HHS Public Access

Author manuscript

Sci Total Environ. Author manuscript; available in PMC 2021 June 30.

Published in final edited form as:

Sci Total Environ. 2021 June 25; 775: 145759. doi:10.1016/j.scitotenv.2021.145759.

Mixtures modeling identifies chemical inducers versus repressors of toxicity associated with wildfire smoke

Julia E. Rager[a,b,c,*], **Jeliyah Clark**[a,b], **Lauren A. Eaves**[a,b], **Vennela Avula**[a,b], **Nicole M. Niehoff**[d], **Yong Ho Kim**[e], **Ilona Jaspers**[b,c,e,f], **M. Ian Gilmour**[g]

[a]Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[b]The Institute for Environmental Health Solutions, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[c]Curriculum in Toxicology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

[d]Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[e]The Center for Environmental Medicine, Asthma and Lung Biology, School of Medicine, The University of North Carolina, Chapel Hill, NC, USA

[f]Department of Pediatrics, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

[g]Public Health and Integrated Toxicology Division, Center for Public Health and Environmental Assessment, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

Abstract

Exposure to wildfire smoke continues to be a growing threat to public health, yet the chemical components in wildfire smoke that primarily drive toxicity and associated disease are largely unknown. This study utilized a suite of computational approaches to identify groups of chemicals induced by variable biomass burn conditions that were associated with biological responses in the

*Corresponding author at: The University of North Carolina at Chapel Hill, 135 Dauer Drive, Chapel Hill, NC 27599, USA. jrager@unc.edu (J.E. Rager).

CRediT authorship contribution statement

Julia E. Rager: Conceptualization, Methodology, Software, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Jeliyah Clark:** Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Lauren A. Eaves:** Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Vennela Avula:** Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Nicole M. Niehoff:** Conceptualization, Formal analysis, Writing – original draft, Writing – review & editing. **Yong Ho Kim:** Conceptualization, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing. **Ilona Jaspers:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **M. Ian Gilmour:** Conceptualization, Investigation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

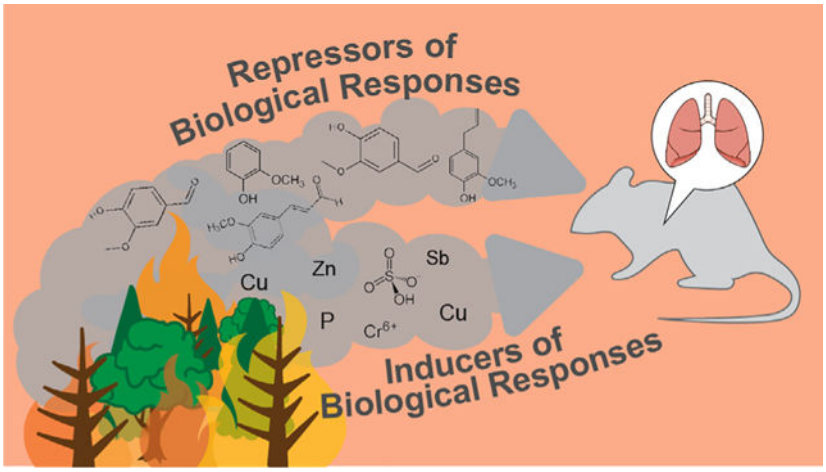
The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.145759.

mouse lung, including pulmonary immune response and injury markers. Smoke condensate samples were collected and characterized, resulting in chemical distribution information for 86 constituents across ten different exposures. Mixtures-relevant statistical methods included (i) a chemical clustering and data-reduction method, weighted chemical co-expression network analysis (WCCNA), (ii) a quantile g-computation approach to address the joint effect of multiple chemicals in different groupings, and (iii) a correlation analysis to compare mixtures modeling results against individual chemical relationships. Seven chemical groups were identified using WCCNA based on co-occurrence showing both positive and negative relationships with biological responses. A group containing methoxyphenols (e.g., coniferyl aldehyde, eugenol, guaiacol, and vanillin) displayed highly significant, negative relationships with several biological esponses, including cytokines and lung injury markers. This group was further shown through quantile g-computation methods to associate with reduced biological responses. Specifically, mixtures modeling based on all chemicals excluding those in the methoxyphenol group demonstrated more significant, positive relationships with several biological responses; whereas mixtures modeling based on just those in the methoxyphenol group demonstrated significant negative relationships with several biological responses, suggesting potential protective effects. Mixtures-based analyses also identified other groups consisting of inorganic elements and ionic constituents showing positive relationships with several biological responses, including markers of inflammation. Many of the effects identified through mixtures modeling in this analysis were not captured through individual chemical analyses. Together, this study demonstrates the utility of mixtures-based approaches to identify potential drivers and inhibitors of toxicity relevant to wildfire exposures.

Graphical Abstract



Keywords

Air pollution; Biomass burns; Complex mixtures; Computational toxicology; Pulmonary effects; Mixtures toxicology

1. Introduction

Wildfires are growing as a public health crisis, with recent events including the California wildfires and Australian bushfires in 2019–2020 elevating this issue to the forefront of concern worldwide (BBC, 2020; TIME, 2019). Wildfire occurrence, duration, and intensity have heightened globally in recent decades in response to an increasingly warm and dry climate (Westerling et al., 2006). This trend is predicted to continue with increasing numbers of wildfires and acreage burned; where in certain regions, wildfire emissions are projected to increase by up to 100% through 2100 (Hurteau et al., 2014). In addition to the destruction of property and land, the degradation of air quality as a result of intensified wildfires is of great public health concern. Exposure to smoke emitted from burning biomass has been associated with exacerbation of a wide variety of diseases (Liu et al., 2015). While it is established that wildfires are associated with increases in morbidity and mortality, the specific role that different components of smoke play towards causing disease remains unknown.

Wildfire smoke from the burning of biomass fuels results in a complex and variable mixture of particulate matter (PM) and toxic gases (Black et al., 2017). Consequently, there is immense variation in the biological responses and subsequent health outcomes that result from these exposures. Specific health effects related to wildfire inhalation exposure include respiratory illness, asthma exacerbations, more severe cardiovascular disease, and even death (Liu et al., 2015); however, the toxicity of specific smoke components remains largely unknown, although health outcomes are likely dependent on the type of biomass that is burned. Identifying which chemicals (or mixtures thereof) are the primary drivers of wildfire smoke-induced respiratory disease is a public health imperative. Once characterized, chemicals present within wildfire smoke that largely dictate toxicity responses can be used as common markers of exposure and measures of baseline risk across geographic regions impacted by wildfire smoke.

Elucidating which chemicals within wildfire smoke drive toxicity represents a challenging task and one that is heavily reliant upon the effective use of mixtures-based statistical approaches. Common to most mixtures-based analyses, some challenges must be addressed in order to adequately parse through such highly variable, high-dimensional datasets. Specific limitations that can occur in mixtures evaluations can include: imbalances between the number of samples vs. number of variables; difficulties differentiating between true predictor variables from correlated variables; difficulties distinguishing between individual vs. joint component effects; and potential compromises to interpretability in favor of identifying potential statistical signals within the noise of high-dimensional data (Agier et al., 2016; Bartel et al., 2013; Hamra and Buckley, 2018). Further, the majority of mixtures-based statistical methods have been generated in the context of epidemiological study designs (Carlin et al., 2013; Hamra and Buckley, 2018; Keil et al., 2020; Niehoff et al., 2020), with limited examples through toxicity designs (Ryan et al., 2019); leaving a current gap in established methodologies within the field of toxicology. Because of these potential limitations, we implemented a suite of statistical methods that each, in part, can address these issues to better understand potential relationships between chemicals, as well as the overall mixture effect of groups of chemicals in wildfire smoke and biological responses.

This study set out to implement mixtures-relevant statistical approaches to characterize relationships between chemicals present in wildfire smoke and associated pulmonary responses. Pulmonary inflammation and injury were the focus of the current analysis, as these represent the most sensitive biological responses resulting from wildfire smoke exposures and play important roles in pulmonary disease outcomes (Reid et al., 2016). We leveraged data previously produced by our research team (Kim et al., 2018), where wildfire smoke exposures were simulated experimentally through the burning of various biomass fuels. Resulting biomass smoke condensate samples were used as exposures and associated pulmonary responses were evaluated in mice. Mixtures modeling methods were used here to further elucidate statistical relationships between these complex exposures and biological responses, to identify groups of inducers and repressors of wildfire-associated toxicity.

2. Materials and methods

2.1. Overview of experimental study design to simulate wildfire smoke exposures

To characterize the chemical components and resulting biological responses associated with wildfire smoke exposures, we have developed laboratory-based approaches to simulate various biomass burn conditions, collect resulting biomass smoke condensate samples from these conditions, and test biological responses to the various smoke condensate samples (Kim et al., 2018). While recognizing that wildfire exposure conditions are complex and comprise of a variety of burn scenarios, we focused on five different biomass fuels, namely eucalyptus, peat, pine, pine needles, and red oak. These fuels were selected for analysis based on previously detailed criteria (Kim et al., 2018), including the representative nature of biomes across different regions of the U.S. as well as the world. For example, eucalyptus, peat, pine, and red oak are significant contributors to wildland fires within the U.S. western coast (and other continent coastal regions); the U.S. midwestern and southeastern regions; the U.S. western regions; and the U.S. eastern and central regions, respectively. These biomasses were evaluated under two combustion conditions, flaming and smoldering, resulting in a total of 10 different biomass smoke conditions.

Biomass smoke was generated from a tube furnace, consisting of a quartz tube and a ring type electric heater that provided a sustained stable flame or smolder condition for 60 min. Biomass smoke samples, consisting of PM and condensable gas-phase semi-volatiles, were collected from a multistage cryotrap system where the biomass smoke samples were trapped in three impingers maintained at different cooling temperatures (−10, −50, −70 °C). This system represents an improvement upon traditional filter-based collection methods, as it allows for the collection of smoke particles, which are typically difficult to extract from filter matrices, and semi-volatile compounds, which typically pass through filters. Samples were extracted from the impingers with acetone, resulting in predominantly less-volatile solid PM samples, and samples were concentrated and dried under nitrogen gas. These samples represented the biomass smoke condensate used in the proceeding analyses. This experiment has been described in detail and previously published (Kim et al., 2018).

2.2. Chemicals within biomass smoke condensate exposures

The current study was aimed at evaluating relationships between individual and/or co-occurring chemicals within the collected biomass smoke condensate samples, that were then used to test pulmonary responses in mice. These substances were previously characterized using various instruments to cover a wide chemical domain. Instruments included, for the organic carbon analyses, a carbon analyzer (107A, Sunset Laboratory, Inc.) and a thermal desorption unit (TD; TDSA2/TDS, Gerstel, Inc.) coupled to a gas chromatograph-mass spectrometer (TD-GC-MS; 6890/5973, Agilent Technologies). For the inorganic elemental analyses, instruments included high-resolution-magnetic sector field inductively coupled plasma mass spectrometry (HR-ICP-MS; ELEMENT™ 2; Thermo Scientific).

All chemistry-based methods have been previously described and results made publicly available (Kim et al., 2018). Chemicals that were characterized included ionic constituents ($N = 5$), inorganic elements ($N = 20$), n-alkanes ($N = 25$), polycyclic aromatic hydrocarbons (PAHs) ($N = 24$), methoxyphenols ($N = 11$), and levoglucosan that were measured across the 10 different biomass burn scenarios. The concentrations of these chemicals were specifically gathered from Kim et al. Supplemental Tables S3–S5 and converted into the final dissolved concentrations that were used during experimental treatments (Kim et al., 2018). In cases in which chemicals were measured at concentrations below the limits of detection (LOD), concentrations were imputed to reflect a LOD divided by the square root of two, paralleling previously published methods (Bailey et al., 2014; Hines et al., 2015; Rager et al., 2014).

2.3. Biological responses in the lung of mice exposed to biomass smoke condensate

Mice were exposed to the collected biomass smoke condensate samples and resulting biological responses within the pulmonary system were measured as previously described (Kim et al., 2018). These experiments were approved by the U.S. EPA Institutional Animal Care and Use Committee. In brief, adult pathogen-free female CD-1 mice were selected for use in this research to parallel experimental designs used in our previous air pollution studies, allowing for more direct comparisons between study findings (Cho et al., 2009; Kim et al., 2014; Kim et al., 2015b; Tong et al., 2010). Furthermore, use of these outbred female mice allows for effective group housing methods and does not promote one specific genotype over another (e.g., C3H/HeJ vs C57BL/6). A vehicle control group was included to acquire background response levels and a lipopolysaccharide (LPS)-treated group to acquire a positive response range to compare against.

Biomass smoke condensate samples underwent a gradual solvent exchange from acetone to saline. This process generated samples of dissolved biomass smoke particles in saline at a resulting concentration of 2 mg/mL while maximizing the maintenance of original chemical formulation. Resulting PM samples were then sonicated for 4 min using an ultrasonicator (Misonix Sonicator S-4000) and administered into the lungs of mice (100 μg of PM in 50 μL) via oropharyngeal aspiration. This dose was selected based on criteria previously described (Kim et al., 2018), including the representation of a peak 24 h exposure condition for a wildfire event relevant to inhaled doses in the human lung. Specifically, PM concentrations near wildfires have been reported to reach peak concentrations between 2 and 2.8 mg/m^3 (Naeher et al., 2007). Thus the PM deposited in human lungs across 24 h from

these peak concentrations would be relevant to the PM dose in mice used in the current investigation. Additional mice were exposed to 2 μg of lipopolysaccharide in 50 μL saline as a positive control, as well as 50 μL saline as a negative control. Biological samples were collected at 4 and 24 h postexposure to evaluate resulting toxicity. Six mice were used per exposure condition, across 10 different biomass burn condensate samples and 2 control samples, with two recovery periods (4 and 24 h postexposure) tested, resulting in a total of 24 groups of mice.

To evaluate pulmonary responses, bronchoalveolar lavage fluid (BALF) samples were collected from the treated mice either 4 or 24 h post-exposure, as previously described (Kim et al., 2018). Cells within the BALF samples were characterized and evaluated for macrophage and neutrophil counts. BALF supernatant samples were used to evaluate cytokine levels and markers of lung injury. Specific cytokines that were measured included interluekin-6 (IL-6), macrophage inhibitory protein-2 (MIP-2), and tumor necrosis factor-α (TNF-α). Lung injury markers were evaluated by measuring albumin, y-glutamyl transferase (GGT), lactate dehydrogenase (LDH), and total protein concentrations, as well as *N*-acetyl-β-D-glycosaminidase (NAG) activity in BALF samples. In this study, we focus on pulmonary biological response data specifically presented in Kim et al. Supplemental Tables S6–S8 (Kim et al., 2018). Data were also filtered to include biological endpoints that showed at least one instance of significant change in association with an exposure condition, for each timepoint, resulting in the following biological responses as the focus of the current evaluation (noted as “Response Type _ Specific Marker _ Timepoint”): Cytokine_IL6_4h, Cytokine_IL6_24h, Cytokine_TNFa_4h, Cytokine_MIP2_4h, Injury_Protein_24h, Injury_Albumin_4h, Injury_Albumin_24h, Injury_NAG_24h, Injury_LDH_24h, Neutrophil_4h, Neutrophil_24h.

2.4. Identifying chemical groups based on co-occurrence and their relationships to biological responses

Chemicals were first grouped in an unsupervised manner to identify those that co-occurred across biomass smoke condensate samples, and resulting chemical groups were then related to biological response profiles. When determining which chemical grouping statistic should be applied, we evaluated whether the underlying chemical distribution profiles were dependent upon fuel type. Specifically, each chemical distribution profile was tested for degree of correlation, and it was determined that the type of fuel used in burning scenarios did not drive sample correlations, and thus a paired statistical approach was not appropriate (Supplementary Material Table S1). This analysis was specifically carried out using an approach originally termed weighted gene co-expression network analysis (WGCNA) and employs data-reduction methods coupled with correlation statistics to describe patterns among high-dimensional datasets (Langfelder and Horvath, 2008). Historically, WGCNA has been used to evaluate co-modulated gene sets across expression signatures derived through microarray technologies (Langfelder and Horvath, 2008), and has since expanded to other applications, including the comparison of cross-platform performances for transcriptomics-based assessments (Wang et al., 2014) and the integration of genomic and epigenomic data to derive benchmark doses that elicit exposure-induced toxicity in humans (Rager et al., 2017). Additional efforts have applied this approach towards chemistry-based

measures, including metabolite profiling datasets (Zhang et al., 2013) and, more recently, chemical distribution profiles within ambient atmospheres (Eaves et al., 2020). Here, we similarly tailored these methods towards high-dimensional chemistry data, and thus refer to the approach as weighted chemical co-expression network analysis (WCCNA).

WCCNA was used in the current study to identify clusters (modules) of highly interconnected chemicals across biomass smoke condensate samples, all correlated with each other, and thus considered cooccurring. The steps involved in this approach have been summarized elsewhere (Langfelder and Horvath, 2008). In brief, Pearson correlation coefficients were first calculated for all pairwise comparisons of chemicals, and the resulting correlation matrix was transformed into an adjacency matrix resulting in a weighted network describing connection strengths between chemical distributions. A signed network was implemented, such that chemicals with positive correlations were grouped together. Correlation values within the adjacency matrix were raised to a power of 9 to maximize the scale-free topology model fit, as suggested (Langfelder and Horvath, 2017). With this, a co-expression network was constructed that was weighted, emphasizing high correlations and de-emphasizing low correlations between chemicals. Modules were then identified as groups of densely interconnected chemicals in the weighted network analysis with high topological overlap, measured using an average linkage hierarchical clustering with a dynamic tree-cutting algorithm. Altogether, resulting modules represented clusters of highly interconnected chemicals with positive correlations. These defined modules were used to calculate module eigenvalues, representing the first principal component of each module.

In the current analysis, all chemicals that were measured in the biomass smoke condensate samples were analyzed collectively using the WGCNA package in R (v1.68) (Langfelder and Horvath, 2008). This first set of analyses included all chemicals, including those that were below detection limits in many samples, in order to identify collective groupings such that if chemicals were present in a portion of the samples together, they were grouped together. The statistical analyses below further filtered the chemicals to prioritize those that were commonly above detection limits across samples. This WCCNA analysis resulted in the derivation of module eigenvalues, representing the collective distribution of chemicals that are co-occurring in biomass smoke. These module eigenvalues were correlated against the biological response data (ln-transformed) using Spearman Rank correlation tests using the Hmisc (v4.4–0) package in R and visualized using the corrplot (v0.84) package (Harrell and Dupont, 2014; Wei and Simko, 2017).

2.5. Integrating chemical mixture effects in relation to biological responses through quantile g-computation

A statistical approach termed quantile-based g-computation was implemented to take into account potential joint effects of chemicals within these complex mixtures. Considering the joint impact is important because, within wildfire-relevant conditions, exposures do not occur to isolated chemical constituents and individual chemicals are often correlated in concentration. Therefore, even if individual chemicals have a small effect, their overall collective impact may be meaningful. This recently developed approach is a generalized-linear-model based implementation of g-computation that provides estimates of the effect of

simultaneously increasing all exposures within a mixture by one quantile, which we refer to as the overall mixture effect (Keil et al., 2020). A strength of this method is that it allows chemicals to have associations with the outcome under evaluation in either direction.

Quantile-based g-computation was implemented here using the R package qgcomp (v2.0.0) (Keil, 2020). Data were preprocessed by scaling all of the chemical variables (mean = 0; SD = 1) and transforming the biological response data by the natural log. All biological replicates were included to capture variation in biological responses. When grouping the individual chemical exposures into quintiles, some chemicals had concentrations that were the same (i.e., tied) between groups, particularly for chemicals that had concentrations below the detection limit in multiple of the biomass smoke condensate samples. In these instances, chemicals were removed from the final statistical model in order to adequately derive exposure quintiles across the mixture. For the remaining chemicals, the *qgcomp.noboot* function was used to estimate exposure effects by categorizing all chemicals into quintiles, giving each chemical a positive or negative weight as defined above, and fitting a linear model for continuous outcomes, while incorporating Bayesian variable penalization. If chemicals have different directions of effect, the positive or negative weights are interpreted as the proportion of the total effect of the exposures that have a positive (or negative) effect on the outcome, and the positive and negative weights together sum to two. Each biological endpoint was evaluated separately in relation to the chemical mixtures, and resulting models were used to derive variable-specific coefficients, scaled effect sizes (i.e., the positive weight or negative weight of each variable contribution to the model), and overall model fit *p*-values. False discovery rate (FDR)-adjusted *p*-values were additionally calculated using the p.adjust function in R. Nine chemical groups were specifically evaluated, including: all chemicals (1 group); chemical modules identified through WCCNA (7 groups); all chemicals except for those contained within one module of high interest, identified as a likely repressor of toxicity (1 group).

An additional analysis was also carried out based on randomly permuted chemical distributions, to further evaluate whether observations were based on chance. Specifically, a random module was generated to include five chemical distribution measures, using the Sample function in R (RDocumentation, 2020), randomly selecting values from five measured chemicals. The same quantile-based g-computation statistics were carried out here, comparing the following three chemical groupings: all chemicals plus those from the randomly permuted module; all chemicals except those from the randomly permuted module; and just the randomly permuted module.

2.6. Individual chemical correlation analyses

Chemicals were evaluated individually in order to compare results identified through more global mixtures-based approaches vs. evaluations based on single chemicals assumed to act independently. Here, the concentrations of each individual chemical were correlated against each biological response profile (ln-transformed), averaged across biological replicates. Calculations were carried out using the Spearman Rank correlation test enabled though the Hmisc (v4.4–0) package in R, with false discovery rate (FDR)-adjusted p-values additionally

calculated using the p.adjust function, and results visualized using the corrplot (v0.84) package (Harrell and Dupont, 2014; Wei and Simko, 2017).

3. Results

3.1. Experimental study overview

This study set out to evaluate relationships between chemicals found in wildfire smoke, as captured through experimental biomass burns, and resulting pulmonary responses in mice. A total of 86 chemicals were detected and characterized across 10 different biomass smoke condensate samples, including n-alkanes, PAHs, methoxyphenols, levoglucosan, inorganic elements, and ionic constituents (Table 1, Supplementary Material Table S2). It is notable that although these individual chemical concentrations varied, the overall mixture dose was constant throughout the evaluations (100 μg of PM in 50 μL saline). Overall, the biomass burn conditions with the greatest concentrations of chemicals in each class included the following: the peat flaming condition produced the greatest overall concentration of inorganics and ionic constituents (followed by the pine needles flaming condition); the eucalyptus smoldering condition produced the greatest concentration of levoglucosan; the pine needles smoldering condition produced the greatest concentration of methoxyphenols and PAHs, and the peat smoldering condition produced the greatest concentration of n-alkanes. These results show that variable biomass burn conditions impart variable chemistries within the resulting condensate samples.

3.2. Chemical groups identified based on co-occurrence across biomass burns

We first set out to define groups of chemicals using an unsupervised approach aimed at identifying which chemicals co-occurred across biomass burn condensate samples. An approach based on WCCNA was employed, allowing for the identification of chemicals that increased and/or decreased in concentration together across exposures. This analysis resulted in the identification of seven different chemical groups (Table 1). These groups are designated as ‘modules’ assigned to different colors, with the following numbers of chemicals: black module (4 chemicals), blue module (17 chemicals), brown module (12 chemicals), green module (8 chemicals), red module (5 chemicals), turquoise module (28 chemicals), and yellow module (12 chemicals). Of note, the green module contained methoxyphenols, which showed the highest relative concentrations emitted from smoldering pine. The brown and yellow modules, which contained inorganic elements and ionic constituents, showed the highest relative concentrations emitted from flaming conditions, including flaming peat, flaming pine needles, and flaming red oak (Fig. 1).

In some instances, chemical modules showed groupings that were largely in-line with chemical structure classes; for example, 24 of the 25 n-alkanes were all grouped together in the turquoise module, and all inorganic elements and ionic constituents were grouped within either the brown or yellow modules. Other chemical structure classes showed assignments that were mixed; for example, the 24 PAHs were grouped to four different modules, namely black, blue, red, and turquoise. For a complete listing of chemicals grouped according to module, see Table 1. These data demonstrate that chemicals may increase or decrease in concentration in concert with other chemicals that may or may not share similar structural

attributes, thus supporting the use of a chemical grouping approach that is not based on a priori classifications.

3.3. Chemical groups showed both positive and negative associations to biological responses

Chemical groups were first related to biological responses in the mouse lung using a data-reduction method followed by correlation analyses. Specifically, the first principal component of each module was calculated (i.e., module eigenvalue) and correlated against each of the biological responses. The brown and yellow modules were identified as the chemical groups showing the most significant, positive association across the largest number of biological responses (Fig. 2, Supplementary Material Table S3). These modules consisted of inorganic elements and ionic constituents, which were positively correlated to cytokine concentrations (i.e., IL-6, MIP2, and TNFα) as well as lung injury markers. Conversely, the green module was identified as the chemical group showing the largest number of negative associations with several biological responses. The green module consisted of methoxyphenols, which were negatively correlated to cytokines (i.e., IL-6, MIP2, and TNFα), lung injury markers, and neutrophil influx, suggesting a potential protective effect or reduction in toxicity associated with these methoxyphenols, either individually or as a group (Fig. 2, Supplementary Material Table S3).

3.4. Mixtures modeling supports chemical groups of inducers vs. repressors of smoke-induced biological responses

A mixture effects modeling approach was implemented to further evaluate whether certain chemical groups showed associated induction vs. repression of biological responses resulting from biomass burn exposures. Specifically, quantile g-computation was used to consider the potential joint effects of chemicals in a group and capture mixture effects-based estimates of chemical contributions to biological responses. This method estimates the expected change in an outcome given that all chemical exposures within a mixture simultaneously increase by one quantile (in this case, one quintile) (Keil et al., 2020). Given that chemicals had to be grouped according to exposure concentration, a portion of the chemicals showed concentration distributions that caused ties to occur between quintiles, impeding adequate model development. These ties occurred for chemicals that had concentrations below the detection limit in multiple biomass smoke condensate samples, which largely included n-alkanes and PAHs. A total of 28 chemicals were able to be effectively grouped into exposure quantiles and were thus included in the final quantile g-computation models (Supplementary Material Table S2). Models were developed in relation to each of the biological responses, separately.

Mixture effect models were developed using the following nine combinations of chemicals: (i) all chemicals (N = 28 chemicals available for quantile groupings), (ii) individual chemical groups (i.e., 7 modules) (N = 1 to 10 chemicals), and (iii) all chemicals except for those contained within the green module (N = 26 chemicals). The last combination of chemicals was included with the goal of assessing the relationships after excluding chemicals that potentially reduce biological responses, as identified through WCCNA. Specifically, the green module was identified through WCCNA to contain chemicals with

collective distributions that were negatively correlated to several biological responses. This module contained two chemicals that could be grouped into exposure quantiles and thus included in the quantile g-computation models: namely, coniferyl alcohol and vanillin. Of the total 88 models tested, 29 models showed significant (FDR $q < 0.05$) relationships between chemical mixtures and a biological response (Supplementary Material Table S4).

Quantile g-computation model results demonstrated the potential for chemicals in the green module to reduce biological responses associated with biomass burn condensate exposures. To detail, in comparison to models using all chemicals, relationships between chemical exposures and biological responses became heightened and increased in significance when the chemicals in the green module were excluded (Fig. 3A). These relationships are quantified through the use of β coefficients, representing quantile g-computation estimates for the change in biological endpoint (ln-scaled) for a one quintile increase in chemical concentration (z-score normalized) in biomass burn condensate samples, alongside associated 95% confidence intervals (CIs). As an example, IL-6 levels collected 4 h post-exposure demonstrated a β coefficient of 0.19 (95% CI: −0.29, 0.66) in association with all chemicals, which then increased to a β coefficient of 0.75 (95% CI: 0.20, 1.30) when chemicals in the green module were excluded. Furthermore, the green module alone demonstrated a decreased β coefficient of −0.36 (95% CI: −0.60, −0.13) in relation to IL-6 levels. Similar trends were apparent for additional cytokine measures (e.g., IL-6 at 24 h, MIP2 at 4 h, TNFα at 4 h), lung injury markers (e.g., albumin at 4 h and 24 h, LDH at 24 h), as well as neutrophil levels at 4 h (Fig. 3). Furthermore, models that just included chemicals in the green module showed an overall average negative relationship averaged across all biological responses in comparison to models excluding the green module (Table 2), suggestive of a potential reduction in toxicity associated with coniferyl alcohol and vanillin.

It is notable that the green module's associated reduction in biological responses was unlikely to be an artifact of potential dilution effects, where chemicals in the green module may simply be inert and exhibiting an associated reduced toxicity when analyzed in combination with toxic chemicals. This was not the case, as the β coefficient largely increased when analyzing all chemicals together in comparison to just those in the green module (Fig. 3A). Analysis of a randomly permuted module additionally supported the likely protective role of chemicals in the green module. Specifically, unlike the green module, the random module was not associated with a reduction in biological responses across multiple endpoints, nor was it associated with a reduction in biological responses across multiple endpoints when analyzed in combination with all chemicals (Fig. 3B, Supplementary Material Table S5). These data further demonstrate that the green module's associated reduction in biological responses was unlikely due to chance or potential dilution effects.

A relationship identified here in the quantile g-computation analysis, but not in the previous group correlation-based analysis, was between the brown and yellow modules and neutrophil responses. This is of particular importance, as the previous analysis identified neutrophil count increases as the most sensitive response to biomass burn condensate exposures, with the greatest increase occurring 24 h postexposure (Kim et al., 2018). Here, the brown module showed the most significant (FDR $q = 0.04$) association with neutrophil

increase at 24 h, followed by the yellow module (FDR q = 0.07). These modules also showed potential associations with neutrophil increase at 4 h, though were less significant at this timepoint. Individual chemicals within these mixture models were found to contribute both positive and negative partial effects (Fig. 4). For example, within the brown module, Antimony (Sb), Sulfate (SO_4), and Zinc (Zn) were positively weighted and Calcium (Ca), Iron (Fe), Magnesium (Mg), Manganese (Mn), and Strontium (Sr) were negatively related in relation to neutrophil count at both 4 and 24 h post-exposure, demonstrating the importance of evaluating the overall mixture exposure effect. Within the yellow module, Copper (Cu), Hexavalent Chromium (Cr), and Phosphorus (P) demonstrated positive effects and Nickel (Ni), Sodium (Na), and Titanium (Ti) demonstrated negative effects in relation to neutrophil count at both 4 and 24 h post-exposure.

3.5. Individual chemical analyses do not capture mixture effects

To evaluate whether similar trends are apparent using individual chemical-based approaches in comparison to the aforementioned mixtures-based approaches, each chemical was correlated separately against each biological response (Fig. 5, Supplementary Material Table S6). These results showed some consistency to the mixtures-based results; though many trends identified through mixture analyses were not captured through individual chemical correlations. For example, many significant correlations were identified between individual inorganic elements and ionic constituents and increased cytokine concentrations and lung injury markers. However, the quantile g-computation models were able to identify relationships with neutrophil influx, and better evaluate the overall effect (including partial positive vs. negative effects) associated with the collective mixtures. Individual correlation analyses also identified negative relationships between chemicals in the green module, including coniferyl aldehyde and vanillin, and several biological responses. However, these correlation-based findings did not demonstrate the effects estimated to occur with versus without these chemicals within the collective mixture. These individual correlation results therefore serve as important points of comparison, demonstrating the utility in implementing mixtures modeling approaches to elucidate joint effects of chemicals with respect to biological outcomes associated with wildfire smoke exposure.

4. Discussion

This study aimed to implement mixtures-relevant statistical approaches to evaluate relationships between chemicals present in wildfire smoke, captured as biomass smoke condensate samples, and associated biological responses in the mouse lung. Chemical clustering identified seven different groups of chemicals based on co-occurrence across exposure samples. One chemical group containing methoxyphenols was highlighted as demonstrating negative relationships to several biological responses, suggesting potential protective effects. The overall mixture effect of the chemicals in the identified groups was evaluated through quantile g-computation, and the collective toxicity of chemicals was further identified as reduced when this chemical group, consisting of coniferyl aldehyde and vanillin, was included in the model. Other groups of chemicals such as those containing inorganic elements and ionic constituents showed collective positive relationships with biological responses indicative of toxicity induction. Many of these joint relationships

identified through mixtures modeling were not captured through individual chemical analyses, demonstrating the utility of mixtures-based statistical approaches.

The mixtures-based modeling approaches employed in this study represent advancements towards evaluating exposure-toxicity relationships within complex exposures. For example, the evaluated exposure conditions included variable atmospheres, with 86 specific chemical constituents measured at variable concentrations. When evaluating these types of complex exposures, most studies define chemical groups using a priori classifications based on chemical structure or class. These a priori classifications may miss additional chemistries and related toxicities that co-occur in combination across different chemical classes. For this reason, we identified chemical groups in an unsupervised manner to potentially elucidate understudied relationships between chemicals in wildfire emissions and their related biological responses. An additional advancement was the implementation of quantile g-computation to better evaluate and quantify the joint impact of chemical groups. Understanding the joint impact is important because chemicals may have small effects that are missed when studied individually but may have a meaningful combined impact (Niehoff et al., 2020; White et al., 2020). Methods to estimate the joint impact of exposure to multiple chemicals may help identify the effects of interventions that affect multiple chemicals. Quantile g-computation builds upon previous mixtures-based regression models and addresses inherent complexities of high-dimensional mixture data. For example, unlike weighted quantile sum, another approach to estimate the overall mixture effect of a weighted exposure index of quantized exposures (Carrico et al., 2015), quantile g-computation does not require a directional homogeneity assumption that all exposures have an effect in the same direction (Keil et al., 2020). Some additional advantages of this approach include that it has a simple implementation and interpretation, it can maintain precision despite strong correlations among the exposures, and it is somewhat insensitive to outliers due to the quantization of exposures. In general, evaluating mixtures in the environment remains difficult, as individual components may always act differently in the presence of additional chemicals, and determining the effects of individual components in every possible mixture combination is not feasible. It is therefore important to continue expanding these types of mixtures-based modeling approaches.

Both the data-reduction and mixtures effect analyses highlighted chemicals within the green module, containing methoxyphenols, as associated with reduced biological responses resulting from biomass smoke exposure. Methoxyphenols such as coniferyl aldehyde, eugenol, guaiacol, and vanillin have previously been shown to have anti-inflammatory and protective effects in the lung, commonly detected during conditions of co-exposures (Houser et al., 2012; Huang et al., 2015; Magalhaes et al., 2010; Murakami et al., 2007; Zin et al., 2012). For example, coniferyl aldehyde has been shown to reduce radiation damage in normal lung tissue by increasing stability of heat shock transcriptional factor 1 (HSF1) (Kim et al., 2015a). Eugenol has been shown to reduce pulmonary inflammation resulting from both diesel exhaust particles exposure and LPS exposure (Huang et al., 2015; Magalhaes et al., 2010; Zin et al., 2012). Guaiacol and vanillin were found to inhibit lipopolysaccharide (LPS)-stimulated nuclear factor kappa B (NF-κB) activation and cyclooxygenase-2 (*COX-2*) gene expression in macrophages (Murakami et al., 2007). Pre-treatment of mice with vanillin has specifically been shown to protect against LPS-induced acute lung injury by

inhibiting macrophage activation and lung inflammation (Guo et al., 2019). Vanillin also has demonstrated general protection against exposure-induced DNA damage and resulting mutagenicity throughout several model systems (Anand et al., 2019). There is currently a lack of data on the potential synergistic effects of co-occurring methoxyphenols, in which effects may be increased in the presence of other chemicals. Future in vivo or in vitro studies could be designed, for instance evaluating a biomass burn sample with individual and co-occurring methoxyphenols added in titration, resulting in a direct evaluation of their impact on biological responses. Still, our findings in combination with existing data support the potential reduction in toxicity that may occur in wildfire smoke exposures mediated through the presence of certain co-occurring chemicals, such as methoxyphenols, that warrant further investigation.

Mixtures modeling also identified inorganic elements and ionic constituents as associated with increased pulmonary neutrophil count. Specific compounds that demonstrated positive relationships with neutrophil influx included copper and hexavalent chromium, which have also been shown to cause increased neutrophil count in the lung (Beaver et al., 2009a; Beaver et al., 2009b; Kim et al., 2011). For example, copper nanoparticle exposure increased neutrophil recruitment to the lungs in mice in response to bacterial infection (Kim et al., 2011). Hexavalent chromium exposure has also been shown to directly induce an airway neutrophilic inflammatory response 24 h after exposure in mice (Beaver et al., 2009a; Beaver et al., 2009b). Ionic constituents are known to play important endogenous roles in neutrophil inflammatory processes (Da Silva-Santos et al., 2002; Gallin and Seligmann, 1984; Northover, 1977; Suri et al., 2008), though studies are limited evaluating the impact of changes in ionic profiles specifically from exogenous sources on neutrophil regulation. These data demonstrate the potential influence of inorganic elements and ionic constituents acting in concert to promote neutrophilic responses associated with wildfire smoke exposures.

This study represents a novel application of mixtures modeling to elucidate potential roles of chemicals in wildfire smoke on biological responses; though this study is not without limitations. For example, this study focused on biological responses resulting from the PM fraction and condensable gas-phase semivolatiles of biomass burn emissions. Future studies could be designed to capture the chemical composition and resulting effects from exposure to more gas-phase constituents (e.g., volatile organic compounds and nitrogen oxides). Our study is also based on mice exposed via oropharyngeal aspiration; where additional studies could evaluate biological responses using different models across different exposure routes, particularly inhalation. Indeed, we recently evaluated a subset of biomass fuels (i.e., peat, eucalyptus and red oak smoke) through inhalation exposures in mice and found some similarities in associated biological responses, including the finding that neutrophil count shows the greatest increase significantly associated with exposures to inorganic elements and ionic components of inhaled PM (Kim et al., 2019).

Future studies could also be designed to evaluate potential sex-related differences in susceptibility to toxicity from chemicals groups within biomass smoke, as there is evidence to support varying inflammatory responses dependent upon sex resulting from wood smoke exposures (Rebuli et al., 2019). It would also be advantageous to investigate impacts in other

target tissues as well as endpoints that are indicative of longer-term health consequences, in addition to the highlighted acute toxicity responses from in vivo BALF samples. Though whole-body systems can provide important findings that demonstrate physiological relevancy, in vitro methods could also be employed, for instance through air-liquid exposure interfaces, to capture biological responses across a wider domain of exposure conditions (Zavala et al., 2020), including additional biomass fuels and potential co-exposures occurring alongside biomass burns in wildfire events. These expanded efforts would reduce reliance upon animal testing while providing critical information needed to more rapidly elucidate chemical drivers (particle and gas phase components) of toxicity in wildfire smoke and ultimately lead to improved safety decision making to protect public health.

In the context of global health, the findings presented here increase our understanding of wildfire exposure conditions that induce the greatest health risks. For example, fuels and burn conditions that emit higher amounts of methoxyphenols in combination with other combustion products may cause less toxicity than those that emit lower amounts of methoxyphenols. Indeed, this was apparent in the conditions evaluated here, where smoldering pine produced the highest relative concentrations of methoxyphenols in combination with other harmful compounds, yet smoldering pine was associated with some of the smallest changes in pulmonary toxicity (Kim et al., 2018). Our data also demonstrate that these trends may become exacerbated when inorganics and ionic constituents are emitted. Here, flaming peat, flaming pine needles, and flaming red oak emitted the highest concentrations of inorganics and ionic constituents, which consistently induced many of the greatest increases in pulmonary toxicity (Kim et al., 2018). We therefore show that coupling mixtures computational modeling with chemical and in vivo biological response profiles extracts meaningful information towards identifying exposure scenarios that pose the highest risk associated with wildfire exposure conditions.

In conclusion, this study employed mixtures modeling approaches to elucidate novel relationships between chemical groups in wildfire smoke and associated biological responses in the mouse lung. Both inducers and repressors of biological responses were identified to likely act collectively within the greater mixture across multiple biological response endpoints, representing relationships that would not have been captured through analyses based on individual chemicals. Additionally, these relationships would not have been captured by analyzing mixtures as one unit, without considering the underlying distribution profiles of chemicals. Inducers of biological responses largely included inorganic elements and ionic constituents, while potential repressors of biological responses included methoxyphenols. This project serves as an important case study on strategies that can be implemented to better understand complex mixture effects. Results such as these can yield information towards the prioritization of harmful chemicals in complex co-exposure conditions and ultimate protection of wildfire smoke-induced health outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank Alexander P. Keil, PhD, for his review of the quantile g-computational methods and associated script in R. The research described in this manuscript has been reviewed by the Center for Public Health and Environmental Assessment, U.S. EPA, and the National Institute of Environmental Health Sciences, NIH, and approved for publication. Approval does not signify that contents necessarily reflect the views and policies of the agency, nor does the mention of trade names or commercial products constitute endorsement or recommendation for use. The authors would like to thank Dr. William Boyes, U.S. EPA, and Dr. Alexandra White, NIH, for providing internal technical review of this manuscript.

Funding

This study was supported by a grant from the National Institutes of Health (NIH) from the National Institute of Environmental Health Sciences (1R21ES031740). Additional support was provided by the Institute for Environmental Health Solutions at the Gillings School of Global Public Health, and the Intramural Research Program of the Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina. Nicole Niehoff is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

Abbreviations:

BALF	bronchoalveolar lavage fluid
Ca	Calcium
COX-2	cyclooxygenase-2
Cu	Copper
Cr	Hexavalent Chromium
Fe	Iron
FDR	false discovery rate
GGT	y-glutamyl transferase
HSF1	heat shock transcriptional factor 1
IL-6	interluekin-6
LDH	lactate dehydrogenase
LOD	limits of detection
LPS	lipopolysaccharide
P	Phosphorus
Mg	Magnesium
MIP-2	macrophage inhibitory protein-2
Mn	Manganese
Na	Sodium
NAG	*N*-acetyl-β-D-glycosaminidase

NF-κB	nuclear factor kappa B
Ni	Nickel
PAHs	polycyclic aromatic hydrocarbons
PM	particulate matter
Sb	Antimony
SO_4	Sulfate
Sr	Strontium
Ti	Titanium
TNF-α	tumor necrosis factor-α
WCCNA	weighted chemical co-expression network analysis
WGCNA	weighted gene co-expression network analysis
WGS	weighted quantile sum
Zn	Zinc

References

Agier L, Portengen L, Chadeau-Hyam M, Basagana X, Giorgis-Allemand L, Siroux V, et al., 2016. A systematic comparison of linear regression-based statistical methods to assess Exposome-health associations. Environ. Health Perspect. 124, 1848–1856. [PubMed: 27219331]

Anand A, Khurana R, Wahal N, Mahajan S, Mehta M, Satija S, et al., 2019. Vanillin: a comprehensive review of pharmacological activities. Plant Archives 19, 1000–1004.

Bailey KA, Laine J, Rager JE, Sebastian E, Olshan A, Smeester L, et al., 2014. Prenatal arsenic exposure and shifts in the newborn proteome: interindividual differences in tumor necrosis factor (TNF)-responsive signaling. Toxicol. Sci. 139, 328–337. [PubMed: 24675094]

Bartel J, Krumsiek J, Theis FJ, 2013. Statistical methods for the analysis of high-throughput metabolomics data. Comput. Struct. Biotechnol. J. 4, e201301009. [PubMed: 24688690]

BBC, 2020. Australia fires: A visual guide to the bushfire crisis. 2020. BBC News.

Beaver LM, Stemmy EJ, Constant SL, Schwartz A, Little LG, Gigley JP, et al., 2009a. Lung injury, inflammation and Akt signaling following inhalation of particulate hexavalent chromium. Toxicol. Appl. Pharmacol. 235, 47–56. [PubMed: 19109987]

Beaver LM, Stemmy EJ, Schwartz AM, Damsker JM, Constant SL, Ceryak SM, et al., 2009b. Lung inflammation, injury, and proliferative response after repetitive particulate hexavalent chromium exposure. Environ. Health Perspect. 117, 1896–1902. [PubMed: 20049209]

Black C, Tesfaigzi Y, Bassein JA, Miller LA, 2017. Wildfire smoke exposure and human health: significant gaps in research for a growing public health issue. Environ. Toxicol. Pharmacol. 55, 186–195. [PubMed: 28892756]

Carlin DJ, Rider CV, Woychik R, Birnbaum LS, 2013. Unraveling the health effects of environmental mixtures: an NIEHS priority. Environ. Health Perspect. 121, A6–A8. [PubMed: 23409283]

Carrico C, Gennings C, Wheeler DC, Factor-Litvak P, 2015. Characterization of weighted Quantile sum regression for highly correlated data in a risk analysis setting. J. Agric. Biol. Environ. Stat. 20, 100–120. [PubMed: 30505142]

Cho SH, Tong H, McGee JK, Baldauf RW, Krantz QT, Gilmour MI, 2009. Comparative toxicity of size-fractionated airborne particulate matter collected at different distances from an urban highway. Environ. Health Perspect. 117, 1682–1689. [PubMed: 20049117]

Da Silva-Santos JE, Santos-Silva MC, Cunha Fde Q, Assreuy J, 2002. The role of ATP-sensitive potassium channels in neutrophil migration and plasma exudation. J. Pharmacol. Exp. Ther. 300, 946–951. [PubMed: 11861802]

Eaves LA, Nguyen HT, Rager JE, Sexton KG, Howard T, Smeester L, et al., 2020. Identifying the transcriptional response of Cancer and inflammation-related genes in lung cells in relation to ambient air chemical mixtures in Houston, Texas. Environ. Sci. Technol. 54, 13807–13816. [PubMed: 33064461]

Gallin JI, Seligmann BE, 1984. Neutrophil chemoattractant fMet-Leu-Phe receptor expression and ionic events following activation. Contemp. Top. Immunobiol. 14, 83–108. [PubMed: 6088175]

Guo T, Su Z, Wang Q, Hou W, Li J, Zhang L, et al., 2019. Vanillin protects lipopolysaccharide-induced acute lung injury by inhibiting ERK1/2, p38 and NF-kappaB pathway. Future Med. Chem. 11, 2081–2094. [PubMed: 31538519]

Hamra GB, Buckley JP, 2018. Environmental exposure mixtures: questions and methods to address them. Curr. Epidemiol. Rep. 5, 160–165. [PubMed: 30643709]

Harrell FE, Dupont MC, 2014. R package Hmisc. R Foundation for Statistical Computing, Vienna, Austria 2020.

Hines EP, Mendola P, von Ehrenstein OS, Ye X, Calafat AM, Fenton SE, 2015. Concentrations of environmental phenols and parabens in milk, urine and serum of lactating North Carolina women. Reprod. Toxicol. 54, 120–128. [PubMed: 25463527]

Houser KR, Johnson DK, Ishmael FT, 2012. Anti-inflammatory effects of methoxyphenolic compounds on human airway cells. J. Inflamm. (Lond) 9, 6. [PubMed: 22414048]

Huang X, Liu Y, Lu Y, Ma C, 2015. Anti-inflammatory effects of eugenol on lipopolysaccharide-induced inflammatory reaction in acute lung injury via regulating inflammation and redox status. Int. Immunopharmacol. 26, 265–271. [PubMed: 25863235]

Hurteau MD, Westerling AL, Wiedinmyer C, Bryant BP, 2014. Projected effects of climate and development on California wildfire emissions through 2100. Environ. Sci. Technol. 48, 2298–2304. [PubMed: 24443984]

Keil AP. qgcomp: Quantile G-Computation. 2020, 2020.

Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ, 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ. Health Perspect. 2020, 47004.

Kim JS, Adamcakova-Dodd A, O'Shaughnessy PT, Grassian VH, Thorne PS, 2011. Effects of copper nanoparticle exposure on host defense in a murine pulmonary infection model. Part Fibre Toxicol. 8, 29. [PubMed: 21943386]

Kim YH, Tong H, Daniels M, Boykin E, Krantz QT, McGee J, et al., 2014. Cardiopulmonary toxicity of peat wildfire particulate matter and the predictive utility of precision cut lung slices. Part Fibre Toxicol. 11, 29. [PubMed: 24934158]

Kim SY, Lee HJ, Nam JW, Seo EK, Lee YS, 2015a. Coniferyl aldehyde reduces radiation damage through increased protein stability of heat shock transcriptional factor 1 by phosphorylation. Int. J. Radiat. Oncol. Biol. Phys. 91, 807–816. [PubMed: 25752395]

Kim YH, Wyrzykowska-Ceradini B, Touati A, Krantz QT, Dye JA, Linak WP, et al., 2015b. Characterization of size-fractionated airborne particles inside an electronic waste recycling facility and acute toxicity testing in mice. Environ. Sci. Technol. 49, 11543–11550. [PubMed: 26332991]

Kim YH, Warren SH, Krantz QT, King C, Jaskot R, Preston WT, et al., 2018. Mutagenicity and lung toxicity of smoldering vs. Flaming Emissions from Various Biomass Fuels: Implications for Health Effects from Wildland Fires. Environ. Health Perspect. 126, 017011. [PubMed: 29373863]

Kim YH, King C, Krantz T, Hargrove MM, George IJ, McGee J, et al., 2019. The role of fuel type and combustion phase on the toxicity of biomass smoke following inhalation exposure in mice. Arch. Toxicol. 93, 1501–1513. [PubMed: 31006059]

Langfelder P, Horvath S, 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559. [PubMed: 19114008]

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. 2017, 2017.

Liu JC, Pereira G, Uhl SA, Bravo MA, Bell ML, 2015. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. Environ. Res. 136, 120–132. [PubMed: 25460628]

Magalhaes CB, Riva DR, DePaula LJ, Brando-Lima A, Koatz VL, Leal-Cardoso JH, et al., 2010. In vivo anti-inflammatory action of eugenol on lipopolysaccharide-induced lung injury. J. Appl. Physiol. (1985) 108, 845–851. [PubMed: 20075264]

Murakami Y, Hirata A, Ito S, Shoji M, Tanaka S, Yasui T, et al., 2007. Re-evaluation of cyclooxygenase-2-inhibiting activity of vanillin and guaiacol in macrophages stimulated with lipopolysaccharide. Anticancer Res. 27, 801–807. [PubMed: 17465205]

Naeher LP, Brauer M, Lipsett M, Zelikoff JT, Simpson CD, Koenig JQ, et al., 2007. Woodsmoke health effects: a review. Inhal. Toxicol. 19, 67–106.

Niehoff NM, Keil AP, O'Brien KM, Jackson BP, Karagas MR, Weinberg CR, et al., 2020. Metals and trace elements in relation to body mass index in a prospective study of US women. Environ. Res. 184, 109396. [PubMed: 32209500]

Northover BJ, 1977. Effect of indomethacin and related drugs on the calcium ion-dependent secretion of lysosomal and other enzymes by neutrophil polymorphonuclear leucocytes in vitro. Br. J. Pharmacol. 59, 253–259. [PubMed: 402167]

Rager JE, Bailey KA, Smeester L, Miller SK, Parker JS, Laine JE, et al., 2014. Prenatal arsenic exposure and the epigenome: altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. Environ. Mol. Mutagen. 55, 196–208. [PubMed: 24327377]

Rager JE, Auerbach SS, Chappell GA, Martin E, Thompson CM, Fry RC, 2017. Benchmark dose modeling estimates of the concentrations of inorganic arsenic that induce changes to the neonatal transcriptome, proteome, and epigenome in a pregnancy cohort. Chem. Res. Toxicol. 30, 1911–1920. [PubMed: 28927277]

RDocumentation. Sample: Random Samples and Permutations. 2020, 2020.

Rebuli ME, Speen AM, Martin EM, Addo KA, Pawlak EA, Glista-Baker E, et al., 2019. Wood smoke exposure alters human inflammatory responses to viral infection in a sex-specific manner. A randomized, placebo-controlled study. Am. J. Respir. Crit. Care Med. 199, 996–1007. [PubMed: 30360637]

Reid CE, Brauer M, Johnston FH, Jerrett M, Balmes JR, Elliott CT, 2016.Critical reviewof health impacts of wildfire smoke exposure. Environ. Health Perspect. 124, 1334–1343. [PubMed: 27082891]

Ryan KR, Huang MC, Ferguson SS, Waidyanatha S, Ramaiahgari S, Rice JR, et al., 2019. Evaluating sufficient similarity of botanical dietary supplements: combining chemical and in vitro biological data. Toxicol. Sci. 172, 316–329. [PubMed: 31504990]

Suri S, Taylor MA, Verity A, Tribolo S, Needs PW, Kroon PA, et al., 2008. A comparative study of the effects of quercetin and its glucuronide and sulfate metabolites on human neutrophil function in vitro. Biochem. Pharmacol. 76, 645–653. [PubMed: 18639531]

TIME. U.S. Natural Disasters: Firefighters Make Progress Against Fires Raging in California. Here's a Map of All the Big Fires Across the State. 2020, 2019.

Tong H, Cheng WY, Samet JM, Gilmour MI, Devlin RB, 2010. Differential cardiopulmonary effects of size-fractionated ambient particulate matter in mice. Cardiovasc. Toxicol. 10, 259–267. [PubMed: 20602262]

Wang C, Gong B, Bushel PR, Thierry-Mieg J, Thierry-Mieg D, Xu J, et al., 2014. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. Nat. Biotechnol. 32, 926–932. [PubMed: 25150839]

Wei T, Simko V, 2017. R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). R Foundation for Statistical Computing, Vienna, Austria 2020.

Westerling AL, Hidalgo HG, Cayan DR, Swetnam TW, 2006. Warming and earlier spring increase western U.S. forest wildfire activity. Science 313, 940–943. [PubMed: 16825536]

White AJ, O'Brien KM, Niehoff NM, Jackson BP, Karagas MR, Weinberg CR, et al., 2020. Toenail metal concentrations and age at menopause: a prospective study. Environ. Epidemiol. 4, e0104. [PubMed: 32832842]

Zavala J, Freedman AN, Szilagyi JT, Jaspers I, Wambaugh JF, Higuchi M, et al., 2020. New approach methods to evaluate health risks of air pollutants: critical design considerations for in vitro exposure testing. Int. J. Environ. Res. Public Health 17.

Zhang G, He P, Tan H, Budhu A, Gaedcke J, Ghadimi BM, et al., 2013. Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. Clin. Cancer Res. 19, 4983–4993. [PubMed: 23918603]

Zin WA, Silva AG, Magalhaes CB, Carvalho GM, Riva DR, Lima CC, et al., 2012. Eugenol attenuates pulmonary damage induced by diesel exhaust particles. J. Appl. Physiol. (1985) 112, 911–917. [PubMed: 22194320]

HIGHLIGHTS

- Mice exposed to biomass burn samples representing different wildfire scenarios
- Biological responses were evaluated including pulmonary immune and injury markers.
- Clustering, data reduction, and quantile g-computational methods addressed mixtures.
- Methoxyphenols suggested to reduce responses in presence of other chemicals
- Inorganics and ionic constituents suggested to induce biological responses

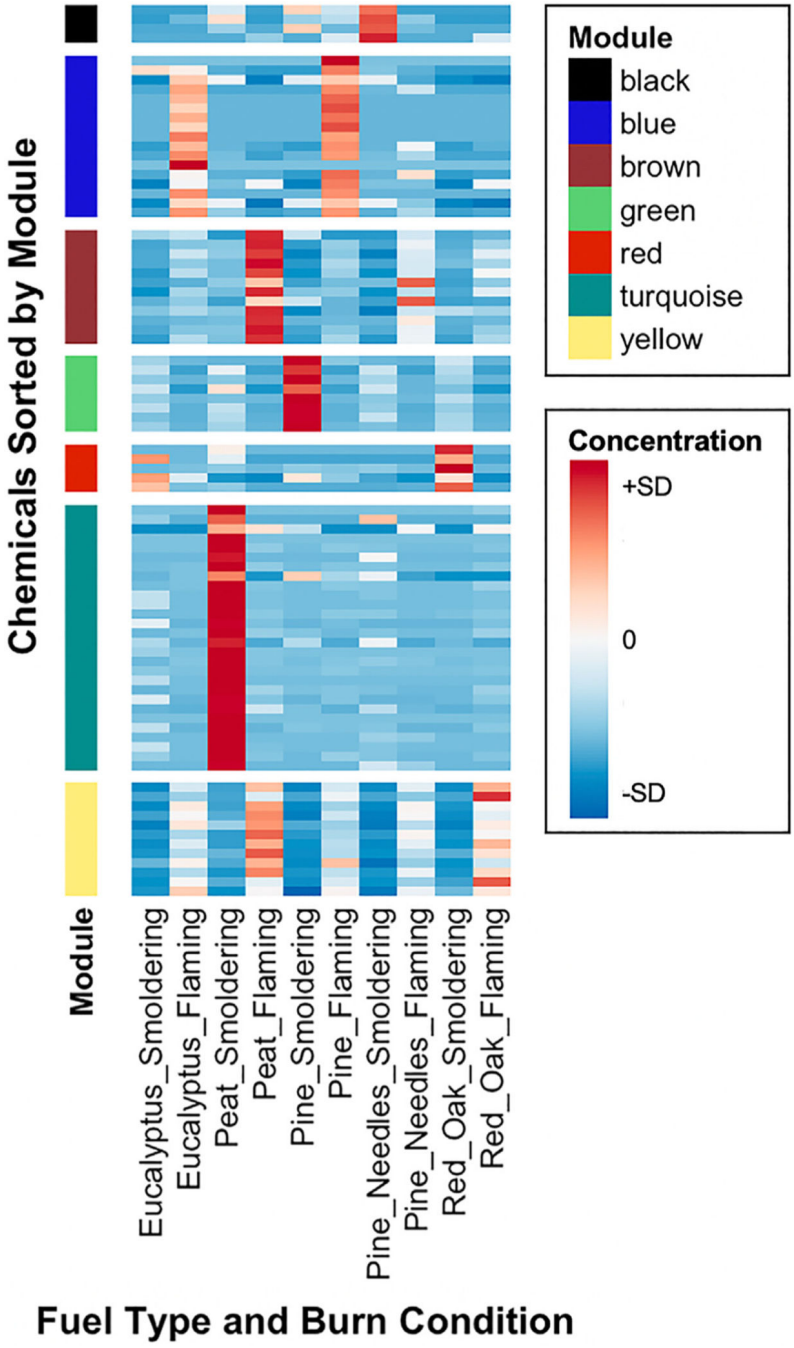


Fig. 1.

Heat map of the chemical concentrations measured in biomass smoke condensate samples, with chemicals organized into modules based on co-occurrence. For specific chemicals organized per module, alongside chemical class, see Table 1. Chemical concentrations are z-score normalized across biomass burn samples. Note that module groups are indicated by color and appear in the same order within the heat map as in the legend.

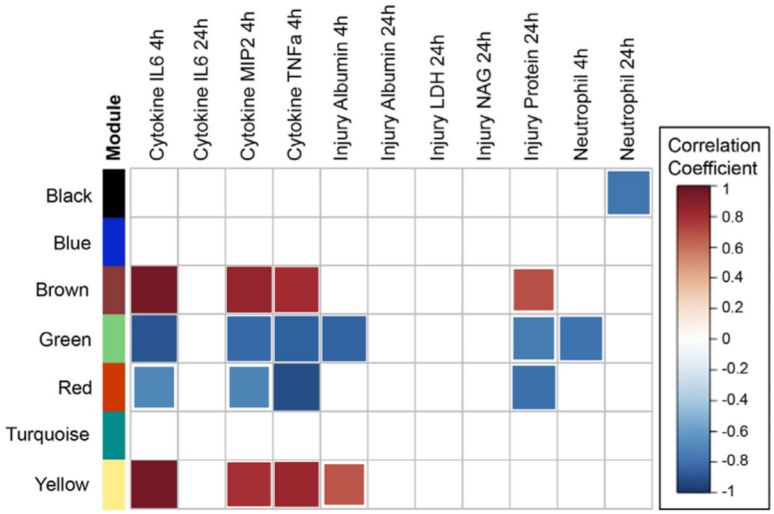


Fig. 2.

Correlations between chemical groups (i.e., module eigenvalues) and biological responses in the mouse lung. Significant ($p < 0.05$) correlations are shown as colored squares, and white squares reflect insignificant correlations.

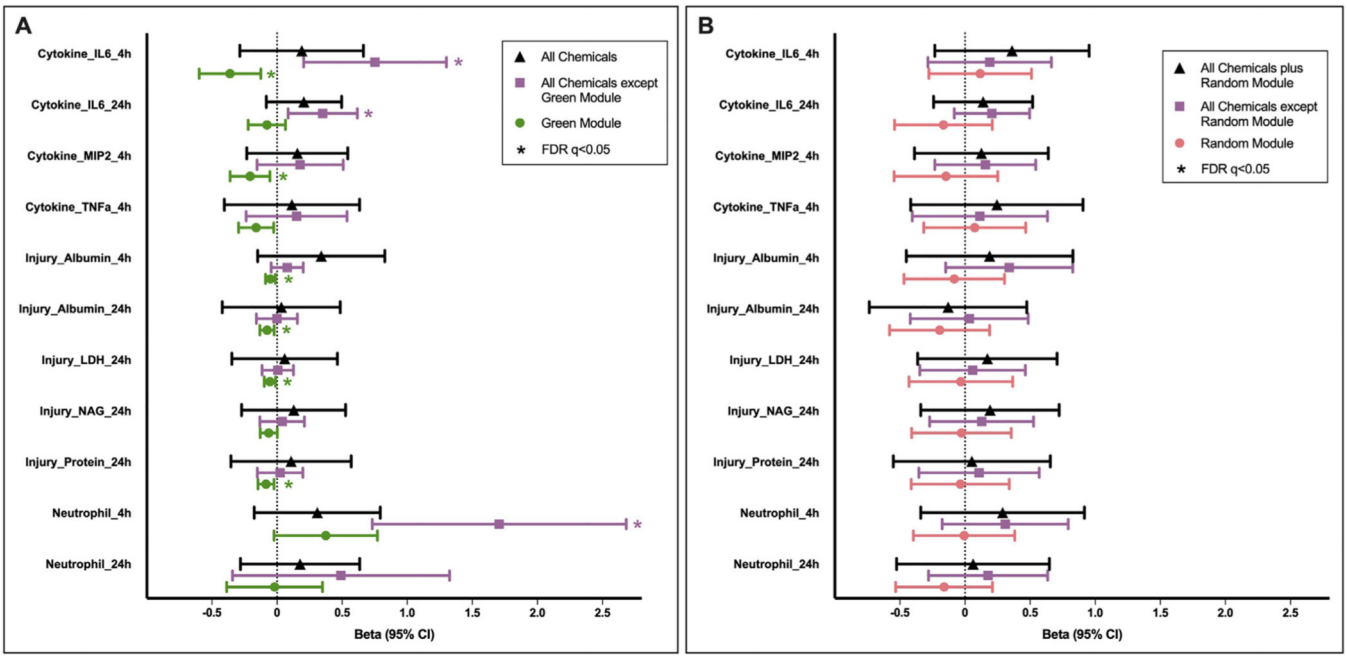


Fig. 3.

Mixtures-based associations between chemical groups and biological responses within the mouse lung, highlighting the impacts of (A) modeling with vs. without chemicals in the green module, and (B) modeling with vs. without random noise, generated as a module of chemical distributions based on random permutations. These values represent quantile g-computation estimates for the change in biological endpoint (ln-scaled) for a one quintile increase in chemical concentration (z-score normalized) in biomass burn condensate samples, summarized as beta coefficients and 95% confidence intervals. Note that significant changes in biological responses are estimated for chemical groups in (A) but not in (B), further supporting the observation of mixtures-based relationships that are unlikely due to chance or dilution effects.

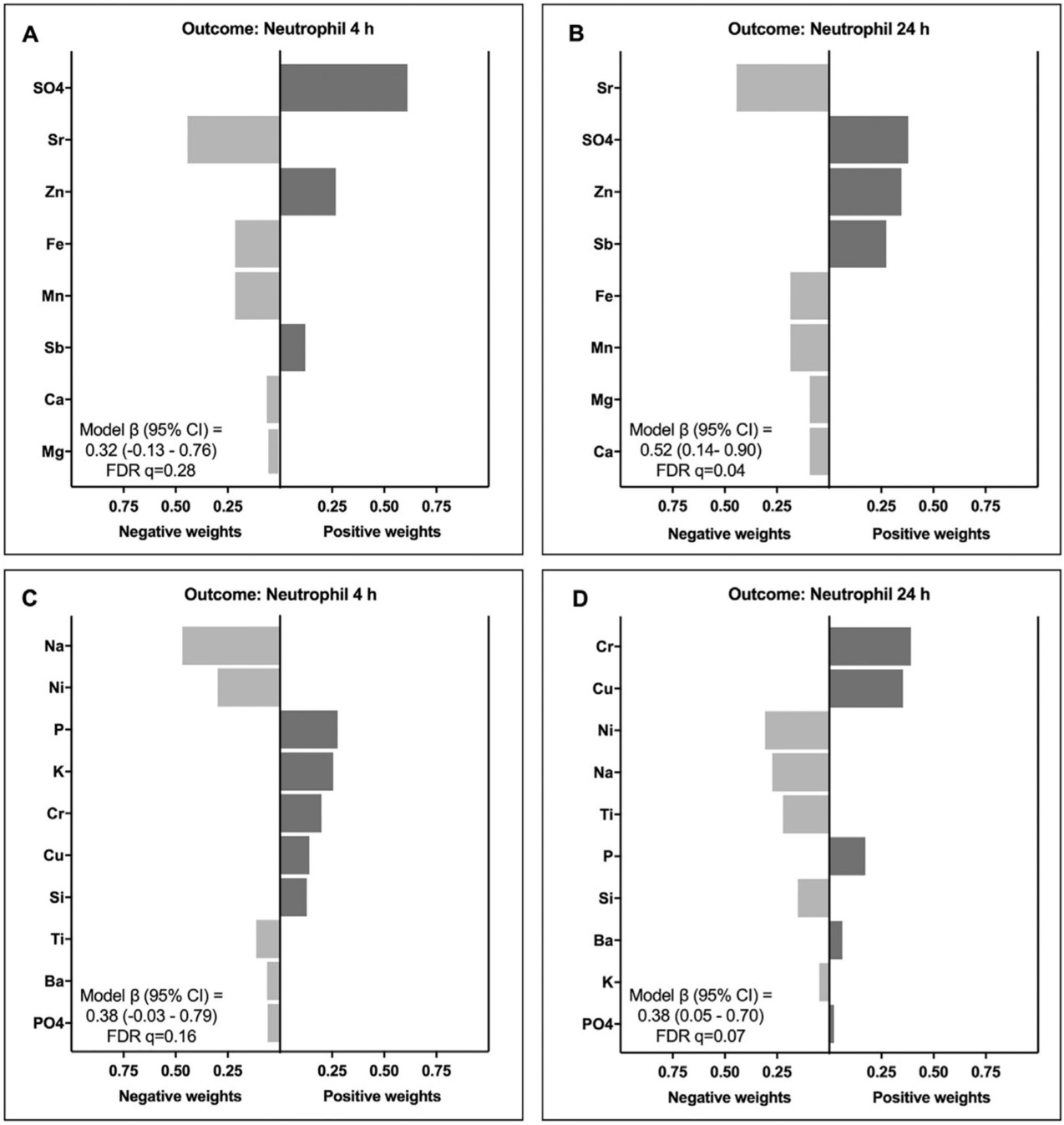


Fig. 4.

Weights representing the proportion of the positive or negative partial effect in the quantile g-computation models for chemicals in the brown module associated with neutrophil count increases (A) 4 h post-exposure and (B) 24 h post-exposure, and chemicals in the yellow module associated with neutrophil count increases (C) 4 h post-exposure and (D) 24 h post-exposure in the mouse lung. Note that chemical concentrations were z-score normalized and biological endpoints were ln-transformed prior to model evaluation.

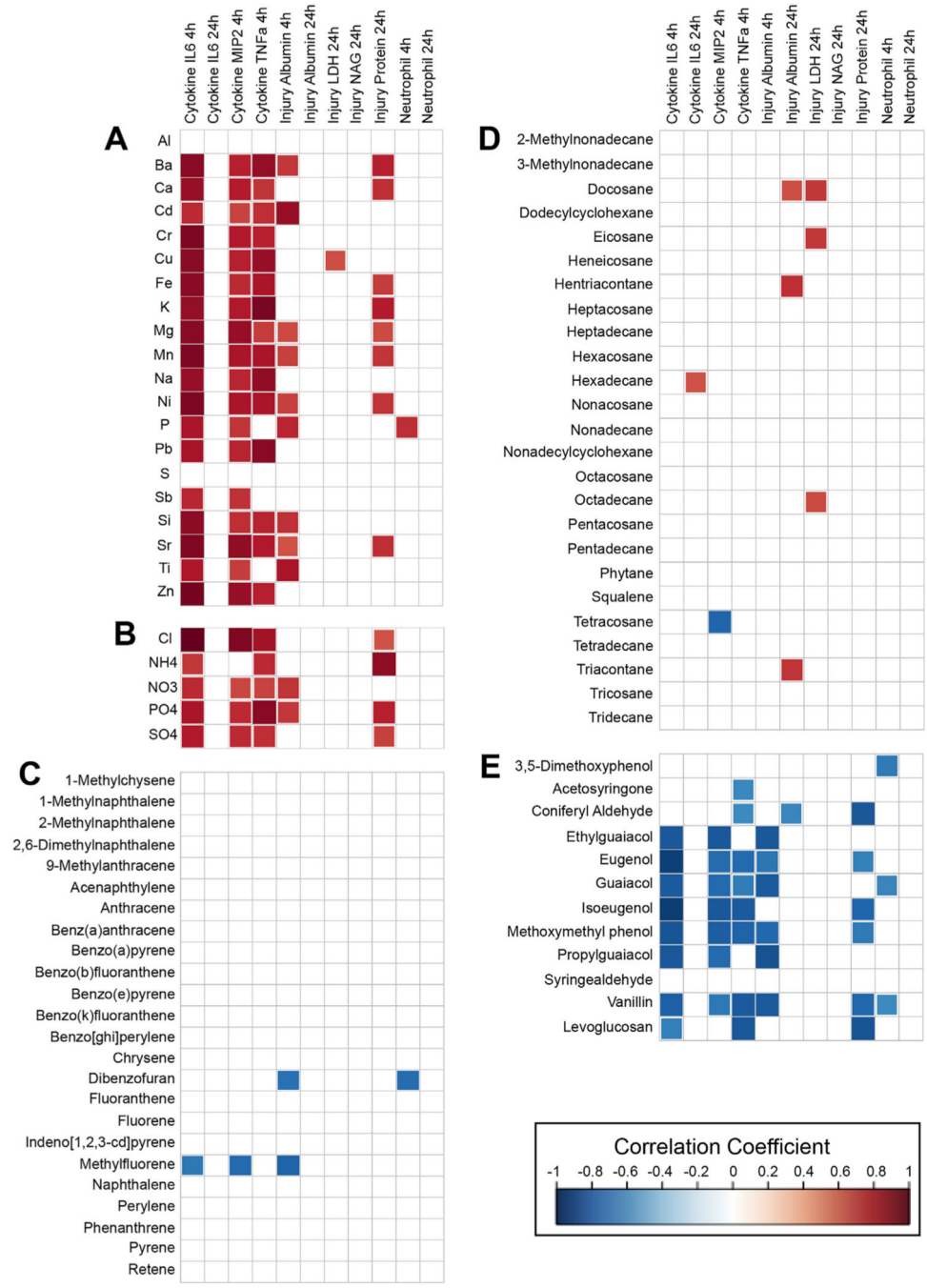


Fig. 5.

Correlations between individual chemicals and biological endpoints, arranged according to the following chemical classes: (A) inorganic elements, (B) ionic constituents, (C) PAHs, (D) n-alkanes, and (E) methoxyphenols and levoglucosan. Significant ($p < 0.05$) correlations are shown as colored squares, and white squares reflect insignificant correlations.

Table 1

Chemicals ($N = 86$) that were measured in biomass smoke condensate samples. Chemical concentrations are summarized alongside chemical grouping results, designated as module assignments assigned to different colors. Associated CASRN are provided in Supplementary Material Table S2.

Chemical	Average (Min - Max)	Module assignment
n-Alkanes (ng/μL)		
2-Methylnonadecane	0.19 (<LOD - 1.36)	Turquoise
3-Methylnonadecane	0.42 (<LOD - 1.13)	Turquoise
Docosane	1.22 (0.16–9.46)	Turquoise
Dodecylcyclohexane	0.07 (<LOD - 0.25)	Turquoise
Eicosane	0.99 (<LOD - 7.55)	Turquoise
Heneicosane	0.93 (<LOD - 6.77)	Turquoise
Hentriacontane	1.13 (<LOD - 8.3)	Turquoise
Heptacosane	1.15 (<LOD - 8.53)	Turquoise
Heptadecane	0.71 (<LOD - 5.05)	Turquoise
Hexacosane	0.84 (<LOD - 5.51)	Turquoise
Hexadecane	0.63 (<LOD - 4.36)	Turquoise
Nonacosane	1.69 (<LOD - 14.14)	Turquoise
Nonadecane	0.78 (<LOD - 5.89)	Turquoise
Nonadecylcyclohexane	0.12 (<LOD - 1)	Turquoise
Octacosane	1.03 (<LOD - 7.75)	Turquoise
Octadecane	0.86 (<LOD - 5.76)	Turquoise
Pentacosane	0.88 (<LOD - 5.57)	Turquoise
Pentadecane	0.64 (<LOD - 3.78)	Turquoise
Phytane	0.11 (<LOD - 0.83)	Turquoise
Squalane	0.12 (<LOD - 0.49)	Black
Tetracosane	1.22 (<LOD - 8.27)	Turquoise
Tetradecane	0.32 (<LOD - 1.6)	Turquoise
Triacontane	1.23 (<LOD - 9.27)	Turquoise
Tricosane	1.16 (<LOD - 7.91)	Turquoise
Tridecane	0.12 (<LOD - 0.38)	Turquoise
PAHs (ng/μL)		
1-Methylchysene	0.02 (<LOD - 0.05)	Blue
1-Methylnaphthalene	0.19 (<LOD - 0.45)	Black
2-Methylnaphthalene	0.13 (<LOD - 0.36)	Black
2,6-Dimethylnaphthalene	0.17 (<LOD - 0.64)	Turquoise
9-Methylanthracene	0.11 (<LOD - 0.15)	Turquoise
Acenaphthylene	0.32 (<LOD - 1.13)	Blue
Anthracene	0.26 (<LOD - 0.59)	Blue
Benz(*a*)anthracene	0.12 (<LOD - 0.39)	Blue
Benzo(*a*)pyrene	0.05 (<LOD - 0.3)	Blue
Benzo(*b*)fluoranthene	0.08 (<LOD - 0.32)	Blue

Chemical	Average (Min - Max)	Module assignment
Benzo(*e*)pyrene	0.06 (<LOD - 0.15)	Blue
Benzo(*k*)fluoranthene	0.07 (<LOD - 0.33)	Blue
Benzo[*ghi*]perylene	0.1 (<LOD - 0.18)	Blue
Chrysene	0.12 (<LOD - 0.39)	Blue
Dibenzofuran	0.13 (<LOD - 0.94)	Red
Fluoranthene	0.96 (0.08–3.54)	Blue
Fluorene	0.39 (<LOD - 1.29)	Turquoise
Indeno[1,2,3-*cd*]pyrene	0.06 (<LOD - 0.12)	Blue
Methylfluorene	0.18 (<LOD - 0.81)	Turquoise
Naphthalene	0.21 (<LOD - 0.81)	Blue
Perylene	0.01 (<LOD - 0.04)	Blue
Phenanthrene	1.02 (0.07–2.23)	Blue
Pyrene	0.33 (<LOD - 1.21)	Blue
Retene	3.42 (<LOD - 17.13)	Black
Methoxyphenols (ng/μL)		
3,5-Dimethoxyphenol	0.67 (<LOD - 2.54)	Red
Acetosyringone	7.26 (<LOD - 28.45)	Red
Coniferyl aldehyde	53.47 (<LOD - 367.36)	Green
Ethylguaiacol	5.99 (<LOD - 26.48)	Green
Eugenol	4.76 (<LOD - 32.46)	Green
Guaiacol	12.38 (<LOD - 45.73)	Green
Isoeugenol	174.27 (<LOD - 1294.16)	Green
Methoxymethyl phenol	14.92 (<LOD - 88.49)	Green
Propylguaiacol	1.42 (<LOD - 7.13)	Green
Syringealdehyde	44.04 (<LOD - 229.37)	Red
Vanillin	27.87 (0.5–188.86)	Green
Levoglucosan (ng/μL)		
Levoglucosan	189.92 (24.22–503.76)	Red
Inorganic elements (ng/mL)		
Al	34.13 (<LOD - 151.07)	Brown
Ba	5.53 (0.25–14.52)	Yellow
Ca	533.78 (53.62–2359.41)	Brown
Cd	0.11 (<LOD - 0.44)	Yellow
Cr	2.07 (0.18–5.46)	Yellow
Cu	21.53 (5.08–49.41)	Yellow
Fe	24.44 (2.13–78.15)	Brown
K	66.89 (1.25–248.75)	Yellow
Mg	50.32 (8.62–195.44)	Brown
Mn	1.02 (0.16–3.49)	Brown
Na	577.74 (31.24–1547.36)	Yellow
Ni	2.48 (0.18–8.51)	Yellow
P	18.44 (3–40.05)	Yellow

Chemical	Average (Min - Max)	Module assignment
Pb	0.14 (<LOD - 0.45)	Brown
S	885.98 (<LOD - 3651.04)	Brown
Sb	1.4 (0.23–4.31)	Brown
Si	590.13 (47.61–1988.07)	Yellow
Sr	3.37 (0.38–15.75)	Brown
Ti	1.21 (<LOD - 2.36)	Yellow
Zn	102.84 (9.51–408.96)	Brown
Ions (μg/mL)		
Cl	0.28 (<LOD - 0.78)	Yellow
NH4	0.04 (0–0.18)	Brown
NO3	0.04 (<LOD - 0.1)	Blue
PO4	4.69 (0–16.31)	Yellow
SO4	1.31 (<LOD - 7.95)	Brown

Table 2

Quantile g-computation estimates for the change in biological response (ln-scaled) for a one quintile increase in chemical concentration (z-score normalized) in biomass burn condensate samples. Results from select chemical groups are shown, highlighting the impacts of modeling with vs. without chemicals in the green module. The β coefficient value summarizes the average across the models evaluating all biological responses. For individual model results, see Fig. 3 and Supplemental Material Table S3.

Wildfire chemical group	Average β coefficient
All chemicals	0.165
All chemicals except green module	0.343
Green module	−0.071