

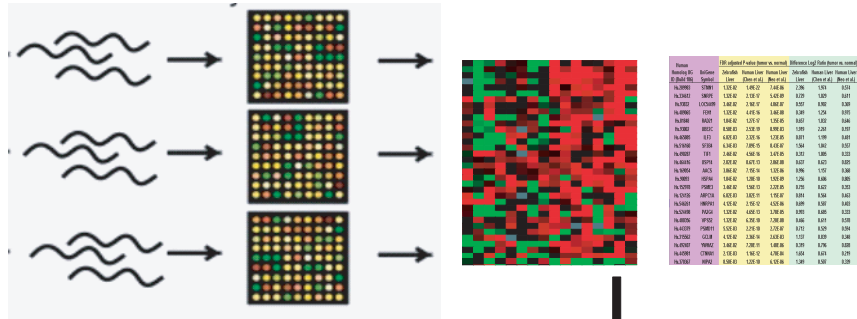
A. Chemical Treatments



16 chemicals
(Experiments A & B)



B. Microarray



RNA
extraction

Exp. A
Datasets
I & II

Exp. B
Datasets
III & IV

C. Differential Gene Expression Analyses for P(H)AH/BAP and ECs/DES groups only

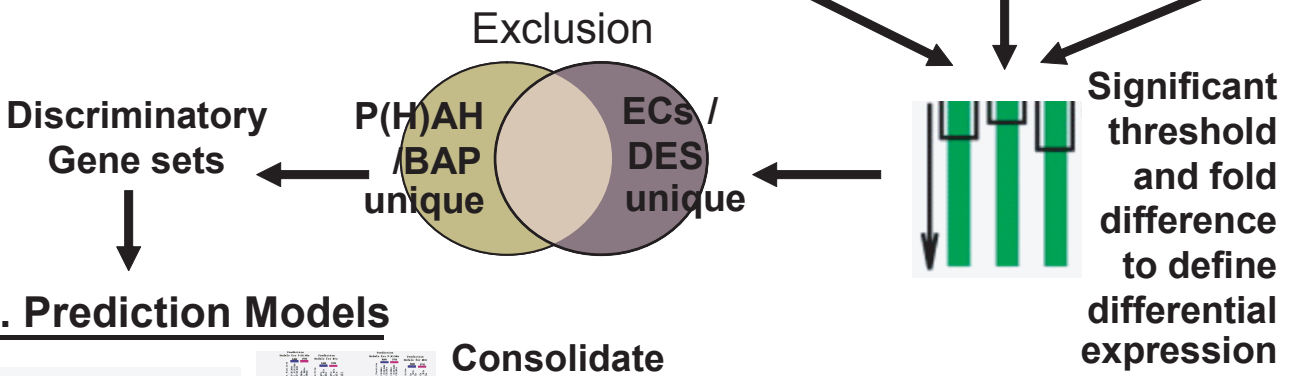
| | A | B | C | D | E | F | G | H | I | J | K | L |
|-----------------|--------|--------|--------|--------|--------|--------|--------|---------|---------|--------|---|--------|
| 1 GenBankGene | | | | | | | | | | | | |
| 2 AF256268.vg | 7.6708 | 8.0293 | 8.826 | 7.0439 | 7.6252 | 8.2518 | 7.2309 | 0.0677 | | | | -0.087 |
| 3 AF408784.vg | 6.92 | 7.1708 | 7.3611 | 7.002 | 7.6841 | 8.1765 | 8.1765 | 0.1689 | -0.1796 | -0.139 | | |
| 4 U08051.ppt | 4.2201 | 5.0350 | 4.8545 | 4.6961 | 4.0280 | 5.4729 | 5.0122 | 0.3344 | -0.134 | -0.161 | | |
| 5 AL078269.mnt | 2.6124 | 3.4532 | 1.9988 | 2.3661 | 3.475 | 4.1693 | 3.2075 | -0.0874 | 0.0967 | 0.419 | | |
| 6 BC088184.cpn1 | 3.023 | 3.3362 | 3.0283 | 2.896 | 3.0939 | 2.9536 | 3.2123 | -0.0543 | 0.3607 | -0.145 | | |
| 7 AF029577.vg | 3.9007 | 1.3200 | 1.4844 | 1.1636 | 1.44 | 1.1179 | 1.1688 | -0.4056 | 0.116 | -0.338 | | |
| 8 BP156109.dh | 2.8124 | 3.4532 | 1.9988 | 2.3661 | 3.475 | 4.1693 | 3.2075 | -0.0874 | 0.0967 | 0.419 | | |

Genes
sorted by
P-values

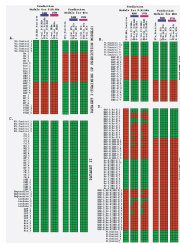
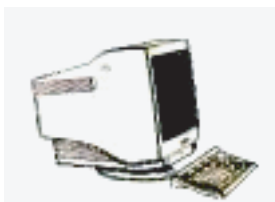
Genes
sorted by
FDR-
values

Genes
sorted by
Q-values
(SAM)

Test of Significance for each
gene (Treated vs. Control)



D. Prediction Models



Training using Supervised Learning Classifiers (kNN and SVM). Assess robustness using leave-one-out approach or independent datasets.

Consolidate
Gene Sets

E. Real-time PCR Validation of Candidate Biomarkers

F. Identification of Biomarkers in Selected Tissues Using Real-time PCR on a New Batch of Fish (Exp.C)

G. Knowledge-based Data Mining for Human Health-risk and Biological Insight Inference

Supplementary Figure 1: An overview of the workflow for large-scale predictive and discovery chemical biology using whole adult zebrafish chemogenomics.