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[5,6], Hâna Chair[5,6], Komivi Dossa[4,5,6,*]

[1]Département de Mathématiques et Informatique, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, BP 5005 Dakar-Fann, 10700, Dakar, Senegal.

[2]Laboratoire Campus de Biotechnologies Végétales, Département de Biologie Végétale, Faculté des Sciences et Techniques, Université Cheikh Anta Diop, BP 5005 Dakar-Fann, 10700, Dakar, Senegal.

[3]Department of Crop Science and Biotechnology, Jeonbuk National University, Jeonju 54896, Republic of Korea.

[4]CIRAD, UMR AGAP Institut, 97170 Petit Bourg, Guadeloupe, France.

[5]UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, F-34398 Montpellier, France.

[6]CIRAD, UMR AGAP Institut, F-34398 Montpellier, France.

*Correspondence: Komivi Dossa (komivi.dossa@cirad.fr)

[+]These authors contributed equally to this work.

Abstract

Background

Microsatellite markers represent a low-cost and efficient tool for rapid genotyping as compared to single nucleotide polymorphism markers in laboratories with limited resources. For the economically important yam species widely cultivated in developing countries, very few microsatellite markers are available and no marker database has been developed to date. Herein, we conducted a genome-wide microsatellite marker development among four yam species, identified cross-species transferable markers, and designed an easy-to-use web portal for the yam breeder community.

Results

The screening of yam genomes resulted in 318,713; 322,501; 307,040 and 253,856 microsatellites in *Dioscorea alata*, *D. rotundata*, *D. dumetorum*, and *D. zingiberensis*, 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, a subset of 17 markers were experimentally validated with good discriminatory power regarding the species and the ploidy levels. Ultimately, we created and deployed a dynamic Yam Microsatellite Markers Database (Y2MD) available at http://yamdb.42web.io/. 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.

Conclusions

The present work is the first comprehensive microsatellite marker mining across several yam species and will contribute to advance yam genetic research and marker-assisted breeding. The released user-friendly database constitutes a valuable platform for yam breeders, especially those in developing countries.

Keywords: genotyping; *Dioscoreacea*; database; microsatellite; transferable markers

Introduction

Microsatellites, also termed simple sequence repeats (SSRs) are tandemly repeated DNA sequences generally spanning 1 to 6 bp (Ellegren, 2004). SSRs are ubiquitous in the genome of higher organisms and exhibit genetic polymorphism at inter- and intra-species levels (Bruford and Wayne 1993; Ellegren, 2004; Amiteye, 2021). 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 (Li et al., 2014; Jha et al., 2021; Kim et al., 2021), forensic investigation (Li et al., 2021; Wang et al., 2022), evolutionary studies (Nowicki et al., 2021; Song et al., 2021; de Freitas et al., 2022), and conservation genetics (Rodríguez-Peña et al., 2018; Dong et al., 2022; Ming et al., 2022).

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 more affordable and implementable in a standard laboratory (Herrera and Ghislain, 2013), especially those from low-income countries. Besides, SSRs can exhibit good discriminatory power, are randomly present in the genome, 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, breeders have preferentially relied on SSRs because they are easier to use than single nucleotide polymorphism (SNP) markers, which demand expensive charges and facilities (Wang and Wang, 2016).

In plant breeding, 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*. (Wang et al., 2018), legumes species (n = 13) (Duan and Kaundal, 2021), horticultural species (n = 112) (Song et al., 2021), and *Anemone coronaria* (Martina et al., 2022). Since the usage is based on a graphical user interface, the database offers the breeders, a user-friendly path for 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 the breeders to design markers related to a trait of interest.

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 (https://www.fao.org/faostat/en/FAOSTAT). The genus contains approximately 650 species (Virouel et al., 2018). As a cash crop, yams contribute to the income of more than 60 million people (Asiedu et al., 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 (cpDNA) and nuclear ribosomal (rDNA) (Terauchi et al., 1992). Later on, Amplified Fragment Length Polymorphism (AFLP) and Random Amplified Fragment DNA (RAPD) were employed for genetic diversity assessment (Malapa et al., 2003, Egesi et al., 2007, Zannou et al., 2009), yam mosaic virus resistance markers identification (Mignouna et al., 2002a), and genetic linkage map construction (Mignouna et al., 2002b).

Furthermore, due to their high level of polymorphism and cost-effectiveness, SSRs were developed from cultivated and wild yam species (Terauchi and Konuma, 1994, Mignouma et al., 2003, Mizuki et al., 2005, Hochu et al., 2006, Tostain et al., 2006, Siqueira et al., 2011, Silva et al., 2014, Tamiru et al., 2015). However, certain yam species including *D. batatas* and *D. opposita* have not yet available SSR markers. More importantly, less than 100 validated SSR markers are available in yams to date and no information on their genome coverage is available. Knowing that yam breeding programs are mostly based in countries of the global South where SSRs are widely used (Tostain et al., 2007, Otto et al., 2015, Zawedee et al., 2014, Loko et al., 2016, Korsa et al., 2022, Nwogha et al., 2022), generating SSR markers will surely empower the breeders in their routine tasks such as paternity testing, hybrid identification, QTL detection, marker-assisted selection, genome-wide association study, etc.

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

To date, the genomes of four species *v.i.z D. alata* (Bredeson et al., 2022), *D. rotundata* (Tamiru et al., 2017), *D. dumetorum* (Siadjeu et al., 2020), and *D. zingiberensis* (Cheng et al., 2021) are publicly available. The genomes presented a high level of collinearity (Bredeson et al., 2022), indicating that SSRs might be highly transferable from one species to another, even likely, to those wild species that have not yet been sequenced. Therefore, having a comprehensive SSR database for the yam breeders is fundamental for pre-breeding and breeding activities.

The present study was designed to (i) mine the genome of four yam species for the identification 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 digital web resource to support yam breeding and genetic research.

Materials and Methods

Genomic sequence resources

A total of four genome sequences were retrieved from the National Center for Biotechnology Information (NCBI) genome portal, including *Dioscorea 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) (**Supplementary Table 1**). Besides, we added a newly assembled chromosome-scale assembly (unpublished data) of another greater yam (*D. alata*) “Kabusa”.

Identification of microsatellites and primer design

SSRs were searched using the Genome-wide Microsatellites Analyzing Tool Package (GMATA) (Wang and 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 FASTA format. The microsatellite search parameters applied for SSR detection are: 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) contains the repeats centred 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. rotuntada* for which genome annotations are available. Subsequently, functional enrichment of genes containing SSRs was performed with KOBAS-i webserver (Bu et al., 2021) in order to find out their biological significance.

In silico Polymerase Chain Reaction

The e-Mapping module was used to run the e-PCR algorithm (Schuler, 1997), to validate markers, generate amplicons and assign marker positions across the *Dioscorea* genomes. 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).

SSR primer validation

Plant materials, DNA isolation and PCR experiment

Leaf samples from 10 yam species (*D. alata*, *D. rotundata*, *D. dumetorum*, *D. bulbifera*, *D. trifida*, *D. transversa*, *D. esculenta*, *D. cayenensis*, *D. abyssinica* and *D. nummularia*) including cultivated and wild species were used for DNA extraction. Six *D. alata* genotypes with different ploidy levels (14M, 74F, Pyramide, Kabusa (2x), CT148 and CT198 (4x)) and two D*. rotundata* (Dr1 and Dr2) were included (**Supplementary Table 2**). Plant samples are available in the germplasm collection of CIRAD, Guadeloupe. DNA was extracted based on the Mixed Alkyl Trimethyl Ammonium Bromide method (Cummings and 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 from 18 chromosomes (**Supplementary Table 3**) which exhibited interspecific amplifications through the e-PCR were selected for the PCR experiment. M13 tail (CACGACGTTGTAAAACGAC) was added to the primers. The PCR reaction was performed using the QIAGEN kit under the conditions as followed 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. 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™).

SSR genotyping data analysis

The genotyping data served to perform an hierarchical clustering to depict the genetic relationships between the 16 accessions (**Supplementary Table 2**). The clustering was conducted in R v4.2.2 (R Core Team, 2022) following the Unweighted Pair Group Method with 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 and Mundt, 2020).

Construction and deployment of Yam Microsatellite Markers Database

The Yam Microsatellite Markers Database (Y2MD) database 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 PHP (LAMP) web application platform. The user-friendly web interface was then developed with the aid of JavaScript, and the HyperText Preprocessor (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/, accessed on 15 March 2022), and InsilicoPCR (http://www.biophp.org/minitools/pcr_amplification/, accessed on 15 March 2022) were embedded. The website is freely accessible at http://yamdb.42web.io/.

Results

Detection, distribution and characteristics of microsatellites detected on yam genomes

The genome sequences of four yam species were used for the identification of simple sequence repeats (SSRs). In total, 318,713; 322,501; 307,040 and 253,856 SSRs were detected in *Dioscorea alata*, *D. rotundata*, *D. dumetorum* and *D. zingiberensis* genomes, respectively, with respective densities of 662.86 SSR/Mbp, 552.22 SSR/Mbp, 633.07 SSR/Mbp and 529.97 SSR/Mbp (**Figure 2; Supplementary Figure 1**). At the intraspecific level, the genome of the *D. alata* cultivar "Kabusa" contains 346,986 SSRs with a density of 693.972 SSR/Mbp, which is more than *D. alata* cultivar TDa95/00328.

The numbers of SSRs per chromosomes are strongly correlated to the chromosome lengths (**Figure 3a**). A total of 658 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 motifs obtained are strongly rich in A/T: A and T mononucleotides are the most represented with an average of 55% of the detected SSRs, 20% of the SSRs are constituted by AT and TA di-nucleotides. For tri-nucleotides, repeats (AAT/ATA/ATT/TAA/TTA) account for 7.5%. Hexa- and penta-nucleotides had relatively low frequencies.

Most of the SSRs (~84%) were localised in the intergenic regions in *D. alata* and *D. rotundata* genomes (**Figure 3d**). This analysis was not performed for other species because of the lack of a proper genome annotation. The genes containing SSRs were analysed for their functional enrichment. They were enriched in a wide panel of biological attributes. Among them, various agronomically relevant gene ontology terms were found 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; Supplementary Table 4**).

Development of genome-wide SSR primers

A total of 864,128 unique primer pairs were created 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). All primers are distributed unevenly across the genomes. The identified primers and their characteristics are presented in **Supplementary Table 5**. All the markers were assigned the identifiers YMa (*D. alata*), YMr (*D. rotundata*), YMd (*D. dumetorum*), YMz (*D. zingiberensis*) followed by the consecutive numbers of the primers in the genomes.

Cross-species transferability of SSRs

First, we performed a pairwise comparison of the microsatellite markers between the four yam species (**Supplementary Table 6)**. *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 evaluate the cross-species transferability of the microsatellite markers, we set *D. alata* as a reference and performed an e-PCR on the other three species. Of the 221,606 SSRs markers from *D. alata* (TDa95/00328), 166,817 SSRs were specific to *D. alata*, while 39,837 SSRs were able to amplify *D. rotundata*, 3473 on *D. dumetorum* and 677 on *D. zingiberensis* (**Figure 4a**). Collectively we identified 1170 core SSR markers able to amplify all the four species and potentially more yam species.

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

Validation of selected SSRs and genotyping analysis

Firstly, to validate the newly developed markers in yam species, we extracted 92 SSR markers from the literature routinely used for genotyping *D. alata* and *D. rotundata*. This validation was done using the e-PCR method. Of the 92 markers, 84 (91%) yielded PCR products on their respective reference genomes.

Next, a set of 18 interspecific markers were selected from 18 chromosomes and used for a wet-lab PCR validation in ten yam species (**Supplementary Table 2; Supplementary Table 3**). Out of the 18 selected markers, 17 produced amplicons (except the marker YMar128). Sixteen markers produced amplicons in at least six species over the ten (**Table 1**). Interestingly, the marker amplifications in diploid and tetraploid *D. alata* cultivars were in the range of maximum two or four alleles, respectively, suggesting the ability of those markers to potentially discriminate cultivars following the ploidy criteria. For instance the marker YMar79 was able to clearly delineate diploid from tetraploid forms of *D. alata* by the expression of the number of allele inferior or equal to the number of ploidy (**Table 1**). Interestingly, four alleles were amplified for *D. transversa* (markers YMar44, YMar79, and YMac82), *D. cayenensis* (markers YMar39, 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 harbouring exclusively *D. alata* cultivars (**Figure 5**). Among *D. alata* cultivars, a clear distinction of 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* formed a distinct group individually 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 pointed out: one consisting of *D. trifida* and *D. esculenta*, and another group formed by *D. dumetorum* and *D. bulbifera*.

Table 1. Allele scoring for 18 selected SSR markers in ten yam species

	D. alata (14M)	*D. alata* (74F)	*D. alata* (CT148)	*D. alata*(CT198)	*D. alata* (Pyramide)	*D. alata* (Kabusa)	*D. bulbifera*	*D. transversa*	*D. rotoundata*	*D. rotoundata*	*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			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	173	154 + 157

YMar44	192 + 206	NA	208 + 212	196 + 208	192 + 204	196 + 206	NA	184 + 192 + 196 + 200	190 + 196	190 + 196	NA	NA	NA	190 + 194 + 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 + 264	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

Development of Yam Microsatellite Markers Database

We built up a comprehensive database by integrating SSR markers and whole genome assemblies and annotations datasets (Figure 1). The Y2MD page contains data of four *Dioscorea* species and over 864,128 primer pairs. Also all SSRs data are downloadable in Excel, PDF, text, CSV format. 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 possibility to choose the SSR of choice according to the type of patterns, the number of repeats, the position, etc. Information on experimentally validated SSRs and cross-species transferable SSRs are available in the "SSRs" tab. On the "Download" page, the links to the different genomes are available.

To enable the database user to perform some fundamental analyses, several functionalities were embedded in Y2MD among which JBrowse (**Figure 6a**), *insilico*PCR (**Figure 6b**), SSR Finder (**Figure 6c**), and Y2MD Blast (**Figure 6d**). Thus, Y2MD empowered the user to find an SSR with SSR Finder, and proceed to a quick amplification check with *insilico*PCR. Besides, the user can also blast his own sequence of interest onto a genome with Y2MD Blast tool. With JBrowse, the user is able to dive into the genome annotation map to find out potential genes related to a targeted SSR. Altogether, Y2MD offers a dynamic platform to the yam breeder to remotely use the available genomic resources in a GUI user-friendly manner. The website is freely accessible at http://yamdb.42web.io/.

Discussion

Developing genome-wide SSR markers in Dioscoreaceae

Microsatellite markers proved to be effective to characterise yam germplasms worldwide (Obidiegwu et al., 2009a; Obidiegwu et al., 2009b; Siqueira et al., 2011; Tostain et al., 2006; Tamiru et al, 2015; Arnau et al., 2017; Cao et al., 2021). The present study enabled the extraction of quite high numbers 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 as compared to the other yam species. Nonetheless, we observed a high

correlation (R^2 =0.99) between the SSR numbers and chromosome lengths (**Figure 3)**. 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 (from 529 to 662 SSR/Mbp) are similar to that of sesame (507 SSR/Mbp) while, lower than other species such as Arabidopsis (*Arabidopsis thaliana*) and rice (*Oryza sativa*) (875 and 807 SSRs/Mbp, respectively) (Lawson and Zhang, 2006; Dossa et al., 2017). 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 (**Figure 2**) and the genetic distances between the two accessions may explain these differences. Among the observed motifs, mononucleotides that account for 50% of SSRs are the most represented (**Figure 3)**. However, they are not ideal targets for PCR markers (Clarke et al., 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 are 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, 2002).

Cross-species transferability of the SSR markers

The identification of polymorphic SSR markers will support research on yam genetics and breeding, especially in developing countries. At the genome scale, we identified 1170 cross-species transferable markers in the studied four Dioscorea species. Previously, Tamiru et al. (2015) showed that cross-amplification of SSR markers from *D. cayenensis*, revealed that 94.4% and 56.7% could be successfully transferred to *D. rotundata* and *D. alata*, respectively. 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, therefore having SSR marker sets applicable on different species is highly useful. We also noticed that the number of cross-species transferable markers decreased significantly when *D. zingiberensis* was included (**Figure 4)**. Indeed, Bredeson et al. (2022) compared *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 at 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* (Dossa et al., 2017; Bhattarai et al., 2021; 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 much 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.

Validating selected SSRs

In *Dioscoreaceae*, very few SSR markers have been experimentally (wet-lab) validated (Obidiegwu et al., 2009a; Obidiegwu et al., 2009b; Andris et al., 2010; Siqueira et al., 2011; Tostain et al., 2006; Tamiru et al., 2015; Arnau et al., 2017; Cao et al., 2021), most of which are from *D. alata* and *D. rotundata*. In other species such as rice (McCouch et al., 2002), sesame (Dossa et al., 2017), and cowpea (*Vigna unguiculata*) (Jasrotia et al., 2019), a large number of markers have been identified, experimentally characterised and available for genetic studies and breeding. In this study, we extracted 92 markers from the literature (wet-lab validated markers) and most of them were found among the markers created in this study. As the e-PCR parameters do allow only two mismatches and one insertion for the primers, this could be partly responsible for the absence of PCR products for the missing wet-lab validated markers. This analysis also provided the exact locations of these SSRs in the yam genomes, which will enable researchers to select markers from different chromosomes in order to have a good discriminatory power.

Out of 18 selected SSR markers, 17 were able to amplify by PCR several yam genomes in this study. The experimentally tested markers were from different chromosomes. They demonstrated a robust discriminatory power by clearly delineating the studied yam species and ploidy levels (**Figure 5**). In

yam species, the basic chromosome number is 20 and most species harbour different cytotypes (Scarcelli et al., 2005; Arnau et al., 2009; Abraham et al., 2013; 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. Besides, the tested SSR markers also exhibited a promising accuracy even at intra-species level with a clear classification of diploid and tetraploid forms of *D. alata* cultivars. Based on the allele numbers, we report that yam species including *D. transversa*, *D. abyssinica*, *D. cayenensis*, and *D. nummularia* have cultivars with tetra- or higher ploidy levels (**Table 1**) (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 check and species differentiation in yams.

Deployment of Yam Microsatellite Markers Database for yam breeders

Many databases dealing with SSR marker data from various species have been made available to the public. However, no database is available to study yam microsatellites. A yam microsatellites marker database will be a useful platform for advancing genetic evaluation, genomic research, and yam breeding. The Pan Species Microsatellite Database (Du et al., 2020) containing SSR data from 18,408 organisms is limited by the lack of features such as ePCR, Jbrowse (Buels et al., 2016) and Blast (Altschul et al., 1990). This is also the case for the Tea Microsatellite Database (Dubey et al., 2020) and SSRome (Mokhtar and Atia, 2019). In this study, we compiled all generated microsatellite data into Yam Microsatellite Markers Database (Y2MD) (**Figure 6**). Y2MD is embedded with various useful tools such as JBrowse, Blast, SSR Finder and *insilico*PCR. These tools will facilitate the localization, SSR detection and polymorphism analysis for existing and new experimental yam SSR markers/nucleotide sequences. Y2MD has been developed to be a very user-friendly database enabling all users, especially those with limited knowledge or resources in bioinformatics to analyse their SSR related data.

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 http://yamdb.42web.io/. To perform the database construction, microsatellites of four *Dioscorea* species were first extracted and primers were designed on the flanking sequences of these microsatellites. In total 1,201,570 microsatellites were extracted from all species with 864,128 designed primer pairs. Then, additional analyses were performed such as cross-species transferability and determination of polymorphic markers which are determining factors in plant genotyping. Experimental validation of selected markers were also performed to provide credence to our datasets. Overall, the SSRs developed in this work will be useful for genetic characterization of yam germplasms, quantitative trait loci (QTL) mapping and marker assisted breeding.

All of this information has been compiled into Y2MD and will be of great use for the yam breeding community, especially in developing countries. With the rapid development in genome sequencing projects, we expect more yam genomes to come. Hence, we will expand Y2MD to include new *Dioscorea* species as soon as their genome sequences are available. In addition, Y2MD will be enhanced by adding new tools and functionalities such as the integration of Primer3, introducing molecular markers associated with QTLs and new wet-lab validated SSRs.

Availability of Data and Materials

All SSR data generated and the genomes used are available online at http://yamdb.42web.io/. 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.

Author Contributions

Conceptualization: Komivi Dossa, Diaga Diouf, Hâna Chair. *SSR validation experiment*: Ronan Rivallan, Erick Malédon, Hâna Chair, Komivi Dossa. *Data acquisition, curation and formal analysis*: Moussa Diouf, Yedomon Ange Bovys Zoclanclounon, Pape Adama Mboup, Diaga Diouf, Komivi Dossa. *Database building*: Moussa Diouf, Pape Adama Mboup, Komivi Dossa. *Funding acquisition*: Komivi Dossa. *Writing ± original draft*: Moussa Diouf, Yedomon Ange Bovys Zoclanclounon, Komivi

Dossa. *Writing ± review & editing*: Diaga Diouf, Hâna Chair. All authors have read and approved the final version of this manuscript.

Acknowledgment

Not applicable.

Funding

This research was financially supported by the AfricaYam Project (Grant OPP1052998-Bill and Melinda Gates Foundation).

Conflicts of Interest

The authors declare no competing financial interests.

Figure captions

Figure 1. A diagram depicting the structure of Yam Microsatellite Markers Database (Y2MD).

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

Figure 3. Correlation between chromosome length and number of 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 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).

Figure 5. Hierarchical clustering tree depicting the relationships between ten yam species using 17 SSR markers. 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.

Figure 6. The yam microsatellite markers database: Y2MD. The website is freely available at http://yamdb.42web.io/. The web interface of Y2MD showcasing JBrowse (a), *insilico*PCR (b), SSR Finder (c), and Blast (d) functionalities.

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., Enkhbileg, D., Fartek, B., Gardner, M.G., Gray, K., Gunasekera, R.M., Hanley, S.J., Havil, N., Hereward, J.P., Hirase, S., Hong, Y., Jarne, P., Jianfei, Q., Johnson, R.N., Kanno, M., Kijima, A., Kim, H.C., Kim, K.S., Kim, W., Larue, E.,

Lee, J.W., Lee, J., Li, C., Liao, M., Lo, N., Lowe, A.J., Malausa, T., Malé, P.G., Marko, M.D., Martin, J., Messing, R., Miller, K.J., Min, B., Myeong, J., Nibouche, S., Noack, A.E., Noh, J.K., Orivel, J., Park, C., Petro, D., Prapayotin-Riveros, K., Quilichini, A., Reynaud, B., Riginos, C., Risterucci, A.M., Rose, H.A., Sampaio, I., Silbermayr, K., Silva, M.B., Tero, N., Thum, R.A., Vinson, C.C., Vorsino, A., Vossbrinck, C.R., Walzer, C., White, J.C., Wieczorek, A., & 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 World 1. Yams. *Food Security*, *2*, 305–315. https://doi.org/10.1007/s12571-010-0085-0

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

Bredeson, J. V., Lyons, J.B., Oniyinde, I.O., Okereke, N.R., Kolade, O., Nnabue, I., Nwadili, C.O., Hřibová, 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., Jamnadass, R., Muchugi, A., Goodstein, D., Egesi, C.N., Featherston, J., Asfaw, A., Simpson, G.G., Doležel, J., Hendre, P.S., Van Deynze, A., Kumar, P.L., Obidiegwu, J.E., Bhattacharjee, 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

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

Denadi, N., Gandonou, C., Missihoun, A.A., Zoundjihékpon, J., & Quinet, M. (2020). Plant sex prediction using genetic markers in cultivated yams (*Dioscorea rotundata* poir.) in benin. *Agronomy*, *10*, 1–13. https://doi.org/10.3390/agronomy10101521

Denadi, N., Yolou, M., Dadonougbo, A.E., Zoundjihékpon, J., Dansi, A., Gandonou, C., & Quinet, M. (2022). Yam (*Dioscorea rotundata* Poir.) Displays Prezygotic and Postzygotic Barriers to Prevent Autogamy in Monoecious Cultivars. *Agronomy*, *12*, 1–17. https://doi.org/10.3390/agronomy12040872

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

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

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

Georgeou, N., Hawksley, C., Wali, N., Lountain, S., Rowe, E., West, C., & Barratt, L. (2022). Food security and small holder farming in Pacific Island countries and territories: A scoping review. *PLOS Sustainability and Transformation*, *1*, e0000009. https://doi.org/10.1371/journal.pstr.0000009

Herrera, M. del 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, Totowa, NJ.

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

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*. 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/https://doi.org/10.21203/rs.3.rs-1905131/v1

Lawson, M.J., Zhang, L., & Tech, V. (2006). Distinct patterns of SSR distribution in the Arabidopsis thaliana and rice genomes *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

Li, J., & Ye, C. (2020). Genome-wide analysis of microsatellite and sex-linked marker identification in *Gleditsia sinensis*. *BMC Plant Biology*, *20*, 338. https://doi.org/10.1186/s12870-020-02551-9

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

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, New York, NY.

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)(a). 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)(b). 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

Mondo, J.M., Agre, P.A., Edemodu, A., Adebola, P., Asiedu, R., Akoroda, M.O., & Asfaw, A. (2020). Floral biology and pollination efficiency in yam (*Dioscorea spp*.). *Agriculture (Switzerland)*, *10*, 1–21. https://doi.org/10.3390/agriculture10110560

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)(a). 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)(b). 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

Padhan, B., & Panda, D. (2020). Potential of Neglected and Underutilized Yams (*Dioscorea spp*.) for Improving Nutritional Security and Health Benefits. *Frontiers in Pharmacology*, *11*, 474–497. https://doi.org/10.3389/fphar.2020.00496

R Core Team. (2022). R: A Language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 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)(a). 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

Song, X., Yang, T., Zhang, X., Yuan, Y., Yan, X., Wei, Y., Zhang, J., & Zhou, C. (2021)(b). Comparison of the Microsatellite Distribution Patterns in the Genomes of Euarchontoglires at the Taxonomic Level. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.622724

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., Asiedu, R., Innan, H., Asfaw, A., Adebola, P., & 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.

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., Bhattacharjee, R., Abberton, M., Kumar, P.L., Rabbi, I., Tsujimura, M., Terachi, T., Haerty, W., Corpas, M., Kamoun, S., Kahl, G., Takagi, H., Asiedu, R., & 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*, *83–83*, 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*. 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, 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

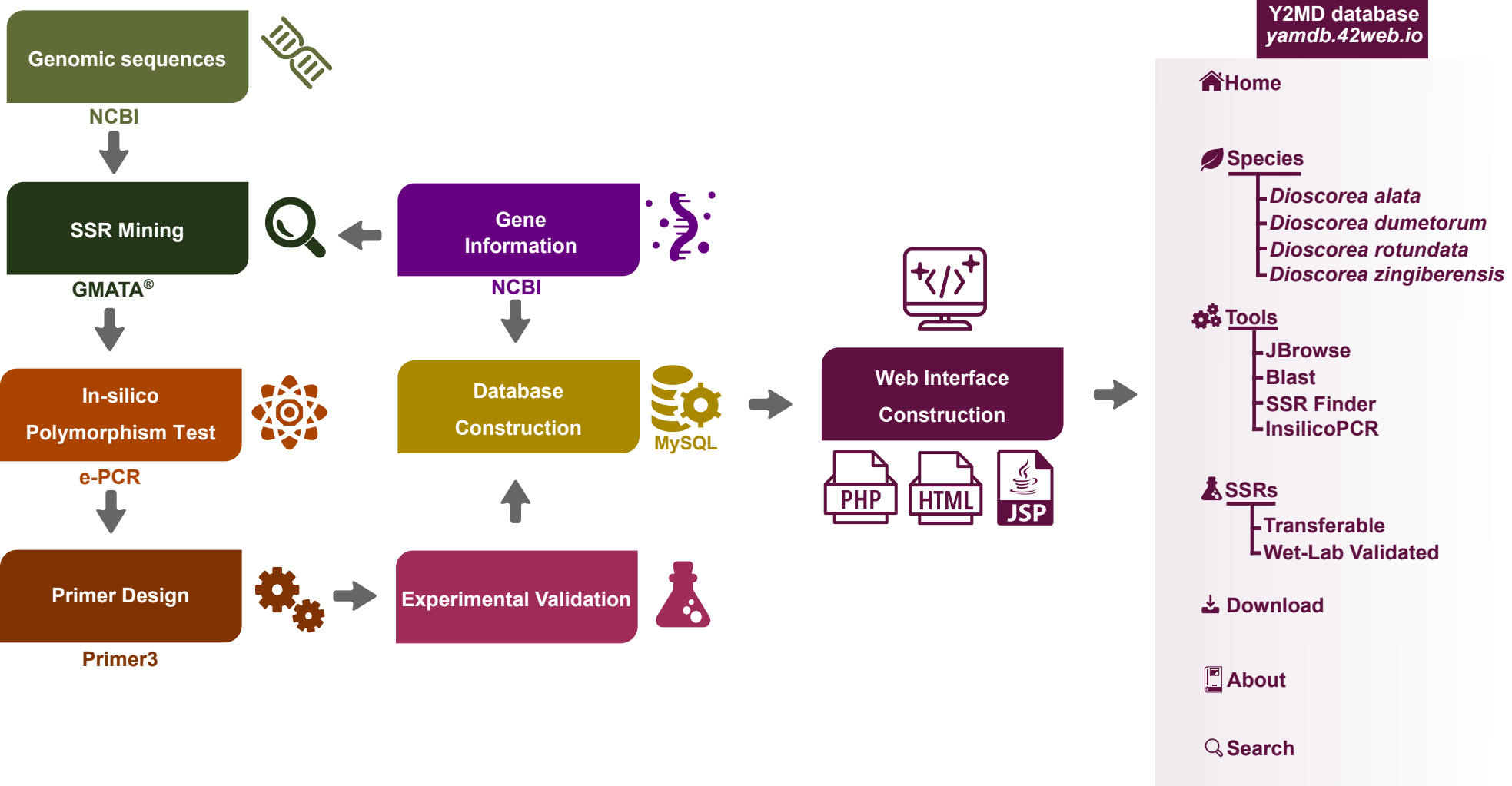
Wang, X., Yang, S., Chen, Y., Zhang, S., Zhao, Q., Li, M., Gao, Y., Yang, L., & Bennetzen, J.L. (2018). Comparative genome-wide characterization leading to simple sequence repeat marker development for *Nicotiana*. *BMC Genomics*, *19*, 500. https://doi.org/10.1186/s12864-018-4878-4

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

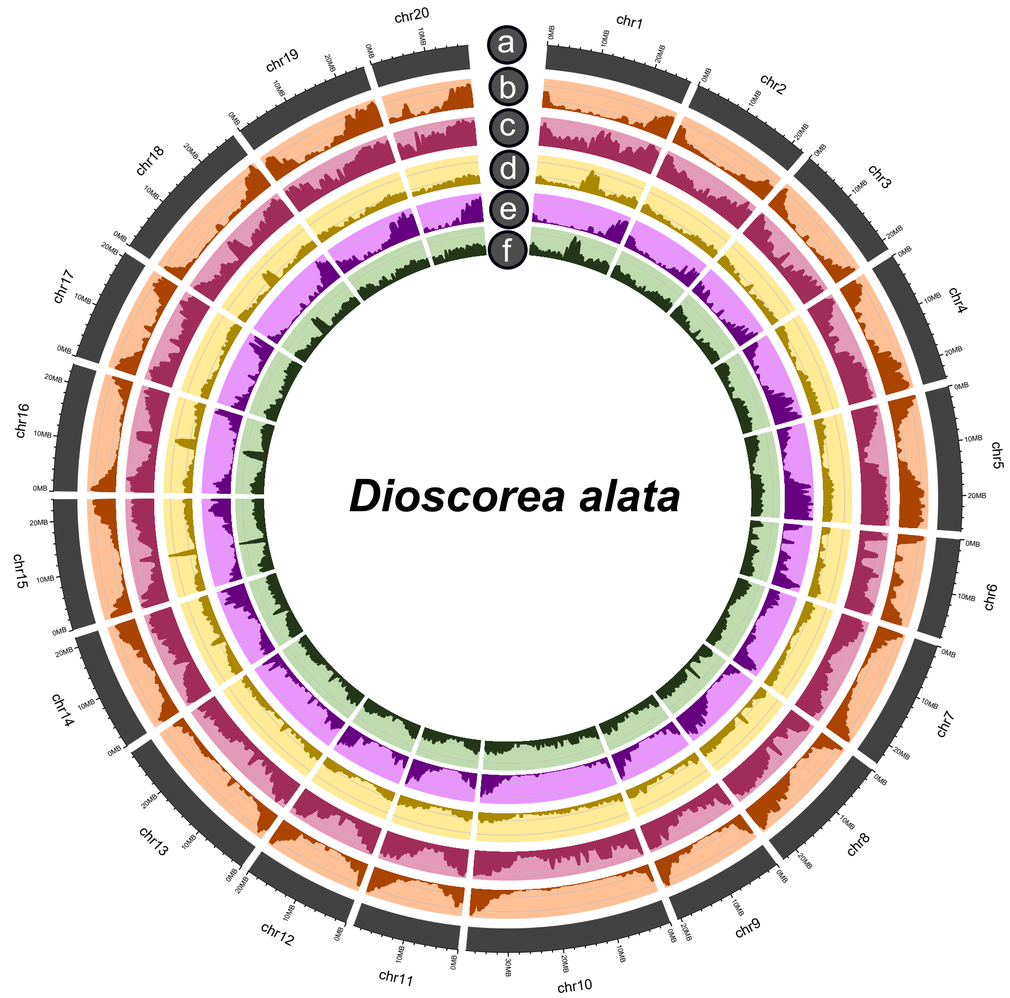
Yonemaru, J. -i., Ando, T., Mizubayashi, T., Kasuga, S., Matsumoto, T., & Yano, M. (2009). Development of Genome-wide Simple Sequence Repeat Markers Using Whole-genome Shotgun Sequences of Sorghum (*Sorghum bicolor* (L.) Moench). *DNA Research*, *16*, 187–193. https://doi.org/10.1093/dnares/dsp005

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

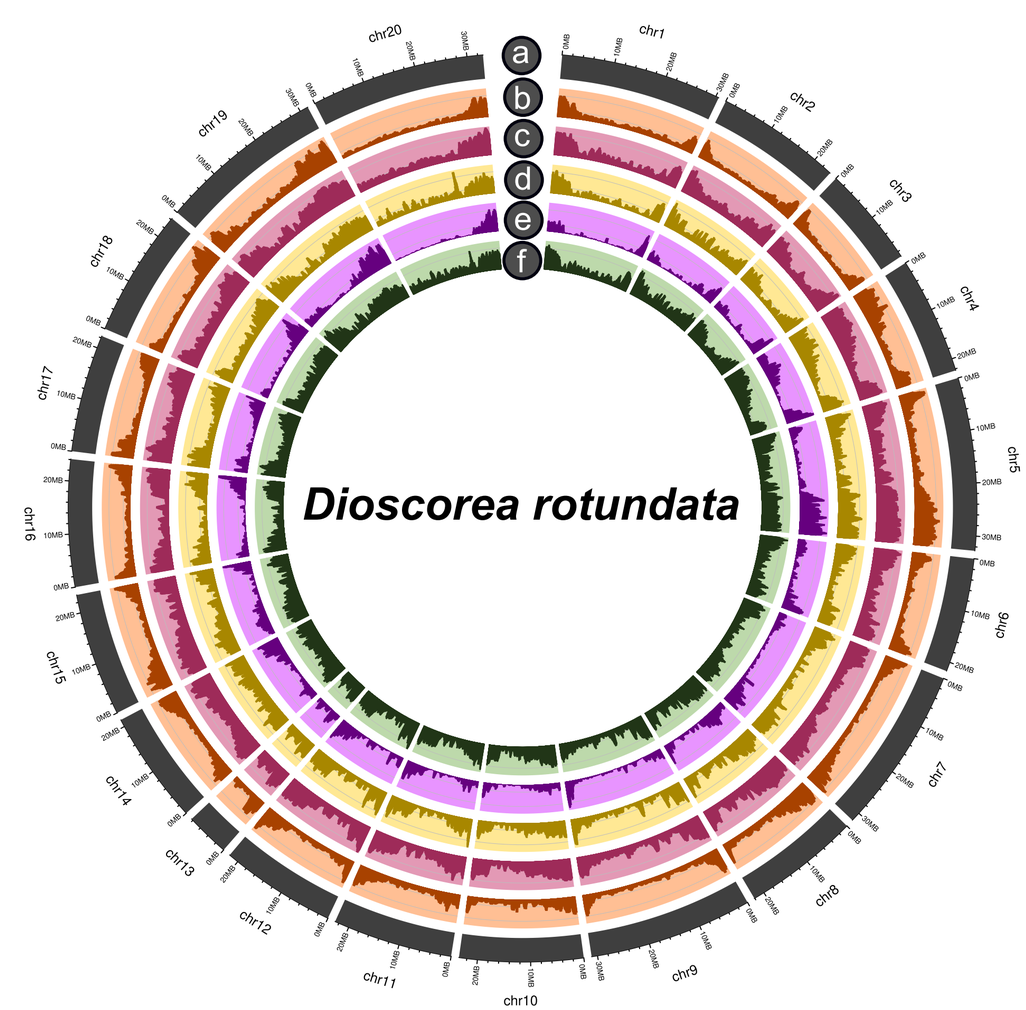
Zawedde, B.M., Ghislain, M., Magembe, E., Amaro, G.B., Grumet, R., & Hancock, J. (2015). Characterization of the genetic diversity of Uganda's sweet potato (*Ipomoea batatas*) germplasm using microsatellites markers. *Genetic Resources and Crop Evolution*, *62*, 501–513. https://doi.org/10.1007/s10722-014-0175-5


Genomic sequences
NCBI
SSR Mining
GMATA®
In-silico Polymorphism Test
e-PCR
Primer Design
Primer3
Gene Information
NCBI
Database Construction
MySQL
Experimental Validation
Web Interface Construction
PHP
HTML
JSP
Y2MD database
yamdb.42web.io
Home
Species
Dioscorea alata
Dioscorea dumetorum
Dioscorea rotundata
Dioscorea zingiberensis
Tools
JBrowse
Blast
SSR Finder
InsilicoPCR
SSRs
Transferable
Wet-Lab Validated
Download
About
Search

a.



b.

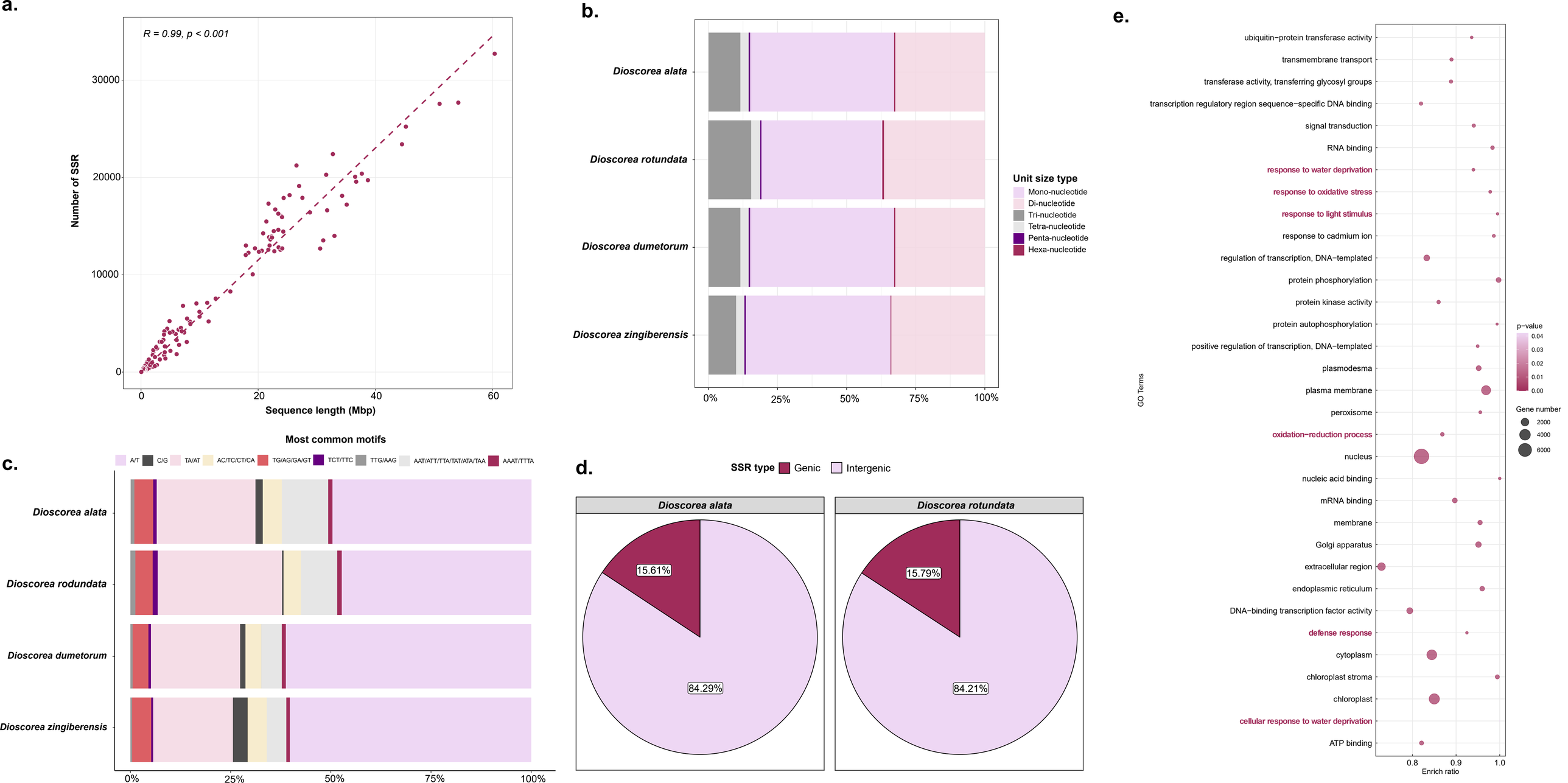


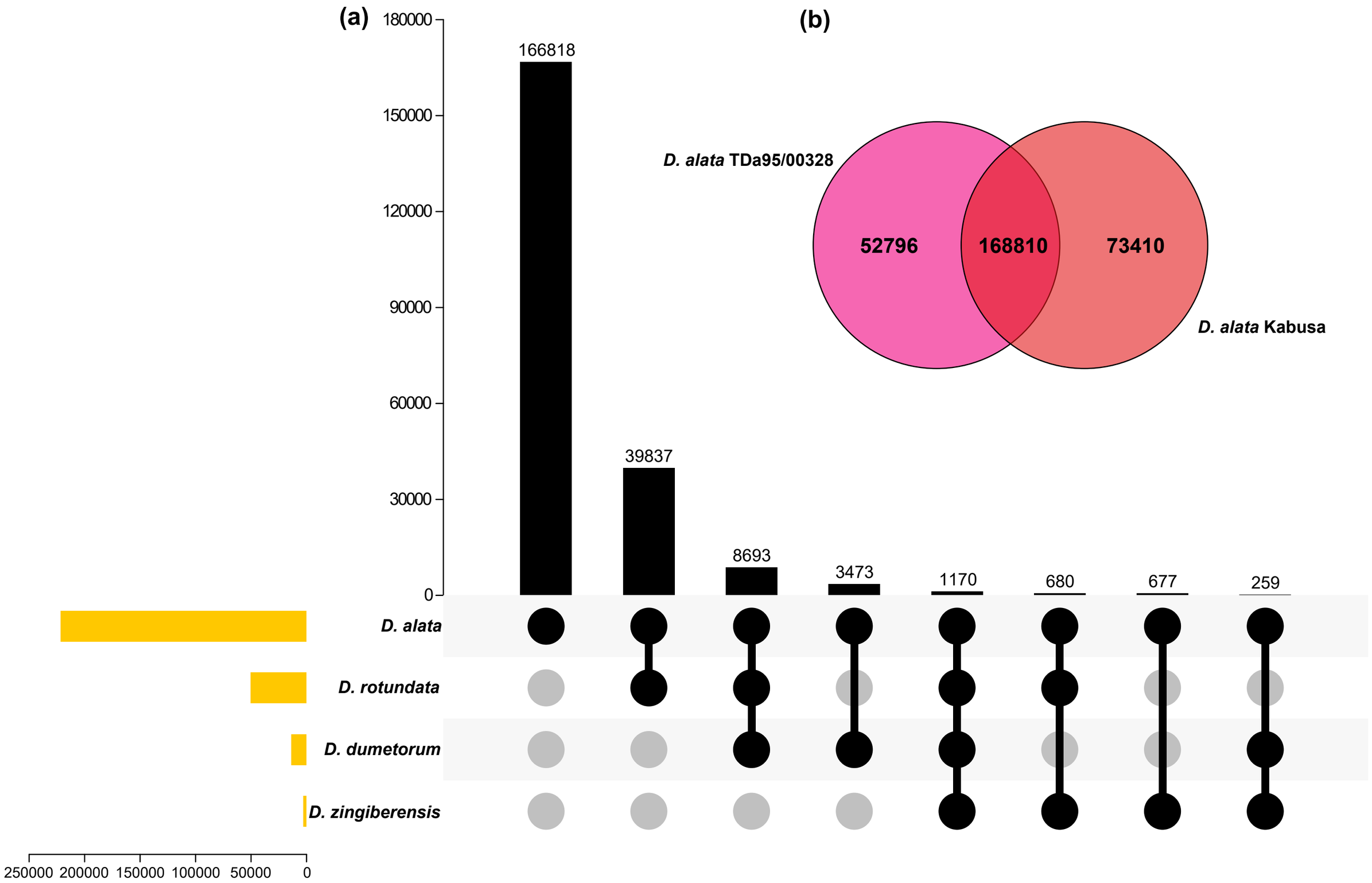
ⓐ Chromosome ⓑ Gene density ⓒ Perfect SSR density ⓓ Compound SSR density ⓔ Genic SSR density ⓕ Intergenic SSR density

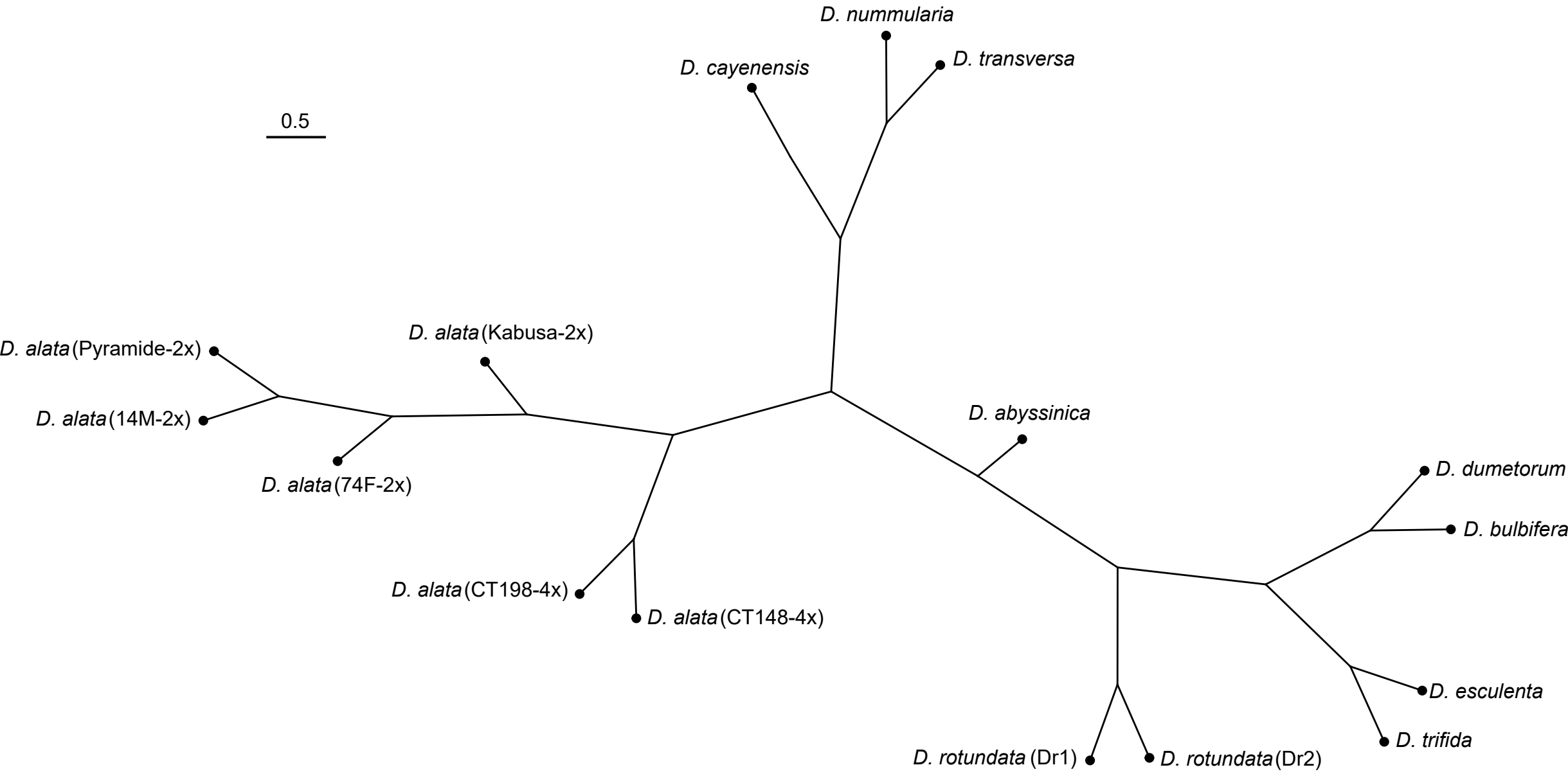
c.

Summary statistics

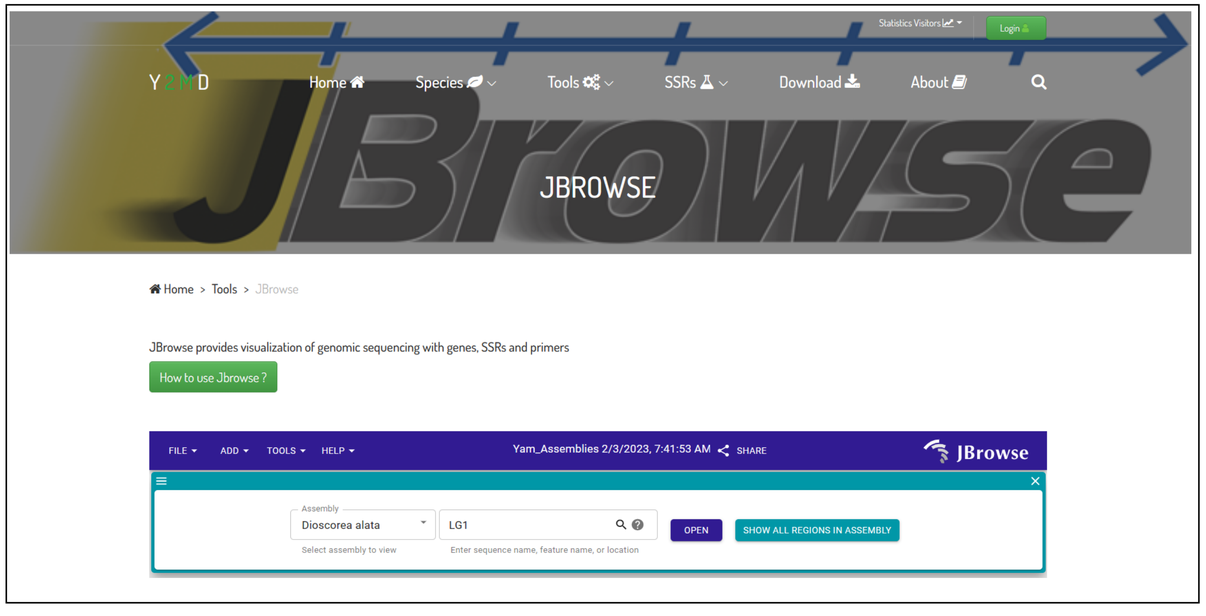
	D. alata (TDa95)	*D. alata* (Kabusa)	*D. rotundata*	*D. dumetorum*	*D. zingiberensis*
Assembly size (Mbp)	475.35	500.39	491.97	485.12	479.42
Total number of sequences analyzed	20	100	20	924	14,408
Total number of SSR	318,858	349,295	276,795	308,654	255,662
Total SSR length size (Mbp)	0.53	0.56	0.50	0.52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


a.
R = 0.99, p < 0.001
Number of SSR
Sequence length (Mbp)
b.
Dioscorea alata
Dioscorea rotundata
Dioscorea dumetorum
Dioscorea zingiberensis
Unit size type
Mono-nucleotide
Di-nucleotide
Tri-nucleotide
Tetra-nucleotide
Penta-nucleotide
Hexa-nucleotide
c.
Most common motifs
A/T
C/G
TA/AT
AC/TC/CT/CA
TG/AG/GA/GT
TCT/TTC
TTG/AAG
AAT/ATT/TTA/TAT/ATA/TAA
AAAT/TTTA
Dioscorea alata
Dioscorea rodundata
Dioscorea dumetorum
Dioscorea zingiberensis
d.
SSR type
Genic
Intergenic
Dioscorea alata
15.61%
84.29%
Dioscorea rotundata
15.79%
84.21%
e.
ubiquitin−protein transferase activity
transmembrane transport
transferase activity, transferring glycosyl groups
transcription regulatory region sequence−specific DNA binding
signal transduction
RNA binding
response to water deprivation
response to oxidative stress
response to light stimulus
response to cadmium ion
regulation of transcription, DNA−templated
protein phosphorylation
protein kinase activity
protein autophosphorylation
positive regulation of transcription, DNA−templated
plasmodesma
plasma membrane
peroxisome
oxidation−reduction process
nucleus
nucleic acid binding
mRNA binding
membrane
Golgi apparatus
extracellular region
endoplasmic reticulum
DNA−binding transcription factor activity
defense response
cytoplasm
chloroplast stroma
chloroplast
cellular response to water deprivation
ATP binding
GO Terms
Enrich ratio
p−value
Gene number

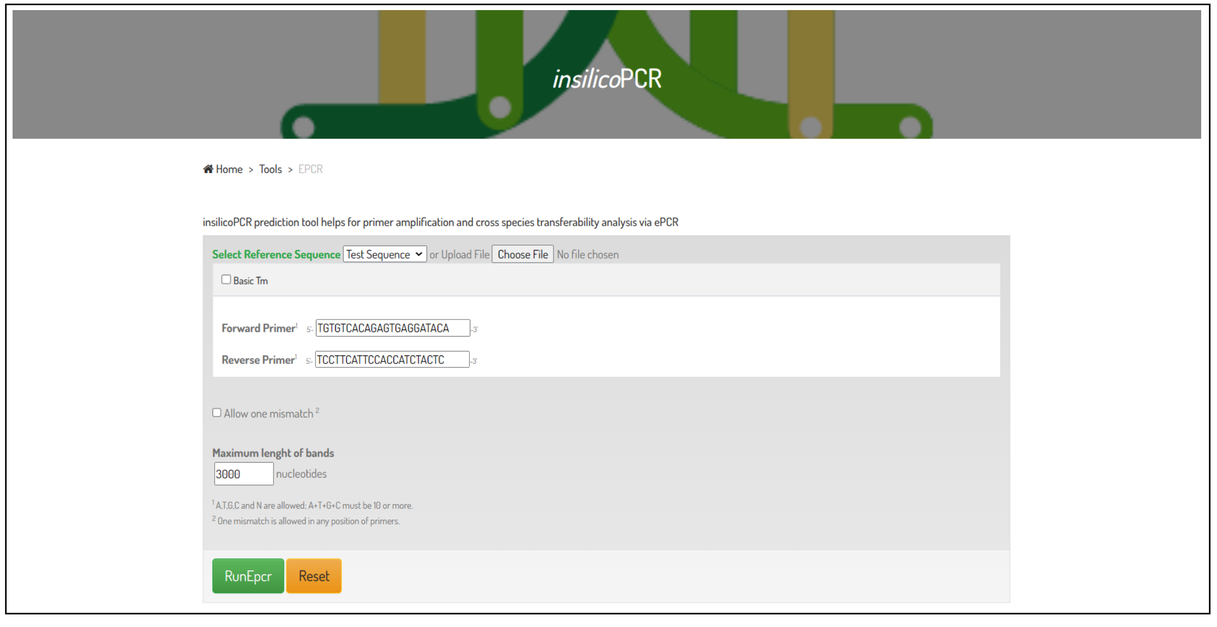

(a)
180000
150000
120000
90000
60000
30000
0
166818
39837
8693
3473
1170
680
677
259
D. alata
D. rotundata
D. dumetorum
D. zingiberensis
250000 200000 150000 100000 50000 0
(b)
D. alata TDa95/00328
52796
168810
73410
D. alata Kabusa


D. nummularia
D. transversa
D. cayenensis
0.5
D. alata (Kabusa-2x)
D. alata (Pyramide-2x)
D. alata (14M-2x)
D. alata (74F-2x)
D. abyssinica
D. dumetorum
D. bulbifera
D. alata (CT198-4x)
D. alata (CT148-4x)
D. esculenta
D. trifida
D. rotundata (Dr1)
D. rotundata (Dr2)

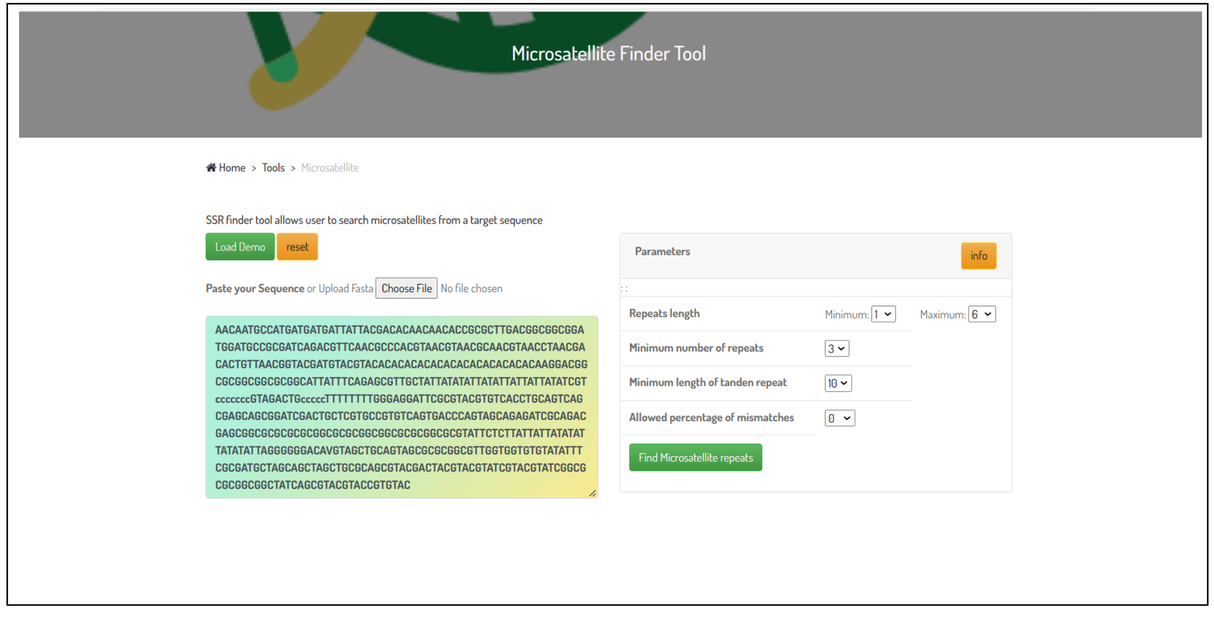
a.


Statistics Visitors
Login
Y2MD
Home
Species
Tools
SSRs
Download
About
JBROWSE
Home > Tools > JBrowse
JBrowse provides visualization of genomic sequencing with genes, SSRs and primers
How to use Jbrowse ?
FILE
ADD
TOOLS
HELP
Yam_Assemblies 2/3/2023, 7:41:53 AM
SHARE
JBrowse
Assembly
Dioscorea alata
LG1
OPEN
SHOW ALL REGIONS IN ASSEMBLY
Select assembly to view
Enter sequence name, feature name, or location

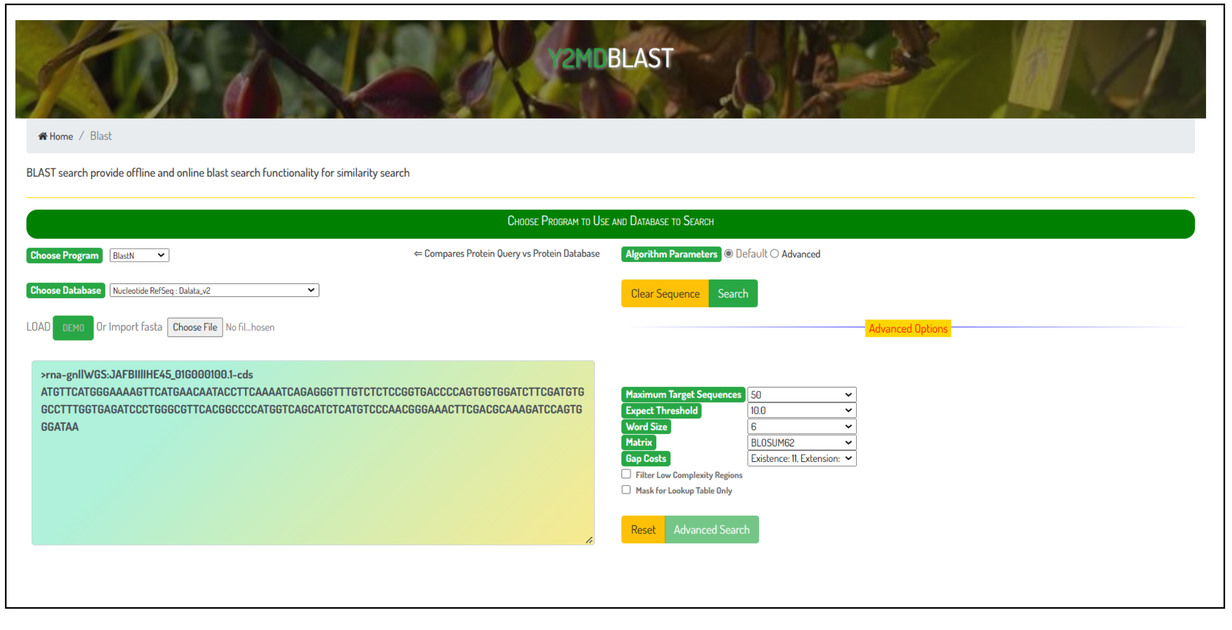
b.


insilicoPCR
Home > Tools > EPCR
insilicoPCR prediction tool helps for primer amplification and cross species transferability analysis via ePCR
Select Reference Sequence
Test Sequence
or Upload File
Choose File
No file chosen
Basic Tm
Forward Primer
TGTGTCACAGAGTGAGGATACA
Reverse Primer
TCCTTCATTCCACCATCTACTC
Allow one mismatch
Maximum lenght of bands
3000
nucleotides
A,T,G,C and N are allowed: A+T+G+C must be 10 or more.
One mismatch is allowed in any position of primers.
RunEpcr
Reset

c.


Microsatellite Finder Tool
Home > Tools > Microsatellite
SSR finder tool allows user to search microsatellites from a target sequence
Load Demo
reset
Paste your Sequence or Upload Fasta
Choose File
No file chosen
AACAATGCCATGATGATGATTATTACGACACAACAACACCGCGCTTGACGGCGGCGGA
TGGATGCCGCGATCAGACGTTCAACGCCCACGTAACGTAACGCAACGTAACCTAACGA
CACTGTTAACGGTACGATGTACGTACACACACACACACACACACACACACACAAGGACGG
CGCGGCGGCGCGGCATTATTTCAGAGCGTTGCTATTATATATTATATTATTATATCGT
cccccccGTAGACTGcccccTTTTTTTTGGGAGGATTCGCGTACGTGTCACCTGCAGTCAG
CGAGCAGCGGATCGACTGCTCGTGCCGTGTCAGTGACCCAGTAGCAGAGATCGCAGAC
GAGCGGCGCGCGCGGCGCGCGGCGGCGCGCGGCGCGTATTCTCTTATTATTATATAT
TATATATTAGGGGGGACAVGTAGCTGCAGTAGCGCGCGGCGTTGGTGGTGTGTATATTT
CGCGATGCTAGCAGCTAGCTGCGCAGCGTACGACTACGTACGTATCGTACGTATCGGCG
CGCGGCGGCTATCAGCGTACGTACCGTGTAC
Parameters
info
Repeats length
Minimum: 1
Maximum: 6
Minimum number of repeats
3
Minimum length of tanden repeat
10
Allowed percentage of mismatches
0
Find Microsatellite repeats

d.


Y2MDBLAST
Home / Blast
BLAST search provide offline and online blast search functionality for similarity search
Choose Program to Use and Database to Search
Choose Program
BlastN
⇐ Compares Protein Query vs Protein Database
Algorithm Parameters
Default
Advanced
Choose Database
Nucleotide RefSeq : Dalata_v2
Clear Sequence
Search
LOAD
DEMO
Or Import fasta
Choose File
No fil...hosen
Advanced Options
>rna-gnl|WGS:JAFBIIIIHE45_01G000100.1-cds
ATGTTCATGGGAAAAGTTCATGAACAATACCTTCAAAATCAGAGGGTTTGTCTCTCCGGTGACCCCAGTGGTGGATCTTCGATGTG
GCCTTTGGTGAGATCCCTGGGCGTTCACGGCCCCATGGTCAGCATCTCATGTCCCAACGGGAAACTTCGACGCAAAGATCCAGTG
GGATAA
Maximum Target Sequences
50
Expect Threshold
10.0
Word Size
6
Matrix
BLOSUM62
Gap Costs
Existence: 11, Extension:
Filter Low Complexity Regions
Mask for Lookup Table Only
Reset
Advanced Search