Plant Physiology and Biochemistry 141 (2019) 183–192

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Genetic and chemical diversity among yacon [*Smallanthus sonchifolius* (Poepp. et Endl.) H. Robinson] accessions based on iPBS markers and metabolomic fingerprinting

Jana Žiarovská[a], Guillermo F. Padilla-González[b], Iva Viehmannová[c], Eloy Fernández[c,*]

[a] Faculty of Agrobiology and Food Resources, Slovak University of Agriculture in Nitra, Tr. A. Hlinku 2, 949 76, Nitra, Slovak Republic
[b] School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Av. Do Café, 14040-903, Ribeirão Preto, Brazil
[c] Faculty of Tropical AgriSciences, Czech University of Life Sciences Prague, Kamýcká 129, 165 00, Prague, Czech Republic

ARTICLE INFO

Keywords:
Genetic diversity
iPBS
Metabolomic fingerprints
Retrotransposons
Smallanthus sonchifolius

ABSTRACT

The present study is focused on the characterization of yacon [*Smallanthus sonchifolius* (Poepp. et Endl.) H. Robinson] accessions from different geographic origins (Bolivia, Ecuador, and Peru) by iPBS markers and metabolomic fingerprinting. The results showed that the number of amplified polymorphic fragment levels ranged from 20 up to 27 with a level of polymorphism ranging from 80 to 100%. Five of the iPBS primers used in this study provided no specific banding pattern able to discriminate between the different yacon accessions. However, two iPBS primer pairs were able to separate Peru accessions from those of Ecuador and Bolivia. The UPLC-HRMS/MS-based metabolomic fingerprinting showed highly similar metabolomic fingerprints characterized by the accumulation of high quantities of sesquiterpene lactones and diterpenes, but no apparent geographic clustering. The present study demonstrates that yacon accessions from different geographical origins maintained *ex situ* (in the Czech Republic) present a rather low chemical and genetic diversity.

1. Introduction

Smallanthus sonchifolius (Poepp. et Endl.) H. Robinson is a plant species native to the Andes. This member of the family Asteraceae has been used for a long time in the traditional medicine of South America (Bolivia and Peru) because of its antidiabetic, nutritious, and fertility enhancing properties (Simonovska et al., 2003; Milella et al., 2005). Modern analytical methods have confirmed the medicinal potential of this species, which is related to its high content of bioactive compounds (Russo et al., 2015).

The variability of yacon has been maintained to date by small farmers *in situ* and the gene resources of this crop are stored mainly in the gene banks in Peru, Bolivia and Ecuador (Svobodová et al., 2013). The yacon was successfully introduced not only to other countries of South America, but also to Central and North America, Europe, New Zealand, and Japan (Fernández et al., 2006).

Yacon is one of the species where still very limited information about its genome exist. Up to date, only two matches for *S. sonchifolius* can be found in the National Centre for Biotechnology Information (NCBI nucleotide database). Both of them are the sequences of internal transcribed spacer (ITS1) with 5.8S rRNA gene and ITS2 (Rauscher, 2002; Žiarovská et al., 2013a). This constitutes a limiting factor for choosing DNA based markers that serve effectively in the differentiation of individual clones of yacon. In fact, different types of non-specific DNA markers can be used for these purposes. Different markers such as amplified fragment length polymorphism (AFLP), random amplification of polymorphic DNA (RAPD) and inter-simple sequence repeat (ISSR) have been used for yacon diversity mapping (Milella et al., 2005, 2011; Mansilla et al., 2006; Svobodová et al., 2013). ISSRs were reported to be informative and able to distinguish yacon clones by Svobodová et al. (2013), but no specific pattern connected to the accessions provenience was obtained for them. Milella et al. (2005) applied RAPD and AFLP markers to analyse the genetic diversity in a group of five yacon accessions. The authors have suggested that RAPD could be applicable for differentiation purposes, as well as AFLP. The RAPD marker later showed that its distinguishing ability is limited when more accessions are used in the study (Mansilla et al., 2006). Moreover, problems with the difficulties connected to low reproducibility, unsatisfactory level of polymorphism, and only a limited possibility of inter-laboratory cross analyses are reported for the RAPDs (Štefúnová et al., 2015). AFLP

* Corresponding author. Czech University of Life Sciences Prague, Prague, Kamýcká 129, 165 00, Prague, Czech Republic.
E-mail addresses: jana.ziarovska@uniag.sk (J. Žiarovská), federicopg@gmail.com (G.F. Padilla-González), viehmann@ftz.czu.cz (I. Viehmannová), eloy@ftz.czu.cz (E. Fernández).

https://doi.org/10.1016/j.plaphy.2019.05.020
Received 18 January 2019; Received in revised form 19 May 2019; Accepted 20 May 2019
Available online 25 May 2019

markers serve as a very good analytical tool, but this method detects the SNP changes in the restriction sites and indels, making this technique more static when compared to the techniques based on retrotransposons, such as inter-primer binding site (iPBS) (Kalendar, 2011; Balážová et al., 2014; Guo et al., 2014). Retrotransposons are mobile genetic elements that transpose, through RNA, intermediates in the genomes of plants leaving a molecular fingerprint there. In fact, retrotransposons are reported to possess many of the characteristics that an ideal DNA marker should have (Trebichalský et al., 2013). They can be used both as the specific as well as non-specific DNA markers, depending on the sequence knowledge about them. Kalendar et al. (2010) has introduced a marker technique that provides a general and non-specific DNA marker based on the retrotransposon for plant species – iPBS. These markers are applicable to any organism that possess primer binding sites of retrotransposons that is complimentary to tRNA. The principle of primer design is based on the presence of the retrotransposon sequences that are described in the literature as to be ubiquitous in plant genomes (Sabot and Schulman, 2006). Actually, the method was successfully used for analysis in many different plant species such as *Linum usitatissimum* L. (Smýkal et al., 2011), *Saussurea esthonica* Baer ex. Rupr. (Gailīte et al., 2011), *Liparis loeselii* (L.) Rich. (Belgorudova et al., 2012), and *Prunus armeniaca* L. (Baránek et al., 2012). The iPBS markers were reported as a reliable DNA marker system which overcomes the problems of other non-specific markers and provides a fully reproducible technique.

The genetic diversity and differences in the chemical profiles of plants are influenced by genetic, ontogenic, and environmental factors (Facanali et al., 2015). Genetic and chemical component diversity information on natural populations is crucial for conservation and breeding programs (Silveira et al., 2009). Previous studies on yacon chemistry report a significant variation in the quantitative composition of sesquiterpene lactones (STLs) in twenty accessions collected in the central region of Peru (Mercado et al., 2014), while the biological, molecular, and chemical variability of five yacon accessions from different geographic origins cultivated under field conditions in the Czech Republic were also determined (Valentová et al., 2006; Milella et al., 2011). However, these studies have focused on monitoring restricted classes of metabolites by classical approaches. Thus, a comprehensive phytochemical characterization of yacon by modern techniques is still lacking in the literature. Metabolomics constitute a relatively new research field that focuses on high-throughput characterization of small molecules in biological matrices (Krastanov, 2010). Ultra-performance liquid chromatography-high resolution tandem mass spectrometry (UPLC-HRMS/MS)-based metabolomics has the advantage of simultaneously detecting hundreds to thousands of metabolites even in minor concentrations, while providing structural information of the metabolites based on their MS^2 fragmentation spectra. Considering that determining the intraspecific variation of yacon may be useful in different breeding programs, in this study iPBS markers and metabolomic fingerprinting were used for characterization and discrimination of yacon germplasm cultivated in the Czech Republic.

2. Materials and methods

2.1. Plant material

Different germplasm accessions of yacon (Table 1) are maintained *ex situ* since 1993 in the Faculty of Tropical AgriSciences, Czech University of Life Sciences in Prague. Plants have been continuously propagated over the years by rhizome and cultivated under greenhouse and field conditions.

For DNA extraction, young leaves were used from 14 accessions (PER 04, PER 05, PER 06, PER 07, PER 09, PER 12, PER 13, PER 14, BOL 22, ECU 41, ECU 42, ECU 43, ECU 44 and ECU 45) and for chemical analyses young leaves of 21 accessions (PER 01, PER 02, PER 03, PER 04, PER 07, PER 08, PER 09, PER 10, PER 11, PER 12, PER 13, PER 14, PER 15, BOL 21, BOL 24, ECU 40, ECU 41, ECU 42, ECU 43, ECU 44, and ECU 45). For the chemical analyses, young leaves were harvested from five -month -old plants cultivated in the greenhouse and immediately frozen to avoid perturbations in the plants' metabolome. After collection, the plants were stored at −70 °C until further use.

2.2. DNA extraction and iPBS fingerprinting

Total genomic DNA extraction was performed from the tissues of five young leaves of each yacon accession, following the modified method of Friar (2005), where 2% cetyl trimethylammonium bromide (CTAB), 1% PVP and mercaptoethanol with double ice-cold 95% ethanol precipitation of DNA was used. Seven different iPBS primers were used for the analysis based on our previous results (Žiarovská et al., 2013b). MyTaq™ Mix with 50 ng of DNA and 400 nM of iPBS primers were used in the PCR and the following amplification conditions were used: 95 °C 4 min followed with 35x of 95 °C 1 min; 52–62 °C 1 min; 72 °C 2 min and the final extension of 72 °C 10 min. Amplification was performed in BIO-RAD C1000™ Thermal Cycler.

2.3. Data analysis

PCR amplicons were evaluated in 6% polyacrylamide gel electrophoresis (PAGE) gels and the obtained fingerprints were captured by GeneBox system. Generated iPBS amplicons were scored as present (1) or absent (0) in the gel using the GeneSnap (Syngene) software. Genetic similarity was calculated on the basis of Jaccard coefficient of genetic similarity. Cluster analysis was performed by unweighted pair-group method with arithmetic average (UPGMA) with statistic program SYNTAX.

2.4. Metabolite extraction and metabolomic fingerprinting

Metabolite extraction of frozen leaf material was carried out based on the detailed protocol for large-scale untargeted metabolomics of plant tissues reported by De Vos et al. (2007) with some modifications. Fifty milligrams of each frozen yacon accession were grinded in a mixer mill (9 Hz, 30 s) with liquid nitrogen and extracted with 1 ml of MeOH:H_2O (7:3, v/v) in an ultrasonic bath for 15 min at 25 °C. Plant extracts were centrifuged at 13000 rpm for 5 min, filtered through a 0.2 μm polytetrafluoroethylene (PTFE) membrane and analyzed by UPLC-HRMS/MS on an Agilent 1290 UPLC system coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Scientific).

Chromatographic separation of plant extracts (5 μl) was performed on a C18 column (Acquity CSH, 1.7 μm, 2.1 × 150 mm, Waters) at a constant temperature of 40 °C, using water with 0.1% formic acid as solvent A and acetonitrile with 0.1% formic acid as solvent B (flow rate: 300 μl/min). The gradient program started with 5% acetonitrile and increased the amount linearly to 50% in the first 15 min. Acetonitrile (50%) remained constant for 5 min (20 min) before increasing to 99% in 30 min. This concentration (99% acetonitrile) was kept constant for 11 min for washing and was subsequently equilibrated back to 5% for a total run time of 42 min.

After chromatographic elution, ions were generated by electrospray ionization under the following conditions: spray voltage (positive mode = 4.2 kV; negative mode = 3.5 kV); capillary temperature = 360 °C. Mass detection was performed in an Orbitrap analyzer using the full scan method (resolution = 70,000; scan range = 140–1200 m/z) and the data-dependent MS^2 (dd-MS^2) method using the following parameters: resolution = 17,500; TopN = 5 and isolation window = 1.5 m/z. This method selects the five most intense peaks in each scan and performs the fragmentation of those ions in the HCD (higher-energy collisional dissociation) cell with normalized collision energy of 30 units.

Table 1
Accessions of yacon used in the study.

No. in dendrogram	Origin	Codes of samples	iPBS markers	Metabolomic fingerprinting	Qualitative characteristics		Year of introduction to the Czech Republic
					Colour of root skin	Colour of root flesh	
1	Peru	PER04	x	x	purplish grey	yellow white	2005
2	Peru	PER05	x		purplish red	yellow white	2005
3	Peru	PER06	x		purplish grey	white	2005
4	Peru	PER07	x	x	greyish orange	yellow white	2005
5	Peru	PER09	x	x	greyish orange	orange yellow	2005
6	Peru	PER12	x	x	purplish grey	white	2005
7	Peru	PER13	x	x	white	white	2005
8	Peru	PER14	x	x	greyish orange	yellow white	2005
9	Bolivia	BOL22	x		purplish red	yellow white	2007
10	Ecuador	ECU41	x	x	white	orange yellow	1993
11	Ecuador	ECU42	x	x	purplish grey	white	1993
12	Ecuador	ECU43	x	x	purplish grey	orange	2007
13	Ecuador	ECU45	x	x	purplish grey	yellow white	1994
14	Ecuador	ECU44	x	x	purplish grey	orange	2007
15	Peru	PER01		x	greyish orange	yellow white	2005
16	Peru	PER02		x	greyish orange	white	2005
17	Peru	PER03		x	purplish red	orange yellow	2005
18	Peru	PER08		x	purplish grey	white	2005
19	Peru	PER10		x	white	orange yellow	2005
20	Peru	PER11		x	greyish orange	yellow white	2005
21	Peru	PER15		x	greyish orange	yellow white	2008
22	Ecuador	ECU40		x	purplish red	yellow white	1994
23	Bolivia	BOL21		x	greyish orange	yellow white	2007
24	Bolivia	BOL24		x	greyish orange	yellow	2007

2.5. Dereplication of plant extracts

Compounds were tentatively identified based on a comparison of the accurate mass measurements, MS^2 fragmentation patterns, and online database searches. Accurate mass comparisons were performed relative to the theoretical monoisotopic mass (< 3 ppm accuracy) of the secondary metabolites previously reported in the family Asteraceae; this was done using an in-house database (AsterDB) with more than one thousand metabolites reported in members of this family (fully available at: www.asterbiochem.org/asterdb), and those from the Dictionary of Natural Products (DNP, http://dnp.chemnetbase.com) and the SciFinder database (Chemical Abstract Service, USA). To confirm peak assignments, the fragmentation pattern of the identified metabolites was proposed based on the MS^2 spectra; to compare accurate mass measurements and MS^2 spectra with those of the detected metabolites, the following reference substances were used: quinic acid, enhydrin, uvedalin and longipilin acetate. Based on the Metabolomics Standard Initiative (Sumner et al., 2007), three levels of confidence were adopted to report the accuracy of the identifications: level 1 (identified compounds), corresponding to the identifications based upon the co-characterization with authentic samples; level 2 (putatively annotated compounds), "without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries"; and level 3 (putatively characterized compound classes), "based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class". In all cases, a minimum of two independent and orthogonal data (e.g. accurate mass and MS^2 fragmentation patterns or retention time and accurate mass, etc.) were considered.

2.6. Data pre-treatment and multivariate analysis

The chromatographic data in raw format was exported to the software MZmine 2.28 (MZmine VTT, Finland), where data pre-processing was performed. The MZmine parameters were set as follows: (*mass detection*) noise threshold: 1.0E5; mass detector: Exact mass; MS level: 1; (*FTMS shoulder peaks filter*) mass resolution: 70,000; peak model function: Lorentzian extended; (*chromatogram builder*) min time span: 0.2 min; min height: 5.0E5; *m/z* tolerance: 0.005 (absolute), 5 ppm (relative); (*chromatogram deconvolution*) algorithm: local minimum search; chromatographic threshold: 50%; search minimum in RT range: 0.2 (min); minimum absolute heigh.5.0E5; minimum relative height: 15%; peak duration range: 0.3–2 min; (*isotopic peaks grouper*) representative isotope: most intense; maximum charge: 2; (*retention time normalizer*) *m/z* tolerance: 0.005 (absolute), 5 ppm (relative); retention time tolerance: 3.5 min; minimum standard intensity: 5.0E5; (*alignment*) algorithm: RANSAC aligner; interactions: 0; minimum number of points: 50%; threshold value: 10; (*gap filling*) algorithm: peak finder; intensity tolerance: 10%; (duplicate peaks filter) *m/z* tolerance: 0.005 (absolute), 5 ppm (relative); RT tolerance: 0.7 min; (*adduct search*) $[M + Na–H]^+$ 21.9825 *m/z*, $[M + K–H]^+$ 37.9559 *m/z*, $[M + Mg–2H]^+$ 21.9694 *m/z*, $[M + NH3]^+$ 17.0265 *m/z*, $[M + CH3CN + H]^+$ 42.0338 *m/z* and $[M + Cl + H]^-$ 35.9774 *m/z*, with the adducts selected according to the ionization mode.

The two data matrices obtained after MZmine pre-processing for each ionization mode (positive and negative) were merged into a single matrix in an MS-Excel file, where the peaks detected in the blank sample (extraction solvent) were subtracted from all plant samples to remove possible interfering variables. This matrix, containing the peak areas of all-detected mass features in both ionization modes, was submitted to unsupervised statistical methods in the software SIMCA P 13.0.3.0 (Umetrics AB, Sweden). A principal component analysis (PCA) was performed to reduce the dimensionality of the data set and to determine trends and outliers, while a hierarchical clustering analysis (HCA) allowed the grouping of plant extracts according to the similarity in their chemical fingerprints, providing a dendrogram-like diagram. In the HCA, the Ward's method was used as the clustering algorithm and a Euclidean distance was employed as the distance measurement. Prior to the statistical analyses, the data matrix was scaled by the Pareto method; this keeps the data structure partially intact (major peaks in each sample remain) while reducing the relative importance of large values.

Table 2
Amplification characteristics of the primers used in the study.

Primer name	Number of totally amplified loci levels	% of polymorphism
1846	20	80
1880	23	86
2078	26	99
1845	23	100
1875	25	81
1886	27	82
2080	20	85

3. Results and discussion

3.1. Genetic diversity based on iPBS markers

Here, the first insight into the yacon genetic fingerprint patterns based on iPBS markers was performed. DNA polymorphism of amplified sequences between different retrotransposons was obtained with the result not only of distinguishing all of the analyzed accessions of yacon, but providing a specific iPBS fingerprints based on the provenience of this underutilized species. iPBS polymorphism of yacon applied in the study was based on the selection of the most appropriate iPBS markers performed previously (Žiarovská et al., 2013b). The number of amplified fragment levels ranged from 20 up to 27 for individual primers and the polymorphism ranged from 80 to 100% (Table 2).

Kalendar et al. (2010) has published the technique of iPBS as a platform that provide both an effective isolation of long terminal repeat (LTR) retrotransposons as well as a universal DNA marker system for the plants, if the very concrete group of retrotransposons presented by the iPBS marker are a part of the genome. This mean that, not all the iPBS are presented in all plant species. For the individual primers used in the study, amplified fragments ranged from 20 up to the 27, with the highest variation for the primer 1886 (Fig. 1). These results are in concordance with the variability of the iPBS fragments amplification that is typical for these techniques. Not all the PBS sites are present in the individual genomes of plants and thus not all the iPBS primers are suitable to use as markers. Lapiña et al. (2012) has used different iPBS markers for the analysis of genetic variability in Latvian populations of alfalfa (*Medicago sativa* L.). Ten primers gave monomorphic products and one gave no amplification product among the 29 tested primers. The same primer (2080) that was used in the study of Lapiña et al. (2012) and gave a monomorphic profile in alfalfa has very similar profile characteristics for yacon. In both cases, the monomorphic profile was obtained among the analyzed accessions, but with differences in the concentration of individual amplified fragments. In the study of Belgorudova et al. (2012), the results for the primer 2080 were similar as those reported by Lapiña et al. (2012) when analyzing the variability

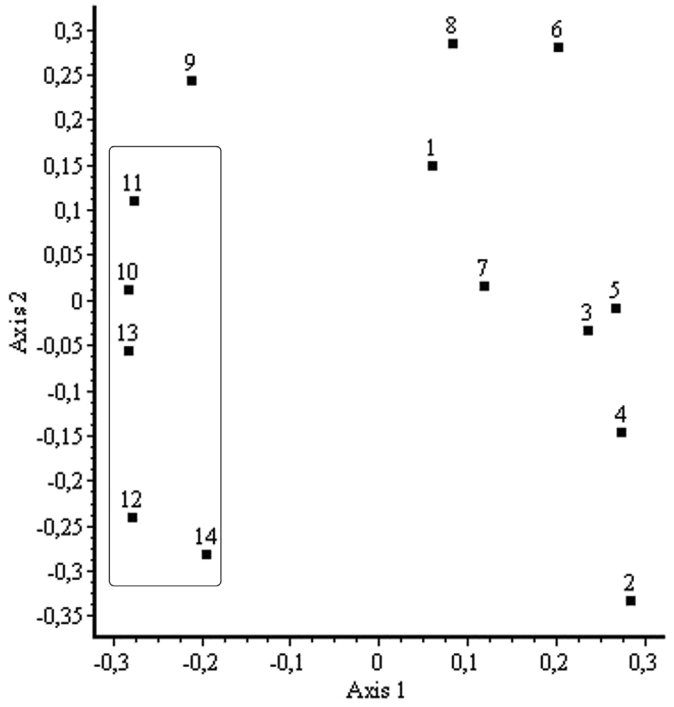

Fig. 2. Scattergram of yacon accession constructed for the provenience-specific iPBS markers. Codes for the accessions are as in the Table 1 and genotypes of Ecuador provenience are marked in rounded rectangle.

of *Liparis loeselii*. Eighteen of the tested iPBS specific primers also gave no results, among them 2080. Only three of tested iPBS markers were selected as useful for examination of genetic diversity in Latvian populations of *L. loeselii*.

A UPGMA scattergram was constructed based on the binary matrix that corresponds to the presence or absence of the separated PCR products in the agarose gels for each of the used primers (Fig. 2). Five of the primers were not connected to some specific banding pattern (primers 1846, 1880, 1845, 1875 and 2080). Two of the used iPBS markers (2078 and 1886) provide patterns where Peru accessions were grouped together according their iPBS amplification profiles (Table 1; Fig. 2).

In silico alignment analysis was performed for the used iPBS primers based on the results of unspecific or provenience-specific manner of yacon accession grouping. All of them were BLASTed against Asteraceae (taxid:4210) in the NCBI (Zhang et al., 2000). Different results were returned that ranged from no matching up to 19 different species matching (Table 3). When comparing the primers that have grouped the yacon accessions in the manner of provenience specificity,

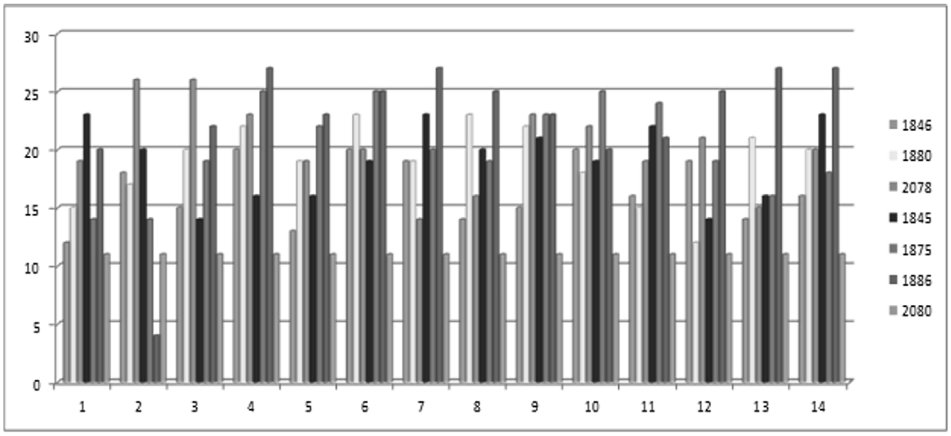

Fig. 1. Number of amplified iPBS loci for individual primers used in the study.

Table 3
Analysis of the used iPBS primers alignment against Asteraceae.

Primer name	Retrotransposon(-like) sequences matched in the alignment	Species matched in the alignment
1846	Ty1-copia-like reverse transcriptase (RT) gene; Ty3/gypsy-like reverse transcriptase-like gene	*Hypochaeris - tenuifolia; sessiliflora; palustris; microcephala; megapotamica; maculata; eremophila; elata; chondrilloides; chillensis; caespitosa; argentina; apargioides; angustifolia; acaulis Helianthus petiolaris; paradoxus; deserticola*
1880	none	none
2078	none	none
1845	Tdv1 gene for CACTA transposable element	*Dahlia pinnata*
1875	none	none
1886	Ty1-copia retrotransposonnonfunctionalpolyprotein gene; Ty1-copia-like reverse transcriptase (RT) gene; Ty3/gypsy-like reverse transcriptase-like gene	*Hypochaerischondrilloides; Helianthus annuus; Helianthus petiolaris; Helianthus paradoxus; Helianthus deserticola; Helianthus anomalus;*
2080	Ty3/gypsy reverse transcriptase-like gene	*Helianthus deserticola*

the results are in concordance with the results for all the tested iPBS primers. No matching for Asteraceae was returned in the case of 2078. Four species of *Helianthus* L. and *Hypochaeris chondrilloides* (A.Gray) Cabrera were returned with the sequences of both Ty1-copia-like and Ty3/gypsy-like reverse transcriptase-like gene for 1886 marker.

Most of retrotransposons are nested, mixed, inverted, or truncated in chromosomal sequences (Kalendar, 2011). Once the iPBS is proved to be suitable for the analysis of the particular species, this means that retrotransposons of a specific class are located near to the other. This is important for LTR retroptransposons mainly, where it subsequently allows the use of such a primer sequence for cloning of new LTR in the analyzed species. The banding patterns obtained in the iPBS profile follow the natural abundance variability of retrotransposon families and their distribution (Kalendar et al., 2010) and provide an excellent marker for the analysis of molecular variability. Retrotransposon markers have been applied successfully in many plant species (Tao et al., 2005; Branco et al., 2007; Lou and Chen, 2007; Belgorudova et al., 2012; Trebichalský et al., 2013). When comparing iPBS to other marker systems that were successfully applied for yacon genome analysis (Milella et al., 2011; Svobodová et al., 2013), iPBS is very similar to the manner of action of RAPD marker system, but its amplification is stringent and the reproducibility is superior (Mehmood et al., 2013). The iPBS markers provide on average 15 to 50 bands that are the length of the whole range of standard polymerases (Kalendar et al., 2010). A scattergram based on Jaccard coefficients of genetic similarity with four main clusters was produced for the analyzed yacon accessions and it showed a close relationship between them, which is in concordance with its vegetative manner of reproduction. This result pointed the potential of iPBS markers for the purpose of yacon germplasm discrimination and characterization, especially because of its vegetative reproduction (Al-Najm et al., 2016) and different ploidy level. From all the up-to-date studies on yacon germplasm it is concluded that the genetic variability of yacon is low.

iPBS can be combined with the other markers to provide a very good base for the whole genome characterization of yacon germplasm when finding suitable and yacon-specific primers. Svobodová et al. (2013) reported 97.3% of polymorphic band for the ISSR markers when applied for the yacon clones in the study, where not only yacon but its three relative species were analyzed.

The highest value of Jaccard genetic similarity coefficient for yacon accessions in the study was 0.24 with an average of 0.056. This supports the results of low genetic variability for vegetative manner reproduction of yacon. Mansilla et al. (2006) collected 30 cultivated accessions of yacon from Peru and scored them for the RADP variability, and Milella et al. (2005) scored five different origin accessions for RAPD variability in pilot studies of yacon germplasm analysis. Using the thirty-four primers, Mansilla et al. (2006) reported only 30.7% polymorphism and using 61 primers by Milella et al. (2005), only 28.7% polymorphism was reported. In both studies, all the accessions were discriminated. Milella et al. (2011) compared AFLP and RADP markers for five different yacon accessions, but the AFLP provided even lower polymorphism (23.4%) compared to the RADP ones. That is why a higher polymorphic technique is suitable to prove and establish successful yacon genetic resources management. Some of the landraces used in this study were used by Russo et al. (2015) to investigate different biological properties such as anticholinesterase activity, antioxidant and antidiabetic characteristics, and phytochemical profiles. In the study of those authors, the compound of 1,5-O-dicaffeoylquinic acid was identified in yacon for the first time. Similar to the iPBS profile, a different profile that was higher than in other yacon landraces was obtained for phenylpropanoid derived in the study of the same authors.

3.2. Chemical diversity based on metabolomic fingerprinting

The UPLC-HRMS/MS-based metabolomic fingerprinting of crude methanol extracts (70%) from 21 yacon accessions allowed assessment of the chemical diversity of this species from individuals grown under the same conditions. Although the analyzed yacon accessions originally came from different geographic localities (Table 1), they display highly similar metabolomic fingerprints characterized by the accumulation of high quantities of sesquiterpene lactones and diterpenes (Fig. 3, Table 4). As observed in the total ion current (TIC) chromatogram of one representative sample from each locality (Peru, Ecuador and Bolivia, Fig. 3), a few major peaks appear common to all samples, suggesting a rather low chemical diversity between the different germplasm accessions of yacon; this is in accordance with previous reports, as well as the semi-domesticated nature and clonal propagation of this species (Mercado et al., 2014). Twenty-one compounds common to most of the analyzed samples were identified (Table 4) with different levels of confidence according to the metabolomics standards initiative (MSI) (Sumner et al., 2007). The identified metabolites belong to five major chemical classes: free or esterified organic acids, glycosylated metabolites, flavonoids, sesquiterpene lactones (STLs), and diterpenes. The chemical structures of the identified metabolites are presented as supplementary data.

Among the free organic acids, peaks 2 and 3 (Fig. 3) were identified based on accurate mass measurements, MS^2 data and database comparisons. Peak 2 (Fig. 3), identified as quinic acid, was unambiguously reported by spectral comparisons with a reference substance corresponding to MSI level 1, while peak 3 (MSI level 2, Fig. 3) was tentatively identified as altraric acid (Table 4). This compound (altraric acid) has been previously reported in *S. sonchifolius* where it occurs in its free form or esterified with one or more caffeoyl units in different positions (Takenaka et al., 2003). This is the case of peak 6 (Fig. 3), identified as a putative di-caffeoylaltraric acid isomer based on its fragmentation pattern, characterized by the neutral loss of two caffeoyl units and a

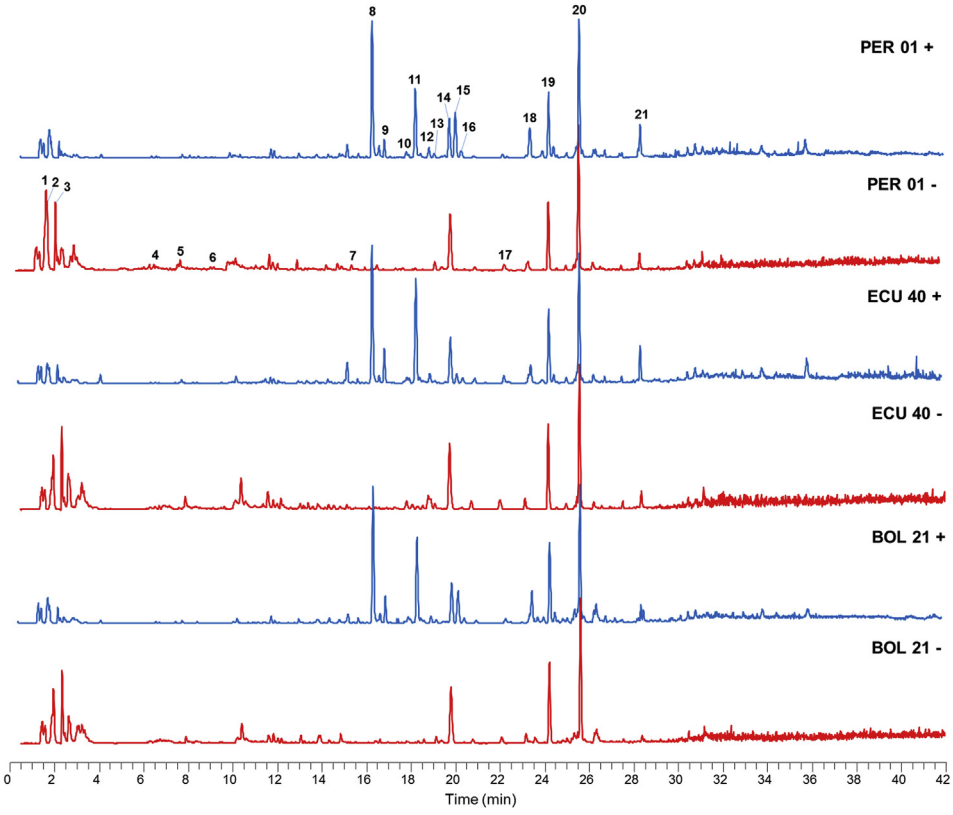

Fig. 3. TIC chromatograms in the positive (blue) and negative (red) ionization mode of three yacon accessions from different geographic origins. Numbers above peaks correspond to the identified metabolites reported in Table 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

base peak at 209.02910 *m/z* in the negative mode of ionization, representing a deprotonated molecule of altraric acid (Table 4) (Takenaka et al., 2003).

Three glycosylated metabolites were tentatively identified in all plant samples (Table 4). Peak 1 (Fig. 3) showed a pseudomolecular ion at 341.10794 *m/z* and a neutral loss of 162.05455 Da in the negative mode (Table 4), corresponding to the loss a hexose unit. Therefore, this compound was tentatively identified as a putative disaccharide with the following molecular formulae: $C_{12}H_{22}O_{11}$. Database searches in the DNP with the accurate mass and molecular formulae obtained for this peak afforded 170 possible hits, all of them disaccharides. Peak 5 (Fig. 3) ($C_{17}H_{30}O_{10}$) was proposed as a putative henexyl diglycoside based on database searches and MS^2 data. This compound showed a neutral loss of 132.04218 Da in the negative mode (Table 4), representing the putative loss of a pentose unit and accurate mass searches in the DNP afforded seven hits for 3-henexyl diglycosides. A similar fragmentation pattern was observed for peak 4 (Fig. 3, Table 4) ($C_{18}H_{26}O_{10}$), proposed as a benzyl diglycoside, for which seventeen hits were obtained in the DNP. Although these two compounds (3-henexyl diglycoside and benzyl diglycoside) have not been previously described in *S. sonchifolius*, their aglycone form (3-hexenol and benzyl alcohol, respectively) corresponds to common volatile alcohols involved in plant-plant communication previously detected in several angiosperm plant families, including Asteraceae (Sugimoto et al., 2016). These volatiles, usually emitted in response to herbivory, are taken up by neighbouring uninfested plants and subsequently undergo glycosylation and glutathionylation, resulting in their conversion to non-volatile compounds that have ecological functions (Sugimoto et al., 2016).

The identity of one flavonoid (peak 7, Fig. 3) was proposed based on accurate mass comparisons with data reported in the AsterDB and spectral comparisons with literature data. This compound, identified as hesperetin, showed a pseudomolecular ion at 301.07101 *m/z* in the negative ionization mode, and two fragment ions at 165.01829 *m/z* and 135.04392 *m/z* (Table 4), which is in accordance with that reported by Tahri et al. (2016). Although this compound has not been reported in *S. sonchifolius*, it is a common metabolite in other Asteraceae species, including members from Espeletiinae (Padilla-González et al., 2017, 2018), a closely related subtribe to the genus *Smallanthus* (Rauscher, 2002; Diazgranados and Barber, 2017). Interestingly, in our experimental conditions we did not detect any of the commonly reported phenolics and flavonoids in yacon (e.g. mono- and di-caffeoylquinic acids and quercetin derivatives). However, based on previous analyses (Padilla-González, unpublished), we believe that the absence of these compounds could be related to the lack of direct solar radiation in the early stages of plant development; this raises concerns about standardizing the cultivation of yacon in the conditions of central Europe to ensure a high concentration of bioactive metabolites, among which flavonoids and phenolics are key substances (Genta et al., 2010; Oliveira et al., 2013; Andrade et al., 2014).

As observed in Fig. 3, most of the peaks detected in high concentrations in the aerial parts of yacon correspond to sesquiterpene lactones (Table 4). This class of metabolites is commonly reported in *S. sonchifolius* and all the identified STLs have been previously reported in this species (Schorr et al., 2007; Mercado et al., 2014). Nine compounds (peaks 8–16, 18, Fig. 3) were identified based on database searches, MS^2 data and spectral comparisons with reference substances. The identity of peak 8 (Enhydrin), 11 (Uvedalin), and 15 (Longipilin acetate) (Fig. 3) was unambiguously reported by spectral comparisons with reference substances, while the identity of peaks 9, 10, 12–14, 16 and 18 (Fig. 3) was proposed based on MS^2 data and chemotaxonomic information (Table 4). All the identified STLs share a common melampolide backbone (Table 4) with a double bound or an epoxyde group between C-4 and C-5 and different ester moieties attached to C-8, C-9 and C-10. Therefore, the identified melampolides share a similar fragmentation pattern characterized by initial losses of the ester moieties and a sequentent peak representing the melampolide backbone, usually at 229.0859 *m/z* or 213.0909 *m/z* (positive mode) depending on the presence of the epoxyde group or double bound between C-4 and C-5, respectively (Table 4). In all cases, enhydrin (peak 8, Fig. 3) was detected as the major STL followed by uvedalin (peak 11, Fig. 3), which is

Table 4
Identified substances in 21 yacon (*Smallanthus sonchifolius*) accessions analyzed by UPLC-HRMS/MS.

#	Rt (min)	Substance	Negative ionization	Negative MS/MS	Positive ionization	Positive MS/MS	Samples	Confidence level[a]
			pseudomolecular ion (*m/z*)		pseudomolecular ion (*m/z*)			
1	1.47	Putative disaccharide ($C_{12}H_{22}O_{11}$)	$[M - H]^-$ 341.10794	341.10867, 179.05412, 89.02292bp	$[M + Na]^+$ 365.10550	–	All	3
2	1.50	Quinic acid[b]	$[M - H]^-$ 191.05504	191.05519, 85.02807	–	–	All	1
3	1.91	Altraric acid	$[M - H]^-$ 209.02916	209.02931, 191.01860, 133.01280, 85.02800bp	–	–	All	2
4	6.15	Benzyl diglycoside	$[M - H]^-$ 401.14447	401.14432, 269.10242bp, 161.04424, 101.02292 71.01235	$[M + Na]^+$ 425.14178, $[M + NH4]^+$ 420.18643, $[M + H]^+$ 403.15930	–	All	2
5	7.53	3-hexenyl diglycoside	$[M - H]^-$ 393.17596	393.17554, 261.13336, 149.04416, 101.02298 89.0229bp 71.01235	$[M + Na]^+$ 417.17319, $[M + NH4]^+$ 412.21780, $[M + H]^+$ 395.19128	–	All	2
6	8.93	di-caffeoylaltraric acid isomer	$[M - H]^-$ 533.09210	371.06085, 209.02910bp, 191.01852, 85.02795	–	–	PER01,03–15, ECU40-45,BOL21,24	2
7	15.25	Hesperetin	$[M - H]^-$ 301.07111	165.01804, 135.04376bp	–	–	All	2
8	16.14	Enhydrin[b]	–	–	$[M + H + ACN]^+$ 506.20227, $[M + Na]^+$ 487.15701, $[M + H]^+$ 465.17517	405.15439, 349.12817, 289.10703bp, 229.08594	All	1
9	16.70	Polymatin A	–	–	$[M + Na]^+$ 413.15701, $[M + H]^+$ 391.17514	373.16446, 291.12259, 273.11206bp, 241.08580, 213.09093	All	2
10	17.83	Fluctuadin	–	–	$[M + Na]^+$ 457.14679, $[M + H]^+$ 435.16495	375.14386, 289.10690, 257.08066, 225.14845bp	All	2
11	18.15	Uvedalin[b]	–	–	$[M + H + ACN]^+$ 490.20724, $[M + Na]^+$ 471.16254, $[M + H]^+$ 449.18048	389.15915, 273.11200bp, 241.08575, 213.09087	All	1
12	18.75	Sonchifolin or isomer	–	–	$[M + Na]^+$ 397.16220, $[M + H]^+$ 375.18011	275.12753, 243.10155bp, 215.10660	All	2
13	19.00	Uvedalin aldehyde	–	–	$[M + H]^+$ 419.16998	359.14859, 259.09644bp, 231.10143, 213.09090	All	2
14	19.70	(1*Z*,4*E*)-8â-Angeloyloxy-germacra-1(10),4,11(13)-trien-6á,12-olide-14-oic acid	$[M - H]^-$ 359.14893	–	$[M + H]^+$ 361.16428	261.11209, 243.10153bp, 215.10666	All	2
15	19.97	Longipilin acetate[b]	–	–	$[M + H]^+$ 449.18039	389.15936, 289.10684bp, 257.08060, 229.08583	All	1
16	20.25	Polymatin B aldehyde	–	–	$[M + H]^+$ 403.17520	343.15390, 243.10156bp, 215.10664	All	2
17	21.86	Smaditerpenic acid C or D	$[M - H]^-$ 349.23755	331.22720, 269.22729, 221.15355, 113.02299, 69.03311	–	–	All	2
18	23.37	Polymatin B	–	–	$[M + H]^+$ 433.18591	373.16452, 291.12250, 273.11203bp, 241.08582, 213.09091	All	2
19	24.01	Smaditerpenic acid E	$[M - H]^-$ 377.23221	59.01239bp	$[M + Na]^+$ 401.22989	–	All	2
20	25.43	Putative smaditerpenic acid E derivative ($C_{23}H_{36}O_5$)	$[M - H]^-$ 391.24783	59.01239bp	$[M + Na]^+$ 415.24557	–	All	3
21	28.31	16,17-epoxy-15-*O*-angeloxy-ent-kauran-18-oic acid	$[M - H]^-$ 415.24817	–	$[M + H]^+$ 417.26370	317.21115, 299.20038, 271.20554bp, 253.19507	All	2

bp, base peak.
[a] Based on the Metabolomics Standards Initiative (Sumner et al., 2007).
[b] Reference substance used to confirm identification.

in accordance with literature reports (Schorr and Da Costa, 2005; Lopes et al., 2013; Mercado et al., 2014). According to previous studies (Mercado et al., 2014), enhydrin usually represents the major STL found in most yacon accessions from Peru, Ecuador, Bolivia and Argentina. However, in such studies three accessions from Peru displayed a different STL profile with uvedalin as the major STL and the enhydrin chemotype was subdivided into three different subtypes: uvedalin, fluctuadin, and uvedalin aldehyde, related to the second major STL accumulated in the studied samples. In our case, all samples displayed a highly consistent STL profile dominated by enhydrin and uvedalin, which is probably a consequence of the standardized cultivation and harvesting conditions to which all yacon accessiones were exposed, rather than a consequence of the selection of several samples belonging to a single chemotype-subtype, as some of the studied samples originally come from the same collection sites reported by Mercado et al. (2014). This result opens an intriguing perspective about the possible quantitative variability in the STL profile of yacon according to the plant's growing conditions or developmental stages, which requires further investigations.

Lastly, the identity of four diterpenes was also proposed based mostly on accurate mass comparisons and MS^2 data with those of the diterpenes previously reported in the genus *Smallanthus*. Peaks 17, 19, and 20 (Fig. 3) were suggested as three different smaditerpenic acids (smaditerpenic acid D, E and F, respectively, Table 4), which correspond to acyclic diterpenes commonly found in the genus *Smallanthus* (Mercado et al., 2010). Peak 21 (Fig. 3) was identified as a putative kaurane-type diterpene (16,17-epoxy-15-*O*-angeloxy-*ent*-kauran-18-oic acid, Table 4) based on accurate mass comparisons with metabolites reported in the Asteraceae family.

Considering that the analyzed yacon accessions display highly similar metabolomic fingerprints with quantitative rather than qualitative differences, the peak area of the detected ions in both ionization modes (positive and negative) was used as input data to perform a Hierarchical Clustering Analysis (HCA) and further explore relationships between yacon accessions based on their metabolomic fingerprints similarity.

Results from the HCA revealed a grouping tendency of the 21 yacon accessions in four primary clusters (Fig. 4). Contrary to the results observed in the molecular markers based-dendrogram (Fig. 2), there is no grouping tendency according to the geographical origin of the different yacon accessions based on their chemical compositions. This result suggests that besides the low intraspecific chemical variability of yacon, the observed quantitative variability is not related to geography, confirming the results of Mercado et al. (2014). However, this is probably a distinctive feature of yacon, considering that a similar study found that twenty wild populations of a closely related *Smallanthus* species [*S. macroscyphus* (Baker ex. Mart.) A.Grau], can be grouped into two major clusters according to their STLs contents and that these two clusters are related to geography (Aráoz et al., 2016). In such studies, authors found a high diversity of STL in populations growing in the tropics and a decrease in chemical diversity of STLs in populations growing in less tropical locations towards the south (Aráoz et al., 2016).

4. Conclusions

The current study presents the first genetic and phytochemical diversity between yacon accessions based on iPBS markers and metabolomic fingerprinting. The results showed that despite rather low chemical and genetic diversity, probably due to vegetative nature of yacon reproduction, two iPBS primer pairs were able to separate Peru accessions from those of Ecuador and Bolivia, and proved to be suitable marker technique to analyse yacon germplasm diversity. The UPLC-HRMS/MS-based metabolomic fingerprinting showed highly similar metabolic fingerprints among all accessions of yacon and allowed the annotation of twenty-one compounds, some of them previously undescribed in yacon. The identified metabolites belong to five major chemical classes: free or esterified organic acids, glycosylated metabolites, flavonoids, sesquiterpene lactones (STLs) and diterpenes.

Contributions

JŽ, GFPG and EF conceived and designed the experiments. JŽ, GFPG performed the experiments. JŽ, GFPG, EF and IV analyzed the data. JŽ, GFPG and EF wrote the manuscript. IV and EF reviewed the manuscript and coordinated the submission processes. All authors read and approved the final version of the paper.

Conflicts of interest

Authors declare no conflict of interest.

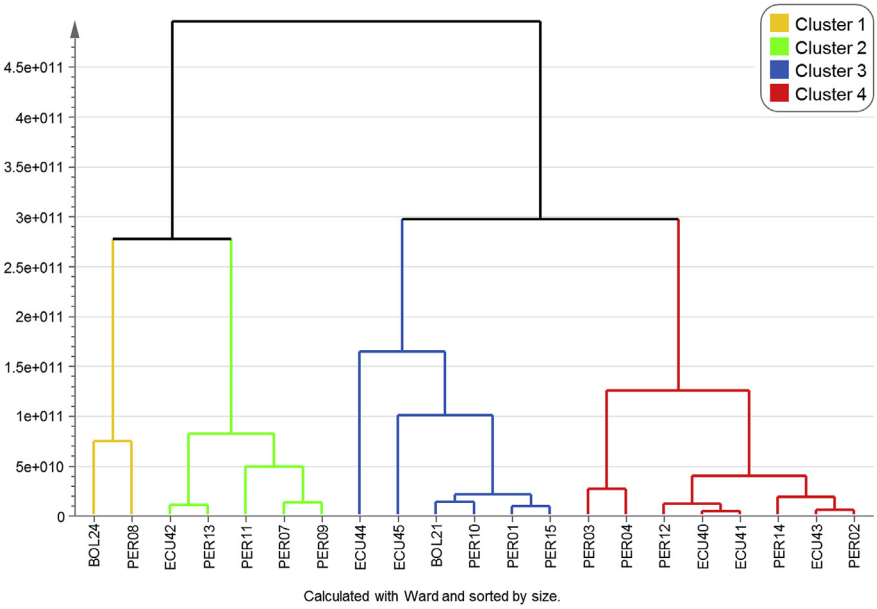

Fig. 4. Hierarchical clustering analysis of 21 yacon accessions based on metabolomic fingerprinting of crude methanol extracts by UPLC-HRMS in the positive and negative ionization modes.

Funding

This work was co-funded by Ministry of Education, Science, Research and Sport of the Slovak Republic under project No 26220220180: Building Research Centre "AgroBioTech" and project No 20185012 of University Internal Grant Agency, Czech University of Life Sciences Prague. The São Paulo Research Foundation (FAPESP, grants No 2014/17702-0, 2014/26866-7 and 2016/21183-4), Brasil, is also acknowledged for financial support of the chemical analyses.

Acknowledgments

Authors are grateful to Prof. Dr. Otmar Spring (Institute of Botany, University of Hohenheim, Germany) for allowing laboratory facilities to prepare plant extracts and to Iris Klaiber (Mass Spectrometry Unit, Core Facility Hohenheim, University of Hohenheim) for UPLC-HRMS/MS measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2019.05.020.

References

Al-Najm, A., Luo, S., Ahmad, N.B., Trethowan, R., 2016. Molecular variability and genetic relationships of Date Palm (*Phoenix dactylifera* L.) cultivars based on inter-primer binding site (iPBS) markers. Aust. J. Crop. Sci. 10, 732–740.

Andrade, E.F., Leone, R.S., Ellendersen, L.N., Masson, M.L., 2014. Phenolic profile and antioxidant activity of extracts of leaves and flowers of yacon (*Smallanthus sonchifolius*). Ind. Crops Prod. 62, 499–506.

Aráoz, M.V.C., Mercado, M.I., Grau, A., Catalán, C.A.N., 2016. Intraspecific variation of sesquiterpene lactones associated to a latitudinal gradient in *Smallanthus macroscyphus* (Heliantheae: Asteraceae). Chemoecology 26, 143–151.

Balážová, Ž., Trebichalský, A., Gálová, Z., Kalendar, R., Schulman, A., Stratula, O., Chňapek, M., 2014. Genetic diversity of triticale cultivars based on microsatellite and retrotransposon-based markers. J. Microbiol. Biotechnol. Food Sci. 3, 58–60.

Baránek, M., Meszáros, M., Sochorová, J., Čechová, J., Raddová, J., 2012. Utility of retrotransposon-derived marker systems for differentiation of presumed clones of the apricot cultivar Velkopavlovická. Sci. Hortic. 143, 1–6.

Belgorudova, I., Grauda, D., Jakobsone, G., Rashal, I., 2012. Usability of retrotransposone-based molecular marker system to assess genetic diversity of *Liparis loeselii*. Acta Biol. Univ. Daugavp. 12, 40–43.

Branco, C.J., Vieira, E.A., Malone, G., Kopp, M.M., Malone, E., Bernardes, A., Mistura, C.C., Carvalho, F.I., Oliveira, C.A., 2007. IRAP and REMAP assessments of genetic similarity in rice. J. Appl. Genet. 48, 107–113.

De Vos, R.C., Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J., Hall, R.D., 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat. Protoc. 2, 778–791.

Diazgranados, M., Barber, J.C., 2017. Geography shapes the phylogeny of frailejones (Espeletiinae Cuatrec., Asteraceae): a remarkable example of recent rapid radiation in sky islands. PeerJ 5, e2968.

Facanali, R., Colombo, C.A., Teixeira, J.P.F., Ming, L.C., Zucchic, M.L., Marques, M.O.M., 2015. Genetic and chemical diversity of native populations of *Ocimum selloi* Benth. Ind. Crops Prod. 76, 249–257.

Fernández, E.C., Viehmannová, I., Lachman, J., Milella, L., 2006. Yacon [*Smallanthus sonchifolius* (Poeppig & Endlicher) H. Robinson]: a new crop in the central Europe. Plant Soil Environ. 52, 564–570.

Friar, E.A., 2005. Isolation of DNA from plants with large amounts of secondary metabolites. Methods Enzymol. 395, 1–12.

Gailīte, A., Ievinsh, G., Ruņģis, D., 2011. Genetic diversity analysis of Latvian and Estonian *Saussurea esthonica* populations. Environ. Exp. Biol. 9, 115–119.

Genta, S.B., Cabrera, W.M., Mercado, M.I., Grau, A., Catalán, C.A., Sánchez, S.S., 2010. Hypoglycemic activity of leaf organic extracts from *Smallanthus sonchifolius*: constituents of the most active fractions. Chem. Biol. Interact. 185, 143–152.

Guo, D.L., Guo, M.X., Hou, X.G., Zhang, G.H., 2014. Molecular diversity analysis of grape varieties based on iPBS markers. Biochem. Syst. Ecol. 52, 27–32.

Kalendar, R., 2011. The use of retrotransposon-based molecular markers to analyze genetic diversity. Field Veg. Crop Res. 48, 261–274.

Kalendar, R., Antonius, K., Smýkal, P., Schulman, A.H., 2010. iPBS: a universal method for DNA Fingerprinting and retrotransposon isolation. Theor. Appl. Genet. 121, 1419–1430.

Krastanov, A., 2010. Metabolomics – the state of art. Biotechnol. Biotechnol. Equip. 24, 1537–1543.

Lapiņa, L., Grauda, D., Rashal, I., 2012. Using retrotransposon-based molecular markers for analysis of genetic variability in Latvian populations of alfalfa. Acta Biol. Univ. Daugavp. 12, 84–87.

Lopes, A.A., Pina, E.S., Silva, D.B., Pereira, A.M.S., Silva, M.F.G.F., Da Costa, F.B., Lopes, N.P., Pupo, M.T., 2013. A biosynthetic pathway of sesquiterpene lactones in *Smallanthus sonchifolius* and their localization in leaf tissues by MALDI imaging. Chem. Commun. 49, 9989–9991.

Lou, Q., Chen, J., 2007. Ty1-copia retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. Genome 50, 802–810.

Mansilla, R.C., López, C., Blas, R., Chia, W.J., Baudoin, J., 2006. Molecular variability analysis of a peruvian *Smallanthus sonchifolius* (Poepp & Endl) H. Robinson "yacon" collection. Ecol. Apl. 5, 75–80.

Mehmood, A., Jaskani, M.J., Ahmad, S., Ahmad, R., 2013. Evaluation of genetic diversity in open pollinated guava by iPBS primers. Pakistan J. Agric. Sci. 50, 591–597.

Mercado, M.I., Araóz, M.V.C., Grau, A., Catalán, C.A.N., 2010. New alicyclic diterpenic acids from yacon (*Smallanthus sonchifolius*) leaves. Nat. Prod. Commun. 5, 1721–1726.

Mercado, M.I., Coll, Aráoz, M.V., Manrique, I., Grau, A., Catalán, C.A.N., 2014. Variability in sesquiterpene lactones from the leaves of yacon (*Smallanthus sonchifolius*) accessions of different geographic origin. Genet. Resour. Crop Evol. 61, 1209–1217.

Milella, L., Salava, J., Martelli, G., Greco, I., Cusimamani, E.F., Viehmannová, I., 2005. Genetic diversity between yacon landraces from different countries based on random amplified polymorphic DNAs. Czech J. Genet. Plant Breed. (CJGPB) 41, 73–78.

Milella, L., Martelli, G., Salava, J., Fernandez, E., Ovesná, J., Greco, I., 2011. Total phenolic content, RAPDs, AFLPs and morphological traits for the analysis of variability in *Smalanthus sonchifolius*. Genet. Resour. Crop Evol. 58, 545–551.

Oliveira, G.O., Braga, C.P., Fernandes, A.A.H., 2013. Improvement of biochemical parameters in type 1 diabetic rats after the roots aqueous extract of yacon [*Smallanthus sonchifolius* (Poepp.& Endl.)] treatment. Food Chem. Toxicol. 59, 256–260.

Padilla-González, G.F., Diazgranados, M., Da Costa, F.B., 2017. Biogeography shaped the metabolome of the genus *Espeletia*: a phytochemical perspective on an Andean adaptive radiation. Sci. Rep. 7, 8835.

Padilla-González, G.F., Diazgranados, M., Oliveira, T.B., Chagas-Paula, D., Da Costa, F.B., 2018. Chemistry of the subtribe Espeletiinae (Asteraceae) and its correlation with phylogenetic data: an in silico chemosystematic approach. Bot. J. Linean Soc. 186, 18–46.

Rauscher, J.T., 2002. Molecular phylogenetics of the *Espeletia* complex (Asteraceae): evidence from nrDNA ITS sequences on the closest relatives of an Andean adaptive radiation. Am. J. Bot. 89, 1074–1084.

Russo, D., Valentão, P., Andrade, P.B., Fernandez, E.C., Milella, L., 2015. Evaluation of antioxidant, antidiabetic and anticholinesterase activities of *Smallanthus sonchifolius* landraces and correlation with their phytochemical profiles. Int. J. Mol. Sci. 16, 17696–17718.

Sabot, F., Schulman, A.H., 2006. Parasitism and the retrotransposon life cycle in plants: a hitchhiker's guide to the genome. Heredity 97, 381–388.

Schorr, K., Da Costa, F.B., 2005. Quantitative determination of enhydrin in leaf rinse extracts and in glandular trichomes of *Smallanthus sonchifolius* (Asteraceae) by reversed-phase high-performance liquid chromatography. Phytochem. Anal. 16, 161–165.

Schorr, K., Mefort, I., Da Costa, F.B., 2007. A novel dimeric melampolide and further terpenoids from *Smallanthus sonchifolius* (Asteraceae) and the inhibition of the transcription factor NF-kB. Nat Prod. Commun. 2, 367–374.

Silveira, D.G., Amorim, E.P., Jesus, O.N., Souza, F.V.D., Pestana, K.N., Santos, V.J., Santana, J.R.F., 2009. Genetic variability estimated among caroá populations through RAPD markers. Pesqui. Agropecu. Bras. 44, 283–290.

Simonovska, B., Vovk, I., Andresek, S., Valentová, K., Ulrichová, J., 2003. Investigation of phenolic acids in yacon (*Smallantus sonchifolius*) leaves and tubers. J. Chromatogr., A 1015, 89–98.

Smýkal, P., Bačová-Kertészová, N., Kalendar, R., Corander, J., Schulman, A.H., Pavelek, M., 2011. Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. Theor. Appl. Genet. 122, 1385–1397.

Štefúnová, V., Bežo, M., Žiarovská, J., Ražná, K., 2015. Detection of the genetic variability of amaranthus by RAPD and ISSR markers. Pakistan J. Bot. 47, 1293–1301.

Sugimoto, K., Matsui, K., Takabayashi, J., 2016. Uptake and conversion of volatile compounds in plant–plant communication. In: Blande, J., Glinwood, R. (Eds.), Deciphering Chemical Language of Plant Communication. Signaling and Communication in Plants. Springer, Cham, pp. 305–316.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.,N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis. Metabolomics 3, 211–221.

Svobodová, E., Dvořáková, Z., Hlásná-Čepková, P., Viehmannová, I., Havlíčková, L., Fernández, E.C., Russo, D., Meza, Z.G., 2013. Genetic diversity of yacon [*Smallanthus sonchifolius* (Poepp. & Endl.) H. Robinson] and its wild relatives as revealed by ISSR markers. Biochem. Syst. Ecol. 50, 383–389.

Tahri, W., Chatti, A., Romero-González, R., López-Gutiérrez, N., Frenich, G.A., Landoulsi, A., 2016. Phenolic profiling of the aerial part of *Chrysanthemum trifurcatum* using ultra high performance liquid chromatography coupled to Orbitrap high resolution mass spectrometry. Anal. Methods-UK 8, 3517–3527.

Takenaka, M., Yan, X., Ono, H., Yoshida, M., Nagata, T., Nakanishi, T., 2003. Caffeic acid derivatives in the roots of yacon (*Smallanthus sonchifolius*). J. Agric. Food Chem. 51, 793–796.

Tao, N.G., Xu, J., Cheng, Y.J., Hong, L., Guo, W.W., Yi, H.L., Deng, X.X., 2005. Isolation and characterization of Copia-like retrotransposons from 12 sweet orange (*Citrus sinensis* Osbeek) cultivars. J. Integr. Plant Biol. 47, 1507–1515.

Trebichalský, A., Kalendar, R., Schulman, A., Stratula, O., Gálová, Z., Balážová, Ž., Chňapek, M., 2013. Detection of genetic relationships among spring and winter triticale (× *Triticosecale* witt.) and rye cultivars (*Secale cereale* L.) by using retrotransposon-based markers. Czech J. Genet. Plant Breed. (CJGPB) 49, 171–174.

Valentová, K., Lebeda, A., Dolezalová, I., Jirovský, D., Simonovska, B., Vovk, I., Kosina, P., Gasmanová, N., Dziechciarková, M., Ulrichová, J., 2006. The biological and chemical variability of yacon. J. Agric. Food Chem. 54, 1347–1352.
Zhang, Z., Schwartz, S., Wagner, L., Miller, W.A., 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. 7, 203–214.
Žiarovská, J., Fernández, E.C., Bežo, M., 2013a. Developing the iPBS strategy for yacon, germplasm evaluation. J. Microbiol. Biotechnol. Food Sci. 2, 1967–1979.
Žiarovská, J., Fernández, E.C., Milella, L., 2013b. A revised ITS nucleotide sequence gives a specifity for *Smallanthus sonchifolius* (Poepp. and Endl.) and its products identification. Genetika 45, 217–226.