Metadata of the study.

Supplementary method S1. Volatilomics metadata.

Metabolomic methods

Based on the recommendation of the Metabolomics Standards Initiative[1], metadata in this study was shown as follows.

Plant context metadata

Biosource:

Soybean cultivar, “Satonofofoemi.”

Source:

Two independent bulk soils in each pot in two fields: Fukushima Agricultural Technology Center in Koriyama city, Fukushima（37° 28' 29.8" N, 140° 23' 01.4" E）in 2019-2021, while Fukushima research station, Tohoku Agricultural Research Center, NARO in Fukushima city, Fukushima (37° 42' 45.8" N, 140° 23' 13.0" E) in 2021. Soil qualities of the two field is as follows: andosol soils[2] in Koriyama, while gray lowland soils[3] in Fukushima. Soybean cultivation has been stated since 2017 in both fields. We applied four conditions to collect soil samples as shown in table_expdesign.docx. We used chemical fertilizer (Kumiai kasei koudo500, JCAM AGRI, Japan) and cow manure fertilizer (Heiwagyu funtai hi, Tachikawa heiwa nouen, Tochigi, Japan).

Amount: Frozen soil (3 g flesh weight).

Experimental condition, Sampling and sampling date:

Seeds were treated with Cruiser MAXX (Syngenta, Japan) before sawing. Soil with two different plots in each field was sampled in concentric circles at 40 cm intervals around the plants. Two replicates sampled in each plot (Table 1). According to root length (Supplementary Figure S1), soil samples (5-15 cm in depth) were collected for sampling at “before sowing”, “flowering periods” and “harvesting periods” (Table S1).

Metabolism quenching method:

All samples were transferred by chilled shipping after sampling. The collected samples were kept in a biofreezer (Nihon Freezer, Japan) at -40°C until use.

Chemical analysis metadata

Sample preparation

Headspace (HS) collection using the SPME fiber:

Volatile organic compounds (VOCs) in soil was extracted using SPME method (Kusano, et al., Metabolites. 2015) with modifications. In brief, thee g of soil was used in the study. We used SPME fiber (50/30 μm, DVB/CAR/PDMS, Stable Flex, 23Ga, Supelco, US). Solid NaCl (0.9 g) and 3 ml of 0.1 mol/l EDTA (Ethylenediaminetetraacetic acid, pH 7.5) in milliQ water was

added to 20-ml of each SPME-GC vial (Agilent Technologies, US). 2EPA524.2 fortification solution (5 mg/l, Supelco, US) was used as internal standard. The SPME device for a CTC CombiPAL auto-sampler (CTC Analytics, Zwingen, Switzerland) was purchased from AMR (Tokyo, Japan). Before analysis, the fiber was conditioned at 250 °C for 340 s in the injection port of an Agilent 6890 N gas chromatograph (Agilent Technologies, Wilmington, USA) equipped with a 30 m × 0.25 mm inner diameter fused-silica capillary column with a chemically bound 0.25-µl film Rxi-5 Sil MS stationary phase (RESTEK, Bellefonte, US). Collection of volatiles was carried out by inserting the SPME-fiber to the vial and by trapping the VOCs for 5 min at 80 °C under continuous agitation. After HS collection it was placed in the injection port of the gas chromatograph that was coupled to a Pegasus III TOF mass spectrometer (LECO, US). The thermo desorption of VOCs occurred for 10 min at 270 °C.

Headspace (HS) collection using the SPME fiber:

A SPME device for a CTC CombiPAL auto-sampler (CTC analytics, Zwingen, Switzerland) was purchased from AMR (Tokyo, Japan). We used an SPME fiber, which was a 65-µm-thick layer of polydimethylsiloxane (PDMS)/divinylbenzene (DVB) -fused silica (FS) fiber/stainless steel (SS) tube for the study. The fiber was conditioned at 250°C for 0.5 h in the injection port of an Agilent 6890N gas chromatograph (Agilent Technologies, Wilmingston, USA) equipped with a 30 m × 0.25 mm inner diameter fused-silica capillary column with a chemically bound 0.25-µl film Rxi-5 Sil MS stationary phase (RESTEK, Bellefonte, USA) before analysis. The fiber was exposed to the vial headspace for 10 min at 80°C under continuous agitation. After HS collection, the fiber was placed into the injection port of the gas chromatograph that was hyphenated with a Pegasus III TOF mass spectrometer (LECO, St. Joseph, USA). The volatile compounds were thermally desorbed for 10 min at 270°C from the fiber. After five injections, the fiber was bake out by Needle heater (CTC analytics) for 15 min at 270°C to clean and ready for the next sample.

GC-TOF-MS conditions

HS sample was injected by an CTC CombiPAL autosampler (CTC analytics) into the injection port of the gas chromatograph (Agilent Technologies) equipped with the Rxi-5 Sil MS column (RESTEK, Bellefonte, USA). Helium was used as the carrier gas at a constant flow rate of 1 ml/min. Septum purge flow and time was set at 20 ml/min and 30 sec. Total flow was set at 19.2 ml/min. Gas saver was used and the flow was set at 15 ml/min. The inlet temperature was 250°C. The temperature program for the HS-VOC analysis started with a 4-min isothermal step at 40°C and this was followed by temperature ramping at 10°C to a final temperature of 250°C, which was maintained for 4 min. The transfer line and the ion source temperatures were 250°C and 200°C, respectively. Ions were generated by a 70-eV electron beam. The acceleration voltage was turned on after a solvent delay of 200 sec.

Data acquisition was performed on a Pegasus III TOF mass spectrometer (LECO) at an acquisition rate of 30 spectra/s in the mass range of a mass-to-charge ratio of m/z = 20–300. Ten µl of EPA524.2 fortification solution (2.5 mg/l in methanol, Supelco, US) was used for data normalization. The alkane standard mixture contains 10 µl of C_8−C_{20}, 40 mg/l each in hexane (SIGMA-ALDRICH, US), was used for calculating the retention indices (RIs)[4,5]. Empty vials were analyzed in every five samples for quality control.

Data processing

Data processing for GC-MS data:

For HS-VOC profiling, non-processed MS data from GC-TOF-MS analysis were exported in the NetCDF format generated by chromatography processing and mass spectral deconvolution software, Leco ChromaTOF version 4.71.00 (LECO, St. Joseph, USA) to MATLAB R2011b (Mathworks, Natick, USA) The chromatograms were preprocessed using the high-throughput data analysis (HDA) method[6]. Peaks in the data matrix was normalized by using unique mass of 1,2-dichlorobenzene-D_4, m/z = 150.

The resolved MS spectra obtained from the HDA method were matched against reference mass spectra using the NIST mass spectral search program for the NIST/EPA/NIH mass spectral library (version 2.2) and our custom software for peak annotation written in JAVA. Peaks were tentatively identified according to the guideline of metabolite identification[1]. When mass spectra showed a match value greater than 799 and the corresponding peaks had RIs with small differences (< |30 unit|) by comparing their resolved mass spectra and RIs to those in the reference libraries, i.e., Adams library (the 3rd and the 4th edition)[7], Terpenoids library (http://massfinder.com/wiki/Terpenoids_Library), VocBinBase[7], NIST11[8], FFNSC 3 Wiley Library (Wiley, US), and own custom library, the peaks were considered to be putatively annotated compounds.

Statistical analysis

SIMCA-P+ 14.0 software (Umetrics AB, Umeå, Sweden) was used for multivariate analysis. The profile data were $\log_{10}$-transformed, centered and scaled to unit variance for the analysis.

The R (version 4.3.1) was used for conducting of PCA and correlation analysis. The profile data (VOC profiles, soil ionome, soil metabolome, rhizosphere chemical, and soil physics) were z-score transformed before statistical analysis. For the microbiome data, we performed principal coordinate analysis (PCoA) based on the weighted UniFrac distance matrix using R cmdscale function instead of PCA. Correlations between the components were calculated and P values were adjusted as False discovery rate according to Benjamini-Hochberg method[9].

Supplementary Method S2. Soil microbiome and rhizosphere microbiome analysis.

Bulk soil samples were freeze-dried using a freeze dryer (TOKYO RIKAKIKAI CO., LTD., Cat. # 63-1397-09). The freeze-dried samples were frozen using liquid nitrogen and immediately ground into a fine powder at 3000 rpm for 15 seconds via a Multi-Beads Shocker (Yasui Kikai Co., Osaka, Japan, Cat. #MB2200(S)). DNA was extracted from the 250-260 mg soil samples using an Extrap Soil DNA Kit Plus ver.2 (BioDynamics Laboratory Inc., Cat. # 212-006) following the instructions provided by the manufacturer. The library for sequencing was prepared through a two-step PCR amplification, focusing on the V4 region of the bacterial 16S rRNA gene using 515f and 806rB primers as detailed in prior studies[10,11]. Libraries were measured for concentration using an Infinite 200 PRO M Nano+ microplate photometer (TECAN Japan Co. Ltd.) and combined in equimolar ratios into a composite library. This library was then sequenced on an Illumina MiSeq system with the MiSeq Reagent Kit v3 (600-cycles, Illumina, CA, USA). Bioinformatic and statistical analyses were performed using the Quantitative Insights Into Microbial Ecology 2 software (QIIME 2 version 2020.11, https://qiime2.org/9). The truncation lengths were set at 200 for forward reads and 140 for reverse reads. Taxonomic classifications were made using the Naive Bayes q2-feature-classifier, trained on the 515F/806R region with 99% operational taxonomic units (OTUs) from the SILVA 138 rRNA database. Sequences from archaeal, eukaryotic mitochondrial, and

chloroplast contaminants were excluded from the final feature table. Following the taxonomic assignment of amplicon sequence variants (ASVs), the selected sequences were aligned using MAFFT and phylogenetic trees were constructed using FastTree. A sequencing depth of 3,165 was chosen, reflecting the lowest sequence count found in any sample and a weighted UniFrac distance matrix was calculated.

Supplementary Method S3. metabolome analysis.

We used 3 g of soil for the experiment. Soil was extracted with milli-Q water with 7 ISs ([$^{2}H_4$]-succinic acid, [$^{13}C_5$, ^{15}N]-glutamic acid, [$^{2}H_7$]-cholesterol, [$^{13}C_5$]-proline, [$^{13}C_{12}$]-sucrose, [$^{2}H_4$]-1,4-butanediamine, [$^{2}H_6$]-2-hydoxybenzoic acid, and [$^{13}C_6$]-glucose). The equivalent of 0.25 g per ml was extracted for 1h at 110℃ with milli-Q water. After centrifugation for 10 min at 1000 g, room temperature. Supernatant was feezed under -80℃, then lyophilized to remove milli-Q water. Three lyophilic ISs ([$^{2}H_7$]-cholesterol, [$^{13}C_3$]-myristic acid, and [$^{13}C_4$]-hexadecanoic acid) was also used after removing milli-Q water. We conducted metabolome analysis by using GC-TOF-MS according to Kusano, et.al[12].

References

1 Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211-221, doi:10.1007/s11306-007-0082-2 (2007).

2 Shoji, S., Nanzyo, M., Dahlgren, R. A. & Quantin, P. EVALUATION AND PROPOSED REVISIONS OF CRITERIA FOR ANDOSOLS IN THE WORLD REFERENCE BASE FOR SOIL RESOURCES. *Soil Science* **161**, 604-615 (1996).

3 Amano, Y. Classification of cultivated soils in Japan. *Japan Agricultural Research Quarterly* **18**, 275-283 (1985).

4 Wagner, C., Sefkow, M. & Kopka, J. Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry* **62**, 887-900, doi:10.1016/s0031-9422(02)00703-3 (2003).

5 Schauer, N. *et al.* GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett* **579**, 1332-1337, doi:10.1016/j.febslet.2005.01.029 (2005).

6 Jonsson, P. *et al.* High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. *Anal Chem* **77**, 5635-5642, doi:10.1021/ac050601e (2005).

7 Sparkman, O. D. Identification of essential oil components by gas chromatography/quadrupole mass spectroscopy Robert P. Adams. *Journal of the American Society for Mass Spectrometry* **16**, 1902-1903, doi:10.1016/j.jasms.2005.07.008 (2005).

8 Stein, S. E. *et al.* Evaluation of the NIST/EPA/NIH Mass Spectral Library. *Abstr Pap Am Chem S* **218**, U368-U368 (1999).

9 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289-300, doi:https://doi.org/10.1111/j.2517-6161.1995.tb02031.x (1995).

10 Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences* **108**, 4516-4522, doi:doi:10.1073/pnas.1000080107 (2011).
11 Ichihashi, Y. *et al.* Multi-omics analysis on an agroecosystem reveals the significant role of organic nitrogen to increase agricultural crop yield. *Proceedings of the National Academy of Sciences* **117**, 14552-14560, doi:doi:10.1073/pnas.1917259117 (2020).
12 Kusano, M. *et al.* Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography-time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. *Journal of Chromatography B* **855**, 71-79, doi:https://doi.org/10.1016/j.jchromb.2007.05.002 (2007).