the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) **118,** 106–123 doi: 10.1111/tpj.16594

Haplotypes at the sorghum *ARG4* and *ARG5* NLR loci confer resistance to anthracnose

Nida Habte[1,†], Gezahegn Girma[1,†], Xiaochen Xu[2], Chao-Jan Liao[1], Adedayo Adeyanju[2], Sara Hailemariam[1], Sanghun Lee[1], Pascal Okoye[1], Gebisa Ejeta[2] and Tesfaye Mengiste[1,*]

[1]*Department of Botany and Plant Pathology, Purdue University, West Lafayette, Indiana 47907, USA, and*
[2]*Department of Agronomy, Purdue University, West Lafayette, Indiana 47907, USA*

Received 16 September 2023; revised 1 December 2023; accepted 6 December 2023; published online 18 December 2023.
*For correspondence (e-mail tmengist@purdue.edu).
†These authors contributed equally.

SUMMARY

Sorghum anthracnose caused by the fungus *Colletotrichum sublineola* (*Cs*) is a damaging disease of the crop. Here, we describe the identification of *ANTHRACNOSE RESISTANCE GENES* (*ARG4* and *ARG5*) encoding canonical nucleotide-binding leucine-rich repeat (NLR) receptors. *ARG4* and *ARG5* are dominant resistance genes identified in the sorghum lines SAP135 and P9830, respectively, that show broad-spectrum resistance to *Cs*. Independent genetic studies using populations generated by crossing SAP135 and P9830 with TAM428, fine mapping using molecular markers, comparative genomics and gene expression studies determined that *ARG4* and *ARG5* are resistance genes against *Cs* strains. Interestingly, *ARG4* and *ARG5* are both located within clusters of duplicate NLR genes at linked loci separated by ~1 Mb genomic region. SAP135 and P9830 each carry only one of the *ARG* genes while having the recessive allele at the second locus. Only two copies of the *ARG5* candidate genes were present in the resistant P9830 line while five non-functional copies were identified in the susceptible line. The resistant parents and their recombinant inbred lines carrying either *ARG4* or *ARG5* are resistant to strains Csgl1 and Csgrg suggesting that these genes have overlapping specificities. The role of *ARG4* and *ARG5* in resistance was validated through sorghum lines carrying independent recessive alleles that show increased susceptibility. *ARG4* and *ARG5* are located within complex loci displaying interesting haplotype structures and copy number variation that may have resulted from duplication. Overall, the identification of anthracnose resistance genes with unique haplotype stucture provides a foundation for genetic studies and resistance breeding.

Keywords: anthracnose, *Colletotrichum sublineola*, *Sorghum bicolor*, NLRs, NLR cluster, R-gene.

INTRODUCTION

Sorghum [*Sorghum bicolor* (L.) Moench] is among the most important cereal crops used for food, feed, biofuels, and alcoholic beverages. It is a staple food crop in developing countries while in the developed world, it is mostly used as livestock feed, a source of biofuels, or for making alcoholic beverages. Sorghum anthracnose caused by the fungal pathogen *Colletotrichum sublineola* is a major biotic constraint to its production. The pathogen mainly affects the leaf tissue where symptoms are more obvious but other plant parts are also affected. Symptoms include chlorotic flecks, acervuli formation, necrotic lesions, and death of leaves. Host resistance is considered the most effective disease control strategy, but its genetic control is less understood.

The sorghum conversion program (Stephens et al., 1967) followed by the identification of lines resistant to anthracnose (Rosenow et al., 1997) were the first set of studies that aimed at understanding the inheritance of genetic resistance to the disease. These early programs led to the identification of sorghum lines such as SC-748-5 that showed resistance in diverse environments (Mehta et al., 2005). The *Cg1* resistance gene in SC-748-5 was mapped to the distal end of chromosome 5 (Ramasamy et al., 2008), but later it was suggested that the resistance in SC-748-5 is not due to a single gene (Burrell et al., 2015). To understand the pathotypes of *C. sublineola*, a set of 18 lines were developed as differentials (Prom et al., 2012) which served as source germplasms for mapping of anthracnose-resistant genes. The sorghum differential lines SC112-14, SC414-12E, SC748-5, and other sorghum lines (BK-7 and SC155-14E) were used to map anthracnose resistance loci (Burrell et al., 2015; Cuevas et al., 2014;

Felderhoff et al., 2016; Patil et al., 2017; Ramasamy et al., 2008). These include three closely linked loci on the distal end of chromosome 5 and one on chromosome 9. The four loci were consistently detected by both bi-parental mapping approaches and recent genome-wide association studies (Cuevas et al., 2018, 2019; Cuevas & Prom, 2020). The identified loci encompass large genomic regions that harbor genes with potential functions in multiple defense response pathways (Cuevas et al., 2019). This has complicated identification of the specific genes underlying resistance. Moreover, the identified QTLs explained only a small proportion of the observed phenotypic variation (Cuevas et al., 2018, 2019). The source germplasms for the identified resistance alleles were mostly of East African origin (Ethiopia and Sudan), and the detected QTLs are known to mediate resistance against pathotypes from USA where the studies were conducted (Patil et al., 2017). Thus, genetic mapping of resistance genes using germplasms with broad-spectrum resistance to the strains of the fungus from different sources may be vital to identify new loci potentially conferring resistance in diverse environments. Moreover, application of whole genome sequences and other genomic tools may improve mapping of anthracnose resistance genes underlying the various QTLs. The recent discovery of anthracnose resistance genes by our group (Lee et al., 2022; Mewa et al., 2023) provided new insights into the complexity and regulation of anthracnose immune response genes in sorghum and revealed the roles of next-generation genomic tools in unraveling the mechanisms of disease resistance in crop plants.

We studied mechanisms of anthracnose resistance in the sorghum lines SAP135 and P9830 that was described as part of PhD thesis (Habte, 2021). These lines showed complete resistance to Cs strains, but the underlining mechanisms were not known. SAP135 (PI 576385 and SC1070) has West African origin with a broad-spectrum resistance to anthracnose and rust and was among resistant accessions used in recent genome-wide association studies (GWAS) (Cuevas et al., 2018, 2019). SAP135 was resistant to five different *C. sublineola* strains (Csgrg, Csgl1, Csgl2, Cs27, and Cs29) whereas P9830 was resistant to all but susceptible to Cs27. In both cases anthracnose resistance was inherited as a dominant trait. Bulked segregant analyses using whole genome resequencing of bi-parental segregating or recombinant inbred lines (RILs) from crosses between SAP135 or P9830 to TAM428 followed by fine mapping identified two tightly linked complex resistance loci, separated by ~1 Mb, here after designated *ARG4* and *ARG5*, harboring a cluster of nucleotide-binding leucine-rich repeat (NLR) genes corresponding to each of the resistant lines. Analysis of DNA sequence polymorphisms of RILs and parental lines identified the specific genes underlying the resistance in SAP135 and P9830. SAP135 carried a functional *ARG4* gene (Sobic.008G166400), while P9830 carried a functional *ARG5* (Sobic.008G177900), both encode highly related NLRs that are phylogenetically aligned with the leaf rust resistance gene *Lr1*. *ARG4*, and *ARG5* were validated using independent alleles from distinct sorghum lines. In sum, we identified two anthracnose resistance genes with interesting genomic organization that can be used for resistance breeding solely or in combination.

RESULTS

Identification of anthracnose resistance loci in sorghum and their genetic inheritance

Two anthracnose-resistant lines SAP135 and P9830 were identified from a genetic screen for enhanced resistance among natural sorghum variants. The responses of SAP135, P9830, and other selected sorghum lines to five strains of *C. sublineola* are summarized (Table 1). Spray inoculation of plants revealed that the sorghum line SAP135 is consistently resistant to all the strains of *C. sublineola* tested (Csgrg, Csgl1, Csgl2, Cs27, and Cs29) while TAM428 is highly susceptible (Figure 1 and Table 1). Observation of leaves of TAM428 after inoculation with Csgrg, Csgl1, and Cs27 indicated high susceptibility with extensive chlorosis whereas SAP135 showed localized resistance responses without any apparent disease symptoms (Figure 1c). P9830 was resistant to Csgrg, Csgl1, Csgl2, and Cs29 but was susceptible to Cs27 (Table 1, Figure 1a–c). In addition, four sorghum lines (RTx430, Ai4, BTx631, and SC23) were susceptible to all the strains tested.

Genetic analyses, based on segregation ratios observed in the F2 populations resulting from the cross between SAP135 and TAM428, indicated that resistance to Csgrg in SAP135 is governed by a single dominant gene. Among the 203 F2 plants, 161 were resistant while the remaining 42 were susceptible to Csgrg (Table 2)

Table 1 Disease reactions of parental lines and variants to anthracnose strains

	Anthracnose strains				
Lines	Csgrg	Csgl1	Csgl2	Cs27	Cs29
SAP135	R	R	R	R	R
P9830	R	R	R	S	R
Rio	R	R	S	S	S
TAM428	S	S	S	S	S
RTx430	S	S	S	S	S
Ai4	S	S	S	S	S
BTx631	S	S	S	S	S
SC23	S	S	S	S	S

R, resistant and S, susceptible, disease symptoms with observed fungal growth.

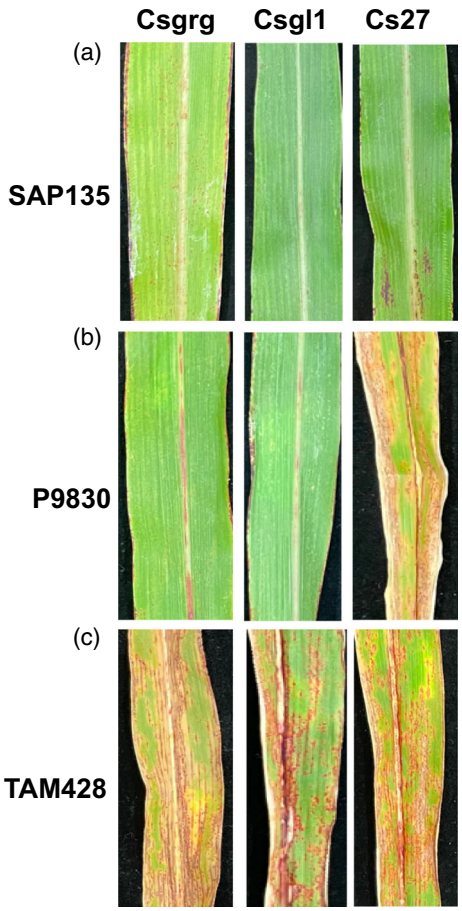

Figure 1. Responses of the sorghum lines SAP135 (a), P9830 (b), and TAM428 (c) to inoculation with *C. sublineolum* strains Csgrg, Csgl1, and Cs27.

Table 2 Inheritance and validation of genetic resistance to *C. sublineola* strain Csgrg in SAP135 based on selected InDel markers in *ARG4* locus. Numbers in boxes represent the number of F2s under each phenotype and genotype groups

		Markers			
Reaction	Genotype	ARG4-5945	ARG4-5998	ARG4-6016	ARG4-6059
Susceptible	1[a]	0	0	0	0
	2	38	42	42	40
	3	4	0	0	2
Sub-total		42	42	42	42
Resistant	1	62	65	65	66
	2	2	0	0	4
	3	97	96	96	91
Sub-total		161	161	161	161
Grand-total		203	203	203	203
		Single gene χ^2 value = 2.1*			

[a]1 = SAP135 allele, 2 = TAM428 allele, 3 = heterozygous.
*= significant at $P < 0.05$.

indicating a 3:1 ratio which fits into a single dominant gene model (χ^2 value = 2.1).

Mapping of *ANTHRACNOSE RESISTANCE GENE 4* (*ARG4*) using bulk segregant analysis

Genetic mapping of the resistance locus in SAP135 was initially conducted using the SAP135 x TAM428 population. The SNP-index from association analysis identified a single major peak for resistance to the *C. sublineola* strain Csgrg on the distal end of chromosome 8 (Figure 2). The SNP-index plot of R-bulk and S-bulk genomic DNA sequences identified a single genomic region associated

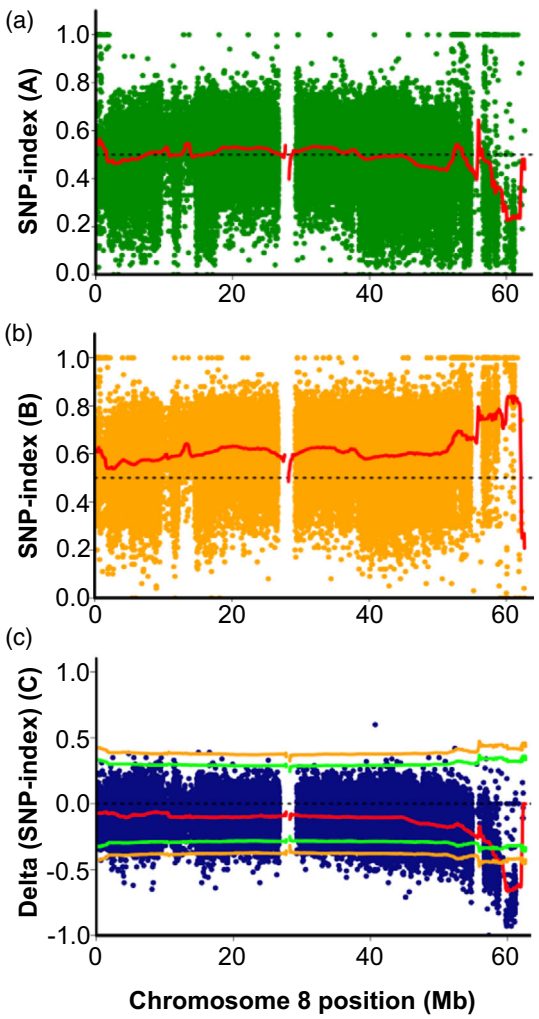

Figure 2. Identification of *ARG4* locus on chromosome 8 through BSA-seq analysis of resistant and susceptible pools of F3 families generated by crossing SAP135 and TAM428. Each dot corresponds to a SNP and the x-axis corresponds to chromosome position. The red line represents the average values of the SNP-index for each bulk or Δ(SNP-index) drawn by sliding windows analysis.
(a) SNP-index plot of R-bulk.
(b) SNP-index plot of S-bulk.
(c) A plot of Δ(SNP-index). The green and orange lines in the Δ(SNP-index) plot indicate a significance threshold ($P < 0.05$ and $P < 0.01$, respectively).

with resistance (Figure 2a,b). The peak region on chromosome 8 was identified as a valley of the Δ(SNP-index) plot below the orange line with a statistical significance of $P < 0.01$ (Figure 2c). This data mapped *ARG4* locus to ~60 Mb region on chromosome 8. The new locus is designated *ANTHRACNOSE RESISTANCE GENE 4* (*ARG4*), following our sequential naming of three other anthracnose resistance loci (*ARG1* to *ARG3*) identified recently using different mapping populations. *ARG1* and *ARG2* have been published (Lee et al., 2022; Mewa et al., 2023) while *ARG3* is pending publication. The segregation pattern of the SAP135 x TAM428 mapping populations varied when plants were inoculated with different strains suggesting SAP135 carries multiple independent loci conferring resistance to different strains. The current mapping is based on responses to Csgrg which was highly virulent to TAM428 and BTx623 but was avirulent to SAP135 (Table 1).

Fine mapping of *ARG4* locus

Fine mapping to identify the *ARG4* gene was conducted on 80 F3 families using 10 InDel markers (Table 3) that spanned about 5 Mb region. To delimit the location of *ARG4* locus, 23 recombinants were identified based on discordance of the genotype and phenotype data. By analyzing the number of recombinants, the locus was further narrowed to a 928 kb region delimited with two molecular markers (ARG4-5945 and ARG4-6038, Figure 3a–d). The two flanking markers are physically located at 59.45 and 60.38 Mb on the BTx623 reference genome (Sorghum bicolor v3.1.1). The tightly linked three InDel markers (ARG4-5998, ARG4-6008, and ARG4-6016) showed no recombination and co-segregated with the phenotype. To further confirm the resistance locus, 203 new F2 plants of SAP135 x TAM428 were genotyped at four of the InDel markers including two of the inner markers and phenotyped using the Csgrg strain. The data revealed that markers ARG4-5998 and ARG4-6016 showed 100% co-segregation with the phenotype confirming that the *ARG4* gene resides in this region defined by the two flanking markers (Table 2).

Identification of candidate genes in *ARG4* locus

The reference sorghum genome BTx623 (v3.1.1) has a total of 91 annotated genes (Table S1) within the mapped *ARG4* region, of which 67 were annotated to have predicted functions. Among the 67 genes, five genes are annotated with a putative function associated with plant disease resistance that include four closely located NBS-LRR disease resistance genes and one gene encoding disease resistance-responsive (dirigent-like protein) family protein (Table S2). The four annotated NBS-LRR genes are Sobic.008G166400, Sobic.008G166550, Sobic.008G167300,

Table 3 Fine mapping of *ARG4* locus: the green cells represent concordance and blue cells indicate discordance between phenotype and genotype data

No	Lines	Reaction	ARG4-5600	ARG4-5890	ARG4-5945	ARG4-5998	ARG4-6008	ARG4-6016	ARG4-6038	ARG4-6059	ARG4-6062	ARG4-6131
1	S6	S	2[a]	2	2	2	2	2	2	3	3	3
2	S7	S	3	3	2	2	2	2	2	2	2	2
3	S9	S	1	2	2	2	2	2	2	2	2	2
4	S12	S	2	2	2	2	2	2	2	2	2	3
5	S15	S	1	2	2	2	2	2	2	2	2	2
6	S17	S	3	2	2	2	2	2	2	2	2	2
7	S18	S	1	2	2	2	2	2	2	2	2	2
8	S23	S	2	2	2	2	2	2	2	2	2	3
9	S25	S	2	2	2	2	2	2	2	2	2	3
10	S26	S	3	2	2	2	2	2	2	2	2	2
11	S27	S	2	2	2	2	2	2	2	3	3	3
12	S28	S	3	3	2	2	2	2	2	2	2	2
13	S29	S	3	2	2	2	2	2	2	2	2	2
14	S30	S	1	1	3	2	2	2	2	2	2	2
15	S49	S	1	1	3	2	2	2	2	2	2	2
16	S52	S	2	2	2	2	2	2	2	2	2	3
17	S57	S	1	2	2	2	2	2	2	2	2	2
18	S58	S	3	2	2	2	2	2	2	2	2	2
19	S59	S	2	2	2	2	2	2	2	3	3	3
20	S64	S	3	3	2	2	2	2	2	2	2	2
21	R27	R	1	1	1	1	1	1	2	3	3	2
22	R67	R	2	3	3	3	3	3	3	3	3	3
23	R79	R	1	1	1	1	1	1	3	3	3	2
No. of recombinants			14	5	2	0	0	0	1	3	3	9

[a]1 = SAP135 allele, 2 = TAM428 allele, 3 = heterozygous.

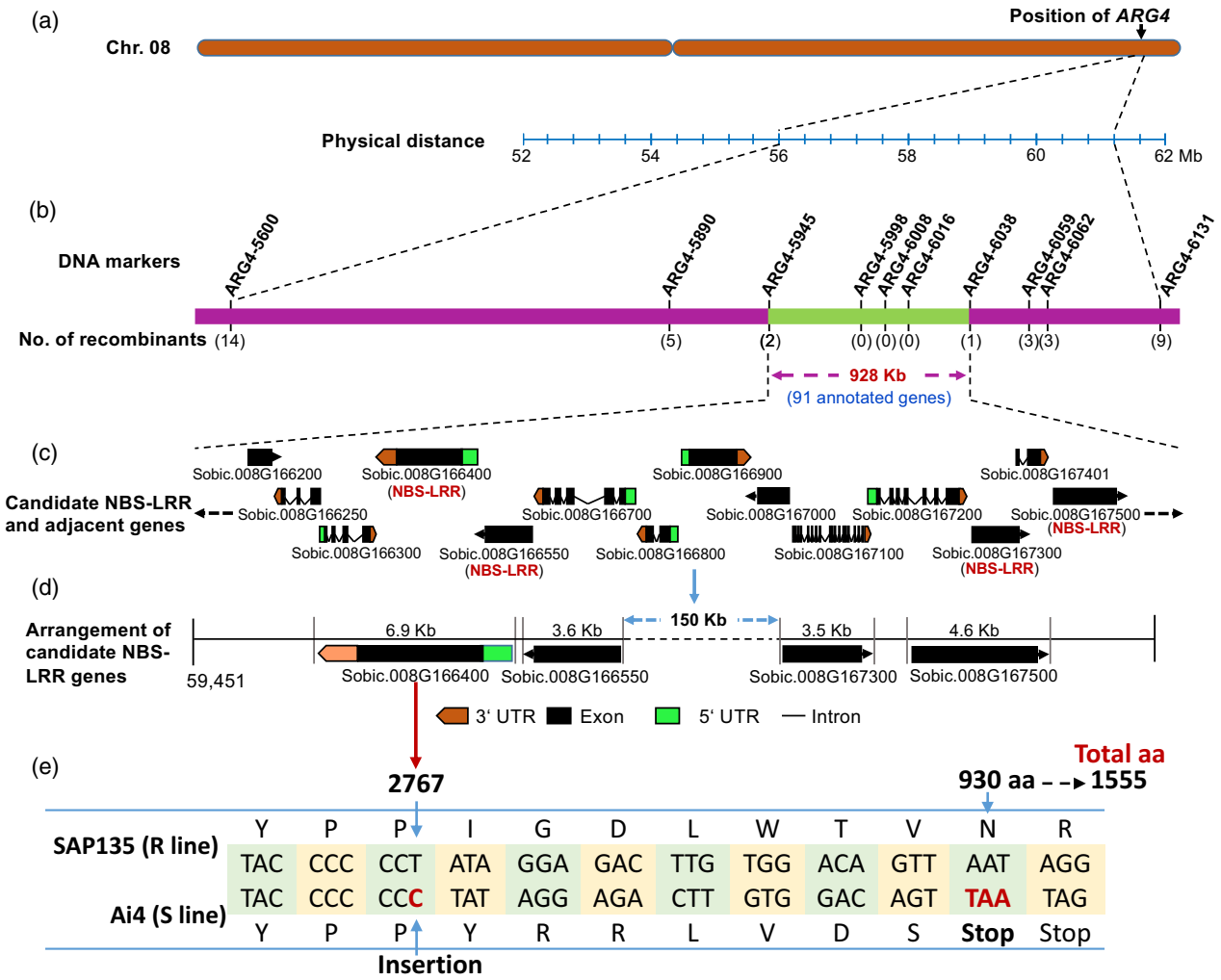

Figure 3. Genetic mapping of the *ARG4* locus.
(a) Physical position of *ARG4* locus on chromosome 8.
(b) Relative position of DNA markers used to narrow down the *ARG4* region and number of recombinants observed among F3 families of the cross between SAP135 and TAM428.
(c) Candidate NBS-LRR genes and partial illustration of adjacent genes with in the target region.
(d) Genomic organization of candidate NBS-LRR genes.
(e) Illustration of an insertion of a single nucleotide in a susceptible line Ai4 that causes a frameshift and pre-mature stop codon in *ARG4* gene.

and Sobic.008G167500 while the dirigent-like gene is annotated as Sobic.008G168800. The four NBS-LRR genes are clustered in two closely located genomic regions separated by about 150 kb flanked by two duplicated copies of NBS-LRR genes on both sides (Figure 3c,d). Visualization of genomic sequences of SAP135, TAM428 and the two bulks indicated that the two parental lines are only polymorphic for three synonymous SNPs located inside the coding region of Sobic.008G168800. The three SNPs, A/C, T/G and G/A are located at nucleotide position, 60 306 054, 60 306 822 and 60 306 879 bp on chromosome 8 of the sorghum reference genome (v3.1.1). By contrast, SAP135 and TAM428 are highly polymorphic for three of the NBS-LRR genes (Sobic.008G166400, Sobic.008G166550 and Sobic.008G167500). One of the NBS-LRR genes (Sobic.008G167300) only has a single synonymous polymorphic SNP (T/G) at position 60 141 356 between SAP135 and TAM428. Therefore, the three canonical disease resistance genes encoding putative NBS-LRR disease resistance protein were considered the most likely candidate genes associated with resistance to the Csgrg strain in SAP135.

Genomic organization, sequencing, and analysis of gene expression of candidate NBS-LRR genes in *ARG4* locus

To determine which of the three candidate NBS-LRR genes (Sobic.008G166400, Sobic.008G166550 and Sobic.008G167500) are responsible for the resistance in SAP135, visualization of predicted gene structures, sequencing, and

gene expression studies were conducted. The predicted gene structure for the NBS-LRR candidates in the sorghum reference genome indicates that Sobic.008G166550 and Sobic.008G167500 have no predicted 5′ and 3′ UTR regions while Sobic.008G166400 has a complete gene structure including both UTR regions (Figure 3c,d). Moreover, Sobic.008G166550 lacked a start codon in the reference genome, and TAM428. These observations suggest that Sobic.008G166550 and Sobic.008G167500 are pseudo genes. However, since the reference line, BTx623 which has 100% sequence similarity to TAM428 for all the candidate genes is susceptible to anthracnose, the lack of predicted UTR sequences in some of the candidates does not necessarily indicate that the genes have the same structure in SAP135. Therefore, all the three candidate genes including the potential 5′ UTR region of Sobic.008G166550 in SAP135 were sequenced to understand the nature of polymorphisms between SAP135 and TAM428. IGV visualization of Sobic.008G166550 region using whole genome sequences from SAP135, TAM428 and the resistant and susceptible bulks, indicated a deletion of large genomic region in TAM428, which was confirmed by sequencing of genomic DNA from SAP135 and TAM428. The deleted region corresponds to the upstream sequence of the gene and 5′ UTR sequences that are intact in the resistant SAP135 but deleted in TAM428. However, prediction of protein sequence from the SAP135 revealed that SAP135 contains a pre-mature stop codon in Sobic.008G166550, which disqualified this gene from consideration as the *ARG4* candidate.

The remaining candidate genes, Sobic.008G166400 and Sobic.008G167500 share over 80% similarity in amino acid sequences although Sobic.008G167500 lacks predicted UTR region. Both IGV visualization and sequencing of genomic DNA of SAP135 and TAM428 revealed that the two lines are highly polymorphic for both genes. The fact that Sobic.008G167500 has no predicted UTRs suggests that its expression might be affected due to lack of upstream regulatory sequences. Consistently, only Sobic.008G166400 is expressed, but no transcripts were detected for Sobic.008G167500 (Figure 4a). Therefore, Sobic.008G166400 is the most likely *ARG4* gene conferring anthracnose resistance in SAP135 and hereafter designated *ANTHRACNOSE RESISTANCE GENE 4* (*ARG4*) gene. Interestingly, *ARG4* is significantly induced at 24 and 36 h after inoculation with *C. sublineola* consistent with its resistance function (Figure 4b).

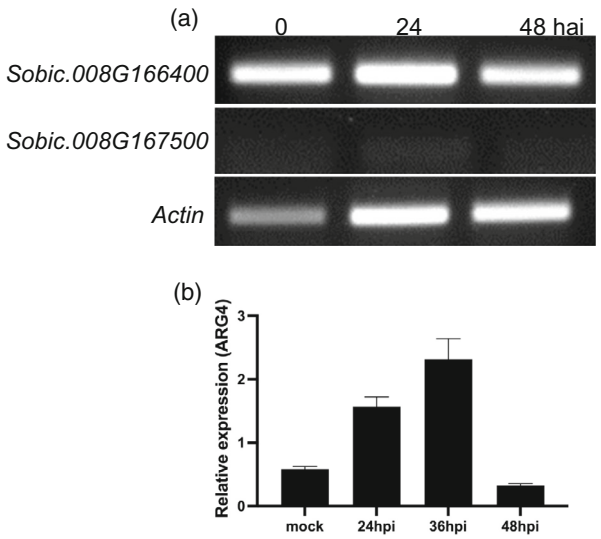

Figure 4. Expression of *ARG4* gene.
(a) Expression of candidate *ARG4* genes.
(b) Expression of *ARG4* gene (Sobic.008G166400) in the resistant parent (SAP135) in response to fungal inoculation. The sorghum Actin gene was used as a constitutive control. hpi, hours post inoculation.

Polymorphisms, protein domain structure, and variants of *ARG4* gene

ARG4 (Sobic.008G166400) contains a single exon that encodes a protein of 1555 amino acids (Figure S1). The *ARG4* alleles from SAP135 and TAM428 share 95.6 and 92.6% nucleotide and amino acid identity, respectively, in the coding regions. ARG4 is predicted to form coiled coil (CC), nucleotide-binding-site (NB-ARC), and leucine-rich-repeat (LRR) motifs (Figure 5) consistent with other NBS-LRR proteins. The 3D structure of ARG4 protein was predicted using Protein Homology/analogY Recognition Engine V 2.0 (Phyre^2) web portal (Kelley et al., 2015). Analyses of the 3D protein structure indicates notable differences between the ARG4 sequences in SAP135 and TAM428 (Figure 5; Figure S2). Obvious structural differences in the LRR domains of ARG4 sequences from SAP135 and TAM428 were observed. The ARG4 protein from SAP135 carried extended beta-sheet structure relative to the susceptible allele from TAM428. The LRR structure may affect effector binding ability (Wang et al., 2019). In addition, the ARC, CC, and NB regions show structural changes in ARG4 variants that may contribute to differences in resistance functions (Figure 5).

To provide further genetic evidence for the function of *ARG4*, additional sorghum variants were explored to identify distinct alleles of the *ARG4* gene. Sequencing of genomic DNA of the sorghum line Ai4, which was highly susceptible to anthracnose strain Csgrg revealed that the line carries a frameshift mutation in *ARG4* gene due to an insertion of a single nucleotide at position 2767 that resulted in a pre-mature stop codon after amino acid position 930 (Figure 3e; Figure S3). Another sorghum line, BTx631, which is also susceptible to Csgrg, carries a different allele of *ARG4* gene with unique polymorphism patterns compared to all the variants studied (Figure S3). BTx631 shares similarity with SAP135 for some of the

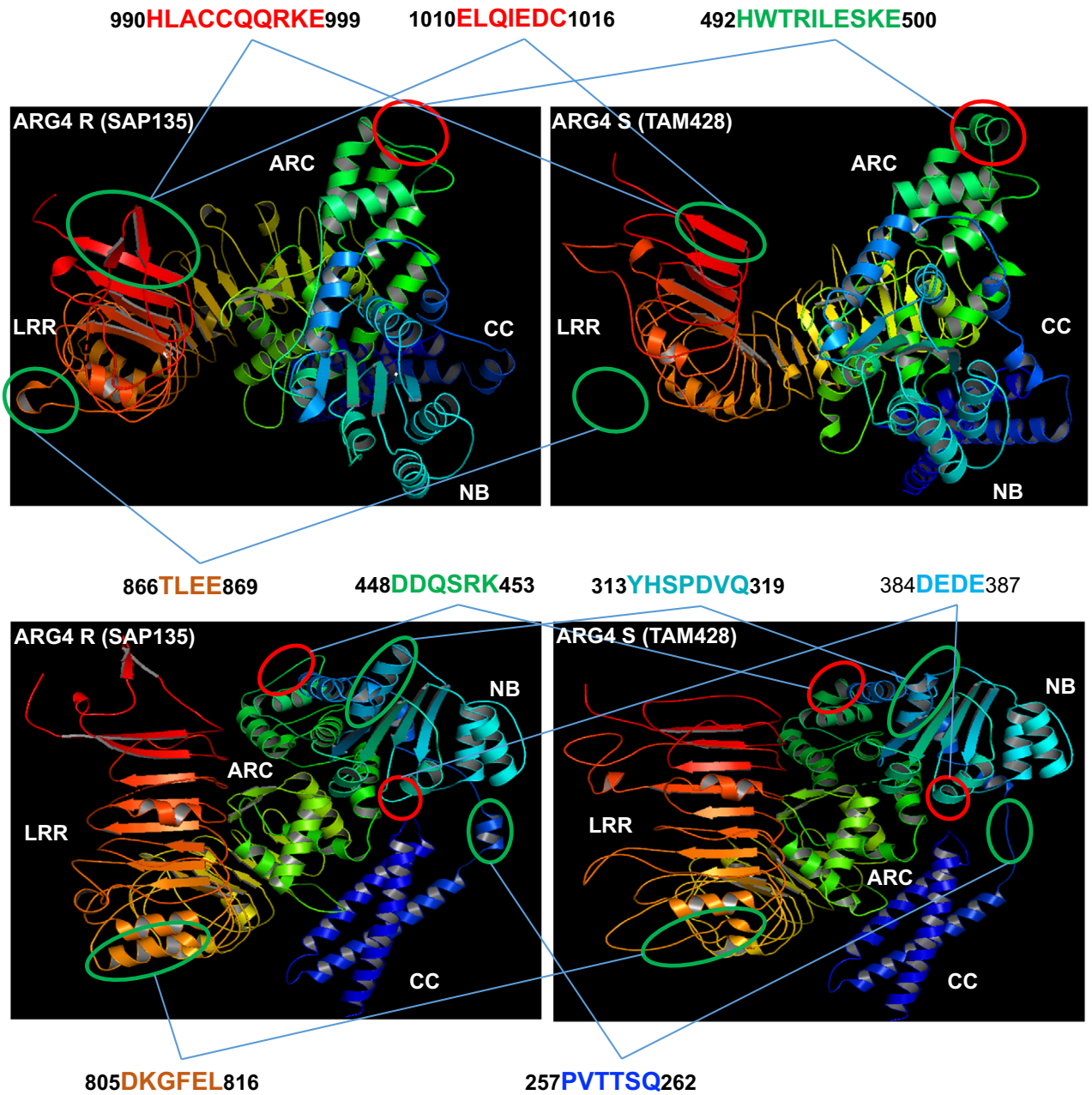

Figure 5. Prediction of tertiary structure of ARG4 proteins from the resistant and susceptible ARG4 alleles.
The tertiary structure of ARG4 was predicted based on the template c6j5tC (RPP13-like protein 4). The ARG4 structure from SAP135 and ATM428 was predicted by Phyre[2] (Kelley et al., 2015). The specific regions of ARG4 were labeled by red or green circles and represent the differences of structures and the color of amino acid residues correspond to the different regions of NB-LRR.

nucleotides that are polymorphic between the variants but also displayed unique sequences at various positions of the gene. The sorghum line RTx430, which is also susceptible to Csgrg, was similar to that of TAM428 although there were variations for a few nucleotides. These data validated that *ARG4* is the gene responsible for the resistance to the *Cs* strains. Based on the high amino acid sequence similarity to ARG4 and the availability of solved structures, the amino acid sequences of ARG4 from SAP135 was compared to HOPZ-ACTIVATED RESISTANCE 1 (ZAR1) (Wang et al., 2019). We found that the residue at amino acid position 730 of SAP135 is correlated to the ZAR1 amino acid residue at position 597. The Serine at this position is shifted to Phenylalanine in sorghum lines susceptible to *Cs* strains (Figure S3). The mutation at this position disrupts the ability to interact with other proteins, reduces the

ability to trigger cell death, and increased susceptibility to the biotrophic bacteria *Xanthomonas campestris* (Wang et al., 2019). The data imply that the amino acid change at this particular position in SAP135 likely affects effector recognition and thus influence its function in plant immunity.

ARG4-related genes are widespread in grass species but with no functional data.

BLASTP search of the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using ARG4 amino acid sequences from SAP135 revealed that disease resistance protein RGA2-like from the grass species green foxtail (*Setaria viridis*) showed high sequence relatedness (67% identity and 77% similarity). Sequence IDs and description of the sequence homologs are provided in Table S3. The next closely related homologs are also from grass species that include hypothetical and disease resistance proteins from foxtail millet (*Setaria italica*), white fonio (*Digitaria exilis*), and weeping lovegrass (*Eragrostis curvula*). Homologs from major crops include NBS-LRR proteins from barley (45% identity and 61% similarity), hypothetical protein from wheat (45% identity and 61% similarity) and a putative disease resistance protein RGA1 from rice (47% identity and 60% similarity). Moreover, the BLASTP search revealed another homologous protein from *Setaria italica* with 45% identity and 60% similarity to the *Arabidopsis* At3g14460 that is the homolog of ARG4 from *Arabidopsis*.

ARG4 co-localized with a resistance locus identified in P9830

The two mapping populations, SAP135 x TAM428 and P9830 x TAM428 identified the same genomic region but it was not clear whether a single locus or tightly linked loci confer anthracnose resistance in SAP135 and P9830. A rough mapping of anthracnose-resistant locus using Csgl1 in P9830 x TAM428 defined a resistance locus within the same region in chromosome 8 (Figure 6). Although two different strains (Csgrg and Csgl1) were used to map the two resistance loci, the fact that both SAP135 and P9830 were resistant to both strains suggest that the two lines carry the same gene conferring resistance to both strains. However, based on IGV visualization and sequencing of the candidate genes at the *ARG4* locus in the three parental lines, impactful polymorphisms were not detected between P9830 and TAM428 while SAP135 is highly polymorphic compared with both P9830 and TAM428. Thus, we concluded that the resistant P9830 line is similar to the susceptible TAM428 at the *ARG4* locus. These observations suggested that the resistance locus in P9830 is different from that in SAP135, which led to fine mapping of the resistance gene in P9830 using P9830 x TAM428 RILs. Some of the molecular markers that were used for fine mapping of *ARG4* using SAP135 x TAM428 mapping population were also used for mapping the resistance gene in P9830.

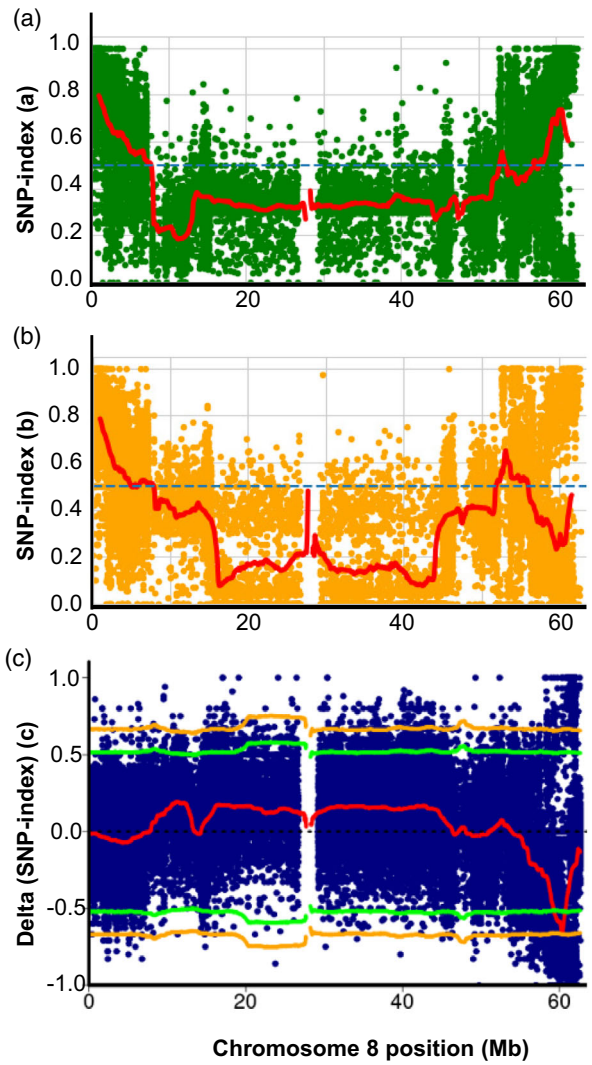

Figure 6. Identification of *ARG5* locus on chromosome 8 through BSA-seq analysis of resistant and susceptible pools of recombinant inbred lines (RILs) generated by crossing P9830 and TAM428. Each dot corresponds to a SNP and the x-axis corresponds to chromosome position. The red line represents the average values of the SNP-index for each bulk or Δ(SNP-index) drawn by sliding windows analysis.
(a) SNP-index plot of R-bulk.
(b) SNP-index plot of S-bulk.
(c) A plot of Δ(SNP-index). The green and orange lines in the Δ(SNP-index) plot indicate a significance threshold ($P < 0.05$ and $P < 0.01$, respectively).

Fine mapping of the resistance locus in P9830 using RILs

Based on fine mapping of the QTL region using 80 F6 RILs of P9830 x TAM428 and eight InDel markers, four of which were shared markers with the SAP135 x TAM428 population, the genomic region associated with resistance in P9830 was found to be a closely located new locus about 1 Mb away from *ARG4* gene. This new resistance locus, conferring anthracnose resistance in P9830 is similarly designated as *ANTHRACNOSE RESISTANCE GENE 5*

(*ARG5*). To identify the *ARG5* gene, 24 recombinants were identified based on recombinants between markers and co-segregation with the phenotype data (Table 4) which mapped *ARG5* to a 415 kb region delimited by InDel markers ARG5-6090 and ARG4-6131. The two flanking InDel markers are physically located at 60.9 and 61.3 Mb, respectively, on the BTx623 sorghum reference genome v3.1.1 and ~ 1 Mb away from the *ARG4* gene.

Identification of candidate genes in *ARG5* locus

Within the *ARG5* locus, a total of 60 genes (Table S4) are predicted in the reference sorghum genome (v3.1.1), of which 40 have predicted functions. Out of the 40 genes with predicted function, seven genes are annotated as having plant disease resistance function which include a cluster of five predicted NBS-LRR disease resistance genes (Sobic.008G177900, Sobic.008G178200, Sobic.008G178300, Sobic.008G178500, and Sobic.008G178600) and two genes (Sobic.008G174966 and Sobic.008G175032) that encode leucine-rich repeat protein kinase family proteins (Table S5). The NBS-LRR genes show high sequence similarity and are arranged in tandem with few genes in between (Figure 7). IGV visualization of the genomic sequences of resistant and susceptible bulks of P9830 x TAM428 population indicated that the two kinase family genes are not polymorphic between the two parental lines, therefore these were not considered as *ARG5* candidates. On the other hand, a major polymorphism for the NBS-LRR genes between resistant and susceptible bulks including copy number variation in the NBS-LRR genes was evident from a poor alignment to the reference of genomic sequences of some of the NBS-LRR genes from the resistant bulk. To determine whether there is variation in the number of the NBS-LRR genes between P9830 and TAM428, de novo assembly of the genomic sequences of the resistant bulk was conducted. A long node of 26 kb sequence was identified that include flanking sequences from both sides of the cluster of the five NBS-LRR genes available in the reference sorghum genome. Subsequent alignment to the reference, revealed that the resistant parent P9830 has only two NBS-LRR genes (Sobic.008G177900 and Sobic.008G178600). Therefore, these two are the only candidate genes from the *ARG5* locus in P9830. Moreover, the sorghum line Rio that is resistant to the same strain of *Cs* also carried only two of the NBS-LRR genes (SbRio.08G196000 and SbRio.08G196500) that correspond to the two in the P9830 genome. However, it is not clear if the resistance in Rio is conferred by the same locus. There are sequence variations in the two candidate genes between P9830 and Rio. Based on both the alignment to the reference genome of the de novo assembled sequence

Table 4 Fine mapping of *ARG5* locus. The green cells represent concordance and blue cells indicate discordance between phenotype and genotype data

No	Lines	Reaction	ARG4-5945	ARG4-5998	ARG4-6059	ARG5-6090	ARG5-6117	ARG4-6131	ARG5-6200	ARG5-6230
1	8-11	S	2[a]	2	2	2	2	2	2	1
2	9-15	S	3	3	3	2	2	2	2	2
3	10-2	S	2	2	2	2	2	2	2	1
4	10-19	S	1	1	1	1	2	2	2	2
5	10-31	S	1	1	2	2	2	2	2	2
6	12-2	S	1	1	2	2	2	2	2	2
7	12-7	S	1	2	2	2	2	2	2	2
8	12-35	S	2	2	2	2	2	2	1	1
9	14-26	S	2	2	2	2	2	1	1	1
10	14-11	S	1	1	2	2	2	2	2	2
11	14-13	S	2	2	2	2	2	3	3	3
12	14-34	S	1	1	1	1	2	2	2	2
13	16-4	S	1	1	2	2	2	2	2	2
14	16-8	S	3	3	3	3	2	2	2	2
15	10-6	R	2	1	1	1	1	1	1	1
16	10-8	R	2	2	1	1	1	1	1	1
17	10-37	R	2	1	1	1	1	1	2	2
18	12-41	R	1	1	1	1	1	1	1	2
19	12-43	R	1	1	1	1	1	2	2	2
20	14-29	R	2	1	1	1	1	1	1	1
21	14-22	R	1	1	1	1	1	2	2	2
22	14-41	R	2	2	2	2	3	3	3	3
23	17-1	R	2	2	2	1	1	1	1	1
24	17-7	R	2	2	2	1	1	1	1	1
No. of recombinants			16	12	7	4	0	4	6	9

[a]1 = P9830 allele, 2 = TAM428 allele, 3 = heterozygous.

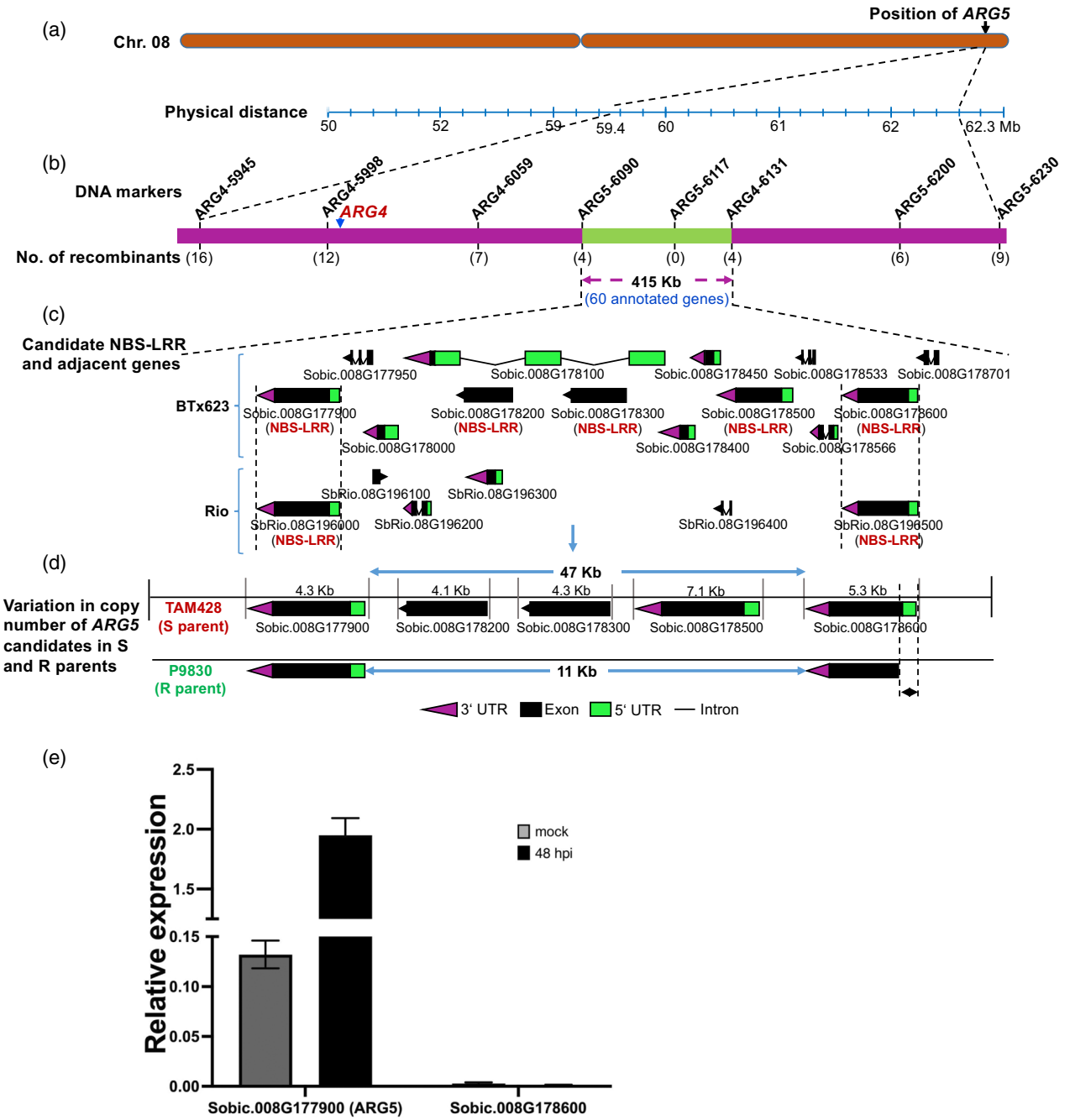

Figure 7. Mapping of the *ARG5* locus on chromosome 8 showing physical linkage to *ARG4*.
(a) Physical position of *ARG5* locus on chromosome 8.
(b) Relative position of DNA markers used to narrow down *ARG5* locus and the number of recombinants identified from a set of recombinant inbred lines generated from a cross between P9830 and TAM428.
(c) Cluster of candidate NBS-LRR genes located within target region and copy number variation of the NBS-LRR genes between BTx623 and Rio reference genomes.
(d) Variation in copy number of *ARG5* candidates in the susceptible (TAM428) and resistant (P9830) parents.
(e) Expression of candidate *ARG5* genes in the resistant parent (P9830) in response to fungal inoculation.

and IGV view of the two candidates from P9830, the second candidate Sobic.008G178600 has a deletion of about 150 bp in its upstream region that includes part of its exon covering the start of the gene and promotor region. Hence, it was apparent that Sobic.008G177900 is the primary candidate *ARG5* gene.

Sequencing and expression of candidate NBS-LRR genes in *ARG5* locus

To further confirm that Sobic.008G177900 is the only candidate associated with resistance to Csgl1 and Csgrg in P9830, gene expression analysis was conducted for both Sobic.008G177900 and Sobic.008G178600. The resistant parent carries a deletion in the Sobic.008G178600 regulatory region and is consistently not expressed in P9830 (Figure 7e) while Sobic.008G177900 is expressed with further induction after inoculation with *Cs*. Thus, the data strongly support that Sobic.008G177900 that encodes NBS-LRR protein, hereafter referred as *ANTHRACNOSE RESISTANCE GENE 5* (*ARG5*), is the most likely gene conferring resistance to Csgl1 and Csgrg strain in P9830.

Polymorphisms, protein domain structure, and sequence comparisons

Similar to *ARG4*, *ARG5* (Sobic.008G177900) has a single exon that encodes a protein of 1421 amino acids (Figure S4). ARG5 has coiled coil (CC), nucleotide-binding-site (NBS-ARC), and leucine-rich-repeat (LRR) domains characteristics of NBS-LRR proteins (Figure 8). The 3D protein structure of ARG5 from P9830 and TAM428 reveals a substantial variation along all the motifs (Figure S5). Based on Phyre^2 analyses, ARG5 from P9830 has several structural differences with ARG5 from TAM428. The structure alteration occurred at ARC and LRR regions, which may impact the ATP binding or ATP hydrolysis (Takken & Tameling, 2009). Genomic DNA sequencing using WideSeq technique from P9830, TAM428, SAP135 and two other variants (BTx631 and SC23) indicated that P9830 carries a unique variant of *ARG5* while TAM428, SAP135, and BTX631 all of which carry recessive alleles of ARG5 encode ARG5 with high sequence similarity (Figure S6).

Protein to protein BLAST search of the NCBI database using the ARG5 amino acid sequences from P9830 (Table S6) revealed the *ARG5* duplicate, Sobic.008G178300 gene, described as "putative disease resistance protein At3g14460 isoform X1" with 91% identity and 93% similarity. Therefore, the *Arabidopsis* At3g14460 was found to be the homolog of both *ARG4* and *ARG5*. A hypothetical protein from weeping lovegrass, *Eragrostis curvula*, was identified as the next most similar (65% identity and 76% similarity).

Phylogenetic analysis between ARG4, ARG5 and functionally characterized NLR proteins revealed that the two proteins are closely clustered but also are in the same clade as many resistance genes such as *Lr1*, *PmR1*, and *Pm60* that confer leaf rust and powdery mildew resistance in wheat (Cloutier et al., 2007; Zou et al., 2018) (Figure S7). *ARG4* and *ARG5* do not cluster with the recently described resistance genes *ARG1* and *ARG2*. *ARG4* and *ARG5* share close phylogenetic relationship that may suggest these two genes resulted from a duplication event consistent with their largely overlapping resistance functions. Moreover, genetic relationship of *ARG4* and *ARG5* genes from a set of resistant and susceptible lines were estimated using in-house generated sequences and from sequences available in public databases (Figure 9). *ARG4* from the resistant line SAP135 forms distinct sub-cluster and it is distantly related with that of SC35C and B35. SC35C is one of the lines that carry a resistant allele of *ARG1* gene (Lee et al., 2022). *ARG5* from P9830 clustered with that of SC10314E.

Haplotypes of *ARG4* and *ARG5* loci and resistance profiles

The resistant parents (SAP135 and P9830) carry dominant resistance alleles at either *ARG4* or *ARG5* loci while the susceptible parent (TAM428) contain recessive allele at both loci (Figure 10a). SAP135 carries *ARG4* but lacks functional *ARG5*. Conversely, P9830 carries *ARG5* and lacks *ARG4*. The fact that SAP135 and P9830 display resistance to both strains despite having only one of the resistance alleles (Figure 10b) suggests either one of the resistance alleles is sufficient to confer resistance to both strains. Furthermore, the disease responses of RILs from the SAP135 and P9830 populations to various strains of *C. sublineola* indicated that lines resistant to either of the Csgrg or Csgl1 strains also exhibit resistance to the other strain. Lines carrying either of the two resistance genes confer resistance to both Csgrg and Csgl1. This is consistent with the reaction of the two resistant parents, SAP135 and P9830, which demonstrate resistance to the two strains. However, the parental lines diverge for their resistance to Cs27 suggesting the possibility that the *ARG4* and *ARG5* may differ in their specificity if other strains are tested.

The sorghum NLR genes

A total of 346 NLR genes were identified in sorghum unevenly distributed across the 10 chromosomes (Mace et al., 2014). Based on *Arabidopsis* and rice functional annotations for the sorghum homologs, we identified a total of 376 NLR domain encoding genes in the BTx623 reference v3.1.1, of which 117 (31%) are located on chromosome 5 followed by 55 (15%) and 50 (13%) on chromosomes 2 and 8, respectively (Figure S8). Many of these are annotated as NBS-LRR indicating they contain both NBS and LRR domains but all of them have at least the NB-ARC domain. Nearly a third of the disease resistance genes are located on chromosome 5 which explains why disease resistance QTLs are frequently detected on this chromosome (Burrell et al., 2015; Cuevas et al., 2014, 2018, 2019; Cuevas & Prom, 2020; Patil et al., 2017; Ramasamy et al., 2008). Chromosome 8, where the two *ARG4* and *ARG5* genes were identified also contains the third highest number of disease resistance genes after chromosome 2.

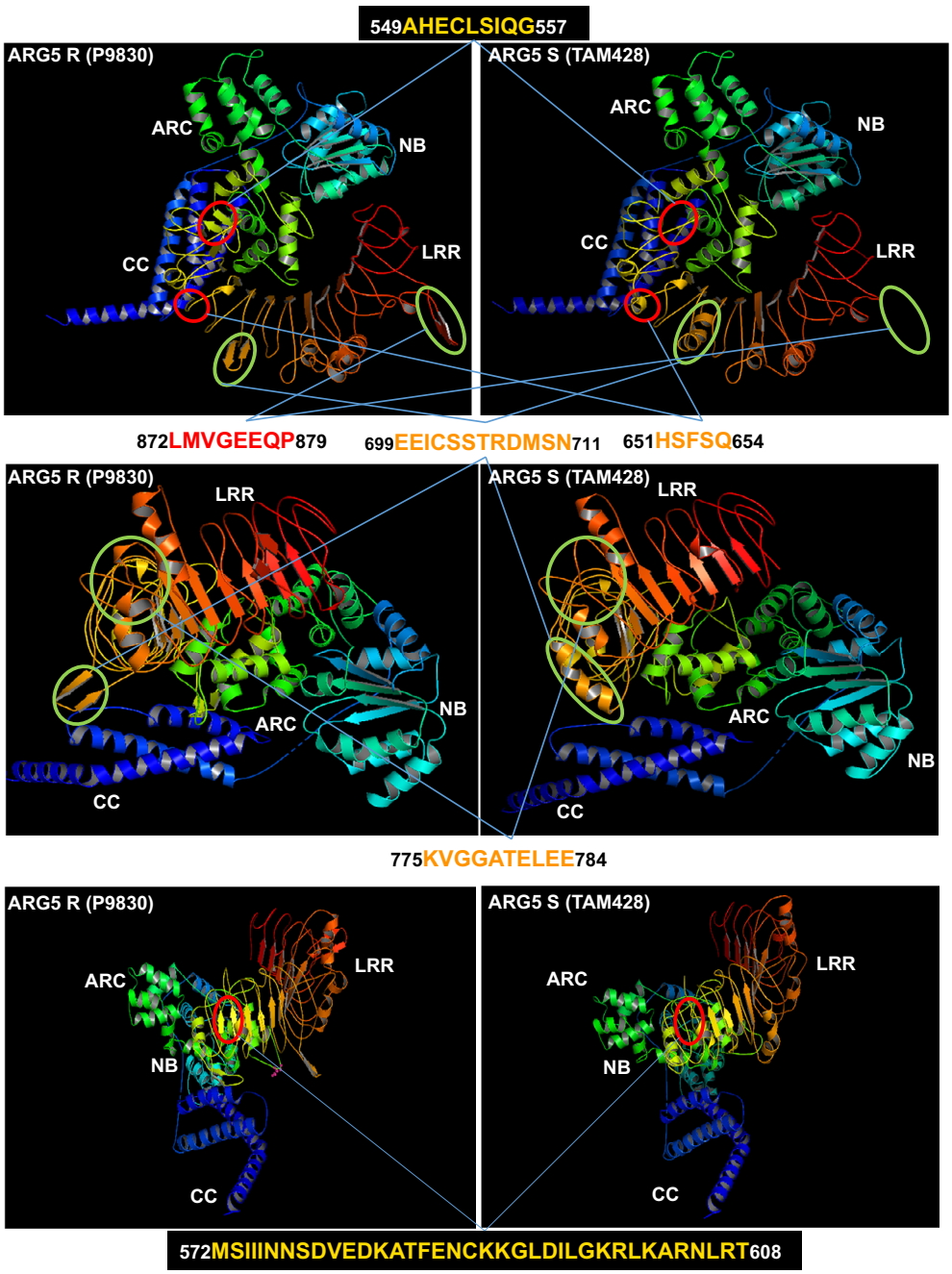

Figure 8. Prediction of tertiary structure of ARG5 protein from the resistant and susceptible ARG5 alleles.
The tertiary structure of ARG5 predicted based on the template c6j5tC (RPP13-like protein 4). The structure of ARG5 from P9830 and TAM428 was predicted by Phyre[2] (Kelley et al., 2015). The specific regions of ARG5 were labeled by red or green circles to represent the differences in structures and the color of amino acid corresponds to the different regions of NB-LRR.

DISCUSSION

Sorghum's immune response to the hemibiotrophic anthracnose fungus *C. sublineola* and their genetic control are poorly understood. Despite the identification of three anthracnose resistance loci which are all located in the distal ends of chromosome 5 (Burrell et al., 2015; Cuevas et al., 2014; Patil et al., 2017; Ramasamy et al., 2008) and chromosome 9 (Felderhoff et al., 2016; Patil et al., 2017) the genes and underlying mechanisms of resistance associated with such loci are not known. Studies using multi-parent mapping populations (Cuevas et al., 2018, 2019, Cuevas & Prom, 2020) redetected the same loci, which

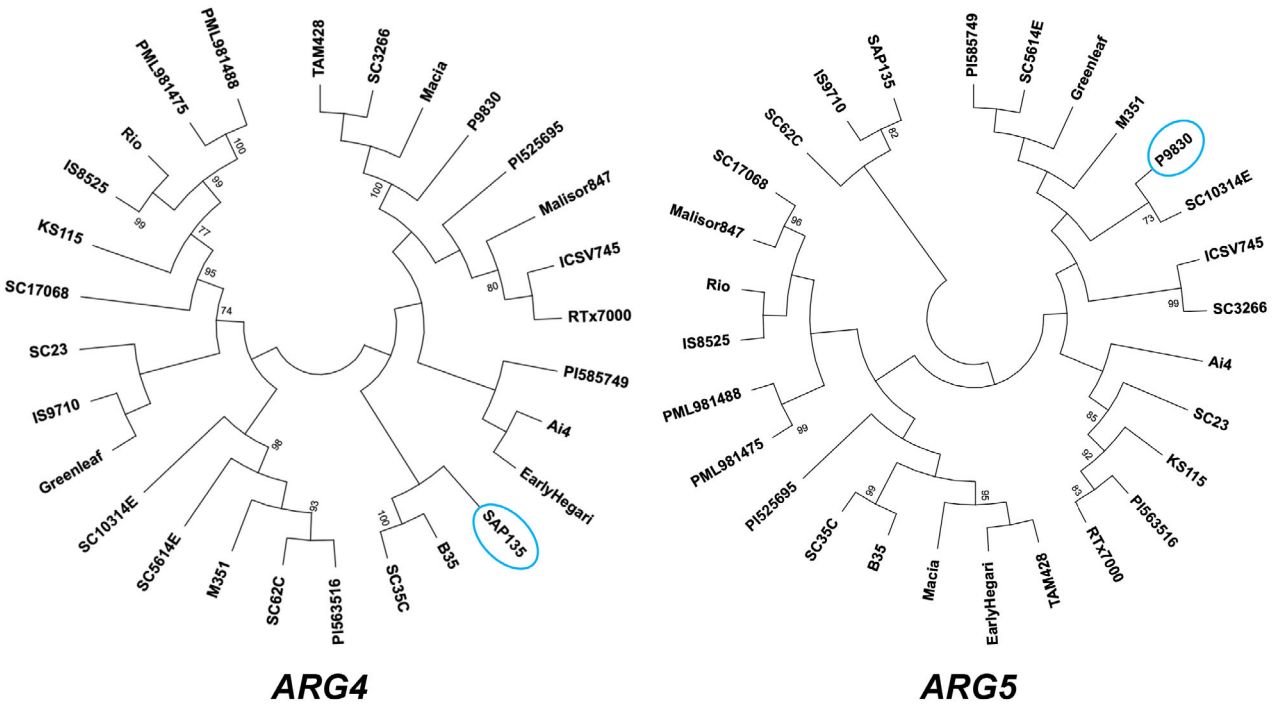

Figure 9. Phylogenetic analysis of *ARG4* and *ARG5* genes from diverse sorghum lines.

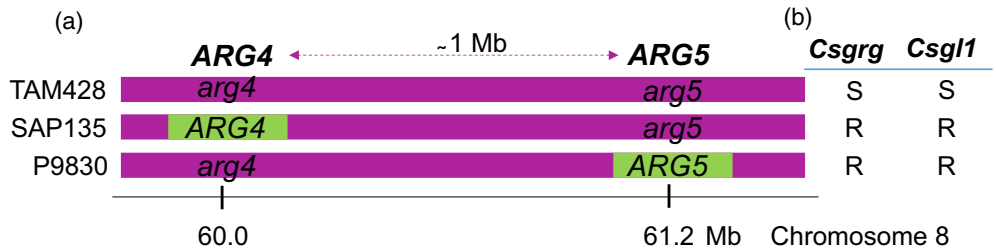

Figure 10. Illustration of *ARG4* and *ARG5* haplotypes of parental lines and the corresponding reaction to *C. sublineola* strains.
(a) *ARG4* and *ARG5* haplotypes of parental lines. Parental lines with upper case letters "*ARG4*" or "*ARG5*" in green background indicates those carrying resistance gene while that with a lower case letter and purple background indicates carrying a susceptible allele.
(b) Disease reaction of parental lines to the *C. sublineola* strains Csgrg and Csgl1.

were initially detected by bi-parental mapping approaches. Hence, the discovery of new and specific anthracnose resistance genes has been slow. Recently, we described a major anthracnose resistance locus in sorghum composed of the nucleotide-binding leucine-rich repeat receptor gene *ANTHRACNOSE RESISTANCE GENE1* (*ARG1*) that is embedded in a natural antisense transcript (Lee et al., 2022). *ARG1* confers broad-spectrum fungal resistance in sorghum. The second gene *ANTHRACNOSE RESISTANCE GENE 2* (*ARG2*) encoding another NLR protein that confers race-specific resistance to *C. sublineola* strains was also identified and characterized in our lab (Mewa et al., 2023). These data, collectively, provided new insights into the genetic control of plant immunity in sorghum. In this study, we used a combination of genomic and genetic approaches and identified two tightly linked resistant genes (*ARG4* and *ARG5*) that encode the NBS-LRR genes (*ARG4*, Sobic.008G166400; *ARG5*, Sobic.008G177900). *ARG4* and *ARG5* genes account for resistance to the Csgl1 and Csgrg strain in SAP135 and P9830, respectively. These genes were identified among clusters of tightly linked and sequence-related paralogous NLR genes. The *ARG5* locus co-localizes with a previously reported leaf rust-resistant locus in sorghum, *Rust locus 4* (Upadhyaya et al., 2013). Moreover, the clusters of disease resistance genes from both loci share significant homology to the wheat rust resistance protein *Lr1* (Upadhyaya et al., 2013). In addition, ARG4 and ARG5 proteins share homology to an *Arabidopsis thaliana* leucine-rich repeat (LRR) protein (At3g14460; AtLRRAC1) that also contain the

adenylyl cyclase core motif and is able to catalyze ATP to cAMP in vitro (Bianchet et al., 2019). The Arabidopsis AtLRRAC mutants are susceptible to fungal and bacterial pathogens. *ARG4* and *ARG5* could provide resistance against anthracnose and provide opportunity to understand mechanisms of anthracnose resistance in sorghum. The following section provides further discussion about the origin of *ARG4*- and *ARG5*-resistant alleles, approaches employed to identify the two genes, and possible mechanisms of anthracnose resistance conferred by *ARG4* and *ARG5*.

The sorghum line SAP135 that was used as source of the *ARG4* gene is also identified by multiple accession numbers, including PI576385, in the GRIN database, IS17209C, SC1070, and NSL365695. This line showed broad-spectrum resistance to anthracnose strains based on our disease assays and data available in the GRIN database. It was also reported as resistant to anthracnose along with other anthracnose-resistant accessions in a study conducted on the sorghum association panel (SAP) (Cuevas et al., 2018). SAP135 is originated from Nigeria. Additional resistant accessions from Nigeria (PI534079, PI533871, PI534071) and Chad (PI534037) were found phylogenetically related to SAP135 (Cuevas et al., 2018, 2019). Therefore, the anthracnose-resistant allele carried by SAP135 has most probably originated from West African sorghum germplasm while the previously detected resistant alleles on chromosomes 5 and 9 have origins mostly in East Africa. On the other hand, P9830, the line that carries the resistant allele of *ARG5* gene was obtained from Gebisa Ejeta's lab sorghum collection at Purdue University but its origin is not known.

SAP135 is a line that is agronomically poor with reduced vigor, short plant height of about 100 cm under field conditions at West Lafayette (Indiana), and limited seed production potential which could be due to the accumulation of several resistance genes as it is resistant to all the strains tested. Although some of the recombinant inbred lines (RILs) showed phenotypes similar to that of SAP135, there was no correlation between resistance and other traits as would be expected if there was a pleiotropic effect of *ARG4*. P9830 is an elite line with a plant height of about 175 cm, vigorous stature, and good grain yield potential. Similar to that of SAP135, we did not observe any pleiotropic effect of the *ARG5* gene in the RILs of P9830. However, to determine the pleiotropic effect of the two resistant genes, Near Isogenic Lines (NILs) are required. Thus, it is worth generating NILs that only differ at the *ARG4* or *ARG5* genes and validate their impact on other traits.

Although SAP135 is among the resistant accessions included in previous GWAS studies (Cuevas et al., 2018, 2019), significant peaks associated with resistance alleles from SAP135 and related accessions failed to detect the *ARG4* and *ARG5* loci. It is possible that such accessions carry resistance alleles in the loci detected by GWAS, but the *ARG4* locus was undetected. *ARG4* was mapped using a virulent strain from Georgia, USA while the QTL on chromosome 9 confers resistance against pathotypes also from Georgia and Texas (Patil et al., 2017). These observations raise questions over efficiency of multi-parent-based mappings to detect rare resistant alleles particularly associated with phenotypes which are mostly qualitative in nature. Those loci detected so far by multi-parent approaches such as GWAS are likely to be QTLs with basal resistance as is evident from the small percentage of explained phenotypic variation associated with resistant alleles (Cuevas et al., 2019). Therefore, bi-parental populations are more powerful to detect infrequent novel alleles. Moreover, mapping populations from bi-parental crosses are vital for fine mapping and identification of specific genes which can be impractical with multi-parent approaches.

The *ARG4* (Sobic.008G166400) and *ARG5* (Sobic.008G177900) genes were identified through genomic approaches that involved QTL-seq, Integrative Genomics Viewer (IGV) visualization of genomic data of the candidate region from parental lines, the resistant and susceptible bulks, and other variants. Moreover, narrowing down the candidate regions through traditional genetic mapping using molecular markers, next-generation sequencing and gene expression analysis of candidate genes were conducted to identify the actual genes among candidates of high sequence similarity. Bulk segregant analysis combined with whole genome resequencing (QTL-seq) is a rapid mapping approach (Takagi et al., 2013) which is becoming common to map loci associated with important quantitative and qualitative traits (Aguado et al., 2020; Kurlovs et al., 2019; Lee et al., 2020; Li et al., 2019; Liang et al., 2020; Pujol et al., 2019). Integrative Genomics Viewer (IGV), which is a high performance and simple tool to interactively visualize genomic data (Thorvaldsdóttir et al., 2013) was used to explore the next-generation sequence data from the parental lines and bulks as well as resequenced sorghum lines from public databases. This has enabled visualization of the candidate region and key polymorphisms between the parental lines for the candidate genes and helped to easily identify InDel markers for marker analysis. Once the candidate regions in both loci were narrowed down and candidate genes identified, the candidates from resistant and susceptible parents and additional variants were sequenced using the WideSeq service at Purdue University. WideSeq is an efficient and inexpensive next-generation sequencing (NGS) approach that involves construction of a NGS library from a target region of up to 100 kb and sequencing and subsequent assembly of the reads. Such capacity to sequence wider genomic regions was instrumental to identify the candidates among a tightly linked and sequence-related cluster of NBS-LRR genes which otherwise would have been

difficult to differentiate among the candidates via the conventional sequencing approaches such as Sanger and even illumina sequencing where the short sequence reads may align to the different genes. Moreover, the gene expression analysis of candidate genes revealed that some of the duplicates lacked expression, which may be due to alterations in regulatory regions due to duplication and pseudogenization of these genes. Therefore, application of genomic tools is essential particularly to identify genes at complex loci.

The fact that both *ARG4* and *ARG5* loci are found in clusters of highly similar NBS-LRR genes is consistent with the reported cluster organization of disease resistance genes (Christie et al., 2015; Cuevas et al., 2019; Jupe et al., 2012; Michelmore & Meyers, 1998; Yi & Richards, 2007). However, except for the *ARG4* (Sobic.008G166400) and *ARG5* (Sobic.008G177900) genes, the other duplicates within each cluster were non-functional either because of sequence truncation or lack of regulatory sequences that caused loss of their expression. Truncated or pseudogenes of NBS-LRR genes have been reported in many plant genomes (Ameline-Torregrosa et al., 2008; Wang, Jia, et al., 2020). Moreover, the *ARG5* locus contains variable number of duplicates between the resistant (P9830) and susceptible (TAM428) parents. Only two copies of the *ARG5* candidates were present in the resistant lines P9830 and the Rio reference genomes while five copies were identified in the susceptible lines TAM428 and BTx623.

The 3D protein structure of ARG4 and ARG5 sequences from the corresponding resistant SAP135 and P9830 lines as well as the susceptible TAM428 revealed variation in 3D conformation. The 3D structure of ARG4 protein from SAP135 and TAM428 showed variation mostly in their LRR region while variation in the other domains was not obvious. The predicted 3D structure of ARG5 from P9830 and TAM428 showed a clear variation. Interestingly, the predicted ARG4 and ARG5 protein structures of the resistant alleles from the two resistant parents SAP135 and P9830 indicated a similar pattern toward their c-terminal region which is different in the susceptible TAM428. However, it was not clear whether such c-terminal modification is associated with resistance/susceptibility to the pathogen or other conformational changes are more important.

Plant NLRs (NBS-LRRs) are widely known for their key role in immune responses particularly through detection of race-specific pathogen effectors (DeYoung & Innes, 2006). The detection occurs indirectly from modifications of the host virulence target by virulence/effector proteins or direct interactions of pathogen effector proteins with the NBS-LRR proteins (DeYoung & Innes, 2006). Conformational changes to the NLRs due to their interaction with pathogen or altered host proteins leads to hydrolysis of ATP to ADP by the NBS domain, which in turn activates downstream process leading to resistance. ARG4 and ARG5 NLRs share homology with the *Arabidopsis* NLRs AtLRRAC1 (At3g14460), characterized as adenylyl cyclase (ACs) that catalyzes the formation of the second messenger cAMP from ATP (Bianchet et al., 2019). Mutants of *Arabidopsis* LRRAC1 are compromised in immune responses to fungal pathogens (Bianchet et al., 2019). cAMP is thought to activate the NBS-LRR downstream signaling (Gehring, 2010). Interestingly, cAMP signaling is involved in elicitor-induced phytoalexin accumulation (Zhao et al., 2004), a known antimicrobial molecule involved in sorghum's defense response against major diseases. cAMP is associated with induction of phenylalanine ammonia lyase (PAL) which in turn is involved in the production of phytoalexins and salicylic acid (SA) (Świeżawska et al., 2014; Wang, Feng, et al., 2020). The amino acid sequences of both ARG4 and ARG5 protein contains motifs that are highly similar to the AC motifs of LRRAC1 (At3g14460). Further studies are required to understand whether ARG4 and ARG5 proteins have AC activities, and such activity contributes to anthracnose resistance.

In summary, two distinct natural variants of sorghum with resistance to anthracnose were used to define two independent resistance loci. Subsequently, the specific resistance genes *ARG4* and *ARG5* encoding NLR proteins underlying these resistance loci were identified by deploying QTL-seq and next-generation whole genomic resequencing. *ARG4* and *ARG5* were linked but independent resistance genes identified in the sorghum lines SAP135 and P9830, respectively, with largely overlapping functions. Although both genes are in clusters of highly similar duplicate copies of NLR genes, most of them were truncated or were pseudogenes. The identification of *ARG4* and *ARG5* genes is supported by multiple and independent recessive alleles for both genes that causes the same disease susceptibility phenotypes. The resistant parent carrying *ARG5*, P9830, has only two duplicates of the NLR gene, of which only one is functional, whereas the susceptible parent TAM428, carried five non-functional copies. ARG4 and ARG5 proteins encoded by the recessive alleles from the susceptible parent TAM428 shows altered 3D protein structure compared with the proteins encoded by the resistant alleles from SAP135 and P9830 consistent with genetic data. The susceptibility of TAM428 might be associated with failure to detect the fungal effector proteins from the Csgl1 and Csgrg strains to which *ARG4* and *ARG5* provide resistance. Each of the line carries resistant allele only in one of the loci while having the recessive allele at the second locus. The function of *ARG4* and *ARG5* for resistance to other strains of *Cs* or other pathogens is unclear but the disease response profiles also indicate that there are possible differences in the spectrum of resistance provided by *ARG4* and *ARG5*. *ARG4* and *ARG5* share homology to rust resistance genes and the NLRs from *Arabidopsis* characterized as an adenylyl cyclase. Finally, we

discovered NLR genes underlying two major anthracnose resistance loci on sorghum chromosome 8 that confer strong resistance to strains of the fungus. The identification of *ARG4* and *ARG5* provides new opportunities for resistance breeding and molecular and genetic studies to understand mechanism of resistance to a widespread and economically damaging fungal pathogen with implications to other pathosystems.

MATERIALS AND METHODS

Plant materials

The sorghum genotype SAP135 is an accession originally collected from Nigeria, a breeding material maintained by Plant Genetic Resources Conservation Unit, Griffin, GA and identified by the accession number PI576385. Other names representing SAP135 include IS17209C, SC1070 and NSL365695 (pre-conversion). The line is described as being resistant to both anthracnose and rust in the Germplasm Resources Information Network (GRIN) data base (https://www.ars-grin.gov/). Disease assays under controlled conditions in the greenhouse identified SAP135 to be resistant to five anthracnose strains available in the lab (Csgrg, Csgl1, Csgl2, Cs27, and Cs29). To map anthracnose resistance loci from SAP135, a mapping population was created by crossing SAP135 to a susceptible inbred, TAM428. The resulting F1 was selfed to generate F2 populations. Based on disease reaction at F2, 71 resistant and 68 susceptible individuals were selected and advanced for further evaluation at F3 stage to identify true-to-type families. A total of 12 plants from each F3 family were planted and phenotyped to identify families that are true-to-type resistant or susceptible. This has identified 30 resistant and 50 susceptible true-to-type F3 families which were subsequently used for QTL-seq analysis and genetic mapping. Additionally, a new F2 population of 203 individuals were phenotyped and genotyped at anthracnose-resistant locus identified in SAP135. These F2s were then advanced to F6 RILs for future mapping studies using different *C. sublineola* strains because the resistant parent SAP135 has broad-spectrum resistance.

Simultaneously, a genomic region was identified in P9830 resistant genotype in an independent project. Likewise, the resistant P9830 was crossed to TAM428 to develop a mapping population. The segregating population was advanced to F6 through single seed descent to generate a recombinant inbred lines (RIL) population. Course mapping was conducted through QTL-seq analyses of pools of 48 resistant and 48 susceptible RILs. Subsequently, fine mapping and identification of candidate genes associated with anthracnose resistance in P9830 was accomplished using 80 F6 RILs.

Fungal culture and anthracnose disease assay

Fungal strains of *C. sublineola* were cultured on potato dextrose agar (BD Difco, Sparks, MD, USA) and conidia harvested from 15-day-old cultures and used for disease assays as described previously (Lee et al., 2022; Mewa et al., 2023). The fungal cultures were kept under a continuous florescence light for 2 weeks, after which spores were collected and used to inoculate 3 weeks old plants. The *C. sublineola* strain Csgrg was used to map resistance in SAP135 x TAM428 population. Plants were sprayed with a spore suspension at a concentration of 1×10^6 spores/ml. Inoculated plants were kept for 48 h in a humidity chamber set at a relative humidity of 70% and then transferred to greenhouse with an overhead misting system. Disease reaction of inoculated plants was scored as resistant or susceptible 4 to 6 days after the plants were transferred to the greenhouse condition. The *C. sublineola* strain Csgl1 was used to map resistance locus in P9830. Therefore, the RILs used to narrow down this locus were inoculated with Csgl1 while other disease assay conditions remained similar to the conditions for SAP135. To study the spectrum of resistance in the parental lines and identify additional alleles, three parental lines and four other sorghum variants (RTx430, Ai4, BTx631, and SC23) were evaluated against a total of five strains including the two used for mapping (Csgrg, Csgl1, Csgl2, Cs27, and Cs29).

Whole genome resequencing and WideSeq of candidate genes

For whole genome resequencing and WideSeq of candidate genes, high quality genomic DNA was isolated from week-old seedlings using DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturers protocol. This includes DNA for the 30 resistant and 50 susceptible F3 families of SAP135 x TAM428, 48 resistant and 48 susceptible RILs of P9830 x TAM428 population, the parental lines, and other variants. For the 80 F6 RILs of P9830 x TAM428 and new F2 population of SAP135 x TAM428, DNA was isolated using a simple and fast high-throughput DNA extraction method developed for PCR (Xin et al., 2003). Equal amount of DNA from each of the resistant and susceptible F3 families of SAP135 x TAM428 and RILs of P9830 x TAM428 were taken and pooled into the corresponding resistant and susceptible groups and used for sequencing along with the parental lines. Whole genome resequencing of the pools, SAP135 and P9830 was conducted at Purdue University Genomics Core Facility, West Lafayette (Indiana) on Illumina HiSeq 2500, while sequence for TAM428 was already available from a previous in-house sequencing effort (Lee et al., 2022). In addition, whole genome resequenced sorghum lines found in publicly available databases were utilized through phenotyping for disease reaction while their genome sequences were visualized for patterns of polymorphism at candidate genes. This helped to identify additional variants that carry alternate alleles at the candidate genes. Candidate genes in the target regions were amplified from the parental lines and other variants by PCR and sequenced using the WideSeq service at Genomics Core at Purdue University. Genomic DNA for PCR amplification was extracted using DNeasy Plant Mini Kit (Qiagen). Primers used to amplify the candidate genes and analysis of expression of selected genes are presented in Table S7.

Bulked segregant analysis and visualization of target regions using IGV

To identify loci associated with resistance in SAP135, the QTL-seq pipeline (Takagi et al., 2013), a method that combines bulk segregant analysis and whole genome resequencing was implemented in the two resistant and susceptible bulks of F3 families of the cross between SAP135 and TAM428. Genome sequences of the two bulks and parental lines of both populations were aligned to the BTx623 reference and binary alignment map (BAM) files were generated for each of the bulks and parental lines. BAM files of target regions were cut out for visualization in integrative genomics viewer (IGV) (Thorvaldsdóttir et al., 2013). For P9830, QTL-seq was conducted with the same approach except RILs were used.

Marker development and fine mapping of target regions

InDel markers spanning the target regions were developed based on visualization of the genomic sequences in BAM format of the

two bulks and the parental lines in each population. Primers flanking the identified InDel markers were designed using the primer-blast tool at NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). PCR template size of 100–200 bp were targeted. The 80 F3 families and 80 F6 RILs corresponding to the two mapping populations, SAP135 x TAM428 and P9830 x TAM428, respectively were genotyped at selected InDel markers to narrow down the target regions. Moreover, new F2 populations of SAP135 x TAM428 were genotyped at selected markers in the QTL region and observed for co-segregation of disease response and DNA markers. PCR products were separated by agarose gel electrophoresis using 3% gel. Primers used in the study are listed in Table S8.

RNA extraction and gene expression analysis of candidate genes

To investigate the expression patterns of candidate genes, we conducted RT-PCR and qPCR analysis. Total RNA was extracted from 100 mg fresh leaf tissues of the parental lines at 0, 48, and 72 h post inoculation (hpi) using TRI Reagent (Molecular Research Center Inc, Cincinnati, OH, USA). cDNA was synthesized from 2 μg of total RNA using the AMV reverse transcriptase (NEB). Sorghum Actin gene was used as an endogenous control.

Prediction of protein structure

The protein structure analysis was carried out by submitting ARG4 and ARG5 amino acid sequences from both the resistant and susceptible lines to the web-based interactive service named Protein Homology/analogY Recognition Engine V 2.0 (Phyre2) (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) based on the normal modeling mode (Kelley et al., 2015). We used PyMOL 2.5.5 software (Schrödinger, 2015) to observe the 3D protein structure presented in a classified publication mode.

ACKNOWLEDGEMENTS

We acknowledge funding from NSF (IOS-1916893). This study was also made possible through funding by the Feed the Future Innovation Lab for Collaborative Research on Sorghum and Millet through grants from American People provided to the United States Agency for International Development (USAID) under cooperative agreement No. AID-OAA-A-13-00047. The contents are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government.

AUTHOR CONTRIBUTIONS

NH conducted the experiments including marker development and fine mapping of *ARG4* and *ARG5* genes, genetic studies, characterization of the loci and identification of the candidate genes, and gene expression. GG developed mapping population, conducted QTL-seq analysis and mapping of *ARG4* locus, bioinformatic and genetic analysis and disease assays. XX conducted BSA-seq analysis and the mapping of *ARG5* locus. C-JL conducted the prediction of tertiary protein structure and conducted comparative analyses of the amino acid sequences of resistant and susceptible parents. AA identified anthracnose-resistant lines, developed mapping populations, and conducted bioinformatic and genetic analysis. SH conducted gene expression analysis. PO conducted disease assays. SL conducted phylogenetic analysis. NH, GG, XX, C-JL, AA, and TM designed the research, analyzed data, and wrote the paper. GE and TM supervised the project.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All relevant data are available within the manuscript and the supporting materials.

SUPPORTING INFORMATION

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

Figure S1. The amino acid sequence of ARG4 from SAP135. The various amino acids in color represent residues that result in structural alteration between SAP135 and ATM428.

Figure S2. Multiple view of 3D protein structure of ARG4 from SAP135 (upper panel) and TAM428 (lower panel). The colors represent the different NLR domains (red and orange = LRR; green = ARC; light blue = NB; deep blue = CC).

Figure S3. Multiple sequence alignment of ARG4 protein from mapping parents and variants.

Figure S4. The amino acid sequences of ARG5 from P9830. The various amino acids in color represent residues that result in structural alteration between P9830 and ATM428.

Figure S5. Multiple view of 3D protein structure of ARG5 from P9830 (upper panel) and TAM428 (lower panel). The colors represent the different NLR domains (red and orange = LRR; green = ARC; light blue = NB; deep blue = CC).

Figure S6. Sequence alignment of ARG5 protein from mapping parents and variants.

Figure S7. Phylogenetic analysis between ARG4, ARG5 and functionally characterized NLR proteins.

Figure S8. Distribution of NBS-LRR genes across the sorghum chromosomes.

Table S1. List of annotated genes located within the mapped *ARG4* region of the reference sorghum genome BTx623 (v3.1.1).

Table S2. Candidate disease resistance genes in *ARG4* locus and their homolog in Arabidopsis and rice.

Table S3. Results of BLASTP analysis of *ARG4* gene showing top hits.

Table S4. List of annotated genes located within the mapped *ARG5* region of the reference sorghum genome BTx623 (v3.1.1).

Table S5. Candidate disease resistance genes in *ARG5* locus and their homolog in Arabidopsis and rice.

Table S6. Results of BLASTP analysis of *ARG5* gene showing top hits.

Table S7. List of primers used for PCR amplification and gene expression study of target genes.

Table S8. InDel markers used to narrow down the QTL region for the *ARG4* and *ARG5* genes.

REFERENCES

Aguado, E., García, A., Iglesias-Moya, J., Romero, J., Wehner, T.C., Gómez-Guillamón, M.L. *et al.* (2020) Mapping a partial *Andromonoecy* locus in *Citrullus lanatus* using BSA-Seq and GWAS approaches. *Frontiers in Plant Science*, **11**, 1243.

Ameline-Torregrosa, C., **Wang, B.B.**, **O'Bleness, M.S.**, **Deshpande, S.**, **Zhu, H.**, **Roe, B.** ***et al.*** (2008) Identification and characterization of nucleotide-binding site-leucine-rich repeat genes in the model plant Medicago truncatula. *Plant Physiology*, **146**, 5–21.

Bianchet, C., **Wong, A.**, **Quaglia, M.**, **Alqurashi, M.**, **Gehring, C.**, **Ntoukakis, V.** ***et al.*** (2019) An Arabidopsis thaliana leucine-rich repeat protein harbors an adenylyl cyclase catalytic center and affects responses to pathogens. *Journal of Plant Physiology*, **232**, 12–22.

Burrell, A.M., **Sharma, A.**, **Patil, N.Y.**, **Collins, S.D.**, **Anderson, W.F.**, **Rooney, W.L.** ***et al.*** (2015) Sequencing of an anthracnose-resistant sorghum genotype and mapping of a major QTL reveal strong candidate genes for anthracnose resistance. *Crop Science*, **55**, 790–799.

Christie, N., **Tobias, P.A.**, **Naidoo, S.** & **Külheim, C.** (2015) The *Eucalyptus grandis* NBS-LRR gene family: physical clustering and expression hotspots. *Frontiers in Plant Science*, **6**, 1238.

Cloutier, S., **McCallum, B.D.**, **Loutre, C.**, **Banks, T.W.**, **Wicker, T.**, **Feuillet, C.** ***et al.*** (2007) Leaf rust resistance gene Lr1, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large psr567 gene family. *Plant Molecular Biology*, **65**, 93–106.

Cuevas, H.E. & **Prom, L.K.** (2020) Evaluation of genetic diversity, agronomic traits, and anthracnose resistance in the NPGS Sudan Sorghum Core collection. *BMC Genomics*, **21**, 88.

Cuevas, H.E., **Prom, L.K.**, **Cooper, E.A.**, **Knoll, J.E.** & **Ni, X.** (2018) Genome-wide association mapping of anthracnose (Colletotrichum sublineolum) resistance in the U.S. Sorghum association panel. *Plant Genome*, **11**.

Cuevas, H.E., **Prom, L.K.** & **Cruet-Burgos, C.M.** (2019) Genome-wide association mapping of anthracnose (*Colletotrichum sublineolum*) resistance in NPGS Ethiopian sorghum germplasm. *G3 (Bethesda)*, **9**, 2879–2885.

Cuevas, H.E., **Prom, L.K.** & **Erpelding, J.E.** (2014) Inheritance and molecular mapping of anthracnose resistance genes present in sorghum line SC112-14. *Molecular Breeding*, **34**, 1943–1953.

DeYoung, B.J. & **Innes, R.W.** (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunology*, **7**, 1243–1249.

Felderhoff, T.J., **McIntyre, L.M.**, **Saballos, A.** & **Vermerris, W.** (2016) Using genotyping by sequencing to map two novel anthracnose resistance loci in Sorghum bicolor. *G3 (Bethesda)*, **6**, 1935–1946.

Gehring, C. (2010) Adenyl cyclases and cAMP in plant signaling - past and present. *Cell Communication and Signaling: CCS*, **8**, 15.

Habte, N. (2021) *Genetic resistance to fungal pathogens in sorghum [Sorghum bicolor (L.) Moench]*. Thesis. West Lafayette: Purdue University Graduate School. Available from: https://doi.org/10.25394/PGS.14484966.v1

Jupe, F., **Pritchard, L.**, **Etherington, G.J.**, **Mackenzie, K.**, **Cock, P.J.**, **Wright, F.** ***et al.*** (2012) Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics*, **13**, 75.

Kelley, L.A., **Mezulis, S.**, **Yates, C.M.**, **Wass, M.N.** & **Sternberg, M.J.** (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, **10**, 845–858.

Kurlovs, A.H., **Snoeck, S.**, **Kosterlitz, O.**, **Van Leeuwen, T.** & **Clark, R.M.** (2019) Trait mapping in diverse arthropods by bulked segregant analysis. *Current Opinion in Insect Science*, **36**, 57–65.

Lee, S., **Fu, F.**, **Liao, C.-J.**, **Mewa, D.B.**, **Adeyanju, A.**, **Ejeta, G.** ***et al.*** (2022) Broad-spectrum fungal resistance in sorghum is conferred through the complex regulation of an immune receptor gene embedded in a natural antisense transcript. *The Plant Cell*, **koab305**, 1641–1665.

Lee, S.B., **Kim, J.E.**, **Kim, H.T.**, **Lee, G.M.**, **Kim, B.S.** & **Lee, J.M.** (2020) Genetic mapping of the c1 locus by GBS-based BSA-seq revealed pseudo-response regulator 2 as a candidate gene controlling pepper fruit color. *Theoretical and Applied Genetics*, **133**, 1897–1910.

Li, R., **Jiang, H.**, **Zhang, Z.**, **Zhao, Y.**, **Xie, J.**, **Wang, Q.** ***et al.*** (2019) Combined linkage mapping and BSA to identify QTL and candidate genes for plant height and the number of nodes on the main stem in soybean. *International Journal of Molecular Sciences*, **21**.

Liang, T., **Chi, W.**, **Huang, L.**, **Qu, M.**, **Zhang, S.**, **Chen, Z.Q.** ***et al.*** (2020) Bulked Segregant analysis coupled with whole-genome sequencing (BSA-Seq) mapping identifies a novel *pi21* haplotype conferring basal resistance to Rice blast disease. *International Journal of Molecular Sciences*, **21**.

Mace, E., **Tai, S.**, **Innes, D.**, **Godwin, I.**, **Hu, W.**, **Campbell, B.** ***et al.*** (2014) The plasticity of NBS resistance genes in sorghum is driven by multiple evolutionary processes. *BMC Plant Biology*, **14**, 253.

Mehta, P.J., **Wiltse, C.C.**, **Rooney, W.L.**, **Collins, S.D.**, **Frederiksen, R.A.**, **Hess, D.E.** ***et al.*** (2005) Classification and inheritance of genetic resistance to anthracnose in sorghum. *Field Crops Research*, **93**, 1–9.

Mewa, D.B., **Lee, S.**, **Liao, C.J.**, **Adeyanju, A.**, **Helm, M.**, **Lisch, D.** ***et al.*** (2023) ANTHRACNOSE RESISTANCE GENE2 confers fungal resistance in sorghum. *The Plant Journal*, **113**, 308–326.

Michelmore, R.W. & **Meyers, B.C.** (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, **8**, 1113–1130.

Patil, N.Y., **Klein, R.R.**, **Williams, C.L.**, **Collins, S.D.**, **Knoll, J.E.**, **Burrell, A.M.** ***et al.*** (2017) Quantitative trait loci associated with anthracnose resistance in sorghum. *Crop Science*, **57**, 877–890.

Prom, L.K., **Perumal, R.**, **Erattaimuthu, S.R.**, **Little, C.R.**, **No, E.G.**, **Erpelding, J.E.** ***et al.*** (2012) Genetic diversity and pathotype determination of Colletotrichum sublineolum isolates causing anthracnose in sorghum. *European Journal of Plant Pathology*, **133**, 671–685.

Pujol, M., **Alexiou, K.G.**, **Fontaine, A.S.**, **Mayor, P.**, **Miras, M.**, **Jahrmann, T.** ***et al.*** (2019) Mapping cucumber vein yellowing virus resistance in cucumber (*Cucumis sativus L.*) by using BSA-seq analysis. *Front Plant Science*, **10**, 1583.

Ramasamy, P., **Menz, M.A.**, **Mehta, P.J.**, **Katilé, S.**, **Gutierrez-Rojas, L.A.**, **Klein, R.R.** ***et al.*** (2008) Molecular mapping of Cg1, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. *Euphytica*, **165**, 597.

Rosenow, D.T., **Dahlberg, J.A.**, **Stephens, J.C.**, **Miller, F.R.**, **Barnes, D.K.**, **Peterson, G.C.** ***et al.*** (1997) Registration of 63 converted sorghum germplasm lines from the sorghum conversion program. *Crop Science*, **37**, 1399–1400. Available from: https://doi.org/10.2135/cropsci1997.0011183X003700040090x

Schrödinger, L.L.C. (2015) The PyMOL molecular graphics system, Version 2.0.

Stephens, J.C., **Miller, F.R.** & **Rosenow, D.T.** (1967) Conversion of alien sorghums to early combine genotypes. *Crop Science*, **7**, 396. Available from: https://doi.org/10.2135/cropsci1967.0011183X000700040036x

Świeżawska, B., **Jaworski, K.**, **Pawełek, A.**, **Grzegorzewska, W.**, **Szewczuk, P.** & **Szmidt-Jaworska, A.** (2014) Molecular cloning and characterization of a novel adenylyl cyclase gene, HpAC1, involved in stress signaling in Hippeastrum x hybridum. *Plant Physiology and Biochemistry*, **80**, 41–52.

Takagi, H., **Abe, A.**, **Yoshida, K.**, **Kosugi, S.**, **Natsume, S.**, **Mitsuoka, C.** ***et al.*** (2013) QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *The Plant Journal*, **74**, 174–183.

Takken, F.L.W. & **Tameling, W.I.L.** (2009) To nibble at plant resistance proteins. *Science*, **324**, 744–746.

Thorvaldsdóttir, H., **Robinson, J.T.** & **Mesirov, J.P.** (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, **14**, 178–192.

Upadhyaya, H.D., **Wang, Y.H.**, **Sharma, R.** & **Sharma, S.** (2013) SNP markers linked to leaf rust and grain mold resistance in sorghum. *Molecular Breeding*, **32**, 451–462.

Wang, J., **Wang, J.**, **Hu, M.**, **Wu, S.**, **Qi, J.**, **Wang, G.** ***et al.*** (2019) Ligand-triggered allosteric ADP release primes a plant NLR complex. *Science*, **364**.

Wang, T., **Jia, Z.H.**, **Zhang, J.Y.**, **Liu, M.**, **Guo, Z.R.** & **Wang, G.** (2020) Identification and analysis of NBS-LRR genes in Actinidia chinensis genome. *Plants (Basel)*, **9**.

Wang, W., **Feng, B.**, **Zhou, J.M.** & **Tang, D.** (2020) Plant immune signaling: advancing on two frontiers. *Journal of Integrative Plant Biology*, **62**, 2–24.

Xin, Z., **Velten, J.P.**, **Oliver, M.J.** & **Burke, J.J.** (2003) High-throughput DNA extraction method suitable for PCR. *BioTechniques*, **34**, 820–826.

Yi, H. & **Richards, E.J.** (2007) A cluster of disease resistance genes in Arabidopsis is coordinately regulated by transcriptional activation and RNA silencing. *Plant Cell*, **19**, 2929–2939.

Zhao, J., **Guo, Y.**, **Fujita, K.** & **Sakai, K.** (2004) Involvement of cAMP signaling in elicitor-induced phytoalexin accumulation in Cupressus lusitanica cell cultures. *New Phytologist*, **161**, 723–733.

Zou, S., **Wang, H.**, **Li, Y.**, **Kong, Z.** & **Tang, D.** (2018) The NB-LRR gene *Pm60* confers powdery mildew resistance in wheat. *New Phytologist*, **218**, 298–309.