Reproducibility and Reliability of High-Throughput Transcriptomics for Chemical Safety Screening

Logan J. Everett, Derik Haggard, Joseph Bundy, Beena Vallanat, Joshua A. Harrill, Richard Judson, Imran Shah

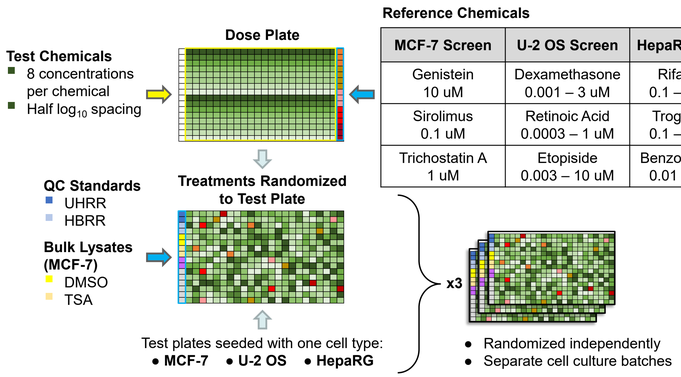
U.S. Environmental Protection Agency, Office of Research and Development, Research Triangle Park, NC, USA

Logan J. Everett | Everett.Logan@epa.gov | 919-541-2495

Background

- New Approach Methodologies (NAMs) aim to replace vertebrate animal testing for chemical safety screening and assessment [1].
- U.S. EPA has proposed a tiered testing strategy using NAMs to broadly identify hazards from chemical exposure and characterize their dose-response relationships, starting with NAMs that are high-throughput and provide broad biological coverage for the first tier of testing [2].
- Targeted RNA-seq of cultured human cells provides a platform for high-throughput transcriptomics (HTTr) screening that covers >20,000 genes and a wide array of biological responses and pathways [3].
- EPA has successfully piloted the targeted RNA-seq approach for HTTr to predict biological pathway altering concentrations (BPACs) using a scalable study design [4].
- Reproducibility and reliability of the HTTr platform are critical factors in its applicability to larger scale chemical safety screening and next generation risk assessment.

Design & Analysis of HTTr Studies



Reference Chemicals

MCF-7 Screen	U-2 OS Screen	HepaRG Screen
Genistein 10 uM	Dexamethasone 0.001 – 3 uM	Rifampicin 0.1 – 100 uM
Sirolimus 0.1 uM	Retinoic Acid 0.0003 – 1 uM	Troglitazone 0.1 – 100 uM
Trichostatin A 1 uM	Etopiside 0.003 – 10 uM	Benzo(a)pyrene 0.01 – 10 uM

Figure 1. Design of large HTTr screens. Upper left: Dose plates are prepared with ~40 test chemical samples at 8 concentrations (half $\log_{10}$ spacing, single replicates). Upper right: For each cell type, a standard set of reference chemicals are included on every dose plate. For MCF-7 cells, reference chemicals are plated at single concentrations in triplicate. For U-2 OS and HepaRG cells, reference chemicals are plated at 7 concentrations (half $\log_{10}$ spacing, single replicates). Lower left: Cells are seeded and expanded on the test plate, and then treated with chemical samples from the dose plate. Positions of each chemical treatment are randomized on every plate by an automated liquid handling system. The left column of each test plate is initially left blank, and QC reference standards are manually added to these wells on every test plate before transcriptomic profiling. Lower right: Each test plate is generated in triplicate with the same test chemicals, but with independent randomization of treatment positions and using separate culture batches.

Reliability of HTTr Assay

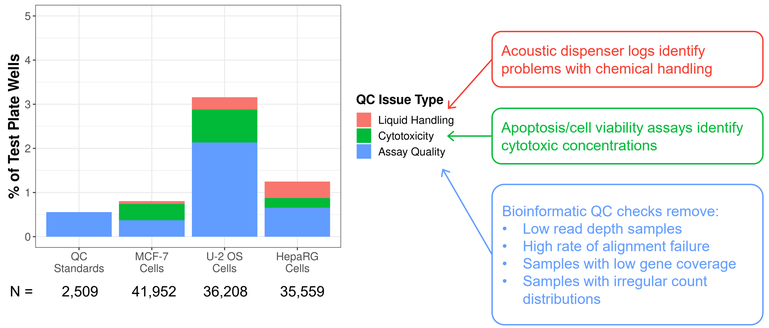


Figure 2. Quality control failure rates by sample and cell type. The QC process described in [4] was applied to >100,000 total samples from HTTr screening of >1,500 chemicals in three cell types. Individual HTTr profiles are excluded from further analysis due to errors in acoustic dispensing of chemicals (red), > 50% cell death indicated by cell viability assay (green), or abnormal parameters from a battery of bioinformatic checks applied to sequencing data results (blue).
Conclusion: >96% of samples pass all QC filters, demonstrating the ability to reliably scale this workflow to studies encompassing thousands of chemicals and samples.

Reproducibility of Differential Expression I

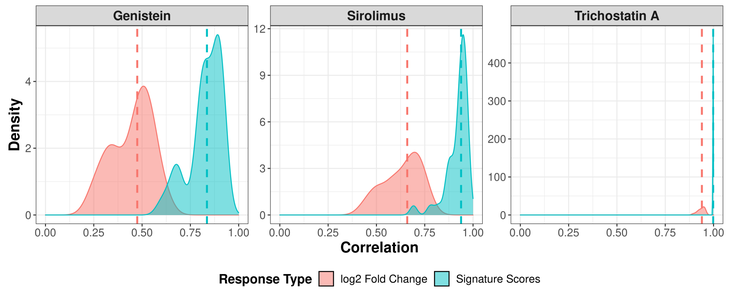


Figure 3. Correlation distributions of differential expression in response to reference chemical treatments. Log2 fold-change profiles (L2FC) against DMSO vehicle controls were computed using DESeq2 [5] for three standard reference chemical treatments included on each set of matching triplicate plates throughout MCF-7 screen (Total of 37 L2FC profiles, each based on 3 replicates x 3 plates). Signature scores were computed using the gene set catalog and method described in [4]. Correlations for each L2FC (red) and signature score (blue) profile were computed against the median profile of the same type and reference chemical. Vertical lines indicate median for each distribution.

Reproducibility of Differential Expression II

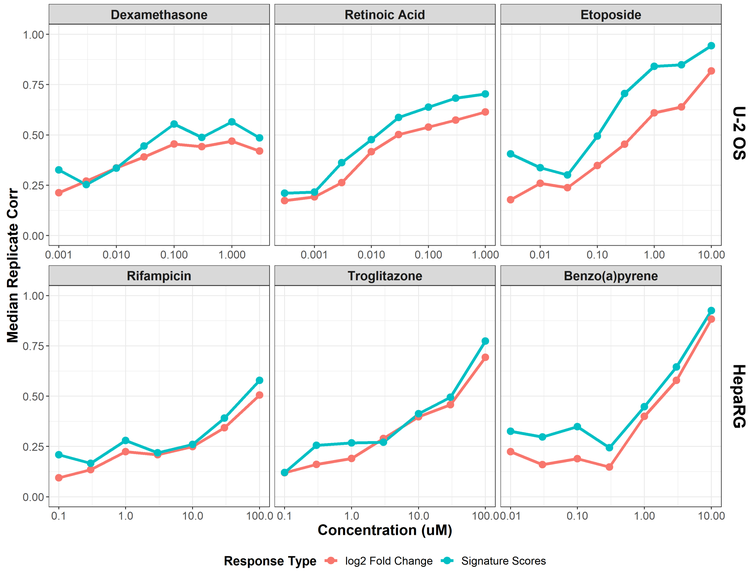


Figure 4. Median replicate correlation of reference chemicals in U-2 OS and HepaRG screens. L2FC and signature score profiles and correlations were computed as in Figure 3. The median correlation across 33 replicate experiments (each run in triplicate) is plotted as a function of concentration for each chemical. Low concentrations tend to be biologically inactive and therefore produce very little reproducible signal, while correlation between replicates tends to increase as a function of the increased transcriptional response at higher concentrations.

Summary & References

- Most samples pass rigorous QC tests, even when screening over 1,000 chemicals.
- Reference chemicals included throughout the large screening studies demonstrate the reproducibility of differential expression profiles for bioactive treatments.

1. 15 U.S.C. §2603(h): Reduction of Testing on Vertebrates
2. Thomas, et al. *The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency.* Toxicological Sciences 2019, 169(2):317-332
3. Yeakley, et al. *A Trichostatin A Expression Signature Identified by TempO-Seq targeted Whole Transcriptome Profiling.* PLoS One 2017, 12(5)e0178302
4. Harrill, et al. *High-Throughput Transcriptomics Platform for Screening Environmental Chemicals.* Toxicological Sciences 2021, doi:10.1093/toxsci/kfab009
5. Love, et al. *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biology 2014, 15(12):550

The views expressed are those of the presenter and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.