Supporting Information for:

(Non)targeted chemical analysis and risk assessment of organic contaminants in darkibor kale grown at rural and urban farms

Christopher L. Brueck[1,a], Xiaoyue Xin[1], Sara N. Lupolt[1,2,3], Brent F. Kim[1,3], Raychel E. Santo[1,3], Q. Lyu[1], Antony J. Williams[4], Keeve E. Nachman[1,2,3], Carsten Prasse[1,2]*

[1] Department of Environmental Health and Engineering, Johns Hopkins University, MD 21205, USA

[2] Risk Sciences and Public Policy Institute, Johns Hopkins University, MD 21205, USA

[3] Center for a Livable Future, Johns Hopkins University, MD 21202, USA

[4] Center for Computational Toxicology and Exposure, Office of Research and Development, U.S. Environmental Protection Agency, NC 27711, USA

[a] current address: Environmental & Earth Sciences Practice, Exponent, Inc., Bellevue, WA 98007, USA

*Corresponding author information: cprasse1@jhu.edu

This document contains 11 pages with 9 figures and 6 equations.

Table of Contents

1. Methods

1.1 Pesticide mix and quantification

For suspect screening and quantitative analysis, a large analytical chemical mixture of pesticides provided by the FDA was used (from hereon, pesticide mix). A complete list of the 1,069 compounds in the pesticide mix is provided in Table S1. The chemicals were stored in 15 different sub-mixtures of approximately ~70 compounds each in acetonitrile at 20 µg/mL. Prior to analysis, sub-mixtures of the standards were combined to make the complete mixture which was serially diluted to obtain 1/X weighted calibration standards for quantitative analysis. Eight calibrants were prepared at 0, 0.15, 0.3, 1.5, 3, 15, 30, and 51 ng/mL in matrix-matched solutions each with 20 ng/mL of the internal standard mixture. Limits of detection (LOD) and quantification (LOQ) for each compound were calculated using the linear regression approach[1], as described in Equations S1-S3. LODs and LOQs were only calculated if the calibration regression had an R^2 of greater than 0.9 and at least 3 sequential calibrants were detected.

1.2 Recovery test

For the recovery test, 500 µL of the pesticide mix (0.75 µg/mL), along with 150 µL of the internal standard solution, were spiked into triplicates of a randomly selected kale sample. This was followed by the addition of 6.85 mL of extraction solvent (to account for the additional volume added by the pesticide mix). All other aspects of the extraction were identical to those described for the non-spiked samples above. Recoveries were calculated as the determined amount in the extract divided by the amount spiked into the kale prior to extraction. Concentrations for compounds that had extraction recoveries between 90-110% were not adjusted.

2. Figures and Equations
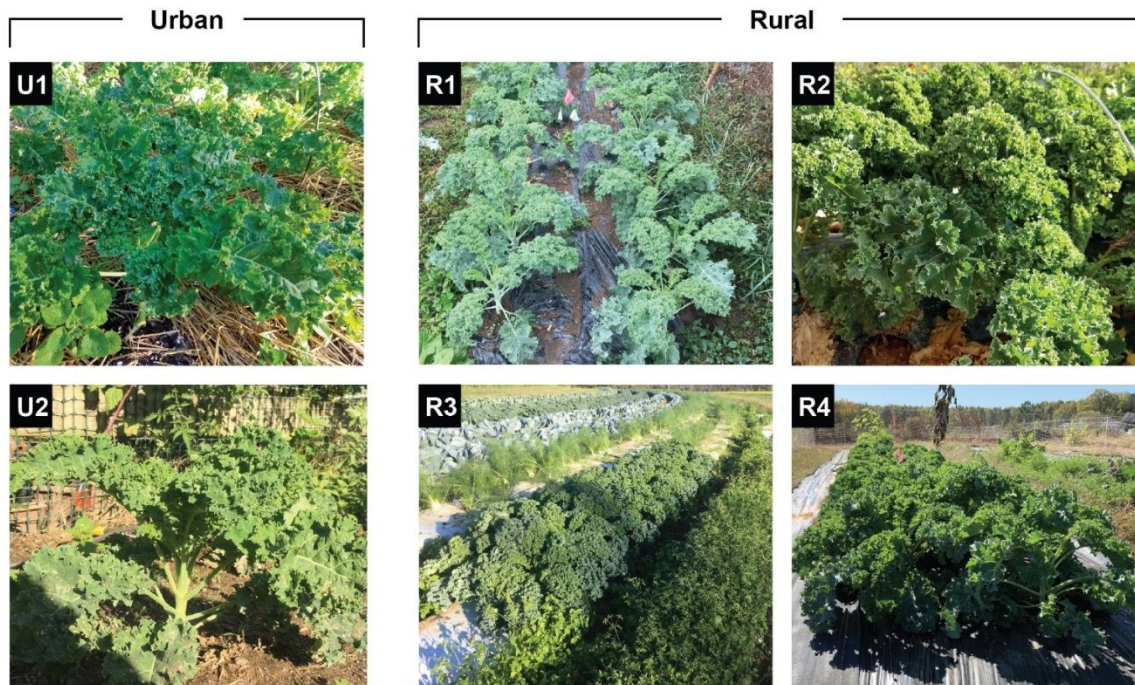
Figure S1. Photos of cultivated kale, by site.

Equation S1 Limit of detection

$$LOD = 3\sigma/s$$

Equation S2 Limit of quantification

$$LOQ = 10\sigma/s$$

Equation S3 Standard deviation of the weighted residuals

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2}{n - 2}}$$

s is the slope of the linear regression for each calibrant i, and w is the weight applied to each point. σ is the standard deviation, n is the sample size, y is the output of the linear regression, and $\hat{y}$ are the regression residuals.

Equation S4 Hazard Score

$$Hazard\ Score = \frac{\sum hazard\ values}{number\ of\ available\ sources}$$

45 Where *hazard values* are numeric scores assigned based on the integrated HCM ranking (i.e..,
46 VH, H, M, L) to indicate the level of hazard associated with a particular chemical, and the
47 *number of available sources* is the count of integrated rankings, excluded "N/A" and
48 "Inconclusive", among the total of selected endpoints.
49
50 **Equation S5** Quality Score

51 $$Quality\ Score = \frac{\sum authority\ values}{number\ of\ available\ sources}$$

52
53 Where *authority values* are numeric scores assigned to different hazard data sources (i.e.,
54 authoritative, screening, and QSAR model), reflecting the confidence level in the hazard data.
55
56 **Equation S6** Completeness Score

57 $$Completeness\ Score = \frac{number\ of\ available\ sources}{number\ of\ total\ sources}$$

58
59 Where *number of total sources* is total number of endpoints included in the framework (in this
60 case, comprises 15 human health relevant endpoints).
61
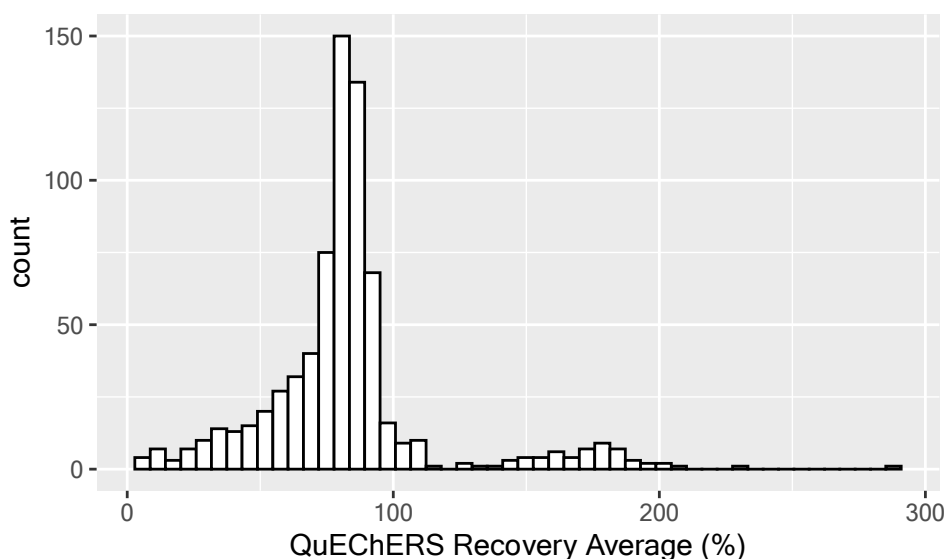
62
63 **Figure S2.** Histogram of QuEChERS recovery averages from 713 compounds from the pesticide mix for
64 compounds that were detected in at least 2 of the 3 experimental replicates.
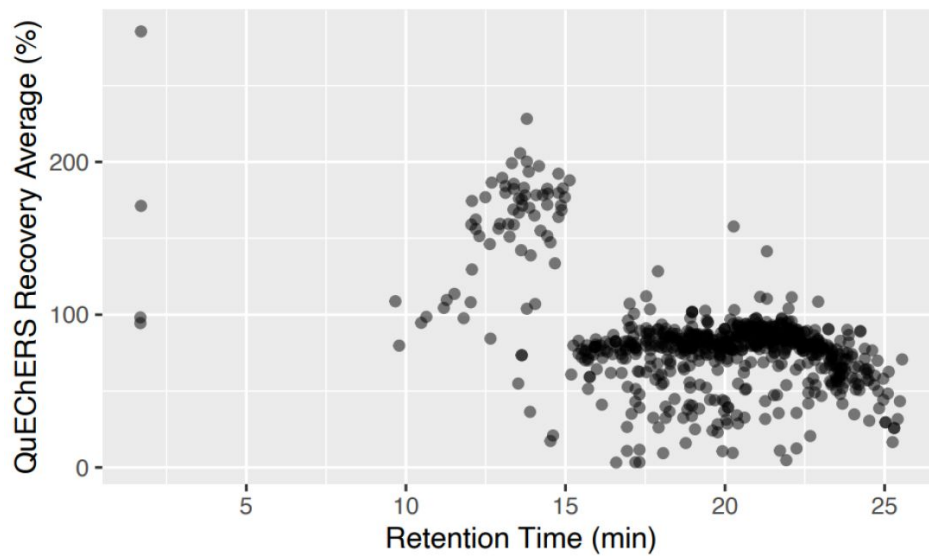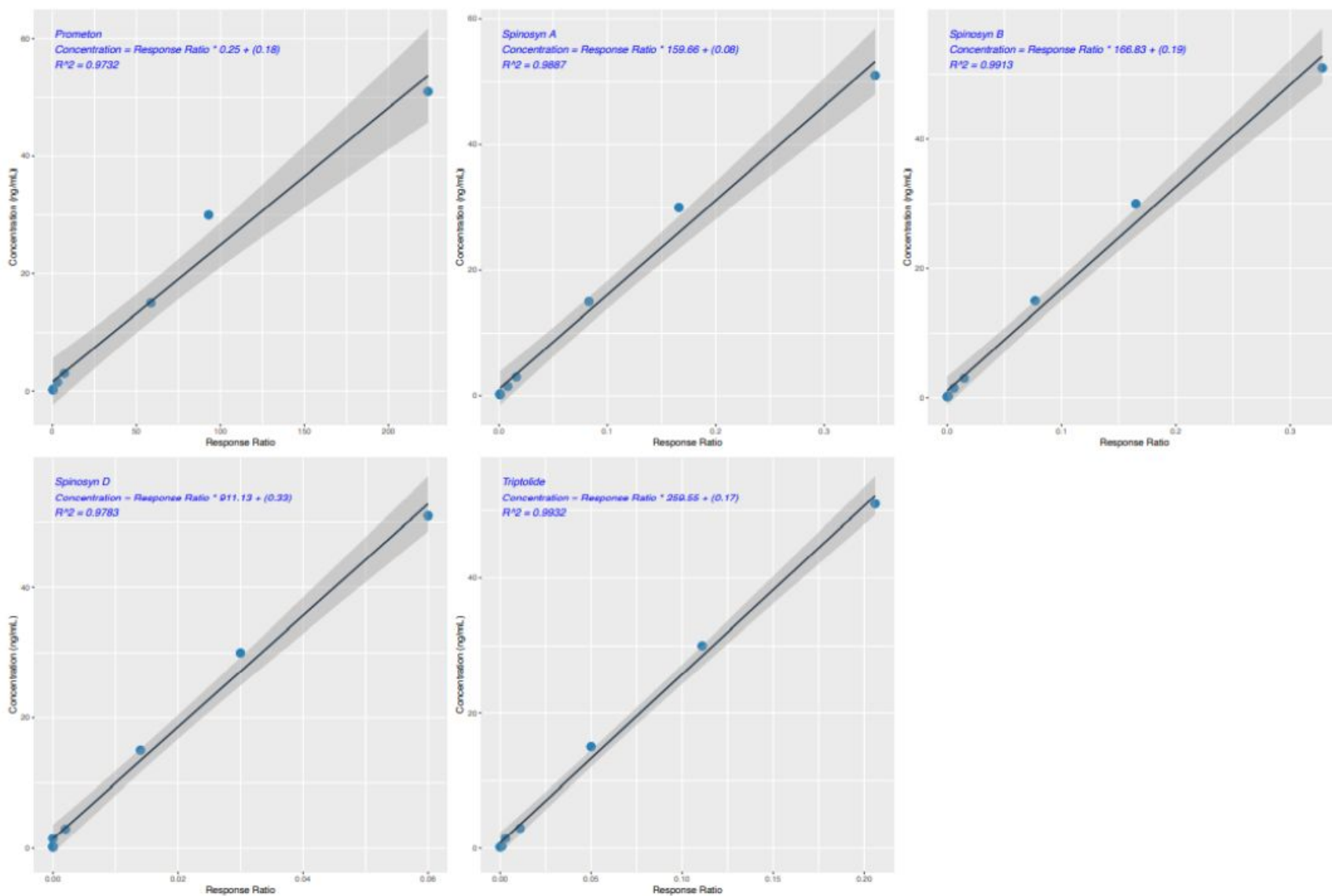
65
66 **Figure S3.** Scatter plot of chromatographic retention time versus QuEChERS recovery averages for 713
67 compounds.

68
69
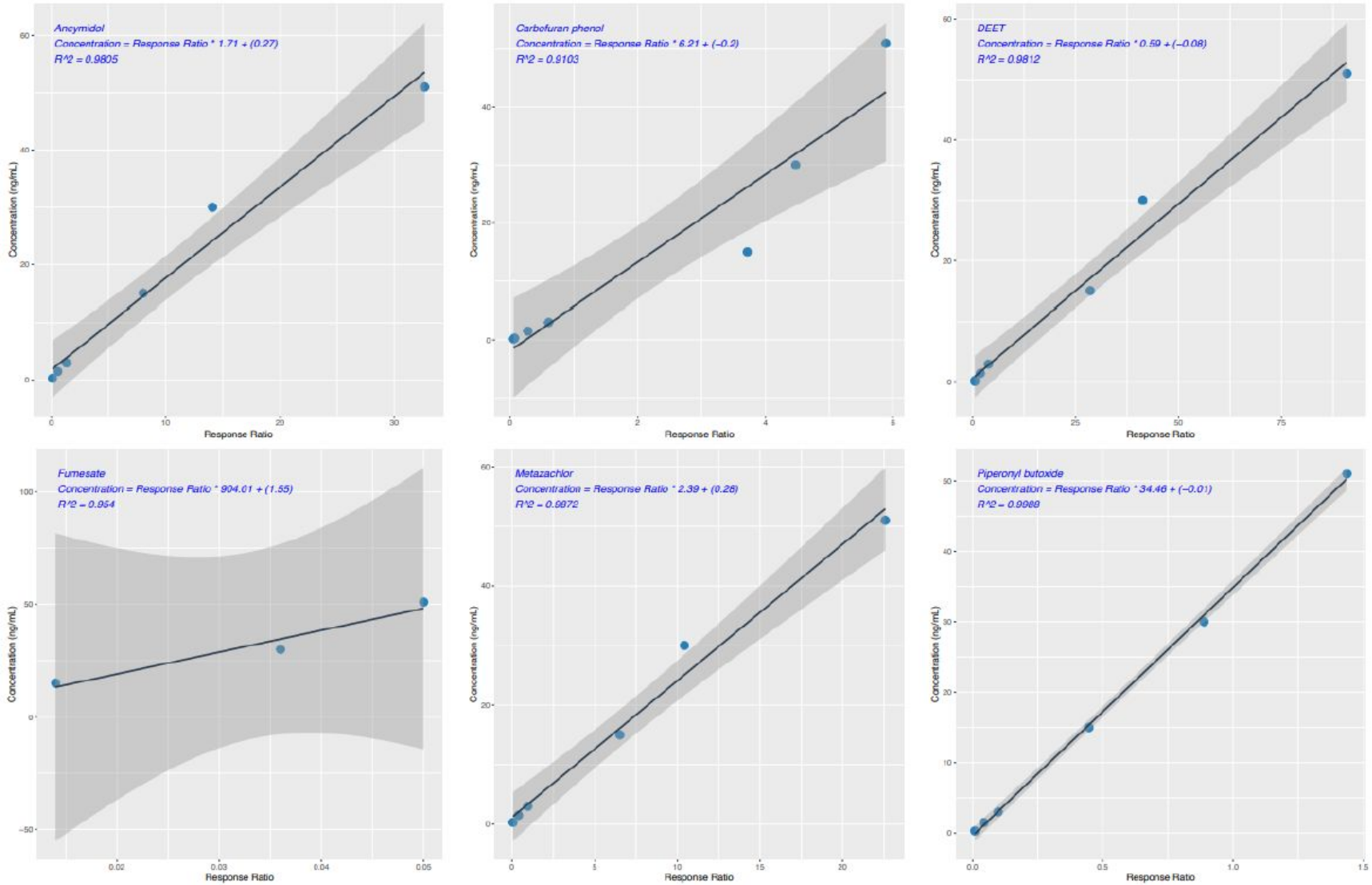
70
71 **Figure S4.** Calibration curves (linear regressions) for compounds quantified using the nDATA method.

72

73
74 **Figure S4 (continued).** Calibration curves (linear regressions) for compounds quantified using the nDATA method.
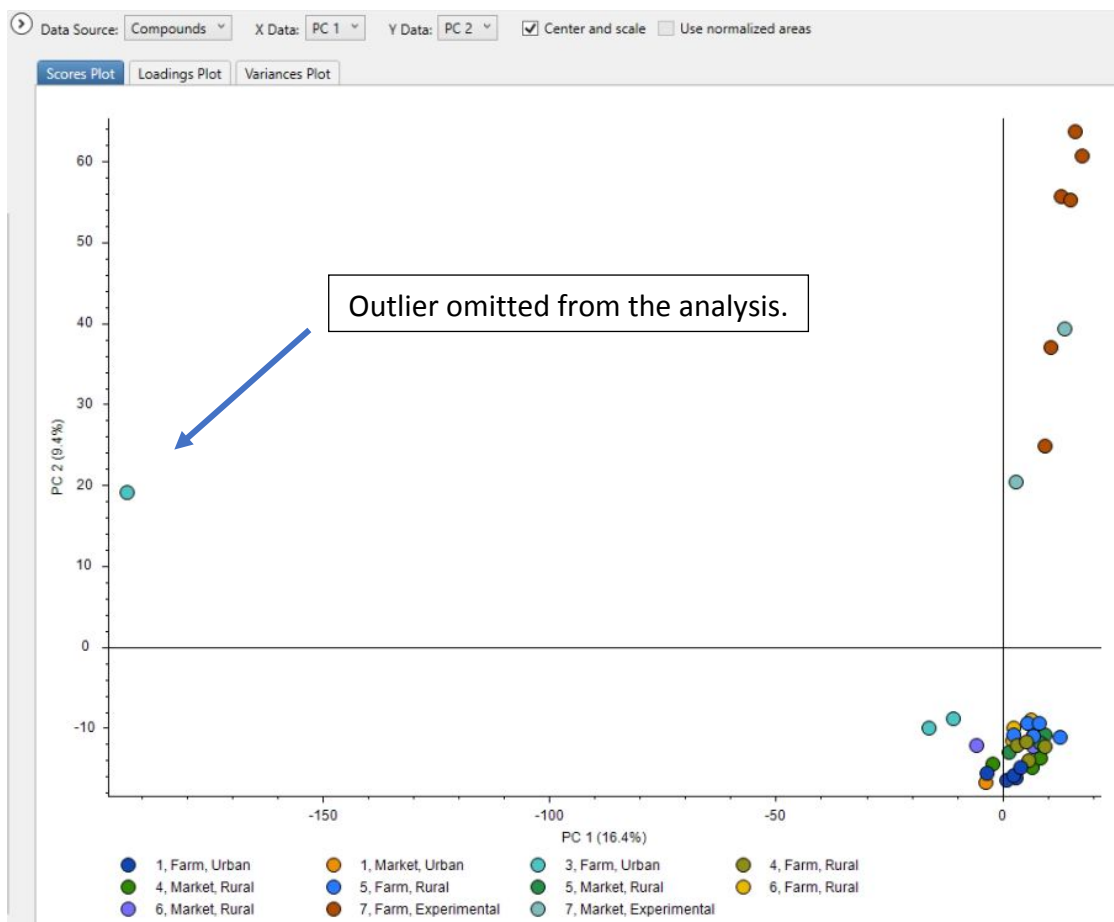
Figure S5. Initial principal component analysis suggesting that one of the samples analyzed using the nontargeted workflow was an outlier.
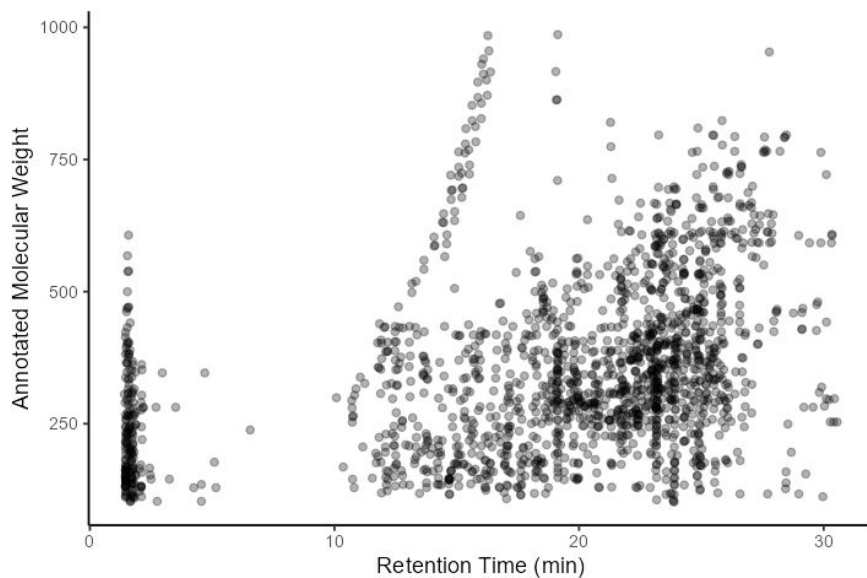

Figure S6. Annotated molecular weight from Compound Discoverer versus chromatography retention time for compounds detected in all kale samples
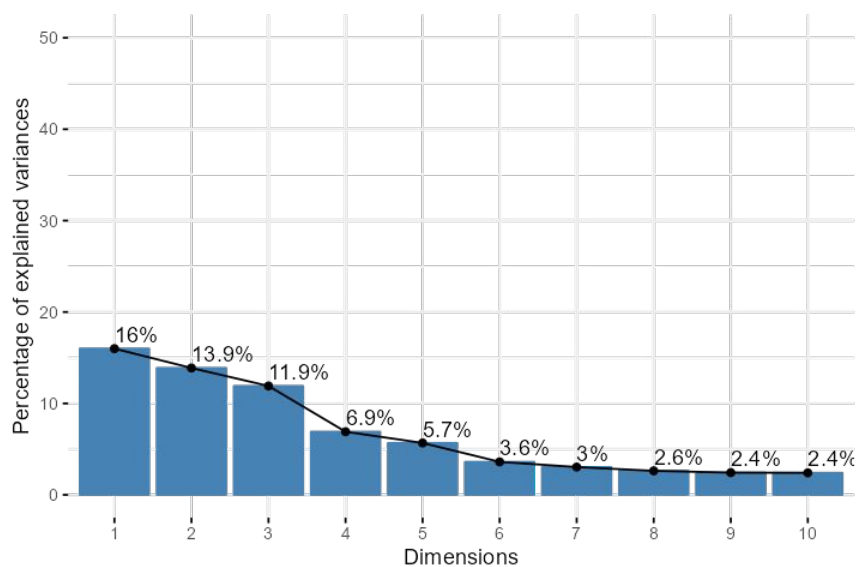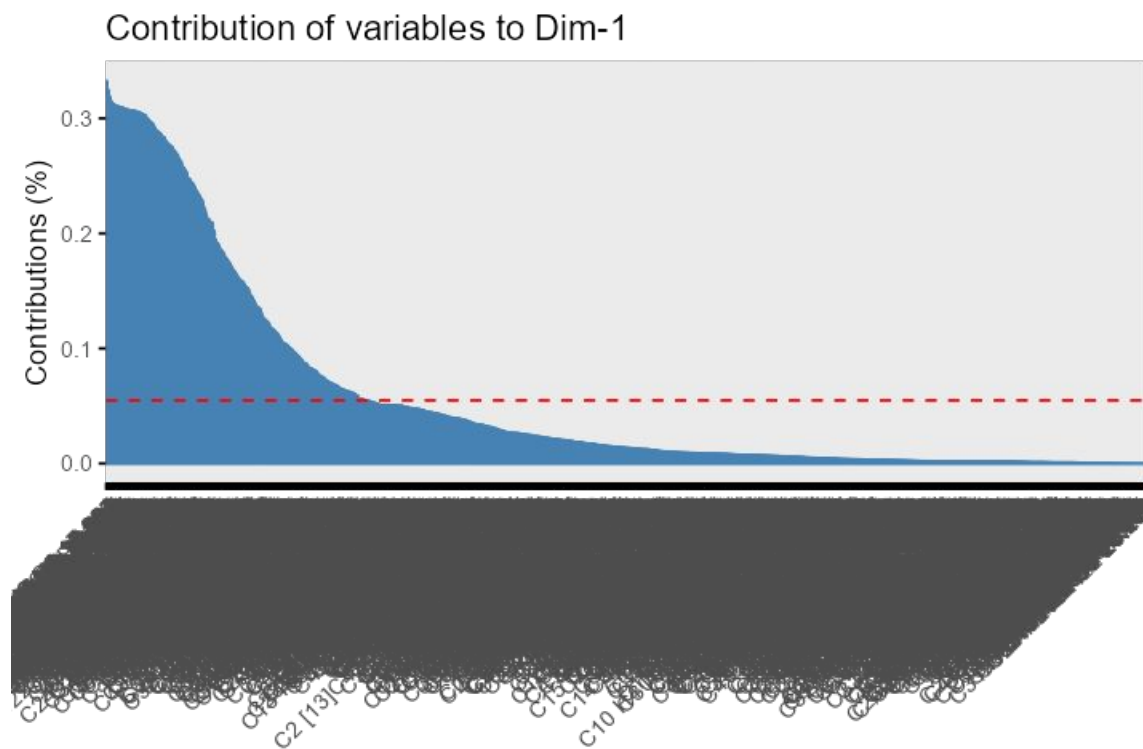
Figure S7. Scree plot.

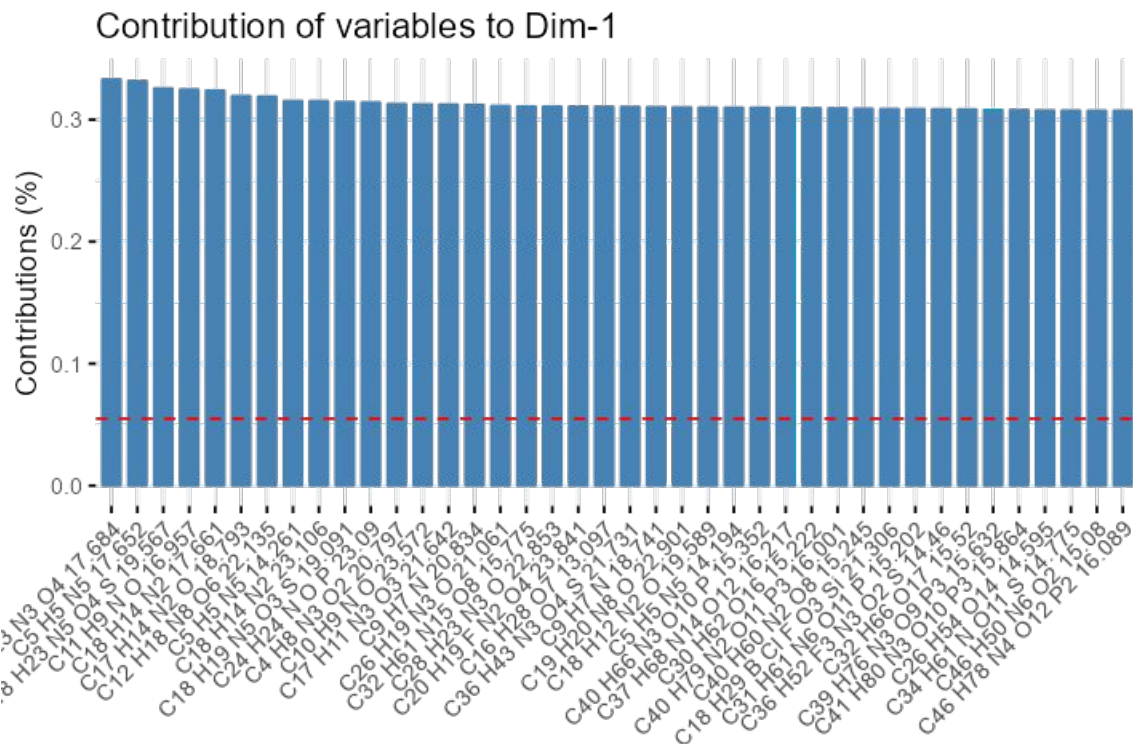**Figure S8.** Contribution of variables (all compounds).

Figure S9. Contribution of variables (top 40 compounds).

3. References

[1] J. Wang, W. Chow, J. Chang, and J. W. Wong, "Ultrahigh-Performance Liquid Chromatography Electrospray Ionization Q-Orbitrap Mass Spectrometry for the Analysis of 451 Pesticide Residues in Fruits and Vegetables: Method Development and Validation," *J. Agric. Food Chem.*, vol. 62, no. 42, pp. 10375–10391, Oct. 2014, doi: 10.1021/jf503778c.