Genome-wide development of intra- and inter-specific transferable SSR markers and construction of a dynamic web resource for yam molecular breeding: Y2MD

Moussa Diouf, Yedomon Ange Bovys Zoclanclounon, Pape Adama Mboup, Diaga Diouf, Erick Malédon, Ronan Rivallan, Hana Chair, Komivi Dossa

▶ **To cite this version:**

HAL Id: hal-04550688

https://hal.inrae.fr/hal-04550688v2

Submitted on 16 Jan 2025

The Plant Genome

ORIGINAL ARTICLE

Special Section: Modern Improvement of Tropical Crops

Genome-wide development of intra- and inter-specific transferable SSR markers and construction of a dynamic web resource for yam molecular breeding: Y2MD

Moussa Diouf[1,2] | Yedomon Ange Bovys Zoclanclounon[3] | Pape Adama Mboup[1] | Diaga Diouf[2] | Erick Malédon[4,5] | Ronan Rivallan[5] | Hâna Chair[5] | Komivi Dossa[4,5]

[1]Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Senegal

[2]Laboratoire Campus de Biotechnologies Végétales, Département de Biologie Végétale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, Dakar, Senegal

[3]Department of Crop Science and Biotechnology, Jeonbuk National University, Jeonju, Republic of Korea

[4]UMR AGAP Institut, CIRAD, Petit Bourg, France

[5]UMR AGAP Institut, University of Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

Correspondence
Komivi Dossa, CIRAD, UMR AGAP Institut, 97170 Petit Bourg, Guadeloupe, France. Email: komivi.dossa@cirad.fr

Funding information
Bill and Melinda Gates Foundation, Grant/Award Number: AfricaYam Project - Grant OPP1052998

Abstract

Microsatellite markers are widely used in population genetics and breeding. Despite the economic significance of yams in developing countries, there is a paucity of microsatellite markers, and as of now, no comprehensive microsatellite marker database exists. In this study, we conducted genome-wide microsatellite marker development across four yam species, identified cross-species transferable markers, and designed an easy-to-use web portal for the yam researchers. The screening of *Dioscorea alata*, *Dioscorea rotundata*, *Dioscorea dumetorum*, and *Dioscorea zingiberensis* genomes resulted in 318,713, 322,501, 307,040, and 253,856 microsatellites, respectively. Mono-, di-, and tri-nucleotides were the most important types of repeats in the different species, and a total of 864,128 primer pairs were designed. Furthermore, we identified 1170 cross-species transferable microsatellite markers. Among them, 17 out of 18 randomly selected were experimentally validated with good discriminatory power, regardless of the species and ploidy levels. Ultimately, we created and deployed a dynamic Yam Microsatellite Markers Database (Y2MD) available at https://y2md.ucad.sn/. Y2MD is embedded with various useful tools such as JBrowse, Blast, *insilico*PCR, and SSR Finder to facilitate the exploitation of microsatellite markers in yams. This study represents the first comprehensive microsatellite marker mining across several yam species and will contribute to advancing yam genetic research and marker-assisted breeding. The released user-friendly database constitutes a valuable platform for yam researchers.

- -

Plain Language Summary

Yams (*Dioscorea* spp.) are vital for developing economies, yet improved varieties are slow to emerge. To accelerate breeding, our study addressed the scarcity of microsatellite markers in yams. We developed 864,128 primer pairs, identifying 1170 transferable markers across species. The validation of 17 markers proved functionality irrespective of species or ploidy. These resources are housed in an online database, fostering collaboration and advancing yam genetic studies and marker-assisted breeding.

1 | INTRODUCTION

Microsatellites, also termed simple sequence repeats (SSRs), are tandemly repeated DNA sequences generally spanning 1–6 bp (Ellegren, 2004). SSRs are ubiquitous in the genome of higher organisms and exhibit genetic polymorphism at inter- and intra-species levels (Amiteye, 2021; Bruford & Wayne, 1993; Ellegren, 2004). Therefore, they have been widely used in population genetics (Chen et al., 2022; Sapkota et al., 2022), genotyping, marker-assisted breeding, genome-wide association analysis (Jha et al., 2021; Kim et al., 2021; C. Li et al., 2014), forensic investigation (D. Li et al., 2021; J. Wang, Sun et al., 2022), evolutionary studies (de Freitas et al., 2022; Nowicki et al., 2021; Song et al., 2021), and conservation genetics (Dong et al., 2022; Ming et al., 2022; Rodríguez-Peña et al., 2018).

Considering the low level of difficulty regarding the technical aspect of setting an SSR-based genotyping core facility, the relatively inexpensive cost of reagents, and the level of throughput, SSRs markers are quite affordable and implementable in a standard laboratory (Herrera & Ghislain, 2013). Besides, SSRs can exhibit good discriminatory power, are randomly present in the genome, multi-allelic, and can be located in the genic region, enabling the genetic dissection of traits of interest (Ellegren, 2004). Microsatellites also exhibit co-dominance with Mendelian inheritance, facilitating their traceability within different breeding populations (Normann et al., 2018). Therefore, owing to their simple identification and high reproducibility, many researchers have relied on SSRs (X. Wang & Wang, 2016).

The development of molecular marker databases has been extensively promoted in economically important plants such as *Sesamum indicum* (Dossa et al., 2017), *Vicia faba* (Mokhtar et al., 2020), *Nicotiana* spp. (X. Wang et al., 2018), legumes species (*n* = 13) (Duan & Kaundal, 2021), horticultural species (*n* = 112), and *Anemone coronaria* (Martina et al., 2022). Since the usage of an online database is based on a graphical user interface, the database offers researchers a user-friendly path for the rapid identification of molecular

markers without a pre-requisite expertise in command line scripting. A web database concentrates all information in one platform and assists researchers in designing markers related to a trait of interest (Dossa et al., 2017).

Yams (*Dioscorea* spp.) represent an important staple food crop, feeding over 300 million people in tropical and subtropical regions (Mignouna et al., 2008). The biggest producer country is Nigeria with 70% of the global production, followed by Ghana and Côte d'Ivoire (FAOSTAT, 2021). The genus contains ~650 species (Virouel et al., 2018). As a cash crop, yams contribute to the income of more than 60 million people (Asiedu & Sartie, 2010), with a production of 74 million tons and an economic market value of USD 21 billion estimated in 2020 (United Nations Statistics Division, 2022). Besides, in West Africa where yam is the most cultivated in the so-called "yam belt" (from Western Cameroon to central Côte d'Ivoire) (Sarcelli et al., 2019), more than five million people directly benefit from yam culture (Mignouna et al., 2020), making this crop a valuable resource for food security and poverty alleviation in this part of the African continent (Darkwa et al., 2020).

To accelerate yam breeding programs, attention has been paid to the development of polymorphic DNA markers. Restriction fragment length polymorphisms (RFLPs) were first developed to investigate the origin and phylogeny of cultivated Guinea yams and wild relatives by using chloroplast DNA and nuclear ribosomal DNA (Terauchi et al., 1992). Later on, amplified fragment length polymorphism (AFLP) and random amplified fragment DNA (RAPD) were employed for genetic diversity assessment (Egesi et al., 2006; Malapa et al., 2005; Zannou et al., 2009), yam mosaic virus resistance markers identification (Mignouna, Abang, et al., 2002), and genetic linkage map construction (Mignouna, Mank, et al., 2002).

Furthermore, due to their high level of polymorphism, SSRs were developed from cultivated and wild yam species (Andris et al., 2010; Hochu et al., 2006; Mignouma et al., 2003; Mizuki et al., 2005; Silva et al., 2014; Siqueira et al., 2011; Tamiru et al., 2015; Terauchi & Konuma, 1994; Tostain

et al., 2006; H. Wang et al., 2022). Nonetheless, less than 100 SSR markers are available in yams to date, and no information on their genome coverage is available (Korsa et al., 2022; Loko et al., 2016; Nwogha et al., 2022; Otto et al., 2015; Tostain et al., 2007). Hence, generating more SSR markers will surely empower researchers in their routine tasks such as paternity testing, variety identification, and quantitative trait loci (QTL) detection.

Meanwhile, with the help of short and long reads, as well as chromosome conformation sequencing technologies, good-quality genomes were recently released in yams. These data have enabled the identification of genomic regions associated with agronomically important traits, such as sex determination systems (Cormier et al., 2019; Sugihara et al., 2020; Tamiru et al., 2017), anthracnose resistance, tuber oxidative browning, and tuber starch content (Bredeson et al., 2022), as well as post-harvest hardening (Siadjeu et al., 2021).

To date, the genomes of four species, *Dioscorea alata* (Bredeson et al., 2022), *Dioscorea rotundata* (Tamiru et al., 2017), *Dioscorea dumetorum* (Siadjeu et al., 2020), and *Dioscorea zingiberensis* (Cheng et al., 2021), are publicly available. The genomes exhibit a high level of collinearity (Bredeson et al., 2022), suggesting that SSRs might be highly transferable from one species to another, even to those wild and minor cultivated species that have not yet been sequenced. Therefore, establishing a comprehensive SSR database for yam researchers is fundamental for genetic studies.

The present study was designed to (i) mine the genome of the four yam species for the identification of single-species and cross-species transferable SSR markers, (ii) create primer sets for the markers, (iii) evaluate the discriminatory power of selected markers, and (iv) build a user-friendly digital web resource to support yam genetic research.

2 | MATERIALS AND METHODS

2.1 | Genomic sequence resources

A total of four genome sequences were retrieved from the National Center for Biotechnology Information genome portal, including *D. alata* cultivar TDa95/00328 (GenBank assembly accession: GCA_020875875.1), *D. dumetorum* (GenBank assembly accession: GCA_902712375.1), *D. rotundata* cultivar TDr96_F1 (GenBank assembly accession: GCA_009730915.2), and *D. zingiberensis* isolate CJ2019 (GenBank assembly accession: GCA_014060945.1) (Table S1). Besides, we added a newly assembled chromosome-scale assembly (unpublished data) of *D. alata* cultivar Kabusa. Kabusa is a widely grown variety in the West Indies.

Core Ideas

- Whole genome assemblies of the four yam species were screened for the detection of microsatellites.
- A set of 1170 cross-species transferable microsatellite markers were identified.
- An experimentally validated subset of 17 markers has shown consistent discriminatory power with respect to species and ploidy levels.
- A user-friendly and dynamic web portal, the Yam Microsatellite Markers Database (Y2MD), has been developed and made freely available at https://y2md.ucad.sn/.
- The database includes a number of useful tools such as JBrowse, *insilico*PCR, SSRFinder, and Blast.

2.2 | Identification of microsatellites and primer design

SSRs were searched using the Genome-wide Microsatellites Analyzing Tool Package (GMATA) (X. Wang & Wang, 2016), and primers were designed from the flanking sequences (200 bp on each side) to probe the genomic sequences. The genome assembly was imported into the GMATA "SSR identification" module in the FASTA format. The microsatellite search parameters applied for SSR detection are as follows: minimum length (nt): 1, maximum length (nt): 6, minimum number of repeats: 5.

Using custom python scripts, only sequences with a size greater than 10 bp were retained. The SSR locus information files generated (.ssr) contain the start and end positions of the SSRs on the chromosomal sequence, the SSR template, and the number of repeat units.

The genome assemblies in FASTA format and the output files (.ssr) generated by the "SSR identification" module were imported into the "Marker Designing" module of GMATA. Primer design parameters were as follows: minimum amplicon size: 100 bp, maximum amplicon size: 400 bp, optimal annealing temperature: 60°C, flanking sequence length: 200 bp, maximum model length (flanking sequence + SSR): 2000 bp. The output file (.seq) included the repeats centered on 200 adjacent sequence bases on both sides. Output files with extensions .mk and .sts were created, containing left and right primer sequences, hybridization temperatures, primer positions on chromosomes, and expected amplicon sizes.

The identified SSRs were classified into genic and non-genic SSRs for *D. alata* and *D. rotundata* for which genome annotations are available. Subsequently, the functional

enrichment of genes containing SSRs was performed with KOBAS-i webserver (Bu et al., 2021) in order to find out their biological significance.

2.3 | In silico polymerase chain reaction

To validate markers, generate amplicons, and assign marker positions across the *Dioscorea* genomes, the e-Mapping module was used to run the e-PCR algorithm (Schuler, 1997). This module was also used for intraspecific (the two *D. alata* genomes) and interspecific marker mapping at the genome scale. In e-PCR, FASTA files containing all genome assemblies and newly developed markers created with the Marker designing module (.sts files) were imported as "Sequence File" and "Marker File," respectively. The gap (-n) and max. indel (-g) parameters were set to 2 and 1, which means that two mismatches and one insertion are allowed. The output file (.emap) provides detailed marker amplification information with calculated amplification sizes and chromosomal targets and identifies single locus and multilocus markers, ultimately producing those that are potentially polymorphic (different sizes of PCR products).

2.4 | Plant materials, DNA isolation, and PCR experiment

Leaf samples from 10 yam species (*D. alata*, *D. rotundata*, *D. dumetorum*, *Dioscorea bulbifera*, *Dioscorea trifida*, *Dioscorea transversa*, *Dioscorea esculenta*, *Dioscorea cayenensis*, *Dioscorea abyssinica*, and *Dioscorea nummularia*) including cultivated and wild species were used for DNA extraction. Six *D. alata* genotypes with different ploidy levels (14 M, 74F, Pyramide, Kabusa [2x], CT148 and CT198 [4x]) and two *D. rotundata* (Dr1 and Dr2) were included (Table S2). Plant samples were available in the germplasm collection of CIRAD, Guadeloupe. DNA was extracted based on the mixed alkyl trimethyl ammonium bromide method (Cummings & Wood, 1989). Prior to the PCR experiment, the quantity and quality of the extracted DNA were consecutively assessed on 1% agarose gel and Invitrogen Qubit Flex Fluorometer. A total of 18 SSR primers were randomly selected from different chromosomes/contigs of the four species (Table 1). In this study, we opted to analyze 18 markers for a practical reason: we aimed to use no more than three PCR plates (each containing 96 samples). These markers exhibited interspecific amplifications through the e-PCR. M13 tail (CACGACGTTGTAAAACGAC) was added to the primers. The PCR reaction was performed using the QIAGEN kit under the conditions as follows: 95°C for 5 min, 10 cycles of 30 s at 95°C, 60°C for 1 min 30 s and 72°C for 30 s, 25 cycles of 30 s at 95°C for 1 min 30 s, and 72°C for 30 s, followed by 30 min at 60°C. The migration of the PCR products was conducted on the ABI 3500xL (Thermo Fisher Scientific

Inc.). Analysis of microsatellite profiles was performed with Genemapper v6.0 (Applied Biosystems).

2.5 | SSR genotyping data analysis

The genotyping data served to perform a hierarchical clustering to depict the genetic relationships between the 16 accessions (Table S2). The clustering was conducted in R v4.2.2 (R Core Team, 2022) following the unweighted pair group method with the arithmetic mean method using the function *hclust()* of the package base. The resulting tree was rendered with the function *fviz_dend()* of the package factoextra v1.0.7 (Kassambara & Mundt, 2020). For comparative analysis purposes, a principal component analysis was also executed with the function *PCA()* of the package FactoMineR (Husson et al., 2020). The eigenvalues following each principal component were rendered with the function *fviz_screeplot()* of the package factoextra v1.0.7 (Kassambara & Mundt, 2020). The species projections in the factorial plan (1, 2), (1, 3), and (1, 4) were rendered with the function *fviz_pca_ind()* of the package factoextra v1.0.7 (Kassambara & Mundt, 2020).

2.6 | Construction and deployment of Yam Microsatellite Markers Database

The Yam Microsatellite Markers Database (Y2MD) conception relied on three major steps detailed in Figure 1: (i) data compilation in SQL database, (ii) back- and front-end web construction, and (iii) tools integration.

In brief, using the generated SSR markers datasets, the Y2MD SQL database was built. Thus, the Y2MD server was implemented using Linux, Apache, MySQL, and the Hypertext Preprocessor (PHP) web application platform. The user-friendly web interface was then developed with the aid of JavaScript and the PHP. To enable the users to conduct their own analyses conveniently, additional functionalities such as JBrowse (Skinner et al., 2009), Blast (Altschul et al., 1990), SSR Finder (http://www.biophp.org/minitools/microsatellite_repeats_finder/), and *insilico*PCR (http://www.biophp.org/minitools/pcr_amplification/) were embedded. The website is freely accessible at https://y2md.ucad.sn/.

3 | RESULTS

3.1 | Detection, distribution, and characteristics of microsatellites detected on yam genomes

The genome sequences of the four yam species were utilized for the identification of SSRs. In total, 318,713, 322,501, 307,040, and 253,856 SSRs were detected in *D. alata*, *D.*

TABLE 1 Characteristics of the simple sequence repeat (SSR) primers used for genotyping and their positions on the reference genome *D. alata* TDa95.

Locus	Repeat	Chromosome	Start	End	Forward_primer (5′−3′)	Reverse_primer (5′−3′)
YMac2	(A)12	LG1	226,041	226,052	AGATGCCACGACGCAAACTT	GCAGCCGATCCTCAAAGGTC
YMar4	(TA)7	LG10	29,804,658	29,804,671	CAGGGCAAGAAGAGGGCAAG	GCTTTTCCAGTCACCTGTCCA
YMac15	(T)12	LG11	11,708,941	11,708,952	AGGCTTGGGCGAACTTTTGA	AGGGCTCGTCGTTTTGACCT
YMar11	(TTG)6	LG12	16,599,891	16,599,908	CTCTGCCAAGGCCAAGAAGG	AAACCAGCTGGCAGCATGTT
YMac33	(A)11	LG13	140,737	140,747	GCTTGCTGATAAAAACCGGAGC	TCTGGAGGAGTGTTGTCGAGA
YMar28	(TA)23	LG14	21,204,635	21,204,680	AGCTCATGGTTGGTTGCACA	GGCCATACACTTTGCCTCCC
YMac45	(AC)9	LG15	15,318,628	15,318,645	TGAGACCACTTCGTGTCTTGAG	GCACCTTCATGATGCAGCCT
YMar39	(A)10	LG17	17,125,279	17,125,288	AGACCCAAGCTGCAAAGCAA	TGACAGAGCCAACACAGAGGA
YMar44	(AT)13	LG18	24,532,026	24,532,051	TCCAGGAGCTTGGGTGATGT	GAAGACCTTCCCCACATGCA
YMac59	(ATC)7	LG19	2,365,155	2,365,175	CGCCCATTCCCTCAAAAGCT	AGCCGGGATGTTGGACAAAG
YMar65	(A)15	LG2	1,704,788	1,704,802	ACTTGCATACCCAACCCCAC	TACTGAGGAAGGTGCTGCCA
YMac82	(AT)18	LG20	13,142,390	13,142,425	GGCTTCATTGGCTCTGGCTT	GCCAGCAAGAAATGGACCCA
YMar79	(CCA)9	LG3	815,571	815,597	ACTATGTCAGCACCCACCCC	AAACTGGGGAGCAGTTGCAC
YMac92	(A)11	LG4	5,139,570	5,139,580	GCAGTTGACACACACGAGACA	CATGTGGATCGCTCATGCCA
YMar90	(GA)9	LG5	13,506,295	13,506,312	AGAGTGCCAGCCTCAGTTACA	AGGAAGGCGCTCTTTGATGC
YMac118	(A)10	LG6	12,894,570	12,894,579	TCATCCTCTCCCTGTCAGCG	GTAGTTGCTGCCATGGTGCA
YMar128	(GA)15	LG7	3,819,341	3,819,370	CGGCGACAAGAGAACAAGCA	CCACCTCCTTGAACATCGCC
YMac135	(TCA)13	LG9	16,767,472	16,767,510	TGCAGAAGTTCACCGCTCAG	CCTTCCTTGCTCAGCTGTGG
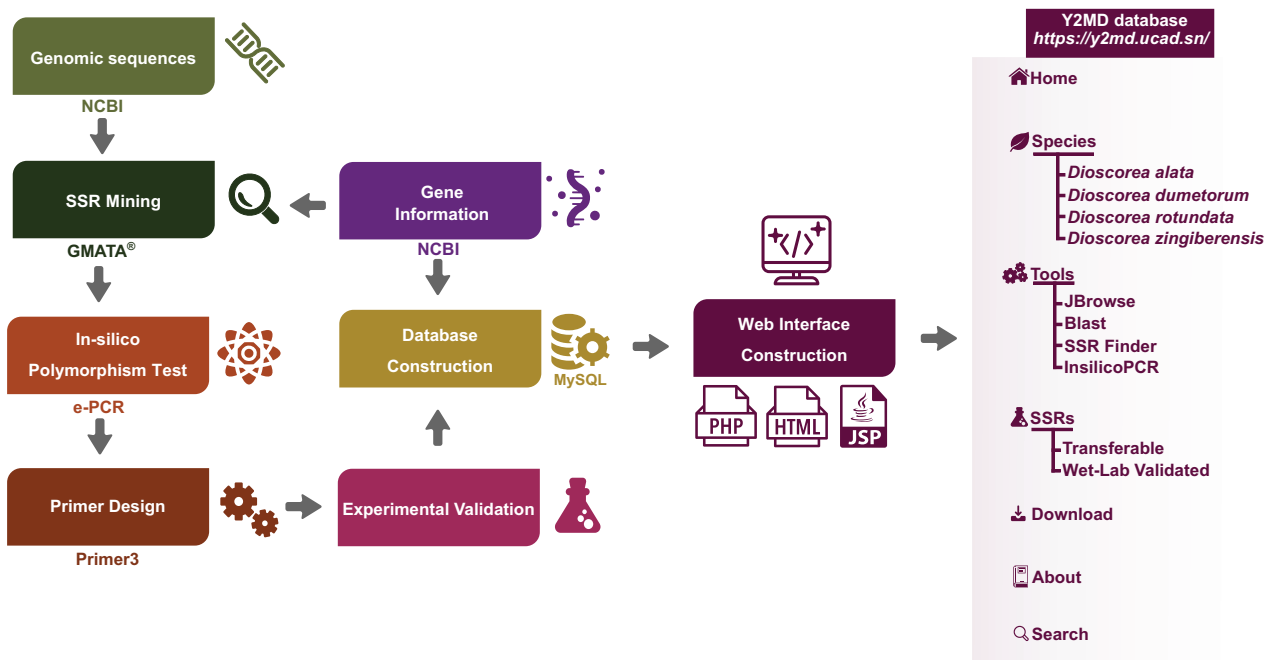
FIGURE 1 A diagram depicting the structure of Yam Microsatellite Markers Database (Y2MD). NCBI, National Center for Biotechnology Information; GMATA, Genome-wide Microsatellites Analyzing Tool Package.

rotundata, *D. dumetorum*, and *D. zingiberensis* genomes, respectively, with corresponding densities of 662.86, 552.22, 633.07, and 529.97 SSR/Mbp (Figure 2; Figure S1). At the intraspecific level, the genome of the *D. alata* cultivar Kabusa contained 346,986 SSRs with a density of 693.972 SSR/Mbp, surpassing that of *D. alata* cultivar TDa95/00328.

The number of SSRs per chromosome is strongly correlated with the chromosome lengths (Figure 3a). A total of 658

(a)

(b)

ⓐ Chromosome ⓑ Gene density ⓒ Perfect SSR density ⓓ Compound SSR density ⓔ Genic SSR density ⓕ Intergenic SSR density

(c)
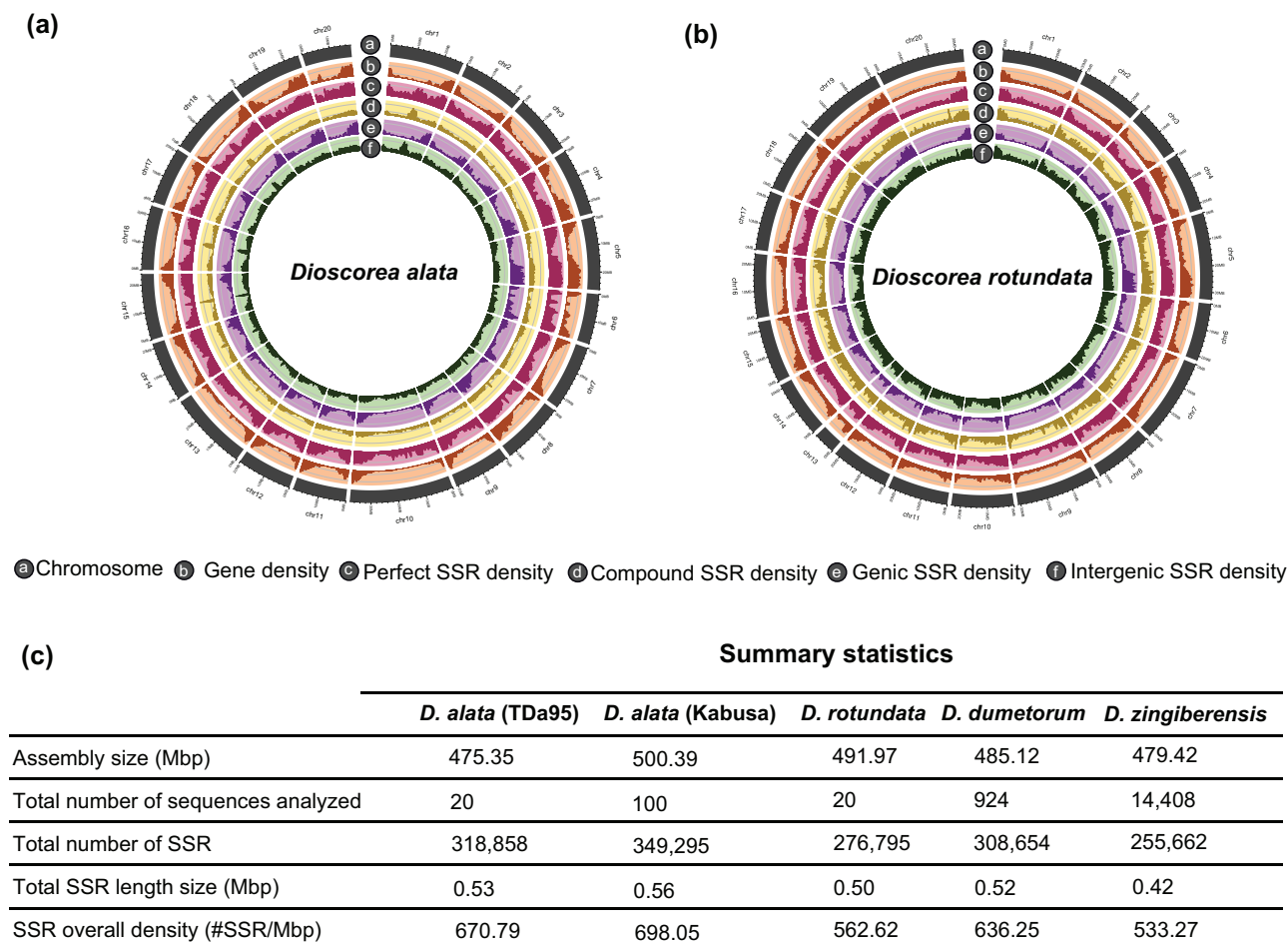
	Summary statistics				
	D. alata (TDa95)	**D. alata (Kabusa)**	**D. rotundata**	**D. dumetorum**	**D. zingiberensis**
Assembly size (Mbp)	475.35	500.39	491.97	485.12	479.42
Total number of sequences analyzed	20	100	20	924	14,408
Total number of SSR	318,858	349,295	276,795	308,654	255,662
Total SSR length size (Mbp)	0.53	0.56	0.50	0.52	0.42
SSR overall density (#SSR/Mbp)	670.79	698.05	562.62	636.25	533.27

FIGURE 2 Distribution and summary statistics of simple sequence repeats (SSRs) identified in yam genomes. Circos plots depicting the SSR density within the chromosome-scale genomes of *Dioscorea alata* (TDa95) (a) and *D. rotundata* (b). Summary statistics of SSR counts in the five studied genomes (c).

different motifs were observed in *D. alata*, 843 in *D. dumetorum*, 762 in *D. rotundata*, and 753 in *D. zingiberensis* (Figure 3b). Mono-, di-, and tri-nucleotides are the most important types of repeats in the different species (Figure 3c).

The obtained motifs are notably rich in A/T; A and T mononucleotides are the most represented, constituting an average of 55% of the detected SSRs. Additionally, 20% of the SSRs are composed of AT and TA dinucleotides. For tri-nucleotides, repeats (AAT/ATA/ATT/TAA/TTA) account for 7.5%. Hexa- and penta-nucleotides exhibit relatively low frequencies. Approximately 84% of the SSRs are localized in the intergenic regions of *D. alata* and *D. rotundata* genomes (Figure 3d). This analysis was not performed for other species due to the lack of proper genome annotation. Genes containing SSRs were analyzed for functional enrichment and showed enrichment in a wide array of biological attributes. Notably, several agronomically relevant gene ontology terms were identified, including response to water deprivation (GO:0009414), response to oxidative stress

(GO:0006979), response to light stimulus (GO:0009416), defense response (GO:0006952), and inflorescence development (GO:0010229) (Figure 3e; Table S3).

3.2 | Development of genome-wide SSR primers

A total of 864,128 unique primer pairs were generated from the 1,201,570 SSRs extracted from the four *Dioscorea* genomes: *D. alata* (221,606), *D. rotundata* (233,593), *D. dumetorum* (208,306), and *D. zingiberensis* (200,623). The distribution of primers across the genomes is uneven. The identified primers and their characteristics are detailed in Table S4. Each marker was assigned an identifier based on the respective *Dioscorea* species: YMa for *D. alata*, YMr for *D. rotundata*, YMd for *D. dumetorum*, and YMz for *D. zingiberensis*, followed by consecutive numbers representing the primer positions in the genomes.
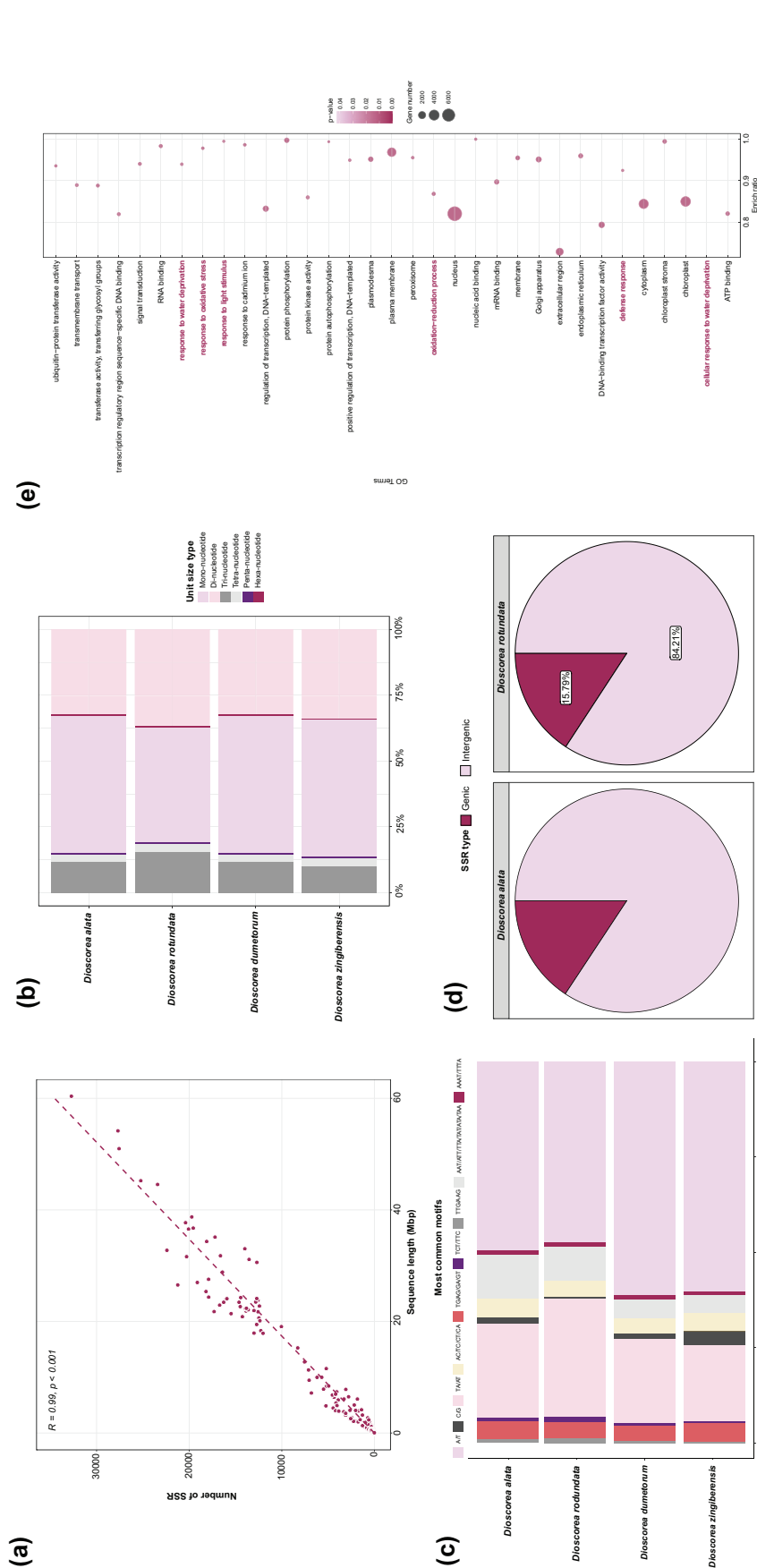
FIGURE 3 Correlation between chromosome length and number of simple sequence repeats (SSRs) (a); proportions of the nucleotide repeats type (b); distribution of the most common patterns in the four species (c); proportions of genic and non-genic SSRs in *D. alata* and *D. rotundata* (d); bubble plot showing significantly enriched gene ontology terms related to genic SSRs regions. The gene ontology terms highlighted in red are those that are agronomically relevant (e).
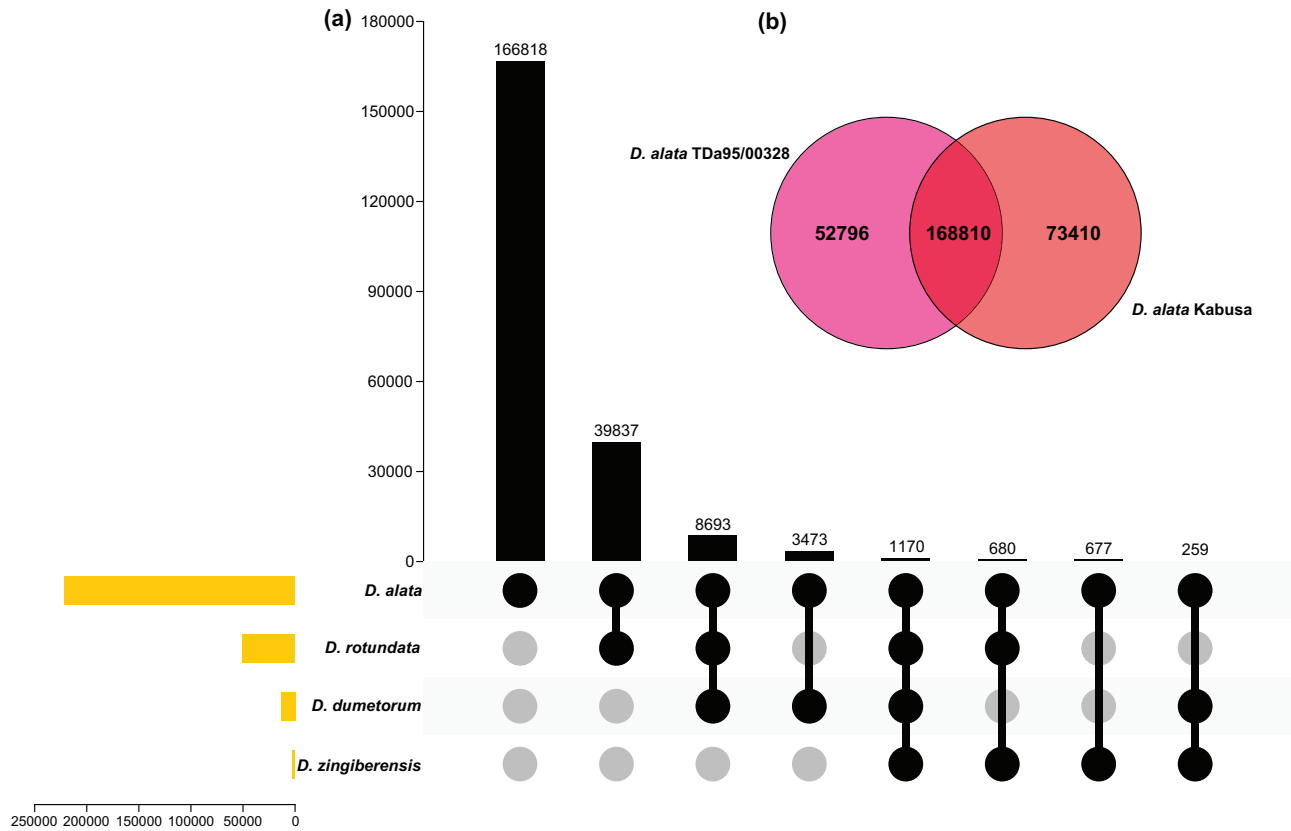
FIGURE 4 Cross-species and transferability of simple sequence repeats (SSRs) in yams. Upset plot showing the transferable markers between four yam species (*D. alata* as the reference genome, *D. rotundata*, *D. dumetorum*, and *D. zingiberensis*) (a); Venn diagram showing the number of common and specific SSR markers detected in two *D. alata* cultivars (TDa95/00328 and Kabusa) (b).

3.3 | Cross-species transferability of SSRs

First, we conducted a pairwise comparison of the microsatellite markers among the four yam species (Table S5). *Dioscorea alata* and *D. rotundata* displayed the highest conservation rate (54.78%), while *D. rotundata* and *D. zingiberensis* had the lowest conservation rate (1.64%). Next, to assess the cross-species transferability of the microsatellite markers, we used *D. alata* as a reference and performed e-PCR on the other three species. Out of 221,606 SSR markers from *D. alata* (TDa95/00328), 166,817 SSRs were specific to *D. alata*, while 39,837 SSRs could amplify *D. rotundata*, 3473 on *D. dumetorum*, and 677 on *D. zingiberensis* (Figure 4a). Overall, we identified 1170 SSR markers capable of amplifying all four species and potentially more yam species.

At the intraspecific level, we compared the SSR marker data identified in the genomes of two *D. alata* cultivars (TDa95/00328 and Kabusa). The results showed that both cultivars shared 77% of the total SSR markers (Figure 4b).

3.4 | Validation of selected SSRs and genotyping analysis

First, to check whether the previously developed SSR markers could be found within our dataset, we extracted a total of 92 SSR markers routinely used for genotyping *D. alata* and *D. rotundata* from the literature (Korsa et al., 2022; Loko et al., 2016; Nwogha et al., 2022; Otto et al., 2015; Tostain et al., 2007; Zawedee et al., 2014). This analysis was performed using the e-PCR method. Out of the 92 markers, 84 (91%) SSRs yielded PCR products on their respective reference genomes.

Next, to further validate our developed SSR markers, 18 markers were randomly selected from the 1170 cross-species transferable SSRs (Figure 4a). These markers are located on different chromosomes/contigs of the four species and were utilized for wet-lab PCR genotyping in 10 yam species (Table S2; Table 1). It is noteworthy to mention that we specifically targeted 18 SSRs for a practical reason: we aimed to use no more than three PCR plates (each containing 96 samples). Out of the 18 selected markers, 17 produced amplicons, with the

exception of the marker YMar128. Sixteen markers produced amplicons in at least six species out of the 10 (Table 2). Interestingly, the marker amplifications in diploid and tetraploid *D. alata* cultivars yielded a maximum of two or four alleles, respectively. For instance, the marker YMar79 was able to clearly delineate diploid from tetraploid forms of *D. alata* (Table 2).

Interestingly, four alleles were amplified for *D. transversa* (markers YMar44, YMar79, and YMac82), *D. cayenensis* (markers YMar39 and YMar44), *D. abyssinica* (markers YMar28 and YMac82), and *D. nummularia* (marker YMar79), indicating that the cultivars of these species used in this study are polyploids (4x or more). Moreover, a maximum of two amplified alleles were found for both cultivars of *D. rotundata*, suggesting that they are diploid types.

Based on the genotyping data, we conducted a hierarchical clustering analysis of the accessions and generated a tree. The tree highlighted a branch exclusively harboring *D. alata* cultivars (Figure 5).

Among *D. alata* cultivars, a clear distinction between diploid and tetraploid types was also noted, implying the robustness of the selected markers for discriminating yam accessions based on both species and ploidy criteria. The species *D. cayenensis* and *D. abyssinica* were singularly separated from the other species, while *D. transversa* and *D. nummularia* grouped together. Interestingly, the two cultivars of *D. rotundata* grouped together as expected. Meanwhile, two other groups were also identified: one consisting of *D. trifida* and *D. esculenta* and another group formed by *D. dumetorum* and *D. bulbifera*.

Interestingly, four alleles were amplified for *D. transversa* (markers YMar44, YMar79, and YMac82), *D. cayenensis* (markers YMar39 and YMar44), *D. abyssinica* (markers YMar28 and YMac82), and *D. nummularia* (marker YMar79), indicating that the cultivars of these species used in this study are polyploids (4x or more). Besides, a maximum of two amplified alleles were found for both cultivars of *D. rotundata*, suggesting that they are diploid types.

Based on the genotyping data, we performed a hierarchical clustering analysis of the accessions and generated a tree. The tree highlighted a branch exclusively harboring *D. alata* cultivars (Figure 5). Among *D. alata* cultivars, a clear distinction between diploid and tetraploid types was also noted, implying the robustness of the selected markers for discriminating yam accessions based on both species and ploidy criteria. The species *D. cayenensis* and *D. abyssinica* were singularly separated from the other species, while *D. transversa* and *D. nummularia* gathered together. Interestingly, the two cultivars of *D. rotundata* grouped together as expected. Meanwhile, two other groups were also identified: one consisting of *D. trifida* and *D. esculenta* and another group formed by *D. dumetorum* and *D. bulbifera*.

Meanwhile, from the principal component analysis (Figure 5b–e), a total of 53.7% of the total variance was retained in the top four factorial dimensions (Figure 5b). Based on the contribution of each species to the construction of factorial plans (1,2) (Figure 5c), plans (1,3) (Figure 5d), and plans (1,4) (Figure 5e), the species *D. nummularia* (Figure 5c), *D. bulbifera* (Figure 5c,d), and *D. trifida* (Figure 5e) clearly separated from their congeners. Moreover, a clear delineation of diploid and tetraploid *D. alata* species was also noticeable (Figure 5d), as observed in the clustering analysis.

3.5 | Development of Yam Microsatellite Markers Database

We established a comprehensive database by integrating SSR markers and whole-genome assemblies and annotation datasets (Figure 1). The Y2MD page contains data for four *Dioscorea* species and over 864,128 primer pairs. Additionally, all SSR data are downloadable in Excel, PDF, text, and CSV formats. The graphical user interface (GUI) menu has six tabs: "Home," "Species," "Tools," "SSRs," "Download," and "About."

The "Species" tab displays information on each species with the option to choose the SSR of choice based on the type of patterns, the number of repeats, the position, and so on. Information on experimentally validated SSRs and cross-species transferable SSRs is available in the "SSRs" tab. On the "Download" page, links to the different genomes are provided.

To enable database users to perform fundamental analyses, Y2MD incorporates several functionalities, including JBrowse (Figure 6a), *insilico*PCR (Figure 6b), SSR Finder (Figure 6c), and Y2MD Blast (Figure 6d). Therefore, Y2MD empowers users to find an SSR with the SSR Finder, conduct a quick amplification check with *insilico*PCR, and blast their own sequence of interest onto a genome using the Y2MD Blast tool. With JBrowse, users can explore the genome annotation map to identify potential genes related to a targeted SSR. Overall, Y2MD offers a dynamic platform for yam researchers to remotely utilize available genomic resources in a user-friendly GUI. The website is freely accessible at https://y2md.ucad.sn/.

4 | DISCUSSION

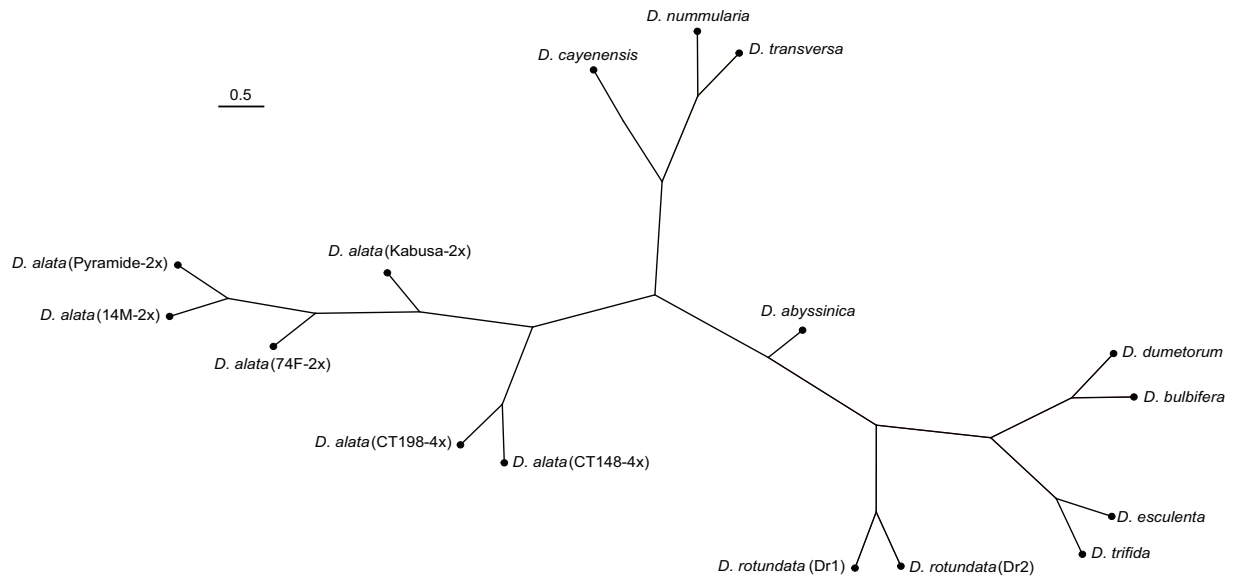### 4.1 | Developing genome-wide SSR markers in Dioscoreaceae

Microsatellite markers have proven to be effective in characterizing yam germplasms worldwide (Arnau et al., 2017;

TABLE 2 Allele scoring for 18 selected simple sequence repeat (SSR) markers in 10 yam species.
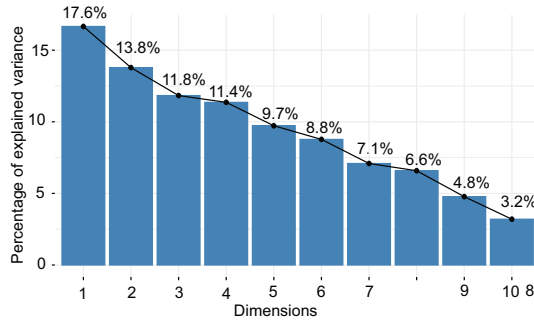
Marker	D. alata (14 M)	D. alata (74F)	D. alata (CT148)	D. alata (CT198)	D. alata (Pyramide)	D. alata (Kabusa)	D. bulbifera	D. transversa	D. rotoundata Dr1	D. rotoundata Dr2	D. trifida	D. dumetorum	D. esculenta	D. cayenensis	D. abyssinica	D. nummularia
YMar02	198	199	197 + 199	197 + 199	198 + 199	196 + 197	212 + 213	198 + 199	198 + 199	198	215	209	204	198 + 199	197 + 200	196 + 199 + 204
YMar04	313 + 325	325	309 + 313	309 + 313	313 + 325	313 + 327	333 + 335	313 + 315	322	322	343	349	322	320 + 322	322	311
YMar11	303	303	303	303	303	303	304	303	317	311 + 317	296	304	308 + 311 + 317	303	314	302
YMar118	196	202	202	202	196	202	194	202	198	196	214	NA	NA	197	197	198
YMar128	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
YMar135	160	160	157 + 160 + 176	157 + 160 + 176	160	176	163 + 169	163 + 166 + 172	160	160	176	160 + 163	NA	157 + 160	163	163 + 166 + 172
YMar15	158	158	158 + 159	158 + 159	158	158	NA	155 + 158	157	157	155	152	155	153 + 157	156 + 157	155 + 158
YMar28	234 + 254	232 + 248	251 + 260	251 + 260	226 + 230	224	234	224	234 + 248	230 + 232	224	NA	NA	236 + 260	232 + 234 + 238 + 240	224
YMar33	181 + 183	183 + 184	181 + 184	181 + 184	183	184	170 + 189	176 + 184	184	184	NA	186	184 + 185	184	182	182
YMar39	180	180	180 + 181	181	180	180 + 182	170	154 + 157 + 181	173 + 176	176	166	171	185	173 + 176 + 1778 + 186	154 + 173	154 + 157
YMar44	192 + 206	NA	208 + 212	196 + 208	192 + 204	196 + 206	NA	184 + 192 + 196 + 200	190 + 192 + 196	190 + 196	NA	NA	NA	190 + 194 + 192 + 198 + 208	192	192 + 196
YMar45	277 + 281	277	281	281	277 + 281	279 + 281	NA	275 + 277 + 281	273	273	NA	281	259	267 + 275	275	275
YMar59	298	298	NA	NA	298	298	310	298 + 301	307 + 310	310 + 319	NA	NA	NA	310	307	301
YMar65	274 + 275	NA	274 + 277	274 + 277	275	273 + 275	NA	274	NA	NA	NA	NA	NA	NA	NA	NA
YMar79	217	208 + 217	207 + 211 + 223	207 + 211 + 220 + 223	217 + 223	220	198	207 + 211 + 221 + 225	215 + 220	214	223	236 + 242	223	208 + 214	212 + 215	221 + 225 + 229 + 236
YMar82	238 + 254	238 + 254	226 + 238 + 256	226 + 256	252 + 254	228 + 234	NA	230 + 238 + 246 + 252	238	238	NA	NA	NA	238 + 248 + 254	232 + 242 + 246 + 252	236 + 246 + 252
YMar90	294	292	292	292	292 + 304	292 + 294	332	294 + 310	306	296 + 298	332 + 340	NA	NA	296 + 314	310	294 + 310
YMar92	215	215	NA	NA	214	215 + 219	209	211 + 214	212 + 215	211 + 212	NA	215 + 223	219	211 + 212	216	211

Abbreviation: NA, not available.

(a)

(b)

(c)

(d)

(e)

● *D. nummularia*   ● *D. esculenta*   ● *D. transversa*   ● *D. dumetorum*   ● *D. rotundata*   ● *D. alata* (4x)

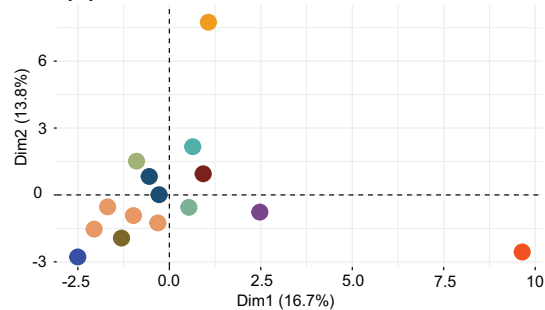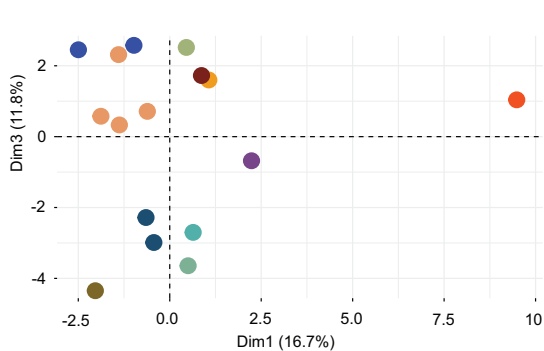● *D. bulbifera*   ● *D. abyssinica*   ● *D. trifida*   ● *D. cayenensis*   ● *D. alata* (2x)
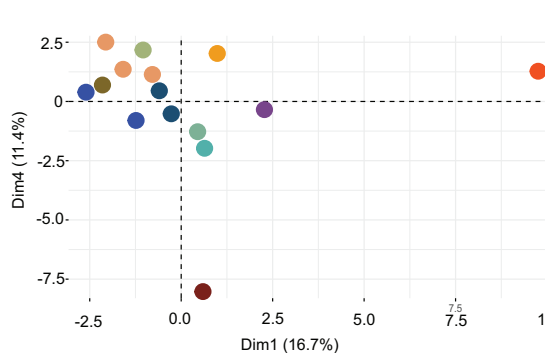
FIGURE 5 Hierarchical clustering tree depicting the relationships between 10 yam species using 17 simple sequence repeat (SSR) markers (a). The tree was constructed based on the Unweighted pair group method with arithmetic mean method. Taxa labels were arranged as follows: scientific name, genotype name, and ploidy. Ploidy information was not mentioned for the species with undetermined ploidy. The scale bar represents the length of the tree branch. Principal component analysis depicting the variability retained on each principal component (b), the projection of each species following the factorial plans (1, 2) (c) (1, 3) (d), and (1, 4) (e).

FIGURE 6 The Yam Microsatellite Markers Database (Y2MD). The website is freely available at https://y2md.ucad.sn/. The web interface of Y2MD showcasing JBrowse (a), *insilico*PCR (b), SSR Finder (c), and Blast (d) functionalities

Cao et al., 2021; Obidiegwu, Asiedu, et al., 2009; Obidiegwu, Kolesnikova-Allen, et al., 2009; Siqueira et al., 2011; Tamiru et al., 2015; Tostain et al., 2006; H. Wang et al., 2022). The present study enabled the extraction of a significant number of microsatellites in different yam genomes. The extracted SSRs were more concentrated in the telomeres, as reported by Jian et al. (2021). *Dioscorea zingiberensis* (253,856 microsatellites) had fewer microsatellites compared to other yam species. Nevertheless, we observed a high correlation ($R^2 = 0.99$) between the SSR numbers and chromosome lengths. Our findings are similar to the report of Dossa et al. (2017) in sesame, where the relationship between chromosome length and the number of SSRs on each chromosome showed a strong correlation ($R^2 = 0.94$).

Overall, the SSR densities observed in yam species (ranging from 529 to 662 SSR/Mbp) are similar to those of sesame (507 SSR/Mbp) but lower than other species such as Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*) (875 and 807 SSRs/Mbp, respectively) (Dossa et al., 2017; Lawson et al, 2006). Comparing two genome assemblies of *D. alata* (Kabusa and TDa95/00328) highlighted major differences in terms of SSR numbers and densities. We speculate that the completeness of the genome assemblies and the genetic distances (data unpublished) between the two accessions may explain these differences.

Among the observed motifs, mononucleotides, which account for 50% of SSRs, are the most represented. However, they are not ideal targets for PCR markers (Clarke, 2001). The most abundant types of SSRs are mono (A/T)n, di (AT/TA)n, poly (ATA/ATT/TAT/TAA)n, and

therefore rich in AT. Similar results were obtained in potato (*Solanum tuberosum*) (Jian et al., 2021), Arabidopsis (Cavagnaro et al., 2010), sesame (Uncu et al., 2015), sorghum (*Sorghum bicolor*) (Yonemaru et al., 2009), and rice (McCouch et al., 2002).

A cost comparison of SSR and SNP genotyping systems in yam is difficult to make, mainly because no SNP array has yet been released for the yam community. Based on our previous genotyping experience, we spent €325 as a full-cost package for genotyping 16 yam samples with 18 SSRs (~€20 per sample). In comparison, we spent €2150 to genotype 120 yam samples with 77 KASPar SNPs (around €17 per sample). Although the cost of SSR genotyping is slightly higher than SNP markers in our case, only 72% of the KASPar SNP genotyping data were usable, compared to 94% for SSR markers. Larsen et al. (2018) revealed that an accurate cost comparison between SSR marker and SNP marker systems is quite impossible. They showed that the costs of generating SNP data were 10 times higher than the laboratory costs of obtaining the SSR data in apple. But they specified that generating SSR data was much more labor-intensive than SNP genotyping since it requires many manual steps. Bernard et al. (2019) showed that the use of only 13 SSR markers was sufficient to cover the genetic structure in the walnut germplasm collection, while hundreds of SNPs were required to achieve the same level of resolution. Similarly, Choudhury et al. (2023) demonstrated that 30 unlinked SSR markers were more powerful than 32,782 SNPs for diversity analysis of Indian rice. In addition to these advantages, the SSR marker system can be fully integrated into laboratories. Nonetheless, SNP

genotyping is still a better option when working on large yam germplasms.

4.2 | Cross-species transferability of the SSR markers

The identification of polymorphic SSR markers will support research on yam genetics and breeding. At the genome scale, we identified 1170 cross-species transferable markers in the studied four *Dioscorea* species. Previously, Tamiru et al. (2015) showed that 94.4% and 56.7% of SSR markers from *D. cayenensis* could be successfully transferred to *D. rotundata* and *D. alata*, respectively. Likewise, Tostain et al. (2006) developed 16 polymorphic SSR markers transferable to *D. alata*, *D. abyssinica*, and *D. praehensilis*. Transferring markers between species is often considered a better investment than developing and testing new markers, especially when available funding is limited (Fan et al., 2013). Because, in most breeding programs, different yam species are grown and evaluated together, having SSR marker sets applicable to different species is highly useful. We also noticed that the number of cross-species transferable markers decreased significantly when *D. zingiberensis* was included. Indeed, Bredeson et al. (2022) compared the *D. alata* reference genome with the genomes of *D. rotundata* and *D. zingiberensis* and revealed substantial conservation of chromosome structures between *D. alata* and *D. rotundata*, but to a lesser extent with *D. zingiberensis*.

When several genome assemblies are available within the same species, in silico detection of polymorphic markers is resource-efficient in terms of time and cost. This approach has been successfully applied in sesame, spinach, and *Perca fluviatilis* (Bhattarai et al., 2021; Dossa et al., 2017; Xu et al., 2022). We compared the common SSR markers from the genomes of *D. alata* (TDa95/00328 and Kabusa) and obtained 93.92% polymorphic SSR markers. In other species, such as potato, tomato, and bell pepper, the levels of intraspecific polymorphisms appear to be rather limited (Stàgel et al., 2008); therefore, more efforts were needed to develop informative SSR markers in these species. In *D. alata*, the high SSR polymorphic rates could be attributed to the dioecy nature of the species.

4.3 | Validating selected SSRs

In *Dioscoreaceae*, only a limited number of SSR markers have undergone experimental (wet-lab) validation (Andris et al., 2010; Arnau et al., 2017; Cao et al., 2021; Malédon et al., 2023; Obidiegwu, Asiedu, et al., 2009; Obidiegwu, Kolesnikova-Allen, et al., 2009; Siqueira et al., 2011; Tamiru et al., 2015; Tostain et al., 2006; H. Wang et al., 2022), most of

which originate from *D. alata* and *D. rotundata*. In contrast, other species such as rice (McCouch et al., 2002), sesame (Dossa et al., 2017), and cowpea (*Vigna unguiculata*) (Jasrotia et al., 2019) have a large number of markers that have been identified, experimentally characterized, and made available for genetic studies and breeding.

In this study, we extracted 92 markers from the literature (markers tested through wet-lab experiments), and most of them were found among the markers identified in our study. The e-PCR parameters, which allow only two mismatches and one insertion for the primers, could be partly responsible for the absence of PCR products for the wet-lab tested markers that are missing. This analysis also provided the exact locations of these SSRs in the yam genomes, enabling researchers to select markers from different chromosomes for enhanced discriminatory power.

Out of the 18 selected SSR markers from this study, 17 were able to PCR amplify genotypes of several yam species. The experimentally tested markers were from different chromosomes and demonstrated robust discriminatory power by clearly delineating the studied yam species and ploidy levels. In yam species, the basic chromosome number is 20, and most species harbor different cytotypes (Abraham et al., 2013; Arnau et al., 2009; Bousalem et al., 2010; Scarcelli et al., 2005; Sughihara et al., 2021). Prior to our work, *D. rotundata* cultivars were reported to be mainly diploid using flow cytometry, isoenzymes, DArTseq SNP, and a set of six microsatellite markers (Gatarira et al., 2021; Scarcelli et al., 2005). Both *D. rotundata* cultivars investigated in this study were found to be diploid. Additionally, the tested SSR markers also showed promising accuracy even at the intra-species level with a clear classification of diploid and tetraploid forms of *D. alata* cultivars. Based on the allele numbers, the yam species, including *D. transversa*, *D. abyssinica*, *D. cayenensis*, and *D. nummularia*, have cultivars with tetra- or higher ploidy levels (Gatarira et al., 2021; Sughihara et al., 2021). Therefore, we provided in the present study SSR markers that can be used at low cost for ploidy checks and species differentiation in yams.

4.4 | Deployment of Yam Microsatellite Markers Database for yam breeders

Several databases containing SSR marker data from various species are available to the public. However, there is currently no database dedicated to the study of yam microsatellites. Establishing a yam microsatellite marker database would serve as a valuable platform for advancing genetic evaluation, genomic research, and yam breeding. While the Pan Species Microsatellite Database (Du et al., 2020) includes SSR data from 18,408 organisms, its utility is limited due to the absence of features such as ePCR, JBrowse (Buels et al., 2016), and Blast (Altschul et al., 1990). This limitation is also observed

in the Tea Microsatellite Database (Dubey et al., 2020) and SSRome (Mokhtar & Atia, 2019).

In this study, we have compiled all generated microsatellite data into the Y2MD. Y2MD incorporates various useful tools, including JBrowse, Blast, SSR Finder, and *insilico*PCR. These tools are designed to facilitate the localization, SSR detection, and polymorphism analysis of both existing and new experimental yam SSR markers/nucleotide sequences. Y2MD has been developed as a highly user-friendly database, enabling all users, especially those with limited knowledge or resources in bioinformatics, to analyze their SSR-related data.

5 | CONCLUSION

In this study, a graphical interface containing the SSR markers of four commercially important *Dioscorea* species was constructed and made available to the public via https://y2md.ucad.sn/. To perform the database construction, microsatellites from four *Dioscorea* species were first extracted, and primers were designed based on the flanking sequences of these microsatellites. In total, 1,201,570 microsatellites were extracted from all species, resulting in 864,128 designed primer pairs. Subsequently, additional analyses were conducted, including cross-species transferability and determination of potential polymorphic markers, which are determining factors in plant genotyping. Experimental validation of selected markers was also performed to lend credence to our datasets. Overall, the SSRs developed in this study will be useful for the genetic characterization of yam germplasms. All these data have been compiled into Y2MD and will be of great use to the yam community.

With the rapid development in genome sequencing projects, we anticipate more yam genomes to become available. Consequently, we plan to expand Y2MD to include new *Dioscorea* species as soon as their genome sequences are accessible. Additionally, Y2MD will be enhanced by incorporating new tools and functionalities, such as the integration of Primer3, introduction of molecular markers associated with QTLs, and inclusion of new wet-lab validated SSRs. We also intend to integrate SNP marker data into Y2MD, evolving it into a new platform with various multi-omics data, such as transcriptome and metabolome data.

AUTHOR CONTRIBUTIONS
Moussa Diouf: Data curation; formal analysis; visualization; writing—original draft. **Yedomon Ange Bovys Zoclanclounon**: Data curation; formal analysis; visualization; writing—original draft. **Pape Adama Mboup**: Data curation; formal analysis. **Diaga Diouf**: Conceptualization; data curation; formal analysis; writing—review and editing. **Erick Malédon**: Formal analysis; validation. **Ronan Rival-

lan**: Formal analysis; validation. **Hâna Chair**: Formal analysis; validation; writing—review and editing. **Komivi Dossa**: Conceptualization; data curation; formal analysis; funding acquisition; methodology; project administration; resources; supervision; validation; writing—review and editing.

CONFLICT OF INTEREST STATEMENT
The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT
All SSR data generated and the genomes used are available online at https://y2md.ucad.sn/. The genome sequence of *Dioscorea alata* cultivar Kabusa has not been made available publicly yet, but it can be shared upon request to the corresponding author.

ORCID
Komivi Dossa https://orcid.org/0000-0003-4894-6279

REFERENCES
Abraham, K., Nemorin, A., Lebot, V., & Arnau, G. (2013). Meiosis and sexual fertility of autotetraploid clones of greater yam *Dioscorea alata* L. *Genetic Resources and Crop Evolution*, *60*, 819–823. https://doi.org/10.1007/s10722-013-9973-4

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Amiteye, S. (2021). Basic concepts and methodologies of DNA marker systems in plant molecular breeding. *Heliyon*, *7*, e08093. https://doi.org/10.1016/j.heliyon.2021.e08093

Andris, M., Aradottir, G. I., Arnau, G., Audzijonyte, A., Bess, E. C., Bonadonna, F., bourdel, G., Bried, J., Bugbee, G. J., Burger, P. A., Chair, H., Charruau, P. C., Ciampi, A. Y., Costet, L., DeBarro, P. J., Delatte, H., Dubois, M., Eldridge, M. D. B., England, P. R., … Wright, M. (2010). Permanent genetic resources added to Molecular Ecology Resources Database 1 June 2010 – 31 July 2010. *Molecular Ecology Resources*, *10*, 1106–1108. https://doi.org/10.1111/j.1755-0998.2010.02916.x

Arnau, G., Bhattacharjee, R., Mn, S., Chair, H., Malapa, R., Lebot, V. K A., Perrier, X., Petro, D., Penet, L., & Pavis, C. (2017). Understanding the genetic diversity and population structure of yam (*Dioscorea alata* L.) using microsatellite markers. *PLoS One*, *12*, e0174150. https://doi.org/10.1371/journal.pone.0174150

Arnau, G., Nemorin, A., Maledon, E., & Abraham, K. (2009). Revision of ploidy status of *Dioscorea alata* L. (Dioscoreaceae) by cytogenetic and microsatellite segregation analysis. *Theoretical and Applied Genetics*, *118*, 1239–1249. https://doi.org/10.1007/s00122-009-0977-6

Asiedu, R., & Sartie, A. (2010). Crops that feed the World1. Yams. *Food Security*, *2*, 305–315. https://doi.org/10.1007/s12571-010-0085-0

Bernard, A., Barreneche, T., Donkpegan, A., Lheureux, F., & Dirlewanger, E. (2019). *Management of a walnut germplasm collection: Which of SSR or SNP markers are most suitable to preserve biodiversity?* https://hal.inrae.fr/hal-02790842

Bhattarai, G., Shi, A., Kandel, D. R., Solís-Gracia, N., da Silva, J. A., & Avila, C. A. (2021). Genome-wide simple sequence repeats (SSR) markers discovered from whole-genome sequence comparisons of multiple spinach accessions. *Scientific Reports*, *11*, 9999. https://doi.org/10.1038/s41598-021-89473-0

Bousalem, M., Viader, V., Mariac, C., Gomez, R.-M., Hochu, I., Santoni, S., & David, J. (2010). Evidence of diploidy in the wild Amerindian yam, a putative progenitor of the endangered species *Dioscorea trifida* (Dioscoreaceae). *Genome*, *53*, 371–383. https://doi.org/10.1139/G10-016

Bredeson, J. V., Lyons, J. B., Oniyinde, I. O., Okereke, N. R., Kolade, O., Nnabue, I., Nwadili, C. O., Hříbová, E., Parker, M., Nwogha, J., Shu, S., Carlson, J., Kariba, R., Muthemba, S., Knop, K., Barton, G. J., Sherwood, A. V., Lopez-Montes, A., Asiedu, R., … Rokhsar, D. S. (2022). Chromosome evolution and the genetic basis of agronomically important traits in greater yam. *Nature Communications*, *13*, 1–16. https://doi.org/10.1038/s41467-022-29114-w

Bruford, M. W., & Wayne, R. K. (1993). Microsatellites and their application to population genetic studies. *Current Opinion in Genetics & Development*, *3*, 939–943. https://doi.org/10.1016/0959-437X(93)90017-J

Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., Wu, Y., Zhao, L., Liu, J., Guo, J., Fang, S., Cao, W., Yi, L., Zhao, Y., & Kong, L. (2021). KOBAS-i: Intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Research*, *49*, W317–W325. https://doi.org/10.1093/nar/gkab447

Buels, R., Yao, E., Diesh, C. M., Hayes, R. D., Munoz-Torres, M., Helt, G., Goodstein, D. M., Elsik, C. G., Lewis, S. E., Stein, L., & Holmes, I. H. (2016). JBrowse: A dynamic web platform for genome visualization and analysis. *Genome Biology*, *17*, 66. https://doi.org/10.1186/s13059-016-0924-1

Cao, T., Sun, J., Shan, N., Chen, X., Wang, P., Zhu, Q., Xiao, Y., Zhang, H., Zhou, Q., & Huang, Y. (2021). Uncovering the genetic diversity of yams (*Dioscorea* spp.) in China by combining phenotypic trait and molecular marker analyses. *Ecology and Evolution*, *11*, 9970–9986. https://doi.org/10.1002/ece3.7727

Cavagnaro, P. F., Senalik, D. A., Yang, L., Simon, P. W., Harkins, T. T., Kodira, C. D., Huang, S., & Weng, Y. (2010). Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics*, *11*, 569. https://doi.org/10.1186/1471-2164-11-569

Chen, J., Liu, Q., Lu, C., Liu, Q., Pan, J., Zhang, J., & Dong, S. (2022). Genetic diversity of *Prunus armeniaca* L. var. ansu Maxim. germplasm revealed by simple sequence repeat (SSR) markers. *PLoS One*, *17*, e0269424. https://doi.org/10.1371/journal.pone.0269424

Cheng, J., Chen, J., Liu, X., Li, X., Zhang, W., Dai, Z., Lu, L., Zhou, X., Cai, J., Zhang, X., Jiang, H., & Ma, Y. (2021). The origin and evolution of the diosgenin biosynthetic pathway in yam. *Plant Communications*, *2*, 100079. https://doi.org/10.1016/j.xplc.2020.100079

Choudhury, D. R., Kumar, R., Maurya, A., Semwal, D. P., Rathi, R. S., Gautam, R. K., Trivedi, A. K., Bishnoi, S. K., Ahlawat, S. P., Singh, K., & Singh, N. K. (2023). SSR and SNP marker-based investigation of Indian rice landraces in relation to their genetic diversity, population structure, and geographical isolation. *Agriculture*, *13*, 823. https://doi.org/10.3390/agriculture13040823

Clarke, L. A. (2001). PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences. *Molecular Pathology*, *54*, 351–353. https://doi.org/10.1136/mp.54.5.351

Cormier, F., Lawac, F., Maledon, E., Gravillon, M. C., Nudol, E., Mournet, P., Vignes, H., Chaïr, H., & Arnau, G. (2019). A reference high-density genetic map of greater yam (*Dioscorea alata* L.). *Theoretical and Applied Genetics*, *132*, 1733–1744. https://doi.org/10.1007/s00122-019-03311-6

Cummings, B., & Wood, T. (1989). A simple and efficient method for isolating genomic DNA from endomycorrhizal spores. *Gene Analysis Techniques*, *6*, 89–92. https://doi.org/10.1016/0735-0651(89)90013-7

Darkwa, K., Olasanmi, B., Asiedu, R., & Asfaw, A. (2020). Review of empirical and emerging breeding methods and tools for yam (Dioscorea spp.) improvement: Status and prospects. *Plant Breeding*, *139*, 474–497. https://doi.org/10.1111/pbr.12783

de Freitas, K. E. J., Busanello, C., Viana, V. E., Pegoraro, C., de Carvalho Victoria, F., da Maia, L. C., & Costa de Oliveira, A. (2022). An empirical analysis of mtSSRs: Could microsatellite distribution patterns explain the evolution of mitogenomes in plants? *Functional & Integrative Genomics*, *22*, 35–53. https://doi.org/10.1007/s10142-021-00815-7

Dong, B., Deng, Z., Liu, W., Rehman, F., Yang, T.-J., Huang, Y., & Gong, H. (2022). Development of expressed sequence tag simple sequence repeat (EST-SSR) markers and genetic resource analysis of tea oil plants (*Camellia* spp.). *Conservation Genetics Resources*, *14*, 41–45. https://doi.org/10.1007/s12686-021-01248-x

Dossa, K., Yu, J., Liao, B., Cisse, N., & Zhang, X. (2017). Development of highly informative genome-wide single sequence repeat markers for breeding applications in sesame and construction of a web resource: SisatBase. *Frontiers in Plant Science*, *8*, 1–10. https://doi.org/10.3389/fpls.2017.01470

Du, L., Liu, Q., Zhao, K., Tang, J., Zhang, X., Yue, B., & Fan, Z. (2020). PSMD: An extensive database for pan-species microsatellite investigation and marker development. *Molecular Ecology Resources*, *20*, 283–291. https://doi.org/10.1111/1755-0998.13098

Dubey, H., Rawal, H. C., Rohilla, M., Lama, U., Kumar, P. M., Bandyopadhyay, T., Gogoi, M., Singh, N. K., & Mondal, T. K. (2020). TeaMiD: A comprehensive database of simple sequence repeat markers of tea. *Database*, *2020*, 1–14. https://doi.org/10.1093/database/baaa013

Duhan, N., & Kaundal, R. (2021). LegumeSSRdb: A comprehensive microsatellite marker database of legumes for germplasm characterization and crop improvement. *International Journal of Molecular Sciences*, *22*, 11350. https://doi.org/10.3390/ijms222111350

Egesi, C. N., Asiedu, R., Ude, G., Ogunyemi, S., & Egunjobi, J. K. (2006). AFLP marker diversity in water yam (*Dioscorea alata* L.). *Plant Genetic Resources*, *4*, 181–187. https://doi.org/10.1079/PGR2006121

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics*, *5*, 435–445. https://doi.org/10.1038/nrg1348

Fan, L., Zhang, M.-Y., Liu, Q.-Z., Li, L.-T., Song, Y., Wang, L.-F., Zhang, S.-L., & Wu, J. (2013). Transferability of newly developed pear SSR markers to other Rosaceae species. *Plant Molecular Biology Reporter*, *31*, 1271–1282. https://doi.org/10.1007/s11105-013-0586-z

FAOSTAT. (2021). https://www.fao.org/faostat

Gatarira, C., Sládeková, L., Němečková, A., Šimoníková, D., Paliwal, R., Asfaw, A., Abberton, M., Badara, G., Asiedu, R., Čížková,

J., Hřibová, E., Doležel, J., DeKoeyer, D., Adetimirin, V., & Bhattacharjee, R. (2021). Cytological and molecular characterization for ploidy determination in yams (*Dioscorea* spp.). *Agronomy*, *11*, 1897. https://doi.org/10.3390/agronomy11101897

Herrera, M. D. R., & Ghislain, M. (2013). Robust and inexpensive SSR markers analyses using LI-COR DNA analyzer. In S. K. Kantartzi (Ed.), *Microsatellites: Methods and protocols* (pp. 197–205). Humana Press.

Hochu, I., Santoni, S., & Bousalem, M. (2006). Isolation, characterization and cross-species amplification of microsatellite DNA loci in the tropical American yam *Dioscorea trifida*. *Molecular Ecology Notes*, *6*, 137–140. https://doi.org/10.1111/j.1471-8286.2005.01166.x

Husson, A. F., Josse, J., Le, S., Mazet, J., & Husson, M. F. (2020). FactoMineR: Multivariate exploratory data analysis and data mining. R package verson 2.3. http://factominer.free.fr100

Jasrotia, R. S., Yadav, P. K., Iquebal, M. A., Bhatt, S. B., Arora, V., Angadi, U. B., Tomar, R. S., Jaiswal, S., Rai, A., & Kumar, D. (2019). VigSatDB: Genome-wide microsatellite DNA marker database of three species of Vigna for germplasm characterization and improvement. *Database*, *2019*, baz055. https://doi.org/10.1093/database/baz055

Jha, U. C., Jha, R., Bohra, A., Manjunatha, L., Saabale, P. R., Parida, S. K., Chaturvedi, S. K., Thakro, V., & Singh, N. P. (2021). Association mapping of genomic loci linked with *Fusarium* wilt resistance (Foc 2) in chickpea. *Plant Genetic Resources: Characterization and Utilization*, *19*, 195–202. https://doi.org/10.1017/S1479262121000228

Jian, Y., Yan, W., Xu, J., Duan, S., Li, G., & Jin, L. (2021). Genome-wide simple sequence repeat markers in potato: Abundance, distribution, composition, and polymorphism. *DNA Research*, *28*, 1–9. https://doi.org/10.1093/dnares/dsab020

Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and visualize the results of multivariate data analyses*. R package version 1.0.7. http://CRAN.R-project.org/package=factoextra

Kim, H. R., Sa, K. J., Nam-Gung, M., Park, K. J., Ryu, S.-H., Mo, C. Y., & Lee, J. K. (2021). Genetic characterization and association mapping in near-isogenic lines of waxy maize using seed characteristics and SSR markers. *Genes & Genomics*, *43*, 79–90. https://doi.org/10.1007/s13258-020-01030-7

Korsa, F., Feyissa, T., Dessalegn, O., & Mekonnen, T. (2022). Genetic diversity and population structure of Yam (*Dioscorea* species) from Western Ethiopia as revealed by simple sequence repeat markers. Research Square. https://doi.org/10.21203/rs.3.rs-1905131/v1

Larsen, B., Gardner, K., Pedersen, C., Ørgaard, M., Migicovsky, Z., Myles, S., & Toldam-Andersen, T. B. (2018). Population structure, relatedness and ploidy levels in an apple gene bank revealed through genotyping-by-sequencing. *PLoS One*, *13*(8), e0201889. https://doi.org/10.1371/journal.pone.0201889

Lawson, M. J., Zhang, L., & Tech, V. (2006). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biology*, *7*, 1–11. https://doi.org/10.1186/gb-2006-7-2-r14

Li, C., Miao, H., Wei, L., Zhang, T., Han, X., & Zhang, H. (2014). Association mapping of seed oil and protein content in *Sesamum indicum* L. using SSR markers. *PLoS One*, *9*, e105757. https://doi.org/10.1371/journal.pone.0105757

Li, D., Pan, S., Zhang, H., Fu, Y., Peng, Z., Zhang, L., Peng, S., Xu, F., Huang, H., Shi, R., Zheng, H., Peng, Y., & Tan, Z. (2021). A comprehensive microsatellite landscape of human Y-DNA at kilobase resolution. *BMC Genomics*, *22*, 76. https://doi.org/10.1186/s12864-021-07389-5

Loko, Y. L., Bhattacharjee, R., Agre, A. P., Dossou-Aminon, I., Orobiyi, A., Djedatin, G. L., & Dansi, A. (2016). Genetic diversity and relationship of Guinea yam (*Dioscorea cayenensis* Lam.–D. rotundata Poir. complex) germplasm in Benin (West Africa) using microsatellite markers. *Genetic Resources and Crop Evolution*, *64*, 1205–1219. https://doi.org/10.1007/s10722-016-0430-z

Malapa, R., Arnau, G., Noyer, J. L., & Lebot, V. (2005). Genetic diversity of the greater yam (*Dioscorea alata* L.) and relatedness to *D. nummularia* Lam. and *D. transversa* Br. as revealed with AFLP markers. *Genetic Resources and Crop Evolution*, *52*, 919–929. https://doi.org/10.1007/s10722-003-6122-5

Malédon, E., Nudol, E., Perrot, C., Gravillon, M.-C., Rivallan, R., Cornet, D., Chair, H., & Dossa, K. (2023). First report of a successful development of yam hybrids (*Dioscorea alata* L.) from lyophilized and long-term stored pollens. *Phyton, Vicente Lopez*, *92*, 2861–2874. https://doi.org/10.32604/phyton.2023.042397

Martina, M., Acquadro, A., Barchi, L., Gulino, D., Brusco, F., Rabaglio, M., Portis, F., Portis, E., & Lanteri, S. (2022). Genome-wide survey and development of the first microsatellite markers database (AnCorDB) in *Anemone coronaria* L. *International Journal of Molecular Sciences*, *23*, 3126. https://doi.org/10.3390/ijms23063126

McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z., Xing, Y., Zhang, Q., Kono, I., Yano, M., Fjellstrom, R., DeClerck, G., Schneider, D., Cartinhour, S., Ware, D., & Stein, L. (2002). Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research*, *9*, 199–207. https://doi.org/10.1093/dnares/9.6.199

Mignouna, D. B., Akinola, A. A., Abdoulaye, T., Alene, A. D., Manyong, V., Maroya, N. G., Aighewi, B. A., Kumar, L. P., Balogun, M., Lopez-Montes, A., Rees, D., & Asiedu, R. (2020). Potential returns to yam research investment in sub-Saharan Africa and beyond. *Outlook on Agriculture*, *49*, 215–224. https://doi.org/10.1177/0030727020918388

Mignouna, H. D., Abang, M. M., & Asiedu, R. (2008). Genomics of yams, a common source of food and medicine in the tropics. *Genomics of tropical crop plants* (pp. 549–570). Springer New York.

Mignouna, H. D., Abang, M. M., & Fagbemi, S. A. (2003). A comparative assessment of molecular marker assays (AFLP, RAPD and SSR) for white yam (*Dioscorea rotundata*) germplasm characterization. *Annals of Applied Biology*, *142*, 269–276. https://doi.org/10.1111/j.1744-7348.2003.tb00250.x

Mignouna, H. D., Abang, M. M., Onasanya, A., Agindotan, B., & Asiedu, R. (2002). Identification and potential use of RAPD markers linked to *Yam mosaic virus* resistance in white yam (*Dioscorea rotundata*). *Annals of Applied Biology*, *140*, 163–169. https://doi.org/10.1111/j.1744-7348.2002.tb00169.x

Mignouna, H. D., Mank, R. A., Ellis, T. H. N., Van den Bosch, N., Asiedu, R., Ng, S. Y. C., & Peleman, J. (2002). A genetic linkage map of Guinea yam (*Dioscorea rotundata* Poir.) based on AFLP markers. *Theoretical and Applied Genetics*, *105*, 716–725. https://doi.org/10.1007/s00122-002-0911-7

Ming, Y., Yu, X., Liu, W., Wang, J., & Liu, W. (2022). The landscape of genome-wide and gender-specific microsatellites in Indo-Pacific humpback dolphin and potential applications in cetacean resource investigation. *Journal of Marine Science and Engineering*, *10*, 834. https://doi.org/10.3390/jmse10060834

Mizuki, I., Tani, N., Ishida, K., & Tsumura, Y. (2005). Development and characterization of microsatellite markers in a clonal plant, *Dioscorea japonica* Thunb. *Molecular Ecology Notes*, *5*, 721–723. https://doi.org/10.1111/j.1471-8286.2005.01020.x

Mokhtar, M. M., & Atia, M. A. M. (2019). SSRome: An integrated database and pipelines for exploring microsatellites in all organisms. *Nucleic Acids Research*, *47*, D244–D252. https://doi.org/10.1093/nar/gky998

Mokhtar, M. M., Hussein, E. H. A., El-Assal, S. E.-D. S., & Atia, M. A. M (2020). Vf ODB: A comprehensive database of ESTs, EST-SSRs, mtSSRs, microRNA-target markers and genetic maps in *Vicia faba*. *AoB PLANTS*, *12*, 1–10. https://doi.org/10.1093/aobpla/plaa064

Norman, P., Asfaw, A., Tongoona, P., Danquah, A., Danquah, E., Koeyer, D., & Asiedu, R. (2018). Can parentage analysis facilitate breeding activities in root and tuber crops? *Agriculture*, *8*, 95. https://doi.org/10.3390/agriculture8070095

Nowicki, M., Hadziabdic, D., Trigiano, R. N., Boggess, S. L., Kanetis, L., Wadl, P. A., Ojiambo, P. S., Cubeta, M. A., Spring, O., Thines, M., Runge, F., & Scheffler, B. E. (2021). "Jumping Jack": Genomic microsatellites underscore the distinctiveness of closely related *Pseudoperonospora cubensis* and *Pseudoperonospora humuli* and provide new insights into their evolutionary past. *Frontiers in Microbiology*, *12*, 1–17. https://doi.org/10.3389/fmicb.2021.686759

Nwogha, J. S., Obidiegwu, J. E., Okereke, R. N., Bhattacharjee, R., & Oselebe, H. O. (2022). Preliminary verification of the adoption status of some yam (*Dioscorea rotundata* and *Dioscorea alata*) varieties in Nigeria using microsatellites markers. *African Journal of Biotechnology*, *21*, 198–207. https://doi.org/10.5897/AJB2021.17425

Obidiegwu, J. E., Asiedu, R., Ene-Obong, E. E., Muoneke, C. O., & Kolesnikova-Allen, M. (2009). Genetic characterization of some water yam (*Dioscorea alata* L.) accessions in West Africa with simple sequence repeats. *Journal of Food, Agriculture and Environment*, *7*, 634–638.

Obidiegwu, J. E., Kolesnikova-Allen, M., Ene-obong, E. E., Muoneke, C. O., & Asiedu, R. (2009). SSR markers reveal diversity in Guinea yam (*Dioscorea cayenensis*/*D. rotundata*) core set. *African Journal of Biotechnology*, *8*, 2730–2739.

Otoo, E., Anokye, M., Asare, P. A., & Tetteh, J. P. (2015). Molecular categorization of some water yam (*Dioscorea alata* L.) germplasm in ghana using microsatellite (SSR) markers. *Journal of Agricultural Science*, *7*, 225–238. https://doi.org/10.5539/jas.v7n10p225

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rodríguez-Peña, R. A., Johnson, R. L., Johnson, L. A., Anderson, C. D., Ricks, N. J., Farley, K. M., Robbins, M. D., Wolfe, A. D., & Stevens, M. R. (2018). Investigating the genetic diversity and differentiation patterns in the *Penstemon scariosus* species complex under different sample sizes using AFLPs and SSRs. *Conservation Genetics*, *19*, 1335–1348. https://doi.org/10.1007/s10592-018-1103-6

Sapkota, S., Boggess, S. L., Trigiano, R. N., Klingeman, W. E., Hadziabdic, D., Coyle, D. R., & Nowicki, M. (2022). Microsatellite loci reveal high genetic diversity, mutation, and migration rates as invasion drivers of callery pear (*Pyrus calleryana*) in the southeastern United States. *Frontiers in Genetics*, *13*, 1–15. https://doi.org/10.3389/fgene.2022.861398

Scarcelli, N., Cubry, P., Akakpo, R., Thuillet, A. C., Obidiegwu, J., Baco, M. N., Otoo, E., Sonké, B., Dansi, A., Djedatin, G., Mariac, C., Couderc, M., Causse, S., Alix, K., Chaïr, H., François, O., & Vigouroux, Y. (2019). Yam genomics supports West Africa as a major cradle of crop domestication. *Science Advances*, *5*, 1–8. https://doi.org/10.1126/sciadv.aaw1947

Scarcelli, N., Daïnou, O., Agbangla, C., Tostain, S., & Pham, J. L. (2005). Segregation patterns of isozyme loci and microsatellite markers show the diploidy of African yam *Dioscorea rotundata* (2n = 40). *Theoretical and Applied Genetics*, *111*, 226–232. https://doi.org/10.1007/s00122-005-2003-y

Schuler, G. D. (1997). Sequence mapping by electronic PCR. *Genome Research*, *7*, 541–550. https://doi.org/10.1101/gr.7.5.541

Siadjeu, C., Mayland-Quellhorst, E., Pande, S., Laubinger, S., & Albach, D. C. (2021). Transcriptome sequence reveals candidate genes involving in the post-harvest hardening of trifoliate yam *Dioscorea dumetorum*. *Plants*, *10*, 787. https://doi.org/10.3390/plants10040787

Siadjeu, C., Pucker, B., Viehöver, P., Albach, D. C., & Weisshaar, B. (2020). High contiguity de novo genome sequence assembly of trifoliate yam (*Dioscorea dumetorum*) using long read sequencing. *Genes*, *11*, 274. https://doi.org/10.3390/genes11030274

Silva, L. R. G., Bajay, M. M., Monteiro, M., Mezette, T. F., Nascimento, W. F., Zucchi, M. I., Pinheiro, J. B., & Veasey, E. A. (2014). Isolation and characterization of microsatellites for the yam *Dioscorea cayenensis* (Dioscoreaceae) and cross-amplification in *D. rotundata*. *Genetics and Molecular Research*, *13*, 2766–2771. https://doi.org/10.4238/2014.April.14.5

Siqueira, M. V. B. M., Marconi, T. G., Bonatelli, M. L., Zucchi, M. I., & Veasey, E. A. (2011). New microsatellite loci for water yam (*Dioscorea alata*, Dioscoreaceae) and cross-amplification for other *Dioscorea* species. *American Journal of Botany*, *98*, e144–e146. https://doi.org/10.3732/ajb.1000513

Skinner, M. E., Uzilov, A. V., Stein, L. D., Mungall, C. J., & Holmes, I. H. (2009). JBrowse: A next-generation genome browser. *Genome Research*, *19*, 1630–1638. https://doi.org/10.1101/gr.094607.109

Song, X., Yang, Q., Bai, Y., Gong, K., Wu, T., Yu, T., Pei, Q., Duan, W., Huang, Z., Wang, Z., Liu, Z., Kang, X., Zhao, W., & Ma, X. (2021). Comprehensive analysis of SSRs and database construction using all complete gene-coding sequences in major horticultural and representative plants. *Horticulture Research*, *8*, 122. https://doi.org/10.1038/s41438-021-00562-7

Stàgel, A., Portis, E., Toppino, L., Rotino, G. L., & Lanteri, S. (2008). Gene-based microsatellite development for mapping and phylogeny studies in eggplant. *BMC Genomics*, *9*, 357. https://doi.org/10.1186/1471-2164-9-357

Sugihara, Y., Darkwa, K., Yaegashi, H., Natsume, S., Shimizu, M., Abe, A., Hirabuchi, A., Ito, K., Oikawa, K., Tamiru-Oli, M., Ohta, A., Matsumoto, R., Agre, P., De Koeyer, D., Pachakkil, B., Yamanaka, S., Muranaka, S., Takagi, H., White, B., … Terauchi, R. (2020). Genome analyses reveal the hybrid origin of the staple crop white Guinea yam (*Dioscorea rotundata*). *Proceedings of the National Academy of Sciences*, *117*, 31987–31992. https://doi.org/10.1073/pnas.2015830117

Sugihara, Y., Kudoh, A., Oli, M. T., Takagi, H., Natsume, S., Shimizu, M., Abe, A., Asiedu, R., Asfaw, A., Adebola, P., & Terauchi, R. (2021). Population genomics of yams: Evolution and domestication of Dioscorea species. *Population genomics* (pp. 1–28). Springer.

Tamiru, M., Natsume, S., Takagi, H., White, B., Yaegashi, H., Shimizu, M., Yoshida, K., Uemura, A., Oikawa, K., Abe, A., Urasaki, N., Matsumura, H., Babil, P., Yamanaka, S., Matsumoto, R., Muranaka, S., Girma, G., Lopez-Montes, A., Gedil, M., … Terauchi, R. (2017). Genome sequencing of the staple food crop white Guinea yam enables

the development of a molecular marker for sex determination. *BMC Biology*, 15, 1–20. https://doi.org/10.1186/s12915-017-0419-x

Tamiru, M., Yamanaka, S., Mitsuoka, C., Babil, P., Takagi, H., Lopez-Montes, A., Sartie, A., Asiedu, R., & Terauchi, R. (2015). Development of genomic simple sequence repeat markers for yam. *Crop Science*, 55, 2191–2200. https://doi.org/10.2135/cropsci2014.10.0725

Terauchi, R., Chikaleke, V. A., Thottappilly, G., & Hahn, S. K. (1992). Origin and phylogeny of Guinea yams as revealed by RFLP analysis of chloroplast DNA and nuclear ribosomal DNA. *Theoretical and Applied Genetics*, 83, 743–751. https://doi.org/10.1007/BF00226693

Terauchi, R., & Konuma, A. (1994). Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. *Genome*, 37, 794–801. https://doi.org/10.1139/g94-113

Tostain, S., Agbangla, C., Scarcelli, N., Mariac, C., Daïnou, O., Berthaud, J., & Pham, J.-L. (2007). Genetic diversity analysis of yam cultivars (*Dioscorea rotundata* Poir.) in Benin using simple sequence repeat (SSR) markers. *Plant Genetic Resources: Characterization and Utilization*, 5, 71–81. https://doi.org/10.1017/S1479262107672323

Tostain, S., Scarcelli, N., Brottier, P., Marchand, J.-L., Pham, J.-L., & Noyer, J.-L. (2006). Development of DNA microsatellite markers in tropical yam (*Dioscorea* sp.). *Molecular Ecology Notes*, 6, 173–175. https://doi.org/10.1111/j.1471-8286.2005.01182.x

Uncu, A. Ö., Gultekin, V., Allmer, J., Frary, A., & Doganlar, S. (2015). Genomic simple sequence repeat markers reveal patterns of genetic relatedness and diversity in sesame. *The Plant Genome*, 8, plantgenome2014.11.0087. https://doi.org/10.3835/plantgenome2014.11.0087

United Nations Statistics Division. (2022). *United Nations Comtrade Database*. https://comtrade.un.org/data/

Viruel, J., Forest, F., Paun, O., Chase, M. W., Devey, D., Couto, R. S., Segarra-Moragues, J. G., Catalán, P., & Wilkin, P. (2018). A nuclear Xdh phylogenetic analysis of yams (Dioscorea: Dioscoreaceae) congruent with plastid trees reveals a new Neotropical lineage. *Botanical Journal of the Linnean Society*, 187, 232–246. https://doi.org/10.1093/botlinnean/boy013

Wang, H., Wang, Y., Xiong, L., Chen, Y., Sun, J., Ouyang, C., Li, B., Zeng, H., Chen, X., & Luan, M. (2022). Development of simple sequence REPEAT markers for genetic diversity analysis based on the cDNA sequences of Chinese yam (*Dioscorea* spp.). *Horticulturae*, 8, 1163. https://doi.org/10.3390/horticulturae8121163

Wang, J., Sun, Z., Jiang, L., & Hu, Y. (2022). Developing microsatellite duplex PCR reactions for sterlet (*Acipenser ruthenus*) and their application in parentage identification. *Scientific Reports*, 12, 12036. https://doi.org/10.1038/s41598-022-16194-3

Wang, X., & Wang, L. (2016). GMATA: An integrated software package for genome-scale SSR mining, marker development and viewing. *Frontiers in Plant Science*, 7, 1–11. https://doi.org/10.3389/fpls.2016.01350

Wang, X., Yang, S., Chen, Y., Zhang, S., Zhao, Q., Li, M., Gao, Y., Yang, L., & Bennetzen, J. L. (2018). Comparative genome-wide characterization leading to simple sequence repeat marker development for *Nicotiana*. *BMC Genomics*, 19, 500. https://doi.org/10.1186/s12864-018-4878-4

Xu, P., Lu, C., Sun, Z., Kuang, Y., Cao, D., Huo, T., Li, C., Jin, H., & Zheng, X. (2022). In silico screening and development of microsatellite markers for genetic analysis in *Perca fluviatilis*. *Animals*, 12, 1809. https://doi.org/10.3390/ani12141809

Yonemaru, J.-I., Ando, T., Mizubayashi, T., Kasuga, S., Matsumoto, T., & Yano, M. (2009). Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (*Sorghum bicolor* (L.) Moench). *DNA Research*, 16, 187–193. https://doi.org/10.1093/dnares/dsp005

Zannou, A., Agbicodo, E., Zoundjihékpon, J., Struik, P. C., Ahanchédé, A., Kossou, D. K., & Sanni, A. (2009). Genetic variability in yam cultivars from the guinea-sudan zone of Benin assessed by random amplified polymorphic DNA. *African Journal of Biotechnology*, 8, 026–036.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.