Taiwania 68(3): 303–313, **2023**
DOI: 10.6165/tai.2023.68.303

Potential of sweet potato (*Ipomoea batatas* (L.) Lam.) breeding populations evaluated based on expressed sequence tags-simple sequence repeats

Kuan-Hung LIN[1,*], Yu-Wen CHANG[2], Yung-Chang LAI[3], Shih-Ying HWANG[4,*]

1. *Department of Horticulture and Biotechnology, Chinese Culture University, Taipei 11114, Taiwan.* ***2.*** *Graduate Institute of Biotechnology, Chinese Culture University, Taipei 11114, Taiwan.* ***3.*** *Chiayi Agricultural Experimental Branch, Taiwan Agricultural Research Institute, Chiayi 60014, Taiwan.* ***4.*** *School of Life Science, National Taiwan Normal University, Taipei 11677, Taiwan.* [*]*Corresponding authors' e-mail: KHL: rlin@ulive.pccu.edu.tw, SYH: hsy9347@ntnu.edu.tw*

(Manuscript received 22 January 2023; Accepted 14 July 2023; Online published 26 July 2023)

ABSTRACT: Expressed sequence tags-derived simple sequence repeats (EST-SSRs) are related to functional genes and can be useful in crop breeding. Parental lines selected in sweet potato (*Ipomoea batatas* (L.) Lam.) breeding including controlled hybridization and polycross may be originated from a limited number of genetic sources. We surveyed genetic variation of three breeding populations (hybrid, landrace, and polycross) of 89 cultivated sweet potato varieties derived from controlled hybridization (41 hybrids) and polycross (38 varieties) breeding population and landrace (10 varieties) maintained by various geographic regions and main agronomic traits of local farmers, using 23 EST-SSRs. Six wild sweet potato (*I. trifida*) samples were also included in the study. High genetic diversity in the three types of breeding populations was detected. High proportions of shared alleles leading to low levels of differentiation between breeding populations were discovered. The neighbor-joining tree and individual assignments showed intermingling of varieties among the three breeding populations. Additionally, genetic clustering showed little power of grouping varieties of different breeding populations into separate clades. These results indicate limits of sweet potato breeding. Nevertheless, ten alleles of nine loci showed variance contributed significantly to the delimitation of varieties in different breeding populations.

KEY WORDS: Breeding, EST, *Ipomoea batatas*, markers, SSR, sweet potato.

INTRODUCTION

Plant domestication or breeding is an evolutionary process driven by human selection for phenotypes favorable for human consumption and cultivation. During the domestication and breeding process, the set of alleles that may be fixed in response to directional selection depending upon the purposes of selection programs in specific crop species (Lavergne and Molofsky, 2007; Zhu *et al.,* 2007). The main purpose of plant genetic resource collection is to harbor the maximum amount of genetic variation. Information on the extent of genetic diversity and trait values of crop germplasm is a prerequisite for developing efficient selection of parents for crossing in breeding programs, providing a more rational basis for crop improvement (Marshall and Brown, 1975; Allard, 1970). However, during the process of breeding, natural genetic variability has been depauperated because only few traits and few genotypes are frequently used in the breeding programs (Keneni *et al.,* 2012; Bhandari *et al.,* 2017). Outstanding lines introduced from different geographic districts are commonly adopted in the breeding programs of different countries that may reduce the level of genetic differentiation between elite lines produced. Therefore, breeding practices can increase genetic similarity and likely raise difficulty in further breeding tasks that aim to fulfill the increasing demand of coping with global climate changes (Ceccarelli *et al.*, 2010; Ceccarelli and Grando, 2020; Cortinovis *et al.*, 2020).

Efficient use of available variation is important and plant breeders may look for stable detectable molecular variation among the parents in their crossing programs for the improvement of crop plants. Many types of molecular makers, such as random amplified polymorphic DNAs, genomic simple sequence repeats (SSRs), and amplified fragment length polymorphisms, have been used for marker-assisted selection in plant breeding (Collard and Mackill, 2008; Meng *et al.*, 2021; Ma *et al.*, 2022). However, these types of molecular markers represent genetic variation mostly in the non-coding regions that may not be suitable accounting for the differences of functional trait values. Although single nucleotide polymorphism is the molecular marker advocated for use in crop breeding programs (Dillon *et al.,* 2007), this type of marker is expensive and required high resource laboratories. Nonetheless, any user could download from the database of the expressed sequence tags (ESTs) and look for SSRs within the ESTs using bioinformatics tools (Thiel *et al.,* 2003; Varshney *et al.,* 2005a). Mining SSRs from ESTs has been popular for SSR development in crop species, including sweet potato (Hu *et al.,* 2004a; Da Silva and Solis-Gracia, 2006; Aggarwal *et al.,* 2007; Chapman *et al.,* 2008; Ercan *et al.,* 2010; Qiu *et al.,* 2010; Wang *et al.,* 2011; Yu *et al.,* 2011; Gupta *et al.,* 2016; Kim *et al.*, 2016; Elshafei *et al.,* 2019; Isabel *et al.,* 2019). A putative function can be deduced for the majority of molecular markers developed from ESTs and they are known as 'functional markers' that provide an added value

(Andersen and Lübberstedt, 2003; Varshney, 2005b). EST-SSRs may be effective in assaying functional genetic variation in the germplasm collection (Kota *et al.,* 2001; Eujayl *et al.,* 2002). EST-SSRs can be useful in distinguishing varieties that differ in functional trait values and could be beneficial for the subsequent breeding program.

The breeding of sweet potato (*Ipomoea batatas* (L.) Lam.) is usually performed through controlled hybridization or polycross breeding strategies. Sweet potato polycross breeding is a method that aims to avoid the time-consuming pre-screening of cross-compatibility of sweet potato varieties, which usually selects 20–30 parental lines in an isolated plot and subjected to open pollination by insects and mass selection (Jones, 1965). Sweet potato breeding via polycross in Taiwan may not be as effective as expected in gaining genotypic combinations from introduced outstanding lines because studies have shown that polycross varieties bred in Taiwan have a closer genetic relationship with Taiwanese landrace varieties than with varieties bred in other geographic regions (Hwang *et al.,* 2002). However, only a small number of varieties were included in the investigation in these two studies (Tseng *et al.,* 2002).

Parental lines selected can have great influence on the efficiency of either controlled hybridization or polycross sweet potato breeding program. To achieve breeding tasks in controlled hybridization and polycross can be difficult because of the self- and cross-incompatibility of sweet potato. Moreover, the parental lines selected in the polycross breeding may not be objective but often subjective. Breeding efficiency could be limited due in part to the overlapping sets of elite varieties used in polycross breeding. Although there may be limits in breeding efficiency, specific genotypic combinations of various alleles of functional genes are attainable using the existing elite varieties produced by the past breeding efforts. By screening a large set of varieties, information of parental lines with agronomic trait values linked to differential sets of EST-SSR functional markers can be generated. In this study, EST-SSR variation of 89 cultivated sweet potato varieties (derived from controlled hybridization, polycross, and landrace varieties) together with six wild sweet potato (*I. trifida*) samples was quantified. Significant genetic structuring and heterogeneous allele frequency distributions of different breeding populations are expected if parental lines selected diverse greatly in their genetic compositions. However, only a minor portion of EST-SSR loci is likely to be heterogeneous in their allele frequency distributions across breeding populations, resulting in shallow genetic structuring because partly overlapping parental lines are generally adopted in different sweet potato breeding programs. Additionally, variation of a portion of alleles of different loci, representing different functional gene groups, may play crucial roles in contributing to the delimitation of varieties between the three breeding populations, and could be useful in the future breeding programs. Our main goals of this study were to (1) assess if there is a difference in the levels of genetic diversity measures among and between breeding populations, including hybrid, landrace, and polycross of 89 cultivated sweet potato varieties using 23 EST-SSRs, (2) estimate the level of the proportion of shared alleles and genetic differentiation between varieties of three breeding populations, and (3) test for the heterogeneous allele frequency distributions between different breeding populations and alleles. Results in this study demonstrate the limits of sweet potato breeding populations and also the potential of applying EST-SSRs in the future breeding practices of sweet potato.

MATERIALS AND METHODS

Plant materials and genotyping

The sweet potato varieties either from polycross or controlled hybridization breeding program and landrace varieties as well from various geographic regions were obtained from the Chiayi Agricultural Experimental Branch of Taiwan Agricultural Research Institute with permission (Supplemental Table S1 and Supplemental Fig. S1). Six wild sweet potato (*I. trifida*) samples were also included in this study. All plant materials were collected complying with relevant institutional, national, and international guidelines and legislation. Silica gel-dried, ground-up plant materials (either fresh, healthy leaves or tissue culture materials) were used in total DNA extraction based on a cetyltrimethyl ammonium bromide procedure (Doyle and Doyle, 1987). Ethanol precipitated DNA was washed with 70% ethanol and dissolved in 200 µl TE buffer (pH 8.0). DNA concentration was quantified using a NanoDrop spectrophotometer (NanoDrop Technology, Wilmington, DE, USA). SSRs were identified based on the *Ipomoea* EST sequences downloaded from the NCBI database using SSRIT (http://www. gramene.org/db/markers/ssrtool). We designed SSR forward and reverse primers using Primer 3 (http://frodo. wi.mit.edu/primer3/ input.htm). In total, 31 EST-SSR primer pairs were obtained, including 22 obtained as stated above and 9 EST-SSR primer pairs designed by Hu *et al*. (2004a), and used in genotyping (Supplemental Table S2). The forward primer of each primer pair was 5'-end-labeled with a 6-carboxyfluorescein, hexachloro-fluorescein, or tetramethyl-6-carboxyrhodamine fluorescent tag. Polymerase chain reaction (PCR) was performed in a 10-µl reaction volume containing 20 ng template DNA, 75 mM Tris-HCl (pH 8.8), 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween-20, 2.5 mM $MgCl_2$, 0.2 mM deoxyribonucleotide triphosphate mix, 75 nM of each primer, 0.8 mg bovine serum albumin, and 0.5 U of Simpler Red Taq DNA polymerase (Thermo Scientific, Waltham, MA, USA).

The condition of PCR was performed as follows: 94 °C for 10 min, 38 cycles of 30s at 94 °C, 30s at the optimal annealing temperature (Supplemental Table S2), 30s at 72 °C, and followed by a final extension of 5 min at 72 °C. PCR products were then separated on a MegaBACE 1000 96-capillary DNA sequencer (Amersham Biosciences, Piscataway, NJ, USA) and sizes of alleles were scored using Genetic Profiler Analyzer v.2.2 (Amersham).

Genetic diversity

The frequency of null allele was estimated with the null function of R package genepop (Rousset, 2008) in the R environment (https://www.R-project.org/R Foundation for Statistical Computing, Vienna, Austria, 2020). Eight EST-SSR primer pairs were found to have significant effect on genotyping errors because of apparent null allele frequencies, and 23 EST-SSR loci remained were used in the following analyses (Supplemental Table S3). Within each breeding population at each locus, potential deviations from HWE and LD between each pair of loci were tested by a modified Fisher's exact test (Guo and Thompson, 1992) based on 10^6 Markov chain iterations, respectively, using the test_HW and test_LD functions of R package genepop. Bonferroni correction was applied for multiple comparisons using the p.adjust function of R stats package. Genetic diversity measures, including A_L, A_R, H_O, and H_E, were analyzed, respectively, using the nb.alleles, allelic.richness, and basic.stats function of R hierfstat package (Goudet and Jombart, 2020). F_{IS} was calculated using the boot.ppfis function of R hierfstat package with 999 bootstrap resampling and *p*-values calculated based on 95% confidence intervals (CIs). Proportions of shared alleles between breeding populations and between breeding populations and wild diploid *I. trifida* were calculated with the pairwise.propShared function of R package PopGenReport (Adamack and Gruber, 2014). Multilocus LD index *r*D (Agapow and Burt, 2001) is an index modified from association index I_A (Brown *et al.,* 1980) to avoid dependence on number of loci. We calculated these two indices using the ia function of R poppr package (Kamvar *et al.,* 2015) and statistical significance of their non-zero values, compared with a null distribution, was tested with 999 permutations. We applied LMM using the lmer function of R package lme4 (Bates *et al.,* 2015) to assess whether genetic diversity measures (A_L, A_R, and H_E) differ significantly between breeding populations. Coefficient in LMM analysis was estimated based on reduced maximum likelihood estimation method. Breeding population and locus was used as a fixed and a random variable, respectively, in LMMs. Overall difference was tested using the Anova function of R package car (Fox and Weisberg, 2011) based on the type-II Wald χ^2-test, and Tukey's post-hoc pairwise comparisons were performed using the emmeans function of R package emmeans (Lenth, 2020).

Genetic relationship, differentiation, and clustering

Genetic delimitation among individual varieties was assessed using an unrooted neighbor-joining (NJ) tree, landscape and ecological association (LEA), and discriminant analysis of principal components (DAPC). The pairwise Nei's genetic distances (Nei, 1978) were calculated using the nei.dist function of R poppr package and used in generating an unrooted NJ tree using the nj function of R ape package (Paradis and Schliep, 2019). The BSV was calculated using the aboot function of R package poppr with 1,000 bootstrap resampling. Missing values were replaced with the mean of the corresponding allele in each population for the calculation of genetic distance matrix. Pairwise population Weir-Cockerham F_{ST} was calculated using the popStructTest function of R package strataG (Archer *et al.,* 2017) and tested the significance (999 permutations). The total dataset and the dataset included only varieties of three breeding populations bred in Taiwan were used for AMOVA using the poppr.amova function of R package poppr and significance tested with the randtest function of R package ade4. (Dray and Dufour, 2007). Genetic homogeneous group of individual varieties were assessed using sNMF algorithm (Frichot and François, 2015) and DAPC (Jombart and Ahmed, 2011). In these analyses, individual assignments with K = 1–10 based on least-squares optimization using the snmf function of R LEA package (Frichot and François, 2015). The regularization parameter, iterations, and repetitions in snmf were set to 100, 200, and 10, respectively, in snmf, and other arguments set to defaults. The find.clusters and dapc functions of R adegenet package (Jombart and Ahmed, 2011) were used in DAPC analysis. The mean CE and BIC were, respectively, used to determine the best K in LEA and DAPC (Supplemental Figures S2 and S3). The optimal number of principal components retained for discriminant analysis in DAPC was determined using the a-score function of R adegenet package.

Linear discriminants 1 and 2 and disparity in allele frequency distribution

To gain insight into the underlying allelic variation that may explain the differences between varieties of breeding populations including hybrid, landrace, and polycross, variance contributions of alleles to linear discriminants 1 and 2 of DAPC clustering were extracted and used to calculate the Z scores. The alleles with Z scores greater than 2.58 and smaller than -2.58 standard deviation (outside of the 99% CIs) were identified as alleles that contributed significantly to the first two linear discriminants of DAPC. Pairwise comparison of allele distributions of the 23 EST-SSR loci across the four breeding populations and wild sweet potato were assessed using the ks.test function of R stats package.

Table 1. Genetic parameters in three sweet potato breeding populations and wild diploid species investigated in this study based on 23 EST-SSR loci.

Genetic background	*N*	A_T	N_P	A_L (SD)	A_R (SD)	H_E (SD)	H_O (SD)	F_{IS} (P)	I_A (P)	r_D (P)
Varieties of three breeding populations and wild sweet potato										
Hybrid	41	138	29	6.000 (1.977)	2.901 (0.505)	0.484 (0.126)	0.614 (0.216)	-0.268 (< 0.001)	0.396 (0.002)	0.018 (0.002)
Landrace	10	86	11	3.739 (1.685)	2.939 (0.986)	0.486 (0.208)	0.588 (0.325)	-0.197 (< 0.001)	0.131 (0.291)	0.007 (0.287)
Polycross	38	128	19	5.565 (1.701)	2.888 (0.468)	0.487 (0.124)	0.620 (0.239)	-0.271 (< 0.001)	0.038 (0.391)	0.002 (0.391)
Wild	6	70	8	3.043 (1.397)	2.870 (1.244)	0.471 (0.266)	0.474 (0.299)	-0.018 (0.842)	2.370 (0.001)	0.139 (0.001)
Varieties of three breeding populations bred in Taiwan										
Hybrid	17	102	40	4.435 (1.376)	3.031 (0.636)	0.478 (0.139)	0.612 (0.230)	-0.278 (< 0.001)	0.039 (0.397)	0.002 (0.397)
Landrace	7	79	16	3.435 (1.376)	3.245 (1.193)	0.496 (0.207)	0.616 (0.331)	-0.235 (< 0.001)	0.409 (0.107)	0.021 (0.105)
Polycross	13	89	20	3.870 (1.100)	3.052 (0.729)	0.486 (0.154)	0.611 (0.282)	-0.290 (< 0.001)	-0.146 (0.733)	-0.007(0.736)

A_L, mean number of alleles per locus; A_R, mean number of standardized allelic richness; A_T, total number of alleles per population; F_{IS}, inbreeding coefficient; H_E, expected heterozygosity; H_O, observed heterozygosity; I_A, index of association; N, sample size; N_P, number of private alleles; r_D, modified index of association.

RESULTS

Genetic diversity

Eight of the 31 EST-SSR loci examined possessed significant null allele frequencies so that they were removed from further use in this study (Supplemental Table S2). Significant departures from Hardy-Weinberg equilibrium (HWE) were detected for hybrid and polycross varieties ($p < 0.0001$), but not significant for landrace varieties and wild sweet potato samples ($p = 0.338$ and 0.676, respectively). Two locus pair comparisons (3710 vs. 2710 and 2100 vs. 3110) were found to have significant linkage disequilibrium (LD) ($p < 0.001$) after applying the Bonferroni correction.

Genetic variability, population structure, and genetic relationships in three sweet potato breeding populations and wild species were investigated based on 23 EST-SSR loci. In total, we observed 171 alleles over 23 loci quantified for the 95 samples used in this study (Table 1). There were 166 alleles detected when only varieties of the three types of breeding populations were assessed, whereas 136 alleles were discovered when 37 varieties bred in Taiwan were examined. Distributions of allele frequencies of the 23 EST-SSR loci in the three types of breeding populations and wild sweet potato were depicted in Supplementary Fig. S4. Kolmogorov-Smirnov (KS) test revealed that disparity in allele frequency distributions was only found in several comparisons with marginal p values including locus 1110 (polycross vs. wild, $p = 0.0337$), locus 3010 (landrace vs. hybrid, $p = 0.0337$), and locus 3210 (hybrid vs. wild, $p = 0.0366$) in 338 pairwise comparisons.

The number of alleles per locus for the retained 23 loci ranged from 2 to 9, from 1 to 8, from 3 to 9, and from 1 to 6, respectively, in hybrid, landrace, polycross, and *I. trifida* (Fig. 1). The mean number of alleles per locus (A_L) ranged between 3.04 (wild) and 6.00 (hybrid) (Table 1). Measures of genetic diversity, including standardized allelic richness (A_R), expected heterozygosity (H_E), and observed heterozygosity (*Ho*), ranged between 2.870 (wild) and 2.939 (landrace), between 0.471 (wild) and 0.487 (polycross), and between 0.474 (wild) and 0.620 (polycross), respectively (Table 1). No significant difference was found between A_R and H_E, using linear mixed effect model (LMM) analysis. Breeding population and EST-SSR locus was used as a fixed and a random variable, respectively. However, LMM revealed overall significant difference in A_L (Wald $\chi^2 = 73.98$, $p < 0.0001$). Pairwise comparisons of difference in A_L between breeding populations were also observed (hybrid vs. landrace, $p < 0.0001$; hybrid vs. wild, $p < 0.0001$; landrace vs. polycross, $p = 0.0002$; and polycross vs. wild, $p < 0.0001$) using LMM.

Inbreeding coefficients (F_{IS}) were all significantly negative for breeding populations of hybrid ($F_{IS} = -0.268$), landrace ($F_{IS} = -0.197$), and polycross ($F_{IS} = -0.271$) (Table 1). When only varieties bred in Taiwan were analyzed, significant negative F_{IS} values were also found. Significant non-zero values of multilocus LD measures, including modified index of association (rD) (Agapow and Burt, 2001) and association index (I_A) (Brown *et al.,* 1980), were found in hybrid and wild sweet potato (hybrid: $I_A = 0.396$, $p = 0.002$ and rD = 0.018, $p = 0.002$; wild: $I_A = 2.370$, $p = 0.001$ and rD = 0.139, $p = 0.001$). However, no significant multilocus LD was found when only varieties of the three types of breeding populations bred in Taiwan were used in the analysis (Table 1).

Genetic differentiation

A population differentiation analysis was performed to analyze the genetic variations among and within groups, as revealed by the population structure, and measures of genetic differentiation were calculated for all pairwise combinations of populations, and statistical significance was assessed using matrix randomization. F_{ST} is based on the infinite alleles model of mutation, its suitability for the analysis of microsatellite variation depends on the spatial and/or temporal scale of divergence under consideration. Slatkin (1993) derived expressions for inbreeding coefficients in terms of allelic

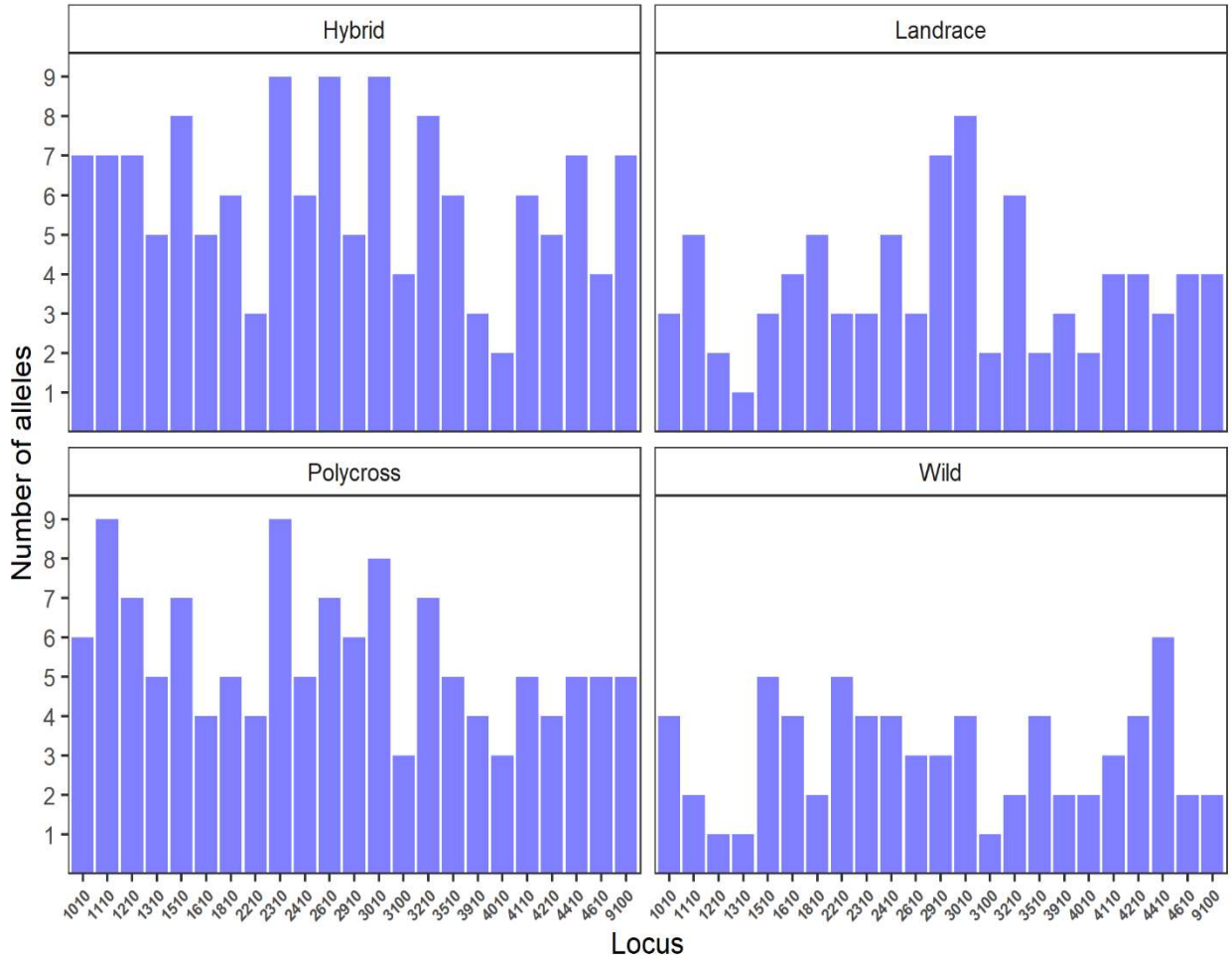

Fig. 1. The number of alleles per locus of the 23 EST-SSR loci examined in the three types of sweet potato breeding populations and wild sweet potato (*Ipomoea trifida*).

genealogies and demonstrated that F_{ST} measures the difference in within and between-population coalescence times scaled by the average coalescence time. Significant level of genetic differentiation was found when comparing wild sweet potato with the three types of breeding populations based on pairwise F_{ST} and analysis of molecular variance (AMOVA) (F_{ST} ranged between 0.0995 and 0.1077, all $p = 0.001$, Table 2; $\Phi_{ST} = 0.243$, $p = 0.0001$, Table 3). However, essentially no genetic differentiation was observed when compared between the three types of breeding populations based on the total samples and varieties bred in Taiwan (average F_{ST} = 0.0025 and 0.0003, respectively, Table 2; $\Phi_{ST} = 0.005$ and 0.000, respectively, Table 3). The low levels of genetic differentiation between breeding populations were also reflected in the high proportions of alleles shared among the three types of breeding populations (total sample: ranged between 0.838 and 0.912; Taiwan sample: ranged between 0.823 and 0.860) (Table 2). However, proportion of shared alleles reduced by 23% on average between wild sweet potato and the three types of breeding populations.

Table 2. Proportion of shared alleles between wild species and varieties of three breeding populations (lower diagonal) and pairwise Weir-Cockerham F_{ST} (upper diagonal).

	Hybrid	Landrace	Polycross	Wild
Varieties of three breeding populations and wild sweet potato				
Hybrid	-	0.002	-0.000	0.1077*
Landrace	0.842	-	0.0054	0.1021*
Polycross	0.912	0.838	-	0.0995*
Wild	0.667	0.654	0.676	-
Varieties of three breeding populations bred in Taiwan				
Taiwan				
Hybrid	-	0.0008	-0.0017	
Landrace	0.823	-	-0.0049	
Polycross	0.863	0.830	-	

Significance of pairwise Weir-Cockerham F_{ST} was tested with 999 permutations, *$p = 0.001$.

Genetic relationship and clustering

New powerful statistical tools and the reduced cost of DNA sequencing make it more feasible to search for or reconfirm the effect of artificially selected genes on past evolution and continued future improvement of sweet potatoes. Nucleotide diversity in sweet potato reflects the history of human selection and migration, combined with

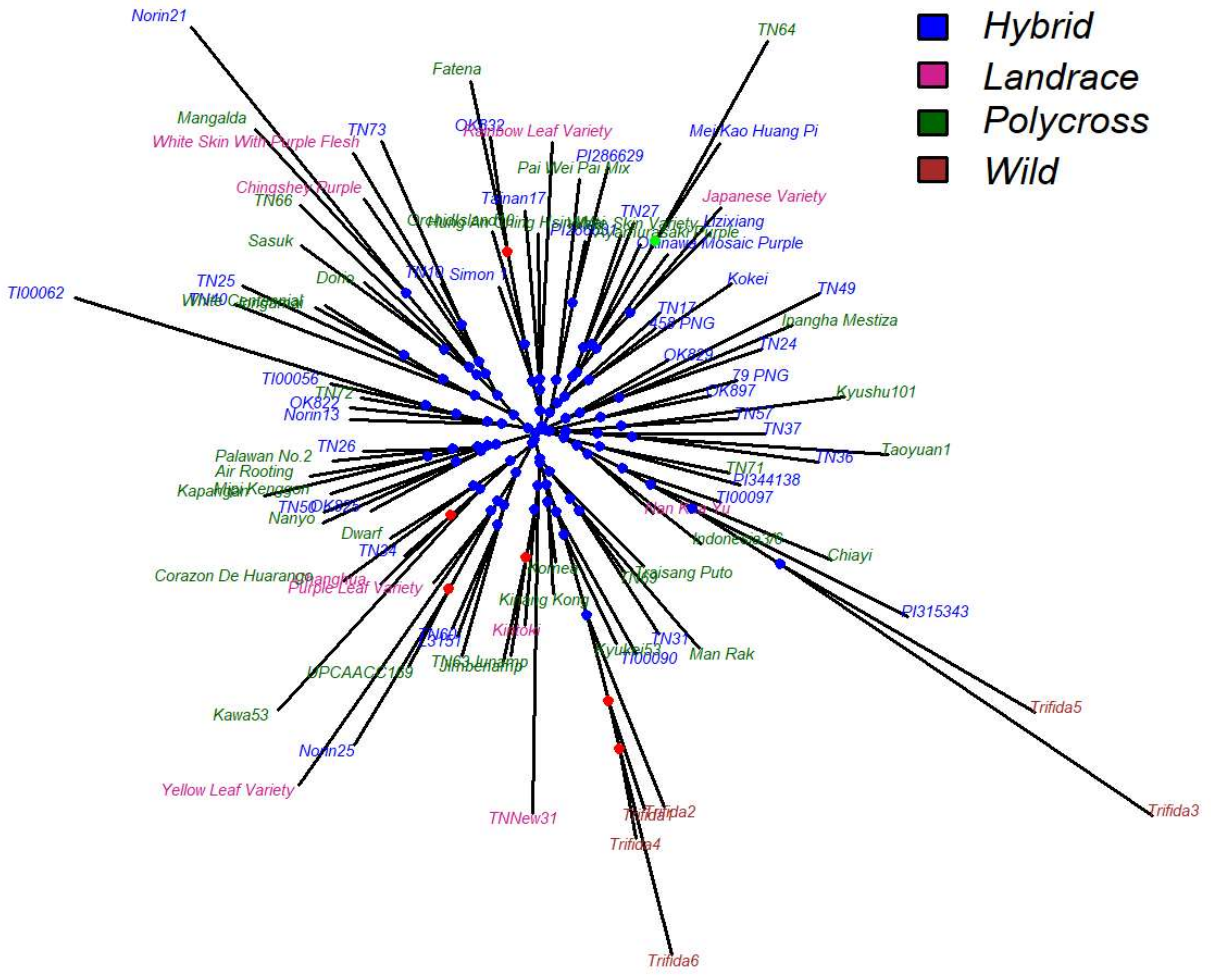

Fig. 2. The unrooted neighbor-joining tree of 95 samples of sweet potato breeding populations and wild sweet potato (*Ipomoea trifida*). Tip labels for individual sample are colored: hybrid (blue), landrace (violet red), polycross (dark green), and *Ipomoea trifida* (brown). For each node, branch bootstrap support values greater than 70%, between 50% and 70%, and smaller than 50% were coded with green, red, and blue circles, respectively.

Table 3. Summary of hierarchical level of genetic differentiation of sweet potato varieties analyzed using analysis of molecular variance (AMOVA).

Source of variation	Df	Sum of square	Percent variation	Φ Statistics (p value)
Between three breeding populations and wild sweet potato				
Between groups	1	19.432	24.32	Φ_{ST} = 0.243 (0.0001)
Within groups	93	391.864	75.68	
Total	94	411.297		
Between three breeding populations grouped by different geographic districts				
Between breeding programs	2	9.282	0.478	Φ_{ST} = 0.005 (0.178)
Within breeding programs	86	354.281	99.522	
Total	88	363.543	100	
Between three breeding populations bred in Taiwan				
Between breeding programs	2	7.983	-0.202	Φ_{ST} = 0.000 (0.547)
Within breeding programs	34	138.971	100.20	
Total	36	146.954	100	

high level of recombination, and out-breeding characteristic of this species (Buckler and Thornsberry, 2002). Phylogenetic analysis based on nucleotide diversity of a genome is a promising strategy for furthering the understanding of sweet potato plant's evolution at the molecular level. The unrooted neighbor-joining (NJ) tree revealed a relationship of intermingling between varieties of different breeding populations with mostly low branch bootstrap support values (BSV < 50%, Fig. 2). Additionally, two separate genetic clades were observed: one clade showed close relationship of breeding varieties with wild sweet potato samples originated from Guatemala and Colombia, and the second clade displayed a close relationship of breeding varieties with wild sweet potato samples originated from Guatemala, Nicaragua, and Venezuela. The genetic homogeneous group displayed using the sparse nonnegative matrix factorization (sNMF) algorithm implemented in the R LEA package (Frichot and François,

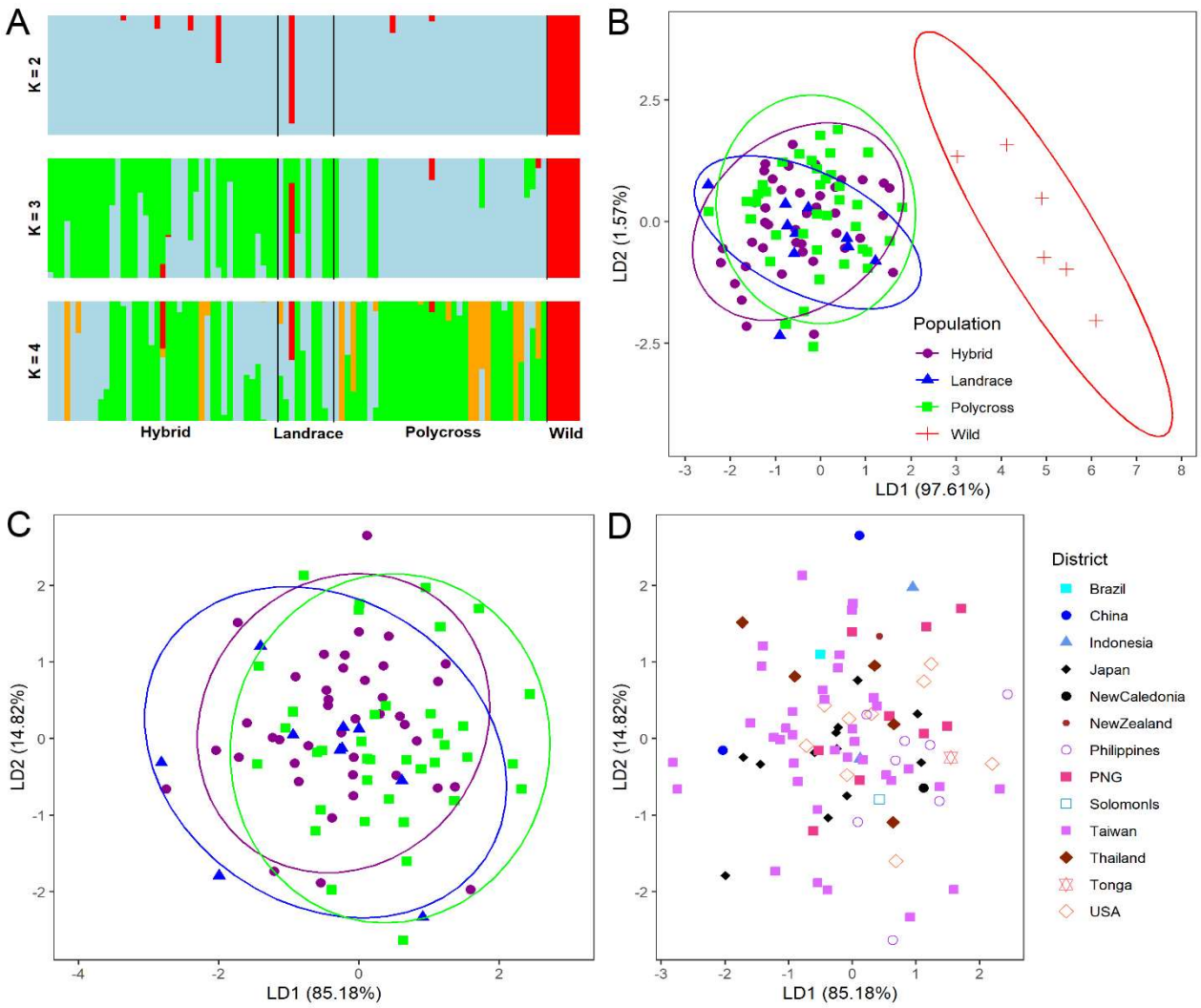

Fig. 3. Analysis of genetic homogeneous groups of 89 varieties of sweet potato derived from three breeding populations (hybrid, landrace, and polycross) and six wild sweet potato (*Ipomoea trifida*) samples based on the 23 EST-SSR loci using LEA **(A)**, DAPC of the total sample **(B)**, DAPC excluding *I. trifida* samples **(C)**, and DAPC excluding *I. trifida* samples with geographic districts coded with different colors **(D).** The clustering scenarios for K = 2–4 were displayed in LEA.

2015) showed a clear distinction of wild sweet potato samples from breeding varieties (Fig. 3A). In LEA, $K = 3$ had the lowest mean cross entropy (CE) value; however, the mean CE values at K = 2–4 may not differ significantly (Supplementary Fig. S2). The Bayesian information criterion (BIC) was minimized at $K = 2$ in discriminant analysis of principal component (DAPC) (Jombart *et al.,* 2010; Jombart and Ahmed, 2011) with and without wild sweet potato samples included in the analysis (Supplemental Fig. 3). The DAPC result showed a clear distinction of wild sweet potato samples from varieties of the three breeding populations (Fig. 3B). Additionally, when only varieties of hybrid, landrace, and polycross breeding populations were used in DAPC analysis, little distinguishing power was observed (Fig. 3C). Nonetheless, some varieties in the respective types of breeding populations were not enclosed within the distribution circles (95% confidence eclipse) of other breeding populations (Fig. 3C). When the DAPC result was displayed with geographic districts colored accordingly, it is clearly shown that varieties bred in Taiwan nearly encompass most of varieties bred in other geographic districts (Fig. 3D). The result of LEA grouping corresponded well with the DAPC clustering result when $K = 2$. Both LEA and DAPC showed that varieties of polycross cannot be distinguished from those of landrace and hybrid breeding populations when $K = 2$.

Individual EST-SSR alleles that contributed significantly to the delimitation of varieties between different breeding populations

In light of finding alleles that contributed significantly to the delimitation of varieties of three breeding populations, we extracted the variance of individual alleles

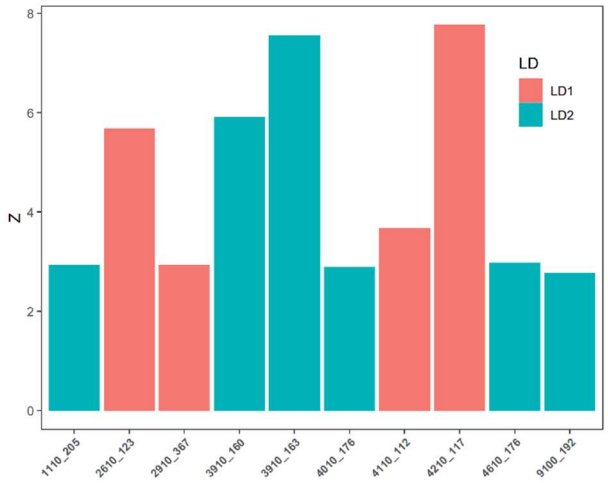

Fig. 4. Ten alleles of nine EST-SSR loci with significant variance contribution to linear discriminants 1 and 2 of DAPC clustering based on 89 varieties of sweet potato breeding populations. LD in the legend of this graph represents linear discriminant.

from the first two linear discriminants of DAPC (Fig. 3C) and calculated the Z scores. Ten alleles of nine loci were found to have significant loadings on linear discriminants 1 and 2 (Z scores greater than 2.58) in the delimitation of varieties of three different breeding populations (Fig. 4). The protein gene coding sequences of the nine loci with significant variance contributing to linear discriminants 1 and 2 were categorized into three functional features including abiotic stress tolerance: 4210 (acetyltransferase complex ard1 subunit), 4610 (proline-rich protein), and 9100 (chloroplast protein 12); transcriptional and translational regulation: 1110 (DNA binding protein), 2610 (DNA binding protein), 2910 (Zinc finger DNA binding protein), 4010 (AP2 domain-containing protein), and 4110 (N-acetyltransferase); and intercellular signalling: 3910 (Rhomboid family protein AtRBL2) (Supplemental Table S4).

DISCUSSION

Genetic diversity of crop species is expected to be lower compared to their wild ancestor (Louwaars, 2018). It is known that hexaploid sweet potato is descended from wild diploid species *I. trifida* via autoploidization and/or alloploidization (Kriegner *et al.,* 2003; Hu *et al.,* 2004b; Muñoz-Rodríguez *et al.,* 2018; Roullier *et al.,* 2013a). However, a recent study proposed a hypothesis of cultivated sweet potato originated from hybridization involving four wild diploid species including *I. trifida*, *I. tenuissima*, *I. littoralis*, and *I. nil* (Gao *et al.,* 2020). Lower level of EST-SSR diversity is expected compared with SSR diversity estimated based on anonymous microsatellites because EST-SSR is derived from coding sequences, which is under strong purifying selection. For example, EST-SSR diversity in the three types of sweet potato breeding populations in this study (average H_E = 0.486, Table 1) were lower compared with that of sweet potato varieties quantified using anonymous SSRs in Tanzania (average H_E = 0.69) (Ngailo *et al.,* 2016), Puerto Rico (average H_E =0.78) (Rodriguez-Bonilla *et al*., 2014), and the USA (average H_E = 0.70) (Rodriguez-Bonilla *et al.* 2014). Nonetheless, the levels of EST-SSR diversity of the three types of breeding populations in this study were comparable to the level of EST-SSR diversity such as sweet potato varieties in sub-Saharan African region (H_E = 0.548) (Quain *et al.,* 2018).

The mean number of alleles per locus was found to be significantly different between the three breeding populations and between the three breeding populations and *I. trifida* (higher in the three types of breeding populations than in *I. trifida*) (Table 1). Additionally, higher total number of alleles and number of private alleles were also found in the three types of breeding populations compared with that of *I. trifida* (Table 1). However, these results are likely caused by the higher number of samples examined in the three types of breeding populations than in *I. trifida*. Nonetheless, the comparable levels of A_R and H_E between *I. trifida* and the three types of breeding populations suggest high levels of genetic diversity in the three types of breeding populations.

Significant negative F_{IS} values observed (Table 1) indicate heterozygote excess in the three types of breeding populations for varieties of different geographic districts or varieties bred in Taiwan. This could be caused by the self-incompatibility of sweet potato which prevents inbreeding and allele fixation, facilitating new genotypic combinations of alleles to occur. Past intensive exchange of genetic materials of parental lines originated from different geographic areas can also be the cause of significant negative F_{IS}. No significant multilocus LD found particularly in the three types of breeding populations bred in Taiwan may have related to past intensive genetic exchange and led to LD decay (Ersoz *et al.,* 2007). The non-significant multilocus LD suggest small LD blocks and may be useful in association study in the future (Wadl *et al.,* 2018; Gemenet *et al.,* 2019). Past intensive hybridization is also reflected in the high degree of allele sharing and the largely homogeneous distribution of allele frequencies in pairwise comparisons analyzed using the KS test, resulting in essentially no genetic differentiation between the three types of breeding populations.

The most important finding from the result of NJ tree (Fig. 2) was that variation of the 23 EST-SSRs can be distinguished into two clades, revealing that varieties of hybrid, landrace, and polycross populations bred in different geographic districts originated from two sources of *I. trifida*. Additionally, the results of NJ tree, LEA (Fig. 3A), and DAPC (Fig. 3B, C) cannot effectively distinguishing varieties derived from different breeding populations probably due to the generally common

multiple progenitors of domestication and/or cultivation centers used for breeding in different geographic districts (Gao *et al.,* 2020; Ting *et al.,* 1953; Nishiyama *et al.,* 1975). The influence of partly overlapping parental lines causing no distinctive pattern of grouping are even more obvious when the DAPC result was depicted with breeding geographic districts coded with different colors (Fig. 3D). This type of grouping is consistent with other studies based on anonymous SSR, AFLP, and SNP (Elameen *et al.,* 2008; Yada *et al.,* 2010; Tumwegamire *et al.,* 2011; Gwandu *et al.,* 2012; Roullier *et al.,* 2013b; Wadl *et al.,* 2018). This phenomenon also suggest that sweet potato gene pool harbored large amount of genetic diversity because similar amount of genetic variation was found compared varieties of three breeding populations with that of *I. trifida* analyzed using LMM.

It is likely that cultivated sweet potato harbored high level of genetic diversity resulted from introgression of alleles between varieties derived from different breeding programs and from landrace varieties that possess variation maintained by natural selection in different geographic areas. However, the general practice in sweet potato breeding adopting a common set of outstanding lines produced in different geographic districts may limit the generation of elite lines with specific traits coping with future needs (Ceccarelli and Grando, 2020; Cortinovis *et al.,* 2020). Nonetheless, we found ten alleles of nine loci that had significant variance contributing to the delimitation of different breeding populations (Fig. 4). These alleles represent functional markers involved in abiotic stress resistance, transcriptional and translational regulation, and intercellular signaling that may be linked to specific agronomic traits (Ercan *et al.,* 2010; Andersen and Lübberstedt, 2003). By screening a large set of elite lines, information related to the contribution of individual EST-SSR functional alleles in distinguishing breeding populations. This highlights the potential of DAPC to go beyond merely serving as a method of group delimitation and will be important in enhancing breeding efficiency of sweet potato breeding. Furthermore, EST-SSR markers can also be used in the construction of sweet potato maps, and provides a basis for QTL mapping, marker-assisted selection and comparative genomics research of sweet potato (Kim *et al.*, 2016; Meng *et al.*, 2021).

In conclusions, varieties of hybrid, landrace, and polycross breeding populations had high genetic diversity based on 23 EST-SSRs. High proportion of allele sharing and low genetic differentiation between breeding populations revealed indicate limits and frontier of sweet potato breeding methods. The result of finding ten EST-SSR alleles contributing to the delimitation of different breeding populations demonstrates usefulness of EST-SSRs that could have linked to specific agronomic traits. EST-SSRs would be valuable for a public sector, such as Chiayi Agricultural Experimental Branch, Taiwan Agricultural Research Institute, to establish a pre-breeding project screening a large set of germplasm. The EST-SSR variation together with information of agronomic trait values can be used in gene-based trait-marker association study for future sweet potato breeding. The screening of germplasm should include wild sweet potato species not only *I. trifida*, but also other wild sweet potato, such as *I. tenuissima*, *I. littoralis*, and *I. nil*. The EST-SSR gene-based selection of various sweet potato varieties and wild diploid species can effectively enhance the efficiency of future allele introgression in the new variety development. The proposed research approach could help breeders to assess gene-based adaptive phenotypes.

ACKNOWLEDGMENTS

This work was funded partly by the Taiwan Ministry of Science and Technology under grant number NSC95-2313-B-034-001 to S.Y.H and the Council of Agriculture under grant number 98 AS-4.2.1.FD-Z2 to K.H.L. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

LITERATURE CITED

Adamack, A.T., Gruber, B. 2014 PopGenReport: simplifying basic population genetic analyses in R. Methods Ecol. Evol. **5(4)**: 384–387.

Agapow, P.M., Burt, A. 2001 Indices of multilocus linkage disequilibrium. Mol. Ecol. Notes **1(1-2)**: 101–102.

Aggarwal, R.K., Hendre, P.S., Varshney, R.K., Bhat, P.R., Krishnakumar, V., Singh, L. 2007 Identification, characterization, and utilization of EST-derived genic microsatellite makers for genome analyses of coffee and related species. Theor. Appl. Genet. **114(2)**: 359–372.

Allard, R.W. 1970 Population structure and sampling methods. In: Frankel, O.H., Bennett, E. (Eds.) Genetic Resources in Plants – Their Exploration and Conservation 97–107pp. Blackwell Scientific Publications, Oxford.

Andersen, J. R., Lübberstedt, T. 2003 Functional markers in plants. Trends Plant Sci. **8(11)**: 554–560.

Archer, F.I., Adams, P.E., Schneiders, B.B. 2017 strataG: An R package for manipulating, summarizing and analysing population genetic data. Mol. Ecol. Resour. **17(1)**: 5–11.

Bates, D., Maechler, M., Bolker, B., Walker, S. 2015 Fitting linear mixed-effects models using lme4. J. Stat. Soft. **67(1)**: 1–48.

Bhandari, H.R. Bhanu, A.N., Srivastava, K., Singh,·M.N., Shreya, Hemantaranjan, A. 2017 Assessment of genetic diversity in crop plants – an overview. Adv. Plants Agricul. Res. **7(3)**: 279–286.

Brown, A. H., Feldman, M. W., Nevo, E. 1980 Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics **96(2)**: 523–536.

Buckler, E.S., Thornsberry J.M. 2002 Plant molecular diversity and applications to genomics. Curr. Opin. Plant Biol. **5(2)**:107–111.

Ceccarelli, S., Grando, S., Maatougui, M., Micheal, M., Slash, M., Haghparast, R., Rahmanian, M., Taheri, A., Al-Yassin, A., Benbelkacem, A., Labdi, M., Mimoun, H.,

Nachit, M. 2010 Plant breeding and climate changes. J. Agricul. Sci. **148(6):** 627–637.

Ceccarelli, S., Grando, S. 2020 Evolutionary plant breeding as a response to the complexity of climate change. iScience **23(12)**:101815.

Chapman, M.A., Pashley, C.H., Wenzler, J., Hvala, J., Tang, S., Knapp, S,J., Burke, J,M. 2008 A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). Plant Cell **20(11)**: 2931–2945.

Cortinovis, G., Di Vittori, V., Bellucci, E., Bitocchi, E., Papa, R. 2020 Adaptation to novel environments during crop diversification. Curr. Opin. Plant Biol. **56:** 203–217.

Collard, B.C.Y., Mackill, D.J. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil. Trans. R. Soc. B **363(1491)**:557–572.

Da Silva, J.A., Solis-Gracia, N. 2006. Development of simple sequence repeat markers from genes related to stress resistance in sugarcane. Subtropical Plant Sci. **58:** 5–11.

Dillon, S. L., Shapter, F.M., Henry, R.J., Cordeiro, G., Izquierdo, L., Lee, L.S. 2007 Domestication to crop improvement: genetic resources for Sorghum and Saccharum (Andropogoneae). Ann. Bot. **100(5):** 975–989.

Doyle, J.J., Doyle, J.L. 1987 A rapid DNA isolation procedure for small quantities of fresh leaf material. Phytochem. Bull. **19**: 11–15.

Dray, S., Dufour, A.B. 2007 The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. **22(4):** 1–20.

Elameen, A., Fjellheim, S., Larsen, A., Rognli, O.A., Sundheim, L., Msolla, S., Masumba, E., Mtunda, K., Klemsdal, S.S. 2008 Analysis of genetic diversity in a sweet potato (*Ipomoea batatas* L.) germplasm collection from Tanzania as revealed by AFLP. Genet. Resour. Crop Evol. **55(3)**: 397–408.

Elshafei, A.A., Afiah, S.A.E.-A., Al-Doss, A.A., Ibrahim, E.I. 2019 Morphological variability and genetic diversity of wheat genotypes grown on saline soil and identification of new promising molecular markers associated with salinity tolerance. Int. J. Plant Inter. **14(1):** 564–571.

Ercan, S. Ertugrul, F., Aydin, Y., Akfirat, F.S., Hasancebi, S., Akan, K., Mert, Z., Bolat, N., Yorgancilar, O., Altinkut-Uncuoglu, A. 2010 An EST-SSR marker linked with yellow rust resistance in wheat. Biol. Planta **54(4):** 691–696.

Ersoz, E.S., Yu, J., Buckler, E.S. 2007 Applications of linkage disequilibrium and association mapping in crop plants. In: Varshney, R.K., Tuberosa, R. (eds) Genomics-Associated Crop Improvement Vol. 1: Genomics Approaches and Platforms 97–107pp. Springer, The Netherlands.

Eujayl, I., Sorrells, M.E., Baum, M., Wolters, P., Powell, W. 2002 Isolation of EST-derived microsatellite markers of phenotyping the A and B genomes of wheat. Theor. Appl. Genet. **104(2):** 399–407.

Fox, J., Weisberg, S. 2011 An R companion to applied regression, 2nd Edn.
http://socserv.socsci.mcmaster.ca/jfox/Books/Companion (Sage, 2011).

Frichot, E., François, O. 2015 LEA: an R package for landscape and ecological association studies. Methods Ecol. Evol. **6(8):** 925–929.

Gao, M., Soriano, S.F., Cao, Q., Yang X., Lu, G. 2020 Hexaploid sweetpotato (*Ipomoea batatas* (L.) Lam.) may not be a true type to either auto- or allopolyploid. PLoS One **15(3):** e0229624.

Gemenet, D.C., Kitavi, M.N., David, M., Ndege, D., Ssali, R.T., Swanckaert, J., Makunde, G., Yencho, G.C., Gruneberg, W., Carey, E., Mwanga, R.O., Andrade, M.I., Heck, S., Campos, H., Kumar, A. 2019 Development of diagnostic SNP markers for quality assurance and control in sweetpotato [I*pomoea batatas* (L.) Lam.] breeding Programs. PLoS One **15(4):** e0232173.

Gupta, D.S., Cheng, P., Sablok, G., Thavarajah, D., Thavarajah, P., Coyne, C.J., Kumar, S., Baum, M., McGee, R.J. 2016 Development of a panel of unigene-derived polymorphic EST-SSR markers in lentil using public database information. The Crop J. **4(5):** 425–433.

Guo, S.W., Thompson, E.A. 1992 Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics **48(2):** 361–372.

Goudet, J., Jombart, T. 2020 hierfstat: Estimation and Tests of Hierarchical F-Statistics. R package version 0.5-7. https://CRAN.R-project.org/package=hierfstat.

Gwandu, C., Tairo, F., Mneney, E., Kullaya, A. 2012 Characterization of Tanzanian elite sweet potato genotypes for sweet potato virus disease (SPVD) resistance and high dry matter content using simple sequence repeat (SSR) markers. Afr. J. Biotechnol. **11(40):** 9582–9590.

Hu, J.J., Nakatani, M., Mizuno, K., Fujimura, T. 2004a Development and characterization of microsatellite markers in sweetpotato. Breed. Sci**. 54(2):** 177–188.

Hu, J.J., Nakatani, M., Lalusin, A. G., Fujimura, T. 2004b New microsatellite markers developed from reported *Ipomoea trifida* sequences and their application to sweet potato and its related wild species. Sci. Horticul. **102(4):** 375–386.

Hwang, S.Y., Tseng, Y.T., Lo, H.F. 2002 Application of simple sequence repeats in determining the genetic relationships of varieties used in sweet potato polycross breeding in Taiwan. Sci. Horticul. **93(3-4):** 215–224.

Isabel, L. P., Park, J.R., Lee, G. S., Park, G.H., Kim, K.M. 2019 Development of EST-SSR markers and analysis of genetic relationship it's resources in hexaploidy oats. J. Crop Sci. Biotech. **22(3):** 243–251.

Jombart, T., Devillard, S., Balloux, F. 2010 Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. **11(1):** 94.

Jombart, T., Ahmed, I. 2011 adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics **27(21)**: 3070–3071.

Jones, A. 1965 A proposed breeding procedure for sweet potato. Crop Sci. **5(2):**191–192.

Kamvar, Z. N., Brooks, J. C., Grünwald, N. J. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front. Genet. **6:** 208.

Keneni, G., Bekele, E., Imtiaz, M., Dagne, K. 2012. Genetic vulnerability of modern crop varieties: causes, mechanism and remedies. Int. J. Plant Res. **2(3)**: 69–79.

Kim, J.H., Jo, W.A., Ham J.G., Chung, K., Kim, K.M. 2016 Characterization and development of EST-SSR markers in sweet potato (*Ipomoea batatas* (L.) Lam). 3 Biotech. **6(2):** 243.

Kota, R., Varshney, R.K., Thiel, T., Dehmer, K.J., Graner, A. 2001 Generation and comparison of EST-derived SSRs

and SNPs in barley (*Hordeum vulgare* L.). Hereditas **135(2-3)**: 145–151.

Kriegner, A., Cervantes, J.C., Burg, K., Mwanga, R.O.M., Zhang, D. 2003 A genetic linkage map of sweetpotato [*Ipomoea batatas* (L.) Lam.] based on AFLP markers. Mol. Breed. **11:** 169–185.

Lavergne, S., Molofsky, J. 2007 Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc. Natl. Acad. Sci. USA **104(10)**: 3883–3888.

Lenth, R. 2020 emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.8. https://CRAN.R-project.org/package=emmeans.

Louwaars, N.P. 2018 Plant breeding and diversity: A troubled relationship? Euphytica **214(7)**: 114.

Ma, J., Chen, J., Zhang, C., Yang, D., Tang, W., Gao, F., Xie, Y., Sun, H. 2022 Development and characterisation of SSR markers in the potato rot nematode *Ditylenchus destructor* Nematology **24(9):** 959–969.

Marshall, D. R., Brown, A. H. D. 1975 Optimum sampling strategies in genetic conservation. In: Frankel, O.H., Hawks, J.G. (eds) Crop Genetic Resources for Today and Tomorrow 53–80pp. Cambridge University Press, Cambridge.

Meng, Y., Zheng, C., Li, H., Li, A., Zhai, H., Wang, Q., He, S., Zhao, N., Zhang, H., Gao, S., Liu, Q. 2021 Development of a high-density SSR genetic linkage map in sweet potato. Crop J. **9(6):** 1367–1374

Muñoz-Rodríguez, P., Carruthers, T., Wood, J.R.I., Williams, B.R.M., Weitemier, K., Kronmiller, B., Ellis, D., Anglin, N.L., Longway, L., Harris, S.A., Rausher, M.D., Kelly, S., Liston, A., Scotland, R.W. 2018 Reconciling conflicting phylogenies in the origin of sweet potato and dispersal to Polynesia. Curr. Biol. **28(8)**: 1246–1256.

Nei, M. 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89(3)**: 583–590.

Ngailo, S., Shimelis, H., Sibiya, J., Amelework, B., Mtunda, K. 2016 Genetic diversity assessment of Tanzanian sweetpotato genotypes using simple sequence repeat markers. South Afr. J. Bot. **102**: 40–45.

Nishiyama, I., Miyazaki, T., Sakamoto, S. 1975 Evolutionary autoploid in the sweet potato (*Ipomoea batatas* (L.) Lam.) and its progenitors. Euphytica **24(1)**: 197–208.

Paradis, E., Schliep, K. 2019 ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics **35(3)**: 526–528.

Qiu, L., Yang, C., Tian, B., Yang, J.B., Liu, A. 2010 Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). BMC Plant Biol. **10(1)**: 278.

Quain, M.D., Adofo, K., Appiah-Kubi, D., Prempeh, R.N., Asafu-Agyei, J., Akomeah, B., Dapaah, H. 2018 Use of expressed sequence tags-derived simple sequence repeat (SSR) markers for population studies of released and elite sweet potato. Int. J. Genet. Mol. Bio. **10(2)**:14–25.

Slatkin, M. 1993 Isolation by distance in equilibrium and non-equilibrium populations. Evolution **47(1)**: 264–279.

Rodriguez-Bonilla, L., Cuevas, H.E., Montero-Rojas, M., Bird-Pico, F., Luciano-Rosario, D., Siritunga, D., Nelson, M.N. 2014 Assessment of genetic diversity of sweet potato in Puerto Rico. PLoS One **9(12)**: e116184.

Roullier, C., Duputié, A., Wennekes, P., Benoit, L., Bringas, V.M.F., Rossel, G., Tay, D., McKey, D., Lebot, V., Joly, S. 2013a Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). PLoS One **8(5)**: e62707.

Roullier, C., Kambouo, R., Paofa, J., McKey, D.B., Lebot, V. 2013b On the origin of sweet potato (*Ipomoea batatas* (L.) Lam.) genetic diversity in New Guinea, a secondary center of diversity. Heredity **110(6)**: 594–604.

Rousset, F. 2008 genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour. **8(1):** 103–106.

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

Ting, C.Y., Kehr, E.A. 1953 Meiotic studies in the sweet potato (*Ipomoea batatas* Lam.). J. Hered. **34:** 207–211.

Tseng, Y.T., Lo, H.F., Hwang, S.Y. 2002 Genotyping and assessment of genetic relationships in elite polycross breeding varieties of sweet potato in Taiwan based on SAMPL polymorphisms. Bot. Bull. Acad. Sin. **43(2):** 99–105.

Tumwegamire, S., Rubaihayo, P.R., LaBonte, D.R., Diaz, F., Kapinga, R., Mwanga, R.O.M., Gruneberg, W.J. 2011 Genetic diversity in white- and orange-fleshed sweet potato farmer varieties from east Africa evaluated by simple sequence repeat markers. Crop Sci. **51(3)**:1132–1142.

Varshney, R. K., Graner, A., Sorrells, M. E. 2005a Genetic microsatellite markers in plants: features and applications. Trends Biotechnol. **23(1)**: 48–55.

Varshney, R.K., Prasad, M., Kota, R., Sigmund, R., Valkoun, Börner, A., Valkoun, J., Scholz, U., Stein, N., Graner, A. 2005b Functional molecular markers in barley: Development and applications. Czech J. Genet. Plant Breed. **41(10)**: 128–133.

Wadl, P.A., Olukolu, B.A., Branham, S.E., Jarret, R.L., Yencho, G.C., Jackson, D.M. 2018 Genetic diversity and population structure of the USDA sweetpotato (*Ipomoea batatas*) germplasm collections using GBSpoly. Front. Plant Sci. **9:** 1166.

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

Yada, B., Tukamuhabwa, P., Wanjala, B., Kim, D.-J., Skilton, R.A., Alajo, A., Mwanga, R.O.M. 2010 Characterizing Ugandan sweet potato germplasm using fluorescent labeled simple sequence repeat markers. Hort. Sci. **45(2)**: 225–230.

Yu, Y., Yuan, D., Liang, S., Li, X., Wang, X., Lin, Z., Zhang, X. 2011 Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC1 population between *Gossypium hirsutum* and *G. barbadense*. BMC Genom. **12(1):** 15.

Zhu, Q., Zheng, X., Luo, J., Gaut, B. S., Ge, S. 2007. Multilocus analysis of nucleotide variation of *Oryza sativa* and its wild relatives: severe bottleneck during domestication of rice. Mol. Biol. Evol. **24(3):** 857–888.

Supplementary materials are available from Journal Website