NTP
National Toxicology Program

# Contract Concept: Bioinformatics Support for DNTP and DIR

B. Alex Merrick, PhD
Biomolecular Screening Branch, DNTP
National Institute of Environmental Health Sciences

NTP Board of Scientific Counselors Meeting
December 1 - 2, 2015

# **Purpose:** Provide bioinformatic support to DNTP/DIR

## What is bioinformatics?

- **What?** Computational analysis, biostatistical treatment and visualization of large, multidimensional datasets in multidisciplinary research fields

- **Includes:** Omics technologies, chemistry, HTS, and literature data extraction

- **Innovation:** Datasets from various disciplines require support for adaptation of software tools or innovative procedures for new informatics approaches

# Need for bioinformatic support

**Large, complex datasets:** NIEHS researchers conduct studies producing large amounts of data, varying in size and complexity.

**Multidisciplinary data** - Fields of scientific study are diverse, require innovative bioinformatics analysis, software tools and data visualization.

**Emerging Bioinformatics Needs:** NTP leadership decided a new contract was needed to address the growing bioinformatics needs of NTP and DIR and to support the magnitude and range of data and bioinformatics analysis for the studies conducted at NIEHS.

**Data Volume and Existing Expertise -** More studies and expertise than can be handled by existing bioinformatics support from Biostatistics and Computational Biology Branch to support NIEHS researchers.

**Limits of Statistical Support Contract** - NTP leadership recognized the informatics component was becoming distinctive from statistical support contract; integration with existing contract support.

# General Requirements

- **Bioinformatic services** – independent of the government
- **Report data analysis** – report data analyses to NIEHS groups
- **New bioinformatic** – adapt or create new, bioinformatic procedures and analysis
- **Frequent interaction** – contract officer representative (COR), DNTP/DIR researchers; onsite and electronic

# Specific Requirements

## Bioinformatic support for:

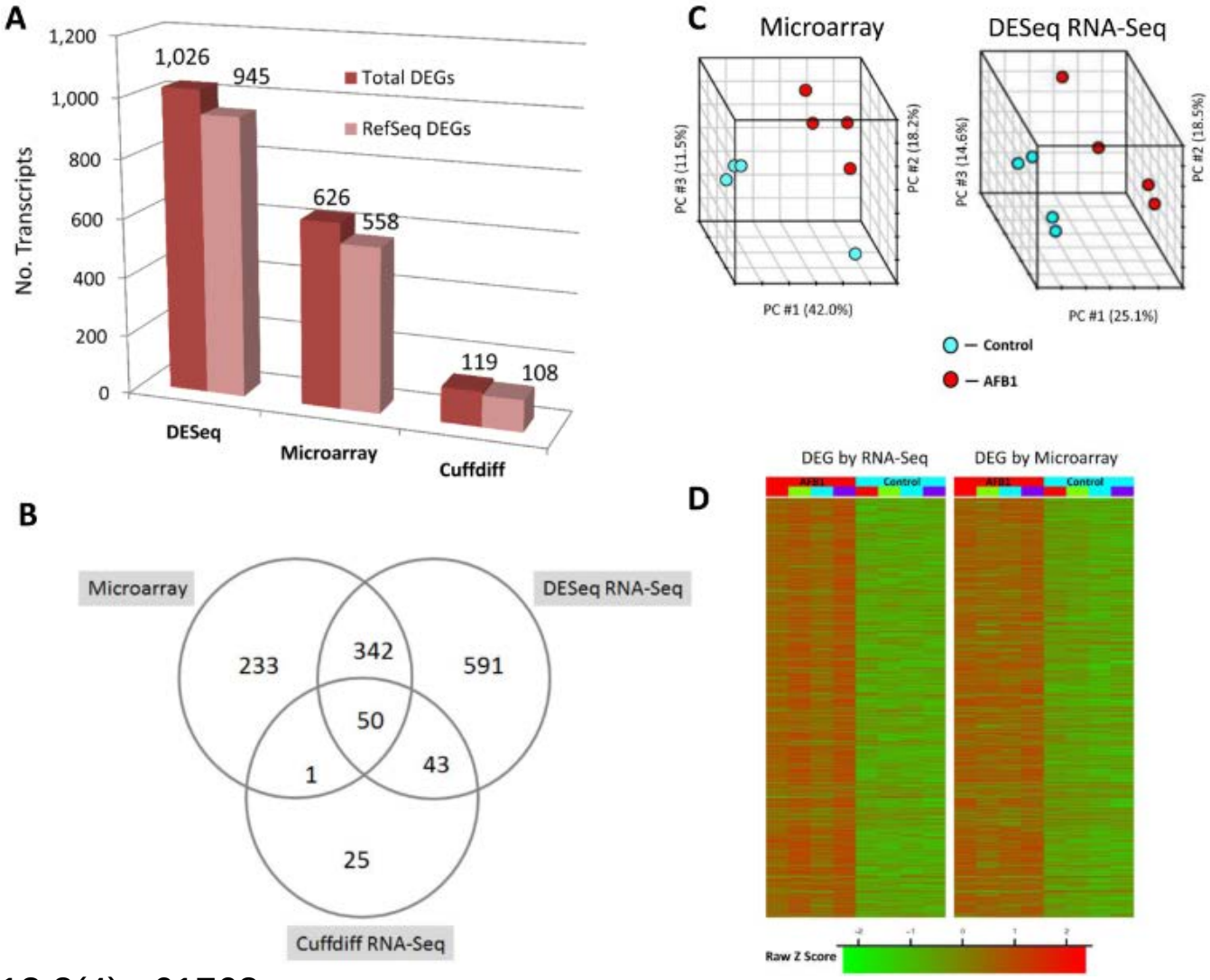
- **Omics** – genomics, transcriptomics, proteomics, metabolomics
- **Genomic sequencing**– NextGen data
- **Epigenetics**– DNA methylation; histone modifications
- **Toxicogenomics** – microarray, RNA-seq
- **High Throughput Transcriptomics** – S1500+
- **Cheminformatics** – Tox21 HTS data, SAR, integration
- **Information Mining** – data or text
- **Impact Research** – NTP research products, impact
- **Informatics data extraction** – Tox21 chemical extraction
- **Study Design** – informatics support for study design teams
- **Innovative Bioinformatics** – adaptation of existing and new data pipelines for new bioinformatic approaches for NTP/DIR

# Bioinformatics Needs of Investigators

- **NextGen Sequencing** – RNASeq
- Tox21 and **Chemoinformatics**
- HTT – **High Throughput Transcriptomics** for Tox21 and toxicity screening projects
- Informatics for **systematic literature** reviews
- Informatics for **impact research**

# Differentially expressed (DEGs) identified from RNA-Seq

## RNA-Seq compared to microarray data




Plos One 2013 8(4):e61768

# High Throughput Transcriptomics

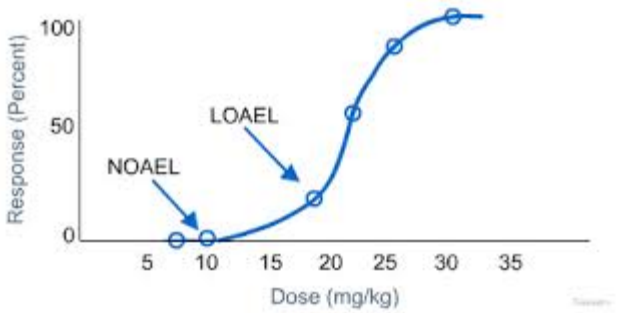
## Gene expression signatures as surrogate markers for chemical toxicity screens – ‘S1500+’

- **Selection of genes** to be used with HTT platform
- **Represent known pathways** - Select highly diverse, correlated and complementary set of genes that fully represent all known pathways
- **Data driven** - use of data to build transcriptome
- **HTT Data** – to recreate transcriptome
- **HTT Chemical Screening** - identify perturbed genes and pathways
- **Informatics** – identify gene modules and pathways impacted by chemical treatment
- **Prioritize** chemicals for further testing
- **Integrate** and analyze data from diverse datastreams

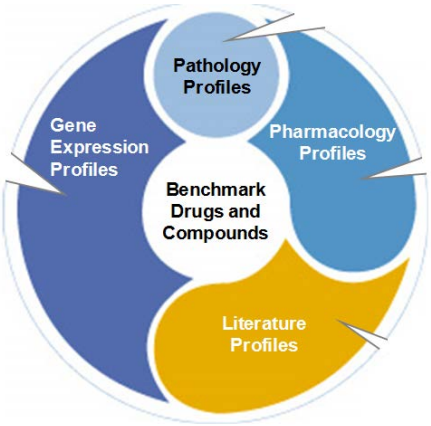
**S1500+ details:**
https://ntp.niehs.nih.gov/results/hts/s1500-gene-set/index.html

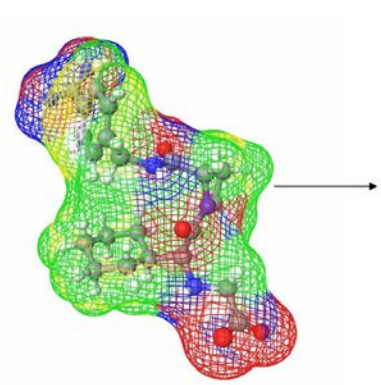
# Tox21 and Cheminformatics

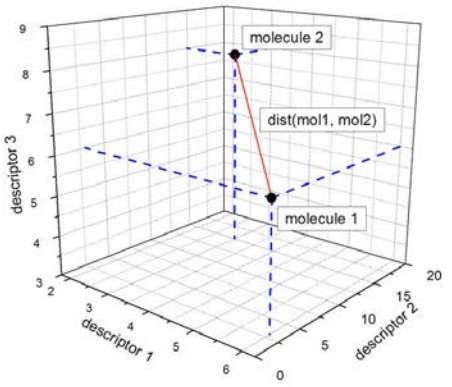
- Integration, Curation and Analysis of large volumes of assay, toxicology and animal study data
- Develop visualization tools to view HTS data
- Research and implement curve fit models




- Maintain Drug Matrix database and provide user support
- User of chemical descriptor, QSAR and cheminformatics approaches for chemical toxicity predictions, integrative analysis and chemical prioritization



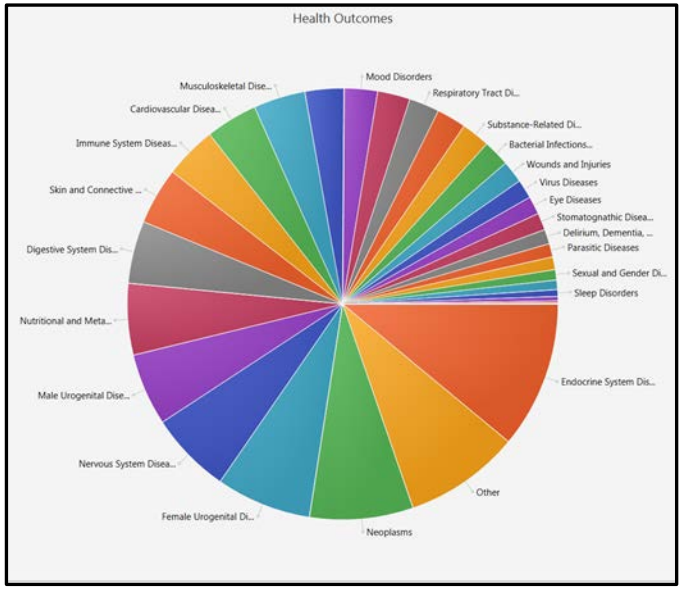


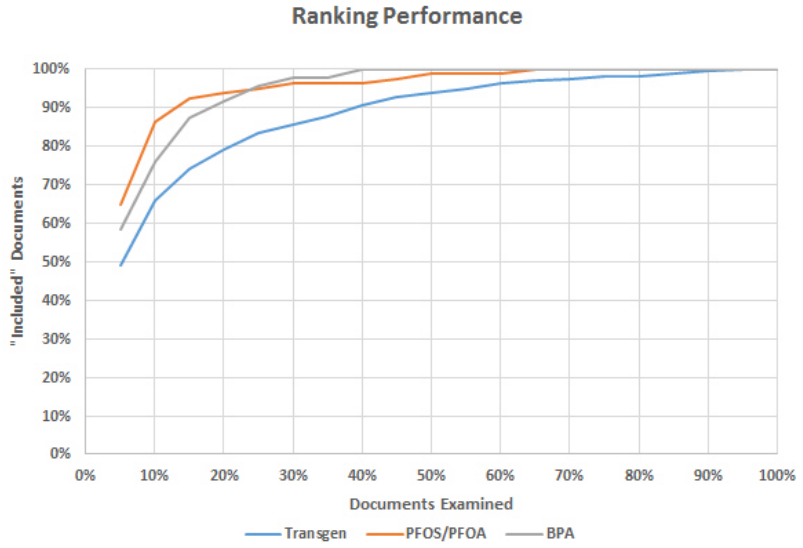

DNTP's *Drug Matrix Database*

# Informatics for Systematic Reviews

## Informatics capabilities and tools for:

- Scoping Reports
- Problem Formulation
- Topic Refinement
- Document Prioritization



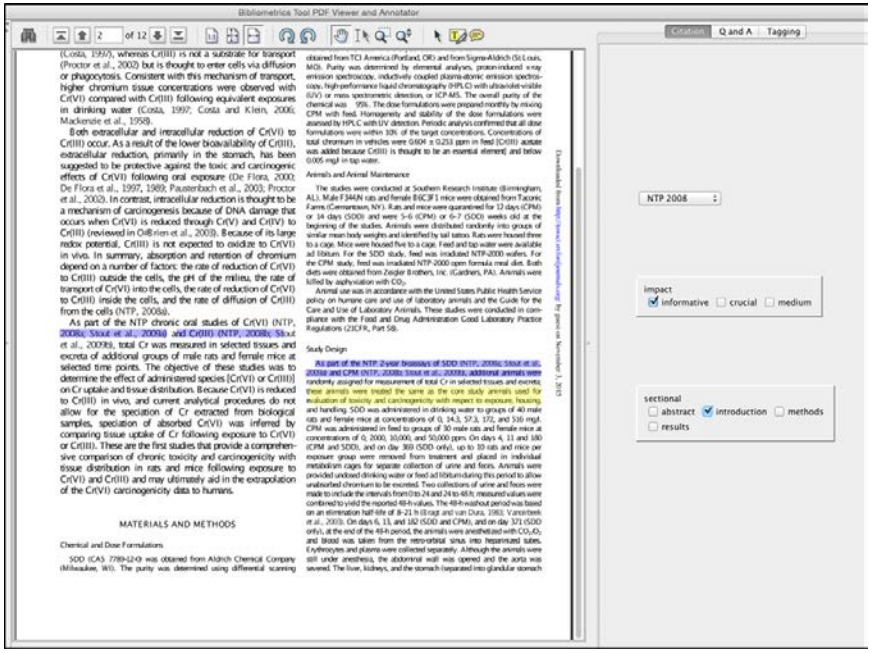




# Informatics for Impact Research

- Informatics expertise and tools to support impact research
- Identify and assess impact of NTP research

## Methodologies under consideration:

- Bibliometrics: easily search, visualize, tag, organize, screen and research articles
- Web mining: perform extensive web mining using high performance computing and sift through results in rapid manner
- News story: organize and summarize news articles



# Contractor Request for Information

## Problem and Scope of Work

- Work plan – informatics approach to areas of need
- Informatics resources – pipelines, biocomputational skills

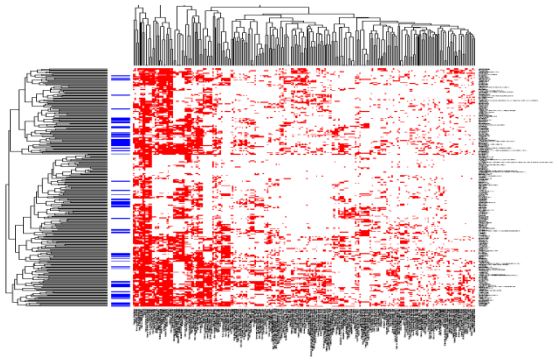
## Relevant Organizational Experience/Management

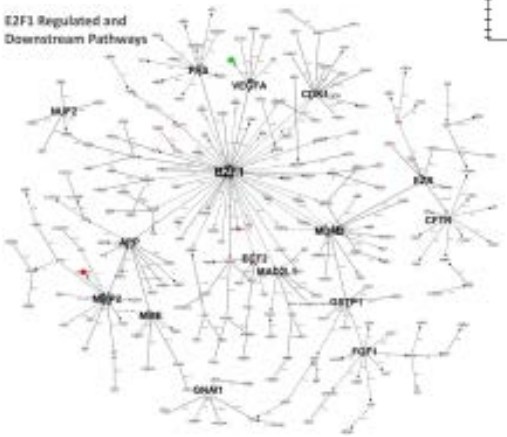
- Bioinformatic project experience
- Management plan
- Facilities and equipment
- Conflict of interest

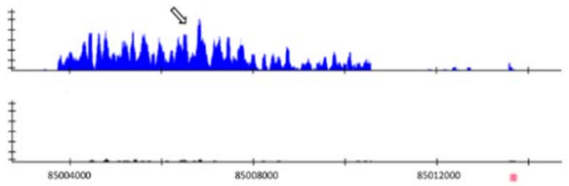
## Quality Assurance

RFI issued on 8-23-2015 – NIHES2015048-RFI

# Questions?




E2F1 Regulated and
Downstream Pathways


85004000
85008000
85012000