Metabolomics-guided comparison of pollen and microalgae-based artificial diets in honey bees

Vincent A. Riciglianoᵃ, Kristof B. Cankᵇ, Daniel A. Toddᵇ, Sonja L. Knowlesᵇ, Nicholas H. Oberliesᵇ

Authors contributed equally to this study

ᵃVincent A. Ricigliano – Honey Bee Breeding, Genetics and Physiology Research, USDA-ARS, Baton Rouge, Louisiana 70820, United States; orcid.org/0000-0002-5167-5812; Email: vincent.ricigliano@usda.gov

ᵇDepartment of Chemistry and Biochemistry, University of North Carolina at Greensboro, Greensboro, NC 27402-6170, Fax: (336) 334-5402, Email address: nicholas Oberlies@uncg.edu
List of figures and tables

**Figure S1.** Schematic overview of honey bee feeding experiment.

**Figure S2.** Honey bee metabolite extraction workflow.

**Figure S3.** Process of feature filtering of raw data from LC-MS and GC-MS analysis.

**Figure S4.** Volcano plots of honey bees fed on four different diets using untargeted metabolomics through GC-MS analysis.

**Figure S5.** Extracted ion chromatograms (XIC) of zeaxanthin (C_40H_56O_2) in the standard and in the bees fed various diets using ESI in the positive mode.

**Figure S6** Extracted ion chromatograms (XIC) of lutein (C_40H_56O_2) in the standard and in the bees fed various diets using ESI in the positive mode.

**Figure S7.** Extracted ion chromatograms (XIC) of linolenic acid (C_{18}H_{30}O_2) in the standard and in the bees fed various diets using ESI in the positive mode.

**Figure S8.** Extracted ion chromatograms (XIC) of quinic acid (C_7H_{12}O_6) in the standard and in the bees fed various diets using ESI in the positive mode.

**Figure S9.** Extracted ion chromatograms (XIC) of α-tocopherol (C_{29}H_{50}O_2) in the standard and in the bees fed various diets using ESI in the positive mode.

**Figure S10.** Extracted ion chromatograms (XIC) of β-carotene in the bees fed various diets using ESI in the positive mode compared to the UV absorbance of β-carotene standard.

**Figure S11.** Extracted ion chromatograms (XIC) of linoleic acid (C_{18}H_{32}O_2) in the bees fed various diets using ESI in the positive mode.

**Figure S12.** Structures of all compounds that are identified through LC-MS standards and putatively identified through GC-MS data comparison to NIST database.

**Table S1.** Amount of extract per sample. Each sample contained 8 bee abdomens. Each diet had 4 biological replicates, resulting in the extraction 8 x 4 bee abdomens per diet condition.

**Table S2.** MZmine2.53 Parameters for LC-MS data Processing.

**Table S3.** MZmine2.53 Parameters for GC-MS data Processing.

**Table S4.** Primers used in this study.

**Table S5.** Identified features from LC-MS volcano plot analysis.

**Table S6.** Relative abundance and average peak areas of compounds identified through LC-MS analysis.

**Table S7.** Relative abundance comparison of metabolites identified through GC-MS library match.

**Table S8.** Benefits and limitations of the subtractive approach used in the study.
**Figure S1.** Schematic overview of the honey bee feeding experiment. Newly emerged worker bees were obtained by incubating sealed brood frames at 35 °C and 50% relative humidity overnight. Three brood frames were obtained from a healthy, vigorous colony that was treated for *Varroa* mites and had no visible signs of disease. Bees (< 24 h old) were collected into a container then randomly assigned to diet treatment groups consisting of sugar, pollen, *Chlorella*, or spirulina diets.
Figure S2. Metabolite extraction of bee abdomens. The samples were ground with mortar and pestles under liquid nitrogen. Then they were transferred to a scintillation vial and then submerged in acetone (5 mL). Samples were shaken for approximately 16 hrs. The acetone layer was transferred to microcentrifuge tubes and centrifuged. Then the acetone layer was removed and saved. The original samples were resuspended in 1:1 MeOH:CHCl₃ (5 mL) and sonicated for 30 min. The solvent was transferred to an Eppendorf and centrifuged. The supernatants from this step and the saved acetone layer were combined and dried under nitrogen to yield the bees fed on diet extract. This extraction procedure was also used to generate the pollen, sugar, and both algae extracts using 1 g of material in each case.
Figure S3. Process of feature filtering of raw data from LC-MS and GC MS analysis.
Figure S4. Volcano plots of honey bees fed on four different diets using untargeted metabolomics through GC-MS analysis.
Figure S5. Extracted ion chromatograms (XIC) of zeaxanthin (C₄₀H₅₆O₂) in the standard and in the bees fed various diets using ESI in the positive mode.
Figure S6. Extracted ion chromatograms (XIC) of lutein (C₄₀H₅₆O₂) in the standard and in the bees fed various diets using ESI in the positive mode. Since lutein and zeaxanthin have similar retention time (under these method conditions) and identical molecular formula, further analysis would be needed to distinguish between these two compounds.
**Figure S7.** Extracted ion chromatograms (XIC) of linolenic acid (C$_{18}$H$_{30}$O$_{2}$) in the standard and in the bees fed various diets using ESI in the positive mode.
Figure S8. Extracted ion chromatograms (XIC) of quinic acid ($C_7H_{12}O_6$) in the standard and in the bees fed various diets using ESI in the positive mode.
Figure S9. Extracted ion chromatograms (XIC) of α-tocopherol (C_{29}H_{50}O_{2}) in the standard and in the bees fed various diets using ESI in the positive mode.

Figure S10. Extracted ion chromatograms (XIC) of β-carotene in the bees fed various diets using ESI in the positive mode compared to the UV absorbance of β-carotene standard. We have attempted to ionize the standard of β-carotene under the same instrument conditions; however,
we were unable to generate ions for XIC-XIC comparison due to the lack of ionization capability of the molecule. Thus, this analysis was recorded as “putative “identification.

Figure S11. Extracted ion chromatograms (XIC) of linoleic acid (C_{18}H_{32}O_{2}) in the bees fed various diets using ESI in the positive mode.
**Figure S12.** Structures of all compounds that are identified through LC-MS standards and putatively identified through GC-MS data comparison to NIST database.

**Table S1.** Amount of extract per sample. Each sample contained 8 bee abdomens. Each diet had 4 biological replicates, resulting in the extraction 8 x 4 bee abdomens per diet condition.

|                   | Sugar Fed Bees | Pollen Fed Bees | Chlorella Fed Bees | Spirulina Fed Bees |
|-------------------|----------------|-----------------|--------------------|--------------------|
| **Amount (mg)**   | 22.2           | 63.0            | 38.6               | 40.2               |
|                   | 25.1           | 59.5            | 30.3               | 53.1               |
|                   | 144            | 72.4            | 51.8               | 40.2               |
|                   | 50.1           | 58.0            | 52.3               | 43.7               |
| **Average Extract (mg)** | **60.5**       | **63.2**        | **43.2**           | **44.3**           |
| **Standard Deviation** | 57.5           | 6.48            | 10.7               | 6.12               |
Table S2. MZmine2.53 Parameters for LC-MS data Processing.

| Module                          | Parameters                                                                 |
|--------------------------------|---------------------------------------------------------------------------|
| Raw Data methods               |                                                                           |
| ➢ Feature Detection            | Scans: MS level: 1                                                        |
| ➢ Mass detection               | Polarity: positive                                                        |
|                                | Mass Detector: centroid                                                   |
|                                | Noise Level: 1.5 E6                                                       |
|                                | Mass list name: masses                                                    |
| Raw Data methods               |                                                                           |
| ➢ Feature detection            | Scans: MS level: 1                                                        |
| ➢ ADAP Chromatogram builder    | Mass list: masses                                                         |
|                                | Min group size # of scans: 5                                              |
|                                | Group intensity threshold: 1.5 E6                                          |
|                                | Min highest intensity: 7.5 E6                                              |
|                                | m/z tolerance: 0.001                                                       |
| Feature list methods           |                                                                           |
| ➢ Feature detection            | Algorithm: Wavelets (ADAP):                                               |
| ➢ Chromatogram deconvolution   | S/N threshold: 10                                                         |
|                                | S/N estimator: Intensity window SN                                         |
|                                | Min feature height: 1                                                     |
|                                | Coefficient/area threshold: 110                                            |
|                                | Peak duration range: 0.00-2.00                                             |
|                                | RT wavelet range: 0.00-0.10                                                |
|                                | m/z center calculation: MEDIAN                                             |
| Feature list methods           |                                                                           |
| ➢ Isotopes                     | m/z tolerance: 0.0015 m/z                                                  |
| ➢ Isotopic peaks grouper       | Retention time tolerance: 0.05                                            |
|                                | Monotonic shape: unchecked                                                |
|                                | Maximum charge: 3                                                         |
|                                | Representative isotope: Most Intense                                      |
| Feature list methods           |                                                                           |
| ➢ Alignment                    | m/z tolerance: 0.0015 m/z                                                  |
| ➢ Join alignment               | Weight for m/z: 2                                                          |
|                                | Retention time tolerance: 0.05                                            |
|                                | Weight for RT: 1                                                           |
|                                | Require same charge state: checked                                        |
|                                | Require same ID: unchecked                                                |
|                                | Compare isotope pattern: checked                                          |
|                                | Isotope m/z tolerance: 010015 m/z                                          |
|                                | Minimum absolute intensity: 7.56 E6                                        |
|                                | Minimum score: 50.0 %                                                      |
| Peak list methods              |                                                                           |
| ➢ Filtering                    | Height: checked, 1.0 E5-1.0 E10                                            |
| ➢ Peak filter                  | # data points: checked, 5-100                                              |
| Export/Import                  |                                                                           |
| ➢ Export to CVS File           | Export common elements:                                                   |
|                                | Export row ID: checked                                                    |
|                                | Export row m/z: checked                                                   |
|                                | Export row retention time: checked                                         |
|                                | Export data file elements:                                                |
|                                | Export peak area: checked                                                 |
### Table S3. MZmine2.53 Parameters for GC-MS data Processing.

| GC-MS Data Processing Steps | Module | Parameters |
|-----------------------------|--------|------------|
| **Raw Data methods**        | Feature Detection | Scans: MS level: 1 |
|                             | Mass detection | Polarity: positive |
|                             |               | Mass Detector: centroid |
|                             |               | Noise Level: 5.0 E2 |
|                             |               | Mass list name: masses2 |
| **Raw Data methods**        | Feature detection | Scans: MS level: 1 |
|                             | ADAP Chromatogram builder | Mass list: masses2 |
|                             |               | Min group size # of scans: 3 |
|                             |               | Group intensity threshold: 5.0 E2 |
|                             |               | Min highest intensity: 5.0 E2 |
|                             |               | m/z tolerance: 0.75 |
| **Feature list methods**    | Feature detection | Algorithm: Wavelets (ADAP): |
|                             | Chromatogram deconvolution | S/N threshold: 1 |
|                             |               | S/N estimator: Intensity window SN |
|                             |               | Min feature height: 500 |
|                             |               | Coefficient/area threshold: 110 |
|                             |               | Peak duration range: 0.05-2.00 |
|                             |               | RT wavelet range: 0.00-0.10 |
|                             |               | m/z center calculation: MEDIAN |
| **Feature list methods**    | Isotopes | m/z tolerance: 0.001 m/z |
|                             | Isotopic peaks grouper | Retention time tolerance: 0.05 |
|                             |               | Monotonic shape: unchecked |
|                             |               | Maximum charge: 3 |
|                             |               | Representative isotope: Most Intense |
| **Feature list methods**    | Alignment | m/z tolerance: 0.75 m/z |
|                             | Join alignment | Weight for m/z: 2 |
|                             |               | Retention time tolerance: 0.05 |
|                             |               | Weight for RT: 1 |
|                             |               | Require same charge state: unchecked |
|                             |               | Require same ID: unchecked |
| **Feature list methods**    | Filtering | # data points: checked, 2-100 |
|                             | Peak filter | |
| **Export/Import**           | Export to CVS File | Export common elements: |
|                             |               | Export row ID |
|                             |               | Export row m/z |
|                             |               | Export row retention time |
|                             |               | Export data file elements: |
|                             |               | Export peak area |
Table S4. Primers used in this study

| Gene (accession number) | Forward 5’-3’ | Reverse 5’-3’ | Annealing temperature (°C) | Study |
|-------------------------|---------------|---------------|-----------------------------|-------|
| actin (XM_623378)       | TGCCAACACTGT  | AGAATTGACCC   | 55.0                        | Alaux et al., 2011 |
|                         | CCTTCTTG      | ACCAATCCA     |                             |       |
| vitellogenin (vg) (AJ517411) | GTTGGAGAGCAA | TCGATCCATTCC  | 57.5                        | Salmela et al., 2016 |
|                         | CATGCGAGA     | TTGATGTG     |                             |       |
| catalase (NM_001178069) | TTCTACTGTGGG  | GTGTGTGTTTAC | 60.0                        | Li et al., 2014 |
|                         | TGGCGAAAG     | CGACCAAATCC  |                             |       |
| CuZn Sod (NM_001178027) | TCAACTTTCAAGG | ATAAACCCACA   | 60.0                        | Li et al., 2014 |
|                         | ACCACATAGTG   | AGCAAGACGAG  |                             |       |
| HSP70 (GB19503)         | GACGCCGGAGC   | AAGCCATAAGC   | 60.0                        | Ramirez et al., 2017 |
|                         | GATAGCAGG     | AATCGCCGCC   |                             |       |
| HSP90 (GB14758)         | ATGCCGGAGGAC  | TTGTGCAATTTTC| 56.0                        | Ramirez et al., 2017 |
|                         | GTCACCAT      | AGCTTGGAAAGCG |                             |       |

Alaux, C., Dantec, C., Parrinello, H., & Le Conte, Y. (2011). Nutrigenomics in honey bees: digital gene expression analysis of pollen’s nutritive effects on healthy and varroa-parasitized bees. BMC genomics, 12(1), 1-14.

Salmela, H., Stark, T., Stucki, D., Fuchs, S., Freitak, D., Dey, A., ... & Sundström, L. (2016). Ancient duplications have led to functional divergence of vitellogenin-like genes potentially involved in inflammation and oxidative stress in honey bees. Genome biology and evolution, 8(3), 495-506.

Li, C., Xu, B., Wang, Y., Yang, Z., & Yang, W. (2014). Protein content in larval diet affects adult longevity and antioxidant gene expression in honey bee workers. Entomologia Experimentalis et Applicata, 151(1), 19-26.

Ramirez, L., Negri, P., Sturla, L., Guida, L., Vigliarolo, T., Maggi, M., ... & Lamattina, L. (2017). Abscisic acid enhances cold tolerance in honeybee larvae. Proceedings of the Royal Society B: Biological Sciences, 284(1852), 20162140.
**Table S5.** Putatively Identified features from LC-MS volcano plot analysis

| RT/m/z value | Identified ion | Accurate mass | Molecular Formula | Example of possible compound Using Dictionary of Natural Products |
|--------------|----------------|---------------|-------------------|---------------------------------------------------------------|
| 5.69/255.23  | [M+H-H2O]+     | 273.2433      | C_{10}H_{18}O_{3} | 2-Hydroxyhexadecanoic acid                                   |
| 5.15/274.275 | [M+H]+         | 274.2752      | C_{16}H_{35}NO_{2} | 2-Amino-1,3-hexadecanediol                                  |
| 7.76/263.238 | [M+H-H2O]+     | 281.2490      | C_{18}H_{32}O_{2} | 4,6-Dimethyl-2,4-hexadecadienoic acid                       |
| 5.03/287.223 | [M+H]+         | 287.2226      | C_{16}H_{30}O_{4} | 2,16-Dihydroxy-6-hexadecenoic acid                         |
| 5.52/335.221 | [M+Na]+        | 313.2386      | C_{18}H_{32}O_{4} | 10,11-Dihydroxy-8,12-octadecadienoic acid                  |
| 4.02/314.270 | [M+H]+         | 314.2702      | C_{18}H_{35}NO_{3} | 2-Amino-4,9-octadecadiene-1,3,8-triol                      |
| 7.99/371.102 | [M+H]+         | 371.1025      | C_{22}H_{44}N_{4}O_{4} | Caulerpinic acid                                      |
| 8.03/386.364 | [M+H]+         | 386.3646      | C_{23}H_{47}NO_{3} | 2-Amino-11-tricosene-1,3,4-triol                           |
| 8.27/463.378 | [M+Na]+        | 441.3958      | C_{27}H_{55}O_{4} | Tricosanedioic acid; Di-Et ester                           |
Table S6. Relative abundance and average peak areas of compounds identified through LC-MS analysis. The closer the p value is to 0, the more significantly different the given metabolite is in a pairwise comparison between diets.

| Compound Name | Avrg T Pollen 1 | Avrg T Pollen 2 | Avrg T Pollen 3 | Avrg T Pollen 4 | Avrg T Chlor 1 | Avrg T Chlor 2 | Avrg T Chlor 3 | Avrg T Chlor 4 | Avrg T Spir 1 | Avrg T Spir 2 | Avrg T Spir 3 | Avrg T Spir 4 | Pollen vs Chlorella p value | Pollen vs Spirulina p value | Chlorella vs Spirulina p value |
|---------------|-----------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------------------|-----------------------------|-----------------------------|
| Zeaxanthin    | 0.00E+00        | 0.00E+00        | 0.00E+00        | 0.00E+00        | 1.66E+07       | 3.13E+07       | 1.56E+07       | 1.61E+07       | 3.72E+06       | 4.29E+06       | 5.13E+06       | 4.49E+06       | 0.01                       | 0.00                        | 0.03                        |
| Lutein        | 0.00E+00        | 0.00E+00        | 0.00E+00        | 0.00E+00        | 1.66E+07       | 3.13E+07       | 1.56E+07       | 1.61E+07       | 3.72E+06       | 4.29E+06       | 5.13E+06       | 4.49E+06       | 0.01                       | 0.00                        | 0.03                        |
| Linolenic acid| 3.12E+07        | 3.35E+07        | 2.19E+07        | 4.27E+07        | 1.54E+07       | 3.04E+07       | 1.43E+07       | 2.05E+07       | 9.67E+06       | 9.80E+06       | 8.51E+06       | 1.10E+07       | 0.07                       | 0.01                        | 0.06                        |
| Linoleic acid | 1.60E+07        | 1.96E+07        | 1.95E+07        | 2.71E+07        | 2.50E+07       | 4.56E+07       | 1.55E+07       | 2.84E+07       | 9.31E+06       | 8.04E+06       | 8.69E+06       | 9.57E+06       | 0.30                       | 0.01                        | 0.05                        |
| Quinic Acid   | 8.90E+05        | 2.74E+05        | 1.55E+06        | 7.98E+05        | 8.35E+05       | 2.28E+05       | 1.92E+05       | 3.60E+05       | 7.02E+05       | 1.18E+06       | 3.15E+05       | 5.36E+05       | 0.18                       | 0.57                        | 0.28                        |
| α-tocopherol  | 3.11E+06        | 2.17E+06        | 1.57E+06        | 2.01E+06        | 3.29E+06       | 4.41E+06       | 2.69E+06       | 3.67E+06       | 2.74E+06       | 2.63E+06       | 1.07E+06       | 2.40E+06       | 0.04                       | 0.99                        | 0.05                        |
| β-carotene    | 2.13E+05        | 9.28E+05        | 5.07E+05        | 2.18E+05        | 6.02E+04       | 6.21E+05       | 1.56E+06       | 5.20E+05       | 1.21E+06       | 6.37E+06       | 8.76E+06       | 9.08E+06       | 0.56                       | 0.05                        | 0.05                        |
Table S7. Relative abundance comparison of metabolites identified through GC-MS library match.

| Number | Compound Name | Retention Time (min) | Bees Fed Pollen | Bees Fed Chlorella | Bees Fed Spirulina | Pollen vs Chlorella | Pollen vs Spirulina | Chlorella vs Spirulina | P value |
|--------|---------------|----------------------|-----------------|-------------------|-------------------|------------------|-------------------|---------------------|---------|
| 1      | phytol acetate | 14.94                | 0.00E+00        | 1.76E+05          | 4.26E+05          | 0.016            | 0                 | 0.197               |         |
| 2      | 1,11,13-octadecatriene | 15.49               | 0.00E+00        | 1.31E+06          | 0.00E+00          | 0.084            | 0                 | 0.084               |         |
| 3      | n-hexadecanoic acid | 15.75               | 9.33E+05        | 2.33E+06          | 4.34E+06          | 0.707            | 0.015            | 0.03                |         |
| 4      | erucic acid    | 16.88                | 2.19E+07        | 1.71E+06          | 7.20E+06          | 0.04             | 0.033            | 0.675               |         |
| 5      | octadecanoic acid | 17.02               | 5.63E+06        | 1.79E+06          | 4.22E+06          | 0.61             | 0.156            | 0.177               |         |
| 6      | 1-heneicosanol | 17.75                | 2.87E+06        | 3.08E+06          | 3.46E+06          | 0.004            | 0.309            | 0.011               |         |
| 7      | heneicosane    | 17.93                | 2.34E+06        | 4.99E+05          | 1.57E+06          | 0.117            | 0.058            | 0.738               |         |
| 8      | n-nonadecanol-1| 18.9                 | 1.92E+06        | 2.44E+05          | 1.99E+06          | 0.035            | 0.757            | 0.024               |         |
| 9      | octacosane     | 19.04                | 1.53E+06        | 6.37E+05          | 1.22E+06          | 0.69             | 0.17             | 0.348               |         |
| 10     | n-tetracosanol | 19.95                | 2.16E+06        | 2.76E+05          | 1.22E+06          | 0.096            | 0.096            | 0.039               |         |
| 11     | hexacosane     | 20.07                | 2.77E+06        | 2.42E+06          | 3.54E+06          | 0.38             | 0.358            | 0.617               |         |
| 12     | tetraeicosane  | 20.25                | 9.27E+05        | 2.51E+05          | 8.37E+05          | 0.879            | 0.507            | 0.744               |         |
| 13     | octacosanol    | 20.95                | 1.55E+08        | 4.13E+06          | 2.85E+06          | 0.423            | 0.063            | 0.129               |         |
| 14     | pentacosan    | 21.03                | 3.23E+06        | 9.86E+05          | 2.04E+06          | 0.332            | 0.093            | 0.67                |         |
| 15     | 2-methylpentacosan | 21.17               | 9.22E+05        | 1.96E+06          | 8.43E+05          | 0.973            | 0.771            | 0.661               |         |
| 16     | 1-heptacosanol | 21.81                | 1.86E+06        | 5.55E+05          | 2.52E+06          | 0.137            | 0.507            | 0.209               |         |
| 17     | (2)-9-Tricosene | 21.85               | 1.32E+06        | 4.57E+05          | 2.46E+06          | 0.528            | 0.305            | 0.216               |         |
| 18     | docosane      | 21.9                 | 2.99E+06        | 3.53E+05          | 7.71E+05          | 0.008            | 0.007            | 0.212               |         |
**Table S8.** Benefits and limitations of the subtractive approach used in the study.

| Feature Subtraction Approach | |
|-------------------------------|---------------------------------------------------------------|
| **What was the goal of the approach?** | To show which features originated from the honey bees as opposed to those from the diet extracts. |
| **How was the subtraction method carried out?** | The features (peak areas and appropriate m/z over retention time values) from bees fed sugar, as well as features from the specific diet samples (i.e., pollen, *Chlorella*, or spirulina), were all subtracted from the feature list acquired from bees fed on the respective diets (See Figure S3 for more details). |
| **What features were identified after subtraction?** | Features remaining after subtraction of the bees fed sugar and diet features were either uniquely made by honey bees or were upregulated by the honeybees when the specific diet was induced. |
| **What are the benefits of the approach?** | **Cost and simplicity** The approach uses subtraction of peak areas originating from the MzMine generated dataset. It is entirely done using Excel tables, and thus, it does not require the use of additional software. **Efficiency** Identified unique features can be pursued further, structurally identified, and compared across diets |
| **What are some limitations of this approach?** | Subtraction of features acquired from GC-MS can be challenging due to the comparison of metabolite features vs fragment features that can lack proper filtering steps. Also, the method is dependent on the filtering and data processing steps prior to subtraction |
