Agr. Nat. Resour. 56 (2022) 455–462

AGRICULTURE AND NATURAL RESOURCES

Journal homepage: http://anres.kasetsart.org

Research article

Evaluation of diversity, population structure and core collection of Thailand *Luffa cylindrica* germplasm accessions

Grimar Abdiel Perez[a], Supasuta Karoojee[b,c], Pumipat Tongyoo[b,c], Julapark Chunwongse[b,c,d], Paweena Chuenwarin[b,c,d,*]

[a] *Program of Agricultural Sciences, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand*

[b] *Center for Agricultural Biotechnology, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand*

[c] *Center of Excellence on Agricultural Biotechnology: (AG-BIO/MHESI), Bangkok 10900, Thailand*

[d] *Department of Horticulture, Faculty of Agriculture at Kamphaeng Saen, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand*

Article Info

Article history:
Received 25 August 2021
Revised 26 November 2021
Accepted 8 March 2022
Available online 17 June 2022

Keywords:
Core collection,
Population structure,
Single nucleotide polymorphisms (SNPs),
Sponge gourd

Abstract

Importance of the work: A *Luffa cylindrica* (sponge gourd) germplasm collection based in Thailand is vital for breeding programs to improve fruit quality and enhance the biotic and abiotic adaptive traits or commercial traits.

Objectives: To explore the population structure and diversity in the germplasm collection.

Materials & Methods: Germplasm consisting of 254 accessions was evaluated using 3,442 single nucleotide polymorphisms based on DArTSeq genotyping-by-sequencing technology.

Results: The collection was divided into two (K = 2) genetic subpopulations based on STRUCTURE analysis. The genetic differentiation between these two clusters (F statistic [F_{ST}] 0.076) and variance based on the analysis of molecular variance analysis (7.6%) revealed a low differentiation between the two subpopulations. These results corresponded with principal coordinates analysis and neighbor-joining analysis in which the accessions were weakly clustered. The total collection displayed moderate diversity (expected heterozygosity = 0.310 and observed heterozygosity = 0.267) accompanied by low inbreeding (inbreeding coefficient = 0.138, and global inbreeding coefficient = 0.205). A range of simulated core collections was generated for practical field evaluation and to reduce the cost of seed maintenance in the gene bank, from which 20% of the total accessions were considered ideal.

Main finding: The moderate-to-low diversity germplasm within the population will be useful for improving sponge gourd in Thailand. Increasing the diversity of sponge gourd germplasm in Thailand should be considered.

* Corresponding author.
E-mail address: paweena.chue@ku.th (P. Chuenwarin)

https://doi.org/10.34044/j.anres.2022.56.3.02

Introduction

Luffa cylindrica (L.) Roem (sponge gourd) is an important cucurbit plant in tropical and subtropical regions with its center of origin and diversification in India (Filipowicz et al., 2014). Young fruits are edible for fresh consumption and are rich in nutrients, phenolic compounds, flavonoids and ascorbic acid or they can be used for their medicinal aspects (Azeez et al., 2013). Smooth luffa is well known in products made from matured, dried fruits having strong fibers and high porosity that are ideal for industrial commodities, such as insulation, filters and packing materials (Vianna et al., 2012). Given the importance of smooth luffa in Asia, this can be a crucial economic crop in Thailand for various applications.

Sponge gourd is one of the well-known agricultural products in Thailand (National Bureau of Agricultural Commodity and Food Standards, 2019) and can be an important economic agricultural product as an exported food or other products. For this reason, plant germplasm collections are essential to guarantee food security through plant breeding programs by preserving genetic resources. However, the increasing size of germplasm collections leads to high maintenance costs and management problems because of the need for proper conservation, regeneration, duplication and evaluation of the germplasm.

Establishing a core collection provides efficiency in germplasm management and better access to plant germplasm resources by selecting subsets that best represent the genetic diversity level of the larger collection. Core collections have been established using various methods such as simple-sequence repeat (SSR) markers in lima bean and chili (Gomes et al., 2020; Mongkolporn et al., 2015) and for phenotyping and geographical origin in sponge gourd (Mavi et al., 2018). Geographical passport data and phenotypic data can be informative but are also susceptible to environmental effects from both biotic and abiotic factors. Therefore, molecular markers provide rapid and reliable data unaffected by environmental factors or the growth stage of the plant and provide a stable source of information for assessing genetic diversity (Kordrostami and Rahimi, 2015). A major disadvantage of SSR markers is the laborious input required and low genome coverage.

Genotyping by sequencing (GBS) is a powerful tool for plant improvement. Advances in next-generation sequencing have provided rapid DNA sequencing at a lower cost while generating large numbers of single nucleotide polymorphism (SNP) markers. Such is the case with the Diversity Arrays Technology Pty Ltd. methodology (DArT) markers generated through a complex reduction method based on polymerase chain reaction (PCR) fragmentation by *Pst*I-*Mse*I and SNP detection using GBS technology (Valdisser et al., 2017). The advantages of a large dataset of SNPs include more comprehensive genome profiling for studying genetic diversity due to their broader genome coverage, genetic linkage, molecular marker discovery and genotyping (Beissinger et al., 2013).

The present study aimed to assess the genetic diversity of *L. cylindrica* germplasm using DArTseq-derived SNPs as the basis for *L. cylindrica* crop improvement and to develop core collections for practical seed stock maintenance and morphological trait evaluation under field conditions.

Materials and Methods

Plant materials

In total, 254 *L. cylindrica* (L.) Rox were used, comprised of accessions from Bangladesh (1), Cambodia (1), China (3), Japan (1), Laos (46), the Philippines (5), Thailand (174) and Vietnam (23), conserved by the Tropical Vegetable Research Center (TVRC), Kasetsart University, Kamphaeng Saen Campus, Thailand and the World Vegetable Center (AVRDC), Taiwan (Table S1).

DNA extraction

Genomic DNA was extracted from 100 mg of pooled young leaves of seedlings age 2 wk from 20 plants per accession using a modified cetyltrimethylammonium bromide method (Doyle and Doyle, 1987). Precipitated DNA was resuspended in TE buffer (10 mM Tris-HCl; 1 mM ethylenediaminetetraacetate pH 8.0) containing 2 µg/mL RNase. DNA quality was evaluated using electrophoresis with a 1% agarose gel run in Tris-borate EDTA buffer at 100 V and was quantified using a NanoDrop 2000c spectrophotometer V 1.6.0. The DNA concentration was adjusted to 50 ng/µL for GBS analysis.

Genotyping of L. cylindrica accessions using DArTseq

The genomic DNA samples were sent to Diversity Arrays Technology Pty Ltd. (Canberra, ACT, Australia) for whole-genome profiling using DArTseq genotype-

based sequencing (Von Mark et al., 2013). Briefly, DNA was digested using *Pst*I-*Mse*I restriction enzymes (Kilian et al., 2012). Then, the digested fragments were ligated to adapters and amplified using PCR (Raman et al., 2014), followed by sequencing using an Illumina Hiseq2000 sequencing system. The single-read sequencing was run for several predetermined cycles; the sequences generated were handled by DArT privately-owned pipelines at Diversity Arrays Technology Pty Ltd. (Canberra, ACT, Australia). Poor-quality sequences were filtered from the files in FASTQ format in the primary pipeline by applying rigorous selection criteria to the barcode region (Barilli et al., 2018). Identified sequences per barcode sample were used for marker calling. Then, these files were used in the second pipeline with DArT P/L's proprietary SNP calling algorithms (DArTsoftseq). The secondary pipeline generated various quality parameters, such as call rate, polymorphic information content and reproducibility.

Population structure and data analysis

The DArTseq-based SNPs were filtered using a call rate of 80% and a minor allele frequency greater than 5%. The population structure of the 254 *L. cylindrica* accessions was determined using the STRUCTURE version 2.3.4 software (Pritchard et al., 2000). Ten repeats were performed for each number of hypothetical clusters (*K*) ranging from 1 to 10. Accessions were assigned to subpopulations based on an admixture model, a burning period of 50,000 steps and 100,000 Markov Chain Monte Carlo chains. The STRUCTURE results were analyzed and visualized using the R package, POPHELPER version 2.3.0 (Francis, 2017). The ideal number of *K* was calculated using the Evanno method (Evanno et al., 2005). The phylogenetic dendrogram was constructed using the weighted neighbor-joining method (Bruno et al., 2000) with the DARwin software version 6.0.021 (Perrier and Jacquemoud-Collet, 2006), and it was visualized using FigTree version 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Principal coordinates analysis (PCoA), expected heterozygosity (H_E), observed heterozygosity (H_O), pairwise F_{ST}, Nei's coefficient of gene differentiation (G_{ST}), gene flow (Nm), the inbreeding coefficient (F_{IS}) and global inbreeding coefficient (F_{IT}) were calculated using ADEGENET version 2.1.3 (Jombart and Ahmed, 2011) and HIERFSTAT version 0.5-7 (Goudet, 2005) in the R statistical environment (R Core Team, 2020). Analysis of molecular variance (AMOVA) and the Shannon-Weiner diversity index (H) were calculated using the POPPR version 2.9.1 package (Kamvar et al., 2014). Jaccard's genetic similarity coefficient (GS) was calculated using the ADE4 version 1.7-16 package in R (Dray and Dufour, 2007).

Establishment of L. cylindrica core collection

Simulations for core collection sizes ranging from 20% to 100% in 5% increments were performed using the 3,442 DArTseq-based SNPs in the COREHUNTER version 3.2.1 package in R (De Beukelaer et al., 2018). The average accession-to-nearest-entry distance was used to generate diverse cores in which selected accessions expressed maximum dissimilarity. The modified Roger's distance was used to measure the relative distance between the chosen accessions of each core collection. The Shannon-Weiner diversity index, H_E, H_O, SNPs and percentage of accessions retained in Cluster 1 and Cluster 2 were used to evaluate the competency of each core collection.

Results and Discussion

Genetic structure of germplasm

The 254 *L. cylindrica* accessions could be divided into subpopulations based on STRUCTURE analysis using the 3,442 SNPs, as indicated by the highest value of ΔK, which indicated grouping into two clusters (*K*), as shown in Fig. 1A. Cluster 1 was composed of 162 (64%) accessions and Cluster 2 of 92 accessions (36%), as shown in Fig. 1B.

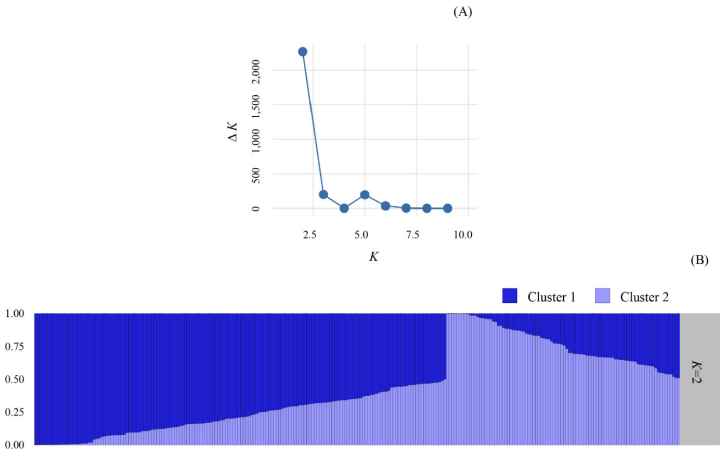

Fig. 1 Population structure of 254 *Luffa cylindrica* accessions: (A) *K* indicated by the highest ΔK; (B) proportion of individuals corresponding to each cluster

PCoA showed the relationship pattern between the two clusters (Fig. 2). The phylogenetic dendrograms displayed accessions almost evenly dispersed from each other, which added to the evenness in similarity across accessions in the *L. cylindrica* germplasm (Fig. 3). However, the dendrogram did not show a definite pattern with geographical origin except for accessions from Vietnam that were clustered together. The structure observed in a population reflects genetic variance amongst subpopulations that arises from collaborative events such as gene flow, genetic drift, and natural selection (Mitton, 2013). In terms of genetic differentiation, F_{ST} or G_{ST} values below 0.05 are defined as low, values in the range 0.05–0.15 as moderate, above 0.15 to 0.25 as high and greater than 0.25 as very high (Xu et al., 2019). The moderate F_{ST} (0.076) and low Nei's G_{ST} (0.037) values indicated that the differentiation between the two subpopulations was not high (Table 1). The substantial genetic similarity (GS = 65%) between Cluster 1 and Cluster 2 supported the low genetic differentiation in the 254 *L. cylindrica* germplasm accessions that was also supported by the low population differentiation statistics (ϕ = 0.08) between the two clusters (Table 2). AMOVA supported highly significant (p = 0.001) STRUCTURE clustering with a low variance between the two clusters (7.6%), as shown in Table 2. The AMOVA analysis indicated highly significant differences at all levels based on STRUCTURE clustering. The total population of 254 *L. cylindrica* accessions could be described as a panmictic population because most of the observed variance occurred within individuals (85.7%).

Table 1 Pairwise genetic differentiation between clusters identified based on STRUCTURE analysis

	F_{ST}	Neis G_{ST}	Nm	GS
STRUCTURE clustering	Cluster 2	Cluster 2	Cluster 2	Cluster 2
Cluster 1	0.076	0.037	3.039	65%

F_{ST} = fixation index; Neis G_{ST} = Nei's coefficient of gene differentiation; Nm = gene flow; GS = Jaccard's genetic similarity coefficient

Genetic diversity in L. cylindrica

The *L. cylindrica* germplasm had a high degree of gene flow between populations. The low genetic differentiation (F_{ST} =0.076) between Cluster 1 and Cluster 2 indicated that the populations were less isolated, with an increased probability of crossing between the populations. The F_{ST} of the *L. cylindrica* germplasm was lower when compared to the F_{ST} of high subpopulation numbers of *L. acutangula* (Perez et al., 2021). It revealed more admixture between the two subpopulations of *L. cylindrica* than those subpopulations in ridge gourd. The present results indicated low inbreeding coefficients in Cluster 1 (F_{IS} = 0.064) and Cluster 2 (F_{IS} = 0.185) and low global inbreeding (F_{IT} = 0.205) in the whole germplasm, indicating that a low degree of inbreeding had occurred in the sponge gourd germplasm collection (Table 3). An overall lower global inbreeding (F_{IT} = 0.057) and inbreeding coefficient (F_{IS} = -0.092) have been reported in the high number of subpopulations in ridge gourd by Perez et al. (2021), indicating a lower degree of inbreeding occurred in ridge gourd.

Fig. 2 Principal coordinate analysis of 254 *Luffa cylindrica* accessions based on DArTseq single nucleotide polymorphisms, where colors correspond to STRUCTURE clustering

Table 2 Analysis of molecular variance of 254 *Luffa cylindrica* accessions based on STRUCTURE clustering

ΔK = 2	Df	Mean Sq	Variance	%	Phi statistic (ϕ)	p-value
Between populations	1	17870.11	72.14	7.6	0.08	0.001
Between samples within populations	252	938.70	63.57	6.7	0.07	0.001
Within samples	254	811.55	811.55	85.7	0.14	0.001
Total	507	908.39	947.26	100		

ΔK = a method used for identification of the most probable number of populations; Df = degrees of freedom; Mean Sq = mean square

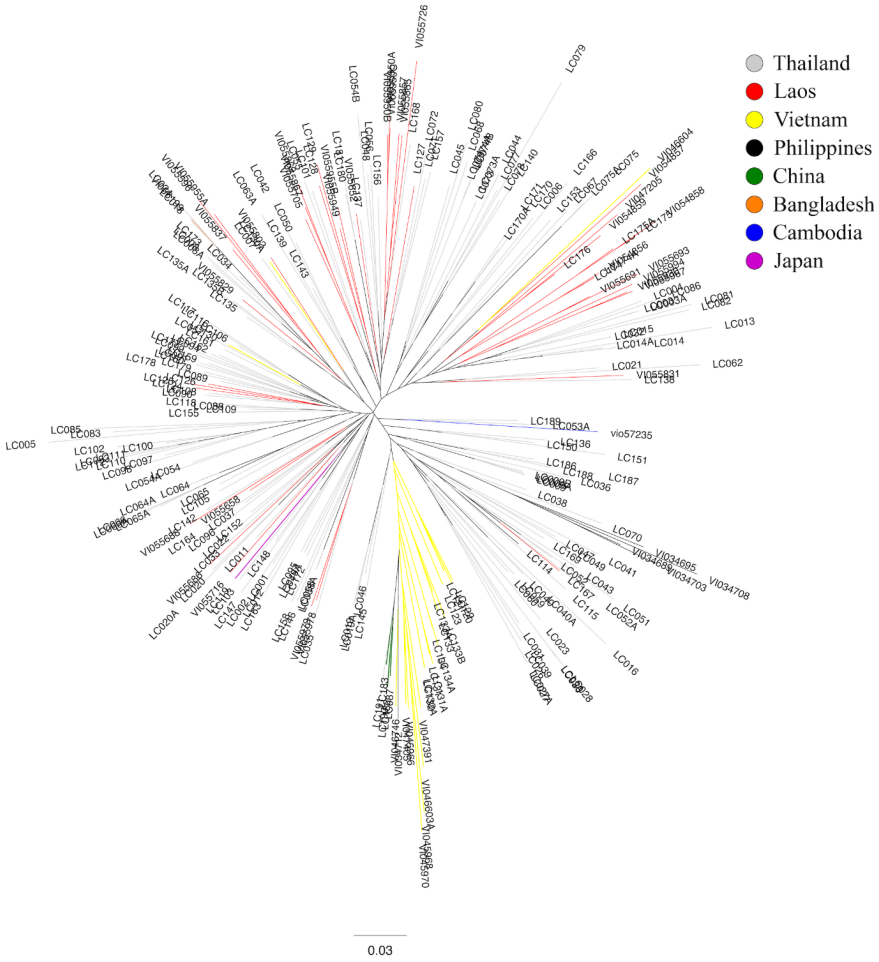

Fig. 3 Radial weighted neighbor-joining phylogenetic dendrogram of 254 *Luffa cylindrica* accessions showing colors of accessions according to country of origin

Table 3 Genetic diversity indices of 254 *Luffa cylindrica* accessions based on STRUCTURE clustering

	N	H	H_E	H_O	F_{IS}	F_{IT}
Cluster 1	162	5.088	0.311	0.291	0.064	-
Cluster 2	92	4.522	0.275	0.224	0.185	-
Total	254	5.537	0.310	0.267	0.138	0.205

N = number of accessions; H = Shannon-Weiner diversity index; H_E = expected heterozygosity, H_O = observed heterozygosity; F_{IS} = inbreeding coefficient; F_{IT} = global inbreeding coefficient

All alleles contributed evenly to the diversity in the entire population (H = 5.537), as shown in Table 3. The total H_E (0.310) and H_O (0.267) showed there was moderate genetic diversity in the whole population and corresponded with the low total F_{IS} and F_{IT} values (Table 3). The value of H_E was greater than for H_O, indicating that germplasm accessions were likely inbreeding subpopulations. However, the low inbreeding coefficients of Cluster 1 (F_{IS} =0.064) and Cluster 2 (F_{IS} = 0.185) may indicate a random mating system in the commonly outcrossing *L. cylindrica* (Misra et al., 2017).

The moderately low diversity may likely stem from a bottleneck in the sponge gourd germplasm collection that is mainly made up of accessions from Southeast Asia, especially from Thailand. A recent study in sponge gourd using ISSR and start codon targeted markers reported moderate genetic variation partially attributed to a sole place of domestication and the use of a few landraces for hybridization (Tyagi et al., 2020). Similar results were observed in a collection of Ethiopian cowpea landraces, which also showed considerable diversity compared to the collection used worldwide (Ketema

et al., 2020). Another possibility may be the outcrossing nature of *L. cylindrica,* which would ensure gene flow throughout the whole population. The gene flow value (Nm = 3.039) was greater than 1 in the entire *L. cylindrica* collection, indicating substantial geneflow (Wright, 1984). Such may be the case in the *L. cylindrica* germplasm, where there is more gene flow than artificial selection. Therefore, the exchange of genotypes across Thailand may occur moderately without selective agricultural pressure, in contrast to the highly structured commercial ridge gourd germplasm (Perez et al., 2021). Perhaps *L. cylindrica* has not been a major source of income for farmers. Therefore, widespread preservation of genotypes has occurred throughout Asia.

Establishment of core collections for practical use

A suitable core collection for taxonomist, geneticist and gene bank curators retains a uniform representation of the whole collection, minimizing the redundancy of the individuals selected and providing the smallest possible size to allow practical use. The simulated core collection at 20% of the total accessions represents adequate genetic diversity and marker quantities suitable for seed management and field evaluation in this germplasm (Table 4). The selected accessions in the 20% core collection is appropriate based on their evenness across the PCoA (Fig. S1). The core collection of 20% retained a high Shannon-Weiner diversity index value of 3.93, included 90% of total SNPs and maintained a corresponding ratio of accessions in Cluster 1 (64.7%) and Cluster 2 (35.4%) as observed in the whole collection (Table 4). A core collection capturing as much diversity and markers as possible is key for choosing appropriate collections for resource management or practical uses. Keeping 97% or more SNPs may be adequate but greatly depends on the purpose of the core collection (Ndjiondjop et al., 2017). A higher modified Roger's distance indicates a higher distance between selected individuals (Wright, 1984). The 20% core collection had a moderate modified Roger's distance value of 0.267. The H_E and H_O values of the 20% core collection captured substantial heterozygosity when compared to the whole germplasm collection (Table 4). Core collections lower than 35% exhibit a higher modified Roger's distance which inversely correlates with a the number of selected accessions (Wright, 1984). However, the addition of accessions offers a more representative core collection.

Table 4 Genetic parameters for sampled cores ranging from 5% to 100%

% Core	N	MR	SNPs	% SNPs	H	H_E	H_O	% Cluster 1	% Cluster 2
5	13	0.347	2,696	78.3	2.56	0.350	0.336	46.2	53.8
10	25	0.318	2,901	84.3	3.22	0.331	0.264	52.0	48.0
15	38	0.291	3,093	89.9	3.64	0.320	0.256	60.5	39.5
20	51	0.267	3,101	90.1	3.93	0.316	0.218	64.7	35.3
25	64	0.245	3,035	88.2	4.16	0.318	0.211	65.6	34.4
30	76	0.225	3,107	90.3	4.33	0.314	0.199	64.5	35.5
35	89	0.204	3,129	90.9	4.49	0.314	0.205	62.9	37.1
40	102	0.184	3,181	92.4	4.62	0.313	0.210	62.7	37.3
45	114	0.166	3,189	92.6	4.74	0.313	0.218	64.0	36.0
50	127	0.147	3,193	92.8	4.84	0.311	0.220	64.6	35.4
55	140	0.129	3,207	93.2	4.94	0.310	0.224	63.6	36.4
60	152	0.113	3,265	94.9	5.02	0.310	0.233	63.8	36.2
65	165	0.096	3,268	94.9	5.11	0.310	0.242	62.4	37.6
70	178	0.079	3,287	95.5	5.18	0.311	0.248	64.0	36.0
75	190	0.065	3,308	96.1	5.25	0.311	0.255	64.7	35.3
80	203	0.050	3,342	97.1	5.31	0.312	0.263	64.5	35.5
85	216	0.036	3,369	97.9	5.38	0.312	0.267	63.9	36.1
90	229	0.022	3,383	98.3	5.43	0.312	0.270	63.8	36.2
95	241	0.011	3,426	99.5	5.48	0.311	0.270	63.5	36.5
100	254	0.002	3442	100.0	5.54	0.310	0.267	63.8	36.2

N = number of accessions; MR = Modified Roger's Distance; SNPs = single nucleotide polymorphisms; H = Shannon-Weiner diversity index; H_E = expected heterozygosity, H_O = observed heterozygosity

Such an observation of a more representative core is supported by the Shannon-Weiner diversity index, H_E and H_O values of the different cores (Table 4). The H_E value remained relatively similar in the simulated cores, while the Shannon-Weiner diversity index increased as the number of accessions increased in the core collections (Gomes et al., 2020). Another important consideration for developing a core collection is capturing as many high-quality SNPs from the whole germplasm (Pereira et al., 2020). The number of SNPs decreased from 3,442 to 3,101 for the whole collection and 20% core collection, respectively. Therefore, keeping more than 90% of the SNPs is critical to increasing the chance of having markers spread throughout the *L. cylindrica* genome for subsequent work.

A sample taken from a whole collection can vary depending on the purpose, size of the collection, quality of phenotype or genotype characterization, evaluation of possible population structure from the whole collection and the sampling method used. In the present study, core collections of 20% and above 30% of the total accessions retained more than 90% of the total SNPs. These simulated core collections would be useful in other experimental designs requiring a different population size.

Conclusion

There was a moderate-to-low genetic diversity in the 254 sponge gourd germplasm accessions that could be divided into two subpopulations. This study also established a suitable core collection consisting of 20% of the whole germplasm for practical usage. The investigation should facilitate future work involving trait evaluation and selection for sponge gourd improvement in Thailand.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The present study was financially supported by the Kasetsart University Research and Development Institute (KURDI), Bangkok, Thailand and by Kasetsart University through the Graduate School Fellowship Program. This research was supported by the Center of Excellence on Agricultural Biotechnology, Office of the Permanent Secretary, Ministry of Higher Education, Science, Research and Innovation (AG-BIO/MHESI). Ms. Hathaiphat Vongthaiwan provided valuable guidance in using bioinformatic tools. Seed samples were provided by: The Tropical Vegetable Research Center (TVRC), Department of Horticulture, Faculty of Agriculture at Kampheang Saen, Kasetsart University, Kamphaeng Saen Campus, Thailand; and the World Vegetable Center, Taiwan. Mr. Akrachai Somakul (TVRC germplasm manager), Mr. Anucha Wongpraneekul and Ms. Waraporn Sinsathapornpong assisted with the seed and seedling management of sponge gourd.

References

Azeez, M.A., Bello, O.S., Adedeji, A.O. 2013. Traditional and medicinal uses of *Luffa cylindrica*: A review. J. Med. Plants Stud. 1: 102–111.

Barilli, E., Cobos, M.J., Carrillo, E., Kilian, A., Carling, J., Rubiales, D. 2018. A high-density integrated DArTseq SNP-based genetic map of *Pisum fulvum* and identification of QTLs controlling rust resistance. Front. Plant Sci. 9: 167. doi.org/10.3389/fpls.2018.00167

Beissinger, T.M., Hirsch, C.N., Sekhon, R.S., et al. 2013. Marker density and read depth for genotyping populations using genotyping-by-sequencing. Genetics 193: 1073–1081. doi.org/10.1534/genetics.112.147710

Bruno, W.J., Socci, N.D., Halpern, A.L. 2000. Weighted neighbor joining: A likelihood-based approach to distance-based phylogeny reconstruction. Mol. Biol. Evol. 17: 189–197. doi.org/10.1093/oxfordjournals.molbev.a026231

De Beukelaer, H., Davenport, G.F., Fack, V. 2018. Core Hunter 3: Flexible core subset selection. BMC Bioinform. 19: 203. doi.org/10.1186/s12859-018-2209-z

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

Dray, S., Dufour, A.B. 2007. The ade4 package: Implementing the duality diagram for ecologists. J. Stat. Softw. 22: 1–20. doi: 10.18637/jss.v022.i04

Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14: 2611–2620. doi.org/10.1111/j.1365-294X.2005.02553.x

Filipowicz, N., Schaefer, H., Renner, S.S. 2014. Revisiting *Luffa* (Cucurbitaceae) 25 years after C. Heiser: Species boundaries and application of names tested with plastid and nuclear DNA sequences. Syst. Bot. 39: 205–215.

Francis, R.M. 2017. Pophelper: An R package and web app to analyse and visualize population structure. Mol. Ecol. Resour. 17: 27–32. doi.org/10.1111/1755-0998.12509

Gomes, R.L.F., Costa, M.F., Alves-Pereira, A., et al. 2020. A lima bean core collection based on molecular markers. Sci. Agric. 77: e20180140. doi.org/10.1590/1678-992X-2018-0140

Goudet, J. 2005. Hierfstat, a package for R to compute and test hierarchical *F*-statistics. Mol. Ecol. Notes 5: 184–186. doi: 10.1111/j.1471-8278.2004.00828.x

Jombart, T., Ahmed, I. 2011. Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27: 3070–3071. doi.org/10.1093/bioinformatics/btr521

Kamvar, Z.N., Tabima, J.F., Grünwald, N.J. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281. doi.org/10.7717/peerj.281

Ketema, S., Tesfaye, B., Keneni, G., Amsalu Fenta, B., Assefa, E., Greliche, N., Machuka, E., Yao, N. 2020. DArTSeq SNP-based markers revealed high genetic diversity and structured population in Ethiopian cowpea [*Vigna unguiculata* (L.) Walp] germplasms. PLoS ONE 15: e0239122. doi.org/10.1371/journal.pone.0239122

Kilian, A., Wenzl, P., Huttner, E., et al. 2012. Diversity arrays technology: A generic genome profiling technology on open platforms. In: Pompanon, F., Bonin, A. (Eds.). Data Production and Analysis in Population Genomics. Springer. Totowa, NJ, USA, pp. 67–89.

Kordrostami, M., Rahimi, M. 2015. Molecular markers in plants: Concepts and applications. Genetics in the Third Millennium 13: 4024–4031.

Mavi, K., Gunduz, K., Yildirim, D., Uzunoglu, F. 2018. Seed characteristics diversity in sponge gourd (*Luffa aegyptiaca* Mill.) germplasms from Hatay region in Turkey. Bull. UASVM Hortic. 75: 40–46.

Misra, S., Srivastava, A.K., Verma, S., Pandey, S., Bargali, S.S., Rana, T.S., Nair, K.N. 2017. Phenetic and genetic diversity in Indian *Luffa* (Cucurbitaceae) inferred from morphometric, ISSR and DAMD markers. Genet. Resour. Crop Evol. 64: 995–1010. doi.org/10.1007/s10722-016-0420-1

Mitton, J. 2013. Gene flow. In: Maloy, S., Hughes, K. (Eds.). Brenner's Encyclopedia of Genetics (2nd ed.). Academic Press. San Diego, CA, USA, pp. 192–196.

Mongkolporn, O., Hanyong, S., Chunwongse, J., Wasee, S. 2015. Establishment of a core collection of chilli germplasm using microsatellite analysis. Plant Genet. Resour-C. 13: 104–110. doi.org/10.1017/S1479262114000768

National Bureau of Agricultural Commodity and Food Standards. 2019. Loofah. Thai Agricultural Standard TAS 1526-2019. Ministry of Agriculture and Cooperatives. https://www.acfs.go.th/files/files/commodity-standard/20191212114212_964500.pdf, 24 August 2021.

Ndjiondjop, M.-N., Semagn, K., Gouda, A.C., et al. 2017. Genetic variation and population structure of *Oryza glaberrima* and development of a mini-core collection using DArTseq. Front. Plant Sci. 8: 1748. doi.org/10.3389/fpls.2017.01748

Pereira, V.M., Filho, J.A.F., Leão, A.P., et al. 2020. American oil palm from Brazil: Genetic diversity, population structure, and core collection. Crop Sci. 60: 3212– 3227. doi: 10.1002/csc2.20276

Perez, G.A., Tongyoo, P., Chunwongse, J., de Jong, H., Wongpraneekul, A., Sinsathapornpong, W., Chuenwarin, P. 2021. Genetic diversity and population structure of ridge gourd (*Luffa acutangula*) accessions in a Thailand collection using SNP markers. Sci. Rep. 11: 15311. doi.org/10.1038/s41598-021-94802-4

Perrier, X., Jacquemoud-Collet, J.P. 2006. DARwin software (Version 6.0.021). Montpellier, France. https://darwin.cirad.fr/, 31 March 2021.

Pritchard, J.K., Stephens, M., Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959. doi.org/10.1093/genetics/155.2.945

R Core Team. 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/, 6 April 2021.

Raman, H., Raman, R., Kilian, A., et al. 2014. Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. PLoS ONE 9: e101673. doi.org/10.1371/journal.pone.0101673

Tyagi, R., Sharma, V., Sureja, A.K., Das Munshi, A., Arya, L., Saha, D., Verma, M. 2020. Genetic diversity and population structure detection in sponge gourd (*Luffa cylindrica*) using ISSR, SCoT and morphological markers. Physiol. Mol. Biol. Plants 26: 119–131. doi.org/10.1007/s12298-019-00723-y

Valdisser, P.A., Pereira, W.J., Almeida Filho, J.E., et al. 2017. In-depth genome characterization of a Brazilian common bean core collection using DArTseq high-density SNP genotyping. BMC Genom. 18: 423. doi.org/10.1186/s12864-017-3805-4

Vianna, M.R., de Melo, G.C., Neto, M.R. 2012. Wastewater treatment in trickling filters using *Luffa cyllindrica* as biofilm supporting medium. J. Urban Environ. Eng. 6: 57–66. doi: 10.4090/juee.2012.v6n2.057066

Von Mark, V.C., Kilian, A., Dierig, D.A. 2013. Development of DArT marker platforms and genetic diversity assessment of the U.S. Collection of the new oilseed crop lesquerella and related species. PLoS ONE 8: e64062. doi.org/10.1371/journal.pone.0064062

Wright, S. 1984. Evolution and the Genetics of Populations, Vol. 4: Variability Within and Among Natural Populations. University of Chicago Press. London, UK.

Xu, Y., Mai, J.W., Yu, B.I., Hu, H.X., Yuan, L., Jashenko, R., Ji, R. 2019. Study on the genetic differentiation of geographic populations of *Calliptamus italicus* (Orthoptera: Acrididae) in Sino-Kazakh border areas based on mitochondrial COI and COII genes. J. Econ. Entomol. 112: 1912–1919. doi.org/10.1093/jee/toz112