Genomic-based epidemiology reveals gene flow and independent origins of glyphosate resistance in *Bassia scoparia* populations across North America

Karl Ravet[1], Crystal Sparks[1], Andrea Dixon[2], Anita Küpper[3], Eric Westra[1], Dean Pettinga[1], Patrick Tranel[4], Joel Felix[5], Don Morishita[6], Prashant Jha[7], Andrew Kniss[8], Phillip Stahlman[9], Paul Neve[2], Eric Patterson[10], Philip Westra[1], and Todd Gaines[1]

[1]Colorado State University
[2]Rothamsted Research
[3]Bayer CropScience AG
[4]University of Illinois at Urbana-Champaign
[5]Oregon State University
[6]University of Idaho
[7]Iowa State University
[8]University of Wyoming
[9]Kansas State University
[10]Michigan State University

November 1, 2020

Abstract

Genomic-based epidemiology can provide insight into the origins and spread of herbicide resistance mechanisms in weeds. We used kochia (Bassia scoparia) populations resistant to the herbicide glyphosate from across western North America to test the alternative hypotheses that 1) a single EPSPS gene duplication event occurred initially in the Central Great Plains and then subsequently spread to all other geographical areas now exhibiting glyphosate-resistant kochia populations or that 2) gene duplication occurred multiple times in independent events in a case of parallel evolution. We used qPCR markers previously developed for measuring the various units of the EPSPS tandem duplication to investigate whether all glyphosate-resistant plants had the same EPSPS repeat structure. We also investigated population structure using simple sequence repeat (SSR) markers to determine the relatedness of kochia populations from across the Central Great Plains, Northern Plains, and the Pacific Northwest. We identified three distinct EPSPS-duplication haplotypes that had geographic associations with the Central Great Plains, Northern Plains, and Pacific Northwest. Population structure revealed a group of populations around the first reported occurrence of glyphosate resistance in the Central Great Plains, a separate group of Pacific Northwest populations, and some relatedness of populations from geographically isolated areas. The results support at least three independent origins of glyphosate resistance in kochia, followed by substantial and mostly geographically localized gene flow to spread the resistance alleles into diverse genetic backgrounds.

INTRODUCTION

The origins and spread of herbicide resistance in a landscape context may result from a single (or few) independent sources followed by dispersal via gene flow versus multiple, localized, independent evolutionary events (Baucom, 2019). A summary of previous research indicates that gene flow often contributes similarly to or more than independent evolution to the occurrence of resistance among populations (Beckie, Busi,

Bagavathiannan, & Martin, 2019). Use of highly effective herbicides, such as glyphosate, across large areas generates a stark selection pressure in which individuals with resistance alleles exhibit high fitness. This widespread selection intensity increases the probability that rare resistance mutations that arise independently will increase in frequency rather than be lost via genetic drift, resulting in a rapid increase in the frequency of a resistance allele following movement to a new location via pollen or seed-mediated dispersal (Beckie et al., 2019). Genomic-based epidemiology enables identification of unique resistance alleles as well as tracking movement of individuals across a landscape when the mutations associated with herbicide resistance are known, providing insight into the origins and spread of resistance mechanisms (Comont & Neve, 2020). Understanding these patterns of independent origins and movement will provide insights for improved management practices, such as focusing area-wide management approaches to prevent sources of seed introduction to new areas when gene flow is moving resistance, or managing selection pressure on a field-by-field basis when resistance is evolving independently and frequently.

To investigate the question of gene flow and independent evolution for herbicide resistance using genomic-based epidemiology, we used kochia [*Bassia scoparia* (L.) A. J. Scott, synonymous with *Kochia scoparia* (L.) Schrad.], an introduced weed that occurs in the semi-arid arable lands of the western United States and Canada (Friesen, Beckie, Warwick, & Van Acker, 2009). Kochia collections from around North America exhibit high levels of genetic diversity in the invaded range and a lack of true population structure (Kumar, Jha, Jugulam, Yadav, & Stahlman, 2019; Mengistu & Messersmith, 2002). This lack of structure is most likely due to several reproductive traits that emphasize cross pollination as well as long-range dispersal of both pollen and seed (Beckie, Blackshaw, Hall, & Johnson, 2016). Kochia has protogynous flowers in which the stigmas emerge first and are receptive to pollen from other flowers before pollen production within the same flower occurs, reducing the self-pollination rate (Guttieri, Eberlein, & Thill, 1995). Additionally, kochia is a well-known tumbleweed species, with some plants dispersing seeds for dozens or even hundreds of miles (Kumar et al., 2019). This dispersal mechanism greatly increases the spread of herbicide-resistance alleles and makes containment extremely difficult (Beckie et al., 2016; Kumar et al., 2019; Stallings, Thill, Mallory-Smith, & Shafii, 1995).

Herbicide-resistance mechanisms can be classified as either target-site (mutations or changes in expression of the gene encoding the protein inhibited by the herbicide) or non-target-site (mechanisms that reduce the concentration of active herbicide reaching the target site protein) (Gaines et al., 2020). Target-site mechanisms can be considered specialist adaptations while non-target-site mechanisms can be considered generalist mechanisms, as they can sometimes confer resistance across different herbicide modes of action (Baucom, 2019). Evolution of resistance to the herbicide glyphosate has included both target-site and non-target-site mechanisms, including the parallel evolution of increased copy number of the gene encoding the glyphosate target enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) in multiple species (Gaines, Patterson, & Neve, 2019; Patterson, Pettinga, Ravet, Neve, & Gaines, 2018).

The first report of glyphosate-resistant (GR) kochia was from Kansas in 2007 (Waite et al., 2013) and reports have since confirmed glyphosate resistance in multiple US states and Canadian provinces (Kumar et al., 2019). The widespread regional evolution of GR kochia has negatively impacted the sustainability of reduced-tillage weed management and moisture and soil conservation during fallow periods in the Central Great Plains of North America (Kumar et al., 2019). The mechanism of glyphosate resistance has been thoroughly investigated in kochia, in terms of physiology and fitness penalty as well as the genetic mechanisms that cause resistance (Beckie et al., 2018; Kumar & Jha, 2015; Martin et al., 2017; Osipitan & Dille, 2017; Wiersma et al., 2015). Increased copy number of the *EPSPS* gene has been identified as the resistance mechanism in all studied kochia populations to date from across seven US states (Montana, Wyoming, Oregon, Idaho, Nebraska, Kansas, and Colorado) (Gaines et al., 2016; Godar, Stahlman, Jugulam, & Dille, 2015; Kumar, Felix, Morishita, & Jha, 2018; Kumar, Jha, Giacomini, Westra, & Westra, 2015; Wiersma et al., 2015). In kochia, *EPSPS* gene duplications are tandem and occur at a single locus (Jugulam et al., 2014; Patterson et al., 2019).

The *EPSPS* locus has been sequenced from a single GR kochia individual using BAC libraries. The *EPSPS*

repeat unit was variable with two units being most common; 1) a full-length repeat containing *EPSPS* and six other flanking genes and 2) a less frequent form containing *EPSPS* and only three other flanking genes (Patterson et al., 2019). A ˜15 kb mobile genetic element (MGE) was found inserted flanking both upstream and downstream of the entire tandem duplication. Additionally, a copy of the MGE was found between every repeat unit, indicating the MGE has been co-duplicated after subsequent crossing over events (Patterson et al., 2019). Once the structure of the repeat was determined, quantitative PCR markers that specifically amplify the two types of repeats and the MGE were developed to confirm the sequence and measure the copy number of each part of the repeat structure.

Due to long distance gene dispersal in kochia and its prevalence in cropping systems where it can be carried in/on agricultural equipment, we tested the hypothesis that a single *EPSPS* gene duplication event occurred initially in Kansas and then subsequently spread to initiate all other GR kochia populations and its alternative, that gene duplication may have occurred multiple times in independent events in a case of parallel evolution. In this study, we used two approaches to address our question. First, we used a genomic-based epidemiology approach with the qPCR markers developed for measuring the various units of the *EPSPS* tandem duplication to investigate whether all GR plants had the same *EPSPS* repeat structure. Second, we used a population genetics approach with simple sequence repeat (SSR) markers to determine the relatedness of GR and glyphosate-susceptible (GS) kochia populations from across the Central Great Plains, Northern Plains, and the Pacific Northwest. Combined, these approaches provide insight into the origins and spread of glyphosate resistance alleles in kochia.

MATERIALS AND METHODS

Plant materials

Seeds were collected from kochia individuals from locations in the western US and western Canada between 2010 and 2015 (Table 1). Crops grown at the sampled locations included winter wheat, no-till fallow, and sugar beet. A total of 44 kochia locations from eight different states in the USA and one province in Canada were used for the analyses (Table 1). A population is defined as the seed from kochia individuals isolated from a single field at a geographically distinct location. The number of individuals sampled per population varied from five to 20, with at least 100 seeds sampled per plant and collecting from plants across at least 0.5 ha. GS populations were collected from several locations in Kansas, Colorado, and Oregon. Populations suspected to contain GR individuals were collected throughout Kansas, Colorado, Oregon, Idaho, Oklahoma, Texas, Montana, and Alberta (Table 1). Populations from Colorado (Westra, Nissen, Getts, Westra, & Gaines, 2019), Oregon and Idaho (Kumar et al., 2018), and Montana (Kumar, Jha, & Reichard, 2014) were previously screened for glyphosate resistance.

Glyphosate resistance screening

A population ID was assigned to each locality (Table 1). Each population ID consists of the abbreviated state/province of origin, a unique identifying number, and its designation as GR or GS (e.g., CO1R = Colorado resistant population 1). To determine glyphosate susceptibility and resistance, a screening was performed in the greenhouse. Seeds from each population were planted in germination flats. After emergence, seedlings were transplanted into 18-insert (7×7 cm pots) flats containing custom mix potting soil (Fafard, Sun Gro Horticulture, Agawam, MA, USA), and grown at 23°C under a 14 h light/10 h dark cycle with supplemental light from sodium halide lamps. Plants were watered daily and fertilized once (Miracle-Gro, Scotts Miracle-Gro Company, Marysville, OH). When plants were 3 weeks old, nine to 18 plants per population (Table 1) were sprayed with commercially formulated glyphosate (Roundup WeatherMax) in distilled water at 0.84 kg ae ha^{-1}. Glyphosate applications were performed using a moving flat-fan nozzle (8002EVS) in a laboratory spray chamber at 156 L ha^{-1} spray volume. Three weeks after herbicide treatment, individual survival for each population was assessed. Populations for which at least one individual survived were classified as GR.

DNA isolation and genotyping

For total genomic DNA extraction, leaf tissue was collected from plants grown as previously described prior to glyphosate treatment. Samples were immediately frozen in liquid nitrogen and stored at -80°C. Tissue was ground using a TissueLyser II (Qiagen; 30 Hz for 2 min). Genomic DNA was extracted from 100 mg fresh weight tissue following a modified cetyltrimethylammonium bromide (CTAB) extraction protocol (Doyle, 1991). DNA quality and concentration were measured using a NanoDrop spectrophotometer (Thermo Scientific ND-1000). All DNA samples were normalized to 5 ng uL^{-1} with deionized water.

Genomic-based epidemiology using *EPSPS* copy number and associated duplication markers

Genomic DNA from 113 individuals representing 27 of the populations evaluated for glyphosate resistance above, along with 36 individuals representing 11 populations used in Gaines et al. (2016) and 58 individuals representing 18 populations collected in Montana, was used for real time PCR to measure the relative copy number of genomic *EPSPS* , as well as the longer type I (56.1 kb) and shorter type II (32.7 kb) segments associated with *EPSPS*duplication, and the mobile genetic element (MGE) Fhy3/FAR1 (Patterson et al., 2019). Primer sequences for these features as well as the normalization gene *Acetolactate Synthase (ALS*) were 1) EPSPS, For (5'-CGCTATATGTTGGATGCTCTAAG-3'), Rev (5'-CACTCCTATTCTCTTTACCAGC-3'); 2) Type I (56.1 kb), For (5'-GACGGAAATACCCTCAATATAGACA-3'), Rev (5'-ACGCCCAAGATGTACATTGATA-3'); 3) Type II (32.7 kb), For (5'-GACGGAAATACCCTCAATATAGACA-3'), Rev (5'-CATGCCTTTGATGTCCAAGTTT-3'); 4) MGE, For (5'-GAAGATAGCGAGACGTTTGAG-3'), Rev (5'-CGGCTTGATCGGTTAAGATAC-3'); and 5) ALS, For (5'-CCAGAAAAGGCTGCGATG-3'), Rev (5'-CTGACTCGCTCTGATTCCA-3'). A GR control (population M32) with high *EPSPS* copy number, presence of both type I and type II *EPSPS* duplication segments, and an increased copy number of MGE was included along with a susceptible control (7710) containing a single copy of *EPSPS* , no copies of type I or II markers, and a low copy number of MGE. The qPCR protocol of Patterson et al. (2019) was used and relative copy number was calculated using the [?]Ct method (Schmittgen & Livak, 2008). The type I and type II*EPSPS* duplication qPCR primers have a forward primer in the MGE and a reverse primer in either the type I or type II sequence, respectively, enabling amplification only when the MGE is located next to the type I or type II repeat segment. Using these markers, we quantified the number of type I (56.1 kb) or type II (32.7 kb) repeats in individuals from multiple populations. Three haplotypes were defined based on the four qPCR markers, using *EPSPS* >1.4 to define increased *EPSPS* , type I and type II >0 to define presence of the two markers of *EPSPS* gene duplication described in Patterson et al. (2019), and MGE <10 defined as 'normal' and [?]10 defined as 'increased' MGE. Populations and their haplotypes and geographic locations were plotted in R using ggplot and sf packages (R, 2019).

SSR genotyping

To develop polymorphic genetic markers for kochia, Roche 454 sequencing technology (Keck Center, University of Illinois) was used to determine partial genomic sequence from a single GR kochia individual. Approximately 75.2 million aligned bases (from 357 million total bases sequenced) with reads having an average length of 557 bases were obtained. This dataset was screened for simple sequence repeats (SSRs) with pentanucleotide repeat units to use as molecular markers for genotyping. Out of a total of 30 SSR markers initially tested, 11 SSR markers (Table 2) exhibiting polymorphisms were selected for genotyping the 44 kochia populations.

Amplification of 100 to 200 bp sequence regions containing the selected SSR markers was carried out using polymerase chain reaction with specific primers (Table 2; together with expected fragment size of the amplified loci). Amplification of 5 ng of genomic DNA was performed using EconoTaq PLUS Master Mix (Lucigen) in a BioRad CFX96 Real-Time System (C1000 Touch Thermal Cycler). After an initial denaturation period of 2 minutes at 94degC, PCR was run for 37 cycles, consisting of denaturation at 94degC for 30 seconds, annealing at either 57degC or 62degC for 30 seconds (Table 2), and extension at 72degC for 45 seconds. A final extension of 2 minutes at 72degC was included. Amplified fragment size analysis was carried out by capillary electrophoresis using an Advanced Analytical Fragment Analyzer, using the 35-500 bp dsDNA method. Fragments were sized by ProSize 2.0 software. Alleles were binned using Flexibin software (Amos,

2005; Amos et al., 2007).

Genetic diversity and population structure

The evaluation of linkage equilibrium and Hardy-Weinberg equilibrium of loci was done using exact tests with the functions test_LD and test_HW, respectively in the "genepop" package (v.1.1.7; Rousset, 2008). Descriptive summaries for each population across loci were calculated using the divBasic function in the "diveRsity" package (v. 1.9.90; Keenan, McGinnity, Cross, Crozier, & Prodohl, 2013). Descriptive summaries for each locus across populations were calculated using locus_table function in the "poppr" package (v. 2.8.6; Kamvar, Brooks, & Grunwald, 2015; Kamvar, Tabima, & Grunwald, 2014).

Missing data were assessed using the info_table function in the "poppr" package (v. 2.8.6; Kamvar et al., 2015; Kamvar et al., 2014) and loci with more than 10% data missing and individuals with more than 20% data missing were removed. Descriptive summaries for populations and loci were performed on the entire data set and again on the data set after removing loci and individuals with unacceptable levels of missing data. A population level phylogeny based on the neighbor-joining clustering method using the Prevosti's genetic distance model was generated with bootstrapped support using 1000 replicates with the aboot function in the "poppr" package (v. 2.8.6; Kamvar et al., 2015; Kamvar et al., 2014) and plotted using the Interactive Tree of Life v4 (Letunic & Bork, 2019).

Model-based putative population clustering was performed using STRUCTURE v2.3.4 (Pritchard, Stephens, & Donnelly, 2000). The number of genetic groups (K) present within the 509 individuals tested was determined running a continuous series of K = 1-22. The program was run with a burn-in of 30,000 and a run-length of 100,000 Markov Chain Monte Carlo (MCMC) replications in 20 independent runs using the LOCPRIOR model (sampling location information included) to account for weak structure signals in the dataset. The most likely number of clusters was determined using the Evanno method (Evanno, Regnaut, & Goudet, 2005) as implemented in STRUCTURE HARVESTER v0.6.94 (Earl & vonHoldt, 2012). The final analysis for K=3 was performed using a burn-in of 50,000 and 500,000 MCMC with 20 independent runs. Runs were summarized using CLUMPP v.1.1.2b (Jakobsson & Rosenberg, 2007) utilizing the Greedy algorithm and visualized with DISTRUCT v1.1 (Rosenberg, 2004).

RESULTS

Genomic-based epidemiology using *EPSPS* copy number and associated duplication markers

Glyphosate resistance in kochia was first detected in Kansas in 2007, followed by Colorado and Alberta in 2012, Oklahoma and Montana in 2013, and Texas, Wyoming, Idaho, and Oregon in 2014 (Heap, 2020). We surveyed a set of populations from across western North America for glyphosate resistance (Table 1). Most populations returned the expected phenotype; however, a few populations were designated as suspected GR based on field observations while all nine individuals tested had GS phenotype in our screening assay, resulting in a classification for the population as GS but with potential for it to be heterogeneous for phenotype (containing GR at a low frequency). For example, both KS5S and KS6S were suspected to be GR when sampled from the field but were classified as S by phenotyping (Table 1) and had GR individuals based on *EPSPS*gene copy number data (Table S1).

We identified three categories of *EPSPS* -duplication haplotypes, defined as follows: 1) increased *EPSPS* , Type I and II, and MGE copy numbers that correspond to increased *EPSPS* copy number ([?] 10) (Central Great Plains); 2) increased *EPSPS* , no Type I or II, MGE [?] 10 (Northern Plains); 3) increased *EPSPS* , no Type I or II, MGE $<$ 10 (Pacific Northwest, north-central Wyoming) (Figure 1). Although all GR individuals from across the continent had increased*EPSPS* copy number, kochia in the Pacific Northwest and Northern Plains did not contain the type I and type II repeats associated with the tandem *EPSPS* duplication in the previously characterized population from Colorado (Figure 1). In contrast, the type I and II repeats were present in the Central Great Plains, and in ratios consistent with those reported by Patterson et al. (2019) (Table S1). In the Central Great Plains, Type I and II always amplified together, and Type I (long repeat) was nearly always present at higher copy number than Type II (short repeat) (Table S1), as

in the originally characterized *EPSPS* repeat structure (Patterson et al., 2019). Some individuals had much higher *EPSPS* copy number than previously reported; for example, individuals collected in MT had between 20-30 copies with no presence of Type I or II and very high (>60 copies) MGE (Table S1). These individuals may represent independent origins via a different molecular genetic mechanism, which will require additional sequencing to assemble this specific *EPSPS* duplication haplotype.

All GS kochia samples in the survey had 1 copy of *EPSPS* , no amplification of Type I or II markers, and 4-6 copies of the MGE (Table S1). The absence of amplification of type I and II markers in GS kochia samples, which requires insertion of the MGE next to duplicated copies of *EPSPS* , confirms that the MGE is not inserted at the start of the *EPSPS* locus in any of the diverse susceptible populations sampled and further supports the hypothesis that MGE insertion occurred prior to the *EPSPS* duplication event, rather than the MGE having been present next to EPSPS ancestrally and co-duplicated with*EPSPS* (Patterson et al., 2019). MGE copy number was increased to >10 in one GS individual from each of four populations (located in CO, NE, and MT) without corresponding increase in*EPSPS* gene duplication (Table S1), suggesting copy number of this MGE varies across populations and it may represent an active element, or that the *EPSPS* duplication locus can segregate away from other MGE duplication sites in the genome.

Genetic diversity and population structure

We next asked if the population genetics data supported the three independent origins identified by genomic-based epidemiology. We developed 11 SSR markers (Table 2) and genotyped kochia populations collected from the three geographic regions to measure population-level genetic diversity and genetic similarity of populations between localities. Across all populations using Fisher's combined probability test, all SSR loci were in linkage equilibrium ($P > 0.05$), but not in Hardy-Weinberg equilibrium (HWE; $P < 0.05$).

Across all loci and populations, 3.98% of the data were missing (Table S2). Of the loci, marker "SSR162" had the most missing data across populations at 13.8%, whereas across loci, populations KS2S and KS8S had the most missing data at 26.3% and 17.2% missing, respectively (Table S3). For descriptive summaries and the neighbor-joining tree, marker "SSR162" was removed as well as five individuals: KS2S_4, KS2S_5, KS2S_7, KS2S_8, KS8S_2, as this locus had more than 10% missing data and the individuals had more than 20% missing data after "SSR162" was removed.

Allele counts, expected heterozygosity, and evenness for all loci across all populations and then after the removal of individuals with missing data are reported in Table S2. Descriptive summaries of 44 populations at ten SSR loci are presented in Table 3 (data for all 11 SSR loci are presented in Table S3). Populations ranged in their percentage of total alleles observed and allelic richness from 57.7% and 1.42 (KS13R) to 24.7% and 2.61 (CO7R). F_{IS} ranged from -0.04 (95% CI = -0.28 – 0.13; KS13R) to 0.58 (95% CI = 0.33 – 0.79; MT3R), while most are in the range of 0.2 to 0.4. A positiveF_{IS} indicates a deficiency of heterozygotes in the population compared to the proportion expected in HWE and a negativeF_{IS} indicates an excess of heterozygotes. TheF_{IS} results should be interpreted with caution noting that loci did not meet the assumptions of Hardy-Weinberg (Waples, 2015) and many confidence intervals spanned from negative to positive and over very large ranges.

As loci and populations did not meet the assumptions of HWE, a neighbor-joining tree was used to assess genetic similarity between populations. This tree showed some expected groups by region, with 12 Central Great Plains populations grouped in a large clade supported 100% by bootstrap values (Figure 2). This clade also contained OR4R (Pacific Northwest) and MT2R (Northern Plains). The STRUCTURE analysis showed that K=3 was the number of clusters or gene pools best supported (Figure S1) and also supported the grouping of OR4R and MT2R with the Central Great Plains populations including CO1R, KS10R, and KS11R (Figure 3). The populations from the Pacific Northwest largely clustered together (OR2R, OR3R, OR6R, OR7R, OR9S, ID1R, and ID2R) with the clade of populations OR9S and ID1R and OR7R and ID2R supported at 61.5% (Figure 2). Populations KS13R, MT3R, and CO6R clustered with this Pacific Northwest group (Figure 2) and had similarity in the STRUCTURE analysis (Figure 3). Some groupings were unexpected, such as a grouping of TX2R, TX3R, TX4R, and TX5R (Central Great Plains) populations

with Alberta, Canada (Northern Plains), as well as OR1R (Figure 2).

DISCUSSION

Genomic-based epidemiology using markers specific to a characterized resistant haplotype of *EPSPS* duplication from the Central Great Plains identified three unique *EPSPS* haplotypes in GR kochia populations from across western North America. Haplotypes 1 and 2 both showed increased copy numbers of the MGE previously identified, but haplotype 2 does not have insertion of the MGE next to *EPSPS* at the same position as identified in haplotype 1. *EPSPS* duplication in haplotype 2 may have occurred through a similar genetic mechanism as haplotype 1 involving insertion of an active mobile genetic element next to *EPSPS* followed by tandem duplication. Resequencing and assembly of the *EPSPS* locus in haplotype 2 will be needed to determine the precise duplication mechanism that occurred and to determine whether the MGE, which had increased copy number in haplotype 2, is associated with the *EPSPS* duplication. We consider haplotype 3 to represent a third independent origin because it had increased *EPSPS* copy number and no increase in MGE, indicating that *EPSPS* duplication may also have occurred through yet another mechanism such as a double-strand break initiated by a different MGE. These haplotypes, based on either the presence or absence of type I and type II elements and ratio of MGE, provide evidence that glyphosate resistance evolved multiple times in geographically distinct locations, with three independent origins supported by the data.

Haplotype 3 was found in populations from OR, ID, and WY. From the neighbor-joining tree, population WY1R (Northern Plains) was closest to OR5R (Pacific Northwest) with support of 41% (Figure 2), but this pair was distant from the other Pacific Northwest populations and not clustered with Pacific Northwest populations in the STRUCTURE plot (Figure 3). The shared haplotype 3 between geographically isolated northern Wyoming and the Pacific Northwest could indicate two separate origins of haplotype 3 (increased-*EPSPS* without increased MGE), or it could indicate lack of resolution in the population genetics data to resolve gene flow from independent origins of resistance. Further characterization of the duplicated resistant locus is needed to determine whether populations from the Northern Plains (WY1R) and the Pacific Northwest have a shared or separate origin of duplicated*EPSPS* genes. Other Northern Plains populations (WY and southern MT) contained individuals with either haplotype 2 or haplotype 3, indicating gene flow via migration between populations and/or dynamic MGE changes in copy number over time.

We predicted that population genetics analysis would show clear geographical structure if the three glyphosate resistance evolutionary events occurred and were followed by rapid dispersal and introgression within a region. Aside from a large grouping of Central Great Plains populations and a second grouping of Pacific Northwest populations, we were not able to identify clear geographical structure for the three regions corresponding to the three *EPSPS* haplotypes. The STRUCTURE analysis supported K=3, although populations were not consistently assigned to three groups corresponding geographically to the regions containing the three *EPSPS* haplotypes and most populations contained some presence of all three groups. Some strong signals of relatedness were detected between populations from geographically isolated locations, such as OR4R and KS10R (Figures 2, 3). While we consider it to be unlikely that the same duplication mechanism is evolving independently multiple times within a region on different genetic backgrounds, the SSR data may have insufficient resolution to identify population genetic relationships and extensive long-distance gene flow via seeds may make regional differences harder to detect. This aligns with previous population genetics studies in kochia that have found high genetic diversity within individuals and little population structure (Friesen et al., 2009; Kumar et al., 2019; Mengistu & Messersmith, 2002).

Haplotype distribution was consistent with division by mountainous geographic barriers, with haplotype 1 distributed across the Central Great Plains associated with the earliest report of glyphosate resistance in kochia, 2007, in Kansas (Waite et al., 2013); haplotype 2 located within the Northern Plains including northern Montana and Alberta; and haplotype 3 located in sugar beet fields in the Pacific Northwest and north-central Wyoming (Figure 1). Since glyphosate resistance evolved recently, the geographic patterns of haplotype distribution (via seed dispersal) also occurred more recently and appear to be introgressing at a more regional scale into diverse genetic backgrounds (via pollen flow).

The presence of more than one *EPSPS* haplotype within populations provides strong evidence for gene flow among populations. This was observed in some northern Colorado populations, with some haplotype 2 individuals present in populations that were mostly haplotype 1 (e.g., CO3R from Cope, CO showing high *EPSPS* , no Type I or II, and high MGE, like samples from northern Montana and Alberta), and some haplotype 3 individuals present in populations mostly containing haplotype 1 (Eaton, CO) (Figure 1, Table S1). These two populations showing admixture in CO indicates gene flow has occurred in this area.

A curious result in the population relatedness data was that several populations from Texas (TX3R, TXR4, TX5R), a population from Oregon (OR1R), and the furthest north population from Alberta, Canada (AB1R) were grouped with relatively low (>20%) bootstrap support in the phylogeny tree (Figure 2). Further research will be needed to define the relationship of these populations and to determine whether long-distance gene flow is occurring in kochia, either through natural tumbling dispersal or through human-mediated seed migration. Some other weed species also demonstrate little population structure. Palmer amaranth (*Amaranthus palmeri*), an obligate outcrossing weed species subjected to widespread glyphosate selection pressure, also maintains high levels of genetic diversity and little population structure (Kupper et al., 2018; Molin, Patterson, & Saski, 2020a). Kochia has protogynous flowers and pollen that is both wind and insect dispersed (Blackwell & Powell, 1981), facilitating high outcrossing levels. Kochia is also one of the primary tumbleweed species of the western US and Canada. This uncommon mode of seed dispersal may serve to further homogenize genetic diversity across long geographic distances.

Single origins of herbicide resistance followed by substantial geographical distribution by pollen and seed mediated gene flow is known to have a major contribution to resistance frequency in multiple weed species (Beckie et al., 2019). Multiple independent origins of glyphosate resistance with little population structure was found in glyphosate-resistant populations of common morning glory (*Ipomoea purpurea*) (Kuester, Chang, & Baucom, 2015). A genomics-based approach in the same species found evidence for parallel genetic responses in genomic regions encoding herbicide detoxification in, while other genomic regions showed divergent patterns of selection (Van Etten, Lee, Chang, & Baucom, 2020). The high sequence similarity of the extrachromosomal DNA containing the *EPSPS* gene in Palmer amaranth (Koo et al., 2018; Molin, Yaguchi, Blenner, & Saski, 2020b) across widespread populations supports the hypothesis of a single origin followed by dispersal (Molin et al., 2020a; Molin, Wright, Lawton-Rauh, & Saski, 2017; Molin et al., 2018). Glyphosate-resistant populations of flaxleaf fleabane from across multiple Australian states were highly related, supporting a high frequency of seed movement (Minati, Preston, & Malone, 2020). Multiple independent origins of glyphosate resistance were detected in horseweed in California, with localized movement of resistant individuals accounting for spread on regional levels correlating with groundwater regulations that encouraged more glyphosate use and less use of other herbicides (Okada et al., 2013).

In summary, we used genomic-based epidemiology to track the mutations underlying one specific origin of glyphosate resistance in kochia and showed that at least three independent origins of glyphosate resistance have evolved in kochia, followed by substantial regional gene flow to spread the resistance alleles to new genetic backgrounds. Due to the tumbling dispersal of kochia, intercepting seed movement across the landscape has high potential to mitigate the negative impact of herbicide resistance spreading from an initial origin. With the kochia reference genome now available (Patterson et al., 2019), the population genomics approach used by Kreiner et al. (2019) can be used in kochia to study population divergence and origins of resistance (Martin et al., 2019). Sequencing and assembly of the duplicated region from haplotypes 2 and 3 would provide insights and new markers to further investigate the evolutionary dynamics of the *EPSPS* tandem duplication in kochia across western North America.

Acknowledgements

The work was supported by USDA National Institute of Food and Agriculture grant number 2012-67013-19347 to PW, Western Sugar Joint Research Committee, Snake River Sugar Beet Research and Seed Alliance, Colorado Wheat Administrative Committee, USDA-NIFA Hatch project COL00783 at Colorado State University, Biotechnology and Biological Sciences Research Council, UK grant number BB/N022319/1 to PN and TAG, and USDA National Institute of Food and Agriculture through the Western Integrated Pest Man-

agement Center, grant number 2018-70006-28881 to support the Western IPM Kochia Work Group. The authors thank Rachel Chayer, Adrian Quicke, Shaheen Bibi, and Darci Giacomini for their contributions to lab work and plant growth.

REFERENCES

Amos, B. (2005). FlexiBin, a program to automate the binning of microsatellite alleles. Available: https://www.zoo.cam.ac.uk/research-groups/molecular-ecology/FlexiBin.pdf.

Amos, W., Hoffman, J., Frodsham, A., Zhang, L., Best, S., & Hill, A. (2007). Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes, 7* , 10-14.

Baucom, R. S. (2019). Evolutionary and ecological insights from herbicide-resistant weeds: what have we learned about plant adaptation, and what is left to uncover? *New Phytologist, 223* , 68-82.

Beckie, H. J., Blackshaw, R. E., Hall, L. M., & Johnson, E. N. (2016). Pollen-and seed-mediated gene flow in kochia (*Kochia scoparia*).*Weed Science, 64* , 624-633.

Beckie, H. J., Blackshaw, R. E., Leeson, J., Stahlman, P. W., Gaines, T., & Johnson, E. A. (2018). Seed bank persistence, germination and early growth of glyphosate-resistant *Kochia scoparia* . *Weed Research, 58* , 177-187.

Beckie, H. J., Busi, R., Bagavathiannan, M. V., & Martin, S. L. (2019). Herbicide resistance gene flow in weeds: Under-estimated and under-appreciated. *Agriculture, ecosystems & environment, 283* , 106566.

Blackwell, W. H., & Powell, M. J. (1981). A preliminary note on pollination in the Chenopodiaceae. *Annals of the Missouri Botanical Garden, 68* , 524-526.

Comont, D., & Neve, P. (2020). Adopting epidemiological approaches for herbicide resistance monitoring and management. *Weed Research* , In press.

Doyle, J. (1991). DNA protocols for plants–CTAB total DNA isolation. In 'Molecular techniques in taxonomy'. (Eds GM Hewitt, A Johnston) pp. 283–293. In: Springer: Berlin.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources, 4* , 359-361.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14* , 2611-2620.

Friesen, L. F., Beckie, H. J., Warwick, S. I., & Van Acker, R. C. (2009). The biology of Canadian weeds. 138. *Kochia scoparia* (L.) Schrad. *Canadian Journal of Plant Science, 89* , 141-167.

Gaines, T. A., Barker, A. L., Patterson, E. L., Westra, P., Westra, E. P., Wilson, R. G., . . . Kniss, A. R. (2016). *EPSPS* gene copy number and whole-plant glyphosate resistance level in *Kochia scoparia* . *PLOS One, 11* , e0168295.

Gaines, T. A., Duke, S. O., Morran, S., Rigon, C. A. G., Tranel, P. J., Kupper, A., & Dayan, F. E. (2020). Mechanisms of evolved herbicide resistance. *Journal of Biological Chemistry, 295* , 10307-10330.

Gaines, T. A., Patterson, E. L., & Neve, P. (2019). Molecular mechanisms of adaptive evolution revealed by global selection for glyphosate resistance. *New Phytologist, 223* , 1770-1775.

Godar, A. S., Stahlman, P. W., Jugulam, M., & Dille, J. A. (2015). Glyphosate-resistant kochia (*Kochia scoparia*) in Kansas: EPSPS gene copy number in relation to resistance levels. *Weed Science, 63* , 587-595.

Guttieri, M. J., Eberlein, C. V., & Thill, D. C. (1995). Diverse mutations in the acetolactate synthase gene confer chlorsulfuron resistance in kochia (*Kochia scoparia*) biotypes. *Weed Science, 43* , 175-178.

Heap, I. (2020). The international survey of herbicide resistant weeds. Available on-line: www.weedscience.com. Accessed April 22, 2020.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23* , 1801-1806.

Jugulam, M., Niehues, K., Godar, A. S., Koo, D.-H., Danilova, T., Friebe, B., . . . Gill, B. S. (2014). Tandem amplification of a chromosomal segment harboring EPSPS locus confers glyphosate resistance in *Kochia scoparia* . *Plant Physiology, 166* , 1200-1207.

Kamvar, Z. N., Brooks, J. C., & Grunwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in genetics, 6* , 208.

Kamvar, Z. N., Tabima, J. F., & Grunwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ, 2* , e281.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodohl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution, 4* , 782-788.

Koo, D.-H., Molin, W. T., Saski, C. A., Jiang, J., Putta, K., Jugulam, M., . . . Gill, B. S. (2018). Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed *Amaranthus palmeri* . *Proceedings of the National Academy of Sciences, 115* , 3332-3337.

Kreiner, J. M., Giacomini, D. A., Bemm, F., Waithaka, B., Regalado, J., Lanz, C., . . . Wright, S. I. (2019). Multiple modes of convergent adaptation in the spread of glyphosate-resistant *Amaranthus tuberculatus* . *Proceedings of the National Academy of Sciences USA, 116* , 21076-21084.

Kuester, A., Chang, S.-M., & Baucom, R. S. (2015). The geographic mosaic of herbicide resistance evolution in the common morning glory, *Ipomoea purpurea* : Evidence for resistance hotspots and low genetic differentiation across the landscape. *Evolutionary Applications, 8* , 821-833.

Kumar, V., Felix, J., Morishita, D., & Jha, P. (2018). Confirmation of glyphosate-resistant kochia (*Kochia scoparia*) from sugar beet fields in Idaho and Oregon. *Weed Technology, 32* , 27-33.

Kumar, V., & Jha, P. (2015). Growth and reproduction of glyphosate-resistant and susceptible populations of *Kochia scoparia* . *PLOS One, 10* , e0142675.

Kumar, V., Jha, P., Giacomini, D., Westra, E. P., & Westra, P. (2015). Molecular basis of evolved resistance to glyphosate and acetolactate synthase-inhibitor herbicides in kochia (*Kochia scoparia*) accessions from Montana. *Weed Science, 63* , 758-769.

Kumar, V., Jha, P., Jugulam, M., Yadav, R., & Stahlman, P. W. (2019). Herbicide-resistant kochia (*Bassia scoparia*) in North America: A review. *Weed Science, 67* , 4-15.

Kumar, V., Jha, P., & Reichard, N. (2014). Occurrence and characterization of kochia (*Kochia scoparia*) accessions with resistance to glyphosate in Montana. *Weed Technology, 28* , 122-130.

Kupper, A., Manmathan, H. K., Giacomini, D., Patterson, E. L., McCloskey, W. B., & Gaines, T. A. (2018). Population genetic structure in glyphosate-resistant and-susceptible Palmer amaranth (*Amaranthus palmeri*) populations using genotyping-by-sequencing (GBS). *Frontiers in Plant Science, 9* , 29.

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research, 47* , W256-W259.

Martin, S. L., Benedict, L., Sauder, C. A., Wei, W., da Costa, L. O., Hall, L. M., & Beckie, H. J. (2017). Glyphosate resistance reduces kochia fitness: Comparison of segregating resistant and susceptible F2 populations. *Plant Science, 261* , 69-79.

Martin, S. L., Parent, J.-S., Laforest, M., Page, E., Kreiner, J. M., & James, T. (2019). Population genomic approaches for weed science.*Plants, 8* , 354.

Mengistu, L. W., & Messersmith, C. G. (2002). Genetic diversity of kochia. *Weed Science, 50* , 498-503.

Minati, M. H., Preston, C., & Malone, J. (2020). Genetic diversity and spread of glyphosate-resistant flaxleaf fleabane. *Bulletin of the National Research Centre, 44* , 24.

Molin, W. T., Patterson, E. L., & Saski, C. A. (2020a). Homogeneity among glyphosate-resistant *Amaranthus palmeri* in geographically distant locations. *PLOS One, 15* , e0233813.

Molin, W. T., Wright, A. A., Lawton-Rauh, A., & Saski, C. A. (2017). The unique genomic landscape surrounding the *EPSPS* gene in glyphosate resistant *Amaranthus palmeri* : a repetitive path to resistance. *Bmc Genomics, 18* , 91.

Molin, W. T., Wright, A. A., VanGessel, M. J., McCloskey, W. B., Jugulam, M., & Hoagland, R. E. (2018). 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 United States. *Pest Management Science, 74* , 1109-1117.

Molin, W. T., Yaguchi, A., Blenner, M. A., & Saski, C. A. (2020b). The eccDNA Replicon: A heritable, extra-nuclear vehicle that enables gene amplification and glyphosate resistance in *Amaranthus palmeri* .*The Plant Cell, 32* , 2132-2140.

Okada, M., Hanson, B. D., Hembree, K. J., Peng, Y., Shrestha, A., Stewart, C. N., Jr., . . . Jasieniuk, M. (2013). Evolution and spread of glyphosate resistance in *Conyza canadensis* in California.*Evolutionary Applications, 6* , 761-777.

Osipitan, O. A., & Dille, J. A. (2017). Fitness outcomes related to glyphosate resistance in kochia (*Kochia scoparia*): What life history stage to examine? *Frontiers in Plant Science, 8* , 1090.

Patterson, E. L., Pettinga, D. J., Ravet, K., Neve, P., & Gaines, T. A. (2018). Glyphosate resistance and *EPSPS* gene duplication: Convergent evolution in multiple plant species. *Journal of Heredity, 109* , 117-125.

Patterson, E. L., Saski, C. A., Sloan, D. B., Tranel, P. J., Westra, P., & Gaines, T. A. (2019). The draft genome of *Kochia scoparia* and the nechanism of glyphosate resistance via transposon-mediated*EPSPS* tandem gene duplication. *Genome Biology and Evolution, 11* , 2927–2940.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155* , 945-959.

R. (2019). R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project/org/.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes, 4* , 137-138.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular ecology resources, 8* , 103-106.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C_T method. *Nature Protocols, 3* , 1101-1108.

Stallings, G. P., Thill, D. C., Mallory-Smith, C. A., & Shafii, B. (1995). Pollen-mediated gene flow of sulfonylurea-resistant kochia (*Kochia scoparia*). *Weed Science, 43* , 95-102.

Van Etten, M., Lee, K. M., Chang, S.-M., & Baucom, R. S. (2020). Parallel and nonparallel genomic responses contribute to herbicide resistance in *Ipomoea purpurea* , a common agricultural weed.*PLoS Genetics, 16* , e1008593.

Waite, J., Thompson, C. R., Peterson, D. E., Currie, R. S., Olson, B. L. S., Stahlman, P. W., & Al-Khatib, K. (2013). Differential kochia (*Kochia scoparia*) populations response to glyphosate. *Weed Science, 61* , 193-200.

Waples, R. S. (2015). Testing for Hardy–Weinberg proportions: have we lost the plot? *Journal of Heredity, 106* , 1-19.

Westra, E. P., Nissen, S. J., Getts, T. J., Westra, P., & Gaines, T. A. (2019). Survey reveals frequency of multiple resistance to glyphosate and dicamba in kochia (*Bassia scoparia*). *Weed Technology, 33* , 664-672.

Wiersma, A. T., Gaines, T. A., Preston, C., Hamilton, J. P., Giacomini, D., Buell, C. R., . . . Westra, P. (2015). Gene amplification of 5-enol-pyruvylshikimate-3-phosphate synthase in glyphosate-resistant*Kochia scoparia* . *Planta, 241* , 463-474.

Data Accessibility Statement:

SSR genotypic data and R scripts used for population genetics analysis have been submitted to the digital repository Mountain Scholar, DOI will be provided.

Author contributions:

KR, CS, AD, PN, EP, PW, and TG designed research; KR, CD, AD, AK, EW, DP, PT, JF, DW, PJ, AK, PS, and EP performed research; KR, CS, AD, AK, EP, AK, and TG analyzed research; KR, CS, AD, AK, EP, PW, and TG wrote the paper; all authors contributed to editing and approved the final version of the paper.

Tables

Table 1. List of kochia (*Bassia scoparia*) populations used in the SSR study of population genetics. Table summarizing name, origin and year of sampling for 44 populations. n corresponds to the number of plants (individuals) sampled for leaf tissue from each population. Glyphosate resistance (R) corresponds to populations with at least one individual surviving treatment with glyphosate at 840 g a.e. ha^{-1}. Populations with no survival were considered susceptible (S).

Population	Country	State/Province	City or County	Year	n (plant)	Resistance
CO1R	USA	Colorado	Akron*	2012	18	R
CO2R	USA	Colorado	Brush	2012	18	R
CO3R	USA	Colorado	Cope	2012	18	R
CO4R	USA	Colorado	Julesburg	2011	18	R
CO5R	USA	Colorado	Kit Carson	2013	18	R
CO6R	USA	Colorado	Otis	2012	18	R
CO7R	USA	Colorado	Strasburg	2012	18	R
CO8R	USA	Colorado	Strasburg	2014	18	R
ID1R	USA	Idaho	Ada	2014	9	R
ID2R	USA	Idaho	Ada	2014	9	R
KS1S	USA	Kansas	Barton	2012	9	S^+
KS2S	USA	Kansas	Finney	2012	9	S^+
KS3R	USA	Kansas	Gray	2012	9	R
KS4R	USA	Kansas	Greeley	2012	9	R
KS5S	USA	Kansas	Meade	2012	9	S^+
KS6S	USA	Kansas	Ness	2012	9	S^+
KS7S	USA	Kansas	Philip	2012	9	S
KS8S	USA	Kansas	Pratt	2012	9	S^+
KS9R	USA	Kansas	Scott	2012	9	R
KS10R	USA	Kansas	Scott	2012	9	R

Population	Country	State/Province	City or County	Year	n (plant)	Resistance
KS11R	USA	Kansas	Stevens	2012	9	R
KS12R	USA	Kansas	Thomas	2012	9	R
KS13R	USA	Kansas	Thomas	2007	9	R
KS14R	USA	Kansas	Wallace	2012	9	R
KS15R	USA	Kansas	Wichita	2012	9	R
MT1R	USA	Montana	Chester	2012	9	R
MT2R	USA	Montana	Gilford	2012	9	R
MT3R	USA	Montana	Joplin	2012	9	R
OK1R	USA	Oklahoma	Cimarron	2012	9	R
OR1R	USA	Oregon	Malheur	2014	9	R
OR2R	USA	Oregon	Malheur	2014	9	R
OR3R	USA	Oregon	Malheur	2014	9	R
OR4R	USA	Oregon	Malheur	2014	9	R
OR5R	USA	Oregon	Malheur	2014	9	R
OR6R	USA	Oregon	Malheur	2014	9	R
OR7R	USA	Oregon	Malheur	2014	9	R
OR9S	USA	Oregon	Malheur	2015	9	S
TX1R	USA	Texas	Hartley		12	R
TX2R	USA	Texas	Hartley		18	R
TX3R	USA	Texas	Hartley		18	R
TX4R	USA	Texas	Hartley		18	R
TX5R	USA	Texas	Hartley		18	R
WY1R	USA	Wyoming	Powell	2015	9	R
AB1R	CANADA	Alberta			18	R

* Population M32, glyphosate-resistant line used to sequence EPSPS duplication region (Patterson et al., 2019; Westra et al., 2019).

+ Population suspected to be R when collected but all tested individuals had S phenotype; population could be heterogeneous with R at low frequency.

Table 2. List of SSR primers used for genotyping. For each locus, primer sequence and melting temperature and targeted repeat motif are provided. PCR amplicon expected size and annealing temperature used for PCR is also indicated.

SSR	Forward primer		Reverse primer		Amplicon	Motif	A
name	sequence (5'-3')	Tm	sequence (5'-3')	Tm	size (bp)	Repeat	T
162	TGATGTGAAAAGAACACCCC	58.4	TGTGATTCCAGGGAGGAGTA	58.1	216	(ATTTG)n	6
1225	GGTCCCAATGACAAACAGTC	57.8	GTTGGGTTTGGTTCTTGTTG	58.0	183	(CCCAA)n	6
1792	AACTAGTCGGATCGAGCCTT	58.0	AATCACACAACTCCGCAAGT	58.2	174	(CCCAA)n	5
2656	AACCAAACCGCACTAAACTG	57.8	GCACAATAGAGAGGGCAAAA	58.0	277	(TGGTT)n	6
2895	GTCATAGCCATCCCTTACCC	58.3	TATTGCCCTGTTCTTCAGGA	58.3	267	(AGTTC)n	6
2916	GTGCCAAAACCAAAGTTGTC	58.1	CCTCTCAACACAGGTTGCTT	57.9	215	(ATTTT)n	6
3332	CATGTACCTCGTGCAATGAA	58.1	TTTAGCTTAGCAATCACGGG	58.1	203	(TGTTG)n	5
5417	AGTGTGCTAAGAATTTGGGC	57.0	ACCATCAATTGTGATCGGAG	58.4	203	(GATAT)n	6
5608	GAGGCAAAGGATAAGGTGGT	58.1	ACGAAGGGAAGAGAAAGGAA	58.0	249	(AGGGA)n	5
5726	GCAGCCAAGCCATTCTATTA	58.0	AGCCCTTCCATGGAGAATTT	59.9	223	(TTATT)n	6
8376	ATGGAGCTGAACTGAACCAA	58.3	TTGTACCAGAATGCCTGTCA	57.7	254	(CTGAA)n	6

Table 3. **Descriptive summaries of 44 populations of kochia (*Bassia scoparia*) genotyped at 10 SSR loci.** Number of individuals genotyped per population, percentage of missing data averaged across loci, number of alleles observed summed across all loci, percentage of the total alleles (n = 70) observed in each population averaged across loci, allelic richness averaged across all loci; H_O, observed heterozygosity (H_o), expected heterozygosity (H_e), and $F_{IS}$and 95% confidence interval (CI).

Population	Number Of individuals genotyped	Percentage missing (%)	Number of alleles observed	Percentage of total alleles observed (%)	Allelic richness	Ho	He	F_{IS} (95% CI)
AB1R	17	4.71	32	48.95	2.22	0.26	0.44	0.39 (0.25 – 0.53)
CO1R	18	3.33	31	45.39	2.15	0.22	0.40	0.45 (0.35 – 0.56)
CO2R	18	2.78	34	50.24	2.52	0.41	0.48	0.16 (0.02 – 0.29)
CO3R	18	1.67	30	47.53	2.01	0.27	0.32	0.14 (-0.04 – 0.32)
CO4R	18	4.44	38	54.94	2.55	0.32	0.50	0.36 (0.23 – 0.48)
CO5R	18	6.11	31	46.02	2.04	0.30	0.37	0.18 (0.02 – 0.32)
CO6R	18	1.11	32	48.03	2.11	0.33	0.37	0.12 (-0.03 – 0.25)
CO7R	18	1.67	38	57.73	2.61	0.44	0.51	0.14 (0.00 – 0.26)
CO8R	18	2.78	32	49.18	2.30	0.41	0.46	0.11 (-0.02 – 0.23)
ID1R	8	0	28	45.61	2.30	0.29	0.42	0.31 (0.07 – 0.49)
ID2R	7	0	20	33.27	1.71	0.17	0.27	0.37 (-0.04 – 0.76)
KS10R	9	0	29	40.27	2.38	0.46	0.45	0.00 (-0.26 – 0.20)
KS11R	9	1.11	32	49.10	2.34	0.33	0.44	0.25 (0.09 – 0.35)

Population	Number Of individuals genotyped	Percentage missing (%)	Number of alleles observed	Percentage of total alleles observed (%)	Allelic richness	Ho	He	F_{IS} (95% CI)
KS12R	9	0	29	43.13	2.37	0.38	0.43	0.12 (-0.07 – 0.27)
KS13R	9	2.22	17	24.74	1.42	0.14	0.13	-0.04 (-0.28 – 0.13)
KS14R	9	2.22	30	42.90	2.21	0.25	0.38	0.35 (0.13 – 0.52)
KS15R	9	5.56	33	48.42	2.51	0.36	0.46	0.22 (0.00 – 0.40)
KS1S	9	1.11	31	47.92	2.44	0.42	0.49	0.14 (-0.08 – 0.32)
KS2S*	5	18.00	18	29.02	1.56	0.18	0.25	0.29 (-0.06 – 0.54)
KS3R	9	0	29	44.46	2.35	0.31	0.51	0.39 (0.17 – 0.56)
KS4R	9	4.44	23	33.93	1.98	0.34	0.37	0.06 (-0.23 – 0.30)
KS5S	9	4.44	28	42.70	2.34	0.34	0.45	0.26 (0.06 – 0.40)
KS6S	9	4.44	32	48.45	2.42	0.33	0.47	0.29 (0.04 – 0.51)
KS7S	9	1.11	26	38.61	2.16	0.32	0.44	0.25 (-0.04 – 0.46)
KS8S*	8	8.75	23	33.61	1.91	0.20	0.36	0.45 (0.14 – 0.75)
KS9R	9	4.44	30	43.88	2.37	0.33	0.46	0.29 (0.02 – 0.48)
MT1R	9	2.22	20	31.46	1.60	0.18	0.19	0.05 (-0.24 – 0.33)
MT2R	9	5.56	26	39.50	2.09	0.31	0.39	0.20 (0.04 – 0.30)

Population	Number Of individuals genotyped	Percentage missing (%)	Number of alleles observed	Percentage of total alleles observed (%)	Allelic richness	Ho	He	F_{IS} (95% CI)
MT3R	9	0	27	39.65	2.11	0.16	0.37	0.58 (0.33 – 0.79)
OK1R	9	3.33	24	36.87	2.00	0.24	0.37	0.33 (0.18 – 0.48)
OR1R	9	0	21	34.87	1.82	0.18	0.39	0.54 (0.21 – 0.80)
OR2R	9	0	36	40.06	2.04	0.21	0.34	0.37 (0.18 – 0.50)
OR3R	9	4.44	25	36.91	1.96	0.20	0.31	0.37 (0.10 – 0.56)
OR4R	9	2.22	19	31.88	1.58	0.14	0.20	0.29 (-0.30 – 0.61)
OR5R	9	3.33	20	32.68	1.68	0.20	0.28	0.29 (0.03 – 0.46)
OR6R	9	1.11	26	39.87	1.99	0.20	0.32	0.37 (0.16 – 0.51)
OR7R	9	0	25	41.55	2.03	0.20	0.40	0.50 (0.25 – 0.75)
OR9S	8	1.25	25	40.44	2.16	0.31	0.41	0.26 (0.06 – 0.42)
TX1R	12	0.83	34	46.33	2.43	0.33	0.41	0.21 (0.00 – 0.40)
TX2R	18	3.33	31	45.45	2.14	0.27	0.37	0.27 (0.11 – 0.43)
TX3R	18	1.67	30	46.70	2.16	0.31	0.39	0.21 (0.07 – 0.32)
TX4R	18	1.67	34	50.86	2.28	0.32	0.44	0.27 (0.12 – 0.40)
TX5R	16	2.50	31	46.67	2.14	0.40	0.41	0.01 (-0.16 – 0.19)

Population	Number Of individuals genotyped	Percentage missing (%)	Number of alleles observed	Percentage of total alleles observed (%)	Allelic richness	Ho	He	F_{IS} (95% CI)
WY1R	9	3.33	32	46.16	2.56	0.36	0.52	0.31 (0.11 – 0.47)

*Populations KS2S (n = 4) and KS8S (n = 1) had individuals removed due to data missing >20%

Figure Legends

Figure 1. Map of three EPSPS haplotypes in glyphosate-resistant kochia (*Bassia scoparia*). 1) increased *EPSPS* copy number, presence of type I and II repeats, MGE [?] 10 (Central Great Plains); 2) increased *EPSPS* copy number, no Type I or II, MGE [?] 10 (Northern Plains); 3) increased *EPSPS* copy number, no Type I or II, MGE < 10 (Pacific Northwest, some Wyoming). Figure legend created using BioRender.

Figure 2. Population level phylogeny tree in glyphosate-resistant kochia (*Bassia scoparia*). Forty-four kochia populations are included in this analysis. The neighbor-joining tree is based on Prevosti's distance. Bootstrap values (%) are shown if >20% and are based on 1000 replicates. Glyphosate-resistant, (R) and -susceptible (S); CGP1, Central Great Plains (green square with diagonal lines); NP2, Northern Plains (yellow square with vertical lines); PNW3, Pacific Northwest (blue square with horizontal lines).

Figure 3. Bayesian clustering analysis (STRUCTURE) in glyphosate-resistant kochia (*Bassia scoparia*). Assignment of 509 kochia individuals from 44 populations to the K=3 genetic clusters inferred by analysis. Populations are sorted by region (Central Great Plains, Northern Plains, and Pacific Northwest) and alphabetized within region. Each horizontal bar corresponds to a distinct individual and its probability of assignment, q , to each cluster.