High-throughput phenotypic profiling (HTPP) to discern putative mechanism-of-action (MOA) for environmental chemicals

Johanna Nyffeler[1,2], Clinton Willis[1,3], Grace Patlewicz[1], Imran Shah[1], Joshua Harrill[1]

[1]US Environmental Protection Agency, Center for Computational Toxicology & Exposure, Office of Research and Development, Research Triangle Park, NC.

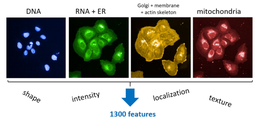
[2]ORISE, Oak Ridge, TN. [3]ORAU, Oak Ridge, TN

ORCiD 0000-0002-6155-9743 | Nyffeler.Johanna@epa.gov

www.epa.gov

Introduction

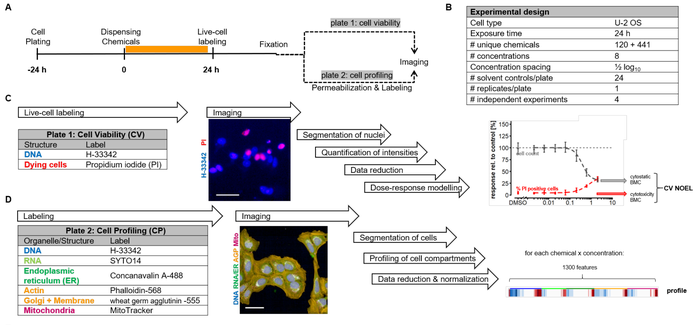
What is phenotypic profiling?



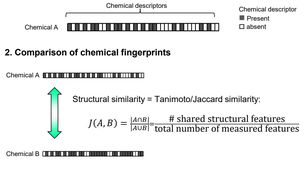
- Image-based phenotypic profiling is a chemical screening method that measures a large variety of morphological features of individual cells in *in vitro* cultures.
- No requirement for *a priori* knowledge of molecular targets.
- **May be used as an efficient and cost-effective method for evaluating chemical bioactivity.**

Methods

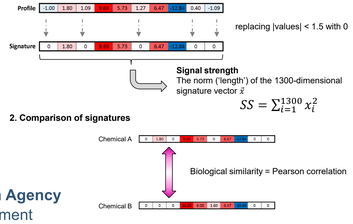
High-throughput phenotypic profiling (HTPP)



Calculation of structural similarity



Calculation of biological similarity



Goal

120 reference chemicals with annotated mode-of-actions

Can we find putative mechanism-of-action for environmental chemicals based on similarities in phenotypic response?

441 environmental chemicals
75% pesticides, drug-like, food additives, industrial chemicals, ...
(Nyffeler *et al.* 2020)
https://comptox.epa.gov/dashboard/chemical_lists/HTPP2019_SCREEN

Category	# of chems
1 actin skeleton modulator	
1.1 destabilizer	4
1.2 stabilizer	2
2 ER modulator	4
3 Golgi modulator	4
4 Mitochondrial fission	
4.1 inhibitor	2
4.2 enhancer	1
5 Microtubuli modulator	
5.1 Colchicin-like	4
5.2 Fenbendazole-like	9
5.3 Taxol-like (stabilizer)	3
6 DNA toxicants	
6.1 alkylators	10
6.2 topoisomerase	8
6.3 antimetabolites	7
6.4 DNA damage	3
6.5 DNA replication inhibitor	1
7 Oxidative stress	4
8 Protein homeostasis	
8.1 Proteasome inhibitor	7
8.2 Protein synthesis inhibitor	2
8.3 RNA polymerase inhibitor	1
9 uncoupler	1
10 mitochondria complex inhibitor	10
11 autophagy	
11.1 inhibitor	4
11.2 activator	7
11.3 Rapamycin-like (activator)	4
12 cell shape modulator	1
13 Berberine chloride group	2
14 Tetrandrine group	2
15 - 20 individual RefChem16 not contained in previous categories	6
21 negative controls	5

Conclusions

1. **Diverse phenotypic profiles were observed across the entirety of the chemical set.**
2. **Biological similarity was measured reproducibly, but only for chemicals that are bioactive (i.e. have a high signal strength)**
3. **Chemicals with shared mechanism-of-action often had a similar profile**
4. **Chemicals with high structural similarity often share the same biological signatures.**

Results: reference chemicals

Optimization of biological similarity calculation

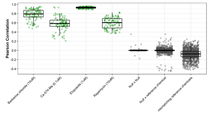


Fig 1.: Biological similarity of QC chemicals. A signature threshold of 1.5 was used. QC reference samples were tested in 15 replicates and compared against each other (green). Comparisons in gray indicate inactive (null) chemicals against each other or mismatching samples.

A range of signature thresholds (0 – 6) and four correlation methods were evaluated.

⇨ A signature threshold of 1.5 and Pearson correlation were used for this study

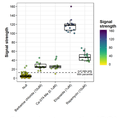


Fig 2.: Signal strength of QC chemicals. A signature threshold of 1.5 was used. QC reference samples were tested in 15 replicates. Signal strength was defined as the norm of the (1300 dimensional) signature vector.

Signal strength (SS) is a measure for the effect size of a condition.

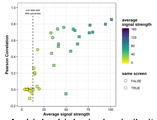


Fig 3.: Signal strength and biological similarity of test chemicals screened in duplicates. A signature threshold of 1.5 was used. Biological similarity was calculated by Pearson correlation. Average signal strength was defined as the mean of the signal strengths of the two duplicates.

A high biological similarity can be measured for chemicals screened in duplicate, but only if they have higher signal strength than null chemicals.

Signatures of reference chemicals

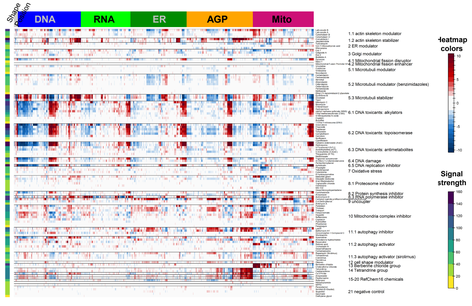


Fig 4.: Signatures of 120 reference chemicals. Chemicals were manually grouped by their known mechanism-of-action. For each chemical, data from the highest non-cytotoxic concentration is displayed. Signatures were generated by flooring all absolute values < 1.5 to 0. Features (in columns) are ordered according to the corresponding channel/organelle. The color key on the left indicates overall signal strength of the corresponding chemical.

⇨ Different signatures are observed

⇨ Different classes of DNA toxicants (group 6) share similar signatures

Finding biological and structural analogues

Example: Benzimidazoles (group 5.2)

The 9 benzimidazoles were used as a 'seed' to retrieve all chemicals with a structural similarity > 0.25 (in green) or biological similarity > 0.6 (in pink):

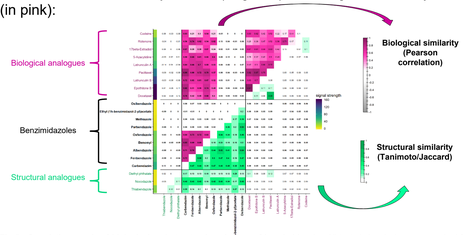


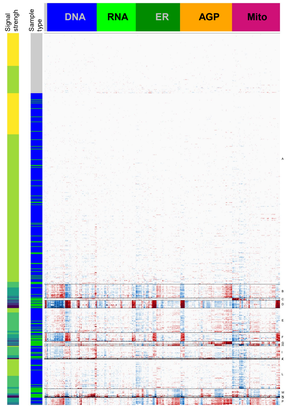
Fig 5.: Correlation matrix of biological and structural analogues of benzimidazoles All tested chemicals were searched for structural and biological similarity to any of the benzimidazoles. Chemicals with a structural similarity > 0.25 or biological similarity > 0.6 to any benzimidazole are displayed. The upper left half of the correlation matrix displays biological similarity (as pearson correlation), while the lower right half of the matrix displays structural similarity (measured as Tanimoto similarity).

⇨ All benzimidazoles are structurally similar, but only 5 had biological similarity (the other 4 have low signal strength)

⇨ Among the biological analogues are microtubule stabilizer (group 5.3) as well as actin cytoskeleton modulators (group 1.1).

Results: environmental chemicals

K means clustering of reference & environmental chemicals



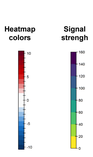


Fig 6.: K means clustering of all chemicals. For each chemical, data from the highest non-cytotoxic concentration was used to generate a signature by flooring all absolute values < 1.5 to 0. Features (in columns) are ordered according to the corresponding channel/organelle. The number of clusters was chosen so that visually different signatures were in different clusters, but replicates of the same chemical were in the same cluster (not shown).

⇨ Approximately 16 signature clusters are observed

⇨ 300/441 environmental chemicals clustered with the null data sets (i.e. have no distinctive signature at the highest non-cytotoxic concentration)

⇨ The remaining environmental chemicals mostly shared signatures with reference chemicals

Examples

Example: Strobilurins

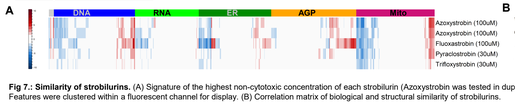


Fig 7.: Similarity of strobilurins. (A) Signature of the highest non-cytotoxic concentration of each strobilurin (Azoxystrobin was tested in duplicate). Features were clustered within a fluorescent channel for display. (B) Correlation matrix of biological and structural similarity of strobilurins.

⇨ Strobilurins have similar signatures with many mitochondrial features affected.

Example: Dieldrin

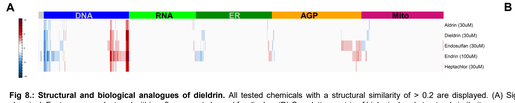


Fig 8.: Structural and biological analogues of dieldrin. All tested chemicals with a structural similarity of > 0.2 are displayed. (A) Signature of the highest non-cytotoxic concentration of each chemical. Features were clustered within a fluorescent channel for display. (B) Correlation matrix of biological and structural similarity.

⇨ Four structural analogues to dieldrin displayed high biological similarity with dieldrin, with changes in the DNA channel.

Future Directions

- Redundant features might distort biological similarity measurements
  → evaluate feature reduction and feature selection approaches prior to similarity calculations

This work does not reflect USEPA policy. Mention of tradenames or products does not represent endorsement for use.