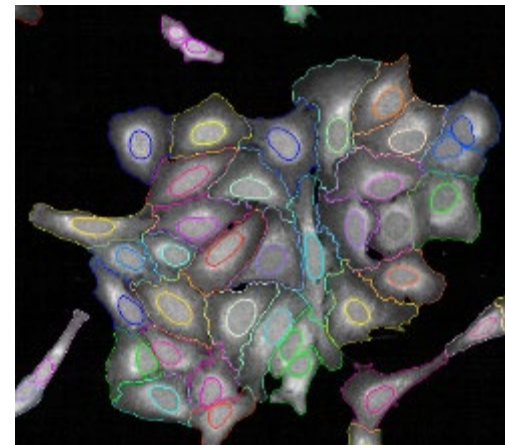
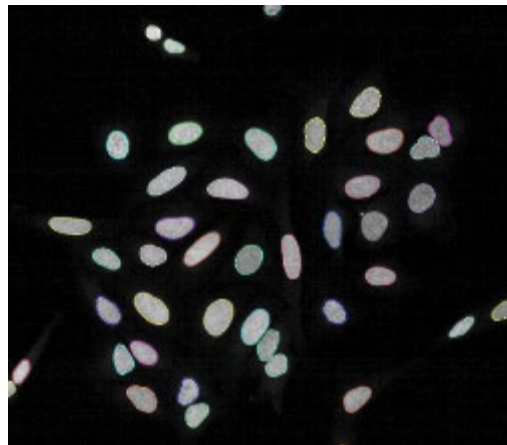
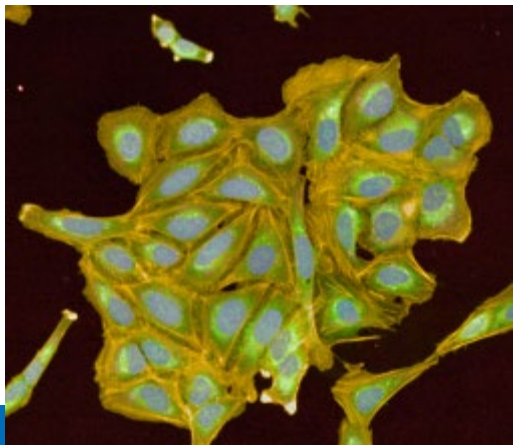
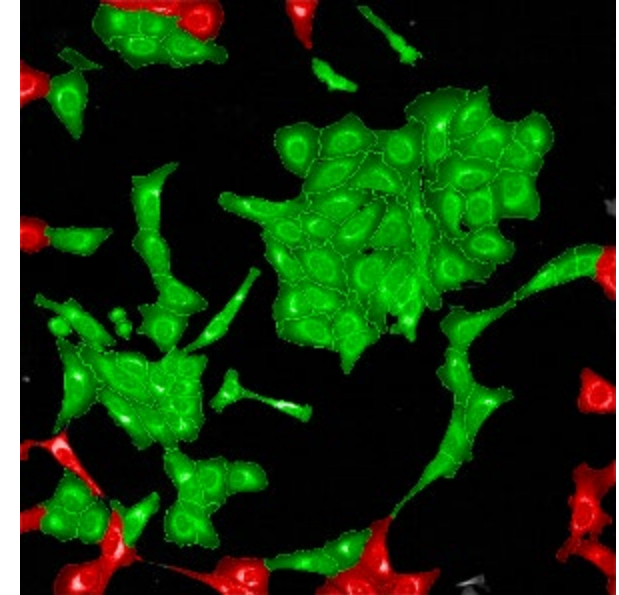
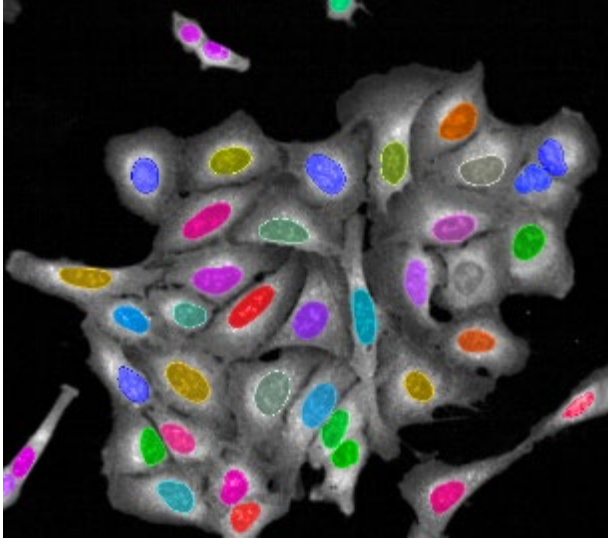
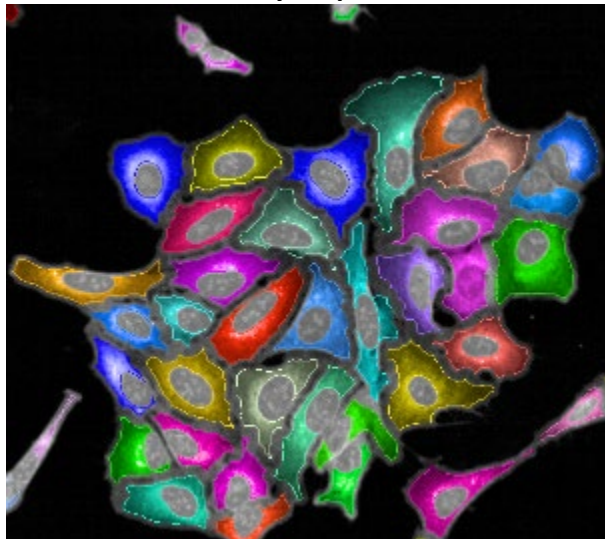


Image analysis workflow: define cellular compartments

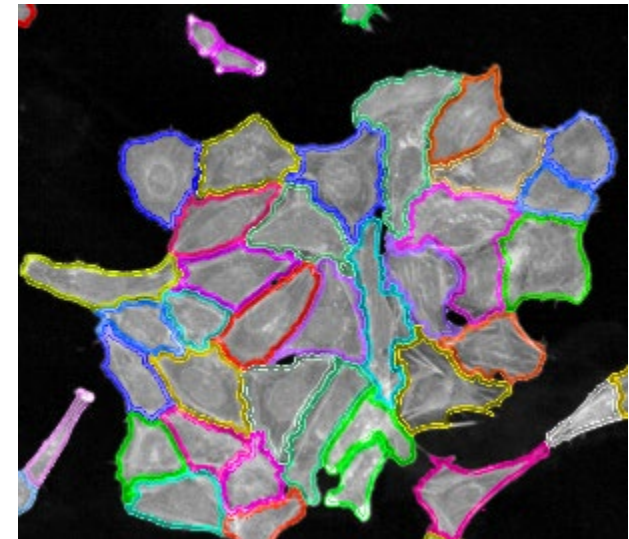
nuclei



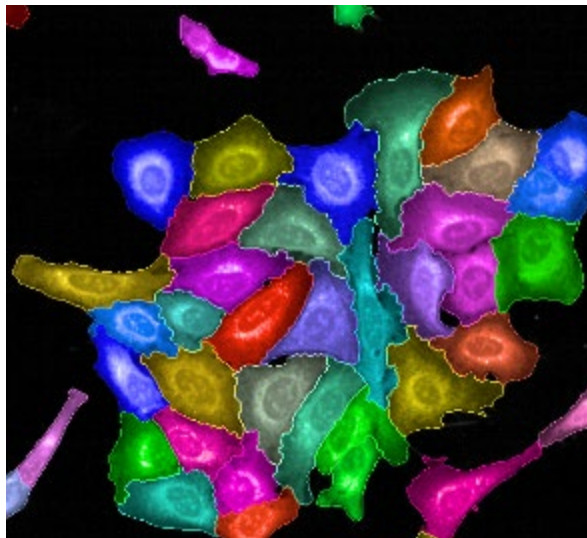
cytoplasm



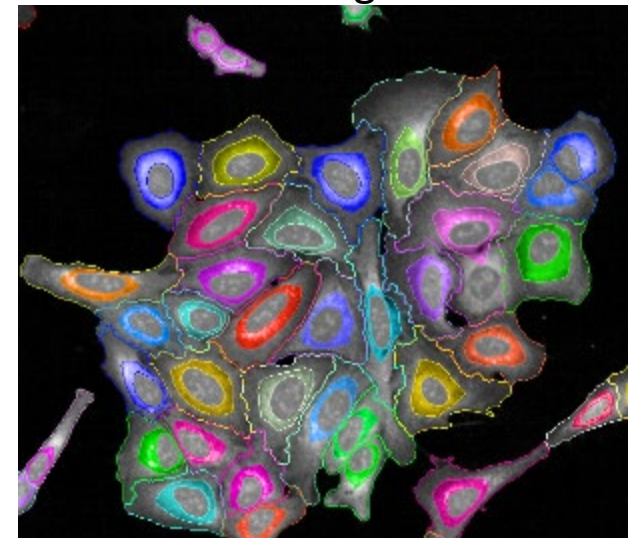
membrane



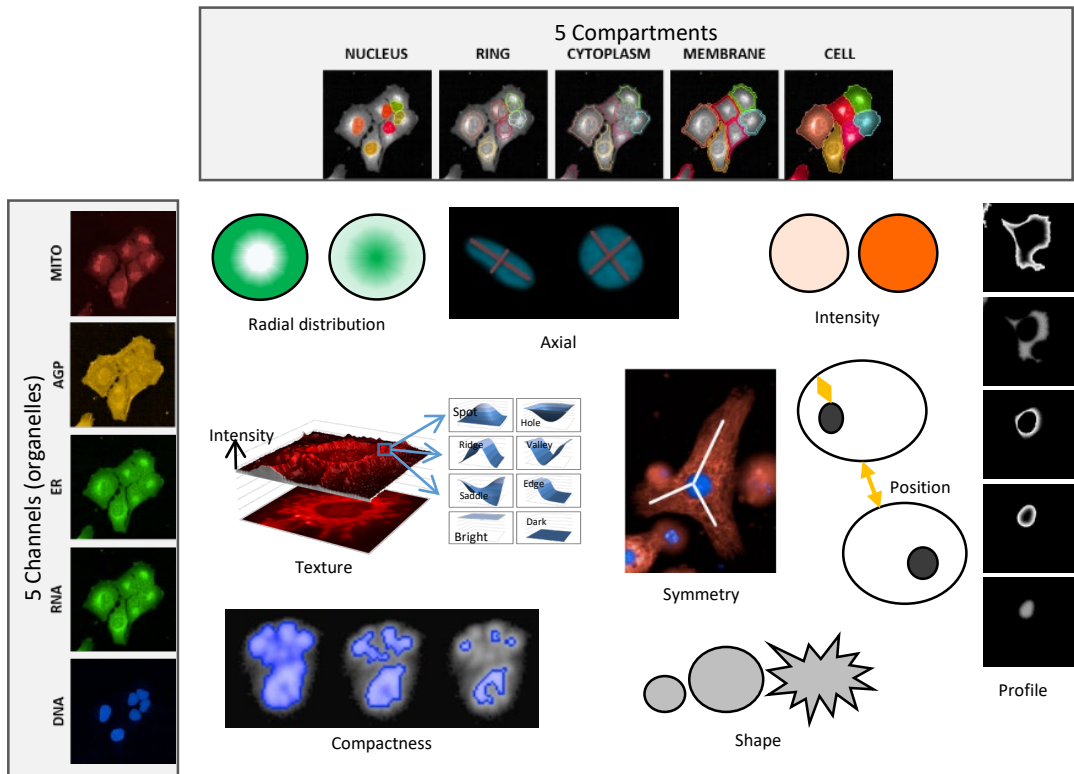
cell



ring



Phenotypic feature extraction



49 feature categories
(ex. Mito_Texture_Cytoplasm)

1300 features / cell

Profile		Module								
		Position [7]	Basic morph- ology [5]	SCARP morphology					Intensity [9]	Texture [14]
				Symmetry [80]	Compactness [40]	Axial [20]	Radial [28]	Profile [20-30]		
Channel	DNA			Nuclei	Nuclei	Nuclei	Nuclei Cell	Nuclei Cytoplasm	Nuclei	Nuclei
	RNA			Nuclei	Nuclei	Nuclei	Nuclei	Nuclei	Nuclei	Nuclei
	ER			Cell	Cell	Cell	Cell	Cytoplasm	Ring Cytoplasm	Ring Cytoplasm
	AGP			Cell	Cell	Cell	Cell	Nuclei Cytoplasm	Ring Cytoplasm Membrane	Ring Cytoplasm Membrane
	Mito			Cell	Cell	Cell	Cell	Nuclei Cytoplasm	Ring Cytoplasm	Ring Cytoplasm
	Not associated with a channel	Nuclei Cell	Nuclei Cell							

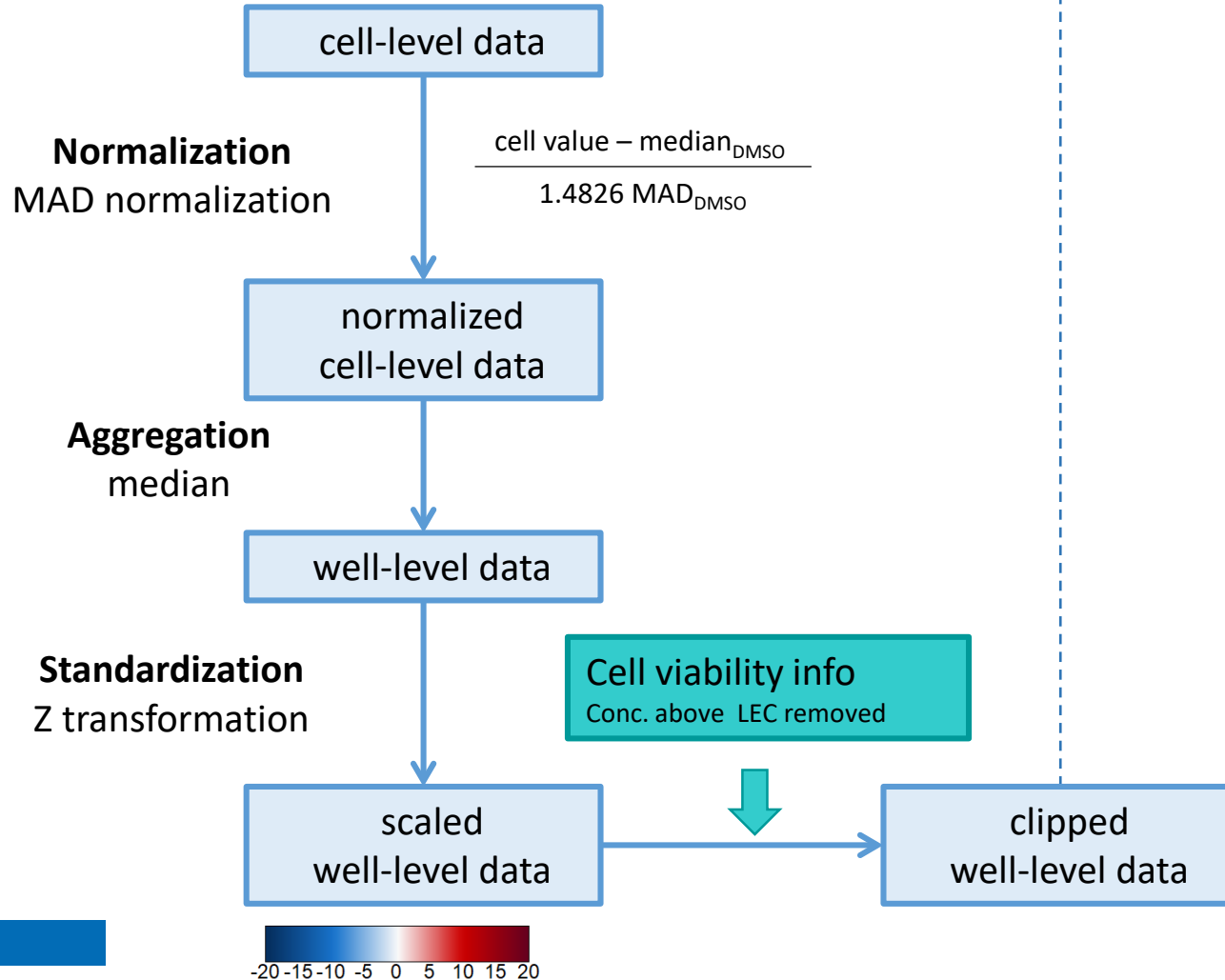
PerkinElmer Opera Phenix

Modality: Confocal (single z)
Objective: 20X Water
Plate: CellCarrier-384 Ultra
Fields: 5 or 9

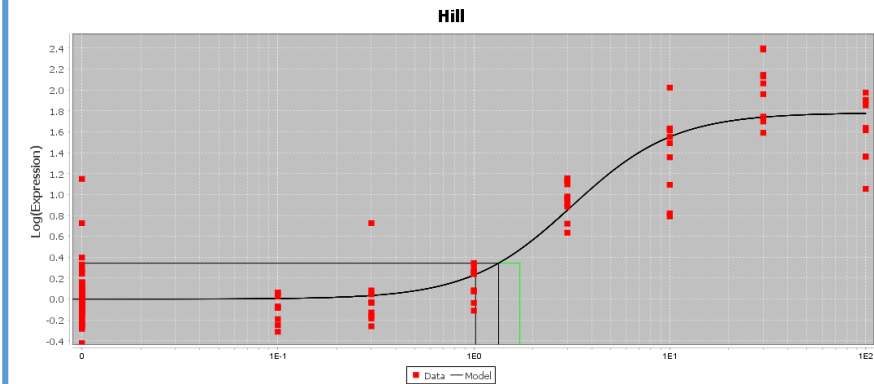
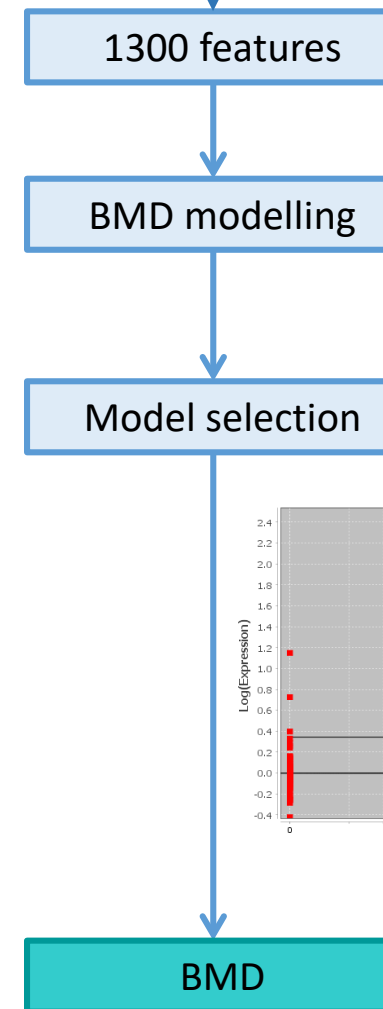


Data processing for profiling plates

Data reduction in R



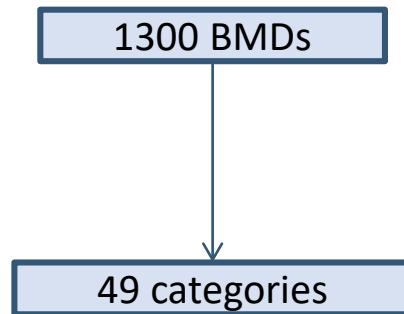
Benchmark dose (BMD) modelling using BMDExpress 2.2



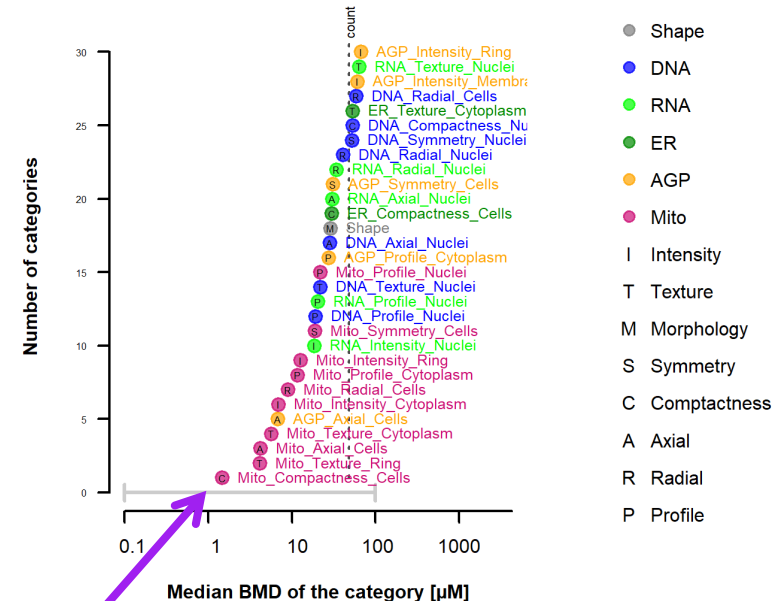
Berberine chloride
Mito_Cells_Morph_STAR

Aggregate the BMDs to a PAC

1. Group the 1300 BMDs into
49 categories



2. Order the categories by potencies



BMD: benchmark dose

=

BMC: benchmark concentration

POD: point-of-departure

=

PAC: phenotype altering concentration

Phenotype altering concentration (PAC):
Median BMD of the most sensitive ontology
(where $\geq 30\%$ ontology elements affected)

Application of HTPP

Toxicology and Applied Pharmacology 389 (2020) 114876

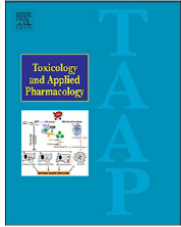


ELSEVIER

Contents lists available at [ScienceDirect](#)

Toxicology and Applied Pharmacology

journal homepage: www.elsevier.com/locate/taap



Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling

Johanna Nyffeler^{a,b}, Clinton Willis^{a,c}, Ryan Lougee^{a,b}, Ann Richard^a, Katie Paul-Friedman^a, Joshua A. Harrill^{a,*}



Check for updates

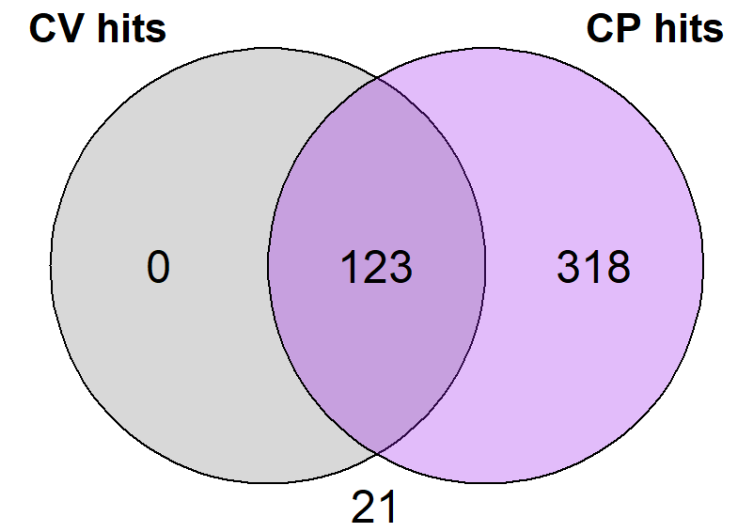
Screen of environmental chemicals

- 462 test chemicals
 - pesticides (~ 75%), drug-like chemicals, food additives, industrial chemicals
 - 448 chemical from the 'APCRA' list
 - available in vivo effect values
 - available toxicokinetic parameters for in vitro to in vivo extrapolation (IVIVE)



Kavlock et al. (2018)
Chem. Res. Tox.; 31(5): 287-290

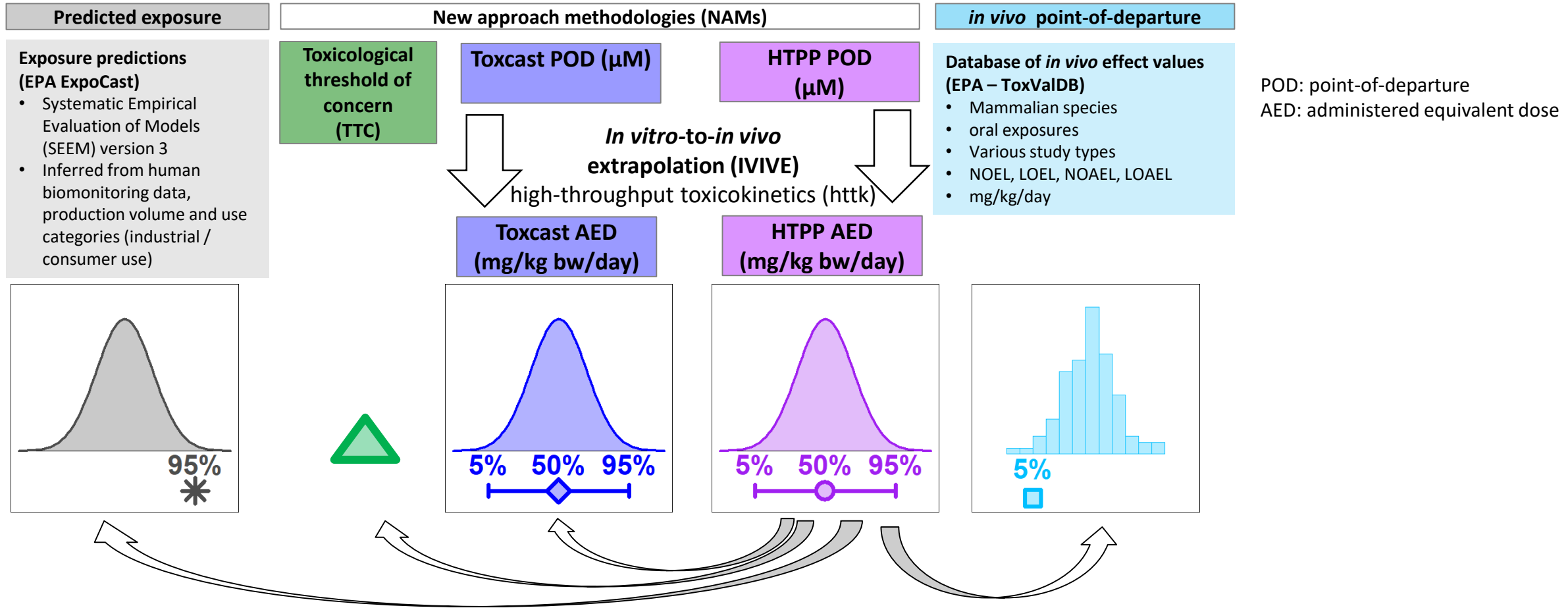
Experimental design	
Cell type	U-2 OS
Exposure time	24 h
Cell seeding density per well	400
# unique chemicals	462
# concentrations	8
Concentration spacing	$\frac{1}{2} \log_{10}$
# solvent controls/plate	24
# replicates/plate	1
# independent experiments	4



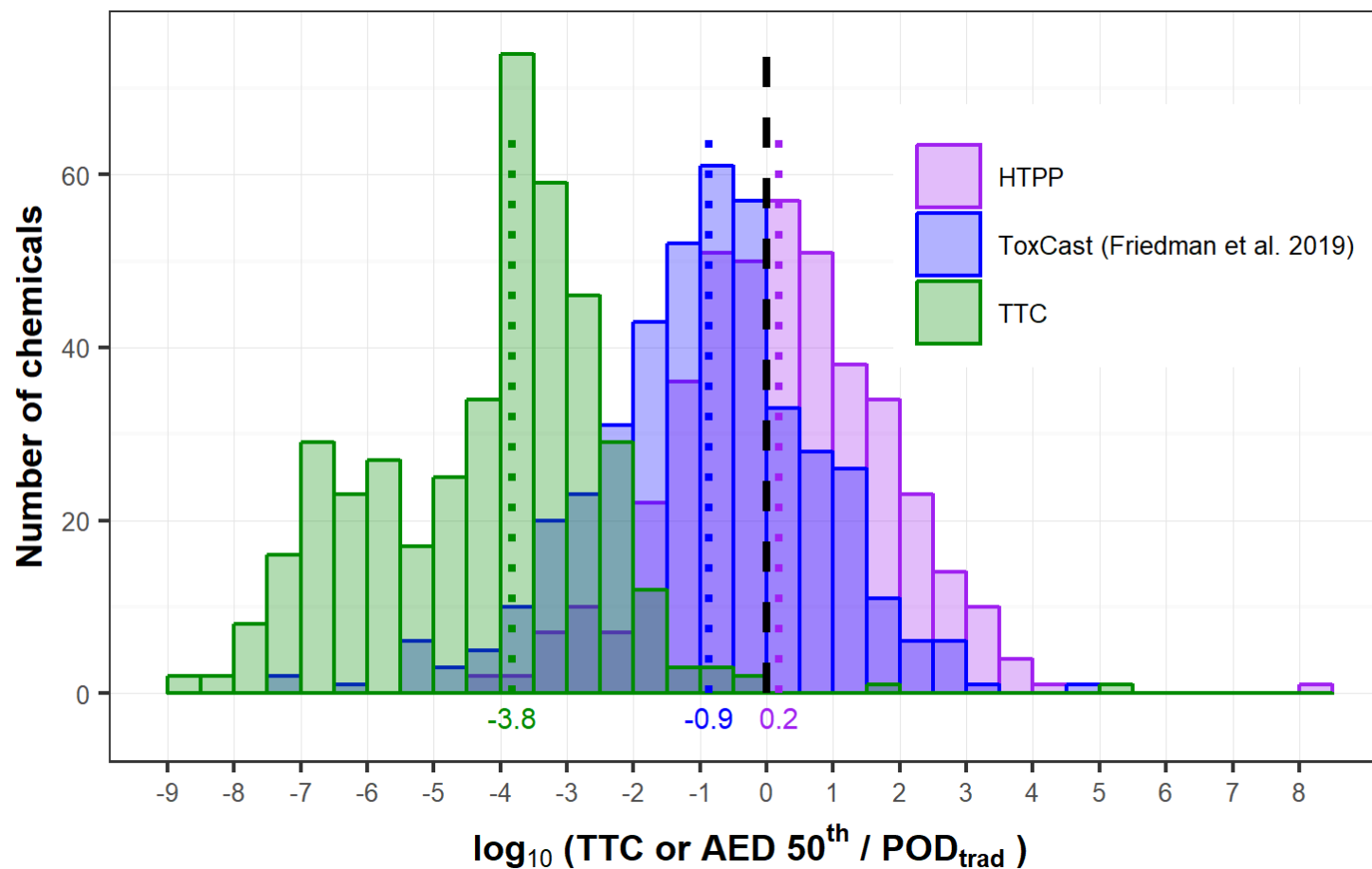
Nyffeler et al. 2020a

⇒ **95% of test chemicals were bioactive in the HTPP assay**

Comparison to *in vivo* data and exposure



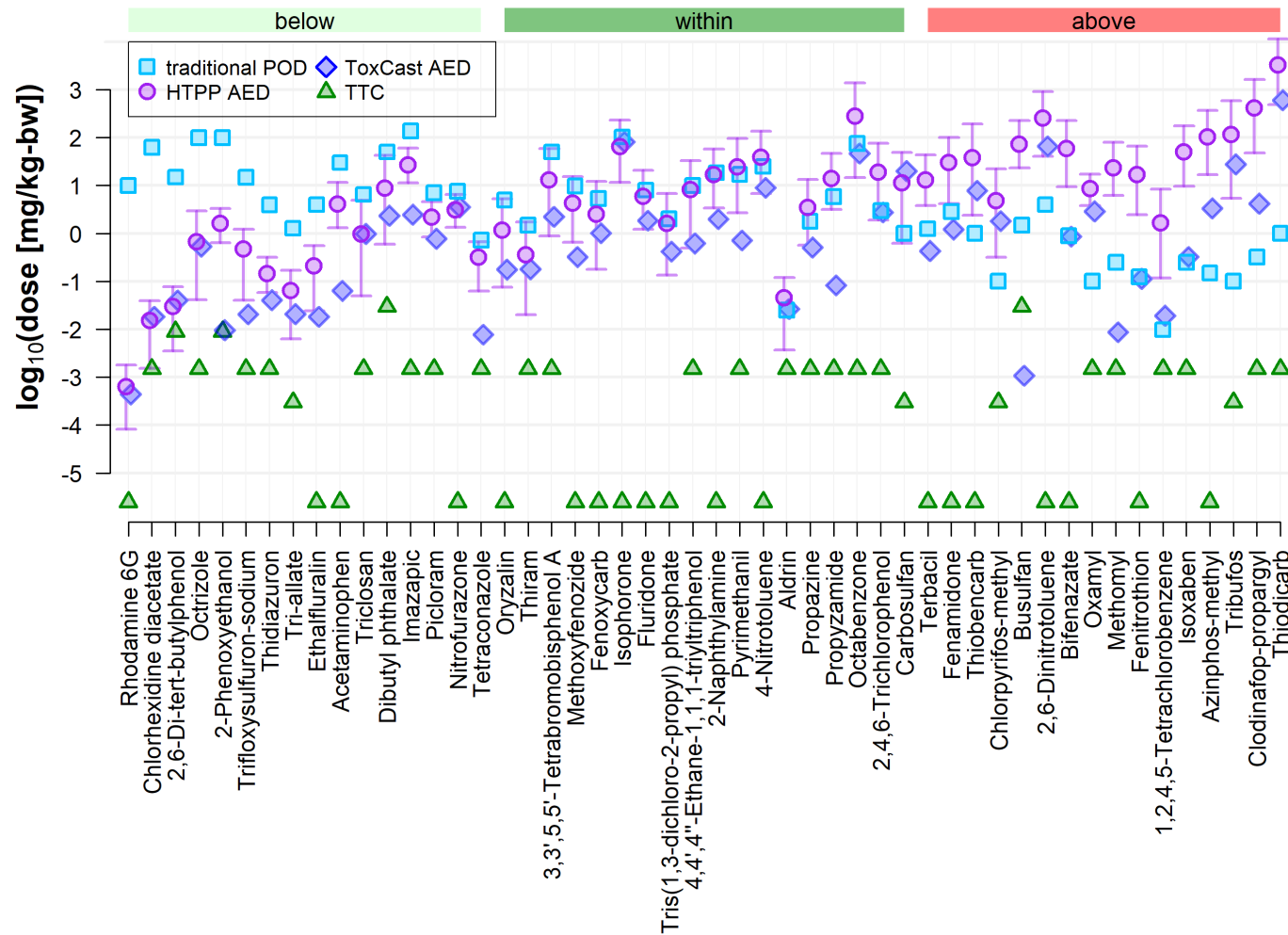
Comparison to *in vivo* effect values & other NAMs (I)



Nyffeler et al. 2020a

- ⇒ HTPP AEDs are higher than ToxCast-derived AEDs and TTC values
- ⇒ 81% of HTPP AED are within 2 orders of magnitude of the *in vivo* POD

Comparison to *in vivo* effect values & other NAMs (II)



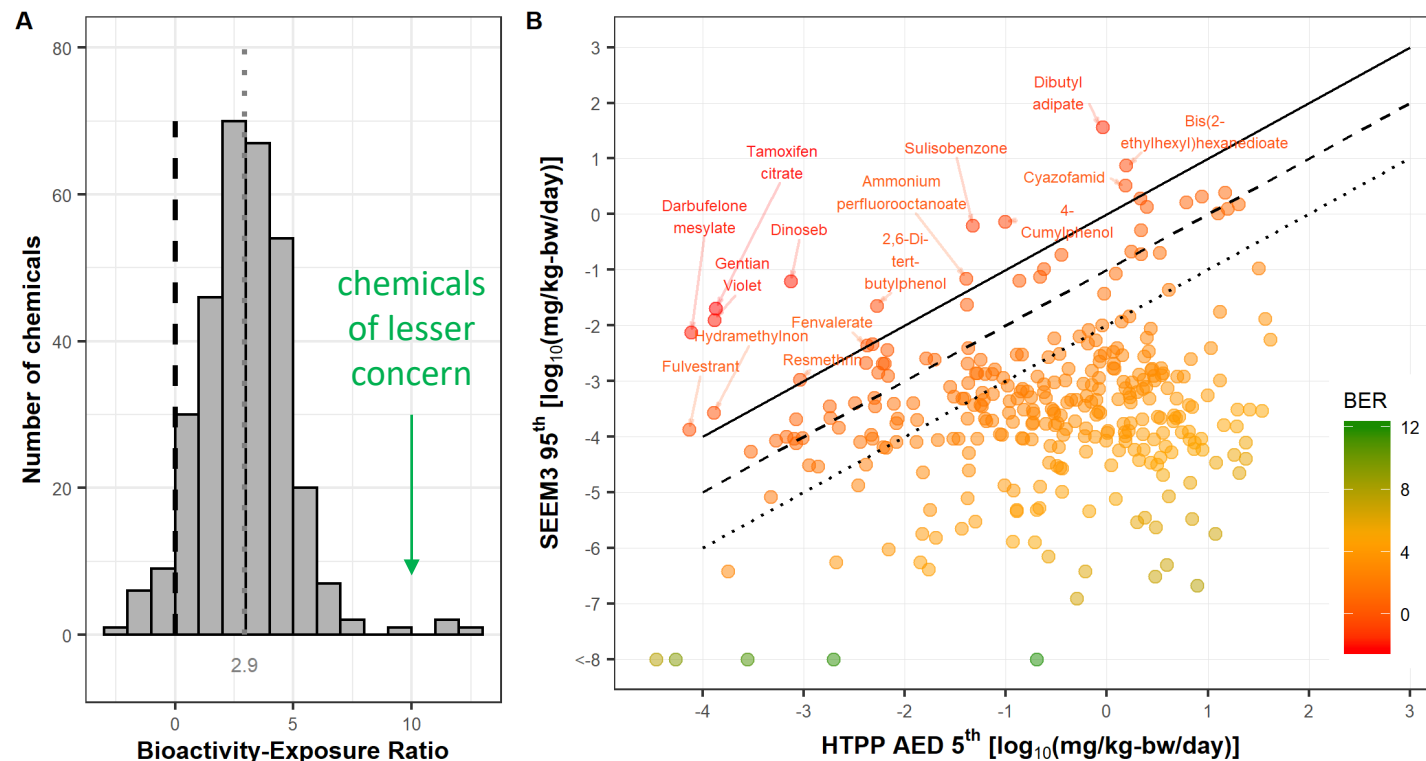
Nyffeler et al. 2020a

⇒ for 68% (285/420) of chemicals, HTPP AEDs led to a conservative or comparable surrogate

Comparison to exposure estimates

HTPP AEDs were compared to exposure predictions and the bioactivity exposure ratio was calculated as follows:

$$\text{Bioactivity exposure ratio (BER)} = \frac{\text{lower bound of HTPP bioactivity}}{\text{upper bound of exposure estimate}} = \log_{10} \left(\frac{\text{HTPP AED } 5^{\text{th}}}{\text{SEEM3 } 95^{\text{th}}} \right)$$



unpublished

- ⇒ for 49% of chemicals, predicted exposure is > 1000x lower than estimated bioactivity
- ⇒ for a small set of chemicals, the BER was negative, indicating a potential for humans to be exposed to bioactive concentrations of these chemicals

Conclusions I



HTPP *in vitro* potencies can be used for bioactivity exposure ratio analysis and prioritizing of chemicals based on inferred bioactivity in relation to predicted human exposure

Next steps:

- Test chemicals in multiple cell types to increase biological coverage

Optimization of Hit Identification

Comparison of Approaches for Determining Bioactivity Hits from High-Dimensional Profiling Data

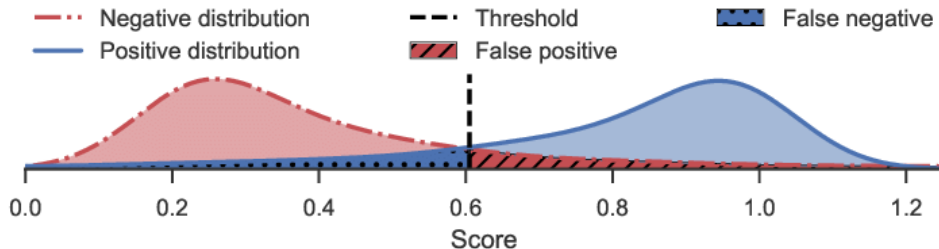
Johanna Nyffeler^{1,2}, Derik E. Haggard^{1,2}, Clinton Willis^{1,3}, R. Woodrow Setzer¹, Richard Judson¹, Katie Paul
Friedman¹, Logan J. Everett¹, Joshua A. Harrill¹

Challenges in analysis of profiling data

Targeted assays

- Response is predictable
- Often have a positive control
- Often have known negative controls

➔ Use of positive and negative controls to set a threshold for hit calls



https://www.researchgate.net/profile/Denis_Reis/publication/327847657/figure/fig1/AS:674446763380738@1537812047280/Threshold-and-score-distribution-for-a-binary-classification-process.png

Profiling assays

- Measure 100s – 1000s of features
➔ not feasible to define a threshold for each feature in an analogous manner to targeted assays.
- Multiple diverse phenotypes can be observed
➔ no single ‘positive control’
- Multiple testing problem can lead to identification of false actives

➔ How should thresholds be chosen to ensure reliable hit calls?

➔ no widely accepted standard practices for hit identification from phenotypic profiling data ➔ potential barrier for regulatory applications

Challenges of environmental chemicals

- Often low expected bioactivity
 - Often lack a specific molecular target in human-based cell models
 - ‘poly-pharmacology’
 - Responses can be associated with general cell stress
- ⇒ **more challenging for hit identification than drug-like chemicals**

- **15 different approaches were compared at a fixed false positive rate of ~10%**

100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	A	B	C	DMSO
30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	A	B	C	DMSO
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	A	B	C	DMSO
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	A	B	C	DMSO
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	A	B	C	DMSO
0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	A	B	C	DMSO
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	A	B	C	DMSO
0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	A	B	C	DMSO
100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	D	Stauro	DMSO	DMSO
30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	D	Stauro	DMSO	DMSO
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	D	Stauro	DMSO	DMSO
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	D	Stauro	DMSO	DMSO
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	D	Stauro	DMSO	DMSO
0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	D	Stauro	DMSO	DMSO
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	D	Stauro	DMSO	DMSO
0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	D	Stauro	DMSO	DMSO

Different approaches to identify hits

multi-concentration approaches

single-concentration approaches

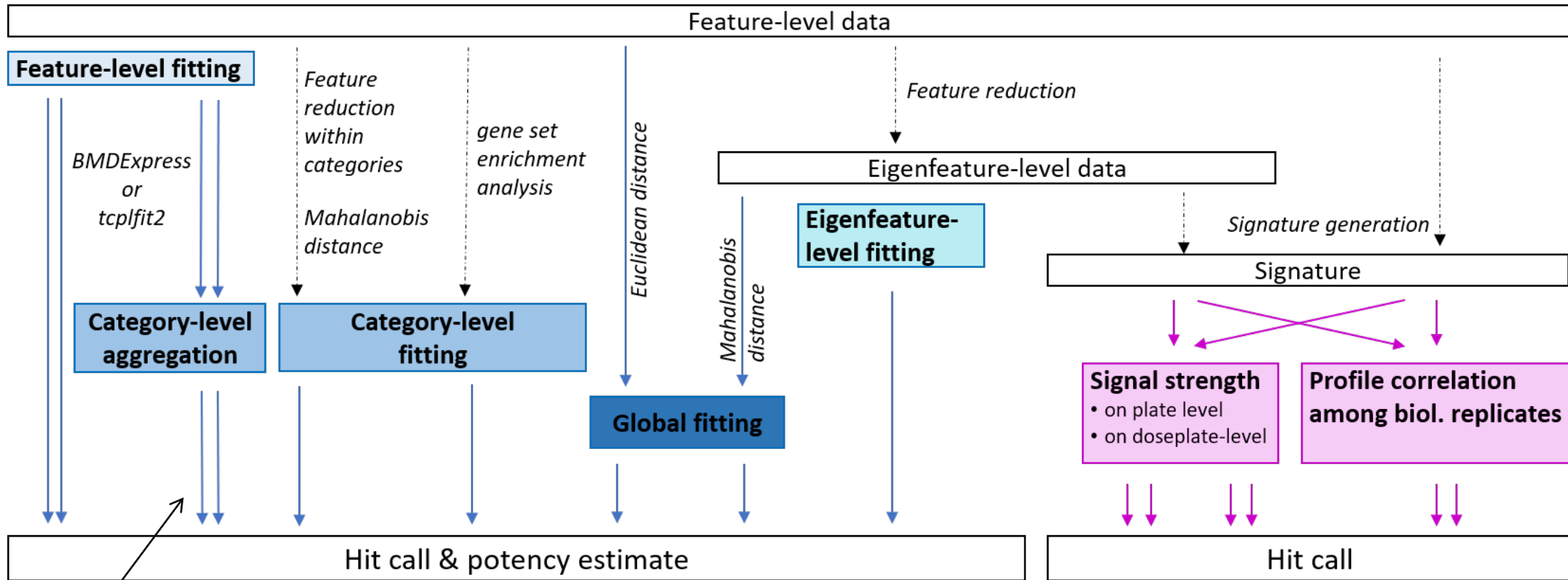
Dimensions

1300

~260

49

1



previously used
approach

Nyffeler et al. 2020b

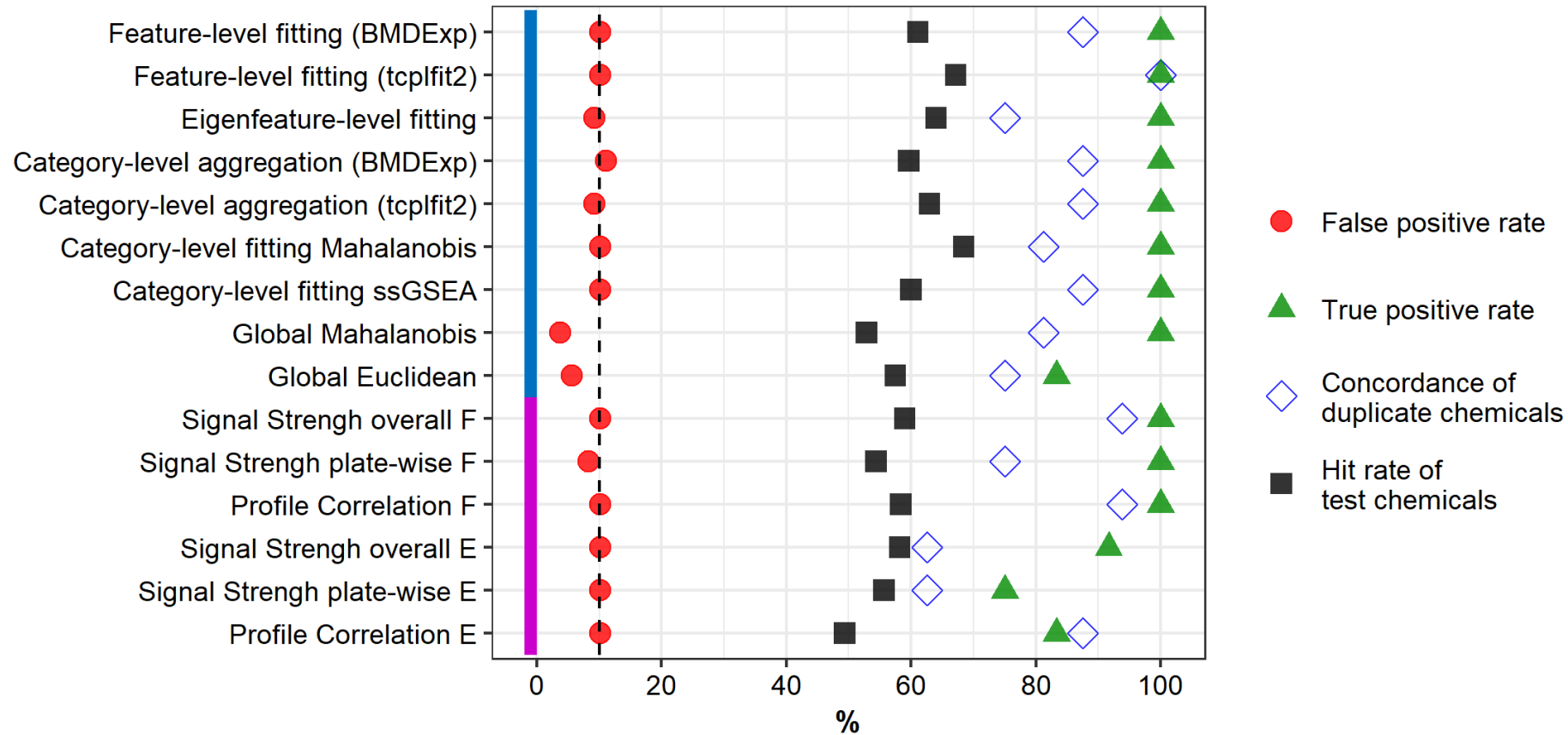
potency estimate = phenotype altering concentration = PAC

- False positive rate (FPR) = % of null chemicals that are positive
 - Null sets are constructed from the lowest 2 concentration of all test chemicals
- True positive rate (TPR) = % of APCRA Berberine that are positive
 - Berberine chloride: weak chemical with specific effects in only 100-200 features
→ most closely resembles expected behavior from positive test chemicals
- Hit rate = % of test chemicals that are active
- Concordance:
 - % of test chemicals with concordant hit calls (all inactive or all active)
 - Number = # chemicals that are active

Thresholds for each approach were individually optimized for

1. False positive rate of ~ 10%
2. Highest true positive rate (100%)
3. Best possible concordance & high hit rate

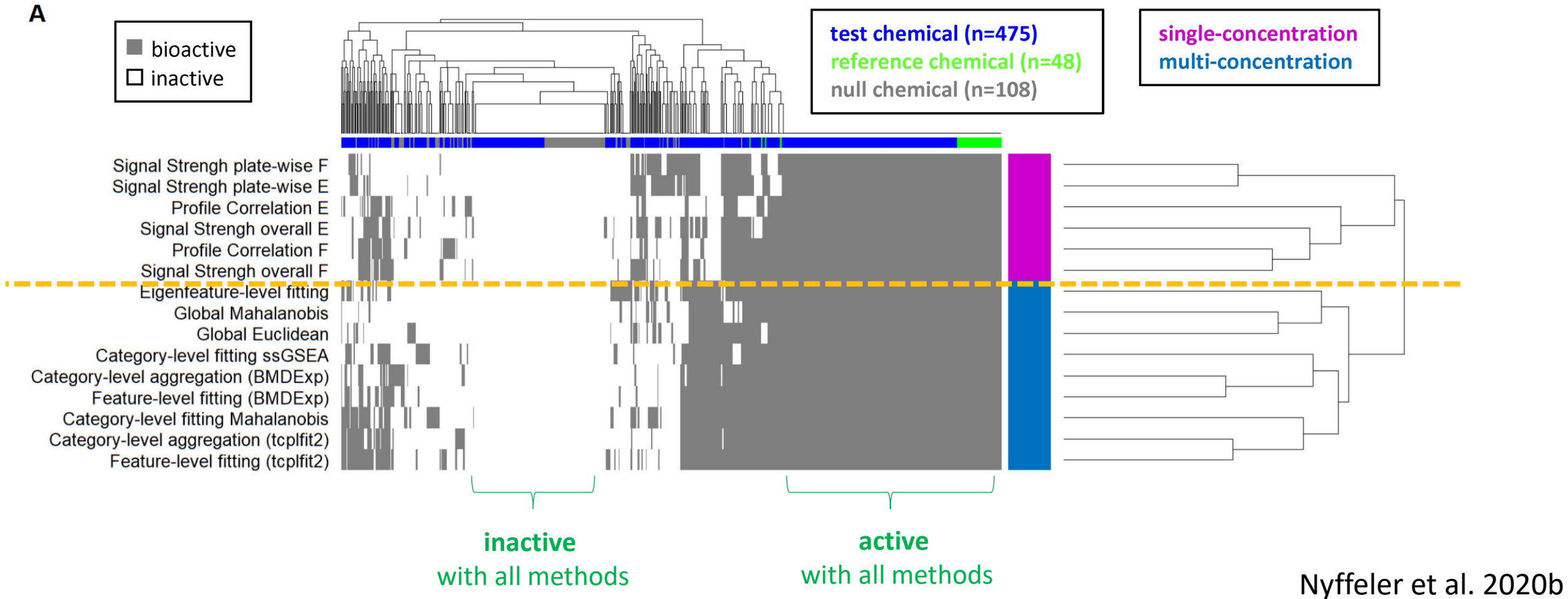
Optimizing approaches to achieve equivalent false discovery rate



Nyffeler et al. 2020b

- ⇒ 11/15 approaches identified 100% of true positives
- ⇒ Hit rate is overall between 50-70%

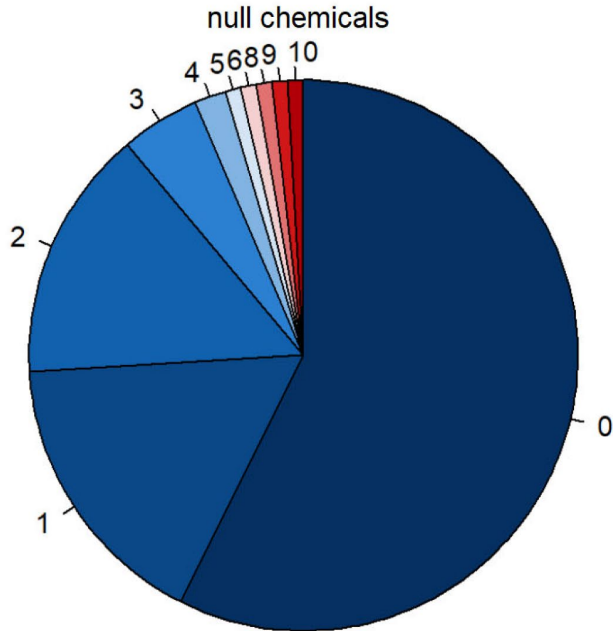
Concordance of hit calls across approaches (I)



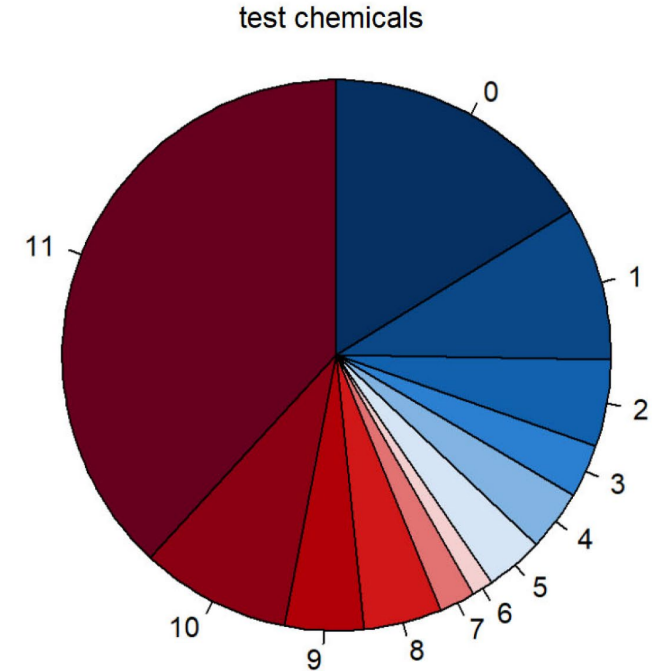
⇒ Large amount of chemicals that are unanimously identified as active/inactive

Concordance of hit calls across approaches (II)

B



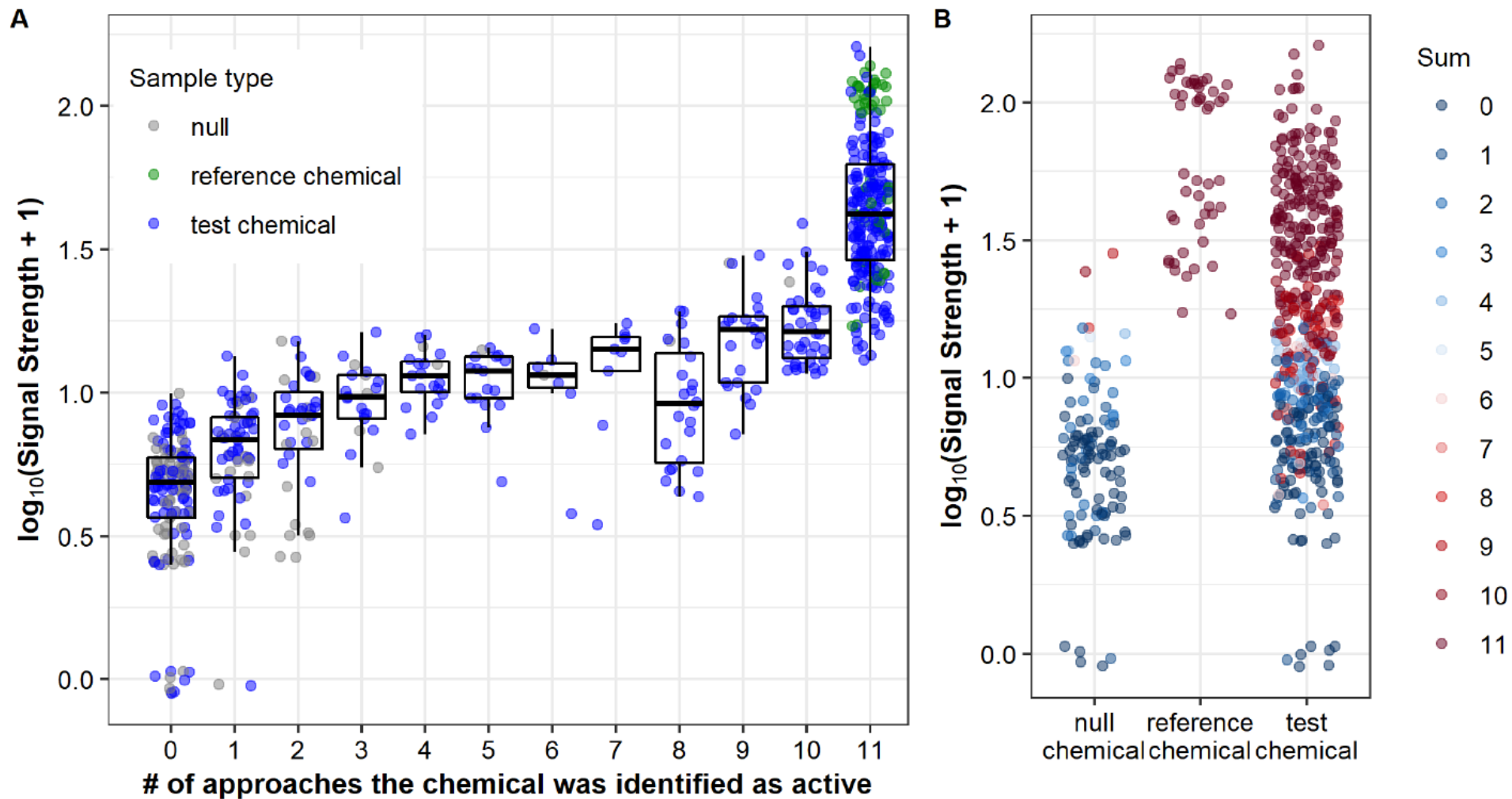
⇒ **87% of null chemicals were inactive in 9 or more approaches**



⇒ **51% of test chemicals were active in 9 or more approaches**

⇒ **30% of test chemicals were inactive in 9 or more approaches**

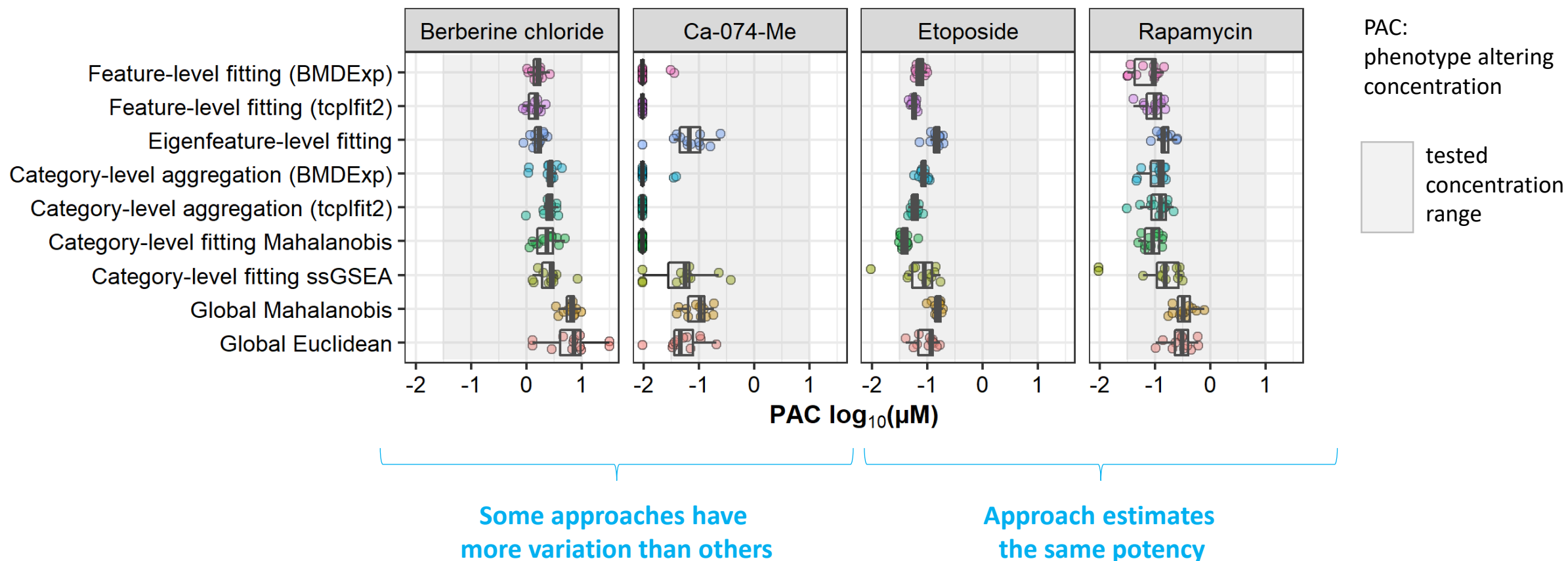
Concordance of hit call associated with signal strength



Nyffeler et al. 2020b

Concordance of potency estimates (I)

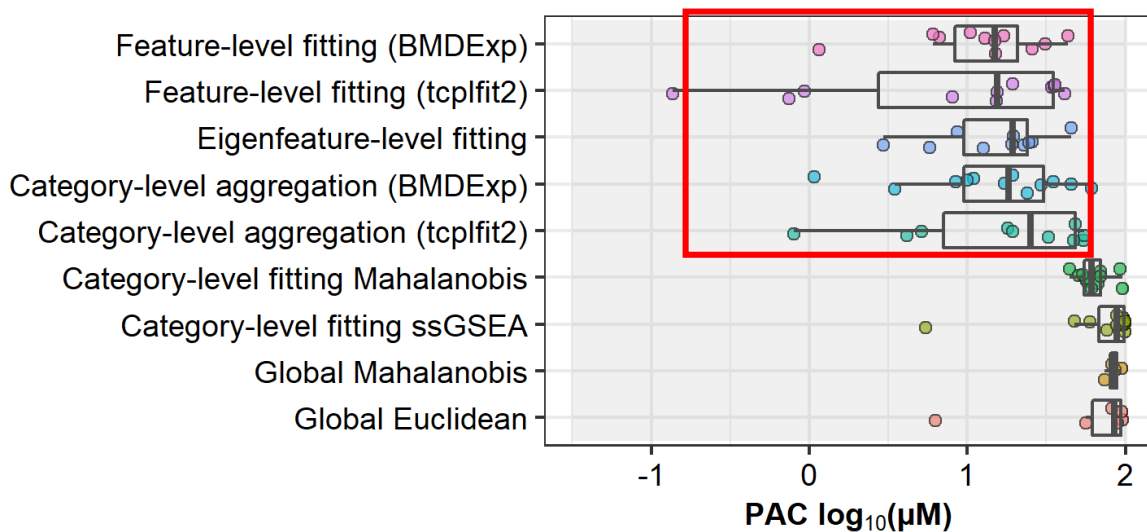
- 12 repetitions of each reference chemical
- *Does the approach always estimate the same potency?*



Concordance of potency estimates (II)

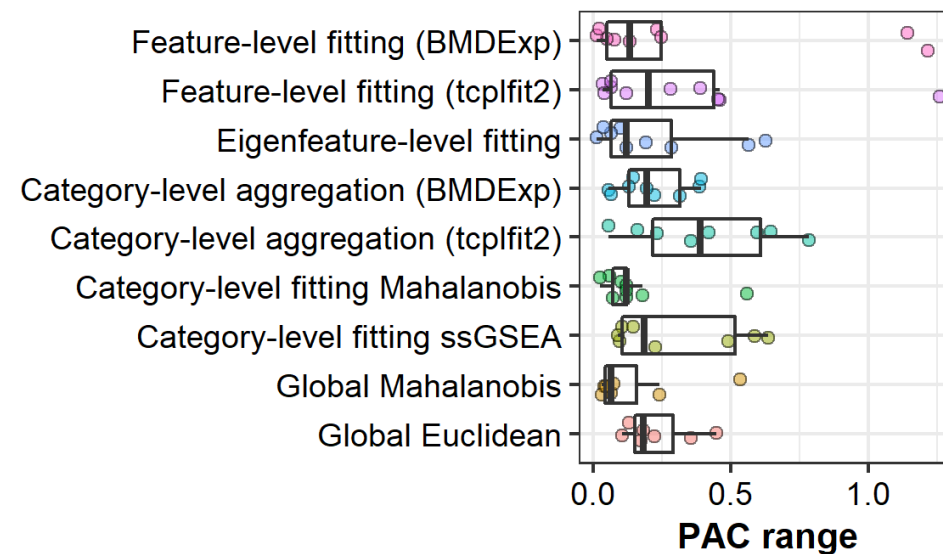
Null chemicals

Does the approach produce many high-potency false positives?



Duplicated chemicals

How much do the potencies of the replicates vary from each other?



⇒ **Feature-based approaches (including category-level aggregation) have a higher risk of false positive, highly potent results**

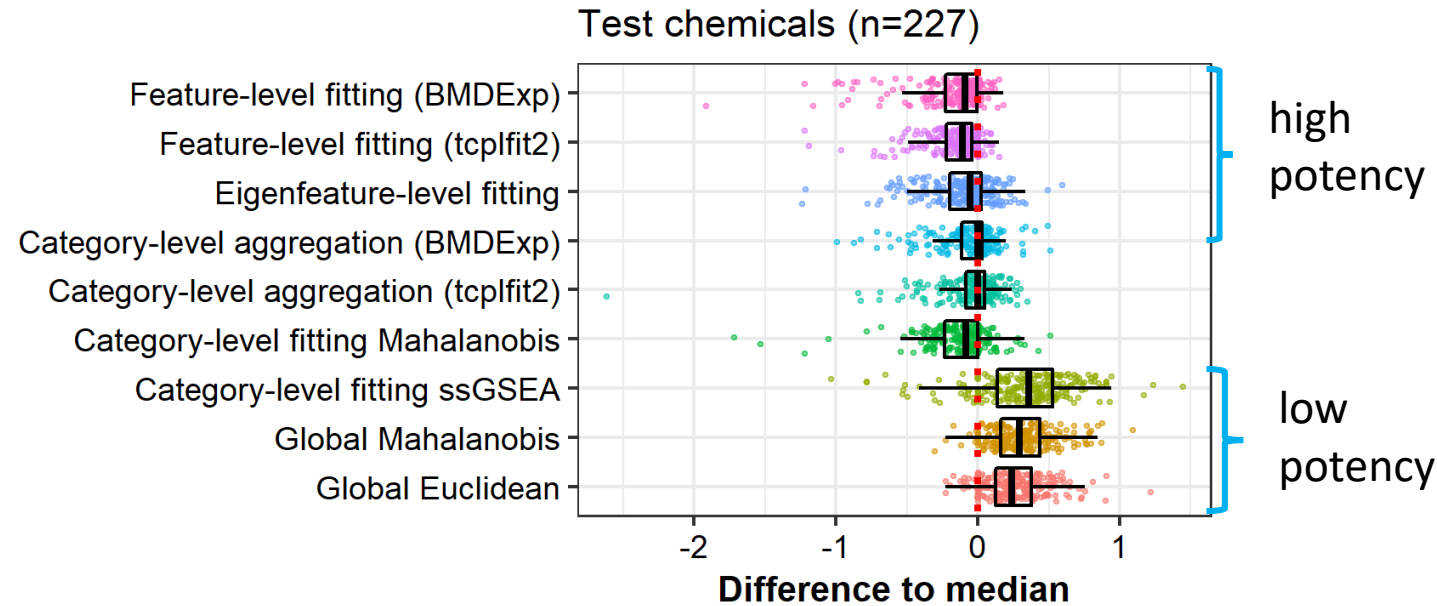
Concordance of potency estimates (III)

Test chemicals

For each chemical

1. Calculate the median potency across all 9 approaches
2. Calculate for each approach the difference to this median

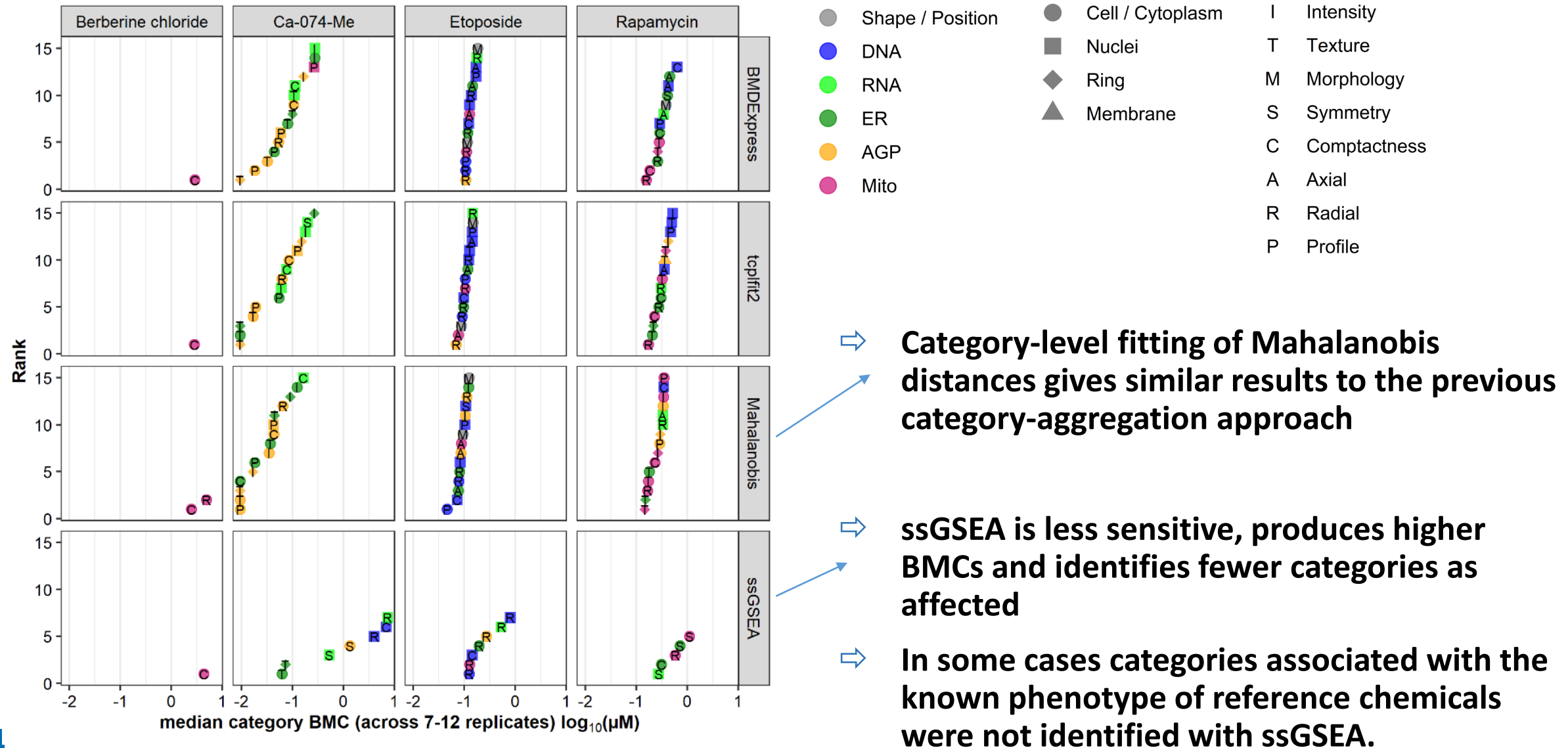
Is an approach rather underpredicting or overpredicting the potency of a chemical?



- ⇒ Feature-level approaches result in highest potency
- ⇒ Global fitting and ssGSEA result in lowest potency

Comparison of bioactivity profiles across category-based approaches

B Category-level approaches



Conclusions on individual approaches

- Feature-level methods:
 - high hit rate, but high risk of high-potent false positives
- Category-level aggregation:
 - alleviate the problem of highly potent false positives slightly
- Category-level fitting Mahalanobis:
 - Worked surprisingly well!
- Category-level fitting ssGSEA:
 - Was not sensitive in picking up one of the reference chemicals; gives lower potencies
- Global Mahalanobis:
 - Computationally fast (only 1 curve modelled), worked well
- Global Euclidean:
 - Computationally fast + simple, but low sensitivity (TPR), low concordance



Conclusions II

- Reanalysis of the APCRA data set increased our confidence in the data
- Constructing a null data set is useful to evaluate method performance
- All approaches had a similar hit rate (50-70%)
- For 81% of chemicals, at least 9/11 approaches agreed
→ we have high confidence in the resulting hit calls

Next steps:

- Analyze a screen of ~1200 chemicals with the category-level Mahalanobis and global Mahalanobis approaches.

Thank you for your attention!

Questions?

Nyffeler.Johanna@epa.gov