

Internal report of activities

Exploring options of genomic selection in Napier grass

Started a genotyping-by-sequencing (GBS) initiative which generated approximately 148,000 genome wide silicoDArT and SNP markers across our collection. These were mapped to the pearl millet genome (used as a reference in the absence of a Napier genome) to determine chromosome position. The map information was used to select a representative subset of high-resolution markers for diversity analysis and to run marker-trait associations to predict genomic regions associated with important agronomic and morphological traits. A selection of 1,000 highly polymorphic markers were used for the diversity analysis and to determine population structure. The diversity analysis revealed a substantial amount of genetic variation was captured in the collection, demonstrating the suitability of the population for further genetic studies, and the structure analysis revealed the presence of up to 4 sub populations.

A set of historical phenotype data collected on 47 accessions from the collection for the traits: plant height; stem thickness; internode length; total stem dry weight; total leaf dry weight, and; yield, were combined with these data. A marker-trait association analysis was performed and detected a total of 13 molecular markers associated with the trait 'plant height', which in turn is highly correlated with 'stem dry weight' and 'yield'. These markers are candidates for use for the selection of improved Napier varieties in the future.

A drought trial, consisting of 84 genotypes, has been initiated to identify accessions with good agronomic performance, nutritional quality and water use efficiency by comparing performance under fully irrigated and water limited conditions. The trial was established at a mid-altitude site (1800 m.a.s.l) in Bishoftu, Ethiopia, with alfisol/vertisol soils and an annual mean rainfall of around 790 mm falling in a bimodal pattern. The trial was set up as a modified augmented design with internal checks and consists of four blocks (two fully irrigated and two water limited). Each block consists of 96 individual plants, 84 genotypes and 12 internal checks, made up of selected best bets, distributed across the block. By employing an augmented design we were able to accommodate as many of the accessions as possible without compromising the statistical power of the experiment. A total of 864 leaf, stem and leaf + stem samples have been taken for dry weight measurement and nutrition analysis.

We have also been strengthening our bioinformatics capabilities. We can remotely access the ILRI-Nairobi high performance computing cluster which contains four super computers and one main login node. Using



this cluster, we can run big data analyses such as sequence mapping, variant calling, sequence annotation, multiple sequence blasting etc. We are also accessing the ILRI-Addis HPC cluster, which has medium processing power and is used for data analysis for intermediate data such as variant calling. We also have a desktop work station with 8 CPU cores, 16 GB RAM and 1 TB storage which is used for routine data analysis, such as diversity analysis and association mapping using R statistical software.