



Molecular and phenotypic characterization of a collection of white grain sorghum [*Sorghum bicolor* (L.) Moench] for temperate climates

Lorenzo Stagnati · Matteo Busconi · Giovanna Soffritti · Michelangelo Martino ·
Alessandra Lanubile · Adriano Marocco

Received: 26 November 2020 / Accepted: 9 March 2021
© The Author(s) 2021

Abstract Sorghum [*Sorghum bicolor* (L.) Moench] is a subsistence crop and the main food for populations in arid or semiarid regions and it is appreciated for the production of gluten-free products, forages, raw materials for industrial transformation and packaging. The end-use of different sorghum purposes having various plant or kernel characteristics require specific breeding programs to develop the desired ideotype. Sorghum grains can be classified according to kernel color, tannins and polyphenols content: white, yellow, red, brown, and black. White sorghum is characterized by a low level of total phenolic content and tannins.

The advantage of using white sorghum is: increased protein digestibility, nutritional composition and consumer acceptance similar to other cereals. A collection of 117 white grain sorghums was characterized using 10 SSRs and preliminary agronomic observations were made for main traits. SSR analysis revealed from 10 to 33 alleles per locus. Observed heterozygosity was lower than expected according to the reproduction system of sorghum. Phylogenetic analysis revealed 6 main groups of genotypes. Only one group is constituted by genotypes with the same geographical origin (Egypt) while other groups are admixtures of different countries. The principal coordinate analysis revealed good correspondence between genetic profiles and groups evidenced by similar agronomic performances.

Stagnati Lorenzo, Busconi Matteo contributed equally to the work.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-021-01166-9>.

L. Stagnati · M. Busconi · G. Soffritti ·
M. Martino · A. Lanubile · A. Marocco (✉)
Department of Sustainable Crop Production, Università
Cattolica del Sacro Cuore, Via Emilia Parmense 84,
29122 Piacenza, Italy
e-mail: adriano.marocco@unicatt.it

M. Busconi · A. Lanubile · A. Marocco
Research Centre for Biodiversity and Ancient DNA,
Università Cattolica del Sacro Cuore, Via Emilia
Parmense 84, 29122 Piacenza, Italy

Keywords White grain sorghum · Temperate sorghum · Genetic characterization · SSR · Microsatellites

Introduction

Sorghum (*Sorghum bicolor* (L.) Moench) belongs to the Poaceae family and is a worldwide widespread C4 crop. It is a subsistence crop and the main food for populations in arid or semiarid regions such as Southern Asia and sub-Saharan Africa where the environment is too stressful for other main cereal crops (Lasky et al. 2015; Boyles et al. 2017). The crop is also extensively grown in different countries like the USA, Australia and Europe for the production of

gluten-free products, forages for animal feeding, raw materials for industrial transformation in alcoholic beverages or bioethanol, backing flowers, pop sorghums, pet foods and packaging materials (Shahwar et al. 2012; Boyles et al. 2017; Li et al. 2018; Xiong et al. 2019). The end-use of different sorghum requires different plant or kernel characteristics needing specific breeding programs to develop the desired plant ideotype (Boyles et al. 2017). It is reported that many of the most interesting agronomical traits are under quantitative inheritance and the identification of genes that influence sorghum grain composition, gross energy content, and forage digestibility would help the selection of superior genotypes through the aid of molecular markers.

Cultivated sorghums were usually classified into five races (bicolor, caudatum, durra, guinea and kafir) on the basis of phenotypic characteristics of panicle and spikelets. The available races can originate different mixed gene-pools lineages and breeding materials can be further divided into “working groups” (de Oliveira et al. 1996; Brown et al. 2011; Motlhaodi et al. 2017).

Sorghum germplasm characterization usually focuses on collections of different materials like cultivar, landraces, inbred lines of contrasting phenotypes and origins (de Oliveira et al. 1996; Abu Assar et al. 2005; Motlhaodi et al. 2017). Nowadays molecular markers are widely used for germplasm identification, breeding and molecular traceability purposes in various crops (de Oliveira et al. 1996; Abu Assar et al. 2005; Brown et al. 2011; Španić et al. 2012; Scarano and Rao 2014; Stagnati et al. 2020).

Simple Sequence Repeats (SSRs) are the among the most used molecular markers, and they have been used extensively. SSR markers have been widely used in studies to inquire the genetic architecture of wild and cultivated sorghums with the aim to elucidate relationships between different germplasms, highlight possible sources of new alleles and phenotypes avoiding the introduction of “wild” characteristics and re-classify mislabeled accessions (de Oliveira et al. 1996; Abu Assar et al. 2005; Brown et al. 2011; Deu et al. 2008; Ganesamurthy et al. 2010; Kumar and Kumar 2009; Motlhaodi et al. 2014; Ng’uni et al. 2011, 2012; Prabhash and Khanna 2009; Singh and Boora 2008).

Grain sorghum is a staple food in many counties and a valid alternative to maize in dry areas and studies

aimed at the characterization of the genetic base underpinning nutritional values of starch, fat, proteins have been reported (Sukumaran et al. 2012; Rhodes et al. 2017; Boyles et al. 2017). Sorghum grains could classify into five groups according to kernel color, tannins and polyphenols content: white, yellow, red, brown, and black sorghums (Xiong et al. 2019). White sorghum has a white or colorless pericarp, a low levels of total phenolic contents, and has very low levels of tannin, 3-deoxyanthocyanidin and flavones whereas other pigmented genotypes have different levels of phenols and tannins (Xiong et al. 2019). The advantage of using white sorghum is an increased protein digestibility and it is reported that breakfast cereals made of white sorghum have nutritional composition and consumer acceptance similar to other cereals. White sorghums can also be used to produce teas, pasta, flour bread to add to wheat bread or pet food. Beers produced from white sorghums have higher phenolic content and antioxidant capacity than barley, providing a real opportunity for producing gluten-free beers for celiac people (Xiong et al. 2019).

In addition to grain quality parameters, it is necessary to take into account also agronomic traits that may distinguish the elite such as: cycle length, plant height, panicle shape, yield and resistance to biotic and abiotic stresses (Boyles et al. 2017). Sorghum germplasm has significant diversity for an array of important traits like stem sugar content, lignin and cellulose content, grain yield, forage/biomass yield, and drought tolerance (Dahlberg et al. 2011; Pasini et al. 2014; Fracasso et al. 2017). This phenotypic diversity has been used by breeders to originate an array of elite and diverse grain, forage and sweet sorghum genotypes with long history of trying to capture heterosis in sorghum hybrids (Dahlberg et al. 2011). It is reported that plant breeders have focused on traits likely to affect productivity, such as yield and/or forage quality thus requiring the characterization of materials under the morphological and productive perspective (Natoli et al. 2002; Perazzo et al. 2017).

Considering the value of white sorghum for the food industry and that the genetic characterization of white grain sorghums collections for temperate climates is lacking, the objective of the present study was the evaluation of genetic diversity within a collection of 117 accessions of white grain sorghum. Accessions were analyzed by using microsatellite markers. In

addition, a preliminary characterization of the different genotypes was performed considering the main agronomic traits to highlight the most suitable ones for direct cultivations and future breeding programs.

Material and Methods

Germplasm collection

The germplasm examined is part of the collection held by the Department of Sustainable Crop Production of Università Cattolica del Sacro Cuore, Piacenza. The entire collection consists of around six hundred different Sorghum accessions. The collection dates back to the beginning of the '80 s and, since then, it has been maintained by selfing of uniform materials within plots (Natoli et al. 2002). The list of the white sorghum accessions with their origin, when available, and main agronomic traits are reported in Supplementary Table 1.

Field experiment and phenotyping

The field experiment was conducted in Azienda Sperimentale Tadini (44°58'51.4"N 9°40'38.6"E) located in Gariga di Podenzano (Piacenza, Italy). The field was sown on 9th May 2015, each plot consisted of two rows 5 m long and spaced 50 cm apart, consecutive blocks were separated by 1 m aisle. The field was managed according to standard agronomic practices.

The following agronomic traits were recorded according to UPOV protocol: stalk diameter, plant height, leaf width, three flowering stages (50 % of plants with panicle still enveloped by the flag leaf, 50 % of plants with emitted panicle and 50 % of plants with visible anthers), panicle shape and weight.

DNA extraction and amplification

Young leaf tissues were sampled from four representative plants per genotype. Sampled leaves were bulked and DNA was extracted according to GenElute DNA Miniprep Kit (Sigma-Aldrich) according to manufacturer instruction with minor modification consisting in the addition of 5 % w/v Polyvinylpyrrolidone (PVP) during the lysis step to help the removal of polyphenols and inhibiting compounds (Stagnati et al. 2020).

Extracted DNA was visualized on 1 % agarose gel electrophoresis stained with Midori Green (NipponGenetics).

PCR reactions were carried out in a final volume of 25 µl. PCR mixture was composed of: 1 µl of crude DNA extract, 1X Reaction Buffer, 12 pmol dNTPs, 4 pmol each primer, 1U Taq polymerase, 2 % PVP and H₂O to final volume. PVP was added also during amplification to improve PCR amplification (Stagnati et al. 2017, 2020).

SSR markers were selected from Bhatramakki et al. (2000), 40 primer pairs were tested on 2–4 genotypes and evaluated for polymorphic amplification to select the 10 most polymorphic pairs as reported in Table 1.

PCR cycle consisted of initial denaturation at 94 °C for 5 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at optimal primer temperature as reported in Table 1 for 30 s, extension at 72 °C for 1 min and a final extension at 72 °C for 10 min.

Fluorescent labelled PCR fragments were visualized using an automated genetic analyzer ABI-Prism 3100 (Applied Biosystem) according to manufacturer's instructions and manually scored.

Statistical analysis

Detected alleles were analyzed with the GenAlEx6 software (Peakall and Smouse 2006) to compute population statistics according to the formulas implemented by the software. The Polymorphic Information Content (PIC) was calculated according to the formula of Botstein et al. (1980).

Collected genetic data were used to construct a phylogenetic tree by the *UPGMA* function of the phangorn package (Schliep 2011) starting from a genetic distance matrix calculated by the *meandistance.matrix* available in the polysat (Clark and Jasieniuk 2011) package of the R software. PCoA was obtained by GenAlEx6 software.

Results and discussion

Phenotypic data

Regarding the agronomic trait measured on the sorghum germplasm collection, plant height ranged from 70 cm (LP155) to 310 cm (IDG93431) with a

Table 1 List of primer pairs used in this work

No.	SSR Marker	Core Motif	Sequence 5'-3', **	T _{ann}	Amplicon Size (bp)
1	XTXP335	(GT) ₁₂	F: TATTCCTCTTGAAAGAATCAGGG R: 6FAM-TATTCATCGAGCAAAAGGCA	33	130–342
2	XTXP298	(AGA) ₂₃	F: 6FAM-GCATGTGTCAGATGATCTGGTGA R: GCTGTTAGCTTCTCTAATCGTCGGT	17	175–221
3	XTXP228	(TC) ₁₂	F: 6FAM-ACAGGTTGGCGATGTTTCTCT R: TTCTTTTTCGAATTCATTCCTTTT	10	216–278
4	XTXP141	(GA) ₂₃	F: 6FAM-TGTATGGCCTAGCTTATCT R: CAACAAGCCAACCTAAA	15	132–166
5	XTXP312	(CAA) ₂₆	F: 6FAM-CAGGAAAATACGATCCGTGCCAAGT R: GTGAACTATTCGGAAGAAGTTTGGAGGAAA	24	121–218
6	XTXP289	(CTT) ₁₆ +(AGG) ₆	F: 6FAM-AAGTGGGGTGAAGAGATA R: CTGCCTTTCGACTC	29	134–345
7	XTXP265	(GAA) ₁₉	F: 6FAM-GTCTACAGGCGTGCAAATAAAA R: TTACCATGCTACCCCTAAAAGTGG	31	76–235
8	XTXP292	(AC) ₁₂	F: GTCTACAGGCGTGCAAATAAAA R: 6FAM-TTACCATGCTACCCCTAAAAGTGG	15	152–366
9	XTXP94	(TC) ₁₆	F: 6FAM-TTTCACAGTCTGCTCTCTG R: AGGAGAGTTGTTTCGTT	17	122–230
10	XTXP24	(TC) ₂₁	F: CCATTGAGCTTCTGCTATCTC R: 6FAM-TTCTAAGCCCACCGAAGTTG	23	107–166

* F means Forward primer; R means Reverse primer; ** 6FAM is 6-Carboxyfluorescein. This molecule is the fluorophore used to label primers for automatic visualization and sizing of PCR amplicons using automated sequencer

mean of 198.7 cm. For grain sorghum, the optimal plant height is around 100–150 cm, 37 accessions have respected this threshold. According to Mutava et al. (2011), farmers prefer shorter plants with a sturdier stalk to resist lodging mechanical harvesting.

Flowering time is an important character to determine the cultivation suitability of crop varieties cultivation in a certain environment. Sorghum flowering times were recorded at three different stages when 50 % of plants reached a particular development point. The first one represented by the barrel-stage during the panicle protruded in the last leaf sheath, which ranged from 38 days after sowing (DAS) for 18ACK60FERT to 113 DAS for IS10050, with a mean of 79.8 DAS. The second flowering-time represented by panicle emergence ranging between 61 DAS of MN55 to 121 DAS of IS10050 with a mean of 86.45 DAS. The last flowering time, measured at 50 % anthesis, varied from 64 DAS of MN55 to 144 DAS of IS3572 with a mean of 92.5 DAS. Globally, 30

genotypes flowered within 80 DAS and harvested in mid-September; genotypes flowering until 98 DAS harvested at the end of September while late flowering genotypes harvested from 6th to 12th October. Sorghum genotypes that flowers within 80 DAS are the most suitable for cultivation in Northern Italy.

Grain sorghum production varied from 0.76 g (304 kg/ha) of IS3572 to 108 g (43.2 t/ha) of MER82_12, with a mean of 40 g per panicle corresponding to a mean of 16 t/ha considering an ideal plant density of 40 pt/m².

Correlations between agronomic traits were calculated and reported in Fig. 1. Flowering times were positively and significantly correlated with plant height. Taller plants flowers generally later than shorter plants as reported in a closely related species as maize (Stagnati et al. 2020b). A very weak (0.16) but significant correlation ($p < 0.01$) between plant height and panicle weight was observed, while a negative weaker (-0.15), but significant correlation

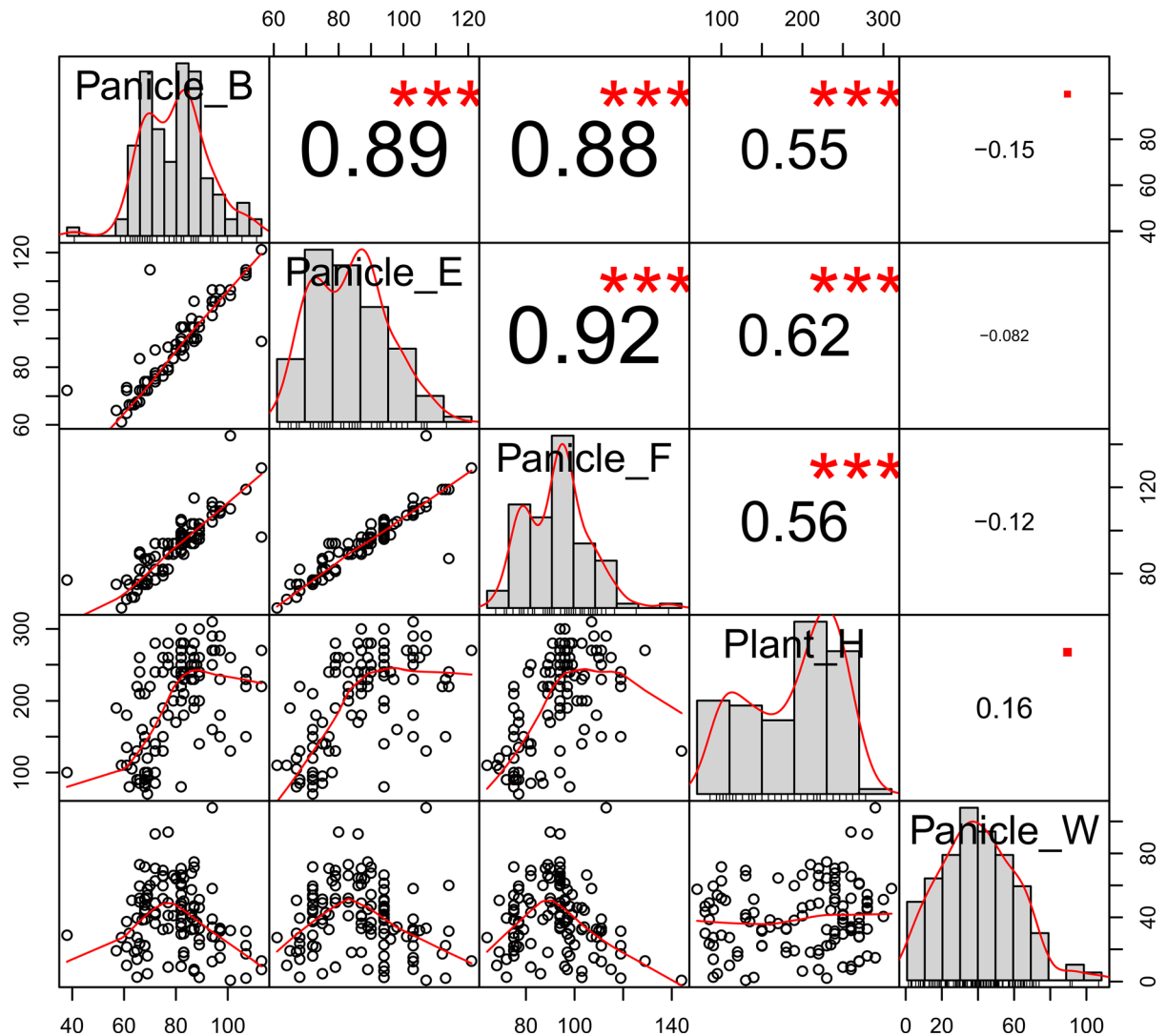


Fig. 1 Correlations among main agronomic traits of the 117 accession of white sorghum collection. Significance are reported: *** (p-value < 10⁻⁴), *(p-value < 0.01)

($p < 0.01$) between barrel stage and panicle weight was detected.

Other morphological traits measured were stem diameters, leaf width and panicle shape. These traits were scored according to class attribution. χ^2 indices were calculated (Supplementary Table S-chi) and were all significant at p -value < $5 \cdot 10^{-3}$ showing that genotypes with medium stem diameters have intermediate leaves width and compact panicles. The possible relation between panicle shape and potential production was examined. χ^2 was found to be significant at p -value < $5 \cdot 10^{-3}$ with 73 genotypes

characterized by low production (less than 60 g per panicle) and a compact panicle shape.

Genetic characterization

Among the 40 microsatellite markers tested, ten were selected and used to analyse the entire germplasm collection. As reported in Table 2 the number of observed alleles (N_a) was generally high varying from 10 for marker XTXP228 to 33 for marker XTXP335 with an overall mean of 21.4.

Different studies detected a reduced number of alleles ranging from 2 to 15 at a single locus

(Bhatramakki et al. 2000; Ali et al. 2008; Motlhaodi et al. 2017). The high number of alleles detected in the present work may be explained by: (1) the higher number of accessions considered than those reported by Bhatramakki et al. (2000), Ali et al. (2008) and Motlhaodi et al. (2017); (2) the more heterogeneous origin of the germplasm (Abu Assar et al. 2005); (3) the use of an automated genetic analyzer to run the PCR products that may have resolved better alleles of similar size. Loci with a high number of alleles showed an increased range of amplicon size: for example XTXP335 (33 alleles) showed a range from 130 to 342 bp, XTXP265 (31 alleles) had a range from 76 to 235 pb, XTXP289 (29 alleles) showed a range from 134 to 345 pb; on the opposite, marker XTXP228, with only 10 alleles, had a range from 216 to 278 bp. The possibility to find an increased amplicon size is in agreement with Abu Assar et al. (2005).

Concerning the number of effective alleles (N_e) ranged from 2.07 for XTXP94 to 9.68 for XTXP289. The Shannon information index (I), which varied from 1.34 of XTXP94 to 2.75 of XTXP289, is used to estimate genetic diversity in terms of a number of alleles at equal frequencies that would be necessary to obtain the same information of the sample set (Sherwin et al. 2006).

Observed heterozygosity (H_o) was found to be lower than expected heterozygosity (H_e) except marker XTXP265 where H_o and H_e were almost the

Fig. 2 Phylogenetic tree showing the relatedness of the 117 accessions of white sorghum collection

same. At this locus, many genotypes showed a heterozygous profile supporting the presence of residual heterozygosity. Similarly, the F_{ST} index was found to be high for all markers, except XTXP265. These results are in agreement with the reproduction system of sorghum where plants are selfed for several generations. Moreover, in the field, accessions showed high phenotypic uniformity typical of materials that have undergone selfing for several generations.

The PIC was between 0.502 of XTXP94 to 0.89 of XTXP289; it was observed that, markers with low PIC are characterized by the presence of few alleles at high frequencies and several alleles at low frequencies. Similar PIC values are reported by Abu Assar et al. (2005); allele frequencies at each locus are reported in Supplementary Table 2.

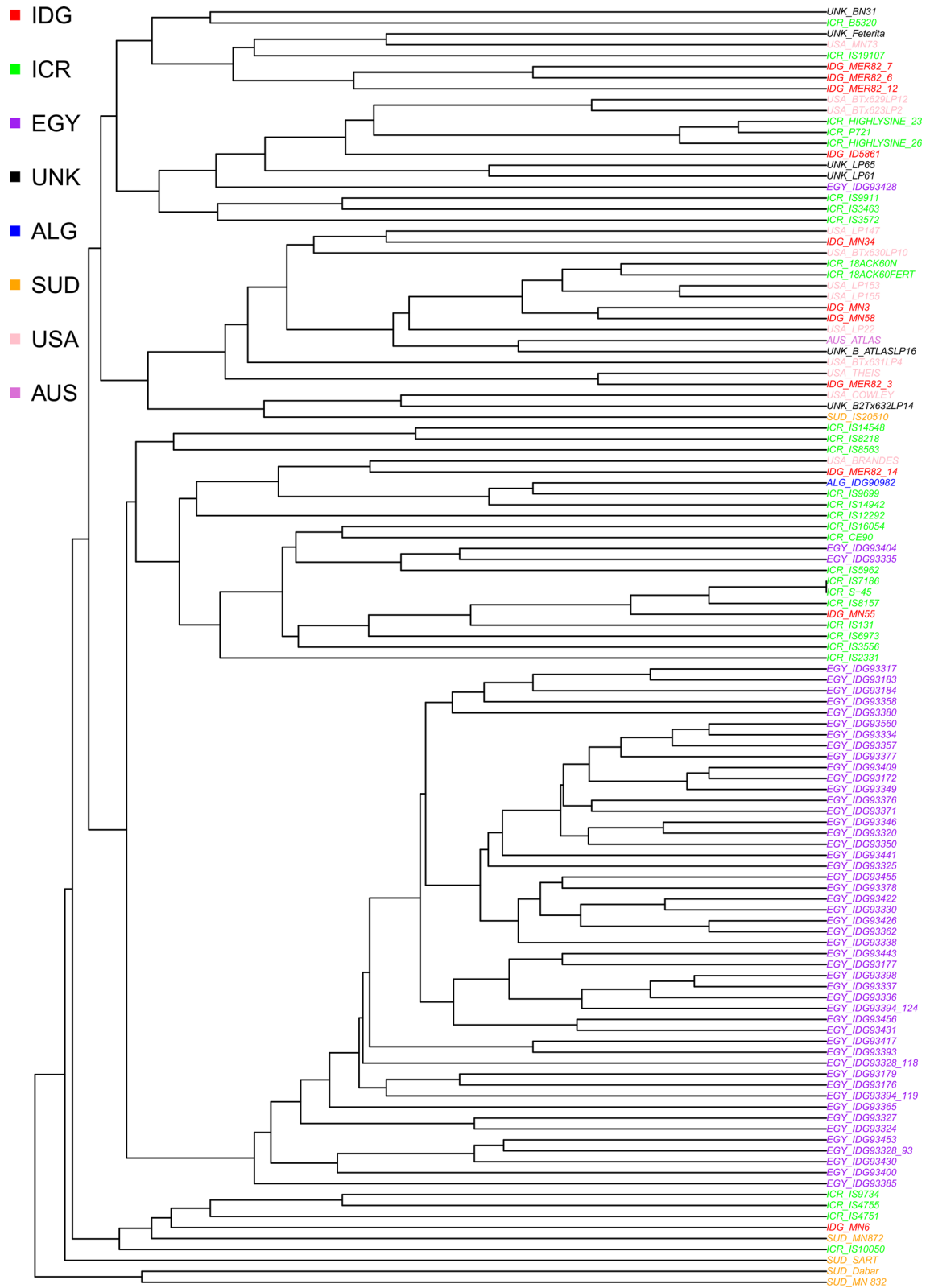
Phylogenetic analysis

The phylogenetic tree was constructed to have a visualization of the genetic variability inside the germplasm collection and to distinguish between groups on the basis of genetic relatedness.

Table 2 Population statistics computed for the sorghum collection

	XTXP 335	XTXP 298	XTXP 228	XTXP 141	XTXP 312	XTXP 289	XTXP 265	XTXP 292	XTXP 94	XTXP 24	Mean
Amplicon length(bp)	130–342	175–221	216–278	132–166	121–218	134–345	102–235	152–366	122–230	107–166	#
N	117	117	117	117	117	115	117	117	117	109	116.00
Na	33	17	10	15	24	29	31	15	17	23	21.4
N_e	7.693	9.072	4.230	2.932	4.929	9.678	6.687	5.608	2.071	9.132	6.203
I	2.721	2.424	1.645	1.686	2.150	2.749	2.520	2.090	1.341	2.510	2.184
H_o	0.419	0.214	0.137	0.111	0.128	0.600	0.838	0.103	0.222	0.257	0.303
H_e	0.870	0.890	0.764	0.659	0.797	0.897	0.850	0.822	0.517	0.890	0.796
F	0.519	0.760	0.821	0.831	0.839	0.331	0.015	0.875	0.570	0.712	0.627
PIC	0.864	0.880	0.728	0.644	0.777	0.890	0.840	0.805	0.502	0.881	0.781

N: sample number, Na: number of observed alleles, N_e : expected allele number, I : Shannon's Information Index, H_o : observed heterozygosity, H_e : expected heterozygosity, F: fixation index



The phylogenetic tree reported in Fig. 2 shows 6 main groups and some unrelated samples. Genotypes from Sudan (Sart, Dabar, MN832, MN872, IS20510) and ICRISAT (IS3463, IS2331, IS19107) are genetically unrelated and dispersed in the entire phylogenetic tree. Nonetheless, some relations exist between Dabar and MN832 that separates immediately in the tree. Genetic diversity of these materials underlines also an interesting phenotypic diversity for main agronomic traits. In Sudanese sorghum landraces, deriving improved cultivars is reported high genetic diversity considering that Sudan is believed to be the diversity origin of Sorghum (Abu Assar et al. 2005). The Ethiopian IS14584 is genetically diverse from Sudanese materials as well as the other African materials, IS14942, IS16054 (both from Cameroon) and IS7186 (from Nigeria). Materials originating from Egypt were clustered together, inside cluster 5, except IDG93428, IDG93404, IDG93335. It is reported that materials from Africa are genetically homogeneous and differentiated from others (Billot et al. 2013).

The genetic diversity was found also from the germplasm of US (Minnesota) origin that, even if adapted to short-season temperate climate, revealed an interesting agronomic variability with short-medium cycle materials characterized by different height, tassel attitude and potential production (Supplementary Table 1). Within these rely the most on productive genotype (MER82_12) of the collection. The dispersion of sorghum genotypes of the same country of origin is reported already in landraces characterization (Motthaodi et al. 2017). Other US genotypes from Texas and Mississippi were generally clustered with other US materials in groups 1–3. The wide distribution of American sorghum is consistent, considering that these accessions were introduced or bred in the US from materials originating elsewhere (Billot et al. 2013).

The high level of variation exhibited by the germplasm indicates the potential application of these accessions for further sorghum breeding (Motthaodi et al. 2017).

Principal coordinates analysis for main agronomic traits

The phenotypic values of the 117 white sorghum accessions plotted against the eigenvalues of the first two principal components.

Three agronomic traits were investigated since they are the main focus for breeders (Mutava et al. 2011) and are easy to measure during progenies selection.

Concerning plant height, the most favorable accessions grouped on the right side of the plot with shorter genotypes mainly grouped in the bottom-right side while accessions 100–150 cm height were mainly on the top-right side of the axis. Taller accessions, which are less useful for mechanical harvesting and grain sorghum, were mainly grouped on the left side of the graph even if some genotypes are interspersed within shorter accessions (Fig. 3).

The correlation between flowering time, considered as 50 % anthesis, and genotypes showed similar distribution. In this case, early and very late flowering genotypes were grouped in the right side of the plot, while intermediate-late genotypes (82 to 109 DAS), which were the most (73 out of 117 genotypes), were distributed mainly on the left part (Fig. 4). The correspondence of the distribution reported in Figs. 3 and 4 was confirmed also by the positive and highly significant correlation between plant height and flowering time. Plant height increased with cycle length until a maximum; for extremely long cycle accessions plant height decreased since these materials were unadapted and unable to develop correctly. The two early flowering genotypes on the left side of the graph (IDG93409 and IDG93385) were of the same origin (Egypt).

Considering the potential yield it was possible to observe that unproductive genotypes are clustered in the bottom-right side of the plot while, intermediate genotypes are evenly distributed (Fig. 5). It was possible to see the tendency of productive accession to move toward the left side of the graph except for the most productive one (MER82.12) which clusters with the unproductive accessions.

Interestingly it was possible to observe a clear correspondence between genotypes and groups highlighted on the basis of agronomic traits of interest. The preliminary observations on this germplasm collection support the identification of 26 genotypes (MN58, MN55, IS7186, B5320, 18ACK60FERT, 18ACK60N, IS131, IS3556, IS12292, BTx623_LP2, BTx631_LP4, BT630_LP10, ID5861, BN31, MN73, MN832, MN872, Feterita, Dabar, IS20510, LP61, LP65, LP147, LP155, LP22 and LP153) as more suitable for grain production in temperate environment. These genotypes, which cluster together on the right side of

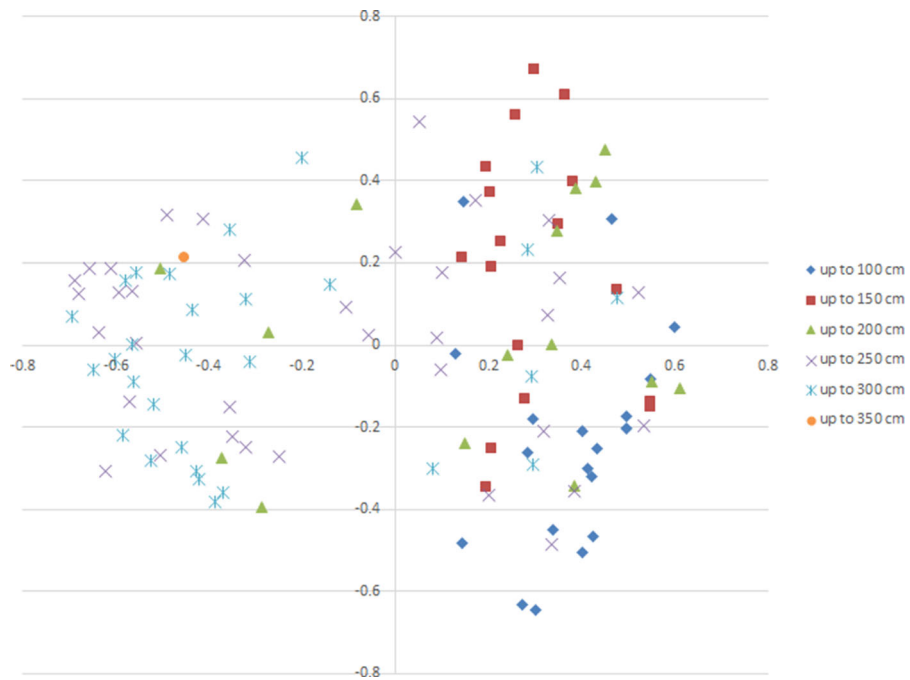


Fig. 3 PCoA analysis of the 117 genotypes of the white sorghum collection with accession grouped according to plant height

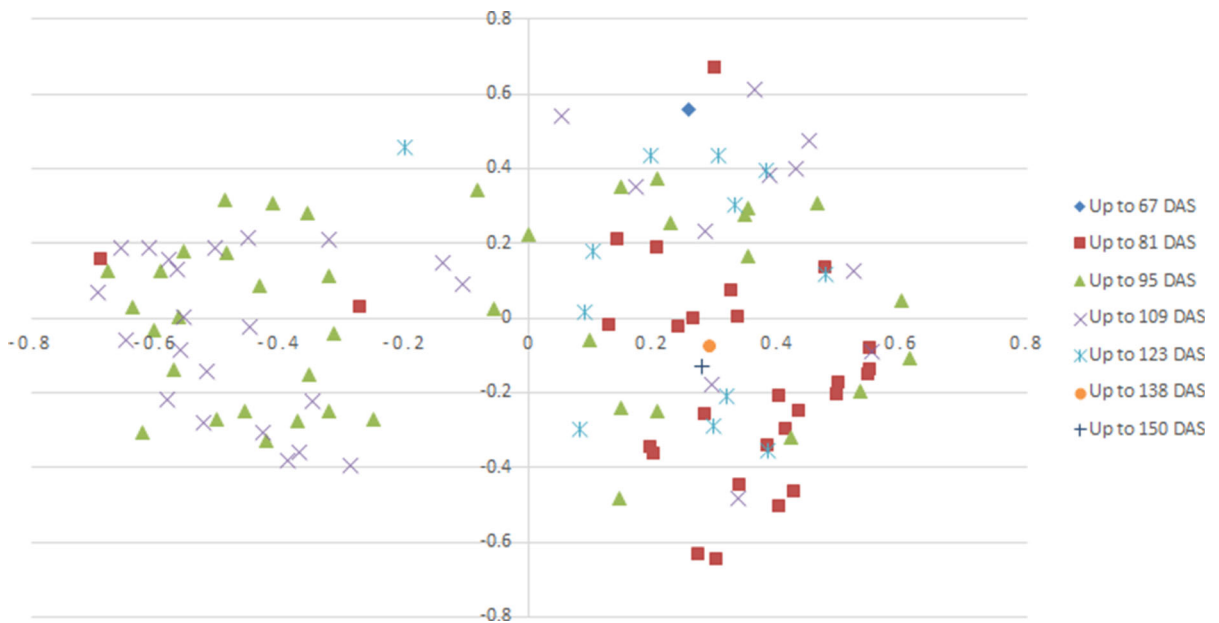


Fig. 4 PCoA analysis of the 117 genotypes of the white sorghum collection with accession grouped according to flowering time (at 50 % anthesis)

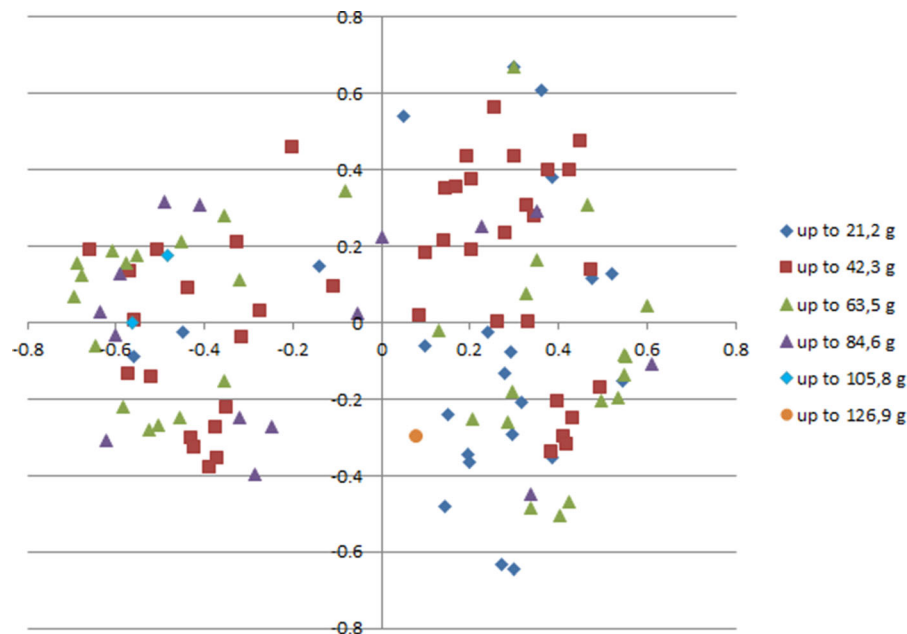


Fig. 5 PCoA analysis of the 117 genotypes of the white sorghum collection with accession grouped according to panicle weight

the genetic coordinates, require a subsequent field characterization before commercial exploitation.

Conclusions

The present work focused on the genetic characterization of a 117 genotypes collection of temperate grain sorghum using 10 SSR markers. Microsatellite analysis revealed an increase in the number and size of alleles comparing previous studies due to the wide distribution of germplasm origin. Observed heterozygosity was found lower than expected heterozygosity, with the exception of one locus, which is consistent with the selfing scheme of sorghum genotype maintenance. Phylogenetic analysis showed that the accessions are grouped in six main groups irrespectively to their origin with the exception of the majority of genotypes originating from Egypt, which form a uniform group. A wide range of phenotypes, for the main agronomic traits, was found in the collection, many 26 genotypes (MN58, MN55, IS7186, B5320, 18ACK60FERT, 18ACK60N, IS131, IS3556, IS12292, BTx623_LP2, BTx631_LP4, BT630_LP10, ID5861, BN31, MN73, MN832, MN872, Feterita, Dabar, IS20510, LP61, LP65, LP147, LP155, LP22 and LP153) genotypes are suitable for cultivation in a

temperate environments. Principal Coordinate Analysis revealed a good correspondence between genetic groups and accessions identified on the basis of their agronomic performance. These results outlined that the collection examined represent a good source of genetic and agronomic variability for breeding programs.

Author contributions Lorenzo Stagnati: Investigation, Formal analysis, Writing - original draft; Matteo Busconi: Investigation, Data curation, Writing; Giovanna Soffritti: Data curation, Review & editing; Michelangelo Martino: Review & editing; Alessandra Lanubile: Review & editing; Adriano Marocco: Conceptualization, Funding acquisition, Review.

Funding Open access funding provided by Università Cattolica del Sacro Cuore within the CRUI-CARE Agreement.. This research has been carried out with the financial support of CREMONA AGRI-FOOD TECHNOLOGIES – CRAFT ID 2018/2757 FONDAZIONE CARIPILO – Delibera Regionale XI/953 Regione Lombardia (Italy).

Code availability Not applicable.

Declarations

Conflict of interest The authors declares no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abu Assar A, Uptmoor R, Abdelmula AA, Salih M, Ordon F, Friedt W (2005) Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). *Crop Sci* 45. doi:<https://doi.org/10.2135/cropsci2003.0383>
- Ali M, Rajewski J, Baenziger P, Gill K, Eskridge K, Dweikat I (2008) Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Mol Breed* 21:497–509. doi:<https://doi.org/10.1007/s11032-007-9149-z>
- Bhatramakki D, Dong J, Chhabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of Sorghum bicolor (L.) Moench. *Genome* 43(6):988–1002
- Billot C, Ramu P, Bouchet S, Chanterreau J, Deu M, Gardes L, Noyer JL, Rami JF, Rivallan R, Li Y, Lu P, Wang T, Folkertsma RT, Arnaud E, Upadhyaya HD, Glaszmann JC, Hash CT (2013) Massive sorghum collection genotyped with SSR markers to enhance use of global genetic resources. *PLoS One* 8(4):e59714. doi:<https://doi.org/10.1371/journal.pone.0059714>
- Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32(3):314–331
- Boyles RE, Pfeiffer BK, Cooper EA, Rauh BL, Zielinski KJ, Myers MT, Brenton Z, Rooney WL, Kresovich S (2017) Genetic dissection of sorghum grain quality traits using diverse and segregating populations. *Theor Appl Genet* 130(4):697–716. doi:<https://doi.org/10.1007/s00122-016-2844-6>
- Brown P, Myles S, Kresovich S (2011) Genetic Support for Phenotype-based Racial Classification in Sorghum. *Crop Sci* 51:224. doi:<https://doi.org/10.2135/cropsci2010.03.0179>
- Clark L, Jasieniuk M (2011) polysat: an R package for polyploid microsatellite analysis. *Molecular Ecology Resources* 11(3):562–566. doi:<https://doi.org/10.1111/j.1755-0998.2011.02985.x>
- Dahlberg J, Wolfrum E, Bean B, Rooney W (2011) Compositional and Agronomic Evaluation of Sorghum Biomass as a Potential Feedstock for Renewable Fuels. *J Biobased Mater Bioenergy* 5:507–513. doi:<https://doi.org/10.1166/jbmb.2011.1171>
- de Oliveira AC, Richter T, Bennetzen JL (1996) Regional and racial specificities in sorghum germplasm assessed with DNA markers. *Genome* 39(3):579–587. <https://doi.org/10.1139/g96-073>
- Deu M, Sagnard F, Chanterreau J, Calatayud C, Héroult D, Mariac C, Pham JL, Vigouroux Y, Kapran I, Traore PS, Mamadou A, Gerard B, Ndjeunga J, Bezançon G (2008) Niger-wide assessment of in situ sorghum genetic diversity with microsatellite markers. *Theor Appl Genet* 116(7):903–913. doi:<https://doi.org/10.1007/s00122-008-0721-7>
- Fracasso A, Magnanini E, Marocco A, Amaducci S (2017) Real-Time Determination of Photosynthesis, Transpiration, Water-Use Efficiency and Gene Expression of Two Sorghum bicolor (Moench) Genotypes Subjected to Dry-Down. *Front Plant Sci* 8:932. <https://doi.org/10.3389/fpls.2017.00932>
- Ganesamurthy K, Punitha D, Elangovan M (2010) Genetic diversity among the land races of sorghum collected in Tamil Nadu. *Electron J Plant Breed* 1:1375–1379
- Kumar MM, Kumar KMH (2009) Estimation of genetic variability among sorghum genotypes using SSR markers. *Mysore J Agric Sci* 43:744–748
- Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE, Hyma K, Acharya C, Mitchell SE, Buckler ES, Brenton Z, Kresovich S, Morris GP (2015) Genome-environment associations in sorghum landraces predict adaptive traits. *Sci Adv* 1(6):e1400218. <https://doi.org/10.1126/sciadv.1400218>
- Li J, Tang W, Zhang YW, Chen KN, Wang C, Liu Y, Zhan Q, Wang C, Wang SB, Xie SQ, Wang L (2018) Genome-Wide Association Studies for Five Forage Quality-Related Traits in Sorghum (*Sorghum bicolor* L.). *Front Plant Sci* 9:1146. doi:<https://doi.org/10.3389/fpls.2018.01146>
- Motlhaodi T, Geleta M, Bryngelsson T, Fatih M, Chite S, Ortiz R (2014) Genetic diversity in ex situ conserved sorghum accessions of Botswana as estimated by microsatellite markers. *Aust J Crop Sci* 8:35–43
- Motlhaodi T, Geleta M, Chite S, Fatih M, Ortiz R, Bryngelsson T (2017) Genetic diversity in sorghum [*Sorghum bicolor* (L.) Moench] germplasm from Southern Africa as revealed by microsatellite markers and agro-morphological traits. *Genet Resour Crop Evol* 64:599–610. doi:<https://doi.org/10.1007/s10722-016-0388-x>
- Mutava R, Prasad PV, Tuinstra M, Kofoid K, Yu J (2011) Characterization of sorghum genotypes for traits related to drought tolerance. *Field Crops Research* 123:10–18. doi:<https://doi.org/10.1016/j.fcr.2011.04.006>
- Natoli A, Gorni C, Chegdani F, Ajmone Marsan P, Colombi C, Lorenzoni C, Marocco A (2002) Identification of QTLs associated with sweet sorghum quality. *Maydica* 47(3–4):311–322
- Ng'uni D, Geleta M, Bryngelsson T (2011) Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) accessions of Zambia as revealed by simple sequence repeats (SSR). *Hereditas* 148:52–62. doi:<https://doi.org/10.1111/j.1601-5223.2011.02208.x>
- Ng'uni D, Geleta M, Hofvander P, Fatih M, Bryngelsson T (2012) Comparative genetic diversity and nutritional

- quality variation among some important Southern African sorghum accessions [*Sorghum bicolor* (L.) Moench]. *Aust J Crop Sci* 6:56–64
- Pasini L, Bergonti M, Fracasso A, Marocco A, Amaducci S (2014) Microarray analysis of differentially expressed mRNAs and miRNAs in young leaves of sorghum under dry-down conditions. *J Plant Physiol* 71(7):537–548. doi:<https://doi.org/10.1016/j.jplph.2013.12.014>
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Perazzo A, Carvalho G, Santos E, Pinho R, Campos F, Macedo C, Tabosa J (2014) Agronomic evaluation of 32 sorghum cultivars in the Brazilian semi-arid region. *Revista Brasileira de Zootecnia* 43:232–237. doi:<https://doi.org/10.1590/S1516-35982014000500002>
- Prabhaskar K, Khanna VK (2009) Assessment of genetic diversity in cultivated sorghum (*Sorghum bicolor* (L.) Moench) accessions using microsatellite markers. *Pantnagar J Res* 7:143–149
- Rhodes DH, Hoffmann L, Rooney WL, Herald T, Bean S, Boyles R, Brenton Z, Kresovich S (2017) Genetic architecture of kernel composition in global sorghum germplasm. *BMC Genom* 18:15. <https://doi.org/10.1186/s12864-016-3403-x>
- Scarano D, Rao R (2014) DNA Markers for Food Products Authentication. *Diversity* 6:579–596. doi:<https://doi.org/10.3390/d6030579>
- Schliep KP (2011) phangorn: phylogenetic analysis in R. *Bioinformatics* 27(4):592–593
- Shahwar D, Noor M, Rahman H, Ullah I, Shah I, Ali F, Mehar S, Shah SM, Mehmood N (2012) Characterization of sorghum germplasm for various morphological and fodder yield parameters. *Afr J Biotechnol*. doi:<https://doi.org/10.5897/AJB11.4000>
- Sherwin WB, Jabot F, Rush R, Rossetto M (2006) Measurement of biological information with applications from genes to landscapes. *Mol Ecol* 15(10):2857–2869. <https://doi.org/10.1111/j.1365-294X.2006.02992.x>
- Singh M, Boora KS (2008) Genetic diversity among forage sorghum [*Sorghum bicolor* (L.) Moench] accessions using simple sequence repeats. *SABRAO J Breed Genet* 40:77–91
- Španić V, Buerstmayr H, Drezner G (2012) Assessment of genetic diversity of wheat genotypes using microsatellite markers. *Periodicum Biologorum* 114:37–42
- Stagnati L, Soffritti G, Lanubile A, Busconi M (2017) Comparison of six methods for the recovery of PCR-compatible microbial DNA from an agricultural biogas plant. *Appl Microbiol Biotechnol* 101(9):3907–3917. <https://doi.org/10.1007/s00253-017-8152-5>
- Stagnati L, Soffritti G, Martino M, Bortolini C, Lanubile A, Busconi M, Marocco A (2020) Cocoa beans and liquor fingerprinting: A real case involving SSR profiling of CCN51 and “Nacional” varieties. *Food Control* 118:107392. <https://doi.org/10.1016/j.foodcont.2020.107392>
- Sukumaran S, Xiang W, Bean SR, Pedersen JF, Kresovich S, Tuinstra MR, Tesso TT, Hamblin MT, Yu J (2012) Association Mapping for Grain Quality in a Diverse Sorghum Collection. *The Plant Genome* 5:126–135. doi:<https://doi.org/10.3835/plantgenome2012.07.0016>
- Xiong Y, Zhang P, Warner RD, Fang Z (2019) Sorghum Grain: From Genotype, Nutrition, and Phenolic Profile to Its Health Benefits and Food Applications. *Comp Rev Food Sci F* 18:2025–2046. <https://doi.org/10.1111/1541-4337.12506>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.