

Development and application of KASP assays to differentiate between *Sorghum bicolor*, *halepense*, and their hybrids

Connor Purvis, Eric L. Patterson  and Erin E. Burns ^{*} 



Abstract

BACKGROUND: *Sorghum bicolor* and *Sorghum halepense* can readily hybridize, creating difficulty in identification. No genetic tools exist to accurately distinguish *S. bicolor*, *S. halepense*, and their hybrids. Detecting hybridization is essential to monitor crop-to-weed introgression. This study utilizes a single nucleotide polymorphism (SNP) in the internal transcribed spacer region between *S. bicolor* and *S. halepense*. This SNP was utilized in Kompetitive allele-specific PCR (KASP) assay to identify *S. bicolor*, *S. halepense*, and their hybrids.

RESULTS: KASP assays were successful in accurately differentiating between *S. bicolor*, *S. halepense*, and their hybrids. The KASP assay performed as well as Oxford Nanopore sequencing for measuring SNP frequency and thus is a perfect proxy for genotyping. Greenhouse crosses confirmed crop-to-weed introgression, with *S. halepense* being more receptive to interspecific pollen. Known and unknown samples assayed displayed misidentification in germplasm lines and significant hybrid frequency in naturally occurring biotypes. Synteny analyses revealed duplications of the ITS region in *S. halepense*.

CONCLUSION: We developed a novel KASP assay targeting a conserved SNP that accurately distinguishes between *S. bicolor*, *S. halepense*, and their hybrids. This assay was validated through Oxford Nanopore sequencing, greenhouse crosses, and diverse germplasm and natural collections.

© 2026 The Author(s). *Pest Management Science* published by John Wiley & Sons Ltd on behalf of Society of Chemical Industry.

Supporting information may be found in the online version of this article.

Keywords: crop-to-weed introgression; hybrid detection; KASP assay; polyploidy; *Sorghum halepense*

1 INTRODUCTION

Species hybridization has played a fundamental role in agronomic crop domestication by accelerating genetic improvements resulting in enhanced quality and higher yields.^{1,2} While controlled hybrids have been essential for crop improvement, hybridization between crops and weeds is recognized as a driver of rapid evolution with potential negative consequences for control of formed hybrids.³ For instance, traits important for rice (*Oryza sativa* L. *ssp.*) and corn (*Zea mays* L. *ssp. mexicana*, *Zea luxurians*) cultivation have been shown to be transferred to weedy varieties and through introgression in several instances.^{4,5} Further, there is evidence that these weed-crop hybrids have increased fitness due to cultivated trait introgression.⁶ Therefore, the flow of agronomic traits into weeds has potential to increase weed resilience to stress.

The *Sorghum* genus is an important model for studying hybridization between crops and weeds. It consists of the domesticated crop grain sorghum (*Sorghum bicolor ssp. bicolor* (L.) Moench) and several destructive weeds including johnsongrass (*Sorghum halepense* (L.) Pers.) and shattercane (*Sorghum bicolor* (L.) Moench *ssp. verticilliflorum* (Steud.) de Wet ex Wiersma & J. Dahlb). *Sorghum bicolor* was domesticated approximately 6000 years ago^{7,8} and

is one of the most important cereal crops globally, with over 60 million metric tons produced annually.⁹ Species within this genus are reproductively compatible due to recent shared ancestry, overlapping phenology, and in some cases, compatible ploidy levels. Polyploidy plays a critical role in reproductive compatibility within *Sorghum*; *S. bicolor* is diploid ($2n = 2x = 20$), while *S. halepense* is a tetraploid ($2n = 4x = 40$). Despite these differences, repeated hybridization events have been shown between these species resulting in triploids, tetraploids, pentaploids, and hexaploids.¹⁰ Long-term gene flow between the crop (*S. bicolor*) and its weedy relatives (*S. bicolor ssp. verticilliflorum* and *S. halepense*) has caused ambiguity in plant identification. Morphologically, hybrids are difficult to distinguish in the field due to taking on intermediate phenotypes between the parents.¹¹ Further complicating the *Sorghum* genus is Columbus grass

^{*} Correspondence to: EE Burns, Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA. E-mail: burnser5@msu.edu

Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA

(*Sorghum alnum* Parodi), a cultivated forage perennial hybrid, which can be mistaken for *S. halepense*.¹²

Climate change will exacerbate the current challenges arising from these hybridization events. Rising temperatures and potentially drier conditions will expand grain sorghum's range to more northern regions and increase its viability under water restrictions. It is well-adjusted to climatic variability and low-input environments, which will be useful as the climate continues to change.¹³ The range expansion for grain sorghum also includes expansion of its weedy relatives to these new ranges, which are also well adapted to climatic variability. Species distribution modeling indicates there will be a northward expansion of *S. halepense*, increasing opportunities for gene flow between the crop and weed.¹⁴ As ranges simultaneously expand, crop-to-weed introgression may increase in these closely related species.

As hybridization is known to happen between these species and their ranges will continue to overlap, a concerning aspect is the transfer of advantageous crop traits (increased growth rate, seed set, stress resistance, etc.) to weedy relatives. Additionally, the transfer of weedy traits such as herbicide resistance or abiotic stress tolerance to crops could make management difficult.¹⁵ Both *S. bicolor* ssp. *verticilliflorum* and *S. halepense* have known widespread cases of herbicide resistance to different sites of action.^{16,17}

Additionally, climate change has driven *S. bicolor* breeding to develop more drought-tolerant lines which increases its potential to handle other abiotic stressors.^{18,19} Pollen-mediated gene flow has been shown at up to 50 m between *S. bicolor* and *S. halepense* biotypes, these traits associated with increased crop performance are at risk to increase weedy competitiveness.²⁰

To effectively monitor gene flow dynamics and to better manage risks associated with crop-weed hybridization, we need reliable tools to distinguish species and detect hybrids. There are currently no genetic markers that have been reported in the literature that can identify hybrids between *S. bicolor* and *S. halepense*. Kompetitive allele-specific PCR (KASP) is a PCR-based genotyping technique that uses single nucleotide polymorphisms (SNPs) for a wide range of applications. It has been used in marker-assisted breeding in crops like rice (*Oryza sativa* L.) and wheat (*Triticum aestivum* L.) to identify valuable alleles^{21,22} and has also been applied to identify interspecific hybrids in plants such as vanilla (*Vanilla spp.*) and watermilfoil (*Myriophyllum spp.*).^{23,24} The KASP method is more cost-effective than other genotyping methods. For example, Oxford Nanopore sequencing can cost approximately \$15 per sample, whereas KASP assays can be run for one cent per sample using standard real-time qPCR thermocyclers at scale.²⁵

Therefore, the goal of this study was to develop a KASP-based genetic assay capable of reliably distinguishing *S. bicolor*, *S. halepense*, and their hybrids. To achieve this goal, we had three objectives: (i) validate our marker and assay with known control samples from pure *S. bicolor* and *S. halepense* individuals and hybrids, (ii) obtain crosses between *S. bicolor*, *S. halepense*, and hybrid in the greenhouse to analyze F₁ populations for hybrid formation, and (iii) assign genotypes to known germplasm biotypes and unknown naturally collected biotypes based on our KASP assay.

2 MATERIALS AND METHODS

2.1 Plant material for KASP

The Michigan State weed seed collection (Supporting Information, Table S1), hereby the MSU collection, was utilized to facilitate

these experiments along with germplasm from the National Plant Germplasm System (NPGS) *Sorghum* collection (Supporting Information, Table S2). To investigate objective one, known *S. bicolor* and *S. halepense* biotypes along with a subset from the MSU collection that spanned a wide range of hybridization phenotypes were grown under greenhouse conditions in the spring of 2023 at MSU. Plants were grown under a 16-h light period at 26 °C and an 8-h dark period at 18 °C. At the three to four leaf stage, the newest leaf was collected, flash frozen in liquid nitrogen and stored at −80 °C until use. Seeds were only used for extraction if a biotype had poor germination. All tissue was collected in the same way for all experiments unless stated otherwise.

2.2 DNA extraction

DNA was extracted from *Sorghum* leaf or seed tissue (80 mg) using a modified CTAB protocol previously described.²⁶ Tissue was first ground in liquid nitrogen with a mortar and pestle before being placed in a 2 mL centrifuge tube. In brief, 1% hot CTAB buffer (1 mL; 65 °C) with 2% β-mercaptoethanol was added to the tube and incubated for an hour at 65 °C. After incubation, samples were centrifuged for 2 min with the supernatant removed to a new tube and RNase (1 µL at 10 mg mL^{−1}) added and incubated for 30 min at 37 °C. All centrifugation steps were performed at 15 000 rpm unless otherwise noted. Chloroform: isoamyl alcohol (600 µL; 24:1) was added and centrifuged for 3 min with the top layer removed to a new tube and chloroform: isoamyl alcohol (400 µL) added and centrifuged for three additional min. The top layer was removed (500 µL) to a new 2 mL centrifuge tube with 3 M sodium acetate (50 µL) and 100% EtOH (1 mL) added. Samples were precipitated by centrifugation for 1.5 min. A sequential pellet cleaning was performed with 70% EtOH (400 µL) for 3 min at room temperature. Excess ethanol was then removed with a Kim-wipe, and the pellet was resuspended in nuclease-free water (50 µL). DNA concentration and quality were quantified using a NanoDrop One spectrophotometer (Thermo Scientific, 5225 Verona Rd., Madison, WI 53711).

2.3 Primer design for KASP

Thirteen sequences from the internal transcribed spacer (ITS) region, including seven from *S. bicolor* and six from *S. halepense*, were obtained from NCBI and aligned using MEGA 11²⁷ to examine SNPs between the two species. A single SNP at position 591 (ITS-SNP591) consistently differentiated the two species with *S. bicolor* accessions having a Thymine and *S. halepense* having a Cytosine at this position (Fig. 1). This SNP was utilized for designing allele-specific KASP primers. Our KASP primer set included two reverse primers: the reverse primer for *S. bicolor* was appended to the 5' end with the complementary sequence of the FAM fluorophore quencher, while the reverse primer for *S. halepense* was appended to the 5' end with the complementary sequence of the HEX fluorophore quencher. A universal forward primer was used in complement with these two primers (Table 1).

2.4 KASP assay

A primer master mix, including the forward and both reverse primers, for ITS-SNP591 was made according to the master mix manufacturer's recommendations (The Universal PACE Genotyping Master Mix 3CR Bioscience, Harlow, England, UK). This PCR master mix is compatible with most allele-specific assays including both PACE and KASP. All primers were suspended in nuclease-free water and a primer master mix was assembled with

Species	590					*	595				
<i>Sorghum bicolor</i> DQ1904201.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> MW911661.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> DQ190421.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> MW911655.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> GQ856358.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> XR002448966	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum bicolor</i> ON201455.1	G	C	C	G	G	T	G	C	A	T	C
<i>Sorghum halepense</i> AY665363.1	G	C	C	G	G	C	G	C	A	T	C
<i>Sorghum halepense</i> DQ005082.1	G	C	C	G	G	C	G	C	A	T	C
<i>Sorghum halepense</i> DQ118116.2	G	C	C	G	G	C	G	C	A	T	C
<i>Sorghum halepense</i> DQ186658.1	G	C	C	G	G	C	G	C	A	T	C
<i>Sorghum halepense</i> DQ186659.1	G	C	C	G	G	C	G	C	A	T	C
<i>Sorghum halepense</i> KF163843.1	G	C	C	G	G	C	G	C	A	T	C

Figure 1. Distinguishing SNP at position 591 in the ITS region differentiating *Sorghum bicolor* (T) and *Sorghum halepense* (C). Asterisk and solid black square notes allele differences at position.

Table 1. Primers for single nucleotide polymorphism in the internal transcribed spacer (ITS) region for KASP PCR assays and PCR amplicon sequencing

Primer name	Primer sequence (5'-3')	Predicted melting temperature (°C)
ITS-SNP591 (A/G)		
Universal FP	CGATTCGTGTCGGGCACAGC	67.4
<i>S. bicolor</i> RP	GGTCCTTAGGGCCGATGCA	65.9
<i>S. halepense</i> RP	GGTCCTTAGGGCCGATGCG	66.5
5' FAM TAG (<i>S. bicolor</i>)	GAAGGTGACCAAGTTCATGCT	
5' HEX TAG (<i>S. halepense</i>)	GAAGGTGCGAGTCAACGGATT	
ITS Primers		
<i>S. halepense</i> FP	GATTCTCAAGCTGGGCTGGTC	64.7
<i>S. halepense</i> RP	AATGGTCCGGTGAAGTGTTTCG	64.6

each of the species-specific primers (18 µL each), the universal forward primer (45 µL), and nuclease-free water (69 µL).

Reactions were assembled in 96-well plates with the primer plus master mix combination (0.14 µL + 5 µL) and genomic DNA (4.86 µL at 10 ng µL⁻¹). Reactions were performed in Bio-Rad CFX96 Real-Time System (Bio-Rad Laboratories, Inc., Hercules, CA, USA). The protocol was performed as follows: activation at 94 °C for 15 min, followed by 10 touchdown cycles at 94 °C for 20 s, 64–56 °C for 1 min 10 s (dropping 0.8 °C per cycle), followed by 37 cycles at 94 °C for 20 s, 63 °C for 1 min 10 s, 37 °C for 1 min. An additional two cycles for insufficient clustering were performed at 94 °C for 20 s and 57 °C for 1 min when necessary. Fluorescence data was chosen for genotyping based on the cycle with the greatest difference between control populations without

background amplification, this was determined to be cycles 26–28 of the amplification phase.

Positive control biotypes included a cultivated grain sorghum line (Rtx430) and a natural collection of *S. halepense*. Both biotypes consistently only amplified the FAM (*S. bicolor*) or HEX (*S. halepense*) fluorophore in the KASP assay, with negligible background fluorescence of the reciprocal fluorophore. KASP assays when typically employed produce clusters of samples corresponding to homozygotes of each allele and heterozygotes²⁴; however, *Sorghum* spp. hybrids are not necessarily diploid, and the ITS region is multicopy in both species which leads to a spectrum of intermediate amplification that depends on the relative contribution of each parental phenotype to the hybrids genome. To account for this, 'pseudohybrids' controls were included which were generated to simulate hybrids between *S. bicolor* and *S. halepense* by mixing DNA from each control sample in varying ratios, ranging from 1:1 to 1:12, with increasing concentrations for both species. Fluorescence data from pseudohybrids were used to set the boundaries for what is called a true hybrid versus purely homozygous individuals.

2.5 Nanopore sequencing for validation

The KASP assay is designed to amplify a single SNP in a sequence; however, it is not normally used to distinguish a range of intermediate genotypes. To validate this assay's ability to correctly assign partial contributions of the two genotypes, we performed amplicon sequencing of the ITS with Nanopore sequencing. To design these primers, a 536 bp ITS sequence (GenBank: DQ005082.1) was analyzed for sequence homology against the *S. halepense* genome (Project ID# 1335880) using BLASTN (version 2.14.1). Coordinates corresponding to the top BLAST hit were extracted using Samtools (version 1.19.1²⁸), and 200 bp of flanking sequence on either side were included that resulted in a 936 bp fragment. This sequence was used to design PCR primers to amplify the 936 bp of the ITS region in the same biotypes that were run on the KASP assay above (Table 2).

The control biotypes, along with a subset from the MSU collection, were amplified using reactions (20 µL) consisting of 2X GoTaq G2 Green Master Mix (16 µL; Promega Corporation,

Table 2. Comparison of KASP and Oxford Nanopore outputs for validation of KASP assay

Biotype*	KASP Output [†]		Nanopore Output [‡]		KASP θ	Nanopore θ	Genotype [§]
	FAM	HEX	T	C			
Arizona-B	3.80	91.79	7	1599	0.0489	0.0044	<i>S. halepense</i>
Michigan-D	4.39	85.05	5	1856	0.0351	0.0027	<i>S. halepense</i>
Ohio-F	18.26	87.11	224	1956	0.2124	0.1140	Hybrid
Texas-F	20.43	60.79	536	2251	0.3213	0.2338	Hybrid
Kansas-H	28.62	55.60	813	1620	0.4816	0.4651	Hybrid
Texas-A	35.34	39.69	1123	2054	0.7525	0.5003	Hybrid
Michigan-H	45.76	39.52	503	437	0.8671	0.8555	Hybrid
Michigan-A	62.60	28.47	1434	881	1.1295	1.0199	Hybrid
Maryland-D	64.27	33.79	1726	1270	1.1010	0.9364	Hybrid
South Carolina-C	82.73	13.85	1166	479	1.4108	1.1810	Hybrid
Rtx430	88.43	6.73	2507	39	1.4876	1.5552	<i>S. bicolor</i>
Pennsylvania-B	96.14	6.17	1981	42	1.5117	1.5496	<i>S. bicolor</i>

*Subset from the MSU collection (Table S1).

[†] Normalized fluorescence quantification from KASP assay.

[‡] Raw nucleotide reads at SNP 591 in ITS region.

[§] Genotype assigned by KASP assay.

Madison, WI, USA), each forward and reverse ITS primer (1 μ L at 10 μ M), and template DNA (2 μ L at 20 ng μ L⁻¹). Cycling consisted of an initial denaturation step of 95 °C for 5 min, followed by 36 cycles of 95 °C for 30 s, 64 °C for 30 s, and 72 °C for 1.5 min, with a final extension of 72 °C for 5 min in a Bio-Rad T100 thermal cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Amplicons were cleaned with Thermo Scientific GeneJET PCR purification kit (Thermo Fisher Scientific, Waltham, MA, USA). Diluted products (≥ 20 ng μ L⁻¹) were sent for Nanopore sequencing (Oxford Nanopore Technologies, Oxford, England, UK). Raw base counts from Nanopore amplicon sequencing at the distinguishing SNP were evaluated and correlated with KASP fluorescence data.

2.6 Validation through greenhouse crosses and known/unknown biotypes

S. bicolor, *S. halepense*, and hybrid individuals were crossed under greenhouse conditions in the spring of 2024 at MSU to generate known hybrids. The known *S. bicolor* (Rtx430) and naturally collected *S. halepense* biotypes described above were used as parents, as well as a naturally collected hybrid. 'Parent set one' consisted of *S. halepense* (male) $\times$ *S. bicolor* (female), resulting seeds were collected from *S. bicolor*. 'Parent set two' consisted of *S. bicolor* (male) $\times$ *S. halepense* (female), resulting seeds were collected from *S. halepense*. 'Parent set three' consisted of *S. bicolor* (male) $\times$ hybrid (female), resulting seeds were collected from the hybrid. Seeds of each parent were sown in 25.5 $\times$ 25.5 cm flats, upon emergence of true leaves, seedlings were transplanted to 10.5 $\times$ 10.5 cm pots. Once seedlings reached adequate size, they were transplanted into 6.5 L pots for crossing. Upon seedhead formation and onset of pollen release, the respective biotypes were tied and seedheads bagged together. Excess individuals of each biotype were grown to ensure plants utilized in crosses were at the same developmental stage. Bagged seed heads were shaken daily to facilitate pollen flow between individuals. Seeds were harvested at maturity and grown for tissue collection and DNA extraction, following the methods outlined above. Due to the number of individuals collected from *S. bicolor* (female) parent in parent set one, DNA extractions were performed in pools of 20 individuals. All

individuals from parent set two and three had DNA extracted from individuals.

Finally, we performed our assay on 62 *S. bicolor*, four *S. alnum* biotypes [obtained from the NPGS *Sorghum* collection (Supporting Information, Table S2)], and 78 naturally collected *S. halepense* biotypes from the MSU collection (Supporting Information, Table S1). All biotypes were grown under greenhouse conditions and followed tissue collection and DNA extraction outlined above. Each individual was subjected to the KASP assay to assign their genotype.

2.7 Data analysis

A synteny analysis was performed using a 536 bp fragment of the ITS region for homology analysis using BLASTN (version 2.14.1) against the *S. bicolor* genome (GenBank: GCA_000003195.3). To determine the genomic location of the ITS sequences in *S. halepense*, sequences from the initial BLAST query against the *S. bicolor* genome, were used as BLAST queries against the *S. halepense* genome. Synteny plot was visualized using Rideo-gram in R (version 4.3.3; Fig. 2).

Florescence data from the KASP assay was normalized by transforming the FAM and HEX florescence for each data point as a percentage of the maximum florescence for each fluorophore within a plate.²⁹ Cutoffs for genotyping calls were made by calculating the geometric mean between the coordinate of outermost pseudohybrid controls used and either the control *S. bicolor* or *S. halepense*, then drawing a line from that point to the origin (0,0; Fig. 3). Additionally, a quarter circle at 30% was defined as the zone of no amplification.²⁴ Sample genotypes were assigned based on the region of the plot where their fluorescence coordinates occurred. Pearson's correlation was conducted in R (version 4.3.2) on arctan transformed KASP (FAM/HEX) and Oxford Nanopore outputs (C/T) to cross-validate KASP results.

3 RESULTS AND DISCUSSION

We have developed a simple, genetic assay that can accurately distinguish between *S. bicolor*, *halepense*, and their hybrids. Our KASP assay utilizes the SNP591 found in the ITS region that

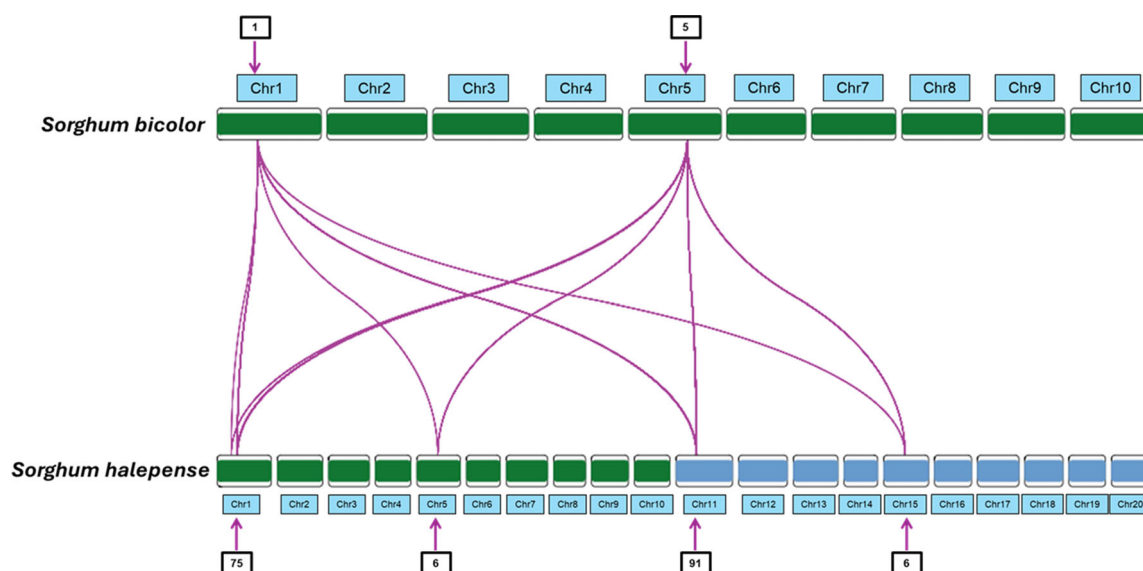


Figure 2. Ideogram of *Sorghum bicolor* and *S. halepense* alignment of the ITS region. Purple links represent shared synteny in the ITS region across chromosomes. Numbers in boxes above and below the ideogram indicate the number of ITS copies at each locus. Green highlighted chromosomes in *S. halepense* represent genome A, derived from *S. bicolor*.

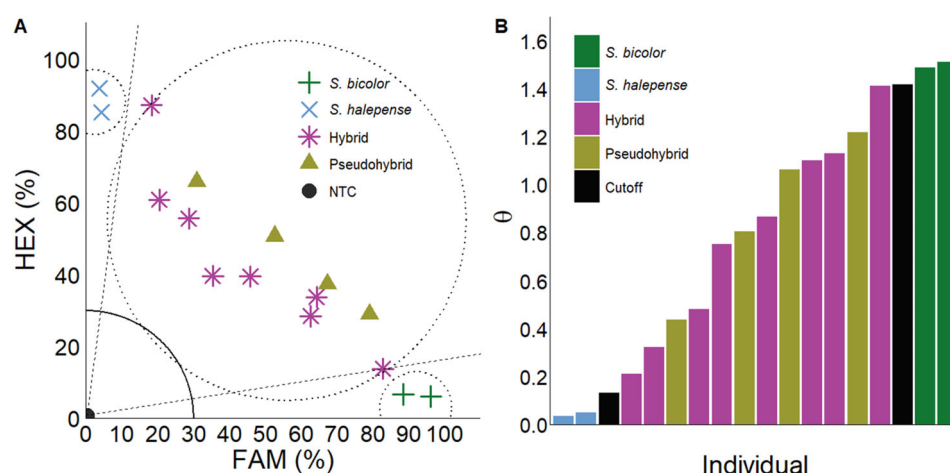


Figure 3. Reference KASP assay (A) of the wide range of genotypes between *S. bicolor* (+), *S. halepense* (x), hybrid (*), pseudohybrid (Δ), and no-template (NTC●) controls. Dashed lines represent cutoffs for making genotyping assignments. Dashed circles enclose groupings based on genotype calls. The solid quarter circle line is the cutoff for no amplification. The arctan transformation (B) from the KASP assay plotted by individual assignment. Black bars in the arctan transformation note the cutoffs from the KASP assay.

distinguished between these species and was consistent over spatial and genetic regions. The KASP assay has been validated through comparative sequencing methods, F_1 individuals from greenhouse crosses ($n = 1389$), known ($n = 66$), and unknown biotypes ($n = 78$). Together, these approaches confirmed the effectiveness of our assay and evaluated it across diverse genetic and biological backgrounds.

3.1 Synteny analysis and KASP assay interpretation

S. bicolor contains ITS clusters with one copy on chromosome one and five copies on chromosome five. Whereas the tetraploid *S. halepense* contains 75 copies on chromosome one and six copies on chromosome five of subgenome A as well as 91 ITS copies on chromosome 11 and six copies on chromosome 15 of subgenome B. Synteny analysis revealed that the ITS in *S. bicolor* and

S. halepense were homologous (Fig. 2). Duplication of ITS on chromosome 11 and 15 is expected for a tetraploid, *S. halepense*, formed from the original diploid *S. bicolor* × diploid *S. propinquum* hybridization event.³⁰ The differences in duplication across chromosomes was surprising as chromosomes one and 11 possess over 93% of the copies compared to chromosomes five and 15. These observations aligned with the variable allele SNP frequencies we noted in our KASP assay (Fig. 3).

Additionally, our KASP primers' binding positions were analyzed in *S. bicolor* and *S. halepense* across all copies in both species. Our *S. halepense* primer in the KASP assay bound within each subgenome in *S. halepense*, is overwhelmingly bound within subgenome B on chromosome 11 with only one copy not binding and three copies on chromosome 15 binding with the *S. halepense* primer. Within subgenome A only one copy on

chromosome one binds and three copies on chromosome five bind. Our *S. bicolor* primer in the KASP assay binds to one copy on chromosome one in *S. bicolor*.

Polyploid species, in this case hybrids between *S. bicolor* and *S. halepense* can display a range of intermediate amplifications of each fluorophore in a KASP assay due to there not being equal contributions of genetic information (i.e., target sequence), which produces a wide range of fluorescence coordinates within KASP assay plots (Fig. 3(A)). We artificially created pseudohybrids that mimic this wide range. Points that fall outside the bounds of the y- or x-axis dashed lines were assigned as *S. halepense* or *S. bicolor*, respectively. Points that fall within the region bounded by the two dashed lines were assigned as a hybrid. The solid quarter circle line represents the zone of no amplification and is considered an unreliable region for genotyping; however, none of our samples ever fell in this region. To simplify KASP scatterplots in easy to interpret single values, arctan transformation graphs (Fig. 3(B)) have the individual sample along the x-axis with the corresponding arctan value derived from the FAM/HEX ratio in the KASP assay. This transformation allows KASP results to be represented as single continuous variables for comparison.²⁹ All KASP assays and arctan transformation graphs will be presented and interpreted using this approach.

3.2 Validation of the KASP assay through nanopore

Our data showed that the KASP assay performed as well as Nanopore sequencing for measuring SNP frequency and thus is a perfect proxy for genotyping *S. bicolor*, *S. halepense*, and their hybrids ($R^2 = 0.969$, P -value < 0.001 , Fig. 4). A similar finding was reported in sunflower (*Helianthus annuus* L.) when comparing KASP marker results and amplicon targeted sequencing,³¹ but did not show the specificity we established in our assay. Our study instead utilized Nanopore sequencing in order to quantify base calls at the distinguishing SNP site for comparison, which is indistinguishable from the fluorescence data from our KASP assay. Additionally, that study³¹ was working within a diploid, crop species with a high quality reference genome. The significance in our correlation in this assay is novel in a polyploid, weedy species.

3.3 Validation through F₁ crosses

Greenhouse crosses revealed striking differences in hybrid formation, with *S. halepense* being more receptive to cross pollination than *S. bicolor*. No hybrids were formed when *S. bicolor* was the female (Fig. 5(A), (B)), while when *S. halepense* was the female, hybrids were formed (Fig. 5(C), (D)). Additionally, the hybrid parent was also more receptive to cross pollination than *S. bicolor* and displayed an allelic shift more towards HEX (*S. halepense*) than FAM (*S. bicolor*) (Fig. 6(A)).

Within parent set 1 (*S. halepense* male x *S. bicolor* female), all of the resulting F₁ individuals were assigned *S. bicolor* (Fig. 5(A), (B)). Within parent set 2 (*S. bicolor* male x *S. halepense* female) 33% of the individuals were assigned as *S. halepense*, while the resulting 67% were assigned as hybrids (Fig. 5(C), (D)). Overall results in crosses between *S. bicolor* and *S. halepense* suggest that the *S. bicolor* parent utilized in this study is less receptive to cross pollination than its weedy relative. These crosses confirm crop-to-weed introgression observed in other studies.^{30,32} Previous studies have shown a wide range of hybridization between *S. bicolor* and *S. halepense* depending on experimental design. Field studies have detected hybrid formation of 2 or 11% when *S. halepense* received pollen at 100 or 5 meters from *S. bicolor*.³³ Much lower hybrid formation (less than 2% across all genotypes) was found in a field study with uniform *S. halepense* infestation in a *S. bicolor* field (0.76 m row spacing) and *S. bicolor* genotypes as the pollen receivers.¹⁰ The *S. halepense* biotype used in our study did show significant hybrid formation when it was the pollen receiver (67%), which was much higher than the previous study reported³³ most likely because the panicles were tied and bagged together from initial pollen release to mature seed.

Within parent set three (*S. bicolor* male x hybrid female), individuals from the female hybrid parent were assigned genotypes through our KASP assay (Fig. 6(A), (B)). Overall, 95% were genotyped as hybrid, 5% were assigned *S. halepense*, and no individuals were assigned as *S. bicolor*. Surprisingly, the hybrids that were formed from this cross demonstrated amplification closer to the HEX fluorophore, which corresponds more to *S. halepense*. One possible explanation involves the sheer overabundance of *S. halepense* ITS copies in comparison to *S. bicolor*.

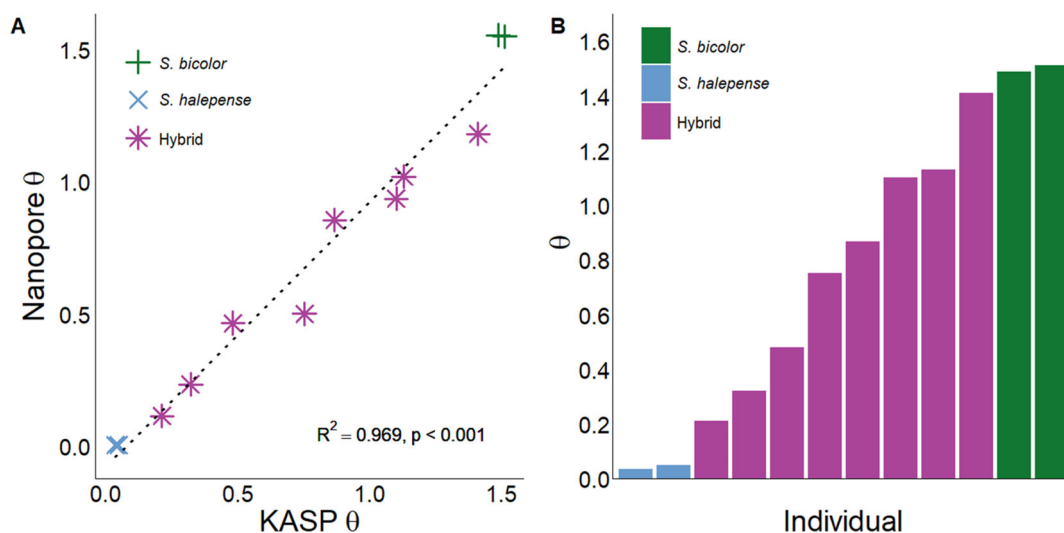


Figure 4. Correlation of arctan transformations (A) of the FAM/HEX ratio (KASP) and nucleotide ratio (Oxford Nanopore). *S. bicolor* (+), *S. halepense* (x), and hybrid (*). Significant positive relationship between the two methods ($R^2 = 0.97$, $P < 0.001$). The arctan transformation (B) from the KASP assay of the same biotypes in (A).

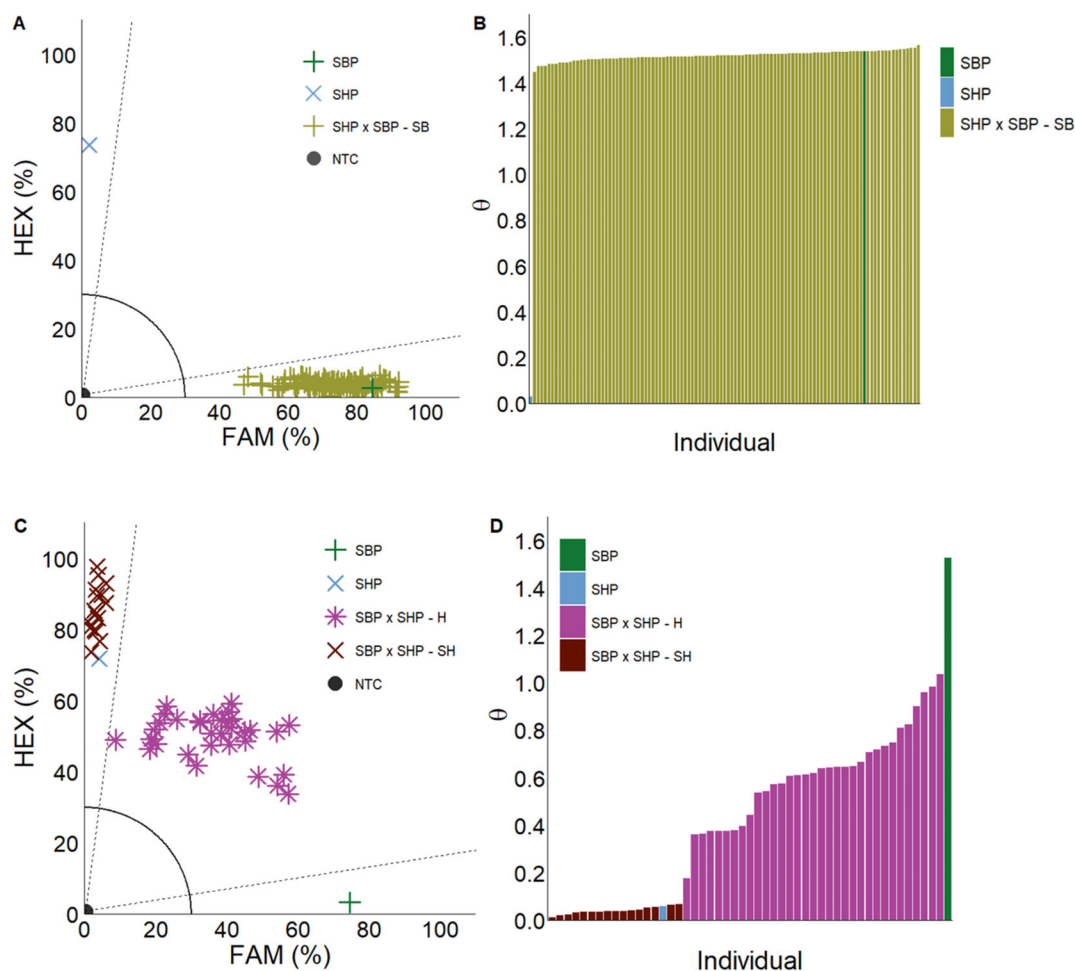


Figure 5. KASP assays and arctan transformation graphs from *S. bicolor* and *S. halepense* greenhouse crosses. Parent set 1 *S. bicolor* female parent (SBP+), *S. halepense* male parent (SHPx), and F_1 individuals assigned *S. bicolor* (SHP x SBP-SB+) through KASP assay (A). Parent set 2 *S. halepense* female parent (SHPx), *S. bicolor* male parent (SBP+), and F_1 individuals assigned hybrid (SBP x SHP-H*) or *S. halepense* (SBP x SHP-SHx) through KASP assay (C). Arctan transformation graph of each individual in the KASP assays, respectively, (B, D). NTC (●): no-template control. Dashed lines represent cutoffs for making genotyping assignments. The solid quarter circle line is the cutoff for no amplification.

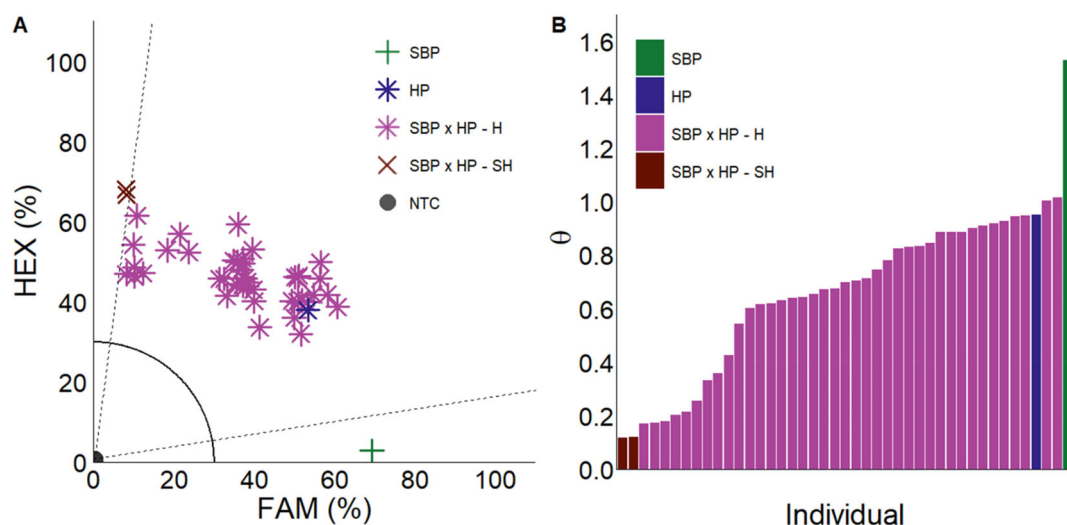


Figure 6. KASP assay and arctan transformation graphs from *S. bicolor* and hybrid greenhouse cross. Parent set 3 hybrid female parent (HP*), *S. bicolor* male parent (SBP+), and F_1 individuals assigned hybrid (SBP x HP-H*) or *S. halepense* (SBP x HP-SHx) through KASP assay (A). Arctan transformation graph of each individual in KASP assay (B). NTC (●): no-template control. Dashed lines represent cutoffs for making genotyping assignments. The solid quarter circle line is the cutoff for no amplification.

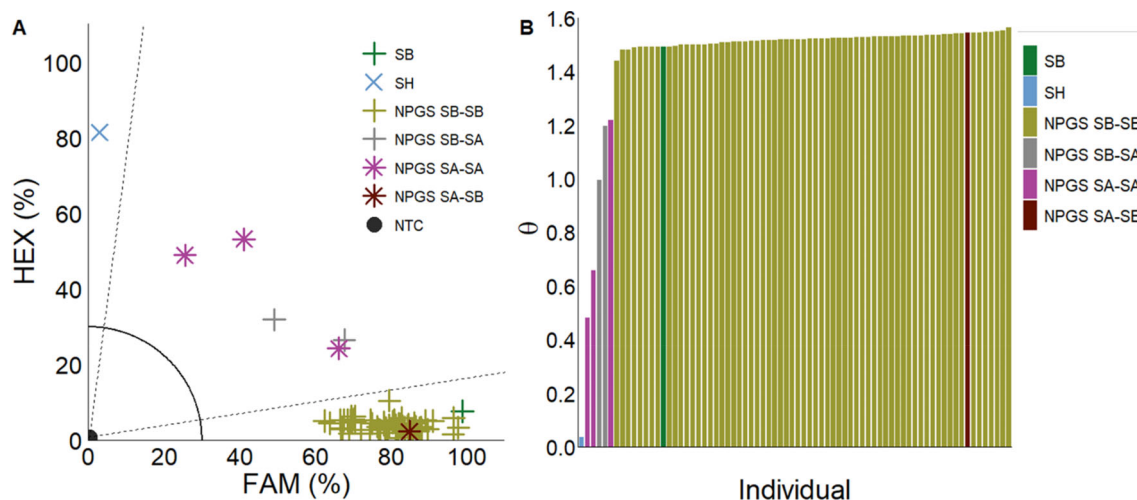


Figure 7. KASP assay (A) and arctan transformation graph (B) for NPGS *Sorghum* collection: *S. bicolor* control (SB+), *S. halepense* control (SHx), NPGS *S. alnum* assigned *S. alnum* (NPGS SA-SA*), NPGS *S. alnum* assigned *S. bicolor* (NPGS SA-SB*), NPGS *S. bicolor* assigned *S. bicolor* (NPGS SB-SB+), NPGS *S. bicolor* assigned *S. alnum* (NPGS SB-SA+), no-template control (NTC●). Dashed lines represent cutoffs for making genotyping assignments. The solid quarter circle line is the cutoff for no amplification.

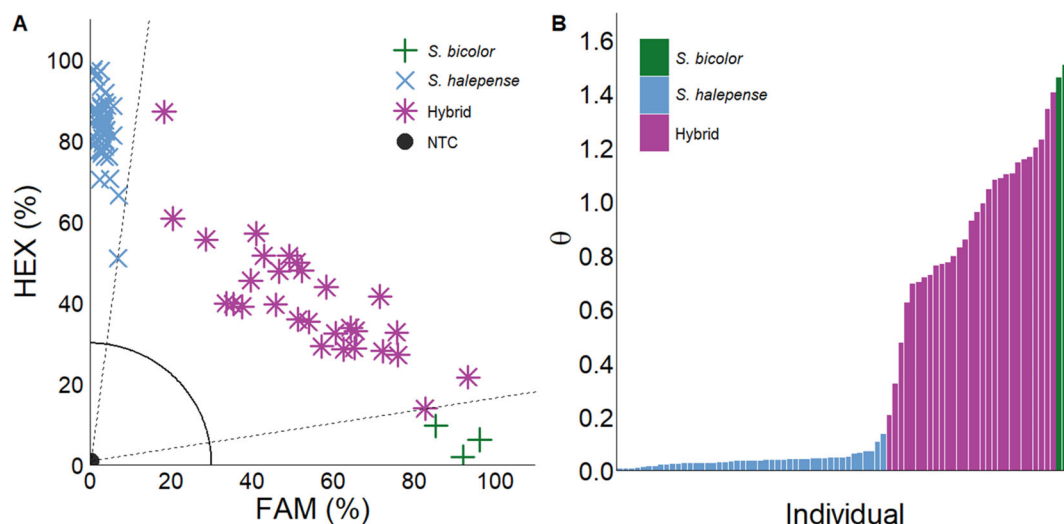


Figure 8. KASP assay (A) and arctan transformation graph (B) for unknown samples in MSU collection: *S. bicolor* (+), *S. halepense* (x), hybrid (*), and no-template controls (NTC●). Dashed lines represent cutoffs for making genotyping assignments. The solid quarter circle line is the cutoff for no amplification.

Previous greenhouse³⁴ and field¹⁰ studies reported that the majority of hybrid seedlings were tetraploids. Given this factor, with the addition that the *S. halepense* KASP primer extensively amplifies subgenome B, hybrids are more biased towards *S. halepense*.

3.4 Validation through known and unknown collections

Sixty-six biotypes from the NPGS *Sorghum* collection were assigned genotypes through our KASP assay (Fig. 7(A), (B)). Of the 62 that were *S. bicolor* lines, 97% aligned with the genotype given by NPGS, while 3% were assigned as hybrid through our KASP assay. The four NPGS *S. alnum* lines, 75% aligned with the genotype given, while 25% were assigned as *S. bicolor* through our KASP assay. KASP results support the fact that hybrids between *S. bicolor* and *S. halepense* can be difficult to distinguish morphologically.¹¹ A lack of genetic tools especially for

outcrossing species may have led to misidentified individuals in germplasm collections. Studies differentiating species often utilize multiple SNPs within conserved regions like the ITS.^{23,24} In contrast, our study successfully distinguished between *S. bicolor*, *S. halepense*, and their hybrids using a single SNP. This was validated by testing accessions that spanned genetic and spatial diversity, including biotypes from the United States to the *Sorghum* center of origin (present day Ethiopia).³⁵ The consistency across these diverse accessions, supports this SNP marker to be of great significance between these two species.

Seventy-eight biotypes were collected across diverse regions of the United States and assigned genotypes through our KASP assay (Supporting Information, Table S2). Of the 78 biotypes collected, 59% were assigned *S. halepense*, 37% were assigned as a hybrid, and 4% were assigned *S. bicolor* (Fig. 8(A), (B)). Biotypes screened through the KASP assay were all collected under the

assumption that they were *S. halepense* or had an intermediate phenotype between *S. bicolor* and *S. halepense*. These assignments show that a significant portion of the samples collected were indeed hybrids (37%), further supporting that hybrids are difficult to identify without genetic tools. This tool will also be valuable for examining the role of phenotypic plasticity in weeds. Often phenotypic plasticity is used to explain intraspecific variation; current genomic advances could help refine how we classify and interpret these differences within species. Many of these biotypes were sent and labelled as *S. halepense* seed stock from other universities and collaborators or collected under the same assumption.

The outcrossing rate under field conditions have been reported as less than 2% in male sterile *S. bicolor* parent lines pollinated by *S. halepense*.¹⁰ *Sorghum halepense* is expected to have a higher outcrossing rate than *S. bicolor* but still considered to be mostly self-pollinated.³⁶ These KASP assigned genotypes from diverse environments also suggest that there could be increased hybridization between these species than typically thought in the United States.

4 CONCLUSION

This study presents a novel, cost-effective KASP assay capable of accurately distinguishing between *S. bicolor*, *S. halepense*, and their hybrids. By validating this assay through Nanopore sequencing, greenhouse crosses, and diverse germplasm, we demonstrated its utility across genetic backgrounds and environments. The detection of crop-to-weed introgression in greenhouse crosses and misidentified accessions notes the need for genetic tools. High frequency of hybridization in natural collections suggests more widespread hybridization than previously thought. This is the first study to provide a genetic tool that can accurately differentiate between *S. bicolor*, *S. halepense*, and their hybrids.

ACKNOWLEDGEMENTS

The authors thank collaborators from the University of Idaho, Purdue University, Southern Illinois University, Virginia Tech University, Texas A&M University, Kansas State University, and Cornell University for providing samples. Additionally, thanks to NPGS for supplying *S. bicolor* and *S. alnum* accessions. Project GREEN (Generating Research and Extension to meet Economic and Environmental Needs) Michigan State University funded this work. Also, to Eduardo Carlos Rudell for initial KASP training and Mohit Mahey for assistance with synteny analysis. We thank the DOE Joint Genome Institute and collaborators for prepublication access to the genome of *Sorghum halepense*.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

- Kopecký D, Martín A and Smýkal P, Interspecific hybridization and plant breeding: from historical retrospective through work of Mendel to current crops. *Czech J Genet Plant Breed* **58**:113–126 (2022).
- Hackauf B, Siekmann D and Fromme FJ, Improving yield and yield stability in winter rye by hybrid breeding. *Plants* **11**:2666 (2022).
- Ellstrand NC, Prentice HC and Hancock JF, Gene flow and introgression from domesticated plants into their wild relatives, in *Horizontal gene transfer*. Elsevier, San Diego, USA, pp. 217–236 (2002).
- Merotto A Jr, Goulart IC, Nunes AL, Kalsing A, Markus C, Menezes VG *et al.*, Evolutionary and social consequences of introgression of non-transgenic herbicide resistance from rice to weedy rice in Brazil. *Evol Appl* **9**:837–846 (2016).
- Le Corre V, Siol M, Vigouroux Y, Tenaillon MI and Delye C, Adaptive introgression from maize has facilitated the establishment of teosinte as a noxious weed in Europe. *Proc Natl Acad Sci* **117**:25618–25627 (2020).
- Campbell LG, Snow AA and Ridley CE, Weed evolution after crop gene introgression: greater survival and fecundity of hybrids in a new environment. *Ecol Lett* **9**:1198–1209 (2006).
- Winchell F, Stevens CJ, Murphy C, Champion L and Fuller DQ, Evidence for sorghum domestication in fourth millennium BC eastern Sudan: spikelet morphology from ceramic impressions of the Butana group. *Curr Anthropol* **58**:673–683 (2017).
- Venkateswaran K, Elangovan M and Sivaraj N, Origin, domestication and diffusion of *Sorghum bicolor*, in *Breeding Sorghum for diverse end uses*. Elsevier, Cambridge, USA, pp. 15–31 (2019).
- Service FA, Production - Sorghum [May, 20 2025].
- Sias C, Subramanian N, Hodnett G, Rooney W and Bagavathiannan M, Rate of crop-weed hybridization in *Sorghum bicolor* × *Sorghum halepense* is influenced by genetic background, pollen load, and the environment. *Evol Appl* **16**:781–796 (2023).
- Dweikat I, A diploid, interspecific, fertile hybrid from cultivated sorghum, *Sorghum bicolor*, and the common johnsongrass weed *Sorghum halepense*. *Mol Breed* **16**:93–101 (2005).
- Ananda GK, Myrans H, Norton SL, Gleadow R, Furtado A and Henry RJ, Wild sorghum as a promising resource for crop improvement. *Front Plant Sci* **11**:1108 (2020).
- Hossain MS, Islam MN, Rahman MM, Mostofa MG and Khan MAR, Sorghum: a prospective crop for climatic vulnerability, food and nutritional security. *J Agric Food Res* **8**:100300 (2022).
- Lakoba VT, Atwater DZ, Thomas VE, Strahm BD and Barney JN, A global invader's niche dynamics with intercontinental introduction, novel habitats, and climate change. *Global Ecol Conserv* **31**:e01848 (2021).
- Jhala AJ, Beckie HJ, Peters TJ, Culpepper AS and Norsworthy JK, Interference and management of herbicide-resistant crop volunteers. *Weed Sci* **69**:257–273 (2021).
- Jhala AJ, Beckie HJ, Mallory-Smith C, Jasieniuk M, Busi R, Norsworthy JK *et al.*, Transfer of resistance alleles from herbicide-resistant to susceptible grass weeds via pollen-mediated gene flow. *Weed Technol* **35**:869–885 (2021).
- Heap I, International Herbicide-Resistant Weed Database [May 14 2025].
- Mwamahonje A, Eleblu JSY, Ofori K, Deshpande S, Feyissa T and Tongoona P, Drought tolerance and application of marker-assisted selection in sorghum. *Biology* **10**:1249 (2021).
- Baloch FS, Altaf MT, Liaqat W, Bedir M, Nadeem MA, Cömertpay G *et al.*, Recent advancements in the breeding of sorghum crop: current status and future strategies for marker-assisted breeding. *Front Genet* **14**:1150616 (2023).
- Maity A, Young B, Subramanian N and Bagavathiannan M, Pollen-mediated transfer of herbicide resistance between johnsongrass (*Sorghum halepense*) biotypes. *Sci Rep* **12**:7663 (2022).
- Song L, Wang R, Yang X, Zhang A and Liu D, Molecular markers and their applications in marker-assisted selection (MAS) in bread wheat (*Triticum aestivum* L.). *Agriculture* **13**:642 (2023).
- Yang G, Chen S, Chen L, Sun K, Huang C, Zhou D *et al.*, Development of a core SNP arrays based on the KASP method for molecular breeding of rice. *Rice* **12**:1–18 (2019).
- Anderson JD, Gastelbondo M and Chambers AH, Diagnostic KASP markers differentiate *Vanilla planifolia*, *V. odorata*, *V. pompona*, and their hybrids using leaf or cured pod tissues. *Mol Biol Rep* **50**:707–717 (2023).
- Patterson EL, Fleming MB, Kessler KC, Nissen SJ and Gaines TA, A KASP genotyping method to identify northern watermilfoil, Eurasian watermilfoil, and their interspecific hybrids. *Front Plant Sci* **8**:752 (2017).
- Ertiro BT, Ogugo V, Worku M, Das B, Olsen M, Labuschagne M *et al.*, Comparison of Kompetitive allele specific PCR (KASP) and

- genotyping by sequencing (GBS) for quality control analysis in maize. *BMC Genomics* **16**:1–12 (2015).
- 26 Aboul-Maaty NA-F and Oraby HA-S, Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. *Bull Natl Res Cent* **43**:1–10 (2019).
 - 27 Tamura K, Stecher G and Kumar S, MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* **38**:3022–3027 (2021).
 - 28 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10**:giab008 (2021).
 - 29 Brusa A, Patterson E and Fleming M, Modifications of kompetitive allele-specific PCR (KASP) genotyping for detection of rare alleles, in *Plant Genotyping: Methods and Protocols*. Springer, New York, NY, pp. 173–189 (2023).
 - 30 Paterson AH, Kong W, Johnston RM, Nabukalu P, Wu G, Poehlman WL *et al.*, The evolution of an invasive plant, *Sorghum halepense* L. ('Johnsongrass'). *Front Genet* **11**:317 (2020).
 - 31 Radanović A, Sprycha Y, Jocković M, Sundt M, Miladinović D, Jansen C *et al.*, KASP markers specific for the fertility restorer locus Rf1 and application for genetic purity testing in sunflowers (*Helianthus annuus* L.). *Genes* **13**:465 (2022).
 - 32 Morrell P, Williams-Coplin T, Lattu A, Bowers J, Chandler J and Paterson A, Crop-to-weed introgression has impacted allelic composition of johnsongrass populations with and without recent exposure to cultivated sorghum. *Mol Ecol* **14**:2143–2154 (2005).
 - 33 Arriola PE and Ellstrand NC, Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense*, and crop sorghum, *S. bicolor*. *Am J Bot* **83**:1153–1159 (1996).
 - 34 Hodnett GL, Ohadi S, Pugh NA, Bagavathiannan MV and Rooney WL, *Sorghum bicolor* x *S. halepense* interspecific hybridization is influenced by the frequency of 2n gametes in *S. bicolor*. *Sci Rep* **9**:17901 (2019).
 - 35 Wondimu Z, Dong H, Paterson AH, Worku W and Bantte K, Genetic diversity, population structure, and selection signature in Ethiopian sorghum [*Sorghum bicolor* L.(Moench)] germplasm. *G3* **11**:jkab087 (2021).
 - 36 Ohadi S, Hodnett G, Rooney W and Bagavathiannan M, Gene flow and its consequences in *Sorghum spp.* *Crit Rev Plant Sci* **36**:367–385 (2017).