

Physicochemical properties, content, composition and partial least squares models of *A. trifoliata* seeds oil

Yicheng Zhong^a, Zhenqian Zhang^b, Jing Chen^a, Juan Niu^a, Yaliang Shi^a, Yue Wang^a,
Tianxin Chen^a, Zhimin Sun^a, Jianhua Chen^{a,*}, Mingbao Luan^{a,*}

^a Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences/Key Laboratory of Stem-Fiber Biomass and Engineering Microbiology, Ministry of Agriculture, Changsha 410205, PR China

^b Agricultural College, Hunan Agricultural University, Changsha 410205, PR China

ARTICLE INFO

Keyword:

Akebia trifoliata

Edible oil

GC-MS

Near-infrared spectroscopy

ABSTRACT

Physicochemical properties, oil content, and fatty acids (FAs) composition are key for determining the value of oil crops. The aim of this study was to illustrate the potential of exploiting *A. trifoliata* as an edible oil crop, and establish a rapid measurement model for the *A. trifoliata* seeds oil (ASO) content and composition. In 130 *A. trifoliata* germplasms, the highest content of ASO was 51.27%, and unsaturated fatty acids (UFAs) mainly accounted for 74–78% of ASO. The partial least squares (PLS) model based on GC-MS and near-infrared spectroscopy was well-suited for the determination of ASO and UFA content; however, the PLS model for oleic acid (OA) and linoleic acid (LA) was not effective. The acid values and peroxide values for ASO also conformed to the Chinese food safety standards. Our findings will provide new insights and guidance for the use of *A. trifoliata* as oil crops.

1. Introduction

The production of edible oil in China cannot meet the increasing consumer demand resulting from population increase and economic development. The supply of a large amount of edible oil relies on imports (Liao, Yin, & Ma, 2018). It is impossible to achieve increased self-sufficiency for edible oil by increasing the planting area of traditional herbaceous oil crops, because China needs large amounts of arable land for other crops to ensure food security (Liao, Li, Yu, Huang, & Yan, 2021). However, woody oil does not compete with food crops for land and improves the greening area. Therefore, the development of woody oil has a huge market potential, and its production is expected to resolve the problem of the global edible oil shortage.

Akebia trifoliata (Thunb.) Koidz is a woody perennial plant. In China, it is mainly found in the Yangtze River Basin, Yellow River Basin, and Shaanxi-Sichuan Area (Xie, Li, Zhang, Ouyang, & Xiao, 2006). It is highly adaptable to natural conditions and can grow in ravines and rock crevices, and other similar environments. Although *A. trifoliata* has been

used in traditional Chinese medicine for a long time, in recent years, it has been found to have excellent oil value. *A. trifoliata* fruits have many seeds (300–800), high seed yield (400 kg/667 m²), and high seed oil content (47.62%) (Niu et al., 2020). *A. trifoliata* seed oil (ASO) contains more than nine kinds of fatty acids (FAs), with the main ingredients of palmitic acid, oleic acid (OA), and linoleic acid (LA). Additionally, ASO contains more than 70% unsaturated fatty acids (UFAs). During the period of shortage of edible oil (before the 1980 s), *A. trifoliata* seed oil (ASO) was used as edible oil in southern China (Zhao et al., 2014). In recent years, research on ASO has been minimal, and most of the research focused on oil extraction methods (Zhou, 2018; Li, Wu, Yu, 2021). Many details still need to be confirmed before *A. trifoliata* may be used as an oil crop. These include its physicochemical properties and composition, the establishment of rapid detection methods, screening of high fat content varieties, and oil synthesis pathways.

Physicochemical properties and composition are vital for the evaluation of plant oil use. The acid value reflects the weight of free FAs (mg) in the oil (1 g), with lower acid value indicating better oil quality. The

Abbreviations: ASO, *A. trifoliata* seed oil; FAs, Fatty acids; UFA, Unsaturated fatty acids; OA, Oleic acid; LA, Linoleic acid; NIRS, Near-infrared spectroscopy; PLS, Partial least squares; SNV, Standard normal variate; MSC, Multiplicative scatter correction; D1, First derivative (Savitzky-Golay); D2, Second derivative (Savitzky-Golay); PCA, Principal component analysis; R2cal, Coefficients of determination for calibration; RMSEC, Root mean square error of calibration; R2cv, Coefficient of determination for cross-validation; RMSEP, Root mean square error of prediction.

* Corresponding authors.

E-mail addresses: chenjianhua@caas.cn (J. Chen), luanmingbao@caas.cn (M. Luan).

<https://doi.org/10.1016/j.fochx.2021.100131>

Received 7 July 2021; Received in revised form 26 August 2021; Accepted 7 September 2021

Available online 21 September 2021

2590-1575/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

peroxide value is an index for the degree of oxidation of the oil (%), reflecting the degree of its rancidity. These are methods of calculating the acid value and peroxide of edible oil, and they are also used in Chinese food safety standards (GB 2716-2018). In China, the acid (KOH) and peroxide values of edible oil cannot be higher than 3.0 mg/g and 0.25 g/100 g, respectively; therefore, it is necessary to determine the physicochemical properties of ASO before it may be considered for use as edible oil. The composition and proportion of FAs are the key factors for evaluating the edible oil quality with edible oils. High OA content and low stearic and erucic acid contents considered to be of a high quality edible oil (Kliem, Shingfield, Humphries, & Givens, 2011; Jarvis et al., 2021). Zhong and Ma (2016) showed that the FAs of *A. trifoliata* seeds were composed of nine different FAs, palmitic acid (21.607 %), OA (47.967 %), LA (25.938 %), stearic acid (3.744 %), and the contents of the other five FA were less than 0.5 %, which was similar to the findings of Zhou et al. (2018) and Zhao et al. (2014). However, based on the current research, we did not know whether there was a significant difference in the oil composition ratio between the samples sourced from different geographical areas and between the different varieties.

The measurement of seed oil content is critical in the breeding of oil crops, and many different measurement methods are currently in use. The measurement methods include Soxhlet extraction, supercritical CO₂ extraction, and methods based on ultrasound extraction. Additionally, spectral analysis and GC-MS are also used. Zhou (2018) showed that the Soxhlet method was superior to both the ultrasonic method and ultrasonic-assisted hydro enzymatic method for the oil extraction of *A. trifoliata*. Both the physical extraction methods (such as Soxhlet extraction) and the chemical analysis methods (such as GC-MS) require pre-treatment of the seeds. As a result, these methods are time-consuming and destroy the activity of the seeds. Thus, it is important to establish a rapid, low-cost method for measuring seed oil content and composition without damaging the seed oil activity (Wang et al., 2020). With the development of computer technology and chemical measurement, near-infrared spectroscopy (NIRS) has become increasingly popular. NIRS examines organic chemicals in the near-infrared spectrum (780–2526 nm) to determine their optical characteristics. It provides a rapid and inexpensive method for assessing one or more chemical components of the samples. Furthermore, in addition to its other advantages, the samples for NIRS do not need to be pre-treated, (Casale, Zunin, Cosulich, Perego, Lanteri, 2010; Allendorf, Subramanian, & Rodriguez-Saona, 2012). NIRS has been used in many fields, such as quality identification (Chen, Yu, Jiang, Zhang, & Wang, 2021), content determination (Ribeiro, Salva, Silvarolla, 2021; Albanell, Martínez, Marchi, & Manuelian, 2021), and material classification (Yusnaini et al., 2021; Roberto et al., 2021). In oil crop science research, NIRS has been widely used to rapidly determine seed fatty acids in rapeseed (Hom, Becker, & Möllers, 2007), soybean (Ferreira, Pallone, & Poppi, 2013), and sunflower (Akkaya, 2018). However, to the best of our knowledge, there have been no reports on the use of NIRS to determine the seed fatty acids of *A. trifoliata*.

The main aim of this study was to analyze the feasibility of NIRS for the determination of oil content and the main components in the whole seed of *A. trifoliata*. In addition, this study explored the potential of exploiting *A. trifoliata* as an edible oil crop based on physicochemical properties, content, composition of ASO.

2. Materials and methods

2.1. Plant materials

The 130 *A. trifoliata* samples were collected from seven different locations (Table S1). JM1-JM6 were taken from the wild, and the other 124 samples were taken from the plantation or scientific research base. First, three ripe fruits were randomly picked from each plant and transported to the laboratory. Seeds were then separated from the flesh and washed with distilled water. Finally, the seeds were dried in a

drying oven at 45°C and stored in a desiccator.

2.2. Seed oil extraction

Based on the results of Zhou et al. (2018), Soxhlet extraction with petroleum ether (boiling range: 30–60 °C) as solvent was used to extract seed oil. A three-factor, five-level experiment was set up to optimize the Soxhlet extraction method (extraction temperature: 35, 40, 45, 50, and 55 °C; extraction time: 1, 2, 3, 4, and 5 h; solid-liquid ratio: 1:30, 1:40, 1:50, 1:60, 1:70). First, single factor tests were carried out, followed by orthogonal optimization. Then, the optimized Soxhlet extraction method was used to determine the ASO content of 130 *A. trifoliata* sample. Seed powder (approximately 3 g) was placed in a filter paper thimble and extracted under corresponding extraction conditions, and the extraction temperature was controlled by a water bath and thermometer. After the extraction, the solvent was evaporated using a rotary evaporator and the round-bottom flask was brought to a constant weight in an oven at 105 °C. The round-bottom flask was dried to constant weight and weighed prior to extraction. The round-bottom flask was weighed and the oil content was calculated based on the weight difference of the round-bottom flask before and after extraction. The oil was collected for the next test. All analyses were performed in triplicate.

2.3. Physicochemical properties analysis

To study the physicochemical properties of the ASO, three varieties were randomly selected from each sample from a different geographical origin.

2.3.1. Acid Value, X_{AV}

ASO (1 g) was weighed precisely and placed in a 250 mL conical flask. Ethanol-diethyl ether (1:1 v/v, 50 mL) was added to the flask before adding phenolphthalein (1.0 mL). The flask was shaken until the mixture dissolved completely. Sodium hydroxide standard solution (0.1 mol/L) was added until the solution turned slightly pink, and the color did not fade for 30 s. Meanwhile, a blank control was performed.

$$X_{AV} = \frac{(V - V_0) * 0.1 * 56.1}{1} \quad (1)$$

Where, X_{AV} : acid value (mg/g), V : measure the standard titrated liquid volume consumed (mL), and V_0 : Ck standard titrated liquid volume consumed (mL).

2.3.2. Peroxide value

ASO (1 g) was placed into a 250 mL conical flask. Trichloromethane - glacial acetic acid (2:3 v/v, 30 mL) was added and the flask was shaken until the mixture dissolved completely. Then, a potassium iodide test solution (0.5 mL) was added. After one min, water (30 mL) was added, and the sodium thiosulfate titrant (0.01 M) was slowly added until the yellow color disappeared. Thereafter, a starch indicator solution (5 mL) was added, and the solution was shaken until the blue color disappeared. A blank control test was also performed.

$$X_1 = \frac{(V - V_0) * 0.01 * 126.9}{1} \quad (2)$$

Where X_1 : peroxide value (g/100 g), V : measure the standard titrated liquid volume consumed (mL), and V_0 : Ck standard titrated liquid volume consumed (mL).

2.4. Fatty acid composition analysis

GC-MS analysis was conducted using an Agilent GC 7890 gas chromatograph and an Agilent 5977 mass spectrometer. The gas chromatograph was equipped with a fused silica capillary column DB-23, and the carrier gas was helium (99.999% purity). The operating conditions of

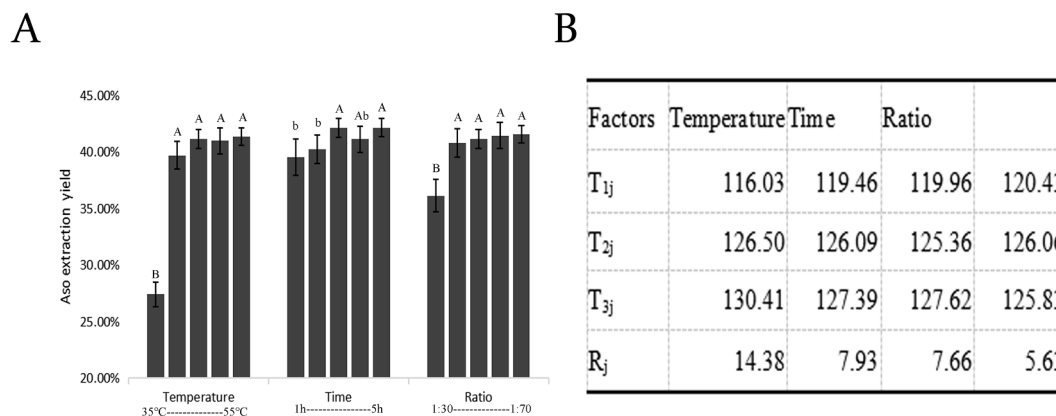


Fig 1. Soxhlet extraction method optimization. **A.** Single factor's influence on ASO extraction rate; **B.** Orthogonal optimization range analysis table. Different letters stands for significance for $p \leq 0.05$ (a,b,c..) or $p \leq 0.01$ (A,B,C..)

the GC-MS are described in Sun et al. (2015). Before GC-MS analysis, FAs need to be derivatized to form FA methyl esters (FAMES). As described in previous studies (Zhang, et al., 2014), seed oil (0.06 g) was diluted with diethyl ether/petroleum ether (1:1 v/v, 2 mL) and 0.4 M KOH-CH₃OH (1 mL), vortexed, and maintained at room temperature (around 25°C) 2.5 h. Then, redistilled water (2 mL) was added, vortexed, and centrifuged at 4500 rpm for 2 min. Finally, the organic phase containing FAMES (100 mL) was collected and diluted with petroleum ether (900 mL).

2.5. NIRS scanning

The spectral data for intact *A. trifoliata* seeds were obtained using a NIRS DS2500 instrument (FOSS, Denmark). Based on the seed size, a small sample cup with a diameter of 58 mm was used in this study. The sample cup had a quartz window that eliminated the drying effect on the surface of the sample. The NIRS DS2500 instrument with an integrating sphere was used to record log (1/R) every 0.5 nm from 850 to 2500 nm, with 3300 data points obtained for each sample. The seeds of each sample were randomly selected and the sample cup was filled to collect the spectrum. The analysis was repeated three times.

2.6. NIRS model establishment and validation

Three repeated data points were averaged, and principal component analysis (PCA) was performed based on the average of the spectral and chemical data. Based on the PCA results, abnormal samples were removed. Seventy percent of the retained samples were used for calibration, and thirty percent of the samples were used as an external source of samples for validation. Standard normal variate (SNV), first derivative (Savitzky-Golay) (D1), second derivative (Savitzky-Golay) (D2), multiplicative scatter correction (MSC), D1 + SNV, and D1 + MSC were applied to the original data matrix to minimize the signal-to-noise ratio and the baseline problem. The near-infrared spectral data were considered as the independent variable X, and the contents of FA, UFA, OA, and LA were considered as the dependent variable Y. PLS was used to establish the determination model of the corresponding content. The best model was determined by comparing different model coefficients of determination for calibration (R^2_{cal}), root mean square error of calibration (RMSEC), coefficient of determination for cross-validation (R^2_{cv}), and root mean square error of prediction (RMSEP). Then, the optimal model corresponding to FA, UFA, OA, and LA was used to predict the samples of the verification set, calculate the difference between the predicted and experimental values, and the *t*-test and

Table 1

Acid value and peroxide value of ASO. Different letters stands for significance for $p \leq 0.05$ (a,b,c..) or $p \leq 0.01$ (A,B,C..).

Group	Sample	Acid value (mg/g)	Mean	National standard	Peroxide value (g/100 g)	Mean	National standard
YC	YC-8	1.234	1.258 a	≤3.0	0.082	0.086b	≤0.25
	YC-13	1.198			0.096		
	YC-20	1.342			0.081		
CS	CS19-16	0.989	1.055 ab		0.103	0.098 ab	
	CSN18	1.134			0.098		
	CS20-2	1.042			0.094		
YJ	YJ1-21	1.342	1.261 a		0.093	0.105 a	
	YJ2-2	1.248			0.114		
	YJ1-19	1.193			0.108		
TJ	TJ1-6	1.245	1.207 ab		0.083	0.095 ab	
	TJ1-13	1.283			0.109		
	TJ1-25	1.093			0.093		
LY	LY-3	0.782	0.998b		0.109	0.108 a	
	LY-4	1.093			0.113		
	LY-6	1.118			0.102		
JM	JM-1	1.193	1.253 a		0.096	0.099 ab	
	JM-3	1.301			0.090		
	JM-4	1.265			0.110		
XA	XA-7	0.832	1.127 ab		0.099	0.109 a	
	XA-11	1.241			0.118		
	XA-15	1.309			0.110		

ASO: *A. trifoliata* seed oil.

National standard: Chinese national standards of edible oil.

Different letters stands for significance for $p \leq 0.05$ (a,b,c..) or $p \leq 0.01$ (A,B,C..).

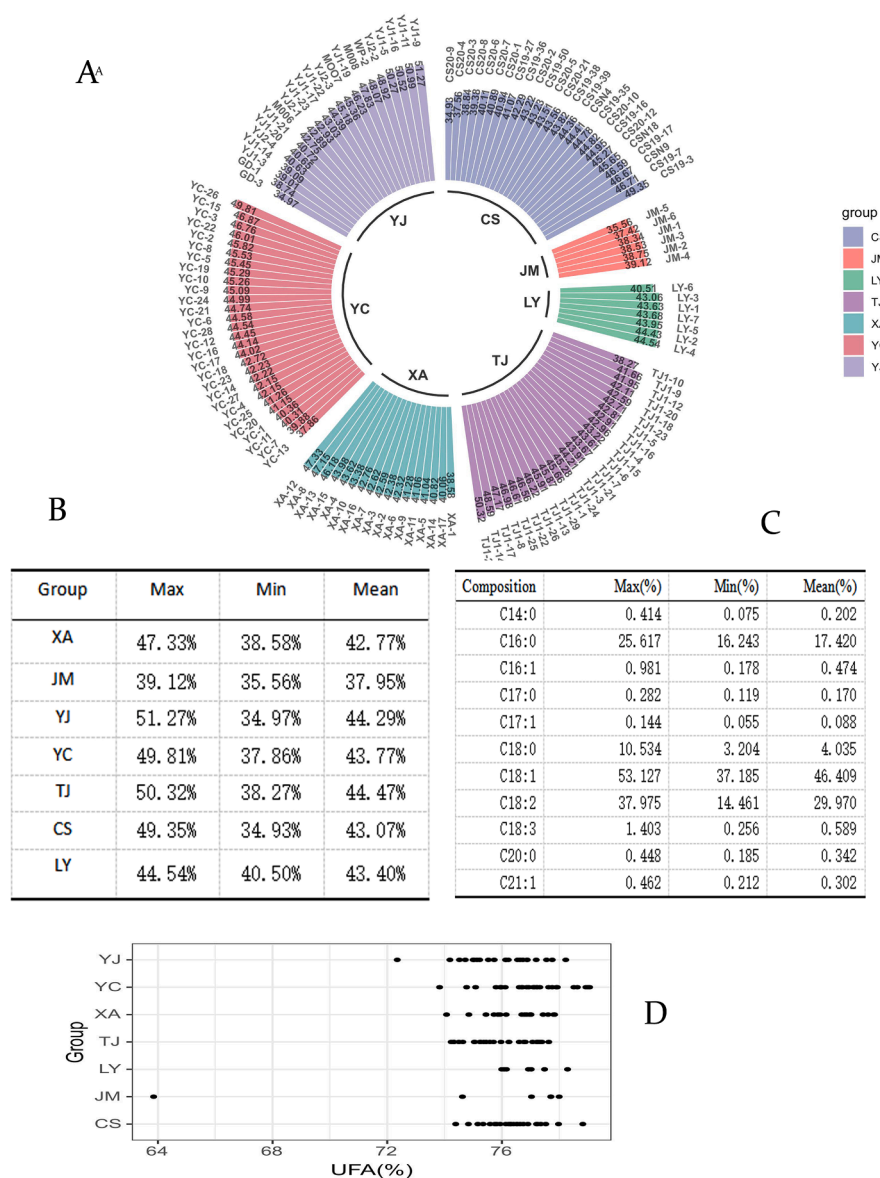


Fig 2. ASO content and composition. **A.** ASO content of 130 samples; **B.** ASO content information of seven groups; **C.** ASO composition average of 130 samples; **D.** UFA accounted for ASO ratio distribution.

correlation analysis were performed with SAS 9.0.

3. Results

3.1. Optimization of the Soxhlet extraction method

Single-factor experiment results (Fig. 1A) showed that the extraction temperature, extraction time, and solid-liquid ratio affected the extraction rate of ASO. When the extraction temperature changed from 35 °C to 40 °C, the ASO extraction efficacy improved substantially (from 27.40% to 39.65%). Subsequently the extraction efficacy gradually stabilized. The effects of extraction time and the solid-liquid ratio were not as considerable as that of the extraction temperature, but the trend of the extraction efficacy was the same: that is, the extraction efficacy first increased and then stabilized. Based on the single factor experiment results and single factor analysis, 30 °C, 40 °C, 50 °C, 2 h, 3 h, 4 h, 1:40, 1:50, and 1:60 were selected as the factors for the $L_9(3^4)$ orthogonal test. Orthogonal test results (Fig. 1B) showed that among the three factors, extraction temperature had the greatest influence on the extraction yield of ASO, followed by extraction time. The solid-liquid ratio had the least

influence on the extraction efficacy. The highest extraction efficacy was obtained at a solid to liquid ratio of 1:60 with extraction at 50 °C carried out for 4 h. The highest extraction efficacy was 45.57%.

3.2. Physicochemical properties of seed oil

The acid and peroxide values of 21 ASO samples from seven different geographical sources were determined. The results (Table 1) showed that the acid and peroxide values of different ASO varied considerably. The acid value ranged from 0.782 to 1.342 mg/g, and the peroxide value ranged from 0.081 to 0.118 g/100 g. All of the values met the Chinese food safety standard. Analysis of different groups suggested that there were also significant differences between groups. The acid value ranged from 0.998 to 1.261 mg/g and the peroxide value from 0.0863 to 0.109 g/100 g. However, the acid and peroxide values do not have obvious north-south regional characteristics.

3.3. FA content and composition

The optimized Soxhlet extraction method was used to determine the

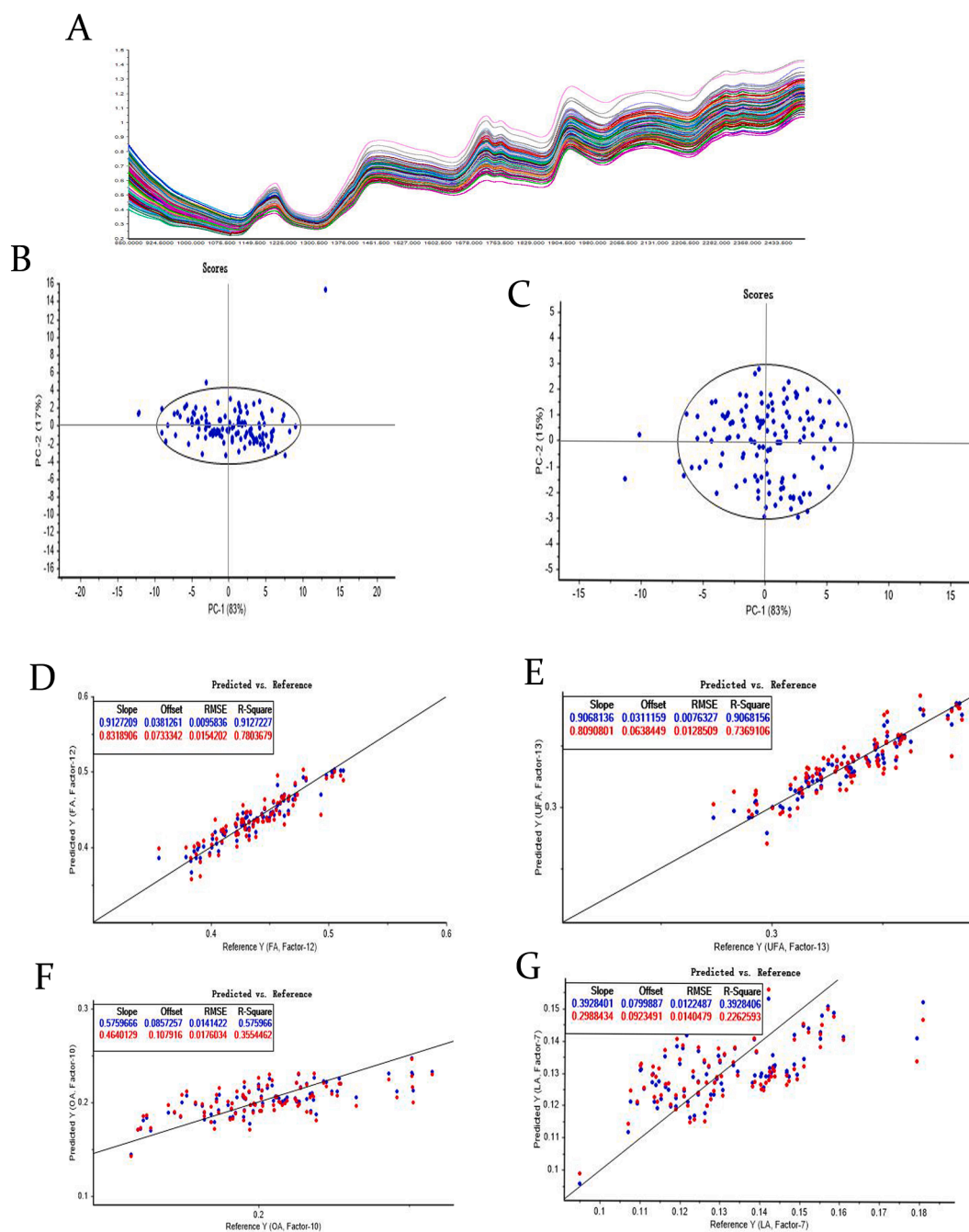


Fig 3. PLS models of ASO. **A.** Near-infrared spectra of 130 samples; **B.** PCA based on the average of the chemical data; **C.** PCA based on the average of the spectral data. **D.** PLS models of FA; **E.** PLS models of UFA; **F.** PLS models of OA; **G.** PLS models of LA.

content of the ASO. The results (Fig. 2A) showed that the maximum and minimum contents were 51.27% and 34.93%, respectively, thus showing a considerable variation in the ASO content of the different varieties. Overall, the ASO content was mainly distributed between 40% and 50%, with only 18 germplasms having ASO contents lower than 40%, and five germplasms having ASO contents higher than 50%. The comparison among different groups suggested that the average ASO values of the other six groups were similar, ranging from 42.77% to 44.47%, except for Jingmen, which had an average ASO content of 37.95% (Fig. 2B). In the five germplasms with ASO content higher than 50%, four germplasms were from Yuanjiang and one was from Taojiang. Within the same group, the ASO content also varied considerably, with the difference between the highest and lowest values in YJ of 16.3%, and differences higher than 10% in four groups.

GC-MS analysis (Fig. 2C) showed that ASO contained 11 types of FAs, including myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1), heptadecanoic acid (C17:0), heptadecenoic acid (C17:1), stearic acid (C18:0), OA (C18:1), LA (C18:2), linolenic acid (C18:3), eicosanoic acid (C20:0), and eicosenoic acid (C20:1). The main components were palmitic acid, stearic acid, OA, and LA, with a total content of approximately 97% of ASO. Palmitic acid and stearic acid concentrations ranged from 16.243% to 25.617% and from 3.204% to 10.534%, respectively. OA ranged from 37.185% to 53.127% and LA ranged from 14.461% to 37.975%, with the exception of JM-6 which had an LA content less than 23%. As for the other seven FAs, only linolenic acid was higher than 1% in a few germplasms, other six FAs, the content was less than 1% or even 0.5%. UFAs are vital for evaluating the quality of edible oil, and there are six types of UFAs in ASO. The

Table 2
PLS models of FA, UFA, OA, LA with different method.

Compound	Pretreatment method	R ² _{cal}	RMSEC	R ² _{cv}	RMSEP
FA	Original	0.867	0.0118	0.762	0.0162
	SNV	0.861	0.0121	0.762	0.0158
	MSC	0.906	0.0099	0.763	0.0156
	D1	0.913	0.0095	0.770	0.0156
	D2	0.853	0.0124	0.681	0.0184
	D1 + SNV	0.913	0.0096	0.780	0.0154
	D1 + MSC	0.906	0.0099	0.763	0.0155
UFA	Original	0.857	0.0095	0.735	0.0131
	SNV	0.833	0.0102	0.709	0.0135
	MSC	0.830	0.0103	0.735	0.0131
	D1	0.907	0.0076	0.737	0.0129
	D2	0.843	0.0099	0.626	0.0155
	D1 + SNV	0.892	0.0082	0.729	0.0134
	D1 + MSC	0.905	0.0077	0.724	0.0132
OA	Original	0.594	0.0138	0.328	0.0179
	SNV	0.570	0.0142	0.302	0.0182
	MSC	0.576	0.0141	0.355	0.0176
	D1	0.376	0.0172	0.249	0.0192
	D2	0.623	0.0133	0.207	0.0193
	D1 + SNV	0.517	0.0151	0.265	0.0188
	D1 + MSC	0.558	0.0144	0.353	0.0179
LA	Original	0.393	0.0122	0.226	0.0140
	SNV	0.378	0.0124	0.222	0.0136
	MSC	0.372	0.0125	0.226	0.0139
	D1	0.301	0.0131	0.169	0.0147
	D2	0.183	0.0142	0.130	0.0149
	D1 + SNV	0.406	0.0210	0.209	0.0140
	D1 + MSC	0.213	0.0139	0.129	0.0148

FA: Fatty acids.

UFA: Unsaturated fatty acids.

OA: Oleic acid.

LA: Linoleic acid.

PLS: Partial least squares.

SNV: Standard normal variate.

MSC: Multiplicative scatter correction.

D1: First derivative (Savitzky-Golay).

D2: Second derivative (Savitzky-Golay).

R²_{cal}: Coefficients of determination for calibration.

RMSEC: Root mean square error of calibration.

R²_{cv}: Coefficient of determination for cross-validation.

RMSEP: Root mean square error of prediction.

UFAs account for 63.849–79.059% of ASO (Fig. 2D), only JM-6, YJ1-23, and YC-8 contained less than 74%, and UFA accounted for 74–78% of ASO in all of the samples.

3.4. Spectral analysis and model checking

The near-infrared original spectra of the seeds of 130 different *A. trifoliata* germplasm (Fig. 3A) showed that the peaks, troughs, and trends of the waves of the 130 materials were similar at different wavelengths. 13 samples scattering out of the principal constituent were removed according to PCA results (Fig. 3B and C) of the near-infrared spectral and chemical data. Ninety samples were used for calibration and twenty-seven samples were used for examination. The FAs, UFAs, OA, and LA models established by the original spectra and six processed spectra (Table 2) showed that the best model of FA was established after D1 + SNV treatment (Fig. 3D); R²_{cal} and R²_{cv} were 0.913 and 0.780, respectively, and RMSEC and RMSEP were 0.0096 and 0.0154, respectively. The UFAs model treated by D1 was the best model (Fig. 3E), R²_{cal} and R²_{cv} were 0.907 and 0.737, respectively, and RMSEC and RMSEP were 0.0076 and 0.0129, respectively. R²_{cal} of the seven FA models ranged from 0.853 to 0.913, R²_{cal} of UFA models ranged from 0.830 to 0.907, R²_{cal} of OA and LA ranged from 0.376 to 0.623 and 0.182 to 0.431, respectively, and the best pretreatment method was MSC and original spectral (Fig. 3F and G).

FA, UFA, OA, and LA of twenty-seven samples were predicted using the optimal model of the corresponding components (Table 3, Table S2). The predicted FA value was between 37.70% and 50.87% (average of 43.68%), and the actual values were between 35.29% and 51.24% (average of 43.26%). The maximum and minimum absolute errors were 2.411% and 0.021%, respectively, and all relative errors were less than 5%. The UFA predicted value was between 28.71% and 39.31% (average of 33.33%), the actual values were between 26.25% and 36.24% (average of 32.97%), and the absolute errors ranged from 0.014% to 3.123% (average of 0.870%); relative errors ranged from 0.039% to 10.057% (average of 2.697%). Only three samples (TJ1-23, CS20-9, and TJ1-14) had relative errors that were higher than 5.0%. The OA and LA predicted values ranged from 15.582% to 24.143% and from 10.652% to 15.062%, respectively; the average of OA and LA was 20.032% and 13.176%, and the actual values were 12.989–23.002 % (average of 20.045%) and 10.678–15.047% (average of 12.723%). The relative errors of OA ranged from 0.345% to 27.341% (average of 6.137 %), and those of LA ranged from 0.742% to 19.694% (average of 7.011%). T-test showed that there were no significant differences between the predicted and actual values of FA and UFA. Additionally, correlation analysis showed that the predicted values were significantly correlated with the actual values of FA and UFA, with the correlation coefficients of 0.946 and 0.909, respectively. However, for OA and LA, the correlation coefficients were only 0.749 and 0.605, respectively. Therefore, FA and UFA can be accurately measured by the model, whereas only preliminary predictions can be obtained for OA and LA.

4. Discussion

High oil content and high yield are prerequisites for oil crops. *A. trifoliata* had a high seed yield (400 kg/667 m²), and the highest oil content was 51.27% (Fig. 2A). The oil content of *A. trifoliata* is higher than that of soybean (Yao et al., 2020) and cottonseed (Quampah et al., 2012), and is similar to those of rapeseed (Tang et al., 2019) and sesame (Were, Onkware, Gudu, Welander, & Carlsson, 2006). Oil content is a complex trait, and is affected by many factors such as temperature (Baux, Hebeisen, & Pellet, 2008), water (Bettaieb, Knioua, Hamrouni, Limam, & Marzouk, 2011), and nutrients (Gong, Guo, Wan, Liang, & Jiang, 2013). In this study, the difference in ASO in different regions was not reflected by the average ASO content of the seven groups, so that the proportion of the ASO higher than the average of the 130 ASO contents was used to compare the differences between different groups (Table S3). The average of the 130 ASO contents was 43.44%, LY had 71.43% of samples with the oil content higher than 43.44%, JM did not have samples with the oil content higher than 43.44%, and the oil contents of TJ, YJ, CS, YC, and XA were 56%, 50%, 52%, 60.71%, and 29.41%, respectively. Therefore, plants from different regions had varied contents of ASO. This conclusion was the same as that of Li et al. (2020). As for other factors that affect the content of ASO, follow-up studies are still needed. To date, *A. trifoliata* has not been bred for high ASO content, however, with future research the value of *A. trifoliata* as an oil crop may be greatly enhanced.

The physicochemical properties of plant oil are an important indicator of whether or not it can be used as edible oil. The main physicochemical properties of edible oil include acid value, peroxide value, saponification value, iodine value, and smoke point. Among these, the acid and peroxide values are the most critical in determining the standard of the edible oil in China (GB 2716-2018). Chinese food safety standards stipulate that the acid and peroxide values cannot be higher than 3.0 mg/g and 0.25 g/100 g, respectively. The maximum acid and peroxide contents of ASO are 1.342 mg/g and 0.114 g/100 g, respectively, which are far below Chinese national standard. Therefore, in terms of acid and peroxide values, ASO can be used as edible oil, which is consistent with the findings of Li et al. (2020). We did not replicate the 21 samples, but we treated three varieties in the same group as replicates, and the results showed that there were significant differences in

Table 3
The predicted value and actual value of FA and UFA in validation set.

Sample	FA					UFA				
	Predicted Value	Actual Value	Absolute Errors	Relative errors	Correlation coefficient	Predicted Value	Actual Value	Absolute Errors	Relative errors	Correlation coefficient
XA-2	43.890%	42.380%	1.510%	3.562%	0.946	32.862%	32.635%	0.227%	0.694%	0.909
XA-7	42.720%	42.615%	0.105%	0.246%		31.857%	31.896%	0.039%	0.122%	
XA-16	43.025%	42.760%	0.265%	0.619%		33.483%	32.914%	0.569%	1.730%	
YC-3	46.284%	45.627%	0.657%	1.440%		35.824%	35.502%	0.323%	0.909%	
YC-5	47.144%	45.894%	1.249%	2.722%		35.931%	34.852%	1.079%	3.097%	
YC-11	40.306%	40.681%	0.375%	0.921%		31.552%	31.069%	0.483%	1.554%	
YC-15	44.713%	45.248%	0.534%	1.181%		34.520%	36.241%	1.721%	4.748%	
YC-16	45.628%	45.648%	0.021%	0.045%		34.651%	33.813%	0.838%	2.477%	
JM-3	38.777%	38.530%	0.247%	0.642%		29.841%	28.752%	1.088%	3.785%	
TJ1-3	42.702%	42.753%	0.052%	0.121%		33.574%	32.929%	0.645%	1.960%	
TJ1-24	46.568%	44.615%	1.953%	4.378%		35.224%	35.041%	0.183%	0.521%	
TJ1-9	42.477%	41.866%	0.610%	1.458%		31.757%	30.963%	0.794%	2.565%	
TJ1-14	50.873%	51.241%	0.368%	0.718%		39.306%	37.042%	2.265%	6.114%	
TJ1-12	42.760%	42.257%	0.503%	1.190%		32.453%	32.327%	0.126%	0.391%	
TJ1-29	45.668%	43.772%	1.896%	4.331%		34.593%	34.847%	0.254%	0.730%	
LY-7	43.311%	43.759%	0.447%	1.022%		32.956%	33.266%	0.310%	0.932%	
YJ1-23	43.782%	44.383%	0.600%	1.353%		34.179%	31.056%	3.123%	10.057%	
YJ1-21	41.400%	39.431%	1.968%	4.992%		31.111%	31.307%	0.196%	0.627%	
YJ1-20	41.037%	42.147%	1.110%	2.633%		31.637%	30.482%	1.155%	3.790%	
YJ2-2	47.340%	46.613%	0.726%	1.558%		36.190%	37.607%	1.417%	3.769%	
YJ1-22	43.520%	42.622%	0.898%	2.107%		33.237%	33.974%	0.737%	2.170%	
CS20-9	37.704%	35.293%	2.411%	6.833%		28.707%	26.253%	2.454%	9.346%	
CS20-21	44.543%	44.964%	0.421%	0.937%		32.625%	32.996%	0.371%	1.125%	
CS20-10	45.091%	47.237%	2.146%	4.544%		34.731%	34.745%	0.014%	0.039%	
CS20-5	43.341%	42.209%	1.132%	2.683%		32.344%	32.982%	0.638%	1.935%	
CS20-4	38.611%	36.515%	2.096%	5.739%		29.737%	28.520%	1.217%	4.268%	
CS19-7	46.119%	46.875%	0.756%	1.613%		34.916%	36.133%	1.216%	3.366%	
Mean	43.679%	43.257%	0.928%	2.207%		33.326%	32.968%	0.870%	2.697%	

the acid and peroxide values among the different geographical regions. This indicates that both species and geographical environment affect the acid and peroxide values of ASO, as has been confirmed by Li et al. (2020).

The FA composition of plant oil is another important indicator of whether it can be used as edible oil. FAs can be divided into saturated FAs and UFAs overall, depending on whether they contain unsaturated oil bonds. UFA content is considered to be an important index for evaluating the quality of edible oil. However, if the content of LA and α -linolenic acid is too high, the oil can deteriorate easily. In oil crop breeding, the quality is improved mainly by increasing the OA content to increase the content of UFAs (Zuil, Izquierdo, Luján, Cantarero, & Aguirrezábal, 2012; Kadirvel et al., 2020). UFA content accounted for 63.849–79.059% of ASO content; OA and LA accounted for 37.19–53.13% and 14.46–37.98% of the ASO content, and JM-6 had the maximum content of palmitic acid and stearic acid and the minimum content of LA in 130 samples. We believe that JM-6 is an outlier and should be excluded; therefore, the UFA, OA, and LA contents should be 72.35–79.06%, 37.15–53.13%, and 23.614–37.975%, respectively. Stearic acid is considered unhealthy, and if the amount of stearic acid is too high, the oil cannot be used as edible oil. After the removal of JM-6 from the analyses, stearic acid accounted for 3.804–7.288% (average of 5.023%) of ASO, which is acceptable for edible oil. Based on the high UFA, OA, and low stearic acid contents, ASO is of a high quality. ASO FA compositions are similar to the composition of normal peanuts (Isleib, Pattee, Sanders, Hendrix, & Dean, 2006), therefore, based on FA compositions, ASO can be used as edible oil.

Modern plant breeding often requires large sample phenotype screening, and quick and accurate identification of the phenotype is important. At present, there are many methods to determine the content of plant oils, such as Soxhlet extraction, ultrasonic-assisted extraction, meteorological chromatography, nuclear magnetic resonance, and NIRS. NIRS is used to measure the oil content of many oil crops. Niewitzki, Tillmann, Becker, and Möllersl. (2010) established a high-

throughput method for the determination of FAs in rapeseed based on NIRS; Ribeiro, Zamora, Maia, Ramos, and Netto (2013) used NIRS and multivariate calibration to measure linolenic and LA content in flax; and Ferreira et al. (2013) used NIRS and mid-infrared spectroscopy to determine quality parameters in soybean. R^2_{cal} and R^2_{cv} , RMSEC, and RMSEP are the indices used to evaluate the PLS model. Data distribution, seed size, seed character, and seed coat color, can influence the model (Wang, Liu, & Ren, 2014; Cui, Cheng, & Peng, 2016). *A. trifoliata* has large seeds, irregular seed shape, black seed coat is not conducive to the construction of the PLS model (Fig. S1). The R^2_{cal} values of the ASO and UFA models were 0.913 and 0.907, and RMSEC was 0.0096 and 0.0076, respectively. The relative error between the predicted and experimental values of the prediction set ranged from 0.045% to 6.833% (average of 2.207%) and 0.039% to 10.057% (average of 2.697%). This shows that the PLS model can accurately predict the ASO and UFA content of *A. trifoliata* seeds. However, the PLS models of OA and LA are not effective, with R^2_{cal} values of 0.576 and 0.393, respectively. Therefore, this model can only be used to preliminarily screen OA and LA contents. Based on NIRS, the contents of ASO and UFA can be accurately determined. Subsequent studies can be carried out using NIRS to determine the protein, cellulose, moisture, and other components of *A. trifoliata* seeds. If *A. trifoliata* seed powder was used in the model, the model's accuracy would be higher; however, as this would destroy the viability of the seed, we did not pursue a high R^2 .

5. Conclusion

In this study, the acid and peroxide values of ASO meet the safety standards for edible oil. Using the optimal Soxhlet extraction conditions, the highest content of ASO in 130 samples was 51.27%, and 11 types of FAs were detected by GC-MS. The four main components were palmitic acid, stearic acid, OA, and LA, which accounted for 97% of the total content of ASO. NIRS could be used for *A. trifoliata* seeds, the R^2_{cal} for the PLS models of ASO and UFA were 0.913 and 0.907, and the models

could accurately predict the content of ASO and UFA. However, the PLS models of OA and LA were not effective, and R^2_{cal} was only 0.576 and 0.393. Thus, the PLS models can be used for preliminary screening. The results of this study show that *A. trifoliata* has excellent prospects for use as edible oil based on the ASO content, physicochemical properties, and ingredients, and provides excellent germplasm. Rapid and low-cost detection methods are also available for the next step of breeding *A. trifoliata* with high ASO content.

CRedit authorship contribution statement

Yicheng Zhong: Data curation, Formal analysis, Writing - original draft. **Zhenqian Zhang:** Writing - review & editing, Supervision. **Jing Chen:** Methodology. **Juan Niu:** Resources. **Yaliang Shi:** Visualization. **Yue Wang:** Supervision. **Tianxin Chen:** Validation. **Zhimin Sun:** Investigation. **Jianhua Chen:** Funding acquisition, Project administration. **Mingbao Luan:** Funding acquisition, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We are very grateful to Zhang Yanxin (Oil Crops Research Institute, Chinese Academy of Agricultural Sciences) and Liu Liangliang (Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences) for their help in this study. We would like to thank Editage (www.editage.cn) for English language editing.

Funding

This research was financially supported by the Central Public-interest Scientific Institution Basal Research Fund (1610242020010) and Agricultural Science, Technology Innovation Program (ASTIP) of CAAS (Grant No. 2017IBFC).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.fochx.2021.100131>.

References

- Akkaya, R. (2018). Prediction of fatty acid composition of sunflower seeds by near-infrared reflectance spectroscopy. *Journal of Food Science and Technology*, 55(6), 2318–2325. <https://doi.org/10.1007/s13197-018-3150-x>.
- Albanell, E., Martínez, M., De Marchi, M., & Manuelian, C. L. (2021). Prediction of bioactive compounds in barley by near-infrared reflectance spectroscopy (NIRS). *Journal of Food Composition and Analysis*, 97, 103763. <https://doi.org/10.1016/j.jfca.2020.103763>.
- Allendorf, M., Subramanian, A., & Rodriguez-Saona, L. (2012). Application of a Handheld Portable Mid-Infrared Sensor for Monitoring Oil Oxidative Stability. *Journal of the American Oil Chemists' Society*, 89(1), 79–88. <https://doi.org/10.1007/s11746-011-1894-9>.
- Baux, A., Hebeisen, T., & Pellet, D. (2008). Effects of minimal temperatures on low-linolenic rapeseed oil fatty-acid composition. *European Journal of Agronomy*, 29(2–3), 102–107. <https://doi.org/10.1016/j.eja.2008.04.005>.
- Bettaieb, I., Knoua, S., Hamrouni, I., Limam, F., & Marzouk, B. (2011). Water-deficit impact on fatty acid and essential oil composition and antioxidant activities of cumin (*Cuminum cyminum* L.) aerial parts. *Journal of agricultural and food chemistry*, 59(1), 328–334. <https://doi.org/10.1021/jf1037618>.
- Casale, M., Zunin, P., Cosulich, M. E., Pitarino, E., Perego, P., & Lanteri, S. (2010). Characterisation of table olive cultivar by NIR spectroscopy. *Food Chemistry*, 122(4), 1261–1265. <https://doi.org/10.1016/j.foodchem.2010.03.080>.
- Chen, J., Yu, H., Jiang, D., Zhang, Y., & Wang, K. (2021). A novel NIRS modelling method with OPLS-SPA and MIX-PLS for timber evaluation. *Journal of Forestry Research*. <https://doi.org/10.1007/s11676-021-01314-y>.
- Cui, H., Cheng, F., & Peng, L. (2016). Determination of the Fatty Acid Composition in Tree Peony Seeds Using Near-Infrared Spectroscopy. *Journal of the American Oil Chemists' Society*, 93(7), 943–952. <https://doi.org/10.1007/s11746-016-2839-0>.
- Ferreira, D. S., Pallone, J. A. L., & Poppi, R. J. (2013). Fourier transform near-infrared spectroscopy (FT-NIRS) application to estimate Brazilian soybean [*Glycine max* (L.) Merrill] composition. *Food Research International*, 51(1), 53–58. <https://doi.org/10.1016/j.foodres.2012.09.015>.
- Gong, Y., Guo, X., Wan, X., Liang, Z., & Jiang, M. (2013). Triacylglycerol accumulation and change in fatty acid content of four marine oleaginous microalgae under nutrient limitation and at different culture ages. *Journal of Basic Microbiology*, 53(1), 29–36. <https://doi.org/10.1002/jobm.v53.1.10.1002/jobm.201100487>.
- Hom, N. H., Becker, H. C., & Möllers, C. (2007). Non-destructive analysis of rapeseed quality by NIRS of small seed samples and single seeds. *Euphytica*, 153(1–2), 27–34. <https://doi.org/10.1007/s10681-006-9195-3>.
- Isleib, T. G., Pattee, H. E., Sanders, T. H., Hendrix, K. W., & Dean, L. O. (2006). Compositional and sensory comparisons between normal- and high-oleic peanuts. *Journal of agricultural and food chemistry*, 54(5), 1759–1763. <https://doi.org/10.1021/jf052353t>.
- Jarvis, B. A., Romsdahl, T. B., McGinn, M. G., Nazareus, T. J., Cahoon, E. B., Chapman, K. D., Sedbrook, J. C. (2021). CRISPR/Cas9-Induced *fad2* and *rod1* Mutations Stacked With *fae1* Confer High Oleic Acid Seed Oil in Pennycress (*Thlaspi arvense* L.). *Front Plant Sci*. Apr 22;12:652319. DOI: 10.3389/fpls.2021.652319.
- Kadirvel, P., Veerajou, C., Senthilvel, S., Yadav, P., Usha Kiran, B., Shaik, M., et al. (2020). Marker-assisted selection for fast-track breeding of high oleic lines in safflower (*Carthamus tinctorius* L.). *Industrial Crops and Products*, 158, 112983. <https://doi.org/10.1016/j.indcrop.2020.112983>.
- Kliem, K. E., Shingfield, K. J., Humphries, D. J., & Givens, D. I. (2011). Effect of replacing calcium salts of palm oil distillate with incremental amounts of conventional or high oleic acid milled rapeseed on milk fatty acid composition in cows fed maize silage-based diets. *Animal*, 5(08), 1311–1321. <https://doi.org/10.1017/S1751731111000310>.
- Liao, B. S., Yin, Y., & Ma, N. (2018). Review and future prospects of oil crops industry development in China. *Journal of Agriculture*, 8(01), 107–112.
- Liao, Y., Li, C. Z., Yu, L. Y. D., Huang, G. W., & Yan, R. L. (2021). Research progress of woody oil resources in China. *Journal of the Chinese Cereals and Oils Association*, 1–16.
- Li, Y. Y., Chen, S. S., Liu, H. C., Zuo, Z. M., Luo, P. G., & Zheng, Z. (2020). Evaluation for *Akebia trifoliata* (Thunb.) Koidz. var. *australis* (Diels) Rehd. germplasm resources based on physicochemical property and fatty acid components of seed oil. *Chinese Journal of Oil Crop Sciences*, 42, 843–850.
- Li, W. Y., Wu, H. S., YU, H. Z. (2021). Comparison of Different Extraction Methods of *Akebia trifoliata* Oil and Optimization of Supercritical CO₂ Extraction Process. *Science and Technology of Food Industry*, 42, 203–208.
- National Health Commission of the People's Republic of China. National food safety standard-plant oil, GB 2716-2018, Beijing, Standards Press of China, 2018.
- Niewietzki, O., Tillmann, P., Becker, H. C., & Möllers, C. (2010). A new near-infrared reflectance spectroscopy method for high-throughput analysis of oleic acid and linolenic acid content of single seeds in oilseed rape (*Brassica napus* L.). *Journal of agricultural and food chemistry*, 58(1), 94–100. <https://doi.org/10.1021/jf9028199>.
- Niu, J., Shi, Y., Huang, K., Zhong, Y., Chen, J., Sun, Z., et al. (2020). Integrative transcriptome and proteome analyses provide new insights into different stages of *Akebia trifoliata* fruit cracking during ripening. *Biotechnol Biofuels*, 13(1). <https://doi.org/10.1186/s13068-020-01789-7>.
- Quampah, A., Huang, Z. R., Wu, J. G., Liu, H. Y., Li, J. R., Zhu, S. J., et al. (2012). Estimation of oil content and fatty acid composition in cottonseed kernel powder using near infrared reflectance spectroscopy. *Journal of the American Oil Chemists' Society*, 89(4), 567–575. <https://doi.org/10.1007/s11746-011-1945-2>.
- Ribeiro, J. S., Salva, T. J. G., & Silvarolla, M. B. (2021). Prediction of a wide range of compounds concentration in raw coffee beans using NIRS, PLS and variable selection. *Food Control*, 125, 107967. <https://doi.org/10.1016/j.foodcont.2021.107967>.
- Ribeiro, L. F., Zamora, P. G. P., Maia, B. H. L. N. S., Ramos, L. P., & Netto, A. B. P. (2013). Prediction of linolenic and linoleic fatty acids content in flax seeds and flax seeds flours through the use of infrared reflectance spectroscopy and multivariate calibration. *Food Research International*, 51(2), 848–854. <https://doi.org/10.1016/j.foodres.2013.01.061>.
- Roberto, M., Hagos, B. D., Ron, P. A. M. H., Peishih, L., Serena, F., Danilo, M., & Riccardo, M. (2021). Pine nut species recognition using NIR spectroscopy and image analysis. *Journal of Food Engineering*, 292, 110357. <https://doi.org/10.1016/j.jfoodeng.2020.110357>.
- Sun, X. M., Zhang, L. X., Li, P. W., Xu, B. C., Ma, F., Zhang, Q., et al. (2015). Fatty acid profiles based adulteration detection for flaxseed oil by gas chromatography mass spectrometry. *LWT - Food Science and Technology*, 63(1), 430–436. <https://doi.org/10.1016/j.lwt.2015.02.023>.
- Tang, M. Q., Zhang, Y. Y., Liu, Y. Y., Tong, C. B., Cheng, X. H., Zhu, W., et al. (2019). Mapping loci controlling fatty acid profiles, oil and protein content by genome-wide association study in *Brassica napus*. *The Crop Journal*, 7(2), 217–226. <https://doi.org/10.1016/j.cj.2018.10.007>.
- Wang, J., Liu, H., & Ren, G. (2014). Near-infrared spectroscopy (NIRS) evaluation and regional analysis of Chinese faba bean (*Vicia faba* L.). *The Crop Journal*, 2(1), 28–37. <https://doi.org/10.1016/j.cj.2013.10.001>.
- Wang, Q. K., Xing, H. X., Liu, X. G., Mao, L. L., Wei, Z., Zhang, H. J., et al. (2020). Estimation of protein and fatty acid composition in shell-intact cottonseed by near infrared reflectance spectroscopy. *Journal of the American Oil Chemists' Society*, 97(4), 331–340. <https://doi.org/10.1002/aocs.v97.4.10.1002/aocs.12312>.
- Were, B. A., Onkware, A. O., Gudu, S., Welander, M., & Carlsson, A. S. (2006). Seed oil content and fatty acid composition in East African sesame (*Sesamum indicum* L.)

- accessions evaluated over 3 years. *Field Crops Research*, 97(2-3), 254–260. <https://doi.org/10.1016/j.fcr.2005.10.009>.
- Xie, J., Li, X. H., Zhang, C. J., Ouyang, H. N., & Xiao, Y. P. (2006). Distribution of *Akebia trifoliata* (Thunb.) Koidz wild resources. *Journal of Shaanxi Normal University: Natural Science Edition*, 34, 272–274.
- Yao, Y., You, Q., Duan, G., Ren, J., Chu, S., Zhao, J., Li, X., Zhou, X., Jiao, Y. (2020). Quantitative trait loci analysis of seed oil content and composition of wild and cultivated soybean. *BMC Plant Biology*, 51. DOI: 10.1186/s12870-019-2199-7.
- Yusnaini, R., Ikhsan, I., Idroes, R., Munawar, A. A., Arabia, T., Saidi, N., et al. (2021). Near-infrared spectroscopy (NIRS) as an integrated approach for rapid classification and bioactive quality evaluation of intact *Feronia limoni*. *IOP Conference Series: Earth and Environmental Science*, 667(1), 012028. <https://doi.org/10.1088/1755-1315/667/1/012028>.
- Zhang, L. X., Li, P. W., Sun, X. M., Hu, W., Wang, W. P., Zhang, Q., et al. (2014). Untargeted fatty acid profiles based on the selected ion monitoring mode. *Analytica Chimica Acta*, 839, 44–50. <https://doi.org/10.1016/j.aca.2014.06.040>.
- Zhao, S. Q., Hu, J. N., Zhu, X. M., Bai, C. Q., Peng, H. L., Xiong, H., et al. (2014). Characteristics and feasibility of trans-free plastic fats through Lipozyme TL IM-catalyzed interesterification of palm stearin and *Akebia trifoliata* variety Australis seed oil. *Journal of agricultural and food chemistry*, 62(14), 3293–3300. <https://doi.org/10.1021/jf500267e>.
- Zhong, W. M., & Ma, Y. H. (2016). Analysis and evaluation of nutritional components in *Akebia trifoliata* seeds. *Southwest China Journal of Agricultural Sciences*, 29(01), 169–173.
- Zhou, N. N. (2018). *Research on extraction and bioactivity of Akebia trifoliata seed oil*. Central South University of Forestry and Technology.
- Zuil, S. G., Izquierdo, N. G., Luján, J., Cantarero, M., L.A.N. Aguirrezábal, L. A. N. (2012). Oil quality of maize and soybean genotypes with increased oleic acid percentage as affected by intercepted solar radiation and temperature. *Field Crops Research*, 127, 203–214. DOI: 10.1016/j.fcr.2011.11.019.