Journal of Agricultural Science; Vol. 9, No. 9; 2017
ISSN 1916-9752 E-ISSN 1916-9760
Published by Canadian Center of Science and Education

Towards the Selection of Superior Sesame Lines Based on Genetic and Phenotypic Characterisation for Uganda

Walter Okello-Anyanga[1,2], Karin Hasel-Hohl[3], Agnes Burg[3], Stephan Gaubitzer[3], Patrick R. Rubaihayo[1], Johann Vollmann[4], Paul T. Gibson[1,5], Silvia Fluch[3] & Eva M. Sehr[3]

[1] Makerere University, Kampala, Uganda

[2] National Semi-Arid Resources Research Institute (NaSARRI), Serere, Uganda

[3] AIT Austrian Institute of Technology GmbH, Tulln, Austria

[4] BOKU Univ. of Natural Resources and Life Sciences, Tulln, Austria

[5] Illinois University, Carbondale, IL, USA

Correspondence: Eva M. Sehr, Center for Health & Environment, AIT Austrian Institute of Technology GmbH, Konrad-Lorenz Str. 24, 3430 Tulln, Austria. E-mail: eva-maria.sehr@ait.ac.at

Received: June 22, 2017 Accepted: July 23, 2017 Online Published: August 15, 2017

doi:10.5539/jas.v9n9p13 URL: https://doi.org/10.5539/jas.v9n9p13

Abstract

Understanding agricultural biodiversity is critical to formulate breeding strategies for crop improvement and it impacts both, conservation and collection activities. Especially germplasm collections serve as valuable resources, thus, their adequate characterisation is of utmost importance. Although Uganda ranks seventh in African sesame production, meagre research was conducted to determine the current genetic diversity among its germplasm. Therefore, in the present study part of the sesame germplasm conserved at the National Semi-Arid Resources Research Institute (NaSARRI) in Uganda focusing on 85 established lines was genetically and phenotypically characterised. Population genetic and structure analyses revealed rather a low extend of genetic diversity (expected heterozygosity [H_E], or gene diversity [D]) ranging from 0 to 0.38 per entry, but a high extend of admixture within and between entries. This decrease of heterozygosity is supported by a fixation index (F_{ST}) of 0.530, indicating a medium genetic differentiation among entries. The analysis of quantitative and qualitative agromorphological traits revealed a great inter-trait variability among the entries and further indicated a certain conservation of some of the traits reflecting the geographic origin of the analysed entries. Based on both, the genetic and phenotypic characterisation, a selection of 26 superior entries is proposed, which may form a valuable basis both for farmers and breeders.

Keywords: genetic diversity, germplasm characterisation, sesame

1. Introduction

Sesame (*Sesamum indicum* L.) is a primarily self-pollinated diploid with 2n = 26 chromosomes. It belongs to the Pedaliaceae (order Lamiales), a small family of 15 genera and 70 species characterised by annual and perennial growth forms. Sesame is an important and ancient crop cultivated in hot, dry climates for its oil and protein-rich seeds (Bedigian et al., 1986). Domesticated in the Indian subcontinent (Bedigian, 2003), currently, sesame is grown throughout the tropical and subtropical regions of the world with Sudan, China, India, and Myanmar being the top producers in 2014, together covering 46% of the world production (FAO, 2015). On the African continent, Uganda with an annual production of 124,300 tonnes ranks seventh in sesame production (FAO, 2015). Sesame, commonly known as simsim in Uganda, was introduced from Kenya in 1910 and since then has been distributed, cultivated, and used (Rubaihayo et al., 1997). Its adaptability to harsh climatic conditions including heat and drought makes it a favourable crop in north-eastern Uganda. Especially in the last decade, sesame has experienced a worldwide boom increase in its production to 158 per cent from 2004 to 2014 (FAO, 2015). Although sesame accounted for 83 per cent of total agricultural sales in 2014 in Uganda (Proctor, 2015), neither its production nor its productivity increased markedly since 2005 (FAO, 2015).

To sustain a level of high productivity and yield, or even to increase it, sesame breeding strongly relies both on genetic diversity and genetic purity. For the former, available germplasm collections serve as an important source for the breeders to combat new pests and diseases, and to produce better adapted varieties for the changing

environments. However, to be able to utilize the wealth of diversity in germplasm collections, their genetic and phenotypic characterisation is indispensable. Currently, genetic diversity is measured by using morphological, biochemical, and molecular markers, whereby the latter marker system became the most attractive one in recent years (Govindaraj et al., 2015). However, phenotypic characterisation is the first step in the classification and description of any germplasm. Several studies have exploited high genetic diversity in populations of sesame by analysing morphological traits only, thereby providing valuable information for cultivar selection to be used in different breeding programs (Arriel et al., 2007; Ansah et al., 2015; Falusi et al., 2015). Other studies were performed using a wide palette of molecular markers such as AFLP, ISSR, SSR and RAPD markers for germplasm diversity analysis and the construction of genetic maps (Laurentin et al., 2006; Sharma et al., 2009; Cho et al., 2011; Kumar et al., 2012; Alemu et al., 2013; Zhang et al., 2013; Dossa et al., 2016). Presently, combination of both morphological and molecular markers is increasingly becoming popular for analysing sesame diversity (Parsaeian et al., 2011; Pandey et al., 2015; Sehr et al., 2016).

Germplasm characterization not only produces valuable agronomic and breeding data, but it is also useful for the identification of duplicates within and between collections. Furthermore, when genetic resources are kept *ex situ*, seeds are frequently regenerated to keep their viability and to replenish seed stocks. During this process, certain extent of gene flow may occur as the result of cross-pollination, as well as through physical mixing of seed lots. As a result, the quality and integrity of the germplasm might get severely reduced. Thus, especially when handling cross-pollinating species, additional planning, care, and special techniques are needed in order to ensure the physical/reproductive isolation of accessions that is required to preserve their genetic identity. For sesame, contradictory outcrossing values have been reported ranging from less than 1 to nearly 70 per cent (Yermanos, 1980; Pathirana, 1994; Andrade et al., 2014) and still, sesame is mainly described as self-pollinated crop. Therefore, determining its regional outcrossing potential is of utmost importance not only for breeding, but also for conservation and collection activities and strategies.

Despite this increasing number of studies characterising sesame germplasm collections, knowledge of the genetic diversity of entries assembled on the African continent at the molecular levels is scarce (Gebremichael et al., 2011; Alemu et al., 2013; Nyongesa et al., 2013; Woldesenbet et al., 2015; Sehr et al., 2016). Common findings were a high amount of genetic diversity within accessions, especially of local origin, and the occurrence of a certain extent of admixture between the accessions, which could probably be attributed to cross-pollination and local seed exchange among farmers. Hence, the two main objectives of the present study were i) to analyse and categorize existing variation in the 85 sesame germplasm entries assembled in Uganda, based on their phenotypic and SSR-related genotypic characteristics, and ii) to select superior lines as a valuable basis both for farmers and breeders. Both objectives intend to impact not only sesame breeding and conservation strategies, but, in the long run, also intend to improve sesame performance and usage for farmers.

2. Materials and Methods

2.1 Plant Material and DNA Extraction

A total number of 85 sesame entries were planted in the first rainy season (month of May) of 2010 in a randomized complete block design with three replications. These entries were comprised of germplasm accessions and breeding lines derived from genotypes and crosses of different countries of origin (China, Ethiopia, Kenya, Korea, Tanzania, Uganda, USA, and Zimbabwe) conserved at the National Semi-Arid Resources Research Institute (NaSARRI) in eastern Uganda (Table 1). Seeds stemming from selfed flowers of each entry were planted in a single-row plot of 2 m in length. Border rows were included at the beginning and the end of each replication to control border effects using the purple-coloured variety Sesim 2. Several flowers of five plants per entry were self-pollinated and two capsules per entry were randomly chosen and taken for further analyses. Seeds from the two capsules were germinated separately in Petri dishes. Eight seedlings from each capsule were picked for DNA extraction resulting in 1,360 samples (85 entries, á 2 capsules, á 8 seedlings). The extraction of genomic DNA was performed using the aerial parts of the seedlings following the protocol described by (van der Beek et al., 1992) with minor modifications for high-throughput handling using robotics. The extracted genomic DNA is deposited at the Repository Centre at the AIT Austrian Institute of Technology and is available upon request (Stierschneider et al., 2016). Detailed sample information is given in Appendix 1.

Table 1. List of countries of origin and the corresponding sesame entries

Country of origin	Total No. of entries	Entry numbers
China	20	036-055
Ethiopia	5	003-007
Kenya (ICRISAT)	14	073-086
Korea	1	034
Tanzania	1	035
Uganda	25	001-002; 008-025; 029-030; 072; 098; 114; 118; 191
USA	16	056-071
Zimbabwe	1	087

2.2 Genotyping and Fragment Analysis

Nine nuclear SSR markers with linkage groups indicated in square brackets (CL297 [LG3], CL569 [LG8], CL78 [LG3], CL93 [LG15], GBssr_sa_08 [n/a], GBssr_sa_108 [n/a], GBssr_sa_123 [LG15], GBssr_sa_184 [n/a], and GBssr_sa_72 [LG5]) were applied as described (Sehr et al., 2016). PCR amplification was performed incorporating the FAM-labelled M13 primer according to (Schuelke, 2000) in a total volume of 25 µl consisting of 2.5 µl of 10x reaction buffer (Qiagen or LGC), 1 mM $MgCl_2$, 0.25 µl of 20 mM dNTPs, 0.4 µl of 4 mM primer forward, 1.2 µl of 4 mM primer reverse, either 0.125 µl of 5 U/µl of HotStarTaq DNA Polymerase (Qiagen) or 0.25 µl of KlearTaq (LGC), 5 µl DNA (1:10) and ddH_2O. $MgCl_2$ was not included when K-Taq was used. The conditions of the PCR amplification were as follows: 94 °C (15 min), followed by 35 cycles including 94 °C (30 sec), 50-55 °C (45 sec), 72 °C (1 min), ending in 72 °C (10 min) with a final halt at 4 °C. The PCR products were electrophoresed on 1% agarose gel. The SSR markers were applied via PCR in all 85 entries, however, the entries 029 and 030 were excluded from further analysis due to PCR failure.

The resulting PCR products were diluted and mixed with Hi-Di Formamide and GeneScan 350/500 ROX dye Size Standard according to the manufacturers protocols (Life Technologies). The size of the fragments was resolved based on capillary electrophoresis using the ABI 3110 XL Genetic Analyzer. Allele calling was performed using GeneMapper® Software 5 (Applied Biosystems). Non-amplified loci were scored as missing data.

2.3 Genetic Data Analysis

To avoid allele frequencies bias due to full/half sibship and to be able to infer population genetic structure over the entire germplasm collection, clonality within the dataset was determined *in silico* by measuring the number of 100 per cent multilocus matches. Repeated matching multilocus genotypes were removed from the data set for subsequent analysis. Genetic variation was investigated on the entire dataset as well as on the reduced dataset using standard genetic diversity estimates per locus and entry including expected heterozygosity (H_E; or gene diversity [D]), observed heterozygosity (H_E), inbreeding coefficients (F, F_{ST}, F_{IS}), gene flow (Nm), and an analysis of molecular variance (AMOVA) among and within the countries of origin with 999 permutations was performed. All computations were done using GenAlEx v. 6. 502 (Peakall et al., 2012). Population structure of the reduced set without repeated matching multilocus genotypes was examined using the Bayesian model-based approach implemented in Structure 2.3.4 (Pritchard et al., 2000; Anderson et al., 2008). The number of subgroups (K) evaluated ranged from 1-30. The analysis was performed using five replicate runs per K value, a burn-in period length of 10,000, and a run length of 50,000. The no admixture model was used to determine the correlated cluster. The R package pophelper (Francis, 2016) was used to determine the final K value based on the delta K algorithm (Evanno et al., 2005). Based on the Nei pairwise genetic distance matrix of the entire dataset, a neighbor-joining (NJ) tree using MEGA 6 (Tamura et al., 2013) was created to visualize genetic diversity and relationships among the genotypes.

2.4 Phenotyping and Trait Statistical Analysis

Seeds from the remaining selfed flowers from the plants used for genotyping that formed capsules were planted in the first rainy season (Mid-April) in 2011 and were phenotyped during the second season of 2011 (September) for evaluating agromorphological diversity on the total set of the germplasm (85 entries) at NaSARRI, Uganda. According to the official descriptors for sesame (IPGRI et al., 2004), the following 10 traits were measured in a quantitative approach and were used for further diversity analysis: days to flowering (DTF), days to maturity (DTM), plant height (PH [cm]), plant height to first capsule (HFC [cm]), plant height to first branch (HFB [cm]), number of branches (NB), length of capsule zone (LCZ [cm]), number of capsules on main stem (NCMS), number

of capsules on branches (NCB), and total number of capsules (TNC). The mean values across three replicates are shown in Appendix 2. Entry number 025, however, was measured only once per trait and exhibits very extreme numbers in comparison to all other lines. Since there was no validation of the traits through replicate measurements, this line was thus excluded from further statistical analysis. For an examination of the overall phenotypic diversity across all 10 traits, box plots were generated for each trait using R (R Core Team, 2015). Furthermore, individual trait values were standardized using a z-transformation for equal mean and standard deviation. Standardized data was then subjected to principal component analysis and principal component scores were determined for each line after applying the varimax rotation procedure using SPSS statistical software (PASW Statistics 18, IBM Corp., Armonk, NY, USA). In addition, 14 qualitative traits were measured per entry on the basis of the official descriptors for sesame (Table 2, Appendix 3).

Table 2. Measured qualitative traits

Qualitative traits	Description	Values
Branching habit	few (1-2 branches)	2
	medium (3-4 branches)	3
Capsule length	very short (23 mm)	1
	short (26 mm)	2
	medium (30 mm)	3
Capsule pubescence	weak	1
	medium	2
	strong	3
Capsule width	narrow	1
	medium	2
	broad	3
Corolla pubescence	weak	1
	medium	2
	strong	3
Flowers per leaf axil	one	1
	more	2
Leaf blade colour at maturity	light green	1
	medium	2
	purple	4
Leaf blade length	short	1
	medium	2
	long	3
Leaf blade width	narrow	1
	medium	2
	broad	3
Petiole length	short	1
	medium	2
	long	3
Position of branches	basal	1
	upper part	2
	middle part	3
	basal and upper part	4
Stem color at maturity	light yellow	1
	light green	2
	green	3
	light purple	5
	deep purple	6
Stem fasciation	absent	1
	present	2
Stem pubescence at maturity	weak	1
	medium	2
	strong	3

2.5 Outcrossing Test

In order to assess the rate of outcrossing present under natural conditions, a sibship analysis was performed. For this, a single farmer's field was chosen where Sesim 2 was cultivated already for some generations and where off-type individuals used to appear. Capsules of individual plants were collected randomly on an area of approx. 100 m². Seeds from each capsule were germinated and DNA was extracted from the seedlings as well as from the capsule tissue reflecting the maternal genetics. This way genomic DNA of seven different mother plants (A–G) and 128 seedlings were investigated (Table 3). Each plant was represented by one sampled capsule. From plants A, C, D, and E eight seedlings per capsule and from plants B, F and G 32 seedlings per capsule were analysed. SSR analysis was performed as described above.

Table 3. Samples from farmer's field analysed for outcrossing by sibship analysis

Mother plant	Capsules	No. of seedlings
A	4-loculed	8
B	4-loculed	32
C	3-loculed	8
D	2-loculed (violet)	8
E	4-loculed (green)	8
F	4-loculed	32
G	4-loculed	32

2.6 Entry Selection

In order to assemble a selection of good-performing entries, the entries were chosen based on their qualitative traits, overall hairiness, and high genetic diversity. For each qualitative treat (n = 10) the mean and the standard deviation was calculated. Only those entries having at least five traits above the single positive standard deviation value were chosen for the selection. The qualitative values for hairiness of stem, corolla and capsule were summed up, the mean and the standard deviation was calculated. The entries with values above the single positive standard deviation value were added to the selection. The same *modus operandi* was applied to genetic diversity. The entries with H_E values above the single positive standard deviation value were considered for the selection.

3. Results and Discussion

3.1 Genetic Diversity and Germplasm Structure

Obtaining unbiased estimates of genetic diversity is particularly critical for management and conservation of species. It has been shown that when full siblings were sampled, the estimates of population genetic parameters were affected, also depending on the software tools used (Anderson et al., 2008; Goldberg et al., 2010; Peterman et al., 2016). Thus, in order to be able to infer population genetic structure of the entire germplasm collection by ruling out a possible bias due to consanguinity, a reduced dataset was created by removing repeated matching multilocus genotypes, resulting in 666 remaining samples (Appendix 1).

Heterozygosity and polymorphism were calculated based on the reduced dataset for each locus (Appendix 2) and for each entry (Appendix 3) separately. Per locus, the number of alleles ranged from 2-15, the H_O values were very low (0.00-0.31), H_E values were in the range of 0.03-0.83. Per entry, the calculated mean H_E values (or gene diversity, D) varied from 0 to 0.378 (grand mean = 0.219), and the H_O values ranged from 0 to 0.489 (grand mean = 0.137), whereby the following entries showed no gene diversity at all (H_E and H_O = 0): 002 (Sesim 2, Uganda), 036 (China), 067 (USA), 078 and 079 (Kenya), and 118 (Local Sesim 2, Uganda). This is in line with previous studies, where a gene diversity (H_E, D) between 0 and 0.440 is described in African sesame lines (Gebremichael et al., 2011; Nyongesa et al., 2013; Sehr et al., 2016). Besides the fact, that in comparison to intronic SSRs, exonic SSRs contain less allelic variability because they are subjected to stronger selection pressure due to their functional significance (Li et al., 2004), low H_E values can further be explained by genetic isolation, historical population bottlenecks, founder effects, inbreeding or selection processes. In the case of the herein analysed germplasm sample subset, the latter effects, inbreeding and selection in breeding processes, might have played a major role in declining heterozygosity, which is further reflected by an inbreeding coefficient (F_{IS}) of 0.329 and a fixation index (F_{ST}) of 0.530, indicating a high extent of homozygote individuals and a medium genetic differentiation among entries, respectively. This is in line with the general knowledge, that

the higher the extent of domestication of a given crop is, the narrower is the range of its genetic diversity (Tanksley et al., 1997; Flint-Garcia, 2013). The relative measure of migration between the entries (Nm) was 0.229, which falls in the range of previously described gene flow values of self-pollinated plant species (Govindaraju, 1989). However, gene flow is also described to occur to a certain extent in germplasm collections (de Vicente, 2005). Whereby it is unclear whether the Nm values of our dataset reflect recent gene flow levels (*e.g.* due to cross-pollination or local seed exchange among farmers) or are caused by the fixation of alleles during the breeding processes in evolutionary time.

After grouping the entries according to their country of origin (n = 8, cf. Table 1), the degree of genetic diversity (H_E, D) within a specific country of origin, ranged from a low value 0.18 up to 0.48 (Table 4). The mean number of alleles ranged from 1.44-4.78, whereby the highest allelic richness was seen in the entries stemming from Kenya and USA. Entries from Korea, Tanzania and Zimbabwe showed lesser number of alleles, which might be due to fact that these countries of origin comprise only one entry each. An AMOVA analysis was used to evaluate the diversity components within and between the individuals, which have been grouped into the respective countries of origin. The majority of the variance occurring among the individuals accounted for 57 per cent of the total variation, and seven per cent of the variation was attributed to differences among the countries of origin (Appendix 6). Similar results were also described, where differences among geographical regions were represented only by five per cent of the total variation in sesame lines (Laurentin et al., 2006; Woldesenbet et al., 2015).

Table 4. Mean values of population genetic parameters per country of origin: number of individuals (N), different alleles per locus (N_A), number of effective alleles per locus (N_E), expected and observed heterozygosity (H_E and H_O), and the fixation index (F)

Population	China	Ethiopia	Kenya	Korea	Tanzania	Uganda	USA	Zimbabwe
Entries	20	5	14	1	1	25	16	1
N	162.22	25.44	87.78	10.89	11.00	207.11	135.56	10.00
NA	4.00	2.67	4.78	1.44	1.78	4.33	4.78	1.89
NE	2.59	1.90	2.80	1.33	1.46	2.39	2.43	1.76
HO	0.18	0.15	0.14	0.15	0.11	0.15	0.17	0.10
HE	0.41	0.37	0.48	0.18	0.25	0.42	0.42	0.33
F	0.58	0.54	0.70	0.14	0.45	0.61	0.56	0.70

In order to resolve the relationships among the entries, a NJ tree based on pairwise population matrix of Nei unbiased genetic distance values was generated (Figure 1). The entries present in the Ugandan germplasm collection showed very little to no relationship with respect to their country of origin. Five entries from Kenya maintained their genetic identity and relationship, but the remaining entries were well intermixed. The Sesim 2-related local selection number 188 and the entry number 016 showed identical marker alleles, in contrast to the related entries number 098, 114, and 191, which were highly divergent from their supposed ancestor, Sesim 2.

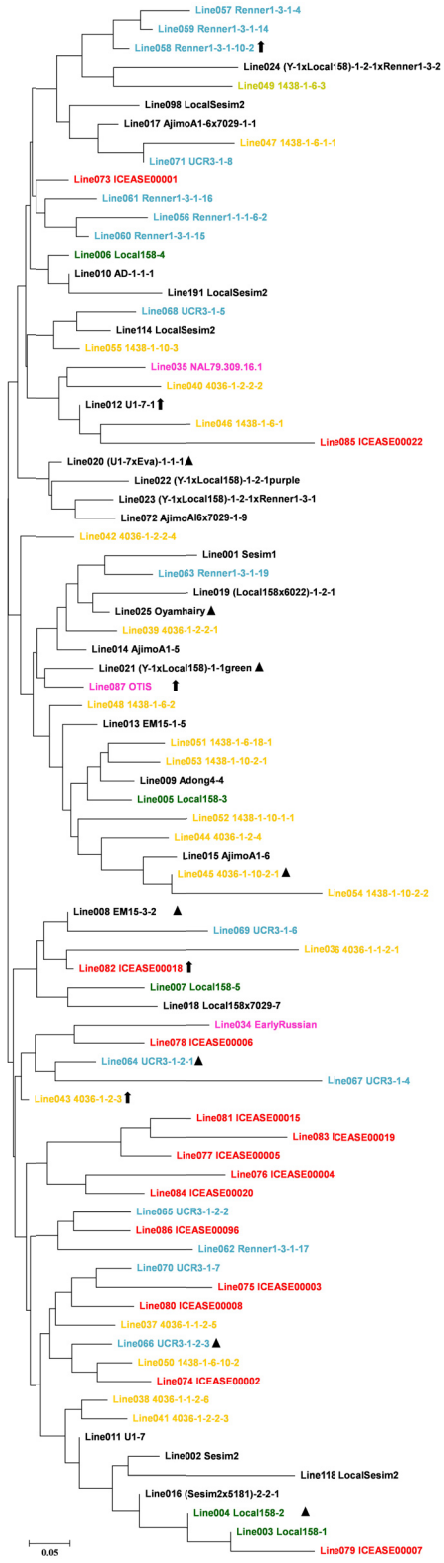

Figure 1. Neighbor-joining (NJ) tree based on pairwise population F_{ST} values of Nei genetic distance matrix of the analysed sesame entries. The entries are colour coded according to their country of origin: China (yellow), Ethiopia (green), Kenya (red), Uganda (black), and USA (turquoise). Korea, Tanzania, and Zimbabwe, each represented only by one entry, are coloured in pink. The 13 most hairy entries are marked with a triangle. The best performing entries characterised by at least five quantitative traits above the single positive standard deviation are marked with an asterisk. The genetically most diverse entries representing H_E values above the positive standard deviation are marked with an arrow

This random distribution of entries without any major clustering based on their origin was also supported by the population structure analysis using the reduced dataset (n = 666) by applying the model-based approach in the Structure software (Anderson et al., 2008). According to the mean L(K) plot the estimated optimum value was K = 22, where the curve started to flatten. However, when taking the delta K into consideration, the optimum value was determined as 14, whereby at K = 2 a second prominent peak was detected. Usually, the highest peak in the delta K plot determines the number of clusters in the population structure, however, in our dataset at that point the highest deviation (noise) was detected in the mean L(K) plot. Neither for K = 2, nor for K = 14 and K = 22 a structure could be identified which can be assigned to the country of origin (Figure 2).

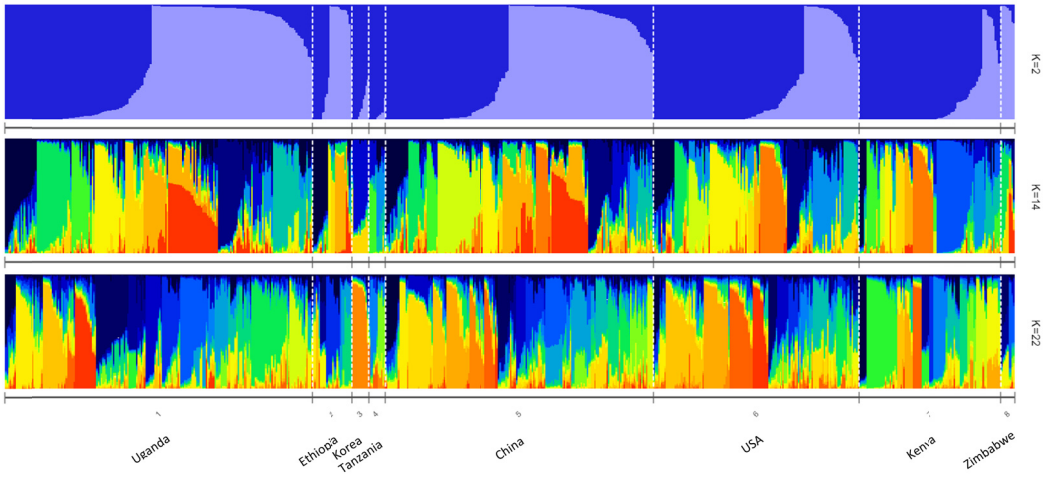

Figure 2. Population structure plots of K = 2, K = 14 and K = 22 based on the reduced data set (n = 666) ordered according to the country of origin. The number of detected gene pools per country of origin at K = 22 were the following: Uganda = 18, Ethiopia = 8, Korea = 1, Tanzania = 3, China = 20, USA = 19, Kenya = 13, and Zimbabwe = 3

When plotting the genetic structure (K = 22) per entry (Figure 3) some entries appear very homogenous in their genetic makeup (*e.g.* entry number 034 or 049). However, most of the entries appear admixed (*e.g.* entry number 012 or 020). This genotypes seem to have a mixed ancestry from parents belonging to different gene pools or geographical origins, which might be due to the breeding process itself or due to material exchange between locations and introduction of novel accessions from different countries (Kim et al., 2002; Dossa et al., 2016).

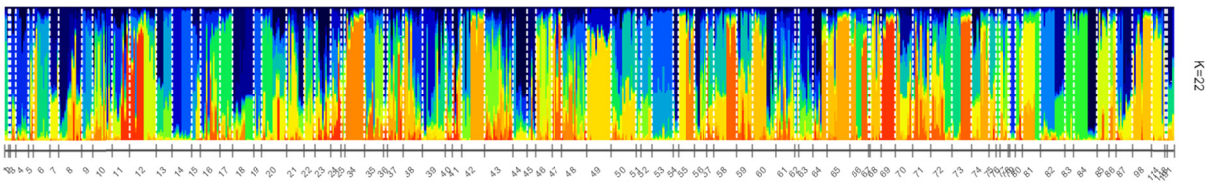

Figure 3. Population structure plot K = 22 based on the reduced data set (n = 666) ordered by entries.

3.2 Diversity Based on Agromorphological Traits

In general, all the entries were characterised with absent stem fasciation and one flower per leaf axil. Although the latter trait is a subject of interest for sesame breeding to increase yield by increasing the number of flowers per axil, none of the entries of the Ugandan germplasm showed more than one flower per axil. Regarding stem characteristics, the majority of germplasm accessions showed light purple stem colour (47%), followed by green (22%), light green (15%), deep purple (13%), and light yellow (3%). Eighty-three per cent of the entries were characterised with medium number of branches (3-4 branches), and 17 per cent with few branches (1-2) along the stem. The position of branches was 65 per cent in the middle part of the stem, 17 per cent in the upper part, 15 per cent in the basal and upper part, and in two per cent in the lower part. Regarding leaf characteristics, 55

per cent of the entries had medium, 36 per cent short, and nine per cent had long petiole length. Forty-three per cent showed medium, 31 per cent short and 26 per cent long leaf blade length. Fifty per cent of the entries had medium, 43 per cent narrow and seven per cent broad leaf blade width. The majority of the entries (62%) had purple leaf blades at maturity, 16 per cent were light green, and 22 per cent had intermediate coloration. Sixty-two per cent showed a short capsule length (26 mm), 36 per cent very short (23 mm), and two per cent medium (30 mm). Seventy-five per cent had medium capsule width, 23 per cent broad and two per cent narrow capsule width.

Among these qualitative traits (Table 2, Appendix 4), especially hairiness is important, since it is reported to be correlated with resistance to gall midge pest and to drought tolerance not only in sesame, but also for other crops (van Rheenen, 1972; Roy et al., 2009). In the herein analysed germplasm, the pubescence of the stem ranged from strong in 54 per cent, over medium in 44 per cent to weak in 45 per cent of the entries. The corolla was strongly pubescent in 27 per cent of the entries, medium in 57 per cent, and weak in 16 per cent. Capsule hairiness was strong in 12 per cent, medium in 42 per cent, and weak in 46 per cent of the entries. Strongest hairiness (by summing up the hairiness values for the stem, corolla and the capsule) was observed in six entries originating from the USA (058, 059, 060, 061, 064, 066), in four Ugandan entries (008, 020, 021, 025), two Chinese entries (045 and 055), and one Ethiopian entry (004). These hairy entries may be preferable for securing healthy plants and therefore could be involved in breeding programs for obtaining plants with moderate resistance to biotic and abiotic stress factors.

In view of the 10 quantitative traits measured in the 85 entries of the germplasm, some traits exhibited great within-trait variability among the lines (Figure 4, Table 5, and Appendix 5). Especially the values for the total number of capsules per plant (TNC) depicted a broad range from 3.6 to 51.3 (grand mean = 23.6) within the analysed germplasm, and the mean plant height (PH) ranged from 81 cm to 150 cm (grand mean = 105 cm). Analysing the traits with respect to their country of origin, some exhibited considerable variation between the countries (Table 5). The breeding lines stemming from Kenya were characterised with an average of 8.7 capsules on the main stem (NCMS) in comparison to the Korean entry having 18.7 average number of capsules. The entries originating from Uganda itself appeared in the upper midfield in all traits. Attention should be paid to those countries, where only one breeding line was analysed. The values might not reflect basic population and thus might be biased in that respect.

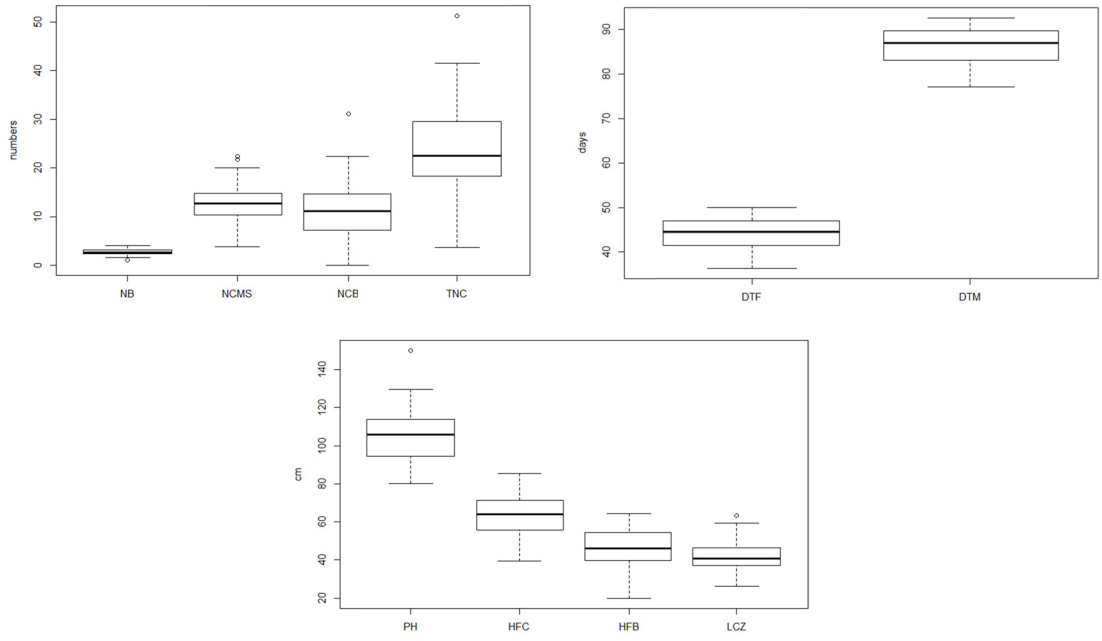

Figure 4. Box plots showing the intra-trait diversity of all 85 entries. Agromorphological traits: number of branches (NB), number of capsules on main stem (NCMS), number of capsules on branches (NCB), and total number of capsules (TNC), days to flowering (DTF), days to maturity (DTM), plant height (PH), plant height until first capsule (HFC), plant height until first branch (HFB), and length of capsule zone (LCZ)

Table 5. Mean values ± standard deviation of 10 agromorphological traits of the sesame germplasm, grouped into their country of origin. Agromorphological traits: number of branches (NB), number of capsules on main stem (NCMS), number of capsules on branches (NCB), and total number of capsules (TNC), days to flowering (DTF), days to maturity (DTM), plant height (PH), plant height until first capsule (HFC), plant height until first branch (HFB), and length of capsule zone (LCZ)

Country (Entries)	NB	NCMS	NCB	TNC	DTF	DTM	PH [cm]	HFC [cm]	HFB [cm]	LCZ [cm]
China (20)	2.5±0.6	12.7±2.7	9.8±3.9	22.0±5.9	42.4±3.3	84.2±3.8	97.7±11.2	57.6±9.6	43.8±8.3	41.2±5.7
Ethiopia (5)	2.5±0.3	15.5±3.1	12.8±5.2	28.3±7.9	45.3±1.2	87.6±1.6	110.4±7.3	71.1±1.0	55.9±2.4	39.5±5.5
Kenya (14)	2.8±0.7	8.7±2.7	7.2±5.4	15.8±7.8	45.0±2.7	87.3±3.1	104.7±10.3	67.1±5.8	48.3±7.3	38.1 ±7.8
Korea (1)	2.3±0.0	18.7±0.0	11.7±0.0	30.3±0.0	46.3±0.0	89.3±0.0	115.0±0.0	56.0±0.0	37.7±0.0	63.3±0.0
Tanzania (1)	3.3±0.0	9.7±0.0	12.3±0.0	22.0±0.0	48.3±0.0	91.7±0.0	119.3±0.0	75.3±0.0	61.3±0.0	36.0±0.0
Uganda (25)	2.9±0.5	13.7±3.0	15.3±4.8	28.8±7.0	46.1±1.9	88.7±2.3	114.8±11.1	71.0±6.4	51.1±6.1	42.8±5.9
USA (16)	2.4±0.5	13.8±2.9	9.3±4.0	22.4 ±.6.0	41.1±3.3	82.7±3.9	95.0±9.8	51.3±8.9	35.6±9.5	44.6±6.4
Zimbabwe (1)	3.3±0.0	11.0±0.0	10.0±0.0	21.0±0.0	49.7±0.0	92.0±0.0	116.7±0.0	74.7±0.0	58.3±0.0	40.0±0.0

The principle component analysis (PCoA) analysis of the quantitative traits of the entries clustered in their respective country of origin revealed a pattern partly associated to the main geographic origins (Figure 5). Clearly, the breeding lines stemming from the USA overlapped with those from China. The entries from Uganda did overlap with those from Ethiopia, whereby the breeding lines from Kenya (ICRISAT) appeared to be more distant. This indicates that some of the entries conserved some morphological traits reflecting their country of origin. The characteristics of crop plants are the product of thousands of years of human management and it is not surprising, that a certain extent of geographical identity can be identified. This separation based on the geographical origin of genotypes might also reflect specific adaptation to different agroecological conditions which affected the diversity in morphological characters through selection processes.

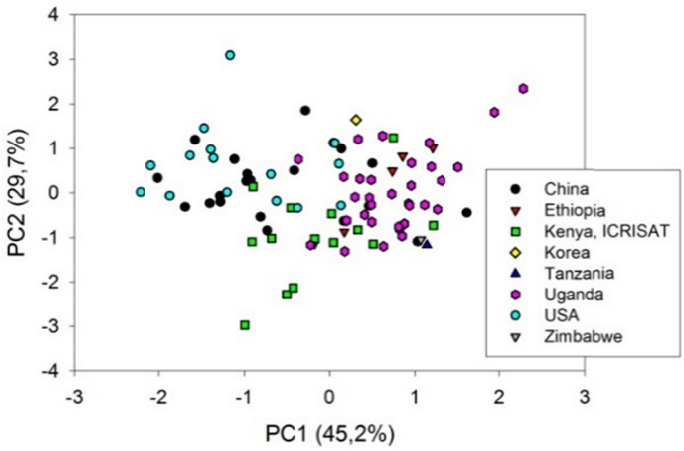

Figure 5. Principle component analysis based on 10 quantitative agromorphological traits of all 85 entries

3.3 Outcrossing Rates of Sesame

Various outcrossing rates have already been described in sesame, ranging from less than 1 per cent to nearly 70 per cent (Yermanos, 1980; Pathirana, 1994; Andrade et al., 2014). Such a high degree of variation necessitates studies under local conditions in order to clarify the outcrossing rate in Uganda. From the seven different mother plants investigated, the progeny showed varying degrees of introgression, most likely by means of foreign pollination. The level ranged from 3.1 per cent (one seedling out of 32 of mother plant B), 12.5 per cent (one seedling out of eight of each of the mother plants A, C, D, and E), 37.5 per cent (12 seedlings out of 32 of F), to 59.4 per cent (19 seedlings out of 32 of G) resulting in a mean value of 28.1 per cent (Table 6). It emphasizes the critical importance of controlled self-pollination or propagation of basic seed material under isolated conditions in order to preserve genetic purity of a line and thereby to maintain a sesame germplasm collection.

Table 6. Outcrossing values by testing 128 individuals from seven mother plants at the nine microsatellite loci

Mother plant	Affected marker	Total individuals	Affected individuals	Outcrossing [%]
A	2	8	1	12.5
B	2	32	1	3.1
C	1	8	1	12.5
D	4	8	1	12.5
E	2	8	1	12.5
F	8	32	12	37.5
G	7	32	19	59.4

3.4 Proposed Selection of Well-Performing Entries

The generated genetic and phenotypic datasets will serve as a valuable knowledge base for the selection of superior genetic material. In order to assemble a selection, the entries were chosen as described above based on their quantitative traits, their overall hairiness, and their genetic diversity.

Taking all 10 quantitative traits into consideration, the best performing entries are 008, 010, and 017 originating from Uganda, 004 from Ethiopia, 035 from Tanzania, 087 from Zimbabwe, and 044 from China (characterised by at least five traits above the single positive standard deviation; marked in Figure 1 with an asterisk (Appendix 2). Based on the same scheme, the least performing entries (characterised with at least five traits below the single negative standard deviation) are coming from China (040, 042, and 050) and USA (057, 067, 068, and 070). The 13 most hairy entries with sums above the positive standard deviation value (sum > 6.5) are composed of six entries from the USA, four from Uganda, two from China, and one from Ethiopia (marked in Figure 1 with a triangle). The genetically most diverse entries representing H_E values above the positive standard deviation value ($H_E \geq 0.31$) are 010 and 012 originating from Uganda; 043 and 055 from China; 058 from USA; 073, 077, 082 and 085 from Kenya, and 087 from Zimbabwe (marked in Figure 1 with an upward arrow). The entries 020 (Uganda), 055 (China), 058 and 060 (both USA), are characterised by both, hairiness and high genetic diversity, whereas the entries 010 (Uganda) and 087 (Zimbabwe) are characterised by the combination of good quantitative trait performance and high genetic diversity. The combination of hairiness and good quantitative trait performance is given in the entries 004 (Ethiopia) and 008 (Uganda). Summarized, a core selection composed of 26 entries is suggested (Table 7).

Table 7. Proposed selection of 26 entries

Entry No.	Entry name	Country of origin	Gene diversity (D, HE)	Hairiness (sum)	Quantitative traits1
004	Local 158-2	Ethiopia	0.12*	7*	PH, HFB, NCMS, NCB, TNC
008	EM 15-3-2	Uganda	0.29*	7*	PH, HFC, HFB, NB, NCB, TNC
010	AD-1-1-1	Uganda	0.38*	6*	PH, NB, LCZ, NCMS, NCB, TNC
012	U 1-7-1	Uganda	0.35*	6*	DTF, DTM
017	Ajimo A1-6 × 7029-1-1	Uganda	0.17*	4*	PH, HFC, LCZ, NCMS, NCB, TNC
020	U1-7 × Eva-1-1-1	Uganda	0.30*	7*	NCMS, TNC
021	Y-1 × Local 158-1-1 green	Uganda	0.29*	7*	DTF, DTM, NB
025+	Oyam hairy	Uganda	0.29*	7*	DTF, DTM, PH, HFC, HFB
035	NAL 79.309.16.1	Tanzania	0.25*	3*	DTF, DTM, PH, HFC, HFB, NB
042	4036-1-2-2-4	China	0.30*	5*	
043	4036-1-2-3	China	0.37*	5*	
044	4036-1-2-4	China	0.14*	5*	DTF, DTM, PH, HFC, HFB, NB
045	4036-1-10-2-1	China	0.20*	8*	
055	1438-1-10-3	China	0.36*	7*	LCZ, NCB
058	Renner 1-3-1-10-2	USA	0.33*	8*	
059	Renner 1-3-1-14	USA	0.27*	9*	
060	Renner 1-3-1-15	USA	0.30*	7*	LCZ
061	Renner 1-3-1-16	USA	0.28*	7*	NB
064	UCR 3-1-2-1	USA	0.30*	8*	LCZ, NCMS
066	UCR 3-1-2-3	USA	0.25*	7*	
073	ICEASE 00001	Kenya	0.35*	4*	DTF, DTM, NB
077	ICEASE 00005	Kenya	0.31*	4*	DTF, DTM, PH, NB
082	ICEASE 00018	Kenya	0.35*	4*	
084	ICEASE 00020	Kenya	0.30*	4*	HFB
085	ICEASE 00022	Kenya	0.35*	3*	NB
087	OTIS	Zimbabwe	0.32*	3*	DTF, DTM, HFC, HFB, NB

Note. *Values above the positive standard deviation ($H_E > 0.31$; hairiness sum > 6.5).

[1]Only traits with values above the positive standard deviation are shown. Number of branches (NB), number of capsules on main stem (NCMS), number of capsules on branches (NCB), and total number of capsules (TNC), days to flowering (DTF), days to maturity (DTM), plant height (PH), plant height until first capsule (HFC), plant height until first branch (HFB), and length of capsule zone (LCZ).

[+]Entry number 025: only one measurement has been done per trait, thus, the values of this entry should be taken with care.

4. Conclusion

Presence of genetic variability in crops is essential for its further improvement by providing opportunities for the breeders to develop new varieties and hybrids. Existing variation in part of the sesame germplasm conserved at NaSARRI in Uganda comprising 85 lines stemming from eight countries of origin was categorized through phenotypic (quantitative and qualitative) and genetic characterization. Despite a rather low genetic diversity (H_E grand mean = 0.219), we detected a strong admixture within and between the entries, which could be the result of the concerted action of several causes such as a differing ancestry (most likely due to the breeding process itself, but also due to cross-pollination) or due to material exchange between locations. Thus, if the maintenance of the genetic integrity of germplasm is attempted, causes of gene flow must be prevented where possible. On the basis of the phenotypic and genetic characterisation, we defined a core selection of 26 superior entries characterised by high genetic diversity, hairiness, and overall good performance of quantitative agromorphological traits. These entries form a valuable repertoire of the sesame germplasm to be used by breeders and farmers in Uganda.

Acknowledgements

The authors thank Dr. Said N. Silim (ICRISAT), Dr. NVPR Ganga Rao (ICRISAT), Dr. Patrick Okori (ICRISAT), and Dr. G.W. Otim-Nape (African Innovations Institute, Uganda) for their scientific support. This work was financially supported by the Austrian Development Agency (ADA) and the AIT Austrian Institute of Technology.

References

Alemu, A., Petros, Y., & Tesfaye, K. (2013). Genetic Distance of Sesame (*Sesamum indicum* L.) Cultivars and Varieties from Northwestern Ethiopia using Inter Simple Sequence Repeat Markers. *East African Journal of Sciences, 7*(1). https://doi.org/10.5897/IJBMBR2014.0219

Anderson, E. C., & Dunham, K. K. (2008). The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources, 8*(6), 1219-1229. https://doi.org/10.1111/j.1755-0998.2008.02355.x

Andrade, P. B. de, Freitas, B. M., Rocha, E. E. de M., Lima, J. A. de, & Rufino, L. L. (2014). Floral biology and pollination requirements of sesame (*Sesamum indicum* L.). *Acta Scientiarum. Animal Sciences, 36*, 93-99. https://doi.org/10.4025/actascianimsci.v36i1.21838

Ansah, K. D., Sintim, H. O., Awuah, S., Ali, J. E., & Oteng, G. (2015). The Adoption of Temperate Selected Sesame Accessions in the Tropics: Selected for Japan and Grown in Ghana. *Journal of Agricultural Science, 7*(5). https://doi.org/http://dx.doi.org/10.5539/jas.v7n5p47

Arriel, N. H. C., Di Mauro, A. O., Arriel, E. F., Unêda-Trevisoli, S. H., Costa, M. M., Bárbaro, I. M., & Muniz, F. R. S. (2007). Genetic divergence in sesame based on morphological and agronomic traits. *Crop Breeding and Applied Biotechnology*, 253-261.

Bedigian, D. (2003). Evolution of sesame revisited: Domestication, diversity and prospects. *Genetic Resources and Crop Evolution, 50*(7), 779-787. https://doi.org/10.1023/A:1025029903549

Bedigian, D., & Harlan, J. (1986). Evidence for cultivation of sesame in the ancient world. *Economic Botany, 40*(2), 137-154. https://doi.org/10.1007/BF02859136

Cho, Y.-I., Park, J.-H., Lee, C.-W., Ra, W.-H., Chung, J.-W., Lee, J.-R., … Park, Y.-J. (2011). Evaluation of the genetic diversity and population structure of sesame (*Sesamum indicum* L.) using microsatellite markers. *Genes & Genomics, 33*(2), 187-195. https://doi.org/10.1007/s13258-010-0130-6

de Vicente, M. (Ed.). (2005). *Gene flow and germplasm management.* Rome: International Plant Genetic Resources Institute.

Dossa, K., Wei, X., Zhang, Y., Fonceka, D., Yang, W., Diouf, D., … Zhang, X. (2016). Analysis of Genetic Diversity and Population Structure of Sesame Accessions from Africa and Asia as Major Centers of Its Cultivation. *Genes, 7*(4). https://doi.org/10.3390/genes7040014

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

Falusi, O. A., Yahaya, S. A., Gado, A. A., Daudu, O. A. Y., Akinbo, O. A., & Teixeira da Silva, J. A. (2015). Morphological evaluation of selected sesame (*Sesamum indicum* L.) genotypes from five states in Northern Nigeria. *African Journal of Agricultural Research, 10*(37), 3657-3661. https://doi.org/10.5897/AJAR2014.8928

FAO. (2015). *FAOSTAT statistics database*. Retrieved July 4, 2016, from http://faostat3.fao.org

Flint-Garcia, S. A. (2013). Genetics and Consequences of Crop Domestication. *Journal of Agricultural and Food Chemistry, 61*(35), 8267-8276. https://doi.org/10.1021/jf305511d

Francis, R. M. (2016). pophelper: An r package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, n/a-n/a. https://doi.org/10.1111/1755-0998.12509

Gebremichael, D. E., & Parzies, H. K. (2011). Genetic variability among landraces of sesame in Ethiopia. *African Crop Science Journal, 19*(1). https://doi.org/http://dx.doi.org/10.4314/acsj.v19i1.68656

Goldberg, C. S., & Waits, L. P. (2010). Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. *Molecular Ecology Resources, 10*(2), 304-313. https://doi.org/10.1111/j.1755-0998.2009.02755.x

Govindaraj, M., Vetriventhan, M., & Srinivasan, M. (2015). Importance of Genetic Diversity Assessment in Crop Plants and Its Recent Advances: An Overview of Its Analytical Perspectives. *Genetics Research International, 2015*, 14. https://doi.org/http://dx.doi.org/10.1155/2015/431487

Govindaraju, D. R. (1989). Variation in gene flow levels among predominantly self-pollinated plants. *Journal of Evolutionary Biology, 2*(3), 173-181. https://doi.org/10.1046/j.1420-9101.1989.2030173.x

IPGRI, & NBPGR. (2004). *Descriptors for Sesame (Sesamum spp.)*. International Plant Genetic Resources Institute, Rome, Italy; and National Bureau of Plant Genetic Resources, New Delhi, India.

Kim, D. H., Zur, G., Danin-Poleg, Y., Lee, S. W., Shim, K. B., Kang, C. W., & Kashi, Y. (2002). Genetic relationships of sesame germplasm collection as revealed by inter-simple sequence repeats. *Plant Breeding, 121*(3), 259-262. https://doi.org/10.1046/j.1439-0523.2002.00700.x

Kumar, H., Kaur, G., & Banga, S. (2012). Molecular Characterization and Assessment of Genetic Diversity in Sesame (*Sesamum indicum* L.) Germplasm Collection Using ISSR Markers. *Journal of Crop Improvement, 26*(4), 540-557. https://doi.org/10.1080/15427528.2012.660563

Laurentin, H. E., & Karlovsky, P. (2006). Genetic relationship and diversity in a sesame (*Sesamum indicum* L.) germplasm collection using amplified fragment length polymorphism (AFLP). *BMC Genetics, 7*, 10-10. https://doi.org/10.1186/1471-2156-7-10

Li, Y.-C., Korol, A. B., Fahima, T., & Nevo, E. (2004). Microsatellites within Genes: Structure, Function, and Evolution. *Molecular Biology and Evolution, 21*(6), 991-1007. https://doi.org/10.1093/molbev/msh073

Nyongesa, B., Were, B., Gudu, S., Dangasuk, O., & Onkware, A. (2013). Genetic diversity in cultivated sesame (*Sesamum indicum* L.) and related wild species in East Africa. *Journal of Crop Science and Biotechnology, 16*(1), 9-15. https://doi.org/10.1007/s12892-012-0114-y

Pandey, S. K., Das, A., Rai, P., & Dasgupta, T. (2015). Morphological and genetic diversity assessment of sesame (*Sesamum indicum* L.) accessions differing in origin. *Physiology and Molecular Biology of Plants, 21*(4), 519-529. https://doi.org/10.1007/s12298-015-0322-2

Parsaeian, M., Mirlohi, A., & Saeidi, G. (2011). Study of genetic variation in sesame (*Sesamum indicum* L.) using agro-morphological traits and ISSR markers. *Russian Journal of Genetics, 47*(3), 314-321. https://doi.org/10.1134/S1022795411030136

Pathirana, R. (1994). Natural Cross-Pollination in Sesame (*Sesamum indicum* L.). *Plant Breeding*, *112*(2), 167-170. https://doi.org/10.1111/j.1439-0523.1994.tb00665.x

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics*, *28*(19), 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Peterman, W., Brocato, E. R., Semlitsch, R. D., & Eggert, L. S. (2016). Reducing bias in population and landscape genetic inferences: The effects of sampling related individuals and multiple life stages. *PeerJ, 4*, e1813. https://doi.org/10.7717/peerj.1813

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945-959.

Proctor, K. (2015). *More Than Markets - Building Resilience in Northern Uganda*. MercyCorps. Retrieved from https://www.mercycorps.org

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Roy, B., & Basu, A. K. (2009). *Abiotic Stress Tolerance in Crop Plants: Breeding and Biotechnology*. New India Pub. Agency. Retrieved from https://books.google.at/books?id=kbTbrDSk7hkC

Rubaihayo, P. R., & Okello-Anyanga, W. (1997). Genetic variation and interrelations among characters of sesame. *African Crop Science Conference Proceedings, 3*(1), 175-181.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nat Biotech, 18*(2), 233-234. https://doi.org/10.1038/72708

Sehr, E. M., Okello-Anyanga, W., Hansel-Hohl, K., Burg, A., Gaubitzer, S., Rubaihayo, P. R., … Fluch, S. (2016). Assessment of genetic diversity amongst Ugandan sesame (*Sesamum indicum* L.) landraces based

on agromorphological traits and genetic markers. *Journal of Crop Science and Biotechnology*. https://doi.org/10.1007/s12892-015-0105-x

Sharma, S. N., Kumar, V., & Mathur, S. (2009). Comparative Analysis of RAPD and ISSR Markers for Characterization of Sesame (*Sesamum indicum* L.) Genotypes. *Journal of Plant Biochemistry and Biotechnology, 18*(1), 37-43. https://doi.org/10.1007/BF03263293

Stierschneider, M., Gaubitzer, S., Schmidt, J., Weichselbaum, O., Kopecky, D., Kremer, A., … Sehr, E. M. (2016). The Evoltree Repository Centre—A central access point for reference material and data of forest genetic resources. *Evolution of Trees and Forest Communities: Ten years of the Evoltree network* (pp. 15-19). Bordeaux: PG Edition.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution, 30*(12), 2725-2729. https://doi.org/10.1093/molbev/mst197

Tanksley, S. D., & McCouch, S. R. (1997). Seed Banks and Molecular Maps: Unlocking Genetic Potential from the Wild. *Science, 277*(5329), 1063. https://doi.org/10.1126/science.277.5329.1063

van der Beek, J. G., Verkerk, R., Zabel, P., & Lindhout, P. (1992). Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: Cf9 (resistance to *Cladosporium fulvum*) on chromosome 1. *Theoretical and Applied Genetics, 84*(1-2), 106-112. https://doi.org/10.1007/BF00223988

van Rheenen, H. A. (1972). *Major problems of growing sesame (Sesamum indicum L.) in Nigeria*. Wageningen University. Retrieved from http://library.wur.nl/WebQuery/wurpubs/421721

Woldesenbet, D. T., Tesfaye, K., & Bekele, E. (2015). Genetic diversity of sesame germplasm collection (*Sesamum indicum* L.): implication for conservation, improvement and use. *International Journal of Biotechnology and Molecular Biology Research, 6*(2), 7-18. https://doi.org/10.5897/IJBMBR2014.0219

Yermanos, D. M. (1980). Sesame. *Hybridization of Crop Plants* (pp. 549-563). Madison, WI: American Society of Agronomy, Crop Science Society of America. Retrieved from http://dx.doi.org/10.2135/1980.hybridizationofcrops.c39

Zhang, Y., Wang, L., Xin, H., Li, D., Ma, C., Ding, X., … Zhang, X. (2013). Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. *BMC Plant Biology, 13*, 141-141. https://doi.org/10.1186/1471-2229-13-141

Appendix

Appendix 1. List of investigated sesame entries

AIT entry no.	Entry name	Country of origin	Type of material	Number of individual samples	Number of remaining samples in reduced dataset*
001	Sesim 1	Uganda	Released 2001, local selection from EM 14	16	2
002	Sesim 2	Uganda	Released 2001, local selection from EM 15	16	1
003	Local 158-1	Ethiopia	Breeding line	16	3
004	Local 158-2	Ethiopia	Breeding line	16	7
005	Local 158-3	Ethiopia	Breeding material	16	3
006	Local 158-4	Ethiopia	Breeding material	16	9
007	Local 158-5	Ethiopia	Breeding material	16	5
008	EM 15-3-2	Uganda	Breeding material	16	13
009	Adong 4-4	Uganda	Breeding material	16	6
010	AD-1-1-1	Uganda	Breeding material	16	12
011	U 1-7	Uganda	Breeding material	16	10
012	U 1-7-1	Uganda	Breeding material	16	15
013	EM 15-1-5	Uganda	Breeding material	16	9

014	Ajimo A1-5	Uganda	Breeding material	16	11
015	Ajimo A1-6	Uganda	Breeding material	16	5
016	Sesim 2 × 5181-2-2-1	Uganda	Selected line from a cross	16	11
017	Ajimo A1-6 × 7029-1-1	Uganda	Selected line from a cross	16	7
018	Local 158 × 7029-7	Uganda	Selected line from a cross	16	12
019	Local 158 × 6022-1-2-1	Uganda	Selected line from a cross	16	5
020	U1-7 × Eva-1-1-1	Uganda	Selected line from a cross	16	14
021	Y-1 × Local 158-1-1 green	Uganda	Selected line from a cross	16	10
022	Y-1 × Local 158-1-2-1 purple	Uganda	Selected line from a cross	16	6
023	(Y-1 × Local 158-1-2-1) × Renner 1-3-1	Uganda	Selected line from a cross	16	9
024	(Y-1 × Local 158-1-2-1) × Renner 1-3-2	Uganda	Selected line from a cross	16	6
025	Oyam hairy	Uganda	Introduced line, unknown origin	16	2
029	Kidetok SPSS purple	Uganda	Landrace	N/A	N/A
030	Otara Awelo-2 purple	Uganda	Landrace	N/A	N/A
034	Early Russian	Korea	Breeding line	16	11
035	NAL 79.309.16.1	Tanzania	Breeding line	16	11
036	4036-1-1-2-1	China	Breeding line	8	2
037	4036-1-1-2-5	China	Breeding line	16	9
038	4036-1-1-2-6	China	Breeding line	16	12
039	4036-1-2-2-1	China	Breeding line	16	13
040	4036-1-2-2-2	China	Breeding line	16	4
041	4036-1-2-2-3	China	Breeding line	16	5
042	4036-1-2-2-4	China	Breeding line	16	13
043	4036-1-2-3	China	Breeding line	16	16
044	4036-1-2-4	China	Breeding line	16	8
045	4036-1-10-2-1	China	Breeding line	16	5
046	1438-1-6-1	China	Breeding line	16	9
047	1438-1-6-1-1	China	Breeding line	16	6
048	1438-1-6-2	China	Breeding line	16	14
049	1438-1-6-3	China	Breeding line	16	14
050	1438-1-6-10-2	China	Breeding line	16	14
051	1438-1-6-18-1	China	Breeding line	16	3
052	1438-1-10-1-1	China	Breeding line	16	6
053	1438-1-10-2-1	China	Breeding line	16	12
054	1438-1-10-2-2	China	Breeding line	16	3
055	1438-1-10-3	China	Breeding line	16	9
056	Renner 1-1-1-6-2	USA	Breeding line	16	8
057	Renner 1-3-1-4	USA	Breeding line	16	4
058	Renner 1-3-1-10-2	USA	Breeding line	16	13
059	Renner 1-3-1-14	USA	Breeding line	16	9
060	Renner 1-3-1-15	USA	Breeding line	16	13
061	Renner 1-3-1-16	USA	Breeding line	16	11
062	Renner 1-3-1-17	USA	Breeding line	16	2
063	Renner 1-3-1-19	USA	Breeding line	16	8
064	UCR 3-1-2-1	USA	Breeding line	16	8
065	UCR 3-1-2-2	USA	Breeding line	16	13
066	UCR 3-1-2-3	USA	Breeding line	16	11
067	UCR 3-1-4	USA	Breeding line	16	1
068	UCR 3-1-5	USA	Breeding line	16	6

069	UCR 3-1-6	USA	Breeding line	16	8
070	UCR 3-1-7	USA	Breeding line	16	10
071	UCR 3-1-8	USA	Breeding line	16	10
072	Ajimo A1-6 × 7029-1-9	Uganda	Selected line from a cross	16	12
073	ICEASE 00001	Kenya, ICRISAT	Breeding line	16	11
074	ICEASE 00002	Kenya, ICRISAT	Breeding line	16	10
075	ICEASE 00003	Kenya, ICRISAT	Breeding line	16	4
076	ICEASE 00004	Kenya, ICRISAT	Breeding line	16	2
077	ICEASE 00005	Kenya, ICRISAT	Breeding line	16	6
078	ICEASE 00006	Kenya, ICRISAT	Breeding line	16	1
079	ICEASE 00007	Kenya, ICRISAT	Breeding line	16	3
080	ICEASE 00008	Kenya, ICRISAT	Breeding line	16	4
081	ICEASE 00015	Kenya, ICRISAT	Breeding line	16	10
082	ICEASE 00018	Kenya, ICRISAT	Breeding line	16	14
083	ICEASE 00019	Kenya, ICRISAT	Breeding line	16	5
084	ICEASE 00020	Kenya, ICRISAT	Breeding line	16	13
085	ICEASE 00022	Kenya, ICRISAT	Breeding line	16	7
086	ICEASE 00096	Kenya, ICRISAT	Breeding line	16	4
087	OTIS	Zimbabwe	Breeding line	16	9
098	Local Sesim 2	Uganda	Selection of Sesim 2, distributed to farmers (2004-07)	16	11
114	Local Sesim 2	Uganda	Selection of Sesim 2, distributed to farmers (2004-07)	16	8
118	Local Sesim 2	Uganda	Selection of Sesim 2, distributed to farmers (2004-07)	16	1
191	Local Sesim 2	Uganda	Selection of Sesim 2, distributed to farmers (2004-07)	16	4

Note. *Repeated matching multilocus genotypes calculated by GenAlEx were removed from the data set resulting in 666 remaining samples.

Appendix 2. Heterozygosity, F-statistics and polymorphism are presented per locus

	CL297 Contig1	CL569 Contig1	CL78 Contig1	CL93 Contig1	GBssr_sa_08	GBssr_sa_108	GBssr_sa_123	GBssr_sa_184	GBssr_sa_72
N	657	647	656	649	653	645	658	660	625
N_A	2	4	2	3	7	15	14	11	3
N_E	1.03	1.86	1.98	1.06	1.88	4.86	5.79	4.26	1.09
H_O	0.00	0.16	0.15	0.03	0.17	0.30	0.31	0.27	0.04
H_E	0.03	0.46	0.50	0.06	0.47	0.79	0.83	0.77	0.08
F	0.91	0.65	0.69	0.44	0.64	0.62	0.63	0.65	0.56

Note. Sample size (N), number of alleles (N_A), number of effective alleles (N_E), observed heterozygosity (H_O), expected heterozygosity (H_E), and the fixation index (F) calculated using GenAlEx.

Appendix 3. Heterozygosity, F-statistics and polymorphism are presented as mean values per entry

Entry	N	N_A	N_E	H_O	H_E	F
Line001_Sesim1	2.00	1.22	1.22	0.11	0.11	0.00
Line002_Sesim2	0.89	0.89	0.89	0.00	0.00	
Line003_Local158-1	2.89	1.33	1.27	0.22	0.15	-0.50
Line004_Local158-2	6.78	1.44	1.19	0.16	0.12	-0.23
Line005_Local158-3	2.78	1.78	1.59	0.19	0.30	0.29
Line006_Local158-4	9.00	1.78	1.50	0.14	0.26	0.46
Line007_Local158-5	5.00	1.56	1.28	0.07	0.18	0.61

Line008_EM15-3-2	12.78	1.78	1.52	0.21	0.29	0.20
Line009_Adong4-4	5.67	1.44	1.39	0.17	0.21	0.21
Line010_AD-1-1-1	11.67	2.89	2.01	0.30	0.38	0.16
Line011_U1-7	9.67	2.22	1.48	0.23	0.26	0.03
Line012_U1-7-1	14.89	2.00	1.80	0.17	0.35	0.47
Line013_EM15-1-5	8.67	1.78	1.62	0.08	0.27	0.61
Line014_AjimoA1-5	10.44	1.78	1.63	0.09	0.26	0.57
Line015_AjimoA1-6	5.00	1.44	1.31	0.04	0.17	0.60
Line016_(Sesim2×5181)-2-2-1	10.67	1.78	1.40	0.19	0.22	0.11
Line017_AjimoA1-6×7029-1-1	6.78	1.44	1.31	0.17	0.17	0.01
Line018_Local158×7029-7	11.67	1.78	1.65	0.20	0.28	0.28
Line019_(Local158×6022)-1-2-1	4.67	1.33	1.15	0.13	0.10	-0.23
Line020_(U1-7×Eva)-1-1-1	13.67	2.00	1.63	0.11	0.30	0.47
Line021_(Y-1×Local158)-1-1green	9.56	2.00	1.60	0.16	0.29	0.37
Line022_(Y-1×Local158)-1-2-1purple	5.78	1.56	1.28	0.13	0.16	0.22
Line023_(Y-1×Local158)-1-2-1×Renner1-3-1	8.33	1.56	1.43	0.11	0.21	0.54
Line024_(Y-1×Local158)-1-2-1×Renner1-3-2	6.00	1.33	1.18	0.07	0.11	0.24
Line025_Oyamhairy	2.00	1.78	1.56	0.28	0.29	-0.02
Line034_EarlyRussian	10.89	1.44	1.33	0.15	0.18	0.14
Line035_NAL79.309.16.1	11.00	1.78	1.46	0.11	0.25	0.45
Line036_4036-1-1-2-1	1.89	1.00	1.00	0.00	0.00	
Line037_4036-1-1-2-5	8.67	1.56	1.53	0.24	0.27	0.11
Line038_4036-1-1-2-6	11.78	1.67	1.44	0.19	0.23	0.18
Line039_4036-1-2-2-1	12.89	1.56	1.44	0.09	0.21	0.58
Line040_4036-1-2-2-2	4.00	1.44	1.30	0.03	0.18	0.87
Line041_4036-1-2-2-3	4.78	1.44	1.34	0.11	0.19	0.36
Line042_4036-1-2-2-4	12.33	1.89	1.73	0.11	0.30	0.62
Line043_4036-1-2-3	15.78	3.00	1.98	0.38	0.37	-0.03
Line044_4036-1-2-4	7.78	1.44	1.25	0.10	0.14	0.26
Line045_4036-1-10-2-1	5.00	1.78	1.35	0.24	0.20	-0.19
Line046_1438-1-6-1	8.89	2.00	1.61	0.15	0.28	0.42
Line047_1438-1-6-1-1	5.89	1.67	1.24	0.11	0.15	0.12
Line048_1438-1-6-2	13.78	1.89	1.70	0.21	0.28	0.25
Line049_1438-1-6-3	13.56	1.56	1.35	0.17	0.18	0.02
Line050_1438-1-6-10-2	13.78	2.00	1.50	0.17	0.23	0.20
Line051_1438-1-6-18-1	2.89	1.33	1.27	0.00	0.15	1.00
Line052_1438-1-10-1-1	5.78	1.33	1.28	0.04	0.15	0.75
Line053_1438-1-10-2-1	11.89	1.44	1.37	0.17	0.17	0.03
Line054_1438-1-10-2-2	2.89	1.44	1.17	0.15	0.12	-0.20
Line055_1438-1-10-3	9.00	2.11	1.86	0.19	0.36	0.51
Line056_Renner1-1-1-6-2	7.67	2.22	1.52	0.21	0.25	0.14
Line057_Renner1-3-1-4	4.00	1.56	1.41	0.19	0.20	0.05
Line058_Renner1-3-1-10-2	12.33	2.44	1.75	0.20	0.33	0.28
Line059_Renner1-3-1-14	8.89	2.00	1.62	0.14	0.27	0.34
Line060_Renner1-3-1-15	12.67	1.89	1.62	0.16	0.30	0.40
Line061_Renner1-3-1-16	10.78	1.89	1.57	0.24	0.28	0.19
Line062_Renner1-3-1-17	2.00	1.44	1.44	0.00	0.22	1.00
Line063_Renner1-3-1-19	7.89	1.56	1.38	0.14	0.17	0.24
Line064_UCR3-1-2-1	8.00	1.89	1.63	0.13	0.30	0.51
Line065_UCR3-1-2-2	13.00	1.78	1.65	0.24	0.24	-0.07
Line066_UCR3-1-2-3	10.78	1.67	1.46	0.09	0.25	0.62
Line067_UCR3-1-4	1.00	1.00	1.00	0.00	0.00	

Line068_UCR3-1-5	5.89	2.22	1.52	0.28	0.29	0.01
Line069_UCR3-1-6	7.89	1.33	1.22	0.07	0.13	0.48
Line070_UCR3-1-7	9.78	1.89	1.47	0.16	0.25	0.37
Line071_UCR3-1-8	10.00	1.78	1.57	0.24	0.28	0.06
Line072_AjimoAl6×7029-1-9	11.89	1.89	1.60	0.21	0.27	0.23
Line073_ICEASE00001	10.78	1.89	1.77	0.14	0.35	0.61
Line074_ICEASE00002	9.67	2.11	1.52	0.12	0.20	0.18
Line075_ICEASE00003	4.00	1.22	1.16	0.03	0.09	0.73
Line076_ICEASE00004	2.00	1.33	1.33	0.00	0.17	1.00
Line077_ICEASE00005	5.89	1.89	1.60	0.13	0.31	0.44
Line078_ICEASE00006	1.00	1.00	1.00	0.00	0.00	
Line079_ICEASE00007	2.78	1.00	1.00	0.00	0.00	
Line080_ICEASE00008	3.89	1.44	1.36	0.03	0.19	0.71
Line081_ICEASE00015	9.67	1.78	1.65	0.10	0.28	0.66
Line082_ICEASE00018	13.67	2.11	1.77	0.14	0.35	0.51
Line083_ICEASE00019	4.78	1.78	1.45	0.14	0.25	0.27
Line084_ICEASE00020	13.00	1.89	1.63	0.16	0.30	0.42
Line085_ICEASE00022	6.89	2.67	1.67	0.49	0.35	-0.37
Line086_ICEASE00096	3.78	1.44	1.44	0.00	0.22	1.00
Line087_OTIS	9.00	1.89	1.70	0.11	0.32	0.65
Line098_LocalSesim2	10.00	1.67	1.55	0.08	0.30	0.72
Line114_LocalSesim2	7.78	1.67	1.48	0.18	0.24	0.34
Line118_LocalSesim2	1.00	1.00	1.00	0.00	0.00	
Line191_LocalSesim2	3.67	1.44	1.29	0.00	0.17	1.00

Note. Sample size (N), number of alleles (N_A), number of effective alleles (N_E), observed heterozygosity (H_E), expected heterozygosity (H_E), and the fixation index (F) calculated as mean values using GenAlEx.

Appendix 4. Values of the qualitative agromorphological traits are listed

Entry	Branching habit	Capsule length	Capsule pubescence	Capsule width	Corolla pubescence	Flowers per leaf axil	Leaf blade colour	Leaf blade length	Leaf blade width	Petiole length	Position of branches	Stem colour	Stem fasciation	Stem pubescence	Sum pubescence
001	3	1	2	1	3	1	4	2	2	2	3	2	1	1	6
002	3	1	1	2	3	1	2	2	2	2	3	6	1	1	5
003	3	2	2	2	2	1	2	2	3	2	3	3	1	3	7
004	2	2	2	2	3	1	4	2	2	2	2	1	1	2	7
005	3	1	1	2	2	1	4	2	2	2	3	3	1	3	6
006	3	2	3	2	2	1	4	3	2	2	3	3	1	3	8
007	3	1	1	2	2	1	2	3	1	2	3	5	1	1	4
008	3	2	1	2	3	1	2	3	1	1	2	6	1	1	5
009	3	2	1	2	3	1	2	3	1	1	3	5	1	1	5
010	3	2	1	2	3	1	2	3	1	2	4	5	1	1	5
011	3	2	2	2	3	1	4	3	2	2	2	5	1	1	6
012	3	1	2	2	3	1	1	2	1	1	3	5	1	1	6
013	3	2	1	2	3	1	1	2	3	2	2	5	1	1	5
014	3	1	2	2	2	1	2	2	2	2	3	5	1	1	5
015	3	2	1	2	2	1	4	3	2	2	2	5	1	1	4
016	3	2	1	2	2	1	4	3	2	2	3	2	1	2	5
017	3	2	1	2	2	1	1	2	2	2	3	2	1	1	4

018	3	1	2	2	2	1	4	2	2	2	3	1	1	2	6
019	3	2	3	2	2	1	4	1	1	1	2	2	1	1	6
020	3	1	3	2	2	1	1	2	2	1	3	5	1	3	8
021	3	2	1	2	3	1	2	2	1	1	2	1	1	1	5
022	3	1	1	2	2	1	2	2	2	1	3	6	1	1	4
023	3	2	2	2	2	1	2	2	2	2	3	2	1	1	5
024	3	2	2	2	2	1	4	3	2	3	3	1	1	2	6
025	2	2	1	2	2	1	4	2	1	1	2	2	1	3	6
029	3	2	1	2	2	1	1	3	2	3	4	2	1	1	4
030	3	2	1	2	2	1	4	2	2	2	4	5	1	1	4
034	3	2	2	2	2	1	4	2	1	2	3	3	1	2	6
035	3	1	2	2	1	1	1	2	1	2	2	5	1	1	4
036	2	1	3	2	1	1	1	2	2	1	3	2	1	3	7
037	3	1	2	2	1	1	1	2	1	1	4	2	1	1	4
038	2	1	1	2	2	1	1	1	1	1	3	5	1	2	5
039	3	2	3	2	2	1	1	2	1	2	3	3	1	2	7
040	2	1	3	2	2	1	2	2	1	3	3	3	1	2	7
041	2	2	3	2	2	1	4	1	1	1	3	1	1	2	7
042	2	1	2	2	1	1	2	2	2	2	3	3	1	3	6
043	2	2	1	2	2	1	4	1	1	1	3	3	1	1	4
044	3	1	2	2	1	1	2	1	1	1	2	3	1	2	5
045	3	2	2	2	2	1	2	3	1	1	3	5	1	3	7
046	2	2	1	3	3	1	4	2	1	1	2	5	1	2	6
047	3	1	1	3	1	1	1	1	1	2	4	2	1	3	5
048	2	2	2	3	1	1	2	2	1	1	3	5	1	3	6
049	3	1	1	3	1	1	4	1	2	1	3	5	1	2	4
050	2	1	3	3	1	1	2	1	2	1	3	5	1	2	6
051	3	1	2	3	2	1	2	1	2	1	2	2	1	1	5
052	3	1	2	3	1	1	1	1	2	2	3	2	1	1	4
053	3	1	1	1	2	1	4	1	2	1	3	3	1	1	4
054	2	2	1	2	1	1	4	1	2	2	1	2	1	1	3
055	3	1	2	2	3	1	1	1	1	1	3	3	1	2	7
056	3	2	3	2	2	1	4	2	1	2	3	2	1	1	6
057	2	1	1	2	2	1	2	1	1	1	3	2	1	2	5
058	3	1	1	2	2	1	4	1	1	1	2	3	1	3	6
059	2	1	1	2	3	1	1	1	1	1	3	2	1	3	7
060	3	2	1	2	3	1	4	1	1	2	3	2	1	3	7
061	3	2	1	2	2	1	1	1	1	2	4	3	1	3	6
062	3	1	1	2	2	1	1	2	2	2	3	3	1	1	4
063	3	1	1	2	2	1	2	1	2	2	3	5	1	1	4
064	3	2	2	2	2	1	4	1	1	1	3	5	1	3	7
065	3	2	1	2	2	1	1	1	1	1	1	2	1	2	5
066	3	1	2	2	2	1	4	2	2	2	1	5	1	3	7
067	2	1	2	2	1	1	1	1	1	1	3	5	1	2	5
068	2	1	2	2	1	1	2	1	1	2	3	3	1	3	6
069	3	2	3	2	2	1	4	1	1	1	4	5	1	2	7
070	2	1	1	2	1	1	1	1	1	1	3	5	1	2	4
071	2	1	1	2	1	1	1	1	1	1	1	5	1	3	5
072	3	1	2	2	1	1	4	1	1	1	1	3	1	1	4
073	3	1	2	2	2	1	2	1	1	1	2	5	1	1	5
074	3	1	2	2	1	1	1	1	1	1	3	3	1	2	5
075	3	1	1	2	1	1	1	1	1	1	2	1	1	1	3

076	3	2	2	2	2	1	2	1	1	1	4	1	1	1	5
077	3	3	2	2	1	1	1	1	1	2	2	3	1	1	4
078	3	2	1	2	1	1	1	1	1	1	3	1	1	1	3
079	3	2	2	2	2	1	1	1	1	1	2	1	1	1	5
080	3	2	1	2	2	1	4	1	1	1	4	3	1	1	4
081	3	2	1	2	1	1	4	1	1	1	3	5	1	2	4
082	3	1	1	3	1	1	4	1	1	1	3	2	1	1	3
083	3	1	2	3	2	1	1	1	1	1	3	2	1	1	5
084	3	1	2	3	2	1	1	1	1	1	3	2	1	1	5
085	3	1	1	3	1	1	1	1	1	1	3	2	1	1	3
086	2	1	2	3	2	1	1	2	3	2	2	3	1	1	5
087	2	1	2	3	1	1	1	1	1	1	2	6	1	1	4
098	3	2	2	1	2	1	4	1	1	1	3	6	1	1	5
114	3	2	1	2	2	1	4	1	1	2	2	6	1	1	4
118	3	1	2	2	1	1	1	2	2	2	3	3	1	2	5
191	3	1	2	3	1	1	1	3	2	3	3	5	1	1	4

Appendix 5. Mean values of the quantitative agromorphological traits are listed: days to flowering (DTF), days to maturity (DTM), plant height (PH), plant height until first capsule (HFC), plant height until first branch (HFB), number of branches (NB), length of capsule zone (LCZ), number of capsules on main stem (NCMS), number of capsules on branches (NCB), and total number of capsules (TNC). *) Entry 25: only one measurement has been done per trait, thus, this entry was excluded from all further statistical analyses. Values above the single positive standard deviation are marked in green, and values below the single negative standard deviation are marked in red. Entries with at least five traits above the single positive standard deviation are marked in green. Entries with at least five traits below the single negative standard deviation are marked in red

Entry	DTF	DTM	PH [cm]	HFC [cm]	HFB [cm]	NB	LCZ [cm]	NCMS	NCB	TNC
001	46.67	89	114.67	70	51	3	43.67	13.33	13	26.33
002	47.67	90.33	115.67	78.33	60.33	2.33	33.33	12	11.67	23.67
003	47	90	104.67	70	56.67	2	35	14.33	6	20.33
004	46.33	89	121	72.67	56.33	2.33	48.33	18.33	18	36.33
005	44.33	87	111.33	71.67	56.67	2.67	38.67	18	14.67	32.67
006	44.33	86	114.67	71.33	51.33	3	42.33	16.67	18	34.67
007	44.33	86	100.33	70	58.33	2.33	33	10	7.33	17.33
008	45	87.33	129.67	81.67	60.33	3.33	46.33	15.67	17.33	33
009	47.33	90	121	75	56	3.33	45.67	15.67	15	30.67
010	46.67	89.67	129	74	49	3.67	54	20	31.33	51.33
011	49	92	111	71	52	2.67	41.67	13	14.67	27.67
012	48.33	90.67	99.67	63.33	38.48	2.87	37.33	10	15.12	20.67
013	44.67	87	107.33	67.33	49.67	3	40.33	12.67	12.33	25
014	41.33	83	122	71.33	53	2.67	48.33	16.67	17	33.67
015	45.33	87.67	116	74	51.33	3.33	40.67	13	17.33	30.33
016	44	86.67	111	70.67	58.67	2.33	43	15.67	12.67	28.33
017	44.33	86	150	86	50	3	51.33	19.33	22.33	41.67
018	47.67	90.33	110.67	69.67	56.67	2.33	39.67	13	9	22
019	47	90	115.67	71.33	59.33	2.67	44.33	18.33	16	34.33
020	42.67	84.33	111.67	63.67	49.33	2.67	48.67	17	15.33	32.67
021	47.67	91.33	108.67	68	46.67	3.67	39.33	12	16.33	28.33
022	47.33	90.67	117	76.33	57	3.67	40.67	12.67	14.67	27.33
023	45.67	88.67	125	63.67	42.33	2	53	13	8.67	21.67
024	46.67	89.33	108.67	63.33	44.33	2.33	45	14	14	28
025*	84	130	130.78	90.52	88.76	2.98	46.23	9.72	3.03	12.6
029	45.33	88.33	94.67	64	43.67	4	32.67	10.33	19.33	29.67

030	48	90.67	109.67	71.67	52.67	3	37.67	8.33	16.33	24.67
034	46.33	89.33	115	56	37.67	2.33	63.33	18.67	11.67	30.33
035	48.33	91.67	119.33	75.33	61.33	3.33	36	9.67	12.33	22
036	39.67	81.33	92.67	56	37	2.33	36.67	9.33	8	17.33
037	41.67	83.33	108.67	62	43	3.33	46.33	14.33	16	30.33
038	41.67	83.67	106	60.67	45	2.33	47.67	13	10	23
039	40.33	82	92.33	54	42.33	2.33	39	12.67	9.67	22.33
040	41.67	83.33	84	43	36.4	2.12	40.67	10.33	3.89	12.33
041	48	90.67	110.33	77.67	54	4	33	11	12	23
042	38.67	80.33	92	44.33	28.67	2	48	14.67	8	22.67
043	40.67	82.33	96.67	53	43.9	2.12	44.33	11	11.89	18.33
044	50	92.67	125.67	76.67	64.33	3.33	48.67	12	13.67	25.67
045	42.33	83.67	97.67	57	45.4	2.62	37	10.33	7.39	14.67
046	46	88.33	101	66.33	55.33	2	35.33	22.33	13	35.33
047	39.33	80.33	86	51.33	37.33	2.33	35.33	15	12.67	27.67
048	41.67	83.67	87.33	48.67	40	2	40.67	11	4.67	15.67
049	41	82.33	94	53	43	2	43.67	12.67	9.67	22.33
050	36.67	77.67	80.67	45.67	37.33	2	36	12.67	5.67	18.33
051	47.67	90.67	96.33	62.67	44.67	3	40	10.33	11.67	22
052	43.33	85.67	99.33	62.67	52	1.67	37	10.67	4.33	15
053	42.33	83.67	83.33	51	42	2	39.33	14	4.33	18.33
054	44.67	87.33	110	70.33	52.67	3.33	39.67	12.67	11	23.67
055	39.67	81.33	110.67	55.33	31.67	3	56.33	13.67	17.67	31.33
056	43.67	85.33	112.33	60.67	39.33	2.33	50.67	15.33	14.33	29.67
057	39	80.33	92.67	47	36	2	47.67	14	8.33	22.33
058	40	81	104.33	67	53.33	2	36.67	11.67	8.33	20.33
059	42.67	84	102.67	63.67	53.62	2.51	40.67	12	7.98	17.33
060	46.67	89.67	106.33	55.67	39.67	2.67	52.33	15.33	10.67	26
061	45	87.67	81.67	45	27.33	3.33	43.33	12.67	12.67	25.33
062	40.67	82.33	108.33	63.33	45.33	2.67	45.33	13.67	18.67	32.33
063	48	91	96.67	57.67	35.67	3.33	40	11.33	13	24.33
064	39.33	81	93.67	41.33	28	2.33	50.33	17.33	6.33	23.67
065	37	78	101.33	42.33	25.33	2	59.33	21.67	13.67	35.33
066	41	82.67	87.33	43	27	3	45.67	15.33	7	19
067	36.33	77.33	81.33	44	29.67	2.33	38.33	13	5.67	18.67
068	38.67	79.67	80.67	43	29.12	1.51	37.67	9.67	5.98	13.67
069	40	81.33	95.33	54.67	39.67	2.33	38.67	11.67	6.33	18
070	37	78	89.33	52.33	41.33	1.67	37.67	10	4.67	14.67
071	42	83.67	86.33	39.33	19.90	2.62	48.67	15.33	4.39	17.67
072	43.33	85	99.33	55.67	42.33	2.33	46	15.67	11.33	27
073	49	91	101	70.33	48	3.33	32.67	10.67	10.67	21.33
074	47	90	100.67	68.33	51.33	3	33.67	12.67	9.67	22.33
075	45.67	88	98.33	70.67	50.33	2.67	27.33	5	2.67	7.67
076	45.33	87.67	93.67	70.67	44.33	3.67	27	5.33	3.67	9
077	48	90.67	124.33	73.67	55.67	4	42.33	10	12.33	22.33
078	43.67	86	105	64.33	45	2.33	42.67	7	3.67	10.67
079	41	82	115	77	60	2.67	37	7.67	6.33	14
080	39.67	81.33	111	59	44	2.67	52.67	9.67	3.67	13.33
081	43	85	113.67	65.67	42	3.67	49.33	13	22.33	35.33
082	44	86.33	101.33	61	38.33	2.67	39.67	10.33	9	19.33
083	45.33	88	111.67	67.67	48.67	2.33	44	10.67	8	18.67
084	45	87	113	73.33	59	2.33	41.33	9	5	14

085	44	86.33	93.67	57.67	34.33	3.33	37	7.67	4	10
086	49.67	92.67	83.78	59.52	54.76	0.98	26.23	3.72	0	3.6
087	49.67	92	116.67	74.67	58.33	3.33	40	11	10	21
098	45	87	125	73	48.67	3	51	14.67	20.67	35.33
114	47.33	89.67	117	79.67	58.33	2.33	37	9.67	9	18.67
118	46	89	104.67	72.67	54.67	2.67	32.33	9.67	8	17.67
191	49.33	92.33	108.67	70	42.67	3.67	39	11	19	30
Mean	44.10	86.27	105.09	63.30	46.28	2.68	41.90	12.73	11.19	23.61
Standard deviation	3.43	4.02	13.23	10.92	9.81	0.61	6.98	3.48	5.43	8.09

Appendix 6. Analysis of molecular variance (AMOVA) based on nine nuclear SSR markers applied on 83 sesame entries which were grouped into eight countries of origin

Source of variation	d.f.	Sum of squares	Mean sum of squares	Percentage of variation
Among countries of origin (n = 8)	7	168.13	24.02	7%
Among individuals	658	1994.58	3.03	56%
Within individuals	666	492.43	0.74	36%
Total	1331	2655.14		100%

Note. d.f. = degrees of freedom.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).