Original Article
Food Science and Technology

ISSN 0101-2061 (Print)
ISSN 1678-457X (Online)

DOI: https://doi.org/10.1590/fst.100821

Fourier-transform infrared spectroscopy and machine learning to predict amino acid content of nine commercial insects

Yinchen HOU[1#], Penghui ZHAO[2], Fan ZHANG[3], Shengru YANG[1], Ahmed RADY[4], Nuwan K. WIJEWARDANE[4], Jihong HUANG[2], Mengxing LI[5,6*#]

Abstract

The nutritional profile, especially amino acid profile, determines the quality and commercial value of insect protein products. Multiple previous studies have used spectroscopy technologies and machine learning algorithms to predict essential amino acid content in various foods and feeds. However, these approaches were not applied for predicting essential amino acid content in insects before. In this study, 200 insect samples containing 9 commercial insect species were collected. Machine learning methods were applied to build the prediction models to predict amino acid content using Fourier-transform infrared spectroscopy (FTIR) raw spectra and first derivative. For all amino acids, partial least square regression, decision tree and radial basis artificial neural network exhibited high performances to predict essential amino acids. Model performances were improved for some amino acids using first derivative than using raw spectra. The highest performance (coefficient of determination: 0.97, root mean square error of prediction: 0.05 g/100 g and ratio of performance: 4.07) was achieved for phenylalanine prediction using radial basis artificial neural network modeling. The high model performance indicates the potential of applying FTIR and subsequent machine learning modeling for fast and non-destructive prediction of amino acid of insect products.

Keywords: mealworm; amino acid; FTIR; machine learning; prediction.

Practical Application: Insects and insect-derived products have been studied actively for exploring their use in food and feed industry. The nutritional profile, especially amino acid profile, determines the quality and commercial value of insect protein products. Multiple previous studies have used spectroscopy technologies and machine learning algorithms to predict essential amino acid content in various foods and feeds. However, these approaches were not applied for predicting essential amino acid content in insects before. In this study, the prediction models of fast and nondestructive prediction of amino acids are developed, which will be helpful to build an automatic system to predict amino acid content in a fast and non-destructive manner.

1 Introduction

Insect-derived products have been widely studied for their use in animal feed and human food industries due to their nutritional facts such as high protein content (Ding et al., 2019; Liu et al., 2021). For labeling declaration and quality control, it is necessary to quantitate the amino acid profile, especially the essential amino acid contents, which determine the quality of insect and insect derived protein products (Li et al., 2018). Amino acid profile differs from different insect diets and across different insect species (Zhang et al., 2019). Traditional chemical analysis methods quantitate the amino acid profile are chromatographic methods such as high pressure liquid chromatographic analysis, which requires the use of chemical solvent and is time-consuming (Liu et al., 2020; Santiago-Saenz et al., 2020). A non-destructive and fast method for determining amino acid composition is of great need.

In food, feed and pharmaceutical industries, spectroscopic study followed by subsequent chemometrics methods has been a hot research topic and was proved to be effective in determining amino acid composition in a non-destructive and fast manner (Farah et al., 2020; Mahboubifar et al., 2016). Research has been conducted on prediction of amino acids contents from cereal, milk, oilseed rape leaves and mammalian cell cultures using spectroscopy spectra, after following multivariate analysis, high prediction performance (R2 > 0.90, RPD > 2) were obtained (Bhatia et al., 2017; Li et al., 2019a; Yuwa-Amornpitak et al., 2020). In previous chemometrics studies, mostly commonly used multivariate analyses are principal component analysis and partial least squares regression (PLSR) (An et al., 2017; Huang et al., 2021; Wang et al., 2020; Zhang et al., 2021). As the mostly commonly and only available multivariate analysis method

Received 26 Oct., 2021
Accepted 21 Nov., 2021
[1] *College of Food and Biological Engineering, Henan University of Animal Husbandry and Economy, Zhengzhou, China*
[2] *College of Biological Engineering, Henan University of Technology, Zhengzhou, China*
[3] *College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, People's Republic of China*
[4] *Faculty of Engineering, University of Nottingham, King's Meadow Campus, Nottingham, UK*
[5] *Department of Biological Systems Engineering, University of Nebraska-Lincoln, Lincoln, USA*
[6] *Department of Statistics, University of Nebraska-Lincoln, Lincoln, USA*
*Corresponding author: mengxing.li@huskers.unl.edu
[#] *Both authors contributed equally to this manuscript.*

in many commercial chemometrics software, sometimes PLSR does not produce good prediction performances (Liu et al., 2021).

In recent years, novel machine learning methods such as decision tree (Ding et al., 2019), and artificial neural networks (Li et al., 2020; Sun et al., 2019) were found to be effective in predicting food ingredients content from spectroscopic data. These machine learning methods may show higher prediction performance than PLSR (Li et al., 2020; Liu et al., 2021). These novel machine learning methods can be built into the commercial chemometrics software and serve as alternative analysis method when PLSR does not perform well. However, to the best of the authors' knowledge, these novel machine learning methods have not been studied in terms of predicting insect amino acid content from spectroscopic data.

In this study, insect samples were scanned with FTIR and the spectra were collected to predict amino acid content using machine learning analysis. The objective of this study was to test the feasibility of predicting amino acid content using FTIR spectra data. Machine learning methods including partial least square regression, decision tree and radial basis artificial neural network were applied to analyze raw and first derivative of the raw spectra. The outcome from this study will be helpful to build an automatic system to predict amino acid content in a fast and non-destructive manner.

2 Materials and methods

2.1 Insect species and amino acid analysis

Fifty insect samples consisting of 9 insect species from multiple vendors were purchased from the online stores at Taobao.com. The 200 samples represented well different insect species commercially available in China. All insect samples were received in air dried or microwave dried form and were stored at 4 °C in air sealed bags before use. Information of the species were shown in Table 1.

Insect protein products are usually defatted on the market. Insect samples were defatted using Soxhlet extration with petroleum ether prior to amino acid analysis (Zhang et al., 2019). Defatted insect samples were freeze-dried and subjected to amino acid composition analysis with an amino acid analyzer (S433D, Sykam, Germany) (Liu et al., 2017; Zhang et al., 2019). Identification and quantification of amino acids were achieved based on their retention times of their peaks (Hou et al., 2020).

Table 1. Life stage and number of 9 insects.

Scientific name	Name	Life stage	#
Tenebrio molitor	Mealworm	Larvae	10
Zophobas morio	Superworm	Larvae	11
Locusta migratoria	Locust	Adult	6
Bombyx mori	Silkworm	Pupae	5
Teleogryllus emma	Cricket	Larvae	3
Acrida cinerea	Grasshopper	Adult	4
Chrysomya megacephala	Oriental latrine fly maggot	Larvae	4
Hermetia illucens	Black solider fly	Larvae	3
Musca domestica	Fly	Larvae	4

Individual free amino acid values were expressed as g/100 g of the freeze-fried sample weight (Li & Wilkins, 2021a). The 11 essential amino acids were studied based on the FAO/WHO requirements (Zhang et al., 2019). The summary statistics for their contents were shown in Table 2.

2.2 FTIR acquisition

FTIR spectra acquisition was conducted according to Bassbasi et al. (2014) and Liu et al. (2021) with some modifications. The transmittance spectra of 200 insect samples were collected from the FTIR spectrometer (WQF-510, Beijing Beifen-Ruili Analytical Instrument Co., Ltd.), equipped with a deuterated triglycerine sulfate KBr detector. Freeze-dried samples were pelleted with KBr powders (IR spectroscopy grade, Kermel Chemical Group) and placed into the sample holder of FTIR instrument for spectral acquisition. For each pellet, spectra were recorded at a resolution of 4 cm^{-1} from 4000 to 400 cm^{-1} using MainFTOS software. The average of 16 scans was used as a raw spectrum for further data analysis. The experiments were conducted in a room that had a controlled ambient temperature (25 °C) and relative humidity (30%). The background air and KBr spectra were subtracted from all sample spectra.

2.3 Machine learning analysis

Three machine learning methods: partial least squares regression (PLSR), decision tree, radial basis artificial neural networks (RBANN) were implemented to build prediction models between FTIR spectral data of defatted insect samples and the amino acid contents determined using chemical analysis. As the most commonly and most successfully used multivariate analysis method, the PLSR algorithm developed by De Jong (1993) was applied in this study. According to a previous study in our lab, the number of latent variables was chosen to be 15 (Liu et al., 2021). For the decision tree model, a regression model is fitted at each node and each regression tree is divided in a binary form where the response variable is partitioned to form homogeneous groups (De'ath & Fabricius, 2000; Li et al., 2018). The numbers of splits was equal to the size of data sample minus one (Liu et al., 2021). RBANN models the relationship between

Table 2. Summary statistics for 11 amino acids.

	Mean	Max	Min	Standard deviation
Lysine	2.92	4.03	1.90	0.50
Phenylalanine	2.08	3.90	1.22	0.56
Methionine	0.74	1.68	0.13	0.37
Threonine	2.01	4.90	1.24	0.51
Isoleucine	2.20	3.08	1.40	0.31
Leucine	3.60	5.49	2.18	0.69
Valine	3.01	4.25	1.87	0.53
Histidine	3.53	5.81	1.80	1.01
Arginine	2.66	3.96	1.58	0.47
Cysteine	0.16	0.39	0.03	0.08
Tyrosine	3.31	4.38	1.93	0.58

Note: all units are in g/100 g.

the predictors and response variables in a nonlinear approach (Rady et al., 2017; Varmuza & Filzmoser, 2009). RBANN is a radial-basis neural network formed with two layers with the first layer contains a number of neurons that is equal to the predictors and the second layer performs linear transformation to the response values based on the criterion of minimizing the mean square error value (Adedeji et al., 2020). For RBANN, the biases were located in both layers and the spread value was set to be 1 (Haykin, 1994). For each of the three models, to increase the robustness of prediction models, 10-fold cross validation was implemented and the optimal parameters were chosen based on the minimum root mean square error of validation (RMSEP).

Machine learning model performances were evaluated using coefficient of determination (R^2), root mean squared error of prediction (RMSEP) and the ratio of performance to deviation (RPD). The calculations of these parameters can be referred to Dai et al. (2014). The more R^2 is close to 1, the better the model performance is. The more root mean square error of prediction (RMSEP) is close to 1, the better the model performance is. By combining the ratio of standard deviation and RMSEP, RPD presents a relative predictive performance of the established model more directly and efficiently than when either R^2 or RMSEP is used separately (Li et al., 2020; Li & Wilkins, 2021b). Generally, the higher the RPD value is, the better and more robust the model is (Dai et al., 2014). A value of RPD above 2 indicates that a good performance of calibration was obtained (Guy et al., 2011; Li et al., 2018). Spectroscopy scientists refer to model reliability as excellent models with RPD > 2 and fair models with 1.4 < RPD < 2 (Li & Wilkins, 2020). All model fitting and model performance evaluation were performed in a MATLAB computational environment (MATLAB R2016, The Mathworks Inc., Natick, MA, USA).

3 Results and discussion

3.1 Insect amino acid statistics

The summary statistics of 11 amino acids is shown in Table 2. Among the 11 amino acids, cysteine, tyrosine and arginine are always required by infants and growing children, therefore they are also included as essential amino acids. Previous studies showed there are variations for amino acids among different insect species and same insect from different diets (Adedeji et al., 2020; Zhang et al., 2019). The variations were verified by the summary statistics in Table 2.

3.2 FTIR spectra

The sample FTIR raw and first derivative of spectra for 9 insects is shown in Figure 1. The transmittance for 11 insect samples in Figure 1 showed overall similar trends with some

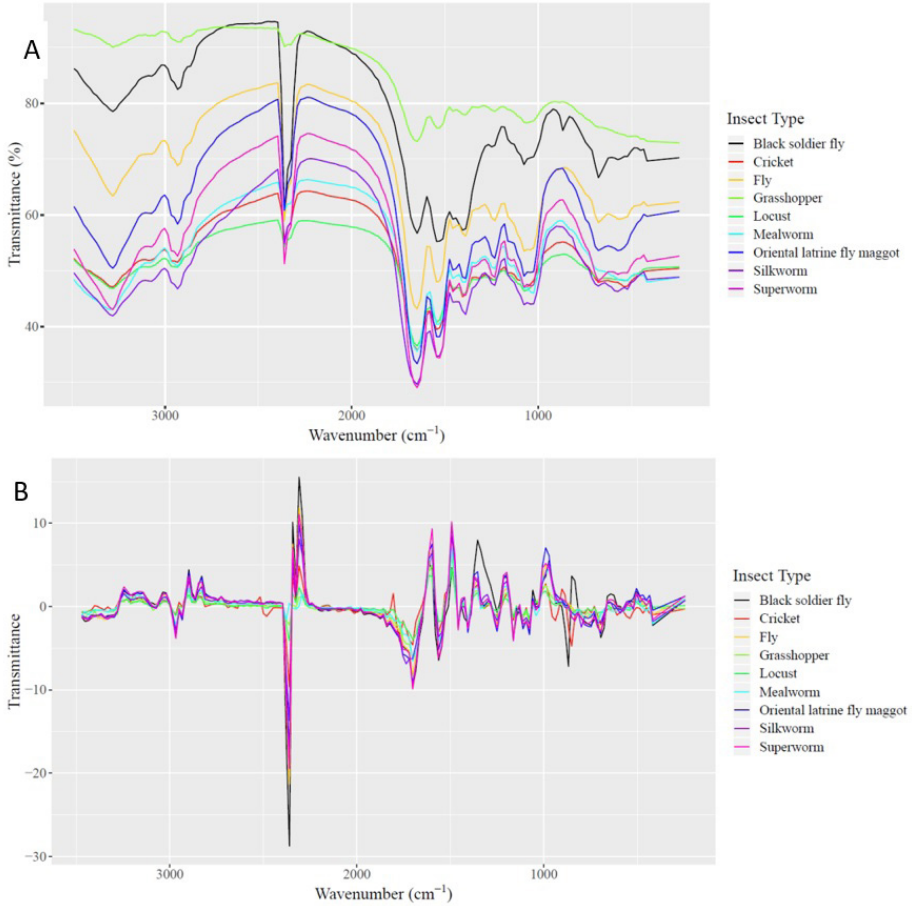

Figure 1. Sample spectra (raw and first derivative) for nine insects. A: raw spectra, B: first derivative spectra.

variations for certain regions. There are a few regions in FTIR spectra that were usually studied to characterize the structures of different proteins and amino acids. For example, Amide A (3225-3280 cm^{-1}) are due to the N-H stretching vibration. The principal Amide I (1700-1600 cm^{-1}) and Amide II (1600-1500 cm^{-1}) regions are mainly associated with the stretching vibrations of peptide carbonyl groups (Sun et al., 2020). The overall trends are similar for different insects but there are variations among the curves, especially for these regions. Compared to raw spectra, first derivative spectra may exhibit more information, and previous studies showed using derivative spectra as the predictor may lead to better prediction than using the raw spectra (Li et al., 2020), therefore, first derivative spectra were also used to predict fatty acid content in this study.

3.3 Machine learning model performance

The model performances using different machine learning methods based on two different input variables are shown in Tables 3 and 4. Using raw spectra as the predictors, different models tend to perform differently for different amino acids. For most amino acids, the highest performing model among 3 models achieved a R^2 of greater than 0.95, a RMSEP less than 0.1 and a RPD greater than 2. The best-performing model is radial basis artificial neural network with a R^2 of 0.0.95, a RMSEP of 0.06 g/100 g, and a RPD of 3.19 for methionine. As the most conventionally and commercially used multivariate analysis method, partial least square regression performing the best among different models is widely reported in previous research using raw spectra (Valdes et al., 2018). In this study, for some amino acids, partial least square regression outperforms other two models.

However, it is worth noting that for some amino acids, partial least square regression model showed bad performance for some amino acids, but other two models had higher performance than partial least square regression. For example, for histidine, R^2 and RPD, and RMSEP were just 0.26, 0.90 g/100 g and 0.95, respectively. However, radial basis artificial neural network produced much higher R^2 and RPD, and lower

Table 3. Model performances using FTIR raw spectra to predict 11 essential amino acids.

Amino Acid	Method	R^2	RMSEP (g/100 g)	RPD
Lysine	PLSR	0.78	0.25	1.59
	Decision Tree	0.93	0.17	2.43
	Radial basis artificial neural network	0.90	0.17	2.31
Phenylalanine	PLSR	0.92	0.23	2.32
	Decision Tree	0.19	0.35	0.58
	Radial basis artificial neural network	0.94	0.08	2.62
Methionine	PLSR	0.80	0.16	1.21
	Decision Tree	0.71	0.17	1.12
	Radial basis artificial neural network	0.95	0.06	3.19
Threonine	PLSR	0.85	0.11	1.95
	Decision Tree	0.66	0.16	1.33
	Radial basis artificial neural network	0.95	0.07	3.04
Isoleucine	PLSR	0.74	0.20	1.51
	Decision Tree	0.45	0.24	1.10
	Radial basis artificial neural network	0.84	0.14	1.86
Leucine	PLSR	0.66	0.50	1.34
	Decision Tree	0.39	0.50	1.10
	Radial basis artificial neural network	0.96	0.18	3.62
Valine	PLSR	0.65	0.38	1.34
	Decision Tree	0.67	0.41	1.26
	Radial basis artificial neural network	0.73	0.32	1.45
Histidine	PLSR	0.26	0.90	0.95
	Decision Tree	0.56	0.95	1.22
	Radial basis artificial neural network	0.83	0.47	1.82
Arginine	PLSR	0.69	0.34	1.34
	Decision Tree	0.86	0.33	1.41
	Radial basis artificial neural network	0.78	0.24	1.58
Cysteine	PLSR	0.68	0.05	1.32
	Decision Tree	0.67	0.05	1.36
	Radial basis artificial neural network	0.85	0.03	1.92
Tyrosine	PLSR	0.82	0.30	1.77
	Decision Tree	0.96	0.17	3.14
	Radial basis artificial neural network	0.91	0.27	2.22

Note: PLSR: partial least squares regression; R^2: coefficient of determination for regression; RPD: the ratio of performance to deviation; RMSEP: root mean squared error of prediction.

RMSEP for histidine prediction. For decision tree, R^2 and RPD are slightly higher, but RMSEP are similar to partial least square regression. In this case, the predictions for these amino acids are not accurate when partial least square regression is the only available prediction model in many commercial chemometrics instruments and software. Therefore, decision tree and radial basis artificial neural network can be built into the software and ensure high prediction performance.

When using first derivative of spectra as predictors, overall, the model performances were improved for most of the amino acids (Table 4). For lysine, phenylalanine, methionine, threonine, isoleucine and leucine, model performances using raw spectra were high, at least for one of the three models. For other amino acids, using first derivative significantly improved the model performances compared to using raw spectra. For example, for valine, the performances of three models improved using first derivative than using raw spectra. For histidine, the performances of partial least square regression and decision tree were improved greatly using first derivative than using raw spectra; the performance of radial basis artificial neural network performances were similar between using raw spectra and first derivative. For arginine, the R^2 and RPD did not change too much but RMSEP decreased a lot. RMSEP is usually regarded as the most important metric to evaluate model performance among R^2, RMSEP and RPD. For cysteine, performances of all three models were improved using first derivative compared to using raw spectra. Similar results were found in many previous studies, model performances were improved using first derivative than using raw spectra for fatty acid samples (Liu et al., 2021).

As the first study to predict amino acid content using FTIR spectra, the model performances are comparable to previous amino acid content prediction using spectroscopic data (Bhatia et al., 2017; Li et al., 2019b). If more data points were available, higher model performance is expected for all of the amino acids. With higher performance, a portable FTIR system coupled with machine learning methods may be developed in the future to realize real-time monitoring amino acid content for large scale manufacturing and logistics, like previous studies (LiEskridge&

Table 4. Model performances using FTIR first derivative of raw spectra to predict 11 essential amino acids.

Amino Acid	Method	R^2	RMSEP (g/100 g)	RPD
Lysine	PLSR	0.92	0.20	2.51
	Decision Tree	0.87	0.28	1.85
	Radial basis artificial neural network	0.93	0.14	2.88
Phenylalanine	PLSR	0.94	0.18	2.85
	Decision Tree	0.80	0.16	1.24
	Radial basis artificial neural network	0.97	0.05	4.07
Methionine	PLSR	0.85	0.12	1.60
	Decision Tree	0.85	0.17	1.14
	Radial basis artificial neural network	0.93	0.07	2.63
Threonine	PLSR	0.89	0.10	2.27
	Decision Tree	0.95	0.08	2.99
	Radial basis artificial neural network	0.85	0.14	1.65
Isoleucine	PLSR	0.77	0.16	1.61
	Decision Tree	0.95	0.09	2.99
	Radial basis artificial neural network	0.79	0.16	1.59
Leucine	PLSR	0.81	0.32	1.72
	Decision Tree	0.79	0.40	1.66
	Radial basis artificial neural network	0.91	0.28	2.41
Valine	PLSR	0.77	0.29	1.61
	Decision Tree	0.98	0.10	4.71
	Radial basis artificial neural network	0.84	0.27	1.91
Histidine	PLSR	0.74	0.71	1.54
	Decision Tree	0.96	0.41	2.69
	Radial basis artificial neural network	0.76	0.60	1.42
Arginine	PLSR	0.65	0.27	1.37
	Decision Tree	0.86	0.19	1.96
	Radial basis artificial neural network	0.74	0.27	1.40
Cysteine	PLSR	0.84	0.04	1.89
	Decision Tree	0.94	0.02	2.73
	Radial basis artificial neural network	0.92	0.03	2.70
Tyrosine	PLSR	**0.97**	**0.13**	**4.41**
	Decision Tree	0.87	0.29	2.07
	Radial basis artificial neural network	0.94	0.18	2.97

Note: PLSR: partial least squares regression.

Wilkins2019). Next step would be to predict amino acid content from whole insect powder. Wavelength selection would be also of interest for reducing memory and computation cost. For example, particular wavelengths within the regions of Amide A (3225-3280 cm^{-1}), principal Amide I (1700-1600 cm^{-1}) and Amide II (1600-1500 cm^{-1}) may provide reliable predictions for amino acid without the need to do full scan from 4000-400 cm^{-1} (Sun et al., 2019).

4 Conclusion

In this study, FTIR followed by machine learning analysis was employed to predict amino acid contents of 9 commercial insects. Machine learning analysis on the spectral data was proved to be effective predicting PHB production. Decision tree and radial basis artificial neural network modeling can produce good prediction performances when partial least square regression does not perform well. Using first derivative of spectra as predictor to predict amino acid led to higher performance compared to using raw spectra as predictors. The highest-performing model was radial basis artificial neural network with a R2 of 0.97, a RMSEP of 0.05 g/100 g, and a RPD of 4.07 using first derivative of raw spectra for phenylalanine. With high-performing prediction models, portable or online spectrometers may be developed to quantitate amino acid content in a fast and non-destructive manner to help label declarations and quality control.

Conflict of interest

This is original work of the authors. There are no conflicts of interest and any related work is not in press or submitted elsewhere. All authors agreed to submit this work to Journal of Food Science and Technology.

Funding

National Key R&D Program of China (2016YFD0400800); Henan province youth talent support project (2019HYTP009); Science and Technology Department of Henan Province (182102110127); Zhongyuan scholars in Henan Province (192101510004) .

Acknowledgements

The authors would like to thank National Key R&D Program of China (2016YFD0400800), Henan province youth talent support project (2019HYTP009), Science and Technology Department of Henan Province (182102110127) and Zhongyuan scholars in Henan Province (192101510004) for funding this research.

References

Adedeji, A. A., Ekramirad, N., Rady, A., Hamidisepehr, A., Donohue, K. D., Villanueva, R. T., Parrish, C. A., & Li, M. (2020). Non-destructive technologies for detecting insect infestation in fruits and vegetables under postharvest conditions: a critical review. *Foods*, 9(7), 927. http://dx.doi.org/10.3390/foods9070927. PMid:32674380.

An, Z., Jiang, X., Xiang, G., Fan, L., He, L., & Zhao, W. (2017). A simple and practical method for determining iodine values of oils and fats by the FTIR spectrometer with an infrared quartz cuvette. *Analytical Methods*, 9(24), 3669-3674. http://dx.doi.org/10.1039/C7AY00727B.

Bassbasi, M., Platikanov, S., Tauler, R., & Oussama, A. (2014). FTIR-ATR determination of solid non fat (SNF) in raw milk using PLS and SVM chemometric methods. *Food Chemistry*, 146, 250-254. http://dx.doi.org/10.1016/j.foodchem.2013.09.044. PMid:24176339.

Bhatia, H., Mehdizadeh, H., Drapeau, D., & Yoon, S. (2017). In-line monitoring of amino acids in mammalian cell cultures using raman spectroscopy and multivariate chemometrics models. *Engineering in Life Sciences*, 18(1), 55-61. http://dx.doi.org/10.1002/elsc.201700084. PMid:32624861.

Dai, Q., Sun, D. W., Xiong, Z., Cheng, J. H., & Zeng, X. A. (2014). Recent advances in data mining techniques and their applications in hyperspectral image processing for the food industry. *Comprehensive Reviews in Food Science and Food Safety*, 13(5), 891-905. http://dx.doi.org/10.1111/1541-4337.12088.

De Jong, S. (1993). SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 18(3), 251-263. http://dx.doi.org/10.1016/0169-7439(93)85002-X.

De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), 3178-3192. http://dx.doi.org/10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2.

Ding, C., Wang, X., & Li, M. (2019). Evaluation of six white-rot fungal pretreatments on corn stover for the production of cellulolytic and ligninolytic enzymes, reducing sugars, and ethanol. *Applied Microbiology and Biotechnology*, 103(14), 5641. http://dx.doi.org/10.1007/s00253-019-09884-y. PMid:31115636.

Farah, J. S., Cavalcanti, R. N., Guimares, J. T., Balthazar, C. F., & Cruz, A. G. (2020). Differential scanning calorimetry coupled with machine learning technique: an effective approach to determine the milk authenticity. *Anais do Congresso Brasileiro de Ciência e Tecnologia de Alimentos*, 121, 107585.

Guy, F., Prache, S., Thomas, A., Bauchart, D., & Andueza, D. (2011). Prediction of lamb meat fatty acid composition using near-infrared reflectance spectroscopy (NIRS). *Food Chemistry*, 127(3), 1280-1286. http://dx.doi.org/10.1016/j.foodchem.2011.01.084. PMid:25214127.

Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Upper Saddle River: Prentice Hall.

Hou, Y., Yang, S., Huang, J., Xu, Q., Liao, A., Zhong, Q., & Li, M. (2020). Nutritional profile and in vitro immunomodulatory activity of protein extract from goat placenta and fermented extraction residual. *Journal of Food Process Engineering*, 44, e13576.

Huang, T., Qin, K., Yan, Y., He, X., Dai, G., & Zhang, B. (2021). Correlation between the storability and fruit quality of fresh goji berries. *Food Science and Technology*. http://dx.doi.org/10.1590/fst.46120.

Li, M., & Wilkins, M. (2020). Fed-batch cultivation and adding supplements to increase yields of polyhydroxybutyrate production by *Cupriavidus necator* from corn stover alkaline pretreatment liquor. *Bioresource Technology*, 299, 122676. http://dx.doi.org/10.1016/j.biortech.2019.122676. PMid:31924491.

Li, M., & Wilkins, M. (2021a). Lignin bioconversion into valuable products: fractionation, depolymerization, aromatic compound conversion, and bioproduct formation. *Systems Microbiology and Biomanufacturing*, 1(2), 166-185. http://dx.doi.org/10.1007/s43393-020-00016-6.

Li, M., & Wilkins, M. R. (2021b). Fed-batch polyhydroxybutyrate production by Paraburkholderia sacchari from a ternary mixture of glucose, xylose and arabinose. *Bioprocess and Biosystems*

Engineering, 44(1), 1-9. http://dx.doi.org/10.1007/s00449-020-02434-1. PMid:32895870.

Li, M., Ekramirad, N., Rady, A., & Adedeji, A. (2018). Application of acoustic emission and machine learning to detect codling moth infested apples. *Transactions of the ASABE*, 61(3), 1157-1164. http://dx.doi.org/10.13031/trans.12548.

Li, M., Eskridge, K., Liu, E., & Wilkins, M. (2019a). Enhancement of polyhydroxybutyrate (PHB) production by 10-fold from alkaline pretreatment liquor with an oxidative enzyme-mediator-surfactant system under Plackett-Burman and central composite designs. *Bioresource Technology*, 281, 99-106. http://dx.doi.org/10.1016/j.biortech.2019.02.045. PMid:30807996.

Li, M., Eskridge, K. M., & Wilkins, M. R. (2019b). Optimization of polyhydroxybutyrate production by experimental design of combined ternary mixture (glucose, xylose and arabinose) and process variables (sugar concentration, molar C: N ratio). *Bioprocess and Biosystems Engineering*, 42(9), 1495. http://dx.doi.org/10.1007/s00449-019-02146-1. PMid:31111213.

Li, M., Wijewardane, N. K., Ge, Y., Xu, Z., & Wilkins, M. R. (2020). Visible/near infrared spectroscopy and machine learning for predicting polyhydroxybutyrate production cultured on alkaline pretreated liquor from corn stover. *Bioresource Technology Reports*, 9, 100386. http://dx.doi.org/10.1016/j.biteb.2020.100386.

Liu, K., Zheng, J., & Chen, F. (2017). Relationships between degree of milling and loss of Vitamin B, minerals, and change in amino acid composition of brown rice. *Lebensmittel-Wissenschaft + Technologie*, 82(Suppl. C), 429-436. http://dx.doi.org/10.1016/j.lwt.2017.04.067.

Liu, E., Li, M., Abdella, A., & Wilkins, M. R. (2020). Development of a cost-effective medium for submerged production of fungal aryl alcohol oxidase using a genetically modified *Aspergillus nidulans* strain. *Bioresource Technology*, 305, 123038. http://dx.doi.org/10.1016/j.biortech.2020.123038. PMid:32120232.

Liu, Z., Rady, A., Wijewardane, N. K., Shan, Q., Chen, H., Yang, S., Li, J., & Li, M. (2021). Fourier-transform infrared spectroscopy and machine learning to predict fatty acid content of nine commercial insects. *Journal of Food Measurement and Characterization*, 15(1), 953-960. http://dx.doi.org/10.1007/s11694-020-00694-9.

Mahboubifar, M., Yousefinejad, S., Alizadeh, M., & Hemmateenejad, B. (2016). Prediction of the acid value, peroxide value and the percentage of some fatty acids in edible oils during long heating time by chemometrics analysis of FTIR-ATR spectra. *Journal of the Indian Chemical Society*, 13(12), 2291-2299. http://dx.doi.org/10.1007/s13738-016-0948-1.

Rady, A., Ekramirad, N., Adedeji, A., Li, M., & Alimardani, R. (2017). Hyperspectral imaging for detection of codling moth infestation in GoldRush apples. *Postharvest Biology and Technology*, 129, 37-44. http://dx.doi.org/10.1016/j.postharvbio.2017.03.007.

Santiago-Saenz, Y. O., López-Palestina, C. U., Gutiérrez-Tlahque, J., Monroy-Torres, R., Pinedo-Espinoza, J. M., & Hernández-Fuentes, A. D. (2020). Nutritional and functional evaluation of three powder mixtures based on mexican quelites: alternative ingredients to formulate food supplements. *Food Science and Technology*, 40(4), 1029-1037. http://dx.doi.org/10.1590/fst.28419.

Sun, X., Atiyeh, H. K., Li, M., & Chen, Y. (2020). Biochar facilitated bioprocessing and biorefinery for productions of biofuel and chemicals: a review. *Bioresource Technology*, 295, 122252. http://dx.doi.org/10.1016/j.biortech.2019.122252. PMid:31669180.

Sun, Y., Yang, S., Li, G., & Li, M. (2019). Preparation of starch phosphate carbamides and its application for improvement of noodle quality. *Czech Journal of Food Sciences*, 37(6), 456-462. http://dx.doi.org/10.17221/159/2019-CJFS.

Valdes, A., Beltran, A., Mellinas, C., Jimenez, A., & Garrigos, M. C. (2018). Analytical methods combined with multivariate analysis for authentication of animal and vegetable food products with high fat content. *Trends in Food Science & Technology*, 77, 120-130. http://dx.doi.org/10.1016/j.tifs.2018.05.014.

Varmuza, K., & Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. Boca Raton: CRC Press.

Wang X, Xing X, Zhao M, Yang J. (2020). Comparison of multispectral modeling of physiochemical attributes of greengage: Brix and pH values. *Food Science and Technology*. In press.

Yuwa-Amornpitak, T., Butkhup, L., & Yeunyaw, P.-N. (2020). Amino acids and antioxidant activities of extracts from wild edible mushrooms from a community forest in the Nasrinual District, Maha Sarakham, Thailand. *Food Science and Technology*, 40(3), 712-720. http://dx.doi.org/10.1590/fst.18519.

Zhang, D., Ji, H.-W., Luo, G.-X., Chen, H., Liu, S.-C., & Mao, W.-J. (2021). Insight into aroma attributes change during the hot-air-drying process of white shrimp using GC-MS, E-Nose and sensory analysis. *Food Science and Technology*. In press. http://dx.doi.org/10.1590/fst.70820.

Zhang, X., Tang, H., Chen, G., Qiao, L., Li, J., Liu, B., Liu, Z., Li, M., & Liu, X. (2019). Growth performance and nutritional profile of mealworms reared on corn stover, soybean meal, and distillers' grains. *European Food Research and Technology*, 245(12), 1-10. http://dx.doi.org/10.1007/s00217-019-03336-7.