Research Article

Received: 31 March 2017 Revised: 28 June 2017 Accepted article published: 7 July 2017 Published online in Wiley Online Library: 11 September 2017

(wileyonlinelibrary.com) DOI 10.1002/ps.4659

Survey of the genomic landscape surrounding the 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*) gene in glyphosate-resistant *Amaranthus palmeri* from geographically distant populations in the USA

William T Molin,[a*] Alice A Wright,[a] Mark J VanGessel,[b] William B McCloskey,[c] Mithila Jugulam[d] and Robert E Hoagland[a]

Abstract

BACKGROUND: Glyphosate resistance in *Amaranthus palmeri*, one of the most prevalent herbicide-resistant weeds in the USA, is attributable to amplification and increased expression of the gene encoding the target site of glyphosate, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). The *EPSPS* gene and the surrounding 287 kilobases (kb) of amplified sequence are unique to glyphosate-resistant plants and termed the *EPSPS* cassette. It has only been sequenced in one *A. palmeri* population from Mississippi. This research compares *EPSPS* cassettes in seven resistant and five sensitive populations from geographically distant locations within the USA, including Mississippi, Arizona, Kansas, Maryland, Delaware and Georgia.

RESULTS: Polymerase chain reaction (PCR) products from 40 primer pairs specific to the cassette were similar in size and sequence in resistant populations. Several primer pairs failed to generate PCR products in sensitive populations. Regions of the cassette sequenced in the resistant populations were found to be nearly identical to those from Mississippi. Gene expression analysis showed that both *EPSPS* and another gene in the cassette, a reverse transcriptase, were elevated in all resistant populations tested relative to the sensitive populations.

CONCLUSION: *EPSPS* cassettes from distant resistant populations were nearly homologous. Considering the complexity of the cassette, and the degree of similarity among some cassette sequences, the results are consistent with the hypothesis that glyphosate resistance probably evolved once and then rapidly spread across the USA.

Keywords: glyphosate; gene amplification; herbicide resistance; *Amaranthus palmeri*; Palmer amaranth

1 INTRODUCTION

With the introduction of glyphosate-resistant crops in the 1990s, the use of glyphosate to control weeds, including *Amaranthus palmeri* (S. Wats), increased and this resulted in heavy reliance on glyphosate for weed control in resistant crops in a given season. Extensive and excessive use of glyphosate on some weeds provided the selection pressure which led to the evolution of glyphosate resistance. Glyphosate resistance was first reported in *Lolium rigidum*[1] in Australia and later in the USA[2] in *Conyza candanesis*[3] and in *A. palmeri*[4] *Amaranthus palmeri* in particular can cause devastating crop losses if left unchecked[5–7] so its control is a high priority. Between 2006 and 2016, glyphosate-resistant *A. palmeri* spread to 27 states,[8] from Georgia to California and as far north as Michigan and Pennsylvania. Compounding the difficulty in controlling this weed is that glyphosate resistance can occur in the same plant with other resistances, most often to acetolactate synthase inhibitor herbicides.[8] Populations with multiple resistances severely reduce control options for farmers, particularly once the crop has emerged.

Glyphosate resistance in *A. palmeri* is attributable to amplification of the gene 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*), which encodes the enzyme inhibited by glyphosate.[9] An increased copy number of *EPSPS* results in an increased amount of enzyme and results in lack of control by a field dose of glyphosate.[9] Amplification has been reported in other species, including *Lolium multiflorum*,[10] *Kochia scoparia*,[11] *Amaranthus tuberculatus*,[12] *Amaranthus spinosus*,[13] and *Bromus diandrus*.[14] Glyphosate-resistant

* *Correspondence to: WT Molin, USDA-ARS, 141 Experiment Station Road, Stoneville, MS 38776, USA. E-mail: william.molin@ars.usda.gov*

a *Crop Production Systems Research Unit, United States Department of Agriculture, Stoneville, MS, USA*

b *Research and Education Center, University of Delaware, Georgetown, DE, USA*

c *School of Plant Science, University of Arizona, Tucson, AZ, USA*

d *Department of Agronomy, Kansas State University, Manhattan, KS, USA*

A. spinosus has only been reported in one location and is a special case as resistance appears to be a result of hybridization with glyphosate-resistant *A. palmeri*.[13]

The mechanism by which gene amplification occurred in *A. palmeri* is unknown. Fluorescence *in situ* hybridization (FISH) analysis with probes for *EPSPS* revealed that the additional copies of the gene were distributed across the genome.[9] Multiple copies of *EPSPS* randomly dispersed throughout the genome may have been the result of a transposon-mediated mechanism.[9] If a transposon-mediated mechanism is the cause of *EPSPS* gene amplification, it is important to know the genomic organization of the amplified *EPSPS* genetic element, its length (size) and the putative regulatory components associated with its amplification.

EPSPS in *A. palmeri* spans a 10-kilobase (kb) sequence with seven introns and eight exons. To explore the genomic landscape surrounding *EPSPS*, a fosmid library was constructed and clones containing *EPSPS* were isolated.[15] Sequencing and alignment of these clones resulted in a 30-kb sequence containing the entire *EPSPS* gene and a putative *Activator* transposase. The entire 30-kb fragment was amplified but the ends of the amplified unit were not determined.[15] In an effort to further determine the genomic content and ends of the cassette, a bacterial artificial chromosome (BAC) library of the *A. palmeri* genome was constructed. BAC clones containing *EPSPS* were sequenced and assembled, which resulted in a 297-kb reference sequence with *EPSPS* centered between two approximately 140-kb sequences.[16] Gene prediction algorithms applied to the cassette identified *EPSPS* and 71 putative genes (27 genes upstream and 44 genes downstream of the *EPSPS* locus), indicating a complex genetic entity. Whole-genome shotgun sequencing was performed for glyphosate-sensitive and -resistant biotypes from Mississippi and the reads were mapped to the cassette. Sequence reads from the resistant biotype mapped precisely to the cassette and completely covered the entire cassette. Reads from the sensitive biotype did not map to the cassette in a contiguous manner; over 400 gaps in the sequence were found. These results indicated that the *EPSPS* cassette was unique to the resistant biotype and that amplification was not just a duplication of *EPSPS* and the surrounding sequence. These results were confirmed by polymerase chain reaction (PCR) analysis.[16] The whole-genome sequencing also revealed that the ends have yet to be determined.[16]

Four potentially full-length genes sufficiently large enough to encode complete proteins were also found in the cassette in addition to *EPSPS*. These were a reverse transcriptase, an NAC domain-containing gene [NAC originates from the first letters of three genes: No apical meristem (*NAM*), Arabidopsis transcription activation factor (*ATAF*), and Cup-shaped cotyledon (*CUC*) superfamily], a heat-shock cognate 70-kDa protein (HCS70), and two copies of a ricesleeper-like transposase (Molin et al. 2017). Transcriptional products of three of the putative genes were confirmed by quantitative PCR in both *EPSPS*-sensitive and -resistant biotypes.[16] If and how these gene products contribute to resistance and/or amplification remain to be determined.

Although the presence of glyphosate-resistant *A. palmeri* populations has been documented in several states, it has yet to be determined if resistance evolved in one location and then spread to many locations or if it evolved independently in multiple locations. Given that the amplicon does not appear as a whole entity in the sensitive biotype, it may have formed piecemeal and is unlikely to have formed in the same way more than once. If the resistance mechanism evolved independently in multiple locations, then it may be expected that PCR products obtained from primer pairs distributed along the 297-kb cassette would differ in size and sequence among the geographically distant locations. Conversely, if resistance evolved in one location and spread, it would be expected that the cassette content would be very similar, if not identical, among populations. The objectives of this research were to test population similarities and differences based on PCR analysis, sequence comparisons of select PCR products, and gene expression analysis in sensitive and resistant populations from Mississippi, Arizona, Georgia, and Kansas, resistant populations from Delaware and Maryland, a resistant *A. palmeri* × *A. spinosus* hybrid from Mississippi, and a sensitive *A. spinosus* from Mississippi.

2 EXPERIMENTAL METHODS

2.1 Plant sources and growth conditions

Resistant populations of *A. palmeri* were collected from Washington County (MS), Macon County (GA), Riley County (KS), Kent County (DE), Wicomico County (MD), and Maricopa County (AZ). Sensitive populations were from Washington County (MS), Decatur County (GA), Riley County (KS), and Pima County (AZ). *Amaranthus spinosus* × *A. palmeri* hybrids were from Yalobusha County (MS). Seeds were planted at 0.25 cm depth in 6 × 4.5 × 6 cm plastic pots with holes that contained a commercial potting mix (Metro-Mix 360; Sun Gro Horticulture, Bellevue, WA, USA). Pots were subirrigated and maintained in a greenhouse set at a temperature regime of 25/20 °C (day/night) and a 12-h photoperiod under natural sunlight conditions supplemented with high-pressure sodium lights providing 400 μmol m^{-2} s^{-1}. All PCR, copy number, and gene expression assays and sequencing were performed using two representatives from each population.

2.2 Confirmation of resistance

Resistant and sensitive plants were confirmed by copy number assays. Briefly, genomic DNA was extracted from one to two leaves per plant by homogenizing the tissue in a mortar and pestle with a buffer developed by Paterson *et al.*.[17] After incubation at 65°C for 1 h, DNA was extracted using a DNeasy plant mini kit (Qiagen, Valencia, CA, USA). DNA quantity and quality were determined by $A_{260/280}$, and agarose gel electrophoresis. *EPSPS* copy number was measured using primers AW146 and AW147 (see Table 1 for primer sequences) with acetolactate synthase (*ALS*) as a reference gene (AW23 and AW24, from Gaines et al.[9]). Reactions consisted of 10 ng of DNA, 200 μM primers, 1 × Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, Pittsburg, PA, USA) containing buffer, dNTPs, thermostable hot start DNA polymerase and SYBR Green dye (Thermo Fisher Scientific), and H_2O to 50 μL. Cycle conditions were 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s and 60°C for 1 min, and 95°C for 15 s. Three technical replicates were performed per sample. Reactions were performed on an ABI 7500 Real Time PCR system (Thermo Fisher Scientific). Data were analyzed using the standard curve method according to Applied Biosystems User Bulletin No. 2 (http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_040980.pdf).

2.3 PCR analysis of *EPSPS* cassette

Primer pairs were designed to amplify regions of the cassette (see Table 1 for primer sequences and Table 2 for primer pairs). Reactions consisted of 10 ng of DNA, 5 μM primers, 1 × Maxima™ Hot Start Green 2 × PCR Master Mix (Thermo Fisher Scientific) containing buffer, dNTPs, thermostable hot start DNA polymerase and

Table 1. Sequences of primers used

Primer	Sequence	Primer	Sequence
AW23	GCTGCTGAAGGCTACGCT	AW426	CATTCGGGTCATCGGGTCAG
AW24	GCGGGACTGAGTCAAGAAGTG	AW429	GACATATAAGGGTGGGCAAAATCG
AW92	GTTCCAACCTTGTTCAACTTTAAG	AW430	GTTGGGAGCTAGGAAGCATCACAT
AW100	GAGTTCTGAGAGCGTCCA	AW433	GTCCACTATCCGTCAATCCTTC
AW119	GAGATTGATTTGTAGTCCTTACAT	AW440	GTGGACGAATTGGACGTTG
AW128	CTCCAAAGTTCATGTCTTTAACCA	AW498	CACTTGTTGCTTATGCGACCT
AW129	CTGGGACAGATGAAGGTTTC	AW501	CGGTAGTCTGATAAATTACGGT
AW130	CTCTCTGGATCGGTTAGTAGC	AW502	GACGTGGCAAGCTCTGATTG
AW131	GTCATTTCAACATACGGTACAGA	AW506	CTCGATCTGTTTCACCTCGAAATC
AW146	CAACAGTTGAGGAAGGATCTG	AW509	CTTGTCGAGTATACCAGTTCA
AW147	CAGCAAGAGAGAATGCCAT	AW510	CAGCAATCCATGAGTTGTACT
AW156	CTGCTCATATTGAATTGATGTTC	AW511	GAGACTGTGAGTGAGTGAGTGAG
AW166	GTATTGAAGCGGCAGTCT	AW512	GCTATGATTCATGTTGGTAGCT
AW168	CATTGCCATTATGTTTATGTCTGT	AW513	GTTCCAACTCCAGTGAGATCAT
AW169	GTTCTGCACAGGAACAAGATGC	AW516	GTCGAGAGACAGGACCTGA
AW176	CATAGTGCCTGAGAAGGATGATA	AW519	GCTCGAAGAATAACAATGTGCAT
AW206	CGTTTCCTGAAGTTTTTAGTGTCA	AW521	GTTCAGCCCTACGTGTAAT
AW211	GTCCAGTGCGAACTCTCTCAAT	AW522	GTATGAAGAACTATGTGATGGGAC
AW216	GACCTGGGTTGTCTTCATTC	AW523	CTCTCGTTTCACAAATCAGA
AW220	GCTCCCTCATTCTTGGTACTC	AW524	GATTGAGAACAAGAAGACCCA
AW225	GATACAGACCCGAACCGA	AW525	CATTCATCAACTGTAATTTCCCA
AW230	GAGAGTGAAGACAACCCAG	AW527	CTATGTTCAACAAGTAAAGACCA
AW247	GTTCCCTCTACCTCAACCTC	AW533	GACGGAAGACATCTAACAGAGTTA
AW253	CGATATGGAGTAATGCGTCAAC	AW534	CACTCACCATGAGAACTTTCCA
AW259	GTTCGGTCCCTGTATACG	AW539	GTATCGACTCTTCTGATTTCCCTC
AW272	GATGGCGTCAAGCGATTC	AW540	CTTGCACAGCTCCACGAAATAG
AW275	CTAGTTGTTTCACTTGTTTGTGTG	AW541	CGATGATCCAACCGTCCA
AW277	GTCCAATTAAGCACCGCA	AW544	CAGATCATAATCCGAATACGGGTC
AW281	GCAGAAGTTTGTTTCGGGATGAGT	AW546	CTCTTTAGACTCCGCTACCAGC
AW291	CTCAACTCAATGAAGAATCCCGTG	AW548	CAGCCTTCTCAATACATGGAT
AW292	CAAATACTCAAGATGAAGGCAAGA	AW549	GCCACACCATGAAATTGTG
AW293	GTTATAGCAGCAATTCACCAG	AW550	CTACCTTGGCATGAATGTGA
AW297	GTGAATGAAGCATTACATCATTGT	AW551	GCAGCCGCAGTTGGT
AW301	GTCTAGGTGGAAGTGTTGAAGA	AW553	CTGATGTCCTTGCTATGTTTCCT
AW303	CATTGGAATTGGTGTTGTGGT	AW554	CAGCAACAACCAGGGAAACAAAT
AW422	GAAGGAGTAATGATGAGTAAATG	AW556	CTCAACAATTAGCGGAAGGC
AW423	CAAGCGAGCGATTTGAATC	AW557	CTTAGCATACTCCGAAAGAACAGA

SYBR Green dye (Thermo Fisher Scientific), and H_2O to 10 µL. See Table 2 for the cycle conditions that were used for each reaction. Cycle condition 1 (CC1) was 94°C for 4 min, 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 90 s, and 72°C for 5 min. In CC2 the extension time was increased to 3 min. In CC3 the annealing temperature was decreased to 53°C and the extension time was set at 3 min. The presence of PCR products was detected by agarose gel electrophoresis.

2.4 Sequencing selected PCR products

Products amplified by primer pairs AW498xAW501, AW301xAW168, and AW554xAW557 were randomly selected for sequencing providing they met the conditions of being short enough to be fully sequenced by the M13 forward and reverse primers (less than 1500 bp of cassette) and amplifying regions that were distant from one another on the cassette. PCR was performed as described above, but scaled up to 25-µL reactions. Products were gel purified using the UltraClean® GelSpin® DNA extraction kit (Mo-bio, Carlsbad, CA, USA). Purified products were checked by gel electrophoresis. PCR products were cloned into the pGEM®-T vector using the pGEM®-T vector system (Promega, Madison, WI, USA). Reactions consisted of a 2:1 insert to vector ratio (50 ng), 1 × buffer, 3 units of ligase, and H_2O to 10 µL. Reactions were incubated at 4°C overnight. Chemically competent cells were prepared and transformed according to Sambrook *et al.*[18] Transformants were screened for the presence of the insert by PCR as described above. Positive transformants were glycerol stocked by adding 800 µL of an overnight culture to 200 µL of 80% glycerol and storing at −80°C. Positive clones were submitted to the Genomics and Bioinformatics Research Unit in Stoneville, MS for plasmid isolation and sequencing. Sequence data were analyzed using Geneious v7.1.9.[19]

2.5 Gene expression analysis

The RNeasy plant mini kit (Qiagen) was used for the RNA extractions according to the manufacturer's instructions. RNA was extracted from one to two leaves per plant by homogenization

Table 2. Results of PCR analysis of the *EPSPS* amplicon in sensitive and resistant *A. palmeri* populations

Forward primer	Reverse primer	Product size	Start	Stop	Cycle	AZ-R A, B	AZ-S A, B	DE-R A, B	MD-R A, B	GA-R A, B	GA-S A, B	KS-R A, B	KS-S A,B	MS-R	MS-S	HYB-R	SA-S
AW546	AW549	484	10751	11234	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW550	AW553	371	28782	29153	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW498	AW501	1437	46061	47498	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	+	+	-
AW502	AW293	2872	55843	58715	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW502	AW272	2432	55843	58275	CC2	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW502	AW303	3061	55843	58904	CC3	+, +	S, -	+, +	+, +	+, +	S, S	+, +	S, S	+	-	+	-
AW293	AW275	1757	77414	79140	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW292	AW291	490	84474	84963	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW506	AW230	2847	92399	95245	CC2	+, +	S, -	+, +	+, +	+, +	-, -	+, +	S, S	+	-	+	S
AW506	AW509	3753	92399	96152	CC2	+, +	S, -	+, +	+, +	+, +	-, -	+, +	S, S	+	-	+	S
AW426	AW511	1350	97125	98474	CC1	+, +	S, S	+, +	+, +	+, +	S, S	+, +	+, S	+	S	+	S
AW510	AW440	1799	98346	100144	CC1	+, +	M, M	+, +	+, +	+, +	+, +	+, +	+, M	+	+	+	+
AW510	AW211	3289	98346	101635	CC1	+, +	S, -	+, +	+, +	+, +	S, S	+, +	+, +	+	S	+	+
AW176	AW513	2436	100279	102715	CC1	+, +	S, +	+, +	+, +	+, +	-, S	+, +	+, S	+	S	+	S
AW176	AW211	1356	100279	101635	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW512	AW225	812	102604	103416	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW301	AW168	1176	103198	104374	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW247	AW166	671	116960	117630	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW169	AW92	1759	117268	119026	CC1	+, +	S, -	+, +	+, +	+, +	S, -	+, +	-, -	+	S	+	-
AW516	AW519	2532	122420	124951	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW297	AW521	1420	127055	128474	CC1	+, +	-, -	+, +	-, -	+, +	-, -	+, +	-, -	+	-	+	-
AW422	AW423	1291	130470	131761	CC1	+, +	M, -	+, +	M, M	+, +	+, +	+, +	+, -	+	-	+	-
AW422	AW523	1953	130470	132423	CC1	+, +	-, -	+, +	-, -	+, +	-, -	+, +	-, -	+	-	(-)	-
AW522	AW525	2753	132290	135043	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW524	AW129	1022	134906	135928	CC1	+, +	+, -	+, +	+, +	+, +	+, +	+, +	+, +	+	-	+	+
AW128	AW100	2697	135818	138515	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW206	AW220	2956	138395	141351	CC1	+, +	+, +	+, +	S, S	+, +	S, S	+, +	+, S	+	S	+	-
AW119	AW131	2344	141108	143452	CC1	+, +	-, +	+, +	-, +	+, +	-, -	+, +	-, -	+	+	+	-
AW130	AW147	2496	143331	145827	CC1	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+, +	+	+	+	+
AW146	AW527	1089	145714	146803	CC1	+, +	S, +	+, +	S, S	+, +	M, M	+, +	+, S	+	S	+	+
AW156	AW533	1856	182654	184510	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW253	AW533	599	183911	184510	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW253	AW272	2278	183911	186190	CC1	+, +	-, -	+, +	+, +	+, +	-, -	-, -	-, -	+	-	+	-
AW281	AW429	1279	210463	211742	CC3	+, +	+, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	L
AW534	AW277	1682	215995	217677	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW430	AW539	3755	225599	229354	CC2	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW216	AW541	1544	233570	235114	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	S
AW540	AW433	3909	235001	238910	CC2	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW544	AW259	1020	276207	277227	CC1	+, +	-, -	+, +	+, +	+, +	-, -	+, +	-, -	+	-	+	-
AW554	AW557	1055	287979	289016	CC3	+, +	-, -	+, +	+, +	+, +	-, +	+, +	-, -	+	-	+	+

AZ-R, Arizona resistant; AZ-S, Arizona sensitive; DE-R, Delaware resistant; MD-R, Maryland resistant; GA-R, Georgia resistant; GA-S, Georgia sensitive; KS-R, Kansas resistant; KS-S, Kansas sensitive; MS_R, Mississippi resistant; MS-S, Mississippi sensitive; HYB-R, *A. spinosus* × *A. palmeri* hybrid resistant; SA-S, *A. spinosus* sensitive; A, B, determinations 1 and 2, respectively; +, expected band is present; –, expected band is absent; S, a smaller band is present; M, multiple bands are present; L, a larger band is present. 'Start' and 'Stop' indicate the position along the replicon. Cycle, PCR cycle conditions. Product Size, number of nucleotide in the PCR product. The primer pairs are listed in order from the 5' to the 3' end of the cassette.

of leaf tissue in the extraction buffer. After elution, RNA was incubated with DNase for 30 min at room temperature and purified on a column according to kit guidelines. The RNA was converted to cDNA using the high-capacity cDNA reverse transcription kit (Thermo Fisher Scientific). Reactions contained 1 µg of RNA, 4 mM dNTPs, 1 × buffer, 1 × random primers, 2.5 units of reverse transcriptase and H_2O to 20 µL. Control reactions lacking reverse transcriptase were included to check for DNA contamination. Cycle conditions were 25°C for 10 min, 37°C for 2 h, and 85°C for 5 s. Using quantitative polymerase chain reaction (qPCR), expression levels were determined for *EPSPS* (AW146xAW147), *HSC70* (AW550xAW551), a NAC domain-containing protein (AW556xAW557), and a reverse transcriptase (AW548xAW549) with *ALS* as a reference (AW23xAW24). Reactions contained 25 ng of cDNA and were performed as described above for the copy number assay.

3 RESULTS AND DISCUSSION

3.1 PCR analysis of the EPSPS cassette

To address questions of common sequence among *EPSPS* cassettes from different populations, primers were designed across the *EPSPS* cassette to probe for the presence of sections of the *EPSPS* cassette in sensitive and resistant populations from Mississippi, Arizona, Georgia, and Kansas, resistant populations from Delaware and Maryland, a resistant *A. palmeri* × *A. spinosus* hybrid from Mississippi, and a sensitive *A. spinosus* from Mississippi. DNA from two plants from each population were probed with 40 primer pairs (Table 2). Overwhelmingly, gel band patterns from the resistant populations were similar in size, pattern and position relative to the DNA ladder and were different from the sensitive populations. Primers that were intragenic, which included primer pairs AW546xAW549 for the reverse transcriptase, AW550xAW553 for *HCS70*, AW510xAW440, AW510xAW211, AW176XAW513, and AW176xAW211 for the ricesleeper-like transposase, and AW524xAW129, AW128xAW100, AW206xAW220, AW119xAW131, and AW130x147 for *EPSPS*, produced PCR products that were similar between the sensitive and resistant populations. However, the primers that fell outside of the gene regions mostly produced the expected band only in the resistant populations. The resistant populations from Arizona, Delaware, Georgia, Kansas, and Mississippi produced identical results in the PCR analysis (Table 2). The *A. palmeri* × *A. spinosus* hybrid failed to produce a band for AW422xAW523 and produced multiple bands for the primer pair AW146xAW527. The resistant population from Maryland differed slightly from the other resistant populations in that two of the primer pairs produced a smaller band than expected, one of the primer pairs produced multiple bands and two primer pairs failed to produce a band. The sensitive populations failed to produce a band, or produced a band differing in size from the bands of the resistant populations for many of the primer pairs (Table 2).

The similarity between the resistant populations and the absence of the cassette as a single unit in the sensitive populations are also evident in Fig. 1. Four primer pairs located in intergenic regions of the *EPSPS* cassette only produced the correct band in the resistant populations (Fig. 1a–c). In Fig. 1(d), a band was visible for *A. spinosus*, but was shorter in length than that observed in the resistant plants and did not produce as robust a signal as that of the resistant samples. This was a typical pattern for many of the primer pairs (Table 2).

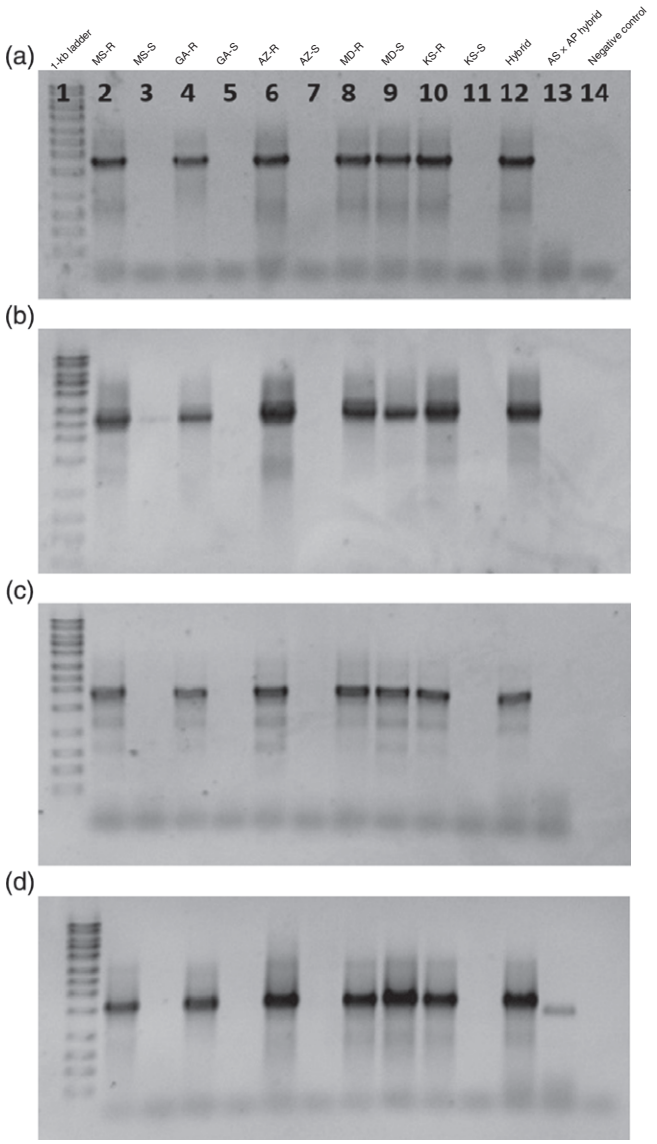

Figure 1. PCR analysis of different R and S populations using primer pairs (a) AW293xAW275, (b) AW516xAW519, (c) AW156xAW533, and (d) AW216xAW541. Lanes include (1) 1-kb ladder (10, 8, 6, 5, 4, 3, 2.5, 2, 1.5, 1, 0.7, 0.5 and 0.3 kb), (2) MS-R, (3) MS-S, (4) GA-R, (5) GA-S, (6) AZ-R, (7) AZ-S, (8) MD-R, (9) DE-R, (10) KS-R, (11) KS-S, (12) *A. spinosus* × *A. palmeri* hybrid (AS × AP), (13) *A. spinosus*, and (14) no template negative control.

3.2 Sequence analysis

To investigate further the similarities between the resistant populations, three PCR products were selected for Sanger sequencing. The sequences were found to be very similar between populations (Fig. 2a–c). In the fragment generated by AW498xAW501 (Fig. 2a), the Arizona population differed at three positions, the Mississippi population was missing a nucleotide at three positions and the Maryland population differed at one position. For the AW301xAW168 fragment (Fig. 2b), the Maryland population differed at one position, Georgia had an additional nucleotide, and the *A. palmeri* × *A. spinosus* hybrid differed at three nucleotide positions. Fragment AW554xAW557 (Fig. 2c) exhibited no sequence differences between populations. The lack of variability in the sequence data and similarity among the resistant populations in the PCR analysis indicate that the *EPSPS* cassette is highly similar among these geographically distant populations.

3.3 Gene expression analysis

Partial sequencing of the *EPSPS* cassette revealed the presence of four potential full-length genes in addition to *EPSPS*.[16] qPCR analysis was performed to determine the expression level of three of these genes. *EPSPS*, as expected, exhibited elevated expression in the resistant plants compared with the sensitive plants (Fig. 3a). An *HSC70* homolog and an NAC domain-containing protein did not exhibit differential expression between the resistant and sensitive populations (Fig. 3b and c). The three plants that exhibited elevated levels of HSC70 were sensitive (MS-SA, MS-SB, and AZ-SD). For the NAC domain-containing protein, some of the resistant plants had elevated levels of expression but this was not consistent among all populations. These homologs encode proteins related to the stress response,[20,21] suggesting that under some conditions, such as in the presence of stressors, these genes have potential to

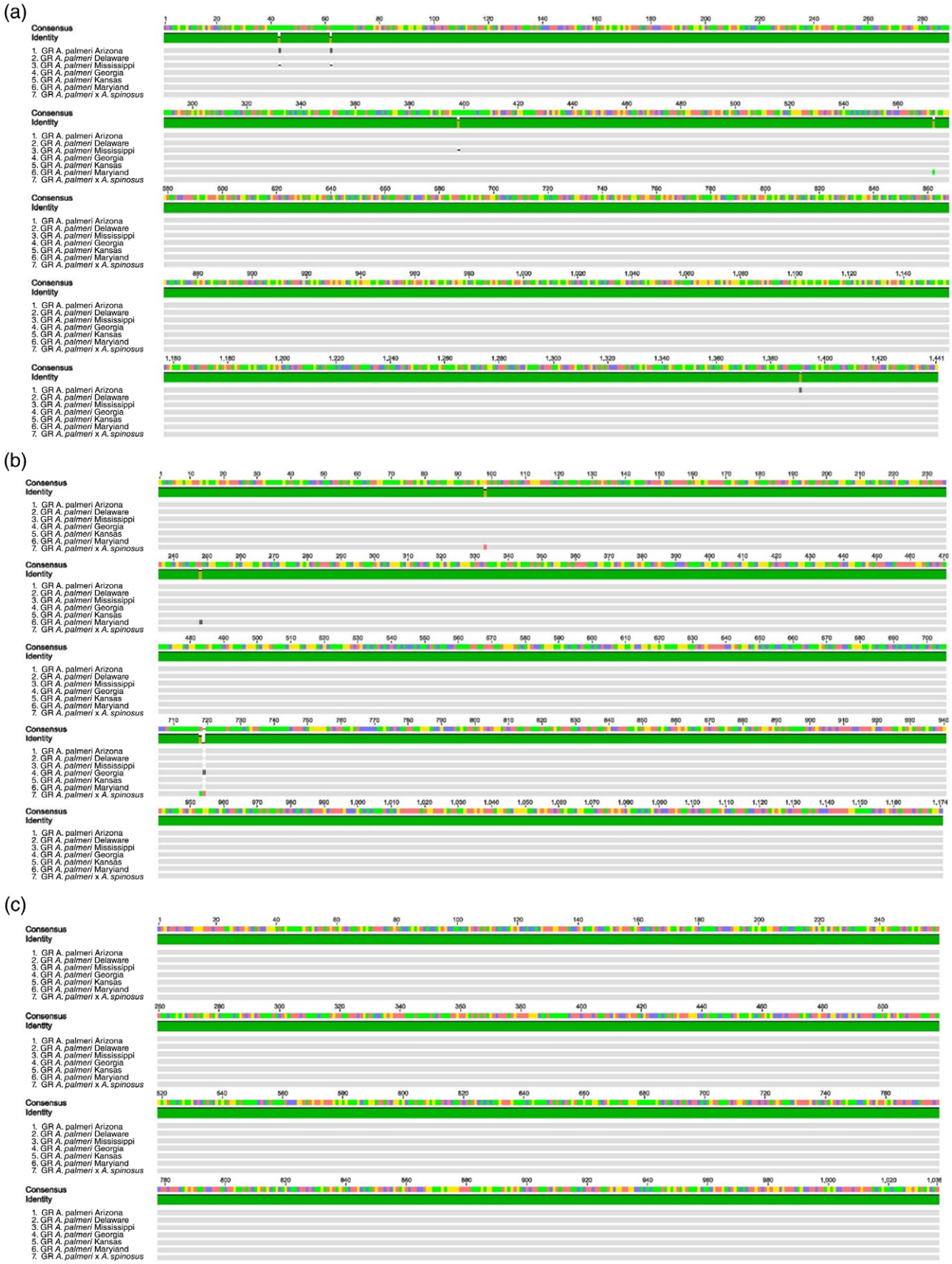

Figure 2. Alignment of sequences for each *A. palmeri* population for primer pairs (a) AW498xAW501, (b) AW301xAW168, and (c) AW554xAW557.

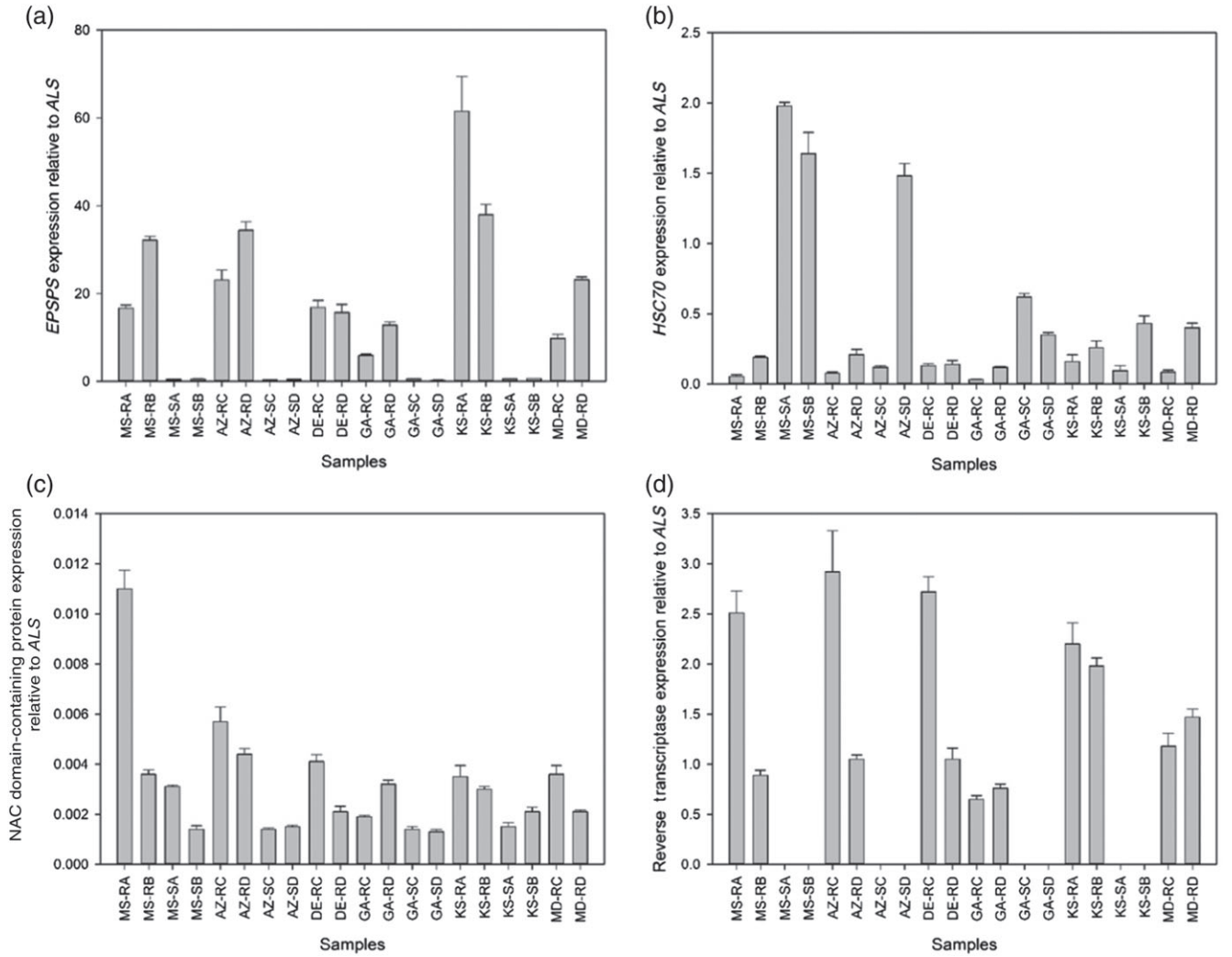

Figure 3. Gene expression analysis of R and S populations for (a) *EPSPS*, (b) *HSC70*, (c) an NAC domain-containing protein, and (d) a reverse transcriptase.

exhibit increased expression in the resistant plants. Interestingly, the reverse transcriptase (Fig. 3d) gene was expressed in the resistant populations but was not detected or was barely detected in the sensitive populations (Table 3).

Glyphosate-resistant *A. palmeri* was the first instance in which amplification of *EPSPS* was determined to be the resistance mechanism.[9] The driving force behind the amplification and how the mechanism evolved remain to be determined. It is known that a large section of genomic DNA, in excess of 297 kb, is amplified and that this amplified DNA does not exist as a single unit within the sensitive plant.[16] The PCR analysis and sequence data presented here suggest that this cassette is not only present in the Mississippi population, but also occurs in the resistant populations from other states. Many of the cassette-specific primers failed to generate a product in the sensitive populations from Mississippi, Arizona, Georgia, and Kansas, but worked well in resistant populations from these states, as well as Delaware and Maryland (Table 2 and Fig. 1). This, and the lack of variability in the sequence data (Fig. 3) suggest that the cassette is very similar among resistant populations. The elevated expression of the reverse transcriptase only in the resistant populations further suggests that this increase in expression is specific to glyphosate-resistant *A. palmeri* and may have a role in the amplification mechanism. Whether the additional amplified putative genes (reverse transcriptase, *HSC70*, and *NAC*) that are also expressed remain contiguous in the different populations is unknown. However, each has been found in all resistant populations and in the *A. palmeri* × *A. spinosus* hybrid.

As glyphosate-resistant *A. palmeri* becomes established in new locations, the gradual evolution of the cassette might be expected, such that with time sequence variations within the cassette from geographically distant locations may become more prevalent. However, at this time the sequence and PCR data indicate that the cassettes are very similar, but additional sequencing would be needed to confirm this. The similarity among cassettes also suggests that they are of recent origin and insufficient time has passed for them to evolve. Considering that there were regions along the cassette that did not produce PCR products in all of the sensitive populations, it seems unlikely that sensitive populations had a precursor which later led to the development of the cassette.

4 CONCLUSION

The origin of glyphosate-resistant *A. palmeri*, whether it was one evolutionary event or many, is an important issue to resolve. The data presented herein support the hypothesis that the glyphosate

Table 3. Means ± standard errors for gene expression relative to expression of *ALS* for Fig. 3

Sample	*EPSPS*	*HSC70*	*NAC*	Reverse transcriptase
MS-RA	16.64 ± 0.75	0.06 ± 0.01	0.0110 ± 0.00075	2.51 ± 0.22
MS-RB	32.15 ± 0.84	0.19 ± 0.01	0.0036 ± 0.00017	0.89 ± 0.05
MS-SA	0.47 ± 0.02	1.98 ± 0.03	0.0031 ± 0.00005	0.000052 ± 0.000008
MS-SB	0.53 ± 0.07	1.64 ± 0.15	0.0014 ± 0.00014	Undetermined
AZ-RC	23.06 ± 2.29	0.08 ± 0.01	0.0057 ± 0.00058	2.92 ± 0.41
AZ-RD	34.44 ± 1.88	0.21 ± 0.04	0.0044 ± 0.00022	1.05 ± 0.04
AZ-SC	0.37 ± 0.01	0.12 ± 0.01	0.0014 ± 0.00005	Undetermined
AZ-SD	0.44 ± 0.01	1.48 ± 0.09	0.0015 ± 0.00006	0.00038 ± 0.00029
DE-RC	16.85 ± 1.56	0.13 ± 0.02	0.0041 ± 0.00028	2.72 ± 0.15
DE-RD	15.69 ± 1.84	0.14 ± 0.03	0.0021 ± 0.00022	1.05 ± 0.11
GA-RC	5.91 ± 0.31	0.03 ± 0.01	0.0019 ± 0.00010	0.65 ± 0.04
GA-RD	12.8 ± 0.68	0.12 ± 0.01	0.0032 ± 0.00016	0.76 ± 0.04
GA-SC	0.59 ± 0.02	0.62 ± 0.03	0.0014 ± 0.00011	0.0011 ± 0.000018
GA-SD	0.25 ± 0.02	0.35 ± 0.02	0.0013 ± 0.00010	0.000067 ± 0.0000026
KS-RA	61.53 ± 7.88	0.16 ± 0.05	0.0035 ± 0.00044	2.20 ± 0.21
KS-RB	37.96 ± 2.34	0.26 ± 0.05	0.0030 ± 0.00010	1.98 ± 0.08
KS-SA	0.59 ± 0.04	0.10 ± 0.04	0.0015 ± 0.00016	0.001 ± 0.0007
KS-SB	0.64 ± 0.05	0.43 ± 0.06	0.0021 ± 0.00018	0.002 ± 0.0007
MD-RC	9.76 ± 0.94	0.08 ± 0.02	0.0036 ± 0.00034	1.18 ± 0.13
MD-RD	23.12 ± 0.63	0.4 ± 0.03	0.0021 ± 0.00007	1.47 ± 0.08

resistance mechanism by amplification in *A. palmeri* evolved once and then rapidly spread across the USA, either by seed[22] or by pollen.[23] The data also indicate that the amplicon is highly conserved across populations. A more complete representation of the singular nature of the *EPSPS* cassette depends upon sequencing plants from additional populations using next-generation sequencing technology and assembly of the reads to the *EPSPS* cassette to assess its entire content and degree of amplification. The persistence of many copies of a large DNA sequence in the *A. palmeri* genome is a testament to the resilience of glyphosate resistance, but it also drives home the point that a single weed escaping control has the potential to wreak devastating consequences upon agriculture.

ACKNOWLEDGEMENTS

The authors are grateful to Fanny Liu and Dr Brian Scheffler of the Genomics and Bioinformatics Research Unit, USDA-ARS, Stoneville, MS, USA for their assistance with Sanger sequencing.

REFERENCES

1 Powles SB, Lorraine-Colwill DF, Dellow JJ and Preston C, Evolved resistance to glyphosate in rigid ryegrass (*Lolium rigidum*) in Australia. *Weed Sci* **46**:604-607 (1998).
2 Simarmata M, Kaufmann JE and Penner D, Potential basis of glyphosate resistance in California rigid ryegrass (*Lolium rigidum*). *Weed Sci* **51**:678-682 (2003).
3 VanGessel MJ, Glyphosate-resistant horseweed from Delaware. *Weed Sci* **49**:703-705 (2001).
4 Culpepper AS, Grey TL, Vencill WK, Kichler JM, Webster TM, Brown SM, York AC, Davis JW and Hanna WW, Glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*) confirmed in Georgia. *Weed Sci* **54**:620-626 (2006).
5 Bensch CN, Horak MJ and Peterson D, Interference of redroot pigweed (*Amaranthus retroflexus*), Palmer amaranth (*A. palmeri*), and common waterhemp (*A. rudis*) in soybean. *Weed Sci* **51**:37-43 (2003).
6 Massinga RA, Currie RS, Horak MJ and Boyer J, Interference of Palmer amaranth in corn. *Weed Sci* **49**:202-208 (2001).
7 Morgan GD, Baumann PA and Chandler JM, Competitive impact of Palmer amaranth (*Amaranthus palmeri*) on cotton (*Gossypium hirsutum*) development and yield. *Weed Technol* **15**:408-412 (2001).
8 Heap I, The International Survey of Herbicide Resistant Weeds. Online. Internet. Tuesday, August 8, 2017. Available www.weedscience.org.
9 Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE and Westra P, Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *P Natl Acad Sci USA* **107**:1029-1034 (2010).
10 Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW, Scott RC and Burgos NR, *EPSPS* gene amplification in glyphosate-resistant Italian ryegrass (*Lolium perenne* ssp. *multiflorum*) from Arkansas. *Pest Manag Sci* **68**:1223-1230 (2012).
11 Jugulam M, Niehues K, Godar AS, Koo D-H, Danilova T, Friebe B, Sehgal S, Varanasi VK, Wiersma A, Westra P, Stahlman PW and Gill BS, Tandem amplification of a chromosomal segment harboring *5-enolpyruvylshikimate-3-phosphate synthase* locus confers glyphosate resistance in *Kochia scoparia*. *Plant Physiol* **166**:1200-1207 (2014).
12 Lorentz L, Gaines TA, Nissen SJ, Westra P, Strek HJ, Dehne HW, Ruiz-Santaella JP and Beffa R, Characterization of glyphosate resistance in *Amaranthus tuberculatus* populations. *J Agr Food Chem* **62**:8134-8142 (2014).
13 Nandula VK, Wright AA, Bond JA, Ray JD, Eubank TW and Molin WT, *EPSPS* amplification in glyphosate-resistant spiny amaranth (*Amaranthus spinosus*): a case of gene transfer via interspecific hybridization from glyphosate-resistant Palmer amaranth (*Amaranthus palmeri*). *Pest Manag Sci* **70**:1902-1909 (2014).
14 Malone JM, Morran S, Shirley N, Boutsalis P and Preston C, *EPSPS* gene amplification in glyphosate-resistant *Bromus diandrus*. *Pest Manag Sci* **72**:81-88 (2016).
15 Gaines TA, Wright AA, Molin WT, Lorentz L, Riggins CW, Tranel PJ, Beffa R, Westra P and Powles SB, Identification of genetic elements associated with *EPSPS* gene amplification. *PLOS ONE* **8**:e65819 (2013).
16 Molin WT, Wright AA, Lawton-Rauh A and Saski CS, The unique genomic landscape surrounding the *EPSPS* gene in glyphosate resistant *Amaranthus palmeri*: A repetitive path to resistance. *BMC Genomics* **18**:91 (2017).
17 Paterson AH, Brubaker CL and Wendel JF, A rapid method for extraction of cotton genomic DNA suitable for RFLP or PCR analysis. *Plant Mol Biol Rep* **11**:122-127 (1993).
18 Sambrook J, Fitsch EF and Maniatis T, *Molecular cloning: a laboratory manual*. Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

19 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Mentijes P and Drummond A, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647-1649 (2012).
20 Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q and Xiong L, Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *P Natl Acad Sci USA* **103**:12987-12992 (2006).
21 Sung DY and Guy CL, Physiological and molecular assessment of altered expression of Hsc70-1 in Arabidopsis. *Evidence for pleiotropic consequences. Plant Phys* **132**:979-987 (2003).
22 Norsworthy, JK, Griffith G, Griffin T, Bagavathiannan M and Gbur EE, In-field movement of glyphosate-resistant Palmer Amaranth (*Amaranthus palmeri*) and its impact on cotton lint yield: Evidence supporting a zero-threshold strategy. *Weed Sci* **62**:237-249 (2014).
23 Sosnoskie LM, Webster TM,, MacRae AW, Grey TL and Culpepper AS, Pollen-mediated dispersal of glyphosate-resistance in Palmer amaranth under field conditions. *Weed Sci* **60**:366-373 (2012).

