

Article

Secondary Metabolite Differences between Naturally Grown and Conventional Coarse Green Tea

Kousaku Ohta ^{1,2,*} , Tatsuya Kawaoka ^{1,3}  and Masatoshi Funabashi ^{1,3} 

¹ Sony Computer Science Laboratories, Inc., Tokyo 141-0022, Japan; t-kawaoka@synecoculture.org (T.K.); masa_funabashi@csl.sony.co.jp (M.F.)

² Laboratory of Sustainable Agriculture, Department of Global Agricultural Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan

³ Synecoculture Association, Tokyo 105-0014, Japan

* Correspondence: kousaku@csl.sony.co.jp; Tel.: +81-3-5443-4380

Received: 31 October 2020; Accepted: 11 December 2020; Published: 14 December 2020



Abstract: Crop culture conditions are one of the important interfaces between food, the environment, and health, and an essential research area for maintaining social-ecological integrity. In recent years, it has been reported that the difference in culture conditions between monoculture with external inputs (*in cultura*) and self-organized ecological niches (*in natura*) is significant for the resulting physiological property of plants. It has also been suggested that there exist metabolic proxies in various foods that can separate these two culture conditions, which does not depend on a single component but on the distribution of various compounds. However, little has been studied in a time series of replicated production to quantify the reproducibility of these metabolomic features associated with culture conditions. In this study, we obtained metabolome data of coarse green tea (*Camellia sinensis*) grown in the same region in Japan under both *in cultura* and *in natura* culture conditions over the course of six years, and constructed a list of multiple components that separated the effects of culture conditions by statistical analysis, and estimated the metabolic functions of the compounds that contributed to the separation. The results suggest that naturally grown samples are rich in allelochemicals, such as phytochemicals, alkaloids, phenylpropanoids, steroids, as well as the compounds related to microorganisms and vitamin B6 that imply the interactions with the soil microbiome. The estimated physiological functions of the distinctive compounds suggest that the *in natura* crop production is not only beneficial with known properties of maintaining ecosystem health such as soil functions and pathogen control, but also for the augmentation of the plant secondary metabolites that support long-term health protective effects.

Keywords: culture condition; secondary metabolites; Synecoculture; *Camellia sinensis*; sustainable diet

1. Introduction

Crop culture conditions represents one of the essential subjects for social-ecological sustainability, which links human and ecosystem health through affecting the plant's metabolic state [1]. According to previous researches, two categories of culture condition are distinguished as a dominant parameter on plant metabolites: (1) *in cultura* condition based on the monoculture with external inputs that maximizes the individual plant growth known as the physiological optimum, and (2) *in natura* condition based on the self-organization of ecological niches known as the ecological optimum [2–4]. Although dominant in present crop production, the *in cultura* culture conditions are inevitably associated with biodiversity loss and imposing a heavy burden on the global environment. This tendency has evolved since prehistoric times, and there are some reports that it could lead to the sixth mass extinction in Earth's history [5,6]. A globally imminent crisis such as the planetary state shift and massive loss

of biosphere integrity is anticipated to occur by 2045 [7]. Recently increasing forest fires, outbreaks of pandemics, and locust plagues could be considered as the symptoms arising from the functional disorder of ecosystems. At the same time, our internal environment such as human health is also sharing the common ground with external ecosystems: systematic disorder of human metabolic state such as non-communicable disease is related to modern dietary habits, which is associated with the development of monoculture exploitation through the use of pesticides, increased consumption of over processed food, and nutrition deficiency due to soil depletion [1,8]. To resolve such an environment-diet-health trilemma, we need a comprehensive approach at the social-ecological system level that contains all inter-related domains, instead of problem-wise symptomatic treatments [3,9].

Agriculture is at the core of such an open complex system problem, in which *in natura* food productions are likely to play a fundamental role in resolving the global trilemma among food production, environmental protection and improvement in public health [1]. Plants that are grown naturally interact with diverse other species and environments through the enhancement of various nutrient cycles and producing various secondary metabolites, which not only serves for plant's own survival through physiological and allelopathic effects, but also exerts health-protective functions in the metabolism of the plant consumers.

On the contrary, when plowing, fertilizers, and chemicals are continuously applied under *in cultura* culture conditions, the interactions between plants, animals, and soil microorganisms are known to be simplified by the loss of diversity in these organisms (e.g., [10]), therefore various regulating services are also reduced [11]. Regulating services provided by the ecosystems include disease control, climate disaster mitigation and pollination, among others [11,12]. *In cultura* culture condition is also known to increase the risk of infectious diseases due to weakened ecological suppression effects of pathogens and virulent viruses [13–19].

Despite its ecological importance and potential effects on the nutritional profile, there are few comparative studies of the metabolic difference induced by the *in natura* and *in cultura* conditions. Although a previous study has pointed out the existence of statistical invariant features of metabolic profiles that distinguish between these culture conditions across different databases [2], there is little study on the reproducibility of such physiological differences over the course of several years under identical culture methods.

Here, we conducted metabolome analysis of coarse green tea (*Camellia sinensis*) with six sets of samples over six years (Syneco 2014–2019), produced under *in natura* culture conditions called Synecoculture [4], and with five sets of samples over five years produced under *in cultura* conditions of a standard conventional farming (Conv 2015–2019). We focused on the compounds that most reflected the differences of culture conditions between *in natura* and *in cultura*, and analyzed their estimated metabolic functions with the use of chemical ontology databases [20,21].

2. Materials and Methods

All coarse green tea samples used in this study were produced by traditional tea farms at Watarai-Cho, Mie prefecture in Japan. Harvest and processing of coarse green tea leaves were homogenized with a conventional method: 2nd cropping (foliage lower than shoots) of the yearly 1st harvest during late May to early June between 2014 and 2019, followed by the standard processing of steaming, kneading, and drying in a local mechanized factory. Leaves from more than 7000 m² culture within 2 km distance were blended for averaging plot-wise variation and to compare samples from different culture methods with a common geographical range.

Two culture conditions of tea were differentiated to examine the metabolomic difference of tea extract: 1. *in cultura* conventional monoculture condition following the standard protocol of Japan Agricultural Cooperatives Ise branch, with the routine application of inter-crop cultivation, synthetic and organic fertilizers, pesticides, fungicides, and herbicides. 2. *in natura* mixed polyculture condition without the application of tillage, fertilizer, and chemicals, defined as Synecoculture [4]. Intercropping

with a variety of vegetables and fruit trees was introduced, along with spontaneous generation of weeds [22].

The conventional condition is controlled in physiological optimum range of production, while Synecoculture is based on ecological optimum dominant in natural vegetation. These qualitatively different growing conditions are biologically termed as *in cultura* and *in natura* conditions of field culture, respectively, in relation to the definition of sustainable diet [2].

Then, eleven sets of samples of dried coarse green tea were produced, from the Synecoculture fields from 2014 to 2019 and conventional tea culture from 2015 to 2019. We hereafter call these samples as Syneco 2014, Syneco 2015, Syneco 2016, Syneco 2017, Syneco 2018, Syneco 2019, Conv 2015, Conv 2016, Conv 2017, Conv 2018 and Conv 2019, respectively.

2.1. Metabolome Analysis

2.1.1. Metabolite Extraction

Briefly, 3.0 g of dried tea leaves stuffed in a teabag was boiled with 1l of ultrapure water (Milli-Q) at 90–93 °C for 10 min in a glass beaker and left at room temperature for 2 h. Each set of samples Syneco 2015–2019 and Conv 2015–2019 were extracted simultaneously in each year, and Syneco 2014 was extracted together with 2015 samples. Organic compounds were further extracted for the samples of metabolome analyses:

For 2014–2017 samples: Each 100 µL sample was mixed with 300 µL methanol and centrifuged with 10,000× g, 10 min, 4 °C. The supernatant was filtered with PTFE filter (Millipore, Cat.SLLGH04NK) and centrifuged through Monospin C18 spin columns with 5000× g, 2 min, 4 °C in order to remove insoluble matters and low polarity components. A mock sample of ultrapure water was prepared with the same procedure, and was used to evaluate and remove background noise contained in the sample preparation and/or LC-MS analysis.

For 2018–2019 samples: Each 100 µL sample was mixed with 300 µL methanol and centrifuged with 15,000 rpm, 10 min. 100% methanol centrifuged through Monospin C18 spin columns with 5000× g, 2 min. Then, 75% methanol centrifuged through the same columns with 5000× g, 2 min. These processes were pretreatment for column equilibration. Then, the supernatant of the sample was centrifuged through the same columns with 5000× g, 2 min. After that, the supernatant of the sample was filtered with a 0.2-µm filter. A mock sample of ultrapure water was prepared with the same procedure, and was used to evaluate and remove background noise contained in the sample preparation and/or LC-MS analysis.

All of the methanol was obtained from FUJIFILM Wako Pure Chemical Corporation.

2.1.2. LC-MS Analysis of 2014–2017 Samples

LC–MS analysis was performed with a combination of Agilent 1200 series [23] and Thermo fisher scientific LTQ ORBITRAP XL [24]. The parameters of measurement are summarized in Supplementary Materials File S1.

After converting raw data (obtained from LTQ ORBITRAP XL) to a text file with the use of ProteoWizard [25], LC-MS data were analyzed using PowerGet ver. 3.5.7 [26] with the following procedure to attribute each MS peak to a chemical formula:

1. Empirical detection of compound peaks, calculation of accurate mass, calculation of compound peak intensity.
2. Differentiation of simultaneous elution peaks with respect to the profile of adduct ion peaks, ionization mode, and natural ¹³C isotopic compound peaks.
3. Matching between MS peaks and MS/MS data, calculation of ¹³C/¹²C isotope ratio with ion intensity in order to estimate C number in each compound, and estimation of ionization mode.
4. Aggregation and sorting of compound peaks with respect to the elution time, accurate mass, and MS/MS patterns for all samples.

- Matching of calculated mean accurate mass with monoisotopic compounds in public databases [20,21] with the use of MF Searcher [27] and derivation of a corresponding chemical formula.
- Truncate the compound peaks with less than 2 times intensity of the mock sample.

The parameters of these analyses are summarized in Supplementary Materials File S1 and Table 1.

Table 1. Summary of the sample replicates and LC-MS analysis method.

Measurement		Parameters w.r.t. Sampling Year					
Replicate	Method	2014	2015	2016	2017	2018	2019
Biological replicate	Plant-wise Area-wise	Averaged	Averaged	Averaged	Averaged	Averaged	Averaged
Product replicate	Teabag-wise	3	3	3	3	3	3
Technical replicate	Photospectrometry (absorbance error)	3 (1/100–1/10,000)	3 (1/100–1/10,000)	3 (1/40–1/1000)	3 (1/50–1/1000)	3 (1/40–1/10,000)	3 (1/100–1/10,000)
	LC-MS (intensity error)	1 (estimated CV: 10–20%)	1 (estimated CV: 10–20%)	1 (estimated CV: 10–20%)	1 (estimated CV: 10–20%)	3 (measured CV: 16.2%)	3 (measured CV: 16.8%)
HPLC			Agilent 1200 series			Ultimate 3000 RSLC	
Tea Sampling and Processing				Same protocol			

2.1.3. LC-MS Analysis of 2018–2019 Samples

LC-MS analysis was performed with a combination of Ultimate 3000 RSLC (Thermo Fisher Scientific) [28] and Q Exactive (Thermo Fisher Scientific) [29]. Samples were analyzed 3 times for each sample.

After converting raw data (obtained from Q Exactive) to a text file with the use of ProteoWizard, LC-MS data were analyzed using PowerGetBatch [30] with the following procedure to attribute each MS peak to a chemical formula:

- Empirical detection of compound peaks, calculation of accurate mass, calculation of compound peak intensity
- Ionization status judgment
- Alignment of compound peaks
- Matching of calculated mean accurate mass with monoisotopic compounds in public database with the use of MF Searcher and derivation of a corresponding chemical formula

The parameters of these analyses are summarized in Supplementary Materials File S1 and Table 2.

Table 2. KEGG BRITE compound classification. The numbers of different chemical formulae (# Formulae) associated with the Syneco- and Conv-distinctive loadings for each category of chemical ontology in KEGG BRITE were shown. Uncertainty Score is the sum of inverse numbers of structural isomers for each chemical formula. The smaller the number, the greater the number of structural isomers, indicating the higher uncertainty of compound identification. See the Supplementary Materials Files S6–S8 for more details.

	Syneco		Conv	
	# Formulae	Uncertainty Score	# Formulae	Uncertainty Score
Flavonoid	10	4.075	13	6.751587
Phytochemical	41	15.98387	25	12.61299
Alkaloid	4	3.125	3	1.47619
Phenylpropanoid	17	3.904167	4	1.821429
Steroid	2	1.083333	0	-
Total	74	28.17137	45	22.662196

As a premise, this metabolome analysis is a result of projecting measured exact masses onto the public database, so there is not sufficient resolution for the accurate distinction between structural isomers.

2.1.4. Integration of Metabolite Data of 2014–2019 Samples

Based on the exact mass detected, MF Searcher [27] was used to match the same compound every year with an error of 1 ppm in KEGG database, and integrated the metabolomic data of all years. The same procedure was also performed in Flavonoid Viewer [31] to compare the total estimated amount of flavonoids (see Supplementary Materials File S5). When multiple structural isomers were detected from the KEGG database, the intensity was taken as the average value for each detected year, because these cannot be distinguished at the present resolution of LC-MS.

2.1.5. Biological and Technical Replicate

The sample tea leaves were obtained from the mixed harvest over entire fields, at a total of 3 times mixing (triple homogenization) at the time of harvesting, kneading, and drying, so it was not possible to sample each small area or individual tea tree as biological replicates. Therefore, this was regarded as a representative of the entire field for each year, and biological variance was greatly averaged through harvesting and processing in this study. Since the tea leaves were inevitably homogenized through the processing, it was not possible to take plant-wise or area-wise biological replicates for the product analyzed.

LC-MS analysis was performed once for the 2014–2017 and three times for the 2018–2019 samples. To assess the homogeneity of the tea products and reproducibility of measurement, the 2014–2019 samples were extracted three times with the same procedure of sample extraction as “product replicate”, and the temporal changes in absorbance of the samples were measured with a spectrophotometer. There was almost no difference in the extraction and no change after a certain period of time, 6 h at the maximum (1/40–1/10,000 error of the absorbance). The summary of the sample replicates and LC-MS analysis method are shown in Table 1.

2.2. Statistical Analysis

Normality of all the formulae were tested with the Shapiro-Wilk test, and 41 out of 342 Syneco formulae and 51 out of 342 Conv formulae had passed under 5% significance threshold. The Welch’s *t*-test (applicable even if the normally distributed data do not have homoscedasticity) and the Brunner-Munzel test [32] (that does not assume either homogeneity of variance nor normality) were performed with Syneco 2014–2019 and Conv 2015–2019 for the LC-MS raw and logarithmic intensity values of the identified compounds. For the 2018 and 2019 samples, the mean values of the three measurements on the same samples were used. Welch’s *t*-test was performed with Microsoft Excel ver. 16.16.21, and a Brunner–Munzel test was performed with statistical analysis software R ver. 3.5.0.

The variance value representing the magnitude of the year-to-year variation of each compound was calculated, and the distribution of the variances was compared by F-test between Syneco and Conv samples. For the 2018 and 2019 samples, the mean values of the three measurements on the same samples were used. F-test was performed with Microsoft Excel ver. 16.16.21. These results were summarized in Figure 1.

In order to investigate the overall effects of culture conditions on the metabolic state, we performed the principal component analysis (PCA) after the renormalization of the intensity values for each compound. Then, PCA was applied repeatedly by increasing the number of compounds having a high positive/negative loading to a PC to see the separability of Syneco and Conv samples on the PC plane (hereafter denoted as LS-PCA, meaning linear separation with PCA). As a result, the top 65 compounds of positive/negative loadings concerning PC3 were extracted (in total 130 compounds), which were enough to thoroughly separate the two culture conditions (Figures 2 and 3). The LS-PCA was performed by a statistical analysis software R ver. 3.5.0.

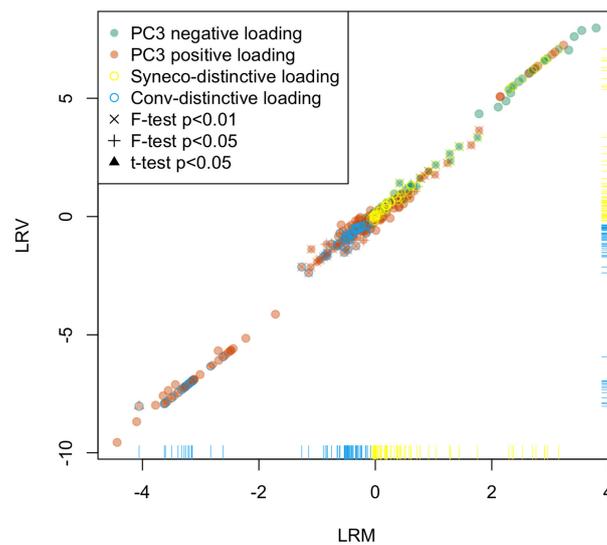


Figure 1. Results of metabolome analysis of coarse green tea samples. The $\log(\text{Syneco}/\text{Conv})$ values of mean intensity (LRM) and variance (LRV) of 342 peaks obtained by LC-MS are plotted. The horizontal axis represents the LRM, and the vertical axis represents the LRV (see Methods). So if these values are positive, it means that the Syneco sample is larger than the Conv sample in the mean intensity or the variance. The green circles represent PC3 negative loadings in PCA (as shown in Figure 3), and the orange circles represent PC3 positive loadings. As an exception, in the t -test of the intensity of Syneco vs. Conv, a compound with $p < 0.05$ is represented by a triangle instead of a circle. The yellow and light blue hollow circles represent Syneco-distinctive loadings and Conv-distinctive loadings, respectively, which are the top 130 loadings that can completely separate the two culture conditions with PCA (see Figure 3). The saltires represent the compounds with $p < 0.01$ in the F-test of Syneco vs. Conv, and the crosses represent the compounds with $p < 0.05$.

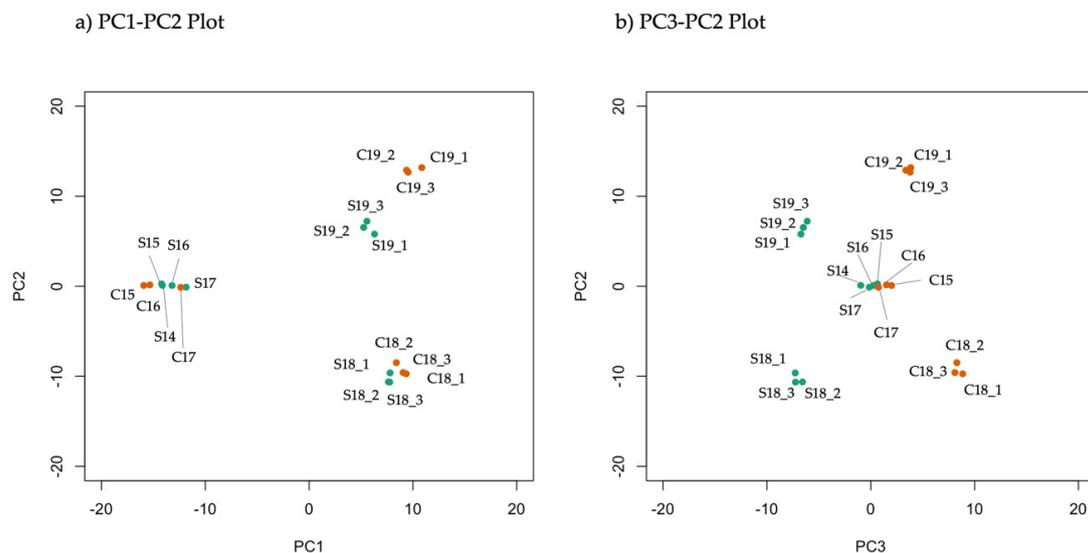


Figure 2. Principal component analysis (PCA) plot based on the intensity of the compounds in coarse green tea samples. (a) PCA plot of PC1-PC2. (b) PCA plot of PC2-PC3. S14, S15, S16, S17, S18, S19, C15, C16, C17, C18, and C19 correspond to the samples Syneco 2014, Syneco 2015, Syneco 2016, Syneco 2017, Syneco 2018, Syneco 2019, Conv 2015, Conv 2016, Conv 2017, Conv 2018, and Conv 2019, respectively. The 2014–2017 samples are based on a single measurement, while the 2018 and 2019 samples consist of 3 different measurements of intensity data.

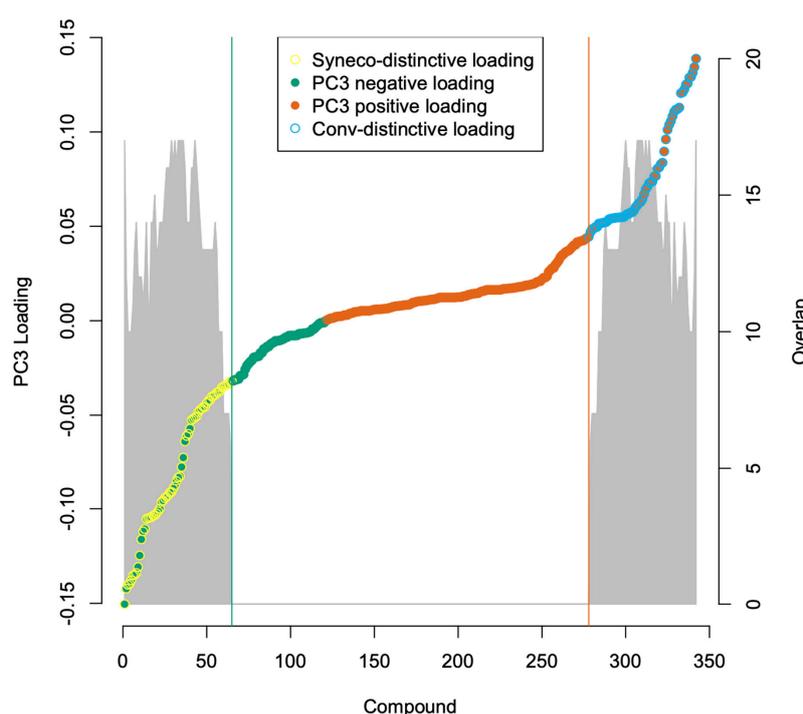


Figure 3. PC3 loading plot of the intensity of the compounds in coarse green tea samples. For negative and positive loadings of PC3 (left Y-axis) aligned in ascending order (X-axis), the 65 smallest and 65 largest loadings of the compounds are separated by 2 vertical lines (green and orange, respectively). The top border of the grey area represents the number of overlaps of Syneco and Conv samples (right Y-axis) when LS-PCA is performed, which means the degree of inseparability between the two culture conditions. When the top 65 negative and positive loadings or more were applied to LS-PCA, the number of overlap is equal to 0, which means the Syneco and Conv samples were completely separated. Hereafter we call these 65 negative and positive loadings “Syneco-distinctive loadings” and “Conv-distinctive loadings” which are plotted with yellow and light blue hollow circles, respectively.

These analyses were performed with 2014–2019 samples, along with the subsets of 2014–2017 and 2018–2019 samples, because LC-MS parameters are different. We call these groupings as three PCA groups 2014–2017, 2018–2019, and 2014–2019. See the PCA plot of 2014–2017 and 2018–2019 samples in Supplementary Materials File S3 for the complete result.

All structural isomers of the 130 distinctive loadings on KEGG were projected onto the “map01110 Biosynthesis of secondary metabolites” of KEGG PATHWAY (see the Supplementary Materials File S17).

In addition, the logarithmic ratio of mean intensity LRM: = $\log(\text{mean intensity of Syneco compounds}) - \log(\text{mean intensity of Conv compounds})$ was calculated for all compounds, and these data were divided into the following three sets in Figure 4.

- D1: Compounds expressed only in Syneco (hereafter Syneco-intrinsic compounds). The mean value of LRM is positive.
- D2: Compounds expressed only in Conv (hereafter Conv-intrinsic compounds). The mean value of LRM is negative.
- D3: Compounds expressed in both Syneco and Conv (hereafter the common compounds). The mean value of LRM can be either positive or negative)

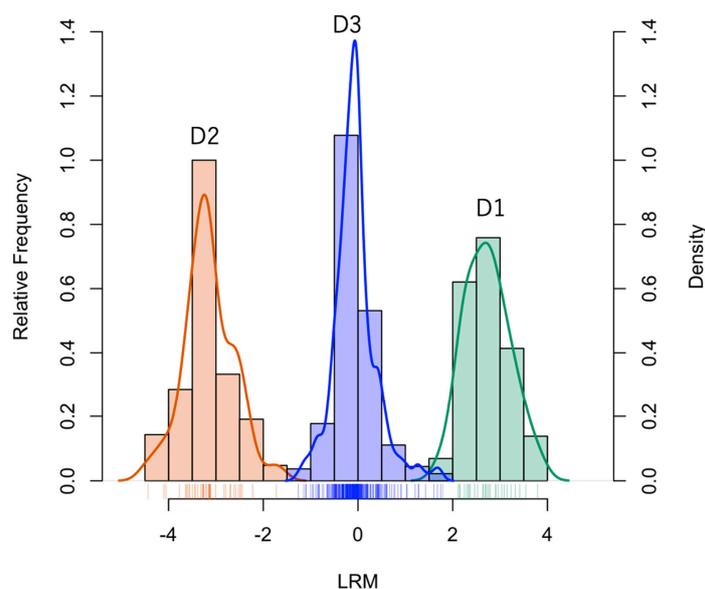


Figure 4. Normalized relative frequency histogram and kernel density estimation of D1 (compounds expressed only in Syneco; green histogram and line), D2 (compounds expressed only in Conv; orange histogram and line), and D3 (compounds expressed in both Syneco and Conv; blue histogram and line). The horizontal axis represents LRM: = $\log(\text{mean intensity of Syneco}/\text{mean intensity of Conv})$. The vertical axes represent normalized relative frequency of the histograms (left) and estimated probability density (right).

The Shapiro-Wilk test was performed on the three distributions D1, D2, and D3, to examine their normality.

Besides LRM, the logarithmic ratio of intensity variance LRV: = $\log(\text{intensity variance of Syneco compounds}) - \log(\text{intensity variance of Conv compounds})$ was calculated for all compounds, and used for the characterization in Figure 1.

2.3. Metabolome Categorization

The list of 130 chemical formulae obtained with LS-PCA was projected to KEGG (Kyoto Encyclopedia of Genes and Genomes) [20,21] databases to annotate possible physiological functions. KEGG API [33] was used to mine the KEGG BRITE [34] database and KEGG PATHWAY [35] in order to categorize the compounds according to the functional classification. Each chemical formula was attributed to the hierarchical ontology of these databases including the matching with structural isomers, as an extensive interpretation of obtained metabolome data on known physiological functions.

Welch's *t*-test and Brunner-Munzel test were performed in each category of KEGG BRITE and KEGG PATHWAY, with Syneco 2014–2019 and Conv 2015–2019 compounds intensity and logarithmic intensity, in order to investigate compound category-wise intensity differences between Syneco and Conv. The mean value of the intensity of the same compound and each value of the intensity were used for each test.

3. Results

3.1. Metabolome Analysis

Exact mass and intensity data of 1055, 815, 714, 1080, and 843 compound peaks were obtained in the Syneco and Conv samples of the years 2015 (including 2014), 2016, 2017, 2018, and 2019, respectively.

3.2. Statistical Analysis

Using the MF searcher, the exact masses of 342 compound peaks were matched with known compounds in the KEGG database (see Supplementary Materials File S4). Among them, the average raw intensity of each 125 compounds was greater for Syneco samples than Conv samples, and 217 compounds for Conv samples than Syneco samples. The result of Welch's *t*-test for the raw intensity showed that only one compound (C19H32O8) was significantly different with a 5% significance level ($p = 0.0143$), but the others were not significant. The result of Brunner-Munzel test for the raw intensity showed that only four compounds (C33H40O21, C19H32O8, C5H11N3O2, C16H18O8) was significantly different with a 5% significance level ($p = 0.0123, 0.00852, 0.0161, 0.0336$, respectively), but the others were not significant.

The result of Welch's *t*-test for the logarithmic intensity showed that only one compound (C5H11N3O2) was significantly different with a 5% significance level ($p = 0.0166$), but the others were not significant. The result of the Brunner-Munzel test for the logarithmic intensity showed that only 4 compounds (C33H40O21, C19H32O8, C5H11N3O2, C16H18O8; the same as for the raw intensity) were significantly different with a 5% significance level ($p = 0.0123, 0.00852, 0.0161, 0.0336$, respectively), but the others were not significant.

The results of F-test between the yearly variances of the samples for each of the 342 compounds showed that 58 compounds were significantly different between the Syneco and Conv samples with the significance level of 5%.

In terms of the variance values representing the yearly fluctuation for each compound, 132 compounds in Syneco were greater than Conv, and 210 compounds in Conv were greater than Syneco. The F-test comparing the variances between Syneco and Conv for the intensity distribution of 342 compounds for all years of sampling showed a statistically significant difference ($p = 7.84 \times 10^{-49}$). Welch's two-sided *t*-test was also performed on the difference between the mean values of the variances, but there was no significant difference ($p = 0.651$).

The result of PCA in Figure 2 revealed that PC3 could linearly separate Syneco and Conv completely (horizontal axis in Figure 2b). PC1 seemed to represent the differences between the two LC-MS analysis conditions that changed between 2014–2017 and 2018–2019 samples (Figure 2a). PC2 mainly represented the yearly fluctuation of all samples that showed three distinctive clusters of 2014–2017, 2018, and 2019 (refer to Supplementary Materials File S2 for the importance and cumulative proportion of the components, and Supplementary Materials File S3 for other PC plots within the subsets 2014–2017 and 2018–2019). In other words, the errors that can occur in our method were aggregated in PC1 and PC2, and PC3 was considered to represent a robust characteristic that separates Syneco and Conv samples.

As shown in Figure 3, The Syneco and Conv samples were completely separated with LS-PCA using the top 65 compounds of PC3 negative/positive loadings (in total 130 compounds). Hereafter we call these top 130 negative and positive loading parameters as Syneco and Conv-distinctive parameters, respectively. We used these 130 distinctive compounds for the ontological categorization using KEGG BRITE and KEGG PATHWAY databases.

Using the MF searcher, 97 flavonoid compounds were listed by matching the metabolome data to the Flavonoid Viewer (see Supplementary Materials File S5). The total intensity of the detected flavonoids in all 2014–2019 samples were compared between Syneco and Conv samples. Although Conv samples tended to contain more flavonoids, it showed no statistically significant difference ($p = 0.786$, two-sided Welch's *t*-test) in total intensity of flavonoids.

With the distribution analysis of LRM, the distribution of Syneco-intrinsic compounds D1 and Conv-intrinsic compounds D2 were regarded as normal distribution with the normality test, and the common compounds D3 was not (p -value of D1: 0.9691, D2: 0.4873, D3: 3.498×10^{-9} , Shapiro-Wilk test). The histograms and estimated probability density of D1, D2, and D3 are drawn in Figure 4. The compounds that were expressed only in Syneco or Conv (D1, D2) showed distinctive offsets of mean values compared to D3 (with Welch's two-sided *t*-test $p = 6.193 \times 10^{-26}$ for D1 > D3; $p = 3.829 \times 10^{-38}$

for $D2 < D3$; Brunner–Munzel test is not applicable because $D1$, $D2$, and $D3$ are completely separated), which implies that $D1$ and $D2$ are based on the qualitative shift of gene expressions compared to the quantitative difference in $D3$.

3.3. Metabolome Categorization

We projected the top 130 (65 positive and 65 negative) of PC3 loadings to KEGG BRITE (Table 2 and Figure 5) and KEGG PATHWAY (Table 3), and looked to the KEGG BRITE compound classification. The negative PC3 loadings that characterized Syneco samples expressed more diversity of allelochemicals than Conv samples such as phytochemicals, alkaloids, phenylpropanoids, and steroids. The “Phytochemicals” include the subcategories of alkaloids, flavonoids, phenylpropanoids, shikimate/acetate-malonate pathway derived compounds, terpenoids, polyketides, fatty acids related compounds, amino acid related compounds, and others, according to the notation in KEGG BRITE database.

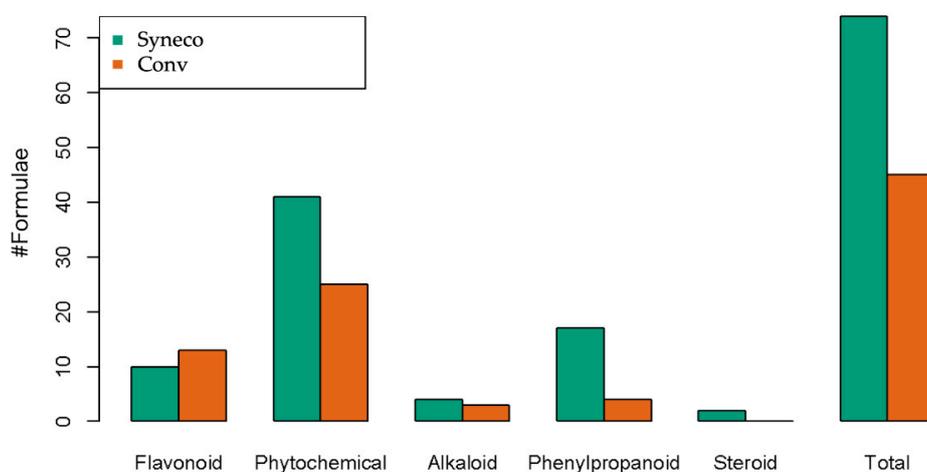


Figure 5. Graphical representation of Table 2.

Among the three PCA groups 2014–2017, 2018–2019, and 2014–2019, we identified common chemical formulae that were sufficient for the separation of Syneco and Conv samples with LS-PCA (see the Supplementary Materials Files S11–S16). Among the three groups, “C13H20O2”, “C14H21NO8”, “C27H30O16” and “C11H16O2” were the common formulae in Syneco samples. Through the metabolome categorization with KEGG databases, these are considered to be heptyloxyphenol, glucosylpyridoxine, rutin, and methylcatechol. Moreover, “C44H34O22”, “C9H11NO3”, “C5H11N3O2”, “C14H16O10”, “C21H20O10”, “C16H18O8”, “C9H11NO2” and “C14H20O3” were detected as common formulae in Conv samples. These were considered to be theasinensin A, l-tyrosine, guanidinobutyric acid, theogallin, isovitexin, coumaroylquinic acid, l-phenylalanine and heptylparaben.

The common 4 compounds of Syneco tended to have larger variance values in Syneco than Conv samples (F-test, $p = 0.0795$), while the common 8 compounds of Conv had significantly larger variance values in Conv than Syneco samples (F-test, $p = 2.576 \times 10^{-13}$).

In addition, epigallocatechin gallate, aromadendrin, pesticide compounds (framprop, aldicarb or butocarboxim), and amino acids that might potentially be derived from fertilizers were detected as Conv-distinctive compounds.

The two-sided Welch’s *t*-test and Brunner–Munzel test were performed in each category of KEGG BRITE and KEGG PATHWAY (KEGG category-wise tests), with the use of total intensity and total logarithmic intensity of Syneco2014–2019 and Conv2015–2019 samples. The mean value of intensity of the same chemical formula and each value of the intensity were used for the tests.

Table 3. KEGG PATHWAY compound classification. The numbers of different chemical formulae (# Formulae) associated with the Syneco- and Conv-distinctive loadings for each category of chemical ontology in KEGG PATHWAY hierarchy were shown. The pathway related to microbial metabolism is typed in red. See the Supplementary Materials Files S6, S7, S9 and S10 for more details.

		Culture Condition	Syneco		Conv	
Category			# Formulae	Uncertainty Score	# Formulae	Uncertainty Score
All Categories	1. Metabolism		152	39.32423687	226	56.47841953
		1.0 Global and overview maps	147	38.19923687	168	42.42704868
		map01100 Metabolic pathways	57	15.5296398	56	17.70079175
		map01110 Biosynthesis of secondary metabolites	18	5.601724664	20	8.437734488
		map01120 Microbial metabolism in diverse environments	21	5.025534188	17	5.276477178
		map01200 Carbon metabolism	9	2.84702381	5	1.382539683
		map01210 2-Oxocarboxylic acid metabolism	1	1	1	0.071428571
		map01230 Biosynthesis of amino acids	2	0.333333333	6	1.166305916
		map01220 Degradation of aromatic compounds	1	0.166666667	6	1.166305916
			5	0.555357143	1	0.2
		1.1 Carbohydrate metabolism	0	0	4	0.285714286
		map00020 Citrate cycle (TCA cycle)	0	0	1	0.071428571
		map00040 Pentose and glucuronate interconversions	0	0	1	0.071428571
		map00053 Ascorbate and aldarate metabolism	0	0	1	0.071428571
		map00630 Glyoxylate and dicarboxylate metabolism	0	0	1	0.071428571
		1.2 Energy metabolism	2	1.125	2	0.182539683
		map00720 Carbon fixation pathways in prokaryotes	1	1	1	0.071428571
		map00680 Methane metabolism	1	0.125	1	0.111111111
		1.3 Lipid metabolism	2	1.5	1	0.2
		map00061 Fatty acid biosynthesis	1	0.5	0	0
		map00073 Cutin, suberine and wax biosynthesis	0	0	1	0.2
		map00140 Steroid hormone biosynthesis	1	1	0	0
		1.4 Nucleotide metabolism	0	0	2	0.666666667
		map00230 Purine metabolism	0	0	2	0.666666667
		1.5 Amino acid metabolism	11	1.476190476	18	6.015151515
		map00250 Alanine, aspartate and glutamate metabolism	0	0	1	0.071428571
		map00260 Glycine, serine and threonine metabolism	0	0	1	0.142857143
		map00270 Cysteine and methionine metabolism	0	0	1	1
		map00310 Lysine degradation	0	0	1	1
		map00220 Arginine biosynthesis	0	0	2	0.75
		map00330 Arginine and proline metabolism	0	0	2	1.25
		map00350 Tyrosine metabolism	5	0.580357143	2	0.611111111
		map00360 Phenylalanine metabolism	3	0.3125	2	0.202020202
map00380 Tryptophan metabolism	1	0.25	3	0.642857143		
map00400 Phenylalanine, tyrosine and tryptophan biosynthesis	2	0.333333333	3	0.344877345		
1.6 Metabolism of other amino acids	0	0	4	1.702020202		
map00410 beta-Alanine metabolism	0	0	1	1		
map00440 Phosphonate and phosphinate metabolism	0	0	1	0.5		
map00460 Cyanoamino acid metabolism	0	0	2	0.202020202		

Table 3. Cont.

Culture Condition		Syneco		Conv			
Category		# Formulae	Uncertainty Score	# Formulae	Uncertainty Score		
1.8 Metabolism of cofactors and vitamins	map00730	Thiamine metabolism	4	2.666666667	3	1.222222222	
	map00770	Pantothenate and CoA biosynthesis	0	0	1	0.111111111	
	map00785	Lipoic acid metabolism	0	0	1	1	
	map00790	Folate biosynthesis	1	0.5	0	0	
	map00670	One carbon pool by folate	1	1	0	0	
	map00130	Ubiquinone and other terpenoid-quinone biosynthesis	1	1	0	0	
1.9 Metabolism of terpenoids and polyketides			8	1.825213675	12	3.258363712	
	map00900	Terpenoid backbone biosynthesis	0	0	1	0.052631579	
	map00902	Monoterpenoid biosynthesis	2	1.076923077	0	0	
	map00909	Sesquiterpenoid and triterpenoid biosynthesis	1	0.011111111	2	0.14354067	
	map00904	Diterpenoid biosynthesis	0	0	1	0.125	
	map00981	Insect hormone biosynthesis	1	0.166666667	2	0.14354067	
	map00908	Zeatin biosynthesis	0	0	1	1	
	map00903	Limonene and pinene degradation	1	0.076923077	0	0	
	map00281	Geraniol degradation	1	0.076923077	0	0	
	map01059	Biosynthesis of enediyne antibiotics	1	0.25	1	0.111111111	
	map01057	Biosynthesis of type II polyketide products	0	0	2	1.5	
	map01053	Biosynthesis of siderophore group nonribosomal peptides	0	0	1	0.071428571	
	map01055	Biosynthesis of vancomycin group antibiotics	1	0.166666667	1	0.111111111	
	1.10 Biosynthesis of other secondary metabolites			33	9.054700855	36	6.420779221
		map00232	Caffeine metabolism	1	1	0	0
map00333		Prodigiosin biosynthesis	1	0.076923077	0	0	
map00940		Phenylpropanoid biosynthesis	8	1.2875	4	0.785353535	
map00945		Stilbenoid, diarylheptanoid and gingerol biosynthesis	0	0	2	0.583333333	
map00941		Flavonoid biosynthesis	5	1.625	6	1.242063492	
map00944		Flavone and flavonol biosynthesis	5	2.45	3	0.485714286	
map00942		Anthocyanin biosynthesis	0	0	1	0.5	
map00943		Isoflavonoid biosynthesis	1	0.125	3	0.325396825	
map00901		Indole alkaloid biosynthesis	0	0	1	0.142857143	
map00950		Isoquinoline alkaloid biosynthesis	3	0.354166667	1	0.111111111	
map00960		Tropane, piperidine and pyridine alkaloid biosynthesis	0	0	1	0.090909091	
map00232		Caffeine metabolism	1	1	0	0	
map00965		Betalain biosynthesis	0	0	1	0.111111111	
map00966		Glucosinolate biosynthesis	0	0	3	0.344877345	
map00332		Carbapenem biosynthesis	0	0	1	0.25	
map00261		Monobactam biosynthesis	1	0.166666667	1	0.111111111	
map00401		Novobiocin biosynthesis	2	0.416666667	1	0.111111111	
map00404		Staurosporine biosynthesis	0	0	1	0.142857143	
map00999		Biosynthesis of various secondary metabolites—part 1	1	0.011111111	0	0	
map00998		Biosynthesis of various secondary metabolites—part 2	4	0.541666667	5	1.011544012	
map00997		Biosynthesis of various secondary metabolites—part 3	0	0	1	0.071428571	

Table 3. Cont.

Culture Condition		Syneco		Conv			
Category		# Formulae	Uncertainty Score	# Formulae	Uncertainty Score		
1.11 Xenobiotics biodegradation and metabolism	map00627	Aminobenzoate degradation	0	0	1	0.5	
	map00623	Toluene degradation	1	0.0625	0	0	
	map00622	Xylene degradation	3	0.305357143	0	0	
	map00633	Nitrotoluene degradation	1	0.166666667	0	0	
	map00642	Ethylbenzene degradation	2	0.205357143	0	0	
	map00643	Styrene degradation	1	0.0625	0	0	
	map00363	Bisphenol degradation	1	0.0625	0	0	
	map00626	Naphthalene degradation	3	0.354166667	1	0.2	
	map00624	Polycyclic aromatic hydrocarbon degradation	1	0.083333333	2	0.311111111	
	map00980	Metabolism of xenobiotics by cytochrome P450	0	0	2	0.342857143	
	map00982	Drug metabolism—cytochrome P450	1	0.5	1	0.166666667	
	1.12 Chemical structure transformation maps		16	3.219444441	23	3.252164502	
		map01060	Biosynthesis of plant secondary metabolites	4	0.854166667	5	1.416305916
		map01061	Biosynthesis of phenylpropanoids	6	1.416666667	6	0.737734488
		map01062	Biosynthesis of terpenoids and steroids	1	0.011111111	1	0.071428571
		map01063	Biosynthesis of alkaloids derived from shikimate pathway	3	0.354166667	4	0.416305916
		map01064	Biosynthesis of alkaloids derived from ornithine, lysine and nicotinic acid	1	0.083333333	2	0.162337662
		map01065	Biosynthesis of alkaloids derived from histidine and purine	0	0	1	0.071428571
		map01066	Biosynthesis of alkaloids derived from terpenoid and polyketide	1	0.5	1	0.071428571
		map01070	Biosynthesis of plant hormones	0	0	3	0.305194805
2. Genetic Information Processing		0	0	3	0.344877345		
	2.2 Translation		0	0	3	0.344877345	
	map00970	Aminoacyl-tRNA biosynthesis	0	0	3	0.344877345	
3. Environmental Information Processing		2	0.208333333	9	2.662337662		
	3.1 Membrane transport		0	0	3	0.924242424	
	map02010	ABC transporters	0	0	3	0.924242424	
	3.2 Signal transduction		1	0.083333333	5	1.404761905	
	map02020	Two-component system	0	0	1	0.071428571	
	map04071	Sphingolipid signaling pathway	0	0	1	0.333333333	
	map04024	cAMP signaling pathway	0	0	1	0.333333333	
	map04022	cGMP-PKG signaling pathway	0	0	2	0.666666667	
	map04152	AMPK signaling pathway	1	0.083333333	0	0	
	3.3 Signaling molecules and interaction		1	0.125	1	0.333333333	
	map04080	Neuroactive ligand-receptor interaction	1	0.125	1	0.333333333	
4. Cellular Processes		0	0	1	0.333333333		
	4.3 Cellular community—eukaryotes		0	0	1	0.333333333	
	map04540	Gap junction	0	0	1	0.333333333	
5. Organismal Systems		3	0.791666667	31	8.173881674		
	5.1 Immune system		0	0	1	0.333333333	
	map04611	Platelet activation	0	0	1	0.333333333	
	5.2 Endocrine system		0	0	9	2.293650794	
	map04922	Glucagon signaling pathway	0	0	1	0.071428571	
	map04923	Regulation of lipolysis in adipocytes	0	0	2	0.666666667	
	map04917	Prolactin signaling pathway	0	0	1	0.111111111	
	map04921	Oxytocin signaling pathway	0	0	1	0.333333333	
	map04916	Melanogenesis	0	0	1	0.111111111	
	map04924	Renin secretion	0	0	2	0.666666667	
	map04925	Aldosterone synthesis and secretion	0	0	1	0.333333333	

Table 3. Cont.

Culture Condition		Syneco	Conv		
Category		# Formulae	Uncertainty Score	# Formulae	Uncertainty Score
5.3 Circulatory system		0	0	2	0.66666667
	map04270 Vascular smooth muscle contraction	0	0	2	0.66666667
5.4 Digestive system		3	0.79166667	8	2.245310245
	map04970 Salivary secretion	0	0	1	0.33333333
	map04976 Bile secretion	2	0.66666667	1	0.33333333
	map04974 Protein digestion and absorption	1	0.125	3	0.344877345
	map04977 Vitamin digestion and absorption	0	0	1	1
	map04978 Mineral absorption	0	0	2	0.233766234
5.6 Nervous system		0	0	4	0.753968254
	map04728 Dopaminergic synapse	0	0	1	0.11111111
	map04726 Serotonergic synapse	0	0	2	0.30952381
	map04730 Long-term depression	0	0	1	0.33333333
5.7 Sensory system		0	0	4	1.071428571
	map04744 Phototransduction	0	0	1	0.33333333
	map04744 Phototransduction—fly	0	0	1	0.33333333
	map04740 Olfactory transduction	0	0	1	0.33333333
	map04742 Taste transduction	0	0	1	0.071428571
5.8 Development and regeneration		0	0	1	0.142857143
	map04361 Axon regeneration	0	0	1	0.142857143
5.10 Environmental adaptation		0	0	2	0.66666667
	map04713 Circadian entrainment	0	0	1	0.33333333
	map04714 Thermogenesis	0	0	1	0.33333333
6. Human Diseases		0	0	14	2.536940837
6.1 Cancer: overview		0	0	5	0.616305916
	map05204 Chemical carcinogenesis	0	0	1	0.2
	map05230 Central carbon metabolism in cancer	0	0	4	0.416305916
6.4 Neurodegenerative disease		0	0	2	0.44444444
	map05012 Parkinson disease	0	0	2	0.44444444
6.5 Substance dependence		0	0	5	1
	map05030 Cocaine addiction	0	0	1	0.11111111
	map05031 Amphetamine addiction	0	0	1	0.11111111
	map05032 Morphine addiction	0	0	1	0.33333333
	map05034 Alcoholism	0	0	2	0.44444444
6.10 Infectious disease: parasitic		0	0	2	0.476190476
	map05143 African trypanosomiasis	0	0	2	0.476190476

The categories of “1. Metabolism,” “1.4 Nucleotide metabolism,” “1.5 Amino acid metabolism,” “1.6 Metabolism of other amino acids,” “Abietanes,” “Drugs with new active ingredients,” “Guaianolide,” “M MUSCULO-SKELETAL SYSTEM,” “M01 ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS,” “M01A ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS, NON-STEROIDS,” “map00300 Lysine biosynthesis,” “map00310 Lysine degradation,” “map00330 Arginine and proline metabolism,” “map00830 Retinol metabolism,” “map01100 Metabolic pathways,” “N-glycosides,” “New drug approvals in Japan [br08318],” “PR0109 Retinoids,” and “Sinapate derivatives,” were significantly different ($p < 0.05$) in at least one of the tests (Tables 4 and 5). Notably, the results showed that Conv samples expressed significantly greater total intensity in the categories of amino acids- and nucleotide- related primary metabolites.

Table 4. Results of KEGG PATHWAY category-wise tests with $p < 0.05$: The difference of intensity in each KEGG PATHWAY category (column “Category in KEGG PATHWAY”) is tested with two-sided Welch’s *t*-test and Brunner-Munzel test (“Test”). The signs of *p*-values (in “*p*-Value”) represent the magnitude relationship of mean intensity between Syneco and Conv (“Magnitude Relationship”); negative sign means Conv was greater than Syneco, and positive sign means Syneco was greater than Conv. The column “Scale” indicates whether the intensity was used in linear or logarithmic scale for the tests. The intensity value was used both as chemical formula-wise aggregated mean and as separated intensity peaks (indicated as Formula and NA, respectively, in the column “Averaging”). The column “# Formulae” represents the number of chemical formulae estimated in each category.

Category in KEGG PATHWAY	# Formulae	Magnitude Relationship	Scale	Test	Averaging	<i>p</i> -Value	
1. Metabolism	199	Syneco < Conv	Logarithmic	Brunner-Munzel	Formula	−0.047539	
1.4 Nucleotide metabolism	6	Syneco < Conv	Linear	Brunner-Munzel	Formula	−0.047815	
1.5 Amino acid metabolism	56	Syneco < Conv	Linear	Brunner-Munzel	Formula	−0.017334	
				Brunner-Munzel	NA	−0.017518	
				Welch	NA	−0.018299	
			Logarithmic	Brunner-Munzel	Formula	−0.010095	
				Brunner-Munzel	NA	−0.017518	
				Welch	NA	−0.007608	
	map00300 Lysine biosynthesis	7	Syneco < Conv	Linear	Brunner-Munzel	Formula	−0.021277
	map00310 Lysine degradation	7	Syneco < Conv	Linear	Brunner-Munzel	Formula	−0.024978
					Brunner-Munzel	NA	−0.014387
					Welch	Formula	−0.016583
				Logarithmic	Welch	NA	−0.01523
					Brunner-Munzel	Formula	−0.000144
Brunner-Munzel					NA	−0.014387	
map00330 Arginine and proline metabolism	8	Syneco < Conv	Linear	Brunner-Munzel	Formula	−0.035047	
			Linear	Brunner-Munzel	NA	−0.026784	
			Logarithmic	Welch	NA	−0.026253	
1.6 Metabolism of other amino acids	17	Syneco < Conv	Linear	Brunner-Munzel	NA	−0.026784	
1.8 Metabolism of cofactors and vitamins	map00830 Retinol metabolism	2	Syneco < Conv	Linear	Welch	Formula	−0.004966
				Logarithmic	Welch	Formula	−0.000541
1.0 Global and overview maps	map01100 Metabolic pathways	127	Syneco < Conv	Logarithmic	Brunner-Munzel	Formula	−0.04356

Table 5. Results of KEGG BRITE category-wise tests with $p < 0.05$: The difference of intensity in each KEGG BRITE category (column “Category in KEGG BRITE”) is tested with two-sided Welch’s t -test and Brunner-Munzel test (“Test”). Other notations follow those of Table 4.

Category in KEGG BRITE						#Formulae	Magnitude Relationship	Scale	Test	Averaging	p-Value		
Compounds and Reactions	Compounds (C numbers)	Phytochemical compounds [BR:br08003]	Terpenoids	Diterpenoids (C20)	Abietanes	2	Syneco < Conv	Linear	Welch	Formula	-0.004966483		
				Sesquiterpenoids (C15)	Guaianolide	3		Logarithmic	Welch	Formula	-0.000540986		
			Phenylpropanoids	Monolignols	Sinapate derivatives	2	Syneco > Conv	Linear	Welch	Formula	0.02822976		
						2	Syneco < Conv	Logarithmic	Welch	Formula	-0.032450817		
		Glycosides [BR:br08021]	N-glycosides			3	Syneco < Conv	Linear	Welch	Formula	-0.042604306		
		Lipids [BR:br08002]	PR Prenol lipids	PR01 Isoprenoids	PR0109 Retinoids	2	Syneco < Conv	Linear	Welch	Formula	-0.004966483		
								Logarithmic	Welch	Formula	-0.000540986		
		Drugs	Drug information (D numbers)	New drug approvals in Japan [br08318]				4	Syneco < Conv	Linear	Brunner-Munzel	Formula	-0.002578598
												NA	-0.019485081
										Formula	-0.014647994		
NA	-0.026673591												
Drugs with new active ingredients	4			Syneco < Conv	Linear	Brunner-Munzel	Formula	-0.002578598					
							NA	-0.019485081					
					Formula	-0.014647994							
					NA	-0.026673591							
Drug classifications (D numbers)	M MUSCULO-SKELETAL SYSTEM				2	Syneco > Conv	Linear	Welch	NA	0.049488163			
							Logarithmic	Welch	Formula	0.022529231			
	M01 ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS				2	Syneco > Conv	Linear	Welch	NA	0.049488163			
							Logarithmic	Welch	Formula	0.022529231			
	M01A ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS, NON-STEROIDS				2	Syneco > Conv	Linear	Welch	NA	0.049488163			
							Logarithmic	Welch	Formula	0.022529231			

4. Discussion

In the pre-analysis with a Welch's *t*-test and Brunner–Munzel test for each compound, only one and four over 342 detected compound peaks had significant differences between Syneco and Conv samples, respectively, concerning the 5% threshold on the *p*-value. Therefore, the simple comparison of compound-wise intensity does not support any statistically significant differences, since about 17 peaks (5% of 342) could stochastically take *p*-values under the 5% threshold in random data. Considering also biological fluctuations and technical errors, the results of compound-wise statistical tests did not seem relevant enough to require further investigation.

However, the variation between different years for each compound measured with F-test was significantly different between the two categories (for 58 compounds, about 17% of 342), suggesting that the *in cultura* samples had larger yearly variance. This result supports the relative consistency of metabolic profile in ecological optimum compared to the larger deviation in monoculture data, as pointed out in a previous study [2]. It means that the human intervention under the *in cultura* culture conditions introduces larger fluctuations in the metabolic state of the crop than the *in natura* self-organized growth under ecological optimum conditions. The observed facts may be relevant to the general relationship between biodiversity and ecological resilience: the *in cultura* culture condition with low biodiversity may behave less stable in the consistency of its metabolic profile because of insufficient ecological feedbacks that are abundant in the *in natura* culture condition.

Previous study has pointed out that there exist statistical invariant features of metabolic profiles that distinguished between the *in cultura* and *in natura* culture conditions [2]. The principal component analysis in this study, which reflected the entire LC-MS measurements over six years of repeated production, also supports this notion. The distinctive metabolic profiles were found with respect to the PC3 loading section, which was classified with the ontologies in KEGG BRITE and PATHWAY databases.

Syneco samples produced in the *in natura* condition contained a larger number of allelochemicals such as phytochemicals, alkaloids, phenylpropanoids, and steroids, according to the classification in KEGG BRITE (Table 2), especially for the compounds with top 130 PC3 loadings that provided complete linear separation between the *in cultura* and *in natura* samples (Figure 2b).

As for the KEGG PATHWAY classification of the top 130 PC3 loadings (Table 3), 9 compounds of *in natura* samples and five compounds of *in cultura* samples were categorized into “map01120 Microbial metabolism in diverse environments” (see the Supplementary Materials Files S9 and S10). This category is related to the metabolism of microorganisms, which implies that the *in natura* samples were raised in more complex microbiological interactions than *in cultura* samples. In this study, the microbiological interaction of tea plants should mainly come from soil microbiota. Indeed, it is reported that Synecoculture largely promotes soil microbial diversity and activities [36]. These results suggest that the *in natura* samples tend to contain more diverse compounds related to the interaction with soil microorganisms.

Plants are known to synthesize repellent and attractant substances called allelochemicals in competitive and cooperative interactions with other plants and insects [37]. Allelochemicals such as alkaloids, phenylpropanoids, steroids, and flavonoids, may be harmful to the human body in excessive amounts, but are known to have anti-inflammatory, anti-cancer, and antioxidant effects with appropriate dose [38]. Our results suggest that the *in natura* culture condition associated with rich interactions between species could support a more diverse production of such health-protective compounds.

Concerning the diversity of allelochemicals, the only exception in this study was the total intensity of flavonoids that was larger in Conv than in Syneco in the total comparison of 2014–2019 samples. This result was different from the previous study that reported larger flavonoids expression in Syneco 2014–2015 samples [39]. Actually, the total amount of flavonoids was superior in Syneco 2014–2016 to Conv 2015–2016 samples (results not shown). For the Syneco 2017–2019 and Conv 2017–2019 samples, the total amount of flavonoids was not effective as a distinctive parameter that constantly separates the two culture conditions. Since flavonoids can be enhanced not only by ecological interactions but also

by physical environmental stress such as sunlight, various flavonoids may have been produced in response to the environmental particularity in Syneco 2014–2015 samples.

Among the three PCA groups (2014–2017, 2018–2019, and 2014–2019), compounds such as glucosylpyridoxine, rutin, heptyloxyphenol, and methylcatechol were mutually estimated in Syneco samples. Glucosylpyridoxine, a glycoside of vitamin B6, was detected as an exclusively characteristic compound of *in natura* samples. Furthermore, rutin is known as an allelopathic chemical that improves plant disease resistances and inhibits the growth of other competing plants [40,41]. These *in natura* distinctive compounds can be interpreted as the results of the enhanced ecological interactions in Synecoculture.

Vitamin B6 is known to be synthesized by soil microorganisms of the genus *Aspergillus*. It is required for controlling the immune response, and its deficiency is known to cause immune system disorders and a decrease in antibody production [42,43]. In recent years, the diversity of human gut microbiota in city environments has been reduced due to the abuse of antibiotics and pesticides [44–46], and non-infectious immune-related diseases have become serious in many countries around the world. Also, the risk of vitamin B2 and B6 deficiency has been reported in vegan populations [47]. Vitamin B6 and B12 have been reported to be effective in treating Alzheimer's disease [48–50]. Our results suggest that vitamin B6 is contained more in the tea samples cultivated under the *in natura* culture condition in which the topsoil ecosystem is not disturbed.

On the other hand, theasinensin A, l-tyrosine, guanidinobutyric acid, theogallin, isovitexin, coumaroylquinic acid, l-phenylalanine and heptylparaben were detected as the characteristic compounds common to conventional samples. Amino acids were detected as a strong characteristic of *in cultura* samples, which may be derived from the application of synthetic fertilizers. This is in line with the previous study that also reported amino acids were one of the characteristics of the conventionally cultured tea product [39]. The results of the KEGG category-wise tests in Tables 4 and 5 also support this notion.

The overall results suggest that the distinction between *in natura* and *in cultura* conditions only becomes possible at the distribution level of metabolome, beyond single-component comparison: in the results of the normality test for the distribution of the intrinsic and common compounds for each culture condition (Figure 4), both *in cultura* and *in natura* intrinsic compounds showed normality in the distribution of intensity, with distinctive offsets of mean values. Each culture condition thus comprised the qualitatively different expression of culture-specific compounds.

5. Conclusions

Culture conditions strongly relate to the health of both environment and humans through our diet as an interface. In this article, the results suggest that the tea plants grown in *in natura* condition produce more diverse and abundant allelochemicals than in *in cultura* conditions, which is expected to contribute to the health of the plants themselves and those consumers. Such enhancement of secondary metabolite in tea plant may be considered as an example of the generic interaction between crops and field biodiversity, especially soil microbiota.

In the formation of ecological niches, it is known that some species specifically rely on the symbiosis with other species and does not tolerate single isolated culture (e.g., [51]). This implies the presence of the “*in natura* effects” that maintain the coexistence of various species in natural environments and associated particular expression patterns of metabolite, which occurs only within the complex interactions of the self-organized plant and animal communities. Similarly, our results suggest that there exists a metabolite-level property that is particularly based on the *in natura* state of culture condition, which has been generally treated as the irreplaceable support of the whole food quality that cannot be produced with synthetic nutritional supplements. Generally, the health effects of plant

products cannot be totally evaluated with a single component, but should be considered within the whole diet composed of many compounds acting synergistically on human health [52,53]. The intrinsic distribution of compounds particular to *in natura* metabolite should further be considered when evaluating the sustainability of food systems toward the resolution of the health-diet-environment trilemma [3].

Supplementary Materials: The following are available online at <http://www.mdpi.com/2077-0472/10/12/632/s1>, Supplementary Material File S1: Parameters of LC-MS analysis 2014–2017 and 2018–2019; Supplementary Material File S2: Importance and Cumulative Proportion of the Components of PCA, and other PC plots; Supplementary Material File S3: PCA plot of 2014–2017 samples and 2018–2019 samples; Supplementary Material File S4: Raw Intensity Data 2014–2019 matched KEGG database; Supplementary Material File S5: Raw Intensity Data 2014–2019 matched Flavonoid Viewer; Supplementary Material File S6: Negative Loadings of PC3 categorized in KEGG PATHWAY and BRITE; Supplementary Material File S7: Positive Loadings of PC3 categorized in KEGG PATHWAY and BRITE; Supplementary Material File S8: KEGG BRITE compound classification in the other hierarchy of BRITE; Supplementary Material File S9: All structural isomers of Syneco-distinctive loadings projected on “map01120 Microbial metabolism in diverse environments” in KEGG PATHWAY; Supplementary Material File S10: All structural isomers of Conv-distinctive loadings projected on “map01120 Microbial metabolism in diverse environments” in KEGG PATHWAY; Supplementary Material File S11: List of chemical formulae with Syneco-distinctive loadings that can completely separate Syneco and Conv samples in 2014–2017 samples; Supplementary Material File S12: List of chemical formulae with Conv-distinctive loadings that can completely separate Syneco and Conv samples in 2014–2017 samples; Supplementary Material File S13: List of chemical formulae with Syneco-distinctive loadings that can completely separate Syneco and Conv samples in 2018–2019 samples; Supplementary Material File S14: List of chemical formulae with Conv-distinctive loadings that can completely separate Syneco and Conv samples in 2018–2019 samples; Supplementary Material File S15: List of chemical formulae with Syneco-distinctive loadings that can completely separate Syneco and Conv samples in 2014–2019 samples; Supplementary Material File S16: List of chemical formulae with Conv-distinctive loadings that can completely separate Syneco and Conv samples in 2014–2019 samples; Supplementary Material File S17: All structural isomers of Syneco-/Conv- distinctive loadings projected on “map01110 Biosynthesis of secondary metabolites” in KEGG PATHWAY.

Author Contributions: K.O. prepared the samples and obtained the intensity data, and carried out the statistical tests including PCA. K.O. and M.F. wrote the manuscript. T.K. carried out PCA and constructed the list of compounds that contribute to the separation of Syneco/Conv samples, and obtained the information of characteristics of the compounds. M.F. directed this experiment and gave essential insights. All authors have read and agreed to the published version of the manuscript.

Funding: Sony Computer Science Laboratories, Inc. funded the research.

Acknowledgments: Authors acknowledge the supports of Shoji Oda for the extraction of the samples, Tomoyuki Minami for the integration of peak intensity data, Godai Suzuki and Kosuke Katano for the literature searching, Kae Miyazawa for comments from the viewpoints of agronomy, editors and anonymous reviewers for giving critical insights, Arnold Ross Falvo, Jordan White Springer and Yuji Kawamura for the revision of manuscript and Sakurashizenjyuku Global Nature Network for the provision of the tea products.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tilman, D.; Clark, M. Global diets link environmental sustainability and human health. *Nature* **2014**, *515*, 518–522. [[CrossRef](#)] [[PubMed](#)]
2. Funabashi, M. Food components as markers linking health and environment: Statistical invariance analysis of *in natura* diet. *Am. J. Biosci. Bioeng.* **2015**, *3*, 183–196.
3. Funabashi, M. Human augmentation of ecosystems: Objectives for food production and science by 2045. *NPJ Sci. Food* **2018**, *2*, 1–11. [[CrossRef](#)] [[PubMed](#)]
4. Funabashi, M. Synecological Farming: Theoretical Foundation on Biodiversity Responses of Plant Communities. *Plant Biotechnol.* **2016**, *33*, 213–234. [[CrossRef](#)]
5. Barnosky, A.D.; Matzke, N.; Tomiya, S.; Wogan, G.O.; Swartz, B.; Quental, T.B.; Marshall, C.; McGuire, J.L.; Lindsey, E.L.; Maguire, K.C.; et al. Has the Earth’s sixth mass extinction already arrived? *Nature* **2011**, *471*, 51–57. [[CrossRef](#)]
6. Ceballos, G.; Ehrlich, P.R.; Raven, P.H. Vertebrates on the brink as indicators of biological annihilation and the sixth mass extinction. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 13596–13602. [[CrossRef](#)]

7. Barnosky, A.D.; Hadly, E.A.; Bascompte, J.; Berlow, E.L.; Brown, J.H.; Fortelius, M.; Getz, W.M.; Harte, J.; Hastings, A.; Marquet, P.A.; et al. Approaching a state shift in Earth's biosphere. *Nature* **2012**, *486*, 52–58. [CrossRef]
8. WHO. *Diet, Nutrition and the Prevention of Chronic Diseases*; World Health Organization: Geneva, Switzerland, 2002.
9. Tokoro, M. Open Systems Science: A Challenge to Open Systems Problems. In *First Complex Systems Digital Campus World E-Conference 2015*; Bourguine, P., Collet, P., Parrend, P., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 213–221.
10. Huang, R.; McGrath, S.P.; Hirsch, P.R.; Clark, I.M.; Storkey, J.; Wu, L.; Zhou, J.; Liang, Y. Plant–microbe networks in soil are weakened by century-long use of inorganic fertilizers. *Microb. Biotechnol.* **2019**, *12*, 1464–1475. [CrossRef]
11. Power, A.G. Ecosystem services and agriculture: Tradeoffs and synergies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2010**, *365*, 2959–2971. [CrossRef]
12. Millennium Ecosystem Assessment. *Ecosystems and human well-being-Synthesis: A report of the Millennium Ecosystem Assessment*; Island Press: Washington, DC, USA, 2005.
13. Daszak, P.; das Neves, C.; Amuasi, J.; Hayman, D.; Kuiken, T.; Roche, B.; Zambrana-Torrel, C.; Buss, P.; Dundarova, H.; Feferholtz, Y.; et al. *Workshop Report on Biodiversity and Pandemics of the Intergovernmental Platform on Biodiversity and Ecosystem Services*; IPBES Secretariat: Bonn, Germany, 2020. [CrossRef]
14. Dobson, A.P.; Pimm, S.L.; Hannah, L.; Kaufman, L.; Ahumada, J.A.; Ando, A.W.; Bernstein, A.; Busch, J.; Daszak, P.; Engelmann, J.; et al. Ecology and economics for pandemic prevention. *Science* **2020**, *369*, 379–381.
15. Morand, S. Emerging diseases, livestock expansion and biodiversity loss are positively related at global scale. *Biol. Conserv.* **2020**, *248*, 108707. [CrossRef] [PubMed]
16. Roossinck, M.J.; García-Arenal, F. Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr. Opin. Virol.* **2015**, *10*, 56–62. [CrossRef] [PubMed]
17. Sacristán, S.; Fraile, A.; García-Arenal, F. Population Dynamics of Cucumber mosaic virus in Melon Crops and in Weeds in Central Spain. *Phytopathology* **2004**, *94*, 992–998. [CrossRef] [PubMed]
18. Pongsiri, M.J.; Roman, J.; Ezenwa, V.O.; Goldberg, T.L.; Koren, H.S.; Newbold, S.C.; Ostfeld, R.S.; Pattanayak, S.K.; Salkeld, D.J. Biodiversity Loss Affects Global Disease Ecology. *BioScience* **2009**, *59*, 945–954. [CrossRef]
19. Keesing, F.; Belden, L.K.; Daszak, P.; Dobson, A.; Harvell, C.D.; Holt, R.D.; Hudson, P.; Jolles, A.; Jones, K.E.; Mitchell, C.E.; et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **2010**, *468*, 647–652. [CrossRef] [PubMed]
20. Kanehisa, M.; Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [CrossRef]
21. KEGG. Available online: <https://www.genome.jp/kegg/> (accessed on 29 March 2020).
22. Funabashi, M.; Hanappe, P.; Isozaki, T.; Maes, A.; Sasaki, T.; Steels, L.; Yoshida, K. Foundation of CS-DC e-laboratory: Open Systems Exploration for Ecosystems Leveraging. In *First Complex Systems Digital Campus World E-Conference 2015*; Bourguine, P., Collet, P., Parrend, P., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 351–374.
23. Agilent. Available online: <http://www.richrom.com/application/v2/public/upload/0/default/145.pdf> (accessed on 29 March 2020).
24. Thermo Fisher Scientific. Available online: https://www.thermoscientific.com/content/dam/tfs/ATG/CMD/CMD%20Marketing%20Material/MassSpectrometryCollateral/Brochures/BR30135_LTQOrbitrapXL_Brochure.pdf (accessed on 29 March 2020).
25. Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24*, 2534–2536. [CrossRef]
26. KOMICS. Available online: <http://www.kazusa.or.jp/komics/ja/home-ja/news-ja/213-powerget-3-5-7.html> (accessed on 29 March 2020).

27. Sakurai, N.; Ara, T.; Kanaya, S.; Nakamura, Y.; Iijima, Y.; Enomoto, M.; Motegi, T.; Aoki, K.; Suzuki, H.; Shibata, D. An application of a relational database system for high-throughput prediction of elemental compositions from accurate mass values. *Bioinformatics* **2013**, *29*, 290–291. [[CrossRef](#)]
28. Thermo Fisher Scientific. Available online: <https://assets.thermofisher.com/TFS-Assets/CMD/brochures/BR-70133-LC-Dionex-UltiMate-3000-RSLC-BR70133-EN.pdf> (accessed on 29 March 2020).
29. Q Exactive. Available online: <https://www.thermofisher.com/order/catalog/product/IQLAAEGAAPFALGMAZR> (accessed on 29 March 2020).
30. KOMICS. Available online: <http://www.kazusa.or.jp/komics/en/tool-en/236-powergetbatch.html> (accessed on 29 March 2020).
31. Flavonoid Viewer. Available online: <http://webs2.kazusa.or.jp/mfsearcher/flavonoidviewer/> (accessed on 29 March 2020).
32. Brunner, E.; Munzel, U. The Nonparametric Behrens-Fisher Problem: Asymptotic Theory and a Small-Sample Approximation. *Biom. J.* **2000**, *42*, 17–25. [[CrossRef](#)]
33. KEGG API. Available online: <https://www.kegg.jp/kegg/rest/keggapi.html> (accessed on 29 March 2020).
34. KEGG BRITE. Available online: <https://www.genome.jp/kegg/brite.html> (accessed on 29 March 2020).
35. KEGG PATHWAY. Available online: <https://www.genome.jp/kegg/pathway.html> (accessed on 29 March 2020).
36. Funabashi, M. Synecological Farming for Mainstreaming Biodiversity in Smallholding Farms and Foods: Implication for Agriculture in India. *Indian J. Plant Genet. Resour.* **2017**, *30*, 99–114. [[CrossRef](#)]
37. Rice, E.L. *Allelopathy*; Academic Press: New York, NY, USA, 1974.
38. Middleton, E.J.R.; Kandaswami, C.; Theoharides, T.C. The effects of plant flavonoids on mammalian cells: Implications for inflammation, heart disease, and cancer. *Pharmacol. Rev.* **2000**, *52*, 673–751. [[PubMed](#)]
39. Funabashi, M.; Ohta, K. Flavonoid-Rich Secondary Metabolites in naturally Grown Green Tea are Correlated with a Higher Shift of the Consumers' Excise Level. *J. Food Sci. Nutr.* **2020**, *6*, 63. [[CrossRef](#)]
40. Golisz, A.; Lata, B.; Gawronski, S.W.; Fujii, Y. Specific and total activities of the allelochemicals identified in Buckwheat. *Weed Biol. Manag.* **2007**, *7*, 164–171. [[CrossRef](#)]
41. Yang, W.; Xu, X.; Li, Y.; Wang, Y.; Li, M.; Wang, Y.; Ding, X.; Chu, Z. Rutin-Mediated Priming of Plant Resistance to Three Bacterial Pathogens Initiating the Early SA Signal Pathway. *PLoS ONE* **2016**, *11*, e0146910. [[CrossRef](#)] [[PubMed](#)]
42. Gross, R.L.; Newberne, P.M. Role of nutrition in immunologic function. *Physiol. Rev.* **1980**, *60*, 188–302. [[CrossRef](#)] [[PubMed](#)]
43. Rail, L.C.; Meydani, S.N. Vitamin B6 and Immune Competence. *Nutr. Rev.* **1993**, *51*, 217–225. [[CrossRef](#)]
44. Bello, M.D.G.; Knight, R.; Gilbert, J.A.; Blaser, M.J. Preserving microbial diversity. *Science* **2018**, *362*, 33–34. [[CrossRef](#)]
45. Defois, C.; Ratel, J.; Garrait, G.; Denis, S.; Goff, O.L.; Talvas, J.; Mosoni, P.; Engel, E.; Peyret, P. Food Chemicals Disrupt Human Gut Microbiota Activity And Impact Intestinal Homeostasis As Revealed By In Vitro Systems. *Sci. Rep.* **2018**, *8*, 11006. [[CrossRef](#)]
46. Clemente, J.C.; Pehrsson, E.C.; Blaser, M.J.; Sandhu, K.; Gao, Z.; Wang, B.; Magris, M.; Hidalgo, G.; Contreras, M.; Noya-Alarcón, Ó.; et al. The microbiome of uncontacted Amerindians. *Sci. Adv.* **2015**, *1*, e1500183. [[CrossRef](#)]
47. Vudhivai, N.; Ali, A.; Pongpaew, P.; Changbumrung, S.; Vorasanta, S.; Kwanbujan, K.; Charoenlarp, P.; Migasena, P.; Schelp, F.P. Vitamin B1, B2 and B6 status of vegetarians. *J. Med. Assoc. Thai.* **1991**, *74*, 465–470.
48. Douaud, G.; Refsum, H.; de Jager, C.A.; Jacoby, R.; Nichols, T.E.; Smith, S.M.; Smith, A.D. Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9523–9528. [[CrossRef](#)] [[PubMed](#)]
49. Bredesen, D.E.; Sharlin, K.; Jenkins, D.; Okuno, M.; Youngberg, W.; Cohen, S.H.; Stefani, A.; Brown, R.L.; Conger, S.; Tanio, C.; et al. Reversal of Cognitive Decline: 100 Patients. *J. Alzheimers Dis. Parkinsonism* **2018**, *8*, 450. [[CrossRef](#)]
50. Shetty, P.; Youngberg, W. Clinical Lifestyle Medicine Strategies for Preventing and Reversing Memory Loss in Alzheimer's. *Am. J. Lifestyle Med.* **2018**, *12*, 391–395. [[CrossRef](#)] [[PubMed](#)]

51. Begum, N.; Qin, C.; Ahanger, M.A.; Raza, S.; Khan, M.I.; Ashraf, M.; Ahmed, N.; Zhang, L. Role of Arbuscular Mycorrhizal Fungi in Plant Growth Regulation: Implications in Abiotic Stress Tolerance. *Front. Plant Sci.* **2019**, *10*, 1068. [[CrossRef](#)] [[PubMed](#)]
52. Liu, R.H. Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. *Am. J. Clin. Nutr.* **2003**, *78*, 517S–520S. [[CrossRef](#)] [[PubMed](#)]
53. Nishi, A.; Ohbuchi, K.; Kushida, H.; Matsumoto, T.; Lee, K.; Kuroki, H.; Nabeshima, S.; Shimobori, C.; Komokata, N.; Kanno, H.; et al. Deconstructing the traditional Japanese medicine “Kampo”: Compounds, metabolites and pharmacological profile of maoto, a remedy for flu-like symptoms. *NPJ Syst. Biol. Appl.* **2017**, *3*, 1–11. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).