High-throughput screening & toxicogenomics space testing concept

HTS-PTGS team: Pekka Kohonen[1,2,3], Vesa Hongisto[2], Penny Nymark[1], Juha Rantala[2] & Roland Grafström[1,2,3]
[1]Karolinska Institutet, Institute of Environmental Medicine, [2]Misvik Biology OY, [3]Predictomics AB

TIER 1

Experimental route

Results

1) High-throughput screening data analysis with primary human hepatocytes (PHH) and cell lines (HepaRG, HepG2 etc.) of the 150 EPAA test chemicals: 384-well plates (1000 cell/well), three end points (cell number, ATP content and apoptosis) 6h, 24h and 72h time points, 8 concentrations (three orders of magnitude range), four biological replicates

Dose response and initial cytotoxicity evaluation ($50x10^3$ data points / cell model)

↓

2) Point-of-departure (PoD) analysis of HTS data (9 methods): perform LOEL/NOEL analysis with t-test (q<0.05) at 10%, 15%, 25% and 50% effect level (2SD & 3SD), benchmark dose (BMD/BMDL) analysis with BMDExpress 3/ToxicR at 10%, 15%, 25% and 50% benchmark response (BMR); AC50 analysis with Hill's equation, including 5% & 95% CIs

4 BMDs, 4 LOELs, AC50

↓

3) Ranking analysis with the HTS data (10 methods): apply the LOEL and BMD PoD measures; utilize Tukey's trend test for ranking of overall activity level

4 BMD ranks, 4 LOEL ranks, AC50 & Tukey's test ranks

↓

4) Grouping analysis with HTS (22 methods): apply the LOEL and BMD PoD measures, as well as t-statistics and Tukey's trend test results across the dose series using hierarchical clustering with bootstrap validation (pvclust) or kmeans clustering with gap statistics

11 groupings with hierarchical clustering, 11 with kmeans

↓

5) Initial Low, Medium, and High toxicity ranking with Tier 1 data by data integration: perform Meta-PoD and meta-ranking analysis with the nine PoDs and ten ranking methods. Assess ranks and lowest/mean/median PoDs across different end points, time points and cell types. Meta-grouping (22 methods) analysis defines consensus groups. Define low, medium and high toxicity hazard chemicals by potency rank and assign toxicity class to derived chemical groupings

9 PoDs, 10 rankings and 22 groupings/clustering

TIER 2

6) High-throughput transcriptomics analyses: concentrations selected based on the HTS data. Apply high-throughput 3'-RNA-seq to assay the full transcriptome under the exact culture protocol (number of cells, well volume, etc.) used for the HTS assessment

DEGs, initial functional analyses (1800 RNA-seq profiles & $36x10^6$ data points / cell model)

↓

7) Predictive Toxicogenomics Space (PTGS) analysis of cytotoxicity and hepatotoxicity: derive PTGS scores, and cytotoxicity and organ toxicity/hepatotoxicity LOELs. Absence of PTGS activation indicates a low-toxicity chemical

2 PTGS predictive scores, PTGS-derived LOELs

↓

8) PTGS component-based MoA and functional classification: utilize competitive gene set testing to derive patterns of PTGS components that serve to delineate toxic MoAs and grouping

14 PTGS component MoA scores

↓

9) Programmatic high-sensitivity adverse outcome pathway (AOP) analyses (4 methods): map PTGS component MoA to AOP/event level in 33 liver AOPs within the AOPWiki. Derive quantitative/qualitative outputs, statistical significance (p-values) and AOP-level statistics. Distinct GO-based qAOP analysis methods serve as controls relative the PTGS concept

4 AOP-level analyses, 4 event-level analyses

↓

10) High-specificity AOP analyses with PTGS component-based correlation analyses (3 methods): perform correlation analysis of PTGS components versus PTGS/GO-annotated AOP events. Modulate specificity from using GO-terms (2000, 1500 or 1000 genes per term)

3 AOP-level analyses, 3 event-level analyses

↓

11) AOP-based grouping (8 methods): Apply kmeans grouping to results from PTGS component activation and the high-sensitivity and high-specificity AOP analysis methods

1 PTGS-, 4 high-sensitivity AOP- and 3 high-specificity AOP grouping analyses

↓

12) Integration across the tier 1 and tier 2 results to derive statistically supported functional groups of chemicals that can be assigned status for Low, Medium, and High toxicity. Apply machine learning to reduce the dimensionality of the testing protocol: Combined ranking and grouping by tier 1 and tier 2 methods, annotate meta-groupings of similar chemicals with toxicity potency class (low, medium and high) with both qualitative and quantitative output; Tier 1 assessment produces ~$100x10^3$ and Tier 2 ~$80x10^6$ derived data points. Absence of PTGS score and component activation indicates a low-toxicity chemical whereas medium and high-toxicity chemicals are assigned based on relative potencies, taking direction eventually from reference chemicals with defined classes.

Uncertainty is modelled as part of the dose response analysis process: BMD analysis includes BMDL and BMDU values that correspond to the confidence intervals modelled by the AC50 analysis process. LOEL/NOEL analyses model uncertainty via p-values and variance analysis. Additional methods that reduce uncertainty are outlier detection and robust summary statistics for PoDs across methods, end points and cell lines (median, weighted means). Multiple types of PoD and MoA analyses counteract uncertainty and map the range of possible responses over the dose-response curves.

The broadness of the PTGS gene cloud and multiple coupled derived assessment concepts ensures high coverage of existing cellular toxicity mechanisms relevant to pathological states. Ultimately, evaluation of the rigorous testing protocols by machine learning approaches should lead to simplification and definition of the key needed assessments for defining the toxic potency classes.

Karolinska Institutet

misvik biology

Predictomics AB

References
Kohonen P, Parkkinen JA, Willighagen EL, Ceder R, Wennerberg K, Kaski S, Grafström RC. A transcriptomics data-driven gene space accurately predicts liver cytopathology and drug-induced liver injury. Nat Commun. 8:15932 (2017)