# Development and Characterization of Simple Sequence Repeat (SSR) Markers from the Genomic sequence of Sweetpotato [*Ipomoea batatas* L. (Lam) cv. Taizhong6]

**Hanna Amoanimaa-Dede**
  Guangdong Ocean University

**Jiacheng Zhang**
  Guangdong Ocean University

**Chuntao Su**
  Guangdong Ocean University

**Hongbo Zhu** 
  Guangdong Ocean University

**Research article**

**Keywords:** genome sequence, SSR markers, sweetpotato, genetic diversity, polymorphism

# Abstract

**Background:** Sweetpotato is a multifunctional root crop with many essential nutrients and bioactive compounds. Due to its genetic complexity and lack of genomic resources, efficient genetic studies, and cultivar development lags far behind other major crops. Simple sequence repeats (SSRs) offer an effective molecular marker technology for molecular-based breeding and for locating important loci in crop plants, but only a few have previously been developed in sweetpotato.

**Results:** To further explore new SSR markers and accelerate its use in sweetpotato genetic studies, genome-wide characterization and development of SSR markers were performed using the recently published genome of sweetpotato cultivar, Taizhong6. In this study, a set of 2,431 primer pairs were developed from 133,727 SSRs identified in the sweetpotato genome using the Perl script MISA software. The average frequency was one SSR per 6.26 kb with dinucleotides (38.5%) being the most dominant repeat motif. The main motif types in all repeats were AT/AT, AAT/ATT, A/T, AAAT/ATTT, AAAAT/ATTTT and AAAAAT/ATTTTT accounting for 78.29% of the total SSRs. 50% of the 100 randomly selected primer pairs amplified 251 alleles and the average number of alleles was 5.02 alleles per locus with a range of 1 to 13 alleles. The UPGMA cluster analysis grouped the 24 sweetpotato materials into four clusters at a similarity coefficient of 0.68.

**Conclusion:** The SSR markers currently developed will provide valuable genetic resources for germplasm identification, genetic diversity analysis, and functional genomics studies in sweetpotato and related species.

# Background

Sweetpotato [*Ipomoea batatas* L. (Lam)] is an essential food crop belonging to the *Convolvulaceae* family. Currently, cultivated sweetpotato is extensively cultured in over 100 countries worldwide but originally native to Central America [1] traversing the Orinoco River of Venezuela and the Yucatán Peninsula of Mexico. Hence, this region is considered as the center of diversity and rich morphological variation of sweetpotato [2]. Economically, sweetpotato is of great importance due to the excellent supply of dietary fiber, vitamins, minerals, and phenolic compounds [3]. Sweetpotato is also appreciated greatly for its health-promoting functions such as anti-carcinogenic and cardiovascular disease-preventing properties in addition to its radical-scavenging activity [4]. Sweetpotato is a highly heterogeneous auto-hexaploid (2n = 6x = 90) plant with complex genetics and a large genome size (~ 3 Gb). These complexities coupled with the lack of genomic resources [5] hinder breeding progress in sweetpotato. Thus, limiting genetic and molecular research as well as cytological studies in sweetpotato compared to other major food crops such as wheat, rice, and maize.

The advancement of next-generation sequencing has generated scores of datasets for many plant species that provide useful genomic materials for developing efficient molecular markers for genetic analyses [6]. Recently, technological advancement in high throughput DNA sequencing offers new information to accelerate the development of molecular markers. Molecular marker technology through marker-assisted

selection has been reported to play an essential role in genetic diversity and relationship assessment in plants [7]. Currently, this technology has been widely used as an effective tool to improve genomic selection and accelerate breeding progress in many plant species. In recent years, molecular markers such as random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), single nucleotide polymorphism (SNP) and microsatellite or SSR have been employed in sweetpotato to study the dispersal and origin of sweetpotato [8], genetic diversity analysis [9] and construction of genetic linkage maps [10]. Interestingly, despite the rapid development of a new generation of molecular markers such as InDels and SNPs (a more stable and abundant type of genetic marker), SSR markers remain indispensable for many genetic-based studies in sweetpotato [11, 12].

SSR is a short repeat sequence composed of 1-6 bases, widely distributed in the genomes of both prokaryotic and eukaryotic organisms [13]. SSR markers are molecular markers extensively used for crop improvement and other genetic-based studies due to their genome-wide coverage, allele specificity, high polymorphism, co-dominant inheritance, relative abundance, and multi-allelic nature [14, 15]. These markers remain superior over many cost-effective third-generation markers (InDels and SNPs) because of their genome-wide distribution, ability to reveal more information per locus than bi-allelic (SNP) markers, and are relatively easy to score [16]. Nonetheless, SSR analyses may present false-positive results that can cause overlapped SSRs or null alleles [17] and polymerase slippage during DNA replication [18]. In sweetpotato, SSR markers are utilized extensively for the construction of genetic maps, diversity analysis, and variety identification [19-21]. They are widely distributed in both transcribed and non-coding sequences, generally described as expressed sequence tag (EST)- and genomic-SSRs respectively [22]. EST-SSRs are mostly restricted to highly conserved transcribed regions and are less polymorphic. In contrast, genomic-SSRs are widely distributed throughout the genome and highly polymorphic [23]. Recent advances in molecular marker technology have accelerated large-scale development of genomic-SSR markers in many plants [24-27]. To a greater extent, SSR markers in sweetpotato are derived from expressed sequences mined from public databases [28], genomic libraries [29], and pyrosequencing [30, 31]. However, the development and application of microsatellite markers based on published genome sequences remain limited in sweetpotato due to the lack of genomic resources. Also, only a few those identified could amplify or show polymorphisms among the diverse sweetpotato varieties [28, 32]. Thus, there is a great need to develop novel genomic-SSR markers to accelerate genomic and genetic studies in sweetpotato. In this study, we evaluated the frequency and relative number of SSRs in the sequenced genome of sweetpotato (cv. Taizhong6), developed a set of SSR markers from the assembled genomic sequences, and assessed the genetic diversity in cultivated sweetpotato. These SSR markers will offer new genetic resources for marker-assisted selection in sweetpotato breeding, adding up to the available resources for analyzing the molecular phylogeny and genetic diversity of sweetpotato and related species.

# Results

### Frequency and distribution of SSRs in the sweetpotato genome

Based on the sweetpotato (cv. Taizhong6) genome sequence, SSRs identified were characterized as mono- to hexanucleotides. The 133,727 SSRs identified had a density of 159.77 SSR/Mb or one SSR per 6.26 kb of sequence on average. Dinucleotides (38.50 %), trinucleotides (31.45%), and mononucleotides (12.77%) were the most abundant repeat motif representing 82.71% of the total SSR while the remaining repeat motifs represented 17.29% (Table 2; Additional file 2: Figure A1). The most dominant motifs accounted for 78.29% of all SSRs, among them, AT/AT (81.24%), AC/GT (9.98%) and AG/CT (8.77%) were the most abundant dinucleotide repeats while A/T (98.75%), AAT/ATT (87.17%) and AAAT/ATTT (56.01%) were the most dominant motifs among mononucleotides, trinucleotides and tetra nucleotides respectively (Table 2; Additional file 2: Figure A1). Of all the repeating sequences, AT/AT were the most common representing 31.27% followed by AAT/ATT (27.42%), A/T (12.61%), AAAT/ATTT (4.67%), AC/GT (3.84%), AG/CT (3.38%), and AAAAT/ATTTT (1.7%; Table 2).

## SSR marker development and primer design

To ascertain the novelty of the designed SSR primer pairs, the SSR-containing sequences from the genomic sequence of sweetpotato cv. Taizhong6 was analyzed. A total of 133,727 SSRs were identified for mono- to hexanucleotide repeats. Of these repeat sequences, dinucleotides (38.50%), trinucleotides (31.45%) and mononucleotides (12.77%) represented 82.71% of the total SSRs (Table 2; Additional file 2: Figure A1). Based on the SSRs obtained, a total of 2,431 SSR primer pairs (Additional file 3: Table A1) were designed using Primer3 ver. 4.0.0 (http://bioinfo.ut.ee/primer3/) of which 100 primers were randomly selected for validation.

## SSR marker validation

The efficiency of the SSR primer pairs was evaluated and validated in 24 sweetpotato materials using the 100 randomly selected primer pairs. Out of these primers, 50 primer pairs (50%) effectively amplified clear bands with 27 primer pairs being polymorphic. The amplification results also showed that the 50 primer pairs amplified 251 alleles in the 24 sweetpotato materials with a mean allele number of 5.02 alleles per locus and a range of 1 to 13 alleles (Table 3). An example of the PCR amplified products in the 24 cultivars is shown in Figure 1. Table 3 shows the characteristics of the amplified 50 randomly selected primers used in the study.

## Cluster analysis

The genetic diversity of 24 sweetpotato varieties as analyzed by the 50 primer pairs revealed a genetic distance range of 0.605 to 1.00 with an average distance of 0.740 among the 24 sweetpotato varieties studied (Additional file 1: Table A1). According to the unweighted pair group method with arithmetic average (UPGMA) clustering results, the genetic similarity coefficient between the 24 sweetpotato germplasms was relatively high ranging between 0.66 and 0.87 with an average of 0.765 (Figure 2). The dendrogram generated grouped the 24 resources into four (4) clusters (cluster I-IV; Figure 2) at a similarity coefficient of 0.68. Among them, dayehong and fushu18 were the same with the closest genetic relationship and the similarity coefficient was 0.745. The clustering results revealed no direct relationship

with the geographical sources of germplasm, indicating a more frequent exchange of germplasm in sweetpotato cultivation and breeding.

## Discussion

Recent advances in high throughput DNA sequencing technology offers new information to accelerate the development of molecular markers. Molecular markers are widely used in many plant genetic and genomic-based studies. Microsatellites or SSR markers are one of the frequently used molecular markers for genetics and molecular studies in diverse plant species. Generally, SSR markers are distributed in both transcribed and non-coding sequences referred to as EST- and genomic-SSRs respectively. With the advantages of being co-dominant, PCR-based, highly polymorphic, chromosome-specific, reproducible and consistent [42] compared to other molecular markers, SSR markers have been widely used in many genetic and molecular-based studies in sweetpotato including variety identification, genetic diversity analysis and construction of linkage maps [19-21]. Although several studies have reported the use of SSR markers in sweetpotato, most of them evolved from investigating transcriptome libraries and ESTs. Again, the published SSR markers do not cover the whole genome in addition to their limited number and availability with only a few being polymorphic compared to other crops. Therefore, the development of novel SSR markers that are highly polymorphic and distributed throughout the genome will be more effective for genetic analysis in sweetpotato.

To identify valuable genomic-SSR markers for sweetpotato genetic improvement, the sweetpotato genome was searched and a total of 2,431 SSR markers (Additional file 3: Table A1) were successfully developed based on the SSR-containing sequences. The distribution density was 159.77 Mb per SSR or 6.26 kb per one SSR on average which was lower than the average density recorded for sweetpotato (7.1 kb), pigeon pea (8.4 kb), cotton (20.0 kb), and soybean (23.80 kb) but almost the same as that of sesame (6.55 kb), and relatively higher compared to that of rice (3.4 kb) and radish (4.93 kb) [28, 43, 44]. However, the differences in frequency and abundance could be attributed to the size of the database, tools for SSR data-mining, the length of repeat motifs and the application of different repeat unit thresholds, hence, it is practically difficult to directly compare the frequency and abundance estimates of different studies [45]. In our current study, mono-, di- and trinucleotides were the most common SSRs with dinucleotides showing the highest frequency (38.50%) followed by trinucleotides (31.46%) and mononucleotide (12.77%; Additional file 2: Figure A1). Feng et al. [46] identified dinucleotides (9439, 51.52%) as the most abundant repeats followed by trinucleotides (7636, 41.68%) in sweetpotato which is consistent with the results of this study. Our findings contrast with previous reports showing trinucleotides as the most dominant repeat motifs in sweetpotato followed by dinucleotides [28, 31]. Other studies also suggested trinucleotides as the second predominant repeat motifs in sweetpotato which is in agreement with our current findings [30].

The main repeat types among the identified SSRs were A/T (12.61%), AT/AT (61.51%), AAT/ATT (27.42%), and AAAT/ATTT (13.32%, Additional 2: Figure A1). In agreement with our current study, Wang et al. [28] identified AAT/ATT as the most dominant SSR motif in sweetpotato. Similarly, Yang et al. [6] identified AAAT/ATTT as the most frequent repeat motif among tetranucleotides in Welsh onion. However, previous

studies identified AG/CT, AAG/CTT, and AT/TA motifs as the most dominant motif types in sweetpotato [47], conflicting with our findings.

In this study, 100 primer pairs were randomly selected for validation of the SSR markers and to assess its usefulness in sweetpotato. Of these, 50 primer pairs (50%) produced clear stable bands. The 50% PCR amplification efficiency recorded in this study was much lower than the reported 75%-90% EST-SSR amplification rate in sweetpotato [28, 29, 32]. However, the amplification efficiency of genomic-SSRs has always been lower than EST-SSRs in sweetpotato which is in line with our results [29, 48]. The reason being that genomic-SSR primers are designed randomly from genomic libraries whereas EST-SSRs are from relatively highly conserved transcribed regions. Due to this reason, EST-SSRs are reported to be highly applicable and transferable to related species but less polymorphic compared to genomic-SSRs [49]. The 50 working primer pairs amplified 251 alleles in the 24 sweetpotato materials (Table 4). The average number of alleles per locus was 5.02 alleles and a range of 1 to 13 alleles. Several studies have also reported a high number of alleles ranging between 2-23 alleles per loci using SSR markers to study the genetic diversity of sweetpotato germplasm, which is similar to that reported in this studies [8, 9, 48, 50, 51]. This indicates a high polymorphism among the sweetpotato accessions studied. Conversely, Hwang et al. [52] had low polymorphism and recorded 1 to 4 alleles per SSR using varied annealing temperatures and SSR primers. The result of our current study confirms the exceptional discriminatory ability of SSR markers [53]. As a hexaploid plant, distinguishing between homozygous and heterozygous sites becomes difficult hence dominant markers are preferred over collinear markers [14, 15]. Previous studies reported the high polymorphism of sweetpotato which is attributed to the large genome size and high heterozygosity [52] influenced by its mating systems (self-incompatibility and outcrossing). Again, the polyploidy (autohexaploid) of sweetpotato combined with the large chromosome number (2n = 6x = 90) makes sweetpotato SSR primers highly polymorphic [21, 54]. Hence, it is likely for sweetpotato genotypes to have huge genetic distances among them even in smaller populations [55]. In this study, we recorded 27 (54%) primer pairs exhibiting polymorphism among the 50 primer pairs. This value was higher than the 41.9% polymorphism recorded by Wang et al. [28] in the eight cultivated sweetpotato varieties tested but lower than the 67.2% and 62.5% polymorphism reported in different sweetpotato test materials [29, 32]. Differences in polymorphism are attributed to the different geographic origins of samples and the number of DNA samples used. For instance, Chavarriaga-Aguirre et al. [56] observed a relatively high polymorphism after increasing the number of samples from the initial 38 to about 500 or more in cassava. Generally, studies involving comparative genomics, genetic linkage mapping, diversity analysis, gene-based association, and evolutionary analysis require polymorphic markers. Thus, the SSR markers in this study could be used for such studies in sweetpotato. Polymorphic information content (PIC) denotes the degree of SSR variation and also assess the discriminatory efficiency of SSR markers [57]. SSR polymorphism of a particular observed size may be derived from two or more homoeologous loci. In other words, a clear single SSR band amplified by one pair of primers may well be from two or more loci (where an allele may mask or override the effect of another allele) and this overlapping problem of SSRs is very severe in sweetpotato due to its polyploid (hexaploid) nature. Thus, in principle, a pair of primers may amplify six alleles making it difficult to distinguish homozygous sites from heterozygous sites. The acclaimed large heterozygosity in sweetpotato is more of homoeoallelic than allelic variations, hence may not be

considered as a true heterogenecity. Therefore, the PIC of the SSRs were not determined because of the aforementioned problem of SSR markers detected in sweetpotato.

The average SSR-based genetic distance among the 24 sweetpotato varieties was 0.740 on average for values ranging between 0.605 and 1.00 (Additional file 1: Table A1). The genetic similarity coefficient range of 0.66 to 0.87 with a mean value of 0.765 recorded in this study is high, indicating a low diversity in the sweetpotato materials studied (Figure 2). The result is consistent with Hwang et al. [52] who recorded a high similarity coefficient of 0.64 on average and thus, concluded a low diversity among the accessions studied. On the contrary, Yada et al. [50] reported an average similarity coefficient of 0.57 by evaluating the genetic diversity of cultivars from Uganda. Zhang et al. [58] observed a low similarity coefficient (0.588) among sweetpotato varieties from South America. Tumwegamire et al. [9] also recorded a similarity coefficient of 0.54 on average when the genetic diversity of farmer varieties of both white- and orange-fleshed sweetpotato from East Africa were assessed. Similarly, David et al. [59] reported a low genetic similarity coefficient of 0.54 on average and concluded a high diversity among the studied accessions. Thus, the differences could be attributed to the number and type of markers used and the genotypic variances. The clustering results revealed no direct relationship between the national and regional sources of germplasm, indicating a more frequent exchange of germplasm in sweetpotato cultivation and breeding. The results from this study provide background information for genomic-SSR markers in sweetpotato.

## Conclusions

To facilitate marker-assisted selection (MAS) and explore new molecular markers in sweetpotato, we developed a set of SSR markers from the reference genome of cultivated sweetpotato (cv. Taizhong6) using MISA software. A total of 133,727 SSRs were identified and 2,431 new SSR markers (Additional file 3: Table A1) were developed. 50% of the randomly selected showed good amplification effects and produced clear stable bands. Findings from this study will help update the sweetpotato genomic-SSR markers database and aid future genetic and genomic studies. Also, the ability of the markers to analogize cultivars qualify them to be utilized as a background data and resource for germplasm identification, genetic relationship studies, and diversity analysis in sweetpotato and other related species. However, the PIC of the SSRs was not determined because sweetpotato exhibits a high overlap of homoeoallelic SSR variations with allelic ones, in that a clear single amplified SSR band may well be from two or more alleles. This overlapping effect of SSRs prompted some researchers to discredit the use of SSR markers for genetic studies in sweetpotato. Nevertheless, this problem could be salvaged by conducting a more detailed comparative sequence analysis of homoeologous loci to distinguish true allelic SSR variations from homoeoallelic ones by comparing their segregation ratios to the expected ones. Again, we suggest that future researches should focus on developing genotyping-by-sequencing markers (such as SNPs) which will expand the amplicon sizes to tag the homoeolog-specific SNPs to the SSRs.

## Methods

### Plant materials and DNA extraction

A total of 24 sweetpotato materials were sampled (Table 1) for this study. The test materials were obtained from the Research Field of Guangdong Ocean University. CTAB (cetyltrimethyl ammonium bromide) protocol was employed for the genomic DNA extraction with little modifications [33]. The amplified PCR fragments were verified on 1% agarose gel to test the quality of the DNA. The genomic sequence of sweetpotato cv. Taizhong6 was downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/).

## SSR mining and primer design

SSRs from the sweetpotato genome were identified using the Perl Script MISA (MIcroSAtellite) identification software (http://pgrc.ipk-gatersleben.de/misa) [34]. The minimum consecutive repeat units for mono- to hexanucleotide sequences were 20, 10, 7, 5, 4, and 4 respectively. Primers were designed based on the selected SSR motifs using Primer3 ver. 4.0.0 [35]. The amplified product size was 100-280 bp, primer length ranged between 18-27 bp (optimum; 20 bp) and the minimum interval between two SSR sequences was 100 bp. The CG contents ranged between 20% and 80%, the mean melting temperature (Tm) was 60°C (57°C to 63°C) and all other parameters were default values.

## SSR primer validation and marker amplification

From the developed primers, 50 primer pairs with good amplification effects and clear stable bands were used to amplify the genomic DNA of 24 sweetpotato materials. The frequency and distribution of all SSRs were analyzed and measured as one SSR/kb of sequence while SSR markers were grouped according to the location of the SSR motifs in the gene. PCR amplification was performed in a reaction volume of 20 μl containing 1.5 μl genomic DNA (20 ng/μl), Taq enzyme 0.4 μl (3 U/μl), 10× PCR Buffer 2 μl, dNTP 0.2 μl (10 mmol/L), 1 μl each of Forward and Reverse primers (2 μmol/L), and sterile double distilled water was added to the final volume [36]. PCR reaction procedure was; pre-denaturation at 94 °C for 5 min, followed by 33 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s and a final extension at 72 °C for 5 min using the Takara PCR Thermal Cycler Dice.

## Allele identification and genetic diversity analysis

To identify alleles, the amplified products were identified and separated on a 6% non-denaturing polyacrylamide gel electrophoresis (PAGE). The electrophoresis buffer contained 1×TBE (100 mM Tris−HCl, 83 mM boric acid, 1 mM $Na_2EDTA$, pH 8.0) [37]. Electrophoresis at 200 V and 100 mA for 120 min was executed after loading of samples. The silver staining method was employed for allele visualization and selection of suitable SSR primers. Markers were scored manually and the polymorphic ones selected for genetic analysis [38].

Each SSR allele for a given primer pair was scored 1 and 0 for samples with and without band respectively, and a "0, 1" binary matrix established. PowerMarker ver.3.25 was used to calculate the number of alleles per locus for the genotypes [39]. To estimate the genetic distance among the sweetpotato test materials, Nei's coefficient (Nei 1973) with the bootstrap procedure of resampling (100) across markers and individuals from allele frequencies was used. The bootstrapping was carried out using the bootstrap function in PowerMarker ver.3.25. The clustering analysis based on the genetic similarity coefficient was

calculated by the UPGMA method [40] and a dendrogram generated using the NTSYS software ver. 2.02 [41].

# Declarations

## Acknowledgments

We are thankful to Mr. Eric Amenyogbe, Ms. Linda Adzigbli, and Mr. Gyan Ray Watson for the critical review and useful suggestions during the manuscript preparation.

## Author Contributions

HAD conceptualized the idea, led the experimental work, developed and screened SSR markers, designed primers, analyzed genotypic data, and drafted the manuscript. JZ and CS collected experimental materials, assisted with the experimental work and statistical analysis. ZH proofread and revised the manuscript. All authors read and approved the final version of the manuscript.

## Funding

This research was funded by the National Natural Science Foundation of China-Guangdong Joint Fund, China, and Studies on Resistance Resources and Molecular Mechanisms of Sweet potato Weevil in South China **(Grant number, U1701234)**.

## Availability of data and materials

All datasets supporting the results of this study are included in this published article and its supplementary information files.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

# Abbreviations

**SSRs:** simple sequence repeats

**RAPD:** random amplified polymorphic DNA

**AFLP:** amplified fragment length polymorphism

**UPGMA:** unweighted pair group method with arithmetic average

**MISA:** MIcroSAtellite

**CTAB:** cetyltrimethyl ammonium bromide

**NCBI:** National Center for Biotechnology Information

**PAGE:** polyacrylamide gel electrophoresis

**InDels:** insertions and deletions

**NTSYS:** Numerical Taxonomy and Multivariate Analysis System

**MAS:** marker-assisted selection

**SNP:** single nucleotide polymorphism

**EST:** expressed sequence tag

**DNA:** deoxyribonucleic acid

**PCR:** polymerase chain reaction

**PIC:** polymorphic information content

# References

1. Rozen S, Skaletsky H: **Primer3 on the WWW for general users and for biologist programmers**. In: *Bioinformatics methods and protocols.* Springer; 2000: 365-386.
2. Roullier C, Kambouo R, Paofa J, McKey D, Lebot V: **On the origin of sweet potato (Ipomoea batatas (L.) Lam.) genetic diversity in New Guinea, a secondary centre of diversity**. *Heredity* 2013, **110**(6):594-604.
3. Lebot V, Michalet S, Legendre L: **Identification and quantification of phenolic compounds responsible for the antioxidant activity of sweet potatoes with different flesh colours using high performance thin layer chromatography (HPTLC)**. *Journal of Food Composition and Analysis* 2016, **49**:94-101.
4. Chandrasekara A, Josheph Kumar T: **Roots and tuber crops as functional foods: a review on phytochemical constituents and their potential health benefits**. *International journal of food science* 2016, **2016**.
5. Yan L, Lai X, Li X, Wei C, Tan X, Zhang Y: **Analyses of the complete genome and gene expression of chloroplast of sweet potato [Ipomoea batata]**. *PLoS One* 2015, **10**(4):e0124083.
6. Yang L, Wen C, Zhao H, Liu Q, Yang J, Liu L, Wang Y: **Development of polymorphic genic SSR markers by transcriptome sequencing in the welsh onion (Allium fistulosum L.)**. *Applied Sciences* 2015,

**5**(4):1050-1063.

7. Kumar P, Gupta V, Misra A, Modi D, Pandey B: **Potential of molecular markers in plant biotechnology**. *Plant Omics* 2009, **2**(4):141.

8. Roullier C, Rossel G, Tay D, McKey D, Lebot V: **Combining chloroplast and nuclear microsatellites to investigate origin and dispersal of New World sweet potato landraces**. *Molecular Ecology* 2011, **20**(19):3963-3977.

9. Tumwegamire S, Rubaihayo P, LaBonte D, Diaz F, Kapinga R, Mwanga R, Grüneberg W: **Genetic diversity in white-and orange-fleshed sweetpotato farmer varieties from East Africa evaluated by simple sequence repeat markers**. *Crop Science* 2011, **51**(3):1132-1142.

10. Zhao N, Yu X, Jie Q, Li H, Li H, Hu J, Zhai H, He S, Liu Q: **A genetic linkage map based on AFLP and SSR markers and mapping of QTL for dry-matter content in sweetpotato**. *Molecular breeding* 2013, **32**(4):807-820.

11. Edwards D, Batley J: **Plant genome sequencing: applications for crop improvement**. *Plant biotechnology journal* 2010, **8**(1):2-9.

12. Liu J, Huang S, Sun M, Liu S, Liu Y, Wang W, Zhang X, Wang H, Hua W: **An improved allele-specific PCR primer design method for SNP marker analysis and its application**. *Plant Methods* 2012, **8**(1):34.

13. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA: **The sequence of the human genome**. *science* 2001, **291**(5507):1304-1351.

14. Yue X-y, Liu G-q, Zong Y, Teng Y-w, Cai D-y: **Development of genic SSR markers from transcriptome sequencing of pear buds**. *Journal of Zhejiang University Science B* 2014, **15**(4):303-312.

15. Silva PI, Martins AM, Gouvea EG, Pessoa-Filho M, Ferreira ME: **Development and validation of microsatellite markers for Brachiaria ruziziensis obtained by partial genome assembly of Illumina single-end reads**. *Bmc Genomics* 2013, **14**(1):17.

16. Xu J, Liu L, Xu Y, Chen C, Rong T, Ali F, Zhou S, Wu F, Liu Y, Wang J: **Development and characterization of simple sequence repeat markers providing genome-wide coverage and high resolution in maize**. *DNA research* 2013, **20**(5):497-509.

17. Wang X, Wang L: **GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing**. *Frontiers in plant science* 2016, **7**:1350.

18. Schlötterer C, Tautz D: **Slippage synthesis of simple sequence DNA**. *Nucleic acids research* 1992, **20**(2):211-215.

19. Yada B, Brown-Guedira G, Alajo A, Ssemakula GN, Mwanga RO, Yencho GC: **Simple sequence repeat marker analysis of genetic diversity among progeny of a biparental mapping population of sweetpotato**. *HortScience* 2015, **50**(8):1143-1147.

20. Yang X, Su W, Wang L, Lei J, Chai S, Liu Q: **Molecular diversity and genetic structure of 380 sweetpotato accessions as revealed by SSR markers**. *J Integr Agric* 2015, **14**(4):633-641.

21. Ngailo S, Shimelis H, Sibiya J, Amelework B, Mtunda K: **Genetic diversity assessment of Tanzanian sweetpotato genotypes using simple sequence repeat markers**. *South African Journal of Botany* 2016, **102**:40-45.

22. Morgante M, Hanafey M, Powell W: **Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes**. *Nature genetics* 2002, **30**(2):194-200.

23. La Rota M, Kantety RV, Yu J-K, Sorrells ME: **Nonrandom distribution and frequencies of genomic and EST-derived microsatellite markers in rice, wheat, and barley**. *BMC genomics* 2005, **6**(1):23.

24. Wen M, Wang H, Xia Z, Zou M, Lu C, Wang W: **Developmenrt of EST-SSR and genomic-SSR markers to assess genetic diversity in Jatropha Curcas L**. *BMC research notes* 2010, **3**(1):42.

25. Nunome T, Negoro S, Kono I, Kanamori H, Miyatake K, Yamaguchi H, Ohyama A, Fukuoka H: **Development of SSR markers derived from SSR-enriched genomic library of eggplant (Solanum melongena L.)**. *Theoretical and Applied Genetics* 2009, **119**(6):1143-1153.

26. Wang Y, Samuels T, Wu Y: **Development of 1,030 genomic SSR markers in switchgrass**. *Theoretical and Applied Genetics* 2011, **122**(4):677-686.

27. Iniguez-Luy FL, Voort AV, Osborn TC: **Development of a set of public SSR markers derived from genomic sequence of a rapid cycling Brassica oleracea L. genotype**. *Theoretical and applied genetics* 2008, **117**(6):977-985.

28. Wang Z, Li J, Luo Z, Huang L, Chen X, Fang B, Li Y, Chen J, Zhang X: **Characterization and development of EST-derived SSR markers in cultivated sweetpotato (Ipomoea batatas)**. *BMC plant biology* 2011, **11**(1):139.

29. Hu J, Nakatani M, Mizuno K, Fujimura T: **Development and characterization of microsatellite markers in sweetpotato**. *Breeding science* 2004, **54**(2):177-188.

30. Xie F, Burklew CE, Yang Y, Liu M, Xiao P, Zhang B, Qiu D: **De novo sequencing and a comprehensive analysis of purple sweet potato (Impomoea batatas L.) transcriptome**. *Planta* 2012, **236**(1):101-113.

31. Tao X, Gu Y-H, Wang H-Y, Zheng W, Li X, Zhao C-W, Zhang Y-Z: **Digital gene expression analysis based on integrated de novo transcriptome assembly of sweet potato [Ipomoea batatas (L.) Lam.]**. *PloS one* 2012, **7**(4):e36234.

32. Schafleitner R, Tincopa LR, Palomino O, Rossel G, Robles RF, Alagon R, Rivera C, Quispe C, Rojas L, Pacheco JA: **A sweetpotato gene index established by de novo assembly of pyrosequencing and Sanger sequences and mining for gene-based microsatellite markers**. *BMC genomics* 2010, **11**(1):604.

33. Porebski S, Bailey LG, Baum BR: **Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components**. *Plant molecular biology reporter* 1997, **15**(1):8-15.

34. Thiel T, Michalek W, Varshney R, Graner A: **Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.)**. *Theoretical and applied genetics* 2003, **106**(3):411-422.

35. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG: **Primer3—new capabilities and interfaces**. *Nucleic acids research* 2012, **40**(15):e115-e115.

36. Xie J, Zhang M, Sun J, Zhang F: **Genome-wide genic SSR marker development for the endangered Dongxiang wild rice (Oryza rufipogon)**. *Plant Genetic Resources* 2017, **15**(6):566-569.

37. Han YC, Teng CZ, Hu ZL, Song YC: **An optimal method of DNA silver staining in polyacrylamide gels**. *Electrophoresis* 2008, **29**(6):1355-1358.

38. Chevallet M, Luche S, Rabilloud T: **Silver staining of proteins in polyacrylamide gels**. *Nature protocols* 2006, **1**(4):1852.

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

40. Nei M, Li W-H: **Mathematical model for studying genetic variation in terms of restriction endonucleases**. *Proceedings of the National Academy of Sciences* 1979, **76**(10):5269-5273.

41. Rohlf F: **NTSYS 2.1: numerical taxonomic and multivariate analysis system**. *New York, Exeter Software* 2000.

42. Karihaloo J: **DNA fingerprinting techniques for plant identification**. In: *Plant Biology and Biotechnology.* Springer; 2015: 205-221.

43. Zhai L, Xu L, Wang Y, Cheng H, Chen Y, Gong Y, Liu L: **Novel and useful genic-SSR markers from de novo transcriptome sequencing of radish (Raphanus sativus L.)**. *Molecular Breeding* 2014, **33**(3):611-624.

44. Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R: **Computational and experimental characterization of physically clustered simple sequence repeats in plants**. *Genetics* 2000, **156**(2):847-854.

45. Dutta S, Kumawat G, Singh BP, Gupta DK, Singh S, Dogra V, Gaikwad K, Sharma TR, Raje RS, Bandhopadhya TK: **Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea [Cajanus cajan (L.) Millspaugh]**. *BMC plant biology* 2011, **11**(1):17.

46. Feng J, Zhao S, Li M, Zhang C, Qu H, Li Q, Li J, Lin Y, Pu Z: **Genome-wide genetic diversity detection and population structure analysis in sweetpotato (Ipomoea batatas) using RAD-seq**. *Genomics* 2020, **112**(2):1978-1987.

47. Zhang K, Wu Z, Tang D, Lv C, Luo K, Zhao Y, Liu X, Huang Y, Wang J: **Development and identification of SSR markers associated with starch properties and β-carotene content in the storage root of sweet potato (Ipomoea batatas L.)**. *Frontiers in plant science* 2016, **7**:223.

48. Buteler M, Jarret R, LaBonte D: **Sequence characterization of microsatellites in diploid and polyploid Ipomoea**. *Theoretical and Applied Genetics* 1999, **99**(1-2):123-132.

49. Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishnakumar V, Singh L: **Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species**. *Theoretical and Applied Genetics* 2007, **114**(2):359.

50. Yada B, Tukamuhabwa P, Wanjala B, Kim D-J, Skilton RA, Alajo A, Mwanga RO: **Characterization of Ugandan sweetpotato germplasm using fluorescent labeled simple sequence repeat markers**. *HortScience* 2010, **45**(2):225-230.

51. Veasey EA, Borges A, Rosa MS, Queiroz-Silva JR, Bressan EdA, Peroni N: **Genetic diversity in Brazilian sweet potato (Ipomoea batatas (L.) Lam., Solanales, Convolvulaceae) landraces assessed with microsatellite markers**. *Genetics and Molecular Biology* 2008, **31**(3):725-733.

52. Hwang SY, Tseng YT, Lo HF: **Application of simple sequence repeats in determining the genetic relationships of cultivars used in sweet potato polycross breeding in Taiwan**. *Scientia Horticulturae*

2002, **93**(3-4):215-224.

53. Gichuru V, Aritua V, Lubega G, Edema R, Adipala E, Rubaihayo P: **A preliminary analysis of diversity among East African sweetpotato landraces using morphological and simple sequence repeats (SSR) markers**. In: *II International Symposium on Sweetpotato and Cassava: Innovative Technologies for Commercialization 703: 2005*; 2005: 159-164.

54. Li R, Zhai H, Kang C, Liu D, He S, Liu Q: **De novo transcriptome sequencing of the orange-fleshed sweet potato and analysis of differentially expressed genes related to carotenoid biosynthesis**. *International journal of genomics* 2015, **2015**.

55. Gruneberg W, Ma D, Mwanga R, Carey E, Huamani K, Diaz F, Eyzaguirre R, Guaf E, Jusuf M, Karuniawan A: **Advances in sweetpotato breeding from 1992 to 2012**. In.: CABI International; 2015.

56. Chavarriaga-Aguirre P, Maya MM, Bonierbale MW, Kresovich S, Fregene MA, Tohme J, Kochert G: **Microsatellites in cassava (Manihot esculenta Crantz): discovery, inheritance and variability**. *Theoretical and Applied Genetics* 1998, **97**(3):493-501.

57. Nachimuthu VV, Muthurajan R, Duraialaguraja S, Sivakami R, Pandian BA, Ponniah G, Gunasekaran K, Swaminathan M, Suji K, Sabariappan R: **Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in Oryza sativa**. *Rice* 2015, **8**(1):30.

58. Zhang D, Carbajulca D, Ojeda L, Rossel G, Milla S, Herrera C, Ghislain M: **Microsatellite analysis of genetic diversity in sweetpotato varieties from Latin America**. *CIP Program report* 2000, **2000**(2000):295-301.

59. David MC, Diaz FC, Mwanga RO, Tumwegamire S, Mansilla RC, Grüneberg WJ: **Gene pool subdivision of East African sweetpotato parental material**. *Crop Science* 2018, **58**(6):2302-2314.

# Additional Files

## Additional file 1

**Table A1:** Genetic distances generated from the SSR marker analysis of sweetpotato germplasm (1-24) in this study.

## Additional file 2

**Figure A1:** Characteristics of SSRs in the sweetpotato genome.

  a. Distribution of different repeat types
  b. Frequency distribution of major repeat motifs
  c. Frequency distribution of main motif sequence in **I.** Dinucleotide, **II.** Trinucleotide and **III.** Tetranucleotide repeats

**Figure A2:** Original uncropped gel of the amplification results generated by IBM 241 in the 24 sweetpotato varieties.

**Additional file 3**

**Table A1:** The 2,431 primer pairs designed for the study.

**Table A2:** Molecular marker analysis results generarted from PowerMarker v.3.25 software.

 **I.:** Summary of bootsrap analysis, **II.:** Allele frequency**, III.:** Frequency 1**, IV.:** Nei's (1973) frequency based genetic distance.

# Tables

**Table 1:** List of 24 sweetpotato materials used in the study.

| No. | Name | Type | Origin |
|---|---|---|---|
| 1. | "Guangshu87" | Cultivar | Guangdong, China |
| 2. | "Pushu32" | Cultivar | Guangdong, China |
| 3. | "Ziluolan" | Cultivar | Guangdong, China |
| 4. | "Wulixiang" | Cultivar | Guangdong, China |
| 5. | "Dayehong" | Cultivar | Guangdong, China |
| 6. | "Tianeshu" | Cultivar | Guangdong, China |
| 7. | "Jishu26" | Cultivar | Shandong, China |
| 8. | "Xinong431" | Cultivar | Shaanxi, China |
| 9. | "Qinshu 8" | Cultivar | Shaanxi, China |
| 10. | "Qinshu 5" | Cultivar | Shaanxi, China |
| 11. | "Shangshu19" | Cultivar | Henan, China |
| 12. | "Longshu 9" | Cultivar | Fujian, China |
| 13. | "Xinnianggao" | Cultivar | Fujian, China |
| 14. | "Simon 1" | Cultivar | Brazil |
| 15. | "Taiwanziyang" | Cultivar | Taiwan, China |
| 16. | "Au-1" | Cultivar | Australia |
| 17. | "AU-2" | Cultivar | Australia |
| | "AU-3" | Cultivar | Australia |

| 18. | | | |
|---|---|---|---|
| 19. | "Yizi138" | Cultivar | Beijing, China |
| 20. | "Longshu 515" | Cultivar | Fujian, China |
| 21. | "Fushu 18" | Cultivar | Fujian, China |
| 22. | "Guijingzi 8" | Cultivar | Guangxi, China |
| 23. | "Guicai 1" | Cultivar | Guangxi, China |
| 24. | "Zheshu 75" | Cultivar | Zhejiang, China |

**Table 2:** Distribution of SSRs in the sweetpotato genome.

| Repeat type | Repeat motif | Number of SSRs | Proportion in all SSRs (%) | Density (SSR/Mb) |
|---|---|---|---|---|
| Mononucleotide | | | | |
| | A/T | 16859 | 12.61 | 20.14 |
| | C/G | 213 | 0.16 | 0.25 |
| | | 17072 | 12.77 | 20.4 |
| Dinucleotide | | | | |
| | AT/AT | 41822 | 31.27 | 49.96 |
| | AC/GT | 5136 | 3.84 | 6.14 |
| | AG/CT | 4516 | 3.38 | 5.4 |
| | CG/CG | 5 | 0 | 0.01 |
| | | 51479 | 38.5 | 61.5 |
| Trinucleotide | | | | |
| | AAT/ATT | 36663 | 27.42 | 43.8 |
| | AAC/GTT | 2303 | 1.72 | 2.75 |
| | AAG/CTT | 1668 | 1.25 | 1.99 |
| | others | 1425 | 1.07 | 1.7 |
| | | 42059 | 31.45 | 50.25 |
| Tetranucleotide | | | | |
| | AAAT/ATTT | 6247 | 4.67 | 7.46 |
| | ACAT/ATGT | 2007 | 1.5 | 2.4 |
| | AATT/AATT | 844 | 0.63 | 1.01 |
| | AAAC/GTTT | 789 | 0.59 | 0.94 |
| | others | 1267 | 0.95 | 1.51 |
| | | 11154 | 8.34 | 13.33 |
| Pentanucleotide | | | | |
| | AAAAT/ATTTT | 2374 | 1.78 | 2.84 |
| | AAAAC/GTTTT | 1369 | 1.02 | 1.64 |
| | AATAT/ATATT | 680 | 0.51 | 0.81 |
| | others | 3077 | 2.3 | 3.68 |
| | | 7500 | 5.61 | 8.961 |
| Hexanucleotide | | | | |
| | AAAAAT/ATTTTT | 716 | 0.54 | 0.86 |
| | AAAAAC/GTTTTT | 569 | 0.43 | 0.68 |
| | others | 3178 | 2.38 | 3.8 |
| | | 4463 | 3.34 | 5.33 |
| Total | - | **133727** | **100** | **159.77** |

Table 3: Characteristics of the 50 SSR primer pairs validated in 24 sweetpotato materials.

| imer | Repeat motif | Primer sequence (5'–3') | Size (bp) | Chromosome | Allele No. |
|---|---|---|---|---|---|
| M 93 | (AT)16 | F: TGCATGTTTGGATGTTACAGG | 248 | Chr 15 | 4 |
| | | R: CAATTACCGGAAAATTTTGGTC | | | |
| M 9 | (AT)21 | F: GGACTAAAATGAGCGCGAAA | 210 | Chr 9 | 1 |
| | | R:TGAAAGAAAAATTCCAACAATCAA | | | |
| M 9 | (TAT)8(TGT)13(TAT)7 | F: GGTCTTTGGCCAGACTATCG | 280 | Chr 9 | 1 |
| | | R: GCGGAGATCCCTTGTCATTA | | | |
| M 2 | (AT)18 | F: GTTCGCAGTTCCAAATTGCT | 254 | Chr 7 | 7 |
| | | R: GCATCAACAACAAAAACAAAACA | | | |
| M 76 | (TA)21 | F: ATGACACTTTGGCAAATGATAGT | 245 | Chr 11 | 1 |
| | | R: AGTAAGGCAGCCCTAACCCG | | | |
| M 4 | (A)24 | F: GATCCCTCGAGGTGTGAAAG | 250 | Chr 8 | 7 |
| | | R: AGGATCATGCTTCACCAACC | | | |
| M 0 | (A)25 | F: CCTGTCCACTGGTCTAAGGC | 278 | Chr 7 | 10 |
| | | R: GCGGTCTTCATCTTCTCTGG | | | |
| M 2 | (T)33 | F: AATCTGTCAGGGAGTGGTGG | 236 | Chr 10 | 2 |
| | | R: AAATGCAACCCAAACAAAGC | | | |
| M 3 | (T)36 | F: CCATATCAATAGGCCGTGCT | 231 | Chr 7 | 7 |
| | | R: CGAACCTCAGTGAAGACGAA | | | |
| M 2 | (A)78 | F: GCCTTTGCTTTTCCTCTCCT | 262 | Chr 6 | 9 |
| | | R: CCGGAAACCAGCTAATCAAA | | | |
| M 95 | (CT)12 | F: AATTCAATGTGGGGTCTTGC | 236 | Chr 12 | 2 |
| | | R: GCTTGATCTAACTCGGTGGC | | | |
| M 5 | (A)25 | F: CTTGAACAACACCTCAGGCA | 199 | Chr 9 | 2 |
| | | R: CGAGAGGAATCAGAGCCAAC | | | |
| M 6 | (ATT)13 | F: ATGCCCACATCATCATCATC | 254 | Chr 9 | 2 |
| | | R: GAATCACATATTTGCCCCTGA | | | |
| M | (TA)11 | F: GGCCTTAGTCTTCGAAACACAT | 236 | Chr 2 | 4 |
| | | R: CGTTTGGTCTTCTGGGGTTA | | | |
| M 8 | (TGA)8 | F: CTCTCTTTCTTCCTTTGCCG | 259 | Chr 7 | 4 |
| | | R: GGTAGAGAAGGGAGGAGAAAGG | | | |
| M 6 | (CA)12 | F: GCGAAAATGTCACCGAGTTT | 191 | Chr 4 | 7 |
| | | R: GCTCTTTTCTCATCGCACCT | | | |

| M 2 | (A)25 | F: TTGTTCTTGCCCAATTTGCT | 211 | Chr 9 | 2 |
| --- | --- | --- | --- | --- | --- |
| | | R: AAACAATCAGCCCACACACA | | | |
| M | (T)26(T)30 | F: CTTAGCGCTTCATGGGAGAC | 259 | Chr 3 | 4 |
| | | R: GGCATAATCAGCTCAATTCCA | | | |
| M | (TTAAAA)4 | F: TGCAATTAGGCTACCGAACC | 180 | Chr 3 | 1 |
| | | R: GTCTCCGGTGAGACGTGTTT | | | |
| M 7 | (TAA)8 | F: TTCATCCTGCAAACACATGC | 266 | Chr 4 | 6 |
| | | R: TTAACGCCAACCCAACTTTC | | | |
| M 5 | (T)27 | F: AAACTTAGGTGATCCCAATCC | 209 | Chr 8 | 13 |
| | | R: AACATAGTTGGTTCGTCGCC | | | |
| M 5 | (T)52 | F: TACGTTGTCTTCCCTTCCCA | 224 | Chr 6 | 8 |
| | | R: TTGGAATTACATCAACCCCC | | | |
| M 66 | (TA)15 | F: TGGGTTGAGGTTGAGGAAAC | 208 | Chr 15 | 3 |
| | | R: CTTCTAAAACCATCGCCCAA | | | |
| M | (AT)12 | F: ATCCAATGACGCTAGTTCGG | 175 | Chr 2 | 3 |
| | | R: CCAAAAACACAGCCATCAGA | | | |
| M 7 | (CTTT)6 | F: TTTGGAGGCCCACTACAAAG | 206 | Chr 8 | 1 |
| | | R: CAGTGCATGATGGACCATTG | | | |
| M 4 | (T)26(T)36 | F: TCTTGAAGGGGTAAGGCAAA | 234 | Chr 4 | 4 |
| | | R: CATAAAATTTTGCTCCACATGC | | | |
| M 60 | (CAA)10(ATA)16 | F: AAGAAATCTTTTTGGAATGCGA | 255 | Chr 14 | 4 |
| | | R: ACCGTACAACGACGGTTCAT | | | |
| M 45 | (ATAC)6(AT)16 | F: TGTGTTTGGTTCAACAAGGAA | 244 | Chr 11 | 7 |
| | | R: ACGAGTTGGGTATGAATCGG | | | |
| M 09 | (TTG)8 | F: ATGGTTTTGTGGGCAAAAGT | 138 | Chr 15 | 1 |
| | | R: ACGCTCTCTTCATGCCAAGT | | | |
| M 10 | (TAT)14(TGTTAT)6 | F: TTAATAAAAGTTTGCGCGGG | 208 | Chr 13 | 7 |
| | | R: ATGCAGATCCCTGATTTTGG | | | |
| M 5 | (TAT)16 | F: AAATTTATTTAGATTTTGGATACGGA | 234 | Chr 8 | 5 |
| | | R: ATTGTTACCATGCACAGGCA | | | |
| M 1 | (A)23(A)21 | F: GACACTGAATTGATCTCCCGA | 207 | Chr 7 | 10 |
| | | R: TCGGTTGTTGTTGTTGTTTTT | | | |
| M 0 | (T)25 | F: GCCAGAATTTTCTGTCAACCA | 180 | Chr 10 | 1 |
| | | R: AAAAGAACGTGGGGAAGGAA | | | |

| | | | | | |
|---|---|---|---|---|---|
| M 74 | (A)25 | F: TGCAACATGCCATAAATGCT | 267 | Chr 10 | 2 |
| | | R: CCTAAAGCTTTCCCGTTTTG | | | |
| M 6 | (T)87 | F: TTGGGGAAGGCTTTTAGGTT | 280 | Chr 3 | 10 |
| | | R: TTGTGATCCTTTCTCAGTTAAGGT | | | |
| M | (TTA)11 | F: GCACTTAGCCACCCCTATCA | 171 | Chr 2 | 10 |
| | | R: AAACAAAATTGTGGGAGAGCA | | | |
| M 1 | (TAA)9(TAT)8 (ATC)7 | F: TGCATTTAAAAACTCCGTAATACA | 218 | Chr 8 | 9 |
| | | R: GAATGAATGCAATTCTAAAAACCC | | | |
| M | (A)24 | F: GGTTTGAATTTGGAGTGAACATC | 217 | Chr 1 | 1 |
| | | R: TGAGTTGTGACGTGTGAGCA | | | |
| M 82 | (AT)16 | F: TATCTACCCAACCGACCTGC | 238 | Chr 14 | 1 |
| | | R: CCGTTAGATCTGAACACGTGAA | | | |
| M 1 | (T)25 | F: CTGCACACATGCAACACAAC | 189 | Chr 8 | 10 |
| | | R: TCAGTATCACAAAGCTCCACAA | | | |
| M 23 | (ATT)9 | F: CAACCAAACCCCCTAATGTG | 211 | Chr 13 | 4 |
| | | R: ACATGGTTTCAGAGGGACCA | | | |
| M 33 | (TA)19 | F: TGATTTTGGATGTTATTTCATCATTT | 270 | Chr 11 | 6 |
| | | R: TCTTGGCTTAAGTTATCGGCA | | | |
| M 9 | (A)24 | F:GGAAACGTTAGTACAAGTTGACACA | 273 | Chr 4 | 4 |
| | | R: TCGCACATTATTAAAAACGGTCT | | | |
| M 95 | (CT)12 | F: AATTCAATGTGGGGTCTTGC | 236 | Chr 12 | 2 |
| | | R: GCTTGATCTAACTCGGTGGC | | | |
| M 1 | (A)30 | F: CCAAGCAAGCACACAACTTT | 277 | Chr 9 | 7 |
| | | R: GCACGCTGTGCTTAAAATGA | | | |
| M | (A)45 | F: CCAAACACCCACATAGACACC | 239 | Chr 2 | 10 |
| | | R: AAGCACACTGATGTGCCACT | | | |
| M 2 | (ATT)9 | F: AAGGTTATGATGGCCGACTG | 157 | Chr 10 | 7 |
| | | R: AAAAACTCCGTTCCCATCAA | | | |
| M 6 | (A)25 | F: ATAAGAAGAGAGCGGGTCGG | 220 | Chr 9 | 5 |
| | | R: TGCACTTTGAATGCACAACA | | | |
| M 84 | (AT)23 | F: TGACATGTGCCGATACTCTAAAA | 250 | Chr 13 | 3 |
| | | R: GCAAAACACTTCTTCATGGG | | | |
| M | (T)29 | F: TCCACATAAGGGAGATGAGGA | 252 | Chr 1 | 10 |

# Figures

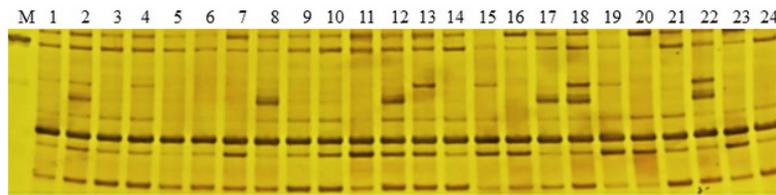
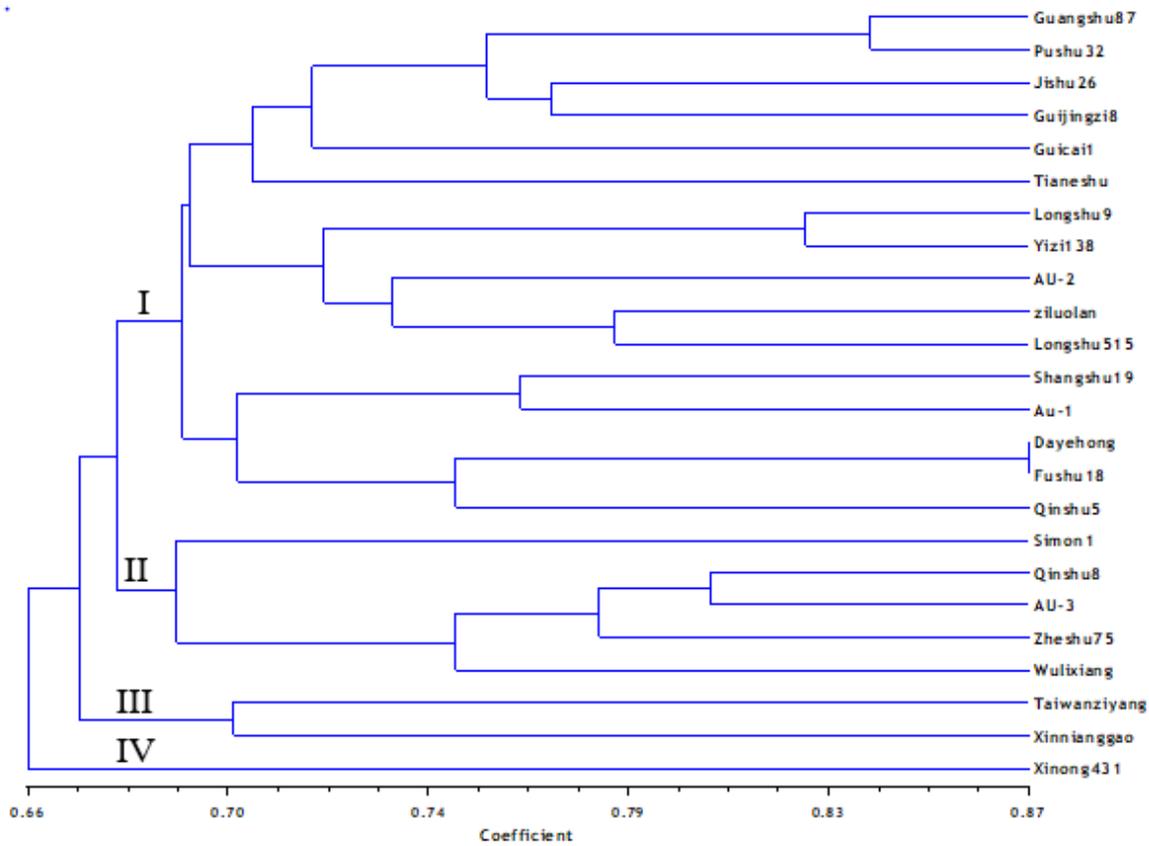



**Figure 1**

Amplification results generated by IBM 241 in the 24 sweetpotato varieties. For clarity, the figure was cropped from the original gel obtained as shown in Additioal file 2: Figure A2. The numbers 1−24 represent sweetpotato materials and M represents the DNA marker.

## Figure 2

UPGMA dendrogram showing the genetic relationships among 24 sweetpotato accessions.