Published in final edited form as:
J Investig Med. 2009 October ; 57(7): 756–764. doi:10.231/JIM.0b013e3181b91a83.

Delineation of a Gene Network Underlying the Pulmonary Response to Oxidative Stress in Asthma

Robert J. Freishtat[1,2,3,*], **Angela S. Benton**[2], **Alan M. Watson**[2,4], **Zuyi Wang**[2,5], **Mary C. Rose**[2,3], and **Eric P. Hoffman**[2,3]

[1]Division of Emergency Medicine, Children's National Medical Center, 111 Michigan Avenue, NW, Washington, DC 20010, United States

[2]Research Center for Genetic Medicine Children's National Medical Center, 111 Michigan Avenue, NW, Washington, DC 20010, United States

[3]School of Medicine and Health Sciences, The George Washington University, 2300 I Street, NW, Washington, DC 20037, United States

[4]Institute of Biomedical Sciences, The George Washington University, 2300 I Street, NW, Washington, DC 20037, United States

[5]The Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, 4300 Wilson Blvd, Arlington, VA 22203, United States

Abstract

Cigarette smoke exposure induces a respiratory epithelial response that is mediated in part by oxidative stress. The contribution of oxidative stress to cigarette smoke-induced responses in asthmatic respiratory epithelium is not well understood. We sought to increase this understanding by employing data integration and systems biology approaches to publicly available microarray data deposited over the last several years. In this study, we analyzed 14 publicly available asthma- or tobacco-relevant data series and found four (two mouse and two human) that fulfilled adequate signal/noise thresholds using unsupervised clustering and F test statistics. Using significance filters and a four-way Venn diagram approach, we identified 26 overlapping genes in the epithelial transcriptional stress response to cigarette smoke and asthma. This test set corresponded to a 26-member gene/protein network containing 18 members that were highly-regulated in a fifth data series of direct lung oxidative stress. Of those network members, two stood out (i.e. tissue inhibitor of metalloproteinases 1 and thrombospondin 1) due to central location within the network and marked up-regulation sustained at later times in response to oxidative stress. These analyses identified key relationships and primary hypothetical targets for future studies of cigarette smoke-induced oxidative stress in asthma.

Keywords

Asthma; Microarray Analysis; Tobacco Smoke Pollution; Reactive Oxygen Species

INTRODUCTION

One of the most common environmental contributors to asthma is cigarette smoke, which has been strongly associated with both an asthma diagnosis and increased morbidity [1–5]. Cigarette smoke contains over 4,000 chemical agents that induce a variety of physiological

*Corresponding Author (RJF); rfreishtat@cnmcresearch.org; Phone: (202) 476-2971; Fax: (202) 476-6014.

and biochemical responses. In particular, cigarette smoke contains high concentrations of oxidants that induce inflammatory cells (e.g. lung macrophages and neutrophils) to release reactive oxygen species [6] and promote inflammation via activation of transcription factors, such as NF-κB [7].

There is accumulating evidence that cigarette smoke also induces an oxidative stress response in respiratory epithelium [8]. Respiratory epithelium is now recognized as a key lung tissue in the pathogenesis of asthma given its simultaneous physical interaction with the airborne and tissue environments [9–11]. Importantly, asthmatic respiratory epithelium may be more susceptible to oxidative stress-induced apoptosis than normal epithelium [12].

We sought to increase the understanding of how oxidative stress impacts asthma- and cigarette smoke-related gene expression using data integration and systems biology approaches. To that end, we integrated multiple publicly available microarray data series to identify gene expression common to both asthmatic and cigarette smoke-exposed lung. This approach strengthens the validity of identified genes by improving signal-to-noise ratios and reduces false positives common in microarray experiments [13]. Herein, we describe the derivation of this list of commonly expressed genes and a resulting network of their known interactions. Further, we tested this network's response to oxidative stress in an additional gene expression data series to identify key relationships and primary targets for future study.

MATERIALS AND METHODS

Acquisition of Genome-wide Expression Profile Data

Fourteen genome-wide expression profile data series of human (i.e. Affymetrix U95A, U133A, and U133+2 platforms) or mouse (i.e. Affymetrix U74Av2, MOE430A, and 430A 2.0 platforms) lung tissues/cells were acquired from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) using keywords of "asthma" and/or "tobacco." Data series in GEO generated with early technology arrays or custom oligonucleotide arrays were not acquired due to their limited number of probe sets and/or incompatibility with our data integration plan. Hierarchical Clustering Explorer (HCE) 3.0 software (http://www.cs.umd.edu/hcil/hce/) was used for unsupervised clustering of these data.

Expression Signal Normalization and Venn Diagram Analysis

Affymetrix .Cel files were imported into GeneSpring GX10 (Agilent Technologies, Santa Clara, CA). Robust Multi-array Average (RMA) signals were generated and normalized to the median within each experiment. GeneSpring GX10 quality control algorithms (i.e. GAPDH 3'/5' ratio >3 and hybridization control outliers) were used to remove poor-quality samples from each experiment [only GSE994 had samples removed (n=26 out of 57 samples; Table 1)]. ANOVA analyses (for time-series or multiple arm studies) or T-tests were used to generate a list of differentially expressed transcripts ($p \leq 0.05$) between experimental conditions without multiple testing corrections but with a fold change ≥1.5 relative to controls. When specified, multiple comparisons were accounted for in GeneSpring GX10 using the Benjamini and Hochberg false discovery rate. To make comparisons between species, transcripts were translated between genomes using the interspecies homology feature of GeneSpring GX10 software.

Ingenuity Pathways Analysis

Networks were generated through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com). To begin, each gene identifier (i.e. Affymetrix Probe Set ID) was mapped to its corresponding gene in the Ingenuity Pathways Knowledge Base. These

genes included in the original probe set list are called focus genes. Focus genes were combined into a molecular network developed from information contained in the Ingenuity Pathways Knowledge Base.

An Ingenuity network is a graphical representation of the molecular relationships between genes/gene products. Genes or gene products are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least 1 reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Pathways Knowledge Base. Human, mouse, and rat orthologs of a gene are stored as separate objects in the Ingenuity Pathways Knowledge Base, but are represented as a single node in the network. The intensity of the node color indicates the degree of up- (red) or down- (green) regulation. Nodes are displayed using various shapes that represent the functional class of the gene product. Edges are displayed with various labels that describe the nature of the relationship between the nodes (e.g., binding, inhibition).

Ingenuity assigns a score to each network it derives only addresses the degree of relevance of the network to the genes in the dataset. The score takes into account the number of genes in the network and its size, as well as the total number of genes analyzed and the total number of molecules in Ingenuity's knowledge base that could potentially be included in networks. The network Score is then calculated using a hypergeometric distribution and a right-tailed Fisher's Exact Test. The resulting score is the negative log of this p-value. It should be noted that this score is not an indication of the quality or biological relevance of the derived network; rather it is simply a calculation of the approximate "fit" between the network and the network eligible molecules. So, network scores were used to focus on a particular network.

The most highly scored preliminary network was selected from those returned by the software. Gene groups in this preliminary network were exploded by adding the individual genes comprising those groups to the network and removing the group icon from the final network. Direct connections were built from the Ingenuity Pathways Knowledge Base while removing indirect connections and isolated or peripheral non-focus genes.

RESULTS

Selection of Asthma- and Tobacco- Relevant Data Series

Of the 14 available asthma- and/or tobacco-relevant genome-wide expression profile data series of human or mouse lung tissues/cells identified in GEO (Table 1), two (i.e. GSE5060 and GSE5372) were excluded from analysis due to technical issues with their data files and three were excluded due to the limited number of genes in their earlier technology platform (i.e. Affymetrix U74Av2). Of the remaining 9 data series, only four (i.e. GSE994, GSE1301, GSE3183, and GSE3184) provided acceptable signal/noise balance as defined by grouping of samples into experimentally-defined sets with F-measure ≥0.5 using methods that we have previously reported [13]. Two of these experiments were mouse models of allergic asthma induced by house dust mite allergen (GSE1301[14] and GSE3184). A third was expression profiling of human respiratory epithelium from smokers and non-smokers (GSE994) [15]. The fourth (GSE3183) was a temporal analysis of a human lung cancer epithelial cell line (i.e. A549 cells) following exposure to interleukin (IL)-13, the Th2 cytokine shown to be a central mediator of mouse allergic asthma [16, 17]. (Table 2)

Derivation of a Network of Commonly Expressed Genes in Asthmatic and Cigarette Smoke-Exposed Lung

Significantly- and highly-regulated ($p \leq 0.05$ with fold-change ≥1.5 up or down from control but not corrected for multiple comparisons) genes from each of the four data series were input into a four-way Venn diagram to identify those genes commonly expressed in asthmatic and cigarette smoke-exposed lung (Figure 1). The four-way Venn diagram contains five areas (identified by superscripts that precede the number of transcripts) representing overlap of at least three of the four data series. The shared response set included a total of 30 transcripts representative of asthma-relevant challenges (e.g. allergens, IL13, or cigarette smoke). The 30 overlapping transcripts corresponded to 26 genes (Table 3). There were no genes lost in our analysis due to lack of homologous genes during interspecies translation.

Ingenuity Pathways Analysis software was then used to identify known interactions between the genes in each overlapping Venn diagram area as described in the Methods section. The resultant top-scoring network for the overlapping genes is shown in Figure 2. The selected network of 26 genes had a score of 35. This corresponds to a Fisher Exact Test p-value of $1x10^{-35}$.

This network contains 26 directly linked members including 11 of the original shared response genes with four as focus genes. These four focus genes represent secreted/ extracellular gene products [i.e. tissue inhibitor of metalloproteinase (TIMP) -1, connective tissue growth factor (CTGF), and insulin-like growth factor binding protein (IGFBP) -3] and a nuclear factor [i.e. endothelial PAS domain protein (EPAS) -1].

Derived Network Response to Oxidative Stress

We then analyzed expression of the members of this 26 gene network in an additional GEO data series, GSE2565 [18]. GSE2565 was identified via a GEO search for “oxidative stress” in lung tissues. Although 17 data series were identified in GEO by these search criteria, only GSE2565 is from non-cancerous, non-immunologic cells or tissues. A further advantage is that this data series contains extensive longitudinal data [i.e. a nine point 72-hour time series following *in vivo* exposure to phosgene (i.e. carbonyl chloride)]. Because phosgene is a potent oxidative agent, it provides a means to test the derived network under oxidative conditions, which can then be extrapolated to cigarette smoke exposure.

The 26 human network genes corresponded to 42 homologous mouse transcripts. We focused on the 30 that showed statistically significant ($p \leq 0.05$ after correcting for multiple comparisons) up- or down-regulation over time in this mouse phosgene data series (Table 4). Inspection of these 30 transcripts in a time-series hierarchical gene cluster showed coordinated expression characterized by several distinct patterns, including early, sustained, and late up- or down-regulation (Figure 3). In particular, 18 genes showed ≥1.5-fold up- or down-regulation at one or more time points relative to time=0, as shown in Figure 4. Notable expression changes included: activating transcription factor 3 (ATF3) which was markedly up-regulated at early times and returned to baseline levels; thrombospondin 1 (THBS1) which was up-regulated early and its expression remained elevated; matrix metalloproteinase 3 (MMP3) which increased only moderately while its antagonist, TIMP1, was markedly increased over time; and IGFBP3 was markedly decreased at later times.

DISCUSSION

In this study, we analyzed 14 publicly available asthma- or tobacco-relevant data series and found four (two mouse and two human) that fulfilled adequate signal/noise thresholds using unsupervised clustering and F test statistics [13]. Using significance filters and a four-way

Venn diagram approach, we identified 26 overlapping genes (i.e. 30 transcripts) in the epithelial transcriptional stress response to cigarette smoke and asthma. This test set corresponded to a 26-member gene/protein network containing 18 members that were highly-regulated in a fifth data series of direct lung oxidative stress. Bioinformatic analyses of these five microarray projects involving two species and model systems defined a network of genes that includes a high density of members regulated by oxidative stress that are also important in the pulmonary response to IL13 (induced by allergens) and cigarette smoke.

Of those network members, two stand out (i.e. TIMP1 and THBS1) due to central location within the network and marked up-regulation sustained at later times in response to oxidative stress. Of note, both of these are also present in a recently published network of genes important in the pathogenesis of COPD, a disease characterized by marked pulmonary fibrosis from cigarette smoke exposure [19].

The first network member, TIMP1, is one of four protease inhibitors that regulates the matrix metalloproteinases (MMP) induced by lung inflammation following exposure to allergens, cigarette smoke, or oxidative stress [20]. TIMP1, a secreted 31 kDa glycoprotein, has been described to inhibit the protease activity of all known MMPs [21]. Notably, there is an abundance of MMP molecules in the network, including MMP1, MMP2, MMP3, MMP7, and MMP9. In particular, MMP9 is one of the most studied MMPs expressed in lung [22, 23].

In lung disease, the ratio of MMPs to TIMP1 is considered more important than expression of either alone. For example, MMP9 (and MMP2) are important for inflammatory cell infiltration into the lung and induction of airway hyperresponsiveness in asthma [24]. Further, sputum MMP9 increases after inhaled allergen exposure in allergic asthmatic patients while there was no change in sputum TIMP1 resulting in an increased MMP9:TIMP1 ratio [25]. The effect of altering this ratio can be seen upon allergen challenge in MMP9 knockout mice, which have less mononuclear cell infiltration of the airways than wild type mice [26]. Conversely, allergen-challenged TIMP1 knockout mice have increased cellular infiltration and decreased lung function than wild type mice [27].

In our analyses, TIMP1 was highly up-regulated in response to oxidative stress while the MMP's were either only moderately up-regulated (i.e. MMP3) or down-regulated. This could be expected to result in lower MMP:TIMP1 ratios, which aside from the above described effects, have also been hypothesized to contribute to extracellular airway remodeling and fibrosis [20].

Another of the centrally-located highly up-regulated network members is THBS1. THBS1 is a 140 kDa extracellular matrix protein that mediates cell-to-cell and cell-to-matrix interactions. Interestingly, THBS1 activates and is itself activated by MMPs [28]. Further, THBS1 is crucial for activating transforming growth factor (TGF)β1[29], an important pro-fibrotic agent, which was down-regulated in our network.

Mouse models of asthma have limitations with regard to translation of results to human asthma. However, they do share similarities and have been invaluable in identifying targets for human study. In particular, the house dust mite model used in the data series included in our study is particularly applicable to human asthma among murine models [30]. Therefore, the inclusion of these experiments in this study, which is ultimately directed toward identifying targets for human validation, is useful. Additionally, interspecies commonality (i.e. evolutionary conservation) of responses to external/environmental stimuli/challenges lends credibility to those responses

The inclusion of GSE3183 in our analysis was initially of concern as A549 cells are a lung cancer cell line that exhibit many properties of respiratory epithelium, including mucin secretion [30], but are often thought to be alveolar in origin. However, the differentially-regulated transcripts from this data series had similar inter-experimental overlap to the other included data series.

In these experiments, we used data integration and systems biology analyses of publicly-available microarray data to derive a list of commonly expressed genes important in asthmatic and cigarette smoke-exposed lung. We then identified a network of their known interactions and tested this network's response to oxidative stress in an additional gene expression data series. These analyses identified key relationships and primary hypothetical targets for future studies of cigarette smoke-induced oxidative stress in asthma.

Acknowledgments

GRANTS

Funding support provided by the National Institutes of Health, Bethesda, Maryland, USA via grant K23-RR-020069 to RJF and institutional funding from Children's National Medical Center.

References

1. Chen Y, Rennie DC, Lockinger LA, Dosman JA. Gender, environmental tobacco smoke, and pulmonary function in rural children and adolescents: the Humboldt study. J Agric Saf Health. 2005; 11:167–73. [PubMed: 15931942]
2. Gilliland FD, Islam T, Berhane K, et al. Regular smoking and asthma incidence in adolescents. Am J Respir Crit Care Med. 2006; 174:1094–100. [PubMed: 16973983]
3. Selgrade MK, Lemanske RF Jr, Gilmour MI, et al. Induction of asthma and the environment: what we know and need to know. Environ Health Perspect. 2006; 114:615–9. [PubMed: 16581555]
4. Sturm JJ, Yeatts K, Loomis D. Effects of tobacco smoke exposure on asthma prevalence and medical care use in North Carolina middle school children. Am J Public Health. 2004; 94:308–13. [PubMed: 14759947]
5. Cunningham J, O'Connor GT, Dockery DW, Speizer FE. Environmental tobacco smoke, wheezing, and asthma in children in 24 communities. Am J Respir Crit Care Med. 1996; 153:218–24. [PubMed: 8542119]
6. Rahman I, MacNee W. Role of oxidants/antioxidants in smoking-induced lung diseases. Free Radic Biol Med. 1996; 21:669–81. [PubMed: 8891669]
7. Yang S-R, Chida AS, Bauter MR, et al. Cigarette smoke induces proinflammatory cytokine release by activation of NF-{kappa}B and posttranslational modifications of histone deacetylase in macrophages. Am J Physiol Lung Cell Mol Physiol. 2006; 291:L46–57. [PubMed: 16473865]
8. Mossman BT, Lounsbury KM, Reddy SP. Oxidants and Signaling by Mitogen-Activated Protein Kinases in Lung Epithelium. Am J Respir Cell Mol Biol. 2006; 34:666–9. [PubMed: 16484683]
9. Holgate ST, Lackie PM, Davies DE, Roche WR, Walls AF. The bronchial epithelium as a key regulator of airway inflammation and remodelling in asthma. Clin Exp Allergy. 1999; 29 (Suppl 2): 90–5. [PubMed: 10421830]
10. Kicic A, Sutanto EN, Stevens PT, Knight DA, Stick SM. Intrinsic biochemical and functional differences in bronchial epithelial cells of children with asthma. Am J Respir Crit Care Med. 2006; 174:1110–8. [PubMed: 16908868]
11. Davies DE, Holgate ST. Asthma: the importance of epithelial mesenchymal communication in pathogenesis. Inflammation and the airway epithelium in asthma. Int J Biochem Cell Biol. 2002; 34:1520–6. [PubMed: 12379273]
12. Bucchieri F, Puddicombe SM, Lordan JL, et al. Asthmatic Bronchial Epithelium Is More Susceptible to Oxidant-Induced Apoptosis. Am J Respir Cell Mol Biol. 2002; 27:179–85. [PubMed: 12151309]

13. Seo J, Bakay M, Chen Y-W, Hilmer S, Shneiderman B, Hoffman EP. Interactively optimizing signal-to-noise ratios in expression profiling: project-specific algorithm selection and detection p-value weighting in Affymetrix microarrays. Bioinformatics. 2004; 20:2534–44. [PubMed: 15117752]

14. Christina CL, Bruce A, John H, et al. Unique and overlapping gene expression patterns driven by IL-4 and IL-13 in the mouse lung. The Journal of allergy and clinical immunology. 2009; 123:795–804.e8. [PubMed: 19249085]

15. Spira A, Beane J, Shah V, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. Proceedings of the National Academy of Sciences. 2004; 101:10143–8.

16. Wills-Karp M, Luyimbazi J, Xu X, et al. Interleukin-13: central mediator of allergic asthma. Science. 1998; 282:2258–61. [PubMed: 9856949]

17. Grunig G, Warnock M, Wakil AE, et al. Requirement for IL-13 independently of IL-4 in experimental asthma. Science. 1998; 282:2261–3. [PubMed: 9856950]

18. Sciuto AM, Phillips CS, Orzolek LD, Hege AI, Moran TS, Dillman JF. Genomic Analysis of Murine Pulmonary Tissue Following Carbonyl Chloride Inhalation. Chem Res Toxicol. 2005; 18:1654–60. [PubMed: 16300373]

19. Wang IM, Stepaniants S, Boie Y, et al. Gene Expression Profiling in Patients with Chronic Obstructive Pulmonary Disease and Lung Cancer. Am J Respir Crit Care Med. 2008; 177:402–11. [PubMed: 17975202]

20. Atkinson JJ, Senior RM. Matrix metalloproteinase-9 in lung remodeling. Am J Respir Cell Mol Biol. 2003; 28:12–24. [PubMed: 12495928]

21. Gomez DE, Alonso DF, Yoshiji H, Thorgeirsson UP. Tissue inhibitors of metalloproteinases: structure, regulation and biological functions. Eur J Cell Biol. 1997; 74:111–22. [PubMed: 9352216]

22. Tanaka H, Miyazaki N, Oashi K, Tanaka S, Ohmichi M, Abe S. Sputum matrix metalloproteinase-9: tissue inhibitor of metalloproteinase-1 ratio in acute asthma. J Allergy Clin Immunol. 2000; 105:900–5. [PubMed: 10808169]

23. Parks WC, Shapiro SD. Matrix metalloproteinases in lung biology. Respir Res. 2001; 2:10–9. [PubMed: 11686860]

24. Kumagai K, Ohno I, Okada S, et al. Inhibition of matrix metalloproteinases prevents allergen-induced airway inflammation in a murine model of asthma. J Immunol. 1999; 162:4212–9. [PubMed: 10201949]

25. Cataldo DD, Bettiol J, Noel A, Bartsch P, Foidart JM, Louis R. Matrix metalloproteinase-9, but not tissue inhibitor of matrix metalloproteinase-1, increases in the sputum from allergic asthmatic patients after allergen challenge. Chest. 2002; 122:1553–9. [PubMed: 12426252]

26. Cataldo DD, Tournoy KG, Vermaelen K, et al. Matrix metalloproteinase-9 deficiency impairs cellular infiltration and bronchial hyperresponsiveness during allergen-induced airway inflammation. Am J Pathol. 2002; 161:491–8. [PubMed: 12163374]

27. Sands MF, Ohtake PJ, Mahajan SD, et al. Tissue inhibitor of metalloproteinase-1 modulates allergic lung inflammation in murine asthma. Clin Immunol. 2009; 130:186–98. [PubMed: 18955015]

28. Qian X, Wang TN, Rothman VL, Nicosia RF, Tuszynski GP. Thrombospondin-1 modulates angiogenesis in vitro by up-regulation of matrix metalloproteinase-9 in endothelial cells. Exp Cell Res. 1997; 235:403–12. [PubMed: 9299165]

29. Crawford SE, Stellmach V, Murphy-Ullrich JE, et al. Thrombospondin-1 Is a Major Activator of TGF-21 In Vivo. 1998; 93:1159–70.

30. Johnson JR, Wiley RE, Fattouh R, et al. Continuous Exposure to House Dust Mite Elicits Chronic Airway Inflammation and Structural Remodeling. Am J Respir Crit Care Med. 2004; 169:378–85. [PubMed: 14597485]

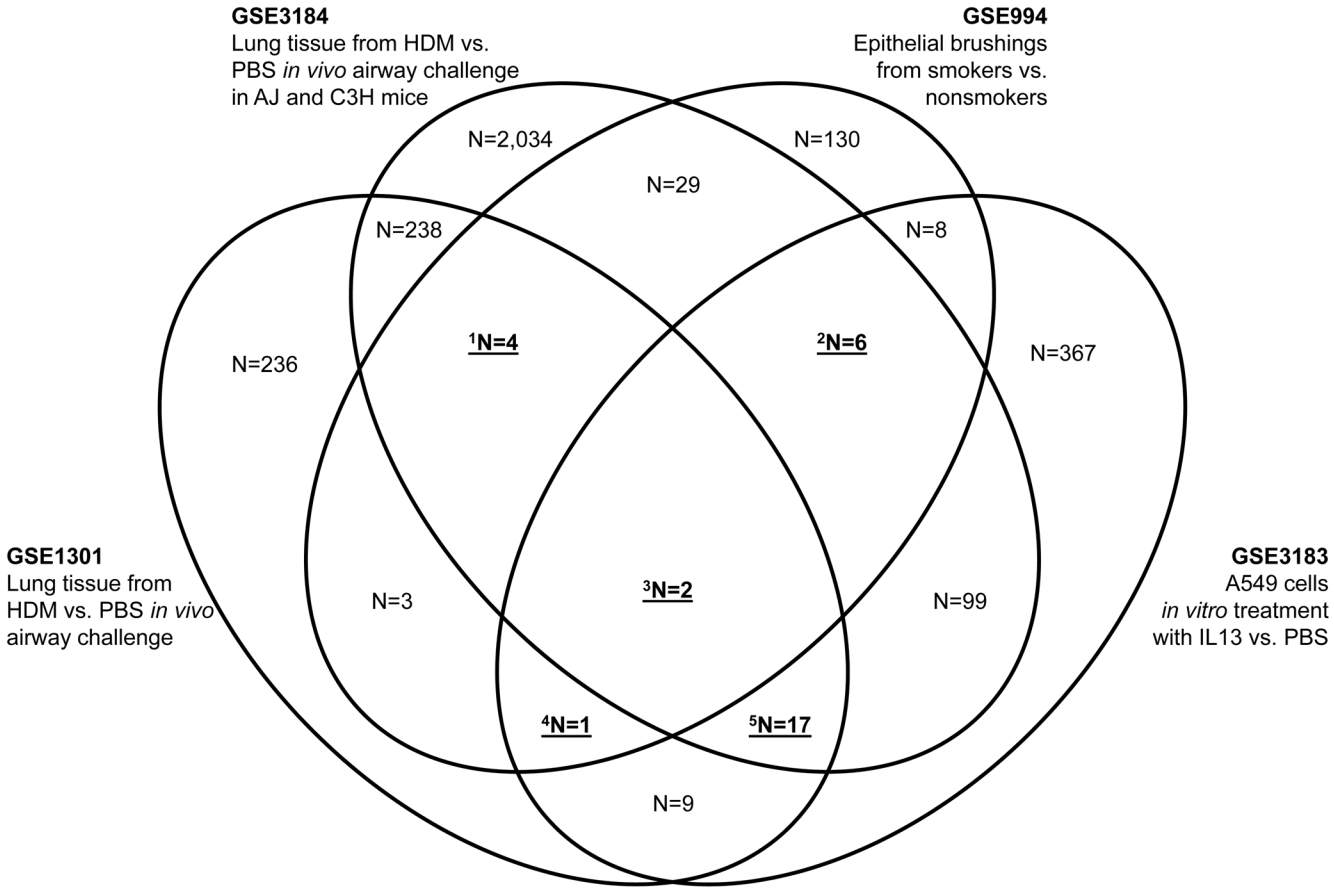

Figure 1.

Venn diagram showing gene expression data integration and interpretation of asthma- and tobacco-relevant data series. As described in detail in the text, the four selected data series identified using appropriate unsupervised clustering and F≥0.5 were used in a four-way Venn diagram to generate a 26 gene (i.e. 30 transcripts) shared response set that showed time- and/or dose-related responses to asthma-relevant challenges. The hypothetical shared response set was comprised of the five Venn diagram areas that are designated by a bold N and a preceding superscript numeral 1–5. These areas represent overlap of at least three of the four data series.

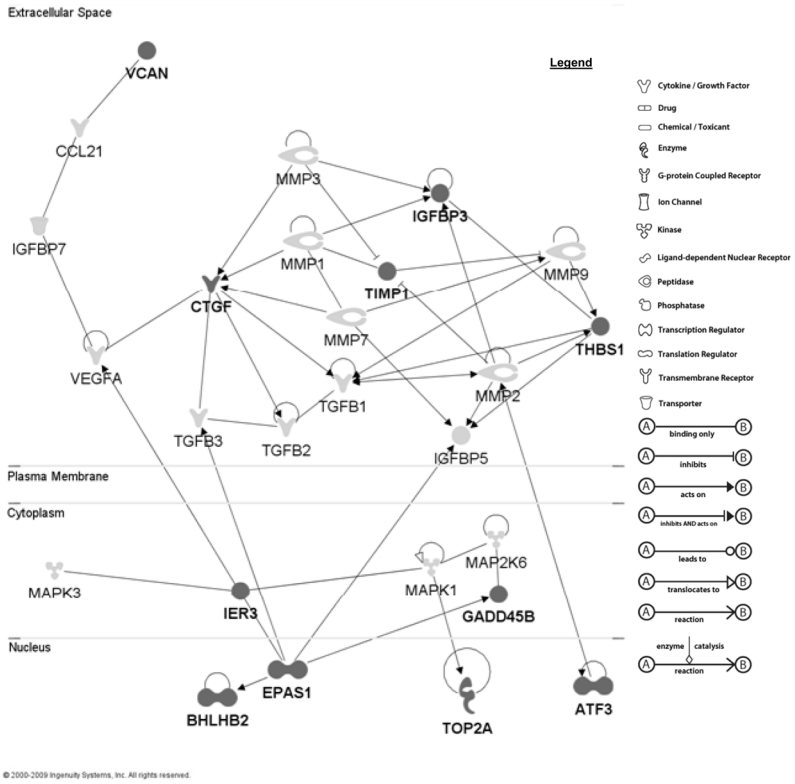

Figure 2.

A network of commonly expressed genes in asthmatic and cigarette smoke-exposed lung. Ingenuity Pathways Analysis software was used to explore the prominent molecular networks represented in the 30 core transcripts (Figure 1, Table 2). The refined (i.e. groups exploded, indirect relationships and disconnected/peripheral non-focus members removed) top-scoring network contains 26 directly linked members centered on the focus genes for tissue inhibitor of metalloproteinases (TIMP)-1, connective tissue growth factor (CTGF), endothelial PAS domain protein 1 (EPAS1) and insulin-like growth factor binding protein 3 (IGFBP3). The genes that are present in the 30 core transcript group (i.e. focus genes) are colored dark gray.

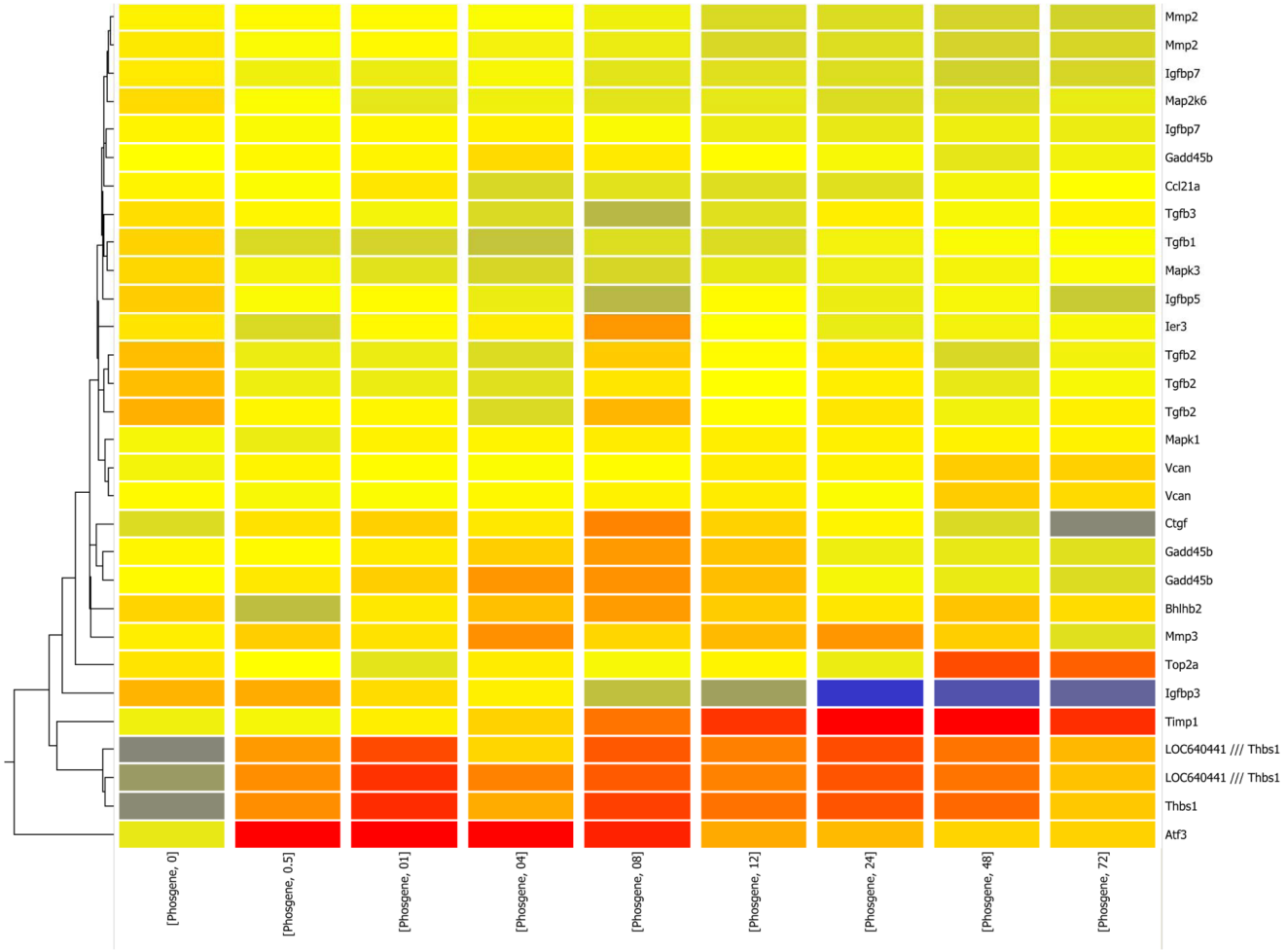

Figure 3.

Time-series hierarchical cluster of phosgene-induced oxidative stress in mouse lung over 72 hours (GSE2565). Mice were exposed to phosgene at time 0 and sacrificed at eight post-exposure time points (0, 0.5, 1, 4, 8, 12, 24, 48, or 72 hours). Lung tissue was isolated and RNA extracted for whole genome expression profiling with the Affymetrix MOE430A microarray platform [18]. A time-series gene cluster was generated in GeneSpring GX10 for the 30 of 42 transcripts (corresponding to 26 network genes) that showed statistically significant (corrected $p \leq 0.05$) up- (red) or down- (blue) regulation over time. Inspection of these 30 transcripts showed coordinated expression characterized by several distinct patterns, including early, sustained, and late up- or down-regulation.

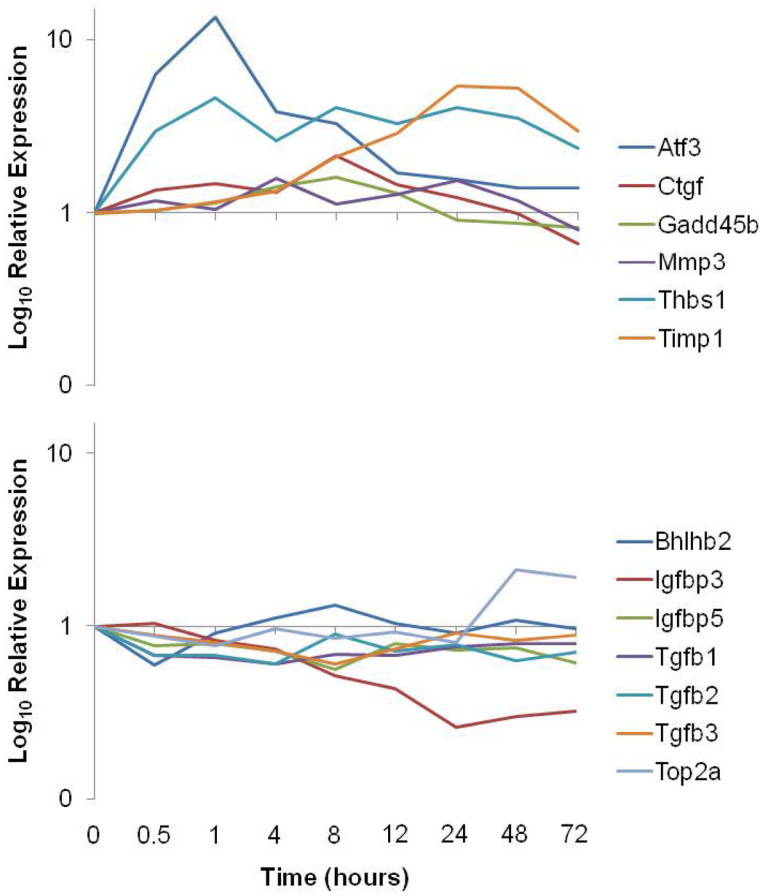

Figure 4.

Highly up- or down- regulated genes in phosgene-induced oxidative stress in mouse lung. Eighteen genes (Figure 3) showed ≥1.5-fold up- or down-regulation at one or more time points relative to time=0 in GSE2565. The graph is split into up-regulated (top) and down-regulated (bottom) genes for clarity. Expression shown for genes with multiple probe sets reflects averaging within each time point.

Table 1

Results of unsupervised clustering for asthma- and cigarette smoke-relevant data series

GEO Record	Contributor	Experiment	Affymetrix Platform	Number of Microarray	Signal-to-Noise Assessment (F-measure [13])	Well-Clustered by Phenotype	Included
GSE470	Spannhake W	Asthmatic vs. normal airway epithelium in vitro challenge with ozone or rhinovirus	U95A	12	0.653	No	No
GSE476	Kleeberger S	Lung tissue from ozone vs. PBS *in vivo* airway challenge in HeJ and OuJ mice	U74Av2	8	0.667	Yes	No
GSE481	Rose M	Lung tissue from IL13 vs. PBS *in vivo* airway challenge	U74Av2	5	0.867	Yes	No
GSE483	Wills-Karp M	Lung tissue from Ragweed vs. PBS *in vivo* airway challenge in BALB/c mice	U74Av2	22	0.867	Yes	No
GSE994	Spira A [15]	Epithelial brushings from smokers vs. nonsmokers	U133A	57	0.705	Yes	Yes
GSE1301	Wills-Karp M 14	Lung tissue from HDM vs. PBS *in vivo* airway challenge	MOE430A	12	0.867	Yes	Yes
GSE3004	Lilly CM	Bronchial epithelium from HDM and cat allergen challenge *in vivo*	U95A	10	0.667	No	No
GSE3183	Scott A	A549 cells with *in vitro* IL13 vs. PBS treatment	U133A	15	0.787	Yes	Yes
GSE3184	Wills-Karp M	Lung tissue from HDM vs. PBS *in vivo* airway challenge in AJ and C3H mice	MOE430A	20	0.821	Yes	Yes
GSE3320	Harvery BG	Epithelial brushings from smokers vs. nonsmokers	U133A	11	0.789	No	No
GSE4498	Heguy A	Epithelial brushings from smokers vs. nonsmokers	U133+2	22	0.72	No	No
GSE5060	Carolan B	Epithelial brushings from smokers vs. nonsmokers	U133+2 and U95A	10	Technical problem		No
GSE5372	Heguy A	Epithelial brushings pre- and post- *in vivo* mechanical injury	U133+2	22	Technical problem		No
GSE6858	Lu X	OVA vs. PBS airway challenge in BALBc and RAG mice	Mouse 430A 2.0	16	0.601	No	No

Table 2

Results of unsupervised clustering for asthma- and cigarette smoke-relevant data series

GEO Record	Organism and Tissue	Experimental Conditions	Number of Microarrays	Probe Sets Screened	Significantly-Regulated Probe Sets (i.e. uncorrected p ≤ 0.05) versus Control Condition	
GSE994	Human lung epithelial cell brushings	Current smokers	19	22,283	3,335	
		Non-smokers (control)	12			
GSE1301	Murine (BALB/c strain) whole lung tissue	House dust mite airway challenge *in vivo*	3	22,690	4,346	
		Phosphate-buffered saline airway challenge *in vivo* (control)	3			
GSE3183	Human A549 cell line	IL13 *in vitro* - 4 hours post exposure	3	22,283	4,931	
		IL13 *in vitro* - 12 hours post exposure	3			
		IL13 *in vitro* - 24 hours post exposure	3			
		Phosphate-buffered saline *in vitro* (control)	6			
GSE3184	Murine (AJ strain) whole lung tissue	House dust mite airway challenge *in vivo* 6 hours post exposure	5	22,690	8,574	4,426 common probe sets
		House dust mite airway challenge *in vivo* 24 hours post exposure	5			
		Phosphate-buffered saline airway challenge *in vivo* (control)	5			
	Murine (C3H strain) whole lung tissue	House dust mite airway challenge *in vivo* 6 hours post exposure	5		7,766	
		House dust mite airway challenge *in vivo* 24 hours post exposure	5			
		Phosphate-buffered saline airway challenge *in vivo* (control)	5			

Table 3

Overlapping Transcripts (N=30) from Venn Diagram Analysis

Probe Set ID	Gene Symbol	Gene Title	Entrez Gene ID	Venn Diagram Area
1419268_at	AGR2	Anterior gradient 2 (Xenopus laevis)	23795	1
1449363_at	ATF3	Activating transcription factor 3	11910	5
1418025_at	BHLHB2	Basic helix-loop-helix domain containing, class B2	20893	5
1423954_at	C3	Complement component 3	12266	2
1422437_at	COL5A2	Collagen, type V, alpha 2	12832	5
1416953_at	CTGF	Connective tissue growth factor	14219	3
1416612_at	CYP1B1	Cytochrome P450, family 1, subfamily b, polypeptide 1	13078	2
1449888_at	EPAS1	Endothelial PAS domain protein 1	13819	2
1449773_s_at	GADD45B	Growth arrest and DNA-damage-inducible 45 beta	17873	5
1455959_s_at	GCLC	Glutamate-cysteine ligase, catalytic subunit	14629	1
1418949_at	GDF15	Growth differentiation factor 15	23886	2
1419905_s_at	HPGD	Hydroxyprostaglandin dehydrogenase 15 (NAD)	15446	2
1417101_at	HSPA2	Heat shock protein 2	15512	2
1419647_a_at	IER3	Immediate early response 3	15937	5
1423062_at	IGFBP3	Insulin-like growth factor binding protein 3	16009	5
1417394_at	KLF4	Kruppel-like factor 4 (gut)	16600	4
1417395_at	KLF4	Kruppel-like factor 4 (gut)	16600	3
1422962_a_at	PSMB8	Proteasome (prosome, macropain) subunit, beta type 8 (large multifunctional peptidase 7)	16913	5
1450699_at	SELENBP1	Selenium binding protein 1	20341	5
1448395_at	SFRP1	Secreted frizzled-related protein 1	20377	5
1423505_at	TAGLN	Transgelin	21345	5
1450377_at	THBS1	Thrombospondin 1	21825	5
1421811_at	THBS1	Thrombospondin 1	21825	5
1460302_at	THBS1	Thrombospondin 1	21825	5
1460227_at	TIMP1	Tissue inhibitor of metalloproteinase 1	21857	1
1454694_a_at	TOP2A	Topoisomerase (DNA) II alpha	21973	5
1460629_at	TRIM16	Tripartite motif-containing 16	94092	1
1420772_a_at	TSC22D3	TSC22 domain family 3	14605	5
1425281_a_at	TSC22D3	TSC22 domain family 3	14605	5
1427256_at	VCAN	Versican	13003	5

Table 4

Significantly-expressed transcripts (N=30) as a function of time in GSE2565 representing 26 genes in the derived network

Probe Set ID	Gene Symbol	Gene Title	Entrez Gene ID	0.5	1	4	8	12	24	48	72	Corrected* P-value
1449363_at	ATF3	Activating transcription factor 3	11910	6.36	13.57	3.87	3.30	1.70	1.56	1.38	1.39	1.72E-22
1418025_at	BHLHB2	Basic helix-loop-helix domain containing, class B2	20893	0.60	0.91	1.11	1.32	1.04	0.92	1.08	0.96	3.62E-10
1419426_s_at	CCL21A	Chemokine (C-C motif) ligand 21a	18829	0.95	1.08	0.79	0.83	0.81	0.82	0.91	0.96	3.44E-03
1416953_at	CTGF	Connective tissue growth factor	14219	1.35	1.48	1.32	2.14	1.46	1.23	0.98	0.66	6.90E-12
1450971_at	GADD45B	Growth arrest and DNA-damage-inducible 45 beta	17873	0.98	1.06	1.22	1.57	1.27	0.89	0.86	0.83	6.81E-06
1449773_s_at	GADD45B	Growth arrest and DNA-damage-inducible 45 beta	17873	1.09	1.24	1.63	1.67	1.33	0.94	0.89	0.83	7.17E-10
1420197_at	GADD45B	Growth arrest and DNA-damage-inducible 45 beta	17873	1.03	1.05	1.20	1.11	1.02	0.97	0.89	0.95	9.45E-03
1419647_a_at	IER3	Immediate early response 3	15937	0.73	0.90	0.96	1.46	0.88	0.80	0.82	0.85	3.97E-05
1423062_at	IGFBP3	Insulin-like growth factor binding protein 3	16009	1.04	0.82	0.74	0.51	0.44	0.26	0.30	0.32	5.13E-12
1452114_s_at	IGFBP5	Insulin-like growth factor binding protein 5	16011	0.77	0.79	0.72	0.56	0.80	0.72	0.75	0.61	4.06E-04
1423584_at	IGFBP7	Insulin-like growth factor binding protein 7	29817	0.85	0.83	0.89	0.80	0.78	0.78	0.73	0.74	2.37E-06
1423585_at	IGFBP7	Insulin-like growth factor binding protein 7	29817	0.93	0.99	1.01	0.93	0.87	0.85	0.88	0.87	2.16E-02
1426850_a_at	MAP2K6	Mitogen-activated protein kinase kinase 6	26399	0.84	0.75	0.79	0.74	0.74	0.71	0.72	0.76	2.18E-08
1419568_at	MAPK1	Mitogen-activated protein kinase 1	26413	0.96	1.11	1.10	1.14	1.13	1.12	1.11	1.11	1.44E-02
1427060_at	MAPK3	Mitogen-activated protein kinase 3	26417	0.78	0.71	0.67	0.67	0.73	0.75	0.78	0.80	2.75E-06
1416136_at	MMP2	Matrix metallopeptidase 2	17390	0.89	0.92	0.85	0.83	0.75	0.77	0.73	0.74	7.45E-07
1439364_a_at	MMP2	Matrix metallopeptidase 2	17390	0.97	0.97	0.94	0.88	0.79	0.80	0.77	0.76	4.85E-07
1418945_at	MMP3	Matrix metallopeptidase 3	17392	1.17	1.06	1.59	1.12	1.29	1.55	1.17	0.80	2.99E-03
1420653_at	TGFB1	Transforming growth factor, beta 1	21803	0.68	0.66	0.60	0.69	0.68	0.76	0.80	0.80	1.24E-06
1450922_a_at	TGFB2	Transforming growth factor, beta 2	21808	0.67	0.67	0.62	0.94	0.74	0.81	0.61	0.68	2.06E-07
1450923_at	TGFB2	Transforming growth factor, beta 2	21808	0.67	0.67	0.63	0.82	0.73	0.78	0.65	0.71	6.07E-06
1423250_a_at	TGFB2	Transforming growth factor, beta 2	21808	0.71	0.71	0.57	0.97	0.69	0.77	0.64	0.73	3.65E-07
1417455_at	TGFB3	Transforming growth factor, beta 3	21809	0.89	0.81	0.72	0.60	0.73	0.92	0.83	0.89	2.18E-05
1450377_at	THBS1	Thrombospondin 1	21825	2.96	4.38	2.20	4.08	3.36	4.34	3.54	2.53	5.23E-05
1421811_at	THBS1	Thrombospondin 1	21825	2.82	4.45	2.99	3.66	3.00	3.78	3.23	2.18	1.53E-09
1460302_at	THBS1	Thrombospondin 1	21825	3.05	4.96	2.64	4.47	3.51	4.07	3.67	2.31	1.83E-07
1460227_at	TIMP1	Tissue inhibitor of metalloproteinase 1	21857	1.02	1.15	1.32	2.10	2.90	5.41	5.25	2.98	4.47E-15

Probe Set ID	Gene Symbol	Gene Title	Entrez Gene ID	0.5	1	4	8	12	24	48	72	Corrected* P-value
1454694_a_at	TOP2A	Topoisomerase (DNA) II alpha	21973	0.88	0.77	0.96	0.85	0.93	0.80	2.12	1.92	2.58E-10
1427256_at	VCAN	Versican	13003	0.95	0.97	1.02	1.04	1.07	0.97	1.26	1.17	7.72E-04
1421694_a_at	VCAN	Versican	13003	1.10	1.06	1.04	1.06	1.15	1.11	1.35	1.32	1.09E-04

Probe sets in **bold** are highly (≥ 1.5 fold) up- or down-regulated from time=0 in GSE2565 [17].

*P-values corrected for multiple comparisons using Benjamimi and Hochberg false discovery rate.