Gene 602 (2017) 16–23

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Carcinogenic effects of oil dispersants: A KEGG pathway-based RNA-seq study of human airway epithelial cells

Yao-Zhong Liu [a,*], Lei Zhang [b,1], Astrid M Roy-Engel [c], Shigeki Saito [d], Joseph A Lasky [d], Guangdi Wang [e], He Wang [f]

[a] Dept. of Global Biostatistics and Data Science, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
[b] Center for Genetic Epidemiology and Genomics, School of Public Health, Medical College, Soochow University, Su Zhou, China
[c] Dept. of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
[d] Dept. of Medicine, Tulane University School of Medicine, New Orleans, LA, USA
[e] Dept. of Chemistry, RCMI Cancer Research Center, Xavier University of Louisiana, New Orleans, LA, USA
[f] Dept. of Chronic Respiratory Diseases, School of Health Sciences, University of Newcastle, Callaghan, Australia

ARTICLE INFO

Article history:
Received 16 October 2016
Accepted 14 November 2016
Available online 16 November 2016

Keywords:
BP oil spill
RNA-seq
Airway epithelial cells
Dispersants
Oil
Carcinogenesis
GAGE
GSNCA
KEGG

ABSTRACT

The health impacts of the BP oil spill are yet to be further revealed as the toxicological effects of oil products and dispersants on human respiratory system may be latent and complex, and hence difficult to study and follow up. Here we performed RNA-seq analyses of a system of human airway epithelial cells treated with the BP crude oil and/or dispersants Corexit 9500 and Corexit 9527 that were used to help break up the oil spill. Based on the RNA-seq data, we then systemically analyzed the transcriptomic perturbations of the cells at the KEGG pathway level using two pathway-based analysis tools, GAGE (generally applicable gene set enrichment) and GSNCA (Gene Sets Net Correlations Analysis). Our results suggested a pattern of change towards carcinogenesis for the treated cells marked by upregulation of ribosomal biosynthesis (hsa03008) ($p = 1.97E-13$), protein processing (hsa04141) ($p = 4.09E-7$), Wnt signaling (hsa04310) ($p = 6.76E-3$), neurotrophin signaling (hsa04722) ($p = 7.73E-3$) and insulin signaling (hsa04910) ($p = 1.16E-2$) pathways under the dispersant Corexit 9527 treatment, as identified by GAGE analysis. Furthermore, through GSNCA analysis, we identified gene co-expression changes for several KEGG cancer pathways, including small cell lung cancer pathway (hsa05222, $p = 9.99E-5$), under various treatments of oil/dispersant, especially the mixture of oil and Corexit 9527. Overall, our results suggested carcinogenic effects of dispersants (in particular Corexit 9527) and their mixtures with the BP crude oil, and provided further support for more stringent safety precautions and regulations for operations involving long-term respiratory exposure to oil and dispersants.

1. Introduction

The year 2016 marked the sixth anniversary of the Deepwater Horizon oil spill, the largest man-made disaster in the history of petroleum industry. The sheer scale of this disaster is marked by its long duration (lasting for ~3 months), the gigantic volume of crude oil (210 million gal) spilled and dispersants (1.8 million gal) applied and the large number of workers (>50.000) involved in the cleaning operation (Hayworth and Clement, 2012; Kujawinski et al., 2011).

The full impacts of the disaster to the environment and marine and human lives have yet to be fully unveiled. Specifically, the long term health impacts of the BP oil spill to the >50.000 workers involved in the cleaning operation have not been well characterized and followed up, although limited data on other smaller scale oil spills (e.g., the Prestige oil spill) did suggest that involvement in oil spill cleaning operations may cause persistent respiratory symptoms (Zock et al., 2012), long-lasting airway oxidative stress (Rodriguez-Trigo et al., 2010), and systemic genetic effects (Laffon et al., 2006; Perez-Cadahia et al., 2007; Perez-Cadahia et al., 2008a; Perez-Cadahia et al., 2008b). In addition, the oil-dispersant mixtures may contain potentially mutagenic/carcinogenic chemicals including PAH, benzene, and benzene derivatives (Rodrigues et al., 2010; Saeed and Al-Mutairi, 1999). More importantly, chemical components in the mixtures may enhance each other to induce harmful effects synergistically. The mechanism for inhalation of hazardous substances during the oil spill was also proposed through models where inhalable aerosols that contain dispersed oil can be

Abbreviations: GAGE, generally applicable gene set enrichment; GSNCA, Gene Sets Net Correlations Analysis; WAF, water accommodated fraction; DAVID, database for annotation, visualization, and integrated discovery.

* Corresponding author at: Dept. of Global Biostatistics and Data Science, Tulane University School of Public Health and Tropical Medicine, 1440 Canal Street, Suite 2001, New Orleans, LA 70112, USA.
E-mail address: yliu8@tulane.edu (Y.-Z. Liu).

[1] Co-first author.

http://dx.doi.org/10.1016/j.gene.2016.11.028
0378-1119/

formed on the sea surface (Ehrenhauser et al., 2014; Middlebrook et al., 2012).

To characterize the effects and mechanisms of oil spill to human lung health at the molecular level, we hypothesize that oil spill chemicals (i.e., oil, dispersant, or their mixtures) may have significant effects on respiratory cells, which can be detected at the transcriptomic level. To test this hypothesis, we performed an RNA-seq study of human airway epithelial cells treated with BP crude oil, oil dispersants (Corexit 9500 and 9527) and their mixtures (Liu et al., 2016). Through that study we identified a large number of genes differentially expressed due to the treatments, suggesting significant transcriptomic perturbations of the cells caused by the toxicological effects of the oil and oil cleaning chemicals. The findings provided a strong support to our hypothesis. Furthermore, by annotating the differentially expressed genes using DAVID analysis (Dennis et al., 2003), our study (Liu et al., 2016) suggested several key biological processes affected by the chemicals, including degradation of the cell junction, enhanced immune response, decreased local steroid biosynthesis and enhanced angiogenesis. These identified biological processes are consistent with some of the pathological features for several common lung diseases, such as COPD (Faner et al., 2013; Holtzman et al., 2014), asthma (Ribatti et al., 2009) and cystic fibrosis (Georas and Rezaee, 2014; Heijink et al., 2014; Rezaee and Georas, 2014). Therefore our study (Liu et al., 2016) not only detected the existence but also preliminarily characterized potential molecular mechanisms for the toxicological effects of oil and oil cleaning chemicals. Overall our findings (Liu et al., 2016) provided compelling evidence for the potential lung health impact of the BP oil spill on those workers involved in the cleaning operation.

One of the key limitations of our previous study (Liu et al., 2016) is that the analysis was largely performed at single gene level. Although we did perform analysis at the level of GO functional terms, the analysis was still based on the results from single gene differential expression analysis. For example, only those genes that achieved a p value of <0.05 in single gene differential expression analysis were submitted to GO analysis. Such a p value threshold, although commonly used, may be arbitrary in a genomic study as those genes that did not achieve the p value <0.05 may also contribute to the effects of biological significance. Hence annotation based only on those genes with a small p value may lose some sensitivity to capture the key signatures of the toxicological effects on the cells. Furthermore, genes often work correlatively and collaboratively in pathways and functional modules. Differential expression analysis at the single gene level as in our previous study (Liu et al., 2016) ignored such correlative relationship, which again may have missed some important transcriptomics signatures and failed to capture some key functional variations of the transcriptome.

To alleviate the problems/limitations of our previous study and take full advantage of this valuable RNA-seq dataset (Liu et al., 2016), here we performed a KEGG pathway-based analysis using the two well-developed software packages, GAGE (generally applicable gene set enrichment) (Luo et al., 2009) and GSNCA (Gene Sets Net Correlations Analysis) (Rahmatallah et al., 2014). While the findings from our new analysis agree with some of the previous study (Liu et al., 2016), the findings provided distinct new clues. A number of pathways, including ribosomal biosynthesis, protein processing, Wnt signaling, neurotrophin signaling and insulin signaling pathways, were all upregulated at the whole pathway level (mainly by the dispersant 9527). Importantly, upregulation of these pathways was all closely related with cancers, including lung cancer, in previous molecular and epidemiological studies (Mazieres et al., 2005; Poloz and Stambolic, 2015; Prakash et al., 2010; Zhou et al., 2015). Moreover, by GSNCA analysis (Rahmatallah et al., 2014), we further identified a number of cancer-related pathways, including the small cell lung cancer pathway, whose gene co-expression was changed due to the exposure to the oil/dispersant treatment, especially the mixture of dispersant 9527 and oil. Overall, our study here by analyzing the transcriptomics signals at the KEGG pathway level has revealed potential carcinogenic effects of dispersant 9527 and its mixture with crude oil. The findings provided further evidence for the health hazards of oil and oil dispersants to the respiratory system.

2. Methods

2.1. Experimental methods

The methods for generating the RNA-seq data were detailed in our previous study (Liu et al., 2016). Briefly, human airway epithelial cells (BEAS-2B cells, ATCC® CRL-9609™) were grown under six treatments using WAF (water accommodated fraction) of the following, i.e., the BP crude oil (abbreviated as "oil"), dispersant Corexit 9500 (abbreviated as "9500"), dispersant Corexit 9527 (abbreviated as "9527"), the mixture of oil + 9500 (abbreviated as "oil + 9500"), the mixture of oil + 9527 (abbreviated as "oil + 9527"), and water (abbreviated as "control"). There were three cell samples (biological replicates) for each treatment and total RNA extraction followed by RNA-seq experiments was performed on each sample. The RNA-seq experiments were performed at Omega Bioservices (Norcross, GA). The RNA-seq data was submitted to GEO under the accession number (GSE70909). Differential expression analysis was performed between each of the first five treatment groups vs. the control group.

The BP crude oil was kindly provided by The Architecture, Engineering, Consulting, Operations and Management Company (AECOM, Los Angeles, CA). This oil was obtained from the site of the Macondo well during the BP Oil Spill disaster. Commercially available Corexit EC9500A and EC9527 dispersants were kindly provided by a contract between Nalco/Exxon Energy Chemicals, L.P. (Sugar Land, TX, USA) and Tulane University (New Orleans, USA). The dispersants were provided as liquid solutions ready for use.

2.2. GAGE analysis

Following the Bioconductor workflow http://www.bioconductor.org/help/workflows/rnaseqGene/, the raw RNA-seq data (GSE70909) was analyzed to generate the raw count matrix, based on which, we used DESeq2 (Love et al., 2014) to perform differential expression analysis between a treatment (e.g., oil) and the control.

DESeq2 package requires "raw" counts of sequencing reads as the starting point for differential expression analysis (Love et al., 2014). Therefore, before submitted to the program for analysis, the count matrix was not normalized (which is explicitly required by the software) (Love et al., 2014). However, during the analysis procedures of DESeq2, normalization did occur in the modeling process, where the read count for gene i and sample j was modeled as a negative binomial distribution with mean μ_{ij} and dispersion α_j, and $\mu_{ij} = s_{ij}q_{ij}$, where q_{ij} is the raw read count and s_{ij} is a size factor that normalizes differences in sequencing depth between samples and other sources of technical biases, such as GC content and gene length (Love et al., 2014). The size factor s_{ij} was estimated with the "estimateSizeFactor()" function and the dispersion α_j estimated with the "DESeq()" function.

We ran the "DESeqDataSetFromMatrix" function, with the count matrix for all 18 samples (3 for each of the 5 treatments and 3 for the control) and the design matrix as input data, which produced an R object for the downstream differential expression analysis. Including count data for all the 18 samples (rather than only 3 samples from a treatment and the 3 control samples) is more robust for estimating parameters (such as the size factor and dispersion). For differential expression analysis to compare a certain treatment with the control, we used "contrast()" function, e.g., by defining "contrast = c ("treatment", "oil", "control")". The result file that contain p values and fold changes for each gene was generated with the "results()" function.

The log2 fold changes for all the genes from the differential expression analysis (which is a column in the result file) were submitted as

an input file for GAGE analysis (Luo et al., 2009). Note that we submitted the log2 fold changes for all the genes, not just those genes that achieved a certain significant p values (such as $p < 0.05$). GAGE (Luo et al., 2009) then uses the information of fold change for each gene to obtain mean and standard deviation of fold changes for a gene set (pathway) as well as for the background (the whole transcriptome) and generate a t-test statistic and p value for a comparison in fold change between a gene set and the background. Essentially, if there is significant difference in fold change between a gene set and the background, an extreme t statistic and a small p value will be achieved. As the fold change involves information on up- or downregulation (i.e., fold change >1 for up- and <1 for downregulation), a pathway identified by GAGE will also be indicated as up- or downregulated. As a nice feature of GAGE, a KEGG pathway identified as significantly differentially expressed can be visualized in a KEGG pathway plot using PathView (Luo and Brouwer, 2013), where an upregulated gene will be shown in red and a downregulated gene in green and an unregulated gene in grey (Fig. 1). Therefore, if, for example, in a pathway plot, the majority of genes are shown in red, the pathway's upregulation can be visually assessed (Fig. 1).

2.3. GSNCA analysis

GSNCA (Rahmatallah et al., 2014) offers multivariate nonparametric statistical methods testing difference in correlation structure of a gene set (pathway) between two conditions. As a key feature, GSNCA quantitatively characterizes the importance of each gene in a correlation network, assigning each gene a weight (w_i for gene i) that is proportional to the gene's cross-correlation with all the other genes. Thus, genes with high cross-correlations (e.g., the hub gene) will normally have a high weight that may indicate their regulatory importance. Based on cross-correlations among different genes, correlation structure of a pathway can be constructed and a statistical test can be performed to test difference of the structure in different conditions (e.g., treatment vs. control). The correlation structure for a pathway under a specific condition usually features a group of high weight genes (including the hub gene), which might suggest the central regulatory roles exerted by these genes.

For GSNCA analysis, we first downloaded gene annotation file from ENSEMBL and extracted gene length information. The raw count matrix was filtered to remove rows of zero counts and rows of no gene annotations. For each element of the count matrix (each gene from each sample), count per million reads (cpm) was calculated to standardize for library size and those rows (genes) with a cpm < 0.1 were further removed, followed by calculation of rpkm (reads per kilobase per million reads) to further standardize for gene length. KEGG pathway list was downloaded and then further partitioned into single KEGG pathway file.

Using the GSAR package (Rahmatallah et al., 2012) we performed GSNCA analysis for each of the KEGG pathways (as a single gene set). The analysis is a multivariate differential co-expression test that accounts for the correlation structure between genes. The test assigns weight factors w_s to genes under one condition and adjusts these weights simultaneously such that equality is achieved between each genes' weight and the sum of its weighted absolute correlations (r_{ij}) with other genes in a gene set of p genes.

$$w_i = \sum_{j \neq i} w_j |r_{ij}| \quad \text{where} \quad 1 \leq i \leq p.$$

The test statistic W_{GSNCA} is given by the 1st norm of difference between the scaled weight vectors $w^{(1)}$ and $w^{(2)}$ (each vector is multiplied by its norm) for the two conditions.

$$W_{GSNCA} = \sum_{i=1}^{p} \left| w_i^{(1)} - w_i^{(2)} \right|$$

This test statistic was used to test the hypothesis: $H_0 : W_{GSNCA} = 0$ *vs*. $H_0 : W_{GSNCA} \neq 0$. The p-values for the test statistic are obtained by comparing the observed value of the test statistic to its null distribution, which is estimated using a permutation approach (through permuting samples' class labels). Specifically, the p value is calculated as $p = \frac{b+1}{nperm+1}$, where b is the number of permutations with a more extreme statistic W_{GSNCA} than the observed statistic and *nperm* is the total number of permutations. For each specific test, we permute for 10,000 times, the most significant empirical p value that can be possibly achieved is $9.99E-5$ (meaning that among 10,000 permutations no permutation has achieved a more extreme test statistic than the observed one).

3. Results

From GAGE analysis, we identified 4 KEGG pathways that were downregulated by the oil treatment, 7 pathways significantly regulated (including 2 up- and 5 downregulated) by 9500 treatment, 27 pathways

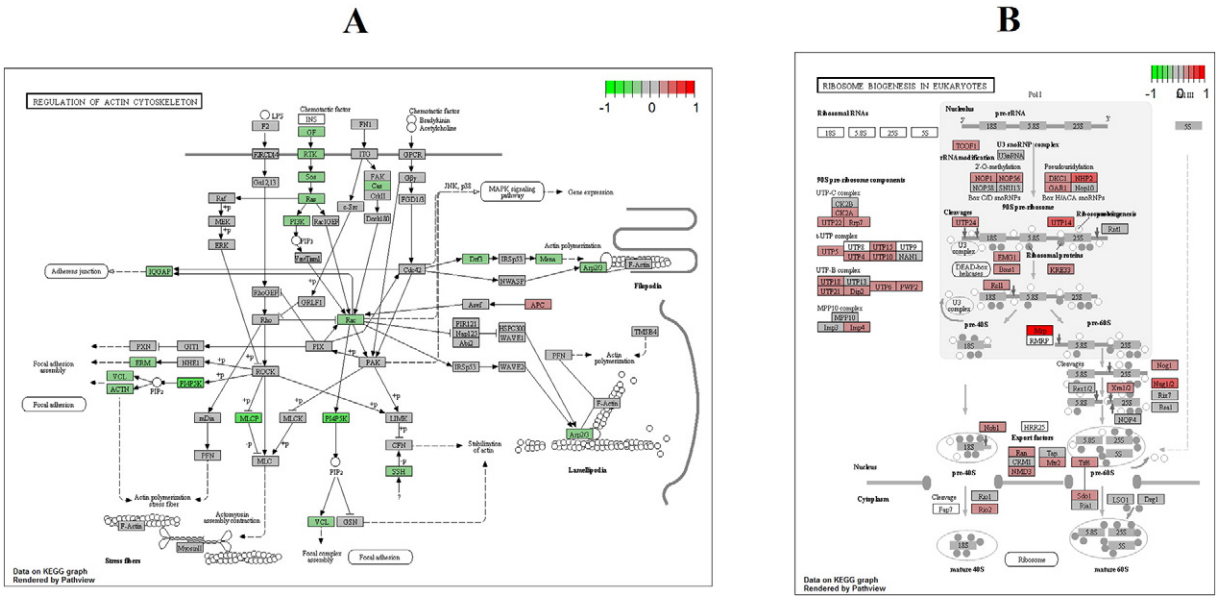

Fig. 1. Downregulation of actin cytoskeleton by oil treatment and upregulation of ribosome biogenesis by 9527 treatment. Note: A: Downregulation of actin cytoskeleton by oil treatment. B: Upregulation of ribosome biogenesis by 9527 treatment.

upregulated by the 9527 treatment, and 8 pathways upregulated by the 9500 + oil treatment. The results are presented in Table 1. For the statistical criterion of significant regulation, we used an FDR q value <0.10 as a cutoff threshold.

With oil treatment, among the most significantly regulated pathways are downregulation of actin cytoskeleton (hsa04810) ($p = 6.53E-5$) and adherens junction (hsa045020) ($p = 9.69E-5$) pathways. Treatment with 9500 caused upregulation of ribosome (hsa03010) ($p = 2.22E-11$) and antigen processing and presentation (hsa04612) ($p = 1.22E-3$) and downregulation of adherens junction (hsa045020) ($p = 1.06E-4$) pathways, among the most significantly regulated pathways. (Table 1)

The 9527 treatment leads to the largest number of pathways that are significantly regulated (with a total of 27 pathways, all upregulated). At the top of the list are the ribosome biogenesis (hsa03008) ($p = 1.97E-13$) and protein processing (hsa04141) pathways. In addition, Wnt signaling (hsa04310) ($p = 6.76E-3$), neurotrophin signaling (hsa04722) ($p = 7.73E-3$) and insulin signaling (hsa04910) ($p = 1.16E-02$) are also among the upregulated pathways.

Overall, in terms of statistical significance, the most significantly upregulated pathways are ribosome biogenesis (hsa03008) and ribosome (hsa03010) under 9527, 9500 or 9500 + oil treatments. The most significantly downregulated pathways are regulation of actin cytoskeleton (hsa04810) and adherens junction (hsa04520) induced by oil or 9500 treatments. As a key feature, a large number of pathways were upregulated under 9527 or 9500 + oil treatments, with the majority of upregulated pathways related to RNA metabolism (Table 1).

To illustrate the significance of the two most regulated pathways (hsa03008 and hsa04810) mentioned above, we used PathView to visualize the up- (in red color) and downregulation (in green color) of each gene in the context of the whole KEGG pathway plots (Fig. 1). Unregulated genes are shown in grey. If the majority of the colored genes are red/green, then that pathway is up/downregulated as a whole. As examples, note the prevailing red color for the ribosome biogenesis pathway upregulated by the 9527 treatment (Fig. 1B) and the prevailing green color for the actin cytoskeleton pathway downregulated by the oil treatment (Fig. 1A).

Using GAGE (Luo et al., 2009), we also performed differential expression analysis at the GO level to corroborate the KEGG pathway results. The detailed GO terms significantly regulated are not presented here due to the large number of terms involved and the space limitation. Instead we present the number of regulated terms and some example GO terms in Table 2.

The GO results are generally consistent with the KEGG results. The impact of oil treatment is featured by downregulation of anchoring junction (GO:0070161) and adherens junction (GO:0005912). The 9500 treatment effect is characterized by upregulation of inflammatory response (GO:0006954) and innate immune response (GO:0045087), in addition to simultaneous downregulation of anchoring junction (GO:0,070,161) and adherens junction (GO:0,005,912). 9527 treatment

Table 1
KEGG pathways significantly regulated by oil and/or dispersants.

KEGG pathway name	*p* value	FDR	Treatment group	Up or down-regulated by treatment
hsa04810 Regulation of actin cytoskeleton	6.53E−05	5.97E−03	Oil	Downregulated
hsa04520 Adherens junction	9.69E−05	5.97E−03		Downregulated
hsa04144 Endocytosis	1.11E−04	5.97E−03		Downregulated
hsa03015 mRNA surveillance pathway	1.61E−03	6.51E−02		Downregulated
hsa03010 Ribosome	2.22E−11	3.60E−09	9500	Upregulated
hsa04612 Antigen processing and presentation	1.22E−03	9.91E−02		Upregulated
hsa04520 Adherens junction	1.06E−04	1.72E−02		Downregulated
hsa00900 Terpenoid backbone biosynthesis	1.84E−03	9.55E−02		Downregulated
hsa04144 Endocytosis	1.95E−03	9.55E−02		Downregulated
hsa03015 mRNA surveillance pathway	2.37E−03	9.55E−02		Downregulated
hsa04810 Regulation of actin cytoskeleton	2.95E−03	9.55E−02		Downregulated
hsa03008 Ribosome biogenesis in eukaryotes	1.97E−13	3.20E−11	9527	Upregulated
hsa03013 RNA transport	2.03E−09	1.64E−07		Upregulated
hsa00970 Aminoacyl-tRNA biosynthesis	3.42E−07	1.66E−05		Upregulated
hsa04141 Protein processing in endoplasmic reticulum	4.09E−07	1.66E−05		Upregulated
hsa03040 Spliceosome	6.09E−07	1.97E−05		Upregulated
hsa03050 Proteasome	4.66E−05	1.26E−03		Upregulated
hsa00240 Pyrimidine metabolism	6.59E−05	1.53E−03		Upregulated
hsa04120 Ubiquitin mediated proteolysis	1.01E−04	2.05E−03		Upregulated
hsa03018 RNA degradation	1.64E−04	2.95E−03		Upregulated
hsa00020 Citrate cycle (TCA cycle)	2.22E−04	3.59E−03		Upregulated
hsa03015 mRNA surveillance pathway	4.09E−04	5.53E−03		Upregulated
hsa04110 Cell cycle	4.10E−04	5.53E−03		Upregulated
hsa04114 Oocyte meiosis	6.65E−04	8.29E−03		Upregulated
hsa00230 Purine metabolism	7.65E−04	8.85E−03		Upregulated
hsa03010 Ribosome	1.07E−03	1.15E−02		Upregulated
hsa03022 Basal transcription factors	1.34E−03	1.35E−02		Upregulated
hsa03020 RNA polymerase	1.75E−03	1.67E−02		Upregulated
hsa00520 Amino sugar and nucleotide sugar metabolism	6.48E−03	5.77E−02		Upregulated
hsa04310 Wnt signaling pathway	6.76E−03	5.77E−02		Upregulated
hsa04722 Neurotrophin signaling pathway	7.73E−03	6.27E−02		Upregulated
hsa04914 Progesterone-mediated oocyte maturation	8.47E−03	6.54E−02		Upregulated
hsa00010 Glycolysis/gluconeogenesis	1.05E−02	7.71E−02		Upregulated
hsa04910 Insulin signaling pathway	1.16E−02	8.16E−02		Upregulated
hsa00290 Valine, leucine and isoleucine biosynthesis	1.37E−02	9.22E−02		Upregulated
hsa03008 Ribosome biogenesis in eukaryotes	6.17E−11	1.00E−08	9500 + oil	Upregulated
hsa00970 Aminoacyl-tRNA biosynthesis	1.11E−07	8.98E−06		Upregulated
hsa03013 RNA transport	3.02E−07	1.63E−05		Upregulated
hsa03010 Ribosome	1.53E−05	6.18E−04		Upregulated
hsa03040 Spliceosome	1.81E−04	5.86E−03		Upregulated
hsa03018 RNA degradation	3.12E−04	8.41E−03		Upregulated
hsa03050 Proteasome	9.20E−04	2.13E−02		Upregulated
hsa04141 Protein processing in endoplasmic reticulum	4.00E−03	8.11E−02		Upregulated

Table 2
GAGE GO analysis results.

Treatment	GO analysis type	Up or downregulated terms	N of significant terms (FDR < 0.1)	Examples of regulated GO terms
Oil	CC	Down	24	GO:0070161 anchoring junction; GO:0005912 adherens junction
	BP	Down	20	GO:0051301 cell division
	MF	Down	5	GO:0000988 protein binding transcription factor activity
9500	CC	Up	8	GO:0022626 cytosolic ribosome
		Down	4	GO:0005912 adherens junction; GO:0070161 anchoring junction
	BP	Up	19	GO:0006954 inflammatory response; GO:0045087 innate immune response
		Down	8	GO:0051301 cell division; GO:0030036 actin cytoskeleton organization
	MF	Up	1	GO:0003735 structural constituent of ribosome
		Down	1	GO:0003779 actin binding
9527	CC	Up	39	GO:0005759 mitochondrial matrix; GO:0005840 ribosome
		Down	1	GO:0034702 ion channel complex
	BP	Up	159	GO:0034660 ncRNA metabolic process; GO:0006412 translation: GO:0022613 ribonucleoprotein complex biogenesis; GO:0042254 ribosome biogenesis
	MF	Up	31	GO:0008135 translation factor activity, nucleic acid binding; GO:0016874 ligase activity; GO:0003735 structural constituent of ribosome
		Down	7	GO:0005261 cation channel activity; GO:0005216 ion channel activity; GO:0022836 gated channel activity
Oil + 9500	CC	Up	14	GO:0005840 ribosome; GO:0005759 mitochondrial matrix
		Down	6	GO:0000323 lytic vacuole; GO:0005764 lysosome
	BP	Up	30	GO:0034660 ncRNA metabolic process; GO:0042254 ribosome biogenesis
		Down	4	GO:0008203 cholesterol metabolic process; GO:0016125 sterol metabolic process
	MF	Up	13	GO:0003735 structural constituent of ribosome; GO:0004812 aminoacyl-tRNA ligase activity

induced the largest significantly regulated GO terms, with the vast majority of the terms being upregulated, especially the upregulation of the BP (biological process) terms ($n = 159$), which again features ribosome biogenesis (GO:0042254), translation (GO:0006412) and the related terms.

Using GSNCA analysis (Rahmatallah et al., 2014), we identified 64 KEGG pathways that have differential co-expression patterns (i.e., the between-gene correlation structure) in treatment and control groups (Table 3), which achieved a p value of 9.99E−5, the smallest p value that can be possibly achieved in 10,000 permutations. This p value is significant even after Bonferroni correction for testing a total of 187 KEGG pathways (as the Bonferroni corrected significance level is 0.05/186 = 2.67E−4). Many of these pathways overlap with those pathways identified by GAGE analysis, such as regulation of actin cytoskeleton (hsa04810), endocytosis (hsa04144), antigen processing and presentation (hsa04612) and insulin signaling pathway (hsa04910). The most interesting pathways identified that are distinct from the GAGE analysis are a number of KEGG pathways related to cancer. These pathways (bolded and italicized in Table 3) include small cell lung cancer (hsa05222), pathways in cancer (hsa05200), prostate cancer (hsa05215), endometrial cancer (hsa05213), and basal cell carcinoma (hsa05217). Among the treatments, oil + 9527 appears to be the one most frequently associated with these cancer-related pathways as oil + 9527 is either the only one or one of the conditions associated with these cancer pathways.

As an example (shown in Fig. 2), the major change in the small lung cancer pathway due to the oil + 9527 treatment is that a number of genes in this pathway that have high or medium weights in controls become genes with a very low weight, suggesting the loss of their correlative relationship with other genes and their regulatory roles. Under the oil + 9527 treatment, the pathway has hence become functionally dominated by a smaller subset of higher weight genes as compared with controls. Using DAVID (Dennis et al., 2003) to annotate the subset of the higher weight genes produces a top annotation cluster (with an enrichment score of 7.27) that contains 8 KEGG cancer pathway terms, e.g., hsa05215:Prostate cancer ($p = 5.60E-14$), hsa05220:Chronic myeloid leukemia ($p = 7.65E-12$), hsa05212:Pancreatic cancer ($p = 1.61E-10$), hsa05223:Non-small cell lung cancer ($p = 2.85E-07$), etc. (note: p values are Bonferroni corrected p values). In contrast, if we annotate those lower weight genes with DAVID (Dennis et al., 2003), the top annotation cluster (with an enrichment score of 4.95) does not contain cancer terms. These results suggest that under the oil + 9527 treatment the pathway is now functionally shifted to a smaller set of genes that are further more enriched with cancer pathways.

Overall, at the whole KEGG pathway level, the most pronounced regulated pathways are ribosomal biosynthesis and its related pathways. Gene co-expression changes were also observed for a number of cancer pathways, most relevantly, the small cell lung cancer pathway. While 9527 treatment has the largest number of pathways significantly regulated, including not only ribosomal biosynthesis but also several signaling pathways, such as Wnt signaling, 9527 + oil appears to be the most common treatment to change gene co-expression patterns of cancer related pathways. Therefore, the overall effects appear to be most pronounced and enriched for 9527 and 9527 + oil treatments.

4. Discussion

In this study, we performed pathway-based analysis on RNA-seq data of the human airway epithelial cells treated with oil and dispersants 9500 and 9527. Pathway-based analysis of gene expression data is more informative and biologically meaningful than individual gene based analysis since very often genes work correlatively and orchestrally as a "team" (a functional module or a functional pathway). For some genes, their moderate changes of expression at the individual gene level may be insignificant under single gene statistical analysis framework yet they may be part of a biologically significant process by contributing to the process in an orchestrated manner. In such case, it is more reasonable that the analytical unit for gene expression analysis be a group of (functionally related) genes rather than individual transcripts. Statistically speaking, it can also alleviate effectively the common problem of multiple testing and "winner's curse" (Ioannidis, 2008) in a genomic study. Therefore, here we used GAGE (Luo et al., 2009) to identify differential expression at the whole pathway level due to oil and/or dispersant treatments.

Given the intrinsic correlation among the genes in a pathway, sometime a pathway is not up or down-regulated at the whole pathway level but may change in a more delicate manner, which is the alteration of the inner correlative structure or the co-expression patterns. For example, under certain conditions, some genes (e.g., the so called hub genes) that have a large number of correlative relationships with other members of a pathway may lose the "network hub" status and be replaced by other genes as new hubs. If that happens, the pathway's function

Table 3
GSNCA analysis results.

KEGG pathways with a significant p value ($p = 9.99E-5$)	Treatments associated with the co-expression changes of a pathway
ABC transporters	9500, oil + 9527
Antigen processing and presentation	9527
Arginine and proline metabolism	9527
Axon guidance	Oil, 9500, oil + 9500,oil + 9527
B cell receptor signaling pathway	oil + 9527
Basal cell carcinoma	***oil + 9527***
Basal transcription factors	oil + 9527
Butanoate metabolism	oil + 9527
Cell adhesion molecules CAMS	9500, 9527
Cytokine cytokine receptor interaction	9500
Cytosolic DNA sensing pathway	Oil
Drug metabolism other enzymes	9500
ECM receptor interaction	Oil, 9527, oil + 9500, oil + 9527
Endocytosis	9527, oil + 9527
Endometrial cancer	***oil + 9527***
Epithelial cell signaling in helicobacter pylori infection	oil + 9527
Fc epsilon RI signaling pathway	Oil, 9527
Fc gamma R mediated phagocytosis	9500
Focal adhesion	Oil, 9500, oil + 9500, oil + 9527
Fructose and mannose metabolism	oil + 9500, oil + 9527
Gap junction	9500, oil + 9500
Glycerolipid metabolism	9500
Glycosaminoglycan biosynthesis chondroitin sulfate	9527
Graft versus host disease	oil
Hedgehog signaling pathway	9500, oil + 9500, oil + 9527
Hematopoietic cell lineage	Oil, oil + 9500, oil + 9527
Histidine metabolism	Oil
Homologous recombination	Oil
Insulin signaling pathway	9527
Intestinal immune network for IGA production	Oil, 9500, oil + 9527
Jak-Stat signaling pathway	Oil
Leukocyte transendothelial migration	9500
Lysine degradation	oil + 9527
MAPK signaling pathway	Oil, oil + 9500
Mismatch repair	oil + 9527
Natural killer cell mediated cytotoxicity	Oil, oil + 9500, oil + 9527
Neuroactive ligand receptor interaction	Oil, 9500, 9527, oil + 9500, oil + 9527
Notch signaling pathway	Oil
Oxidative phosphorylation	Oil
P53 signaling pathway	9527
Pathways in cancer	***Oil, 9500, oil + 9500, oil + 9527***
Phosphatidylinositol signaling system	9500, oil + 9500
PPAR signaling pathway	Oil, oil + 9500
Primary immunodeficiency	oil + 9527
Prostate cancer	***oil + 9527***
Proximal tubule bicarbonate reclamation	oil + 9500
Pyrimidine metabolism	Oil
Regulation of actin cytoskeleton	9500, oil + 9500, oil + 9527
Regulation of autophagy	9500
Selenoamino acid metabolism	Oil
Small cell lung cancer	***Oil, oil + 9527***
Snare interactions in vesicular transport	Oil
T cell receptor signaling pathway	9500, 9527, oil + 9500, oil + 9527
Terpenoid backbone biosynthesis	Oil, oil + 9527
TGF beta signaling pathway	Oil
Tight junction	Oil, 9500
Toll like receptor signaling pathway	Oil, 9500, 9527
Tryptophan metabolism	Oil
Type II diabetes mellitus	9500
Tyrosine metabolism	Oil
Ubiquitin mediated proteolysis	oil + 9527
Valine leucine and isoleucine degradation	9500
Vascular smooth muscle contraction	9500, 9527, oil + 9527
VEGF signaling pathway	oil + 9527

may also changes as the regulatory roles asserted by different hub genes may be different. Here we used GSNCA (Rahmatallah et al., 2014) to capture such delicate changes at the whole pathway level and identified a large number of pathways with changed co-expression patterns under the oil and/or dispersant treatments.

Some of the results from the present study are consistent with what our previous study reported (Liu et al., 2016) based on individual genes. For instance, downregulation of the cell junction (e.g., adherens junction and anchoring junction) and upregulation of inflammatory response and immune response due to oil and/or dispersants treatment were observed in both studies. These findings again indicated the potential detrimental effects on cellular structure integrity and activation of immune response by the oil/dispersant treatment, lending mechanistic support for the observed respiratory symptoms in workers involved in previous oil cleaning operations (Zock et al., 2012).

However, the key distinct features observed in this study are upregulated ribosomal biosynthesis, protein processing, Wnt signaling, neurotrophin signaling and insulin signaling pathways upon treatment with the dispersant 9527. Importantly the observation of such a coordinated upregulation of these pathways may indicate a tendency towards cancer development for the cells.

Increased ribosomal biosynthesis has been linked to oncogenic signals. For example, a study showed that c-Myc oncogenic activity required robust ribosome biogenesis and protein synthesis and when the upregulated protein synthesis was reduced to normal level by deleting one allele of RPL24 (Ribosomal Protein L24), the oncogenic potential of c-Myc was markedly abolished (Barna et al., 2008). Several tumor suppressors, such as RB and p130, inhibit tumor cell growth and proliferation by interfering with ribosome biogenesis and thus restricting global protein synthesis (Cavanaugh et al., 1995; Ciarmatori et al., 2001; Hannan et al., 2000). Ribosomal biosynthesis pathway has been proposed as a potential target for cancer treatment (Zhou et al., 2015).

Activation of Wnt signaling pathway was associated with lung cancer development (Mazieres et al., 2005; Stewart, 2014) and inhibition of Wnt signaling was shown to inhibit lung cancer cell growth both in vitro and in vivo (Kim et al., 2007). Neurotrophin signaling has also been shown to promote lung cancer development (Prakash et al., 2010) as high levels of TrkA and TrkB have been found in human lung adenocarcinoma and squamous cell carcinomas (Ricci et al., 2001), and TrkB has been found in small cell lung cancers and atypical carcinoids (Ricci et al., 2005).

Insulin signaling has been well known for its crucial role in cancer development (as reviewed in (Poloz and Stambolic, 2015)), including lung cancer (Dziadziuszko et al., 2008; Ferguson et al., 2012; Frisch et al., 2015). Hyperinsulinemia is linked to the development of lung cancer (Ferguson et al., 2012) as well as other types of cancers (Calle et al., 2003; Goodwin et al., 2002; Renehan et al., 2008; Stewart et al., 2009). Insulin signaling through insulin receptor A has direct oncogenic and anti-apoptotic effects on cancer cells (Cox et al., 2009; Frasca et al., 1999; Kalli et al., 2002; Law et al., 2008; Novosyadlyy et al., 2010; Papa et al., 1990). Insulin signaling has also been a target for cancer treatment as hyperinsulinemia drugs, such as metformin, have proven to be valuable for cancer therapy (Bosco et al., 2011; Evans et al., 2005; Goodwin et al., 2011; Vissers et al., 2015).

Through GSNCA analysis we also identified several cancer-related pathways, including small cell lung cancer pathway, which have gene co-expression changes due to oil/dispersant treatment. For example, under oil + 9527 treatment, the higher weight genes (those with more important regulatory roles) were reduced to a smaller subset of genes that are further enriched for cancer development (Fig. 2).

Not surprisingly, our findings indicate that exposure to Corexit 9527-oil mixture has the most nocive effect on cells with significant upregulation of cancer pathways. The toxic effect of Corexit 9527 has been previously established. In particular, Corexit 9527 contains the toxin 2-Butoxyethanol (NIH US National Library of Medicine Toxicology Data Network https://toxnet.nlm.nih.gov/cgi-bin/sis/search/a?dbs+hsdb:@

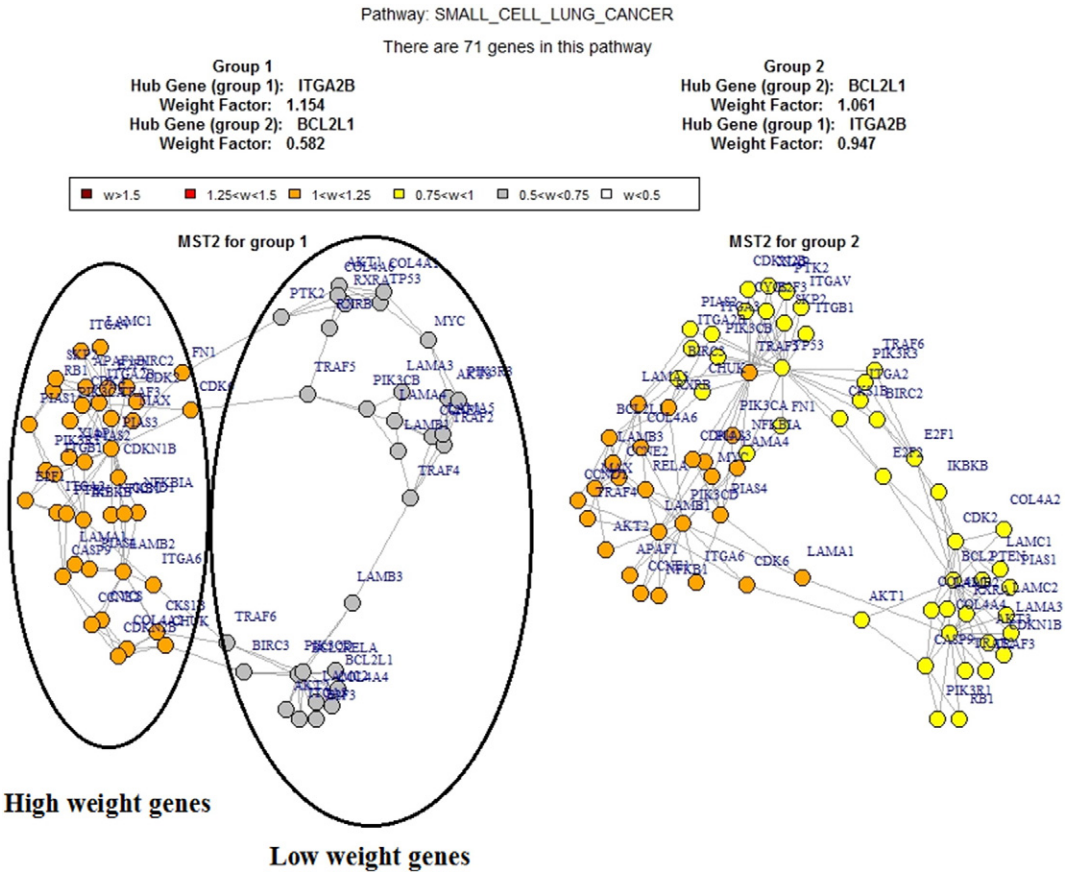

Fig. 2. Gene co-expression change of small cell lung cancer pathway under oil + 9527 treatment. Note: DAVID (Dennis et al., 2003) annotation for the high weight genes resulted in the top annotation cluster (with an enrichment score of 7.27) that contains 8 KEGG cancer pathways, e.g., hsa05215:Prostate cancer ($p = 5.60E-14$), hsa05220:Chronic myeloid leukemia ($p = 7.65E-12$), hsa05212:Pancreatic cancer ($p = 1.61E-10$), hsa05223:Non-small cell lung cancer ($p = 2.85E-07$), etc. (note: p values are Bonferroni corrected p values), suggesting that the high weight genes are highly enriched for these cancer pathways. In contrast, DAVID (Dennis et al., 2003) annotation for the low weight genes resulted in the top annotation cluster (with an enrichment score of 4.95) that does not contain cancer terms. The difference in annotation results may indicate that under treatment of oil + 9527 this pathway's regulatory genes (as manifested by those high weight genes) are more concentrated to a subset of genes highly enriched for cancer development.

term+@DOCNO+7838) that was shown to cause liver, kidney, lung, nervous system, and blood disorders among cleanup crews in Alaska following the 1989 Exxon Valdez oil spill (http://www.takepart.com/article/2013/04/17/corexit-deepwater-horizon-oil-spill). Although approved by the EPA, formulations of Corexit 9500 and 9527 were banned from use in the United Kingdom in 1998 because laboratory tests found them harmful to marine life that inhabits rocky shores (http://www.biologicaldiversity.org/programs/public_lands/energy/dirty_energy_development/oil_and_gas/gulf_oil_spill/dispersants.html). Regardless of the risks, these dispersants were still used in the BP oil spill site, and by June 1, 2010, 800,000 to 900,000 gal of Corexit 9527 had already been sprayed (https://www.whistleblower.org/gulftruth).

The above findings, both from differential expression analysis at the pathway level using GAGE (Luo et al., 2009) and from gene co-expression analysis using GSNCA (Rahmatallah et al., 2014), suggest a significant pattern of change towards cancer development in human airway epithelial cells under oil/dispersant treatment, especially the 9527 treatment. This, together with our previous study (Liu et al., 2016), provides compelling evidence for the toxicological effects, especially carcinogenic effects, of oil and dispersants and further reveals the potential health impacts of the recent BP oil spill on the lung health for those workers involved in the rescue mission. Accordingly, more strict regulations should be established to protect the personnel routinely working on the sites of oil spill and even those working in the refineries who may be exposed to oil products on a daily basis. Precautions should also be exercised on using oil dispersants, especially the Corexit 9527, to contain an oil spill in the future.

Acknowledgements

This study was supported by NIOSH (T42 OH008421) (National Institute for Occupational Safety and Health) Pilot Projects Research Training Program of Southwest Center for Occupational and Environmental Health. AMR-E is also supported by the by National Institutes of Health (NIH) P20GM103518/P20RR020152. GW is supported by 2G12MD007595. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIH.

References

Barna, M., Pusic, A., Zollo, O., Costa, M., Kondrashov, N., Rego, E., Rao, P.H., Ruggero, D., 2008. Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. Nature 456 (7224), 971–975.

Bosco, J.L., Antonsen, S., Sorensen, H.T., Pedersen, L., Lash, T.L., 2011. Metformin and incident breast cancer among diabetic women: a population-based case-control study in Denmark. Cancer Epidemiol. Biomark. Prev. 20 (1), 101–111.

Calle, E.E., Rodriguez, C., Walker-Thurmond, K., Thun, M.J., 2003. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N. Engl. J. Med. 348 (17), 1625–1638.

Cavanaugh, A.H., Hempel, W.M., Taylor, L.J., Rogalsky, V., Todorov, G., Rothblum, L.I., 1995. Activity of RNA polymerase I transcription factor UBF blocked by Rb gene product. Nature 374 (6518), 177–180.

Ciarmatori, S., Scott, P.H., Sutcliffe, J.E., McLees, A., Alzuherri, H.M., Dannenberg, J.H., te Riele, H., Grummt, I., Voit, R., White, R.J., 2001. Overlapping functions of the pRb family in the regulation of rRNA synthesis. Mol. Cell. Biol. 21 (17), 5806–5814.

Cox, M.E., Gleave, M.E., Zakikhani, M., Bell, R.H., Piura, E., Vickers, E., Cunningham, M., Larsson, O., Fazli, L., Pollak, M., 2009. Insulin receptor expression by human prostate cancers. Prostate 69 (1), 33–40.

Dennis Jr., G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003. DAVID: database for annotation, visualization, and integrated discovery. Genome Biol. 4 (5), 3.
Dziadziuszko, R., Camidge, D.R., Hirsch, F.R., 2008. The insulin-like growth factor pathway in lung cancer. J. Thorac. Oncol. 3 (8), 815–818.
Ehrenhauser, F.S., Avij, P., Shu, X., Dugas, V., Woodson, I., Liyana-Arachchi, T., Zhang, Z., Hung, F.R., Valsaraj, K.T., 2014. Bubble bursting as an aerosol generation mechanism during an oil spill in the deep-sea environment: laboratory experimental demonstration of the transport pathway. Evnviron. Sci. Process. Impacts 16 (1), 65–73.
Evans, J.M., Donnelly, L.A., Emslie-Smith, A.M., Alessi, D.R., Morris, A.D., 2005. Metformin and reduced risk of cancer in diabetic patients. BMJ 330 (7503), 1304–1305.
Faner, R., Cruz, T., Agusti, A., 2013. Immune response in chronic obstructive pulmonary disease. Expert. Rev. Clin. Immunol. 9 (9), 821–833.
Ferguson, R.D., Novosyadlyy, R., Fierz, Y., Alikhani, N., Sun, H., Yakar, S., Leroith, D., 2012. Hyperinsulinemia enhances c-Myc-mediated mammary tumor development and advances metastatic progression to the lung in a mouse model of type 2 diabetes. Breast Cancer Res. 14 (1), R8.
Frasca, F., Pandini, G., Scalia, P., Sciacca, L., Mineo, R., Costantino, A., Goldfine, I.D., Belfiore, A., Vigneri, R., 1999. Insulin receptor isoform A, a newly recognized, high-affinity insulin-like growth factor II receptor in fetal and cancer cells. Mol. Cell. Biol. 19 (5), 3278–3288.
Frisch, C.M., Zimmermann, K., Zillessen, P., Pfeifer, A., Racke, K., Mayer, P., 2015. Non-small cell lung cancer cell survival crucially depends on functional insulin receptors. Endocr. Relat. Cancer 22 (4), 609–621.
Georas, S.N., Rezaee, F., 2014. Epithelial barrier function: at the front line of asthma immunology and allergic airway inflammation. J. Allergy Clin. Immunol. 134 (3), 509–520.
Goodwin, P.J., Ennis, M., Pritchard, K.I., Trudeau, M.E., Koo, J., Madarnas, Y., Hartwick, W., Hoffman, B., Hood, N., 2002. Fasting insulin and outcome in early-stage breast cancer: results of a prospective cohort study. J. Clin. Oncol. 20 (1), 42–51.
Goodwin, P.J., Stambolic, V., Lemieux, J., Chen, B.E., Parulekar, W.R., Gelmon, K.A., Hershman, D.L., Hobday, T.J., Ligibel, J.A., Mayer, I.A., Pritchard, K.I., Whelan, T.J., Rastogi, P., Shepherd, L.E., 2011. Evaluation of metformin in early breast cancer: a modification of the traditional paradigm for clinical testing of anti-cancer agents. Breast Cancer Res. Treat. 126 (1), 215–220.
Hannan, K.M., Hannan, R.D., Smith, S.D., Jefferson, L.S., Lun, M., Rothblum, L.I., 2000. Rb and p130 regulate RNA polymerase I transcription: Rb disrupts the interaction between UBF and SL-1. Oncogene 19 (43), 4988–4999.
Hayworth, J.S., Clement, T.P., 2012. Provenance of Corexit-related chemical constituents found in nearshore and inland Gulf Coast waters. Mar. Pollut. Bull. 64 (10), 2005–2014.
Heijink, I., van Oosterhout, A., Kliphuis, N., Jonker, M., Hoffmann, R., Telenga, E., Klooster, K., Slebos, D.J., ten Hacken, N., Postma, D., van den Berge, M., 2014. Oxidant-induced corticosteroid unresponsiveness in human bronchial epithelial cells. Thorax 69 (1), 5–13.
Holtzman, M.J., Byers, D.E., Alexander-Brett, J., Wang, X., 2014. The role of airway epithelial cells and innate immune cells in chronic respiratory disease. Nat. Rev. Immunol. 14 (10), 686–698.
Ioannidis, J.P., 2008. Why most discovered true associations are inflated. Epidemiology 19 (5), 640–648.
Kalli, K.R., Falowo, O.I., Bale, L.K., Zschunke, M.A., Roche, P.C., Conover, C.A., 2002. Functional insulin receptors on human epithelial ovarian carcinoma cells: implications for IGF-II mitogenic signaling. Endocrinology 143 (9), 3259–3267.
Kim, J., You, L., Xu, Z., Kuchenbecker, K., Raz, D., He, B., Jablons, D., 2007. Wnt inhibitory factor inhibits lung cancer cell growth. J. Thorac. Cardiovasc. Surg. 133 (3), 733–737.
Kujawinski, E.B., Kido Soule, M.C., Valentine, D.L., Boysen, A.K., Longnecker, K., Redmond, M.C., 2011. Fate of dispersants associated with the Deepwater horizon oil spill. Environ. Sci. Technol. 45 (4), 1298–1306.
Laffon, B., Fraga-Iriso, R., Perez-Cadahia, B., Mendez, J., 2006. Genotoxicity associated to exposure to Prestige oil during autopsies and cleaning of oil-contaminated birds. Food Chem. Toxicol. 44 (10), 1714–1723.
Law, J.H., Habibi, G., Hu, K., Masoudi, H., Wang, M.Y., Stratford, A.L., Park, E., Gee, J.M., Finlay, P., Jones, H.E., Nicholson, R.I., Carboni, J., Gottardis, M., Pollak, M., Dunn, S.E., 2008. Phosphorylated insulin-like growth factor-i/insulin receptor is present in all breast cancer subtypes and is related to poor survival. Cancer Res. 68 (24), 10238–10246.
Liu, Y.Z., Roy-Engel, A.M., Baddoo, M.C., Flemington, E.K., Wang, G., Wang, H., 2016. The impact of oil spill to lung health-insights from an RNA-seq study of human airway epithelial cells. Gene 578 (1), 38–51.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (12), 550.
Luo, W., Brouwer, C., 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29 (14), 1830–1831.
Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., Woolf, P.J., 2009. GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinf. 10, 161.
Mazieres, J., He, B., You, L., Xu, Z., Jablons, D.M., 2005. Wnt signaling in lung cancer. Cancer Lett. 222 (1), 1–10.
Middlebrook, A.M., Murphy, D.M., Ahmadov, R., Atlas, E.L., Bahreini, R., Blake, D.R., Brioude, J., de Gouw, J.A., Fehsenfeld, F.C., Frost, G.J., Holloway, J.S., Lack, D.A., Langridge, J.M., Lueb, R.A., McKeen, S.A., Meagher, J.F., Meinardi, S., Neuman, J.A., Nowak, J.B., Parrish, D.D., Peischl, J., Perring, A.E., Pollack, I.B., Roberts, J.M., Ryerson, T.B., Schwarz, J.P., Spackman, J.R., Warneke, C., Ravishankara, A.R., 2012. Air quality implications of the Deepwater Horizon oil spill. Proc. Natl. Acad. Sci. U. S. A. 109 (50), 20280–20285.
Novosyadlyy, R., Lann, D.E., Vijayakumar, A., Rowzee, A., Lazzarino, D.A., Fierz, Y., Carboni, J.M., Gottardis, M.M., Pennisi, P.A., Molinolo, A.A., Kurshan, N., Mejia, W., Santopietro, S., Yakar, S., Wood, T.L., Leroith, D., 2010. Insulin-mediated acceleration of breast cancer development and progression in a nonobese model of type 2 diabetes. Cancer Res. 70 (2), 741–751.
Papa, V., Pezzino, V., Costantino, A., Belfiore, A., Giuffrida, D., Frittitta, L., Vannelli, G.B., Brand, R., Goldfine, I.D., Vigneri, R., 1990. Elevated insulin receptor content in human breast cancer. J. Clin. Invest. 86 (5), 1503–1510.
Perez-Cadahia, B., Lafuente, A., Cabaleiro, T., Pasaro, E., Mendez, J., Laffon, B., 2007. Initial study on the effects of Prestige oil on human health. Environ. Int. 33 (2), 176–185.
Perez-Cadahia, B., Laffon, B., Valdiglesias, V., Pasaro, E., Mendez, J., 2008a. Cytogenetic effects induced by Prestige oil on human populations: the role of polymorphisms in genes involved in metabolism and DNA repair. Mutat. Res. 653 (1–2), 117–123.
Perez-Cadahia, B., Mendez, J., Pasaro, E., Lafuente, A., Cabaleiro, T., Laffon, B., 2008b. Biomonitoring of human exposure to prestige oil: effects on DNA and endocrine parameters. Environ. Health Insights 2, 83–92.
Poloz, Y., Stambolic, V., 2015. Obesity and cancer, a case for insulin signaling. Cell Death Dis. 6, e2037.
Prakash, Y., Thompson, M.A., Meuchel, L., Pabelick, C.M., Mantilla, C.B., Zaidi, S., Martin, R.J., 2010. Neurotrophins in lung health and disease. Expert Rev. Respir. Med. 4 (3), 395–411.
Rahmatallah, Y., Emmert-Streib, F., Glazko, G., 2012. Gene set analysis for self-contained tests: complex null and specific alternative hypotheses. Bioinformatics 28 (23), 3073–3080.
Rahmatallah, Y., Emmert-Streib, F., Glazko, G., 2014. Gene Sets Net Correlations Analysis (GSNCA): a multivariate differential coexpression test for gene sets. Bioinformatics 30 (3), 360–368.
Renehan, A.G., Tyson, M., Egger, M., Heller, R.F., Zwahlen, M., 2008. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet 371 (9612), 569–578.
Rezaee, F., Georas, S.N., 2014. Breaking barriers. New insights into airway epithelial barrier function in health and disease. Am. J. Respir. Cell Mol. Biol. 50 (5), 857–869.
Ribatti, D., Puxeddu, I., Crivellato, E., Nico, B., Vacca, A., Levi-Schaffer, F., 2009. Angiogenesis in asthma. Clin. Exp. Allergy 39 (12), 1815–1821.
Ricci, A., Greco, S., Mariotta, S., Felici, L., Bronzetti, E., Cavazzana, A., Cardillo, G., Amenta, F., Bisetti, A., Barbolini, G., 2001. Neurotrophins and neurotrophin receptors in human lung cancer. Am. J. Respir. Cell Mol. Biol. 25 (4), 439–446.
Ricci, A., Graziano, P., Mariotta, S., Cardillo, G., Sposato, B., Terzano, C., Bronzetti, E., 2005. Neurotrophin system expression in human pulmonary carcinoid tumors. Growth Factors 23 (4), 303–312.
Rodrigues, R.V., Miranda-Filho, K.C., Gusmao, E.P., Moreira, C.B., Romano, L.A., Sampaio, L.A., 2010. Deleterious effects of water-soluble fraction of petroleum, diesel and gasoline on marine pejerrey *Odontesthes argentinensis* larvae. Sci. Total Environ. 408 (9), 2054–2059.
Rodriguez-Trigo, G., Zock, J.P., Pozo-Rodriguez, F., Gomez, F.P., Monyarch, G., Bouso, L., Coll, M.D., Verea, H., Anto, J.M., Fuster, C., Barbera, J.A., 2010. Health changes in fishermen 2 years after clean-up of the Prestige oil spill. Ann. Intern. Med. 153 (8), 489–498.
Saeed, T., Al-Mutairi, M., 1999. Chemical composition of the water-soluble fraction of the leaded gasolines in seawater. Environ. Int. 25, 117–129.
Stewart, D.J., 2014. Wnt signaling pathway in non-small cell lung cancer. J. Natl. Cancer Inst. 106 (1), djt356.
Stewart, S.T., Cutler, D.M., Rosen, A.B., 2009. Forecasting the effects of obesity and smoking on U.S. life expectancy. N. Engl. J. Med. 361 (23), 2252–2260.
Vissers, P.A., Cardwell, C.R., van de Poll-Franse, L.V., Young, I.S., Pouwer, F., Murray, L.J., 2015. The association between glucose-lowering drug use and mortality among breast cancer patients with type 2 diabetes. Breast Cancer Res. Treat. 150 (2), 427–437.
Zhou, X., Liao, W.J., Liao, J.M., Liao, P., Lu, H., 2015. Ribosomal proteins: functions beyond the ribosome. J. Mol. Cell Biol. 7 (2), 92–104.
Zock, J.P., Rodriguez-Trigo, G., Rodriguez-Rodriguez, E., Espinosa, A., Pozo-Rodriguez, F., Gomez, F., Fuster, C., Castano-Vinyals, G., Anto, J.M., Barbera, J.A., 2012. Persistent respiratory symptoms in clean-up workers 5 years after the Prestige oil spill. Occup. Environ. Med. 69 (7), 508–513.