Data in Brief 69 (2026) 113417



Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib



Data Article

Phenotypic, carotenoid and SSR marker-band dataset for VNUA sweetpotato germplasm



Pham Quang Tuan, Nguyen Thi Thu, Doan Thu Thuy, Vu Thi Xuan Binh, Nguyen Van Loc, Nguyen Trung Duc*

Vietnam National University of Agriculture, Trau Quy Town, Gia Lam District, Hanoi, Vietnam

ARTICLE INFO

Article history:
Received 4 August 2026
Revised 20 October 2026
Accepted 2 November 2026
Available online 5 November 2026

Keywords:
Ipomoea batatas
Phenotyping
Storage-root yield
Total carotenoid
SSR marker bands

ABSTRACT

This data article describes a curated *Ipomoea batatas* (L.) Lam. dataset generated from a sweetpotato germplasm evaluation conducted at Vietnam National University of Agriculture, Hanoi, Vietnam, provided as a single multi-sheet Excel workbook. The workbook contains a VNUA-KL accession registry for 50 materials, unreplicated Stage I screening records for all 50 accessions, validation-stage plot-level records for 26 genotypes across three replicates, and a genotype-level summary covering 22 phenotypic and quality traits, including yield components, root architecture classes, spectrophotometric total carotenoid estimates, breeder-assessed eating-quality descriptors and pest/disease tolerance scores. Carotenoid data are provided for the 26-genotype validation panel. Binary SSR marker-band profiles and a pairwise Jaccard similarity matrix for 26 genotyped accessions, 20 of which are in the validation panel, are included, generated from ten microsatellite markers (S0940, S0911, S0782, S0819, SPGS2, S0871, S0694, S0922, S0338 and SPGS3) yielding 21 scored marker bands in total. A dedicated Variable_Dictionary sheet defines every trait code, full column name and unit. VNUA-KL codes identify each accession and genotype consistently throughout the workbook; the confirmed validation-stage control clone

* Corresponding author.
E-mail address: ntduc@vnua.edu.vn (N.T. Duc).

https://doi.org/10.1016/j.dib.2026.113417
2352-3409/© 2026 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

is VNUA-KL050. The data can be reused for germplasm characterisation, accession identity checks and multi-trait selection studies.

© 2026 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Specifications Table

Subject area	Agricultural and Biological Sciences
More specific subject area	Sweetpotato breeding; quantitative genetics; SSR fingerprinting; carotenoid data
Type of data	Excel workbook (.xlsx) containing phenotypic, carotenoid, and SSR marker-band data; supplementary gel photographs (.pdf) and daily weather series (.csv)
How data were acquired	Field evaluation, spectrophotometric measurements, and SSR marker scoring
Data format	Curated and processed tabular data with raw and analysis sheets retained
Data source location	Vietnam National University of Agriculture, Hanoi, Vietnam
Accessibility	Repository name: Mendeley Data Data identification number: 10.17632/g43zrny7m3.2 Direct URL to data: https://data.mendeley.com/datasets/g43zrny7m3/2 Instructions for accessing these data: publicly downloadable under a CC BY 4.0 licence; no registration required [1].
Related research article	No related research article has been published yet. Two companion manuscripts have been submitted: a research article describing phenotypic diversity, MGIDI-based multi-trait selection and SSR fingerprinting of this VNUA sweetpotato panel, and a method article on that workflow.

1. Value of the Data

- Sweetpotato is a priority crop for provitamin A biofortification programmes, yet public datasets that combine phenotypic, carotenoid and molecular-identity data for the same clonal-crop accessions remain scarce [2]. This workbook lets researchers test associations between agronomic performance, carotenoid content and genetic identity within one accession set, instead of reconciling separate phenotyping and fingerprinting studies.
- The paired unreplicated Stage I screening (50 accessions) and replicated Stage II validation (26 genotypes) records let other researchers benchmark selection strategies on the same panel, such as testing whether unreplicated screening ranks predict replicated performance, or applying multi-trait selection indices to the Stage II trait matrix.
- The binary marker-band matrix and pairwise Jaccard similarity scores support identity checks of the 26 genotyped accessions, 20 of which have Stage II phenotypes, when the material is re-grown or exchanged. New samples of these accessions can be scored with the same ten markers and scoring convention.
- A dedicated Variable_Dictionary sheet defines every trait code, full column name and unit independently of this manuscript, so the workbook can be reused, filtered or merged with other sweetpotato datasets without needing to consult the article text.

2. Background

Sweetpotato breeding datasets frequently combine field phenotyping, quality measurements and marker-based identity checks. The data layers are often scattered across reports, spreadsheets and supplementary files. This dataset was compiled to provide a single reusable data object for a VNUA sweetpotato germplasm panel, with public analysis identifiers, trait-code metadata, plot-level records and marker-band files connected through the same accession registry. Recently published stand-alone sweetpotato trait datasets in this journal provide direct prece-

dent for this kind of data-only descriptor [3], including one accompanied by combined analysis of variance, correlation, principal component and multi-trait ideotype-index analysis [4].

The dataset was prepared against the background of sweetpotato breeding priorities that include storage-root yield, root architecture, nutritional quality, sensory traits and reliable germplasm identity [2,5–10]. The molecular component uses SSR marker-band data for identity control rather than genome-wide inference [11–14]. This distinction is important because the SSR panel is compact and was scored as binary marker-band presence/absence data.

3. Data Description

The dataset is provided as a single Excel workbook, Paper_1_Data_in_Brief.xlsx, with eleven sheets, described in Table 1. Two supplementary files accompany this article: the daily NASA POWER weather series (Supplementary File S1, .csv) and the SSR gel photographs with the lane-to-accession key (Supplementary File S2, .pdf).

The official accession code is VNUA-KL001 to VNUA-KL050, used consistently to identify every accession and genotype throughout the workbook. The confirmed control clone is VNUA-KL050. The original field or laboratory code is retained only in the Genotype registry and the raw plot-level sheet (2025_N26_genotypes_plot), for traceability, and plays no role in analysis or reporting elsewhere in the workbook.

Table 2 summarises the Stage I (22 variables) and Stage II (23 variables) genotype-level sheets. Measured variables are given as range and mean ± SD, and scores as the number of accessions or genotypes at each value. The 22 traits of the Abstract are the Stage II variables except LeafCurlScore, which is invariant.

4. Experimental Design, Materials and Methods

4.1. Plant material and accession coding

The data were generated from sweetpotato germplasm managed by Vietnam National University of Agriculture. All data come from institutional project T2024-41-25TD and have not been published previously. A standard VNUA-KL code was assigned to 50 accessions and is used consistently throughout the workbook. The original field or laboratory code is retained only in the Genotype registry and the raw plot-level sheet, for traceability. The validation panel contained 26 genotypes chosen by the project team from the 50-accession registry after the Stage I screening and marked in the SelectedForStageII column of the Stage I sheet. No accession was excluded for poor performance or missing data. Selected and non-selected accessions had the same mean Stage I yield (23.5 t ha^{-1}). VNUA-KL050 is the confirmed control clone in the validation stage.

4.2. Phenotypic and quality data

The screening-stage dataset contains the 50-accession panel evaluated from October 2024 to March 2025 in a sequential, non-replicated field screening, using the same 28 m^2 plot layout as the validation stage (two single ridges of 10 m x 1.4 m, 100 vines planted per plot at 5 vines m^{-1}) but with a single, unreplicated plot per accession. The validation-stage dataset contains 26 genotypes evaluated from May 2025 to December 2025 in a randomised complete block design with three plot replicates; the randomisation layout was not archived. These dates span each project stage as a whole. Crop duration, from planting to harvest, is recorded as GrowthDays. Each validation plot covered 28 m^2 including furrows and contained two single ridges of 10 m x 1.4 m; 100 vines were planted per plot at approximately 5 vines m^{-1}. Harvest was conducted at physiological maturity, indicated by natural yellowing of approximately one-third of basal

Table 1
Summary of workbook sheets.

Sheet	Content	Records	Recorded or derived	Missing data
Read_Me	Documentation of the deposit: sheet descriptions, codes and units, the SSR gel and lane record, and how to re-derive the processed columns	Documentation	Metadata	None
Genotype	VNUA-KL registry: project code and holding institution	50 accessions	Recorded	None
2024_50_genotypes	Stage I screening: phenology, yield components, eating quality, morphology, storage-root colour and pest, disease and cold scores	50 accessions	Recorded, from the project Stage I report; YieldRank derived	None
2025_N26_genotypes_plot	Stage II plot-level records, 26 genotypes in three blocks	78 plots	As compiled in the project Stage II report; ComponentRootsPerPlant and PlotYield_kg_from_Yield derived	None
2025_N26_genotypes_mean	Stage II genotype summary, including the total carotenoid estimate	26 genotypes	Means of the plot sheet; TasteMeanScore, StressMeanScore and YieldRank derived	None
2025_SSR_Binary_MarkerBand	Binary marker-band calls for ten SSR markers	21 bands, 26 profiles	Laboratory scoring matrix; profile columns labelled with VNUA-KL codes from the laboratory lane order	Two unreadable lanes, four blank band cells, counted in MissingBinaryCalls
2025_SSR_Jaccard_Similarity	Pairwise Jaccard similarity of the 26 profiles	26×26	Derived from the binary matrix over the bands scored in both profiles	None
Variable_Dictionary	Code, column name, Vietnamese name, unit and definition of every column, with the formula of each derived column	53 entries	Metadata	None
Changelog	Changes from version 1 to version 2 of the deposit	Documentation	Metadata	None
SSR_Lab_Scoring_Matrix	Laboratory marker-scoring matrix with its original sample labels, in lane order	21 bands, 26 samples	Recorded, as supplied by the laboratory	Four blank cells, the two unreadable lanes
SSR_Band_Informativeness	Presence frequency and informativeness of each scored band	21 bands	Derived from the binary matrix	None

Table 2
Per-trait descriptive statistics for the numeric variables of the Stage I screening panel (50 accessions) and the Stage II validation panel (26 genotypes).

Trait	Code	Unit	Stage	n	Min	Max	Mean ± SD or score (n)
RootingDays	ROOTD	days	I	50	4	6	4.98 ± 0.68
BranchingDays	BRD	days	I	50	8	15	10.42 ± 1.70
CanopyCoverDays	CANOPY	days	I	50	17	26	21.34 ± 2.45
GrowthDays	TGST	days	I	50	120	135	129.00 ± 3.91
LargeRoots	LROOT	roots	I	50	0.0	3.0	1.56 ± 0.82
MediumRoots	MROOT	roots	I	50	0.0	5.4	3.10 ± 1.17
SmallRoots	SROOT	roots	I	50	0.0	3.4	1.25 ± 0.91
RootsPerPlant	RPP	roots	I	50	4.7	6.7	5.91 ± 0.47
RootWeight_g_perPlant	RWPP	g	I	50	539.3	731.3	652.25 ± 48.42
Yield_t_ha	YLD	t/ha	I	50	19.3	27.2	23.47 ± 1.93
SweetnessScore	SWEET	points	I	50	1	5	1 (7); 3 (33); 5 (10)
MealinessScore	MEAL	points	I	50	3	5	3 (41); 5 (9)
InternodeLength_cm	INTERN	cm	I	50	2.2	8.0	4.62 ± 1.46
InternodeLength_Dispersion_cm	INTDSP	cm	I	50	0.3	1.8	0.94 ± 0.43
StemDiameter_mm	STEMD	mm	I	50	3.3	6.7	4.69 ± 0.85
StemDiameter_Dispersion_mm	STMDSP	mm	I	50	0.1	1.4	0.58 ± 0.32
LeafWormScore	LWORM	points	I	50	1	2	1 (38); 2 (12)
StemBorerScore	SBORER	points	I	50	1	2	1 (43); 2 (7)
WeevilScore	WEEVIL	points	I	50	1	2	1 (25); 2 (25)
LeafCurlScore	LCURL	points	I	50	1	1	1 (50)
BlackRotScore	BROT	points	I	50	1	2	1 (43); 2 (7)
ColdToleranceScore	COLD	points	I	50	1	2	1 (37); 2 (13)
GrowthDays	TGST	days	II	26	120	132	127.58 ± 3.42
RootingDays	ROOTD	days	II	26	4	6	4.88 ± 0.59
BranchingDays	BRD	days	II	26	8	14	10.15 ± 1.46
CanopyCoverDays	CANOPY	days	II	26	17	25	19.88 ± 1.97
InternodeLength_cm	INTERN	cm	II	26	2.2	8.0	4.58 ± 1.66
StemDiameter_mm	STEMD	mm	II	26	3.8	5.7	4.73 ± 0.56
LeafWormScore	LWORM	points	II	26	1	2	1 (22); 2 (4)
StemBorerScore	SBORER	points	II	26	1	2	1 (22); 2 (4)
WeevilScore	WEEVIL	points	II	26	1	2	1 (17); 2 (9)
LeafCurlScore	LCURL	points	II	26	1	1	1 (26)
BlackRotScore	BROT	points	II	26	1	2	1 (21); 2 (5)
ColdToleranceScore	COLD	points	II	26	1	2	1 (21); 2 (5)
LargeRoots	LROOT	roots	II	26	0.0	3.0	1.80 ± 0.75
MediumRoots	MROOT	roots	II	26	0.0	4.4	2.71 ± 1.02
SmallRoots	SROOT	roots	II	26	0.0	3.4	1.38 ± 0.94
RootsPerPlant	RPP	roots	II	26	5.1	6.8	5.90 ± 0.48
RootWeight_g_perPlant	RWPP	g	II	26	549.8	716.1	659.05 ± 45.00
Yield_t_ha	YLD	t/ha	II	26	20.3	28.3	23.60 ± 1.89
SweetnessScore	SWEET	points	II	26	1	5	1 (5); 3 (18); 5 (3)
MealinessScore	MEAL	points	II	26	3	5	3 (22); 5 (4)
Carotenoid_mg_100g_DW	CAR	mg/100 g DW	II	26	4.8	6.4	5.70 ± 0.43
StressMeanScore	STRS	points	II	26	1.0	1.5	1.00 (8); 1.17 (11); 1.33 (5); 1.50 (2)
TasteMeanScore	TASTE	points	II	26	2	5	2 (4); 3 (17); 4 (4); 5 (1)

n, number of accessions or genotypes with a recorded value; SD, sample standard deviation. For scores (points), the last column lists each observed value with the number of accessions or genotypes in parentheses. TasteMeanScore and StressMeanScore are means of scores. The carotenoid unit is discussed under Phenotypic and quality data. No numeric cell is missing in either sheet. The Dispersion variables hold the ± value given in the project Stage I report, which does not state whether it is a standard deviation or a standard error.

leaves. Crop management followed the national sweetpotato VCU testing guideline QCVN 01–60:2011/BNNPTNT [15] as recorded in the project method reports. For regional climate context (not project-measured data; no weather station was operated at the trial site), NASA POWER reanalysis for the grid cell containing the trial site (21.01°N, 105.94°E) [16] indicates a mean temperature of 19.5 °C (range 5.0–33.1 °C) and 100 mm total precipitation during Stage I (October 2024–March 2025), and a mean temperature of 25.3 °C (range 7.2–37.2 °C) and 2031 mm total precipitation during Stage II (May–December 2025), consistent with Hanoi's cool dry winter and warm monsoon-season climate; the full daily series is provided as Supplementary File S1 (.csv). All 50 accessions have complete, directly measured Stage I field records.

Phenotypic variables include growth traits, root number, root weight per plant, storage-root yield and root-size classes, matching the yield-architecture logic used in sweetpotato breeding and quality assessment [2,6,7,10]. GrowthDays ranges from 119 to 133 days across the 78 validation plots and from 120 to 132 days across genotype means (Table 2). Root-size classes are counts of large, medium and small storage roots per plant, as recorded in the project reports. The class limits printed in those reports (Stage I > 200, 150–200 and <150 g root^{-1}; Stage II >250, 200–250 and <200 g root^{-1}) do not fit the recorded root weight per plant, which averages about 110 g per root at both stages. The counts are therefore relative size classes and should not be converted to mass. Categorical morphology variables use category terms matching the national sweetpotato VCU testing guideline QCVN 01–60:2011/BNNPTNT [15] for vine growth habit, vine colour, storage-root skin colour and flesh colour; leaf colour and leaf shape have no counterpart in that standard and instead follow the international CIP/AVRDC/IBPGR sweetpotato descriptor standard [17]. The original Vietnamese category terms are retained alongside their English translation in the Variable_Dictionary sheet for traceability. Quality variables include spectrophotometric total carotenoid estimate, sweetness score, mealiness score and mean taste score. Sweetness and mealiness were scored by tasting boiled storage roots 7–10 days after harvest on the QCVN 01–60:2011/BNNPTNT raw scale [15], on which a lower score denotes more desirable eating quality (1 = sweetest/mealiest, higher values progressively blander/less mealy); TasteMeanScore, derived for Stage II only, is their mean. Total carotenoid is expressed as mg 100 g^{-1} DW. The project carotenoid protocol, following Ooi et al. [18], used storage roots at harvest: slices were dried at 60 °C to constant mass, milled to powder, extracted from 2 g dry powder with 8 mL acetone, vortexed for 3 min, mixed with 2 mL 20% NaCl, centrifuged at 5000 rpm for 10 min at 25 °C and read at 450 nm on a UV-2700 Shimadzu spectrophotometer. Individual absorbance readings were not retained. The project-recorded calculation was: total carotenoid estimate (mg 100 g^{-1} DW) = $(A450 \times V \times 10^4)$ / $(2150 \times W)$, where $A450$ is absorbance at 450 nm, V is extract volume in mL and W is dry sample mass in g. This expression gives µg g^{-1}, but the values were reported in mg 100 g^{-1} DW. Without the absorbance readings, the tenfold difference cannot be resolved, so the values should only be used to compare genotypes within this panel. No absorption spectrum of the extracts was recorded. The coefficient 2150 is the β-carotene coefficient used by Ooi et al. [18] and is applied to the whole extract, so the estimate assumes that all carotenoids share it. The eating-quality variables are breeder-assessed project descriptors rather than consumer sensory-panel data. No personal information or identifiable human-subject data were collected. Pest, disease and cold-tolerance status were recorded as LeafWormScore, StemBorerScore, WeevilScore, LeafCurlScore, BlackRotScore and ColdToleranceScore at both stages, on two-level bands defined in the project reports (observed range 1–2). The five pest and disease scores (LeafWorm, StemBorer, Weevil, LeafCurl and BlackRot) used a uniform incidence threshold: score 1 = <5% of leaves, vines or plants affected, score 2 = 5–15%. ColdToleranceScore has no incidence measure and was scored qualitatively as 1 = very good and 2 = good.

4.3. SSR marker data

SSR marker-band data were compiled from project marker-scoring files and converted into marker-level summaries and pairwise similarity outputs, following the operational use of SSR

fingerprints for sweetpotato identity checking and germplasm comparison [11–13]. Ten markers were used: S0940, S0911, S0782, S0819, SPGS2, S0871, S0694, S0922, S0338 and SPGS3, scoring a total of 21 marker bands across 26 genotyped accessions, 20 of which are in the validation panel. Eight markers (GDAAS0940, GDAAS0911, GDAAS0782, GDAAS0819, GDAAS0871, GDAAS0694, GDAAS0922 and GDAAS0338) were developed by the Guangdong Academy of Agricultural Sciences research group [19]; SPGS2 and SPGS3 were developed by Meng et al. [11]. Young leaf tissue was sampled 45 days after transplanting and genomic DNA was extracted using a modified CTAB method. PCR amplification was performed in 20 µL reactions containing 10 µL 2× Master Mix, 1 µL DNA template, 4 µL 10 µM primer solution and 5 µL nuclease-free water, following the PCR conditions described by Chen et al. [19]: initial denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at 54 °C for 35 s and extension at 72 °C for 40 s; and a final extension at 72 °C for 3 min. The same conditions were used for all ten markers. PCR products were resolved on 3% agarose gels at 150 V and 200 mA for 50–70 min, stained with GelRed, sized against a 100 bp Vegazyme ladder, visualised under UV and scored as binary presence/absence marker-band data. A band was scored present (1) when a discrete product of the expected size was visible against the ladder and absent (0) otherwise. Weak or unclear lanes were re-amplified and re-run before the final call, and two lanes that could not be read were left blank (four missing band calls). The gel photographs used for scoring are provided as Supplementary File S2. The first page of that file gives the lane-to-accession key (lane, laboratory label and VNUA-KL code). Marker performance was summarised using band-level polymorphic information content (PIC), marker-band rows, missing binary calls and polymorphism status. For a binary marker-band row, PIC was calculated from the band-presence frequency p and absence frequency q as $1 - p^2 - q^2$, which for a dominant marker equals $2p(1 - p)$ [20]. Across the 21 bands it ranged from 0.000 to 0.500 (mean 0.432), and 20 of the 21 bands were polymorphic. Pairwise similarity among the 26 profiles was calculated using the Jaccard coefficient ($J = a / (a + b + c)$, where a, b and c count bands shared, unique to the first profile and unique to the second), computed over the bands scored in both profiles; band-scoring and similarity-coefficient conventions for binary marker data follow established genetic-marker methodology [14]. The SSR sheets should be reused as binary marker-band fingerprints. They are not dosage-aware SSR genotypes and should not be used for allele dosage estimation, model-based ancestry inference, phylogenetic inference or formal population-genetic structure analysis.

4.4. Data curation

The raw and processed sheets should be read together, since the raw plot-level sheet preserves original field entries while the genotype-level summary applies the curated VNUA-KL identifiers. The workbook was compiled with Python (pandas). For version 2 of the deposit, the Stage I sheet was checked against the project Stage I report, and the SSR calls against the laboratory scoring matrix and the gel photographs. Every correction is listed in the Changelog sheet.

Limitations

All data come from one location. Stage I (October 2024–March 2025) and Stage II (May–December 2025) fell in different seasons and years, so stage effects cannot be separated from season and year effects. Stage I had a single unreplicated plot per accession. The 26 Stage II genotypes are a purposive subset of the 50 accessions, chosen by the project team. In the Stage II records, seven genotypes have yield-component means identical to the Stage I record of a different accession, and the three plot values of each genotype average exactly to the recorded genotype mean. The plot sheet is therefore best read as the breakdown given in the project report, and its plot values should not be treated as independent replicates. The SSR data cover 26 accessions, 20 of them with Stage II phenotypes. Ten markers and 21 dominant bands are enough for identity checks within this set but too few to estimate genetic diversity or population structure.

Ethics Statement

The authors have read and follow the ethical requirements for publication in Data in Brief. The current work does not involve human subjects, animal experiments or data collected from social media platforms. The eating-quality descriptors were breeder-assessed project descriptors recorded during germplasm evaluation. No consumer survey, personal information or identifiable human data were collected.

CRediT Author Statement

Pham Quang Tuan: Project administration, Investigation, Resources, Funding acquisition, Methodology, Writing – review & editing. **Nguyen Thi Thu:** Investigation, Data curation, Writing – review & editing. **Doan Thu Thuy:** Investigation, Methodology, Data curation, Writing – review & editing. **Vu Thi Xuan Binh:** Investigation, Data curation, Writing – review & editing. **Nguyen Van Loc:** Investigation, Data curation, Writing – review & editing. **Nguyen Trung Duc:** Conceptualization, Investigation, Methodology, Formal analysis, Validation, Writing – original draft, Writing – review & editing.

Declaration of Generative AI and AI-assisted Technologies in the Manuscript Preparation Process

During the preparation of this work the authors used AI-assisted tools (Microsoft Copilot 365) in order to support manuscript language editing. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Acknowledgements

Funding: This work was supported by Vietnam National University of Agriculture, Hanoi, Vietnam (institutional project T2024-41-25TD). The funder had no role in the preparation of the data article beyond institutional and financial support.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2026.113417.

References

[1] Q.T. Pham, T.T. Nguyen, T.T. Doan, T.X.B. Vu, V.L. Nguyen, T.D. Nguyen, Phenotypic, carotenoid and SSR marker-band dataset for VNUA sweetpotato germplasm, Mendeley Data (2026) V2, doi:10.17632/g43zrny7m3.2.

[2] J.W. Low, et al., Nutrient-dense orange-fleshed sweetpotato: advances in drought-tolerance breeding and understanding of management practices for sustainable next-generation cropping systems in Sub-Saharan Africa, Front. Sustain. Food Syst. 4 (2020) 50, doi:10.3389/fsufs.2020.00050.

[3] Z. Alam, et al., Stability and performance analysis of storage root yield in a dataset of sweet potato varieties (*Ipomoea batatas* L.), Data Brief 54 (2024) 110493, doi:10.1016/j.dib.2024.110493.
[4] Z. Alam, et al., Ideotype based genotype selection in a multivariate dataset of sweet potato (*Ipomoea batatas* L.), Data Brief 55 (2024) 110575, doi:10.1016/j.dib.2024.110575.
[5] M. Modak, et al., Genotype × harvest time effects on yield, root size distribution, and sensory quality in sweet potato (*Ipomoea batatas* L.), BMC Plant Biol. 26 (2026) 849, doi:10.1186/s12870-026-08693-6.
[6] J. Swanckaert, et al., Sweet Potato (*Ipomoea batatas* (L.) Lam.) Breeding, in: J.M. Al-Khayri, et al. (Eds.), Advances in Plant Breeding Strategies: Vegetable Crops, Springer, Cham, 2021, pp. 513–546, doi:10.1007/978-3-030-66965-2_12.
[7] V.-D. Truong, et al., Sweetpotato production, processing, and nutritional quality, in: M. Siddiq, M.A. Uebersax (Eds.), Handbook of Vegetables and Vegetable Processing, 2nd ed., Wiley-Blackwell, 2018, pp. 811–838, doi:10.1002/9781119098935.ch35.
[8] Y. Xiao, et al., Genetic and genomic research on sweet potato for sustainable food and nutritional security, Genes 13 (2022) 1833, doi:10.3390/genes13101833.
[9] M. Yan, et al., Exploring and exploiting genetics and genomics for sweetpotato improvement: status and perspectives, Plant. Commun. 3 (2022) 100332, doi:10.1016/j.xplc.2022.100332.
[10] C. Tang, et al., Comprehensive quality assessment of 296 sweetpotato core germplasm in China: a quantitative and qualitative analysis, Food Chem.: X 24 (2024) 102009, doi:10.1016/j.fochx.2024.102009.
[11] Y.-S. Meng, et al., SSR fingerprinting of 203 sweetpotato (*Ipomoea batatas* (L.) Lam.) varieties, J. Integr. Agric. 17 (2018) 86–93, doi:10.1016/S2095-3119(17)61687-3.
[12] Z. Luo, et al., Genetic fingerprint construction and genetic diversity analysis of sweet potato (*Ipomoea batatas*) germplasm resources, BMC Plant Biol 23 (2023) 355, doi:10.1186/s12870-023-04329-1.
[13] D.C. Gemenet, et al., Development of diagnostic SNP markers for quality assurance and control in sweetpotato [*Ipomoea batatas* (L.) Lam.] breeding programs, PLoS ONE 15 (2020) e0232173, doi:10.1371/journal.pone.0232173.
[14] M. Nei, W.-H. Li, Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc. Natl. Acad. Sci. USA 76 (1979) 5269–5273, doi:10.1073/pnas.76.10.5269.
[15] QCVN 01-60:2011/BNNPTNT, National Technical Regulation on Testing for Value of Cultivation and Use of Sweet Potato Varieties, Ministry of Agriculture and Rural Development, Vietnam, 2011.
[16] NASA Langley Research Center (LaRC) POWER Project, Prediction Of Worldwide Energy Resources (POWER) Daily 2.9.5, https://power.larc.nasa.gov/, accessed 2026-08-01.
[17] CIP, AVRDC, IBPGR, Descriptors for Sweet Potato, Z. Huaman (Ed.), International Board for Plant Genetic Resources, Rome, 1991 ISBN 92-9043-204-7.
[18] S.F. Ooi, et al., Carotenoids, phenolics and antioxidant properties of different sweet potatoes (*Ipomoea batatas*) varieties, IOP Conference Series: Earth and Environmental Science, 756, 2021, 012077, doi:10.1088/1755-1315/756/1/012077.
[19] Y. Chen, et al., Genetic diversity and population structure of sweet potato landraces based on phenotypic traits, carotenoid content, and SSR molecular markers, Res. Square (preprint) (2023), doi:10.21203/rs.3.rs-2417254/v1.
[20] I. Roldán-Ruiz, et al., AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.), Mol. Breed. 6 (2000) 125–134, doi:10.1023/A:1009680614564.