**Opportunities from the genetic diversity of the ILRI genebank forage germplasm collection**

***Alieu Mortuwah Sartie*[1]*, Alemayehu Teressa Negawo*[1]*, Ermias Habte*[1]*, Meki Shehabu Muktar*[1]*, Abel Teshome*[2]*, Jean Hanson*[1]*, Habib Olumide Akinmade*[1]*, Ki-Won Lee*[3] *and Chris S. Jones*[1]

[1]Feed and Forage Development, international Livestock Research institute, Addis Ababa, Ethiopia.

[2]teagasc|ceLUP Crop Research, Oak Park, Carlow, R93 XE12, Ireland.

[3]Grassland and forages Division, national institute of Animal Science, Rural Development Administration, Cheonan, 31000, Republic of Korea.

**Abstract**:

The International Livestock Research Institute (ILRI) maintains a collection of 18,662 forage germplasm accessions of grasses, herbaceous legumes and browse species at its genebank in Addis Ababa, Ethiopia. Most of the collection was acquired from different regions, in partnership with and the consent of national genebanks, while some were donations from other institutes, notably the Commonwealth and Scientific Industrial Research Organization (CSIRO) in Australia. The focus of the forage germplasm activities in ILRI is on the conservation, characterization and use as animal feed of these resources in smallholder livestock systems. To this end, the determination of genetic diversity in the collection is essential, underpinning the development of trait-based subsets of accessions and for the identification of genotypes that can be used as parents to develop new germplasm containing specific traits of interest. Here we report on the use of genotyping-by-sequencing (GBS), a robust and affordable genotyping method which uses a combination of genome complexity reduction using restriction endonucleases and next generation sequencing to identify large numbers of high-quality genome-wide genetic markers. GBS is a particularly useful technique to use on species with limited genomic information and we have applied this technique to assess genetic diversity in a range of our forage germplasm collections, including: Napier grass (*Cenchrus purpureus*); Buffel grass (*Cenchrus*

*ciliaris*); Rhodes grass (*Chloris gayana*); Lablab *(Lablab purpureus*); and Sesbania (*Sesbania sesban*). These data provide a significant resource for genetic and marker-trait association studies and genomic prediction, enhancing the prediction accuracy of superior genotypes and the efficiency of selection of new varieties, supporting improved animal production, using marker assisted breeding. Furthermore, the subsets are of a manageable size and can act as reference sets for distribution and evaluation in different agro-ecologies and production systems.

**Key words:** forage germplasm. genetic diversity. genotyping by sequencing. subset development and evaluation. trait phenotyping

**Introduction:**

The International Livestock Research Institute (ILRI) maintains a collection of 18,662 forage germplasm accessions of grasses, herbaceous legumes and browse species at its Genebank in Addis Ababa, Ethiopia. Most of the collection was acquired from different regions, in partnership with and the consent of national genebanks. Others were donations from other institutes, notably the Commonwealth and Scientific Industrial Research Organization (CSIRO) in Australia. The species range from short-lived annuals to long-lived perennial plants that are adapted to the tropics and Mediterranean areas. ILRI maintains this collection on behalf of the international community, under the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) framework (FAO, 2009).

A Global Tropical and Sub-Tropical Forages (TSTF) Strategy was developed in 2015 based on a survey of major TSTF national and international genebanks and the input from a workshop of genebank managers and forage specialists (Pengelly and Maass, 2017). The strategy attempted to balance the higher-level objectives of improving both the conservation and use of forage germplasm for forage or environmental use via three themes: (1) Building a global community of genebanks and forage

utilisation researchers; (2) Achieving greater efficiency and security in genebank operations to ensure conservation of the most important TSTF genera and species and better understanding of their diversity; and (3) Attaining greater utilisation of this valuable collection in the developed and developing world. Amongst the recommendations for the implementation of this strategy was that the genebanks should adopt a species prioritisation (table 1) and apply this prioritisation to accession management, including the appropriate removal of duplicates and the archiving or equivalent of low priority species, so that limited resources can be better applied to taxa most likely to contribute to impacts in improved livestock production and resource management (Pengelly and Maass, 2017). The number of ILRI accessions in each of the categories is shown in table 2.

Most of the forage accessions in the ILRI genebank collections have rarely or never been requested and distributed, and this is possibly so, at least in part, because most of these accessions have not been effectively characterised to better understand their forage value. This could be the reason for moving some of those uncharacterised accessions into the low prioritisation categories, for example, the more than 4000 accessions in category 5 (table 2). The focus of the forage germplasm activities in ILRI is on the conservation, characterization and use as animal feed of these resources in smallholder livestock systems. The determination of genetic diversity in the collection is essential for the development of trait-based subsets of accessions, and for the identification of genotypes that can be used as parents to develop new germplasm containing specific traits of interest. Here we report on the use of genotyping-by-sequencing (GBS), a robust and affordable genotyping method which uses a combination of genome complexity reduction with restriction endonucleases, enzymes that cut the DNA at specific sites, and next generation sequencing to identify large numbers of high-quality genome-wide genetic markers that are suitable for diversity analysis, marker-trait associations and genomic prediction.

**Materials and methods:**

*Genotypic analysis*

Grass and legume species with good or potential forage value were selected and subjected to GBS using the DArTseq platform (Kilian et al., 2012). Three grass and two legume species from the ILRI forage Genebank have been included in the study to date. The grasses included: 105 accessions of Napier grass (*Cenchrus purpureus*); 185 accessions of Buffel grass (*Cenchrus ciliaris*) and; 104 accessions of Rhodes grass (*Chloris gayana*). The Napier grass collection consisted of 60 accessions from ILRI and 45 accessions from EMBRAPA, Brazil. The legumes included: 145 accessions of Hyacinth bean (*Lablab purpureus*) and; 171 accessions of Egyptian pea (*Sesbania sesban*).

Genotypic data were generated through the application of GBS on the DArTseq platform that combines genome complexity reduction using a combination of restriction enzymes and next-generation sequencing (Kilian et al., 2012). DNA was isolated from one plant per accession of the grass species and 15 plants per accession, to assess the level of diversity contained both within and between accessions, of the legumes. The genomic DNA (approximately 50 ng) was digested with a combination of *PstI/HpaII* restriction endonucleases and the resulting fragments were ligated to a *PstI* overhang compatible oligonucleotide adapter and sequenced on an Illumina HiSeq 2500 platform (Illumina Inc.) using *PstI* site-specific primers. Short sequence fragments, SilicoDArT (presence/absence) and SNP markers were generated following the DArTseq protocol. Data were analysed using R tools and other statistical analysis software to identify diversity, population structure and subsets, as described by Muktar et. al., (2019).

*Phenotypic analysis*

Field phenotyping of the Napier grass collection for agronomic traits is as described by Habte et al., 2019 (*in preparation*). Napier grass collections consisting of 84 (59 ILRI and 25 EMBRAPA) accessions were planted at the ILRI field genebank in Bishoftu, Ethiopia using a partial replication design with four replications. After establishment at the beginning of the main rains and an initial harvest six months later in the dry season of 2018, drought stress was imposed on the established field plants in such a

way that two replicates were watered using a drip irrigation system to replace the loss of water due to evapotranspiration (volumetric soil moisture content of 20 %), i.e., optimal water (OW) and the other two replicates were irrigated with a limited amount of water (volumetric soil moisture content of 10 %), i.e., water deficit (WD). The soil moisture content of both water regimes was monitored using a Delta soil moisture probe (HD, England). The physiological drought stress effect was also monitored using a portable chlorophyll fluorescence meter Handy PEA (Hansatech, UK) to analyze the photochemical efficiency of leaves growing under stress. The trial was harvested and data on morphological traits, agronomic performance and feed quality were collected following every 8 weeks of regrowth.

*Subset identification*

To select a subset of representative accessions of Napier grass, the R package Core Hunter v. 3.2.1 (De et al., 2018) was used. This program identifies core subsets using diverse allocation strategies by optimizing many genetic parameters simultaneously. The modified Roger's distance (RD), Shannon's information index (SH), average entry-to-nearest-entry distance (EN), expected proportion of heterozygous loci (He) and allele coverage (CV), each with an equal weight, were used to define a core subset representing the entire collection.

Field phenotyping of the other species is currently being planned following results of genotypic analysis.

**Results:**

**GBS analysis and molecular marker development**

Results are as described by Muktar et. al., (2019) and Negawo et al., (2018). Both SNP and SilicoDArT genome-wide markers were generated for the different forage species (table 3). The short sequences of the generated markers were aligned with reference genomes of closely related species (table 3). In

Napier grass, a total of 85,452 SNP and 116,190 SilicoDArT genome-wide markers were called on the 105 accessions with an average call rate of 87 % for SNPs and 95 % for SilicoDArT markers. Missing values ranged from 0 to 59 % for SNPs, and from 0 to 30 % for the SilicoDArT markers, with an average value of 15 % in both marker sets. Accession ILRI_16621 had the highest missing value content (74 %) and was excluded from further analysis. Approximately 42 % (48,536) of the SilicoDArT and 20 % (17,086) of the SNP markers had a polymorphic information content (PIC) value above 0.25. The short sequence reads, averaging 55 nt in length, corresponding to each marker were mapped on to the pearl millet reference genome and genomic position information was generated for 17 % of the SNP and 33 % of the SilicoDArT markers.

The genotypic data for Buffel grass and Rhodes grasses were analysed as described for Napier grass above. For Buffel grass, 111,917 SilicoDArT and 93,501 SNP markers were obtained for 185 accessions. Out of those markers, 8,053 (7 %) SilicoDArT and 15,465 (16 %) SNP markers were aligned with *Setaria italica* as a reference genome. For Rhodes grass, a preliminary analysis has generated 93,128 SilicoDArT and 65,529 SNP markers from 94 accessions. Of the three selected reference genomes, more markers (0.74 % SilicoDArT and 5.86 % SNP markers) were able to be aligned on the teff (*Eragrostis tef*) reference genome followed by Manila grass (*Zoysia matrella*) (0.56 % Silico DArT and 5.13 % SNP markers). The least number of markers (0.23 % Silico DArT and 2.07 % SNP markers) were mapped to the *Setaria italica* reference genome.

For *Sesbania sesban* 34,798 SilicoDArT and 47,609 SNP markers were generated. Relatively few markers (1,168 SilicoDArT and 2,460 SNP markers) were aligned with the *Glycine max* reference genome.

For *Lablab purpureus* a total of 1,843 samples generated from 142 accessions (1 to 29 plants per accession) were genotyped. The genotyping produced a total of 38,824 SNP and 64,793 silicoDArT genome-wide and high-density markers. Out of the 142 accessions, 108 were represented by 10 or more plants per accession, and 72 were represented by 15 or more plants per accession. These will

be used for the study of within and between accession diversity, an analysis which is currently being undertaken. Distribution of PIC and He values of the markers are shown in figure 5.

**Diversity and population structure**

In the Napier grass population, diversity and population structure have previously been evaluated and presented using 980 highly polymorphic and independent SNP markers (pruned for LD at $r^2$ = 0.5) distributed across the genome, which were selected from the 85,452 genome-wide SNP markers (Muktar et al., 2019). The presence of subpopulations within the accessions was analysed with the 980 SNP markers described above, using the software STRUCTURE, PCA and UPGMA clustering methods. The analysis revealed the presence of between 2 and 7 sub-groups in the Napier grass population. The STRUCTURE analysis detected two major groups, with the collection from ILRI predominantly represented in Group I and most of the EMBRAPA collections assigned to Group II. However, this analysis also indicated the presence of up to 5 possible sub-groups, described in detail in Muktar et al., 2019. UPGMA further clustered the accessions into seven sub-groups (figure 1a), and Groups I, II, III, V, and VI were highly consistent with the STRUCTURE classification (figure 1b). Group IV and VI mainly consist of materials from ILRI and Groups I, II and III are mainly EMBRAPA materials, with Groups V and VII containing material from both collections. The eight C*. purpureus* x *P. glaucum* hybrids were distributed across groups IV (ILRI_16835 and ILRI_16837), V (ILRI_16834 and ILRI_16838), and VI (ILRI_15357, ILRI_16840, ILRI_18662 and ILRI_14982).

For Buffel grass, diversity and population structure analysis using 1,000 selected SNP markers distributed across the reference genome revealed the presence of two main groups with further sub-groups in the collection (figure 2 a and b).

For Rhodes grass, cluster analysis using 10,111 SNP markers with no missing data clearly showed two differentiated groups (figure 3).

A preliminary cluster analysis of the *Sesbania sesban* data with SNP markers filtered for polymorphic information content (PIC) ≥0.2 and missing value ≤30% is shown in figure 4.

**Phenotyping morphological traits:**

The performance of Napier grass genotypes for agro-morphological and feed quality traits were assessed over three wet and three dry seasons harvests during 2018. Significant variations were observed among the genotypes for plant height, leaf size, stem diameter, tiller number, biomass yield and water use efficiency, that indicated the existence of phenotypic variability among the experimental accessions (figures 6, 7 and 8). Similarly, results from forage quality analysis from leaf and stem tissues showed significant differences among genotypes, particularly for Acid detergent fibre (ADF), Neutral detergent fibre (NDF), Acid detergent lignin (ADL), Organic matter (OM), Dry matter (DM), Total nitrogen (N), Crude protein (CP), *In vitro* organic matter digestibility (IVOMD) and Metabolizable energy (Me). These results indicate a substantial opportunity for the improvement of different forage quality traits in Napier grass (figure 9).

**Identification of sub-sets**

Mini core subsets of Napier grass were identified using a combined analysis of genotypic and phenotypic data based on 68 accessions (Muktar et al., 2019). The initial phenotypic trait data (table 4) were used to complement the selected 980 genome-wide SNP marker data in the analysis. UPGMA analysis clustered the 68 accessions into seven sub-groups, and each sub-group was well represented in the subsets. Forage biomass traits, total fresh weight per plant (TFWPP) and total dry weight per plant (TDWPP), were highly variable among accessions in the sub-groups. Groups II and IV had higher mean values while groups I and VII had lower mean values for both traits when grown under optimal-

water conditions. A similar trend was observed when grown under water-deficit conditions, except that group IV had an average mean value in this case.

A subset of 14 (20%) accessions representing the range of phenotypic and genetic diversity in the 68 Napier grass accessions was identified for both optimal-water and water-deficit conditions and seven accessions are common between the two subsets (table 5).

Mini core subsets of Buffel grass were also created based on the genotypic (silicoDArT and SNP) data generated from the collection and some historical feed quality data (not shown). The 'corehunter' R package was used for the purpose. Subset analysis for the other species will be undertaken following completion of the full genetic diversity analysis and evaluating the accessions for agronomic performance and nutritional qualities in the field.

**Discussion:**

ILRI maintains a collection of more than 18,600 forage germplasm accessions of grasses, herbaceous legumes and browse species at its genebank in Addis Ababa, Ethiopia, as a public good for use for agriculture, research and education for food security. Understanding of genetic diversity in this collection is essential for its management and utilisation in plant breeding, other research and direct use for feed production and security. The determination of genetic diversity in the collection is also essential for the development of trait-based subsets of accessions, and for the identification of genotypes that can be used as parents to develop new germplasm containing specific traits of interest. Furthermore, identifying heterozygosity within selected collections of accessions held in trust will generate information that will facilitate the establishment of a baseline for the diversity of the collections across multiple crops, will be useful for exploring crop evolution, and will support forage plant breeding and genebank management. Genotyping is perceived as a tool to comprehensively characterize collections and reveal the diversity and population structure within and among

germplasm accessions. Here we report on the use of genotyping-by-sequencing (GBS), a robust and affordable genotyping method which uses a combination of genome complexity reduction with restriction enzymes and next generation sequencing to identify large numbers of high-quality genome-wide genetic markers that are suitable for diversity analysis, marker-trait associations and genomic prediction. GBS is a particularly useful technique to use on species with limited genomic information and we have applied this technique to asses genetic diversity in a range of our forage germplasm collections, including: Napier grass; Buffel grass; Rhodes grass; Lablab; and Sesbania. The above species are amongst the germplasm that are most frequently requested and distributed from the ILRI genebank, and they are all in category 1 of the Global Tropical and Sub-Tropical Forages (TSTF) Strategy species prioritisation list (Tables 1 and 2).

GBS analysis revealed a significant amount of diversity held in a small collection of Napier grass accessions (Muktar et al., 2019) and also in the Buffel grass collection (Negawo et al., 2019). The data for Rhodes grass, Sesbania and Lablab are currently being analysed, and a preliminary analysis has indicated the presence of diversity in the collections of Rhodes grass (figure 3) and Sesbania (figure 4).

In Napier grass, genetic diversity and population structure analyses revealed the existence of a substantial amount of variation in the collection (figure 1). This supports previous work by Negawo et al., (2018) who previously identified genetic diversity in this collection using microsatellite markers. The presence of two to seven groups was observed by STRUCTURE, principal component (PCA) and phylogenetic analyses and most of the materials from ILRI and EMBRAPA grouped separately. Analysis of molecular variance (AMOVA) indicated that the seven groups detected are significantly different from each other, with up to 14 % variation among the groups. The high level of diversity and population stratification observed could be attributed to the outcrossing, self-incompatibility (Hanna et al., 2004) and polyploid nature of Napier grass. Furthermore, selection, breeding systems, and variation in geographical origin may also be contributing to the variation seen between the materials derived from the ILRI and EMBRAPA collections.

Phenotypic analysis of morphological traits, agronomic performance and feed quality characteristics in Napier grass indicated the existence of phenotypic variability among the experimental accessions that would potentially identify highly productive accessions for promotion in support of livestock production both in water stressed and irrigated forage production areas. This adds to a previous study by Wouw et al. (1999) that also identified phenotypic variation in the ILRI collection.

In the Buffel grass collection, diversity and population structure are shown in figure 2. Two main groups were identified, with the possibility of identifying additional sub-groups, in the collection. Here too a previous study by Ricardo et al. (2017) had identified phenotypic variation in this collection and efforts are underway to put these various sources of data together for a more comprehensive analysis of the collection.

*Sub-setting*

In Napier grass, subsets were identified based on a combined analysis of GBS and phenotypic data. Only a few accessions were selected for the subsets but they well represent the overall genetic and phenotypic diversity of the collections. Based on a stress tolerance index (Fernandez, 1992) and water use efficiency analysis in the water deficit experiment in the 2018 dry season, accessions 16791, 16792, 16800, BAGCE100, 16801, 16802, 16819, 18438, 16786 and CNPGL 93-37-5 showed higher WUE (figures 8 & 9) and as such offer potential candidates for improved performance under dry conditions. For high biomass production during irrigation accessions 16791, 16819, 16802, 14983, 16814, 16783, BAGCE 100, BAGCE 30, CNPGL 00-1-1 and CNPGL 92-198-7 were identified as potential promising lines. The high biomass producing genotypes were either tall plants or they produced many tillers, indicating that high biomass production is associated with plant height and tiller number in Napier grass. These accessions offer prime candidates for further evaluation in different areas and production systems. Evaluating these subsets, which consist of only a few genotypes, will save time

and resources when compared to evaluating the whole collection for target traits. The subsets can also serve as reference sets, representing the genetic diversity of the whole collection.

Trait specific subset identification for the other species will follow the analysis of GBS data, field phenotyping and feed quality analysis and the combined analysis of genotypic and phenotypic data.

**Conclusion**:

We have used GBS to identify the diversity held in forage germplasm collections of Napier grass and Buffel grass. Phenotypic assessment of the Napier grass collection on the field for agronomic traits enabled the identification of subsets of accessions for drought tolerance and biomass productivity. Evaluation of these subsets across ecologies and seasons may further lead to the identification of best-bet accessions for those environments. A large number of SNP and SilicoDArT markers were generated that could be useful for diversity assessment in germplasm collections of other collections and related forage species.

**Acknowledgements:**

The authors gratefully acknowledge the support of: Federal Ministry for Economic Cooperation and Development (BMZ), Genebank uplift Funding from Germany; Germany-GIZ-Deutsche Gesellschaft für Internationale Zusammenarbeit, Gap Funding for Forage Selection and Breeding Activities; Rural Development Administration (RDA) of the Republic of Korea, project on the development of new forage genetic resources and their utilization; The CGIAR Research Program on Livestock. The authors would also like to thank EMBRAPA for making their germplasm and breeding lines available for the study through the Africa Brazil Agricultural Innovation Marketplace Project.

**References:**


De Beukelaer, H., Davenport, G. F. & Fack, V. Core Hunter 2018. Flexible core subset selection. *BMC Bioinformatics* **19,** 1–12

Fernandez, G.C.J. 1992. Effective selection criteria for assessing plant stress tolerance. p. 257-270. In Adaptation of food crops to temperature and water stress: Proceedings of an International Symposium, Taiwan, China. 13-18 August. Publication nr 93-410. 531 p. Asian Vegetable Research and Development Center, Taipei, China.

Food and Agriculture Organization of the United Nations (FAO) 2009. The International Treaty on Plant Genetic Resources for Food and Agriculture

Galili, T. 2015. Dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **31,** 3718–3720.

Habte Ermias, Meki Shehabu and Chris S. Jones 2019. Performance of Napier grass genotypes under optimum and water stress conditions (*In preparation*).

Hanna, W. W., Chaparro, C. J., Mathews, B. W., Burns, J. C. & Sollenberger, L. E. 2004. *Perennial Pennisetums* (eds. Moser, L. E., Burson, B. L. & Sollenberger, L. E.) 503–535 (American Society of Agronomy Monograph).

Kilian, A. *et al.* 2012 Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol Biol.* **888**, 67–89. doi: 10.1007/978-1-61779-870-2_5

Muktar, S. Meki, Abel Teshome, Jean Hanson, Alemayehu T. Negawo, Ermias Habte, Jean-Baka Domelevo Entfellner, Ki-Won Lee, and Chris S. Jones 2019. Genotyping by sequencing provides new insights into the diversity of Napier grass (*Cenchrus purpureus*) and reveals variation in genome-wide LD patterns between collections, Scientific Reports 9:6936.

Negawo, A. T., Jean Hanson, Yilikal Assefa, Alieu Sartie and Chris S. Jones 2019. The first in-depth insight into the genetic diversity of a buffel grass (*Cenchrus ciliaris L*.) collection using genotyping-by-sequencing analysis. Poster presented at the International Forage and Turf Breeding Conference (IFTBC), Lake Buena Vista, FL, USA, March 24-27, 2019.

Negawo, A. T., Alexandra Jorge, Jean Hanson, Abel Teshome, Meki S. Muktar, Ana Luisa S. Azevedo, Francisco J.S. Lédo, Juarez C. Machado and Chris S. Jones 2018. Molecular markers as a tool for germplasm acquisition to enhance the genetic diversity of a Napier grass (*Cenchrus purpureus* syn. *Pennisetum purpureum*) collection. *Trop. Grasslands-Forrajes Trop.* **6,** 58–69.

Pengelly Bruce and Brigitte Maass 2015. Report on 2016 Implementation of the Tropical and Sub Tropical Forages Strategy.

Ricardo Alonso Sánchez Gutiérrez, Carlos Raúl Morales Nieto, Jean Hanson3, Eduardo Santellano Estrada1, Pedro Jurado Guerra4, José Francisco Villanueva Avalos5 y Alicia Melgoza Castillo 2017. Forage characterization of ecotypes of buffel grass under temporary conditions in Debre Zeit, Ethiopia. Revista Mexicana de Ciencias Agrícolas Vol.8 Núm.1 01 de enero - 14 de febrero, p. 13-26.

Wouw MVD, Hanson J, Luethi S 1999. Morphological and agronomic characterisation of a collection of Napier grass (*Pennisetum purpureum*) and *P. purpureum* × *P. glaucum*. Tropical Grasslands 33: 150–158.

| Table 1. Prioritisation categories and their definition. | |
|---|---|
| **Category** | **Definition/explanation of species' category** |
| **1** | Species of known high value, included in the Tropical forages database (www.tropicalforages.info) or commercially useful somewhere |
| **2** | Identified as high potential for further development towards commercial use or emerging as one of high value somewhere |
| **3** | Often thought of as being interesting, but never with enough value to advance to category 1 or 2 |
| **4** | Recognized anywhere as being of importance through its taxonomic affinity to (even minor) crop species (crop wild relatives, CWR) |
| **5** | Widely recognized as being of low value for forage or environmental use |

Table 2: Distribution of ILRI accessions in the Tropical and Sub-Tropical Forages Strategy prioritisation categories

| Prioritisation Category | Number of accessions |
|---|---|
| 1 | 6132 |
| 2 | 551 |
| 3 | 2003 |
| 4 | 771 |
| 1, 2, 3 | 13 |
| 1, 4 | 724 |
| 2, 4 | 195 |
| 3, 4 | 170 |
| 3, 5 | 94 |
| 4, 5 | 3 |
| 5 | 4203 |
| Temperate /Mediterranean | 1182 |
| Trees | 370 |
| Others | 2229 |
| **Total** | **18640** |

Table 3. Markers generated GBS studies in forage crops and percentage of the mapped markers onto the selected reference genomes.

| **Species** | **Number of accessions genotyped** | **Number of markers** | | **Number of Mapped markers (%)** | | **Reference genome** |
|---|---|---|---|---|---|---|
| | | Silico DArT | SNP | Silico DArT | SNP | |
| *Cenchrus purpureus* | 105 | 116,190 | 85,452 | 17 % | 33% | *Penissetum glaucum* (pearl millet) |
| *Cenchrus ciliaris* | 185 | 111,917 | 93,501 | 7.2 % | 16.3 % | *Setaria italica* (foxtail millet) |
| *Chloris gayana* | 94 | 93,128 | 65,529 | 0.23 % | 2.07 % | *Setaria italica* (foxtail millet) |
| | | | | 0.74 % | 5.86 % | *Eragrostis tef* (teff) |
| | | | | 0.56 % | 5.13 % | *Zoysia matrella* (Manila grass) |
| *Sesbania sesban* | 41 | 34,798 | 47,609 | 3.36 % | 5.17 % | *Glycine max* (Soybean) |

Table 4. Mean phenotypic data of accessions used for subsetting the Napier grass collection (Muktar et al., 2019)

| | Fv/Fm | | PI | | TFWPP(g/plant) | | TDWPP(g/plant) | |
|---|---|---|---|---|---|---|---|---|
| Genotype | OW | WD | OW | WD | OW | WD | OW | WD |
| ILRI_1026 | 0.72 | 0.61 | 1.7 | 1.25 | 122.9 | 58.76 | 34.62 | 19.81 |
| ILRI_14355 | 0.74 | 0.69 | 2.85 | 2.11 | 310.67 | 224.67 | 94.06 | 62.24 |
| ILRI_14389 | 0.75 | 0.72 | 3.06 | 1.77 | 194.93 | 104.16 | 56.24 | 29.25 |
| ILRI_14982 | 0.74 | 0.68 | 2.99 | 1.94 | 307.7 | 149.82 | 79.19 | 41.41 |
| ILRI_14983 | 0.73 | 0.69 | 3.41 | 1.97 | 304.76 | 190.68 | 72.94 | 50.11 |
| ILRI_14984 | 0.71 | 0.7 | 2.58 | 2.16 | 387.9 | 189.24 | 111.13 | 57.27 |
| ILRI_15357 | 0.75 | 0.73 | 3.55 | 3.25 | 294.03 | 167.24 | 81.51 | 48.42 |
| ILRI_15743 | 0.75 | 0.7 | 3.38 | 2.26 | 221.7 | 167.34 | 55.6 | 44.5 |
| ILRI_16782 | 0.74 | 0.76 | 3.25 | 4.7 | 187.87 | 139.94 | 51.21 | 35.1 |
| ILRI_16783 | 0.56 | 0.67 | 1.09 | 1.78 | 284.47 | 72.89 | 77.67 | 19.8 |
| ILRI_16784 | 0.74 | 0.69 | 3.01 | 2.06 | 184.81 | 181.25 | 49.64 | 47.83 |
| ILRI_16785 | 0.7 | 0.69 | 1.64 | 1.37 | 315.95 | 194.94 | 84.61 | 63.68 |
| ILRI_16786 | 0.71 | 0.72 | 1.69 | 2.78 | 322.39 | 203.85 | 94.42 | 61.98 |
| ILRI_16787 | 0.74 | 0.68 | 2.14 | 1.38 | 262.44 | 83.55 | 72.11 | 23.94 |
| ILRI_16788 | 0.69 | 0.69 | 1.27 | 1.61 | 142.69 | 145.63 | 39.17 | 44.41 |
| ILRI_16789 | 0.71 | 0.68 | 2.18 | 2.44 | 342.49 | 201.72 | 97.59 | 59.6 |
| ILRI_16790 | 0.72 | 0.66 | 2.84 | 2.37 | 76.02 | 37.01 | 17.99 | 9.97 |
| ILRI_16791 | 0.75 | 0.7 | 3.77 | 2.71 | 291.99 | 313.16 | 79.01 | 87.85 |
| ILRI_16792 | 0.73 | 0.68 | 2.35 | 2.38 | 347.47 | 291.4 | 102.81 | 83.71 |
| ILRI_16793 | 0.73 | 0.73 | 3.25 | 3.71 | 375.99 | 181.42 | 111.22 | 50.12 |
| ILRI_16794 | 0.77 | 0.73 | 5.37 | 4.82 | 264.5 | 161.28 | 75.71 | 48.91 |
| ILRI_16795 | 0.74 | 0.71 | 2.79 | 3.02 | 322.89 | 203.99 | 93.51 | 60.72 |
| ILRI_16796 | 0.76 | 0.71 | 4.71 | 2.43 | 135.45 | 65.7 | 63.1 | 19.53 |
| ILRI_16797 | 0.7 | 0.69 | 2.36 | 2.26 | 13.78 | 47.71 | 3.29 | 12.06 |
| ILRI_16798 | 0.73 | 0.71 | 2.58 | 2.53 | 325.68 | 236.78 | 88.49 | 69.74 |
| ILRI_16799 | 0.72 | 0.7 | 1.91 | 1.75 | 120.12 | 64.36 | 29.91 | 17.15 |
| ILRI_16800 | 0.75 | 0.7 | 2.9 | 2.11 | 413.65 | 251.75 | 127.19 | 74.77 |
| ILRI_16801 | 0.72 | 0.72 | 1.72 | 2.71 | 434.76 | 202.42 | 126.66 | 58.48 |
| ILRI_16802 | 0.74 | 0.71 | 3.6 | 2.85 | 276.79 | 287.27 | 75.02 | 77.49 |
| ILRI_16803 | 0.7 | 0.7 | 1.69 | 1.83 | 380.05 | 200.15 | 111.7 | 62.73 |
| ILRI_16804 | 0.73 | 0.71 | 2.43 | 1.71 | 391.56 | 82.84 | 101.7 | 22.05 |
| ILRI_16805 | 0.73 | 0.72 | 2.49 | 2.33 | 39.55 | 46.94 | 8.36 | 13.58 |
| ILRI_16806 | 0.73 | 0.67 | 2.93 | 2.32 | 416.31 | 155.13 | 117.92 | 51.68 |
| ILRI_16807 | 0.7 | 0.69 | 1.96 | 2.49 | 183.95 | 137.32 | 47.14 | 35.78 |
| ILRI_16808 | 0.75 | 0.71 | 3.87 | 3.11 | 104.16 | 65.47 | 31.6 | 19.18 |
| ILRI_16809 | 0.73 | 0.64 | 2.93 | 2.29 | 161 | 107.06 | 47.19 | 32.36 |
| ILRI_16810 | 0.69 | 0.72 | 2.11 | 2.25 | 185 | 128.67 | 40.28 | 38.37 |
| ILRI_16811 | 0.74 | 0.68 | 2.91 | 1.75 | 291.16 | 108.54 | 78.34 | 30.91 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ILRI_16812 | 0.69 | 0.73 | 2.06 | 4.15 | 190.4 | 86 | 52.28 | 21.98 |
| ILRI_16813 | 0.74 | 0.64 | 5.25 | 1.5 | 137.93 | 153.37 | 26.53 | 40.48 |
| ILRI_16814 | 0.7 | 0.71 | 2.39 | 3.57 | 291.55 | 139.95 | 73.11 | 36.63 |
| ILRI_16815 | 0.75 | 0.64 | 3.9 | 1.23 | 239.01 | 100.2 | 65.49 | 27.11 |
| ILRI_16816 | 0.76 | 0.76 | 3.75 | 4.94 | 118.12 | 62.08 | 34.35 | 17.49 |
| ILRI_16817 | 0.75 | 0.69 | 3.56 | 1.72 | 198.56 | 82.35 | 51.45 | 23.84 |
| ILRI_16818 | 0.75 | 0.71 | 3.72 | 2.5 | 109.2 | 96.1 | 31.71 | 29.81 |
| ILRI_16819 | 0.75 | 0.71 | 5.02 | 1.88 | 366.63 | 266.45 | 104.47 | 73.15 |
| ILRI_16821 | 0.72 | 0.72 | 2.5 | 2.97 | 123.08 | 106.05 | 33.53 | 31.03 |
| ILRI_16822 | 0.72 | 0.74 | 2.28 | 3.24 | 96.01 | 103.92 | 26.19 | 31.8 |
| ILRI_16834 | 0.75 | 0.73 | 3.1 | 2.47 | 194.79 | 53.76 | 54.93 | 15.97 |
| ILRI_16835 | 0.71 | 0.69 | 1.92 | 2.43 | 137.03 | 44.21 | 33.24 | 12.76 |
| ILRI_16836 | 0.76 | 0.69 | 2.63 | 1.84 | 306.78 | 150.31 | 71.3 | 47.23 |
| ILRI_16837 | 0.73 | 0.66 | 2.55 | 1.11 | 225.16 | 164.55 | 60.73 | 41.64 |
| ILRI_16838 | 0.75 | 0.75 | 3.89 | 3.74 | 131.98 | 111.62 | 33.96 | 31.83 |
| ILRI_16839 | 0.7 | 0.7 | 2.22 | 2.56 | 411.17 | 128.22 | 94.35 | 35.18 |
| ILRI_16840 | 0.71 | 0.68 | 1.89 | 1.11 | 241.16 | 104.71 | 59.81 | 29.37 |
| ILRI_16902 | 0.77 | 0.72 | 3.97 | 3.23 | 207.22 | 173.02 | 61.07 | 50.36 |
| ILRI_18438 | 0.7 | 0.73 | 2.79 | 3.28 | 248.14 | 214.48 | 70.41 | 57.58 |
| ILRI_18448 | 0.7 | 0.72 | 3.65 | 3.62 | 141.79 | 76.48 | 42.49 | 23.24 |
| ILRI_18662 | 0.76 | 0.63 | 3.96 | 1.89 | 4.55 | 38.94 | 1.7 | 7.92 |
| BAGCE_100 | 0.67 | 0.73 | 2.25 | 2.77 | 240.39 | 179.33 | 65.59 | 61.91 |
| BAGCE_17 | 0.71 | 0.59 | 2.07 | 1.34 | 144.31 | 108.33 | 38.16 | 27.61 |
| BAGCE_30 | 0.69 | 0.75 | 1.92 | 3.57 | 368.22 | 254.29 | 89.49 | 70.84 |
| BAGCE_343 | 0.76 | 0.77 | 3.36 | 5.03 | 241.19 | 111.35 | 65.92 | 31.58 |
| BAGCE_53 | 0.74 | 0.72 | 2.73 | 2.54 | 344.48 | 91.93 | 83.73 | 24.93 |
| BAGCE_81 | 0.74 | 0.68 | 2.73 | 1.56 | 229.01 | 62.61 | 56.36 | 17.59 |
| BAGCE_86 | 0.73 | 0.71 | 3.32 | 2.34 | 274.05 | 123.33 | 69.15 | 32.93 |
| BAGCE_90 | 0.72 | 0.71 | 2.69 | 2.36 | 385.86 | 133.88 | 93.74 | 35.84 |
| BAGCE_97 | 0.74 | 0.7 | 3.06 | 2.32 | 191.56 | 126.51 | 45.31 | 34.67 |
| **Maximum** | **0.77** | **0.77** | **5.37** | **5.03** | **434.76** | **313.16** | **127.19** | **87.85** |
| **Minimum** | **0.56** | **0.59** | **1.09** | **1.11** | **4.55** | **37.01** | **1.7** | **7.92** |
| **Average** | **0.73** | **0.70** | **2.86** | **2.49** | **239.40** | **139.71** | **65.01** | **39.86** |

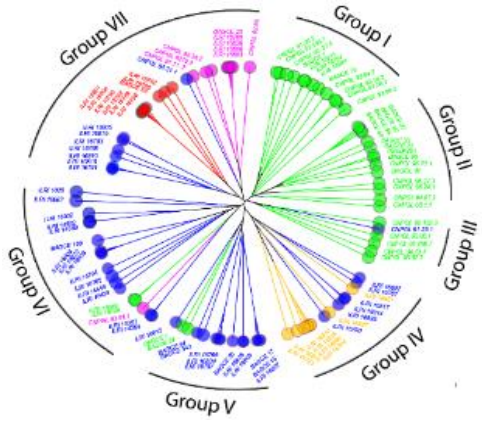
Fv/Fm =quantum efficiency of photosystem II, PI = performance index, TDWPP=total dry weight per plot, TFWPP= total fresh weight per plot.

Table 5. Napier grass subsets representing the diversity in the collection from the ILRI genebank for evaluation under irrigated and water deficit conditions (Muktar et al., 2019)

| Optimal water | | | | Water-deficit | | | |
|---|---|---|---|---|---|---|---|
| **NAME** | **Species** | **Origin** | **Collection** | **NAME** | **Species** | **Origin** | **Collection** |
| ILRI_1026* | *C. purpureus* | Burundi | ILRI | ILRI_1026* | *C. purpureus* | Burundi | ILRI |
| ILRI_16840* | *C. purpureus x P. glaucum* | Zimbabwe | ILRI | ILRI_14389 | *C. purpureus* | Nigeria | ILRI |
| ILRI_14982 | *C. purpureus x P. glaucum* | USA | ILRI | ILRI_14983 | *C. purpureus* | USA | ILRI |
| ILRI_14984 | *C. purpureus* | USA | ILRI | ILRI_16811 | *C. purpureus* | USA | ILRI |
| ILRI_16793* | *C. purpureus* | Cuba | ILRI | ILRI_16791 | *C. purpureus* | Swaziland | ILRI |
| ILRI_16794 | *C. purpureus* | Mozambique | ILRI | ILRI_16793* | *C. purpureus* | Cuba | ILRI |
| ILRI_16814* | *C. purpureus* | USA | ILRI | ILRI_16816 | *C. purpureus* | USA | ILRI |
| ILRI_16839 | *C. purpureus* | Zimbabwe | ILRI | ILRI_16796 | *C. purpureus* | Zimbabwe | ILRI |
| ILRI_16819 | *C. purpureus* | USA | ILRI | ILRI_16806* | *C. purpureus* | USA | ILRI |
| ILRI_16797 | *C. purpureus* | Zimbabwe | ILRI | ILRI_16782 | *C. purpureus* | Tanzania | ILRI |
| ILRI_16806* | *C. purpureus* | USA | ILRI | ILRI_16814* | *C. purpureus* | USA | ILRI |
| ILRI_16822 | *C. purpureus* | Malawi | ILRI | ILRI_16840* | *C. purpureus x P. glaucum* | Zimbabwe | ILRI |
| BAGCE_30* | *C. purpureus* | Brazil | EMBRAPA | BAGCE_30* | *C. purpureus* | Brazil | EMBRAPA |
| BAGCE_97* | *C. purpureus* | Brazil | EMBRAPA | BAGCE_97* | *C. purpureus* | Brazil | EMBRAPA |

*accession selected in both subsets

a




b

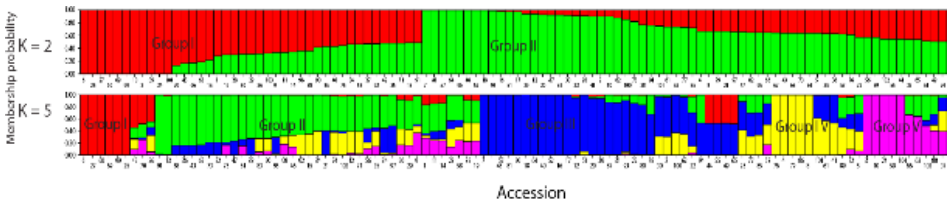



*Figure 1. Clusters of 104 Napier grass accessions using selected SNP markers: (a) UPGMA tree showing seven groups; (b) Bar plots based on the admixture model in STRUCTURE, for K = 2 and K = 5. Colours in (a) are according to the STRUCTURE analysis with k = 5. (Muktar et al., 2019).*

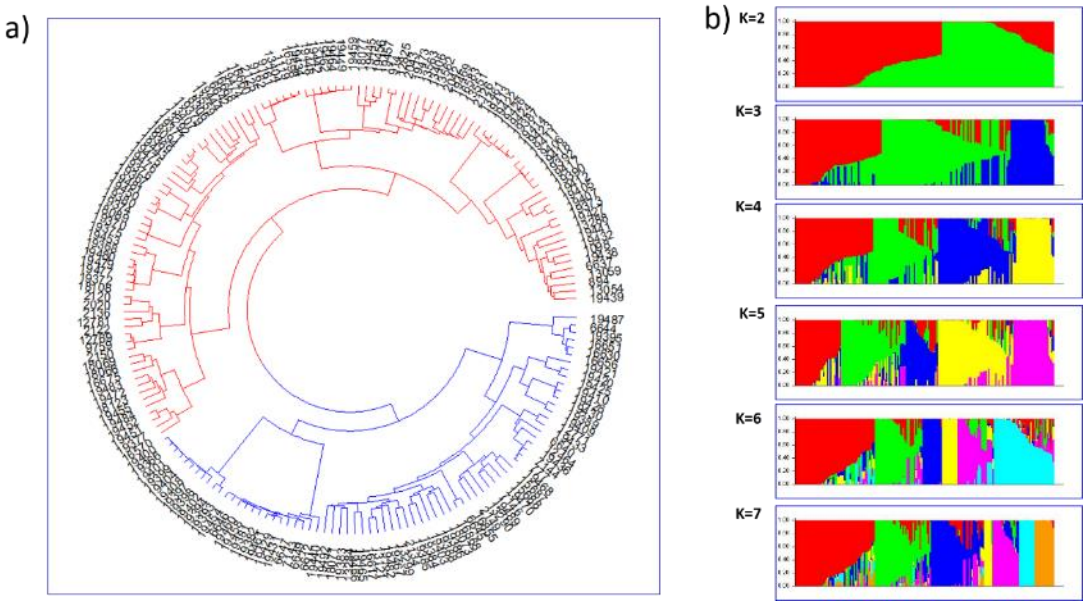



*Figure 2. Population structure of Buffel grass using 1,000 selected SNP markers with (a) cluster analysis of the 185 accessions and; (b) bar plots showing the suggested subpopulations.*

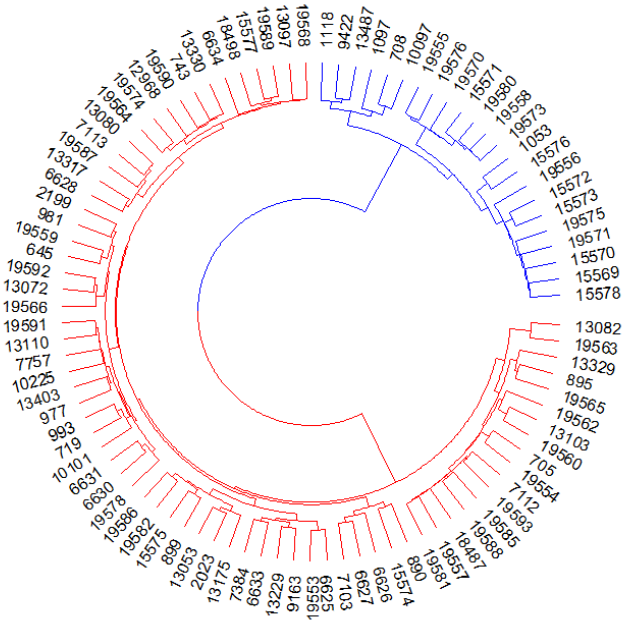



*Figure 3. Cluster analysis of the 94 Rhodes grass accessions using SNP markers clearly showing two differentiated groups.*

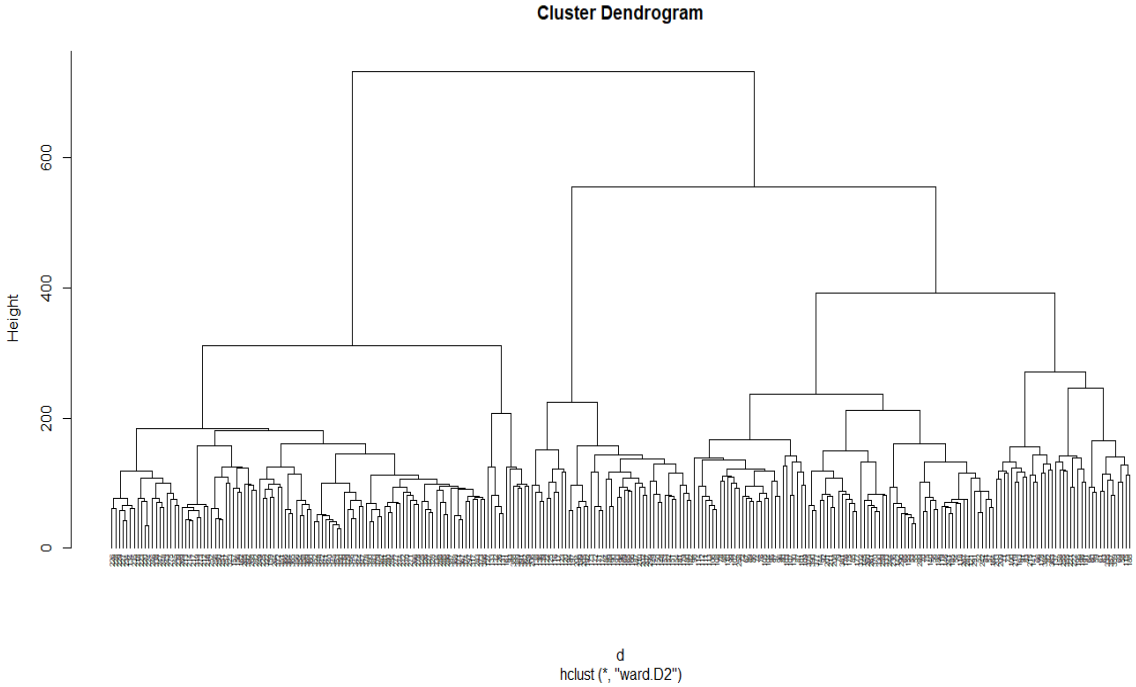



*Figure 4. Preliminary cluster analysis of the Sesbania sesban collection using SNP markers.*

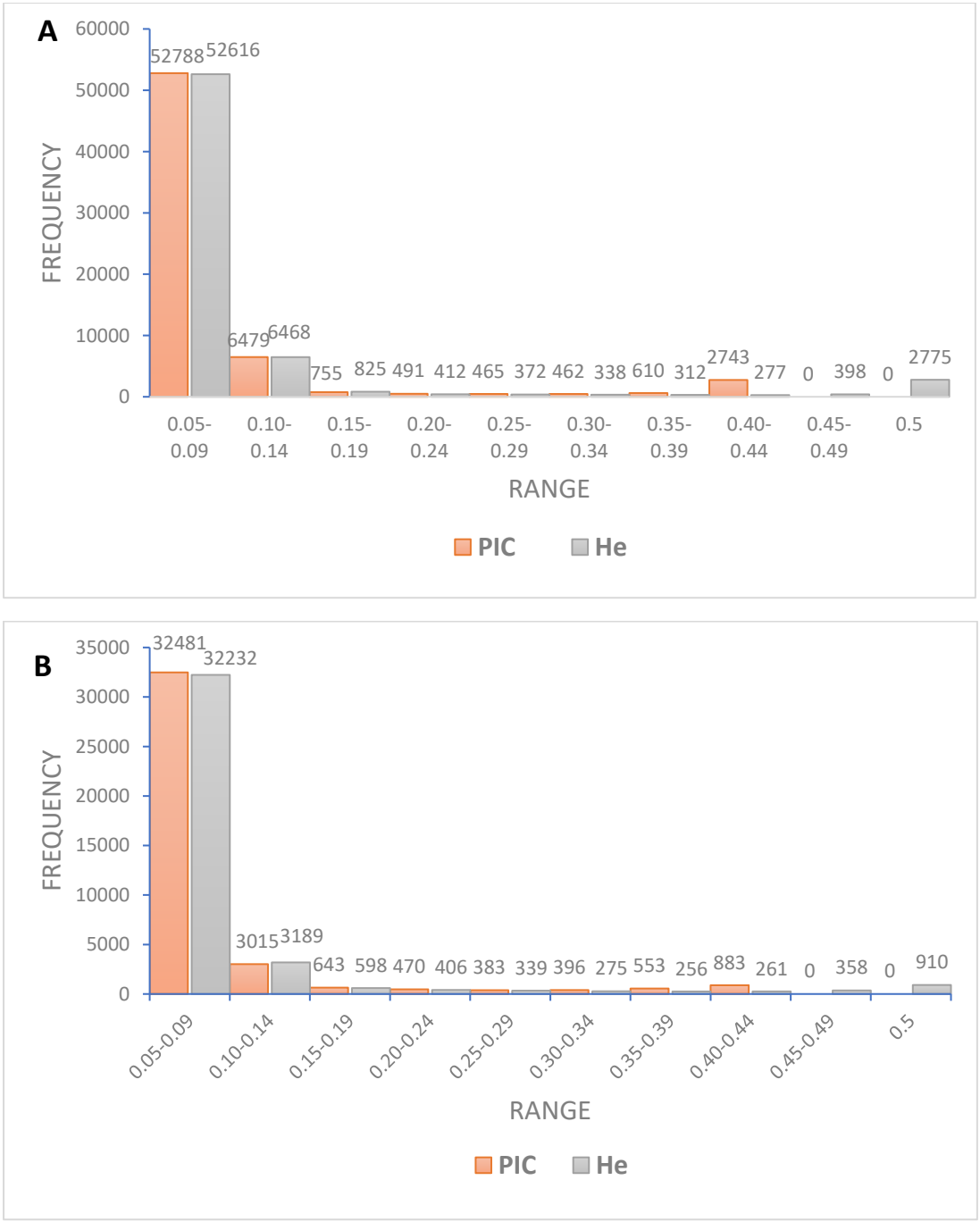



*Figure 5. Distribution of polymorphic information content (PIC) and heterozygosity (He) for the silicoDArT (A) and SNP (B) markers in Lablab purpureus.*

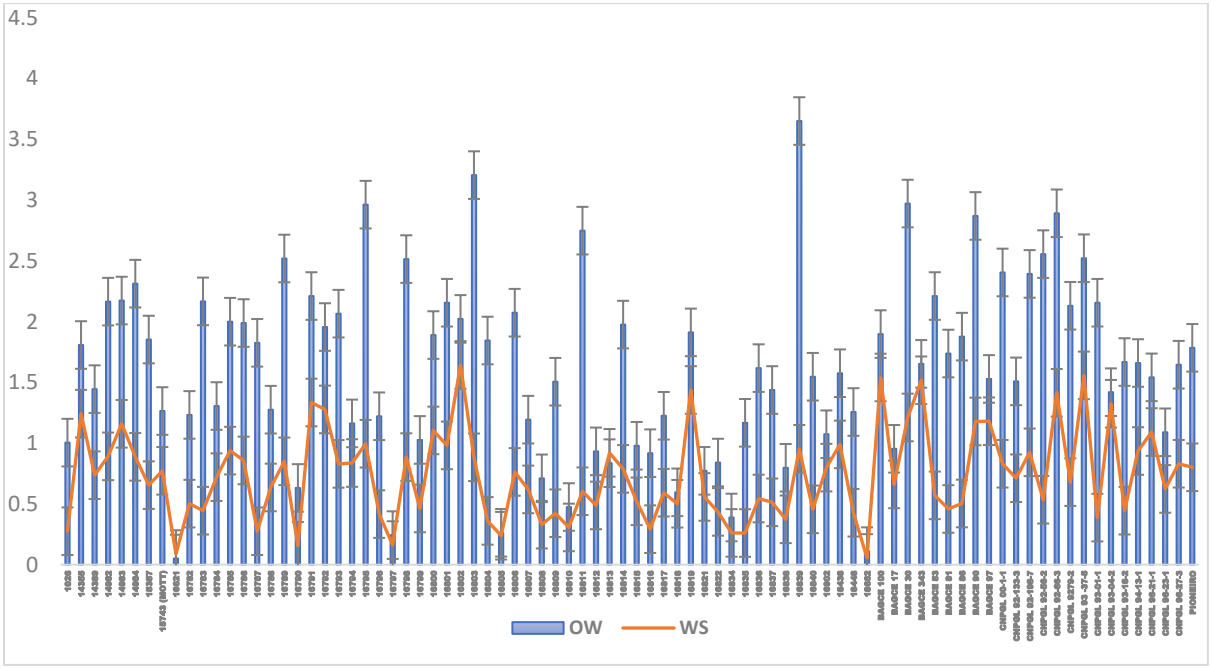



*Figure 6. Total dry weight (t/ha) of Napier grass genotypes grown under optimum water (OW) and water stress (WS) conditions across dry season harvests.*

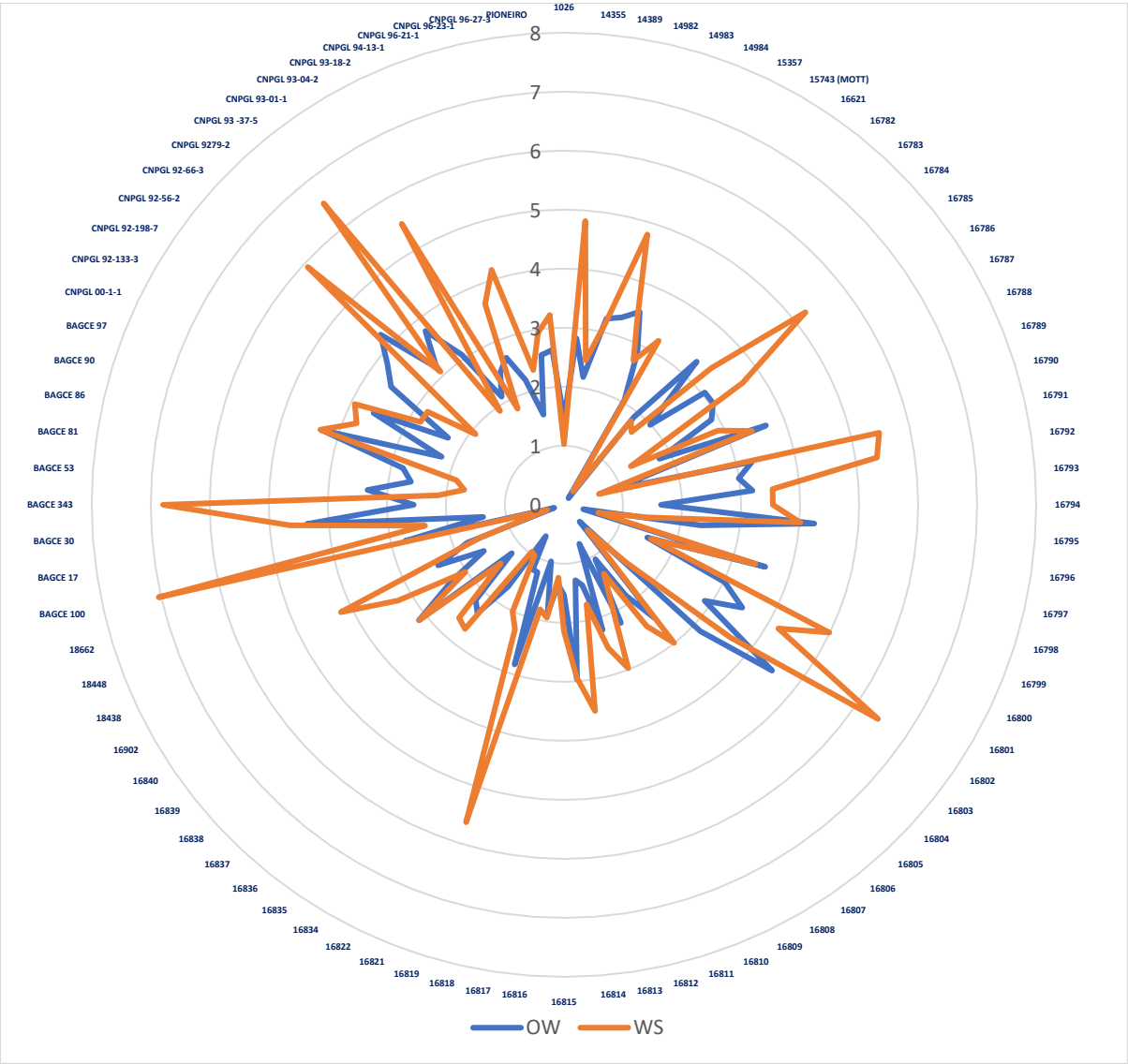



*Figure 7. Water use efficiency (g/m$^3$) of Napier grass genotypes grown under optimum water (OW) and water stress (WS) conditions across dry season harvests.*

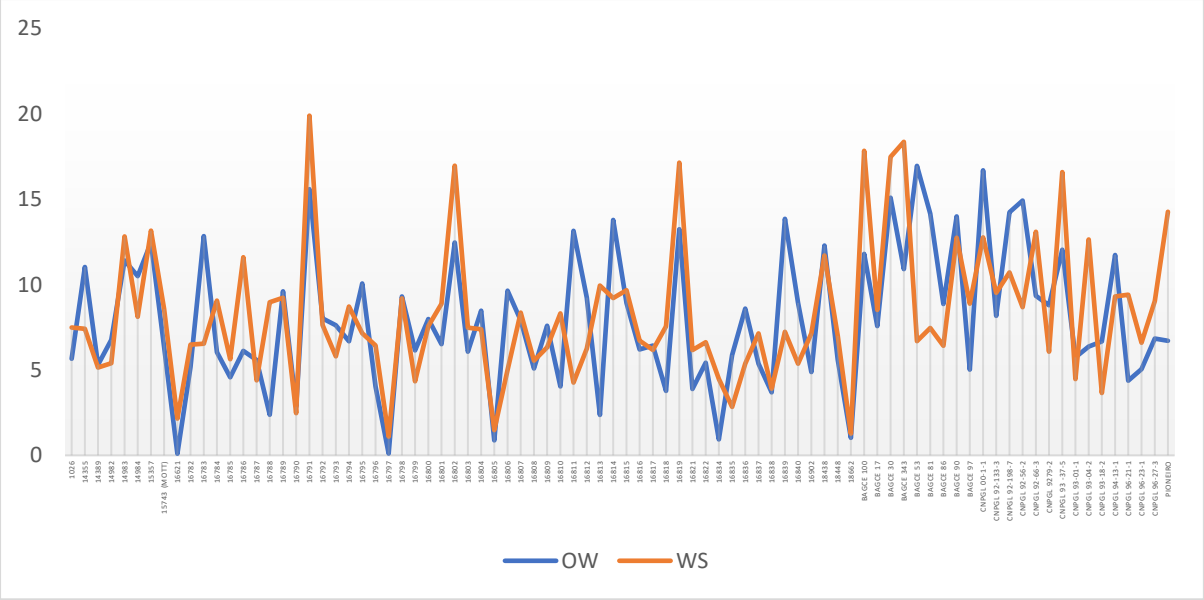



*Figure 8. Total dry weight (t/ha) of Napier grass genotypes grown under optimum water (OW) and water stress (WS) conditions across wet season harvests.*

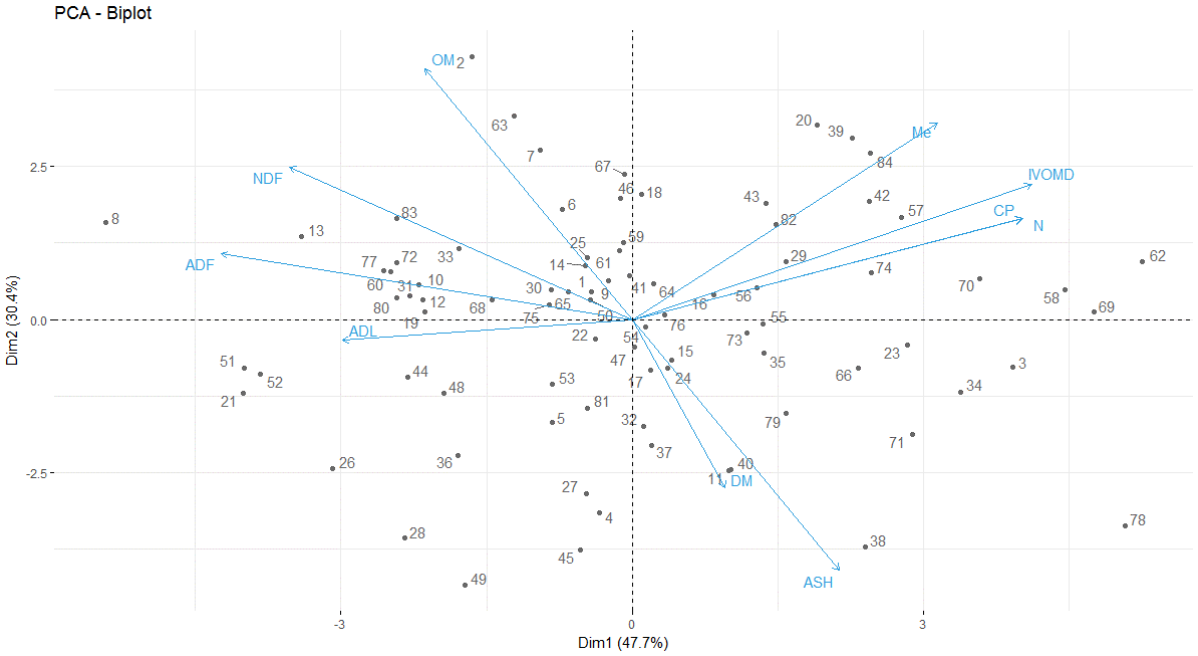



*Figure 9. PCA biplot showing feed quality traits of Napier grass accessions: Acid detergent fibre (ADF); Neutral detergent fibre (NDF); Acid detergent lignin (ADL); Metabolizable energy (Me); Organic matter (OM); Dry matter (DM); Total nitrogen Crude protein (CP); In vitro organic matter digestibility (IVOMD).*