RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2024, **14**, 29319

Safety evaluation of *Plukenetia volubilis* seeds: a metabolomic profiling and network toxicology approach†

Vinh-Tuyen T. Le,[ab] Thanh Hao Huynh,[c] Lo-Yun Chen,[a] Muhammad Riki Shindi Prariastiya,[ad] Hung-Yu Lin,[ef] Kuei-Hung Lai,[agh] Ya-Lin Lee,[i] Lih-Geeng Chen[hj] and Ching-Chiung Wang *[acgh]

Sacha Inchi (*Plukenetia volubilis*) seeds and oil have been integrated into daily diets. However, scientific reports have raised concerns regarding potential health risks associated with saponins and alkaloids in this seeds. This study employed a combination analysis using proton-NMR, GC-MS, LC-QTOF, and GNPS molecular networking to evaluate the chemical composition of these seeds. *In silico* toxicology analysis and *in vitro* cytotoxicity assays were conducted to investigate the potential toxicity effects of Sacha Inchi seeds and their contained metabolites. The results revealed that major components of these seeds are oils (linoleic, linolenic, and oleic acids) and sugars, with minor amounts of phytosterols and trigonelline, a pyridine alkaloid. GNPS analysis suggested the absence of saponins, instead, it identified trigonelline and a few other nitrogen-containing metabolites (amino acids and oligopeptides). *In silico* toxicology analysis indicated that this sample did not exhibit toxicity. Furthermore, *in vitro* cytotoxicity screening demonstrated no cytotoxic effects against NIH-3T3 cells, even at 400 μg mL^{-1}. In general, these findings collectively indicated the absence of saponins, the presence of phytosterols and trigonellin (a pyridine alkaloid), and a low safety risk related to saponin and alkaloid content in the Sacha Inchi seeds.

Received 22nd May 2024
Accepted 5th August 2024

DOI: 10.1039/d4ra03767g

rsc.li/rsc-advances

Introduction

Sacha Inchi (*Plukenetia volubilis* L.) is a perennial plant belonging to the Euphorbiaceae family and primarily found in the Amazon rainforest. Traditionally, the seeds are roasted and leaves are cooked and integrated into daily diets. In the Amazon region, Sacha Inchi seeds have been used as a traditional remedy for the treatment of rheumatic problems and muscle aches. In the pharmaceuticals, Sacha Inchi oil has historical use in skincare treatments, primarily for skin softening, wound healing, managing insect bites, and addressing skin infections.[1] Over recent years, it has gained more recognition and popularity in global markets, often in the form of encapsulated oil. Owing to its suitable adaption to the diverse ecosystems of the rainforest, Sacha Inchi was also widely spread and cultivated in other parts of the world, such as Southeast Asia and even parts of China, making itself a substantial potentially economically valuable crop.[2]

The Sacha Inchi chemical composition primarily contains lipids, proteins, and carbohydrates, found in the seeds, whereas the seed shells and leaves are predominantly characterized by the presence of phenols, flavonoids, tannins, steroids, and terpenoids. Raw Sacha Inchi seeds have approximately 22–30% protein content; however, the defatted seeds, obtained after oil extraction, are notably rich in protein, with up to 53–59% protein.[3,4] The primary soluble protein components found in seeds are albumins, glutelins, globulins, and prolamins. Additionally, these seeds comprise essential amino acids such as leucine, tyrosine, isoleucine, lysine, and tryptophan, with approximate concentrations of 64, 55, 50, 43, and 43 mg per g of protein, respectively.[5] Lipids represent the primary component

[a]*PhD Program in Clinical Drug Development of Herbal Medicine, College of Pharmacy, Taipei Medical University, Taipei 110, Taiwan*
[b]*Department of Pharmacognosy – Traditional Pharmacy – Pharmaceutical Botany, College of Pharmacy, Can Tho University of Medicine and Pharmacy, Can Tho 941, Vietnam*
[c]*School of Pharmacy, College of Pharmacy, Taipei Medical University, Taipei 110, Taiwan. E-mail: crystal@tmu.edu.tw*
[d]*Pharmacy Program – College of Health Sciences Darul Azhar Batulicin, Tanah Bumbu, South Borneo 722, Indonesia*
[e]*Department of Applied Chemistry, Chaoyang University of Technology, Taichung 413, Taiwan*
[f]*Department of Food Science, Tunghai University, Taichung 407, Taiwan*
[g]*Graduate Institute of Pharmacognosy, College of Pharmacy, Taipei Medical University, Taipei 110, Taiwan*
[h]*Traditional Herbal Medicine Research Center, Taipei Medical University Hospital, Taipei 110, Taiwan*
[i]*Crop Genetic Resources and Biotechnology Division, Taiwan Agricultural Research Institute, Taichung 413, Taiwan*
[j]*Department of Microbiology, Immunology and Biopharmaceuticals, College of Life Sciences, National Chiayi University, Chiayi 600, Taiwan*

† Electronic supplementary information (ESI) available: ^{1}H NMR, GC-MS, and MS^2 spectra, and GNPS classification diagrams. See DOI: https://doi.org/10.1039/d4ra03767g

in Sacha Inchi seeds, ranging from 33 to 54%. The oil fraction comprises neutral lipids (97.2%), free fatty acids (1.2%), and phospholipids (0.8%). These lipids exhibit high unsaturation, with only 6.8–9.1% of fatty acids being saturated. The seed fatty acid composition includes 77.5–84.4% polyunsaturated fatty acids and 8.4–13.2% monounsaturated fatty acids. The predominant fatty acid in oil is α-linolenic acid (46.8–50.8%), followed by linoleic acid (33.4–36.2%), and oleic acid (8.7–9.6%).[3,6,7] The carbohydrate content of Sacha Inchi seeds varies between 13.4 and 30.9%, with limited available information on the carbohydrate composition of Sacha Inchi.[3]

Many plants used in human diets contain varying levels of secondary metabolites, which, if consumed in large quantities, can potentially have adverse effects. While these substances are often effective against insect herbivores, they are typically not present in quantities that would lead to acute toxic effects in humans who maintain a diverse and balanced diet. However, the safety and potential toxicity of any food consumed must be thoroughly evaluated. Although not all anti-nutritional components in plant-based foods are necessarily harmful, some can reduce the effective absorption of essential inorganic micronutrients and hinder the digestion of macronutrients.[8] Certain chemical substances in Sacha Inchi seeds and leaves, particularly saponins and alkaloids, may pose potential risks to human health.[9,10] Previous studies have indicated that saponins possess cytotoxic and hemolytic effects and can inhibit protease activities.[11,12] The toxicity of saponins varies widely depending on the animal and the conditions under which they are consumed. Saponins are highly toxic to fish. For warm-blooded animals, the toxicity of orally administered saponins is much lower. Research confirms that saponins affect the permeability of intestinal mucosa cells, suggesting that they may alter the absorption and excretion processes in the small intestine. This can lead to disruptions in the absorption of essential micronutrients or the uptake of certain undesirable compounds with allergenic effects. An example of the harmful effects of saponins is pseudoaldosteronism caused by glycyrrhizin.[13] Despite that alkaloids are among the most important drugs for human disease treatment, certain plant alkaloids can cause systemic effects like gastrointestinal, kidney toxicity, and genotoxicity. Notably, alkaloids such as pyrrolizidine alkaloids undergo biotransformation, leading to DNA interactions that can cause mutations and cancer.[14] It is well-established that pyrrolizidine alkaloids cause both acute and chronic liver toxicity in humans and animals. Symptoms of acute poisoning include abdominal pain, nausea, vomiting, diarrhea, and edema. The primary toxic effects of pyrrolizidine alkaloids are highly toxic, carcinogenic, and genotoxic.[15]

Bueno-Borges reported that the raw Sacha Inchi seed contained 7.02 ± 0.2 mg per g of saponins;[10] Srichamnong reported that the Sacha Inchi raw seed contained 27 ± 4 mg per kg DW of saponins and 485 ± 35 mg per kg DW of alkaloids;[9] in an application submitted to European Food Safety Authority (EFSA), the raw Sacha Inchi seed was reported to contain 11.2–17.3 mg per kg of alkaloids and 6.3–13.7 mg per kg of saponins.[16] This raises concerns about the safety of Sacha Inchi seeds, especially regarding the alkaloid content in *P. volubilis*. Furthermore, the available quantitative data on total saponins and alkaloids are insufficient to make a comprehensive assessment, as well as no information regarding the type of quantified alkaloids was provided. Therefore, the EFSA has expressed reservations about the safety of *P. volubilis*-derived products and raised objections to their introduction into the European Union market.[16]

In previous studies, the fatty acid contents of the Sacha Inchi seeds and oil, and their benefits were reported comprehensively.[1,3,17] In this study, with an attempt to verify the presence of saponins and alkaloids in Sacha Inchi seeds, we designed an extensive strategy to analyse the chemical composition and toxicity of these seeds. Briefly, the raw Sacha Inchi seeds were extracted and separated into different polarity fractions. These extracts and fractions were analysed using proton-NMR, GC-MS, and LC-QTOF experiments. Subsequently, GNPS molecular networking was performed to analyse the MS/MS results and predict the metabolites present in this sample. For the toxicity analysis, the extract and fractions were evaluated for their *in vitro* cytotoxicity using the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. Additionally, the predicted metabolites were screened for their toxicity using *in silico* network toxicology analysis. Additionally, the total saponin and alkaloid contents and trigonelline concentration were also quantified. Based on the combination of the chemical profile and toxicity results, we can make judgments regarding the toxic effects of Sacha Inchi seeds, to prove that these seeds are non-toxic and edible.

Results

^{1}H-NMR-based chemical profile

By interpreting the proton-NMR results, the spectrum of the MeOH extract (Fig. S1†) revealed the predominant presence of sugar components, indicated by multiple signals of oxygenated protons in the range of δ_H 4.00–3.23 ppm and the signals of a doublet oxygenated proton at δ_H 5.43 (J = 6.0 Hz) ppm. Additionally, triglycerides or/and fatty acids were identified by signals of the terminal methyl protons at δ_H 0.98–0.86 ppm, the acyl chains at δ_H 1.60–1.20 ppm, –**CH$_2$**–COOH protons at δ_H 2.50–2.30 ppm, –**CH$_2$**–CH=CH protons around δ_H 2.0 ppm, as well as

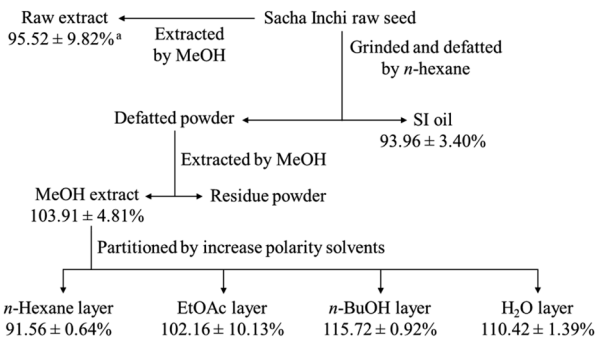

Fig. 1 The separation process and cytotoxicity effects of SI oil, MeOH, *n*-hexane, EtOAc, *n*-BuOH, and H_2O extracts against the NIH-3T3 cell line. [a]Cell viability. The cells are treated with the test samples at a concentration of 400 μg mL^{-1} for 24 h. Data are presented as mean ± standard deviation (SD) (n = 3).

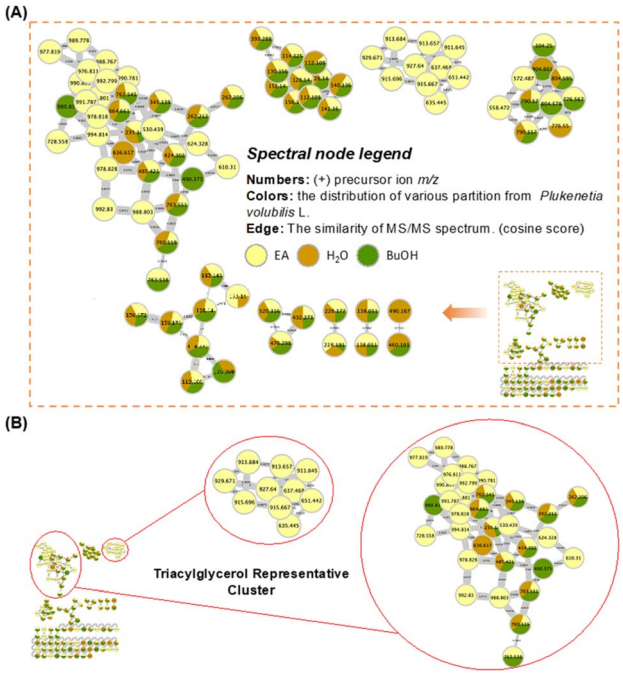

Fig. 2 The metabolomic diversity of the EtOAc, *n*-BuOH, and H_2O extracts illustrated by LC-QTOF coupled with molecular networking. (A) The classical molecular network spectral nodes coloured according to the extracts; (B) the representative distributions of constituents in the EtOAc extract.

the olefinic protons of unsaturated fatty acids at δ_H 5.37 ppm.[18,19] The ^{1}H-NMR spectroscopy was confirmed is a common analytical method used presently in this field, which can provide valuable information on the composition of the oil sample.[19] Notably, the spectrum also exhibited signals characteristic of trigonelline, a pyridine alkaloid, at δ_H 9.07 (1H, s), 8.78 (2H, t, $J = 6.5$), and 8.03 (1H, t, $J = 6.5$).[20] Analysis of the proton spectrum of the EtOAc extract indicated a predominance of triglycerides or/and fatty acids (Fig. S2†). Further, the proton spectra of the *n*-BuOH and H_2O extracts (Fig. S3 and S4†) revealed that the *n*-BuOH extract primarily contained sugars, triglycerides or/and fatty acids, and trigonelline. In contrast, the H_2O extract was found to exclusively contain sugars and trigonelline.

The *n*-BuOH and H_2O extracts were subjected to further fractionation using silica gel MPLC and polyamide open column, respectively. The proton-NMR spectra of fractions PV-Bu-1 to PV-Bu-9, separated from the *n*-BuOH extract, revealed that PV-Bu-1, PV-Bu-2, and PV-Bu-3 predominantly consisted of triglycerides or/and fatty acids, and plasticizer residues (Fig. S5–S7†). Whereas the fractions PV-Bu-6, PV-Bu-7, PV-Bu-8, and PV-Bu-9 were found to be rich in triglycerides or/and fatty acids (Fig. S10–S13†). Notably, fractions PV-Bu-4 and PV-Bu-5 exhibited extraordinary signals beyond typical triglycerides and fatty acids, such as the triplet protons at δ_H 3.45, singlet protons at δ_H 3.57, 3.80, 4.02, and olefinic protons at δ_H 6.11, 7.04, and 8.10 (Fig. S8 and S9†). Additionally, fractions PV-H_2O-1 to PV-H_2O-4, separated from the H_2O extract, showed predominant signals indicative of sugars and trigonelline in their proton-NMR spectra (Fig. S14–S17†). Whereas fractions PV-H_2O-5 and PV-H_2O-6 displayed a more complex array of unrecognized signals in their proton-NMR spectra (Fig. S18 and S19†). The proton-NMR interpretation of these extracts and fractions indicated that the major chemical compositions of Sacha Inchi seeds were oil, including *n*-hexane and EtOAc extracts, and sugar contents, the majority of *n*-BuOH and H_2O layers.

GC-MS-based chemical profile

The GC-MS results of SI oil, *n*-hexane, and EtOAc extracts FAME products are shown in Table 1. These results indicated that the fatty acid contents are similar in all three extracts, predominantly comprising linoleic acid (34.83–38.26%) and linolenic acid (33.28–40.02%). Additionally, oleic acid (10.06–17.41%), palmitic acid (5.86–7.72%), and stearic acid (3.42–4.39%) were also identified. These findings are consistent with previous studies on the chemical composition of Sacha Inchi seeds, which reported that the predominant fatty acids were linoleic and linolenic acids (polyunsaturated fatty acids), along with moderate concentrations of oleic acid (monounsaturated fatty acid), and palmitic and stearic acids (saturated fatty acids).[3,21] The results presented in Table 2 further confirm the presence of a minor amount of phytosterols in all the SI oil, and *n*-hexane and EtOAc extracts of Sacha Inchi seeds, including β-sitosterol, campesterol, stigmasterol, and clionasterol. To analyse the phytosterol content in those extracts, the saponification reaction was conducted to saponify the triglycerides, allowing for the removal of triglycerides from the samples. The remaining non-saponified compositions were then acetylated and analysed by the GC-MS. The presence of phytosterols in this sample aligns with data from previous publications, indicating the

Table 1 FAMEs in SI oil, *n*-hexane, and EtOAc extracts identified by GC-MS analysis

			Peak area ratio (%)		
RT (min)	FAME[a]	CN : DB[b]	SI oil	*n*-Hexane extract	EtOAc extract
18.10	Palmitic acid	16 : 0	5.86	6.26	7.72
23.02	Linoleic acid	18 : 2	38.26	34.83	37.01
23.19	Linolenic acid	18 : 3	33.28	37.28	40.02
23.22	Oleic acid	18 : 1	17.41	17.07	10.06
23.94	Stearic acid	18 : 0	4.39	3.62	3.42
Total			99.20	99.06	98.23

[a] The fatty acid name of FAMEs. [b] Carbon number/double bond.

Table 2 Phytosterols in non-saponified SI oil, *n*-hexane, and EtOAc extracts identified by GC-MS analysis

		Peak area ratio (%)		
RT (min)	Compound name[a]	SI oil	*n*-Hexane extract	EtOAc extract
10.04	β-Sitosterol	4.44	3.46	7.77
16.13	Campesterol	3.91	3.21	4.52
16.96	Stigmasterol	25.15	17.42	23.56
19.16	Clionasterol	33.07	37.59	33.96
Total		66.57	61.68	69.81

[a] The original sterol of the acetylated-sterol products.

containing of β-sitosterol, campesterol, and stigmasterol in Sacha Inchi seeds.[21,22]

GNPS-based chemical profile

The obtained high-resolution MS/MS data was subjected to the GNPS molecular networking analysis, which resulted in an overview of the metabolomic spectral network. The following molecular network diagrams were established based on different extraction and separation methods. The MeOH extract was partitioned with EtOAc, *n*-BuOH, and H_2O in equal volumes to obtain three different polarity layers. The established molecular network is shown in Fig. 2A, with a pale-yellow colour represented for the EtOAc layer, while the *n*-BuOH and H_2O layers are represented by green and brown, respectively. The result indicated that the predominant metabolites distributed in the EtOAc layer (Fig. 2B) containing precursor ions with high values (>500 *m*/*z*), in combination with the proton-NMR and GC-MS results, it is inferred that its main chemical compositions were triglycerides, and the contained metabolites in the *n*-BuOH and H_2O layers have a greater chemical diversity.

The cluster formed by the fractions separated from the *n*-BuOH layer showed its chemical structure diversity in Fig. 3 and S26.† The classification analysis using Network Annotation Propagation (NAP) and NPClassifier indicated that these fractions are predominantly composed of fatty acids, carbohydrates, and amino acids, oligopeptides, and analogues (Fig. S26†). Three compounds were annotated in the built-in database of GNPS, including trigonelline (**1**), phenylalanine (**2**), and L-alanyl-5-oxo-L-proline (**3**). The mass spectrometry information about these nodes is listed in Table 3 and Fig. S28.† The presence of trigonelline can be further revealed by the characteristic signals of the proton-NMR analysis (^{1}H-NMR-based chemical profile section). The GNPS network diagram demonstrated that fraction 1 of the *n*-BuOH layer (BFr. 1), represented by a light-green colour, predominantly contains large molecular metabolites (>500 *m*/*z*). The proton-NMR spectrum of this fraction further supports this observation, indicating a significant presence of triglycerides (^{1}H-NMR-based chemical profile section). This finding indicated that three major clusters in the *n*-BuOH layer, which contained nodes from fraction 1 (light-green colour), were also predominantly composed of triglycerides and their derivatives. Whereas the pyridine alkaloid and nitrogen-containing compounds are primarily found in fractions 7, 8, and 9, distinguished by their brown colour (BFr. 7–9). The chemical

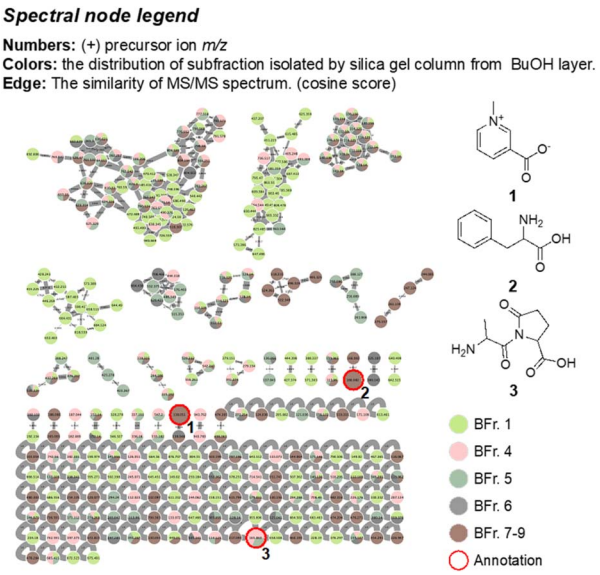

Fig. 3 The metabolomic diversity of fractions PV-Bu-1–9 illustrated by LC-QTOF coupled with molecular networking and three annotations from GNPS libraries.

composition predicted through molecular networking analysis of the *n*-BuOH layer aligns with the results obtained from the proton-NMR interpretation, in overall, this layer still contained the oil content, in addition to sugars and a minor amount of pyridine alkaloid and nitrogen-containing metabolites, no signals of the saponin content were predicted.

The metabolomic diversity of fractions PV-H_2O-1 to PV-H_2O-6 from the H_2O layer and their predicted contained metabolites are shown in Fig. 4. The classification analysis using NAP and NPClassifier of these fractions is shown in Fig. S27.† The diagram indicated that the distributions of fractions PV-H_2O-1 to PV-H_2O-4 are relatively similar. According to the comparison results of NAP, fractions PV-H_2O-1 to PV-H_2O-4 are mainly classified as carbohydrate compounds. This result is consistent with the proton-NMR results of these fractions, which showed the signals of sugars throughout four fractions. Fractions PV-H_2O-5 and PV-H_2O-6 (marked in light orange and dark blue) contained much more diverse chemical compositions because they almost contributed in all clusters.

Five nitrogen-containing compounds were annotated in the built-in database of GNPS, including trigonelline (**1**), L-γ-

Table 3 The information for annotations from GNPS libraries and MassHunter

No.	Formula	RT (min)	Precursor ion (*m*/*z*)	Name
1	$C_7H_7NO_2$	0.49	138.051	Trigonelline
2	$C_9H_{11}NO_2$	0.85	166.083	DL-Phenylalanine
3	$C_8H_{12}N_2O_4$	2.30	165.063	L-Alanyl-5-oxo-L-proline
4	$C_{11}H_{20}N_2O_5$	11.84	261.145	L-γ-Glutamyl-L-leucine
5	$C_{10}H_{17}NO_3$	12.78	241.155	*N*-Methacryloyl-L-leucine
6	$C_5H_{11}NO_2$	13.02	118.086	3-Aminopentanoic acid
7	$C_{18}H_{27}N_5O_{10}$	19.39	515.214	Asp–Glu–Pro–Asn

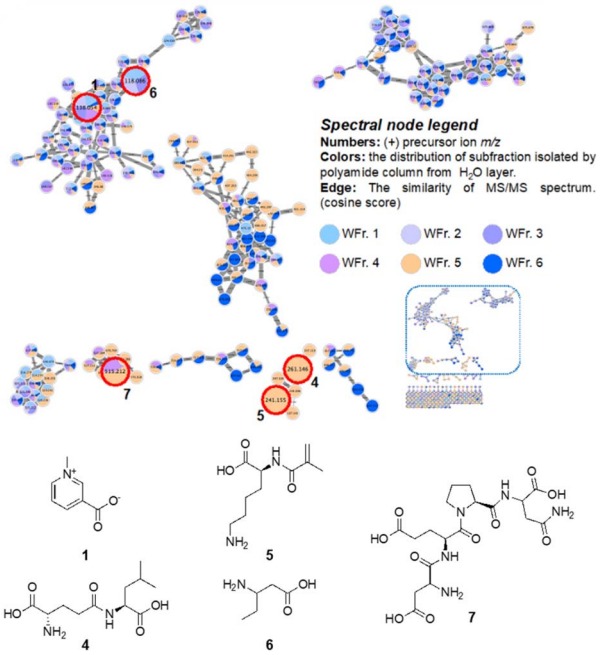

Fig. 4 The metabolomic diversity of fractions PV-H_2O-1–6 illustrated by LC-QTOF coupled with molecular networking and five annotations from global natural products social (GNPS) libraries.

glutamyl-L-leucine (**4**), *N*-methacryloyl-L-leucine (**5**), 3-aminopentanoic acid (**6**), and Asp–Glu–Pro–Asn (**7**), the mass spectrometry information about these nodes was listed in Table 3 and Fig. S28 and S29.† Overall, the combination results of GNPS molecular networking analysis and proton-NMR analysis indicated that the H_2O layer primarily contained sugars and nitrogen-containing compounds, no more oil contents, and also no signals of the saponins were observed.

Quantification of total saponin, total alkaloid, and trigonelline

The total saponin and total alkaloid concentrations of this sample were quantified as 3.59 ± 0.07 mg g^{-1} and 20.46 ± 0.20 μg g^{-1}, respectively ($n = 3$). These results were quite in agreement with those results in previous publications, which reported that the total saponin concentrations were 7.02 ± 0.2 mg g^{-1}.[10] or 6.3–13.7 mg kg^{-1};[16] and the total alkaloid concentration was 11.2–17.3 mg kg^{-1}.[16] However, it showed high variability with the results of Srichamnong *et al.*, which indicated concentrations of 27 ± 4 mg kg^{-1} for total saponin and 485 ± 35 mg kg^{-1} for total alkaloid.[9]

Trigonelline is a pyridine alkaloid found in various plants, particularly in the seeds of fenugreek (*Trigonella foenum-graecum*) and coffee beans, and known for its bitter taste.[23,24] Trigonelline is the only alkaloid metabolite predicted by GNPS analysis in this study, the presence of this compound was also confirmed by its proton-NMR signals in MeOH extract. This compound exerts antibacterial, antidiabetic, and antitumor effects without causing cytotoxicity to normal lung cell lines.[23,24] To determine the concentration of trigonelline in Sacha Inchi seed samples collected from various regions of Taiwan, an HPLC-UV experiment equipped with a Hypercarb column was utilized and validated for linearity, precision, and accuracy. The method exhibited a wide quantification range of 3.9 to 250 μg mL^{-1}, with a calibration curve demonstrating high correlation coefficients ($r^2 = 1.000$), expressed as $y = 20849x + 12735$. The LOD and LOQ were determined to be 1.0 and 2.0 μg mL^{-1}, respectively. The intraday and interday precision, expressed as the relative standard deviation, were found to be 0.407% and 0.873%, respectively. Recovery rates for three different concentrations of the spiked standard solution ranged from 99.4% to 110.3%. These validation results confirm that the developed HPLC method is suitable for quantifying trigonelline content in Sacha Inchi seeds. Subsequently, the established method was applied to quantify trigonelline in nine Sacha Inchi seed samples. Analysis of the results (Table 4) revealed that trigonelline concentrations in the raw samples ranged from 526.57 to 848.75 μg g^{-1}, with no discernible systematic differences observed among the various regions of Taiwan.

However, the concentrations of trigonelline are not in agreement with the result of the total alkaloid content reported in this sample, which revealed that the total alkaloid content was 20.46 ± 0.20 μg g^{-1} for the raw seeds. This also contradicts the findings of previous studies, which indicated a total alkaloid content ranging from 11.20 to 17.30 μg g^{-1}.[16] These disparities in total alkaloid concentration between the reports, despite using the same quantification method, raise questions about the reliability and sensitivity of the colorimetric method employed to quantify the total alkaloid content.

In vitro cytotoxicity

The Sacha Inchi raw extract and the SI oil, MeOH, *n*-hexane, EtOAc, *n*-BuOH, and H_2O layers were tested for their cytotoxicity against the mouse fibroblast NIH-3T3 cell line using the MTT assay at a concentration of 400 μg mL^{-1} for 24 h. The results shown in Fig. 1 indicated that all the extract and layers showed no cytotoxic effects against the NIH-3T3 cells, even at a concentration of 400 μg mL^{-1}. A previously published study also revealed that raw and roasted Sacha Inchi seeds are safe for some normal cell lines, including embryonic mouse fibroblast (3T3-L1), human embryonic kidney cells (HEK293), human hepatic stellate cells (LX-2), and human peripheral blood mononuclear cells

Table 4 Trigonelline contents in Sacha Inchi seeds collected in different areas[a]

Sample	Place of collection		Trigonelline (μg g^{-1})
PV-1	Eastern Taiwan	Hualien County	641.97 ± 4.00
PV-2		Hualien County	675.29 ± 1.36
PV-3		Hualien County	592.38 ± 0.31
PV-4*	Middle Taiwan	Taichung City	548.52 ± 1.58
PV-5		Changhua County	650.26 ± 0.76
PV-6		Nantou County	799.56 ± 3.57
PV-7	Southern Taiwan	Chiayi County	526.57 ± 0.73
PV-8		Chiayi County	754.14 ± 0.13
PV-9		Tainan City	848.75 ± 0.26

[a] Data was presented as mean ± SD ($n = 2$). *PV-4 is the sample used for all other experiments in this study.

Table 5 Prediction results and confidence level of five toxicity categories

Compound	Prediction	Toxicity categories and confidence value[a]				
		1	2	3	4	5
Trigonelline (**1**)	Inactive	0.60	0.66	0.98	0.94	0.75
DL-Phenylalanine (**2**)	Inactive	0.81	0.82	0.99	0.87	0.71
L-Alanyl-5-oxo-L-proline (**3**)	Inactive	0.80	0.63	0.99	0.83	0.75
L-γ-Glutamyl-L-leucine (**4**)	Inactive	0.76	0.67	0.99	0.90	0.63
N-Methacryloyl-L-leucine (**5**)	Inactive	0.82	0.60	0.99	0.60	0.65
3-Aminopentanoic acid (**6**)	Inactive	0.79	0.68	0.99	0.80	0.66
Asp–Glu–Pro–Asn (**7**)	Inactive	0.89	0.64	0.99	0.80	0.83

[a] 1: Organ toxicity (hepatotoxicity); 2: carcinogenicity; 3: immunotoxicity; 4: mutagenicity; 5: cytotoxicity.

(PBMC).[9] No significant differences were observed in any tested cells at 500 μg mL^{-1} for 48 h, although mild cytotoxicity was specifically observed in hepatic stellate cells (LX-2) with the IC_{20} and IC_{50} values of 250 and 500 μg mL^{-1}, respectively.[9]

In silico toxicology

The *in silico* analysis of the toxicity for seven predicted compounds, assessed using the Protox-II computational platform, revealed a consistently inactive result across five crucial categories of toxicity, substantiated by the high confidence values throughout the analysis (Table 5). The toxicity class indicated whether a compound was active (toxic) or inactive (non-toxic). The confidence score ranged from 0.50 to 1.00, with higher values indicating more reliable predictions for toxicity classification.[25] In particular, the confidence levels for organ toxicity measured by hepatotoxicity risk ranged from 0.60 for trigonelline to 0.89 for Asp–Glu–Pro–Asn, thus showing a low possibility of any hepatotoxic effects. The confidence values in the crucial region of carcinogenicity ranged from 0.60 for *N*-methacryloyl-L-leucine to 0.82 for DL-phenylalanine, which also suggested a negligible carcinogenic risk, if any. Comparably, every compound showed remarkably high confidence levels in the assessment of immunotoxicity, 0.98 for trigonelline and 0.99 for the remaining compounds, which strongly supports the negative result of immunotoxic effects. The mutagenicity predictions also showed very high confidence values, 0.60 for *N*-methacryloyl-L-leucine to 0.94 for trigonelline, which clearly showed a negligible chance of mutagenic effects. Finally, the confidence values ranged reasonably in the category of general cytotoxicity, ranging from 0.63 for L-γ-glutamyl-L-leucine to 0.83 for Asp–Glu–Pro–Asn, once again suggesting a negligible cytotoxic potential. Overall, these results suggested that the seven nitrogen-containing compounds possess a non-toxicological profile. The high confidence values of these metabolites indicate the negligible predicted risks across five crucial toxicity categories.

Discussion

The raw Sacha Inchi seeds were defatted by *n*-hexane to afford SI oil, which occupied approximately 40% of the raw material. The defatted residue was further extracted by MeOH, and then partitioned into different polarity layers, including *n*-hexane, EtOAc, *n*-BuOH, and H_2O. The proton-NMR spectra of these layers indicated that the *n*-hexane and EtOAc layers still contained large amounts of oil, whereas the *n*-BuOH and H_2O layers were predominant in sugar contents. The trigonelline alkaloid was found to present with remarkable quantity in the MeOH extract and *n*-BuOH and H_2O layers, evidenced by its ^{1}H-NMR signals. The GC-MS analysis of the SI oil, *n*-hexane, and EtOAc layers confirmed that the predominant compositions of oil content were linoleic acid (34.83–38.26%), linolenic acid (33.28–40.02%), oleic acid (10.06–17.41%), palmitic acid (5.86–7.72%), and stearic acid (3.42–4.39%), and the presence of four phytosterols, β-sitosterol, campesterol, stigmasterol, and clionasterol. The fatty acid contents analysed by GC-MS were in agreement with those data in previous publications, which predominated in linolenic, linoleic, and oleic acids.[1,3,17] This result further confirmed the beneficial effects of the Sachi Inchi seeds on human health by containing large amounts of the ω-3 (linolenic) and ω-6 (linoleic) fatty acids, which can prevent several diseases such as cancer, coronary heart disease, and hypertension.[1] The LC-QTOF-based GNPS molecular networking analysis suggested that the EtOAc extract predominantly contained oil. Detailed analysis of *n*-BuOH extract indicated that this layer still contained oil, in addition to sugars and minor amounts of nitrogen-containing metabolites, while the H_2O layer contained sugars and nitrogen-containing metabolites. The GNPS molecular networking also predicted the presence of some nitrogen-containing metabolites, including trigonelline (**1**), phenylalanine (**2**), L-alanyl-5-oxo-L-proline (**3**), L-γ-glutamyl-L-leucine (**4**), *N*-methacryloyl-L-leucine (**5**), 3-aminopentanoic acid (**6**), and Asp–Glu–Pro–Asn (**7**); however, not show any signals for the saponin compounds. The trigonelline concentration in the raw Sacha Inchi seeds then was quantified by analytical HPLC, which showed remarkable amounts of 548.52 to 848.75 μg g^{-1}.

Previous studies have indicated that the Sacha Inchi seeds contain a certain amount of saponins.[9,10,16] However, the combination analysis of GC-MS and LC-QTOF-based GNPS molecular networking performed in this study, indicated no saponin content presented in this Sacha Inchi seed sample. This discrepancy can be explained by inspecting the total saponin quantification method employed. The total saponin quantification methods used in the reported studies involve the formation of a complex between saponins and vanillin, sulfuric acid reagents, resulting in a measurable colour change.[26]

However, these reagents lack specificity and can react with other substances possessing the same skeleton, such as sterols, triterpenes, and bile acids.[26] The GC-MS analysis of the Sacha Inchi seed chemical composition in this study confirmed the presence of four steroidal compounds, β-sitosterol, campesterol, stigmasterol, and clionasterol. Previous studies also indicated that Sacha Inchi seeds contain sterols, which include β-sitosterol, campesterol, and stigmasterol.[21,22] Therefore, based on the results of GC-MS experiments and LC-QTOF-based GNPS molecular networking, suggested that the saponin content reported in Sacha Inchi seeds may not be saponins but

rather sterols, as determined by the unspecific colorimetric methods. The discrepancy was also observed in the alkaloid quantification between total alkaloid concentration and trigonelline concentration. The total alkaloid content of raw seed was quantified as 20.46 ± 0.20 μg g^{-1} using the BCG colorimetric method, whereas the trigonelline content for the corresponding sample was 548.52 ± 1.58 μg g^{-1} using the analytical-HPLC method, much larger than total alkaloid concentration. These results have raised concerns regarding the specificity of the total alkaloid quantification using colorimetric method, which is not specific to trigonelline, a pyridine alkaloid. Therefore, we suggested that the total saponin and alkaloid quantifications in Sacha Inchi seeds using colorimetric methods should be reconsidered regarding their specificity.

The major concern about the toxic alkaloid contents present in plant-derived products is pyrrolizidine alkaloids.[27] Pyrrolizidine alkaloids are a large group of natural toxic secondary metabolites present in many plant species. Certain pyrrolizidine alkaloids have been identified as posing significant toxicity risks to both humans and animals owing to their occurrence within the food chain.[27,28] In 2020, the Herbal Medicinal Products Committee of the European Medicines Agency released a draft of the updated public statement concerning pyrrolizidine alkaloids. This statement suggested maximum daily intake limits for pyrrolizidine alkaloids. For the presence of pyrrolizidine alkaloids in medicinal products, a proposed cap of 1.0 μg per day for adults using them orally would be enforced. This limit was similarly referenced in the preliminary version of the new European Pharmacopoeia chapter 2.8.26.[27,29] Although the total alkaloids content in Sacha Inchi seeds has been reported, the type of alkaloids contained was not specified. In this study, using LC-QTOF-based molecular network analysis and analytical-HPLC quantification, trigonelline, a pyridine alkaloid, was confirmed to be the only alkaloid present in this Sacha Inchi seed sample with significant quantities ranging from 548.52 to 848.75 μg g^{-1}. Previous publications have confirmed trigonelline to be safe for human health, in addition to its benefits.[30,31] Trigonelline showed no cytotoxicity at concentrations up to 100 μM in human neuroblastoma SK-N-SH cells,[32] human hepatocellular carcinoma (Hep3B) cells,[33] human immortalized dermal keratinocytes (HaCaT), and human foreskin fibroblasts (Hs68)[34] after being treated for six days, 48 h, and 24 h, respectively. The trigonelline is predominantly present in coffee beans and is known for its bitter taste.[23,24] This study represents the first documentation of trigonelline's presence in Sacha Inchi seeds with specific content's range,[2,3] suggesting its potential as a chemical marker for quality control purposes.

The *in vitro* cytotoxicity evaluation of the SI oil, MeOH, *n*-hexane, EtOAc, *n*-BuOH, and H_2O extracts showed no cytotoxic effects against the NIH-3T3 cell line, even at a concentration of 400 μg mL^{-1}. These results in agreement with a previously published study indicated that raw and roasted Sacha Inchi seeds showed no cytotoxicity against some normal cell lines, including 3T3-L1, HEK293, LX-2, and PBMC cells.[9] The *in silico* toxicity analysis of the GNPS-predicted nitrogen-containing compounds presented in this sample against five toxicity categories (hepatotoxicity, carcinogenicity, immunotoxicity, mutagenicity, and cytotoxicity), showed that these compounds possess negligible toxicity across these five crucial toxicity categories. In previous publications, by using *in vivo* animal models, Rodeiro reported that there is no toxicity was observed from the rats and mice administrated with a single dose of 2000 mg per kg body weight of Sacha Inchi seed defatted powder, or 500 mg kg^{-1} for 90 days.[35] Gorriti indicated no signs of toxicity in rats and mice administrated with 0.5 mL per kg body weight of SI oil for 60 days.[36] In summary, in this study, the chemical composition of Sacha Inchi seeds was investigated, and the results revealed no saponin in this sample, instead, it confirmed the presence of trigonelline, a pyridine alkaloid, and some other nitrogen-containing compounds; however, these compounds also showed no toxicity effects evidenced by *in silico* experiment. Therefore, through this study, we suggested that the total saponin and alkaloid concentrations contained in the Sacha Inchi seeds, quantified by using colorimetric methods should be reconsidered, as well as the potential health risks associated with saponins and alkaloids, towards a comprehensive approval for the use of this seed as food.

Conclusion

This study indicated that the chemical composition of Sacha Inchi seeds was predominantly oil and sugars, with minor amounts of phytosterols and nitrogen-containing metabolites. The oil content comprised linoleic, linolenic, oleic, palmitic, and stearic acids, while the nitrogen-containing content included pyridine alkaloid, amino acids, and oligopeptides. No saponin content was identified, instead, β-sitosterol, campesterol, stigmasterol, and clionasterol were detected. In this study, the presence of trigonelline, a non-toxic pyridine alkaloid, was confirmed in the alkaloid content of Sacha Inchi seeds for the first time. Given its identification, trigonelline should be regarded as a valuable chemical marker for ensuring the quality control of Sacha Inchi seeds. The nitrogen-containing metabolites possessed a non-toxicological profile, as evidenced by *in silico* analysis using the Protox-II computational platform. Extracts of Sacha Inchi seeds showed no *in vitro* cytotoxicity against the NIH-3T3 cell line, as evaluated by MTT assay. The chemical analysis of saponin and alkaloid contents in this study highlighted the absence of saponin, the low specificity of total saponin and alkaloid quantifications using colorimetric methods and suggested the low safety risk of Sacha Inchi seeds regarding the total saponins and total alkaloid content.

Experimental

Chemicals and reagents

The ACS-grade acetone, ethyl acetate (EtOAc), methanol (MeOH), *n*-butanol (*n*-BuOH), and *n*-hexane were purchased from Echo Chemical (Miaoli, Taiwan). The HPLC-grade acetonitrile (ACN), MeOH, *n*-hexane, and toluene were purchased from Macron Fine Chemicals (Radnor, PA, US). Sulfuric acid (H_2SO_4), hydrochloric acid (HCl), formic acid, acetic anhydride, vanillin, and *d*-solvents for the NMR experiment were obtained

from Sigma-Aldrich (St. Louis, Missouri, US). Atropine, diosgenin, and trigonelline were purchased from the Tokyo Chemical Industry (Tokyo, Japan), Sigma-Aldrich, and Wuhan ChemFaces Biochemical (Wuhan, China), respectively. Column chromatography was carried out using silica (40–63 μm, Merck) or polyamide (PA 6, Sigma-Aldrich).

Sacha Inchi seed materials

The Sacha Inchi seeds (without husk) used in this study were provided by the Crop Genetic Resources and Biotechnology Division, Taiwan Agricultural Research Institute (Taichung, Taiwan). The shells of these seeds were removed manually to provide raw seeds. The voucher specimens of these seed samples were deposited in the Crop Genetic Resources and Biotechnology Division, Taiwan Agricultural Research Institute, and the Graduate Institute of Pharmacognosy, Taipei Medical University.

Total saponin and alkaloid quantification

The total saponin quantification assay was modified from the method described by Hiai *et al.*[26] Briefly, the raw Sacha Inchi powder was defatted by *n*-hexane in the Soxhlet system for 3 h. The solvent was discarded and the sample was dried at 40 °C to yield the defatted powder. 0.5 g defatted powder was extracted by 20 mL MeOH and then filtered. 5 mL of the filtrate was dried and suspended in 5 mL of H_2O, and stepwise partitioned with *n*-hexane and *n*-BuOH. The combined *n*-BuOH layer was prepared to 10 mL, then 1 mL of *n*-BuOH extract was transferred to a glass bottle and dried under vacuum. 1000 μL of 72% H_2SO_4 and 100 μL of 8% vanillin in 95% ethanol were added, mixed well, heated at 60 °C for 10 min, and cooled in ice. 200 μL of the reaction solution was transferred into a 96-well plate and the absorbance was measured at 530 nm. The total saponin concentration in the sample was calculated according to the diosgenin standard curve. For the diosgenin standard curve, a series of diosgenin concentrations from 5 to 200 μg mL^{-1} were prepared and the same procedure was performed to establish the standard curve. The total saponin concentration in the original sample was calculated using the following equation:

$$\text{Total saponin (mg g}^{-1}) = (\text{Sap} \times 4 \times 10 \times R)/(m \times 100)$$

where Sap (mg) is saponin concentration calculated on the calibration curve; R (%) is the percentage of sample residue after defatting; and m (g) is sample weight.

The total alkaloid quantification assay was modified from the method described by Shamsa *et al.*[37] Briefly, the defatted powder (2 g) was extracted two times using 40 mL of MeOH, and the combined extract was dried and suspended in 4 mL of 2 N HCl, then partitioned with *n*-hexane. The HCl layer was filtered through a pre-wet 70 mm diameter filter paper number 2. The sample tube and filter paper were rinsed using 2 N HCl and then with water, and the combined filtrates were neutralized with 3 N NaOH to pH 5–6. Add 5 mL of PBS pH 4.7 and 5 mL of bromocresol green (BCG) 0.03% were added and mixed well in an ice bath. The mixture was partitioned with 10 mL of chloroform in the ice bath. The absorbance of the chloroform layer was measured in a 1 cm quartz cuvette at 417 nm using a UV-Vis spectrometer. The total alkaloid concentration in the sample was calculated according to the atropine standard curve. For the atropine standard curve, a series of atropine concentrations from 10 to 125 μg mL^{-1} were prepared and the same procedure was performed to establish the standard curve. The total alkaloid concentration in the original sample was calculated using the following equation:

$$\text{Total alkaloid (μg g}^{-1}) = (\text{Alk} \times R)/(m \times 100)$$

where Alk (μg) is the alkaloid concentration based on the calibration curve; R (%) is the percentage of sample residue after defatting; and m (g) is the sample weight. For the preparation of 0.03% BCG solution, 3 mL 2 N NaOH and 5 mL water were added to 69.8 mg BCG, and heated to completely dissolve, then diluted to 1000 mL with water. For the preparation of PBS at pH 4.7, 1000 mL water was added to 71.6 g Na_2HPO_4, and 2 N citric acid was used to adjust the pH to 4.7.

Extraction and separation

The raw Sacha Inchi seeds (3600 g) were grounded and defatted using *n*-hexane, then filtered and dried to yield 1900 g of defatted powder. The *n*-hexane filtrate was dried under a vacuum to obtain 1.5 L of SI oil. The defatted powder was subsequently extracted with MeOH, resulting in 137 g of MeOH extract. This extract was suspended in H_2O and stepwise partitioned with *n*-hexane, EtOAc, and *n*-BuOH to afford *n*-hexane extract (40 g), EtOAc extract (8 g), *n*-BuOH extract (12 g), and an H_2O layer (80 g). The *n*-BuOH extract was loaded onto a silica gel MPLC column and eluted stepwise with gradients of *n*-hexane/EtOAc and EtOAc/MeOH, resulting in nine fractions, PV-Bu-1 (330 mg), PV-Bu-2 (82 mg), PV-Bu-3 (42 mg), PV-Bu-4 (458 mg), PV-Bu-5 (807 mg), PV-Bu-6 (4870 mg), PV-Bu-7 (2189 mg), PV-Bu-8 (1480 mg), and PV-Bu-9 (1223 mg). Part of the H_2O layer (11 g) was applied to a polyamide open column and eluted with mixtures of acetone/H_2O, resulting in six fractions, PV-H_2O-1 (3290 mg), PV-H_2O-2 (3310 mg), PV-H_2O-3 (3400 mg), PV-H_2O-4 (750 mg), PV-H_2O-5 (340 mg), and PV-H_2O-6 (50 mg).

Proton-NMR spectra recording

These extracts and fractions were recorded for their proton-NMR spectra, using a Bruker Fourier 300 MHz FT-NMR spectrometer (Bruker, Billerica, MA, USA), at 25 °C in MeOH-d_4 or $CDCl_3$.

Low polarity content analysis using GC-MS spectrometry

The fatty acid compositions were determined in the form of their corresponding fatty acid methyl esters (FAMEs). The preparation of FAMEs followed the modified International Organization for Standardization methods.[38,39] Briefly, 0.1 g of the samples (SI oil, *n*-hexane, and EtOAc extracts) were suspended in 10 mL of anhydrous MeOH containing 0.2 M sulfuric acid. Esterification was carried out by refluxing the mixture for 30 min. After cooling to room temperature, 10 mL of petroleum

ether (PE) was added, followed by 10 mL of deionized water. The mixture was left to settle until the upper PE layer became clear. The upper layer of FAMEs in the PE layer was separated into a vial. The PE was subsequently removed under a vacuum to obtain the FAMEs.

The saponification was carried out following the method described by Chen *et al.*,[40] to analyse components other than triglycerides. Briefly, 1 g of sample (SI oil, *n*-hexane, and EtOAc extracts) was weighed into a 250 mL flask, and then 50 mL of 7.5% KOH in MeOH was added. The mixture was refluxed at 80 °C for 60 min. After cooling to room temperature, the saponified mixture was transferred to a separating funnel and partitioned with *n*-hexane three times. The combined upper layer was then washed three times with 40 mL of H_2O in the separating funnel until it reached a neutral pH (pH = 7). The solvents were removed under a vacuum to afford the non-saponified sample. The acetylation method described by Jia *et al.*,[41] was carried out to acetylate the non-saponified sample. Briefly, 6 mL of a mixture of toluene/acetic anhydride (1 : 10) and 0.1 mL pyridine were added to 20 mg of the sample. The reaction mixture was stirred and heated for 2 h then cooled to room temperature. The combined organic layer was washed with 0.1% HCl three times, dried over anhydrous sodium sulphate, filtered, and concentrated under reduced pressure. The product was dissolved in toluene to a concentration of 1 mg mL^{-1}, and filtered through a 0.22 μm membrane.

The esterification and non-saponified products were then analysed using the GC-MS. The GC-MS was carried out using an Agilent 5977B GC-MSD system (Agilent Technologies, Santa Clara, CA, USA) equipped with a split injector (5 : 1 split ratio) and a Zebron ZB-5ms capillary column (Phenomenex, 30 m × 0.25 mm i. d.; 0.25 μm film thickness). Helium was used as the carrier gas at a flow rate of 1 mL min^{-1}. The oven temperature for FAME samples was 150 °C and held for 2 min, then increased from 150 to 280 °C at a rate of 3 °C min^{-1} and held for 5 min, and increased from 280 to 310 °C at a rate of 10 °C min^{-1} and held for 5 min. The injection volume was 1 μL for each sample. The mass spectrometer operated with an ion source temperature of 230 °C and MS-quadrupole temperature of 150 °C. Solvent delay was set at 3 min. The electron impact energy was set at 70 eV acquiring a total ion chromatogram mode from *m*/*z* 29–1000. The oven temperature for acetylated non-saponified samples was 280 °C and held for 25 min, then increased from 280 to 310 °C with a rate of 3 °C min^{-1} and held for 10 min. The injection volume was 1 μL for each sample. The mass spectrometer operated with an ion source temperature of 300 °C and MS-quadrupole temperature of 200 °C. Solvent delay was set at 3 min. The electron impact energy was set at 70 eV acquiring a total ion chromatogram mode from *m*/*z* 29–1000.

MS/MS non-targeted fragment ions collection using LC-QTOF spectrometry

For MS/MS data collection, the extracts and fractions were re-dissolved in 50% MeOH at a concentration of 5 mg mL^{-1} and analysed using an Agilent 1200 LC system coupled with an Agilent 6530A LC-QTOF mass spectrometer (Agilent Technologies, Santa Clara, CA, USA). Prior to injection, the samples were filtered through a 0.22 μm membrane, with 10 μL of injection volume.

The MeOH, EtOAc, *n*-BuOH, and H_2O extracts and fractions PV-Bu-1,2,3, PV-Bu-4, PV-Bu-5, PV-Bu-6, and PV Bu-7,8,9 were separated using a Luna Omega 1.6 μm Polar C18 column (2 × 50 mm, 1.6 μm, Phenomenex, Torrance, CA, USA) with a flow rate of 0.4 mL min^{-1} and a column temperature of 40 °C. The mobile phase consisted of A (H_2O containing 0.1% formic acid) and B (ACN containing 0.1% formic acid). The gradient elution conditions were as follows: 0–1 min, 3% B; 1–16 min, 3–100% B; 16– 26 min, 100% B; 26–26.2 min, returning to 3% B; 26.2–28 min, 3% B. The fractions PV-H_2O-1, PV-H_2O-2, PV-H_2O-3, PV-H_2O-4, PV-H_2O-5, and PV-H_2O-6 were separated using a Poroshell 120 HILIC column (3 × 150 mm, 2.7 μm, Agilent, Santa Clara, CA, USA) with a flow rate of 0.6 mL min^{-1} and a column temperature of 40 °C. The mobile phase consisted of A (H_2O containing 0.1% formic acid) and B (ACN containing 0.1% formic acid). The gradient elution conditions were as follows: 0–5 min, 95% B; 5–30 min, 95–45% B; 30–35 min, 45% B; 35–36.00 min, returning to 95% B; 36–40 min, 95% B.

The QTOF mass spectrometer was equipped with an electrospray ionization (ESI) source. Source parameters were set as follows: capillary voltage at 2800 V, cone voltage at 20 V, with a source temperature of 120 °C and desolvation temperature at 350 °C, and desolvation gas flow at 500 L h^{-1}. Data were collected at a scan rate of 5 spectra per s. To enhance the coverage of compounds targeted for MS/MS fragmentation, an iterative data-dependent acquisition (DDA) approach was employed. The top five highest-intensity precursor ions per MS^1 scan were selected and fragmented, with collision energy set at 35 V.

GNPS-based molecular networking analysis

The LC-QTOF data was converted to mzML format using MSConvert software,[42] followed by performing a sequence of data processing steps in MZmine software.[43] The GNPS web-based platform (**https://gnps.ucsd.edu**) was used to generate visualized networks.[44] First, the preprocessed LC-QTOF data were uploaded to the system for molecular network building.[45] The major parameter settings were set as follows: precursor ion mass tolerance of 0.02 Da, fragment ion mass tolerance of 0.02 Da, minimum pair cosine score of 0.7, and minimum matched fragment ions of 6. After running the molecular networking algorithm on the GNPS server, the generated molecular networks were downloaded from the job status pages as Cytoscape data and then imported to Cytoscape software for molecular network visualization and editing.[46] The component identification mark was carried out with network annotation propagation (NAP); NAP is built on the molecular fragments calculated by MetFrag and searches for relevant molecules in databases, including GNPS, Super Natural II, CHEBI, and DNP. Its parameters were set as follows, *n*-first: 10, error quality <5 ppm.[46] Structure prediction was carried out using MS2LDA, an unsupervised method that extracts ion fragments and neutral loss molecular substructures from spectra, improving the interpretation and classification of molecular fragments without the reference spectra. Its parameters were set as

follows, *m*/*z* tolerance: 5.0 ppm, tR tolerance: 10.0 s, minimum MS^1 intensity: 0 au, minimum MS^2 intensity: 20.0 au, no duplicate filtering, number of iterations: 1000, and number of Mass2Motifs: 100.[47] Chemical annotation was carried out using MolNetEnhancer, a software package composed of Python and R languages, which combines multiple independent calculation functions and annotation functions, including molecular network, MS2LDA, NAP and automatic classification through ClassyFire to provide a more comprehensive chemical overview of secondary metabolites.[48]

Quantification of trigonelline

The GNPS molecular networking and proton-NMR results suggested the presence of trigonelline, a pyridine alkaloid, in this sample. To quantify the concentration of this alkaloid, the raw Sacha Inchi powder was defatted by *n*-hexane in the Soxhlet system for 3 h, and then 250 mg of defatted powder was extracted using 10 mL of MeOH 50%. Subsequently, 10 μL of the extract was injected into an analytical HPLC system (LC-2050C 3D HPLC, Shimadzu, Tokyo, Japan) with a Hypercarb column (Thermo Scientific, 100 × 4.6 mm, 5 μm). Detection was performed using UV at 265 nm, with a flow rate of 1 mL min^{-1}. The mobile phase conditions were set as follows: 0–1 min, 2% B; 1–10 min, 2–15% B; 10–12 min, 15% B; 12–13 min, 15–100% B; 13–16.5 min, 100% B; 16.5–16.6 min, returning to 2% B; 16.6–20 min, 2% B. Solvent A was H_2O, and solvent B was ACN. Trigonelline was eluted at 7.15 min. The HPLC method was developed and validated regarding ICH Q2 (R1) 2005 guideline from the *International Conference on Harmonization*. The validation included linearity, the limit of detection (LOD), the limit of quantification (LOQ), precision, and accuracy. The concentration of trigonelline in the sample was calculated based on the established calibration curve.

In vitro cytotoxicity assay

The cytotoxic effect of Sacha Inchi seed extracts against the mouse embryonic fibroblast NIH-3T3 cell line (American Type Cell Culture) was determined using the MTT assay, following the method described in a previous publication.[49] Briefly, the test cells (2×10^5 cells per mL) were seeded in 96-well plates and incubated for 24 h in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS), 1% penicillin–streptomycin, and 1% L-glutamine. The samples were dissolved in 100% DMSO, then diluted to 10% DMSO by water to yield the stock solution at 10 mg mL^{-1}. The cells were then treated with test samples at a concentration of 400 μg mL^{-1} for 24 h. After the treatment, 200 μL of MTT (0.5 mg mL^{-1}) in DMEM medium was added to each well and incubated at 37 °C for 4 h in a 5% CO_2 incubator. The supernatants were then removed, and the formazan crystals were dissolved in 200 μL of DMSO. The experiment was repeated in triplicate, and absorbance was recorded at a wavelength of 570 nm.

In silico toxicology analysis

The ProTox-II offers a freely accessible web server for predicting *in silico* toxicity. Users can visit **http://tox.charite.de/protox_II** to utilize the webserver.[25] In this study, the structures of seven targeted compounds were drawn using ChemDraw and then converted into Simplified Molecular-Input Line-Entry System (SMILES) strings. The generated SMILES strings were systematically inputted into Protox-II. Organ toxicity, carcinogenicity, immunotoxicity, mutagenicity, and cytotoxicity are five major toxicity endpoints that Protox-II thoroughly analysed in this work. These endpoints are important indicators for the overall safety of a substance; from these data, it uses machine learning algorithms to create QSAR models, and the toxicity profiles of seven compounds are then predicted by applying these models. Compared to analysing individual endpoints separately, this *in silico* technique provides a more comprehensive picture of hazards by assessing many adverse health consequences concurrently through these QSAR models. The results include toxicity status (active or inactive), confidence scores, and an overall toxicity radar chart.

Statistical analysis

Experimental data are presented as mean ± standard deviation (SD) ($n = 3$). Statistical analysis performed using multiple *t*-test comparisons in Prism software.

Data availability

The data supporting this article have been included as part of the ESI.†

Author contributions

V.-T. T. L.: methodology, investigation, and writing – original draft; T. H. H.: methodology, investigation, and writing – original draft; L.-Y. C.: investigation; M. R. S. P.: investigation; H.-Y. L.: investigation; K.-H. L. conceptualization and methodology; Y.-L. L.: conceptualization and project administration; L.-G. C. conceptualization and methodology; C.-C. W.: conceptualization, project administration, and writing – review and editing.

Conflicts of interest

The authors declare there is no any conflict of interest related to this study.

Acknowledgements

This work was supported by the Taiwan Agricutural Research Institute, Ministry of Agriculture, Taiwan (Grant numbers 110-5802-040-300 and 111-5802-054-300). The authors also would like to thank Prof. Shoei-Sheng Lee (School of Pharmacy, National Taiwan University) for the academic consultation, the Core Facility Center (Taipei Medical University) and the Department of Applied Chemistry (Chaoyang University of Technology) for providing the NMR spectrometer, GC-MS and LC-QTOF facilities.

Notes and references

1 C. Fanali, L. Dugo, F. Cacciola, M. Beccaria, S. Grasso, M. Dacha, P. Dugo and L. Mondello, *J. Agric. Food Chem.*, 2011, **59**, 13043–13049.
2 N. Kodahl and M. Sørensen, *Agronomy*, 2021, **11**, 1066.
3 S. Wang, F. Zhu and Y. Kakuda, *Food Chem.*, 2018, **265**, 316–328.
4 L. A. Follegatti-Romero, C. R. Piantino, R. Grimaldi and F. A. Cabral, *J. Supercrit. Fluids*, 2009, **49**, 323–329.
5 S. K. Sathe, H. H. Kshirsagar and G. M. Sharma, *Plant Foods Hum. Nutr.*, 2012, **67**, 247–255.
6 L. F. Gutiérrez, L. M. Rosada and Á. Jiménez, *Grasas Aceites*, 2011, **62**, 76–83.
7 M. D. Guillén, A. Ruiz, N. Cabo, R. Chirinos and G. Pascual, *J. Am. Oil Chem. Soc.*, 2003, **80**, 755–762.
8 S. M. Colegate, D. R. Gardner and S. T. Lee, *Handbook of Food Chemistry*, Springer, Berlin/Heidelberg, Germany, 2015, pp. 753–783.
9 W. Srichamnong, P. Ting, P. Pitchakarn, O. Nuchuchua and P. Temviriyanukul, *Food Sci. Nutr.*, 2018, **6**, 962–969.
10 L. B. Bueno-Borges, M. A. Sartim, C. C. Gil, S. V. Sampaio, P. H. V. Rodrigues and M. A. B. Regitano-d'Arce, *J. Food Sci. Technol.*, 2018, **55**, 4159–4166.
11 M. Takechi, K. Doi and Y. Wakayama, *Phytother. Res.*, 2003, **17**, 83–85.
12 J. M. Wierenga and R. M. Hollingworth, *Nat. Toxins*, 1992, **1**, 96–99.
13 R. Lásztity, M. Hidvégi and Á. Bata, *Food Rev. Int.*, 1998, **14**, 371–390.
14 A. M. Bode and Z. Dong, *Cancer Prev. Res.*, 2015, **8**, 1–8.
15 H. N. Matsuura and A. G. Fett-Neto, *Plant Toxins*, 2015, **2**, 1–15.
16 European Food Safety Authority, Technical Report on the notification of roasted seeds from Plukenetia volubilis L. as a traditional food from a third country pursuant to Article 14 of Regulation (EU) 2015/2283, 2020.
17 S. Kittibunchakul, C. Hudthagosol, P. Sanporkha, S. Sapwarobol, P. Temviriyanukul and U. Suttisansanee, *Horticulturae*, 2022, **8**, 344.
18 E. Alexandri, R. Ahmed, H. Siddiqui, M. I. Choudhary, C. G. Tsiafoulis and I. P. Gerothanassis, *Molecules*, 2017, **22**, 1663.
19 P. Siudem, A. Zielińska and K. Paradowska, *J. Pharm. Biomed. Anal.*, 2022, **212**, 114658.
20 F. Wei, K. Furihata, F. Hu, T. Miyakawa and M. Tanokura, *Magn. Reson. Chem.*, 2010, **48**, 857–865.
21 N. Mhd Rodzi and L. Lee, *Heliyon*, 2022, **8**, e10572.
22 R. Chirinos, G. Zuloeta, R. Pedreschi, E. Mignolet, Y. Larondelle and D. Campos, *Food Chem.*, 2013, **141**, 1732–1739.
23 C. Fouzder, A. Mukhuty, S. Mukherjee, C. Malick and R. Kundu, *Toxicol. in Vitro*, 2021, **70**, 105038.
24 J.-Y. Zhou and S.-W. Zhou, *Fitoterapia*, 2012, **83**, 617–626.
25 P. Banerjee, A. O. Eckert, A. K. Schrey and R. Preissner, *Nucleic Acids Res.*, 2018, **46**, W257–W263.
26 S. Hiai, H. Oura and T. Nakajima, *Planta Med.*, 1976, **29**, 116–122.
27 E. P. o. C. i. t. F. Chain, H. K. Knutsen, J. Alexander, L. Barregård, M. Bignami, B. Brüschweiler, S. Ceccatelli, B. Cottrill, M. Dinovi and L. Edler, *EFSA J.*, 2017, **15**, e04908.
28 R. Moreira, D. M. Pereira, P. Valentão and P. B. Andrade, *Int. J. Mol. Sci.*, 2018, **19**, 1668.
29 B. Steinhoff, *Food Chem. Toxicol.*, 2019, **130**, 262–266.
30 N. Mohamadi, F. Sharififar, M. Pournamdari and M. Ansari, *J. Diet. Suppl.*, 2018, **15**, 207–222.
31 N. Konstantinidis, H. Franke, S. Schwarz and D. W. Lachenmeier, *Molecules*, 2023, **28**, 3460.
32 C. Tohda, N. Nakamura, K. Komatsu and M. Hattori, *Biol. Pharm. Bull.*, 1999, **22**, 679–682.
33 J. C. Liao, K. T. Lee, B. J. You, C. L. Lee, W. T. Chang, Y. C. Wu and H.-Z. Lee, *Food Nutr. Res.*, 2015, **59**, 29884.
34 S. Naikoo and S. A. Tasduq, *J. Photochem. Photobiol., B*, 2020, **202**, 111720.
35 I. Rodeiro, D. Remirez and D. Flores, *J. Pharm. Pharmacogn. Res.*, 2018, **6**, 17–26.
36 A. Gorriti, J. Arroyo, F. Quispe, B. Cisneros, M. Condorhuamán, Y. Almora and V. Chumpitaz, *Rev. Peru. Med. Exp. Salud Publica*, 2010, **27**, 352–360.
37 F. Shamsa, H. Monsef, R. Ghamooshi and M. Verdian-rizi, *Thai J. Pharm. Sci.*, 2008, **32**, 17–20.
38 V. Kostik, S. Memeti and B. Bauer, *J. Hyg. Eng. Des.*, 2013, **4**, 112–116.
39 ISO 12966-1, Animal and vegetable fats and oils—Gas chromatography of fatty acid methyl esters—Part 1: Guidelines on modern gas chromatography of fatty acid methyl esters, in *International Organization for Standardization*, Geneva, Switzerland, 2014.
40 Y.-Z. Chen, S.-Y. Kao, H.-C. Jian, Y.-M. Yu, J.-Y. Li, W.-H. Wang and C.-W. Tsai, *J. Food Drug Anal.*, 2015, **23**, 636–644.
41 M. Jia, R. Zhao, B. Xu, W. Yan, F. Chu, H. Gu, T. Xie, H. Xiang, J. Ren and D. Chen, *RSC Med. Chem.*, 2017, **8**, 148–151.
42 M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt and J. Egertson, *Nat. Biotechnol.*, 2012, **30**, 918–920.
43 T. Pluskal, S. Castillo, A. Villar-Briones and M. Orešič, *BMC Bioinf.*, 2010, **11**, 1–11.
44 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapono and T. Luzzatto-Knaan, *Nat. Biotechnol.*, 2016, **34**, 828–837.
45 L.-F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, I. Protsyuk, M. Ernst, H. Tsugawa and M. Fleischauer, *Nat. Methods*, 2020, **17**, 905–908.
46 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.*, 2003, **13**, 2498–2504.
47 J. Wandy, Y. Zhu, J. J. van der Hooft, R. Daly, M. P. Barrett and S. Rogers, *Bioinformatics*, 2018, **34**, 317–318.
48 M. Ernst, K. B. Kang, A. M. Caraballo-Rodríguez, L.-F. Nothias, J. Wandy, C. Chen, M. Wang, S. Rogers, M. H. Medema and P. C. Dorrestein, *Metabolites*, 2019, **9**, 144.
49 H.-S. Shiah, C.-J. Lee, F.-Y. Lee, S.-H. Tseng, S.-H. Chen and C.-C. Wang, *Front. Pharmacol*, 2023, **14**, 1106030.