Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Alesha Thompson Date

Unregulated Heavy Metals in United States Public Water Systems: An Assessment of Contaminant Co-Occurrence and Human Health Risk

By

Alesha Thompson
Master of Public Health

Global Environmental Health

_____

Matthew O. Gribble, PhD DABT
Committee Chair

_____

Tim Frederick, MPH
Committee Member

_____

Michele Monti, MS MPH
Committee Member

Unregulated Heavy Metals in United States Public Water Systems: An Assessment of Contaminant Co-Occurrence and Human Health Risk

By

Alesha Thompson

B.S.

University of Oklahoma

2017

Thesis Committee Chair: Matthew O. Gribble, PhD DABT

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2019

Abstract

Unregulated Heavy Metals in United States Public Water Systems: An Assessment of Contaminant Co-Occurrence and Human Health Risk

By Alesha Thompson

Background: The United States Environmental Protection Agency monitors unregulated contaminants in drinking water and consolidates these results in the National Contaminant Occurrence Database. Our objective was to assess contaminant co-occurrence and human health risk of unregulated metals (chromium, chromium-6, molybdenum, vanadium, cobalt and strontium) over 2013-2015.

Methods: Multilevel Tobit regressions with state and water system random intercepts were used to estimate geometric means of each contaminant in public water systems. Human health risk was assessed by using the Environmental Protection Agency Regional Screening Level calculator. Co-occuring contaminant gene interactions were examined by using the Comparative Toxicogenomics Database and the Database for Annotation, Visualization and Integrated Discovery.

Results: Public water supplies' geometric means of vanadium and chromium were positively associated in the water samples recorded ($r = 0.45$, $p < 0.01$), and these contaminants co-occurred in individual water samples. The geospatial maps of these contaminants in the top 5 cm of soil in the contiguous United States are similar. There were 24 overlapping genes that interact with these chemicals, and the most affected pathway was the HIF-1 pathway.

Conclusion: This assessment is a preliminary step in toward understanding the potential health implications of unregulated contaminants in United States drinking water. Certain public water system samples returned values that represented potentially negative health effects and they should be examined further to understand what is causing the high values. Further research needs to address the cumulative human health risk of ingesting more than one contaminant in the drinking water.

Keywords: water pollutants; water quality; drinking water; risk assessment; toxicogenetics; medical geology; GeoHealth

Unregulated Heavy Metals in United States Public Water Systems: An Assessment of Contaminant Co-Occurrence and Human Health Risk

By

Alesha Thompson

B.S.

University of Oklahoma

2017

Thesis Committee Chair: Matthew O. Gribble, PhD DABT

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Global Environmental Health
2019

TABLE OF CONTENTS

1. INTRODUCTION

There are thousands of chemicals in United States drinking water; some are potentially beneficial, and some are potentially harmful (Guidotti, 2009). Very few of these contaminants are regulated. Some of the chemicals found in drinking water are naturally occurring and although human activities can increase the concentrations available, fully eliminating these exposures from the environment is impossible (Gong & Schaubel, 2018). A number of heavy metals are naturally occurring contaminants and occur frequently in United States public water systems at high levels (Simic, 2017). Understanding which heavy metals pose health risks when ingested, and at what doses, is a priority for environmental health and environmental engineering professionals because these contaminants will be an exposure challenge regardless of polluting activities. The scientific community has begun to embrace the challenge of addressing how these naturally occurring contaminants may affect human health (Almada, Golden, Osofsky, & Myers, 2017; Bundschuh et al., 2017).

The United States has a surveillance system in place for these unregulated contaminants and years of data stored in the National Contaminant Occurrence Database. The Environmental Protection Agency (EPA) monitors unregulated contaminants in drinking water through two processes: The Contaminant Candidate List (CCL) and the Unregulated Contaminant Monitoring Rule (UCMR) (OW US EPA, 2015b). Through a multi-year process, organic contaminants, hormones, synthetic compounds and heavy metals are selected on criteria set forth by the Safe Drinking Water Act (OW US EPA, 2015b) for monitoring under the CCL (US EPA, 2015b). Contaminants from the CCL are then chosen for monitoring under the UCMR (US EPA, 2015) if they satisfy two conditions: first, that the contaminant has potential to cause health effects, and second, the contaminant occurs in water systems at concentrations that would constitute a public

health concern (OW US EPA, 2015b). This is a three-to-four-year process wherein selected public water systems are regularly sampled for the contemporary list of UCMR contaminants (*The Third Unregulated Contaminant Monitoring Rule (UCMR 3); Searching for Emerging Contaminants in Drinking Water*, 2015).

The objective of this secondary data analysis of the heavy metal measurements collected for the third wave of the EPA's UCMR process (UCMR 3) is to assess the co-occurrence of contaminants and explore potential dimensions of the health risk that these chemicals may pose.

2. METHODS

2.1. Data

The Environmental Protection Agency Unregulated Contaminant Monitoring Rule 3 dataset contained detailed information about each sample obtained by each water system in the National Contaminant Occurrence Database on the US EPA website (OW US EPA, 2015a). The third wave of UCMR monitoring took place over 2013-2015, and all water systems serving > 10,000 people were required to monitor for the assessment list. A number of heavy metals [chromium (Cr), chromium-6 (Cr (VI)), strontium (St), molybdenum (Mo), vanadium (V) and cobalt (Co)] were included on the UCMR 3 assessment list. Public water systems were required to sample groundwater twice in one consecutive 12-month period and these samples were obtained five to seven months apart. The water was sampled at the entry point to the distribution system as well as distribution system maximum-residence time-sampling locations and resulted in over 185,000 water samples for these heavy metals (*The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Fact Sheet for Assessment Monitoring (List 1 Contaminants)*,

2016). Samples were then distributed to EPA-approved laboratories and results of the monitoring are stored in the National Drinking Water Contaminant Occurrence Database (Simic, 2017).

Public Water System ID and Public Water System Name identify each large water system that sampled for the contaminants. There were at least four samples per Public Water System in each state. Sample Method identifies the analytical method deemed appropriate for the specific contaminant by the EPA. Collection Date and Contaminant Name were recorded for each sample. The sample's chemical concentration value was also documented as whether it was above or below the Minimum Reporting Level (MRL) (Simic, 2017). The MRL is based on the ability what the specific analytical method for that contaminant can detect (Simic, 2017). For any sample below the MRL, it was recorded as missing in the dataset along with the symbol "<".

While cleaning the dataset, only the heavy metal samples were selected. Samples were also refined to only include results from large water systems (serving > 10,000 people) (Table S1).

2.2. Summary of Individual Contaminant Reported Values

Box-and-whiskers plots (Krzywinski & Altman, 2014) were used to summarize the distribution of reported values for each contaminant separately. Many reported values were below the MRL, so these distributions may be unreliable.

2.3. Assessment of Contaminant Co-Occurrence

Tobit regression models account for censored data (Gong & Schaubel, 2018) and mixed-effect Tobit models account for clustering of observations in a dataset (Lu, 2018). Mixed-effect Tobit regression models of the log-transformed contaminants, left-censored at the log-transformed MRL, with normally distributed random intercepts for community water system and

state, but no fixed effects, were used to predict the mean log-concentrations of each contaminant for each community water system. Models were fitted for each chemical separately. The predicted mean log-concentrations were exponentiated to obtain estimated geometric means of each chemical for each water system. In order to calculate co-occurrence of contaminants in public water systems, the predicted community water system-level geometric means were compared pair-wise across chemicals with scatterplots. A Pearson correlation (de Winter, Gosling, & Potter, 2016) of the predicted values was estimated to understand the strength of the relationships showing in public water systems sampled. Correlations were considered Bonferroni significant accounting for 15 hypothesis tests at $p < 0.0033$ (Dunn, 1961).

In a secondary analysis, to evaluate whether the associations observed between aggregated data summaries for chemicals modeled separately were consistent with the patterns of chemical co-occurrence in individual water samples (Sedgwick, 2015), we augmented these mixed-effect Tobit regression models with fixed effects allowing quartiles of one chemical to predict another chemical. Quartiles of predictor chemicals were used to avoid MRL issues. We fit models with vanadium predicting chromium, and of chromium predicting vanadium (Table 2).

All mixed-effect Tobit regression models were implemented using Stata 15.1 S/E.

2.4. Non-carcinogenic and Carcinogenic Health Risk Calculations

To assess human health risk as a result of the exposure to the levels sampled in drinking water recorded by the UCMR 3, the EPA Regional Screening Level (RSL) calculator was used in the calculations. The RSL calculator is used by EPA risk assessors, Remedial Project Managers and On Scene Coordinators to determine to make decisions about CERCLA hazardous waste sites and whether the associated levels need further investigation or cleanup (“Regional

Screening Levels," 2018.). The RSL calculator uses updated toxicity information from multiple sources and is a strong tool to assess potential human health risk.

The geometric means for each community water system tested along with the minimum and maximum values of samples per state were input into the RSL calculator used by the EPA in Superfund risk assessments ("Regional Screening Levels", 2018) . The calculator was set to use the resident tap-water equations with reference dose set as chronic which resulted in the equations as follows (ORD US EPA, 2015) This was repeated with each value (µg/L) for each contaminant by state.

Noncarcinogenic – Adult Equation for Tapwater (ORD US EPA, 2015):

$$SL_{res\text{-}wat\text{-}nc\text{-}ing\text{-}a}(\mu g/L) = \frac{THQxAT_{res-a}(\frac{365\ days}{year}\ x\ ED_{res}(26\ years)xBW_{res-a}(80\ kg)x(1000\ \frac{\mu g}{mg})}{EF_{res-a}(\frac{350\ days}{year})xED_{res}(26\ years)\ x\ (\frac{1}{RfD_0(\frac{mg}{kg}-d)})xIRW_{res-a}(2.5\ \frac{L}{day})}$$

Tapwater Ingestion equation for Carcinogenic Effects (ORD US EPA, 2015):

$$SL_{res\text{-}wat\text{-}ca\text{-}ing}(\mu g/L) = \frac{TRxAT_{res}(\frac{365\ days}{year}\ x\ LT(70\ years)x(1000\ \frac{\mu g}{mg})}{CSF_0\left(\frac{mg}{kg-day}\right)^{-1}\ x\ (IFWres-adj\left(327.95\frac{L}{kg}\right))}$$

where:

$$IFW_{res-adj}\left(\frac{327.95\ L}{kg}\right) = (\frac{EF_{res-c}\left(\frac{350\ days}{year}\right)xED_{res-c}(6\ years)=IRW_{res-c}\left(\frac{0.78\ L}{day}\right)}{BW_{res-c}(15\ kg)} +$$

$$\frac{EF_{res-a}\left(\frac{350\ days}{year}\right)x(ED_{res}(26\ years)-ED_{res-c}(6\ years))xIRW_{res-a}(\frac{2.5\ L}{day})}{BW_{res-a}(80\ kg)})$$

The RSL calculator displayed toxicity information, recommended screening levels, hazard indexes for adults and children, and carcinogenic risk when applicable. Using the risk information generated, box-and-whisker plots (Krzywinski & Altman, 2014) were created using

Stata 15.1 S/E to visually represent the distribution of the human health risks for each contaminant.

Chromium-6 was the only metal included on UCMR 3 with potential carcinogenic risk. A separate box-and-whisker plot was generated to visualize the levels associated with the exposure levels recorded from the community water systems tested.

2.5. Toxicogenomic Pathway Investigation

The Comparative Toxicogenomic Database (CTD) is a public database created to better understand the impact of environmental exposures on human health. It consists of information on chemical-gene/protein interactions, chemical-disease and gene-disease relationships. By combining this data with genetic pathway information, it helps answer research questions about the mechanisms behind diseases linked to environmental exposures (Davis et. al, 2018). The Database for Annotation, Visualization and Integrated Discovery (DAVID) was also used to understand the information gained from using the CTD (Huang, Sherman, & Lempicki, 2009). Genes were input in DAVID with species selected as Homo sapiens.

Chemicals with a significant correlating relationship were examined with the Venn Viewer tool that compared chemical databases from CTD with each other. The Venn Viewer results were presented as genes related to each chemical on its own and as genes that relate to both chemicals. The overlapping gene list from CTD was then input in DAVID as official gene symbols with species was selected as Homo sapiens. Using the Functional Annotation Tool allowed visualization of the toxicogenomic pathways from the KEGG PATHWAY database, which is a collection of pathway maps that represent information on the molecular level for

processes such as metabolism, genetic information processing and human diseases (Kanehisa & Goto, 2000).

2.6. Mapping of Co-Occurring Contaminants

The chemicals that were significantly correlated were mapped using geochemical and mineralogical maps for soil in the United States that were publicly available ("Geochemical and mineralogical maps for soils of the conterminous United States - Data.gov," 2019). Files were obtained for the specific chemicals and used to create maps in ArcGIS 10.6.

3. RESULTS

3.1. Assessment of Contaminant Co-Occurrence

The correlations of community level geometric means are summarized in Table 1 and significance was marked by an asterisk. There is a positive relationship between chromium and chromium-6 ($r = 0.984$, $p < 0.01$). The most noteworthy positive correlations are between vanadium and chromium-6 ($r = 0.445$, $p < 0.01$) and between vanadium and chromium ($r = 0.448$, $p < 0.01$).

Four of the six Pearson correlations with cobalt are negative: as cobalt increases, chromium ($r = -0.017$, $p < 0.01$), strontium ($r = -0.024$, $p < 0.01$), chromium-6 ($r = -0.021$, $p < 0.01$) and vanadium ($r = -0.046$, $p < 0.01$) decrease slightly.

Scatterplots of the relationships between each pair of geometric mean of contaminants is summarized in Figure 1. The strongest linear relationship visualized in Figure 1 is between

chromium and chromium-6. This leads to the conclusion that high levels of chromium-6 would indicate high levels of chromium in the same community water system being tested.

3.2. Non-carcinogenic and Carcinogenic Health Risk Calculations

The adult hazard indexes for all states were summarized in Figure 2. Extreme outlier values were removed from the graphs for simpler visualization which resulted in one result deleted from the cobalt graph. Results > 1 are considered indicators of potential non-carcinogenic health effects. Using the highest contaminant values in tested community water systems returns hazard indexes mostly < 1, which means the samples obtained during the UCMR testing period do not show hazardous levels of contaminants except for cobalt. Chromium does not have toxicity data associated with it in the RSL calculator, so health risk cannot be assessed with this tool.

Chromium-six is the only UCMR 3 heavy metal that has information for carcinogenic effects. These results are summarized in Figure 3. Acceptable risk is considered one-in-a-million (10^{-6}) (Fewtrell & Bartram, 2001). Most of the results are closer to what is considered acceptable risk. However, the highest carcinogenic risk was recorded at 1 in 1,000 (10^{-3}).

3.5. Toxicogenomic Pathway Investigation

CTD showed that the correlation between chromium and vanadium also has mutual genes that are affected with the two chemicals. There are twenty-four genes affected that overlap between the two contaminants (Table 3). These overlapping genes affect several pathways according to DAVID and the top pathway that DAVID returned was the Hypoxia-Inducible Factor (HIF)-1 Signaling Pathway (p-value = 8.2E-9) which is a crucial process that regulates

homeostasis in response to hypoxia (Ziello, Jovin, & Huang, 2007). Figure 4 represents the HIF-1 pathway with the affected genes signified by red stars.

3.6. Mapping of Co-Occurring Contaminants

Mapping the location of vanadium and chromium in the top 5 cm of soil show very similar distributions across the United States and explains the significant correlation seen in the analysis (Figure 5) ("Geochemical and mineralogical maps for soils of the conterminous United States - Data.gov," 2019).

4. DISCUSSION

This paper presents an innovative method of assessing emerging heavy metals in the United State community water systems. The extensive datasets available from the EPA UCMR have been largely unused in risk assessments so this paper attempted to use the data in a new way to understand potential human health risk that results from chronic exposure to the unregulated contaminants in United States' drinking water.

First, geometric means of contaminants were calculated for each community water system and they were graphed using scatterplots of each pairwise correlation. This allowed the relationship of each contaminant to be assessed by pairing it with each other contaminant. Chromium and chromium-six had a strong correlation from the samples collected. There was a strong significant relationship between vanadium and chromium and vanadium and chromium-6. Previous studies show overlapping health effects between vanadium and chromium. One paper by Scibior and Zaporowska (2007) described vanadium increasing kidney and liver weight in rats. When these rats were dosed with vanadium and chromium (III) together in their drinking

water, there were significantly higher levels of iron and zinc in kidneys and liver which leads to substantial decrease in renal L-ascorbic acid concentration, which is a regulator of renal cell growth and metabolism. Another interesting finding described in previous studies is how vanadium affects hemostasis, which is the physiological process in a body that stops bleeding.

The studies show common health effects from chromium and vanadium; however, there were no identified studies that studied the joint toxicology effects from these two metals. In order to further investigate the potential for interaction, the genes that interact with either chemical were obtained from the Comparative Toxicogenomic Database and input in DAVID to determine genetic interactions. Because vanadium and chromium both have effects on the HIF-1 pathway, this can potentially lead to disruption of homeostasis and result in cascading health effects such as cardiometabolic disorders (Ziello et al., 2007) and ischemic stroke (Skalny et al., 2017). This interaction on the HIF-1 pathway demonstrates the need to examine joint exposure to contaminants in drinking water.

Second, the ranges of values obtained during the UCMR sampling process were used to assess human health risk. By using the EPA RSL calculator, the results were translated from sample values to hazard indexes which presented a visualization of the potential human health risk that exposure to the tap-water in the tested community water systems represent. This method shows that the samples obtained during the UCMR 3 monitoring do not represent substantial non-carcinogenic health risks according to the EPA RSL calculations. However, this method does show that there are chromium-6 values in the tested public water systems that represent potentially concerning levels if chromium-6 is eventually classified as a carcinogen via oral exposure when ingested via tap-water.

4.1 Limitations

While the included datasets are beneficial, there are limitations for this project. Using the CTD is useful; however, it relies on studies that use cell lines and animal models (Davis et. al, 2018). Using these studies is not always generalizable to the human population due to differences in dosing. Also, the fundamental human biological system is different than animal models (Sasso & Schlosser, 2015). In the CTD, hexavalent chromium interactions with GJA1 is based on evidence that comes from one rat cell line. This gives very limited information for a chemical-gene interaction (Carette et al., 2013). It is difficult to make the connection between this cell line and a human system because the dose given in the referenced study is different than the target organ dose in humans. Also, in the case of hexavalent chromium, systemic biology in humans leads to reduction of chromium into the beneficial form of chromium (III) (Sasso & Schlosser, 2015). Hexavalent chromium data is largely uncertain because most studies are based on animal models. The RSL calculator user's guide does acknowledge the uncertainty of risk calculations based on the CalEPA toxicity data ("Regional Screening Levels", 2015). While the CTD is based on studies that are statistically significant and increases specificity, the evidence included from studies that do not address human biology can decrease specificity.

Another limitation for the CTD and the RSL calculator is that gene-environment interactions being studied are often determined by the Matthew Effect (Grandjean et. al, 2011). For example, certain chemicals, such as chromium, are going to be studied more frequently and more in-depth than other chemicals. This effect is also reflected in how laboratories receive funding. Larger labs often receive more of the funding that they often use to study the chemicals that correspond with their focus (Cristea & Ioannidis, 2018). This results in certain research topics receiving more time and funding and the creation of specific communities that address

topics such as HIV, cancer and cardio-metabolic disease. So, while these are critical issues for public health, they are studied more largely due to the greater resources given to these areas.

Finally, the CTD only uses papers that were considered statistically significant. Using this as a criterium is beneficial; however, this can result in potential selection biases. For example, a p-value might be classified as non-significant due to difference in study sample size (Cristea & Ioannidis, 2018). An analysis tool like DAVID has potential bias in the algorithm that it uses to test for gene interaction in the pathways (Timmons, Szkop, & Gallagher, 2015).

5. CONCLUSIONS

This study investigates the potential health implications that the unregulated heavy metals represent in public water systems. Using Tobit regression (Gong & Schaubel, 2018) allowed assessment of chemicals with large portions of the recorded values under the MRL. The EPA RSL calculator was efficient in calculating the potential risks of non-carcinogenic and carcinogenic effects while using the most up-to-date toxicity information from many credible sources. Results showed that chromium and chromium-6 are frequently present in the same samples and showed frequent co-occurrence of vanadium and chromium. The vanadium and chromium relationship has implications for more complex health effects on hemostasis as past research shows an interaction between the two contaminants (Filler et al., 2017). The calculated health risks show potential for negative health effects to occur in select public water systems for most of the heavy metals sampled for in the UCMR 3. This study shows a need for further research in order to better understand the risk represented by these levels of contaminants sampled.

5.1 Recommendations for Future Studies

It is recommended that the methods presented in this paper are replicated with the data from the first two UCMRs and any future assessments conducted by the EPA. It would also be beneficial to combine an assessment of unregulated contaminants in drinking water with data from the United States Geologic Survey (USGS) to gain another perspective on the state of water in the United States. Based on the findings of the toxicogenomic interactions, it is highly recommended to conduct follow up studies on the effect of these chemicals on genetic pathways to fully understand the potential health risk they pose.

REFERENCES

Almada, A. A., Golden, C. D., Osofsky, S. A., & Myers, S. S. (2017). A case for Planetary Health/GeoHealth. *Reviews of Geophysics*, 75–78. https://doi.org/10.1002/2017GH000084@10.1002/(ISSN)1944-9208.CGH1

Amanambu, A. C., & Egbinola, C. N. (n.d.). Geogenic contamination of groundwater in shallow aquifers in Ibadan, south-west Nigeria. *Management of Environmental Quality: An International Journal*, *26*, 327–341.

Bartell, S. M. (2019). Understanding and Mitigating the Replication Crisis, for Environmental Epidemiologists. *Current Environmental Health Reports*, *6*(1), 8–15. https://doi.org/10.1007/s40572-019-0225-4

Bundschuh, J., Maity, J. P., Mushtaq, S., Vithanage, M., Seneweera, S., Schneider, J., … Chen, C.-Y. (2017). Medical geology in the framework of the sustainable development goals. *The Science of the Total Environment*, *581–582*, 87–104. https://doi.org/10.1016/j.scitotenv.2016.11.208

Carette, D., Perrard, M.-H., Prisant, N., Gilleron, J., Pointis, G., Segretain, D., & Durand, P. (2013). Hexavalent chromium at low concentration alters Sertoli cell barrier and connexin 43 gap junction but not claudin-11 and N-cadherin in the rat seminiferous tubule culture model. *Toxicology and Applied Pharmacology*, *268*(1), 27–36. https://doi.org/10.1016/j.taap.2013.01.016

Cristea, I. A., & Ioannidis, J. P. A. (2018). P values in display items are ubiquitous and almost invariably significant: A survey of top science journals. *PLOS ONE*, *13*(5), e0197440. https://doi.org/10.1371/journal.pone.0197440

Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegers J, Wiegers TC, Mattingly CJ The Comparative Toxicogenomics Database: update 2019. Nucleic Acids Res. 2018 Sep 24.

de Winter, J. C. F., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods*, *21*(3), 273–290. https://doi.org/10.1037/met0000079

Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association*, *56*(293), 52–64. https://doi.org/10.1080/01621459.1961.10482090

Fewtrell, L., & Bartram, J. (Eds.). (2001). *Water quality: guidelines, standards, and health: assessment of risk and risk management for water-related infectious disease*. Geneva: World Health Organization.

Filler, G., Belostotsky, V., Kobrzynski, M., Huang, S.-H. S., & Yang, L. (2017). High prevalence of elevated molybdenum levels in pediatric CKD patients. A cross-sectional and longitudinal study. *Clinical Nephrology*, *88*(8), 79–85. https://doi.org/10.5414/CN109015

Filler, G., Kobrzynski, M., Sidhu, H. K., Belostotsky, V., Huang, S.-H. S., McIntyre, C., & Yang, L. (2017). A cross-sectional study measuring vanadium and chromium levels in paediatric patients with CKD. *BMJ Open*, *7*(5), e014821. https://doi.org/10.1136/bmjopen-2016-014821

Geochemical and mineralogical maps for soils of the conterminous United States - Data.gov. (n.d.). Retrieved April 15, 2019, from https://catalog.data.gov/dataset/geochemical-and-mineralogical-maps-for-soils-of-the-conterminous-united-states

Ghaderpoori, M., Kamarehie, B., Jafari, A., Ghaderpoury, A., & Karami, M. (2018). Heavy metals analysis and quality assessment in drinking water - Khorramabad city, Iran. *Data in Brief*, *16*, 685–692. https://doi.org/10.1016/j.dib.2017.11.078

Glassmeyer, S. T., Furlong, E. T., Kolpin, D. W., Batt, A. L., Benson, R., Boone, J. S., … Wilson, V. S. (2017). Nationwide reconnaissance of contaminants of emerging concern in source and treated drinking waters of the United States. *The Science of the Total Environment*, *581–582*, 909–922. https://doi.org/10.1016/j.scitotenv.2016.12.004

Gong, Q., & Schaubel, D. E. (2018). Tobit regression for modeling mean survival time using data subject to multiple sources of censoring. *Pharmaceutical Statistics*, *17*(2), 117–125. https://doi.org/10.1002/pst.1844

Grandjean, P., Eriksen, M. L., Ellegaard, O., & Wallin, J. A. (2011). The Matthew effect in environmental science publication: a bibliometric analysis of chemical substances in journal articles. *Environmental Health: A Global Access Science Source*, *10*, 96. https://doi.org/10.1186/1476-069X-10-96

Guidotti, T. L. (2009). Emerging contaminants in drinking water: what to do? *Archives of Environmental & Occupational Health*, *64*(2), 91–92. https://doi.org/10.3200/AEOH.64.2.91-92

Hypoxia-Inducible Factor (HIF)-1 Regulatory Pathway and its Potential for Therapeutic Intervention in Malignancy and Ischemia. (n.d.). Retrieved April 15, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2140184/

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Kolpin, D. W., Glassmeyer, S. T., & Furlong, E. T. (2017). An introduction to joint research by the USEPA and USGS on contaminants of emerging concern in source and treated drinking waters of the United States. *The Science of the Total Environment*, *579*, 1608–1609. https://doi.org/10.1016/j.scitotenv.2016.03.052

Krzywinski, M., & Altman, N. (2014). Points of Significance: Visualizing samples with box plots. *Nature Methods*, *11*, 119–120. https://doi.org/10.1038/nmeth.2813

Lei, M., Zhang, L., Lei, J., Zong, L., Li, J., Wu, Z., & Wang, Z. (2015). Overview of Emerging Contaminants and Associated Human Health Effects. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/404796

Lentini, P., Zanoli, L., Granata, A., Signorelli, S. S., Castellino, P., & Dell'Aquila, R. (2017). Kidney and heavy metals - The role of environmental exposure (Review). *Molecular Medicine Reports*, *15*(5), 3413–3419. https://doi.org/10.3892/mmr.2017.6389

Lu, T. (2018). Mixed-effects location and scale Tobit joint models for heterogeneous longitudinal data with skewness, detection limits, and measurement errors. *Statistical Methods in Medical Research*, *27*(12), 3525–3543. https://doi.org/10.1177/0962280217704225

Mitrov, D., Hadzi-Petrushev, N., Stojkovski, V., Gjorgievska, E., Gagov, H., & Mladenov, M. (2014). Influence of chronic chromium exposition on the processes of lipid peroxidation inflammation and platelet activation in rats. *Journal of Biological Regulators and Homeostatic Agents*, *28*(3), 531–535.

Multiple sources of bias confound functional enrichment analysis of global -omics data | Genome Biology | Full Text. (n.d.). Retrieved April 15, 2019, from https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0761-7

Olin, K. L., Cherr, G. N., Rifkin, E., & Keen, C. L. (1996). The effects of some redox-active metals and reactive aldehydes on DNA-protein cross-links in vitro. *Toxicology*, *110*(1–3), 1–8.

Passlack, N., Mainzer, B., Lahrssen-Wiederholt, M., Schafft, H., Palavinskas, R., Breithaupt, A., & Zentek, J. (2015). Concentrations of strontium, barium, cadmium, copper, zinc, manganese, chromium, antimony, selenium, and lead in the liver and kidneys of dogs according to age, gender, and the occurrence of chronic kidney disease. *Journal of Veterinary Science*, *16*(1), 57–66.

Regional Screening Levels (RSLs) - User's Guide | Risk Assessment | US EPA. (n.d.). Retrieved February 8, 2019, from https://www.epa.gov/risk/regional-screening-levels-rsls-users-guide

Report on NIH Funding vs. Global Burden of Disease -NIH Research Portfolio Online Reporting Tools (RePORT). (n.d.). Retrieved April 15, 2019, from https://report.nih.gov/info_disease_burden.aspx

Sasso, A. F., & Schlosser, P. M. (2015). An evaluation of in vivo models for toxicokinetics of hexavalent chromium in the stomach. Toxicology and Applied Pharmacology, 287(3), 293–298. https://doi.org/10.1016/j.taap.2015.06.016

Scibior, A., & Zaporowska, H. (2007). Effects of vanadium(V) and/or chromium(III) on L-ascorbic acid and glutathione as well as iron, zinc, and copper levels in rat liver and kidney. *Journal of Toxicology and Environmental Health. Part A*, *70*(8), 696–704. https://doi.org/10.1080/15287390601187906

Scibior, A., Zaporowska, H., Ostrowski, J., & Banach, A. (2006). Combined effect of vanadium(V) and chromium(III) on lipid peroxidation in liver and kidney of rats. *Chemico-Biological Interactions*, *159*(3), 213–222. https://doi.org/10.1016/j.cbi.2005.11.008

Scibior, A., Zaporowska, H., Wolińska, A., & Ostrowski, J. (2010). Antioxidant enzyme activity and lipid peroxidation in the blood of rats co-treated with vanadium (V(+5)) and chromium (Cr (+3)). *Cell Biology and Toxicology*, *26*(6), 509–526. https://doi.org/10.1007/s10565-010-9160-8

Simic, M. (2017a). *The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Data Summary, January 2017*. 12.

Simic, M. (2017b). *The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Data Summary, January 2017*. 12.

Skalny, A. V., Klimenko, L. L., Turna, A. A., Budanova, M. N., Baskakov, I. S., Savostina, M. S., … Tinkov, A. A. (2017). Serum trace elements are associated with hemostasis, lipid spectrum and inflammatory markers in men suffering from acute ischemic stroke. *Metabolic Brain Disease*, *32*(3), 779–788. https://doi.org/10.1007/s11011-017-9967-6

Spatial analysis and health risk assessment of heavy metals concentration in drinking water resources | SpringerLink. (n.d.). Retrieved February 8, 2019, from https://link-springer-com.proxy.library.emory.edu/article/10.1007/s11356-017-0102-3

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

The Fourth Unregulated Contaminant Monitoring Rule (UCMR 4): Data Summary, October 2018. (n.d.). 14.

The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Fact Sheet for Assessment Monitoring (List 1 Contaminants). (n.d.-a). 4.

The Third Unregulated Contaminant Monitoring Rule (UCMR 3): Fact Sheet for Assessment Monitoring (List 1 Contaminants). (n.d.-b). 4.

The Third Unregulated Contaminant Monitoring Rule (UCMR 3); Searching for Emerging Contaminants in Drinking Water. (2015). 2.

Timmons, J. A., Szkop, K. J., & Gallagher, I. J. (2015). Multiple sources of bias confound functional enrichment analysis of global -omics data. Genome Biology, 16(1), 186. https://doi.org/10.1186/s13059-015-0761-7

Tobit Analysis | Stata Data Analysis Examples. (n.d.). Retrieved February 14, 2019, from https://stats.idre.ucla.edu/stata/dae/tobit-analysis/

Toccalino, P. L., & Norman, J. E. (2006). Health-based screening levels to evaluate U.S. Geological Survey ground water quality data. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, *26*(5), 1339–1348. https://doi.org/10.1111/j.1539-6924.2006.00805.x

Toxicological risk assessment and prioritization of drinking water relevant contaminants of emerging concern - ScienceDirect. (n.d.). Retrieved February 8, 2019, from https://www-sciencedirect-com.proxy.library.emory.edu/science/article/pii/S0160412018302721

Tsai, C.-C., Wu, C.-L., Kor, C.-T., Lian, I.-B., Chang, C.-H., Chang, T.-H., … Chiu, P.-F. (2018). Prospective associations between environmental heavy metal exposure and renal outcomes in adults with chronic kidney disease. *Nephrology (Carlton, Vic.)*, *23*(9), 830–836. https://doi.org/10.1111/nep.13089

US EPA, ORD. (2015, September 3). Regional Screening Levels (RSLs) - Equations [Data and Tools]. Retrieved April 1, 2019, from US EPA website: https://www.epa.gov/risk/regional-screening-levels-rsls-equations

US EPA, ORD. (n.d.). Chromium(VI) CASRN 18540-29-9 | IRIS | US EPA, ORD. Retrieved April 3, 2019, from https://cfpub.epa.gov/ncea/iris2/chemicalLanding.cfm?substance_nmbr=144

US EPA, OW. (2015a, September 1). Occurrence Data for the Unregulated Contaminant Monitoring Rule [Data and Tools]. Retrieved February 14, 2019, from US EPA website: https://www.epa.gov/dwucmr/occurrence-data-unregulated-contaminant-monitoring-rule

US EPA, OW. (2015b, September 1). Third Unregulated Contaminant Monitoring Rule [Data and Tools]. Retrieved February 14, 2019, from US EPA website: https://www.epa.gov/dwucmr/third-unregulated-contaminant-monitoring-rule

Wahls, W. P. (2018). The NIH must reduce disparities in funding to maximize its return on investments from taxpayers. *ELife*, *7*. https://doi.org/10.7554/eLife.34965

Table 1

Pairwise correlations of geometric means of community water systems by contaminant with p-value for the correlation on the second line

Variables	Cr	St	Cr (VI)	Mo	Co	V
Cr	1.000					
St	0.036* <0.01	1.000				
Cr (VI)	0.984* <0.01	0.032* <0.01	1.000			
Mo	0.008 0.10	0.097* <0.01	0.010 0.05	1.000		
Co	-0.017* <0.01	-0.024* <0.01	-0.021* <0.01	-0.003 0.53	1.000	
V	0.448* <0.01	0.011 0.02	0.445* <0.01	0.045* <0.01	-0.046* <0.01	1.000

* shows significance at the .0033 level

Table 2

Tobit regression results for Chromium and Vanadium models

logva	Coef.	logcr	Coef.
Chromium Quartile 1	0.35	*Vanadium*	0.35
Quartile 2	1.31		1.12
Quartile 3	2.17		2.13
Constant	-0.03		-0.95
PWSName			
Var(_cons)	0.79		0.44
PWSName>State			
var (_cons)	0.25		0.18
var (e.logva)	0.19		0.21

Table 3

Genes that interact with both Vanadium and Chromium

GENE	GENE NAME	POSITION
AHR	Aryl Hydrocarbon Receptor	Chr 7: 17.30-17.35
AKT1	AKT Serine/Threonine Kinase 1	Chr 14: 104.77-104.8
ATP2B1	ATPase Plasma Membrane Ca2+ Transporting 1	Chr 12: 89.59-89.71
CXCL8	C-X-C Motif Chemokine Ligand 8	Chr 4: 73.74-73.74
CYP1A1	Cytochrome P450 Family 1 Subfamily A Member 1	Chr 15: 74.72 - 74.73
CYP1B1	Cytochrome P450 Family 1 Subfamily B Member 1	Chr 2: 38.07-38.11
GJA1	Gap Junction Protein Alpha 1	Chr 6: 121.44 - 121.45
HIF1A	Hypoxia Inducible Factor 1 Alpha Subunit	Chr 14: 61.7-61.75
IL10	Interleukin 10	Chr 1: 206.77 - 206.77
IL4	Interleukin 4	Chr 5: 132.67 - 132.68
IL6	Interleukin 6 leptin	Chr 7: 22.73-22.73
INS1	Insulin 1	Chr 11: 2.16-2.16
LEP	Leptin	Chr 7: 128.24-128.26
MAPK1	Mitogen-activated Protein Kinase 1	Chr 22:21.75 - 21.87
MAPK3	Mitogen-activated Protein Kinase 3	Chr 16: 30.11 - 30.12
MMP1	Matrix Metallopeptidase 1	Chr 11: 102.79 - 102.8
NOS3	Nitric Oxide Synthase 3	Chr 7: 150.99 - 151.01
NQO1	NAD(P)H Quinone Dehydrogenase 1	Chr 16: 69.71-69.73
RPS6KB1	Ribosomal Protein S6 Kinase B1	Chr 17: 59.89-59.95
SLC2A4	Solute Carrier Family 2 Member 4	Chr 17: 7.28-7.29
TIMP1	TIMP Metallopeptidase Inhibitor 1	Chr X: 47.58-47.59
TNF	Tumor Necrosis Factor	Chr 6: 31.58-31.58
TP53	Tumor Protein p53	Chr 17: 7.66-7.69
UCP1	Uncoupling Protein 1	Chr 4: 140.56-140.56

Fig. 1
Scatterplot matrix of predicted public water system-level geometric means of contaminants. Water system-level geometric means were obtained from mixed-effect Tobit regression models fitted to each contaminant separately, with two random intercepts accounting for nesting of water samples within water systems within states.

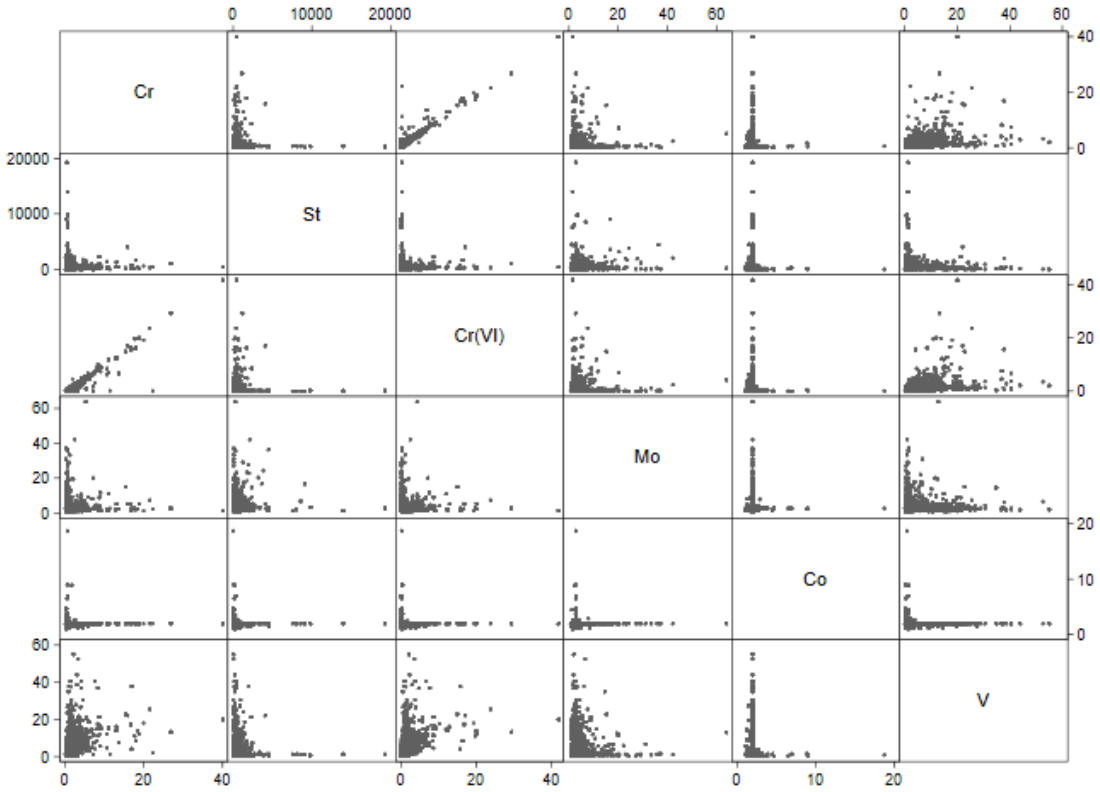

Fig. 2
Hazard Index (Adult) by Contaminant. Adult Hazard Indexes were obtained from the EPA Regional Screening Level calculator. Results were combined in an Excel file and boxplots were generated for each group of metals across all states (n=50).

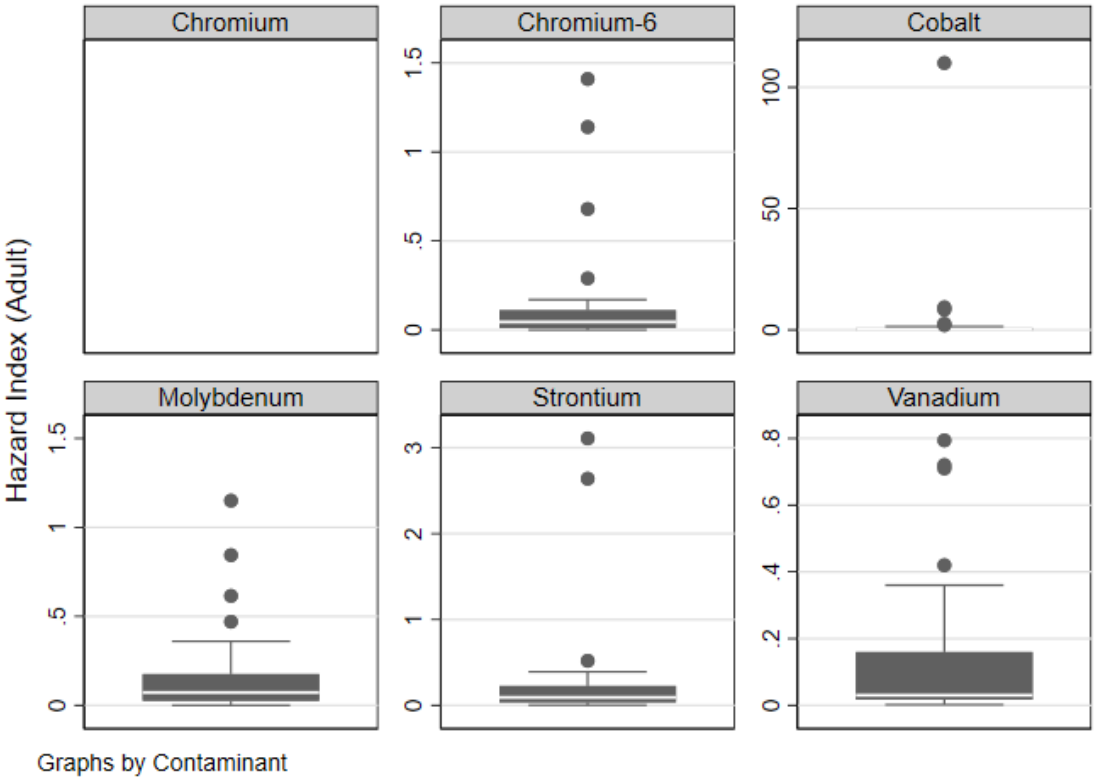

Fig. 3
Carcinogenic Risk for Chromium-6. Risk of carcinogenic effects were obtained from the EPA Regional Screening Level calculator. Results were combined in an Excel file and a boxplot was generated for the carcinogenic risk across all states (n=50).

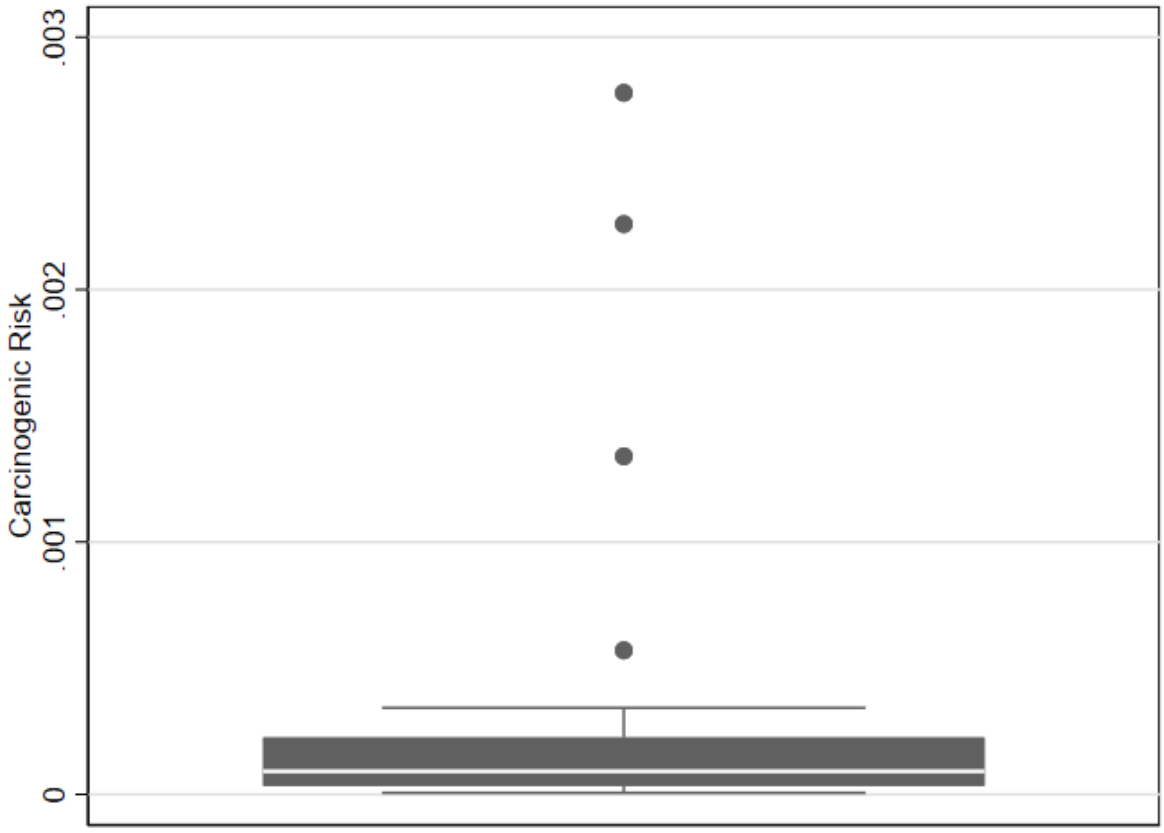

Fig. 4

Vanadium and Chromium Interaction Pathway. Overlapping genes affected by both vanadium and chromium were obtained from the Comparative Toxicogenomic Database (CTD) and input in the Database for Annotation, Visualization and Integrated Discovery (DAVID) which generated the HIF-1 pathway as affected by the overlapping genes.

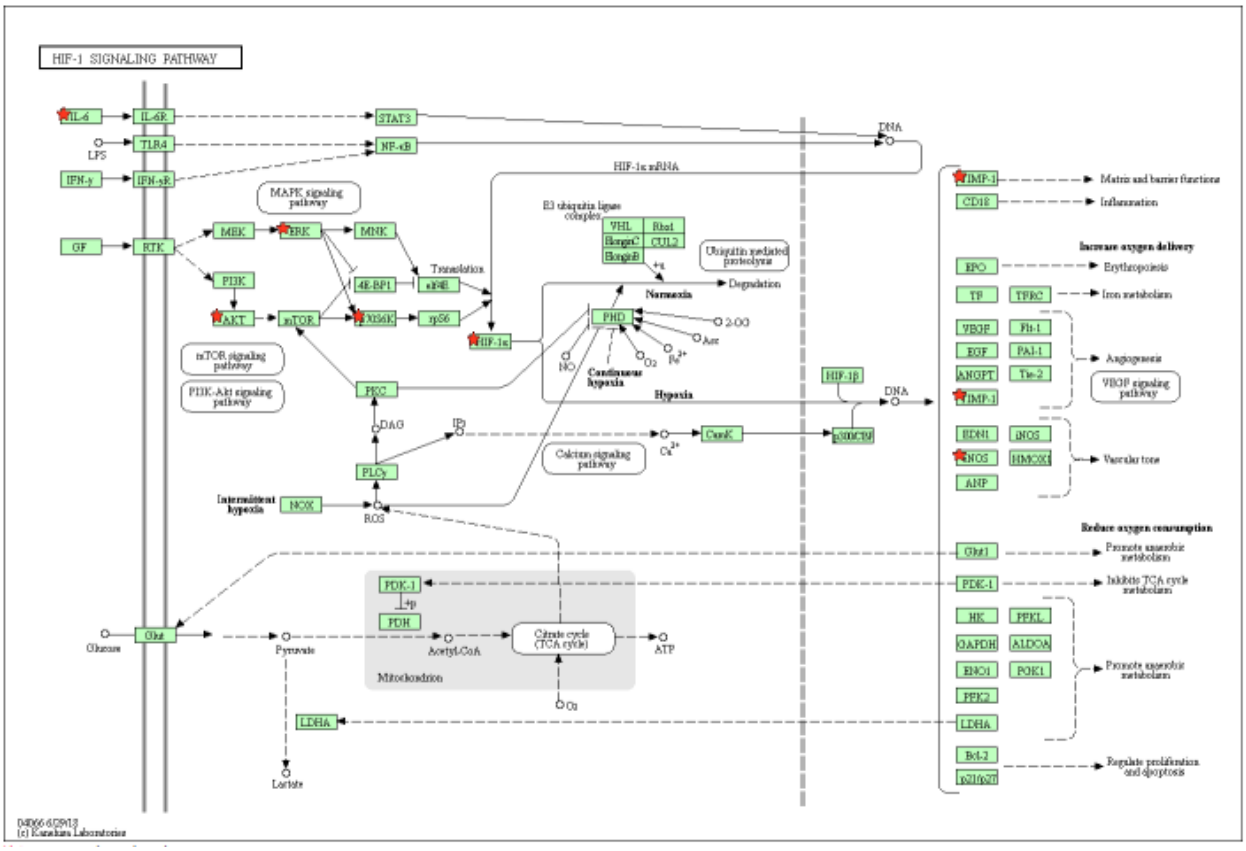

Fig. 5

Maps of Vanadium and Chromium in the Top 5 cm of Soil in the contiguous United States. These maps were publicly available from the United States Geological Survey ("Geochemical and mineralogical maps for soils of the conterminous United States - Data.gov," 2019.)