Tox21 Phase 3: High-Throughput Transcriptomics and the S1500+ Initiative

NTP Board of Scientific Counselors Meeting
December 08, 2017

Richard S. Paules, Ph.D.
Acting Chief, Biomolecular Screening Branch
Division of the National Toxicology Program
National Institute of Environmental Health Sciences
The Toxicology in the 21st Century ("Tox21") Federal Partnership

NTP Board of Scientific Counselors Meeting
December 14, 2016

Richard S. Paules, Ph.D.
Acting Chief, Biomolecular Screening Branch
Division of the National Toxicology Program
National Institute of Environmental Health Sciences

NTP Tox21 Phase III: Improving on Biological Coverage and Human Relevance
The Assumption:

- Global “Omic” (Whole System) approaches can link genomic perturbations with alterations in biological processes that result in toxicity and/or disease.

The Goal:

- A rapid and affordable quantitative High-Throughput Transcriptomic (HTT) measurement of expression levels of genes for large numbers of samples (chemicals x doses x times x cells/tissues)

  1. Low depth coverage, “whole transcriptome” targeted gene expression analysis of ~22k genes (EPA-led)

  2. Targeted gene expression analysis using a set of representative or “Sentinel” genes, S1500+, to determine pathway and network perturbations (NTP-led)
Outline

• Development of the S1500+ Gene Set

• Evaluation of Performance of the Human S1500+ Gene Set

• Evaluation of the S1500+ on the TempO-Seq HT-Transcriptomics (HTT) Platform

• HTT Proof of Concept Study with an In Vitro Human Liver Organotypic Model

• Future Directions
Tox21 S1500+ Gene Selection Workgroup Members

- Scott Auerbach, Biomolecular Screening Branch, DNTP, NIEHS
- Pierre Bushel, Biostatistics Branch, DIR, NIEHS
- Jennifer Collins, Exposure, Response & Technology Branch, DERT, NIEHS
- Agnes Forgacs, National Center for Computational Toxicology, US EPA
- David Gerhold, Genomic Toxicology Group, National Center for Advancing Translational Sciences (NCATS)
- Richard Judson, National Center for Computational Toxicology, US EPA
- Elizabeth Maull, Biomolecular Screening Branch, DNTP, NIEHS
- Deepak Mav, SciOme, Inc.
- Alex Merrick, Biomolecular Screening Branch, DNTP, NIEHS
- Rick Paules, Biomolecular Screening Branch, DNTP, NIEHS
- Ruchir Shah, SciOme, Inc.
- Dan Svoboda, SciOme, Inc.
- Donna Mendrick, National Center for Toxicological Research, US FDA
- Weida Tong, National Center for Toxicological Research, US FDA
- Rusty Thomas, National Center for Computational Toxicology, US EPA
Attributes of a Tox21 S1500+ Gene Set

1. **Diversity**: Capture the maximal expression variability and dynamics

2. **Co-Expression**: Capture the *Sentinel* genes with maximal co-expression information to represent members of nodes or networks

3. **Maximal Pathway Coverage**: Genes are included to ensure maximal biological pathway coverage

   1500 “*Sentinel*” Genes Informatic Derived – S1500

4. **Inclusion of toxicity and disease related genes**: Specific genes selected for their reported roles in toxicity-related and disease-related processes (Includes L1000 gene set being used in the NIH LINCS program)

   *Hybrid Informatic Derived + Nominated Genes – S1500*

5. **“Extrapolatability”**: This property refers to the ability to extrapolate or infer or impute with some accuracy the expression changes of unmeasured genes from those measured in this reduced set of sentinel genes

X-axis — the percentage of the total principal components (eigengenes); Y-axis — percentage of variability captured. The red line represents the expected relationship given statistically independent gene expression, whereas the blue curve shows the observed relationship.
5. **Extrapolatability**: This property refers to the ability to extrapolate or infer or **impute** with some accuracy the expression changes of unmeasured genes from those measured in this reduced set of sentinel genes.

Human S1500+ Gene Set Pathway Coverage

Molecular Signature DB Canonical Pathways Coverage

<table>
<thead>
<tr>
<th>Pathways Covered</th>
<th>Mean Coverage</th>
<th>Median Coverage</th>
<th>Mean Multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Total 1320)</td>
<td>(Proportion of Pathway Genes)</td>
<td>(Proportion of Pathway Genes)</td>
<td></td>
</tr>
<tr>
<td>1500 data-driven gene list (after Steps 1-3)</td>
<td>659</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>S1500 gene set (after Step 1-4)</td>
<td>1320</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>S1500+ (2739 genes)</td>
<td>1320</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>Random 1500</td>
<td>541 (443, 695)</td>
<td>0.07 (0.05, 0.09)</td>
<td>0.07 (0.05, 0.09)</td>
</tr>
<tr>
<td>Random 2739</td>
<td>852 (759, 946)</td>
<td>0.13 (0.11, 0.15)</td>
<td>0.12 (0.11, 0.14)</td>
</tr>
<tr>
<td>L1000</td>
<td>906</td>
<td>0.17</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Pathways covered are calculated relative to the **1320 canonical pathways** (MSigDB genesets) in MSigDB version 4.0.

The pathway level coverage is defined as fraction of genes from pathway that overlap selected gene set.

Mean and Median coverage values are derived from pathway level coverage of 1320 canonical pathways from MSigDB (v4.0).

The gene level multiplicity metric represents number of pathways a gene is part of. Mean multiplicity is computed using gene level multiplicity metrics across all selected genes.

For “Random 1500” and “Random 2739” gene sets parenthesized values represent mean and range (min, max) across 20 alternative randomizations.
Human S1500+ Pathway Coverage – NCATS BioPlanet

Hosts the universe of public, curated human pathways (~ 2000 Pathways)

- 77% of pathways have 4 or more genes
- 84% have 3 or more genes
- 91% have 2 or more genes
- 96% have 1 or more genes
- 4% have no gene representation
Outline

- Development of the S1500+ Gene Set
- Evaluation of Performance of the Human S1500+ Gene Set
- Evaluation of the S1500+ on the TempO-Seq HT-Transcriptomics (HTT) Platform
- HTT Proof of Concept Study with an In Vitro Human Liver Organotypic Model
- Future Directions
Human S1500+ Gene Set Extrapolation Performance

Performance of Extrapolated Gene Sets Using an Independent GEO Test Set

<table>
<thead>
<tr>
<th>Gene Level Performance (Fold Change Values)</th>
<th>Pearson Correlation(^a)</th>
<th>Concordance Rate(^b)</th>
<th>Significance Overlap(^c)</th>
<th>Mean Squared Error(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1500+ (2739 genes)</td>
<td>0.75</td>
<td>0.94</td>
<td>0.37</td>
<td>0.20</td>
</tr>
<tr>
<td>Random 2739</td>
<td>0.76 (0.75, 0.76)</td>
<td>0.93 (0.93, 0.93)</td>
<td>0.38 (0.37, 0.38)</td>
<td>0.22 (0.22, 0.22)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pathway Level Performance (GSEA Scores)</th>
<th>Pearson Correlation(^a)</th>
<th>Concordance Rate(^b)</th>
<th>Significance Overlap(^c)</th>
<th>Mean Squared Error(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1500+ (2739 genes)</td>
<td>0.87</td>
<td>0.90</td>
<td>0.60</td>
<td>0.05</td>
</tr>
<tr>
<td>Random 2739</td>
<td>0.78 (0.77, 0.79)</td>
<td>0.86 (0.86, 0.86)</td>
<td>0.44 (0.42, 0.46)</td>
<td>0.08 (0.08, 0.08)</td>
</tr>
</tbody>
</table>

Random values represent mean and range (min, max) across 20-fold cross validation.

\(^a\) Pearson correlations reflect agreement between extrapolated and measured values

\(^b\) Concordance rates reflect the agreement between the extrapolated and the measured data calculated as \((TP + TN)/(TP+TN+FP+FN)\)

\(^c\) Significance overlap relays the proportion of genes/pathways having values (i.e. fold change or GSEA scores) in the top 1% in both the measured and extrapolated datasets

\(^d\) Mean squared error measures the average squared difference between the extrapolated and measured values
Human S1500+ Gene Set Extrapolated Performance

Pathway Level Performance on a Case Study

Follicular Lymphoma vs. Tonsillectomy – Concordance Venn Diagrams

All significantly enriched pathways were identified using enrichment score >0.5 and Kolmogorov Smirnov p-value < 0.001 for this analysis.

Recall is the percentage of the observed up-/down-regulated Pathways (Obs-Up and Obs-Down) that were also correctly predicted as up-/down-regulated (Pred-Up/Down).

Precision is the percentage of the predicted up- and down-regulated Pathways that were observed as up- and down-regulated.
Outline

• Development of the S1500+ Gene Set

• Evaluation of Performance of the Human S1500+ Gene Set

• Evaluation of the S1500+ on the TempO-Seq HT-Transcriptomics (HTT) Platform

• HTT Proof of Concept Study with an In Vitro Human Liver Organotypic Model

• Future Directions
RASL-Seq Based Assay

NextGen Sequencing: $300 \times 10^6$ reads / lane (HiSeq 2500)
- Detector Oligo Target Sequence – Gene ID
- Sample Specific Index Sequence – Sample ID
- NTP Target Average Read Count ~ 500 Reads / Gene
- S1500+ Gene Set – ~ 200 samples / lane
- Whole Transcriptome (22K) – ~ 30 samples / lane
- RNA-Seq – ~ 10 samples / lane

## TempO-Seq Attenuation

<table>
<thead>
<tr>
<th>Fold Attenuator Added</th>
<th>Expected Signal</th>
<th>Probe ID</th>
<th>Max fold</th>
</tr>
</thead>
<tbody>
<tr>
<td>0X</td>
<td>100%</td>
<td>IL8_14324</td>
<td>10</td>
</tr>
<tr>
<td>1X</td>
<td>50%</td>
<td>NAMPT_10869</td>
<td>4</td>
</tr>
<tr>
<td>2X</td>
<td>33%</td>
<td>TTR_7397</td>
<td>4</td>
</tr>
<tr>
<td>4X</td>
<td>25%</td>
<td>ALB_217</td>
<td>3</td>
</tr>
<tr>
<td>9X</td>
<td>10%</td>
<td>CALR_943</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXCL1_28308</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FN1_27231</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ND6_4508</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CYP3A5_27090</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FN1_2459</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GDF15_18329</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HP_3085</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IFITM3_17816</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MT1G_26225</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MT2A_4334</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAMPT_28431</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PGK1_5094</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPIA_27705</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SAT1_6103</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SERPINE1_6253</td>
<td>2</td>
</tr>
</tbody>
</table>

20 Genes Attenuated in Human HepaRG Experiments

In this example, by adding attenuator at 2 times the amount of functional detector, the expected reads for an RNA should decrease by 3, due to competitive inhibition.
DrugMatrix Rat Liver RNA Samples

- Total of 105 RNA samples from livers of rats treated with 27 chemicals with 7 Modes of Action (MOAs) for technology evaluation

**Study design**

- Training set
  - 45 treated rats and 18 matched controls
  - 15 chemicals with 3 rats per chemical
  - 5 MOAs with 3 chemicals per MOA
    - PPARA, CAR/PXR, AhR, Cytotoxic, DNA damage

- Test set
  - 36 treated rats and 6 matched controls
  - 12 chemicals with 3 rats per chemical
  - 4 MOAs with 3 chemicals per MOA
    - PPARA, CAR/PXR, ER, HMGCOA

**Comparing RNA-seq with microarray**

- Ability to recognize the biological replicates
- Concordance in DEGs and pathways
- Agreement to determine MOA-specific biology
- Classifiers from the training set for prediction of the test set samples by MOA
Gene Level Performance

**Measured Genes Only**: Rat “S1500+ beta” - ~ 2300 genes
**Extrapolated Genes Only**: Rat Transcriptome - ~ 15,000 genes

<table>
<thead>
<tr>
<th>Stat</th>
<th>Threshold</th>
<th>Pearson Correlation</th>
<th>MSE Measured</th>
<th>Significance Overlap</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Concordance Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foldchange</td>
<td>2</td>
<td>0.8314</td>
<td>0.3641</td>
<td>0.6875</td>
<td>0.6785</td>
<td>0.9370</td>
<td>0.8857</td>
</tr>
<tr>
<td>Nominal.Pvalue</td>
<td>0.06</td>
<td>0.7477</td>
<td>0.6192</td>
<td>0.2270</td>
<td>0.3270</td>
<td>0.6644</td>
<td>0.7249</td>
</tr>
<tr>
<td>FDR.Broad (mean)</td>
<td>0.1</td>
<td>0.7242</td>
<td>0.1063</td>
<td>0.5231</td>
<td>0.6000</td>
<td>0.9814</td>
<td>0.9789</td>
</tr>
<tr>
<td>Global.Pvalue</td>
<td>0.05</td>
<td>0.8133</td>
<td>0.7095</td>
<td>0.6875</td>
<td>0.8114</td>
<td>0.7792</td>
<td>0.7873</td>
</tr>
<tr>
<td>Connection</td>
<td>0.5</td>
<td>0.9799</td>
<td>0.0033</td>
<td>0.7500</td>
<td>0.8795</td>
<td>0.9776</td>
<td>0.9544</td>
</tr>
<tr>
<td>Foldchange</td>
<td>2</td>
<td>0.6156</td>
<td>0.4167</td>
<td>0.2220</td>
<td>0.9876</td>
<td></td>
<td>0.9094</td>
</tr>
<tr>
<td>Nominal.Pvalue</td>
<td>0.05</td>
<td>0.6007</td>
<td>0.9740</td>
<td>0.9743</td>
<td>0.9740</td>
<td>0.9734</td>
<td>0.9664</td>
</tr>
<tr>
<td>FDR.Broad (mean)</td>
<td>0.1</td>
<td>0.5813</td>
<td>0.1011</td>
<td>0.3726</td>
<td>0.1643</td>
<td>0.9968</td>
<td>0.9953</td>
</tr>
<tr>
<td>Global.Pvalue</td>
<td>0.05</td>
<td>0.6588</td>
<td>0.5538</td>
<td>0.4167</td>
<td>0.3950</td>
<td>0.9358</td>
<td>0.8622</td>
</tr>
<tr>
<td>Connection</td>
<td>0.5</td>
<td>0.9625</td>
<td>0.0055</td>
<td>0.8333</td>
<td>0.9773</td>
<td>0.9121</td>
<td>0.9202</td>
</tr>
</tbody>
</table>
Gene vs. Pathway Level Performance

All Genes = Measured + Extrapolated Rat genes

<table>
<thead>
<tr>
<th>Stat</th>
<th>Threshold</th>
<th>Pearson Correlation</th>
<th>MSE</th>
<th>Significance</th>
<th>Overlap</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Concordance Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foldchange</td>
<td>2</td>
<td>0.6922</td>
<td>0.3578</td>
<td>0.4992</td>
<td>0.3675</td>
<td>0.9787</td>
<td></td>
<td>0.9048</td>
</tr>
<tr>
<td>Nominal P-value</td>
<td>0.5</td>
<td>0.5153</td>
<td>0.9058</td>
<td>0.1841</td>
<td>0.5040</td>
<td>0.7899</td>
<td></td>
<td>0.9242</td>
</tr>
<tr>
<td>FDR Broad (mean)</td>
<td>0.1</td>
<td>0.6170</td>
<td>0.1021</td>
<td>0.4258</td>
<td>0.3662</td>
<td>0.9938</td>
<td></td>
<td>0.9921</td>
</tr>
<tr>
<td>Global P-value</td>
<td>0.05</td>
<td>0.7052</td>
<td>0.5840</td>
<td>0.4992</td>
<td>0.5232</td>
<td>0.9088</td>
<td></td>
<td>0.8477</td>
</tr>
<tr>
<td>Connection</td>
<td>0.5</td>
<td>0.9709</td>
<td>0.0041</td>
<td>0.8889</td>
<td>0.9786</td>
<td>0.9238</td>
<td></td>
<td>0.9316</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stat</th>
<th>Threshold</th>
<th>Pearson Correlation</th>
<th>MSE</th>
<th>Significance</th>
<th>Overlap</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Concordance Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathway Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>0.5</td>
<td>0.6988</td>
<td>0.1415</td>
<td>0.6566</td>
<td>0.6625</td>
<td>0.8923</td>
<td></td>
<td>0.9242</td>
</tr>
<tr>
<td>NES</td>
<td>2</td>
<td>0.7108</td>
<td>0.8612</td>
<td>0.5815</td>
<td>0.5840</td>
<td>0.9887</td>
<td></td>
<td>0.9814</td>
</tr>
<tr>
<td>Nominal P-value</td>
<td>0.5</td>
<td>0.6651</td>
<td>0.4238</td>
<td>0.5940</td>
<td>0.6957</td>
<td>0.8999</td>
<td></td>
<td>0.9316</td>
</tr>
<tr>
<td>FDR Broad (mean)</td>
<td>0.1</td>
<td>0.8232</td>
<td>0.1442</td>
<td>0.5795</td>
<td>0.7044</td>
<td>0.9471</td>
<td></td>
<td>0.9303</td>
</tr>
<tr>
<td>Global P-value</td>
<td>0.05</td>
<td>0.8198</td>
<td>0.5277</td>
<td>0.5815</td>
<td>0.6722</td>
<td>0.8897</td>
<td></td>
<td>0.8539</td>
</tr>
<tr>
<td>Connection</td>
<td>0.5</td>
<td>0.9355</td>
<td>0.0121</td>
<td>0.8056</td>
<td>1.0000</td>
<td>0.9257</td>
<td></td>
<td>0.9373</td>
</tr>
</tbody>
</table>

ES = Enrichment Score
NES = Normalized Enrichment Score
• Development of the S1500+ Gene Set
• Evaluation of Performance of the Human S1500+ Gene Set
• Evaluation of the S1500+ on the TempO-Seq HT-Transcriptomics (HTT) Platform
• HTT Proof of Concept Study with an In Vitro Human Liver Organotypic Model
• Future Directions
HT Transcriptomics – Initial Human In Vitro Model System

Steve Ferguson & Sreeni Ramaiahgari, NTP

HT Transcriptomics – Initial Study

Steve Ferguson & Sreeni Ramaiahgari, NTP

- HepaRG *in vitro* Model Treatments – 384 well format
  - 2D proliferating HepaRG (moderate metabolism)
  - 2D HepaRG differentiated, confluent (induced metabolism)
  - 3D HepaRG (robust metabolism) – *data yet to come*
  - 24 Compounds
  - 10 Concentration levels, each with triplicate wells
  - 3 Independent runs (different days)
  - 96 hr treatment period with dosing on Day 1 and Day 3 (media changes)
  - Endpoints: morphology, LDH, ROS, metabolomics, targeted P450 metabolism
  - Lysis for gene expression analysis using the Tox21 human S1500+ gene set
  - ~9000 samples analyzed by HTT, generating ~25 Million data points
S1500+ HT Transcriptomics – Gene Read Counts

Trey Saddler, NTP

Transcripts Per Million

Compounds / Samples

Genes
Sreeni Ramaiyahgari, NTP

<table>
<thead>
<tr>
<th>Compounds / Samples</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sreeni Ramaiahgari, NTP

Compounds / Samples

Aspirin  Caffeine  Sucrose  KCl  Levo  Trova  Floxacin  BaP  AF-B1

Genes
Moving Tox21 towards Quantitative Systems Toxicology

Hypothesis

- Transcriptome profiling of *in vitro* treated human cells can provide an approximation of human *in vivo* responses to chemical exposures

Goal

- To evaluate High-Throughput Transcriptomic analysis of *in vitro* cell models for providing *Bench Mark Dose (BMD)* information relevant to human BMD values and begin to address best practices.
Signature-level Genomic Benchmark Dose

BMDExpress – Rusty Thomas, Hamner, et al.
BMDExpress 2.0 – Scott Auerbach, NTP, et al.
Accumulation of significantly affected Genes with their BMD values
Accumulation of significantly affected Pathways with their BMD values

Each dot is a gene or pathway BMD value. The BMDs from a given treatment are linked by a line. The BMDs are ordered from most “potent” to least “potent”. The plot shows the accumulation from rank 1 to the total number of significantly affected genes or pathways. The most “potent” perturbed gene or pathway BMDs are on the left and the least “potent” are on the right.
Inter-Run Repeatability (2D Confluent HepaRG)

Steve Ferguson, NTP – Gene Level Median BMD Values
Rifampicin Activation of PXR Signaling via CYP3A4 Induction

Steve Ferguson & Nisha Sipes, NTP – Gene Level Median BMD Values

- PXR/CAR classical target genes painted in red
  - PXR target gene median BMD: 33 nM
- ER stress, oxidative stress, & apoptosis pathway gene BMDs largely higher than the human therapeutic (oral) peak plasma concentration
  - C\textsubscript{max} ~ 13 μM
  - ER Stress/Oxidative Stress/Apoptosis

CYP3A4 with Increasing RIF
Case Study Comparison of Drug Analogues

Steve Ferguson, NTP – Gene Level Median BMD Values

Trovafoxacin vs. Levofloxacin

Trovafoxacin Runs 1,2,3

Levofloxacin Runs 1,2,3
• Development of the S1500+ Gene Set

• Evaluation of Performance of the Human S1500+ Gene Set

• Evaluation of the S1500+ on the TempO-Seq HT-Transcriptomics (HTT) Platform

• HTT Proof of Concept Study with an In Vitro Human Liver Organotypic Model

• Future Directions
NTP HT Transcriptomics QC Workflow (under construction)

Step 1: Sequencing Depth > X?
- YES
- NO → Flag the sample

Step 2: >X% of reads align full-length to the probes?
- YES
- NO → Flag the sample

Step 3: >X% of the reads align uniquely to the probes?
- YES
- NO → Flag the sample

Step 4: Aligned reads > X?
- YES
- NO → Flag the sample

Step 5: >X% of the probes have at least 5 reads?
- YES
- NO → Flag the sample

Step 6: Per base sequence quality is high?
- YES
- NO → Flag the sample

Step 7: Per base N content is low?
- YES
- NO → Flag the sample

Passes QC → Executive summary

Flag the sample

Detailed Summary

Tox21
Moving Forward

- S1500+ development for human (complete), rat & mouse (near completion), zebrafish (initiated)

- HTT screening for prioritization – BioSimilarity, Biological Read-Across
  - e.g. Perfluorinated Compounds (PFOA, PFOS), Botanicals, Polycyclic Aromatic Compounds, Flame Retardants, etc.

- Integrate High Content Imaging with qHTS & HT-T – Phenotypic Anchoring

- Integrate Metabolomics data from NTP Labs with HT-T data – PBPK, etc.

- Application to biomaterial from NTP rat and mouse studies
  - Acute (5 Day) Rat Transcriptomics Studies
  - Diversity Outbred mouse models for screening for Genetic Variation in responses to chemical exposures
  - Rat and Mouse Archived FFPE material
Tox21: A Collaboration of Many...

NTP BSB
Scott Auerbach
Steve Ferguson
Sreenivasa Ramaiyahgari
Nisha Sipes
Alex Merrick
Kristine Witt
Stephanie Smith-Roe
Alison Harrill
Julie Foley
Jui-Hua Hsieh
Fred Parham
Trey Saddler
Ray Tice (Retired)

BioSpyder Team

NTP / NIEHS
Mike DeVito
Warren Casey
Elizabeth Maull
Nicole Kleinstreuer
Suramya Waidyanatha
Windy Boyd
Mamta Behi
Brad Collins
Jennifer Fostel
Keith Shockley
Grace Kissling
John Bucher
Linda Birnbaum

SciOme
Deepak Mav
Ruchir Shah
Dan Svoboda

Tox21 Colleagues
NCATS
Anton Simeonov
Chris Austin
Menghang Xia
Ruili Huang
Dave Gerhold
Anna Rossoshek

US EPA
Rusty Thomas
Kevin Crofton
Richard Judson
Ann Richard
Bob Kavlock

US FDA
Suzy Fitzpatrick
Dan Acosta
Donna Mendrick
Weida Tong
Thank you!

Questions?