The role of population and quantitative genetics and modern sequencing technologies to understand evolved herbicide resistance and weed fitness

Ramon G Leon, Jeffrey C Dunne and Fred Gould

Abstract

Evolution of resistance to multiple herbicides with different sites of action and of nontarget site resistance (NTSR) often involves multiple genes. Thus, single-gene analyses, typical in studies of target site resistance, are not sufficient for understanding the genetic architecture and dynamics of NTSR and multiple resistance. The genetics of weed adaptation to varied agricultural environments is also generally expected to be polygenic. Recent advances in whole-genome sequencing as well as bioinformatic and statistical tools have made it possible to use population and quantitative genetics methods to expand our understanding of how resistance and other traits important for weed adaptation are genetically controlled at the individual and population levels, and to predict responses to selection pressure by herbicides and other environmental factors. The use of tools such as quantitative trait loci mapping, genome-wide association studies, and genomic prediction will allow pest management scientists to better explain how pests adapt to control tools and how specific genotypes thrive and spread across agroecosystems and other human-disturbed systems. The challenge will be to use this knowledge in developing integrated weed management systems that inhibit broad resistance to current and future weed-control methods.

INTRODUCTION

Herbicide resistance challenges

Weeds are one of the most pervasive and constant threats to food security, and also one of the most challenging to address due to limited control tools and widespread herbicide resistance (HR). This problem has recently become particularly serious because there are weed species that have evolved resistance to herbicides affecting almost all sites of action, and herbicides with new sites of action have not been registered in almost 30 years. During the second half of the twentieth century, most of the reported resistance cases were the result of target site mutations that reduced the ability of the herbicide to prevent critical biochemical reactions by reducing the affinity of enzyme and protein binding sites to the herbicide. In recent years, there has been a dramatic change in HR, with more species exhibiting multiple resistance (up to seven sites of action), which is often the result of the action of several genes controlling different target sites or conferring nontarget site resistance (NTSR). NTSR can result from single or multiple mechanisms, which can include increased metabolic degradation of the herbicide (frequently mediated by cytochrome P450 or glutathione-S-transferases), herbicide sequestration to the vacuole, reduced uptake and/or translocation, gene amplification, and changes in gene expression. NTSR is also commonly expressed as a quantitative trait (i.e. polygenic), and in many cases the genetic variation conferring NTSR is already present in the population, so selection can increase resistance levels without the need for new mutations. Identifying the genetic mechanisms controlling NTSR, particularly those that provide resistance to herbicides with different mechanisms of action, is more challenging than for target site resistance, but this is critical for developing predictive and diagnostic systems before the NTSR alleles become predominant within and across populations.

The literature has abundant cases of target site resistance, while fewer studies on NTSR have been published. This is due to the fact that single genes (i.e. monogenic) commonly associated with target site resistance often cause large phenotypic differences...
Fitness in absence of the herbicide

When the herbicide is absent, resistant and wild type individuals compete with each other.20,21 If weeds with and without the HR allele are equally fit under this scenario (including individuals that are homozygous and heterozygous for the HR allele), it is said that the HR alleles do not have a ‘fitness cost’. The fitness of all three genotypes is then taken as 1.0 (all else in the genome being identical). If there is a fitness cost, then individuals with HR alleles will have fitness less than 1.0 relative to the competing individuals without the HR allele. Gressel (2002) defines this as Haldanian fitness.20

Haldane recognized that during an episode of selection a new allele could be increasing in frequency even if it had a fitness cost in the absence of the selection event, and that this may select for alleles at other loci that would ameliorate this problem. If the alleles at the other loci increased the relative fitness of the weeds with HR alleles more than those without the HR alleles, this would result in an epistatic interaction between alleles at the different loci.22,23 It is possible that when the weeds with HR alleles first begin to increase, their fitness relative to weeds without the HR alleles is less than 1.0 in the absence of the herbicide, but once the other epistatically interacting alleles increase in frequency, their presence in the genome of weeds with HR alleles ameliorates or compensates for the negative effects of the HR alleles to the point where plants with and without HR alleles have equal fitness.

Genetic knowledge of the basis for overall weed fitness in the presence and absence of herbicide is important for understanding weed success in agroecosystems and the evolution of HR. Frequently, once HR has been confirmed, studies are conducted to determine whether there are any fitness costs to the resistant individuals, but these studies are typically conducted in the absence of competition with other weed genotypes under field conditions. There are several possible explanations for the limited number of reports of fitness costs associated with HR. The simplest is that most existing HR mutations cause no easily measurable fitness costs. But let us consider the possibility that there might be fitness costs,6 although those have not been detected due to the use of inappropriate methods.24 Also, evolutionary processes that allow mutations in other parts of the genome might compensate for the penalties caused by the resistance allele through epistatic interactions.23,25

Compensatory mechanisms have only rarely been studied, and the case of Allopecurus myosuroides populations carrying the Gly2078 mutation conferring resistance to acetyl-CoA-carboxylase inhibiting herbicides is noteworthy. This mutation causes reduced vegetative growth and reproduction when present in lines that did not evolve resistance naturally (i.e. controlled crosses to wild-type sensitive biotypes). Conversely, Gly2078 was associated with higher fitness under field conditions when present in populations that evolved resistance when compared with wild populations without this mutation.26 In insect pests, where resistance genetics has been an important focus,27 there is only one clearly documented case of modifier alleles evolving and decreasing the cost of insecticide resistance.28

Fitness is often studied as a quantitative trait (e.g. biomass production, reproductive biomass, seed number), while HR is commonly considered a qualitative/categorical trait (e.g. resistant versus susceptible). Even when using quantitative indicators such as the herbicide rate that causes 50% growth reduction (i.e. GR50), the number of lines studied is so small that the ultimate analysis is categorical, and often limited in the statistical power of detection.
For target site resistance, the interaction between the strong directional selection exerted by the herbicide, and mutation rate, gene flow, and genetic drift is the driving mechanism that explains the potential rapid changes in the frequency and even fixation of the HR allele in the population, while the overall fitness of the population in a given environment is likely the result of a less dramatic balancing selection mechanism maintaining intermediate frequencies of alleles of multiple genes affecting fitness.29,30 Thus, changes in fitness could be subtler than the arbitrary HR categories, so just comparing means and ignoring variance components or changes in population distribution might restrict the ability to identify those changes.

Some HR studies do provide quantitative measurements of single or multiple traits related to fitness, but they are insufficient to characterize fitness in population genetics terms because they do not consider demographic changes of the genotypes over time, especially when using genetic lines that are not representative of population variability.17 Weed fitness in a specific environment with or without herbicide use is a complex polygenic trait including multiple physiological processes that could be redundant and/or compensatory. A single (nonlethal) novel mutation is unlikely to dramatically affect the Haldanian fitness (sensu Gressel) of the population if enough time is allowed for adaptation to occur from standing allelic variation and recombination events generating novel alleles.31

Beyond herbicides, response to fluctuating selection pressures (in direction and strength) due to varying environmental factors and agricultural practices not only can cause changes from year to year in the frequency of alleles that might affect weed fitness in the field, they also may maintain intrapopulation genetic variation.32 For example, in a fallow year, alleles favoring growth under low fertility, as well as physically linked alleles controlling other traits present in the same individuals, will increase in frequency. Conversely, in years with intensive fertilization, alleles favoring growth under these nutrient-rich conditions will increase in frequency, while those for low fertility will decrease.33

Intensive directional selection, such as that exerted by herbicides, should drastically reduce genetic variation within the population (i.e. in the absence of gene flow), fixing the alleles more important for survival. However, in the short term at the onset of selection, genetic variation could increase if many low frequency alleles are selected to intermediate frequencies.34 Thus, as long as heterozygosity exists at adaptive loci favoring fitness, only a few genotypes are needed in the population to maintain enough genetic diversity to respond to selection and favor adaptation. Understanding this dynamic balance between the directional selection of a specific trait (e.g. HR) and a more complex set of polygenic traits influencing overall fitness in an agricultural landscape should be a key component for the design of integrated weed management systems intended to be effective for a long time while preventing/delaying weed evolution. This is especially important considering that existing allelic variation in the population might be a more important source of adaptation than new mutations, including for HR.9

Evolution of not only NTSR but also compensatory mechanisms is likely to occur in cross-pollinating species that can, under selection pressure, combine the necessary alleles from the multiple genes controlling those mechanisms.3 An example of a weed adapting to multiple environmental factors besides evolving HR is Amaranthus palmeri. This weed species exhibits very little genetic differentiation among populations due to extensive gene flow, so, in theory, all populations should behave similarly, unless there is differential selection due to local environmental conditions in those populations. It is interesting that glyphosate-resistant populations in the southeastern USA tended to be taller and produce more biomass than glyphosate-susceptible populations under controlled and field conditions. This differential growth was not the result of pleiotropic effects of the HR allele (S), but instead was due to genetic factors controlling plant size and also partially explained by differences in nutrient use efficiency. Higher nutrient use efficiency and taller plants were found in populations collected from areas with a history of high fertilization rates as well as production of high-statured crops, but this was not unique to glyphosate resistant populations.33,36 Therefore, a population that is able to fix both the HR allele as well as the high nutrient use efficiency and size because of multiple selection forces occurring simultaneously will likely have higher fitness than populations only carrying the HR allele. Alopecurus myosuroides has also been reported exhibiting inconsistent fitness costs resulting from NTSR (i.e. enhanced metabolism). Although resistant individuals produced 27% less seed than susceptible individuals under low fertility conditions, the former competed and produced the same seed amount as the latter when growing in competition with wheat (Triticum aestivum). This illustrates how genetic changes in other parts of the genome can impact how well HR populations will be able to compete within the crop. In other words, the effect on fitness of the HR allele on individual plants provides only a limited view of how the HR allele will affect the fitness of the population. Therefore, weed fitness research must always consider physiological processes that depend not only on the genetic makeup of a resistant individual, but also on the dynamics at the population level where genetic diversity, gene flow, reproduction strategies, dominance, and epistatic interactions can play major roles in determining changes in HR, fitness, and pest behavior. Evolutionary and quantitative genetic approaches are great tools to conduct such research.37

NEW APPROACHES TO SEE THE BIG PICTURE
Learning from breeders and evolutionary biologists

Studying quantitative/polygenic traits has been a major focus in animal and plant breeding, and has historically relied heavily on statistical approaches. More recently, the development of molecular markers and especially high-throughput next-generation sequencing technologies, as well as powerful bioinformatic techniques, has dramatically increased the ability to study variation in crop traits and to map the loci that underpin that variation. These new techniques are now also being used to study evolutionary processes in natural populations. Evolutionary biologists and ecologists have been successful in combining population genetics typically used to study natural populations with quantitative genetics and genetic mapping used predominantly by breeders. This has enabled them to decipher how plants and animals have evolved adaptation to their environments based on ‘DNA sequence signatures’ of natural selection discernible from genomic sequences and nucleotide polymorphisms.38,39 Biologists who study invasive species have started to recognize the importance of paying more attention to the evolutionary characteristics that drive invasiveness.40 Weed scientists and pest specialists can and should use these tools to better describe resistance evolution as well as pest fitness and pest behavior.41 This could generate a more comprehensive and better informed understanding of how pests adapt to human disturbance.
The ‘population genetic uncertainty principle’

Most of the weed species that have evolved resistance to multiple herbicides are obligate outcrossers, exhibiting large within-population genetic diversity. Examples include Alopecurus myosuroides, Amaranthus palmeri, Amaranthus tuberculatus, Kochia scoparia, Lolium rigidum, and Papaver rhoesas. In such species, in every generation, populations will have individuals with different genetic composition. Even if a target site resistance allele has been fixed in a population, the individuals in this ‘resistant’ population will not have the same genetic background, and if there are other alleles interacting with alleles at the target site locus, the fitness of the resistant individuals might differ from one another in the presence and absence of herbicides. Abundant studies suggest that epistasis may play a key role in explaining differences in fitness among populations of the same species. In other words, the probability of having good combinations of alleles for multiple genes will determine the maximum fitness of the resistant individuals. That probability is affected by the frequency of alleles of epistatic, compensatory genes in the population, the size of the population, and selection pressure.17

Most of the HR studies do not report how seed was collected from the putative resistant population (e.g. how many plants, how far apart, how many seeds per plant, etc.). This information must always be provided, especially when studying polygenic resistance. Selecting only the more resistant individuals to generate experimental ‘homogenous lines’ to study the physiological and biochemical processes explaining resistance reduces our ability to properly characterize the importance of the HR allele and the actual fitness changes occurring in the population. This is especially true when epistatic interactions affect the impact of the HR allele on fitness.44

Studying genetic diversity and phenotypic variation can provide a more comprehensive view of the changes in genetic architecture (i.e. the characteristics of genetic sequence variation that are responsible for heritable phenotypic variability) caused by HR, but there might be limits to our ability to identify the gene(s) responsible for HR with high accuracy. This has been a long-standing conundrum between functional and evolutionary biologists, which is in some ways analogous to Heisenberg’s uncertainty principle of quantum mechanics, which supported the idea that probabilistic approaches were needed to describe combinations of particle characteristics such as position and momentum. This has been termed the ‘population genetic uncertainty principle’.48 To gain a more accurate understanding of the genetic basis for a population’s fitness, you would need to measure a large proportion of the genetic variation in the population. The closer you wanted to get to a precise understanding, the higher the number of measured individuals needed. To totally understand the genetic underpinnings of the fitness of a population, you would need to measure all individuals in a population, and in measuring them you would change the population.

One cannot afford to study every detail about the genetics of a resistant population, and one does not want to focus on a single locus and miss important knowledge. The reality is that we need both approaches complementing each other. We need a functional understanding of resistance with other adaptive traits, but we also need a clear understanding of how those traits and the frequency of their alleles vary geographically (e.g. within and among populations) and over time to design monitoring strategies and reduce the spread of HR alleles and individuals.

More than a single gene and a single genotype

To illustrate how frequencies of life-history alleles within and across populations can influence fitness and HR evolution, let us consider several scenarios. First, a population with no or very little genetic diversity will have individuals with the same genetic composition (Fig. 1(a),(b)). This can occur because it is either a clonal population or a population that recently went through a genetic bottleneck (e.g. recent introduction or strong selection pressure such as herbicides and only a few resistant individuals survived, so genetic diversity is dramatically and quickly reduced). This is the scenario presumed in many target site resistance-fitness studies, in which researchers expect to identify pleiotropic effects due to the HR allele with the underlying assumption that no other genes changed during the selection process.

The assumption of genetic uniformity has proven to reduce the ability of demographic models to properly predict invasiveness over extensive temporal and spatial scales because in order to maintain such a level of population genetic uniformity either...
Thus, speciﬁc adaptive alleles that provide a ﬁtness gain in the new environment created by the introduction of the herbicide (e.g. elongation within a crop canopy) or are physically linked to the HR allele, then selection can occur in the absence of gene flow (e.g. isolation by distance) or genetic drift (Fig. 1(d)). Even when an allele for that trait is selected and ﬁxed in the population, that does not mean that all individuals will be the same because they might differ in other traits or at other loci controlling the same trait (Fig. 2). For example, despite the strong selection exerted by herbicide applications on the obligate outcrossing weed Alopeceus myosuroides in winter cereal ﬁelds in France, the genome-wide variability determined with ampliﬁed fragment length polymorphisms (AFLPs) was not affected, and little to no differentiation in overall genetic diversity among populations was found.49 Therefore, speciﬁc alleles under selection might be present in individuals with different genetic backgrounds within the same population, and the ﬁtness of those individuals might be determined by epistatic interactions.44,45,50 Similarly, other selection forces can act on alleles of other genes, increasing the frequency of beneﬁcial alleles until the overall ﬁtness of the population changes.31

Genetically diverse populations with a few rare HR alleles will exhibit increased frequency of HR alleles in response to herbicide selection even if the HR allele confers a minor ﬁtness penalty (e.g. shorter plants, Fig. 2(a)). This type of selection can also occur for adaptive traits like plant height, such as weeds competing with taller crop canopies selecting for biotypes with taller plants (Fig. 2(b)).36 Because not all individuals have the same alleles due to existing genetic diversity, not all individuals carrying the HR allele exhibit the same ﬁtness, and there might be ‘harmonious combinations’ creating adaptive peaks.50 For example, if there are other adaptive alleles that provide a ﬁtness gain in the new environment created by the introduction of the herbicide (e.g. elongation within a crop canopy) or are physically linked to the HR allele, then their frequency will increase over time, changing the overall ﬁtness of the population (Fig. 2(c)). Furthermore, gene flow can nearby populations must have the same genetic composition in the presence of gene flow (Fig. 1(a); i.e. low within and among populations genetic diversity) or no gene flow exists between populations, resulting in different allelic frequencies (Fig. 1(b); i.e. low within and high among populations genetic diversity).

Another scenario involves populations with high within-population genetic diversity, meaning that individuals of the same population have different alleles (Fig. 1(c),(d)). If there is gene flow, then nearby populations will exhibit low among-populations genetic diversity, but high within-population diversity (Fig. 1(c)). This is the case of obligate outcrossers such as Alopecurus myosuroides. In this latter case, it is very difﬁcult to have a uniform population, and very intense selection is needed to ﬁx a speciﬁc trait (Fig. 1(c)).

For obligate outcrossers differentiation among populations can occur in the absence of gene flow (e.g. isolation by distance) or genetic drift (Fig. 1(d)). Even when an allele for that trait is selected and ﬁxed in the population, that does not mean that all individuals will be the same because they might differ in other traits or at other loci controlling the same trait (Fig. 2). For example, despite the strong selection exerted by herbicide applications on the obligate outcrossing weed Alopeceus myosuroides in winter cereal ﬁelds in France, the genome-wide variability determined with ampliﬁed fragment length polymorphisms (AFLPs) was not affected, and little to no differentiation in overall genetic diversity among populations was found.49 Thus, speciﬁc alleles under selection might be present in individuals with different genetic backgrounds within the same population, and the ﬁtness of those individuals might be determined by epistatic interactions.44,45,50 Similarly, other selection forces can act on alleles of other genes, increasing the frequency of beneﬁcial alleles until the overall ﬁtness of the population changes.31

Figure 2. Schematic representation of how selection for herbicide resistance (HR; blue circles) and other adaptive/ﬁtness (plant height; pink crescents) alleles modiﬁes the genetic diversity of a weed population and its ﬁtness (assuming taller plants are better competitors with the crop). Each column represents individual haploids within a population, and each symbol within columns represent alleles at separate loci. The height of the black columns represents the height of single plants. (a) The HR allele confers a small ﬁtness penalty (shorter plants), but it is rapidly ﬁxed in the population due to intensive herbicide selection decreasing the average height of the population. (b) A population varying in plant height increases the average height by the selection exerted by tall crops. (c) A population that ﬁrst ﬁxes the HR allele, decreasing the average height, but then selects for HR individuals carrying the tall plant allele by growing within tall crops, which increases the population height. Gene ﬂow between populations (a) and (b) can also promote this increased ﬁtness result.
favor the combination of HR and increased fitness alleles. This type of post-HR evolution genomic change has the potential to counter any fitness penalty resulting from the HR trait (Fig. 2).26

Change of mindset and methods
The acknowledgement that a resistance trait(s) might be quantitative and polygenic will potentially change many aspects of how the population is studied. For example, sampling must be done to characterize differences among individuals (e.g. avoid bulking seed). Ideally the kinship of the sampled individuals is determined (e.g. half-sibs, full-sibs), and the sample size is large enough to properly quantify the frequency of the alleles controlling the traits and potentially mapping them using molecular markers.13 When phenotyping for a specific trait(s), careful quantification of the variability among individuals is important, so an adequate sample size to accurately capture that variability of the population is critical. Furthermore, the stability of the phenotypes across environments (both locations and years) can help quantify what proportion of the phenotypic variance is due to the genetic variance. If narrow sense heritability could be determined, it might be possible to also estimate how much response can be expected based on selection intensity.15,51

Adaptive evolution is the result of phenotypic selection and response to selection. Phenotypic selection is determined by how trait values influence fitness, while the selection response depends on the additive genetic variances and covariances among traits.39,52 Thus, even before the gene(s) responsible for the HR trait can be identified, it is possible to project changes in fitness and HR. Comparing multiple populations is also a very important step to understand the fitness of the population and its potential to influence other populations through gene flow. In many HR studies, a single susceptible population is used as a control, which can create a bias in our interpretation of the effect of the HR trait on the fitness of the weed. The inclusion of multiple susceptible populations differing in genetic background is particularly important to provide the necessary genetic context to understand the dynamics that drive the potential success of a HR trait or a given biotype in a new location or system.

CHASING QTLs VERSUS SINGLE GENES
After characterizing within and among population diversity for the trait of interest, the researcher will have to decide whether to (i) focus on approaches that simplify the study of HR using biochemical approaches or (ii) dig deeper into the genetic architecture of the HR trait in order to map regions in the genome and even genes that are responsible. The first approach focuses on breeding strategies that generate lines that can be used for both mapping and biochemical studies, although the mapping accuracy might be limited. In this category, we find QTL analysis and mapping using biparental lines resulting from planned crosses from resistant and susceptible individuals. The second approach is based on genome association techniques that link phenotypic variation observed in segregating populations to single-nucleotide polymorphisms (SNPs) frequency based on statistical models. These association techniques are frequently used for fine mapping of the genes controlling the QTLs, but rely on existing genetic variability across individuals. Existing genetic variability can be an important source of resistance vis-à-vis nontarget site mechanisms that are triggered by stress response networks.53 Both approaches are valuable and can complement techniques such as proteomics, transcriptomics, and metabolomics (and other ‘omics’) to increase the probability of finding the factors that explain the HR trait. However, due to evolutionary changes in weeds, and the constant changes in selection pressures, it might be more cost-effective to start with association approaches using existing genetic variation than with predetermined crosses using homogeneous lines.

QTL analysis
The two main goals of QTL analysis and mapping are to find the physical region(s) in the genome where the loci controlling the QTL are located and to quantify how much of the phenotypic variation is due to each QTL. Many quantitative traits, such as NTSR, have been addressed by ‘omics’ approaches,5 but genotype by environment interactions can make it difficult to identify the genetic factors that are truly the main drivers of the expression of the trait. For this reason, it is critical to assess traits under different environments, and preferably using field experiments. In many cases, QTLs identified in the field are not detected in the laboratory or greenhouse and vice versa.38 Also, by determining the proportion between the genetic variation and the effects of environmental interactions with the QTL, it is possible to explain the stability of the genetic architecture affecting the trait across environments avoiding confounding metabolic variability.25 As done for characterizing single-gene herbicide resistance, QTL analysis requires generating a segregating population by crossing resistant (R) and susceptible (S) parents. Commonly, this is done by using either inbred or highly homozygous lines as parental lines,55 and then generating a segregating F1 and/or F2 population. Due to their high levels of heterozygosity in outcrossing species, F2 populations exhibit high levels of segregation, similar to F2 populations of self-pollinating species, so they can sometimes be used for mapping purposes. The F2 individuals are phenotyped under several environments and genotyped using molecular markers. For single-gene resistance, most studies measure phenotypes of a relatively small number of F2 individuals (e.g. n < 12). In contrast, QTL mapping typically requires more than 200 segregating F2 individuals. The analysis should include enough polymorphic markers to have good coverage of the genome, so QTLs can be mapped accurately. Today, whole-genome sequencing has made it possible to use thousands of SNPs as markers, which has increased our ability to have adequate coverage of the genome with a high resolution. Recombinant inbred lines (RILs) and near-isogenic lines (NILs), if the number of lines is large enough, provide ideal populations to conduct QTL analysis,38 but they require multiple inbreeding or backcross cycles to reach genetic stability before phenotyping, which makes these genetic populations difficult to generate for studying many weed species.

The problem with QTL-mapping is that it is intended to identify all of the loci with alleles that significantly impact the HR trait, but it has a high probability of missing the importance of HR alleles that have minor effects on the phenotype. Also, because QTL mapping is based on segregation of specific traits of interest, in many cases segregating populations must be generated for each trait of interest. If there are other secondary traits that segregate, those can be used for QTL mapping as well, although this is not always possible. Therefore, if relying on this technique, weed scientists must have a more structured strategy, as breeders and geneticists do, to continuously generate the germplasm needed to study HR traits as well as traits that influence weed fitness, but this will also require lines with large phenotypic differences that might not be readily available.
Several studies have identified QTLs associated with important adaptive traits such as flower morphology and timing, although the relationship of those QTLs to the heritability of the trait was not always strong.\(^{30,34,36}\) Three QTLs associated with traits favoring ‘weediness’ such as root growth ability and rhizome and tiller production were identified in *Sorghum halepense*.\(^{27}\) That work was conducted almost 25 years ago, and similar research has not been common, with a few exceptions such as the identification of QTLs of traits controlling clonality and hybrid vigor associated with ‘invasiveness’ in *Helianthus tuberosus*.\(^{58}\) Unfortunately, markers for those QTLs have not been used to study how differences in ‘weediness’ and fitness might be explained by allelic variation. In the case of NTSR, there is evidence that it is sometimes controlled by alleles of a few genes or major QTLs.\(^{5,9,59–61}\)

**GWAS**

The development of genome-wide association studies (GWAS) and genomic prediction have been great scientific advancements that have opened a wide range of possibilities to study the genetics of HR and other adaptive traits. GWAS uses large numbers of genotyped and phenotyped individuals to find correlations between the frequency of SNPs, which are used as markers, and the genetic variance of the trait with which they are hypothesized to be associated.\(^{62}\) The advantage of GWAS over QTL mapping is that the pedigree of the lines does not have to be known and many populations can be compared to characterize the allelic diversity controlling the expression of the trait. For this reason GWAS should be able to detect main effect QTLs across a range of genetic backgrounds.\(^{38,55,63}\)

The power of the detection of the precise location of the sequence of interest (e.g. resistance, plant height, seed dormancy) is limited by the recombination rate in the genomic regions of interest. If the recombination rate is high and there is rapid decay of linkage disequilibrium, the resolution provided by GWAS might improve the ability to accurately detect the regions responsible for trait variability. Therefore, because GWAS uses natural populations with much larger genetic diversity and historic recombination than biparental segregating populations, the resolution of mapping HR loci is higher than with QTL mapping, and in some cases almost to the gene level (given that the marker density is high enough).\(^{38,55}\) Furthermore, the same populations can be used to study multiple traits that might or might not be related. The cost of GWAS which requires large numbers of individuals might be problematic. Approaches that use pooled samples to scan the genome for changes in SNP or haplotype frequencies could offer a more affordable route.\(^{64}\)

GWAS has the limitation that it can produce false positives due to population structure artificially associating SNPs to the trait of interest.\(^{58}\) Thus, the association between SNPs and phenotypic variation can be affected such that biases might be introduced by rare variants with large phenotypic effects and common variants with very small effects on the phenotype.\(^{66,67}\) This problem can be significantly reduced by using the same SNPs to determine kinship and population structure to estimate how much of the phenotypic variability can be explained by those factors.\(^{58–70}\) Also, statistical approaches have been developed to detect correlations between environmental factors and allele frequencies driving adaptive responses.\(^{71}\) Mixed models can be used to identify artificial associations by considering the phenotypic covariance resulting from kinship or genetic relatedness.\(^{70}\)

GWAS was used to map QTLs involved in seed dormancy (an important weedy trait) in *Oryza sativa*, confirming that several of those QTLs contained, or were associated with, genes such as Gaox3 and Sdr4 that had been previously demonstrated as critical for seed dormancy.\(^{69}\) Furthermore, combining ecology-based studies with GWAS can help elucidate the evolutionary path and forces acting on resistance alleles. This has been done to track how polymorphisms of resistance genes to *Pseudomonas syringae* in natural populations of *Arabidopsis thaliana* are selected and their frequency affected by the action of other pathogen species.\(^{72}\) Another example of the applications of this technology to NTSR is how GWAS has been demonstrated to be a valuable tool to identify markers associated with metabolic pyrethroid resistance in the East African mosquito *Anopheles gambiae* resulting from existing genetic variation.\(^{11,72}\) This study illustrated how allelic variation outside the insecticide target site was important to determine the level and stability of the resistance phenotype.

Van Etten *et al.*,\(^{13}\) provide a recent example in which multiple genomic technologies were used to assess the basis for herbicide resistance. In their study of *Ipomoea purpurea* resistance to glyphosate, they found that target site resistance was not present. Instead a number of genomic regions had signatures of selection in resistant populations. Because overall genomic scans of resistant and susceptible populations did not show two distinct clusters, there was no support for a single origin of resistance that spread as a unique strain. The genomic regions of differentiation between susceptible and resistant populations included detoxification genes. Interestingly, one genomic region changed in parallel in all resistant populations while for other regions only some populations showed changes in haplotype frequencies. These populations vary in fitness costs associated with resistance, and the authors speculated that some of the genomic changes could compensate for these fitness costs. Overall, the study demonstrates the utility of genomic approaches and also emphasizes the potentially complex genetic architecture of herbicide resistance.

Genomic prediction is a new technique that has been rapidly adopted by plant breeders to use GWAS to predict the performance of crosses and progeny lines. Genomic prediction uses large phenotypic data sets in combination with detailed SNP data from many lines to train random models to predict breeding values.\(^{30,73}\) In other words, genomic prediction can be used to anticipate how genotype by environment interactions, recombination, dominance, and epistatic interactions may affect the performance of the progeny generated by a given cross without any formal, field-based phenotypic testing. In simple terms, genomic prediction assigns a proportion of the phenotypic variability to individual SNPs, and once the progeny is genotyped and the SNPs are identified, genomic prediction estimates the potential performance of each line for the trait of interest. Breeders have successfully used this technique in several crops to eliminate lines from evaluation and selection programs because they are not likely to meet the desired standards, increasing the efficiency of the breeding process. Also, they have used genomic prediction to predict the performance of lines in new environments and/or under different production practices.\(^{74}\)

Current HR evolution models are predominantly based on traits controlled by single dominant/codominant genes and use probabilistic methods to determine how soon selection pressure will increase the frequency of resistant alleles in a uniform background.\(^{75–77}\) However, there are no models that explain how the HR trait will behave in other populations and new environments where genetic diversity is different (i.e. higher) and genotype by environment interactions can greatly modify the phenotype of
individuals receiving HR alleles. As stated before, this is critical for NTSR.

The crop-breeding literature has demonstrated that genomic prediction loses accuracy when a training crop population is used to predict the responses of other populations with somewhat different genetic backgrounds. If the predictions from a training weed population are only useful for prediction in weed populations within a very local distance from the training population, then it will not be an effective weed resistance management tool. However, to date this tool has never been tested with a set of weed populations. As the expense of doing such work declines, it would certainly be worth testing the hypothesis that the predictions of this tool are of more than local value.

**TACKLING BIG QUESTIONS IN WEED BIOLOGY AND HR**

There are several big challenges that the weed scientific community has not been able to properly study, and that consequently have impeded our ability to manage weed populations efficiently while preventing adaptations that result in control tools becoming ineffective.

**Resistance**

As discussed above, HR and especially NTSR and multiple HR, are among the most immediate concerns that should be addressed using population and quantitative genetics. It can be argued that, at this point, HR management could be considered a moot point for several species because they have evolved resistance to multiple (almost all) sites of action registered for specific crops. However, the recent reliance on tank mixtures combining herbicides with different sites of action and herbicide tolerance crops with stacked traits might favor selection of NTSR to multiple sites of action. Furthermore, herbicide stewardship is still of great value for most weed species as well as situations where herbicides have not been used. In the eventuality that new sites of action are registered and because of changes in herbicide use due to the introduction of crops with resistance to existing sites of action, more detailed characterization of existing genetic variability in sensitivity to the new herbicide in the populations and a more aggressive monitoring of changes in tolerance will be needed to preserve the efficacy of the new herbicide(s). Expecting to find dramatic shifts and clear differences in sensitivity to the herbicide to recognize that a population has become resistant has been demonstrated to be a failed strategy with the serious consequence that most HR events have not been eradicated or contained on time.

**Adaptation to mechanical and cultural control**

In cropping systems with high incidence of HR, growers are already incorporating into their management programs more mechanical weed control as well as cultural practices that reduce weed interference and seed banks. Although properly documented examples of adaptations to herbicide control actions are few, ignoring that these types of adaptations exist could result in increased weediness and invasiveness. Traits such as plant height, flowering time, seed shattering, emergence time and pattern, root architecture, photosynthetic efficiency, and nutrient use efficiency could be affected by changes in cultivation, crop planting density and arrangement, fertilization strategy, harvest seed destruction, and cover cropping. A main risk of evolution to these adaptive traits is that they can broadly impact fitness and thereby change the niche of the weed. Although adaptations to mechanical and cultural control might occur by the action of a single gene (e.g. dwarfing), they are more likely to involve polygenic variation in traits. The rate of adaptation could be considerably slower compared to HR, especially if changes in multiple morphological and physiological traits were needed. For example, if a population evolves reduced internode length and become shorter in response to mowing, it might also evolve modified leaf angle and photosynthetic efficiency to optimize growth under more shade. In order to develop strategies to counter that selection process, a clear understanding of specific allelic frequency changes will be needed.

**Climate change and geographic distribution of weeds**

Another important consideration for understanding weediness is how climate change in combination with human dispersal can increase the geographic and ecological range in which a weed species can establish and develop populations, and how those factors could also help increase the economic importance of weed species that were not a serious problem in the past. Although many weed species appear to have long been present across wide latitudinal and altitudinal ranges, others have in a short time dramatically increased their niche and climatic range. For example, A. palmeri, which is native to the southwestern USA (i.e. Sonoran desert), has invaded northern latitudes, establishing populations as far north as Michigan and Canada. A pronounced latitudinal change like this one requires adaptations in traits such as seed overwintering and survival potential, emergence timing, and likely flowering timing as influenced by photoperiod. By characterizing the genetic diversity of multiple populations to key adaptive traits, it might be possible to anticipate potential changes for the environmental and climatic ranges to which a weed species will be adapted after selection. Genomic prediction is a key and effective tool for this purpose, especially when building the models on multienvironment phenotypic data because it does not require knowing either specific genes or QTLs a priori. Instead, it requires large phenotypic data and extensive in-depth sequencing data from the lines or populations studied. The former is constantly generated by weed scientists, but it will require more coordination among researchers. The latter data is easier to generate and process, and its cost has declined dramatically, so it has become a more accessible tool.

**Complex genetic traits and the future of molecular weed control**

In the last 10 years, there have been dramatic advances in molecular approaches to modify gene and protein activity in vivo that can provide new weed control options. Techniques such as RNAi, gene-drive, CRISPR-Cas9, and PROTAC are examples of tools that can be used for targeting weed metabolism and growth; all of these techniques require knowing in detail the molecular target and its function in the plant because SNPs can determine the specificity and activity of the tools. Therefore, the level of conservation of the target sequence (nucleotide or amino acid) in the population(s) is a key component for the design of those tools. If their use were to become widespread for weed management, it would be critical to know how population genetic diversity could affect control efficacy and also to avoid making the same mistakes that lead to target site resistance with herbicides. From what we know today, multiple targets with different mechanisms should be attacked simultaneously, with complementary as well as redundant functions, and strategies for managing adaptations in
the field should be considered to be also a core component of the technology. With the knowledge generated by use of the methods discussed in this article, much more targeted approaches could be developed for specific cropping/weed systems.

CONCLUSIONS

The rapid rates of adaptation by weeds, pest, and pathogens to control practices and to changes in the environment that we are observing must be a serious reminder that a unidimensional (i.e. single-gene) mindset has and will limit our ability to generate relevant scientific information to adequately understand and manage pests. Individual genetic and physiological processes do not necessarily provide an accurate representation of the dynamics that drive resistance evolution and changes in fitness in the field. This is certainly the case for those weed species that have demonstrated the fastest rates of evolution and (not coincidentally) are outcrossers with high levels of genetic diversity. Important advances in whole-genome sequencing technologies as well as the development of bioinformatic tools and new statistical methods have opened the possibility to study at the field, local, and regional levels how resistance and fitness shape the evolution of weed populations. By adopting these new tools, pest management scientists will generate better and more reliable knowledge about the rate of resistance evolution and resistance dispersal, and may be more likely to identify the genetic and biochemical factors responsible for resistance. Furthermore, it should be possible to identify genes playing important roles in weediness and fitness in agricultural and disturbed systems, which would transform our understanding of weeds, opening new alternatives for their management.

ACKNOWLEDGEMENTS

Support to R.G. Leon was partially provided by USDA-NIFA Grants 2017-6505-26807, 2018-70006-28933, 2019-68012-29818, and Hatch Project NC02653.

REFERENCES

1. Oerke EC, Crop losses to pests. J Agric Sci [Biol] 144:31–43 (2006).
2. Duke SO, Stithman MA and Dayan FE, A novel genomic approach to herbicide and herbicide mod of action discovery. Pest Manag Sci 75:314–317 (2019).
3. Deluy C, Jasiomik M and Le Corr V, Deciphering the evolution of herbicide resistance in weeds. Trends Genet 29:649–658 (2013).
4. Yuan JS, Tranal PJ and Stewart CN, Non-target-site herbicide resistance: a family business. Trends Plant Sci 12:6–13 (2007).
5. Heap I, The international herbicide-resistant weed database. Available: www.weedsociety.org (27 March 2020).
6. Kreiner JM, Stinchcombe JR and Wright SL, Population genomics of herbicide resistance: adaptation via evolutionary rescue. Annu Rev Plant Biol 69:611–635 (2018).
7. Cocker KM, Moss SR and Coleman JOD, Multiple mechanisms of resistance to fenoxaprop-P-ethyl in United Kingdom and other European populations of herbicide-resistant Alopecurus myosuroides (black-grass). Pestic Biochem Physiol 65:169–180 (1999).
8. Hall LM, Moss SR and Powles SB, Mechanism of resistance to chlorotoluron in two biotypes of the grass weed Alopecurus myosuroides. Pestic Biochem Physiol 53:180–192 (1995).
9. Busi R, Neve P and Powles S, Evolved polygenic herbicide resistance in Lolium rigidum by low-dose herbicide selection within standing genetic variation. Evol Appl 6:231–242 (2013).
10. Donnelly MJ, Isaacs A and Weetman D, Identification, validation, and application of molecular diagnostics for insecticide resistance in malaria vectors. Trends Parasitol 32:197–206 (2016).
11. Weetman D, Wilding CS, Neaffsey DE, Muller P, Ochomu E, Isaacs AT et al., Candidate-gene based GWAS identifies reproducible DNA markers for metabolic pyrethroid resistance from standing genetic variation in East African Anopheles gambiae. Sci Rep 8:2920 (2018).
12. Lucas ER, Rockett KA, Lynd A, Essandoh J, Grisales N, Kemie B et al., A high throughput multi-locus insecticide resistance marker panel for tracking resistance emergence and spread in Anopheles gambiae. Sci Rep 9:13335 (2019).
13. Van Etten M, Lee KM, Chang S-M and Baucoum RS, Parallel and nonparallel genomic responses contribute to herbicide resistance in Ipomoea purpurea, a common agricultural weed. PLoS Genet 16:e1008593 (2020).
14. Darmency H, Does genetic variability in weeds respond to non-chemical selection pressure in arable fields? Weed Res 59:260–264 (2019).
15. Darmency H and Uludag A, A quantitative genetic examination of non-target-site resistance applied to Avana species. Weed Res 58:69–75 (2018).
16. Falconer DS and Mackay TFC, Introduction to Quantitative Genetics, 4th edn. Pearson Education Limited, Harlow, Essex, UK (1996).
17. Silvertown JW and Charlesworth D, Introduction to Plant Population Biology, 4th edn. Blackwell Science, Ames, IA (2009).
18. Darwin C, The Origin of Species by Means of Natural Selection. Barnes and Noble, New York, NY (1859).
19. Gressel J, Molecular Biology of Weed Control. Taylor & Francis, London UK (2002).
20. Haldane JBS, The Causes of Evolution. Longman, Green, and Co., London, UK (1932).
21. Haldane JBS, The cost of natural selection. J Genet 55:511–524 (1957).
22. Smith JM, “Haldane’s dilemma” and the rate of evolution. Nature 219:114–116 (1968).
23. Cousens RD and Fournier-Level A, Herbicide resistance costs: what are we actually measuring and why? Pest Manag Sci 74:1539–1546 (2018).
24. Simmons MJ and Crow JF, Mutations affecting fitness in Drosophila populations. Annu Rev Genet 11:49–78 (1977).
25. Darmency H, Menchant Y, Le Corr V and Deluy C, Fitness cost due to herbicide resistance may trigger genetic background evolution. Evolution 69:271–278 (2015).
26. Gould F, Brown ZS and Kuzma J, Wicked evolution: can we address the sociobiological dilemma of pesticide resistance? Science 360:728–732 (2018).
27. Bories AG, Game AY, Chen Z, Williams TJ, Goodall S, Yen JL et al., Scalloped wing is the Lucilia cuprina Notch homologue and a candidate for the Modifier of fitness and asymmetry of diazinon resistance. Genetics 141:321–1337 (1996).
28. Johnson T and Barton NH, Theoretical models of selection and mutation on quantitative traits. Philos Trans R Soc Lond B Biol Sci 360:1411–1425 (2005).
29. Scoville AG, Lee YW, Willis JH and Kelly JK, Explaining the heritability of an ecologically significant trait in terms of individual quantitative trait loci. Biol Lett 7:896–898 (2011).
30. Remington DL, Alleles versus mutations: understanding the evolution of genetic architecture requires a molecular perspective on allelic origins. Evolution 69:3025–3038 (2015).
31. Pritchard JK, Prickrell JK and Coop G, The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr Biol 20:R208–R215 (2010).
32. Bravo W, Leon RG, Ferrell JA, Mulvaney MJ and Wood CW, Evolutionary adaptations of Palmer amaranth (Amaranthus palmeri) to nitrogen fertilization and crop rotation history affect morphology and nutrient-use efficiency. Weed Sci 66:180–189 (2018).
33. Troth A, Fuzey JR, Kim RS, Willis JH and Kelly JK, Selective trade-offs maintain alleles underpinning complex trait variation in plants. Science 361:475–478 (2018).
34. Chandi A, Milla-Lewis SR, Jordan DL, York AC, Burton JD, Zuleta MC et al., Use of AFLP markers to assess genetic diversity in Palmer amaranth (Amaranthus palmeri) populations from North Carolina and Georgia. Weed Sci 61:136–145 (2013).
35. Bravo W, Leon RG, Ferrell JA, Mulvaney MJ and Wood CW, Differentiation of life-history traits among Palmer amaranth populations (Amaranthus palmeri) and its relation to cropping systems and glyphosate sensitivity. Weed Sci 65:339–349 (2017).
