Journal of Ginseng Research 48 (2024) 149–162

Contents lists available at ScienceDirect

Journal of Ginseng Research

journal homepage: www.sciencedirect.com/journal/journal-of-ginseng-research

Review Article

Mass spectrometry-based ginsenoside profiling: Recent applications, limitations, and perspectives

Hyun Woo Kim [a,b], Dae Hyun Kim [b], Byeol Ryu [b], You Jin Chung [b], Kyungha Lee [c], Young Chang Kim [d], Jung Woo Lee [e], Dong Hwi Kim [e], Woojong Jang [f], Woohyeon Cho [g], Hyeonah Shim [g], Sang Hyun Sung [b,1], Tae-Jin Yang [g], Kyo Bin Kang [b,c,*]

[a] College of Pharmacy and Integrated Research Institute for Drug Development, Dongguk University, Seoul, Republic of Korea
[b] Research Institute of Pharmaceutical Sciences, College of Pharmacy, Seoul National University, Seoul, Republic of Korea
[c] College of Pharmacy and Drug Information Research Institute, Sookmyung Women's University, Seoul, Republic of Korea
[d] Future Agriculture Strategy Team, Research Policy Bureau, Rural Development Administration, Jeonju, Republic of Korea
[e] Ginseng Division, Department of Herbal Crop Research, National Institute of Horticultural & Herbal Science, Rural Development Administration, Eumseong, Republic of Korea
[f] Herbal Medicine Resources Research Center, Korea Institute of Oriental Medicine, Naju, Republic of Korea
[g] Department of Agriculture, Forestry and Bioresources, Plant Genomics and Breeding Institute, College of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Chemical profiling
Genetic variation
Ginseng
Ginsenoside
Mass spectrometry

ABSTRACT

Ginseng, the roots of *Panax* species, is an important medicinal herb used as a tonic. As ginsenosides are key bioactive components of ginseng, holistic chemical profiling of them has provided many insights into understanding ginseng. Mass spectrometry has been a major methodology for profiling, which has been applied to realize numerous goals in ginseng research, such as the discrimination of different species, geographical origins, and ages, and the monitoring of processing and biotransformation. This review summarizes the various applications of ginsenoside profiling in ginseng research over the last three decades that have contributed to expanding our understanding of ginseng. However, we also note that most of the studies overlooked a crucial factor that influences the levels of ginsenosides: genetic variation. To highlight the effects of genetic variation on the chemical contents, we present our results of untargeted and targeted ginsenoside profiling of different genotypes cultivated under identical conditions, in addition to data regarding genome-level genetic diversity. Additionally, we analyze the other limitations of previous studies, such as imperfect variable control, deficient metadata, and lack of additional effort to validate causation. We conclude that the values of ginsenoside profiling studies can be enhanced by overcoming such limitations, as well as by integrating with other -omics techniques.

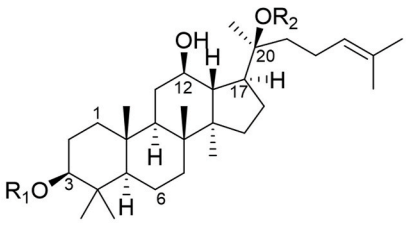
1. Introduction

Panax species are reputed medicinal plants with roots known as ginseng, which is not only a critical ingredient in East Asian traditional medicine [1], but also a widely used dietary supplement in numerous countries, including the US and European countries. Although defining the bioactivity of ginseng is challenging, it is commonly considered a tonic. A systematic review in 2015 reported that 29 out of 44 randomized controlled clinical trials showed the positive efficacies of ginseng in terms of cardiovascular, sexual, and psychomotor functions, glucose metabolism, antioxidation, and anti-fatigue effects [2].

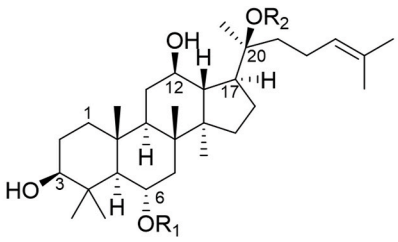
Ginsenosides, which are the major specialized metabolites of *Panax* species, are considered major contributors to the bioactivities of ginseng. Ginsenosides are triterpenoidal saponins that are further categorized based on their aglycone structures. Dammarane-type ginsenosides are major members, with tetracyclic structures, and they are subdivided into protopanaxadiol (PPD), -triol (PPT), and ocotillol types. Pentacyclic oleanane (OA) type ginsenosides are relatively minor components, but they are also relevant to the diverse bioactivities of ginseng. Natural ginsenosides are glycosides with aglycone structures and 2–5

* Corresponding author. College of Pharmacy and Drug Information Research Institute, Sookmyung Women's University, Cheongpa-ro 47 gil 100, Yongsan-gu, Seoul, 04310, Republic of Korea.
E-mail address: kbkang@sookmyung.ac.kr (K.B. Kang).
[1] deceased on July 24, 2018.

https://doi.org/10.1016/j.jgr.2024.01.004
Received 23 November 2023; Received in revised form 9 January 2024; Accepted 14 January 2024
Available online 19 January 2024



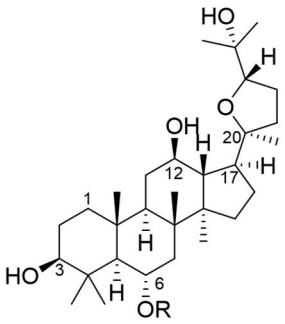
protopanaxadiol-type (PPD-type)



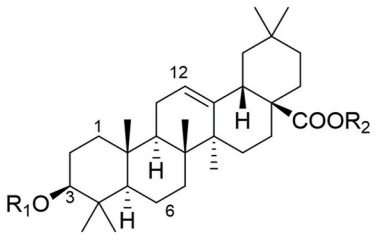
protopanaxatriol-type (PPT-type)

	R_1	R_2
ginsenoside Rb1	Glc^{1}-^{2}Glc-	Glc^{1}-^{6}Glc-
ginsenoside Rb2	Glc^{1}-^{2}Glc-	$Ara(p)^{1}$-^{6}Glc-
ginsenoside Rc	Glc^{1}-^{2}Glc-	$Ara(f)^{1}$-^{6}Glc-

	R_1	R_2
ginsenoside Re	Rha^{1}-^{2}Glc-	Glc-
ginsenoside Rf	Glc^{1}-^{2}Glc-	H
ginsenoside Rg1	Glc-	Glc-



ocotillol-type



oleanane-type (OA-type)

	R
pseudoginsenoside F_{11}	Rha^{1}-^{2}Glc-

	R_1	R_2
chikusetsusaponin IVa	GlcA-	Glc-
ginsenoside Ro	Glc^{1}-$^{2}GlcA$-	Glc-

Fig. 1. Structures of the four major types of ginsenosides. The major sites of glycosylation of each type of aglycone are represented as R, and representative ginsenosides for each type are indicated with their sugar moieties.

saccharide moieties, and Fig. 1 shows the structures of the major ginsenoside aglycones. In addition to sugars, malonyl groups are attached to the glycosyl chains of ginsenosides in fresh ginseng [3,4]. Further details regarding the structures of ginsenosides and their distribution in the genus *Panax* may be found in previous reviews [5–10].

As ginsenosides are the major bioactive constituents of ginseng, the chemical profiling of these compounds is a crucial methodology in ginseng research. Since the first application of liquid chromatography-electrospray ionization mass spectrometry (LC-ESI-MS) in analyzing ginsenosides by van Breemen et al., in 1995 [11], LC-MS has been the most used analytical method in ginsenoside profiling. MS has been the optimal choice in ginsenoside profiling because the detection of ginsenosides using UV detectors is challenging owing to the lack of chromophores in their structures. Remarkably, even the earliest studies were based on fragmentation spectra acquired via tandem mass spectrometry (MS/MS) in characterizing isomeric compounds [11–13]. The early application of MS/MS in structural annotation in ginsenoside profiling may be due to the ease of data interpretation. Owing to the glycosidic structures of ginsenosides, MS/MS fragmentation spectra exhibit sequential neutral losses of 162 or 132 Da, corresponding to the losses of hexoses or pentoses. The aglycone structures were deduced based on their fragment ions at *m/z* 459 (PPD), 475 (PPT), 491 (ocotillol), and 455 (OA) observed in the negative ion mode spectra. In 2012, W. Yang et al. used 2D orthogonal column chromatography to putatively annotate 623 ginsenosides from the roots of *P. ginseng* Meyer, *P. quinquefolius* L., and *P. notoginseng* Chen, highlighting the applicability of LC-MS/MS in ginsenoside profiling [14].

In this review, we summarize the previous applications of MS-based ginsenoside profiling. We categorized the previous studies based on their application goals: discrimination of different species, geographical origins, and ages, and the monitoring of processing and biotransformation. We focused mainly on an overview of the biological insights into ginseng provided by the studies, but significant technical advances are also briefly analyzed. This review categorizes ginsenoside profiling into two subcategories: targeted and untargeted. The definitions of these two subcategories follow those used for targeted and untargeted metabolomics in metabolomics communities [15]. The targeted approach refers to a method that uses a set of standard compounds and yields data regarding the absolute quantities of the target compounds. The untargeted method provides a global view of the entire metabolome, and the molecules of interest (ginsenosides in most cases described here) are putatively annotated based on the fragmentation spectra. In addition to summarizing the recent studies, we briefly introduce our data obtained via untargeted and targeted analyses of the ginsenoside contents of

different accessions of *P. ginseng*, which emphasize the effects of genetic differences on the chemotypes. Based on these findings, we analyze the current limitations and future perspectives of ginsenoside profiling.

2. Ginsenoside profiling of different *Panax* species

The World Flora Online Plant List includes 23 species in the genus *Panax* [16], only four of which, i.e., *P. ginseng* (Korean ginseng), *P. japonicus* (Japanese ginseng), *P. notoginseng* (Chinese ginseng), and *P. quinquefolius* (American ginseng), are commercially circulated. Accurate discrimination among *Panax* spp. is critical not only in preventing adulteration but also in ensuring the efficacies and safety of ginseng products and ginseng-based formulations in the pharmaceutical industry. MS/MS-based analytical techniques have been used to determine the chemical compositions of the *Panax* species. Advances in MS technology have enabled highly sensitive and specific analyses of complex mixtures of metabolites, thereby facilitating the identification of characteristic markers that may be used to differentiate between *P. ginseng, P. quinquefolius,* and *P. notoginseng*.

Early studies quantified a few specific ginsenosides via multiple reaction monitoring and used them as markers in discriminating between *P. ginseng* and *P. quinquefolius*. Wang et al. analyzed the contents of ginsenosides Rb1, Rb2, Rc, Rd, Re, Rf, and Rg1 in *P. ginseng* and *P. quinquefolius*. They suggested that *P. ginseng* contains higher amounts of ginsenosides Rf and Rg1, whereas *P. quinquefolius* contains higher amounts of the other ginsenosides [12]. This result was reproduced in a quantitative study of the commercial products by Ji et al. [17]. Chan et al. quantified a different set of ginsenosides, including Rb1, Rb2, Rc, Rd, Re, Rf, Rg1, Ro, and 24(*R*)-pseudoginsenoside F_{11}, to discriminate between the crude extracts of *P. ginseng* and *P. quinquefolius* and their commercial products. They reported the exclusive presence of ginsenoside Rf in *P. ginseng*, whereas 24(*R*)-pseudoginsenoside F_{11} was exclusively detected in *P. quinquefolius* [18]. Li et al. analyzed these two ginsenosides, and ginsenoside Rf was absent in *P. quinquefolius*, whereas a trace amount of 24(*R*)-pseudoginsenoside F_{11} (<0.0001 % w/w) was present in *P. ginseng* [19]. Based on these results, the authors suggested the presence and ratio of ginsenoside Rf and 24(*R*)-pseudoginsenoside F_{11} as discrimination and authentication markers for *P. ginseng* and *P. quinquefolius*.

Advances in chromatography and MS improved the sensitivity and resolution of ginsenoside profiling, enabling the detection of a larger number of compounds in a single sample run. Park et al. used ultra-high performance liquid chromatography-quadrupole/time-of-flight MS (UHPLC-Q/TOF-MS) for targeted and untargeted analyses to discriminate between processed *P. ginseng* (Korean Red Ginseng) and *P. quinquefolius* (American red ginseng) [20]. In addition to Rf and F_{11}, ginsenosides Ra1, F2, and 20-*gluco*-ginsenoside Rf were proposed as potential chemical markers of the processed ginseng samples. Yuk et al. performed untargeted analyses of the roots of three species (*P. ginseng, P. quinquefolius,* and *P. notoginseng*) and their commercial products using UHPLC-Q/TOF-MS [21]. Ginsenosides Ra1, Ra2, Rb2, and Rf were suggested as respective chemical markers for *P. ginseng*, while ginsenosides Rd and Re, pseudoginsenoside F_{11}, and gypenoside XVII were for *P. quinquefolius*, and notoginsenosides R1, R4, and Fa were for *P. notoginseng*. Yang et al. putatively annotated 87 ginsenosides from the roots of *P. ginseng, P. quinqeufolius, and P. notoginseng*, and then selected 17 chemical markers based on the results of the untargeted analyses of 85 root samples [22]. The group led by Wu and Guo introduced analytical methods that were optimized to maximize the variety of ginsenosides observed in the untargeted analysis. In 2017, they introduced a method specific for determining malonyl-ginsenosides using successive losses of CO_2 (44 Da) and an entire malonyl group (86 Da) [23,24]. In 2020, they suggested a method of determining ginsenosides without carboxyl groups, using a neutral loss of 46 Da as a diagnostic signal representing formic acid adducts, followed by further MS^3 fragmentation analysis performed using selected product ions of sapogenins [25]. These methods were evaluated using a sample set comprising the roots of *P. ginseng*, *P. quinquefolius*, and *P. notoginseng*, resulting in more suggested chemical markers.

Direct infusion MS (DI-MS) has not been favored in ginsenoside profiling because of the presence of numerous isobaric compounds, but it may be useful in practical sample authentication owing to the scalability of the method. Kim et al. investigated the applicability of DI-MS/MS in species discrimination of *P. ginseng, P. notoginseng, P. quinquefolius,* and *P. vietnamensis* [26]. They suggested that the four target ions at *m/z* 783.5, 945.5, 1107.5, and 1149.2 could be used as chemical markers in DI-MS/MS-based fingerprinting for species discrimination.

3. Ginsenoside profiling in distinguishing geographical origins

As geographical origins may critically affect the quality of an agricultural product, the authentication of geographical origin is a crucial application of metabolomics studies with commercial plants [27]. Korean ginseng is protected by Geographical Indications in numerous countries, and thus, ginsenoside profiling has been widely conducted to distinguish geographical origins. Song et al. utilized LC-MS-based untargeted analysis and orthogonal partial least squares-discriminant analysis (OPLS-DA) to distinguish and predict the geographical origins of the roots of *P. ginseng* cultivated in six different regions of South Korea [28]. They applied the same method in distinguishing *P. ginseng* cultivated in Korea and China; then ginsenoside Rf and an isomer of notoginsenoside R3 were suggested as markers for Korean products and ginsenoside Ro and chikusetsusaponin IVa for Chinese products [29]. Similarly, *P. ginseng* samples cultivated in three regions of China were distinguished and their geographical origins were predicted by Zhang et al. [30], where a support vector machine (SVM) was used to interpret the data. Chen et al. performed targeted analyses of 21 ginsenosides in *P. ginseng* cultivated in New Zealand, China, and South Korea. Samples from New Zealand displayed higher contents of ginsenosides Re, Rf, and Rg1 than those of samples from China and South Korea, and the volcanic pumice soil of New Zealand was suggested as the cause of the higher ginsenoside contents [31]. Yoon et al. applied a multiplatform-based metabolomics approach in distinguishing the geographical origins of *P. ginseng* from Korea, China, and Japan [32]. LC-MS was used in the untargeted and targeted analyses of ginsenosides in this study, whereas NMR spectroscopy and GC-MS were used in analyzing primary metabolite contents.

Distinguishing the geographical origin of *P. quinquefolius* has also been of interest in numerous studies because of its market size as a commercial product. Shuai et al. performed HPLC-based targeted analyses of ginsenosides, along with headspace-GC-MS (HS-GC-MS)-based untargeted analyses of volatile compounds to distinguish between *P. quinquefolius* cultivated in the US, Canada, and two provinces of China [33]. The results of principal component analysis (PCA) suggested 25 volatile metabolites and 8 ginsenosides as chemical markers. Additionally, linear discriminant analysis- and random forest-based discriminant models were used to predict the geographical origin of *P. quinquefolius* with a high accuracy based on the contents of the markers. An LC-MS-based untargeted analysis of *P. quinquefolius* to distinguish its geographical origin was performed by Pang et al. using samples from the US, Canada, and four provinces of China [34]. The analytic results were used to tentatively annotate 382 metabolites, including ginsenosides, amino acids, organic acids, and lipids; and among them, 20 potential chemical markers were suggested.

4. Ginsenoside profiling of ginseng cultivated for different years

The contents of ginsenosides in *P. ginseng* are significantly affected by age. Generally, the roots of *P. ginseng* are harvested after 4, 5, or 6 y of cultivation, and the six-year-old ginseng is considered a high-quality product. Ginsenoside profiling was conducted to validate the relationship between cultivation age and quality, and it has been suggested as a

method of sample authentication to avoid potential mislabeling or deceptive marketing of ginseng products. The targeted analyses of nine ginsenosides (Rb1, Rb2, Rb3, Rc, Rd, Re, Rf, Rg1, and Rg2) by Wang et al. may be the first study regarding ginsenoside content according to cultivation age [35]. They analyzed *P. ginseng* grown for 2–6 y, and the total ginsenoside content reached the maximum level after 4 y. This is partially consistent with the results of the untargeted analysis of ginseng grown for 1–6 y by Kim et al. [36]. This study mainly aimed to develop a classification model for the cultivation age using random forest and partial least squares-discriminant analyses, and prediction analysis of microarray. However, PCA revealed that the ginsenoside profiles of the samples grown for 4–6 y were similar to each other compared to those of ginseng grown for 1–3 y. In a follow-up study, they showed that minimal amounts of fine roots could be used to predict the cultivation age with a similar accuracy [37]. Seven MS ions, representing ginsenosides Rb1, Rd, Re, and Rg2, malonyl-ginsenoside Rb1, and two unknown compounds were suggested as chemical markers for use in discrimination. Similarly, Huang et al. identified *P. ginseng* cultivated for 2–6 y using the results of untargeted LC-MS data and OPLS-DA and suggested >50 compound markers [38].

Most of these studies analyzed homogenized roots, and only a few studies have investigated the localization and spatial distributions of ginsenosides and applied these data in distinguishing cultivation age. Bai et al. traced the localization of 31 ginsenosides in *P. ginseng* cultivated for 2, 4, or 6 y using matrix-assisted laser-desorption/ionization (MALDI)-TOF-MS imaging, and suggested that cork tissue exhibits the most significant difference according to age [39]. In the PCA of the spectra of the cork tissues, ginsenosides Ra8/Ra9 (meaning that the MS ion was arbitrarily annotated as ginsenosides Ra8 or Ra9, which display isobaric molecular masses), Ro, Rd/Re, and Rb1 and malonyl-ginsenoside Rb2/Rc displayed the largest variations among the groups, whereas the spectra of the whole tissues did not exhibit significant differences. Similarly, Lee et al. performed MALDI-MS analyses to visualize the spatial distributions of 14 ginsenosides in ginseng roots and determined the localization patterns of ginsenosides Rh1, Rg2, and Rc (or Rb2 and Rb3 with the same m/z). They also suggested that the content of ginsenoside Rb1 is affected by the ages of xylem, cortex, and periderm [40]. Yang et al. utilized LC-MS and desorption electrospray ionization mass spectrometry imaging to develop a rapid, solvent-saving method of analyzing ginseng root slides and suggested 18 markers for growth age [41].

5. Ginsenoside profiling of wild-simulated ginseng

Ginseng is primarily obtained following cultivation, and wild ginseng is traded at extremely high prices owing to its rarity. The growth rate of wild ginseng is comparatively slower than that of cultivated ginseng, and it may reach an age of hundreds of years. Wild-simulated ginseng (also known as mountain-cultivated ginseng) is an alternative to wild ginseng. In contrast to cultivated ginseng, wild-simulated ginseng is harvested after 10–20 y of growth, and several studies compared the ginsenoside profiles of cultivated, wild-simulated, and wild ginseng to reveal their differences and authenticate them. Xu et al. suggested that malonyl-ginsenosides are more abundant in cultivated ginseng. Conversely, based on untargeted analyses, wild-simulated ginseng exhibits higher contents of minor ginsenosides, such as ginsenosides Rs6/Rs7, Ra2, Ra3/isomer, and Ra7, notoginsenoside Fe, quinquenoside R1, and gypenoside XVII [42]. Zhu et al. performed untargeted analyses of cultivated (four-, five-, and six-year-old) and wild-simulated ginseng (twelve- and twenty-year-old) collected from China [43]. Multivariate analysis suggested that cultivated and wild-simulated ginseng exhibited significant differences in their metabolome. In this study, fatty acids, such as α-linolenic acid, 9-octadecenoic acid, linoleic acid, and panaxydol, not ginsenosides, were suggested as marker compounds for use in differentiating wild-simulated and cultivated ginseng. Guo et al. analyzed the ginsenoside contents of wild-simulated ginseng grown for 5, 10, 15, 20, and 25 y [44]. They suggested that ginsenosides generally accumulated in wild-simulated ginseng for 15 y, and the chemical profiles did not change significantly after 15 y of growth. Qu et al. attempted to maximize the observation window in their untargeted analyses of wild-simulated and cultivated ginseng using offline 2D LC separation consisting of hydrophilic interaction and reversed phase LC [45]. The putatively annotated 559 ginsenosides via integration of the positive and negative ion modes. They also quantified 14 ginsenosides via targeted analyses and integrated the results with the relative intensities of 199 putative ginsenosides before comparing the quasi-quantitative data acquired using 57 batches of wild-simulated and cultivated ginseng. Wild-simulated ginseng grown for an extended duration displayed higher malonyl-ginsenoside and total ginsenoside contents.

6. Ginsenoside profiling of steamed and microbially fermented ginseng

Korean Red Ginseng is a representative example of processed ginseng. Red ginseng is prepared by steaming fresh ginseng. The chemical profile of red ginseng differs considerably from that of fresh or dried ginseng (also denoted as white ginseng), because steaming causes the denaturation of ginsenosides. Generally, heat and the organic acids in ginseng induce the hydrolyses of the saccharide chains, and thus, ginsenosides in red ginseng generally exhibit shorter glycosyl moieties than those of the ginsenosides in white ginseng. Red ginseng may display a superior bioavailability compared to that of white ginseng, and the enhanced lipophilicity due to the shortening of the saccharide chains may contribute to the enhanced absorption. Studies performed in the late 1990s support this hypothesis, suggesting that most ginsenosides are poorly absorbed from the gut, whereas compound K, which is a metabolite formed via intestinal microbial biotransformation, is absorbed well [46–49]. Steaming also modifies the alkyl side chain at C-17 of the dammarane-type scaffold, and further details regarding Korean Red Ginseng and chemical transformation may be found in a review published in 2015 [50]. A recent review published in 2023 describes the chemical and biological details of black ginseng, which is another processed ginseng product manufactured via nine-time steaming [51].

Ginsenoside profiling has contributed considerably to current knowledge regarding the chemistry of red ginseng. An untargeted analysis by Zhang et al. provided an overview of the chemical differences between white and red ginseng [52]. In this study, ginsenosides Rg3 and 20(*R*)-Rh1 were suggested as the characteristic components of red ginseng, while malonyl-ginsenosides Rb1 and Rg1 were the characteristic components of white ginseng. Xie et al. performed targeted analyses of 12 ginsenosides in commercial white and red ginseng samples, in addition to untargeted analyses. The contents of ginsenosides Rb1, Rb2, Rc, Rd, Re, Rf, Rg1, and Ro and malonyl-ginsenosides Rb1, Rb2, Rc, and Rd were significantly lower in red ginseng [53]. They also mimicked the steaming process in the laboratory to investigate the chemical conversion and suggested that hydrolysis, dehydration, isomerization, and decarboxylation at C-20 and hydrolysis at C-3 or C-6 were the major reactions during the production of red ginseng. Sun et al. used an untargeted method to monitor the chemical changes in *P. quinquefolius* during nine-time steaming to produce black American ginseng [54]. Among the 29 annotated ginsenosides, 18 were newly generated during steaming, mainly via hydrolysis, dehydration, decarboxylation and addition reactions. Decarboxylation (of the malonyl groups) and dehydration of ginsenosides were also observed in an untargeted study by Chu et al. [55].

Microbial biotransformation is an alternative method of modifying the chemical composition of ginseng. The major bioconversion during fermentation is the shortening of the sugar chains by microbial β-glucosidase, and thus, similar to those of steamed ginseng, fermented ginseng also displays higher contents of ginsenosides with shorter glycosyl chains. The microorganisms used in biotransformation and

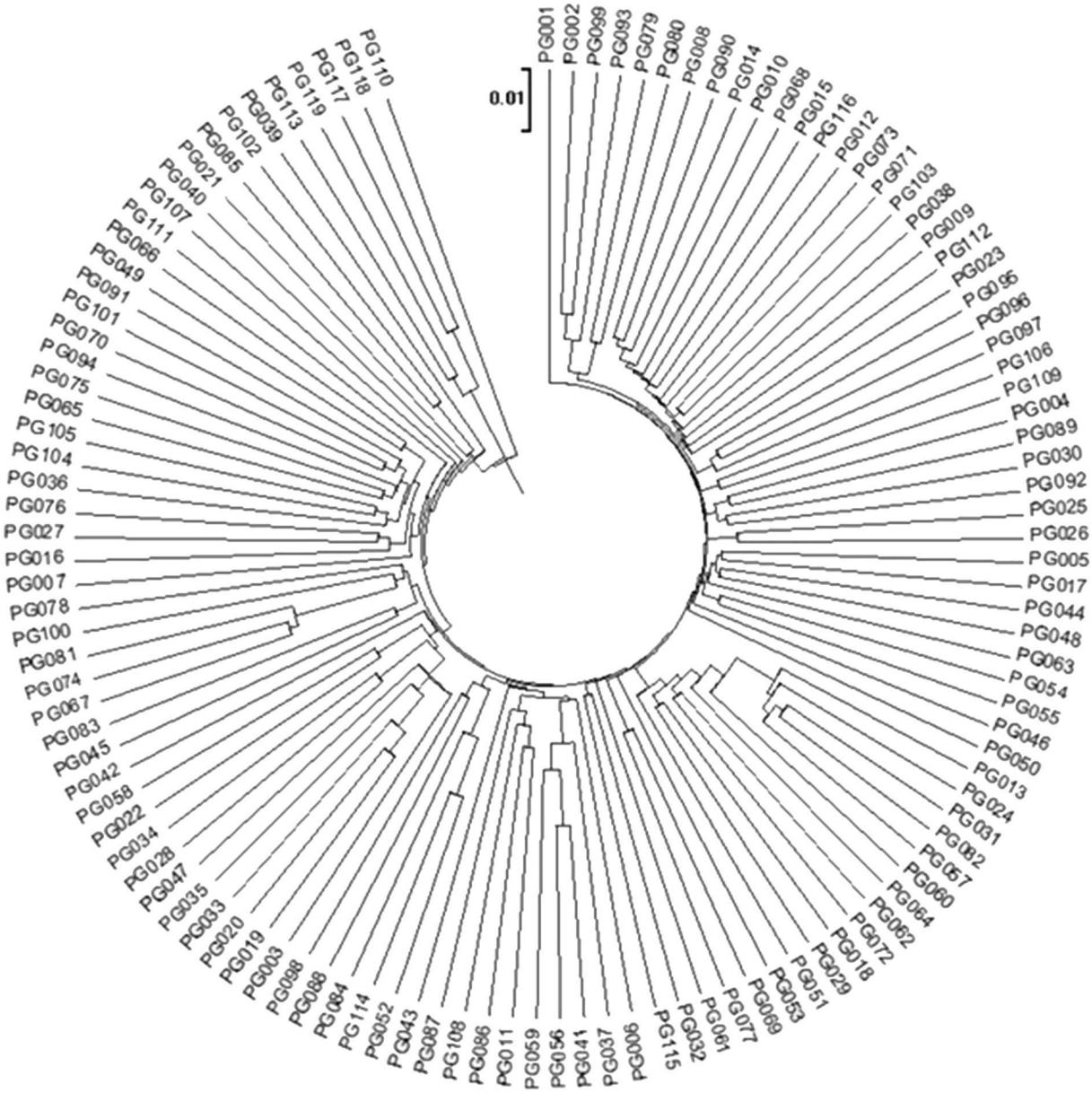


Fig. 2. Phylogenetic analysis of 119 ginseng accessions, utilizing 249,885 single nucleotide polymorphisms identified via genotyping-by-sequencing.

their catalytic activities are summarized in a review published in 2018 [56]. Bai et al. quantified 14 ginsenosides in ginseng fermented by *Lactobacillus plantarum*. They suggested that the removal of the glucosyl moieties at the C-20 positions of ginsenosides Rb1, Rd, and Re produced racemic mixtures of products, such as ginsenosides Rg3, Rk1, and Rg5 [57]. Xiao et al. performed an untargeted analysis of the fermentation process using *Paecilomyces hepiali*, and the glycosidic groups were generally hydrolyzed [58].

7. Ginsenoside profiling of the other components of the plant

Ginseng roots are the only commercially used components of the plant, and thus, the other components are wasted as byproducts. Recent chemical profiling of the other components of ginseng revealed the presence of numerous ginsenosides and suggested their potential for use as sources of functional foods or cosmetics, e.g., ginseng flower extract has been developed as a cosmetic agent for use in anti-aging and whitening of the skin [59]. Li et al. compared the ginsenoside profiles of the flowers of *P. ginseng, P. quinquefolius,* and *P. notoginseng* [60]. They also investigated the correlations between the compound contents and immune-enhancing activity and suggested that PPT-type ginsenosides and malonyl-ginsenosides in *P. notoginseng* are critical in the bioactivity. Jia et al. conducted untargeted profiling of the flower buds of *P. ginseng, P. quinquefolius*, and *P. notoginseng* using LC-ion mobility-MS. They suggested six distinguishing markers that were annotated as ginsenosides Rb3, Ra1, Ra1/Ra2, Rb1, and Ra3 and malonyl-ginsenoside Rc/Rb2/Rb3 [61]. Yoon et al. described the ginsenoside compositions of ginseng berries from seven *P. ginseng* cultivars and annotated 26 ginsenosides [62]. Chang et al. analyzed the leaves of wild-simulated ginseng cultivated for 6–18 y using LC-MS and suggested 39 compounds as chemical markers for use in age discrimination without destroying the roots [63].

8. Underestimated metadata - intraspecific genotype

In preparing this literature review, we realized that most of the cited studies missed a crucial factor affecting ginsenoside contents: genetic variation. We previously acquired targeted and untargeted ginsenoside

Table 1
Putative annotation of 71 ginsenosides in our untargeted LC-MS/MS analysis of the berries and roots of 60 *P. ginseng* accessions.

no.	identification	t_R (min)	Measured *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	Molecular formula (neutral)	Structure (PubChem CID)	Major fragment ions (*m/z*)
G1	ginsenoside Re3	3.45	961.5363	961.5378	$C_{48}H_{82}O_{19}$	10605931	799.48 $[M-Glc-H]^-$ 637.43 $[M-2Glc-H]^-$ 475.38 $[M-3Glc-H]^-$
G2	ginsenoside Re4 or isomer	4.50	931.5253	931.5272	$C_{47}H_{80}O_{18}$	162861378 or its isomer	799.48 $[M-Ara-H]^-$ 637.43 $[M-Ara-Glc-H]^-$ 475.38 $[M-Ara-2Glc-H]^-$
G3	floralginsenoside N or O	4.76	1077.5854	1077.5851	$C_{53}H_{90}O_{22}$	101423541 or 101423542	945.54 $[M-Ara-H]^-$ 931.53 $[M-Rha-H]^-$ 799.48 $[M-Rha-Ara-H]^-$ 783.49 $[M-Ara-Glc-H]^-$ 637.43 $[M-Rha-Ara-Glc-H]^-$ 475.38 $[M-Rha-Ara-2Glc-H]^-$
G4	20-*O*-glucosyl-ginsenoside Rf	4.93	961.5372	961.5378	$C_{48}H_{82}O_{19}$	24721561	799.48 $[M-Glc-H]^-$ 637.43 $[M-2Glc-H]^-$ 475.38 $[M-3Glc-H]^-$
G5	ginsenoside Re4 or isomer	5.39	931.5258	931.5272	$C_{47}H_{80}O_{18}$	162861378 or its isomer	637.43 $[M-Ara-Glc-H]^-$ 475.38 $[M-Ara-2Glc-H]^-$
G6	notoginsenoside R1	5.77	931.5236	931.5272	$C_{47}H_{80}O_{18}$	441934	799.48 $[M-Xyl-H]^-$ 637.43 $[M-Xyl-Glc-H]^-$ 475.38 $[M-Xyl-2Glc-H]^-$
G7	floralginsenoside N or O	5.94	1077.5854	1077.5851	$C_{53}H_{90}O_{22}$	101423541 or 101423542	945.54 $[M-Ara-H]^-$ 931.53 $[M-Rha-H]^-$ 799.48 $[M-Rha-Ara-H]^-$ 783.49 $[M-Ara-Glc-H]^-$ 637.43 $[M-Rha-Ara-Glc-H]^-$ 475.38 $[M-Rha-Ara-2Glc-H]^-$
G8	ginsenoside Re1	6.36	961.5349	961.5378	$C_{48}H_{82}O_{19}$	122397102	799.48 $[M-Glc-H]^-$ 637.43 $[M-2Glc-H]^-$ 475.38 $[M-3Glc-H]^-$
G9	ginsenoside Re2	7.45	961.5356	961.5378	$C_{48}H_{82}O_{19}$	101717751	637.43 $[M-2Glc-H]^-$ 475.38 $[M-3Glc-H]^-$
G10	ginsenoside Rg1[a]	8.47	799.4848	799.4849	$C_{42}H_{72}O_{14}$	441923	637.43 $[M-Glc-H]^-$ 475.38 $[M-2Glc-H]^-$
G11	ginsenoside Re[a]	8.88	945.5432	945.5423	$C_{48}H_{82}O_{18}$	441921	783.49 $[M-Glc-H]^-$ 637.43 $[M-Glc-Rha-H]^-$ 475.38 $[M-2Glc-Rha-H]^-$
G12	6′-*O*-acetyl-ginsenoside Rg1 or isomer	11.69	841.4950	841.4955	$C_{44}H_{74}O_{15}$	N/A	799.48 $[M-acetyl(Ac)-H]^-$ 637.43 $[M-Ac-Glc-H]^-$ 475.38 $[M-Ac-2Glc-H]^-$
G13	malonyl-ginsenoside Re isomer	12.09	1031.5435	1031.5432	$C_{51}H_{84}O_{21}$	N/A	987.55 $[M-CO_2-H]^-$ 945.54 $[M-malonyl(m)-H]^-$ 637.43 $[M-(m\text{-}Glc)-Rha-H]^-$ 475.38 $[M-(m\text{-}Glc)-$ $Rha-Glc-H]^-$
G14	6‴-*O*-acetyl-ginsenoside Re or isomer	12.66	987.5521	987.5534	$C_{50}H_{84}O_{19}$	N/A	945.54 $[M-Ac-H]^-$ 799.48 $[M-Ac-Rha-H]^-$ 783.49 $[M-Ac-Glc-H]^-$ 637.43 $[M-Ac-Glc-Rha-H]^-$ 475.38 $[M-Ac-2Glc-Rha-H]^-$
G15	malonyl-ginsenoside Re	12.89	1031.5435	1031.5432	$C_{51}H_{84}O_{21}$	N/A	987.55 $[M-CO_2-H]^-$ 945.54 $[M-m-H]^-$ 799.48 $[M-m-Rha-H]^-$ 783.49 $[M-(m\text{-}Glc)-H]^-$ 637.43 $[M-(m\text{-}Glc)-Rha-H]^-$ 475.38 $[M-(m\text{-}Glc)-$ $Rha-Glc-H]^-$
G16	notoginsenoside A	12.96	1123.5919	1123.5906	$C_{54}H_{92}O_{24}$	6451129	637.43 $[M-3Glc-H]^-$ 475.38 $[M-4Glc-H]^-$
G17	unknown 1	13.68	883.5037	883.5061	$C_{46}H_{76}O_{16}$	N/A	637.43 475.38
G18	floralginsenoside P	13.80	1093.5785	1093.5800	$C_{53}H_{90}O_{23}$	101423543	961.54 $[M-Ara-H]^-$ 799.48 $[M-Ara-Glc-H]^-$ 637.43 $[M-Ara-2Glc-H]^-$ 475.38 $[M-Ara-3Glc-H]^-$
G19	floralginsenoside C	13.95	815.4809	815.4799	$C_{42}H_{72}O_{15}$	16655212	637.43 $[M-Glc-H]^-$ 475.38 $[M-2Glc-H]^-$
G20	notoginsenoside N	14.87	961.5374	961.5378	$C_{48}H_{82}O_{19}$	101717750	799.48 $[M-Glc-H]^-$ 637.43 $[M-2Glc-H]^-$ 475.38 $[M-3Glc-H]^-$
G21	unknown 2	15.35	929.5466	929.5479	$C_{48}H_{82}O_{17}$	N/A	783.49 637.43

(*continued on next page*)

Table 1 (*continued*)

no.	identification	t_R (min)	Measured *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	Molecular formula (neutral)	Structure (PubChem CID)	Major fragment ions (*m/z*)
G22	ginsenoside Rf[a]	15.87	799.4847	799.4849	$C_{42}H_{72}O_{14}$	441922	637.43 $[M-Glc-H]^-$ 475.38 $[M-2Glc-H]^-$
G23	ginsenoside Ra3	16.42	1239.6366	1239.6379	$C_{59}H_{100}O_{27}$	73157064	1107.60 $[M-Xyl-H]^-$ 1077.59 $[M-Glc-H]^-$ 945.54 $[M-Xyl-Glc-H]^-$
G24	ginsenoside F3	16.86	769.4733	769.4744	$C_{41}H_{70}O_{13}$	46887678	637.43 $[M-Ara-H]^-$ 475.38 $[M-Ara-Glc-H]^-$
G25	ginsenoside Ra0	17.25	1269.6455	1269.6485	$C_{60}H_{102}O_{28}$	102601548	1107.60 $[M-Glc-H]^-$ 945.54 $[M-2Glc-H]^-$
G26	ginsenoside F5	17.90	769.4740	769.4744	$C_{41}H_{70}O_{13}$	46887590	637.43 $[M-Glc-H]^-$ 475.38 $[M-Ara-Glc-H]^-$
G27	20(*S*)-ginsenoside Rg2[a]	18.29	783.4881	783.4900	$C_{42}H_{72}O_{13}$	12912322	637.43 $[M-Rha-H]^-$ 475.38 $[M-Rha-Glc-H]^-$
G28	ginsenoside Ra2	18.74	1209.6260	1209.6274	$C_{58}H_{98}O_{26}$	100941543	1077.59 $[M-Xyl-H]^-$ 1047.57 $[M-Glc-H]^-$ 945.54 $[M-Xyl-Ara-H]^-$ 783.49 $[M-Xyl-Ara-Glc-H]^-$
G29	notoginsenoside R2	18.81	769.4756	769.4744	$C_{41}H_{70}O_{13}$	21599925	475.38 $[M-Xyl-Glc-H]^-$
G30	ginsenoside Rb1[a]	19.21	1107.5928	1107.5957	$C_{54}H_{92}O_{23}$	9898279	945.54 $[M-Glc-H]^-$ 783.49 $[M-2Glc-H]^-$ 621.44 $[M-3Glc-H]^-$ 459.38 $[M-4Glc-H]^-$
G31	malonyl-ginsenoside Ra2	19.58	1295.6262	1295.6278	$C_{61}H_{100}O_{29}$	N/A	945.54 $[M-m-Xyl-Ara-H]^-$ 783.49 $[M-m-Xyl-Ara-Glc-H]^-$
G32	malonyl-ginsenoside Rb1	20.00	1193.5964	1193.5961	$C_{57}H_{94}O_{26}$	118987129	1149.61 $[M-CO_2-H]^-$ 1107.60 $[M-m-H]^-$ 1089.59 $[M-m-H_2O-H]^-$ 945.54 $[M-m-Glc-H]^-$ 783.49 $[M-m-2Glc-H]^-$
G33	ginsenoside Ro[a]	20.17	955.4896	955.4908	$C_{48}H_{76}O_{19}$	11815492	793.44 $[M-Glc-H]^-$ 631.39 $[M-2Glc-H]^-$ 613.38 $[M-2Glc-H_2O-H]^-$ 455.35 $[M-2Glc-GlcA-H]^-$
G34	ginsenoside Rc[a]	20.56	1077.5833	1077.5851	$C_{53}H_{90}O_{22}$	12855889	945.54 $[M-Ara-H]^-$ 783.49 $[M-Ara-Glc-H]^-$ 621.44 $[M-Ara-2Glc-H]^-$ 459.38 $[M-Ara-3Glc-H]^-$
G35	Ginsenoside Ra1	20.88	1209.6263	1209.6274	$C_{58}H_{98}O_{26}$	100941542	1077.60 $[M-Xyl-H]^-$ 1047.57 $[M-Glc-H]^-$ 945.54 $[M-Xyl-Ara-H]^-$ 915.53 $[M-Xyl-Glc-H]^-$ 783.49 $[M-Xyl-Ara-Glc-H]^-$
G36	malonyl-ginsenoside Rc	21.37	1163.5824	1163.5855	$C_{56}H_{92}O_{25}$	N/A	1119.60 $[M-CO_2-H]^-$ 1077.59 $[M-m-H]^-$ 945.54 $[M-m-Ara-H]^-$ 783.49 $[M-m-Ara-Glc-H]^-$ 621.44 $[M-m-Ara-2Glc-H]^-$
G37	malonyl-ginsenoside Rb1 isomer	21.70	1193.5979	1193.5961	$C_{57}H_{94}O_{26}$	an isomer of 118987129	1149.61 $[M-CO_2-H]^-$ 1107.60 $[M-m-H]^-$ 945.54 $[M-m-Glc-H]^-$ 783.49 $[M-m-2Glc-H]^-$
G38	ginsenoside Rb2[a]	22.18	1077.5826	1077.5851	$C_{53}H_{90}O_{22}$	6917976	945.54 $[M-Ara-H]^-$ 915.53 $[M-Glc-H]^-$ 783.49 $[M-Ara-Glc-H]^-$ 621.44 $[M-Ara-2Glc-H]^-$ 459.38 $[M-Ara-3Glc-H]^-$
G39	ginsenoside Rb3[a]	22.63	1077.5818	1077.5851	$C_{53}H_{90}O_{22}$	12912363	945.54 $[M-Xyl-H]^-$ 915.53 $[M-Glc-H]^-$ 783.49 $[M-Xyl-Glc-H]^-$ 621.44 $[M-Xyl-2Glc-H]^-$ 459.38 $[M-Xyl-3Glc-H]^-$
G40	malonyl-ginsenoside Rb2	22.85	1163.5830	1163.5855	$C_{56}H_{92}O_{25}$	N/A	1119.60 $[M-CO_2-H]^-$ 1077.59 $[M-m-H]^-$ 1059.57 $[M-m-H_2O-H]^-$ 945.54 $[M-m-Ara-H]^-$ 915.53 $[M-m-Glc-H]^-$ 783.49 $[M-m-Ara-Glc-H]^-$ 621.44 $[M-m-Ara-2Glc-H]^-$ 459.38 $[M-m-Ara-3Glc-H]^-$
G41	malonyl-ginsenoside Rb3	23.10	1163.5836	1163.5855	$C_{56}H_{92}O_{25}$	N/A	1119.60 $[M-CO_2-H]^-$ 1077.59 $[M-m-H]^-$ 1059.57 $[M-m-H_2O-H]^-$ 945.54 $[M-m-Xyl-H]^-$

(*continued on next page*)

Table 1 (*continued*)

no.	identification	t_R (min)	Measured *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	Molecular formula (neutral)	Structure (PubChem CID)	Major fragment ions (*m/z*)
							783.49 $[M-m-Xyl-Glc-H]^-$ 621.44 $[M-m-Xyl-2Glc-H]^-$ 459.38 $[M-m-Xyl-3Glc-H]^-$
G42	ginsenoside Rb2 or Rb3 isomer	23.32	1077.5829	1077.5851	$C_{53}H_{90}O_{22}$	an isomer of 6917976 or 12912363	945.54 $[M-pentose-H]^-$ 783.49 $[M-pentose-hexose-H]^-$ 621.44 $[M-pentose-2hexose-H]^-$
G43	quinquenoside R1	23.49	1149.6052	1149.6062	$C_{56}H_{94}O_{24}$	101679657	1107.60 $[M-acetyl(Ac)-H]^-$ 945.54 $[M-Ac-Glc-H]^-$ 783.49 $[M-Ac-2Glc-H]^-$
G44	malonyl-ginsenoside Rb2 or Rb3 isomer	23.84	1163.5834	1163.5855	$C_{56}H_{92}O_{25}$	N/A	1119.60 $[M-CO_2-H]^-$ 1077.59 $[M-m-H]^-$ 1059.57 $[M-m-H_2O-H]^-$ 945.54 $[M-m-pentose-H]^-$ 915.53 $[M-m-hexose-H]^-$ 783.49 $[M-m-pentose-hexose-H]^-$
G45	malonyl-ginsenoside Rb2 or Rb3 isomer	24.01	1163.5862	1163.5855	$C_{56}H_{92}O_{25}$	N/A	1119.60 $[M-CO_2-H]^-$ 1077.59 $[M-m-H]^-$ 1059.57 $[M-m-H_2O-H]^-$ 945.54 $[M-m-pentose-H]^-$ 915.53 $[M-m-hexose-H]^-$ 783.49 $[M-m-pentose-hexose-H]^-$
G46	ginsenoside Rd[a]	24.18	945.5409	945.5428	$C_{48}H_{82}O_{18}$	11679800	783.49 $[M-Glc-H]^-$ 621.44 $[M-2Glc-H]^-$ 459.38 $[M-3Glc-H]^-$
G47	malonyl-ginsenoside Rd	24.55	1031.5415	1031.5432	$C_{51}H_{84}O_{21}$	14162967	987.55 $[M-CO_2-H]^-$ 945.54 $[M-m-H]^-$ 927.53 $[M-m-H_2O-H]^-$ 783.49 $[M-m-Glc-H]^-$ 621.44 $[M-m-2Glc-H]^-$ 459.38 $[M-m-3Glc-H]^-$
G48	malonyl-ginsenoside Rd isomer	24.78	1031.5402	1031.5432	$C_{51}H_{84}O_{21}$	an isomer of 14162967	987.55 $[M-CO_2-H]^-$ 945.54 $[M-m-H]^-$ 927.53 $[M-m-H_2O-H]^-$ 783.49 $[M-m-hexose-H]^-$ 621.44 $[M-m-2hexose-H]^-$
G49	ginsenoside Rs1 or Rs2	24.98	1119.5929	1119.5957	$C_{55}H_{92}O_{23}$	85044013 or 162343294	1077.59 $[M-acetyl(Ac)-H]^-$ 945.54 $[M-Ac-Ara-H]^-$ 783.49 $[M-Ac-Ara-Glc-H]^-$ 621.44 $[M-Ac-Ara-2Glc-H]^-$
G50	dimalonyl-ginsenoside Rd	25.01	1117.5414	1117.5436	$C_{54}H_{86}O_{24}$	N/A	1073.55 $[M-CO_2-H]^-$ 1029.56 $[M-2CO_2-H]^-$ 987.55 $[M-m-CO_2-H]^-$ 945.54 $[M-2\ m-H]^-$ 783.49 $[M-2\ m-Glc-H]^-$ 621.44 $[M-2\ m-2Glc-H]^-$ 459.38 $[M-2\ m-3Glc-H]^-$
G51	notoginsenoside O	25.11	1047.5724	1047.5745	$C_{52}H_{88}O_{21}$	154497111	915.53 $[M-Xyl-H]^-$ 753.48 $[M-Xyl-Glc-H]^-$
G52	ginsenoside Rb2 or Rb3 isomer	25.28	1077.5837	1077.5851	$C_{53}H_{90}O_{22}$	an isomer of 6917976 or 12912363	945.54 $[M-pentose-H]^-$ 783.49 $[M-pentose-hexose-H]^-$ 621.44 $[M-pentose-2hexose-H]^-$
G53	malonyl-ginsenoside Rd isomer	25.29	1031.5422	1031.5432	$C_{51}H_{84}O_{21}$	an isomer of 14162967	987.55 $[M-CO_2-H]^-$ 945.54 $[M-m-H]^-$ 927.53 $[M-m-H_2O-H]^-$ 783.49 $[M-m-hexose-H]^-$ 621.44 $[M-m-2hexose-H]^-$
G54	gypenoside XVII	25.38	945.5420	945.5428	$C_{48}H_{82}O_{18}$	44584555	783.49 $[M-Glc-H]^-$ 621.44 $[M-2Glc-H]^-$ 459.38 $[M-3Glc-H]^-$
G55	acetyl-ginsenoside Rd	25.57	987.5525	987.5534	$C_{50}H_{84}O_{19}$	N/A	945.54 $[M-Ac-H]^-$ 783.49 $[M-Ac-Glc-H]^-$ 621.44 $[M-Ac-2Glc-H]^-$ 459.38 $[M-Ac-3Glc-H]^-$
G56	acetyl-malonyl-ginsenoside Rd isomer	25.76	1073.5502	1073.5538	$C_{53}H_{86}O_{22}$	N/A	1029.56 $[M-CO_2-H]^-$ 945.54 $[M-Ac-m-H]^-$ 783.49 $[M-Ac-m-Glc-H]^-$ 621.44 $[M-Ac-m-2Glc-H]^-$
G57	notoginsenoside O isomer	25.77	1047.5713	1047.5745	$C_{52}H_{88}O_{21}$	an isomer of 154497111	915.53 $[M-pentose-H]^-$ 753.48 $[M-pentose-hexose-H]^-$

(*continued on next page*)

Table 1 (*continued*)

no.	identification	t_R (min)	Measured *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	Molecular formula (neutral)	Structure (PubChem CID)	Major fragment ions (*m/z*)
G58	acetyl-malonyl-ginsenoside Rd isomer	26.12	1073.5552	1073.5538	$C_{53}H_{86}O_{22}$	N/A	1029.56 $[M-CO_2-H]^-$ 945.54 $[M-Ac-m-H]^-$ 783.49 $[M-Ac-m-Glc-H]^-$ 621.44 $[M-Ac-m-2Glc-H]^-$
G59	ginsenoside compound-Mc1	26.13	915.5294	915.5323	$C_{47}H_{80}O_{17}$	90657714	783.49 $[M-Ara-H]^-$ 621.44 $[M-Ara-Glc-H]^-$ 459.38 $[M-Ara-2Glc-H]^-$
G60	ginsenoside compound-O	26.41	915.5323	915.5323	$C_{47}H_{80}O_{17}$	21672569	783.49 $[M-Ara-H]^-$ 621.44 $[M-Ara-Glc-H]^-$ 459.38 $[M-Ara-2Glc-H]^-$
G61	vinaginsenoside R16	26.55	915.5322	915.5323	$C_{47}H_{80}O_{17}$	131751558	783.49 $[M-Xyl-H]^-$ 621.44 $[M-Xyl-Glc-H]^-$ 459.38 $[M-Xyl-2Glc-H]^-$
G62	gypenoside IX	26.69	915.5306	915.5323	$C_{47}H_{80}O_{17}$	46887681	783.49 $[M-Xyl-H]^-$ 621.44 $[M-Xyl-Glc-H]^-$ 459.38 $[M-Xyl-2Glc-H]^-$
G63	ginsenoside F2[a]	27.66	783.4900	783.4900	$C_{42}H_{72}O_{13}$	9918692	621.44 $[M-Glc-H]^-$ 459.38 $[M-2Glc-H]^-$
G64	chikusetsusaponin IVa	27.78	793.4370	793.4380	$C_{42}H_{66}O_{14}$	13909684	613.37 $[M-Glc-H_2O-H]^-$ 455.35 $[M-Glc-GlcA-H]^-$
G65	20(*S*)-ginsenoside Rg3[a]	28.12	783.4877	783.4900	$C_{42}H_{72}O_{13}$	9918693	621.44 $[M-Glc-H]^-$ 459.38 $[M-2Glc-H]^-$
G66	ginsenoside Rs3	28.13	825.4987	825.5006	$C_{44}H_{74}O_{14}$	100937823	783.49 $[M-Ac-H]^-$ 765.48 $[M-Ac-H_2O-H]^-$ 621.44 $[M-Ac-Glc-H]^-$ 459.38 $[M-Ac-2Glc-H]^-$
G67	ginsenoside compound-Mc	28.36	753.4810	753.4795	$C_{41}H_{70}O_{12}$	9896928	621.44 $[M-Ara-H]^-$ 459.38 $[M-Ara-Glc-H]^-$
G68	calenduloside E	28.45	631.3840	631.3852	$C_{36}H_{56}O_9$	176079	455.35 $[M-GlcA-H]^-$
G69	ginsenoside compound-Y	28.50	753.4796	753.4795	$C_{41}H_{70}O_{12}$	21672570	459.38 $[M-Ara-Glc-H]^-$
G70	20(*S*)-ginsenoside Rh2[a]	29.09	621.4357	621.4372	$C_{36}H_{62}O_8$	119307	459.38 $[M-Glc-H]^-$
G71	20(*R*)-ginsenoside Rh2[a]	29.28	621.4351	621.4372	$C_{36}H_{62}O_8$	14081290	459.38 $[M-Glc-H]^-$

[a] Identification of these compounds was confirmed using reference standards.

profiling data using multiple accessions of *P. ginseng* produced for use in plant breeding by the Korean Rural Development Administration. As all accessions were cultivated at the same research farm over the same period, the effects of most environmental variables on chemical composition were minimized. The results highlight the influence of genetic variation on the specialized metabolome of *P. ginseng*. The genetic variation among 119 accessions of *P. ginseng* was evaluated via genotyping by sequencing, which revealed 249,885 single nucleotide polymorphisms within the samples [64]. Phylogenetic analysis highlights the extensive genetic diversity among ginseng resources (Fig. 2), which is consistent with the high genetic diversity within this species observed in other studies [65–68].

The untargeted metabolomics dataset was acquired using the berries and roots of 60 accessions, which were a subset of the 119 accessions. 71 mass spectral features were putatively annotated as ginsenosides from this dataset based on their fragmentation spectra and a comparison of elution order with those reported in previous studies (Table 1). As only a few ginsenosides are commercially available, this putative method is the most common strategy used in structural annotation. Most relevant knowledge could be only found in previous studies, but recent technical advances are accelerating the use of computational tools in compound annotation in untargeted metabolomics projects [69]. Chemical knowledge should be digitized into a computer-readable format and stored in publicly available databases for use in computational toolkits. A search query with "ginsenoside" in the compound name provided 7509 experimental spectra from the library of Global Natural Product Social Molecular Networking (https://gnps.ucsd.edu) [70], along with 193 suspect spectra [71], as of November 6, 2023. This suggests that the knowledge regarding the mass spectral fragmentation of ginsenosides has been digitized well. However, the relative retention time, which is extraordinarily helpful in distinguishing isomeric ginsenosides, has not been digitized, mainly because of the lack of a standardized method for reporting chromatographic metadata [72]. Organizing Table 1, we recognized that the nomenclature of ginsenosides, particularly minor ones, is arbitrary, and several compounds have numerous synonyms, whereas others are challenging to structurally characterize using the names provided in previous studies. The metabolomics community recently began to raise this nomenclature issue [73,74], although a standardized nomenclature system has not yet been established. To clarify the compound structures, the PubChem IDs of the annotated compounds are included in Table 1, which may be a possible solution for use in future profiling studies to overcome the nomenclature confusion.

Fig. 3 shows the variations in the peak intensities of the 71 putative ginsenosides, which are normalized as *Z*-scores, and the contents of most ginsenosides largely vary based on genetic variation. Outlier samples of numerous ginsenosides are observed outside the whiskers representing the 1.5 interquartile ranges, indicating that certain accessions may contain extremely high or low contents of specific ginsenosides. We also quantified the contents of 13 ginsenosides (ginsenosides Rb_1, Rb_2, Rb_3, Rc, Rd, Re, Rf, Rg_1, Rg_2, Rg_3, Rh_1, Ro, and F_2) in the roots of 87 accessions of *P. ginseng*, which was another subset of the 119 accessions, and the results support the large variations among the different genotypes. The total content of the 13 quantified ginsenosides varies from 0.68 % to 2.01 %, and every ginsenoside exhibits a large variation among the 87 samples, as shown in Fig. 4.

Considering the entirety of the genome used in the genetic analysis, it is anticipated that even more variation may exist among the accessions, and these genetic differences may profoundly influence the ginsenoside contents and composition. Recent pan-genome studies of major crops suggested that >30 % of the total genes are variable genes unique to certain genotypes within the same species [75]. Several genes, which are responsible for determining key traits, occur exclusively in specific individuals. These findings suggest that specific functional genes may occur in specific ginseng accessions, and these unique genotypes may

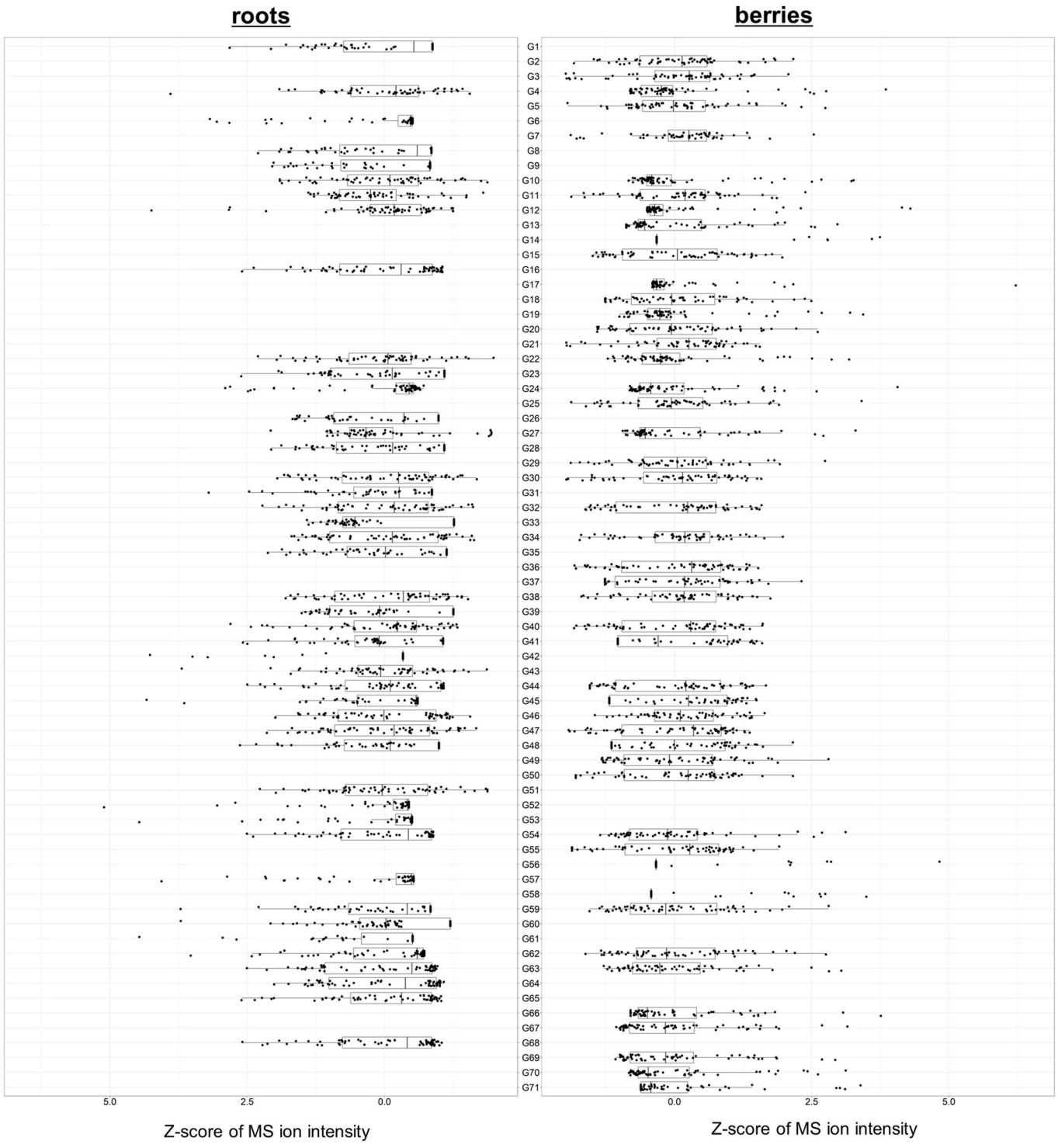


Fig. 3. Untargeted LC-MS/MS analysis reveals the relative contents of 71 putative ginsenosides in the roots and berries of 60 *P. ginseng* accessions. Due to the different mass spectral intensities of the ions, the values were normalized to yield *Z*-scores. The blanks mean each compound was not detected. The putative annotations of ginsenosides **G1**–**G71** are summarized in Table 1.

contribute to forming distinct ginsenoside profiles.

Such large variations in the chemical compositions of different genotypes suggest that the results of previous ginsenoside profiling studies should be critically accepted. Most studies summarized above did not report the genotypes of the analyzed samples, and thus, whether the conclusions may be extrapolated to all individuals is unclear owing to the unclear levels of representativeness of the samples. For example, Yuk et al. suggested that ginsenosides Rd and Re are present at lower levels in *P. ginseng* compared to those in *P. quinquefolius* [21]. However, our data suggest that different genotypes of *P. ginseng* display a wide range of contents of ginsenosides Rd and Re (0.02–0.10 % for Rd and 0.06–0.42 % for Re). We cannot be certain that the marker compounds suggested by Yuk et al. are valid for all *P. ginseng* genotypes, because validating this issue is extremely challenging when the exact amounts of marker compounds are not provided.

Uncertainty in variable control is another issue encountered when

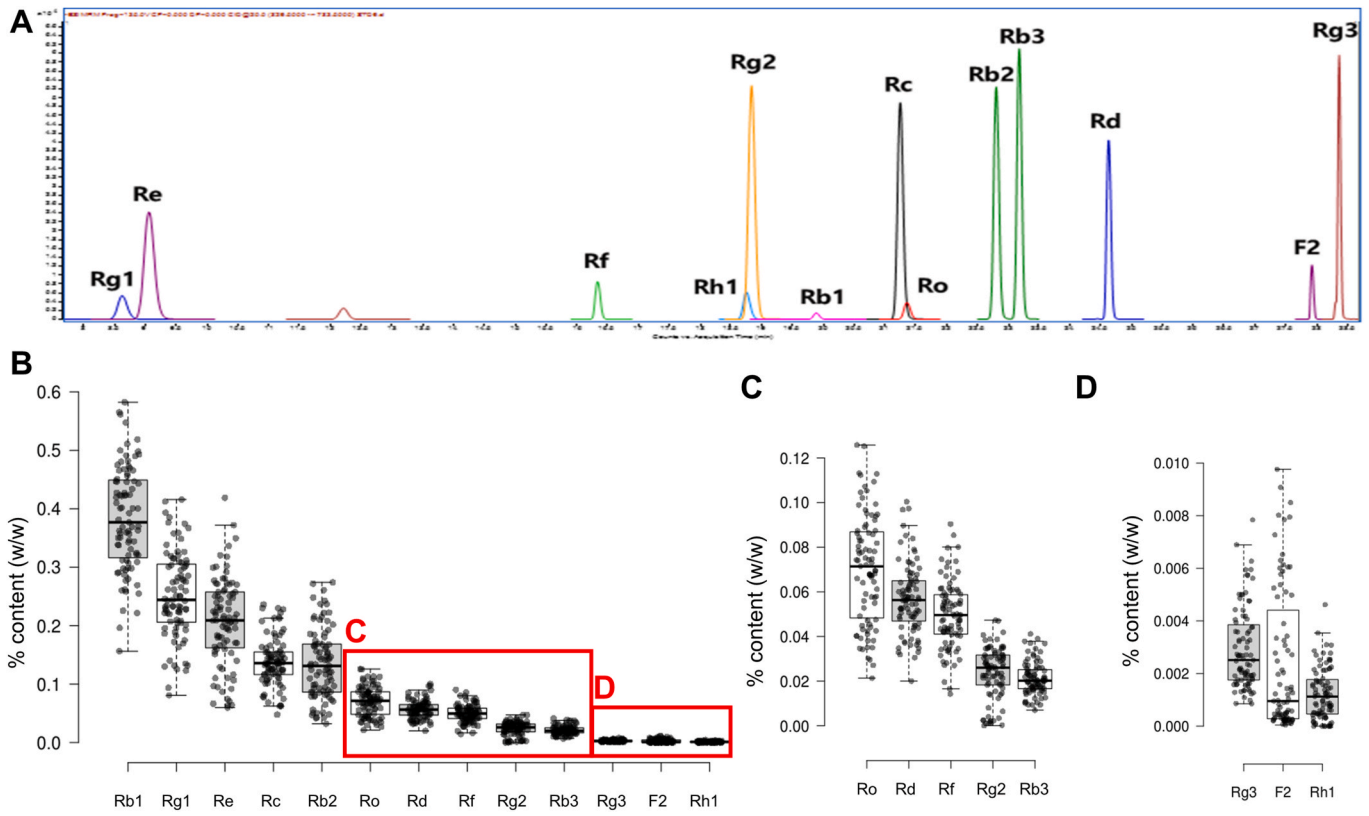


Fig. 4. Targeted LC-MS/MS analysis reveals the variety in the contents of 13 ginsenosides in the roots of 87 *P. ginseng* accessions. (A) The multiple reaction monitoring chromatogram of the reference standard mixture. (B–D) Absolute contents of the analyzed ginsenosides in the samples.

profiling studies do not consider the genetic variation in the samples. In numerous study designs, different genotypes may be unexpected dependent variables that may affect the analytical result. For instance, numerous studies reported that ginseng grown in different geographical locations displayed different chemical profiles, as summarized above. However, most of these studies did not consider which genotype or cultivar was mainly cultivated at each location. Although many previous studies described the chemical differences between ginseng grown in Korea and China, whether these differences originate from environmental or genetic differences is unclear.

9. Perspective and conclusion

Ginsenoside profiling has provided significant insights into the chemical and biological aspects of ginseng. However, in preparing this review, we also identified the common limitations in most of the previous studies analyzed here. The most critical issue is that most studies analyzed the samples, observed the differences between the chemical profiles, and concluded the investigation, providing only plausible conclusions. Similar to other -omics approaches, metabolomics is an unbiased observation-based method for hypothesis generation [76]. Ginsenoside profiling revealed correlations between the contents of certain ginsenosides and variables such as species and geographical origin, but most of these correlations did not proceed to concrete conclusions describing causation. Profiling studies provided numerous hypotheses over the last 25 y. Table 2 summarizes some of the 'marker' compounds for interspecific diversity, geographical origin, and steaming process suggested by the previous studies, and they should now be validated via investigations with a well-controlled experimental design.

We already described the issue regarding the uncertainty of variable control in ginsenoside profiling studies, but genetic variation is not the only factor that renders the variable control arbitrary. Geographical origin is another arbitrary variable. Song et al. distinguished ginseng from Korea and China based on chemical profiles [29], and Zhang et al. suggested the chemical heterogeneity of ginseng grown in three different regions of China [30]. Is the chemical diversity of Korean and Chinese ginseng larger than that of ginseng from different regions of China? We cannot answer this question because no further details regarding the geographical origins of Chinese ginseng were provided by Song et al., Yoon et al. [32], or other researchers. To mitigate this ambiguity, more detailed metadata (structured information regarding the data) should be collected and provided in future ginseng profiling studies. In this regard, a recent study by Sun et al. shows which types of data should be provided, although whether these types are optimal is unclear [77]. In addition to the names of the provinces, the authors provided GPS coordinates and the climates of the collection sites, which enabled them to suggest a climate-related hypothesis.

"Multiomics" is becoming a buzzword in multiple subfields of biology, and ginseng research is no exception. Integration with other -omics techniques, especially genomics and transcriptomics, should drastically expand the value of metabolomics data. Recent studies regarding the genome, transcriptome, and metabolome of wild and cultivated tomatoes (*Solanum* spp.) are excellent cases highlighting the potential of plant multiomics studies focusing on specialized metabolism [78–81]. Despite such potential, the application of multiomics strategies, particularly large-scale metabolome-based genome-wide association studies, began only a few years ago, and most of them have been performed using model organisms (e.g. *Arabidopsis thaliana*) or major crops. There are several hurdles to overcome before multiomics strategies may be applied in ginseng research, but it has become considerably easier since the draft genome sequence of *P. ginseng* was fully assembled in 2018 [82]. Such a large-scale multiomics study regarding *Panax* species has not yet been reported. However, small-scale studies integrating transcriptome and ginsenoside profiles were recently conducted and provided many insights into ginsenoside biosynthesis [83–87], the tissue-level distributions of ginsenosides and their biosynthesis [88,89],

Table 2
Summary of compounds suggested to be significantly different along with the interspecific diversity, geographical origin, and steaming process by previous studies.

		Interspecific difference (*P. ginseng*)		geographical origin (grown in Korea)		processing (red ginseng)
		vs. *P. quinquefolius*	vs. *P. notoginseng*	vs. grown in China	vs. grown in Japan	vs. white ginseng
PPD-type	ginsenoside Ra1	H [20–22]	H [21]	H [32]		
	ginsenoside Ra2			H [32]		
	ginsenoside Ra3	H [22]		H [32]		
	ginsenoside Rb1	L [12,17,20]		H [31,32]		L [53]
	malonylginsenoside Rb1	L [12]				L [53]
	ginsenoside Rb2	L [12], H [17,20–22]	H [22]	L [32]		L [53]
	malonylginsenoside Rb2	L [12]	H [22]			L [53]
	ginsenoside Rb3			H [32]		
	ginsenoside Rc	L [12]	H [22]	H [32]		
	malonylginsenoside Rc	L [12]	H [22]			L [53]
	ginsenoside Rd	L [12,21]				H [53]
	20®-ginsenoside Rh2			H [32]	H [32]	
	20®-ginsenoside Rg3	L [20]				
	20(*S*)-ginsenoside Rg3	L [20]				
	ginsenoside Rg5	L [20]				H [55]
	ginsenoside Rk1	L [20]		H [32]		H [55]
	ginsenoside Rs1	H [22]	H [22]			
	ginsenoside Rs3					H [55]
	notoginsenoside Fa		L [21]			
	notoginsenoside Fe					H [53]
	notoginsenoside R4		L [21]			
	gypenoside XVII	L [21]		H [32]	H [32]	L [55]
PPT-type	ginsenoside Re	L [20,21]				L [53]
	ginsenoside Rf	H [12,17–22]	H [21,22]	H [29,31,32]	H [32]	L [53]
	ginsenoside Rg1	H [12,17,20]		H [32]	H [32]	L [53]
	ginsenoside Rg2	H [22]				
	20®-ginsenoside Rg2	L [20]		H [32]	H [32]	H [55]
	20(*S*)-ginsenoside Rg2	L [20]				
	ginsenoside Rg6	L [20]				H [55]
	ginsenoside Rh1				H [32]	H [55]
	20(*S*)-ginsenoside Rh1	H [20]		H [32]	H [32]	
	ginsenoside Rk3			H [32]	H [32]	
	ginsenoside F3			H [32]	L [32]	
	ginsenoside F4	L [20]				H [55]
	ginsenoside F5			H [32]		
	20-*O*-glucosyl-ginsenoside Rf	H [22]				
	24(*R*)-pseudoginsenoside F11	L [18,19,21,22]				
	notoginsenoside R1	H [20]	L [21,22]			
	notoginsenoside R2	H [22]		L [29]		
	vinaginsenoside R4			H [32]	H [32]	
OA-type	ginsenoside Ro	L [20]	H [22]	L [29,32]	H [32]	L [53]

'L' denotes lower while 'H' denotes higher. e.g. ginsenoside Ra1 was suggested to be lower in *P. ginseng* than in *P. quinquefolius* by references 20–22.

and the effects of microorganisms [90,91]. We anticipate that further multiomics studies with *Panax* species will be conducted in the near future.

LC-MS/MS-based ginsenoside profiling has provided valuable insights to expand our understanding of ginseng. Untargeted analysis is useful in expanding the observational window to ginsenosides that have not yet been isolated, whereas targeted analysis provides the absolute quantitative data required to answer various questions regarding the chemistry and biology of this valuable medicinal plant. Meanwhile, numerous previous studies are limited. Clearly, our objective is not to criticize the studies summarized here, and they should still be valuable cornerstones of future research on ginseng and its chemical composition.

Declaration of competing interests

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean Government (Ministry of Science and ICT; 2022M3H9A2082952, 2018R1A5A2023127, and RS-2023-00211868) and a grant (21173MFDS561) from the Ministry of Food and Drug Safety of Korea in 2023.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgr.2024.01.004.

References

[1] Park H-J, Kim D-H, Park S-J, Kim J-M, Ryu J-H. Ginseng in traditional herbal prescriptions. J Ginseng Res 2012;36:225–41.
[2] Kim Y-S, Woo J-Y, Han C-K, Chang I-M. Safety analysis of *Panax ginseng* in randomized clinical trials: a systematic review. Medicines 2015;2:106–26.
[3] Kitagawa I, Taniyama T, Yoshikawa M, Ikenishi Y, Nakagawa Y. Chemical studies on crude drug processing. VI.: chemical structures of malonyl-ginsenosides Rb1, Rb2, Rc, and Rd isolated from the root of *Panax ginseng* C. A. Meyer. Chem Pharm Bull 1989;37:2961–70.
[4] Liu Z, Li Y, Li X, Ruan C-C, Wang L-J, Sun G-Z. The effects of dynamic changes of malonyl ginsenosides on evaluation and quality control of *Panax ginseng* C. A. Meyer. J Pharm Biomed Anal 2012;64–65:56–63.
[5] Shin B-K, Kwon SW, Park JH. Chemical diversity of ginseng saponins from *Panax ginseng*. J Ginseng Res 2015;39:287–98.
[6] Wang Y, Choi H-K, Brinckmann JA, Jiang X, Huang L. Chemical analysis of *Panax quinquefolius* (North American ginseng): a review. J Chromatogr A 2015;1426:1–15.
[7] Wang T, Guo R, Zhou G, Zhou X, Kou Z, Sui F, Li C, Tang L, Wang Z. Traditional uses, botany, phytochemistry, pharmacology and toxicology of *Panax notoginseng* (Burk.) F.H. Chen: a review. J Ethnopharmacol 2016;188:234–58.
[8] Xu C, Wang W, Wang B, Zhang T, Cui X, Pu Y, Li N. Analytical methods and biological activities of *Panax notoginseng* saponins: recent trends. J Ethnopharmacol 2019;236:443–65.

[9] Yang Y, Ju Z, Yang Y, Zhang Y, Yang L, Wang Z. Phytochemical analysis of *Panax* species: a review. J Ginseng Res 2021;45:1–21.
[10] Li X, Liu J, Zuo T-T, Hu Y, Li Z, Wang H-D, Xu W-Y, Yang W-Z, Guo D-A. Advances and challenges in ginseng research from 2011 to 2020: the phytochemistry, quality control, metabolism, and biosynthesis. Nat Prod Rep 2022;39:875–909.
[11] van Breemen RB, Huang C-R, Lu Z-Z, Rimando A, Fong HHS, Fitzloff JF. Electrospray liquid chromatography/mass spectrometry of ginsenosides. Anal Chem 1995;67:3985–9.
[12] Wang X, Sakuma T, Asafu-Adjaye E, Shiu GK. Determination of ginsenosides in plant extracts from *Panax ginseng* and *Panax quinquefolius* L. by LC/MS/MS. Anal Chem 1999;71:1579–84.
[13] Fuzzati N, Gabetta B, Jayakar K, Pace R, Peterlongo F. Liquid chromatography–electrospray mass spectrometric identification of ginsenosides in *Panax ginseng* roots. J Chromatogr A 1999;854:69–79.
[14] Yang W-Z, Ye M, Qiao X, Liu C-F, Miao W-J, Bo T, Tao H-Y, Guo D-A. A strategy for efficient discovery of new natural compounds by integrating orthogonal column chromatography and liquid chromatography/mass spectrometry analysis: its application in *Panax ginseng, Panax quinquefolium* and *Panax notoginseng* to characterize 437 potential new ginsenosides. Anal Chim Acta 2012;739:56–66.
[15] Patti GJ, Yanes O, Siuzdak G. Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol 2012;13:263–9.
[16] Borsch T, Berendsohn W, Dalcin E, Delmas M, Demissew S, Elliott A, Fritsch P, Fuchs A, Geltman D, Güner A, et al. World Flora Online: placing taxonomists at the heart of a definitive and comprehensive global resource on the world's plants. Taxon 2020;69:1311–41.
[17] Ji QC, Harkey MR, Henderson GL, Gershwin ME, Stern JS, Hackman RM. Quantitative determination of ginsenosides by high-performance liquid chromatography-tandem mass spectrometry. Phytochem Anal 2001;12:320–6.
[18] Chan TWD, But PPH, Cheng SW, Kwok IMY, Lau FW, Xu HX. Differentiation and authentication of *Panax ginseng, Panax quinquefolius*, and ginseng products by using HPLC/MS. Anal Chem 2000;72:1281–7.
[19] Li W, Gu C, Zhang H, Awang DVC, Fitzloff JF, Fong HHS, van Breemen RB. Use of high-performance liquid chromatography-tandem mass spectrometry to distinguish *Panax ginseng* C. A. Meyer (Asian ginseng) and *Panax quinquefolius* L. (North American ginseng). Anal Chem 2000;72:5417–22.
[20] Park H-W, In G, Kim J-H, Cho B-G, Han G-H, Chang I-M. Metabolomic approach for discrimination of processed ginseng genus (*Panax ginseng* and *Panax quinquefolius*) using UPLC-QTOF MS. J Ginseng Res 2014;38:59–65.
[21] Yuk J, Patel DN, Isaac G, Smith K, Wrona M, Olivos HJ, Yu K. Chemical profiling of ginseng species and ginseng herbal products using UPLC/QTOF-MS. J Braz Chem Soc 2016;27:1476–83.
[22] Yang W, Qiao X, Li K, Fan J, Bo T, Guo D-A, Ye M. Identification and differentiation of *Panax ginseng, Panax quinquefolium*, and *Panax notoginseng* by monitoring multiple diagnostic chemical markers. Acta Pharm Sin B 2016;6:568–75.
[23] Shi X-J, Yang W-Z, Qiu S, Yao C-L, Shen Y, Pan H-Q, Bi Q-R, Yang M, Wu W-Y, Guo D-A. An in-source multiple collision-neutral loss filtering based nontargeted metabolomics approach for the comprehensive analysis of malonyl-ginsenosides from *Panax ginseng, P. quinquefolius*, and *P. notoginseng*. Anal Chim Acta 2017;952:59–70.
[24] Shi X, Yang W, Huang Y, Hou J, Qiu S, Yao C, Feng Z, Wei W, Wu W, Guo D. Direct screening of malonylginsenosides from nine Ginseng extracts by an untargeted profiling strategy incorporating in-source collision-induced dissociation, mass tag, and neutral loss scan on a hybrid linear ion-trap/Orbitrap mass spectrometer coupled to ultra-high performance liquid chromatography. J Chromatogr A 2018;1571:213–22.
[25] Yang W-Z, Shi X-J, Yao C-L, Huang Y, Hou J-J, Han S-M, Feng Z-J, Wei W-L, Wu W-Y, Guo D-A. A novel neutral loss/product ion scan-incorporated integral approach for the untargeted characterization and comparison of the carboxyl-free ginsenosides from *Panax ginseng, Panax quinquefolius*, and *Panax notoginseng*. J Pharm Biomed Anal 2020;177:112813.
[26] Kim S, Shin B-K, Lim DK, Yang T-J, Lim J, Park JH, Kwon SW. Expeditious discrimination of four species of the *Panax* genus using direct infusion-MS/MS combined with multivariate statistical analysis. J Chromatogr B 2015;1002:329–36.
[27] Cassago ALL, Artêncio MM, de Moura Engracia Giraldi J, Da Costa FB. Metabolomics as a marketing tool for geographical indication products: a literature review. Eur Food Res Technol 2021;247:2143–59.
[28] Song H-H, Kim D-Y, Woo S, Lee H-K, Oh S-R. An approach for simultaneous determination for geographical origins of Korean Panax ginseng by UPLC-QTOF/MS coupled with OPLS-DA models. J Ginseng Res 2013;37:341–8.
[29] Song H-H, Moon JY, Ryu HW, Noh B-S, Kim J-H, Lee H-K, Oh S-R. Discrimination of white ginseng origins using multivariate statistical analysis of data sets. J Ginseng Res 2014;38:187–93.
[30] Zhang C, Liu Z, Lu S, Xiao L, Xue Q, Jin H, Gan J, Li X, Liu Y, Liang X. Rapid discrimination and prediction of ginsengs from three origins based on UHPLC-Q-TOF-MS combined with SVM. Molecules 2022;27:4225.
[31] Chen W, Balan P, Popovich DG. Analysis of ginsenoside content (*Panax ginseng*) from different regions. Molecules 2019;24:3491.
[32] Yoon D, Shin W-C, Oh S-M, Choi B-R, Lee DY. Integration of multiplatform metabolomics and multivariate analysis for geographical origin discrimination of *Panax ginseng*. Food Res Int 2022;159:111610.
[33] Shuai M, Yang Y, Bai F, Cao L, Hou R, Peng C, Cai H. Geographical origin of American ginseng (*Panax quinquefolius* L.) based on chemical composition combined with chemometric. J Chromatogr A 2022;1676:463284.
[34] Pang S, Piao X, Zhang X, Chen X, Zhang H, Jin Y, Li Z, Wang Y. Discrimination for geographical origin of *Panax quinquefolius* L. using UPLC Q-Orbitrap MS-based metabolomics approach. Food Sci Nutr 2023;11:4843–52.
[35] Wang Y, Pan J-Y, Xiao X-Y, Lin R-C, Cheng Y-Y. Simultaneous determination of ginsenosides in *Panax ginseng* with different growth ages using high-performance liquid chromatography-mass spectrometry. Phytochem Anal 2006;17:424–30.
[36] Kim N, Kim K, Choi BY, Lee D, Shin Y-S, Bang K-H, Cha S-W, Lee JW, Choi H-K, Jang DS, et al. Metabolomic approach for age discrimination of *Panax ginseng* using UPLC-Q-Tof MS. J Agric Food Chem 2011;59:10435–41.
[37] Kim N, Kim K, Lee D, Shin Y-S, Bang K-H, Cha S-W, Lee JW, Choi H-K, Hwang BY, Lee D. Nontargeted metabolomics approach for age differentiation and structure interpretation of age-dependent key constituents in hairy roots of *Panax ginseng*. J Nat Prod 2012;75:1777–84.
[38] Huang B-M, Zha Q-L, Chen T-B, Xiao S-Y, Xie Y, Luo P, Wang Y-P, Liu L, Zhou H. Discovery of markers for discriminating the age of cultivated ginseng by using UHPLC-QTOF/MS coupled with OPLS-DA. Phytomedicine 2018;45:8–17.
[39] Bai H, Wang S, Liu J, Gao D, Jiang Y, Liu H, Cai Z. Localization of ginsenosides in *Panax ginseng* with different age by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry imaging. J Chromatogr B 2016;1026:263–71.
[40] Lee JW, Ji S-H, Lee Y-S, Choi DJ, Choi B-R, Kim G-S, Baek N-I, Lee DY. Mass spectrometry based profiling and imaging of various ginsenosides from *Panax ginseng* roots at different ages. Int J Mol Sci 2017;18:1114.
[41] Yang Y, Yang Y, Qiu H, Ju Z, Shi Y, Wang Z, Yang L. Localization of constituents for determining the age and parts of ginseng through ultraperfomance liquid chromatography quadrupole/time of flight-mass spectrometry combined with desorption electrospray ionization mass spectrometry imaging. J Pharm Biomed Anal 2021;193:113722.
[42] Xu X-F, Cheng X-L, Lin Q-H, Li S-S, Jia Z, Han T, Lin R-C, Wang D, Wei F, Li X-R. Identification of mountain-cultivated ginseng and cultivated ginseng using UPLC/oa-TOF MSE with a multivariate statistical sample-profiling strategy. J Ginseng Res 2016;40:344–50.
[43] Zhu H, Lin H, Tan J, Wang C, Wang H, Wu F, Dong Q, Liu Y, Li P, Liu J. UPLC-QTOF/MS-based nontargeted metabolomic analysis of mountain- and garden-cultivated ginseng of different ages in Northeast China. Molecules 2018;24:33.
[44] Guo N, Yang Y, Yang X, Guan Y, Yang J, Quan J, Yan H, Hou W, Zhang G. Growth age of mountain cultivated ginseng affects its chemical composition. Ind Crops Prod 2021;167:113531.
[45] Qu H, Wang J, Yao C, Wei X, Wu Y, Cheng M, He X, Li J, Wei W, Zhang J, et al. Enhanced profiling and quantification of ginsenosides from mountain-cultivated ginseng and comparison with garden-cultivated ginseng. J Chromatogr A 2023;1692:463826.
[46] Hasegawa H, Sung JH, Matsumiya S, Uchiyama M. Main ginseng saponin metabolites formed by intestinal bacteria. Planta Med 1996;62:453–7.
[47] Hasegawa H, Sung JH, Benno Y. Role of human intestinal *Prevotella oris* in hydrolyzing ginseng saponins. Planta Med 1997;63:436–40.
[48] Akao T, Kanaoka M, Kobashi K. Appearance of compound K, a major metabolite of ginsenoside Rb1 by intestinal bacteria, in rat plasma after oral administration: measurement of compound K by enzyme immunoassay. Biol Pharm Bull 1998;21:245–9.
[49] Akao T, Kida H, Kanaoka M, Hattori M, Kobashi K. Drug metabolism: intestinal bacterial hydrolysis is required for the appearance of compound K in rat plasma after oral administration of ginsenoside Rb1 from Panax ginseng. J Pharm Pharmacol 1998;50:1155–60.
[50] Lee SM, Bae B-S, Park H-W, Ahn N-G, Cho B-G, Cho Y-L, Kwak Y-S. Characterization of Korean Red ginseng (*Panax ginseng* Meyer): history, preparation method, and chemical composition. J Ginseng Res 2015;39:384–91.
[51] Huang L, Li H-J, Wu Y-C. Processing technologies, phytochemistry, bioactivities and applications of black ginseng – a novel manufactured ginseng product: a comprehensive review. Food Chem 2023;407:134714.
[52] Zhang H-M, Li S-L, Zhang H, Wang Y, Zhao Z-L, Chen S-L, Xu H-X. Holistic quality evaluation of commercial white and red ginseng using a UPLC-QTOF-MS/MS-based metabolomics approach. J Pharm Biomed Anal 2012;62:258–73.
[53] Xie Y-Y, Luo D, Cheng Y-J, Ma J-F, Wang Y-M, Liang Q-L, Luo G-A. Steaming-induced chemical transformations and holistic quality assessment of red ginseng derived from Panax ginseng by means of HPLC-ESI-MS/MS^n-based multicomponent quantification fingerprint. J Agric Food Chem 2012;60:8213–24.
[54] Sun B-S, Xu M-Y, Li Z, Wang Y-B, Sung C-K. UPLC-Q-TOF-MS/MS analysis for steaming times-dependent profiling of steamed *Panax quinquefolius* and its ginsenosides transformations induced by repetitious steaming. J Ginseng Res 2012;36:277–90.
[55] Chu C, Xu S, Li X, Yan J, Liu L. Profiling the ginsenosides of three ginseng products by LC-Q-TOF/MS. J Food Sci 2013;78:C653–9.
[56] Eom SJ, Kim K-T, Paik H-D. Microbial bioconversion of ginsenosides in *Panax ginseng* and their improved bioactivities. Food Rev Int 2018;34:698–712.
[57] Bai Y, Gänzle MG. Conversion of ginsenosides by *Lactobacillus plantarum* studied by liquid chromatography coupled to quadrupole trap mass spectrometry. Food Res Int 2015;76:709–18.
[58] Xiao D, Xiu Y, Yue H, Sun X, Zhao H, Liu S. A comparative study on chemical composition of total saponins extracted from fermented and white ginseng under the effect of macrophage phagocytotic function. J Ginseng Res 2017;41:379–85.
[59] Kim K. Effect of ginseng and ginsenosides on melanogenesis and their mechanism of action. J Ginseng Res 2015;39:1–6.
[60] Li F, Lv C, Li Q, Wang J, Song D, Liu P, Zhang D, Lu J. Chemical and bioactive comparison of flowers of *Panax ginseng* Meyer, *Panax quinquefolius* L., and *Panax notoginseng* Burk. J Ginseng Res 2017;41:487–95.

[61] Jia L, Zuo T, Zhang C, Li W, Wang H, Hu Y, Wang X, Qian Y, Yang W, Yu H. Simultaneous profiling and holistic comparison of the metabolomes among the flower buds of *Panax ginseng, Panax quinquefolius*, and *Panax notoginseng* by UHPLC/IM-QTOF-HDMS^{E}-based metabolomics analysis. Molecules 2019;24:2188.
[62] Yoon D, Choi B-R, Kim Y-C, Oh SM, Kim H-G, Kim J-U, Baek N-I, Kim S, Lee DY. Comparative analysis of *Panax ginseng* berries from seven cultivars using UPLC-QTOF/MS and NMR-based metabolic profiling. Biomolecules 2019;9:424.
[63] Chang X, Zhang J, Li D, Zhou D, Zhang Y, Wang J, Hu B, Ju A, Ye Z. Nontargeted metabolomics approach for the differentiation of cultivation ages of mountain cultivated ginseng leaves using UHPLC/QTOF-MS. J Pharm Biomed Anal 2017;141:108–22.
[64] Cho W-H. Establishment of high-throughput digital genotyping system for *Panax ginseng* and *Triticum aestivum* [dissertation]. Seoul: Seoul National University; 2021.
[65] Ma K-H, Dixit A, Kim Y-C, Lee D-Y, Kim T-S, Cho E-G, Park Y-J. Development and characterization of new microsatellite markers for ginseng (*Panax ginseng* C. A. Meyer). Conserv Genet 2007;8:1507–9.
[66] Choi H-I, Kim NH, Kim JH, Choi BS, Ahn I-O, Lee J-S, Yang T-J. Development of reproducible EST-derived SSR markers and assessment of genetic diversity in *Panax ginseng* cultivars and related species. J Ginseng Res 2011;35:399–412.
[67] Jang W, Jang Y, Kim N-H, Waminal NE, Kim YC, Lee JW, Yang T-J. Genetic diversity among cultivated and wild *Panax ginseng* populations revealed by high-resolution microsatellite markers. J Ginseng Res 2020;44:637–43.
[68] Lee KJ, Lee J-R, Sebastin R, Cho G-T, Hyun DY. Molecular genetic diversity and population structure of ginseng germplasm in RDA-genebank: implications for breeding and conservation. Agronomy 2020;10:68.
[69] Beniddir MA, Kang KB, Genta-Jouve G, Huber F, Rogers S, van der Hooft Jjj. Advances in decomposing complex metabolite mixtures using substructure- and network-based computational metabolomics approaches. Nat Prod Rep 2021;38:1967–93.
[70] Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, et al. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. Nat Biotechnol 2016;34:828–37.
[71] Bittremieux W, Avalon NE, Thomas SP, Kakhkhorov SA, Aksenov AA, Gomes PWP, et al. Open access repository-scale propagated nearest neighbor suspect spectral library for untargeted metabolomics. Nat Commun 2023;14:8488..
[72] Harrieder E-M, Kretschmer F, Dunn W, Böcker S, Witting M. Critical assessment of chromatographic metadata in publicly available metabolomics data repositories. Metabolomics 2022;18:97.
[73] Fahy E, Subramaniam S. RefMet: a reference nomenclature for metabolomics. Nat Methods 2020;17:1173–4.
[74] Koistinen V, Kärkkäinen O, Keski-Rahkonen P, Tsugawa H, Scalbert A, Arita M, Wishart D, Hanhineva K. Towards a Rosetta stone for metabolomics: recommendations to overcome inconsistent metabolite nomenclature. Nat Metab 2023;5:351–4.
[75] Wang S, Qian Y-Q, Zhao R-P, Chen L-L, Song J-M. Graph-based pan-genomes: increased opportunities in plant genomics. J Exp Bot 2023;74:24–39.
[76] Editorial. Defining the scientific method. Nat Methods 2009;6:237.
[77] Sun Y, Liu X, Fu X, Xu W, Guo Q, Zhang Y. Discrepancy study of the chemical constituents of *Panax ginseng* from different growth environments with UPLC-MS-based metabolomics strategy. Molecules 2023;28:2928.
[78] Zhu G, Wang S, Huang Z, Zhang S, Liao Q, Zhang C, et al. Rewiring of the fruit metabolome in tomato breeding. Cell 2018;172. 249–261.e12.
[79] Garbowicz K, Liu Z, Alseekh S, Tieman D, Taylor M, Kuhalskaya A, Ofner I, Zamir D, Klee HJ, Fernie AR, et al. Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato. Mol Plant 2018;11:1147–65.
[80] Szymański J, Bocobza S, Panda S, Sonawane P, Cárdenas PD, Lashbrooke J, Kamble A, Shahaf N, Meir S, Bovy A, et al. Analysis of wild tomato introgression lines elucidates the genetic basis of transcriptome and metabolome variation underlying fruit traits and pathogen response. Nat Genet 2020;52:1111–21.
[81] Tieman D, Zhu G, Resende Jr MFR, Lin T, Nguyen C, Bies D, Rambla JL, Beltran KSO, Taylor M, Zhang B, et al. A chemical genetic roadmap to improved tomato flavor. Science 2017;355:391–4.
[82] Kim N-H, Jayakodi M, Lee S-C, Choi B-S, Jang W, Lee J, Kim HH, Waminal NE, Lakshmanan M, Nguyen Bv, et al. Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*. Plant Biotechnol J 2018;16:1904–17.
[83] Kang KB, Jayakodi M, Lee YS, Nguyen VB, Park H-S, Koo HJ, Choi IY, Kim DH, Chung YJ, Ryu B, et al. Identification of candidate UDP-glycosyltransferases involved in protopanaxadiol-type ginsenoside biosynthesis in *Panax ginseng*. Sci Rep 2018;8:11744.
[84] Koo H, Lee YS, Nguyen VB, Giang VNL, Koo HJ, Park H-S, Mohanan P, Song YH, Ryu B, Kang KB, et al. Comparative transcriptome and metabolome analyses of four *Panax* species explore the dynamics of metabolite biosynthesis. J Ginseng Res 2023;47:44–53.
[85] Lee YS, Park H-S, Lee D-K, Jayakodi M, Kim N-H, Koo HJ, Lee S-C, Kim YJ, Kwon SW, Yang T-J. Integrated transcriptomic and metabolomic analysis of five *Panax ginseng* cultivars reveals the dynamics of ginsenoside biosynthesis. Front Plant Sci 2017;8:1048.
[86] Di P, Yan Y, Wang P, Yan M, Wang Y-P, Huang L-Q. Integrative SMRT sequencing and ginsenoside profiling analysis provide insights into the biosynthesis of ginsenoside in *Panax quinquefolium*. Chin J Nat Med 2022;20:614–26.
[87] Zhang S, Wang G, Zuo T, Zhang X, Xu R, Zhu W, You J, Wang R, Chen P. Comparative transcriptome analysis of rhizome nodes and internodes in *Panax. japonicus* var. *major* reveals candidate genes involved in the biosynthesis of triterpenoid saponins. Genomics 2020;112:1112–9.
[88] Wei G, Yang F, Wei F, Zhang L, Gao Y, Qian J, Chen Z, Jia Z, Wang Y, Su H, et al. Metabolomes and transcriptomes revealed the saponin distribution in root tissues of *Panax quinquefolius* and *Panax notoginseng*. J Ginseng Res 2020;44:757–69.
[89] Wei G, Dong L, Yang J, Zhang L, Xu J, Yang F, Cheng R, Xu R, Chen S. Integrated metabolomic and transcriptomic analyses revealed the distribution of saponins in *Panax notoginseng*. Acta Pharm Sin B 2018;8:458–65.
[90] Ran Z, Chen X, Li R, Duan W, Zhang Y, Fang L, Guo L, Zhou J. Transcriptomics and metabolomics reveal the changes induced by arbuscular mycorrhizal fungi in *Panax quinquefolius* L. J Sci Food Agric 2023;103:4919–33.
[91] Deng L, Luo L, Li Y, Wang L, Zhang J, Zi B, Ye C, Liu Y, Huang H, Mei X, et al. Autotoxic ginsenoside stress induces changes in root exudates to recruit the beneficial *Burkholderia* strain B36 as revealed by transcriptomic and metabolomic approaches. J Agric Food Chem 2023;71:4536–49.