CrossMark

RESEARCH ARTICLE

Assessment of genetic relationships between cultivated arracacha (*Arracacia xanthorrhiza* Bancr.) and its wild close relatives in the area of domestication using microsatellite markers

Eduardo Morillo[1,3] · Steen Randers Knudsen[2] · Gérard Sécond[3]

Abstract Arracacha (*Arracacia xanthorrhiza* Bancr.) is an asexual propagated root crop domesticated in the Andean highlands, which exists naturally with polycarpic and monocarpic forms. Wild *A. xanthorrhiza* are present in the area of domestication and can occasionally be mistaken in the same field for a crop or a weed. To study genetic relationships between cultivated arracacha and the wild forms, we surveyed the diversity of 178 plant samples at 11 microsatellite (SSR) loci. As expected, wild *A. xanthorrhiza* forms showed a significantly higher allelic diversity for all the examined SSR markers. The cultivated pool showed an excess of heterozygosity as opposed to a deficit found in the wild compartment. High *Fst* values and AMOVA analysis suggest that the cultivated variety has genetically differentiated from the wild forms and is more related to the wild polycarpic than to the monocarpic. Both the wild forms were well distinguished from the cultivars. Nevertheless, among a set of F1 experimental hybrids (cultivated×wild polycarpic), some other genotypes were revealed, also being admixed. Our results highlight a large genetic base available in the wild populations of *A. xanthorrhiza* with potential implications for the utilization and breeding of this promising crop.

Keywords *Arracacia* · Arracacha · SSRs · Gene flow · Andean region

Gérard Sécond: Deceased.

✉ Eduardo Morillo
   eduardo.morillo@iniap.gob.ec

1 Instituto Nacional de Investigaciones Agropecuarias (INIAP), Quito, Ecuador

2 Botanical Section, Royal Veterinary and Agriculture University, Copenhagen, Denmark

3 Institut de Recherche pour le Développement (IRD), Montpellier Cedex 5, France

Introduction

Arracacha (*Arracacia xanthorrhiza* Bancr.) is an interesting, little-known crop cultivated in the Andean region of South America. As other native Andean crops, arracacha is cultivated for its edible root, probably since pre-Columbian times (Hermann 1997). Arracacha has been catalogued by many authors as one of the most promising new crops because of its economic value and possibilities for agronomic expansion. For example, arracacha was introduced into the southern states of Brazil at the beginning of nineteenth century and today Brazil exports arracacha to Colombia, Ecuador and Venezuela (Hermann 1997; Lim 2015).

The most recent review of the genus confirms that the genus *Arracacia* is particularly diverse in Mexico (Knudsen 2003), but the wild species most closely resembling the crop arracacha are reported in Peru and Ecuador (Hermann 1997; Knudsen 2003; Blas et al. 2008b). Knudsen (2003) proposed that arracacha belongs to an *Arracacia xanthorrhiza* complex consisting of three varieties, which include: the cultivated *A. xanthorrhiza* var. *xanthorrhiza* and two wild varieties; a polycarpic (*A. xanthorrhiza* var. *andina*); and a monocarpic (*A. xanthorrhiza* var. *monocarpa*). This was later supported by Blas et al. (2008b). The wild *A. xanthorrhiza* varieties or forms are similar in terms of morphology to the cultivated type; however, as argued by

🖄 Springer

Hermann (1997), they are sufficiently distinct to be recognized as wild. Nevertheless, the polycarpic wild type has been proposed as more closely related to the crop (Knudsen 2003; Blas et al. 2008b; Morillo and Sécond 2016).

In Ecuador the most frequently observed arracacha are the polycarpic populations, which have a preference for disturbed habitats such as roadsides and the boundaries of cultivated fields. This variety is also commonly found along roadsides in northern Peru and is frequently associated with ancient human settlements (Valderrama and Seminario 2002; Knudsen 2003). This variety is perennial like the cultivated variety but becomes dormant during the dry period emerging again at the beginning of the rainy season (Hermann 1997; Knudsen 2003). Polycarpic populations share the same ecological and geographical distribution throughout the Inter-Andean valleys, possibly having constant gene flow among them (Valderrama and Seminario 2002; Blas et al. 2008a). In the southern region of Ecuador and in northern Peru the monocarpic type can be found in denser vegetation growing generally in calcareous soils, but it is not as abundant as the polycarpic type. The two varieties can occasionally be found within a few 100 m of each other; however mixed populations, have not been reported (Knudsen 2003; pers. obs). Although cultivated and wild *Arracacia* are rarely observed at the same location, the polycarpic variety can occasionally be found growing as weeds close to the domesticated type (Valderrama and Seminario 2002; Knudsen 2003; Blas 2005). It is not known if spontaneous genetic events occur or if such hybridizations could be a possible source for genetic variation in arracacha (Blas et al. 2008a; Morillo and Sécond 2016).

To establish genetic relationships between cultivated and wild arracacha, we developed a genetic diversity analysis using microsatellite markers of a selection of cultivated and wild *A. xanthorrhiza* samples from Ecuador and Peru, and a set of F1 presumed hybrids between the cultivated and the wild polycarpic form.

Materials and methods

Plant material

A set of 178 samples of cultivated and wild *A. xanthorrhiza* were used in this survey: 72 cultivated accessions, 76 plant samples preliminarily identified as wild polycarpic (from 36 localities), and 30 plant samples preliminarily identified as wild monocarpic (from 10 different localities) (Fig. S1). In addition, we included 13 presumed F1 hybrids obtained from two experimental crosses using as parents two cultivars and two wild polycarpic plants. We also included two samples from the closely related tuberous species *A. incisa*

(Supplementary Table 1). Ecuadorian accessions were conserved in the National Germplasm Bank of INIAP in the Santa Catalina Experimental Station and wild *A. xanthorrhiza* populations were sampled from prospections in the central and southern provinces of Ecuador. The Peruvian and F1 hybrid samples (including the parents) were provided by Knudsen as young leaf tissues dried in silica gel.

SSR genotyping

We used the microsatellite markers reported by Morillo et al. (2004). Genomic DNA was extracted from dried leaf samples using the same method reported. DNA samples were quantified by spectrophotometry and normalized to 5 ng/μl. SSR genotyping was performed on an IR^2 Automated DNA Sequencer (LI-COR, model 4300S, Lincoln, NE, USA) using the following primers: AxD82, AxC27, AxD43, AxD34, AxC38, AxD13, AxC85, AxD72, AxD55, and AxD85. Amplifications were performed in a 10 μl final volume with 20 ng of genomic DNA, 0.02 μM of the M13-labelled primer, 0.3 μM of the reverse primer, 0.06 μM of M13 primer-fluorescent dye IR700 or IR800, 2.5 mM MgCl$_2$, 0.2 mM dNTP, and 0.5 U Taq DNA Polymerase (PROMEGA). A BIOMETRA thermocycler was used with the following cycling conditions: 94 °C (1 min), 30 cycles at 94 °C for 60 s, 45–60 °C for 60 s, 72 °C for 60 s, and a final elongation step at 72 °C for 7 min. IR700 or IR800-labeled PCR products were diluted tenfold in the appropriate loading buffer, subjected to electrophoresis in a 25 cm 6.5% polyacrylamide gel, and then sized by the IR fluorescence scanning system of the LI-COR. Scoring of allele size was performed using the SAGA GT™ software (LICOR Biotech) using a 50–350 bp ladder as a reference (Cat. No. LI-COR 829–05343/44).

Data analysis

We developed our genetic analysis under the assumption of diploidy of *A. xanthorrhiza*, since the SSR markers used revealed a classic diploid profile. *A. xanthorrhiza* is presumed to be an autotetraploid species (Ishiki et al. 2001), so up to four alleles can be expected in a microsatellite amplification. Among the SSRs analysed, only the marker AxD55 showed more than two amplified bands per individual (Morillo et al. 2004). Here we assume that this marker amplifies two independent loci, each one being bi-allelic. Diversity parameters, including the number of alleles per locus, allele frequencies and heterozygosis, were calculated with POWER MARKER ver. 3.23 (Liu and Muse 2005). To determine the observed and expected heterozygosities (Ho, HE), an estimation of the fixation index *Fis* was calculated according to Weir and Cockerham (1984). Heterozygote deficit or excess hypotheses were both tested assuming

the random union of gametes as the null hypothesis using the exact HW test in Genepop ver. 3.1 (Raymond and Rousset 1995). An estimation to determine genetic differentiation was calculated with the Fst statistic and its significance using FSTAT ver.2.9.3.2 (Goudet 2001). We also calculated the analysis of molecular variance (AMOVA, Michalakis and Excoffier 1996) using GenAlex ver. 6.5 (Peakall and Smouse 2012). For establishing genetic relationships, a cluster analysis was obtained in POWER MARKER using the Chord and DAS distances, which are both suggested as more appropriate for recently diverged populations (Chakraborty and Jin 1993). Clustering trees were generated by the Neighbor Joining (NJ) method. A Principal Coordinate Analysis (PCoA) was obtained with NTSYS ver.2.1 (Rohlf 2002) using the SMC coefficient (Sokal and Michener 1958).

We also used the STRUCTURE ver. 2.0 software to infer the population number (K) from allelic frequencies using the Bayesian method (Pritchard et al. 2000). This method provides a probability for each individual for being assigned to each population (value q). The "*Correlated Allele Frequencies Model*" was used for the assignment of individuals with and without *a priori* information on the membership of the individuals (cultivated or wild-type), using a given number of populations (K) of 2 and 3. As a result, admixed genotypes that are jointly assigned to two or more clusters are defined by their individual genetic

proportion of each compartment. Estimations of the parameter values were examined with a run of 1,000,000 burn-in steps, followed by 100,000 MCMC iterations as indicated in Pritchard et al. (2004).

Results

Microsatellite variability in *A. xanthorrhiza*

A total of 108 alleles were observed in the 11 SSR loci, of which 31 were shared between the wild and the cultivated *A. xanthorrhiza* populations (Fig. S2). The mean number of alleles per locus was 9.3 in the wild polycarpic types; the cultivated types have a mean of 2.1 alleles per locus, and this value decreases to 4.9 in the monocarpic types. Polymorphism values (PIC) were lower in the cultivated types (0.35) compared to the wild-types where this value doubles (0.7 and 0.57 in the wild polycarpic and monocarpic types, respectively) (Table 1). The PIC values of each examined locus ranged from 0.37 to 0.85, of which most of the loci showed PIC values >0.5 (Table 2). The 31 alleles observed in the cultivated types were distributed among 48 different multilocus genotypes over the 72 crop accessions, indicating 34% were duplicates among this collection.

Negative *Fis* (inbreeding coefficient) values in cultivated types suggest outbreeding, whereas positive Fis values in

Table 1 Number of alleles observed over 11 SSR loci in 177 samples of cultivated and wild *A. xanthorrhiza*

A. xanthorrhiza type	N	Aleles	A/locus	PIC	Ho	He
Cultivated	72	31	2.1	0.35	0.46	0.41
Wild polycarpic	76	103	9.3	0.7	0.43	0.68
Wild monocarpic	30	54	4.9	0.57	0.37	0.53

Table 2 Fis values obtained of the survey of 11 SSR in cultivated and wild *A. xanthorrhiza*

Locus	Cultivated		Wild polycarpic		Wild monocarpic		TA	PIC
	NA	Fis	NA	Fis	NA	Fis		
AxC27	3	0.024	5	0.526***	2	−0.021	5	0.59
AxC38	2	−0.015	6	0.690***	6	0.609***	7	0.71
AxC85	3	−0.057	16	0.291***	6	0.530***	17	0.72
AxD13	1	–	5	0.378**	4	0.753***	6	0.37
AxD34	3	−0.132	8	0.250***	6	0.165	8	0.74
AxD43	3	−0.061	11	0.270***	5	0.455**	13	0.73
AxD72	5	−0.122	14	0.353***	6	0.176	14	0.60
AxD82	4	−0.215**	12	0.182***	8	−0.106	12	0.85
AxD85	3	−0.154	8	0.299**	3	−0.016	8	0.65
AxD55A	2	0.060	10	0.545***	4	0.544**	10	0.67
AxD55B	2	−0.363**	8	0.376***	4	0.435	8	0.67

Polymorphism values (PIC) and total number of SSR alleles (TA) are also indicated

NA number of observed alleles in each *A. xanthorrhiza* form

*p < 0.05, **p < 0.01, ***p < 0.001

the wild populations suggest inbreeding under HW equilibrium (Table 2). The excess of heterozygotes in the cultivated types was statistically significant for two loci (AxD82 and AxD55B) when the excess of heterozygotes hypothesis was tested (p < 0.01). In most cases, the probability was highly statistically significant when the hypothesis of deficit was tested.

Among the SSR loci, locus AxD82 showed the highest number of alleles in the cultivated and the monocarpic form, whereas for the wild monocarpic the locus AxC85 was the most polymorphic. The frequency distribution of the 31 shared alleles between the cultivated and both wild forms showed important differences in some cases. In most cases, the allele in question occurred more frequently in the cultivated type than in its wild homologue, and in fewer cases a frequent allele in the wild types had a lower presence in the cultivated types (Fig. S2). Alleles unique to a particular type (private alleles) were not observed in the domesticated types, whereas 46 alleles were private to the polycarpic types and four were observed for the monocarpic. Concerning the allelic size, for all the examined loci the difference between the smaller and the larger allele ranged between 10 and 50 bp. Thus, the smallest difference in allele size was observed in the AxC27 and AxD13 loci with 10 bp, and the largest difference was observed at locus AxC85 with 50 bp.

Genetic structure and relationships

The analysis of genetic diversity (NJ and PCO) showed a differentiation between the wild and cultivated samples of *A. xanthorrhiza*, revealing the existence of intermediate individuals between the two forms. Figure 1 shows the NJ tree based on Chord distance (which was similar to the DAS tree), with artificial hybrids forming an intermediate group that also includes other individuals (wild individuals S36 and S37, and cultivars C60 and C66). The PCoA analysis (Fig. 2), which represents 24.8% of the total variation, similarly shows the artificial hybrids in a central position relative to the first axis of variation, which separates the wild and cultivated *A. xanthorrhiza* compartments. The wild polycarpic and the monocarpic types are then distinguished by the second axis of variation. Among the polycarpic types, there is also a differentiation according to the geographical origin of the populations (southern and central regions) through this same axis.

The obtained Fst values were all highly significant (p < 0.01) in the comparisons between the cultivated and the polycarpic and monocarpic types, suggesting a higher genetic differentiation between the cultivated *A. xanthorrhiza* and the monocarpic wild types. A lower value (Fst = 0.240) was obtained with the polycarpic type than with the monocarpic (Fst = 0.388). In contrast, both wild-types were less differentiated between them (Fst = 0.125). The AMOVA analysis indicated that 75% of the total genetic variance was found within the groups and that 25% was found among the identified groups: cultivated, wild polycarpic, wild monocarpic and experimental hybrids (Table 3). This information indicated significant structuring within *A. xanthorrhiza*.

The analysis of genetic structure, with or without (results not shown) the experimental hybrids between *A. x. var. andina* and the cultivated types, suggests the existence of spontaneous intermediate individuals. The genotype membership data, or Q values of the assignment test, show that when K = 2, individuals are classified as wild and cultivated types, except for eight genotypes which have intermediate values (data not shown). As shown in the plot of the Q values (Fig. 3) when K = 3, the three *A. xanthorrhiza* types were differentiated in each angle, suggesting admixed genotypes between the cultivated types and the wild polycarpic types and between the two wild arracacha types. Among the admixed genotypes between the cultivated types and the wild polycarpic types, one was found among the cultivated accessions with 80% belonging to the cultivated group and 20% to the wild polycarpic group. Three other admixed genotypes were found among the wild collections (Id. populations S9, S37, and S182). The first two plants exhibited 15% of the cultivated genotype, whereas the third showed 50% of the cultivated and wild genotypes. Similar values (36–41%) were observed for five of the presumed plant hybrids, whereas one was a cultivated genotype and the seven remaining plants had the wild polycarpic genotype.

Discussion

As indicated by statistical parameters, the wild pool of *A. xanthorrhiza* shows a genetic diversity that is not observed in the cultivated group. The cluster and the PCoA analysis (Figs. 1, 2 respectively) showed that wild diversity is related to the taxonomic classification and geographical origin of the *A. xanthorrhiza* populations. In the wild variety a distinction between the polycarpic or monocarpic types is suggested by the existence of intermediate genotypes. Some varieties, however, were ambiguously classed based on morphological classification (see Supplementary Table 1). Similar results were observed in the assignment test that detected admixed genotypes between the polycarpic and the monocarpic groups. These results and the fact that both wild types can be easily crossed (Knudsen 2003) suggest that gene flow between both wild types is occurring in nature. As a result, high morphological variation in the wild has been observed and has likely troubled past taxonomic studies, as previously argued by Constance

Fig. 1 NJ tree based on Chord (Cavalli-Sforza and Edwards 1967) distance representing the genetic relationships between 178 cultivated and wild *A. xanthorrhiza* samples based on 11 SSR loci. The tree was rooted using the species *A. incisa* (S48) as an outgroup. Wild types includes the polycarpic and the monocarpic forms (distinguished by *). *Arrows* indicate wild or cultivated plants grouped with artificial hybrids. For label of samples refer to supplementary material (Id. column)

and Affolter (1995), Knudsen (2003), and Blas et al. (2008a).

Allelic diversity in the cultivated group represents one-third of the total observed microsatellite variation in the *A. xanthorrhiza* wild group; however, a high degree of outbreeding in the cultivated types was supported by genetic parameters (Tables 1, 2). SSR allelic diversity observed in the cultivated forms suggests that; (a) domestication of arracacha could be set up from a limited number of wild progenitors and/or, (b) systematic integration

Fig. 2 Plot of the two principal axes of variation (PCoA) of SSR diversity observed in 178 samples of cultivated, wild, and experimental hybrids of *A. xanthorrhiza. C* cultivated accessions, *W* wild arraca-cha (monocarpic types are distinguished by *), *H* artificial hybrids
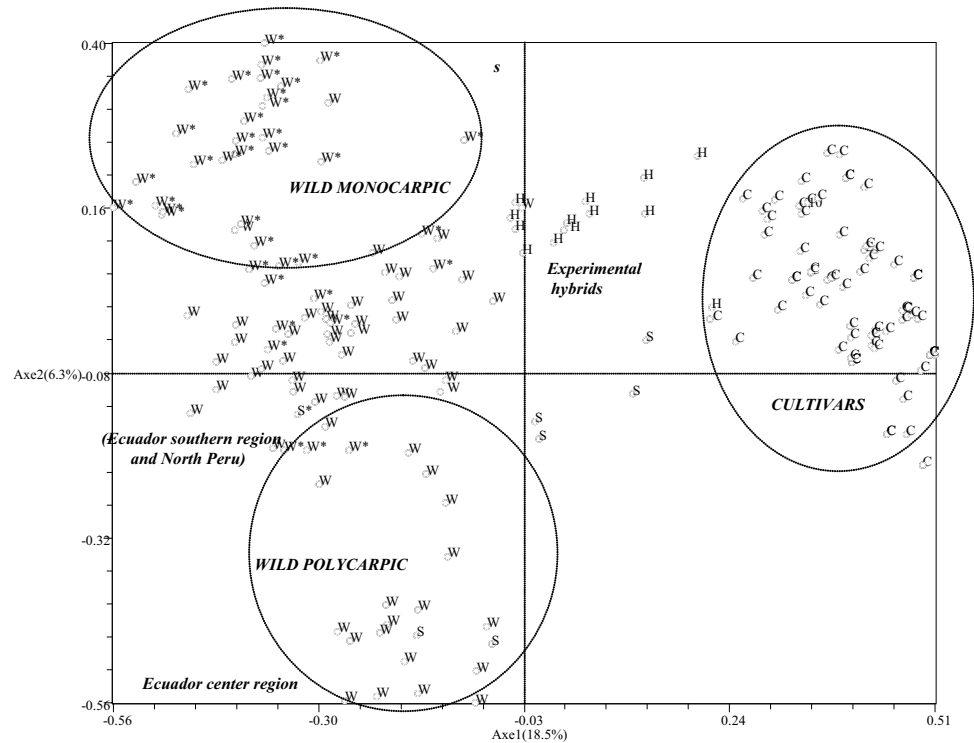

Table 3 Analysis of molecular variance components (AMOVA) and their significance (999 permutations) for the analyzed *A. xanthorrhiza* groups (cultivated, wild polycarpic and monocarpic, and experimental hybrids)

Source	df	SS	MS	Est. var.	% of total variation	Φst (PhiPt)
Among groups	3	354.46	118.15	2.86	24.8	0.248*
Within groups	174	1448.23	8.67	8.67	75.2*	
Total	177	1802.70		11.53	100%	

*p < 0.001

Fig. 3 Triangle plot of 178 samples of cultivated and wild *A. xanthorrhiza* obtained without prior status information in STRUCTURE ver. 2.0 of the polymorphism surveyed at 11 SSR loci (K = 3). (Color figure online)
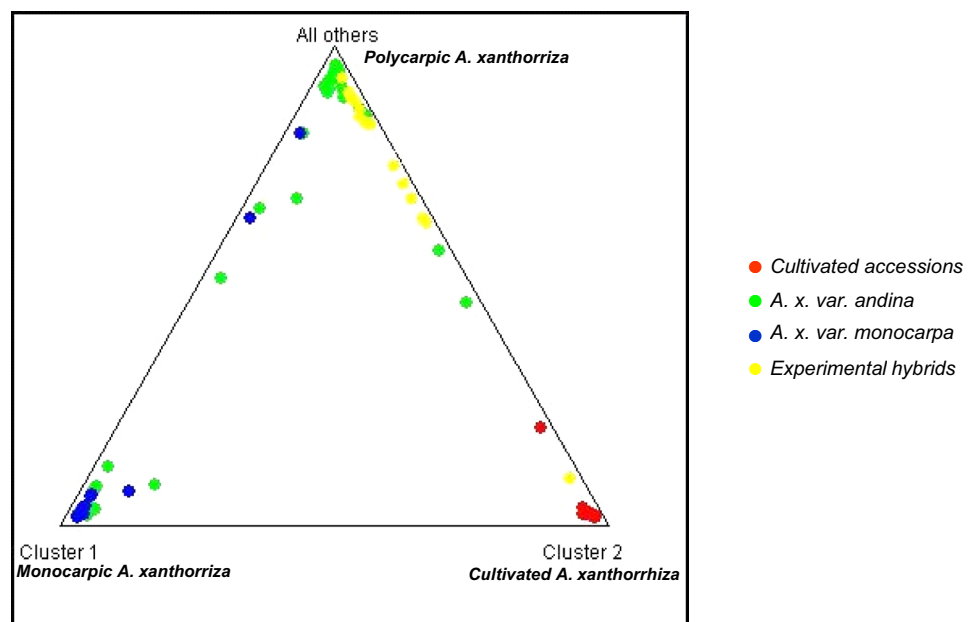
of diversity from the wild was probably occurring but with low incidence. Other authors concluded that cultivated arracacha exhibit a narrow genetic base (Erazo et al. 1996; Mazon et al. 1996; Castillo 1997) but more recent studies have reported higher genetic variability using a greater number of molecular markers (Blas et al. 2008a; Morillo and Sécond 2016).

The moderate genetic diversity observed in the cultivated form could be, among other factors, a consequence of asexual reproduction which would be the selected trait(s) applied for arracacha domestication (Morillo and Sécond 2016). In fact, despite the low rates of flowering observed for the domesticated crop (Hermann 1997; pers. obs), our results suggest a moderate but recombinant diversity. In fact, 66% of the original genotypes were found among the 72 characterized accessions, and the 31 alleles observed for the 11 SSR loci were evenly distributed among 48 multilocus genotypes. This result, which is higher than previously reported estimates (57 and 51%; Mazon et al. 1996; Blas 2005 respectively), could be explained by (a) the polymorphism rate of SSR loci, instead of isoenzymes which revealed very little polymorphism (Mazon et al. 1996; Erazo et al. 1996); (b) a good sampling (although the majority of accessions were collected in Ecuador and the northern region of Peru), where arracacha probably was domesticated (Knudsen 2003), and (c) the existence of genetic recombination only possible by residual sexuality as reported in other clonal crops.

However, residual sexuality is unlikely to have a major role in the domestication of arracacha. Actual flowering in the Andean region is relatively rare, and in many cases flowering plants are discarded, a practice seen in the production areas in Ecuador and Peru (pers. obs). On the contrary, in Brazil, where the crop was introduced, flowering is more frequent and sexuality allows new cultivars to be generated through selection from self-seeding progenies (Giordano et al. 1995 cited by Hermann 1997). Thus, the Brazilian cultivars that were originated from seed selection of a very few introduced clones, showed a high degree of heterozygosity in the F1 generation, i.e. color variation in the leaves, roots, and petioles (Sediyama et al. 1990b cited by Hermann 1997).

A similar scenario could also be proposed in the area of domestication, in which wild polycarpic genotypes would have been domesticated (Morillo and Sécond 2016), and the different morphological variants reported in the different germplasm collections (Mazon et al. 1996; Rosso et al. 2002; Blas et al. 2008a) could have been generated by residual sexuality. In contrast to other Andean clonal crops such as potatoes, or more recently reported in oca (*Oxalis tuberosa*), where a functional sexual reproduction system leading to productive plants has been found (Quiros et al. 1992; Bonnave et al. 2014), in arracacha no recruitment

of sexual seedlings has been documented. Interestingly, in Colombia a greater diversity of cultivars maintained by the Sibundoy Indians has been suggested as a result of sexual reproduction (Bristol 1988; Vasquez et al. 2004). In Ecuador, on the other hand, Andean farmers are not accustomed to using seed produced through sexual recombination. Although occasionally sexual seed can be observed in remnant plants, farmers are not familiar with seed plants (pers. obs.). Even if arracacha is mainly a cross-compatible species, outcrossing is not favored in the absence of pollinator insects, which seems to be the case in the Andean region. Besides, self-seeding presence is affected by inbreeding depression, which affects the germination capacity (Knudsen 2003). In Brazil, germination rates of only 10% were observed in self-seeding plants in the field (Santos et al. 1990).

Gene flow in arracacha?

Bidirectional hybridization between the cultivated types and their wild relatives has been reported in an increasing number of crop species (Ellstrand et al. 1999). New methods and analyses have demonstrated that plant gene flow rates vary tremendously (from nil to very high) depending on the species and specific populations involved (Ellstrand 2014). In the case of the related "old world" carrot for instance, it is well known that wild carrots may pollute the seed crops of cultivated carrots (Wijnheijmer et al. 1989; Hauser and Bjorn 2001). In the case of minor Andean crops, high diversity exists in the ex situ collections. However, gene flow has been poorly studied and little is known about the mechanisms that generate diversity within clonal crop species (Bonnave et al. 2016; Moscoe et al. 2017). In the case of arracacha, Blas et al. (2008a) argue a constant gene flow between the species *A. equatorialis* (here *A. x.* var. *andina*) and *A. xanthorrhiza*, since both share the same ecological and geographical distribution throughout all of the Peruvian Inter-Andean valleys.

Our results confirm the detection of spontaneous admixed genotypes, confirming that genetic flow occurs in the domestication area. The diversity analysis shows that among the genotyped samples, at least five wild plants had more affinity with the cultivated group. The assignment test made by STRUCTURE using F1 hybrid plants as a reference confirmed that at least three samples had admixed genotypes. Interestingly among these, the accession AM, which is a cultivar, was identified as admixed. This result would explain why this accession could be crossed easily with wild materials (Knudsen 2003). Admixed plants come from populations found in a region where the association of crop-weed complexes occurred with higher frequency. In contrast, among Ecuadorian cultivars no intermediate genotypes were detected.

Nevertheless, the identification of admixed plants, and the occurrence of spontaneous gene flow should have low incidence. The low number of spontaneous admixed genotypes detected between the cultivated and the wild types of *A. xanthorrhiza* corroborate this assumption. Even among the experimental 13 F1 hybrids, only five presented an admixed genotype between the cultivated and the wild polycarpic, whereas the seven remaining showed a wild genotype, and one was grouped as a cultivated type. This low occurrence of gene flow could be because it is very rare in Ecuador to find cultivated and wild plants in close proximity in the field (pers. obs.). However, in northern Peru this association has been reported as "occasional" (Valderrama and Seminario 2002; Knudsen 2003; Blas 2005), and it is in this area where a set of admixed genotypes has been detected by SSRs analysis. In addition, other constraints must also be surmounted, such as synchronized flowering, absence of pollinator insects, and the low flowering ability of many cultivars. The discovery of an admixed clone among the cultivated accessions demonstrated that integration has occurred, even if the gene flow incidence in the domesticated crop was low. SSR results also confirmed the true status of wild arracacha populations in most cases, since its habitat has been frequently related to ancient human presence and some wild populations are presumed to be associated with escaped cultivated plants (Valderrama and Seminario 2002).

Implications for arracacha breeding

Several authors, such as Hermann (1997), advocate the need to characterize the genetic diversity of arracacha ex situ collections for the evaluation of genetic erosion and the establishment of good strategies for the management and conservation of these genetic resources. Our study provides new information about these topics, revealing and quantifying the genetic diversity that exists in the wild populations, and validating the hybridization of plants obtained from experimental crosses. Our results corroborated those found by Knudsen (2003) that breeding of new cultivars incorporating genetic diversity from the wild pool of *A. xanthorrhiza* is possible as arracacha is both self and cross-compatible. A screen for traits such as resistance to drought and decay of storage roots in wild *A. xanthorrhiza* populations, as claimed by Blas et al. (2008a), will be useful for an eventual introgression of these desirable traits into the existing cultivars. It therefore becomes important to prioritize the conservation of *A. xanthorrhiza* wild populations through both ex situ and in situ strategies, and the emergence of a regional arracacha breeding program, using wild genetic resources, appears promising.

References

Blas R (2005) Diversity of *Arracacia* species in Peru. Ph. D. Thesis. Gembloux Agricultural University, Belgium p 154

Blas R, Ghislain M, Herrera M, Baudoin JP (2008a) Genetic diversity analysis of wild *Arracacia* species according to morphological and molecular markers. Genet Resour Crop Evol Int J Org Evol 55:625–642

Blas R, Hermann M, Baudoin JP (2008b) Analysis of the geographic distribution and relationships among Peruvian wild species of *Arracacia*. Genet Resour Crop Evol 55(5):643–655

Bonnave M, Bleeck G, Beltrán JR, Maughan P, Flamand MC, Terrazas F, Bertin P (2014) Farmers' unconscious incorporation of sexually-produced genotypes into the germplasm of a vegetatively-propagated crop (*Oxalis tuberosa* Mol.). Genet Resour Crop Evol 61(4):721–740

Bonnave M, Bleeckx T, Terrazas F, Bertin P (2016) Effect of the management of seed flows and mode of propagation on the genetic diversity in an Andean farming system: the case of oca (*Oxalis tuberosa* Mol.). Agric Hum Values 33(3):673–688

Bristol ML (1988) Edible arracachas of the Sibundoy. Rev Academia Colombiana de Ciencias Exactas, Físicas y Naturales 16(63):107–110

Castillo RO (1997) Caracterización molecular de 29 morfotipos de arracacha (*Arracacia xanthorrhiza* Bancroft) de la colección ecuatoriana. IX Congreso Internacional de Cultivos Andinos, University of Nacional de Sa, Antonio Abad del Cusco. Libro de resumenes, p 42 (**abstract**)

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Am J Hum Genet 19:233–257

Chakraborty R, Jin L (1993) A unified approach to study hypervariable polymorphisms: statistical consideration of determining relatedness and populations distances. In: Chakraborty R et al (ed), DNA fingerprinting: state of science. Basel, Boston, pp. 153–175

Constance L, Affolter JM (1995) Three new species and a new combination in *Arracacia* Bancroft (Umbelliferae/Apiaceae). Brittonia 47(3):320–327

Ellstrand N (2014) Is gene flow the most important evolutionary force in plants? Am J Bot 101(5):737–753

Ellstrand N, Prentice H, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. Ann Rev Ecol Syst 30:539–563

Erazo M, Dorregaray F, Hermann M (1996) Electroforesis de proteinas e isoenzimas en *Arracacia xanthorrhiza, Canna edulis y Oxalis tuberosa*. In: Memorias del II Congreso Ecuatoriano de Botanica, PUCE, Quito

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)

Hauser TP, Bjorn GK (2001) Hybrids between wild and cultivated carrots in Danish carrot fields. Genet Resour Crop Evol 48:499–506

Hermann M (1997) Arracacha. In: Hermann M, Heller J (eds) Andean roots and tubers: ahipa, arracacha, maca and yacon. International Plant Genetic Resources Institute, Rome, pp 75–172

Ishiki K, Salazar C, Salgado V, Galarza R, Vinueza J (2001) Plant type and starch content of arracacha (*Arracacia xanthorrhiza* B.) cultivated in Ecuador. Breed Res 3(Extra issue 2):270

Knudsen SR (2003) Reproduction biology of the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancroft var. *xanthorrhiza*) and the taxonomic status of the South American *Arracacia* Bancroft species with special emphasis on the position of the cultivated arracacha and related wild species. PhD. Thesis. Department of Ecology, KVL University, Denmark

Lim TK (2015) *Arracacia xanthorrhiza*. In: Edible medicinal and non medicinal plants. Springer, Dordrecht, pp 361–366

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21(9):2128–2129

Mazon N, Castillo R, Hermann M, Espinosa P (1996) La zanahoria blanca o arracacha *(Arracacia xanthorrhiza* Bancroft) en Ecuador. Instituto Nacional Autónomo de Investigaciones Agropecuarias (INIAP), Publicación Miscelánea No. 67, Quito, p 41

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064

Morillo E, Sécond G (2016) Tracing the domestication of the Andean root crop arracacha (*Arracacia xanthorrhiza* Bancr.): a molecular survey confirms the selection of a wild form apt to asexual reproduction. Plant Genet Resour 1–8. doi:10.1017/S1479262116000046

Morillo E, Second G, Pham JL, Risterucci AM (2004) Development of DNA microsatellite markers in the Andean root crop arracacha: *Arracacia xanthorrhiza* Banc. Apiaceae Mol Ecol Notes 4:680–682

Moscoe LJ, Blas R, Masi DH, Masi MH, Emshwiller E (2017) Genetic basis for folk classification of oca (*Oxalis tuberosa* Molina; Oxalidaceae): implications for research and conservation of clonally propagated crops. Genet Resour Crop Evol 64:867. doi:10.1007/s10722-016-0407-y

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539

Pritchard JK, Stephens M, Donelly P (2000) Inference of populations structure using multilocus genotype data. Genetics 155:945–959

Pritchard JK, Wen W, Falush D (2004) Documentation for the structure software, version 2. Department of Human Genetics, University of Chicago, Chicago

Quiros CF, Ortega R, Van Raamsdonk L, Herrera Montoya M, Cisneros P, Schmidt E, Brush SB (1992) Increase of potato genetic resources in their center of diversity: the role of natural outcrossing and selection by the Andean farmer. Genet Resour Crop Evol 39:107–113

Raymond M, Rousset F (1995) GENEPOP Ver.0.2. Population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rohlf FJ (2002) NTSYSpc: numerical taxonomy system, ver. 2.1. Exeter Publishing, Ltd., Setauket

Rosso CA, Medina I, Lobo M (2002) Morphologic characterization and agronomic evaluation of a Colombian collection of arracacha (*Arracacia xanthorrhiza* Bancroft). PGR Newslett 132:22–29

Santos FF, Vieira JV, Reis N (1990) Ocorrencia de germinaçao de sementes de mandioquinha-salsa (*Arracacia xanthorrhiza* Bancroft) a nivel de campo, em Brasilia. Hortic Bras Bras, 8(n1):65 (**Resumo**)

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kansas Sci Bull 38:1409–1438

Valderrama M, Seminario J (2002) Los parientes silvestres de la arracacha (*Arracacia xanthorrhiza* Bancroft) y su uso en medicina tradicional, en el norte peruano. Arnaldoa Trujillo Perú 9(1):67–91

Vásquez N, Medina C, Lobo M (2004) Caracterización morfológica de la colección colombiana (Tolima, Huila, Boyacá, Cauca) de arracacha (*Arracacia xanthorrhiza*). In: Raíces andinas: contribuciones al conocimiento y a la capacitación. Serie: conservación y uso de la biodiversidad de tubérculos y raíces andinos: una década de investigación para el desarrollo (1993–2003), pp. 165–178

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evol Int J Org Evol 38:1358–1370

Wijnheijmer EHM, Brandenburg WA, Ter Borg SJ (1989) Interactions between wild and cultivated carrots (*Daucus carota* L.) in the Netherlands. Euphytica 40:147–154