Development of Efficient Genotyping Workflow for Accelerating Maize Improvement in Developing Countries

Queen Nkem Obi

International Institute of Tropical Agriculture <https://orcid.org/0000-0002-1317-219X>

Abebe Menkir

International Institute of Tropical Agriculture

Deborah F. Babalola

International Institute of Tropical Agriculture

Melaku Gedil (✉ M.Gedil@cgiar.org)

International Institute of Tropical Agriculture <https://orcid.org/0000-0002-6258-6014>

Methodology

Keywords: Molecular breeding, KASP, Genotyping workflow, Marker-assisted selection, Quality Control, National Agricultural Research Systems (NARS), Developing countries

Posted Date: December 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-129326/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Molecular breeding has been recognized as one of the pillars to accelerate the rate of genetic gain in crop improvement towards meeting the need to feed an ever-growing world population. Establishing low-cost and flexible genotyping platforms in small and public laboratories and regionally can stimulate the application of molecular breeding in developing countries where many plant breeding projects require low to medium density markers for genomics-assisted selection and quality control (QC) activities.

Results

Here we present an optimization of the entire genotyping workflow (from sample collection to genotyping and data analysis) to accelerate the QC and genomic-assisted selection process, which can readily be adopted by National agricultural research system (NARS) partners in developing countries to fast-track molecular marker-based genotyping for crop improvement. An in-house KASP genotyping system combined with optimized sample collection and DNA preparation processes expedited the genotyping workflow from over five weeks (when outsourcing) to approximately three weeks (5 days per week) for a total of 637 samples. The QC experiment using a subset of 28 KASP SNPs validated for maize revealed the genetic identity of 4 maize varieties taken from 5 seed sources. Another subset of 10 KASP SNPs was sufficient in verifying the parentage of 388 F₁ lines. The marker-based selection of high PVA maize lines identified nine lines harboring the favorable allele of the *crtRB1* gene, which could serve as donor lines for the maize PVA breeding program. The ongoing marker-assisted backcrossing experiment to introgress resistance to aflatoxin accumulation in elite tropical maize lines has so far identified twenty-four maize lines harboring the favorable alleles associated with resistance to aflatoxin accumulation, and are currently undergoing field evaluation. The optimized genotyping workflow has so far generated over 1700 datapoints.

Conclusion

The result of this work could serve as a prototype to fast-track maize improvement activities of IITA's MIP group, and facilitate DNA fingerprinting for adoption tracking of improved crop varieties. It will also provide public and small laboratories in developing countries with the knowledge of efficient genotyping workflow to accelerate crop improvement activities.

Background

Agriculture is the mainstay of millions of low-income households in Sub-Saharan Africa (SSA). However, productivity is way below the yield potential of significant crops because several interacting factors contribute to the yield reduction. The paucity of nutritionally-improved and resilient varieties is a crucial constraint that can be mitigated by the rapid development of cultivars adapted to specific agroecology zone. The current yield gain trend in major food crops has shown that relying on conventional breeding

alone is insufficient to meet the food need of an estimated 9 billion people in 2025 [1]. There is a need to accelerate genetic gain by deploying new breeding strategies [2, 3]. This need has led to the scientific community's massive investment in developing genomic resources and support systems to provide useful tools to accelerate the breeding process and develop high yielding cultivars [4]. In the form of marker-assisted selection, including genomic selection (GS), molecular markers are one of such genomic tools that are commonly used for molecular breeding. The availability of DNA-based molecular markers has provided a means for better discrimination and faster determination of genetic purity of inbreds and hybrid verification. They are useful for varietal identification, diversity analysis, and genomics-assisted breeding.

Molecular markers are advantageous over morphological markers because they are unaffected by environmental conditions or the plant's physiological stage.

Simple sequence repeats (SSR) have been the preferred marker for assessing genetic purity because of their multi-allelic and co-dominant nature, high polymorphism, and wide distribution throughout the genome [5]. However, the latest advances in molecular biology technologies have stressed the use of single nucleotide polymorphic (SNP) markers [6]. SNP markers have become powerful tools for genetic discovery research and marker-assisted breeding for crop improvement [7, 8]. They are abundant in the genome, co-dominant in nature, require a simple documentation system, low genotyping cost, and amenable to high throughput analysis [9]. Besides their cost-effectiveness, speed, and flexibility, a unique advantage of SNP markers is the availability of various medium to high throughput genotyping platforms that address different breeding needs for varying marker densities and cost per sample. High throughput platforms such as the sequence-based Next-Generation Sequencing (NGS) technologies (genotyping by sequencing, GBS) [10], array-based sequencing technologies (diversity arrays technology, DArTseq, and Sequenom) [11], and chip-based technologies have made SNPs very popular for various applications. The application of molecular markers in crop breeding has significantly impacted the development of major food crops, including maize [12, 13, 14, 15, 16]. The Maize Improvement Program (MIP) of the International Institute of Tropical Agriculture (IITA) is using molecular markers in various stages of the breeding pipeline including, quality control (QC) for hybrid verification, fingerprinting parental lines to assess genetic diversity, and using DNA fingerprinting for adoption tracking of improved varieties [17]. Breeders start their work by assessing the extent of genetic variation in their germplasm collection for the valid selection of parents for various breeding objectives. Marker-based diversity analysis is commonly used to assess genetic variation and heterotic grouping [17]. Once crosses are made between selected parents, hybrid verification can be performed with molecular markers rapidly and at a low cost. Progeny selection based on various molecular breeding schemes requires a different set of markers. For instance, GS is done with NGS-based genome-wide markers such as GBS or DArTseq, while marker-assisted backcrossing (MABC) or forward breeding utilizes selected predictive markers strongly linked to priority traits.

However, various bottlenecks have hindered the substantial impact of molecular breeding for crop improvement, particularly in developing countries [18, 19]. The major limiting factors are lack of

infrastructure and capacity for genomics resources and poor information flow, resulting in reduced access to operational and decision support tools [17]. Because private companies in developed countries own the proprietary rights to many of the emerging genomics resources and systems, it is difficult for public research sectors, non-profit research institutes, and small laboratories in developing countries to have direct access. These challenges are being curbed by various international initiatives such as the Excellence in Breeding (EiB) platform, which coordinates its activities with the Genomic and Open-source Breeding Informatics Initiative (GOBii), and High Through-Put Genotyping (HTPG). There is also the Integrated Breeding Platform (IBP)-hosted Generation Challenge Program (GCP) and the Breeding Management System (BMS) [20], which target the development and adoption of molecular breeding in developing countries. These and other Consultative Group-hosted initiatives and platforms galvanize worldwide partners drawn from public, private, and governmental institutions towards the common goal of increasing agricultural productivity through efficient tools, technologies, and data management systems.

Despite the availability of many low-cost genotyping platforms and resources, it is not easy to meet the genotyping needs of many users who work on different crops and in different locations. There is the need to complement these international initiatives by providing in-house or local (regional) genotyping platforms, where possible, to accelerate the genotyping workflow. One such regional initiative in Africa is the Integrated Genotyping Support Services (IGSS) genotyping facility at BeCA/ILRI (Biosciences eastern and central Africa/International Livestock Research Institute), Kenya. Because these international initiatives are grant driven and short-term projects, it is imperative to devise a sustainable strategy for routine, cost-effective, and easily accessible genotyping service. Such a strategy will provide breeders with either outsourcing to genotyping service providers or setting up a core facility in-house.

One factor that influences breeders' choice of genotyping platform is the level of throughput. Other factors considered are the data turnaround time, ease of data analysis (available informatics), reproducibility, flexibility, and cost per datapoint or cost per sample [21, 22]. For high and ultra-high throughput markers, breeders outsource to an array- and sequenced-based genotyping service providers. These platforms are suitable for discovery applications and for approaches requiring hundreds to thousands of samples to be genotyped with tens to thousands of markers, such as Genome-wide association studies (GWAS), gene mapping, and large-scale genomic selection. They are also suitable for genotyping a few samples with many markers (multiplexing), such as genetic diversity analysis or background selection. While multiplex platforms provide higher throughput with lower reagent consumption, it limits scientists to using a multiplexed set of several thousand SNPs per assay. The high per sample cost of highly multiplexed platforms can be problematic for crop improvement applications, which usually requires low- to medium-density markers, and the time consumed for initial assay development [22]. They are also costly in informatics needs and presently produce datasets with a significant percentage of missing data [23]. Multiplex platforms are not suitable for small MAS projects, which genotype small-volume samples with few target markers. For example, in MABC, few markers may be used to introgress a particular gene. Also, for genetic identity and purity analysis of inbreds, a small volume of samples is genotyped with tens to a hundred markers. In hybrid verification studies, a subset

of carefully developed SNP markers can be used. Also, in variety identification and adoption tracking studies, where a library of released varieties fingerprinted with a selected set of SNPs is required, the multiplex platforms may not be cost-effective. For these low- to mid-density genotyping approaches, a uniplex SNP genotyping platform is appropriate.

Uniplex genotyping assays are small-scale high throughput genotyping systems that are ideally flexible regarding assay design, ease of running, and cost-effectiveness. These systems provide plant breeders with the flexibility to mix and match different SNPs for a given sample set. They allow breeders to use a smaller subset of informative SNPs such as functional SNPs and trait-specific haplotypes, thereby eliminating the generation of unintended datapoints when using fixed array SNPs. There exists a range of uniplex SNP genotyping assays developed, especially for the medical field. However, the most competitive uniplex systems that have been successfully applied in crop improvement research are TaqMan [24, 25, 26], KASP [22, 27], Amplifluor [28], and rhAmp [29] assays. These uniplex genotyping systems vary in reaction chemistry, detection method, and reaction format. Uniplex systems can either be outsourced or installed in-house.

This study utilized the Kompetitive allele-specific PCR (KASP) assay because it is one of the most used assays among plant breeders and biologists. KASP is an endpoint PCR-based SNP genotyping method from KBiosciences, now LGC Biosearch Technologies, UK. KASP uses fluorescently labeled allele-specific primers for the bi-allelic discrimination of SNPs and INDELs. KASP was developed to reduce cost, mainly from probe design, and improve genotyping efficiency and has become a preferred alternative to TaqMan [30, 22]. The KASP genotyping system has been successfully applied in crops such as maize [22, 31], wheat [21, 27, 32], rice [33], Soybean [34], peanut [35], amongst others. KASP has developed into a global benchmark technology for genotyping crop plants [22, 35, 36, 37, 38] following the validation of KASP markers across crops of global importance (such as maize- 1250 markers, wheat-1864 markers, and rice-2015 markers) by the Generation Challenge Program of the Integrated Breeding Platform [20]. CIMMYT has successfully utilized the 1250 maize KASP markers for various genetic applications, including quantitative trait loci (QTL) mapping, Marker-assisted recurrent selection (MARS), allele mining, and QC analysis [22].

The Maize Improvement Program of IITA has generated over 17,000 datapoints using KASP for different genotype analyses, including QC and MAS. However, some bottlenecks in the genotyping workflow slow down the genotyping process, delaying crop improvement: (1) method of sample collection and processing (freeze-drying, grinding, and sample prep), (2) DNA quality and quantity analysis, (3) DNA-based genotyping, and (4) data analysis. Gedil and Menkir (2019) provided a thorough review of the MIP group's molecular marker-based crop improvement activities. However, reports are lacking in accelerating the entire genotyping process by minimizing these bottlenecks and providing a cost-effective genotyping workflow suitable for small scale breeders and laboratories in developing countries. This work aimed to show how deploying an in-house genotyping platform alongside optimized molecular techniques can accelerate maize improvement activity for a quick result turnaround—providing a workflow that can also

be adopted by less sophisticated breeding laboratories in developing countries for efficient crop improvement.

Hence, our objective was to develop an efficient and cost-effective genotyping workflow that removes or minimizes significant bottlenecks in the molecular marker-based genotyping process viz: optimized leaf sample collection, high-quality DNA extraction, optimized high throughput DNA quantitation, and low-cost in-house KASP genotyping system; and to deploy the workflow to ascertain genetic identity among adopted maize varieties and verify parentage of F₁ maize lines using a panel of QC SNP markers, and trait markers for MAS of high PVA and aflatoxin resistant maize lines..

Results

Optimizing in-house genotyping workflow

Our laboratory's routine sampling procedure, from manual plant tagging to leaf storage and cutting of leaf discs for DNA extraction, spans seven days. Here we introduce methods that expedite every step of the procedure from plant sampling to DNA quantification and ensures a good sample tracking system. Firstly, barcoding software, barcode readers, barcode labels, and barcode printers were introduced to facilitate sample tracking and data management. Waterproof/tear-proof tags and labels designed using BarTender barcoding software (Seagull Scientific) were printed upon using ZT230 Printer (Zebra, USA) and attached to plants before sample collection. Plate maps created in the BarTender software were linked to the sample location on the field and in the lab storage facility. Next, young plant leaf tissues were collected by punching leaf discs directly into the 96-well 1.2 mL polypropylene cluster tubes in wet-ice cooler bags, which reduced the sampling time and the time required for freeze-drying. Following this process, the total time of sampling and processing before DNA extraction of the 637 leaf samples was reduced from seven to four days.

Nucleic acid extraction is one of the rate-limiting steps in many DNA-based procedures, especially in routine genotyping in a breeding pipeline where DNA needs to be extracted from a large number of plant samples. The KASP genotyping system requires high-quality DNA obtained using the widely-used standard DNA extraction protocols such as the cetyl trimethyl ammonium bromate (CTAB)- based and the sodium dodecyl sulfate (SDS)-based protocols and their modifications. Here we used a modified CTAB-based DArT DNA extraction protocol, which includes reagents such as polyvinylpyrrolidone-40 (PVP-40), sarcosyl sorbitol in the extraction buffer. These reagents help remove common DNA contaminants such as polyphenols and polysaccharides, which hinder PCR. Two main modifications were made to the DArT protocol. Firstly, we used 96-well 1.2 ml cluster tubes with strip caps instead of the described 15 ml Sarsted and 2 ml Eppendorf tubes.

Consequently, the reagents' volume was optimized to extract maize DNA from a smaller amount of leaf tissue (16–20 leaf-discs, 4.0 mm). Secondly, like most conventional DNA extraction methods, the DArT protocol requires the manual grinding of fresh plant tissue in liquid nitrogen using mortar and pestle.

However, we used freeze-dried leaf tissues, and grinding was performed using an automatic high throughput tissue homogenizer, Geno/Grinder 2010, which can grind 384 samples (4 × 96-sample plates) in two minutes. Using our method, DNA was extracted from 637 maize plant leaf tissues in three days (8 hr per work-day).

The UV absorbance method of the FLUOstar Omega microplate reader (BMG LABTECH) was used to measure the concentration and purity of the 637 DNA samples. By using this method, the DNA samples were quantified in less than 20 minutes. The DNA purity (A_{260}/A_{280} ratio) ranged from 1.7 to 2.0, with an average concentration of 1915 ng/μl.

To optimize and use the KASP system for in-house genotyping, KASP assays and allele-calling software (KlusterCaller) were purchased from LGC Biosearch Technologies, UK. The amplification parameters on the compatible PCR (GeneAmp 9700) and real-time PCR machines (Roche LightCycler 480 II) were optimized. Microtiter 96- and 384-well plates compatible with the different machines were acquired from Roche, Germany. We also optimized the FLUOstar Omega microplate reader for fluorescence measurement of amplified products following the manufacturer's manual. Following KASP set up, we genotyped plant samples for QC and MAS in-house, with low to medium density markers.

Application of the optimized genotyping workflow

The optimized workflow was subsequently applied in two QC and MAS experiments. The QC genotyping ensured on-time identification of errors and mislabeling in inbred lines and false hybrids in F₁ maize breeding populations. The MAS experiments reduced the breeding cycles in a maize breeding program.

The genetic identity test was performed using a subset of 28 KASP assays to identify a set of twenty well-adapted maize varieties originating from IITA but regenerated at different locations (suffice 1, 2, 3, 4, and IITA). Genetic identification was performed using the original maize varieties' molecular marker profile and the genetic distance approach. Seed sources having < 5% genetic distance were considered the same. The genetic distance among the four original maize lines and between lines from IITA and each of the four seed sources was calculated using PowerMaker 3.25 statistical software. The genetic distance among the four designation lines from IITA ranged from 0.0563 to 0.1239, indicating that the lines were different. The genetic distance among the different seed sources of the same line designation was: 0.0105–0.0314 (SAMMAZ15), 0.0105–0.0418 (SAMMAZ16), 0.0105–0.0837 (SAMMAZ27), and 0.000–0.0563 (SAMMAZ39). SNPviewer image showed that the four designated lines from three seed sources grouped with lines from IITA (Fig. 1). The dendrogram image (Fig. 3) also showed a grouping of different seed sources of the same line designation except for SAMMAZ39-1, SAMMAZ16-3, and SAMMAZ27-4. This clustering pattern indicates that all seed of the same line has a common origin. SAMMAZ27-4 appeared to be genetically distant from SAMMAZ27-IITA by 0.0837. However, it grouped with SAMMAZ15 (Fig. 3: blue circle), suggesting a possible mislabeling or mix-up of seeds during harvesting and storage. SAMMAZ39-1 grouped with SAMMAZ16 (Fig. 3: red circle), indicating possible pollen contamination or seed mix-up during handling.

In another QC experiment using our workflow, we screened two groups of F_1 plants for hybrid verification, including their parental inbred lines, with 10 KASP SNP markers. The parental inbred lines were screened with an initial 50 KASP SNP taken from a defined panel of maize QC KASP markers to identify polymorphic markers. Only 10 KASP markers were polymorphic between the parental lines and were used to screen the F_1 plants to verify their parentage. Cluster analysis of Group1 F_1 s (Fig. 2) grouped the genotypes into three clusters. The heterozygous F_1 progenies were in the middle of the plot, and the homozygous parental inbred lines diverged from each other (along the X- and Y-axis of the plot) for all ten markers. This clustering pattern indicates that the F_1 progenies were true hybrids. Similar clustering was observed in Group 2 F_1 s except in Set- 3b, where 38 F_1 s grouped with parental genotypes. The homozygous F_1 s could be due to contamination from foreign pollens during the crossing in the field or seed mix-up during storage or planting.

By applying our workflow in MAS projects, we performed multiple field selections annually, which accelerated the maize breeding process. For instance, in the MABC project, a set of trait-specific KASP SNPs was used to select 24 BC_1S_2 maize lines potentially introgressed with resistance to aflatoxin accumulation after four selection-cycles in less than two years. Potentially introgressed lines are now undergoing field evaluation for resistance to aflatoxin accumulation. The result of the MAS of high PVA lines, on the other hand, identified nine out of 70 inbred maize lines harboring favorable alleles of the *crtRB1* gene, which is associated with high PVA content in maize.

Table 1.

Comparing Cost for KASP Genotyping in-house versus outsourcing

Comparing cost for KASP Genotyping in-house versus outsourcing				
KASP Genotyping In-House				
Reagents/Consumables	Quantity	Unit Price (\$)	Price for 637 Samples	
384-well PCR Plate	50 plates/pack + Foil	269.24		
For 637 samples	2 plates +2 Foil		10.77	
Pipette Tips	96 tips/box	7.50		
For 637 samples	7 boxes		52.50	
KBD Primer Mix	1.2 mL tube (2500 *10 µL reactions)	72.43		
For 637 samples	44.59 µL (0.07 * 637 samples)		2.69	
KASP Master Mix, Low Rox	2.5 mL tube (500 * 10 µL reactions)	175.56		
For 637 samples	1592.5 µL (2.5 * 637 samples)		111.83	
Total Genotyping Cost			<u>177.79</u>	
KASP Genotyping Outsourced				
Reagent/Service	Quantity	Unit Price (\$)	Price for 637 Samples (\$)	
			KOD (New assay)	KBD (Panel/online assay)
Assay Setup and Validation	1 Assay	192.00		
For 637 samples	1 Assay		192.00	0.00
Assay Restock and Handling	1 Restock	80.00		
For 637 samples	1 Restock		0.00	80.00
Genotyping Service	1 Datapoint	0.33		
For 637 samples + 2 NTC	637 Datapoints		210.21	210.21
Total Genotyping Cost Ex. Works			<u>402.21</u>	<u>290.21</u>

LEGEND: KOD = KASP on Demand; KBD = KASP by Design; \$ = US Dollar; μ L = Microliter.

For higher volume of samples, price per datapoint can go as low as USD \$ 0.13.

Discussion

There are different methods of plant tissue sampling, including collecting samples in silica gel [40], NaCl/CTAB [41], alcohol [42], blotter paper, gel pack, dry ice, and liquid nitrogen [43]. These methods provide reasonably good quality and quantity of DNA for molecular marker genotyping; however, deciding which method to use is based on the number of samples and distance from field to the laboratory [43]. We routinely use wet ice in Styrofoam boxes and cooler bags. It is cost-effective and suitable for the close-proximity sample collection, and leaf samples are preserved by freeze-drying [44] before DNA extraction. We collected leaf tissues directly into 96-well extraction tubes, rather than jute bags, which means our procedure provides high throughput sampling. This sampling process also ensured that sample DNA is not degraded by prolonged exposure to moisture as obtainable in post-freeze drying cutting of leaf tissues stored in jute bags.

Our protocol aimed to extract high-quality DNA suitable for KASP genotyping from a smaller amount of leaf tissues, which reduced the cost of reagents and the time for DNA extraction. The automated grinding in 96-well plates increased throughput and minimized the time required for manual grinding. Thus, this method would be highly beneficial for MAS breeding programs that often screen thousands of plant samples each season [45]. A similar high throughput result was achieved by Anderson *et al.* (2018). The latter optimized the DNA extraction method of Whitlock *et al.* (2008), used a 96-well plate for extraction, and achieved a consistent yield across the plate with a low failure rate. Jain *et al.* (2013), used an optimized DArT DNA extraction protocol and were able to extract suitable quality DNA from honey that was amplifiable by PCR.

The success of the KASP genotyping experiment is dependent on the quality and quantity of genomic DNA. Usually, for maize, pure genomic DNA in the amount of 5–50 ng is required to generate clear and consistent allele calls using KASP assay [46]. Commonly used methods for DNA quality and quantity analysis include agarose-gel electrophoresis, fluorescence-, and Ultraviolet (UV) absorbance-based measurement [43]. Fluorescence-based measurement using DNA-binding dyes such as PicoGreen is fast, sensitive, and dsDNA-specific; however, it comes with the DNA-binding reagent's added cost. Agarose gel electrophoresis is laborious and carries the risk of exposure to hazardous chemicals like ethidium bromide. The UV absorbance measurement is the most common DNA quantitation method; based on the fact that DNA absorbs UV light at a specific wavelength, DNA concentration is calculated by measuring the absorbance at 260 nm and using the relationship A_{260} of 1.0 equals 50 μ g/ml pure dsDNA. DNA purity is estimated based on two UV absorbance ratios: $A_{260}/A_{280} \geq 1.7$ and $A_{230}/A_{260} \geq 1.5$ for pure DNA [47].

Using a microplate reader and a 384-well plate, the nucleic acid quantification method was optimized for high throughput. Here, we used the UV absorbance method for high throughput DNA quality and quantity analysis in the FLUOstar Omega microplate reader. The FLUOstar microplate reader uses ultrafast UV/Vis spectrometers for absorbance measurements, measuring 96 samples (96-well plate) to 384 samples (384-well plate) simultaneously within one second per well. It combines speed and the acquisition of complete absorbance spectra (220 to 1000 nm), making it ideal for nucleic acid quantification [48].

Although outsourcing KASP offers a lower cost per datapoint, this lower genotyping cost is usually driven by a high volume of samples, impracticable for most MAS projects genotyping smaller sample volumes with select markers. Our in-house genotyping system provides reduced cost (Table), mainly from logistics, and faster data turnaround time, which ultimately accelerates the genotyping workflow (Figure).

A few studies serve as the benchmark for QC analysis in maize using the KASP genotyping system. Semagn *et al.* (2012) suggested using a subset of 50 to 100 KASP markers for routine QC; Chen *et al.* (2016) used a smaller subset of markers (10 markers) to assess mislabeling of entries across a panel of CIMMYT Maize Lines (CMLs) achieving up to 99% detection probability. The latter also proposed using a rapid QC approach, with a smaller subset of markers, to ensure effective QC, lower genotyping costs, and shorten data turnaround time during seed production. Using a subset of markers, we were able to identify seed mix-up and labeling errors. For instance, the grouping of SAMMAZ27-4 with SAMMAZ15 suggests a possible mislabeling or mix-up of seeds during harvesting and storage. Also, the grouping of SAMMAZ39-1 with SAMMAZ16 indicates possible pollen contamination or seed mix-up during handling. Similar errors due to seed mix-up and contamination were reported in Semagn *et al.* (2012), where 50 KASP SNPs were used to determine genetic identity among 2 to 4 seed sources of the same inbred line. Ertiro *et al.* (2019) also reported a high discrepancy in genetic purity and identity by the origin of seed sources irrespective of the genotyping platform used. They concluded that using a small subset of pre-selected high-quality markers was sufficient for performing QC analysis using low marker density genotyping platforms like KASP. This study showed that the rapid QC method using 28 KASP SNPs efficiently distinguished the four maize varieties taken from five seed sources.

Hybrid verification is often performed during seed production or population breeding to confirm that a particular hybrid is derived from the intended parental lines (free from contamination by foreign pollens). Reducing the data turnaround time is essential to ensure that an accurate hybrid is selected to be carried forward in breeding programs or dissemination to farmers in seed production. A reduced turnaround time also saves the cost of inputs applied to undesired genotypes since they can be discarded as soon as they have been identified upon genotyping. Our expedited workflow was able to achieve this. The possibility of contamination by self-pollination or foreign pollen exists; as such, hybrid verification is necessary to enable a seed producer to check whether accurate crosses are made for the production of the hybrid; this increases the confidence of the end-users on the quality and integrity of seeds produced [49]. Our result showed that 10 KASP markers were sufficient in distinguishing between maize parental inbred lines and identified true hybrid lines and residual contaminations, and possible sampling errors. A small subset of KASP markers has also been used to verify hybrids in other plant species. Osei *et al.* (2020) used 38

KASP markers to screen tomato genotypes to identify true F₁ hybrids for the possible development of inbreds with long shelf life through MABC. Patterson *et al.* (2017) achieved a highly accurate picture of *Myriophyllum* species distribution dynamics in North American lakes by genotyping 39 individuals from both parental watermilfoil and their hybrids using a subset of 3 KASP markers.

Following our optimized workflow, we were able to identify high PVA maize lines harboring the favorable allele of the *crtRB1* gene, which could serve as donor lines for the maize PVA breeding program. The ongoing marker-based selection of maize lines introgressed with resistance to aflatoxin accumulation promises to provide elite tropical maize lines with resistance to aflatoxin resistance, which will contribute to genetic gain in maize production. Similar success was achieved by the Biotechnology Center of the University of California, Davis, USA, which used KASP SNPs associated with *Phytophthora capsici* resistance to identify and selectively breed pepper strains [39]. So far, we have generated over 1700 datapoints using our in-house genotyping workflow. Applying our optimized workflow to the QC and MAS experiments outlined above reduced the volume of reagents and consumables used, shortened the data turnaround, and ultimately accelerated the crop improvement process.

The optimized sample collection method helped reduce the lyophilization time, and the sample tracking system was useful for storing sample information. The modified DArT DNA extraction protocol provided a fast and cost-effective method for extracting a large amount of high-quality genomic DNA from maize leaf tissue. The FLUOstar Omega Microplate Reader provided a quick method for high throughput measurement of DNA absorbance. Performing the KASP genotyping in-house (our laboratory) has provided a means of genotyping a small volume of samples, which are not cost-effective for outsourcing to large-scale high throughput genotyping service providers. This in-house genotyping accelerated the data turnaround time and cost. The subset of KASP markers used here was sufficient to identify maize genotypes taken from different seed sources, distinguish the parental genotypes, and was useful in verifying true hybrids of F₁ progenies from two groups of bi-parental crosses.

Conclusion

Our work clearly illustrates an improvement to the conventional DNA-based genotyping workflow utilized by MIP. We acknowledge the initial capital investment to procure some of these instruments. However, it is not always necessary to equip each lab or breeding program. The use of shared facilities locally and regionally helps to overcome the high cost of essential instruments. The improved genotyping workflow promises to accelerate the marker-assisted selection process and push crop improvement activities to attain the yield potential in a shorter period. The result of this work can be readily adopted by national institutions, public and small plant breeding laboratories in developing countries to fast-track molecular marker-based genotyping for crop improvement activities, including QC and MAS. The result will also be useful to accelerate the QC activities of seed producers and facilitate cultivar identification and adoption-tracking studies.

Methodology

Plant Materials

The genetic identity test was performed using four well-adopted maize varieties taken from five seed sources, including the origin (Table 2). Maize seedlings were raised in pots for about two weeks, until they reached the 3-4-leaf stage, in a screen house at the Bioscience Center of the International Institute of Tropical Agriculture (IITA) Ibadan, Nigeria.

For the hybrid verification test, 60 maize F₁ progenies originating from 5 bi-parental crosses ((Set 1a: KS23-3 X IITATZI1653; Set 2a: KS23-5 X IITATZI1653; Set 3a: KS23-6 X IITATZI1653; Set 4a: KS23-3 X IITATZI1667; and Set 5: KS23-5 X IITATZI1667) were used. Lines KS23-3, KS23-5, and KS23-6, are Maize Lethal Necrosis (MLN)-resistant maize lines, while IITATZI1653 and IITATZI1667 are IITA-adopted elite maize lines with high PVA content. Another 328 F₁ plants originating from 4 sets of bi-parental crosses (Set 1b: TZEEI 29 x TZdEEI 99; Set 2b: TZdEEI 4 x TZEEI 79; Set 3b: TZEEI 79 x TZdEEI 13; and Set 4b: TZdEEI 102 x TZEEI 29) involving *Striga* susceptible (TZdEEI 102, TZdEEI 99, TZdEEI 4, and TZdEEI 13) and *Striga* resistant (TZEEI 29, and TZEEI 79) parents were also screened to identify true hybrids. Seedlings for the F₁ plants were raised in a maize field at IITA Ibadan, Nigeria.

A total of 70 PVA-QPM enriched, early maturing maize inbred lines were genotyped to select lines harboring the favorable allele for the *crtRB1* gene, which is associated with PVA content in maize. Seedlings of the inbred maize lines were raised in a maize field at IITA Ibadan, Nigeria. In the 4th breeding cycle of the maize enrichment project using MABC to introgress resistance to aflatoxin accumulation in elite tropical maize lines, we genotyped a total of 159 BC₁S₂ maize lines and applied a 15% selection intensity to identify lines harboring the favorable alleles of the 5 major QTLs (1.05, 3.03, 3.05, 5.03 and 5.07 originating from the donor maize line Mp715) associated with resistance to aflatoxin accumulation. Mp715 is a temperate inbred maize line with known resistance to aflatoxin accumulation. These plants were raised in a maize field at IITA Ibadan, Nigeria

Table 2
List of maize varieties used for genetic identity study.

Variety Name	Original Name	Outstanding characteristics/Potential Yield	Year of Release	Origin
SAMMAZ 15	IWDC2SynF2	Medium maturing, good seed quality, high yield potential, tolerance to <i>Striga hermonthica</i> . (6.9t/ha)	2008	IITA-Ibadan
SAMMAZ 16	TZLComp1SynW-1	Late maturing, good seed quality, high yield, resistance to <i>Striga hermonthica</i> . (6.4t/ha)	2008	IITA-Ibadan
SAMMAZ 27	EV99DT-W-STR	Drought tolerant and <i>Striga</i> resistant. (5.5t/ha)	2009	IITA-Ibadan
SAMMAZ 39	PVA SYN8	Intermediate level of pro-vitamin A content (6.4'g/g), high yield potential. (6.8t/ha)	2013	IITA-Ibadan

Sample Collection

Plants from which tissues are to be collected were tagged. Tags typically contained the line name, serial number, and pedigree. A total of 16 to 20 leaf-discs were collected from young leaves of each tagged plant, directly into Corning 96-well Polypropylene 1.2 ml cluster tubes with strip caps (Merck, Germany) using Haris Uni-core 4.0 mm puncher and cutting mat (Merck, Germany). Before sample collection, two 4.0 mm stainless steel grinding balls (SPEX SamplePrep) were placed in each tube. Plant tissues were preserved on ice from the field to laboratory where they were stored at -80°C freezer before lyophilizing. The leaf tissues were lyophilized for 48 hours using FreeZone Freeze Dryer (Labconco) following the manufacturer's manual.

DNA Extraction, and Quality and Quantity Analysis

Lyophilized leaf tissues were ground into powder by shaking at 1500 strokes per min for 1.5 min using an automated high-throughput tissue homogenizer, Geno/Grinder 2010 (SPEX SamplePrep). Genomic DNA was extracted from plant leaf tissues using the CTAB-based Diversity Array Technology (DArT) DNA extraction protocol [50] with some modification (Additional file 1). This protocol is similar to the modified CTAB protocol [51] but excludes mercaptoethanol. However, other reagents such as sarcosyl, sodium metabisulfite, and sorbitol were added. It also has shorter steps and provides suitable quality DNA for different molecular applications. Our protocol used freeze-dried leaf tissues instead of the fresh leaf tissues used in the DArT protocol. The DNA quality and quantity were confirmed by spectrophotometry using the FLUOstar Omega Microplate Reader (BMG LABTECH) following the manufacturer's manual.

KASP Genotyping and Data Analysis

The isolated genomic DNA was diluted to a working concentration of 50 ng/μl and used as template DNA for the KASP genotyping reaction. Twenty-eight KASP SNPs were used to determine the selected maize varieties' genetic identity, while 10 KASP SNPs were used to verify true hybrids among the F₁ maize lines. The SNPs (Additional file 2. Table S1) were taken from a maize QC SNP panel [20] recommended by CIMMYT [49, 8], and chosen for their high polymorphic information content (PIC) and uniform maize genome coverage. Trait-specific KASP markers (Additional file 2. Table S2) were used to screen for BC₁S₂ lines carrying the favorable allele for resistance to aflatoxin accumulation, and identify inbred lines with high PVA content. The KASP reaction was performed in a 384-well plate with a total reaction volume of 5 μl consisting 2.5 μl template DNA and 2.5 μl of the prepared genotyping mix (2 × KASP master mix and primer mix) according to the KASP manual [46]. The KASP assay kit was purchased from LGC Genomics (LGC Group) and the KASP amplification reaction was run in-house using the LightCycler 480 II real time-PCR System (Roche Life Sciences, Germany) and GeneAmp PCR System 9700 (Applied Biosystems, USA) at the Bioscience Centre of IITA Ibadan, Nigeria. The amplification condition included 1 cycle of KASP special Taq activation at 94 °C for 15 min, followed by 36 cycles of denaturation at 94 °C for 20 s, and annealing and elongation at 60 °C (dropping 0.6 °C per cycle) for 1 min. Endpoint detection of the fluorescence signal was acquired for 1 min at 30 °C when using the LightCycler 480 II real time-PCR System or read using the FLUOstar Omega Microplate reader (BMG Labtech, SA) when using the

GeneAmp PCR System 9700. The genotype calls were analyzed using the KlusterCaller analysis software (LGC Biosearch Technologies). Genotypes of the amplified products were displayed in Cluster plot and called as homozygous for allele 1 (allele reported by FAM, X-axis), homozygous for allele 2 (allele reported by HEX, Y-axis), heterozygous (alleles reported by FAM and HEX, between X- and Y-axis), or inconclusive. The cluster plot image was viewed using the SNPviewer software (LGC Biosearch Technologies). Data generated from KlusterCaller was used to calculate the genetic distance using PowerMaker 3.25 statistical software [52].

Abbreviations

BMS: Breeding Management System

CG: Consultative Group

CIMMYT: International Maize and Wheat Improvement Center

CTAB: Cetyl trimethyl ammonium bromate

DArTseq: Diversity Arrays Technology Sequencing

DNA: Deoxyribonucleic acid

EiB: Excellence in Breeding

GBS: Genotyping by sequencing

GCP: Generation Challenge Programme

GOBii: Genomic and Open-source Breeding Informatics Initiative

GS: Genomic Selection

GWAS: Genome-wide Association study

Hr: Hour

HTPG: High Through-Put Genotyping

IBP: Integrated Breeding Platform

InDels: Insertions/deletions

IITA: International Institute of Tropical Agriculture

KASP: Kompetitive Allele-Specific PCR

MABC: Marker-assisted backcrossing

MARS: Marker-assisted recurrent selection

MAS: Marker-assisted selection

μL: Microliter

Min: Minutes

MIP: Maize Improvement Programme

mL: Milliliter

MLN: Maize Lethal Necrosis

MSV: Maize Streak Virus

ng/μL: Nanogram per microliter

NGS: Next Generation Sequencing

PVA: Provitamin A

QC: Quality control

SNP: Single nucleotide polymorphism

SSA: Sub-Saharan Africa

SSR: Simple sequence repeats

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

All relevant data associated with this study are included in this published article and its supplementary information files. Any other detailed info can be made available by the corresponding author upon a reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the CGIAR Research Program on Maize (MAIZE) and by the Bill & Melinda Gates Foundation and USAID funded project titled “Stress Tolerant Maize for Africa” (grant number OPP1134248). The MAIZE program receives financial support from the governments of Australia, Belgium, Canada, China, France, India, Japan, Korea, Mexico, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the UK, and the United States, and from the World Bank.

Authors’ contributions

QO and MG conceived and designed the experiment; QO conducted the laboratory experiment analyzed the data, and drafted the manuscript; MG supervised the experiment, contributed to the write-up, and reviewed the final draft. AM supervised the germplasm development, and DB participated in the revision of the manuscript.

Acknowledgments

The authors are grateful to the support provided by colleagues in MIP and Bioscience Center, IITA. The authors thank Oluyinka Ilesanmi for designing and printing the barcode labels for tagging plant materials on the field, and Dr. Ryo Matsumoto for providing the images.

References

1. Ray DK, Mueller ND, West PC, and Foley JA. Yield trends are insufficient to double global crop production by 2050. *PLoS One*. 2013; 8(6): e66428. DOI: 10.1371/journal.pone.0066428.
2. Xu Y, Li P, Zou C, Lu Y, Xie C, Zhang X, Prasanna BM, Olsen MS. Enhancing genetic gain in the era of molecular breeding. *J. Exp. Bot.* 2017; 68: 2641–2666. DOI: 10.1093/jxb/erx135
3. Barabaschi D, Tondelli A, Desiderio F, Volante A, Vaccino P, Vale G, Cattivelli L. Next generation breeding. *Plant Science*. 2016; 242:3–13. DOI: 10.1016/j.plantsci.2015.07.010.
4. Varshney RK, Singh VK, Hickey JM, Xu X, Marshall DF, Wang J, Edwards D, Ribaut JM. Analytical and decision support tools for genomics-assisted breeding. *Trends in Plant Science*. 2016; 21: 354–363.

5. Daniel IO, Adetumbi JA, Oyelakin OO, Olakojo SA, Ajala MO, Onagbesan SO. Application of SSR markers for genetic purity analysis of parental inbred lines and some commercial hybrid maize (*Zea mays* L.). *American Journal of Experimental Agriculture*. 2012;2(4):597-606.
6. Hamblin MT, Warburton ML, Buckler ES. Empirical comparison of simple sequence repeats and single nucleotide polymorphisms in the assessment of maize diversity and relatedness. *PLoS One*. 2007;2: e1367.
7. Hao C, Wang Y, Chao S, Li T, Liu H, Wang L, Zhang X. The iSelect 9 K SNP analysis revealed polyploidization induced revolutionary changes and intense human selection, causing strong haplotype blocks in wheat. *Scientific Reports*. 2017; 7(1). DOI: 10.1038/srep41247.
8. Semagn K, Beyene Y, Makumbi D, Mugo S, Prasanna BM, Magorokosho C, Atlin G. Quality control genotyping for assessment of genetic identity and purity in diverse tropical maize inbred lines. *Theoretical and Applied Genetics*. 2012; 125:1487–1501.
9. Schlotterer C. The evolution of molecular marker - just a matter of fashion? *Nature Reviews Genetics*. 2004; 5(1): 63. DOI:10.1038/nrg1249.
10. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A Robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*. 2011; 6: e19379.
11. Sansaloni C, Petrolis C, Jaccoud D, Carling J, Detering F, Grattapaglia D, Kilian A. Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. *BMC Proceedings*. 2011;5 Suppl 7:54. DOI:10.1186/1753-6561-5-s7-p54.
12. Prasanna B, Pixley K, Warburton M, Xie CX. Molecular marker-assisted breeding options for maize improvement in Asia. *Mol Breed*. 2010; 26:339–356.
13. Miah G, Rafii M, Ismail M, Puteh A, Rahim H, Islam K, Latif MA. Review of Microsatellite Markers and Their Applications in Rice Breeding Programs to Improve Blast Disease Resistance. *International Journal of Molecular Sciences*. 2013; 14(11): 22499–22528. DOI:10.3390/ijms141122499.
14. Goron TL, Raizada MN. Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Front. Plant Sci*. 2015; 6:157. DOI:10.3389/fpls.2015.00157.
15. Pandey MK, Khan AW, Singh VK, Vishwakarma MK, Shasidhar Y, Kumar V, Garg V, Bhat RS, Chitikineni A, Janila P, Guo B, Varshney RK. QTL-seq approach identified genomic regions and diagnostic markers for rust and late leaf spot resistance in groundnut (*Arachis hypogaea* L.). *Plant Biotechnology Journal*. 2017;15(8): 927–941. DOI:10.1111/pbi.12686.
16. Liu LW, Wang Y, Gong YQ, Zhao TM, Liu G, Li XY, Yu FM. Assessment of genetic purity of tomato (*Lycopersicon esculentum* L.) hybrid using molecular markers. *Scientia Horticulturae*. 2007; 115:7-12.
17. Gedil M, and Menkir A. An Integrated Molecular and Conventional Breeding Scheme for Enhancing Genetic Gain in Maize in Africa. *Front. Plant Sci*. 2019; 10:1430. DOI: 10.3389/fpls.2019.01430.
18. Gedil M, Tripathi L, Ghislain M, Ferguson M, Ndjioudjop MN, Kumar L, Ribaut JM. Biotechnology Success Stories by the Consultative Group on International Agriculture Research (CGIAR) System. *Science Policy Reports*. 2014; 95–114. DOI: 10.1007/978-3-319-04001-1_6.

19. Ribaut JM, de Vicente MC, Delannay X. Molecular breeding in developing countries: challenges and perspectives. *Curr Opin Plant Biol.* 2010; 13:213–218.
20. Integrated Breeding Platform. www.integratedbreeding.net. Accessed on 30 June 2020.
21. Rasheed A, Hao Y, Xia X, Khan A, Xu Y, Varshney R K, He Z. Crop breeding chips and genotyping platforms: progress, challenges, and perspectives. *Molecular Plants.* 2017; 10(8): 1047-1064. DOI: 10.1016/j.molp.2017.06.008.
22. Semagn K, Olsen M, Babu R, Hearne S. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the Technology and its Application in Crop Improvement. *Mol Breed.* 2013; 33:1 – 14. DOI: 10.1007/s11032-013-9917-x.
23. Beissinger TM, Hirsch CN, Sekhon RS, Foerster JM, Johnson JM, Muttoni G, de Leon N. Marker Density and Read Depth for Genotyping Populations Using Genotyping-by-Sequencing. *Genetics.* 2013; 193(4):1073–1081. DOI: 10.1534/genetics.112.147710.
24. Giancola S, McKhann HI, Berard A, Camilleri C, Durand S, Libeau P, et al. Utilization of the three high-throughput SNP genotyping methods, the GOOD assay, Amplifluor and TaqMan, in diploid and polyploid plants. *Theor Appl Genet.* 2006; 112:1115–24.
25. Helguera M, Khan IA, Kolmer J, Lijavetzky D, Zhong-qi L, Dubcovsky J. PCR Assays for the Cluster of Rust Resistance Genes and Their Use to Develop Isogenic Hard Red Spring Wheat Lines. *Crop Science.* 2003; 43(5):1839. DOI:10.2135/cropsci2003.1839.
26. Stevanato P, Broccanello C, Biscarin F, DelCorvo M, Sablok G, Panella L, Stella A, Concheri G. High-throughput RAD-SNP genotyping for characterization of sugar beet genotypes. *Plant Mol Biol Rep.* 2014; 32:691 – 6.
27. Wu J, Liu S, Wang Q, et al. Rapid identification of an adult plant stripe rust resistance gene in hexaploid wheat by high-throughput SNP array genotyping of pooled extremes. *Theor Appl Genet* 2018; 131:43–58. DOI:10.1007/s00122-017-2984-3.
28. Myakishev M. V, Khripin Y, Hu S, and Hammer D.H. High-Throughput SNP Genotyping by Allele-Specific PCR with Universal Energy-Transfer-Labeled Primers. *Genome Research.* 2001; 11:163–169. DOI:10.1101/gr.157901.
29. Broccanello C, Chiodi C, Funk A, *et al.* Comparison of three PCR-based assays for SNP genotyping in plants. *Plant Methods.* 2018; 14:28. DOI:10.1186/s13007-018-0295-6.
30. Ayalew H, Tsang PW, Chu C, Wang J, Liu S, Chen C, et al. Comparison of TaqMan, KASP, and rhAmp SNP genotyping platforms in hexaploid wheat. *PLoS ONE.* 2019; 14(5): e0217222. DOI: 10.1371/journal.pone.0217222.
31. Ertiro BT, Ogugo V, Worku M, Das B, Olsen M, Labuschagne M, Semagn K. Comparison of Kompetitive Allele-Specific PCR (KASP) and genotyping by sequencing (GBS) for quality control analysis in maize. *BMC Genomics.* 2015; 16(1). DOI: 10.1186/s12864-015-2180-2.
32. Kaur B, Mavi GS, Gill MS, and Saini DK. Utilization of KASP technology for wheat improvement. *Cereal Research Communications.* 2020. DOI:10.1007/s42976-020-00057-6.

33. Steele KA, Quinton-Tulloch MJ, Amgai RB, Dhakal R, Khatiwada SP, Vyas D, Heine M, Witcombe JR. Accelerating public sector rice breeding with high-density KASP markers derived from whole genome sequencing of indica rice. *Mol Breed* 38:38.
34. Chen J, Zavala C, Ortega N, Petroli C, Franco J, Burgueno J, Hearne SJ. The Development of Quality Control Genotyping Approaches: A Case Study Using Elite Maize Lines. *PLOS ONE*. 2016;11(6): e0157236. DOI: 10.1371/journal.pone.0157236.
35. Neelam K, Brown-Guedira G, and Huang L. Development and validation of a breeder friendly KASPar marker for wheat leaf rust resistance locus Lr21. *Molecular Breeding*. 2013; 31(1):233-237. DOI:10.1007/s11032-012-9773-0.
36. Khera P, Upadhyaya HD, Pandey MK, Roorkiwal M, Sriswathi M, Janila P, Varshney RK. Single Nucleotide Polymorphism–based Genetic Diversity in the Reference Set of Peanut (spp.) by Developing and Applying Cost-Effective Kompetitive Allele Specific Polymerase Chain Reaction Genotyping Assays. *The Plant Genome*. 2013; 6(3), 0. DOI:10.3835/plantgenome2013.06.0019.
37. He C, Holme J, Anthony J. SNP Genotyping: The KASP Assay. *Crop Breeding*. 2014;75–86. DOI: 10.1007/978-1-4939-0446-4_7.
38. Pariasca-Tanaka J, Lorieux M, He C, McCouch S, Thomson MJ, Wissuwa M. Development of an SNP genotyping panel for detecting polymorphisms in *Oryza glaberrima*/O. *sativa* interspecific crosses. *Euphytica*. 2014; 201(1): 67–78. DOI:10.1007/s10681-014-1183-4.
39. Majeed U, Darwish E, Rehman SU, Zhang X. Kompetitive Allele-Specific PCR (KASP): A Singleplex Genotyping Platform and Its Application. *Journal of Agricultural Science* 11. 2018;11.
40. Chase MW, Hills HH. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon*. 1991; 40:215–220.
41. Storchova H, Hrdlickova R, Chrték J Jr, Tetera M, Fitze D, Fehrer J. An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. *Taxon* 2000; 49:79–84.
42. Fournoy LE, Adams RP, Pandey RN. Interim and archival preservation of plant specimens in alcohols for DNA studies. *Biotechniques*. 1996; 20:657–660.
43. Semagn K. Leaf Tissue Sampling and DNA Extraction Protocols. In: Besse P, Editor. *Molecular Plant Taxonomy: Methods and Protocols*. vol. 1115. New York: Human Press; 2014. p. 53 – 67.
44. Till BJ, Jankowicz-Cieslak J, Huynh OA, Beshir MM, Laport RG, Hofinger BJ. Sample Collection and Storage. *Low-Cost Methods for Molecular Characterization of Mutant Plants*. 2015; 9–11. DOI:10.1007/978-3-319-16259-1_3.
45. Gu W, Weeden N, Yu J, Wallace D. Large-scale, cost-effective screening of PCR products in marker-assisted selection applications. *Theoretical and Applied Genetics*. 1995; 91:465–470.
46. LGC Biosearch Technologies. <https://biosearch-cdn.azureedge.net/assetsv6/KASP-genotyping-chemistry-User-guide.pdf>. Accessed on 24 June, 2020.
47. BMG LABTECH Floustar Omega Application Notes. https://www.bmg-labtech.com/fileadmin/06_Support/Download_Documents/App_Note/AN263.pdf.

Accessed on 2 June 2020.

48. BMG LABTECH Floustar Omega Operating Manual. <https://med-fom-avgaylab.sites.olt.ubc.ca/files/2018/04/0415B0001K-Operating-Manual-Omega.pdf>. Accessed on 2 June 2020.
49. Gowda M, Worku M, Nair SK, Palacios-Rojas N, Huestis G, Prasanna BM. Quality Assurance/ Quality Control (QA/QC) in Maize Breeding and Seed Production: Theory and Practice. CIMMYT: Nairobi. 2017.
50. Diversity Array Technology Pty Ltd (DARt). Plant DNA extraction protocol for DARt. <http://www.diversityarrays.com>. Accessed on 24 June 2020.
51. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. Focus. 1990; 12: 13–15.
52. Liu K, Muse SV. Power marker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005;21: 2128–2129.

Figures

Figure 1

SNPviewer image showing clustering of IITA's Sammaz15 and Sammaz39 maize lines with same lines from three out of four seed sources. (a) Sammaz15-2, -3, and -4 grouped with IITA's Sammaz15 (blue dots) using SNP PZA02746_2; (b) Sammaz39-2, -3, and -4 grouped with IITA's Sammaz39 (blue dots)

using SNP PZB01658_1. For each SNP marker, blue dots represent homozygous genotypes, green dots represent heterozygote genotypes, and the black dots represent no-template controls (NTC) as indicated on the left side of each image.

Figure 1

SNPviewer image showing clustering of IITA's Sammaz15 and Sammaz39 maize lines with same lines from three out of four seed sources. (a) Sammaz15-2, -3, and -4 grouped with IITA's Sammaz15 (blue dots) using SNP PZA02746_2; (b) Sammaz39-2, -3, and -4 grouped with IITA's Sammaz39 (blue dots) using SNP PZB01658_1. For each SNP marker, blue dots represent homozygous genotypes, green dots represent heterozygote genotypes, and the black dots represent no-template controls (NTC) as indicated on the left side of each image.

Figure 2

SNPviewer image showing result of hybrids verification in two sets of F1 Plants. (a) Genotyping of 12 F1 lines produced from a cross between KS23-5 and IITATZI1653, using SNP PZA03135_4. (b) Genotyping of 12 F1 lines produced from a cross between KS23-5 and IITATZI1667, using SNP PZA02779_1. For each SNP marker, blue dots represent homozygous parental genotype reported by FAM, red dots represent homozygous parental genotype reported by HEX, green dots represent heterozygous hybrid genotypes, and the black dots represent no-template controls (NTC).

Figure 2

SNPviewer image showing result of hybrids verification in two sets of F1 Plants. (a) Genotyping of 12 F1 lines produced from a cross between KS23-5 and IITATZI1653, using SNP PZA03135_4. (b) Genotyping of 12 F1 lines produced from a cross between KS23-5 and IITATZI1667, using SNP PZA02779_1. For each SNP marker, blue dots represent homozygous parental genotype reported by FAM, red dots represent homozygous parental genotype reported by HEX, green dots represent heterozygous hybrid genotypes, and the black dots represent no-template controls (NTC).

Figure 3

Neighbor-joining tree for four Maize varieties taken from five seed sources based on genetic distance, performed with 1000 bootstrap. Bootstrap values are indicated on the tree branches. The suffix "-1", "-2", "-3", "-4", and "-IITA", after line name indicate seed source 1, 2, 3, 4, and IITA.

Figure 3

Neighbor-joining tree for four Maize varieties taken from five seed sources based on genetic distance, performed with 1000 bootstrap. Bootstrap values are indicated on the tree branches. The suffix "-1", "-2", "-3", "-4", and "-IITA", after line name indicate seed source 1, 2, 3, 4, and IITA.

Figure 4

Diagram showing improvement to minimize bottlenecks in the genotyping workflow.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1QueenN.Obi.docx](#)
- [Additionalfile1QueenN.Obi.docx](#)
- [Additionalfile2QueenN.Obi.docx](#)
- [Additionalfile2QueenN.Obi.docx](#)