Prioritization of Chemicals for Effects on Steroidogenesis Using an Integrated Statistical Approach to High-throughput H295R Data

Derik E. Haggard[1,2], R. Woodrow Setzer[2], Richard S. Judson[2], Katie Paul Friedman[2]

[1]ORISE Postdoctoral Research Participant [2]National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency

Derik E. Haggard | haggard.derik@epa.gov | 919-541-0745

www.epa.gov

Introduction

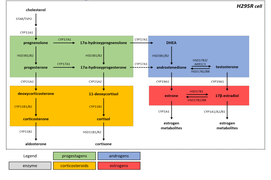
Background:

- Endocrine disruption is a toxicity of both physiological and regulatory importance; as steroid hormones regulate reproduction, development, and other biological processes, it is a priority to identify chemicals that may interact with production of these hormones.
- A high-throughput H295R assay (HT-H295R) was developed as part of the U.S. EPA's ToxCast program that includes measurement of 11 hormones across the steroid hormone biosynthesis pathway in H295R cells, including progestagens, corticosteroids, androgens, and estrogen.
- HT-H295R has been used to screen a total of 2012 chemicals in single-concentration and 656 chemicals in multi-concentration.

Mahalanobis Distance Metric:

- We previously developed a statistical metric using the mean Mahalanobis distance (mMd) to quantify the effect of a chemical on the overall steroidogenesis pathway in HT-H295R.
- Mahalanobis distance incorporates the effect size for each steroid hormone measure after adjusting for covariance between the steroid hormone measures (Equation 1).
- The maximum mMd (maxmMd) value for each chemical was selected to indicate the magnitude of steroidogenesis pathway perturbation.

Steroidogenesis in H295R Cells



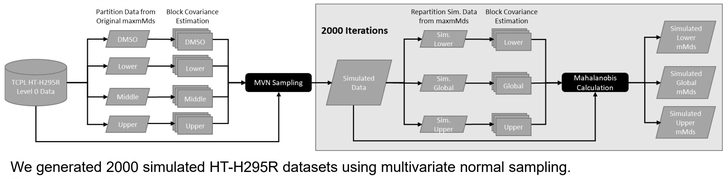
Eq. 1 $$mMD = \sqrt{(\mathbf{y_c} - \mathbf{y_1})' \boldsymbol{\Sigma}^{-1} (\mathbf{y_c} - \mathbf{y_1}) / N_h}$$

Where y_j is the vector of log-transformed hormone concentrations for the j^{th} concentration, N_h is the number of hormones with measurements for this chemical, and $\sum$ is the estimate of the covariance matrix

Primary Questions:

- Is the covariance structure stable when you only use low or high responding chemicals for the mMd calculation *(i.e.* if the dataset changed), and how does this affect the mMd values?
- How well did we approximate the type I error rate, and what is the sensitivity of the mMd approach to detect significant perturbations at different effect sizes and steroid hormone response types?
- Do putative aromatase inhibitors have a unique steroid hormone pattern in HT-H295R?
- Can HTS markers of cytotoxicity and mitochondrial toxicity be used to contextualize the mMd as a prioritization metric?

Data Simulation to Test Stability of the Approach



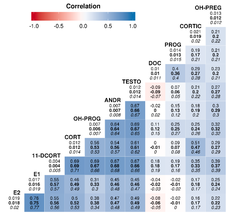
We generated 2000 simulated HT-H295R datasets using multivariate normal sampling.
For each simulated dataset:

- Estimate three different covariance matrices based on weak responders (lower), all data (global), and strong responders (upper).
- Derive new mMd values for a total of three sets of mMds to compare.

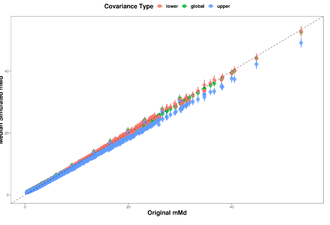
The Mean Mahalanobis Distance (mMd) Approach is Stable

Changes in covariance matrices from data simulation suggest that the covariance structure is stable, leading to reproducible mMds.

A.



B.



A. Comparison of the correlation of the three estimated covariance matrices.
- Correlation of the covariance of steroid hormone measures are largely the same.
- Variance of the steroid hormone measures is also similar.

B. Scatterplot of the difference between the original mMd and the median of the simulated mMd for the three covariance types.
- Practically no difference was observed between the simulated mMds and the original mMds (dashed identity line).

What is the true type I error rate (false-positive rate) and the power to detect fold-changes in steroid hormones?

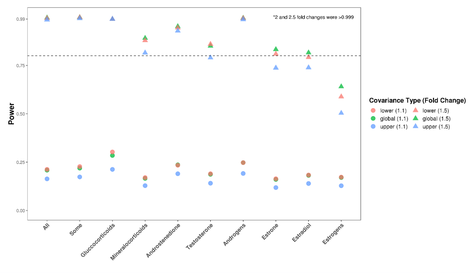
A.

	Covariance Type		
Error Rate	**Lower**	**Global**	**Upper**
0.05	0.015	0.019	0.012
0.01	0.007	0.008	0.006

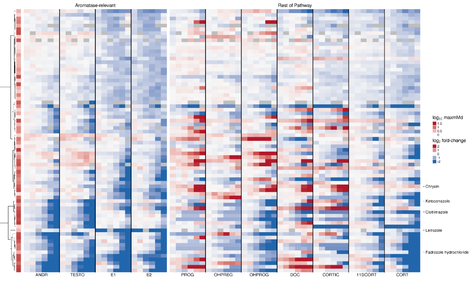
A. Estimated type I error rate of a simulate true negative HT-H295R response profile at two α levels.

B. The power to detect experimental significance using the mMd approach.
- Fold-change effect sizes ≥1.5-fold resulted in sufficient power to detect significant effects on steroidogenesis.
- Power was smallest for steroid hormone changes in both E1 and E2.

B.



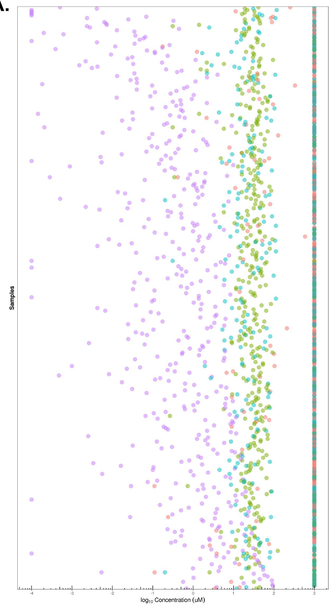
Steroid Hormone Profiling of Putative Aromatase Inhibitors



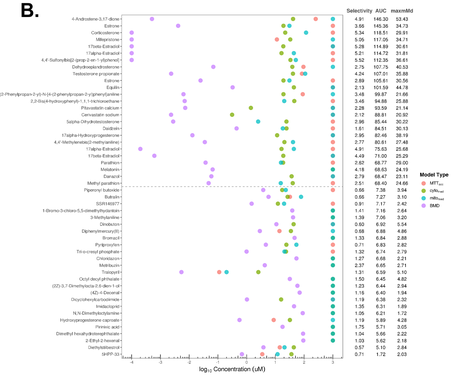
Steroid response heatmap for putative aromatase inhibitors.
- 72 out of 766 chemical samples had a putative aromatase inhibitory signature (1.5-fold decreases in both E1 and E2).
- Hierarchical clustering of androstenedione, testosterone, estrone, and estradiol responses identified four predominant clusters.
- Known aromatase inhibitors decrease estrogens *and* androgens but have myriad effects on upstream steroid hormone patterns.

Development of a HT-H295R Prioritization Metric

A.



B.



We developed a rank prioritization metric using a selectivity score which accounted for cytotoxicity and mitochondrial toxicity from ToxCast assays.

A. 432 of 635 (68%) chemical samples with sufficient ToxCast assay data were considered selective in HT-H295R and were ordered by efficacy in the mMd response.

B. Top and bottom 25 most selective and efficacious chemicals samples in HT-H295R
- Most selective chemicals included steroid hormones and known steroidogenesis disruptors

Summary and Future Directions

- The mMd appears to be robust to changes in the covariance matrix derived from different effect size partitions of HT-H295R chemical space.
- The observed type I error rates were conservative and similar to the approximated values used in the critical threshold calculation, suggesting that the hit-call method used for the mMd approach is appropriate and protective.
- An effect size of a 1.5-fold increase in hormone levels results in a maxmMd response value greater than the critical threshold with a power of 0.8 or higher, except for changes in both estrogen steroid hormones.
- The HT-H295R assay as implemented can identify aromatase inhibitors, but the myriad responses across the majority of the steroid hormone measures for these chemicals suggests that proper mechanistic evaluation requires multiple timepoints to better model the complex and dynamic steroidogenesis pathway present in H295R cells.
- Removing chemical samples that had BMDs within the cytotoxicity/mitochondrial toxicity 'burst' range using the selectivity score, and then ordering by the AUC estimates, appeared to efficiently rank active chemicals with steroid hormones and known steroidogenesis disrupts ranked the highest..

This poster does not reflect US EPA policy. D.E.H. was supported by appointment to the Research Participation Program of the U.S. Environmental Protection Agency, Office of Research and Development, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. EPA.