

Gene Expression Profiles of Rodent Pulmonary Tissue Following Sulfur Mustard Exposure Using ArrayTrack

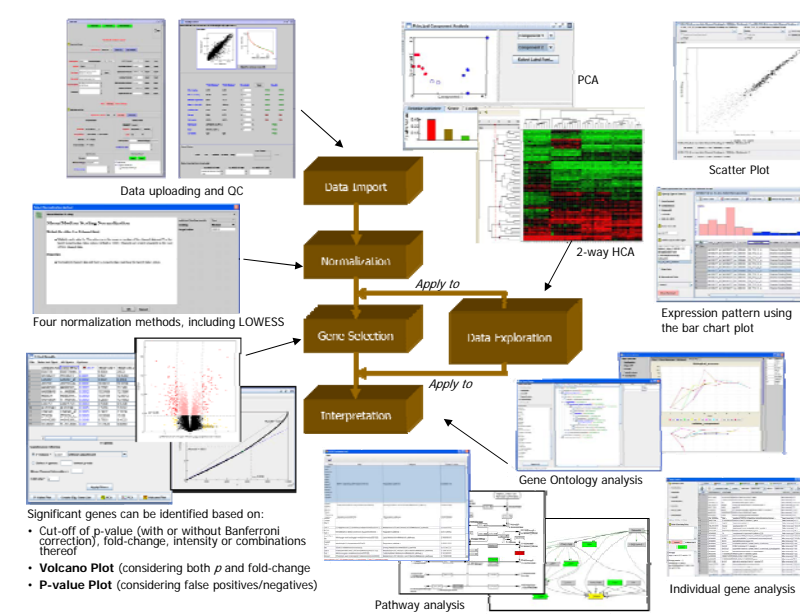
Minjun Chen^{1,2}, Weida Tong⁴, William Welsh^{1,3}, Panos G. Georgopoulos^{1,2} - ¹Environmental & Occupational Health Sciences Institute of UMDNJ-RWJMS/Rutgers University; ²Dept. of Environmental and Occupational Medicine and ³Dept. of Pharmacology, UMDNJ-RWJMS; ⁴USFDA National Center for Toxicological Research, Center for Toxicoinformatics

Abstract

ArrayTrack is a software platform for bioinformatic studies, available from USFDA, that provides an integrated solution for managing, analyzing and interpreting microarray data. Currently ArrayTrack incorporates a Gene Ontology for Function Analysis (GOFA) component ("tool") for biological interpretation based on gene ontology. On-going work at our Center (ebCTC) aims to expand ArrayTrack into a new resource, ebTrack, which will incorporate new components for the analysis of proteomic and metabolomic data as well as additional "systems toxicology" resources.

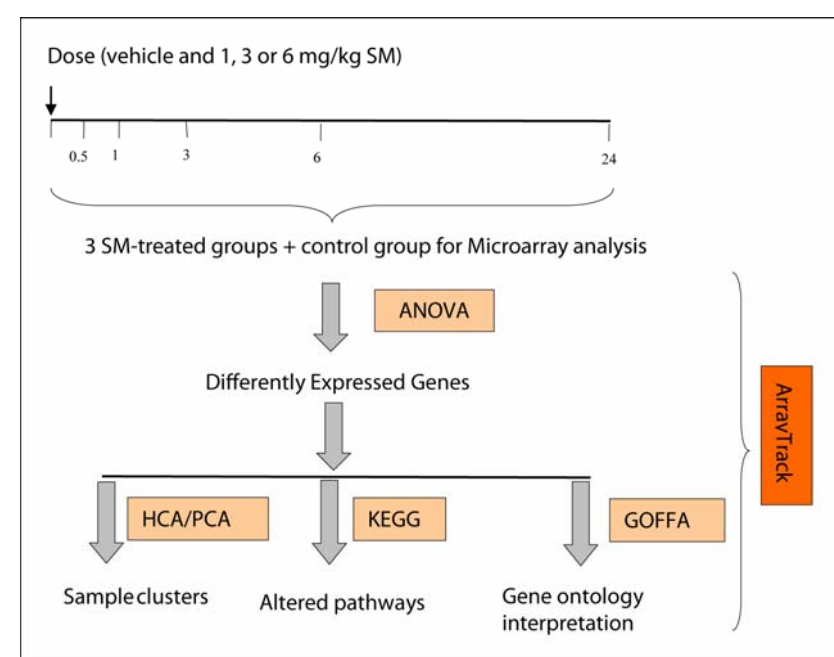
The present example of ArrayTrack application utilizes microarray data of rodent pulmonary tissue expression following Sulfur Mustard (SM) exposure, downloaded from NCBI. Principal component analysis and hierarchical cluster analysis of the differently expressed genes, performed with ArrayTrack, showed that the samples were grouped together according to the dose and harvested time points. Pathways analysis employing the KEGG library identified various pathways (such as Glycolysis) that were involved in the response to SM treatment. Also, the analysis with the GOFA tool indicated that the biological process related to immune response, defense response, and response to biotic stimuli, were significantly altered by SM exposure in the lung. These results demonstrate the utility of ArrayTrack as a platform for microarray data processing, that allows direct linking of analysis results with function information to facilitate biological interpretation.

An example application of using ArrayTrack/ebTrack for microarray data analysis, presenting all the important steps involved



Experimental design

Data source: NCBI's gene expression omnibus (GEO, www.ncbi.nlm.nih.gov/geo/)^{1,2}

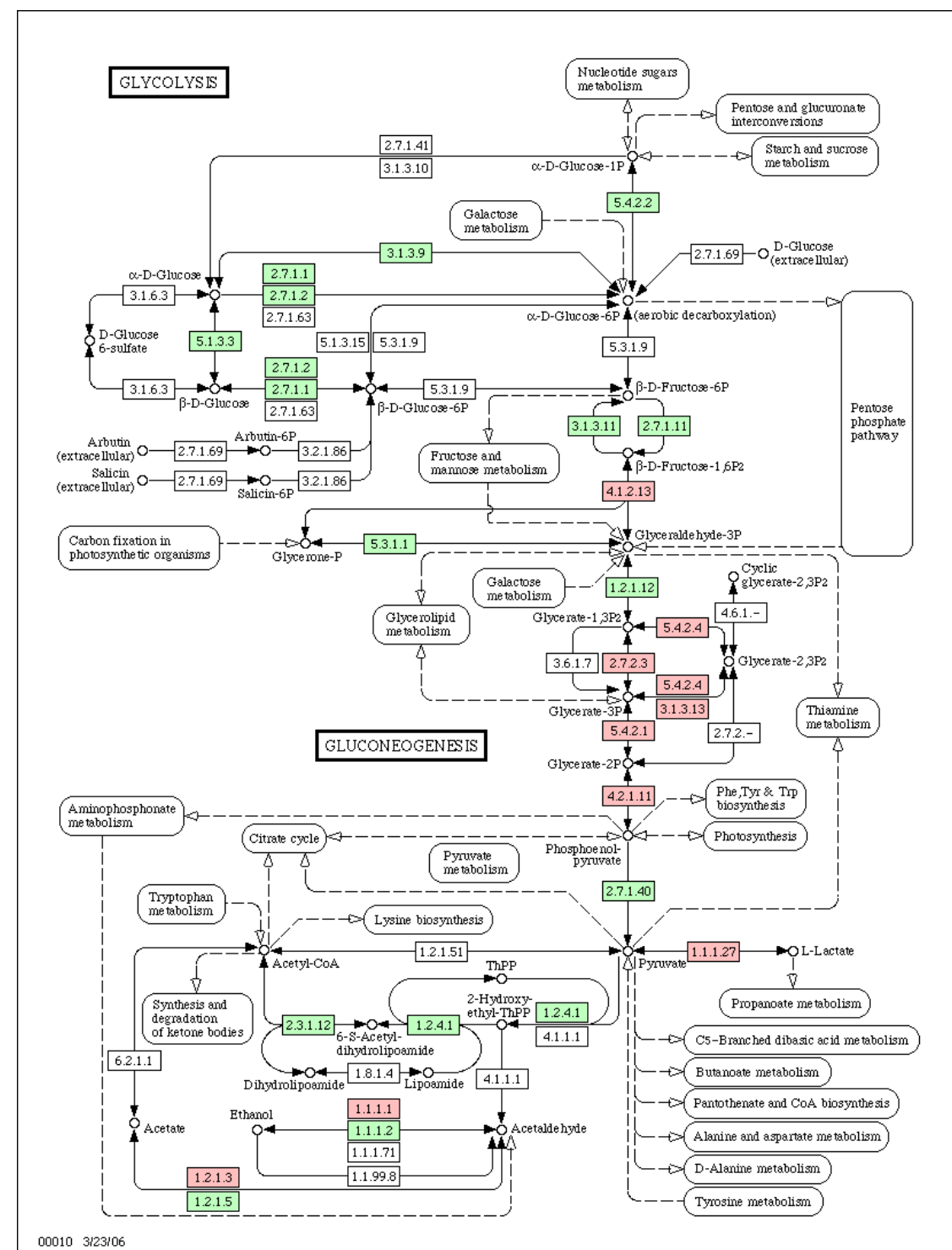


Schematic of the microarray experimental design and subsequent data analysis using ArrayTrack for interpretation of gene expression profiles induced by sulfur mustard (SM).

Pathway analysis using KEGG

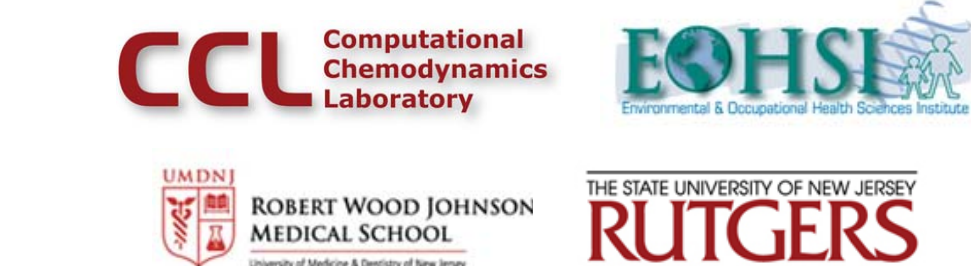
pathway	number of genes involved
Cell adhesion molecules (CAMs)(mo04514)	22
ECM-receptor interaction(mo04512)	13
Type 1 diabetes mellitus(mo04540)	13
Antigen processing and presentation(mo04612)	13
Inositol phosphate metabolism(mo00562)	7
Glioma(mo05214)	9
Cell Communication(mo01430)	13
Apoptosis(mo04210)	7
Glycolysis / Gluconeogenesis(mo00010)	7
Phosphatidylinositol signaling system(mo04070)	8
Hematopoietic cell lineage(mo04640)	9
GnRH signaling pathway(mo04912)	10
Rheumatoid arthritis(mo04210)	9
Glutamate metabolism(mo00251)	4
Alzheimer's disease(mo05010)	4
Neurodegenerative Disorders(mo01510)	5
Nitrogen metabolism(mo00910)	3
Dipeptide biosynthesis(mo01220)	4
1- and 2-Methylsaphthalene degradation(mo00624)	2
Natural killer cell mediated cytotoxicity(mo04650)	9
Prion disease(mo05060)	2
Focal adhesion(mo04510)	15
TGF-beta signaling pathway(mo04350)	7
Toll-like receptor signaling pathway(mo04620)	7
Gap junction(mo04540)	8
MAPK signaling pathway(mo04010)	20

KEGG pathway analysis of the dose-response gene expression alteration after SM treatment



KEGG pathway map (mo00010) on Glycolysis / Gluconeogenesis involved in the SM treatment. The red and green represent the up- and down- regulation of gene expression.

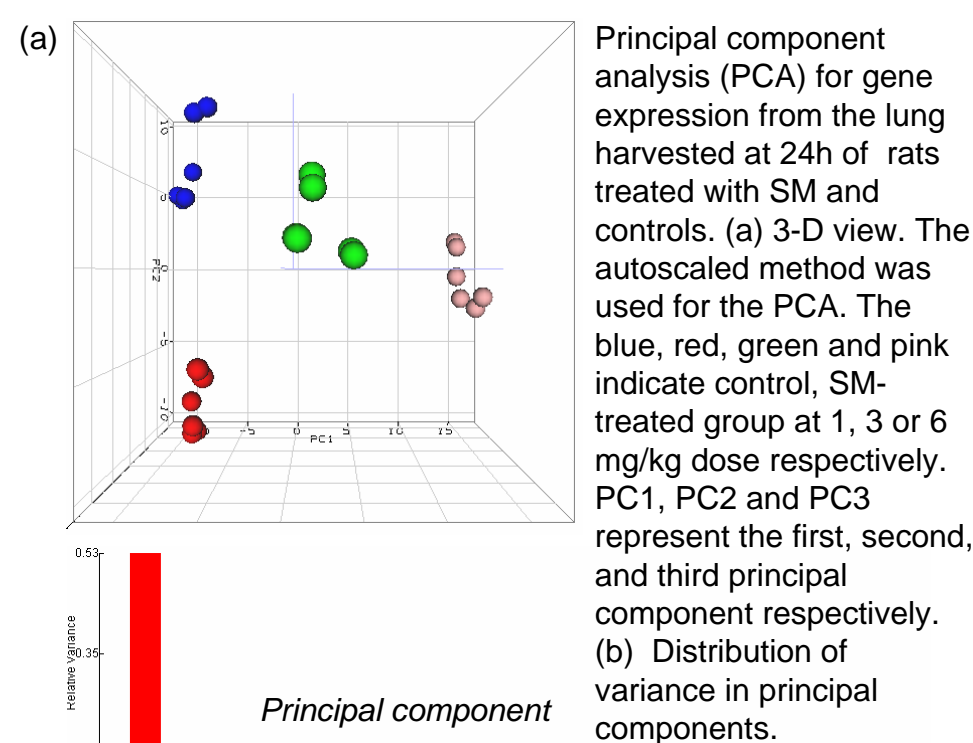
ebCTC environmental bioinformatics and Computational Toxicology Center



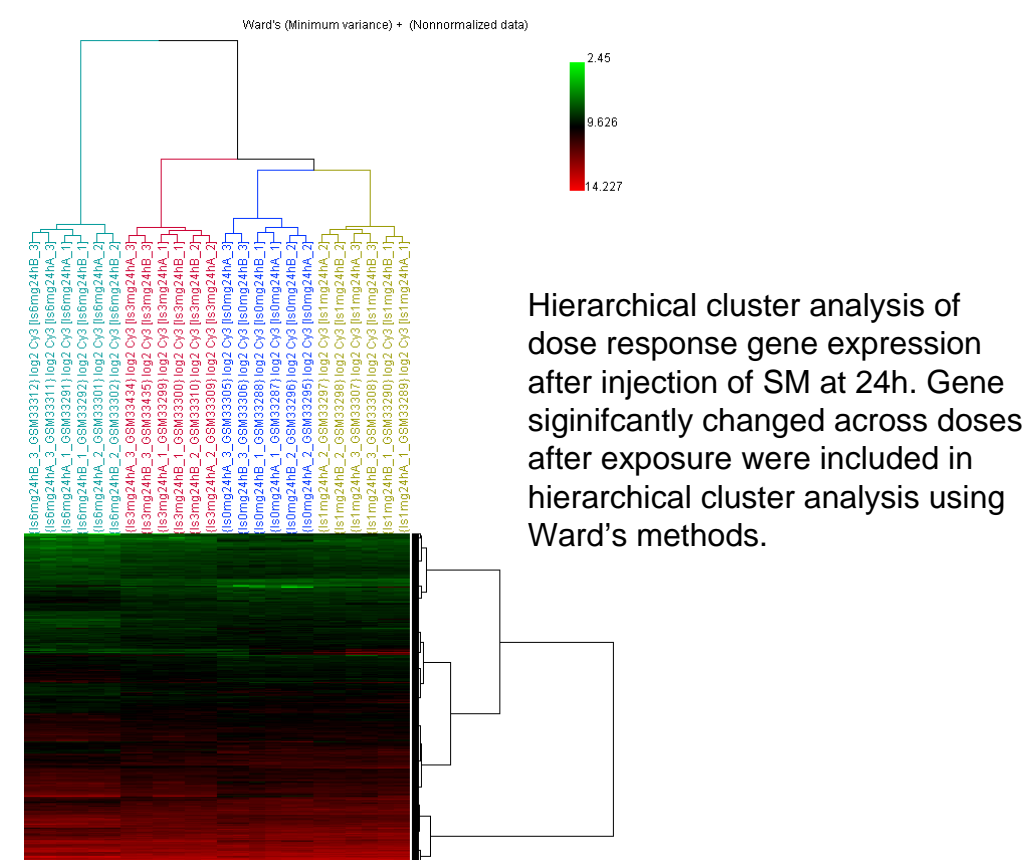
Acknowledgements

Support for this work has been partially provided by the USEPA-funded Environmental Bioinformatics and Computational Toxicology Center (ebCTC), under STAR Grant number GAD R 832721-010. This work has not been reviewed by and does not represent the opinions of the funding agency.

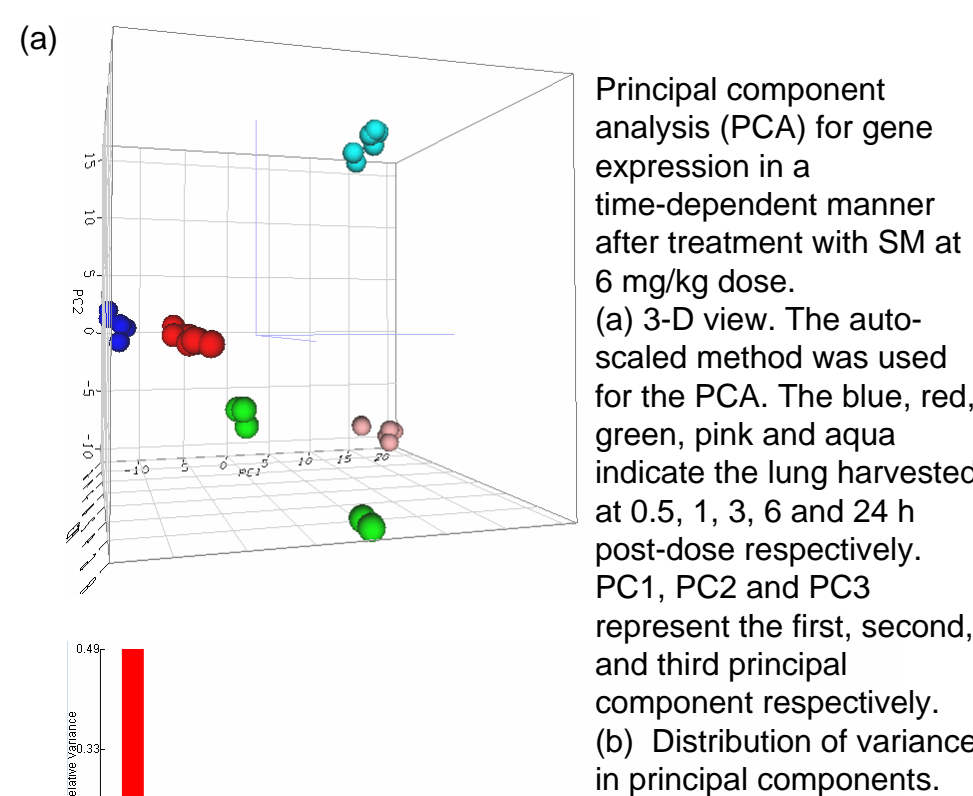
Principal component analysis (PCA) and hierarchical cluster analysis (HCA) of the dose/time response gene expression alteration after SM treatment



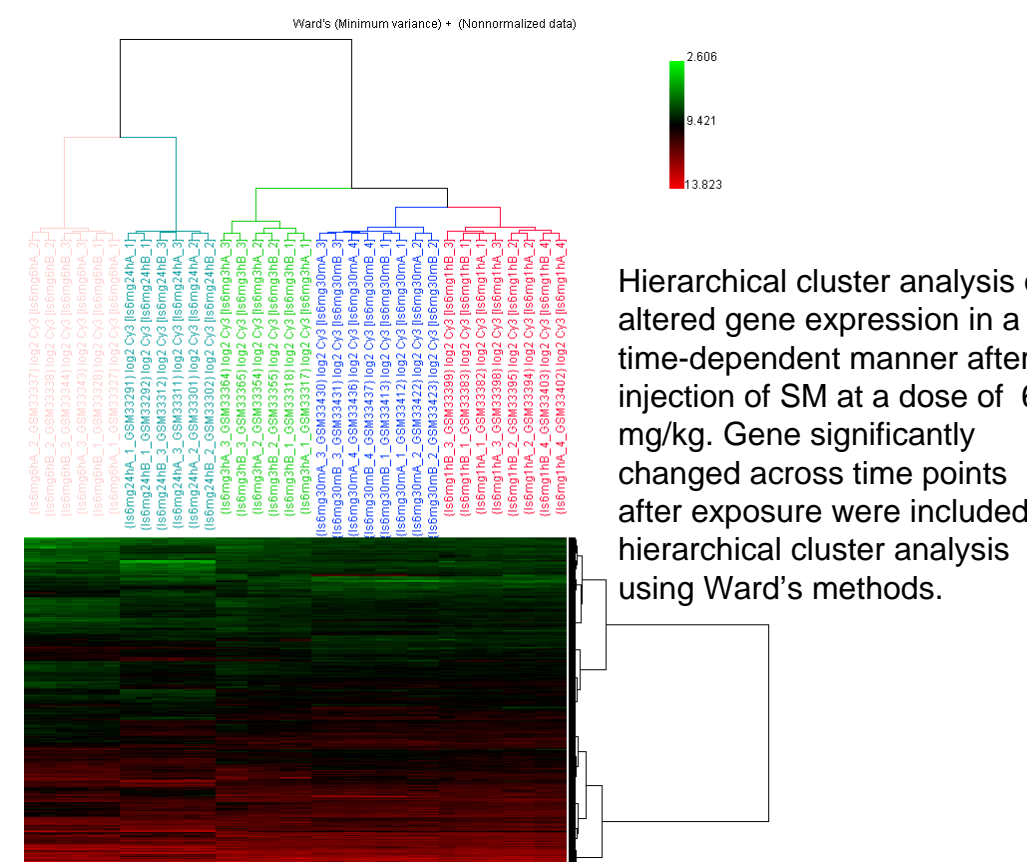
Principal component analysis (PCA) for gene expression from the lung harvested at 24h of rats treated with SM and controls. (a) 3-D view. The autoscaled method was used for the PCA. The blue, red, green and pink indicate control, SM-treated group at 1, 3 or 6 mg/kg dose respectively. PC1, PC2 and PC3 represent the first, second, and third principal component respectively. (b) Distribution of variance in principal components.



Hierarchical cluster analysis of dose response gene expression after injection of SM at 24h. Gene significantly changed across doses after exposure were included in hierarchical cluster analysis using Ward's methods.

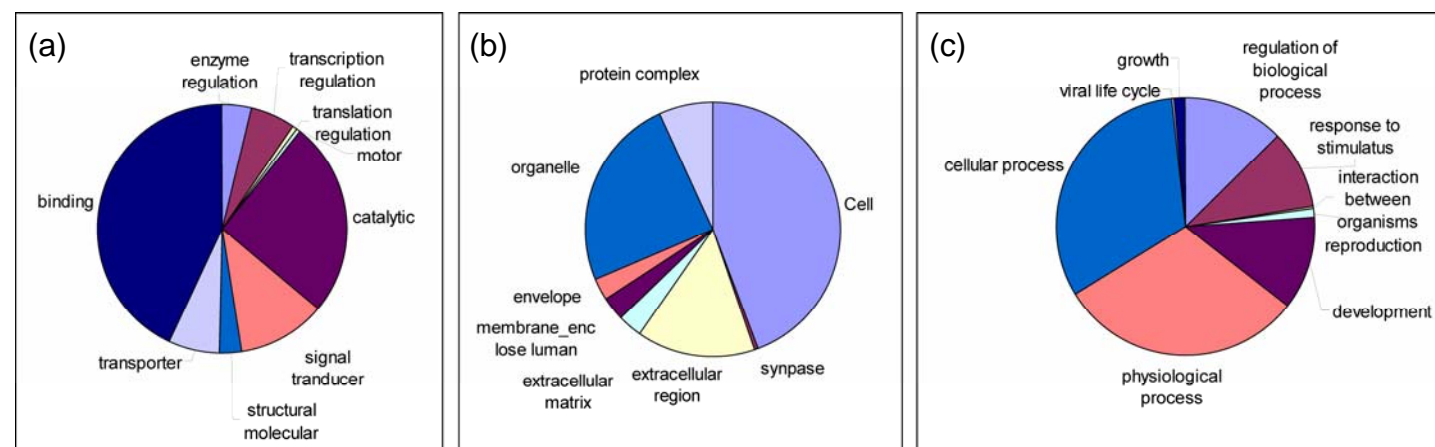


Principal component analysis (PCA) for gene expression in a time-dependent manner after treatment with SM at 6 mg/kg dose. (a) 3-D view. The autoscaled method was used for the PCA. The blue, red, green and aqua indicate the lung harvested at 0.5, 1, 3, 6 and 24 h post-dose respectively. PC1, PC2 and PC3 represent the first, second, and third principal component respectively. (b) Distribution of variance in principal components.



Hierarchical cluster analysis of altered gene expression in a time-dependent manner after injection of SM at a dose of 6 mg/kg. Gene significantly changed across time points after exposure were included in hierarchical cluster analysis using Ward's methods.

Gene Ontology for Function analysis (GOFA)³ of the dose response gene expression alteration after SM treatment



Gene ontology Biological Process, Molecular Function and Cellular Component associated with SM-regulated gene expression sets. Pie charts showing GO Biological Process: (a) Molecular Function; (b) Cellular Component; (c) Annotations associated with SM-regulated gene expression sets. Size of slices is indicative of number of items with that annotation.

List of significant biological processes generated with GOFA Terms in lung of rats treated with SM

GO term	GO ID	P value (average)	Number of gene involved
immune response	GO:0006955	0.000001	60
defense response	GO:0006952	0.000002	61
response to biotic stimulus	GO:0009607	0.000002	64
transmembrane receptor protein serine/threonine kinase signaling pathway	GO:0007178	0.000044	14
organ development	GO:0048513	0.000046	83
localization of cell	GO:0051674	0.000058	38
cell motility	GO:0006928	0.000058	38
locomotion	GO:0040011	0.000082	38
humoral defense mechanism (sensu Vertebrata)	GO:0016064	0.000152	13
release of cytochrome c from mitochondria	GO:0001836	0.000157	4
monovalent inorganic cation homeostasis	GO:0030004	0.000206	6
development	GO:0007275	0.000212	142
cell adhesion	GO:0007155	0.000231	39
response to pest, pathogen or parasite	GO:0009613	0.000409	39
transforming growth factor beta receptor signaling pathway	GO:0007179	0.000511	10
BMP signaling pathway	GO:0030509	0.000525	6
positive regulation of I-kappaB kinase/NF-kappaB cascade	GO:0043123	0.000755	9
response to other organism	GO:0051707	0.000906	39
regulation of I-kappaB kinase/NF-kappaB cascade	GO:0043122	0.000949	9
T cell activation	GO:0042110	0.00104	13
cell migration	GO:0016477	0.001081	26
regulation of immune response	GO:0050776	0.001256	16
lymphocyte activation	GO:0046649	0.001416	16
ATP metabolism	GO:0046034	0.00194	5
cytokine and chemokine mediated signaling pathway	GO:0019221	0.002054	7
angiogenesis	GO:0001525	0.002426	15
death	GO:0016265	0.00252	49
cell death	GO:0008219	0.00252	49
morphogenesis	GO:0009653	0.0028	70
muscle development	GO:0007517	0.002801	17
ribonucleoside triphosphate metabolism	GO:0009199	0.002831	5
purine ribonucleoside triphosphate metabolism	GO:0009205	0.002831	5
positive regulation of immune response	GO:0050778	0.00289	11
humoral immune response	GO:0006959	0.003292	14
enzyme linked receptor protein signaling pathway	GO:0007167	0.003348	24
apoptosis	GO:0008151	0.00347	47
response to stress	GO:0006950	0.003472	73
regulation of apoptosis	GO:0042981	0.003568	38
vasculature development	GO:0001944	0.003595	19
I-kappaB kinase/NF-kappaB cascade	GO:0007249	0.003864	10
pathogenesis	GO:0009405	0.003919	3
purine nucleoside triphosphate metabolism	GO:0009144	0.003983	5
response to stimulus	GO:0050896	0.004283	118

List of significant cellular component generated with GOFA terms in lung of rats treated with SM

GO term	GO ID	P value (average)	Number of gene involved
extracellular matrix (sensu Metazoa)	GO:0005578	0	28
extracellular matrix	GO:0031012	0.000001	28
MHC protein complex	GO:0042611	0.000031	6
extracellular matrix part	GO:0044420	0.000037	45
extracellular region part	GO:0044421	0.000131	124
MHC class II protein complex	GO:0042613	0.000157	4
extracellular region	GO:0005576	0.000252	128
extracellular space	GO:0005615	0.000258	119
immunological synapse	GO:0001772	0.000365	8
collagen	GO:0005581	0.000581	7
plasma membrane	GO:0005886	0.001827	112
fibillar collagen	GO:0005583	0.004833	4
basement membrane	GO:0005604	0.004846	8
collagen type I	GO:0005584	0.005799	2
cell surface	GO:0009986	0.006149	16
plasma membrane part	GO:0044459	0.008136	89
actin filament	GO:0005884	0.009391	5

List of significant molecular function generated with GOFA terms in lung of rats treated with SM

GO term	GO ID	P value (average)	Number of gene involved
hematopoietin/interferon-class (D200-domain) cytokine receptor signal transducer activity	GO:0005062	0.000033	4
hydrolase activity	GO:0016787	0.002886	88
aminopeptidase activity	GO:0004177	0.003841	6
cation-transporting ATPase activity	GO:0019829	0.005436	5
lamin binding	GO:0005521	0.005799	2
hydrogen-transporting ATPase activity, rotational mechanism	GO:0048961	0.007395	3
cyclin-dependent protein kinase inhibitor activity	GO:0004861	0.007395	3
ATPase activity, coupled to transmembrane movement of substances	GO:0042626	0.00781	10
ATPase activity, coupled to movement of substances	GO:0043492	0.008879	10
ATPase activity, coupled to transmembrane movement of ions	GO:0042625	0.009414	8
hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances	GO:0016820	0.010058	10

References:

- Dillman, J. F., 3rd, C. S. Phillips, et al. (2005). "Genomic analysis of rodent pulmonary tissue following bis-(2-chloroethyl) sulfide exposure." *Chem Res Toxicol* 18(1): 28-34.
- Yu, X., W. C. Griffith, et al. (2006). "A system-based approach to interpret dose- and time-dependent microarray data: quantitative integration of gene ontology analysis for risk assessment." *Toxicol Sci* 92(2): 560-77.
- Sun H.M., Fang H.,(2006). "GOFA: gene ontology for function analysis- A FDA gene ontology tool for analysis of genomic and proteomic data" *BMC bioinformatics*, 7(Suppl 2): S23.