GMR
The Original

Development of simple sequence repeat (SSR) markers in Yacon (*Smallanthus sonchifolius*)

T. Bekaert[1], P. Hulpiau[2], M. Vandekerckhove[2], C. Hermans[2], M. Vanhee[3], I. Vanherpe[1], J. Van Doorsselaere[1] and A. Dewitte[1]

1) Research Group Smart Farming, Vives University of Applied Sciences, Campus Roeselare Wilgenstraat 32, 8800 Roeselare, Belgium
2) BioInformatics Knowledge Center (BiKC), Campus Station Brugge, Howest University of Applied Sciences, Rijselstraat 5, 8200 Brugge, Belgium
3) Department of Laboratory Medicine, AZ Sint Jan Brugge-Oostende, 8000 Brugge, Belgium

*Corresponding Author: Bekaert Tim – Tim.bekaert@vives.be

ABSTRACT
Yacon (Smallanthus sonchifolius (Poepp. & Endl.) H. Robinson) is a tuberous crop originating from the Andes region. Interest in yacon is increasing because of its high level of fructo-oligosaccharides (FOS) in the tubers, which offers health promoting benefits. Different varieties, mostly landraces, have been spread worldwide over the past decades and have been grown by several farmers or gardeners. The origin of these varieties is mostly lost and identification based on genetic information is necessary. Starting from raw EST cDNA sequencing data, 46 SSR markers were designed and tested on a collection of 11 yacon varieties that have been collected in Belgium. A total of 17 polymorphic markers were identified, resulting in a total number of 49 polymorphisms. Marker Y849 had the strongest discriminative power with up to 5 alleles per variety and 4 polymorphisms within the tested population. Using the combination of Y849 and Y2423 or Y10620, all varieties could be identified unambiguously, with the exception of 3 couples: 101-108, 103-107 and 105-109. These couples could not be discriminated against by the other tested markers. The use of these SSR markers may now be helpful to characterize yacon varieties in larger populations worldwide.

Keywords: *Yacon; SSR; Polymorphism; Nanopore Sequencing*

INTRODUCTION

Yacon (*Smallanthus sonchifolius* (Poepp. & Endl.) H. Robinson) is a tuberous crop grown by local farmers or communities scattered throughout the Andes, from Ecuador to north-western Argentina (Lebeda et al., 2012). The greatest germplasm diversity was found in a strip stretching along the eastern Andean slopes in Peru and Bolivia (Grau & Rea 1997). The crop is propagated vegetatively from the rhizomes and because of its high ploidy level (octoploid or dodecaploïd) and presumed hybrid character (Grau & Rea, 1997), fertility and seed set are reported to be very low (Lebeda et al., 2012).

Despite its high yield, sweet taste, crispy and juicy texture of the tubers and other attractive agronomical traits, the rather low caloric value of the tuber resulted in a diminished interest of agronomists in this crop, in favor of many other competing crops that became widely available in the region because of improved global transport. This may have contributed to the disappearance of many landraces over the years (Grau

& Rea, 1997).

Nowadays, many sugars such as glucose and starch have become general commodities with low prices in many parts of the world. The overconsumption of these carbohydrates resulted in bad dietary habits and the search for alternative products low in carbohydrates even become an important topic for many citizens and companies worldwide. Yacon may play an important role in the conversion to low-calorie diets because of its sweet taste and high level of fructo-oligosaccharides (FOS), the characteristic compound of the tubers. These FOS are mainly composed of GF2 to GF6 oligomers, with a fructose (F) chain as a basis and a terminal glucose (G). The length of these chains, and also the total FOS content, varies strongly between varieties (Ohyama et al., 1990; Goto et al., 1995; Lebeda et al., 2012; Kamp et al., 2019).

The FOS cannot be digested by the intestinal tract and does not cause an increase in the blood sugar level. Therefore, the consumption of yacon products with high levels of FOS offers health-promoting benefits such as reduced blood glucose levels and even prebiotic effects (Goto et al. 1995; Lachmann et al, 2003; Campos et al., 2012; De Almeida Paula et al., 2015; Caetano et al., 2016; Yan et al., 2019). In addition to the high FOS level, yacon also contains phenolic acids, which are responsible for a strong antioxidant activity (Reis et al., 2021).

Yacon has been spread to most parts of the world during the past century, including Europe. According to the distribution map of Fernandez (2005), yacon was already distributed from the Dominican Republic to Italy in 1927. From the early 2000s on, most of the yacon research in Europe was concentrated in the Czech Republic (Lachman et al., 2003; Fernandez et al., 2006), but a real yacon market has not been developed yet, and the crop is still unknown for most consumers.

However, slowly, the application of yacon in various ways is increasing. The increased interest in low-calorie, health-promoting crops may offer new chances for the food and pharmaceutical industry (Lorenzoni, 2017). This also triggers other European countries, such as Germany (Kamp et al., 2019) and Belgium (Tack, 2021; Tack, 2022), to perform variety trials.

In Belgium, the first variety trials already started about 10 years ago. Since breeding programs on yacon are lacking or in progress, most of these varieties are still landraces without commercial names. Some exceptions were released in Japan (Sugiura et al., 2007; Fujino et al., 2008) or Peru (cv. Morado) or more recently in the US (cv. Quinault).

Mostly, the name of the landrace refers to the country where the plant was collected (‘Peru’), the person who spread the plant material (‘Dimi’; New Zealand ‘Richard’), or the colour of the skin or the flesh (‘Rojo’, ‘Blanco’). During material exchange between gardeners, farmers, or institutes, plants are often mislabelled since it is difficult to distinguish the different varieties based on phenotypical characteristics. Belgian yacon farmers are mostly not aware of the variety they are cultivating, although the variety is important for further processing characteristics. It is therefore important to correctly identify the different varieties.

For correct identification, techniques based on molecular markers are recommended since they are not subjected to environmental fluctuations. Unfortunately, marker description in yacon is limited, mainly because of the lack of knowledge of the yacon genome. Consequently, the few reports on marker development were based on techniques that do not require knowledge of the yacon DNA sequence, such as RAPD (Milella et al., 2005; Mansilla et al., 2006), AFLP (Milella et al., 2011; Ovesná et al., 2018) or ISSR (Svobodová et al., 2013; Lorenzoni, 2017). Ziarovská et al. (2019) also described the use of iPBS and metabolic fingerprinting to describe genetic and chemical variation in yacon. In general, these studies

concluded that the genetic (and chemical) variation within the genus and between several yacon accessions was very low.

Over the past decades, SSR (simple sequence repeats), also known as microsatellites, have become a commonly used and popular marker technique to determine genetic diversity, species identification, parental assessment and genetic linkage mapping (Goncalves-Vidigal & Rubiano, 2011; Feng et al., 2016; Vieira et al., 2016). SSR has some advantages over other molecular markers, like ISSR, such as genetic co-dominance. Furthermore, they are multi-allelic, highly reproducible, relatively abundant, widely dispersed across the genome, highly variable and easily and automatically scored (Powell et al., 1996; Aitken et al., 2005; Miah et al., 2013).

Raw sequence data of yacon became available from a general study within the Asteraceae (Hodgins et al., 2013), which offers opportunities for the development of SSR markers in yacon. Hence, the main objective of this study was to develop SSR markers within yacon to unambiguously differentiate the different varieties present in Belgium. These SSR markers may also play an important role in our yacon breeding program, which was started a few years ago.

MATERIAL AND METHODS

Plant material

Eleven varieties were obtained from the PCG research institute in Kruishoutem (Belgium). These varieties have been evaluated in variety trials for many years, since 2012. An overview of the different varieties is listed in Table 1.

Table 1: Overview of the different varieties as collected by the PCG research institute.

Number	Variety	Main tuber skin colour	Flesh colour after harvest/after storage
101	Blanco	Reddish-purple	White
102	Dimi	light brown with sometimes red blush	White-yellow to yellow
103	Most common	Yellow-brown	White to white-yellow
104	Morado	Reddish/purple	White to white-yellow
105	Peru	Yellow-brown	White-yellow to yellow
106	Rojo	Yellow-brown to red	White
107	Cajamarca Peru	Yellow-brown	White
108	New-Zealand	Yellow-brown to red	White
109	New-Zealand II 'Richard'	Yellow-brown with sometimes red blush	White-yellow
110	Russia	Yellow-brown	White-yellow
111	White	Yellow-brown with sometimes red blush at the tops	White-yellow to yellow

SSR tracking and Primer design

Raw EST cDNA sequencing data from yacon were obtained from Hodgins et al. (2013). These data consist of about 19 million nucleotides, assembled in 12768 contigs. The sequences were screened for SSRs using SciRoKO 3.4 (Kofler et al., 2007). Only perfect SSRs consisting of di-, tri-, or tetranucleotide motifs and a minimal total length of 24 nucleotides were selected. SSR regions with flanking sequences of 200 nucleotides upstream and downstream of each SSR were extracted using SciRoKo's Little Helper module. Next, primers were designed for each extracted region using the 'SSR screening and primer design' tool in BatchPrimer3 (You et al., 2008). Optimal melting temperature was set at 60°C (+/- 5°C) and optimal product size at 150 bp (min. 100 bp and max. 300 bp). Finally, a total of 46 primer sets were selected to screen for polymorphisms in the SSR region.

Polymerase chain reaction

Genomic DNA was extracted from yacon leaves using the Nucleospin Plant II kit from Machery-Nagel, according to the manufacturer's protocol. PCR amplifications were performed in 25 µL reaction mixtures existing of 1X PCR buffer, 0.4 µM of each primer, 200 µM of each dNTP, 2.5 mM of $MgCl_2$ (50 mM), 0.5 U of Taq polymerase and 1 µL of DNA template. The PCR was performed on a Biorad T100 Thermocycler. The program included an initial denaturation of 94°C for 10 minutes, followed by 38 cycles of 30 seconds at 94°C, 30 seconds at 55°C, 30 seconds at 72°C and a final extension at 72°C for 20 minutes. PCR products were finally analysed on a 2% agarose gel and visualized by SYBR Safe, using an Imagequant ™ LAS500 (General Electric Company GE) detection system.

Fragment analysis

All primer combinations with an obvious PCR fragment on the agarose gel were selected for further fragment analysis. Therefore, the forward primer was replaced by a 5'-FAM-labeled primer and a new PCR was initiated. After PCR, 1 µL of amplicon was mixed with 8.5 µL of HiDi formamide and 0.5 µL of 400 HD ROX standard and a denaturation step was performed by heating the mix to 95°C for 5 minutes, followed by immediate cooldown to 4°C. Next, the samples were run in capillary electrophoresis on an ABI 3130XL genetic analyser (Applied Biosystems). The length of the DNA fragments was obtained and analysed via Peak Scanner (Thermo Fisher). A dendrogram was constructed using UPGMA with MEGA software (Kumar et al., 2016).

DNA sequencing using Oxford Nanopore Sequencing

Library preparation and barcoding ligation were carried out by using respectively the Ligation Sequencing Kit (SQK-LSK109) and barcoding kit (EXP-NBD104, EXP-NBD114) according to the Native Barcoding Genomic DNA protocol using the flongle flow cell (Oxford Nanopore Technologies). Briefly, the PCR fragments (34 ng of total DNA per sample) were first end-repaired and dA-tailed using NEBNext End-Repair and NEBNext dA-Tailing Modules (New England Biolabs) according to the manufacturer's protocol. Next, each sample was barcoded with a unique barcode number using Blunt/TA Ligase Master Mix (New England BioLabs). The barcoded samples were pooled and subsequently cleaned using Agencourt AMPure XP beads at a 2:1 bead-to-DNA ratio (to retain small DNA fragments). Finally, adaptors were added to the barcoded samples using the Adapter Mix II (AMII from the SQK-LSK109 kit) and subsequently purified using AMPure XP Beads again at a 1:1 bead-to-DNA ratio. For both cleaning steps, the library eluted from the beads was quantified by Qubit prior to the next steps. The samples were sequenced on a MinION device (Oxford Nanopore Technologies), making use of an FLO-FLG001 flongle flow cell for 30 hours. Basecalling was performed using the high-accuracy model.

Data analysis and identification of variants

During basecalling, reads were filtered based on their quality score (Q = 9). NanoFilt was then used to further filter the reads by length. High coverage, ranging from 350 to 3500 sequences, was achieved for each variety of yacon. Since no reference consensus was available for the expected variants for each yacon variety, custom Python scripts were used to perform the data analysis. In a first step, by k-means clustering, the reverse complement sequence of one of two clusters was calculated to enable analysis of all sequences in the same orientation. In a second step, after k-means clustering, different variants were identified based on cluster significance. MUSCLE alignments were then generated and a correction step was applied to retain only conserved positions with a threshold of 0.2. This resulted in putative consensus reference sequences for each variant. Finally, all reads were mapped against these consensus sequences using Minimap2 to eliminate potential errors. The final results were visually inspected using the Integrative Genomics Viewer (IGV).

RESULTS

Initially, 46 primer pairs were designed based on the expressed-sequence tag (EST) regions of yacon with BatchPrimer3. These primer pairs were tested on DNA extracted from all 11 available varieties using PCR and standard gel electrophoresis. In the case of 21 primer pairs, strong DNA amplification was observed in all varieties, and these primers were selected for further fragment analysis. Again, for all of these 21 primer pairs, PCRs were performed on all 11 varieties but with a FAM-labelled forward primer. Finally, the amplicons were run on capillary gel electrophoresis with a 400HD ROX standard and analysed with Peak Scanner.

Of the 21 marker regions analysed, only 4 showed no polymorphisms in the tested population (Y950, Y1233, Y6031 and Y7595). The other 17 markers showed polymorphisms, with the number of peaks within one genotype varying between 1 and 7 (Figure 1, Table 2), and the number of polymorphisms between the varieties for one single marker ranging from 1 to 5, as indicated in Table 2. A total of 49 polymorphisms were observed. Of all tested markers, Y849 had the highest discriminative power. With this marker, 4 varieties had a unique peak pattern (104, 106, 110 and 111) while identical peak patterns were observed within the varieties 101-102-108, 103-107 and 105-109. By using the other markers, no further discrimination could be obtained except for 102, which could be identified unambiguously when marker Y849 was combined with markers Y2423 or Y10620. Unfortunately, for all primers used, we were not able to distinguish the couple's 101-108, 103-107 and 105-109, which is also visible in the UPGMA dendrogram (Figure 2).

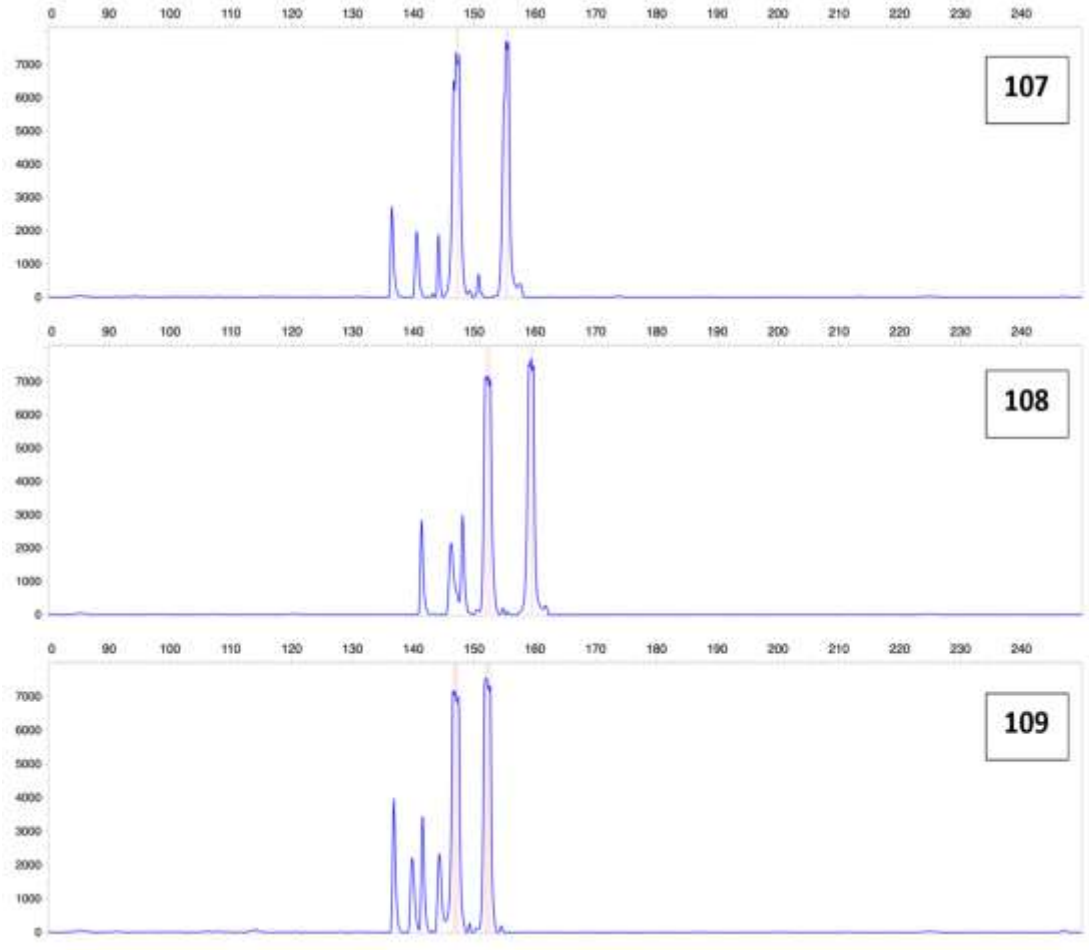

Figure 1: Capillary gel electrophoresis results for three varieties after amplification with primer set Y10620. The image shows relative fluorescence units as a function of the number of basepairs. Example of a peak pattern for the varieties 107 (up), 108 (middle), 109 (below). In this set, mostly 2 reproducible dominant peaks were observed with a fragment length around 150 bp.

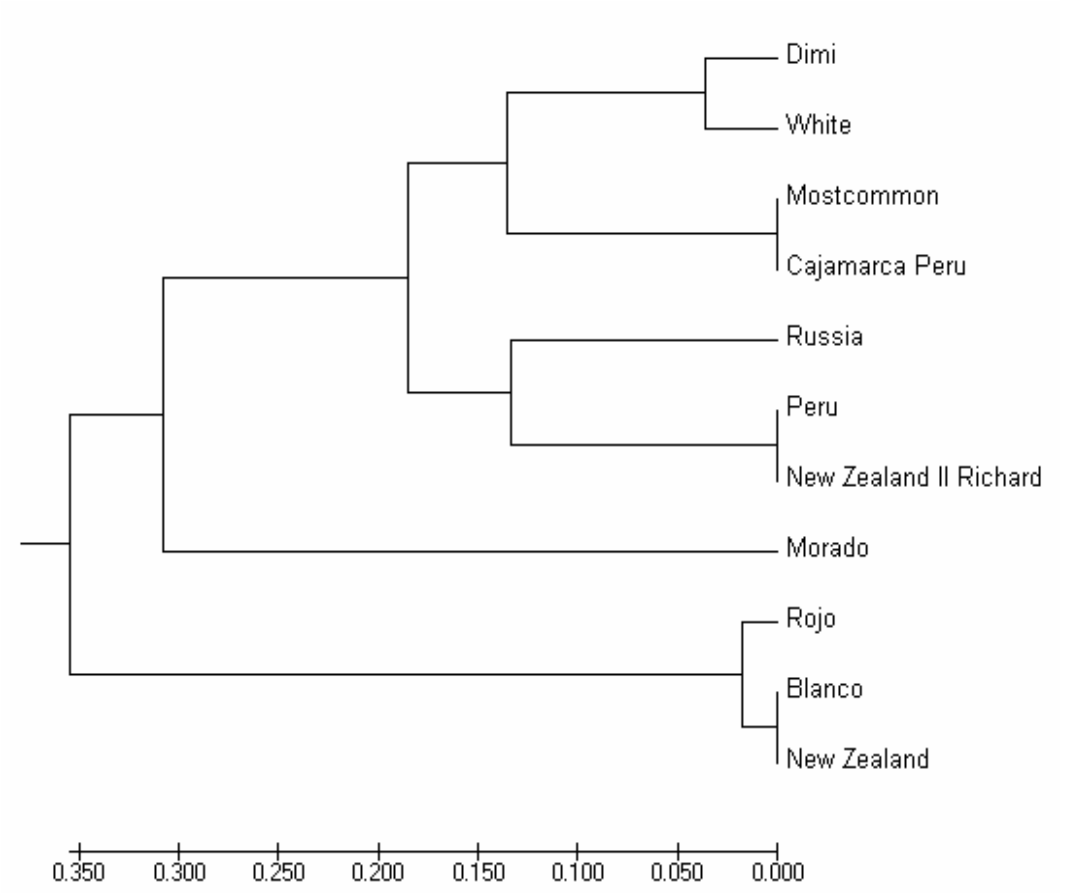

Figure 2: Dendrogram of the 11 assessed varieties, using UPGMA. Three couples could not be distinguished with the tested markers.

Table 2: Overview of the tested primer combinations and SSR motifs.

	Forward sequence	Reverse sequence	Motif	SSR length	N° of peaks	N° of polymorphisms
Y259	TCACCTCGAGCTCATCATCTT	ATGTTGAACTACCTGCCTCCA	TTC	24	2-3	2
Y294	AAGTCCCCCAAAATTGAAGAA	CCTCCTCAGGTGTCAATTCAA	ATG	24	-	-
Y399	CTTCGGCTAAAGTCTCCTTGC	TTTGAAGCCCTGATTTCACTG	CAT	30	-	-
Y584	CCCCAGAAAACAAAGCAAAAT	AAGAAATAAACCGGGTTCGTG	TCT	24	-	-
Y849	CCCTAATCCACCCAACAATTT	CGTGTGGGAAATTAAGGGAAT	ATC	27	3-5	4
Y950	GGGGGAGAAACTGAGAACAAG	CCATCGTCATCAACCATAAGC	AG	26	2	0
Y1110	AACATCTTCTCCCCGTCGTT	GTATGGTGAGGATGCCTTGAA	ACC	27	6-7	2
Y1233	GCTTTTCATCACCAAAATGGA	GGATCTCTGTGCAGCATTTCT	ATG	27	2	0
Y1245	CAATGAGCTCTCTCTTTGGACA	TCCCCTTGTACCTTGCTTAAA	CAT	30	-	-

Y2094	TGAAAAGGCATCAATCTTTGG	TCAATATTCTTGGCCTTGTGG	GAT	27	-	-
Y2134	CCTCTGTCAACTCCGATTTTG	TTTAACAAACCACAGGGACCA	TTA	24	-	-
Y2168	AAAATTTTCCGGTTTCCAGAAT	TCAGAAGCATCTGCAACCTCT	TAT	24	-	-
Y2184	GGGAAGGCGATCTTATCGTT	TCCTCCGTTTGGATTCTTCTT	CTT	36	-	-
Y2423	CCATTCTTCATCATTGAAGTTCC	ACAACAAACACACGGAACAGG	TCT	24	2-3	5
Y2490	ACCCTCTGTCCTCCCAAATAA	TCCTTCGAGCTCATAAACCAA	CT	28	-	-
Y2824	AAACTGATTTGCTGGTGGTTG	GGCCACCAAAGATCATAGCTC	GTG	24	-	-
Y3034	CCCCAAATTTACACTCTGAACC	GTGCTACGACCGCCATTTTAG	GTTA	24	1-2	1
Y3075	GCACGGAAATCACTTTGAAGA	TCATCCAGTGGTCACAACAGA	ATC	30	2-4	2
Y3476	TGGCTTTACAGAAAGGAAGCA	GGGCTAACAATGAGAAAACCA	ATG	27	2-4	5
Y3508	GGGGTCTCCAAAACAAAGAAA	TCGCGAGCATTTTCTACAGAT	CT	24	2-4	2
Y3525	CCGTCTCCTAATTCTCCGTTT	GATGCTGGTGGTGGTGATACT	TCA	24	2-3	3
Y4187	CACCTTCACCTTCACCTTCAC	CAAGAAGCTGCCTTGTTATTTTG	TTTG	24	-	-
Y4570	GGTGGATGAGGATAGGAGGTC	GCACGCAAAATTCTTTACCAA	CAT	24	-	-
Y5476	TCCTGTGGAGAGTTTTGATGG	ATGACCACCACAGTCCAGAAA	TTG	24	-	-
Y5852	GGACTCACTGCTGCTGACTTT	TGATGATGATGATGAGGGACA	CAT	36	-	-
Y6031	GGAAAAGATTGGGTTTTGAGG	CCAAATTGTTCAACCCTCTCA	TGG	24	3	0
Y6650	GGAAGTCACACAGAACCCTCTC	TCAGTTGGATCATGTTGAGCA	CAT	33	3	4
Y6831	CGGACACCACCTTTTGCTAT	TGGTGAGACAAGTGGATGGAT	TTC	27	-	-
Y6872	GCCAGAATTCATCTTCGTTGA	ATCCTCAAAGGGCAAAAAGAA	ATC	30	2-3	3
Y7001	CATGCTTGTGCCATAATGTTCT	CACTGCTACTTTTGACATGATGA	ATC	27	-	-
Y7261	GGCCTCAAAGAAATCTTCTCAT	ATTTGGGTACACCTGCGGTA	CAC	24	-	-
Y7286	CCCTCACTCAAAACAGTCTGG	CGGATACTGGGAAACAAGAGA	TCT	24	-	-
Y7372	CTTGTGGCTAAACCCTTGGAT	CAACAGTCGCGAATCTTCTTC	CAT	27	2-3	3
Y7566	GGGTTCGATTTCAATACACCA	AAGAGTAACCAGGGACACACG	GAA	27	-	-
Y7595	GCTCACAACCGTCTCATCCTA	CGGAGTGTACCGTCGTTAATC	CCA	24	1	0
Y7842	CAACACCCTTCCTGAACAATG	CGGGGGTTTGATACTGTTTCT	CTT	24	2-3	1
Y7862	CCATGTTTGAATCAGGAGATGA	CATGCTCCAATGAGTGATGTG	TTC	30	-	-
Y8815	ACTCGATGCATTCATCACCTT	CATCCAGTCCCTCAAACTGAA	TGA	30	-	-
Y9548	TTTCAAAAGGATGGGAAGAAAA	CAACAGTCAATTAACACGAGAGA	TGG	24	-	-
Y10084	TAGATCCGAACGAGGAAGGTT	GTCCTGCAATTCCACTTGTTC	CTG	24	2-3	2
Y10246	CTCGTTTCGTTTGACCTCATC	CAAGTTCAGGGAGGGATGAAT	TCA	24	-	-
Y10295	CAGATGACAGACCTGCTTTCC	CAGCAGATGAAAGTGACAAAGC	GA	24	-	-
Y10620	GGCCCTAAACAATCGGAATAC	GCCGAGAGTTGTCTGAGTACG	TCCT	24	1-2	5
Y12127	AGGCTGATGATCTTGTTCTCG	GGGGACTGCTTTGCAGTTTAC	GAT	24	1-3	3
Y12202	GCTGGCCTACTTTGTGCATTA	CTGGTGAAGTGGAGAAGATGAA	CAT	24	2-4	2
Y12730	TTCCAATCAAACCCAAATCAA	ACGATACACCGTCTGCTCATC	TG	24	-	-

The selected primers set after the PCR reaction are indicated in bold. '-': no fragment analysis performed.

Finally, the DNA sequences of the SSR regions after amplification with primer pairs Y2423 and Y10620

were analysed using Nanopore sequencing for all yacon varieties. Figure 3 shows the genetic sequence of variety 107 after using Y10620, including the adaptors used for sequencing. This sequence confirmed the 2 alleles with a gap of 8 nucleotides, as is observed in the peak analysis in Figure 1 (peaks at 148 and 156 nucleotides). Figure 4 shows all polymorphisms in all varieties after using Y2423, confirming the peak analysis of this marker. In each variety, 2 or 3 alleles were observed (Table 2, Figure 4, Figure 5).

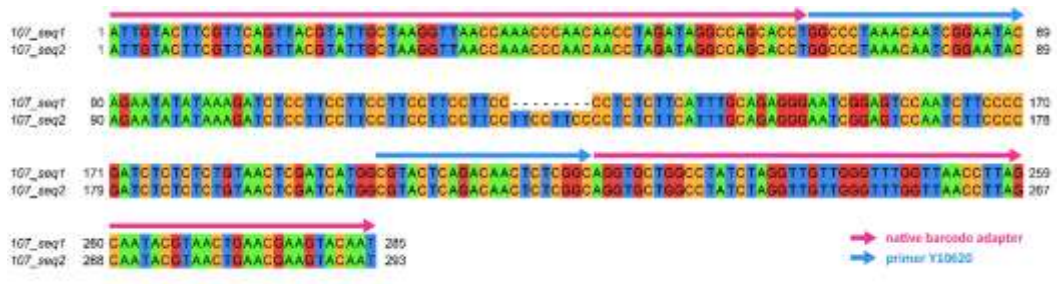

Figure 3: DNA analysis of variety 107 using primer pair Y10620 and Nanopore sequencing. 2 alleles were observed with a difference of 8 nucleotides. The barcode sequences added during the nanopore sequencing protocol are indicated. The SSR region is located between the primer sequences.

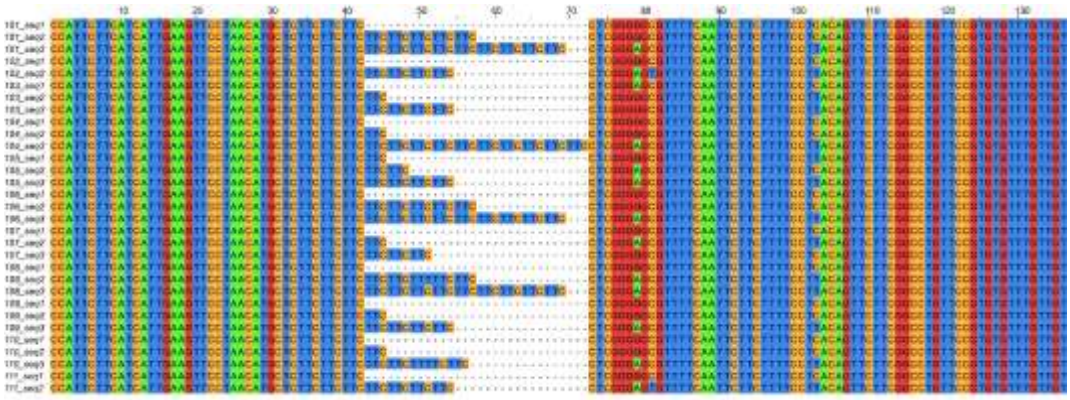

Figure 4: DNA analysis of all varieties after using primer pair Y2423 and Nanopore sequencing. For each genotype, 2 or 3 alleles were present and 5 different peak patterns could be observed between the genotypes. This resulted in the opportunity to form 3 distinct groups; 101-106-108, 102-111, 103-105-107-109 and to uniquely identify the genotypes 104 and 110.

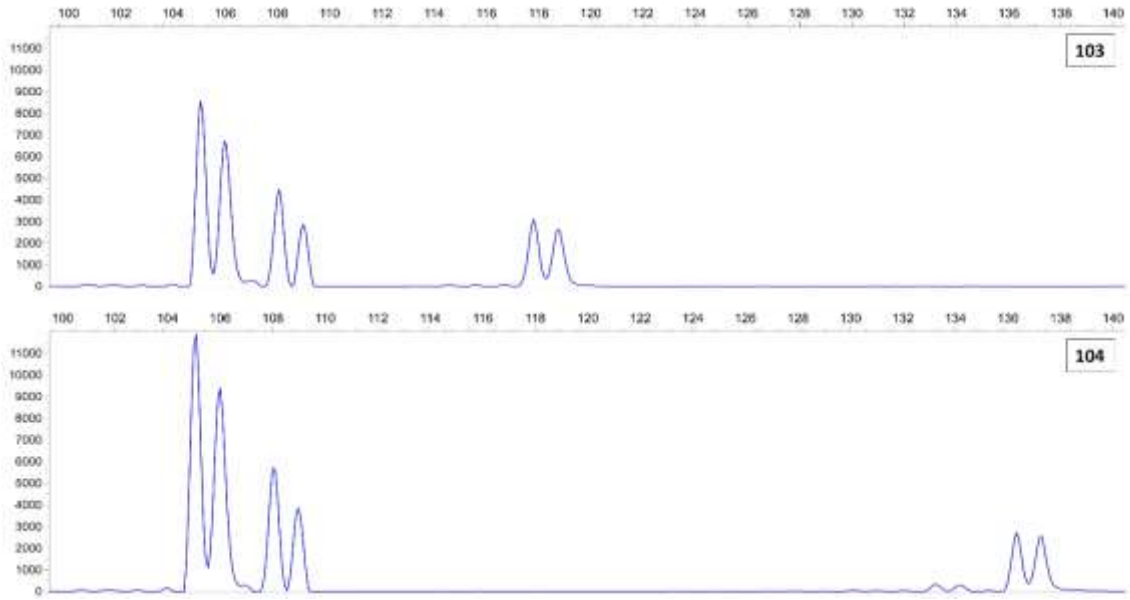

Figure 5: Capillary gel electrophoresis results of two varieties after amplification with primer pair Y2423. Image shows relative fluorescence units in function of number of basepairs. Example of peak pattern for the varieties 103 (top) and 104 (bottom).

DISCUSSION

Few studies have focused on marker development in yacon. The available reports mainly focused on techniques such as RAPD (Milella et al., 2005; Mansilla et al., 2006) or AFLP (Milella et al., 2011) since no DNA sequence is available. However, the increasing interest in inulin, an important reserve polysaccharide in Asteraceae, results in an increasing number of reports on yacon cultivation and processing. Gradually, genomic data also becomes available (e.g., Fan et al., 2022). In this report, we describe SSR markers in yacon for the first time, starting from the genomic data of Hodgins et al. (2013), who assembled data from several Compositae crops, including yacon.

According to the study of Zhao et al. (2012) in peanuts, EST trinucleotide SSRs show more polymorphisms than dinucleotide polymorphisms. In this study, from all potential SSR regions, 46 primer combinations were selected for further testing, in particular those within SSR regions with a trinucleotide motif. Of these 46 primer sets tested, 21 of them gave satisfying results after PCR amplification and were further analysed with fragment analysis. Finally, 17 of them resulted in a total number of 49 polymorphisms. Despite these polymorphisms, we were not able to distinguish all 11 varieties unambiguously. The best results were obtained with the combination of the markers Y849 and Y10620 or Y2423. This resulted in unique patterns for 5 of the 11 varieties, but 3 couples (101 – 108, 103 – 107 and 105 – 109) could not be discriminated. Other marker combinations did not result in further discrimination of these couples. Supplementary (mostly morphological) data were sometimes necessary for the final identification of the varieties. Variety 109 for instance, has a typical red blush on its skin, which is absent in 105.

In general, fragment analysis resulted in one to four peaks per variety. Rather exceptional, 6 or 7 peaks were observed. The presence of multiple peaks could be expected for a polyploid, because the discriminative power of a locus is higher for polyploids than for diploids (Pfeiffer, 2011). Due to the co-

dominant character of SSR markers, one or 2 bands or peaks can be expected in diploids. In polyploids, more than two bands may be obtained but it is difficult to assign the correct allelic dosage. For that reason, it is more difficult to describe genetic diversity within and between populations of polyploid plant species (Sampson and Byrne, 2012; Garcia-Verdugo et al., 2013; Bhandawat et al., 2019). In *Atriplex nummularia* Lindl. for instance, a polyploid perennial shrub, individual phenotypes consisted of one to eight bands per locus, as expected for an octoploid genome (Sampson and Byrne, 2012). In *Prunus lustitanica* L., including subspecies with 8x and 16x ploidy levels, the number of alleles varied between 2 to 6, what could be expected from polyploids. However, most loci showed fixed heterozygosity for only 2 allele variants in a great proportion of individuals and populations. This fixed heterozygosity resulted in levels of allelic diversity much lower than those expected for a high-order polyploid. So, despite its polyploid constitution, *Prunus lusitanica* showed very low levels of genetic variation that can be explained by the predominance of clonal population recruitment (Garcia-Verdugo et al., 2013).

Since yacon is an octoploid (or 12x in some cases), up to 8 alleles per genotype can be expected. In our data set, in general 2 to 4 alleles per genotype were observed, except 7 alleles in Y1110. These data rather suggest a low allelic diversity within the genotypes. This is not surprising since yacon has been described as a hybrid, with low seed set and low seed fertility (Lebeda et al., 2012), so a low gene flow between populations can be expected. This low genetic variation within yacon was already suggested by previous marker-associated studies (Svobodová, 2013).

In conclusion, for the first time, a set of polymorphic SSR markers on yacon is described. The best discriminative power was obtained after a combination of the markers Y849 and either Y2423 or Y10620 and these markers can be used now on larger populations for the correct identification of varieties or to study genetic relationships between them. Meanwhile, in Belgium, we could identify several varieties cultivated by yacon farmers.

ACKNOWLEDGMENTS

The authors would like to thank Annelien Tack and Helena Arents from the Viaverda research institute in Kruisem for providing all the yacon material.

REFERENCES

Aitken, K. S., Jackson, P. A., & McIntyre, C. L. (2005). A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theoretical and Applied Genetics, 110*(5), 789–801. https://doi.org/10.1007/s00122-004-1906-5

Bhandawat, A., Sharma, V., Singh, P., Seth, R., & Sharma, R. K. (2019). Discovery and utilization of EST-SSR marker resources for genetic diversity and population structure analyses of a subtropical bamboo (*Dendrocalamus hamiltonii*). *Biochemical Genetics, 57*(5), 652–672. https://doi.org/10.1007/s10528-019-09914-4

Caetano, B. F. R., de Moura, N. A., Almeida, A. P. S., Díaz, M. C., & Sivieri, K. (2016). Yacon (*Smallanthus sonchifolius*) as a food supplement: Health-promoting benefits of fructooligosaccharides. *Nutrients, 8*(7), 436. https://doi.org/10.3390/nu8070436

Campos, D., Betalleluz-Pallardel, I., Chirinos, R., Aguilar-Gálvez, A., Noratto, G., & Pedreschi, R. (2012). Prebiotic effects of yacon (*Smallanthus sonchifolius*), a source of fructooligosaccharides and phenolic

compounds with antioxidant activity. *Food Chemistry, 135*(3), 1592–1599. https://doi.org/10.1016/j.foodchem.2012.05.088

De Almeida Paula, H. A., Abranches, M. V., & de Luces Fortes Ferreira, C. L. (2015). Yacon (*Smallanthus sonchifolius*): A food with multiple functions. *Critical Reviews in Food Science and Nutrition, 55*(1), 32–40. https://doi.org/10.1080/10408398.2011.608925

Fan, W., Wang, S., Wang, H., Wang, A., Jiang, J., & Zhang, Z. (2022). The genomes of chicory, endive, great burdock and yacon provide insights into Asteraceae palaeo-polyploidization history and plant inulin production. *Molecular Ecology Resources, 22*(8), 3124–3140. https://doi.org/10.1111/1755-0998.13607

Feng, S., He, R., Lu, J., Jiang, M., Shen, X., & Wang, H. (2016). Development of SSR markers and assessment of genetic diversity in medicinal *Chrysanthemum morifolium* cultivars. *Frontiers in Genetics, 7*, 113. https://doi.org/10.3389/fgene.2016.00113

Fernandez, E. C. (2005). *Yacon (Smallanthus sonchifolius (Poepp. & Endl.) H. Robinson)* (Habilitation thesis). Czech University of Life Sciences Prague, Institute of Tropics and Subtropics.

Fernandez, E. C., Viehmannová, I., Lachman, J., & Milella, L. (2006). Yacon (*Smallanthus sonchifolius*) (Poepp. & Endl.) H. Robinson: A new crop in Central Europe. *Plant, Soil and Environment, 52*(12), 564–570.

Fujino, M., Nakanishi, T., Ishihara, J., Ono, S., & Sato, K. (2008). New yacon cultivars 'Andes no Yuki' and 'Salad Okame'. *Bulletin of the National Agricultural Research Center for Western Region (Japan), 7*, 131–143.

Garcia-Verdugo, C., Calleja, J. A., Vargas, P., Silva, L., & Moreira, O. (2013). Polyploidy and microsatellite variation in the relict tree *Prunus lusitanica* L.: How effective are refugia in preserving genotypic diversity of clonal taxa? *Molecular Ecology, 22*(6), 1546–1557. https://doi.org/10.1111/mec.12194

Gonçalves-Vidigal, M. C., & Rubiano, L. B. (2011). Development and application of microsatellites in plant breeding. *Crop Breeding and Applied Biotechnology, 11*(Special Issue 1), 66–72.

Goto, K. K., Fukai, K., Hikida, J., Nanjo, F., & Hara, Y. (1995). Isolation and structural analysis of oligosaccharides from yacon (*Polymnia sonchifolia*). *Bioscience, Biotechnology, and Biochemistry, 59*(12), 2346–2347.

Grau, A., & Rea, J. (1997). Yacon – *Smallanthus sonchifolius* (Poepp. & Endl.) H. Robinson. In M. Hermann & J. Heller (Eds.), *Andean roots and tubers: Ahipa, arracacha, maca and yacon* (pp. 199–242). International Plant Genetic Resources Institute (IPGRI).

Hodgins, K. A., Lai, Z., Oliveira, L. O., Still, D. W., Barker, M. S., & Rieseberg, L. H. (2013). Genomics of Compositae crops: Reference transcriptome assemblies and evidence of hybridization with wild relatives. *Molecular Ecology Resources, 14*(1), 166–177. https://doi.org/10.1111/1755-0998.12163

Kamp, L., Hartung, J., Mast, B., & Graeff-Hönninger, S. (2019). Tuber yield formation and sugar composition of yacon genotypes grown in Central Europe. *Agronomy, 9*(6), 301. https://doi.org/10.3390/agronomy9060301

Kofler, R., Schlötterer, C., & Lelley, T. (2007). SciRoKo: A new tool for whole genome microsatellite search and investigation. *Bioinformatics, 23*(13), 1683–1685. https://doi.org/10.1093/bioinformatics/btm157

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution, 33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Lachman, J., Fernandez, E. C., & Orsák, M. (2003). Yacon (*Smallanthus sonchifolius*): Chemical composition and use—A review. *Plant, Soil and Environment, 49*(6), 283–290.

Lebeda, A., Doleželová, I., Fernandez, E. C., & Viehmannová, I. (2012). Yacon (*Smallanthus sonchifolius*). In R. J. Singh (Ed.), *Genetic resources, chromosome engineering, and crop improvement: Medicinal plants* (pp. 641–702). CRC Press.

Lorenzoni, M. M., Menine, F., Marques Junior, O. G., Oliveira, F. L., & Borem, A. (2017). Genetic diversity of yacon accessions using ISSR markers. *Genetics and Molecular Research, 16*(3). https://doi.org/10.4238/gmr16039762

Mansilla, S., Roberto, C., López, B., & César, G. (2006). Molecular variability analysis of a Peruvian *Smallanthus sonchifolius* (Yacon) collection. *Ecología Aplicada, 5*(1–2), 75–80.

Miah, G., Rafii, M. Y., Ismail, M. R., Puteh, A. B., Rahim, H. A., Islam, K. N., & Latif, M. A. (2013). A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. *International Journal of Molecular Sciences, 14*(11), 22499–22528. https://doi.org/10.3390/ijms141122499

Milella, L., Salava, J., Martelli, G., & Fernandez, E. C. (2005). Genetic diversity between yacon landraces based on random amplified polymorphic DNAs. *Czech Journal of Genetics and Plant Breeding, 41*(2), 73–78.

Milella, L., Martelli, G., Salava, E., Fernandez, E. C., & Greco, I. (2011). Total phenolic content, RAPDs, AFLPs and morphological traits for the analysis of variability in *Smallanthus sonchifolius*. *Genetic Resources and Crop Evolution, 58*(4), 545–551. https://doi.org/10.1007/s10722-010-9595-0

Ohyama, T., Ito, O., Yasuyoshi, S., Ikarashi, T., & Baba, T. (1990). Composition of storage carbohydrates in tuber roots of yacon (*Polymnia sonchifolia*). *Soil Science and Plant Nutrition, 36*(1), 167–171.

Ovesná, J., Russo, D., Frescura, D., Fernandez, E. C., & Viehmannová, I. (2018). Assessment of genetic diversity of *Smallanthus sonchifolius* landraces using AFLP markers. *Genetika, 50*(3), 803–816. https://doi.org/10.2298/GENSR1803803O

Pfeiffer, T., Roschanski, A. M., Pannell, J. R., Korbecka, G., & Schnittler, M. (2011). Characterization of microsatellite loci and reliable genotyping in a polyploid plant, *Mercurialis perennis*. *Journal of Heredity, 102*(4), 479–488. https://doi.org/10.1093/jhered/esr024

Powell, W., Machray, G. C., & Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Science, 1*(7), 215–222.

Reis, F. R., Marques, C., Sales de Moraes, A. C., & Masson, M. L. (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.2021.05.005

Sampson, J. F., & Byrne, M. (2012). Genetic diversity and multiple origins of polyploidy in *Atriplex nummularia*. *Biological Journal of the Linnean Society, 105*(2), 218–230.

Sugiura, M., Nakanishi, T., Kameno, T., & Doi, Y. (2007). A new yacon cultivar 'Sarade Otome'. *Bulletin of the National Agricultural Research Center for Western Region (Japan), 6*, 1–13.

Svobodová, E., Dvořáková, Z., Cepková, P. H., Viehmannová, I., & Fernandez, E. C. (2013). Genetic diversity of yacon and its wild relatives revealed by ISSR markers. *Biochemical Systematics and Ecology, 50*, 383–389. https://doi.org/10.1016/j.bse.2013.06.012

Tack, A. (2021). *Demonstrative varietal trial of yacon*. Provincial Research Centre for Vegetable Production (PCG).

Tack, A. (2022). *Evaluation of varieties and cultivation techniques in yacon: Effects of mulching and irrigation*. Provincial Research Centre for Vegetable Production (PCG).

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. F. (2016). Microsatellite markers: What they mean and why they are so useful. *Genetics and Molecular Biology, 39*(3), 312–328. https://doi.org/10.1590/1678-4685-GMB-2016-0027

Yan, M. R., Welch, R., Rush, E. C., Xiang, X., & Wang, X. (2019). A sustainable wholesome foodstuff: Health effects and potential dietotherapy application of yacon. *Nutrients, 11*(11), 2632. https://doi.org/10.3390/nu11112632

You, F. M., Huo, N., Gu, Y. Q., Luo, M.-C., Ma, Y., Hane, D., & Lazo, G. R. (2008). BatchPrimer3: A high-throughput web application for PCR and sequencing primer design. *BMC Bioinformatics, 9*, 253. https://doi.org/10.1186/1471-2105-9-253

Zhao, Y., Prakash, C. S., & He, G. (2012). Characterization and compilation of polymorphic simple sequence repeat (SSR) markers of peanut from public databases. *BMC Research Notes, 5*, 362. https://doi.org/10.1186/1756-0500-5-362

Ziarovská, J., Padilla-González, G. F., Viehmannová, I., & Fernandez, E. C. (2019). Genetic and chemical diversity among yacon accessions based on iPBS markers and metabolomic fingerprinting. *Plant Physiology and Biochemistry, 141*, 183–192. https://doi.org/10.1016/j.plaphy.2019.05.026