ORIGINAL ARTICLE

Phenotypic variability of *Smallanthus sonchifolius* germplasm of Peru

Angel Esteban Santa Cruz-Padilla*, [a], Jorge Luis Vásquez-Orrillo*, [a], Silvia Yanina Rodríguez López[a], Araceli Eugenio Leiva[a], Ricardo Manuel Bardales-Lozano[b], Juan F. Seminario[c] and Hipolito Murga-Orrillo[d]

[a] *Instituto Nacional de Innovación Agraria (INIA), Estación Experimental Agraria Baños del Inca, Subdirección de Recursos Genéticos, Jirón. Wiracocha s/n, 06004, Los Baños del Inca, Perú*

[b] *Universidad Nacional de la Amazonía Peruana (UNAP), Departamento Académico de Suelos y Cultivos, Facultad de Agronomía, Jirón Nauta, 16002, Iquitos, Perú*

[c] *Universidad Nacional de Cajamarca. Facultad de Ciencias Agrarias, Programa de Raíces y Tubérculos Andinos. Av. Atahualpa 1050. C.P. 06003. Cajamarca, Cajamarca, Perú*

[d] *Universidad Nacional Autónoma del Alto Amazonas (UNAAA). Escuela Profesional de Ingeniería Agrónoma, Prol. Libertad 1220 - 1228 - Yurimaguas, Alto Amazonas, Loreto, Perú*

Abstract: *Smallanthus sonchifolius* (yacon) is a functional food native to the South American Andes. Its tuberous root and leaves are the main parts consumed; however, few studies have been carried out on its phenotypic variability. This study aimed to characterize 214 yacon accessions from the Germplasm Bank of the Instituto Nacional de Innovación Agraria (INIA), Peru. Twelve qualitative and seven quantitative variables were used. Accession Y-74 showed the largest leaf dimensions, while Y-28 showed the highest productivity per plant. Multiple correspondence analysis and principal component analysis revealed that the variables propagule color, leaf shape, root pulp color, leaf length and width, root weight per plant, and yield contributed significantly to the discrimination and identification of promising accessions. The geographical grouping of the accessions showed differences between accessions from the north and south of Peru. The qualitative phylogenetic tree showed 12 morphological groups discriminated mainly by leaf morphology and root characteristics, while the dendrogram analysis identified four clusters, with Cluster II standing out with an average yield of 73.5t/ha of tuberous roots. These results are important, as they allowed the identification of promising accessions and useful traits that can contribute to improving productivity and promoting the expansion of yacon cultivation at national and international levels.

Keywords: Germplasm, phenotypic, functional food, yacon, Andes

Introduction

Smallanthus sonchifolius, known as yacon, is a perennial species native to the Andes of South America (Caballero and Colonia, 2018; De Sales *et al*, 2021). It is cultivated from Venezuela to northern Argentina, between 900 and 3,500masl (Huaycho *et al*, 2016). However, its remarkable plasticity has facilitated its adaptation to climates outside the Andes (Seminario *et al*, 2003; Mansilla *et al*, 2006; Wagner *et al*, 2019), in countries such as the Czech Republic, the United States, Brazil (Quaresma *et al*, 2020), New Zealand and Germany (Lachman *et al*, 2007). In the Peruvian Andes, tuberous root yields vary from 8 to 96t/ha depending on the genotype (Santa Cruz and Vásquez-Orrillo, 2023).

Yacon maintains a historical and cultural value, as it has been an important functional food for Andean populations since pre-Columbian times (Huaycho *et al*, 2016; Lopera-Marín, 2020). The roots and leaves have benefits for human health. The roots are usually eaten fresh, and the leaves in infusions (Lebeda *et al*, 2004; Moreira *et al*, 2020). The health benefits of yacon are due to its antioxidant, antimicrobial, hypolipidemic, antidiabetic and even anticancer properties

---

* Corresponding authors: Angel Esteban Santa Cruz-Padilla (asantacruz@inia.gob.pe), Jorge Luis Vásquez-Orrillo (jorge.vasquez.orrillo@gmail.com)

(Baek *et al*, 2018; Myint *et al*, 2019; Adriano *et al*, 2019; Minchola-Castañeda *et al*, 2022). The consumption of extracts from the leaves and tuberous roots of yacon regulates the glycemic state and increases the concentration of insulin in the blood (Santos *et al*, 2017; Ferraz *et al*, 2020); it also contributes to the reduction of body weight (Honoré *et al*, 2018). In this context, the therapeutic benefits of yacon highlight the need to conserve and document the genetic heritage of this species (García *et al*, 2022; Wagner *et al*, 2019).

Yacon germplasm exhibits phenotypic variability among different accessions. Variations in shape, weight and oligofructan content have been revealed in tuberous roots (Valentová *et al*, 2006), and differences in isoenzymes and phenolic content in leaves (Valentová *et al*, 2006; Mansilla *et al*, 2006). Morphological characterization has identified multiple morphotypes and ecotypes of both cultivated and wild yacon (Polanco and García, 2013; Ignacio *et al*, 2017). Genetic diversity analysis using molecular markers has shown distinct groupings among accessions, with the highest diversity observed in central Peru (Mansilla *et al*, 2006). Variations in reproductive biology, including flowering time and pollen viability, have also been reported among accessions (Mansilla *et al*, 2010). Furthermore, studies have found differences in total phenolic content, antioxidant activity and chemical composition among local yacon phenotypes (Lachman *et al*, 2007; Russo *et al*, 2015). This phenotypic variability of yacon makes it a valuable resource for breeding programmes and agroindustrial applications.

In Peru, the Instituto Nacional de Innovación Agraria (INIA) conserves yacon germplasm from 11 regions distributed throughout the Andes, currently counting 214 accessions. This diversity has highlighted the need to update phenotypic characterization studies. In this context, it is hypothesized that this germplasm has significant phenotypic variability, which will allow the identification of accessions with superior agronomic characteristics, suitable for use in future genetic improvement programmes. This study aimed to characterize the phenotypic variability of 214 yacon accessions from the INIA Germplasm Bank, conserved at the Estación Experimental Agraria Baños del Inca, Cajamarca, Peru.

Materials and methods

Plant material

The study was carried out between June 2021 and March 2022. A total of 214 yacon accessions were used, originating from the regions of Piura, Cajamarca, Amazonas, La Libertad, Ancash, Pasco, Junín, Cusco, Ayacucho, Apurímac and Puno (Figure 1 and Supplemental Table 1) and conserved *ex situ* since 1986. These accessions are part of the Andean Roots Germplasm Bank of INIA, Estación Experimental Agraria Baños del Inca, Cajamarca, Peru.
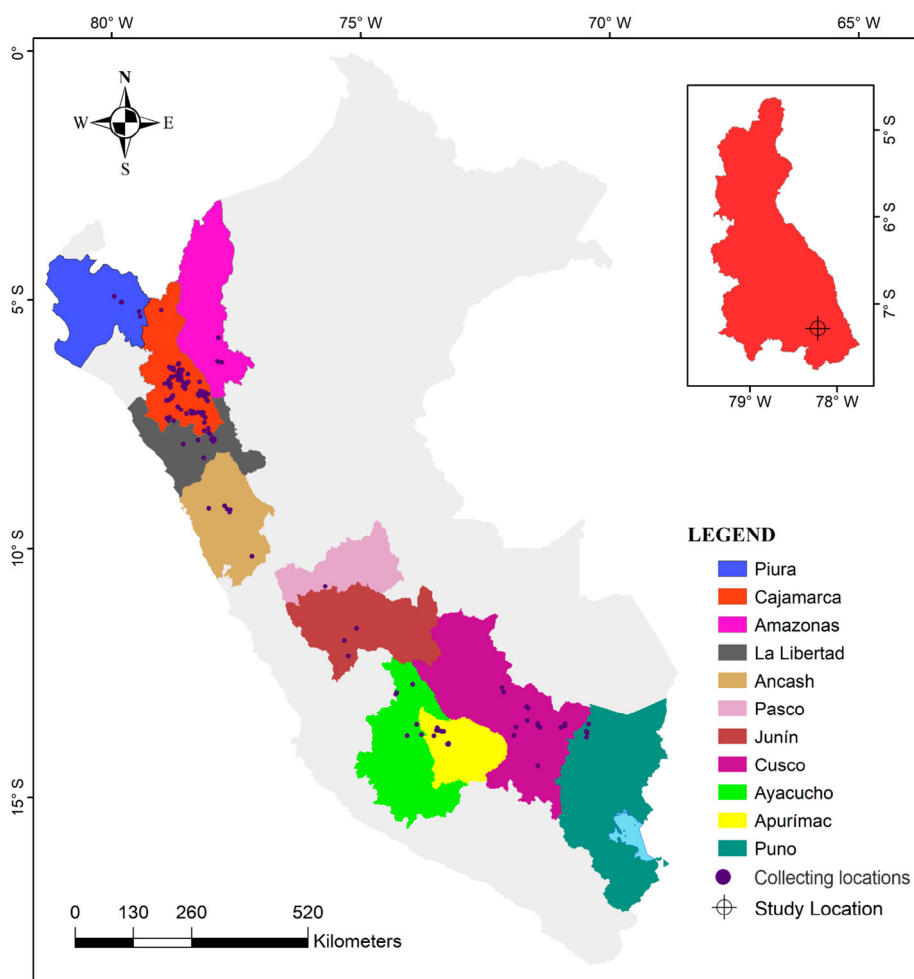

Figure 1. Map of collecting locations of 214 accessions of *Smallanthus sonchifolius* collected in 11 regions of Peru and conserved in the INIA Germplasm Bank.

Experimental area

In the INIA Cochamarca Experimental Annex, located at 7.28137 S, 78.21987 W and 2,835masl, in the province of San Marcos, Cajamarca region (Figure 1). The experiment was installed in an area of 1,796m², prepared to a depth of 30cm, where island guano (1,200kg/ha), diammonium phosphate (164kg/ha) and potassium chloride (73kg/ha) were applied before sowing. Each experimental unit had an area of 6.3m² (0.9m x 7m). Planting was done in a systematic way (each accession in a row), at distances of 0.90m between rows and 0.50m between plants (22,220 plants per hectare). Each experimental unit consisted of 15 plants, 10 plants being evaluated and 5 plants as borders.

Soil analysis of the experimental plot indicated the following: pH of 5.7, organic matter 2.62%, 33.39ppm of phosphorus and 165ppm of potassium, and sandy loam texture (Laboratorio de suelos, aguas, abonos y foliares de la Estación Experimental Agraria Baños del Inca - INIA). During this period of study execution, a mean temperature of 14.3ºC, a minimum of 7.7ºC and a maximum of 20.8ºC were recorded; and the mean monthly rainfall was 117.9mm (SENAMHI, 2024).

Qualitative and quantitative descriptors used in characterization

Twelve qualitative morphological descriptors and seven quantitative ones were used (Table 1), which were assessed over several agricultural campaigns. Colors were defined using the Royal Horticultural Society colour chart (RHS, 2001).

Table 1. Qualitative and quantitative descriptors used in the characterization of yacon accessions

Descriptor type	Descriptor	Acronym	Period of assessment
Qualitative	Secondary stem color and its distribution	SSCD	Preflowering
	Stem branching	SBRA	Preflowering
	Pigmentation of the vein on the underside of the leaf	PVUL	Preflowering
	Leaf blade shape	LBSH	At 50% flowering
	Shape of the leaf base	STLB	At 50% flowering
	Leaf blade edge	LBED	At 50% flowering
	Ray flower shape	RFSH	At 50% flowering
	Petal tooth slit depth	PTSD	At 50% flowering
	Storage root surface color	SRSC	At harvest
	Flesh color of the storage root	FCSR	At harvest
	Clefts in the storage roots	CSRO	At harvest
	Color of the propagules	COPR	At harvest
Quantitative	Number of stems per plant	NSPL	At the end of flowering
	Plant height (cm)	PLHE	At the end of flowering
	Leaf length (cm)	LELE	At 50% flowering
	Leaf width (cm)	LEWI	At the end of flowering
	Weight of storage roots per plant (kg)	WSRP	At harvest
	Number of storage roots per plant	NSRP	At harvest
	Yield of storage roots (t/ha)	YOSR	At harvest

Data analysis

The characterization data were subjected to multivariate statistical analyses. Initially, the overall structure of the dataset was explored using a factor analysis of mixed data (FAMD), which simultaneously integrated both qualitative and quantitative traits. To further investigate specific patterns, a multiple correspondence analysis (MCA) was applied to the qualitative variables, and a principal component analysis (PCA) to the quantitative ones. In addition, associations among the quantitative traits were assessed using Pearson's correlation coefficient.

To define phenotypically differentiated groups among the accessions, a hierarchical cluster analysis was performed, employing Euclidean distance as the dissimilarity measure and the complete linkage method for clustering. The optimal number of clusters was determined by inspecting the resulting dendrograms, selecting the cut-off point that maximized within-group homogeneity. The biological validity of the clusters was confirmed by evaluating their phenotypic coherence based on the descriptors analyzed.

To assess quantitative differences among the defined clusters, mean comparisons of the traits were conducted using Tukey's HSD test ($p < 0.05$). The analyses were conducted using the Factoextra (Kassambara & Mundt, 2020) and FactoMineR (Lê *et al*, 2008) packages for MCA and PCA, respectively. Dendrograms were generated using the cluster (Maechler *et al*, 2021) and circlize (Gu *et al*, 2014) packages, while visualization of results was carried out with ggplot2

(Wickham, 2016) and iTOL: Interactive Tree Of Life (Letunic & Bork, 2024). Mean comparisons were performed using the AgroR package (Shimizu *et al*, 2023). All analyses were conducted in the RStudio statistical software (R Core Team, 2024).

Results

The phenotypic characterization data recorded for the 214 yacon accessions conserved in the INIA Germplasm Bank are presented in Supplemental Table 2. These include 12 qualitative and 7 quantitative traits, which were used to assess phenotypic variability.

Factor analysis of qualitative and quantitative traits

Figure 2 shows a joint analysis using the 'scree plot' and 'variable contribution' on the first two principal components, based on a mixed data set including 19 traits (qualitative and quantitative) analyzed for the 214 accessions. The scree plot (Figure 2A) shows the values of the first ten components, with the first and second components having the highest values adding up to 8.55, out of a total value of 22.53. In the contribution of variables (Figure 2B), those above the mean contribution line stand out: COPR, FCSR, LELE, LEWI, PTSD, SRSC, LBED and SSCD (for explanation of acronyms see Table 1). Of these, COPR and FCSR are the most relevant qualitative variables, while LELE and LEWI are the most prominent among the quantitative variables. Due to these findings, and considering that the traits WSRP and YOSR are relevant for breeding programmes, further analysis was conducted to explore in detail the specific contribution of each type of variable to obtain a more complete understanding of their influence on the dataset.

Qualitative characteristics of yacon

The characters showed a cumulative variability between Dim1 and Dim2 of 19.9%. The COPR and FCSR traits showed the highest similarity in the MCA (Figure 3A), emphasizing their high contribution to the observed differentiation. In addition, a pattern of grouping of the accessions according to their geographical origin was evident, showing that accessions from Piura and La Libertad were associated with those from Cajamarca; similarly, accessions from Puno and Cusco showed proximity. Individual associations were also identified, such as those from Junín and Pasco (Figure 3B).

Hierarchical analysis of qualitative yacon characters

The hierarchical analysis presented in Figure 4 shows the formation of 12 groups based on qualitative morphological traits, which are structured according to vegetative and reproductive characteristics of the accessions, as well as geographical and possibly environmental factors. The groups are mainly distinguished by leaf morphology (shape, base and margin), branching pattern, as well as surface and pulp coloration of the reservoir root and propagules.

The analysis of Table 2 and the information in Supplemental Table 1 shows that the regions with the lowest representation are Junín (9 accessions), Puno (14 accessions), La Libertad (9 accessions) and Pasco (1 accession), with only two clusters for Junín and Puno, three for La Libertad, and one for Pasco. This low frequency is possibly due to the fact that the accessions from these regions exhibited lower morphological diversity, were restricted to specific characteristics, or there was limited availability of accessions for analysis. The presence of accessions from Junín and Puno in only a few clusters suggests reduced variability, which could indicate local adaptations that have occurred over a shorter evolutionary period or under more homogeneous environmental conditions.
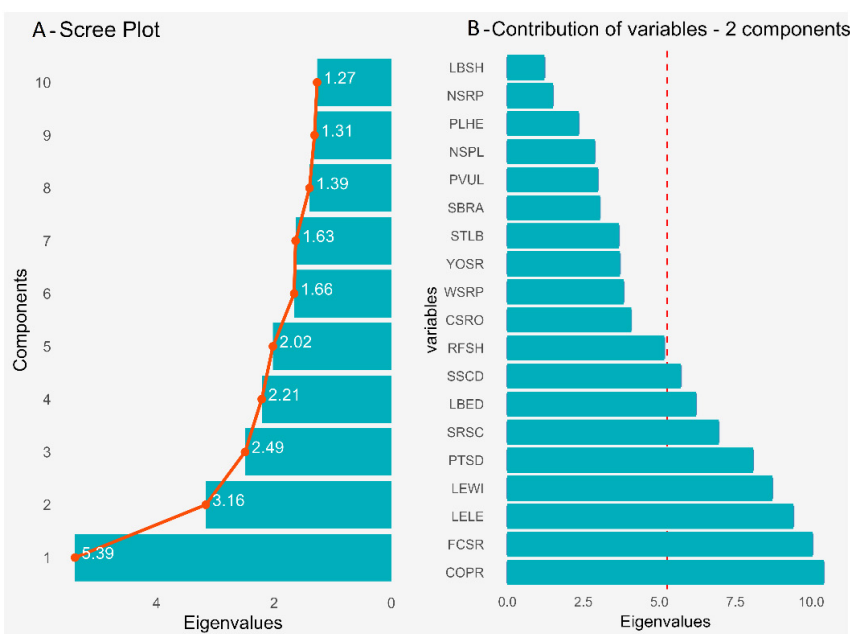

Figure 2. Factor analysis of a mixed data set in 214 yacon (*Smallanthus sonchifolius*) accessions from INIA, Peru. A, Scree plot of ten principal components; B, Contribution of qualitative and quantitative variables. Acronyms are the same as those in Table 1.
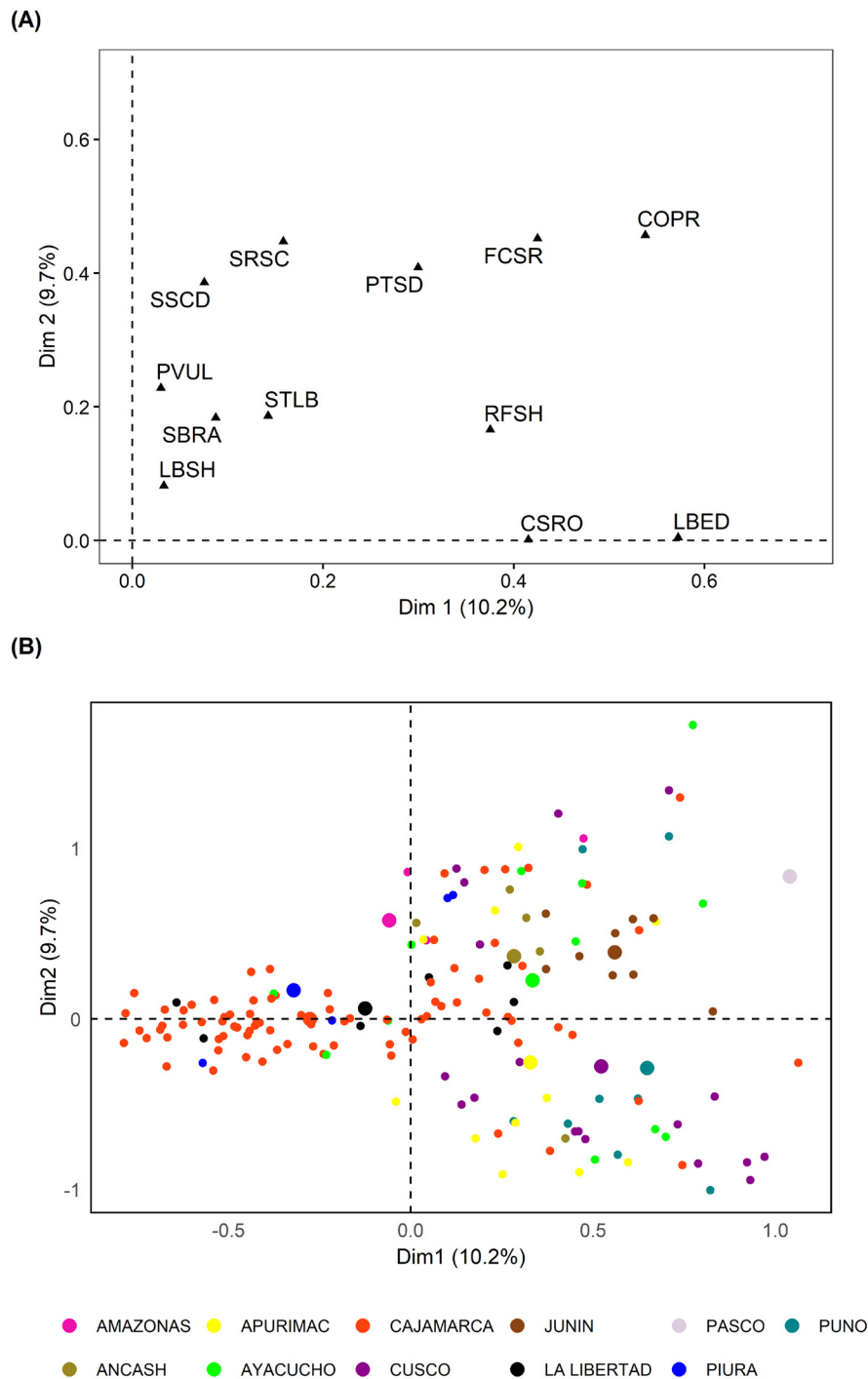
(A)

(B)

Figure 3. Multiple correspondence analysis (MCA) of twelve qualitative traits in 214 yacon (*Smallanthus sonchifolius*) accessions from INIA, Peru. A, Contribution of qualitative variables to the MCA; B, Geographical clustering based on collection locations. Acronyms correspond to those listed in Table 1. Small dots represent individual accessions; larger dots indicate the centroid of each geographic group.

On the other hand, regions such as Cajamarca and Apurímac, with a high frequency of accessions (presence in seven clusters), showed greater morphological diversity. This is reflected in a wide range of characteristics such as leaf shape, flesh color and stem structure, indicating a wide genetic variability and adaptive capacity.

Quantitative characteristics

Plants had between 1 and 5 stems (NSPL), with a mean of 2.71 (Supplemental Table 2). Fourteen accessions reached the minimum value, while three accessions reached the

maximum. In the case of plant height (PLHE), a minimum value of 41.6cm and a maximum value of 164cm were observed, corresponding to accessions Y-206 and Y-154, respectively. On the other hand, accession Y-74 had the longest leaves (LELE) with 25.4cm, while Y-174 recorded the lowest value with 9.6cm. The latter also had the lowest leaf width (LEWI) value with 9.2cm, while the highest value was measured in accession Y-65 with 29cm.

Accessions Y-111, Y-114, Y-137, and Y-182 coincided in the lowest storage root weight (WSRP) with 0.5kg, while accession Y-28 reached the maximum value with 4.7kg, positioning itself as a promising accession due to its high

Table 2. Qualitative morphological group distribution in relation to the main traits of the yacon accessions, as identified through hierarchical cluster analysis.

Cluster	N° of accessions	Regions of provenance	Characteristics
I	74	Cajamarca, Amazonas, La Libertad, Piura	Stem with purple secondary color at nodes and internodes; triangular leaves, with predominant basal branching; surface of storage roots light yellow; yellow-orange pulp; dark greyish purple propagules.
II	40	Apurímac, Ayacucho, Cajamarca, Cusco, La Libertad, Piura	Triangular leaves; elliptic to oblong ligulate flower; variability in flesh color: yellowish white, yellow-orange and light orange.
III	9	Junín	Stem with purple secondary color in internodes; triangular leaves with sagittate base; absence of pigmentation on the veins on the underside of the leaf; light greyish purple storage root surface ; light orange and orange-yellow flesh.
IV	7	Amazonas, Apurímac, Cajamarca, Cusco, Puno, Junín	Secondary color absent on stem; triangular leaves with subhastate base; dark greyish purple storage root surface; white flesh; purplish red and white propagules.
V	15	Amazonas, Ancash, Apurímac, Ayacucho, Cajamarca, Piura	Secondary color absent on stem; triangular leaves; light greyish purple storage root surface; yellowish white or orange-yellow flesh color of the storage roots; purplish red and dark greyish purple propagules.
VI	40	Ancash, Apurímac, Ayacucho, Cajamarca, Cusco, Puno	Stem with purple secondary color at nodes and internodes; triangular leaves with hastate base; light yellow storage root surface, with white flesh mottled with greyish purple or reddish purple tones; white and purplish red propagules.
VII	7	Apurímac, Cajamarca, Cusco	No secondary color on stem; triangular leaves with subhastate base; light greyish purple storage root surface; yellowish white or orange-yellow flesh color of the storage roots; purplish red and dark greyish purple propagules.
VIII	10	Amazonas, Ancash, Apurímac, Ayacucho, Cajamarca, La Libertad	Green stem with purple tinges at nodes and internodes; triangular leaves with subhastate base; light yellow storage root surface; purplish red to dark greyish purple propagules.
IX	6	Cajamarca, La Libertad	Branching along the entire length of the stem; triangular leaves with truncate or subhastate base; light yellow storage root surface; yellow-orange flesh color of the storage roots; white and dark greyish purple propagules.
X	3	Ayacucho, Cusco	Branching along the entire length of the stem; deltoid leaves with truncate base; oblong ligulate flower; light yellow storage root surface; white flesh color of the storage roots; violet-blue propagules.
XI	1	Cajamarca	Stem with purple secondary color at nodes and internodes; basal branching; leaves cordate with lobed base; light yellow storage root surface; light orange flesh color of the storage roots; purplish red with white propagules .
XII	2	Pasco, Ayacucho	Stem with purple secondary color at nodes and internodes; branching along stem; triangular leaves with sagittate base; light greyish purple storage root surface; light orange or yellowish white with reddish purple pits flesh color of the storage roots; purplish red propagules.

productivity. Regarding NSRP, accessions Y-42, Y-103, Y-126, Y-181 presented the lowest number of storage roots (NSRP) with six units, while accession Y-144 reached the maximum production with 32 units. The accessions ranged from 10t/ha to 94t/ha of storage root yield (YOSR). Finally, the coefficients of variation ranged between 18.9% and 45.7%; NSPL, WSRP, NSRP and YOSR were the characters with a variability higher than 30%.

Principal component analysis and correlation

The PCA results (Figure 5) show that the first two principal components together account for 63.80% of the total variance (PC1 = 41.60%; PC2 = 22.20%). Using this analysis, the accessions were grouped into six categories based on the flesh color of the storage root. Accessions with white and
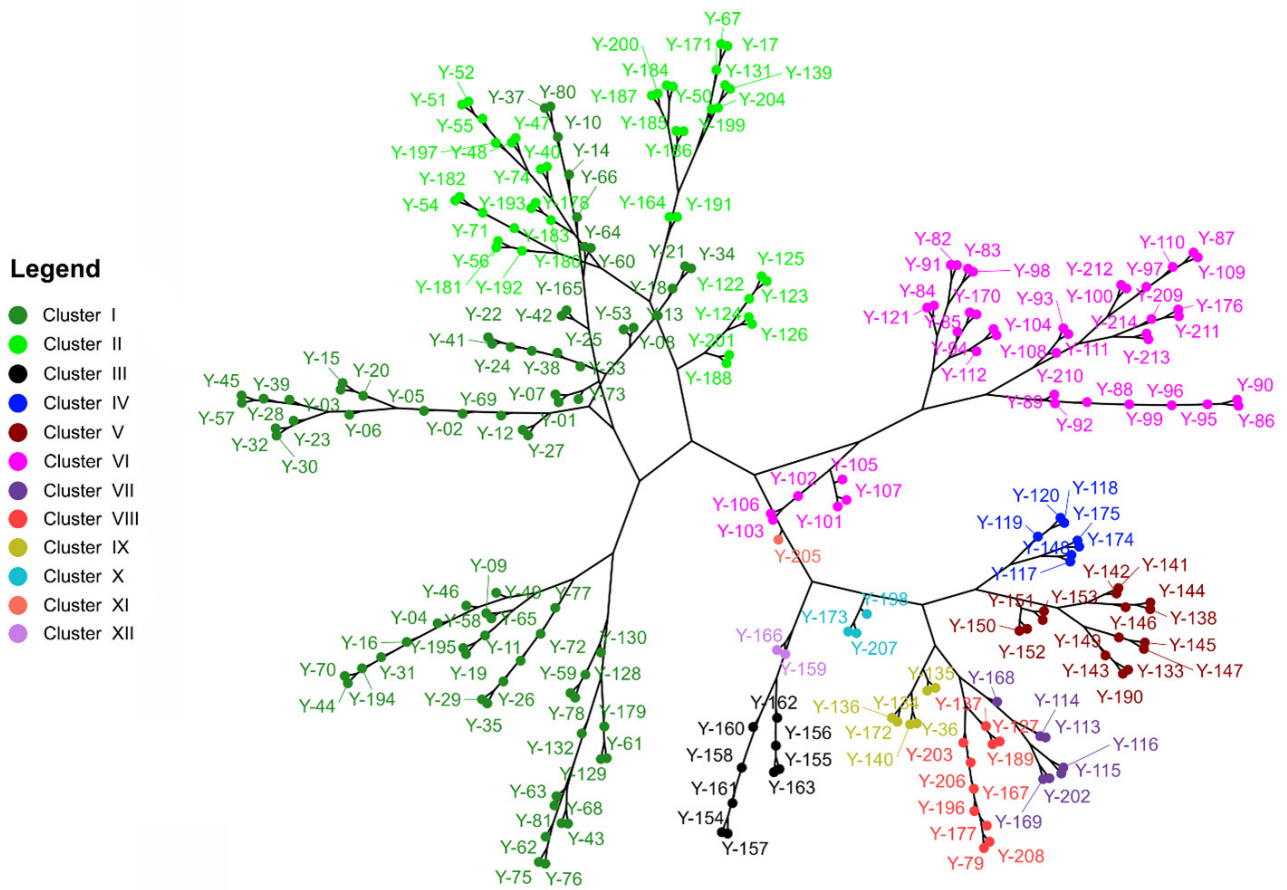
Figure 4. Phylogenetic hierarchical tree of 214 accessions of yacon (*Smallanthus sonchifolius*) based on 12 qualitative characters of the germplasm of INIA, Estación Experimental Agraria Baños del Inca, Cajamarca. Clusters are colour-coded and numbered as in Table 2.
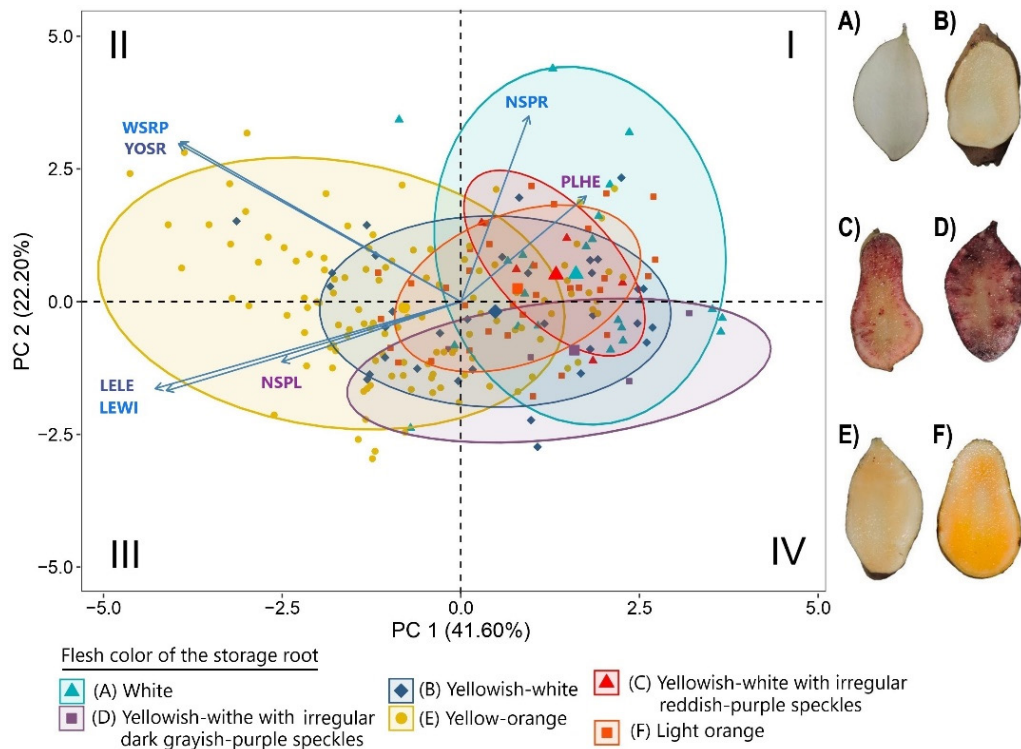

Figure 5. Principal component analysis on 214 accessions of yacon (*Smallanthus sonchifolius*) from INIA, Peru, grouped according to the flesh color of the storage root (with 95% confidence ellipses): A, white; B, yellowish-white; C, yellowish-white with irregular reddish-purple speckles; D, yellowish-white with irregular dark greyish-purple speckles; E, yellow-orange; F, light orange. Acronyms are the same as those in Table 1.

light orange flesh are mainly distributed in quadrants I and IV; accessions with yellowish-white flesh are distributed in the four quadrants; accessions with yellowish-white flesh with irregular reddish-purple speckles are mainly located in quadrant I; accessions with yellowish-white flesh with irregular dark greyish-purple speckles are found in quadrant IV; and accessions with yellowish-orange flesh are distributed in quadrants II and III.

In this case, WSRP and YOSR vectors point towards the upper left quadrant suggesting a strong association of accessions with negative PC1 and positive PC2 values, mainly related to yellow-orange pulp color. On the other hand, PLHE and NSRP show a higher correlation with positive PC1 and PC2 values, being associated with pulp color accessions A, B, C and F.

The variables LELE, LEWI and NSPL correlate with negative values in both PC1 and PC2, suggesting their relationship with accessions in the lower left quadrant.

Correlation between quantitative traits

The correlation matrix for quantitative traits is presented in Table 3. A strong positive correlation was observed between leaf length and leaf width. Both traits also showed moderate positive correlations with the weight of storage roots per plant and with total root yield. The weight of storage roots per plant exhibited a high correlation with total yield. In contrast, the number of storage roots per plant did not show a significant correlation with either root weight or yield.

Plant height was negatively correlated with leaf length, leaf width, and the number of stems per plant. Additionally, the number of storage roots per plant showed negative correlations with leaf dimensions.

Hierarchical analysis of quantitative yacon characters

The circular dendrogram in Figure 6, along with the corresponding information in Table 4, shows a grouping of the accessions in four clusters. Cluster I, with 120 accessions, has a mean of three stems per plant and 116.3cm plant height. This group exhibits a mean leaf length of 20.1cm and leaf width of 20.9cm. The weight of storage roots per plant is 1.8kg, with 13.1 storage roots per plant and a mean yield

of 36.91t/ha. This cluster is characterized by mean values for vegetative development and root production compared to the other clusters.

Cluster II is composed of 19 accessions that together have the most outstanding agronomic characteristics of all the groups. With a plant height of 104.3cm and a mean of three stems per plant, the accessions in this group have the largest leaf dimensions, with 21.1 and 21.9cm length and width, respectively. This cluster is particularly distinguished by its high root productivity, with 3.6kg of storage root weight per plant, 15.2 storage roots per plant, and a mean yield of 73.5t/ha.

Cluster III groups a total of 55 accessions with a plant height of 115.4cm and a mean of 2.3 stems per plant. The leaf dimensions of this group are 16.0 cm leaf length and 15.9 cm leaf width, indicating moderate leaf development compared to Clusters I and II. Furthermore, in comparison with Clusters I, II and IV, it presents the lowest values in storage root weight per plant (1.3kg), number of storage roots per plant (11.6) and yield (26.5t/ha).

Cluster IV includes 20 accessions, which represent 9.3% of the total. These accessions are distinguished by a plant height of 131.5cm, the highest among the clusters. In addition, it has a mean of 1.7 stems per plant; together with a reduced leaf development in length and leaf width, with 14.9 and 14.7cm, respectively. In terms of production, the storage root weight per plant is 1.46kg, with 21 storage roots per plant and a yield of 37.37t/ha. This cluster stands out for its high plant size and a higher number of roots per plant.

Descriptive analysis and comparison of means of quantitative traits of yacon

The analysis of the mean values of the quantitative traits (Table 4) revealed significant differences ($p < 0.05$) among the clusters, indicating a clear structuring of the yacon accessions into four groups with distinct agronomic profiles. These groups enable the identification of materials with potential for different objectives: selection aimed at high yields (Cluster II), balance between growth and productivity (Cluster I), or evaluation of accessions with agronomic limitations that may require specific improvements (Clusters III and IV).

Table 3. Correlation matrix among quantitative characters. Significant correlations at *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; ns: not significant. The character acronyms are the same as those used in Table 1.

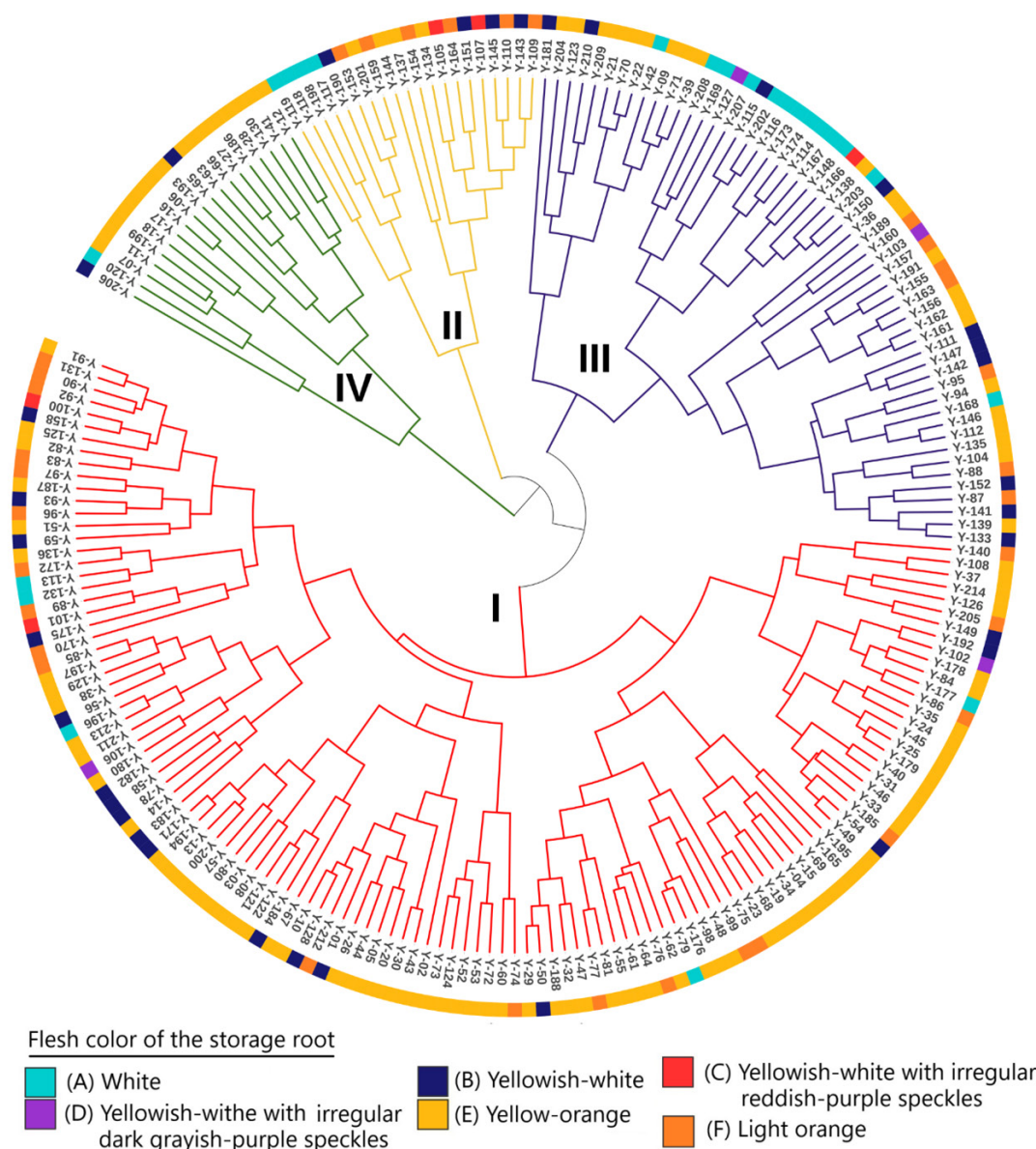| Character	NSPL	PLHE	LELE	LEWI	WSRP	NSRP	YOSR
NSPL	1.00						
PLHE	-0.23 ***	1.00					
LELE	0.34 ***	-0.23 ***	1.00				
LEWI	0.30 ***	-0.18 **	0.90 ***	1.00			
WSRP	0.21 **	-0.10 ns	0.39 ***	0.35 ***	1.00		
NSRP	-0.09 ns	0.22 **	-0.23 ***	-0.24 ***	0.10 ns	1.00	
YOSR	0.20 **	-0.11 ns	0.37 ***	0.35 ***	0.98 ***	0.11 ns	1.00

Figure 6. Circular dendrogram of yacon (*Smallanthus sonchifolius*) germplasm of INIA, Peru

Table 4. Descriptive analysis and comparison of means between quantitative traits of the clusters. The character acronyms are the same as those used in Table 1. SD, standard deviation; CV, coefficient of variation. *, means followed by the same letter in the rows do not differ statistically from each other, by Tukey's test (p < 0.05).

Character	Mean character values* and SD				CV(%)
	Cluster I	Cluster II	Cluster III	Cluster IV	
NSPL	3.00±0.77 a	3.05±0.52 a	2.29±0.71 b	1.75±0.55 c	26.51
PLHE (cm)	116.33±20.86 b	104.31±25.5 b	115.45±28.05 b	131.5±16.97 a	19.78
LELE (cm)	20.15±2.27 a	21.42±3.21 a	16.03±3.65 b	14.93±2.09 b	14.77
LEWI (cm)	20.96±3.34 a	21.99±4.01 a	15.92±4.25 b	14.76±2.72 b	18.82
WSRP (kg)	1.87±0.59 b	3.68±0.56 a	1.34±0.50 c	1.46±0.66 c	30.94
NSRP	13.18±3.52 bc	15.26±3.33 b	11.64±3.27 c	21.05±3.95 a	25.43
YOSR (t)	36.91±11.93 b	73.58±11.11 a	26.58±9.98 c	29.2±13.18 c	31.29

Discussion

Qualitative characteristics

The qualitative traits contributed heterogeneously to the phenotypic variability of 214 yacon accessions (Figure 3A), explaining 19.9 % of the total variability in the first two dimensions of the analysis. Although the contribution was moderate, it was observed that the color of the propagules and flesh color of the storage root traits stand out for their high discriminatory capacity between accessions, indicating their relevance in group differentiation. In particular, the relevance of FCSR is supported by previous studies that identified it as one of the three most important traits for evaluating yacon hybrids (Vegas *et al*, 2015).

The geographical analysis in Figure 3B revealed that phenotypic variability is influenced by the adaptation of the accessions to specific environmental conditions, observing clustering patterns according to their geographical proximity. Accessions from Piura, La Libertad and Cajamarca formed particular and related groups, probably due to ecological, anthropogenic and genetic conditions, suggesting a strong relationship of the accessions with the environment where they thrive (Da Silva *et al*, 2019). This finding is consistent with those obtained by Polanco and García (2013) who noted that yacon genotypes are adapted and specialized to specific agro-ecological conditions.

Complementarily, the hierarchical analysis presented in Figure 4 and detailed in Table 2 provides a more detailed view of the morphological diversity of the accessions, classifying them into 12 qualitative morphological groups according to their vegetative and reproductive characteristics. This grouping reflects the genetic complexity and adaptation of the plants to different ecological conditions. The differences observed in branching patterns, leaf morphology (including shape, base and margin), as well as the coloration of the storage root and propagules, support a grouping based on their phenotypic characteristics.

The Junín and Pasco accessions showed phenotypic characteristics differentiated from the rest, probably due to the influence of unique microenvironments and genetic factors, which would have driven the evolution and differentiation of these accessions. This finding is supported by molecular studies conducted by Mansilla *et al* (2006) and Soto (2012), who identified accessions specific to central and southern Peru, while in the north, they showed greater homogeneity. The results suggest the existence of important centres of diversity for the conservation, genetic improvement and sustainable use of yacon.

Quantitative characteristics

The descriptive analysis of quantitative traits revealed variability among the accessions with evident differences in plant height and yield traits. PCA (Figure 5) showed the greatest contribution of leaf size and yield traits in the phenotypic differentiation of the yacon accessions. A significant proportion of accessions with yellow-orange flesh color were associated with the WSRP and YOSR vectors, showing a phenotypic differentiation centred on the storage root, suggesting that these accessions were adapted

to optimize the accumulation of reserves. This result is congruent with Polanco and García (2013) who determined that yacon has been subjected to anthropogenic selection aimed at obtaining highly productive storage root genotypes.

Accessions with white and light orange flesh were grouped with those exhibiting yellowish-white flesh marked by irregular reddish-purple speckles or irregular dark greyish-purple speckles. This grouping, as observed in Figure 5, shared morphological traits related to PLHE and, to a lesser extent, to NSRP. PLHE was inversely correlated with LELE and LEWI (Table 3). This indicates that accessions with higher plant height had smaller leaf dimensions, while those with lower plant height had larger leaf dimensions (Table 4). In contrast, accessions with yellowish-white flesh exhibited greater dispersion in the four quadrants, indicating greater variability, probably associated with their phenotypic plasticity. Given the relationship of the traits assessed in the PCA, we can select LELE, LEWI, WSRP and YOSR as valuable traits to discriminate accessions within the species.

According to the quantitative traits, the accessions were distributed into four clusters (Figure 6). The analysis of the distribution of accessions suggested that accessions with larger leaf dimensions were associated with higher yields, since a larger leaf area implies greater light uptake, a larger surface area for gas exchange, and greater accumulation of water and nutrients. Consequently, photosynthate production increased, leading to a greater biomass in the storage roots. Leaf dimensions and their relationship with yield have been correlated in other crops such as potato (*Solanum tuberosum* L.) and tomato (*Solanum lycopersicum* L.) (León-Burgos *et al*, 2021), as well as in common bean (*Phaseolus vulgaris* L.) (Warnock *et al*, 2006). These studies suggest that increased photosynthate accumulation in sink organs is related to optimal development of the source organs. However, further studies are required to determine the direct impact of leaf size on yield.

On the other hand, clusters grouping lower-yielding accessions showed limitations in biomass mobilization to storage roots, possibly attributable to vegetative or environmental factors influencing the phenotype. This finding is consistent with Douglas *et al* (2007) who established a significant positive relationship between yield and both planting time and climatic conditions. This observation suggests the need to investigate genotype–environment interaction to identify accessions that maximize the translocation of assimilates to storage roots under different environmental conditions.

Descriptive analysis and comparison of means of yacon quantitative traits (Table 4) provided valuable information on variability and performance of the accessions. These results showed significant relationships between WSRP, YOSR and leaf dimensions. This indicates that selection of promising individuals should focus on clusters with significant and outstanding traits in yield and associated traits (foliage) to maximize productivity in future breeding programmes. This finding coincides with the study by Rodríguez López *et al* (2022), who identified promising genotypes based on their yield and morphological characteristics such as leaf area and number of stems, among others.

Conclusions

The qualitative traits COPR and FCSR, together with the quantitative traits LELE and LEWI, were key determinants in the phenotypic differentiation of the 214 yacon accessions.

Morphological variability exhibited a clear geographical structuring. Accessions from northern Peru (Piura, La Libertad and Cajamarca), the south (Cusco and Puno) and the central region (Junín and Pasco) formed well-defined groupings based on phenotypic similarity.

The hierarchical analysis based on quantitative traits identified Cluster II, comprising 19 accessions, as having the greatest agronomic potential, with an average yield of 73.5t/ha, a storage root weight of 3.6kg per plant, and an average of 15.2 storage roots per plant.

Positive correlations were observed between YOSR and both WSRP, LELE and LEWI, suggesting that foliar development may serve as a reliable predictor of yield performance.

Supplemental data

Supplemental Table 1. Geographical origin and coding of 214 accessions of yacon from the INIA Germplasm Bank, Cajamarca, Peru.

Supplemental Table 2. Agromorphological characterization data of 214 yacon accessions from the INIA Germplasm Bank, Peru.

Acknowledgements

Author contributions

Angel Esteban Santa Cruz-Padilla and Jorge Luis Vásquez-Orrillo: conceptualization, formal analysis, writing – original, research, data curation, resources, methodology, proofreading and editing; Silvia Yanina Rodríguez López and Araceli Eugenio Leiva: research, data curation, resources, methodology, proofreading and editing; Ricardo Manuel Bardales-Lozano and Hipolito Murga-Orrillo: formal analysis, writing – original, research, methodology, proofreading and editing; Juan F. Seminario: research, resources, methodology, proofreading and editing.

Conflict of interest statement

The authors have declared that no competing interests exist.

References

Adriano, L., Dionísio, A., De Abreu, F. Carioca, A., Zocolo, G., Wurlitzer, N., De Oliveira Pinto, C., de Oliveira, A., and De Carvalho Sampaio, H. (2019). Yacon syrup reduces postprandial glycemic response to breakfast: A randomized, crossover, double-blind clinical trial. *Food Research International*, 126, 108682. https://doi.org/10.1016/j.foodres.2019.108682

Baek, S., Choi, N., Lee, K.-P., Jhun, H., and Kim, J. (2018). *Smallanthus sonchifolius* leaf attenuates neuroinflammation. *Journal of Exercise Nutrition & Biochemistry*, 22(2), 31-35. https://doi.org/10.20463/jenb.2018.0014

Caballero, L. and Colonia, A. (2018). Yacón como planta promisoria en el manejo de enfermedades. *Investigaciones Andina*, 20(36), 145-157. https://dialnet.unirioja.es/servlet/articulo?codigo=9359322

Da Silva, D., De Oliveira, F., Quaresma, M., Erlacher, W., and Mendes, T. (2019). Yacon production at different planting seasons and growing environments. *Bioscience Journal*, 35(4), 992-1001. https://doi.org/10.14393/BJ-v35n4a2019-42091

De Sales, R., De Oliveira, E., Xavier, A., De Oliveira, F., Pezzopane, J., Da Silva, D., and Da Silva Berilli, S. (2021). Base temperature, cycle duration, and thermal constant for yacon culture. *Acta Scientiarum. Agronomy*, 44, e52623. https://doi.org/10.4025/actasciagron.v44i1.52623

Douglas, J., Follett, J., Douglas, M., Deo, B., Scheffer, J., Littler, R. and Manley-Harris, M. (2007). Effect of environment and time of planting on the production and quality of yacon (*Smallanthus sonchifolius)* storage roots. *New Zealand Journal of Crop and Horticultural Science*, 35(1), 107-116. https://doi.org/10.1080/01140670709510174

Ferraz, A., Garcia, J., Costa, M., De Almeida, C., Gregolin, C., Alves, P., Hasimoto, F., Berchieri-Ronchi, C., Dos Santos, K., and Corrêa, C. (2020). Yacon (*Smallanthus sonchifolius*) use as an antioxidant in diabetes. En V. R. Preedy (Ed.), Pathology (pp. 379-386). Academic Press. https://doi.org/10.1016/B978-0-12-815972-9.00036-6

García, D., Sotelo, A., Malpica, E., Álvarez, H., Norabuena, E., Gonzáles, T., and Sumarriva, L. (2022). Impacto del helado dietético con yacón (*Smallanthus sonchifolius*) en la hipoglicemia y aceptabilidad. *Nutrición Clínica y Dietética Hospitalaria*, 42(2), 142-149. https://doi.org/10.12873/422garcia

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics* 30(19), 2811–2812. doi: https://doi.org/10.1093/bioinformatics/btu393

Honoré, S., Grande, M., Gomez, J., and Sánchez, S. (2018). Smallanthus sonchifolius (Yacon) Flour Improves Visceral Adiposity and Metabolic Parameters in High-Fat-Diet-Fed Rats. *Journal of Obesity*, 2018(1), 5341384. https://doi.org/10.1155/2018/5341384

Huaycho, H., Aruquipa, R., Mercado, G., Trigo, R., Bosque, H. y Condori, J. (2016). Conocimientos tradicionales en yacón o aricoma (*Smallanthus sonchifolius*) en comunidades de Mocomoco, Coroico e Irupana de La Paz. *Revista de Investigación e Innovación Agropecuaria y de Recursos Naturales*, 3(2), 152-165. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2409-

16182016000200005&lng=es&tlng=es

Ignacio, S., Camarena, F., Baudoin, J. y Blas, R. (2017). Ethno-Botany and in-situ conservation of the genetic diversity of arracacha (*Arracacia xanthorrhiza* Bancroft), yacon (*Smallanthus sonchifolious* H. Robinson), and wild relatives. *Peruvian Journal of Agronomy*, 1(1), 21-31. http://dx.doi.org/10.21704/pja.v1i1.1064

Kassambara, A. and Mundt, F. (2020). Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package Version 1.0.7. url: https://CRAN. R-project.org/package=factoextra

Lachman, J., Fernández, E., Viehmannová, I., Šulc, M. and Èepková, P. (2007). Total phenolic content of yacon (*Smallanthus sonchifolius*) rhizomes, leaves, and roots affected by genotype. *New Zealand Journal of Crop and Horticultural Science*, 35(1), 117-123. https://doi.org/10.1080/01140670709510175

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* 25(1), 1–18. doi: http://dx.doi.org/10.18637/jss.v025.i01

Lebeda, A., Dolezalová, I. and Dolezal, K. (2004). Variation in morphological and biochemical characters in genotypes of maca and yacon. *Acta Horticulturae*, 629, 483-490. https://doi.org/10.17660/ActaHortic.2004.629.62

León-Burgos, A., Beltrán-Cortes, G., Barragán-Pérez, A., and Balaguera-López, H. (2021). Distribution of photoassimilates in sink organs of plants of Solanaceas, tomato and potato. A review. *Ciencia y Agricultura*, 18(3), 79-97. https://doi.org/10.19053/01228420.v18.n3.2021.13566

Letunic, I., and Bork P. (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool, *Nucleic Acids Research*, Volume 52, Issue W1, 5 July 2024, Pages W78–W82, https://doi.org/10.1093/nar/gkae268

Lopera-Marín, J., Angulo-Arizala, J., Murgueitio-Restrepo, E. y Mahecha-Ledesma, L. (2020). Producción de tubérculos y biomasa aérea del yacón, *Smallanthus sonchifolius* (Poepp.) H. Rob. (Asteraceae), para alimentación animal en el trópico alto colombiano. *Livestock Research for Rural Development*. 32 (135). http://www.lrrd.org/lrrd32/8/jjlop32135.html

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., and Hornik, K. (2021). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.2. url: https://CRAN.R-project.org/package=cluster

Mansilla, R., López, C., Blas, R., Chia, J. y Baudoin, J. (2006). Análisis de la variabilidad molecular de una colección peruana de *Smallanthus sonchifolius* (Poepp & Endl) H. Robinson "YACÓN". *Ecología Aplicada*, 5(1-2), 75-80. https://doi.org/10.21704/rea.v5i1-2.320

Mansilla, R., López, C., Flores, M. y Espejo, R. (2010). Estudios de la biología reproductiva en cinco accesiones de *Smallanthus sonchifolius* (Poepp. & Endl.) Robinson. *Ecología Aplicada*, 9(2), 167-175. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-22162010000200012&lng=es&tlng=es.

Minchola-Castañeda, K., Luzuriaga-Tirado, E., Montalvo-Rodríguez, A., Moncada-Carrera, J., Morales-Ibañez, F. y Gil-Reyes, W. (2022). Propiedades beneficiosas del yacón (*Smallanthus sonchifolius*) en la salud. *Más Vita*, 4(3), 87-98. https://doi.org/10.47606/ACVEN/MV0135

Moreira, R., Redko, F., Ulloa, J., Flor, S., Tulino, M., Muschietti, L., and Carballo, M. (2020). Toxicogenetic evaluation of *Smallanthus sonchifolius* (yacon) as a herbal medicine. *Journal of Ethnopharmacology*, 257. 112854. https://doi.org/10.1016/j.jep.2020.112854

Myint, P., Dao, T., and Kim, Y. (2019). Anticancer Activity of *Smallanthus sonchifolius* Methanol Extract against Human Hepatocellular Carcinoma Cells. *Molecules*, 24(17), 3054. https://doi.org/10.3390/molecules24173054

Polanco, M. and García, M. (2013). Caracterización morfológica y molecular de materiales de yacón (*Smallanthus sonchifolius* Poep. & Endl) H. Robinson colectados en la ecorregión Eje Cafetero de Colombia. *Revista de Investigación Agraria y Ambiental*, 4(2), 97-116. https://doi.org/10.22490/21456453.981

Quaresma, M., De Oliveira, F., Amaral, J., Parajara, M., Dalvi, L., and Teixeira, A. (2020). Planting methods and depths for yacon (*Smallanthus sonchifolius*) crops. *Revista Colombiana de Ciencias Hortícolas*, 14(2), 249-256. https://doi.org/10.17584/rcch.2020v14i2.9562

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/

Reis, F., Marques, C., Sales de Moraes, A., and Masson, M. (2021). Effect of processing methods on yacon roots health-promoting compounds and related properties. *Trends in Food Science & Technology*, 113, 346-354. https://doi.org/10.1016/j.tifs.2021.05.010

RHS (2001). The Royal Horticultural Society Colour Chart.

Rodríguez, S., Seminario, A., Vásquez, V. y Seminario, J. (2022). Rendimiento agronómico de ocho cultivares de yacón [*Smallanthus sonchifolius* (Poepp. & Endl.) H. Rob.] del norte peruano. *Siembra*, 9(1). e3630 https://doi.org/10.29166/siembra.v9i1.3630

Russo, D., Malafronte, N., Frescura, D., Imbrenda, G., Faraone, I., Milella, L., Fernandez, E., and De Tommasi, N. (2015). Antioxidant activities and quali-quantitative analysis of different *Smallanthus sonchifolius* [[(Poepp. and Endl.) H. Robinson] landrace extracts. *Natural Product Research*, 29(17), 1673-1677. DOI: https://doi.org/10.1080/14786419.2014.990906

Santa Cruz Padilla, A. E. y Vásquez Orrillo, J. L. (2023). Catálogo de yacón del Banco de Germoplasma del INIA. En Instituto Nacional de Innovación Agraria. Instituto Nacional de Innovación Agraria. https://repositorio.inia.gob.pe/handle/20.500.12955/2237

Santos, K. dos, Bueno, B., Pereira, L., Francisqueti, F., Braz, M., Bincoleto, L., Da Silva, L., Ferreira, A., Nakamune, A. C. de M. S., Chen, C.-Y. O., Blumberg, J., and Corrêa, C. (2017). Yacon (*Smallanthus sonchifolius*) Leaf Extract Attenuates Hyperglycemia and Skeletal Muscle Oxidative Stress and Inflammation in Diabetic Rats. *Evidence-Based Complementary and Alternative Medicine*, 2017(1), 6418048. https://doi.org/10.1155/2017/6418048

Seminario, J., Valderrama, M. y Manrique, I. (2003). El Yacon Fundamentos para el Aprovechamiento de un Recurso Promisorio. https://cipotato.org/wp-content/uploads/2014/07/Yacon_Fundamentos_password.pdf

SENHAMI (2020). Servicio Nacional de Meteorología e

Hidrología del Perú-Datos hidrometeorológicos. url: https://www.senamhi.gob.pe/servicios/?p=estaciones

Shimizu, G., Marubayashi, R., and Goncalves, L. (2023). AgroR: Experimental Statistics and Graphics for Agricultural Sciences. url: https://agronomiar.github.io/AgroRpackage/index.html

Soto Torres J. V. (2012). Evaluación de la diversidad genética de colecciones de *Smallanthus sonchifolius* (Poepp. & Endl.) "Yacón" en el Perú. Tesis de Maestría. La Molina – Perú. url: https://hdl.handle.net/10568/126005

Valentová, K., Lebeda, A., Doležalová, I., Jirovský, D., Simonovska, B., Vovk, I., Kosina, P., Gasmanová, N., Dziechciarková, M., and Ulrichová, J. (2006). The biological and chemical variability of yacon. *Journal of agricultural and food chemistry*, 54(4), 1347-1352. DOI: https://doi.org/10.1021/jf052645u

Vegas, D., Bracamonte, O. y Valladolid, A. (2015). Caracterización morfológica de seis variedades parentales de yacón (*Smallanthus sonchifolius*) y trece cruzas obtenidas de un plan de hibridación. *Revista Peruana de Biología*, 22(2), 175-192. https://doi.org/10.15381/rpb.v22i2.11352

Wagner, M., Kamp, L., Graeff-Hönninger, S., and Lewandowski, I. (2019). Environmental and Economic Performance of Yacon (*Smallanthus sonchifolius*) Cultivated for Fructooligosaccharide Production. *Sustainability*, 11(17), 4581. https://doi.org/10.3390/su11174581

Warnock, R., Valenzuela, J., Trujillo, A., Madriz, P. y Gutiérrez, M. (2006). Área foliar, componentes del área foliar y rendimiento de seis genotipos de caraota1. *Agronomía Tropical*, 56(1), 21-42. url: https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0002-192X2006000100002&lng=es&tlng=es

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. url: https://ggplot2.tidyverse.org