



www.epa.gov

Toxicological LINCS to Stress: Measuring Stress Response Pathways in Transcriptomic Data

Bryant A. Chambers and Imran Shah

Center for Computational Toxicology and Exposure, US EPA



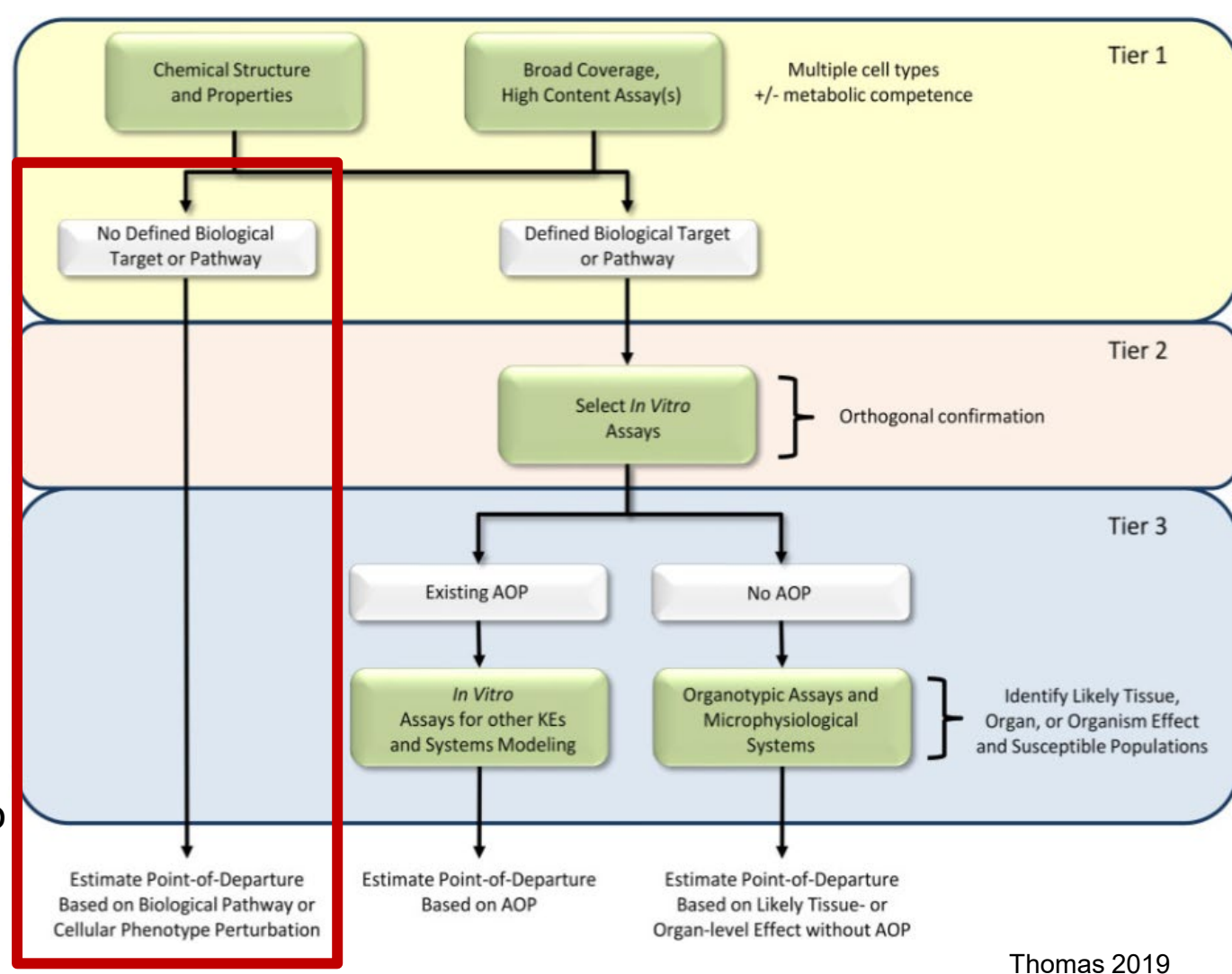
The views expressed in this presentation are those of the author[s] and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency

Bryant A. Chambers | chambers.bryant@epa.gov | 919.541.4268

Background and Hypothesis

Rationale

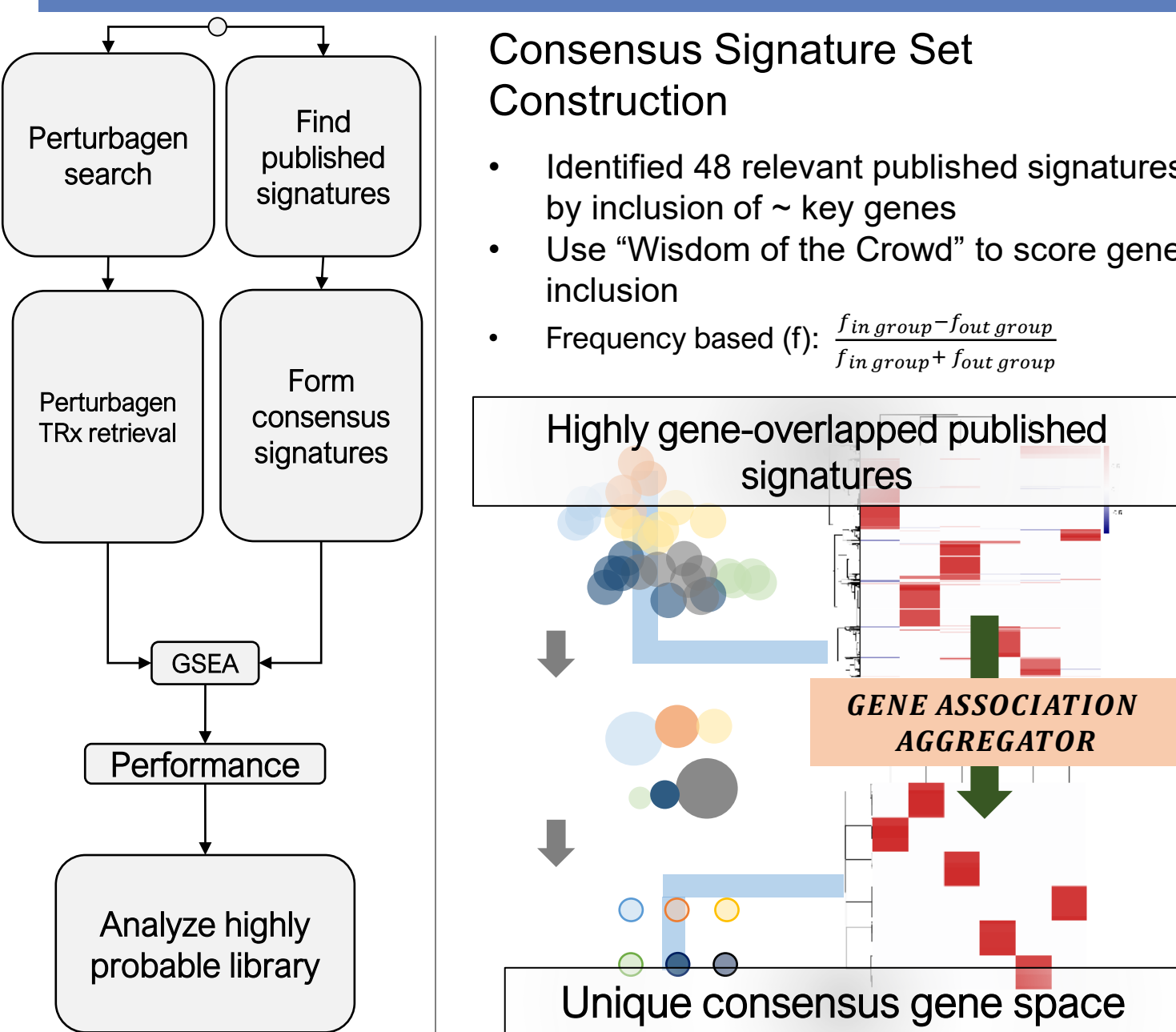
- Many environmental chemicals act via non-specific mechanisms: they do not activate molecular initiating events (MIEs) and cannot be related to adverse outcomes (Ankley 2010).
- Non-specific chemicals can disrupt homeostasis and result in DNA damage, misfolded proteins, hypoxia, and altered cellular reductive potential (Simmons 2009, Judson 2016).
- Cells sense disruption and induce stress response pathways (SRPs) to restore homeostasis.
- With the availability of high-throughput transcriptomics (HTTr) data, it may be feasible to characterize non-specific chemicals by SRP activation.



Hypothesis:

- Gene signatures can quantify SRP activation and provide insight about chemical hazards

Constructing Unique Stress Signatures



Consensus Signature Set Construction

- Identified 48 relevant published signatures by inclusion of ~ key genes
- Use "Wisdom of the Crowd" to score gene inclusion
- Frequency based (f): $f = \frac{f_{in\ group} - f_{out\ group}}{f_{in\ group} + f_{out\ group}}$

Reference Chemicals

- 49 reference chemicals were identified in the literature
- 2000 GSEs found in the Gene Expression Omnibus matching
- Only 35 GSEs proved sufficiently annotated

SRP	Abbreviation	Reference Chemical
DNA Damage Response	DDR	Benzo(a)pyrene Glycinamide Lasicoarpine Methylmethanesulfonate
Unfolded Protein Response	UPR	Brefeldin A Thapsigargin Tunicamycin
Heat Shock Response	HSR	Geldanamycin Heat with recovery Radicalol
Response to Hypoxia	HPX	Oxygen VU-0418946-1 VU-0418946-2
Metal Response	MTL	Silver Nitrate Zinc
Oxidative Stress Response	OSR	Hydrogen peroxide Tert-butyl hydrogen peroxide

Signature performance analyzed with receiver operator analysis (ROC)

Identifying best performing stress signatures

Consensus signatures are accurate

- 72% accuracy for highest scoring signature (88% for second)

Consensus signatures perform better than published signatures

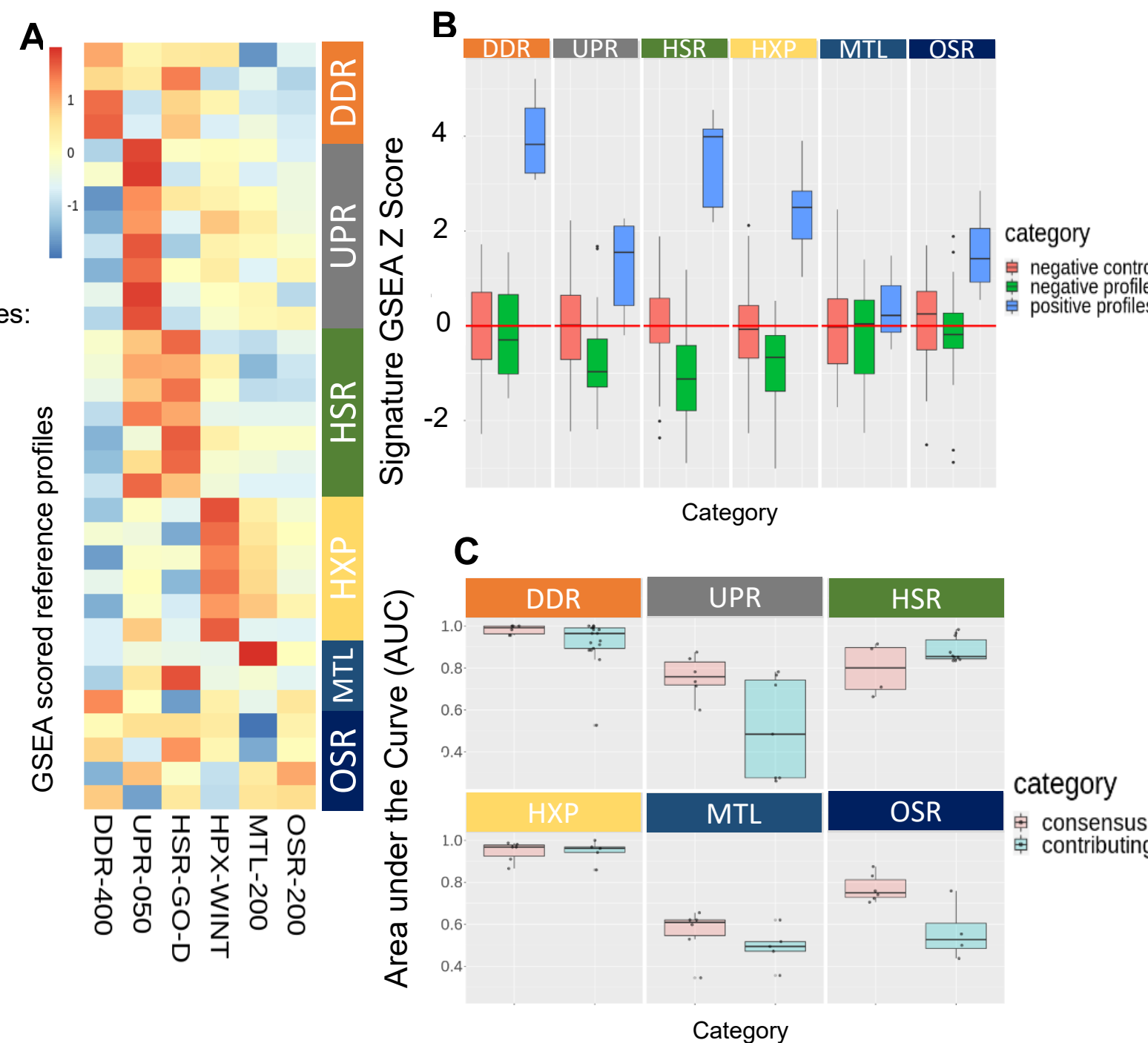
- % performance relative to median of published signatures:
DDR 0.99 (109%) HTS 0.91 (103%)
UPR 0.89 (172%) MTL, 0.66 (133%)
HPX 0.99 (104%) OSR 0.87 (155%)

SRP signal distinct from random null

- 4x SD for DDR, HSR, HPX
- OSR and UPR are least discernable
- Minimal overlap negative chemicals

Accuracy of Signatures in training data set

SR Pathway	Accuracy in 1st	Accuracy by 2nd	Accuracy by 3rd
DDR-400	75%	100%	100%
UPR-050	100%	100%	100%
HSR-GO	57%	100%	100%
HPX-WINT	100%	100%	100%
MTL-200	33%	33%	67%
OSR-200	25%	50%	50%
Mean Accuracy	72%	88%	91%

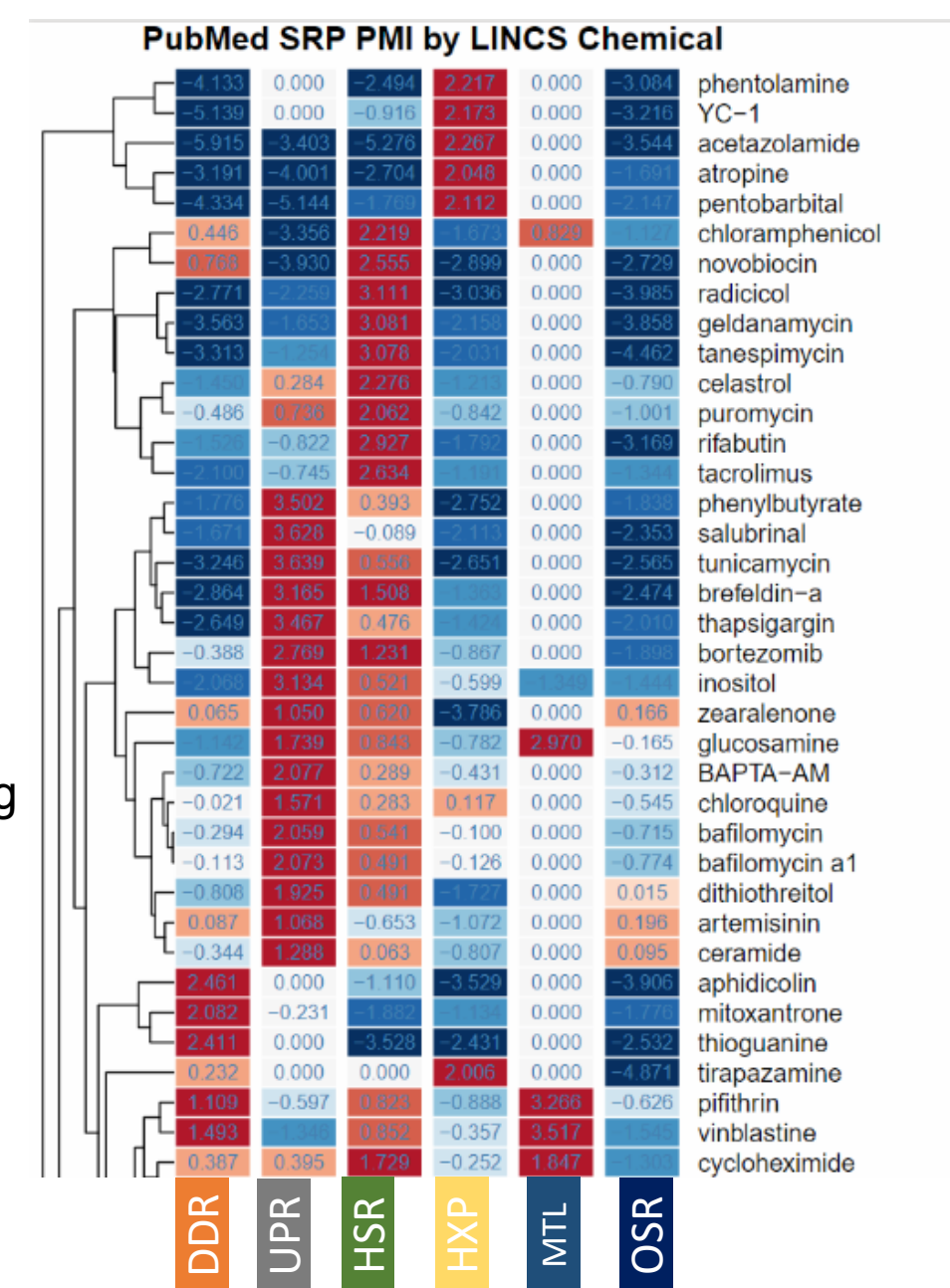


A) Activity of SRPs in training data set B) Performance of SRP signatures against randomized null C) Performance of SRP signatures against published signatures

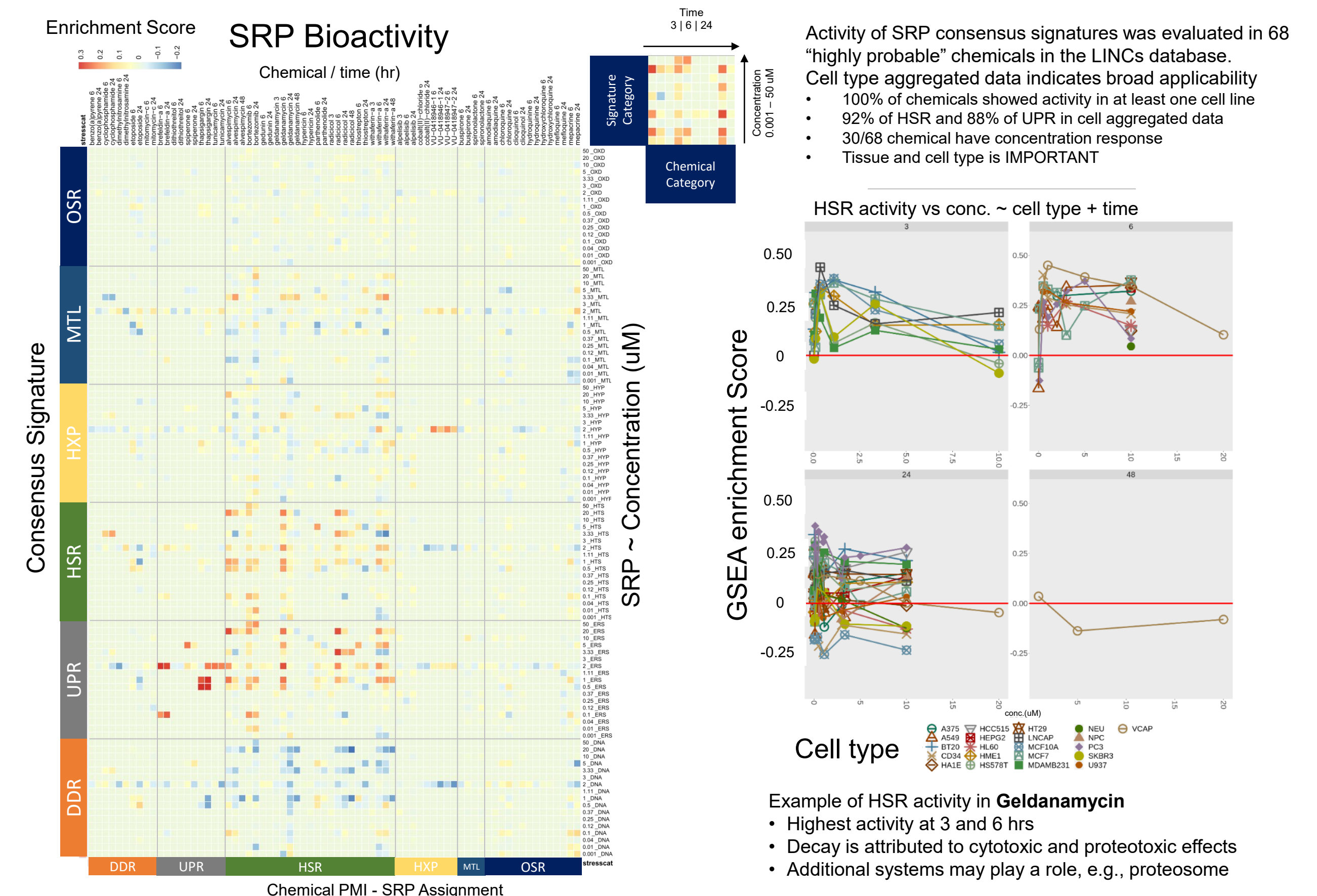
Finding a test set within the Library of Integrated Network-Based Cellular Signatures (LINCS)

Highly Probable Library of Integrated Network-Based Cellular Signatures (LINCS)

- Searched all LINCS chemicals in pub med against stress systems
 - 7 terms (e.g., 'DNA damage', 'er stress', unfolded protein response')
 - Totaled 32,679 searches
- Pairwise mutual information (PMI) score calculated for each chemical
 - $PMI(chemical, stress) = \log \frac{F(chemical, stress)}{F(chemical)F(stress)}$
- Filtered to chemical with PMI > 1 and references >200 yielding a subset of 90 perturbagens
- Pulled first 500 abstracts for each match and validated 68 high probability chemicals by review
- Transcriptomic Dataset includes:
 - 11000 Profiles
 - 81 cell types



Stress signature activity within LINCS test set



Example of HSR activity in Geldanamycin

- Highest activity at 3 and 6 hrs
- Decay is attributed to cytotoxic and proteotoxic effects
- Additional systems may play a role, e.g., proteasome

Conclusions and future directions

Key Conclusions

- SRP can be accurately identified using gene signatures in transcriptomic data.
- Distinct signals arise from well known chemical agents.
- Consensus SRP signatures perform better than published signatures.
- Consensus SRP signatures activity identified in LINCS data test set.

Future Directions

- Concentration-response analysis to estimate benchmark concentrations (BMC)
 - The lack of available methods to test non-specific chemical points of departure can be expanded to include a new approach method derived from general stress response transcriptomic assays.
- Tissue and cell specificity
 - Expression of stress response systems is partially dependent on cell and tissue type; as such, a deeper understanding of tissue dependency must be achieved.

References

- Ankley (2010) *Environmental Toxicology and Chemistry*, 29: 730-741
- Simmons (2009) *Toxicological Sciences* 111(2): 202-225
- Judson (2016) *Toxicological Sciences* 152(2):323-339
- Shah (2016) *Environ Health Perspect.* 124(7):910-9
- Thomas (2019) *Toxicological Sciences* 169(2):317-332
- Stathias (2020) *Nuc. Acids Res.* 48(D1):431-439