ORIGINAL ARTICLE



Identification of chemotypes in bitter melon by metabolomics: a plant with potential benefit for management of diabetes in traditional Chinese medicine

Shuaizhen Zhou^{1,3} · Pierre-Marie Allard¹ · Christian Wolfrum² · Changqiang Ke³ · Chunping Tang³ · Yang Ye³ · Jean-Luc Wolfender¹

Received: 12 October 2018 / Accepted: 6 July 2019 / Published online: 18 July 2019 © Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Introduction Bitter melon (*Momordica charantia*, Cucurbitaceae) is a popular edible medicinal plant, which has been used as a botanical dietary supplement for the treatment of diabetes and obesity in Chinese folk medicine. Previously, our team has proved that cucurbitanes triterpenoid were involved in bitter melon's anti-diabetic effects as well as on increasing energy expenditure. The triterpenoids composition can however be influenced by changes of varieties or habitats.

Objectives To clarify the significance of bioactive metabolites diversity among different bitter melons and to provide a guideline for selection of bitter melon varieties, an exploratory study was carried out using a UHPLC-HRMS based metabolomic study to identify chemotypes.

Methods Metabolites of 55 seed samples of bitter melon collected in different parts of China were profiled by UHPLC-HRMS. The profiling data were analysed with multivariate (MVA) statistical methods. Principle component analysis (PCA) and hierarchical cluster analysis (HCA) were applied for sample differentiation. Marker compounds were identified by comparing spectroscopic data with isolated compounds, and additional triterpenes were putatively identified by propagating annotations through a molecular network (MN) generated from UHPLC-HRMS & MS/MS metabolite profiling.

Results PCA and HCA provided a good discrimination between bitter melon samples from various origins in China. This study revealed for the first time the existence of two chemotypes of bitter melon. Marker compounds of those two chemotypes were identified at different MSI levels. The combined results of MN and MVA demonstrated that the two chemotypes mainly differ in their richness in cucurbitane versus oleanane triterpenoid glycosides (CTGs vs. OTGs).

Conclusion Our finding revealed a clear chemotype distribution of bioactive components across bitter melon varieties. While bioactivities of individual CTGs and OTGs still need to be investigated in more depth, our results could help in future the selection of bitter melon varieties with optimised metabolites profile for an improved management of diabetes with this popular edible Chinese folk medicine.

Keywords Bitter melon \cdot *Momordica charantia* \cdot Molecular networking \cdot PCA \cdot UHPLC-HRMS \cdot Triterpenoids \cdot Metabolomics

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11306-019-1565-7) contains supplementary material, which is available to authorized users.

Yang Ye yye@simm.ac.cn

Jean-Luc Wolfender Jean-Luc.Wolfender@unige.ch

Extended author information available on the last page of the article

1 Introduction

Herbs used in traditional medicine are known to contain potent active principles and may represent a valuable alternative notably for the treatment of chronic disease. It is however also known that it is quite common that the compositions of medicinal plants could notably change along with varieties, seasons, habitats, species misidentifications, etc. (Kaiser et al. 2016; Pan et al. 2016; Soro et al. 2016). The existence of different chemotypes could lead to differences in efficacy, or even safety of derived drugs (Bhatia et al. 2013). Because of the inherent complexity of plant composition, metabolomics is well suited for assessing globally changes in metabolites that can be of importance in assessing quality of herbs (Yang et al. 2006; Ye et al. 2007; Bhatia et al. 2013).

Bitter melon (Momordica charantia Linn.), also known as bitter gourd, is a world widely cultivated edible and medicinal vine from the Cucurbitaceae family. It has been used as a botanical dietary supplement in middle and far east Asia for treatment of diabetes (Leung et al. 2009), obesity (Popovich et al. 2011; Bao et al. 2013; Xu et al. 2014) and a number of reported diseases (Singh et al. 1998a, b; Sathishsekar and Subramanian 2005). In depth investigation of this plant have revealed over 70 triterpenoid saponins whose structures are categorized in two types of triterpene skeletons, known as cucurbitane-type triterpene glycosides (CTGs) and oleanane-type triterpene glycosides (OTGs) (Tomita et al. 2005; Nakamura et al. 2006; Tan et al. 2008). In previous bioactivity studies, CTGs, an exclusive type of tetracyclic triterpenoids in Cucurbitaceae family, have demonstrated anti-diabetic and hypoglycaemic effects (Harinantenaina et al. 2006; Keller et al. 2011). Our team has proved that CTGs play a critical role in bitter melon's anti-diabetic activity through stimulation of GLUT4 translocation and activation of AMPK signaling pathway. CTGs further enhance fatty acid oxidation and glucose disposal in mice (Tan et al. 2008; Iseli et al. 2013). Meanwhile, Oleanolic acid, one of the OTGs aglycones, is widely distributed in plants and its anti-diabetic activities have been documented (Ayeleso et al. 2017). OTGs from other species than bitter melon are known to exhibit hypoglycemic properties as well (Duan et al. 2016; Luyen et al. 2018). Moreover, we demonstrated that oleanolic acid showed anti-hyperglycemia effect on type 2 diabetic mice (Zeng et al. 2012). Those discoveries confirmed ancient records and recent pharmacology results of bitter melon's efficacy at the molecular level.

Our previous phytochemical investigations also demonstrated the complex composition of such plants. Bioactivity variations and changes in concentrations of the bioactive metabolites along with the varieties or habitats were noticed in sample to sample comparison. In order to obtain an overview of the metabolome variations that are likely to occur, and to lay the chemical foundation in view of a future possible evidenced based therapeutic usage of bitter melon, 55 bitter melon varieties were collected in different parts of China and all samples were submitted to a MS-based metabolomics investigation.

Our strategy consisted in the profiling of all extracts with an optimized UHPLC-HRMS for discrimination by PCA and HCA followed by pooling of samples for extensive metabolite annotation by molecular networking (MN) (Pan et al. 2016; Wang et al. 2016; Mohimani et al. 2017; Wolfender et al. 2019b).

2 Materials and methods

2.1 Chemicals and reagents

Reference compounds goyasaponins I, II, momordicosides A, B, S and T from bitter melon were isolated and identified in Prof. Yang Ye's lab at Shanghai Institute of Materia Medica, Chinese Academy of Sciences (Tan et al. 2008).

2.2 Plant material, extraction and sample preparation

55 bitter melon seed samples were purchased from different local seed companies in China (Table S1). 10 g of each dried sample were used for extraction. The remaining material was kept as voucher specimen samples deposited in the Herbarium at the Shanghai Institute of Materia Medica, Chinese Academy of Sciences. The seeds were grounded and extracted with 70% ethanol (1:10 w/v) using ultra-sonication for 45 min. Process was repeated three times. Each solution was centrifuged at 4000 rpm for 10 min. The supernatant was filtered and evaporated under reduced pressure to remove solvents and yield dried extract. Ten mg of extract were dissolved in 85% methanol–water as a 1 mg/ml solution. Filtration on a 0.22 μ m membrane was applied before UHPLC analysis. In addition, a pool of all extracts solutions was prepared and used as pooled sample mixture.

2.3 UHPLC-HRMS profiling

All extract samples were subjected in a random sequence for metabolite profiling on a Waters Acquity UPLC system coupled with a Waters/Micromass LCT Premier time of flight (TOF) mass spectrometer. A mixture of reference saponins was inserted into the sequence every 10 injections as reference sample. Goyasaponins I (labeled as **A9**) and II (**A11**), as well as momordicosides A (**B5**), B (**B9**), S (**B2**) and T (**B3**) were concomitantly analyzed. The structures of the reference compounds see Fig. 1.

The injection volume was 2 μ l. Temperature of sample chamber was kept at 10 °C. Water (UPLC grade, Waters) with 0.1% formic acid and acetonitrile (UPLC grade, Waters) with 0.1% formic acid were used for phase A and B, respectively. An Acquity UPLC BEH RP18 column (1.7 μ m, 100×2.1 mm; Waters) was used at a column temperature of 45 °C. The chromatographic method started from an isocratic elution with 20% B for 0.5 min, continued with a linear gradient from 23 to 43% B in 7.0 min then increased to 95% B in another 3 min. A washing procedure was followed using 95% B for 2 min. The column was then equilibrated using 20% B for 2.5 min before next injection. The



Fig. 1 Chemical structures of reference compounds isolated from M. charantia extracts

flow rate was 0.8 ml/min. ESI negative full scan mode was used for its high ionization efficiency, the data were acquired from 200 to 2000 m/z with the following parameters: capillary voltage, 2.4 kV; sample cone voltage, 40.0 V; source temperature, 120.0 °C; desolvation gas temperature (N₂), 300.0 °C; desolvation gas flow, 700 unit; aperture 1 voltage, 35.0; resolution, 7500. A real-time calibration was performed using 20 ng/ml leucine-enkephalin solution in 50/50 acetonitrile–water as lock spray (lock mass m/z 554.2615).

2.4 UHPLC-HRMS-HRMS/MS profiling

A pooled sample of all extracts and reference compounds were analyzed on a Thermo Dionex Ultimate 3000 UHPLC interfaced to a Thermo Q-Exactive Plus mass spectrometer with a heated electrospray ionization (HESI) source. The LC method used for detailed metabolite profiling with higher resolution was as follows: column, Acquity UPLC BEH RP18 column ($1.7 \mu m$, $150 \times 2.1 mm$; Waters), column temperature, 45 °C, mobile phase, water (A) and acetonitrile (B) both with 0.1% formic acid, flow rate 0.5 ml/min, injection volume, 2 µl. The gradient started from a 20% isocratic elution for 1.5 min then followed by a linear gradient of 20% to 43% B until 18 min, and continued increasing to 95% B in 7 min. The column was washed using 95% B for 5 min and equilibrated 10 min before next injection. The analyses were performed in negative mode with following parameters: source voltage, 2.5 kV; capillary temperature, 256.25 °C; sheath gas flow (N₂), 47.50 units; auxiliary gas flow, 11.25; spare gas flow, 2.25; probe heater temperature, 412.50 °C; S-Lens RF Level, 50. Full MS resolution, 70,000; scan range, 300-2000 m/z. The data-dependent MS/ MS experiments were performed on the four most intense ions and the resolution was 17,500. The stepped normalized collision energy (NCE) was set as 15-30-45 units, and the isolation window was 2.0 m/z. A dynamic exclusion was applied on parent ions for 2.0 s after being fragmented.

2.5 Molecular networking

Molecular networks were created by global natural products social (GNPS, http://gnps.ucsd.edu) online workflow (Wang et al. 2016) with the following parameters: precursor ion mass tolerance, 1.0 Da; MS/MS fragment ion mass tolerance, 0.8 Da. Filters were applied to remove all edges which cosine score was below 0.7 and match fragment peaks <6. The maximum number of neighbor nodes for one single node was set to 10. The final results were visualized both online at GNPS and using Cytoscape 3.4.0. All self-loops were removed at the data import step.

2.6 Multivariate statistical analysis

The UHPLC-TOF/MS profiling dataset was processed using the open-source software MZmine 2 to extract and align peaks (Pluskal et al. 2010). The peak detection was performed using 2D Grid Mass algorithm using following parameters: minimum height 100, m/z tolerance 0.06, min-max width time 0.03–0.4. Isotopic peaks grouping was using m/z tolerance 0.05, max charge 2 and representative isotope with lowest m/z. RANSAC aligner was applied using 50% min number of points and threshold 0.05. Duplicate peak filter and peak finder gap filling were applied then and a peak list of 511×55 matrix was obtained.

The matrix was normalized and then imported to SIMCA 13.0 (Umetrics, Sweden) and pretreated with mean-centered and Pareto variance scaling. A PCA model was generated with the data matrix. A Hierarchical clustering analysis (HCA) was performed using PCA algorithm but with > 95% ($R^2X = 0.952$, 20 PCs) information of original feature of samples.

Potential markers were selected using a PCA "group to group comparison" scores contribution plot. A threshold score > 1.2 was applied to select variables that contributed to the differences between the two groups. Isotopes and adducts were removed manually from the final marker list by checking UHPLC-MS raw data.

Heat-maps were created using MeV 4.9.0. The data set was processed following normalization, divided by SD and meancentered. Hierarchical clustering analysis was performed based on Euclidean distance.

2.7 Quantitation of four marker compounds using UHPLC-MS/MS

Four compounds, goyasaponin I (A9), goyasaponin II (A11), momordicoside S (B2) and momordicoside A (B5), were quantitated against an external standard calibration. All stock solutions of reference compounds and extract were prepared in 85% MeOH (HPLC grade, Thermo) at a concentration of 1 mg/ml.

Quantitation was performed using an Agilent 1290 infinity LC-system interfaced to an Agilent 6460 triple quadrupole mass spectrometer equipped with ESI source. Two µl of each sample were injected in UHPLC under following conditions: column, Acquity UPLC BEH RP18 column (1.7 µm, 100×2.1 mm; Waters); column temperature, 40 °C; mobile phase, water (A) and acetonitrile (B) both with 0.1% formic acid, flow rate 0.5 ml/min. The elution was performed as a linear gradient from 30 to 50% B in 5 min, and followed by a 2 min washing procedure with 95% B. Then the column was equilibrated 3 min before next injection. Quantitation was performed in multiple reaction monitoring (MRM) in ESI negative ion mode using following parameters: capillary voltage, 3.5 kV; gas flow, 10 l/min; gas temperature, 320 °C; nebulizer, 50 psi; sheath gas temperature, 350 °C; sheath gas flow, 11 l/ min; collision gas, helium. The MRM method was optimized by injection of a standard compound solution, as shown in Table S2. The dwell time for each monitored transition was 120 ms. Each sample was injected as three analytical replicates. All accuracies and precisions were within the range of $\pm 15\%$, which is in accordance with the bioanalytical method validation guidelines of the US Food and Drug Administration (FDA 2013) and recovery of spiked compounds were above 95%.

Mann–Whitney tests were performed using followed by Dunnett's multiple comparisons test was performed using GraphPad Prism version 6.0c, USA.

3 Results

Fifty-five purchased bitter melon seed samples (Table S1) were grinded and extracted with 70% ethanol–water. MVA specimen differentiation metabolite profiling of these crude extracts was performed by full scan UHPLC-HRMS on a TOF/MS instrument with an optimized reversed phase gradient. In parallel, the pool of all samples was profiled on an Orbitrap with a data-dependent acquisition using a longer LC method with similar gradient to acquire MS/MS spectra automatically on all main triterpenoids for annotation purposes. For improved marker identification two OTGs, goyasaponins I and II, as well as four CTGs, momordicosides A, B, S and T, which were previously isolated from bitter melon, were concomitantly analyzed. Database searching and MS/MS spectra analyses were performed as well for marker compounds dereplication and identification.

3.1 Differentiation of bitter melon extracts based on UHPLC-HRMS profiling

Each bitter melon seed sample was profiled by UHPLC-HRMS to investigate trends in composition modifications across varieties and locations. All samples were preprocessed by MZmine and a matrix of 55×511 features was generated and normalized. To view the differences across this complex dataset, the data were visualized using a heat-map of the HCA results that provided an overview of all observed metabolites (Fig. 2a). As shown in the heat map, clear chemical constituent's composition differences between different bitter melon samples were observed, which may suggest differences in bioactivities.

Differentiation of the samples was performed by PCA and HCA after mean centering and pareto scaling. The PCA gave a model that represented 73.1% (PC1 + PC2 + PC3) of original dataset information. The parameters of this model ($R^2X = 0.731$ and $Q^2 = 0.529$) validated the quality of this PCA. The PCA score plot (Fig. 2b) provided a good visualization of the sample set relationship and further confirmed that the PCA model was reliable for representing bitter melon accessions.

To investigate the differentiation among bitter melon samples, the HCA was analyzed in more details. The HCA dendrogram (Fig. 2c) revealed two main branches corresponding to two groups of bitter melon samples with distinct chemical compositions. Based on this, samples were labeled with different colors to show the groups identified in the PCA



Fig. 2 Multivariate statistical analysis of UHPLC-HRMS profiling data acquired from *M. charantia* extracts. **a** Heat-map of *M. charantia* extracts corresponding to relative peak areas. Columns represent 55 varieties and rows represent 511 detected metabolites. **b** PCA score plot derived from profiling data of all samples, **c** HCA den-

drogram, Group A colored in green and Group B in blue were cut off at first branch, **d** score plot of PCA marked CA and CB. **e** PCA "group to group comparison" contribution scores plot of CB-CA, weight=p1p2. Threshold score>|1.2| were used to select marker compounds (colored in red) (Fig. 2d). These combined PCA/HCA results provided a good discrimination of all samples and clearly evidenced the presence of two chemotypes of bitter melon varieties. Hereinafter, the samples colored in green in Fig. 2d are referred as chemotype A (CA) while the blue ones are described as chemotype B (CB). According to PCA/HCA, 24 bitter melon samples were assigned as CA while 31 were assigned as CB (see Table S1). The discovery of two chemotypes of bitter melon could lead finally in different biological activities.

3.2 Selection of marker compounds

To understand which specific metabolites cause the differences, the PCA contribution plot of scores was analyzed to identify the features significantly contributing to the chemotype differentiation.

In the contribution plot (Fig. 2e) a contribution score above 1.2 was chosen as a threshold and after filtering extra adducts and isotopes, a marker compound list was obtained

as shown in Table 1. Compounds A1 to A13 were found to be potential markers of CA and B1 to B15 as potential markers of CB. These compounds represent the characteristics of two types of bitter melon, and identification of these marker compounds allows us to document at the molecular level the trends in composition that clearly differentiate those bitter melon chemotypes.

3.3 Marker compounds annotation via molecular networking

By comparing UHPLC-TOF/MS retention time and HRMS¹ data with available reference compounds from previous studies, (Okabe et al. 1980; Murakami et al. 2001; Tan et al. 2008) six marker compounds were unambiguously identified (LEVEL1) (Table 1). This included two OTGs (**A9** goyasaponin I and **A11** goyasaponin II) for CA and four CTGs (**B2** momordicoside S, **B3** momordicoside T, **B5** momordicoside A and **B9** momordicoside B) for CB. However, this was

Table 1 Selected marker compounds of two bitter melon chemotypes

INO.	m/z	∆ppm	t _R (min)	Molecular formula	Ion description	Classification/identification	MSI level ^a
A1	1657.6994	-9.7	3.50	C ₇₅ H ₁₁₈ O ₄₀	[M-H] ⁻	OTG	2
A2	1511.6402	-9.3	3.63	C ₆₉ H ₁₁₀ O ₃₆	$[M - H]^{-}$	OTG	2
A3	1525.6578	-7.9	4.36	C ₇₀ H ₁₁₀ O ₃₆	$[M - H]^{-}$	OTG	2
A4	1219.5671	-6.3	4.90	C ₅₈ H ₉₂ O ₂₇	$[M - H]^{-}$	OTG	2
A5	1393.6156	-8.6	5.13	C ₆₅ H ₁₀₂ O ₃₂	$[M - H]^{-}$	OTG	2
A6	762.3272	-5.7	5.21	$C_{70}H_{110}O_{36}$	$[M - 2H]^{2-}$	OTG	2
A7	1567.6650	-9.8	5.69	C ₇₂ H ₁₁₂ O ₃₇	$[M - H]^{-}$	OTG	3
A8	1509.6594	-10.3	5.79	C ₇₀ H ₁₁₀ O ₃₅	$[M - H]^{-}$	OTG	2
A9	1377.6242	-6.2	6.18	C ₆₅ H ₁₀₂ O ₃₁	$[M - H]^{-}$	Goyasaponin I	1
A10	1363.6080	-6.6	6.22	$C_{64}H_{100}O_{31}$	$[M - H]^{-}$	OTG	2
A11	1509.6676	-4.8	6.39	C ₇₀ H ₁₁₀ O ₃₅	$[M - H]^{-}$	Goyasaponin II	1
A12	1479.6519	-8.4	6.55	$C_{69}H_{108}O_{34}$	$[M - H]^{-}$	OTG, new compound	3
A13	476.2741	-4.2	9.10	C25H39NO6	$[M - H]^{-}$	Unknown	4
B1	1271.6163	-8.6	1.84	$C_{59}H_{100}O_{29}$	$[M - H]^{-}$	CTG, new compound	3
B2	977.5288	-3.4	2.21	$C_{48}H_{82}O_{20}$	$[M - H]^{-}$	Momordicoside S	1
B3	1109.5669	-6.8	2.31	$C_{53}H_{90}O_{24}$	$[M - H]^{-}$	Momordicoside T	1
B4	961.5332	-4.2	2.64	$C_{48}H_{82}O_{19}$	$[M - H]^{-}$	CTG	2
B5	815.4786	-0.9	3.02	$C_{42}H_{72}O_{15}$	$[M - H]^{-}$	Momordicoside A	1
B6	947.5192	-2.5	3.13	$C_{47}H_{80}O_{19}$	$[M - H]^{-}$	Momordicoside B	1
B7	799.4823	-2.6	3.96	$C_{42}H_{72}O_{14}$	$[M - H]^{-}$	CTG	2
B8	977.5291	-3.1	4.05	$C_{47}H_{80}O_{18}$	$[M + HCOOH - H]^{-}$	CTG	3
B9	813.4613	-2.8	4.06	$C_{42}H_{70}O_{15}$	$[M - H]^{-}$	CTG	2
B10	695.3987	-2.9	4.14	$C_{37}H_{60}O_{12}$	$[M - H]^{-}$	Unknown	4
B11	845.4878	-2.5	4.65	$C_{42}H_{72}O_{14}$	$[M + HCOOH - H]^{-}$	CTG	2
B12	781.4711	-3.5	5.01	$C_{42}H_{70}O_{13}$	$[M - H]^{-}$	CTG	2
B13	959.5189	-2.8	5.14	$C_{47}H_{78}O_{17}$	$[M + HCOOH - H]^{-}$	CTG	3
B14	797.4655	-4.0	5.61	$C_{42}H_{70}O_{14}$	$[M - H]^{-}$	CTG	2
B15	829.4925	-2.9	6.28	$C_{42}H_{72}O_{13}$	$[M + HCOOH - H]^{-}$	CTG	2

not sufficient for a comprehensive survey of all compounds responsible for the differences between bitter melon chemotypes. In order to annotate the largest possible number of marker compounds, MS/MS spectra of the pool of all samples and a mixture of references were recorded by Orbitrap with a similar UHPLC gradient and processed by molecular networking (MN) analysis (Wang et al. 2016).

A MN composed by 1405 nodes was generated from our dataset which included 24'724 MS/MS spectra (Fig. S1). In the first step of dereplication, six nodes were unambiguously annotated from the whole network by comparing the spectral data (retention time, MS and MS/MS) of the six reference compounds, goyasaponin I (A9), goyasaponin II (A11), momordicoside S (B2), momordicoside T (B3), momordicoside A (B5) and momordicoside B (B9). As highlighted in Fig. S1, this allowed the assignment of one cluster to OTG- and another to CTG-type triterpenoids. This was used to propagate the annotation to other triterpenoid derivatives in both clusters. The results indicated that these two clusters (referred as clusters A and B in the following text) indeed corresponded to OTGs or CTGs, respectively (Fig. 3).

In the second step, dereplication of all metabolites related to nodes connected in these two clusters was carried out against the dictionary of natural products (DNP) (Hall 2017) based on UHPLC-HRMS¹ data restricted to compounds isolated from Curcurbitaceae family. For this a peak list was constructed after peak picking by retrieving chromatographic data regarding to parent ion masses and retention times of each nodes. The masses were compared with calculated exact masses of compounds in DNP with mass tolerance of ± 3 ppm. As a result, 15 nodes were characterized as OTGs within cluster A in addition to the references, including nine marker compounds of CA (Fig. 3a). Similarly, in cluster B, 25 nodes in total were annotated as CTGs containing 14 markers of CB (Fig. 3b). Most of the corresponding features appeared as $[M - H]^-$ ions while some of them also occurred as $[M - HCOO]^-$ (Table 1). The results of the DNP search confirm the classification by MN.

Combination of these results enabled the annotation of 6 triterpene glycosides at level 1, 36 at level 2 and 3 in cluster A and B (Sumner et al. 2007). This also revealed the presence of still unreported OTGs and CTGs according the molecular ions features detected (around 37% of all nodes detected on the two cluster assignments made). Most of the marker compounds of CA and CB were found and annotated in the MN except A4, A6, A8, A13 and B10 (Fig. 3). As observed in the MN, markers of CA were mainly found in cluster A, while almost all the markers of CB were included in cluster B. These results clearly indicate that bitter melon sample belonging to chemotype A represent OTGs rich varieties and those belonging to CB are CTGs rich. This information should provide the basis for an appropriate selection of bitter melon varieties based on their OTGs versus CTGs composition in future, once a clear involvement of the different constituents in the overall antidiabetic effects will be established.

The MN of bitter melon can be used in LC–MS guided isolation to find derivatives of bioactive compounds such as CTGs **B2** and **B3** (Tan et al. 2008; Iseli et al. 2013) in future studies. This analysis also showed that MN is capable of identifying a significant number of previously unknown triterpenoids analogues from *M. charantia* compared to those previously isolated from this species. Overall, MN navigation combined with LC–MS guided isolation will permit to efficiently obtain unreported OTGs or CTGS derivatives and



Fig. 3 a Clusters A corresponding to OTGs and b clusters B corresponding to CTGs in the molecular network of *M. charantia* extract. Dereplication was made using precursor ion mass against DNP

thus expand our bitter melon compound library, for documenting their bioactivities.

3.4 Identification of new marker compounds in two bitter melon chemotypes

Markers not annotated in MN were investigated for further dereplication by searching their molecular formula against the DNP, A4, A6 and A8 were putatively identified as OTGs. A13 and B10 remained unknown as no dereplication hits was found in DNP. Moreover, A12 and B1 were annotated as OTG and CTG by MN but their molecular formula was still unreported. This indicated that A12 and B1 are probably new natural products (Table 1). Their structures were further investigated by interpretation of their HRMS-MS/ MS spectra.

The MS/MS spectrum of A12 was compared with that of A11, which identified unambiguously as goyasaponin II $(m/z 1509.7, cal. C_{70}H_{109}O_{35}[M-H]^{-})$. A11 showed major fragment ions m/z 807.4, 627.4, 583.4 and 469.3 in MS/ MS spectra, which indicated its characteristic fragmentation pathway started with a neutral loss of saccharide moiety at C-28 (Fig. S2a) forming goyasaponin III (Shibuya et al. 1990), another OTG reported from M. charantia. The fragmentation continued with the loss of one or two glycosyl groups at C-3 acquired m/z 627.4 or 469.3 as described in Fig. S2a. Comparison of MS/MS spectra of A12 and A11 revealed that they had high similarity, except the $[M - H]^{-}$ mass m/z 1479. 6 and a high abundance fragment peak m/z 777.4 in A12 that were 30 Da less compared to A11 (1509.7 and 807.4). This suggested that they share a similar structure except for the glycoside moiety attached to the glucuronic acid group which is a pentose instead of a galactose in A12. The elemental compositions of all fragments were calculated based on exact masses with mass tolerance of ± 2 ppm. Thus, the most probable structure of compound A12 was proposed as gypsogenin-3-O-[pentosylglucuronopyranoside],28-O-[xylopyranosyl-xylopyranosylrhamnopyranosyl-(rhamnopyranosyl)-fucopyranosyl] ester (Fig. S2).

The MS/MS spectrum of **B1** was highly similar to that of **B3**, momordicoside T (Fig. S3). All major peaks in the spectrum of **B3** including the deprotonated molecular ion could be found in that of **B1**, which revealed that **B1** shared the same moiety of **B3** after a characteristic glucosyl neutral loss of 162 Da ($C_6H_{10}O_5$). **B1** was thus and analogue of **B3** with an additional hexose motley as shown in Fig. S3b. The elemental compositions of all fragments were calculated with mass tolerance within ±2 ppm. Thus, two new compounds **A12** and **B1** were identified. Their structure types revealed by MS/MS analysis also verified that the chemotype differentiation of bitter melon and compounds classification model are reliable. As a summary 12 out of 13 markers of CA were identified as OTGs while 13 out of 14 markers of CB were assigned to CTGs. This suggested that the two chemotypes of bitter melon mainly differ in in their richness either in cucurbitane or oleanane triterpenoids. It is particularly worth to mention that previous bioactivity studies (Harinantenaina et al. 2006; Tan et al. 2008; Keller et al. 2011; Iseli et al. 2013) indicated that CTGs are most probably involved in the antidiabetic activity of bitter melon. The existence of two well distinct chemotypes revealed by this study may influence the pharmacological effect of the extract and such results need to be taken account into consideration for further quality control studies for the usage of bitter melon in diabetes management.

3.5 Content comparison of marker compounds between chemotypes A and B

To verify contribution of the marker compounds related to the differentiation, a heat-map with HCA in both rows and columns was generated based on UHPLC-TOF/MS peak area data (Fig. 4). As shown, the contents of markers in samples obviously matched with the chemotypes discrimination in Table S1, all CA varieties are OTGs rich while CB varieties are CTGs rich. As an example, "E05 Feicuilvzao" is a CA bitter melon and it has higher content of OTG markers than "E30 Xiangbaiyu" which is assigned as CB and is richer in CTG markers.

To quantitatively inspect the differences between those bitter melon chemotypes, a quantitation of the four of the most abundant marker compounds was performed. Two OTGs A9 and A11, as well as two CTGs—B2 and B5 were chosen, and their concentrations in 55 bitter melon extracts were quantified through UHPLC-MS/MS in Multiple Reaction Monitoring (MRM) mode. The calibration curve range and equation of each compound was: 0.1, 0.2, 0.5, 1.0 and 2.0 μ g/ml for A9 (y=0.5456x+13.8353); 0.2, 0.5, 1.0, 2.0 and 5.0 μ g/ml for A11 (y = 0.2356x + 37.9698); 0.05, 0.1, 0.5, 1.0 and 5.0 μ g/ml for **B2** (y = 0.6543x - 3.9605); 0.05, 0.1, 0.5, 1.0 and 5.0 μ g/ml for **B5** (y = 1.2231x - 6.1481). The concentrations of compounds in each sample are shown in Table S1. Mann-Whitney tests were performed on the quantitation results. Figure 5 showed that the concentrations of A9 (p = 0.0423) and A11 (p = 0.0004) in CA were significantly higher than that in CB. In contrast, CTGs B2 and **B5** were significantly more abundant in CB (p < 0.0001for both compounds). Since **B5** showed low *p* value and the most significant inter-group difference, B5 was selected as a good marker to distinguish those two chemotypes. Both heat-map and quantitation results proved that the selected marker compounds contribute to chemotype discrimination. Quantitation of **B5** provided an efficient way to preliminary assign the bitter melon chemotypes.



Fig. 4 Heat-map of selected marker compounds in bitter melon extracts. Columns represent 55 varieties and rows represent 28 marker compounds of two chemotypes. Colors correspond to profiling peak areas



Fig. 5 Marker compounds' concentrations in CA and CB dry samples: A9 goyasaponin I, A11 goyasaponin II, B2 momordicoside S and B5 momordicoside A. $*P \le 0.05$, $**P \le 0.01$

Furthermore, by analyzing bitter melon samples across various regions in China, PCA score plot showed that there is no consistent correlation between geographical distributions and chemotype distributions. For example, bitter melons collected from Shandong, Hunan and Guangdong provinces belong to one of the chemotypes respectively. In many provinces however, the varieties from a given location do not belong to the same chemotypes, and significant chemical component differences between different provinces were observed (Fig. S4). The differences among these locations indicate that it is still difficult to correlate bitter melon chemotype in habitats and morphological traits. Therefore, identification of marker compounds in bitter melon varieties would indeed be necessary for chemotype assignment.

4 Discussion

Bitter melon is a traditional Chinese medicine recorded in ancient documents. Nowadays it has also been studied as an anti-diabetic traditional medicine in clinical practice. Many previous studies have reported the anti-diabetic and hyperglycemic effects of CTGs. Our team has also previously conducted thorough chemical investigations, isolated various compounds and characterized their biological properties in vitro and in vivo assays. Several CTGs, represented by B2 and B3, showed promoting GLUT4 translocation effect associated with activation of AMPK, a key pathway leading to upregulation of fatty acid oxidation and glucose disposal. Those findings demonstrated that CTGs might played a therapeutic role in combating diabetes and obesity. In addition, OTGs from other plants have been reported to exhibit hyperglycemia effects (Duan et al. 2016; Luyen et al. 2018). To our knowledge no anti-diabetic activities for OTGs in bitter melon have been reported. We have however demonstrated that oleanolic acid, which is one of the aglycones of OTGs widely distributed in plants, showed anti-hyperglycemia effect on type 2 diabetic mice (Zeng et al. 2012). These studies indicated that through different mechanisms, both CTGs and OTGs may be involved at the molecular level in the effects of bitter melon. This also proposed the necessity of differentiation between CTGs and OTGs rich bitter melon varieties. In the present study, we successfully established a reliable and practical method of identifying bitter melon varieties with different chemotypes.

One of the issues related to the characterization of bitter melon bioactivity is their composition. This vegetable has been artificially cultivated all over the world, with nearly 100 varieties in China alone. To our knowledge our study establishes for the first time the presence of two chemotypes of bitter melon. At present however, no study explored the relation between the varieties and the impact on the possible health benefit of bitter melon. Based on our results, the impact of chemical differences between CA and CB on bitter melon's bioactivities still requires further investigations. It is of great significance to reveal the relationship between CTGs/OTGs content and the activities of bitter melon extracts, and our research has laid the chemical foundation for future exploration of different therapeutic potential in bitter melon variations.

The PCA score plot also showed the chemotype distribution of bitter melon varieties in different regions of China (Fig. S4). Some geographical distribution trends could be found. For example, the varieties from Shandong, Hunan and Guangdong provinces are belong to the same chemotype, respectively. However, it is worth mentioning that in many cases the chemotype cannot be linked to a given province. These might be caused by the long-term artificial planting and cultivation that have randomized the distribution of various varieties, and it makes the relationship between bitter melon varieties and their activities more complicated and random. Thus, having marker to characterize chemotype will be key in future. The result also illustrated the importance of metabolomics to tackle such issue.

To the best of our knowledge this is the first comprehensive metabolomics study on bitter melon. The integration UHPLC-HRMS¹ and MN provided a good strategy for highlighting composition traits across varieties and for the annotation of large numbers of triterpenoids. The MN that was established expanded the pool of triterpenoids in bitter melon, and it can be also used for tracing more bioactive derivatives from bitter melon extracts. Further study on mass targeted isolation of CTGs based on MN exploration will be carried to identify new bioactive derivatives and we have recently established efficient high resolution preparative strategies for this (Queiroz et al. 2019; Wolfender et al. 2019a).

The multivariate models and sample set we built in this study not only have provided the differentiation of bitter melon varieties, but also can be useful on classify more bitter melon varieties by PCA model fitting. In addition, **B5** is a known marker compound with the most significant difference of content between CA and CB. Quantification of **B5** and could be related to the CA and CB sample sets and can thus provide a simplified method for the preliminary chemotype assignment of different bitter melon varieties.

Our results should help to select plant sources with defined chemical composition and thus reproducible efficacy/safety profile. Such comprehensive chemical characterization is valuable contribution of modern research to traditional Chinese medicine.

5 Conclusion

In this study, it was found for the first time that bitter melon on the market could be roughly divided into two chemotypes. The identification of marker compounds showed that one of the bitter melon chemotype is richer in CTGs, while the other contains mainly OTGs. The results were further validated by UHPLC-MRM quantitation. We have proved in previous studies that CTGs have good anti-diabetic effects. The discovery of these two defined chemotypes is important for the selection of bitter melon varieties rich in bioactive components for the development of diabetes therapeutics or dietary supplements. Further studies on the bioactivities of both CTGs and OTGs are required to investigate in depth in order to verify if the final antidiabetic effects can be linked to a given chemotype.

Acknowledgements We are thankful for the financial support of the Sino-Swiss Science and Technology Cooperation (SSSTC) program (Grant No. EG 08-032014). Financial support from the National Natural Science Foundation of China (Grant Nos. 81573305, 81673327) is also acknowledged.

Author contributions YY and JLW designed the project and supervised the data analysis. SZ performed the chemical experiments and UHPLC-HRMS analysis. SZ and PMA performed data analysis. SZ, CQ and CT performed the quantitation analysis. SZ and CW wrote the manuscript. All authors edited and approved the manuscript.

Data availability The data in this study are deposited in MetaboLights (https://www.ebi.ac.uk/metabolights/) study identifier MTBLS701.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human and animal participants This article does not contain any studies with human and/or animal participants performed by any of the authors.

References

- Ayeleso, T. B., Matumba, M. G., & Mukwevho, E. (2017). Oleanolic acid and its derivatives: Biological activities and therapeutic potential in chronic diseases. *Molecules*, 22(11), 1915. https:// doi.org/10.3390/molecules22111915.
- Bao, B., Chen, Y. G., Zhang, L., Na Xu, Y. L., Wang, X., Liu, J., et al. (2013). *Momordica charantia* (Bitter Melon) reduces obesityassociated macrophage and mast cell infiltration as well as inflammatory cytokine expression in adipose tissues. *PLoS ONE*, 8(12), e84075. https://doi.org/10.1371/journal.pone.0084075.
- Bhatia, A., Bharti, S. K., Tewari, S. K., Sidhu, O. P., & Roy, R. (2013). Metabolic profiling for studying chemotype variations in *Withania* somnifera (L.) Dunal fruits using GC-MS and NMR spectroscopy. *Phytochemistry*, 93, 105–115. https://doi.org/10.1016/j.phyto chem.2013.03.013.
- Duan, J., Yin, Y., Cui, J., Yan, J., Zhu, Y., Guan, Y., et al. (2016). Chikusetsu saponin IVa ameliorates cerebral ischemia reperfusion injury in diabetic mice via adiponectin-mediated AMPK/ GSK-3beta pathway in vivo and in vitro. *Molecular Neurobiol*ogy, 53(1), 728–743. https://doi.org/10.1007/s12035-014-9033-x.
- FDA. (2013). Bioanalytical method validation. Guidance for industry. http://www.fda.gov/Drugs/GuidanceComplianceRegulat

oryInformation/%20Guidances/ucm064964.htm. Accessed February 2017.

- Hall, C. (2017). *Dictionary of natural products on CD-ROM*. London: Chapman & Hall.
- Harinantenaina, L., Tanaka, M., Takaoka, S., Oda, M., Mogami, O., Uchida, M., et al. (2006). *Momordica charantia* constituents and antidiabetic screening of the isolated major compounds. *Chemi*cal and Pharmaceutical Bulletin (Tokyo), 54(7), 1017–1021.
- Iseli, T. J., Turner, N., Zeng, X. Y., Cooney, G. J., Kraegen, E. W., Yao, S., et al. (2013). Activation of AMPK by bitter melon triterpenoids involves CaMKKbeta. *PLoS ONE*, 8(4), e62309. https://doi.org/10.1371/journal.pone.0062309.
- Kaiser, S., Carvalho, A. R., Pittol, V., Dietrich, F., Manica, F., Machado, M. M., et al. (2016). Genotoxicity and cytotoxicity of oxindole alkaloids from *Uncaria tomentosa* (cat's claw): Chemotype relevance. *Journal of Ethnopharmacology*, 189, 90–98. https://doi.org/10.1016/j.jep.2016.05.026.
- Keller, A. C., Ma, J., Kavalier, A., He, K., Brillantes, A. M., & Kennelly, E. J. (2011). Saponins from the traditional medicinal plant *Momordica charantia* stimulate insulin secretion in vitro. *Phytomedicine*, 19(1), 32–37. https://doi.org/10.1016/j.phyme d.2011.06.019.
- Leung, L., Birtwhistle, R., Kotecha, J., Hannah, S., & Cuthbertson, S. (2009). Anti-diabetic and hypoglycaemic effects of *Momordica charantia* (bitter melon): a mini review. *British Journal of Nutrition*, 102(12), 1703–1708. https://doi.org/10.1017/S0007 114509992054.
- Luyen, N. T., Dang, N. H., Binh, P. T. X., Hai, N. T., & Dat, N. T. (2018). Hypoglycemic property of triterpenoid saponin PFS isolated from *Polyscias fruticosa* leaves. *Anais da Academia Brasileira de Ciências*, 90, 2881–2886. https://doi. org/10.1590/0001-3765201820170945.
- Mohimani, H., Gurevich, A., Mikheenko, A., Garg, N., Nothias, L. F., Ninomiya, A., et al. (2017). Dereplication of peptidic natural products through database search of mass spectra. *Nature Chemical Biology*, 13(1), 30–37. https://doi.org/10.1038/nchem bio.2219.
- Murakami, T., Emoto, A., Matsuda, H., & Yoshikawa, M. (2001). Medicinal foodstuffs. XXI. Structures of new cucurbitane-type triterpene glycosides, goyaglycosides-a, -b, -c, -d, -e, -f, -g, and -h, and new oleanane-type triterpene saponins, goyasaponins I, II, and III, from the fresh fruit of Japanese Momordica charantia L. Chemical & Pharmaceutical Bulletin, 49(1), 54–63.
- Nakamura, S., Murakami, T., Nakamura, J., Kobayashi, H., Matsuda, H., & Yoshikawa, M. (2006). Structures of new cucurbitanetype triterpenes and glycosides, karavilagenins and karavilosides, from the dried fruit of *Momordica charantia* L. in Sri Lanka. *Chemical & Pharmaceutical Bulletin (Tokyo)*, 54(11), 1545–1550.
- Okabe, H., Miyahara, Y., Yamauchi, T., Miyahara, K., & Kawasaki, T. (1980). Studies on the constituents of *Momordica charantia* L. I. Isolation and characterization of momordicosides A and B, glycosides of a pentahydroxy-cucurbitane triterpene. *Chemical & Pharmaceutical Bulletin, 28*(9), 2753–2762. https://doi. org/10.1248/cpb.28.2753.
- Pan, Y., Zhang, J., Shen, T., Zhao, Y. L., Zuo, Z. T., Wang, Y. Z., et al. (2016). Investigation of chemical diversity in different parts and origins of ethnomedicine *Gentiana rigescens* Franch using targeted metabolite profiling and multivariate statistical analysis. *Biomedical Chromatography*, 30(2), 232–240. https://doi. org/10.1002/bmc.3540.
- Pluskal, T., Castillo, S., Villar-Briones, A., & Oresic, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, 11, 395. https://doi. org/10.1186/1471-2105-11-395.

- Popovich, D. G., Lee, Y., Li, L., & Zhang, W. (2011). Momordica charantia seed extract reduces pre-adipocyte viability, affects lactate dehydrogenase release, and lipid accumulation in 3T3-L1 cells. Journal of Medicinal Food, 14, 201–208. https://doi.org/10.1089/ jmf.2010.1150.
- Queiroz, E. F., Alfattani, A., Afzan, A., Marcourt, L., Guillarme, D., & Wolfender, J. L. (2019). Utility of dry load injection for an efficient natural products isolation at the semi-preparative chromatographic scale. *Journal of Chromatography A*, 1598, 85–91. https://doi.org/10.1016/j.chroma.2019.03.042.
- Sathishsekar, D., & Subramanian, S. (2005). Antioxidant properties of *Momordica Charantia* (bitter gourd) seeds on Streptozotocin induced diabetic rats. *Asia Pacific Journal of Clinical Nutrition*, 14(2), 153–158.
- Shibuya, H., Kawashima, K., Sakagami, M., Kawanishi, H., Shimomura, M., Ohashi, K., et al. (1990). Sphingolipids and glycerolipids. I. Chemical structures and ionophoretic activities of soya-cerebrosides I and II from soybean. *Chemical & Pharmaceutical Bulletin*, 38(11), 2933–2938.
- Singh, A., Singh, S. P., & Bamezai, R. (1998a). Momordica charantia (Bitter Gourd) peel, pulp, seed and whole fruit extract inhibits mouse skin papillomagenesis. Toxicology Letters, 94(1), 37–46.
- Singh, A., Singh, S. P., & Bamezai, R. (1998b). Postnatal efficacy of Momordica charantia peel, pulp, seed and whole fruit extract in the detoxication pathway of suckling neonates and lactating mice. Cancer Letters, 122(1–2), 121–126.
- Soro, L. C., Munier, S., Pelissier, Y., Grosmaire, L., Yada, R., Kitts, D., et al. (2016). Influence of geography, seasons and pedology on chemical composition and anti-inflammatory activities of essential oils from *Lippia multiflora* Mold leaves. *Journal of Ethnopharmacology*, 194, 587–594. https://doi.org/10.1016/j.jep.2016.10.047.
- Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) metabolomics standards initiative (MSI). *Metabolomics*, *3*, 211–221. https://doi.org/10.1007/s11306-007-0082-2.
- Tan, M. J., Ye, J. M., Turner, N., Hohnen-Behrens, C., Ke, C. Q., Tang, C. P., et al. (2008). Antidiabetic activities of triterpenoids isolated from bitter melon associated with activation of the AMPK pathway. *Chemistry & Biology*, 15, 263–273. https://doi.org/10.1016/j. chembiol.2008.01.013.
- Tomita, K., Tamiya, G., Ando, S., Kitamura, N., Koizumi, H., Kato, S., et al. (2005). AICAR, an AMPK activator, has protective effects on alcohol-induced fatty liver in rats. *Alcoholism, Clinical and Experimental Research*, 29(12 Suppl), 240S–245S.

Affiliations

Shuaizhen Zhou^{1,3} · Pierre-Marie Allard¹ · Christian Wolfrum² · Changqiang Ke³ · Chunping Tang³ · Yang Ye³ · Jean-Luc Wolfender¹

- ¹ School of Pharmaceutical Sciences, EPGL, University of Geneva, University of Lausanne, CMU – Rue Michel-Servet 1, CH-1206 Geneva 4, Switzerland
- ² Swiss Federal Institute of Technology, ETH Zürich, Institute of Food Nutrition and Health, Schorenstr. 16, 8603 Schwerzenbach, Switzerland

- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology*, 34(8), 828–837. https://doi. org/10.1038/nbt.3597.
- Wolfender, J. L., Litaudon, M., Touboul, D., & Queiroz, E. F. (2019a). Innovative omics-based approaches for prioritisation and targeted isolation of natural products—New strategies for drug discovery. *Natural Products Reports*, 36(6), 855–868. https://doi. org/10.1039/c9np00004f.
- Wolfender, J. L., Nuzillard, J. M., van der Hooft, J. J. J., Renault, J. H., & Bertrand, S. (2019b). Accelerating metabolite identification in natural product research: Toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry and NMR profiling, in silico databases, and chemometrics. *Analytical Chemistry*, 91(1), 704–742. https://doi.org/10.1021/acs.analc hem.8b05112.
- Xu, J., Cao, K., Li, Y., Zou, X., Chen, C., Szeto, I. M., et al. (2014). Bitter gourd inhibits the development of obesity-associated fatty liver in C57BL/6 mice fed a high-fat diet. *Journal of Nutrition*, 144(4), 475–483. https://doi.org/10.3945/jn.113.187450.
- Yang, S. Y., Kim, H. K., Lefeber, A. W., Erkelens, C., Angelova, N., Choi, Y. H., et al. (2006). Application of two-dimensional nuclear magnetic resonance spectroscopy to quality control of ginseng commercial products. *Planta Medica*, 72(4), 364–369. https://doi. org/10.1055/s-2005-916240.
- Ye, M., Liu, S. H., Jiang, Z., Lee, Y., Tilton, R., & Cheng, Y. C. (2007). Liquid chromatography/mass spectrometry analysis of PHY906, a Chinese medicine formulation for cancer therapy. *Rapid Communications in Mass Spectrometry*, 21(22), 3593–3607. https:// doi.org/10.1002/rcm.2832.
- Zeng, X. Y., Wang, Y. P., Cantley, J., Iseli, T. J., Molero, J. C., Hegarty, B. D., et al. (2012). Oleanolic acid reduces hyperglycemia beyond treatment period with Akt/FoxO1-induced suppression of hepatic gluconeogenesis in type-2 diabetic mice. *PLoS ONE*, 7(7), e42115. https://doi.org/10.1371/journal.pone.0042115.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

³ State Key Laboratory of Drug Research, & SIMM/CUHK Joint Research Laboratory of Promoting of Traditional Chinese Medicines, Natural Products Chemistry Department, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zu-Chong-Zhi Road, Zhangjiang High-Tech Park, Shanghai 201203, People's Republic of China