

# Interchromosomal linkage disequilibrium and linked fitness cost loci associated with selection for herbicide resistance

Sonal Gupta<sup>1,2\*</sup> , Alex Harkess<sup>3,4\*</sup> , Anah Soble<sup>1</sup>, Megan Van Etten<sup>5</sup>, James Leebens-Mack<sup>6</sup> and Regina S. Baucom<sup>1</sup> 

<sup>1</sup>Ecology and Evolutionary Biology Department, University of Michigan, 4034 Biological Sciences Building, Ann Arbor, MI 48109, USA; <sup>2</sup>Center for Genomics and Systems Biology, New York University, New York, NY 10003, USA; <sup>3</sup>Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL 36849, USA; <sup>4</sup>HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA; <sup>5</sup>Biology Department, Pennsylvania State University, Dunmore, PA 18512, USA; <sup>6</sup>Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

Author for correspondence:  
Regina S. Baucom  
Email: [rsbaucom@umich.edu](mailto:rsbaucom@umich.edu)

Received: 15 November 2022  
Accepted: 20 January 2023

New Phytologist (2023)  
doi: 10.1111/nph.18782

**Key words:** cost, detoxification, genetic hitchhiking, interchromosomal linkage disequilibrium, NTSR resistance.

## Summary

- The adaptation of weeds to herbicide is both a significant problem in agriculture and a model of rapid adaptation. However, significant gaps remain in our knowledge of resistance controlled by many loci and the evolutionary factors that influence the maintenance of resistance.
- Here, using herbicide-resistant populations of the common morning glory (*Ipomoea purpurea*), we perform a multilevel analysis of the genome and transcriptome to uncover putative loci involved in nontarget-site herbicide resistance (NTSR) and to examine evolutionary forces underlying the maintenance of resistance in natural populations.
- We found loci involved in herbicide detoxification and stress sensing to be under selection and confirmed that detoxification is responsible for glyphosate (RoundUp) resistance using a functional assay. We identified interchromosomal linkage disequilibrium (ILD) among loci under selection reflecting either historical processes or additive effects leading to the resistance phenotype. We further identified potential fitness cost loci that were strongly linked to resistance alleles, indicating the role of genetic hitchhiking in maintaining the cost.
- Overall, our work suggests that NTSR glyphosate resistance in *I. purpurea* is conferred by multiple genes which are potentially maintained through generations via ILD, and that the fitness cost associated with resistance in this species is likely a by-product of genetic hitchhiking.

## Introduction

Pesticide and herbicide use has reshaped ecological networks and induced strong selective pressures in the anthropogenic era. How species may adapt to strong selection is a fundamental question in evolution with great importance to the control of pesticide-resistant organisms. A striking feature of pesticide resistance evolution is that there are a number of different genetic solutions that can lead to resistance (Liu, 2015; Hawkins *et al.*, 2018). In herbicide-resistant plants, for example, resistance can be due to single gene mutations, often found in the herbicide's target protein (target site resistance, TSR), or due to changes in multiple genes, often underlying nontarget-site herbicide resistance (NTSR) mechanisms (Mithila & Godar, 2013; Yu & Powles, 2014). A growing body of work has produced a better understanding of resistance controlled by single genes across a variety of species (Mithila & Godar, 2013; Yu & Powles, 2014; Murphy & Tranel, 2019). However, we currently lack a deep understanding of both the genetic basis and evolutionary

potential of nontarget-site resistance mechanisms genome wide (Délye, 2013; Baucom, 2019; Beckie, 2020; Leon *et al.*, 2021).

This gap is due in part to the broad nature of nontarget-site herbicide resistance mechanisms more generally. NTSR can be caused by reduced herbicide uptake or penetration, altered translocation or sequestration, and/or enhanced metabolism (Délye *et al.*, 2013; Vemanna *et al.*, 2017; Gaines *et al.*, 2019, 2020; Pan *et al.*, 2019) – mechanisms that likely rely on a complex genetic basis (Busi *et al.*, 2013; Scarabel *et al.*, 2015; Kreiner *et al.*, 2018; Baucom, 2019). For example, detoxification is hypothesized to involve three steps – uptake of the herbicide by phosphate transporters, chemical modification, and transport to vacuoles by ABC transporters and other sugar transporters where the molecule is stored and/or inactivated (Yuan *et al.*, 2007; Gaines *et al.*, 2020). Combinations of genes involved in steps of this process have been shown to underlie resistance to a number of herbicides (Liu *et al.*, 2018; Yannicari *et al.*, 2020; Amrhein & Martinoia, 2021; Dimaano & Iwakami, 2021; Huang *et al.*, 2021; Pan *et al.*, 2021).

While some investigations have pinpointed a single gene conferring NTSR (Cummins *et al.*, 2013; Pandian *et al.*, 2021), gene

\*These authors contributed equally to this work.

expression surveys or whole-genome resequencing assays in a small handful of resistant weeds are beginning to shed light on the complexity of nontarget-site resistance mechanisms (Yuan *et al.*, 2007; Kreiner *et al.*, 2021b; Van Etten *et al.*, 2020). Because we lack a deep understanding of the genetic basis of NTSR in most weeds, however, we lack a firm grasp on the underlying forces that influence the maintenance of resistance in natural populations, such as the prevalence of alleles that may contribute to fitness costs of resistance, or the influence of interchromosomal linkage disequilibrium (ILD). The presence of ILD could implicate the potential for correlational selection and/or coadaptation between alleles underlying either resistance or its cost, or could also reflect historical associations.

*Ipomoea purpurea* is a common agricultural weed in the Southeast and Midwest United States. Populations of this species have consistently been exposed to glyphosate-based herbicides since the late 1990s, with some populations exhibiting low and others high survival postherbicide application (Kuester *et al.*, 2015). There is a fitness cost associated with this resistance: resistant populations show lower germination and deteriorated seed quality compared with susceptible populations (Van Etten *et al.*, 2016). Furthermore, populations show evidence of genetic admixture, with low genetic differentiation for a mixed-mating plant species ( $F_{ST} = 0.11\text{--}0.14$ ) as well as recent genetic connectivity (Alvarado-Serrano *et al.*, 2019). Previous work has identified separate regions of the genome housing detoxification genes (cytochrome P450s, glycosyltransferases, and ABC transporter genes) to be under selection for glyphosate resistance, and that furthermore, these regions were not similarly under selection by other environmental agents (reduced water availability, temperature, and elevation; Van Etten *et al.*, 2020). Although this suggests multiple detoxification genes across different chromosomes likely contribute to resistance, the work relied on low-coverage RAD-Seq sequencing without the benefit of a contiguously assembled genome. Thus, loci that may contribute to NTSR or its cost were likely missed (Lowry *et al.*, 2017), implying that we lack a thorough understanding of NTSR, the genomic context of NTSR alleles, and the potential for interactions among NTSR alleles in this species – all crucial to understanding the evolution of resistance more broadly.

Here, we implemented a genome-wide selection screen using whole-genome resequencing data from natural populations along with a gene expression survey to characterize the genetic architecture of glyphosate resistance and its cost in *I. purpurea*. We complemented our survey with a functional assay to test the potential that resistant *I. purpurea* individuals detoxify the herbicide more rapidly. For this, we used malathion, a pesticide that inhibits some cytochrome P450s, a key component of the detoxification pathway (following Christopher *et al.*, 1994; Preston *et al.*, 1996; Yu & Powles, 2014; Yannicari *et al.*, 2020; Pandian *et al.*, 2021). Given previous evidence that multiple loci likely contribute to glyphosate resistance in this species, and evidence of fitness cost of resistance, we had three main expectations regarding resistance in *I. purpurea*. First, we expected that regions of the genome showing high differentiation and evidence of selection when comparing herbicide-resistant and susceptible individuals

would contain loci with strong functional links to either herbicide resistance or its cost. Second, we predicted that resistant individuals would become susceptible in the presence of malathion, supporting the idea that detoxification underlies glyphosate resistance in this species. Third, we anticipated that linkage disequilibrium, the nonrandom association of alleles at different loci, should be evident among regions of the genome harboring resistance loci. Although ILD has been identified in other systems for ecologically relevant traits such as mate choice and coloration (Petkov *et al.*, 2005; Long *et al.*, 2013; Hench *et al.*, 2019), and recently has even been implicated in TSR (Kreiner *et al.*, 2021a), it is unknown whether loci underlying NTSR that are found across chromosomes exhibit long-distance or ILD, as would be expected if adaptation to herbicide is facilitated by multilocus genotypes favored by selection (Wallace, 1953; Dobzhansky, 1971; Schluter, 2000; Yeaman *et al.*, 2016).

## Materials and Methods

### Genome sequencing, assembly, and annotation

We used an *Ipomoea purpurea* (L.) Roth (Convolvulaceae) line originally sampled from NC in 1985 by M. Rausher and selfed for > 18 generations in the laboratory for genome sequencing (seeds available upon request). High molecular weight DNA was isolated using a modified large-volume CTAB protocol (Doyle & Doyle, 1990) and sequenced on a PacBio Sequel at the University of Georgia. Raw PacBio subreads from nine cells of Sequel chemistry were error-corrected with CANU (v.1.7.1; Koren *et al.*, 2017) with default parameters for raw PacBio reads (–pacbio-raw). The corrected and trimmed reads from CANU were assembled with Flye (v.2.4-release; Kolmogorov *et al.*, 2019) and anchored onto pseudomolecules by nearly 81 million read pairs of Phase Genomics Hi-C (Seattle, WA, USA) of leaf tissue using Sau3AI cut-sites. Within-genome and across-genome synteny was visualized using the CoGE SYNMAP platform (Lyons, 2008), with DAG-CHAINER options ‘-D20 -A 5’, as well as with JCVI with default parameters (<https://github.com/tanghaibao/jcvi>). *Ipomoea purpurea* pseudomolecules were numbered and oriented according to chromosome synteny against *Ipomoea nil* pseudomolecules.

Raw 50 nt single-end RNA-Seq reads (Josephs *et al.*, 2021) were aligned using STAR (v.2.7.0; Dobin *et al.*, 2013) with default single-pass parameters. Repetitive elements were annotated with REPEAT-MODELER (v.1.0.11). Long terminal repeat (LTR) retrotransposons were annotated with LTRHARVEST (v.1.6.1) with options -similar 85 -mindistlr 1000 -maxdistlr 15000 -mintsd 5 -maxtsd 20". REPEATMODELER annotations were combined with all Viridiplantae repeats from REPBASE and used as a species-specific repeat database built using REPEATMODELER with default options.

Genome annotation was performed using a diverse set of evidence. First, a set of 12 RNA-Seq libraries from leaf tissue (Van Etten *et al.*, 2021) was aligned with STAR (v.2.7.0), and transcripts assembled with STRINGTIE (v.2.1.3). MAKER2 (Holt & Yandell, 2011) was initially run with evidence from the RNA-Seq alignments, as well as peptides from *Ipomoea trifida*, *Ipomoea*

*triloba*, and *I. nil*. The resulting gene set was used to train SNAP (v.2013-11-29). AUGUSTUS (v.3.3.2) was trained with evidence from BUSCO (v.4.1.0) against the eudicot odb10 set. with default options. MAKER2 was re-run with the *ab initio* SNAP and AUGUSTUS training sets, in addition to the homologous protein and RNA-Seq evidence, to build a final gene annotation set.

## Sampling and sequencing

We selected eight populations to investigate the genetic basis of glyphosate resistance and its cost following (Van Etten *et al.*, 2020) – 4 low and high resistance each, hereon referred to as the susceptible and resistant populations, respectively (S: <20% population survival and R: >70% population survival, at 1× the field dose of RoundUp; Supporting Information Table S1). We extracted DNA from leaf tissue of 10 maternal lines per population using the Qiagen Plant DNeasy Kit and performed 150 paired-end sequencing using Illumina HiSeq4000 and NovaSeq6000. We sequenced two populations (WG, resistant, and RB, susceptible) at high coverage (average of 28.84×) and the remaining six populations at relatively low coverage (average of 14.66×).

## Variant calling

Variant calling was performed using our draft genome following Van Etten *et al.* (2020). Briefly, reads were aligned using BWA (Li & Durbin, 2009), and SNPs were called and filtered using the GATK pipeline (McKenna *et al.*, 2010). Reads were further filtered using VCF- and BCF tools (Danecek *et al.*, 2011; Li, 2011; see Methods S1 for more details). This gave us a total of 3942 549 high confidence SNPs for downstream analyses.

We performed a PCA analysis using these SNPs to investigate the population structure using the package BIGSNPR v.1.4.4 (Privé *et al.*, 2018) and found that the populations did not segregate into two separate genetic clusters (Figs S1a,b, S2a). Furthermore, PCA and admixture analysis on SNPs from regions under selection (to be described later) separated individuals into distinct resistant and susceptible groups, except for a resistant population, BI, which as we have previously shown (Alvarado-Serrano *et al.*, 2019; Van Etten *et al.*, 2020) clustered between the susceptible and other resistant populations (Figs S1c, S2b). Previous work suggests this population has evolved resistance through a different genetic mechanism compared with the other populations (Alvarado-Serrano *et al.*, 2019; Van Etten *et al.*, 2020). We thus focus our analyses on individuals from three resistant populations (DW, WG, and SPC) but present results of selection screens including BI in Tables S2–S6.

## Selection analysis

We used a two-step approach to identify the regions associated with resistance. First, we used BAYENV2 which is designed to identify candidate outliers while taking population structure into account (Günther & Coop, 2013), following (Qian *et al.*, 2020; Van Etten *et al.*, 2020). Briefly, we ran BAYENV2 on a pruned

dataset of SNPs located 5 kb apart to reduce the effect of linkage among loci (86 648 SNPs) and identified robust SNP candidates associated with resistance if they belonged in the top 5% of both Bayes factors (BFs) and Spearman's  $\rho$ , following suggested cut-offs (Qian *et al.*, 2020; Van Etten *et al.*, 2020). Genes under selection were defined as those found 5 kb up- and downstream of the regions housing the robust SNP candidates.

Second, we employed a Md-rank- $P$  approach which integrates multiple selection statistics to identify the regions under selection, a method that has been shown to considerably reduce false positives (Lotterhos *et al.*, 2017; Yurchenko *et al.*, 2018). We used the 3.9 M identified SNPs to calculate diversity and selection statistics  $G_{ST}$  (Weir, 1996),  $\pi$ , Tajima's  $D$  (Tajima, 1989), Fu and Way's  $H$  (Fay & Wu, 2003) using a custom script from (Badel *et al.*, 2018) in a 300 SNP window, for the resistant and susceptible individuals. Furthermore, using phased variants (BEAGLE v.5.1; Browning & Browning, 2007), we detected hard sweeps by estimating the H12 statistic (Garud *et al.*, 2015). For regions above the 95 percentile  $G_{ST}$ , we used  $\pi$ , Tajima's  $D$ , Fu and Way's  $H$ , and H12 and calculated a composite rank-based statistic (Md-rank- $P$ ) computed as the Mahalanobis distance of the negative  $\log_{10}$  transformation of raw statistics into rank  $P$ -values (Lotterhos *et al.*, 2017). To identify potential regions of selection, we chose bins with >95 percentile Md-rank- $P$ . Genes that were identified via both approaches were considered as the genes associated with glyphosate resistance in *I. purpurea*.

## Linkage analysis

We calculated linkage disequilibrium ( $r^2$ ) at three different levels using all the sampled individuals. First, to estimate the background genome-wide long-distance (and ILD), we calculated  $r^2$  values for randomly chosen SNPs separated by at least 100 kb (number of SNPs = 5842) using VCFtools v.0.1.15 (Danecek *et al.*, 2011). Second, we estimated the  $r^2$  for SNPs separated by at least 1 kb in and between broad regions (0.75 Mb upstream and downstream) around the four focused regions ( $G_{ST} > 0.39$ ) under selection obtained from Md-rank- $P$  approach as above. Since one would expect higher LD between regions with high differentiation, we examined the  $r^2$  values of four randomly chosen regions of similar lengths showing high differentiation ( $G_{ST} > 0.39$ ) but no signs of selection as a control.

## Differential gene expression via RNA-Seq

To identify transcripts associated with resistance and its cost, we sequenced transcriptomes of 16 individuals belonging to four treatments; resistant control, susceptible control, resistant herbicide sprayed, and susceptible herbicide sprayed (Table S7). These individuals were generated via selfing and were grown in a controlled environment. Twenty days after planting, we sprayed glyphosate (1.54 kg ai ha<sup>-1</sup>) on the treatment plants and collected the second and fourth leaf 8 h postspray. We extracted RNA using Qiagen RNeasy Plant Mini Kit and had libraries created and sequenced using Illumina NovaSeq 6000 at 150 bp paired-end sequencing by Novogene (Sacramento, CA, USA).



We processed the raw reads to remove adapters using CUTADAPT v.1.18 (Martin, 2011), mapped them to the assembled genome using STAR v.2.7.5 (Li, 2011) and estimated per gene read counts using HTSEQ v.0.11.1 (Anders *et al.*, 2015; Table S8). Transcripts with >0.5 counts-per-million in at least two samples were retained after filtering with edgeR v.3.18.1 (Robinson *et al.*, 2010). The libraries were normalized (using the trimmed mean of *M*-values method) followed by differential gene expression analysis using edgeR. For the plants exposed to herbicide, we controlled for population-specific differential gene expression by contrasting individuals from each resistant population to each susceptible population (i.e. DW (R) vs HA (S) and DW (R) vs RB (S); WG (R) vs HA (S) and WG (R) vs RB (S)). We then identified population-specific DETs associated with resistance as those that were common among separate contrasts of each resistant population to the two susceptible populations (i.e. DW (R) vs HA (S) and DW (R) vs RB (S) for DW, and WG (R) vs HA (S) and WG (R) vs RB (S) for WG; see Table S9). Using these population-specific contrasts, which controls for population of origin, we determined which loci were differentially expressed in common in each resistant vs susceptible comparison, and consider these a 'core' set of putative resistance genes. We likewise compared gene expression of all control (nontreated) resistant and susceptible individuals to one another to determine whether potential resistance loci were differentially expressed in the absence of herbicide. Due to sample size limitations, we did not perform population-specific contrasts using nontreated plants. Significance of differentially expressed transcripts (DET) was considered as  $FDR \leq 0.05$ .

### Malathion experiment

On 10 January 2022, we planted 215 replicate seeds from multiple resistant and susceptible lineages per population (Table S10) in Cone-Tainers (Stewe and Sons) in a controlled growth room. After 30 d, we subjected plants to four treatments: malathion (1 kg ai ha<sup>-1</sup> OrthoMAX; Scotts Co., Marysville, OH, USA), glyphosate (3.4 kg ai ha<sup>-1</sup> RoundUp PowerMax, Bayer), glyphosate and malathion, or a control. Before this experiment, we exposed 20 plants to five rates of malathion ( $N = 100$ ) and identified 1 kg ai ha<sup>-1</sup> as a rate that did not influence *I. purpurea* biomass or cause damage on its own. Malathion was applied using a handheld CO<sub>2</sub> sprayer, and glyphosate was applied 1 h later, again using a hand-held CO<sub>2</sub> sprayer (Spraying Systems Co., Wheaton, IL, USA) calibrated to deliver 187 l ha<sup>-1</sup>. Twenty-four days post-treatment, we recorded death, harvested and dried plants for 3 d at 7°C, and weighed each sample for an estimate of biomass.

To compare death and biomass between resistant and susceptible plants, we used log-likelihood tests and generalized linear models. Biomass was log-transformed (transformTukey function; RCOMPANION v.2.0.0; Mangiafico, 2015) and used as the dependent variable with population type (R/S) and treatment as independent variables (biomass ~ population type + treatment). We assessed significance using the ANOVA function of the CAR package v.3.0.10 (Fox & Weisberg, 2018), and performed a pairwise

comparison between groups using the lsmeans function from package LSMEANS v.2.30.0 (Lenth, 2016), adjusted for multiple tests using Tukey correction. Using the same general model, we also compared biomass between treatments for each population type (R vs S). To control for inherent differences in plant size across lineages, we standardized the biomass of each individual grown in the treatment environment by dividing individual biomass by the average of plants from the same maternal line grown in the control environment.

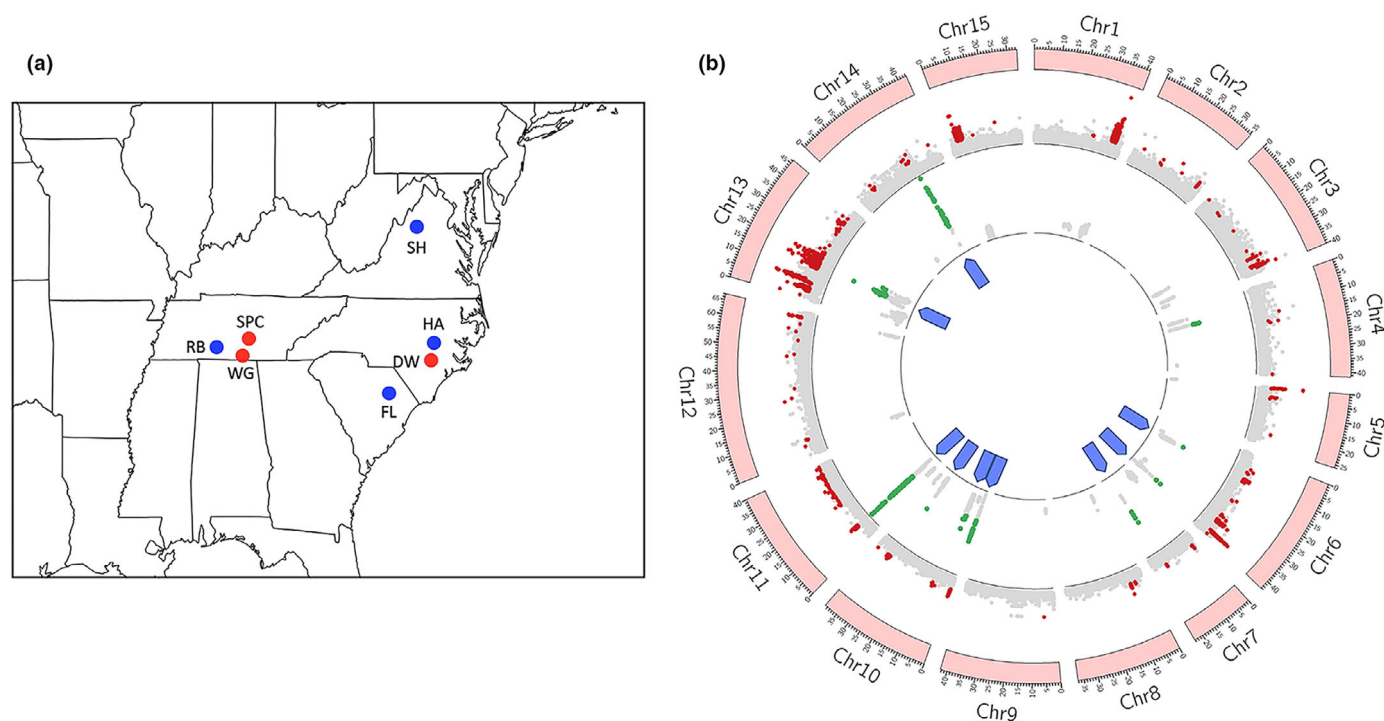
## Results

### A chromosome-scale genome assembly for common morning glory

We generated a total of 48 gigabases of PacBio Sequel whole-genome shotgun data for genome assembly (Fig. S3a), which based on flow cytometry genome size (Benaroya Institute, Seattle, WA, USA), amounts to roughly 59× genome coverage for an estimated haploid genome size of 814 Mb. We used 34.79 gigabases of trimmed and self-corrected reads for assembly and scaffolding, which produced a 602 Mb assembly in 402 scaffolds (434 contigs; scaffold N50 = 5.77 Mb). Pseudomolecule scaffolding produced 15 scaffolds corresponding to the haploid chromosome number (Fig. S3b). We oriented and numbered the chromosomes to match the highly syntenic *Ipomoea nil* genome (Hoshino *et al.*, 2016; Fig. S3c). No PacBio assembly contig misjoins were identified based on the Hi-C linkage data. BUSCO scores on the unannotated assembly show 97.5% completeness against the Viridiplantae odb10 gene set (Fig. S3d). Approximately 63% of the assembly was masked as repetitive DNA, with a significant proportion of recently expanded LTR retrotransposons (Fig. S3e). Given the high degree of synteny with the *I. nil* genome, the discrepancy between the flow cytometry genome size (814 Mb) and the assembled size (602 Mb) is likely due to young retrotransposon proliferation. We annotated 53 973 genes by combining *ab initio* gene predictions and RNA sequencing data from leaf tissue. The assembly showed high synteny with genomes across the Convolvulaceae, including *I. nil*, *I. trifida*, and *I. triloba* (Fig. S4).

### Detecting loci under selection

Whole-genome analysis of 69 individuals (from three resistant and four susceptible populations, Fig. 1a) identified genes under selection for glyphosate resistance. BAYENV2 identified 2016 SNP outliers (Table S11). Within 5 kb flanking regions of these outliers, 1908 genes were present, of which 1485 genes could be functionally annotated (Table S12). The Md-rank-*P* approach identified 15 regions (total 4.47 Mb) that exhibited signs of selection (Table S13). These housed 358 genes, 202 of which could be functionally annotated (Table S14). There were 128 genes located on six different chromosomes (Fig. 1b; Table S15) found in common between the two selection scans. These genes were broadly involved in the process of detoxification, environmental sensing, and stress signaling and response (Table 1).



**Fig. 1** Overview of the *Ipomoea purpurea* populations examined in this work and the genomic context of selection. (a) Ten individuals were sampled from each population, with resistant populations (70–100% survival postglyphosate) indicated in red and susceptible populations (<20% survival postglyphosate) indicated by blue markers (Supporting Information Table S8). (b) CIRCOS plot depicting the regions of the genome that show signs of selection associated with herbicide resistance. The genome assembly resulted in 15 scaffolds which are represented here by pink bars. Significant values of Bayes factor (BF) from BAYENV2 (top 5% BF and top 5% Spearman's Rho) indicating outlier SNPs for herbicide resistance are depicted by red dots (for ease of representation  $\log_{10}(\text{BF})$  are plotted here;  $\max \log_{10}(\text{BF}) = 4.1$ ), and the significant Md-rank-P-values (top 5% Md-rank-P-value) identifying signatures of selection are presented in green dots ( $\max \text{Md-rank-P} = 5.72$ ). Regions of the genome that were significant using both BAYENV2 and Md-rank-P are identified by the blue arrows.

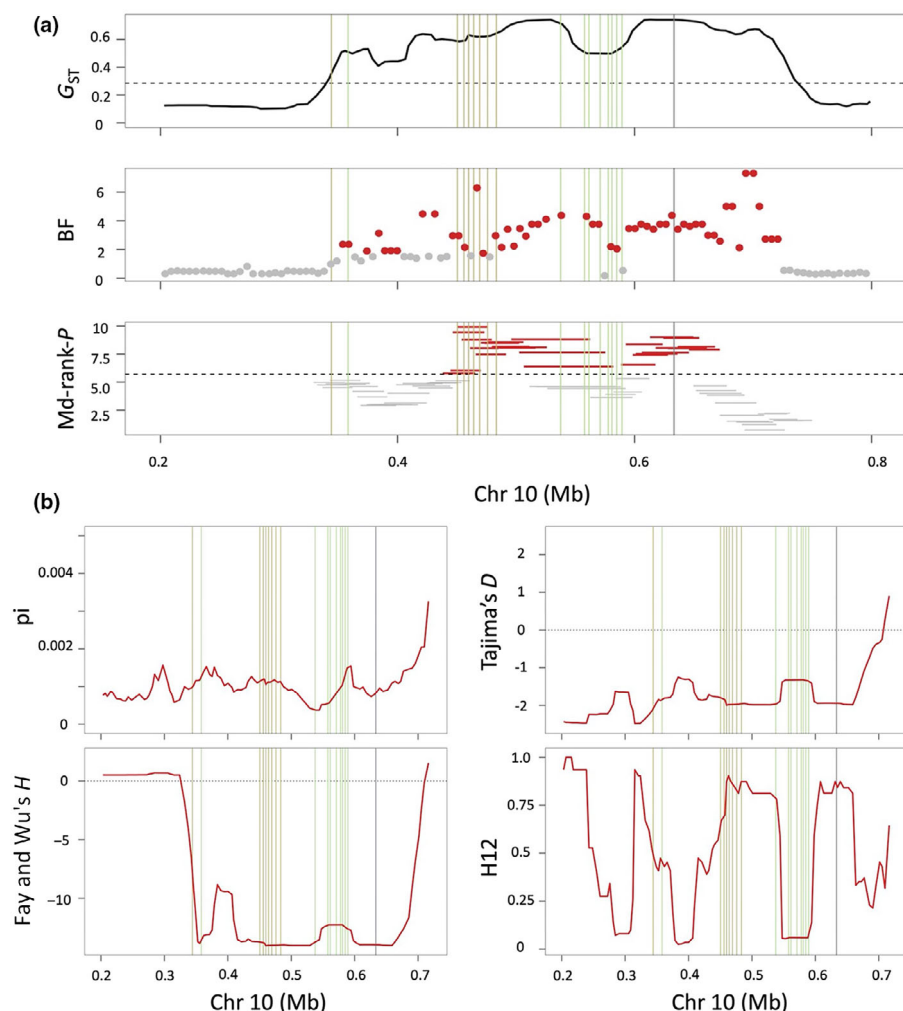
**Table 1** Overview of the genes under potential selection for glyphosate resistance in *Ipomoea purpurea* that are involved in the process of detoxification, environmental sensing, and stress signaling and response.

Chr	Location (Mb)	Functional annotations for genes of interest	Process
6	12.82–12.84	trihelix transcription factor GT-3b-like	Response to stress
6	40.40–40.57	MOR1-like	Microtubules organization during mitosis and cytokinesis
7	16.78–16.83	AP2-like ethylene-responsive transcription factor PLT1, zinc finger A20 and AN1 domain-containing stress-associated protein 8	Responses to environmental stimuli
10	0.43–0.66	Glycosyltransferase (x7), CYP76A1 (x3), CYP76A2 (x3)	Detoxification
11	6.45–7.31	serine/threonine-protein kinase CTR1	Stress signaling
11	6.45–7.31	CYP736A12	Detoxification
11	6.45–7.31	phosphate transporter PHO1 homolog 3-like (x5)	Transport of glyphosate into the cell
11	6.45–7.31	NAC domain-containing protein 92, LOB domain-containing protein 41, Leaf Rust 10 Disease-resistance locus Receptor-like Protein Kinase (x6)	Stress response
13	16.45–17.73	serine/threonine-protein kinase-like protein At5g23170, AT-hook motif nuclear-localized protein 24-like, zinc finger A20, and AN1 domain-containing stress-associated protein 1	Stress sensing and response
13	20.91–21.22	CYP87A3	Detoxification

For a comprehensive list of genes under potential selection, see Supporting Information Table S15.

The strongest candidate loci underlying nontarget-site glyphosate resistance in *I. purpurea* were present on Chr6, Chr10, Chr11, and Chr13 (Table 1). Among these, Chr10 harbors especially promising candidate genes associated with glyphosate

resistance (Fig. 2). Within a 0.23 Mb region of this chromosome showing signals of selection, we identified 13 genes – six copies of cytochrome P450 genes (CYP) and seven copies of glycosyltransferases – both of which are gene families previously



**Fig. 2** Region of chromosome 10 showing signs of selection. Shown is the (a)  $G_{ST}$  (top), Bayes factor (BF; middle) estimated between the resistant and the susceptible individuals, and Md-rank-P (bottom) for the resistant individuals, and statistics used to estimate Md-rank-P (b) clockwise starting from upper left,  $\pi$ , Tajima's  $D$ , H12 and Fay and Wu's  $H$ . Red lines indicate respective values for the resistant populations. Khaki vertical lines represent copies of glycosyltransferases, green vertical lines are the cytochrome P450, and the gray vertical line represents CTR1 (to be described later). The black dashed line in (a) represents 95 percentile values.

implicated in herbicide detoxification. The six CYPs (three CYP76A1 and three CYP76A2) were present in tandem within 53 kb, which also contained two additional copies of CYP76. Of these eight CYPs, four exhibited multiple nonsynonymous mutations that were almost fixed in the resistant individuals (resistant allele frequency = 0.95; susceptible allele frequency = 0.31; Table S16). Furthermore, two CYPs (CYP76A2) in this block exhibited either a premature stop codon and/or a splice site donor variant (G → C) in the first intron (allele 1 susceptible frequency = 0.68, resistant frequency = 0.05) in the majority of the susceptible individuals. The seven glycosyltransferases were also found in a tandem block; one glycosyltransferase copy showed the loss of a stop codon (susceptible frequency = 0.64, resistance frequency = 0.05), whereas the others exhibited multiple nonsynonymous mutations close to fixation in the resistant individuals (resistant frequency = 0.95, susceptible frequency = 0.60; Fig. S5; Table S16). The block of glycosyltransferases in this region showed evidence of a hard sweep (H12 = 0.87).

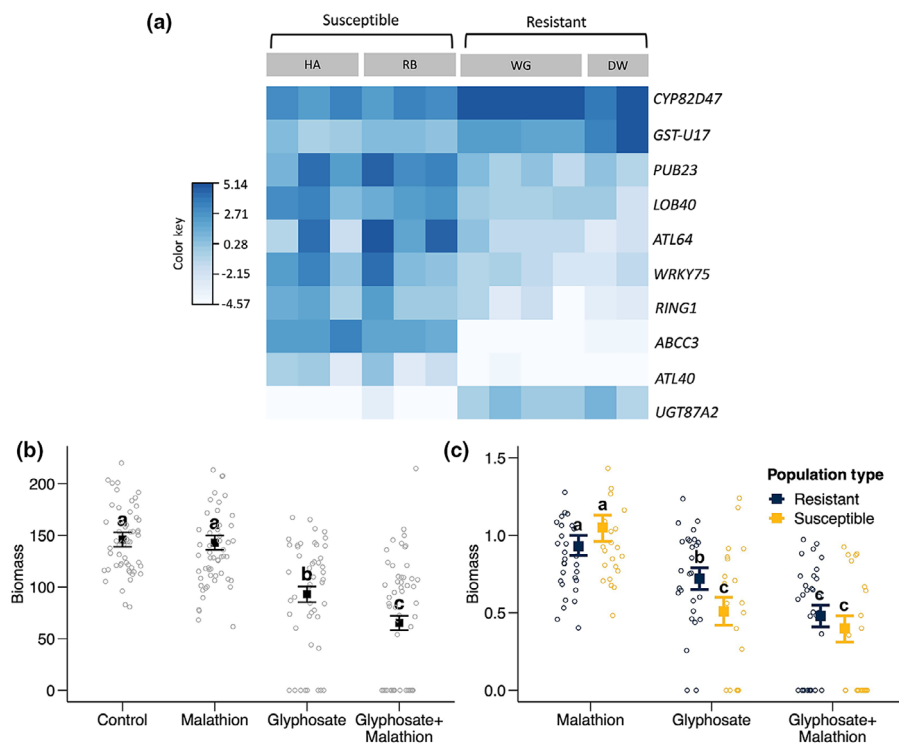
On chromosome 11, we found a 6.4–7.3 Mb region exhibiting signs of selection with 33 genes (Table 1; Fig. S6). Among these genes, were five copies of a phosphate transporter gene (*PHO1*) containing almost fixed nonsynonymous mutations in resistant individuals (resistant frequency = 0.99; Table S16). This region

also contained a copy of a CYP (*CYP736A12* family, Table S15). Interestingly, multiple stress response genes were also present in this region, with similar stress response genes found under selection on chromosome 13 (Table 1; Fig. S7). An additional short stretch under selection on chromosome 13 contained a copy of cytochrome P450 (*CYP87A3*, Table 1; Fig. S7).

Another region of note showing strong signals of selection was found on chromosome 6 (average Md-rank-P = 6.55; average  $G_{ST}$  = 0.782, Figs 1b, 5), with evidence of strong differentiation continuing further upstream and downstream (40.23–40.81 Mb; mean  $G_{ST}$  = 0.727). Within this region, we identified loci involved in stress response and loci potentially related to the cost of resistance in this species, as is described further in the 'Signs of selection on potential cost loci' section below.

Overall, we identified highly differentiated genomic regions under selection containing genes involved in herbicide detoxification (cytochrome P450s, glycosyltransferases), translocation (phosphate transporters), environmental sensing (serine/threonine kinases), and stress response (SAPs, PLT1, AHL24, and GT-3B transcription factor; Tables 1, S15). Thus, our study extends our previous work which found detoxification genes to be under selection (Van Etten *et al.*, 2020) by providing strong evidence that glyphosate resistance in *I. purpurea* is a polygenic

**Fig. 3** Gene expression variation associated with herbicide resistance, and results of a functional assay supporting the idea that resistance in *Ipomoea purpurea* is due to detoxification. (a) Loci associated with glyphosate resistance identified by differential expression analysis with  $FDR < 0.05$ . Color key represents log counts-per-million values. (b) Least square means of above-ground biomass (mg) according to treatment (malathion, glyphosate, glyphosate plus malathion, and a control (no treatment)) and (c) summarized according to resistance type (R/S), normalized over the control. Letters in (b, c) indicate significant differences between treatment environments. The addition of the cytochrome P450 inhibitor malathion reverses glyphosate resistance (glyphosate vs glyphosate+malathion contrast,  $t = 2.31$ ,  $df = 44$ ,  $P$ -value = 0.03), with the resistant individuals showing the same phenotype as the susceptible individuals in the presence of glyphosate and malathion but not in the presence of glyphosate only. Error bars represent the one SD from the least square means. Dots represent normalized biomass.



NTSR mechanism involving herbicide detoxification and response to environmental stimuli and stress.

### Gene expression differences implicate herbicide detoxification

We compared gene expression between herbicide-treated resistant and susceptible plants and found support for the idea that herbicide detoxification, plant signaling, and stress response underlie resistance. Population-specific contrasts yielded 197 (90 up- and 107 downregulated) and 597 (197 up- and 400 downregulated) DETs when contrasting either the DW or WG population to both HA and RB, respectively (Tables S17–S19). Eighty-five of these DETs were observed in pairwise comparisons of each resistant population to each susceptible population (42 up- and 43 downregulated), suggesting this set of 85 DETs represents a core set of DETs associated with glyphosate resistance irrespective of the population identity considered here (Tables S19, S20). Within this core set of DETs, we found a cytochrome P450 (CYP82D7), glycosyltransferase (UGT87A2), ABC transporter (ABCC3), and glutathione S-transferase gene (GSTU17; Fig. 3a) to be differentially regulated.

We likewise uncovered differences in the expression of genes previously associated with environmental stress responses. Among notable genes were E3 Ubiquitin ligases and transferases (PUB23, RING1, ATL40, and ATL64; Lee & Kim, 2011; Song *et al.*, 2016) and WRKY transcription factor (WRKY75; Lai *et al.*, 2008; Jiang *et al.*, 2017; Zhao *et al.*, 2019; Fig. 3a). These genes potentially represent the difference in response to the herbicide stress between the resistant and susceptible populations.

When examining population-specific gene expression differences (Tables S17, S18, S21) we found that the DW population

had an additional copy of CYP (CYP76A2) upregulated in the resistant individuals. This CYP76A2 is present within the 53-kb tandem region of eight copies of cytochrome P450 under selection on Chr10 and has a premature stop codon compared with both the susceptible populations. Furthermore, there were two additional copies of glycosyltransferases in the DW population that were differentially expressed (Fig. S8a; Table S18a). In comparison, in the WG population, we observed four more copies of CYP (three up- and one downregulated), and a copy each of ABC and GST that were downregulated in the resistant individuals (Fig. S8b; Table S17b). We identified more stress response genes showing differential expression in the WG population as compared to the DW population (Table S21).

In the control (nonherbicide) environment, we found 623 DETs when comparing resistant and susceptible individuals (319 up- and 304 downregulated; Table S22). We identified multiple copies of cytochrome P450s, glycosyltransferases, and ABC transporters that were differentially expressed, indicating that glyphosate resistance through detoxification could be constitutive rather than induced in this species. Due to low sample size in the control (i.e. nonherbicide) environment ( $n = 4$ ), we were unable to perform population-specific contrasts for control plants as we did for samples from the treatment environment. We note however that the high number of DETs in the control environment comparison ( $N = 623$ ), which does not control for population of origin, is not meaningfully different compared with the number of DETs we identified from population-specific contrasts of samples from treatment conditions ( $N = 794$ ).

Interestingly, and with the exception of the CYP76A2 cytochrome P450 identified from the DW population, the specific



genes that exhibited signs of selection from our whole-genome scan were not the same as those that exhibited differential expression. However, several of them belong to the same gene families. Beyond the possibility of detecting false positives, the most likely explanation for this result is that we examined gene expression at one time point (8 h postglyphosate application). In horseweed, different members of the ABC transporter gene family have been shown to be differentially regulated at different time points after glyphosate exposure (Piasecki *et al.*, 2019), supporting the idea of transcriptional sampling time caveat in *Ipomoea*. That we identified both signs of selection and differential expression on genes involved in detoxification strongly supports a functional role for detoxification in this species.

### Functional assay supports herbicide detoxification as a mechanism of resistance

We performed an assay to determine whether the functional mechanism of resistance in *I. purpurea* was herbicide detoxification (following Christopher *et al.*, 1994; Preston *et al.*, 1996; Yu & Powles, 2014; Yannicari *et al.*, 2020; Pandian *et al.*, 2021). We applied malathion, a pesticide that inhibits some cytochrome P450s to multiple resistant and susceptible *I. purpurea* individuals from the same populations used in the WGS resequencing and gene expression studies. We found a significant overall treatment effect ( $F$ -value = 32.05,  $df = 3$ ,  $P < 0.0001$ ; Fig. 3b) with individuals treated with both glyphosate and malathion showing lower biomass compared with individuals treated with just glyphosate (as well as just malathion or untreated controls, Fig. 3b; Table S23). Plant death followed the same pattern, with significantly more plants dying in the glyphosate plus malathion treatment compared with just glyphosate treatment (39% vs 17% death; Log-likelihood ratio test,  $G = 19.30$ ,  $df = 1$ ,  $P$ -value  $< 0.001$ ). Furthermore, the biomass of plants subjected to malathion only did not differ from the controls (Fig. 3b), showing no phytotoxicity of malathion itself on plant size.

As expected, resistant individuals showed a lower proportion of death (7% R vs 32% S; Log-likelihood ratio test,  $G = 4.55$ ,  $df = 1$ ,  $P$ -value = 0.03) and greater biomass compared with the susceptible individuals in the presence of glyphosate ( $F$ -value = 3.61,  $df = 1$ ,  $P$ -value = 0.06; Fig. 3c). However, the biomass of resistant individuals in the presence of both malathion and glyphosate was significantly lower, and proportion death significantly higher, than that of resistant individuals treated only with glyphosate (resistant plants, malathion+glyphosate *vs* glyphosate:  $t = 2.31$ ,  $df = 44$ ,  $P$ -value = 0.03; Fig. 3c; 7% vs 31% dead; Log-likelihood ratio test,  $G = 4.55$ ,  $df = 1$ ,  $P$ -value = 0.03), indicating that the presence of malathion inhibits resistance. In fact, there was no significant difference in biomass between glyphosate-resistant and susceptible genotypes when malathion was applied with glyphosate (malathion+glyphosate treatment: resistant *vs* susceptible plants:  $t = 0.77$ ,  $df = 54$ ,  $P$ -value = 0.45; Fig. 3c). Furthermore, there was no significant difference in the proportion death between resistant and susceptible plants in the glyphosate plus malathion treatment (Log-likelihood ratio test,  $G = 0.58$ ,  $df = 1$ ,  $P$ -value = 0.45). Thus, the presence of a

cytochrome P450 inhibitor lowers the level of resistance in *I. purpurea*, supporting the idea that the detoxification pathway underlies glyphosate (or RoundUp) resistance in this species.

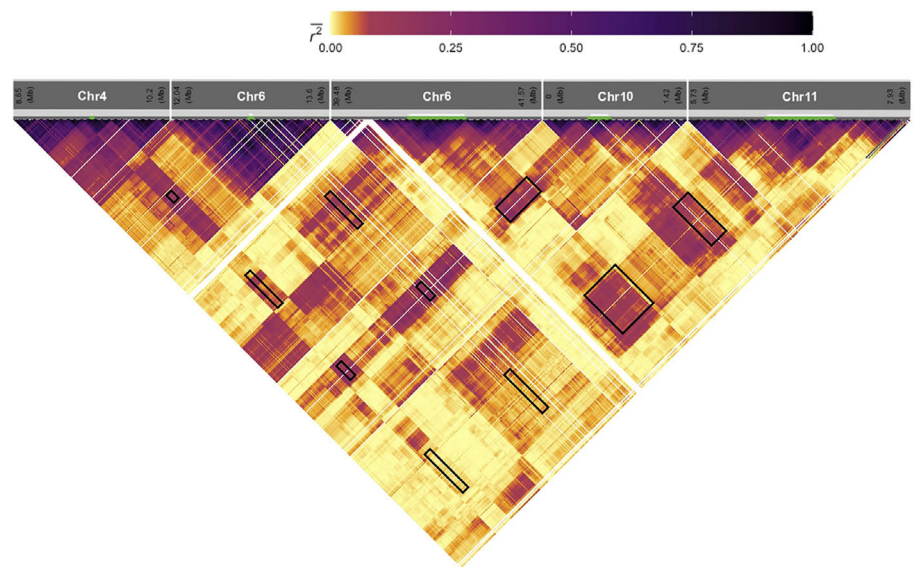
### Long-distance and interchromosomal linkage disequilibrium among NTSR alleles

We next sought to determine whether there was evidence of linkage disequilibrium between regions of the genome housing genes under selection, as would be expected for polygenic selection leading to local adaptation (Yeaman *et al.*, 2016, 2018). We calculated a measure of linkage disequilibrium ( $r^2$ ) between long-distance and interchromosomal SNPs that showed the most extreme level of differentiation and selection (98<sup>th</sup> percentile,  $G_{ST} > 0.39$ ) – regions on Chr 6 (two regions, hereon referred to as 6.1 and 6.2), 10, and 11, and compared it to the whole-genome measure. We found that the four highly differentiated regions under selection showed islands of elevated ILD (Table S24) in a backdrop of nearly zero genome-wide ILD (background interchromosomal  $r^2$  mean = 0.00096; Fig. 4). The four regions with high differentiation between resistant and susceptible populations that exhibited signs of selection showed higher LD with one another (99<sup>th</sup> percentile  $r^2 = 0.22 \pm 0.0002$  SE) compared with four randomly pulled, but highly differentiated regions between resistant and susceptible populations not under selection (99<sup>th</sup> percentile  $r^2 = 0.11 \pm 0.0001$  SE). The region under selection on Chr10 exhibited the strongest LD to other chromosomal regions under selection (99% ILD Chr6.1-Chr10 = 0.257, Chr6.2-Chr10 = 0.22, Chr11-Chr10 = 0.17). Interestingly, the highest  $r^2$  values (within the top 1%) within these regions were observed for putative resistance genes identified above. For instance, multiple glycosyltransferases and cytochrome P450s under selection on Chr10 showed high ILD with SNPs on Chr11 (Table S25).

### Signs of selection on potential cost loci

Costs of resistance can be due to resistance loci with negative pleiotropic fitness effects, or arise from physical linkage between resistance loci and other loci with negative fitness effects (Bergelson & Purrington, 1996; Zhong *et al.*, 2005). Here, we examined the latter idea. We found alternate alleles close to fixation in each population type within the 585 kb highly differentiated region on chromosome 6 (40.23–40.81 Mb; mean  $G_{ST} = 0.727$ , Fig. 5a). This region contained the nuclear fission defective 6 (*NFD6*) and NAC transcription factor 25 (*NAC25*) genes, both of which function in seed development. *NFD6*, a protein required for nuclear fusion in the embryo sac during the production of the female gametophyte (Porteriko *et al.*, 2006), contained six missense variants in the resistant individuals (mutant allele resistant frequency = 0.88, susceptible frequency = 0.21). The resistant haplotype of this gene also contained 10 SNPs in the promoter region which could potentially alter its expression. Indeed, we found this protein to be downregulated in the presence of the herbicide, with a log-fold change of  $-3.97$  across resistant individuals as compared to the susceptible individuals





**Fig. 4** Long-distance linkage disequilibrium and interchromosomal linkage disequilibrium (ILD) among the four highly differentiated ( $G_{ST} > 0.39$ ) regions under selection associated with glyphosate resistance. The four intervals displayed islands of increased linkage disequilibrium as estimated by  $r^2$  for SNPs separated by at least 1 kb in and between broad regions under selection. The white lines represent absence of SNPs (missing data), whereas the black boxes represent linkage between the five selection intervals.  $r^2$  values are averaged over two-dimensional bins of  $10 \times 10$  kb.

(Fig. 5b). Thus, our data suggest that these genes may be responsible for the lower and abnormal germination leading to the observed fitness cost in this species (Van Etten *et al.*, 2016).

Furthermore, we found that *NFD6* is within 83 kb from, and thus physically linked to, and in high LD with ( $r^2 = 0.70$ ), the potential regulatory region under selection for glyphosate resistance on Chr6.2. Similarly, the *NAC25* gene is found in close proximity to serine/threonine kinases on Chr6.2 (i.e. 82 kb away), indicating another potential gene involved in the cost phenotype is in LD ( $r^2 = 0.67$ ) with potential NTSR loci. Furthermore, this highly differentiated region containing seed development genes is linked to other regions under selection that harbor resistance alleles (99% ILD value = 0.22; Fig. 4), indicating the potential role of both linkage and linkage disequilibrium in maintaining the cost.

We compared gene expression of resistant and susceptible individuals with the expectation that putative cost loci should be differentially expressed in the absence of herbicide (i.e. the environment in which fitness costs are assessed), but may or may not be differentially expressed in the presence of herbicide, depending on whether the gene expression is constitutive or herbicide dependent. In the absence of herbicide, we identified five differentially expressed genes that play a role in fertilization and seed maturation and are thus potentially related to the cost (Table S22). Of special interest, the bud-site selection protein 31 (*BUD31*) was highly upregulated ( $\log FC = 7.22$ ) in resistant plants in the absence of herbicide, whereas its homolog *BUD13*, critical for early embryo development (Xiong *et al.*, 2019), was highly significantly downregulated in resistant individuals in the presence of herbicide ( $\log FC = -11.39$ ). We likewise found two genes involved in megagametogenesis to be downregulated in the resistant individuals – *NFD4* ( $\log FC = -3.61$ ) and *AGL61* ( $\log FC = -4.98$ ). A loss of function mutation in *AGL61* has been shown to cause abnormal morphology and over 50% seed abortion upon fertilization in *Arabidopsis* (Steffen *et al.*, 2008). Finally, we also found that E3 ubiquitin-protein ligase BRE1 (*HUB1*), a protein involved in seed germination, was strongly

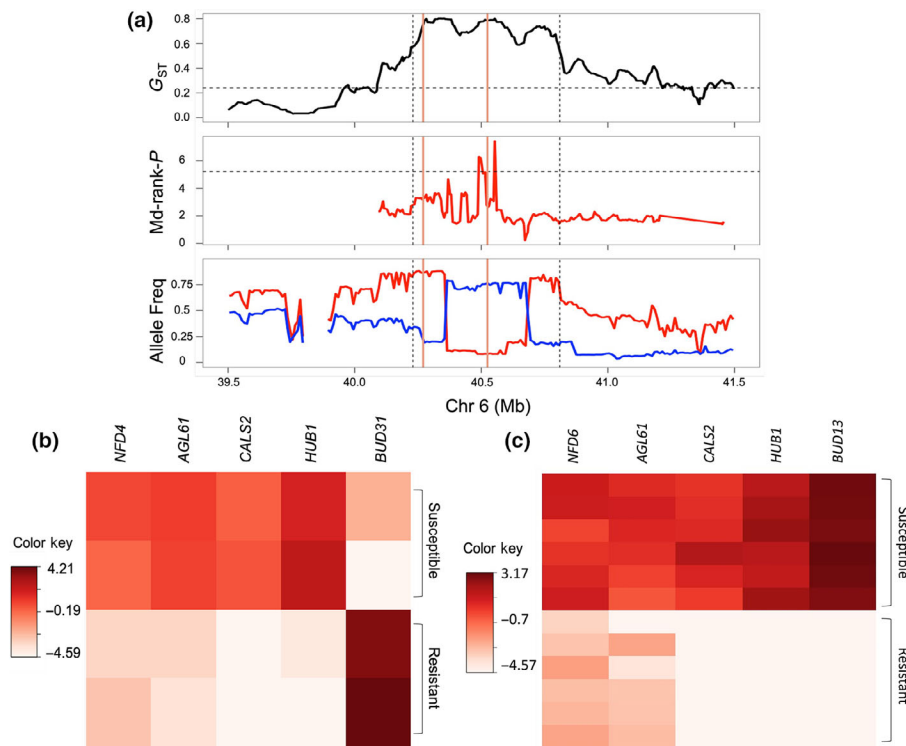
downregulated among the resistant individuals ( $\log FC = -8.27$ ). *HUB1* has been shown to control chromatin remodeling during seed development and leads to alterations in seed dormancy (Liu *et al.*, 2011).

Interestingly, three of these five candidate genes (*AGL61*, *CALS2*, and *HUB1*), and homologs of other two (*BUD13* and *NFD6*) were also significantly downregulated in the resistant populations in the presence of herbicide (Table S20). These candidate cost genes, all highly downregulated, are essential for plant reproduction and can potentially explain the phenotypic costs of glyphosate resistance in *I. purpurea* seen by Van Etten and colleagues (Van Etten *et al.*, 2016).

## Discussion

While there is an increasing appreciation for the role of nontarget-site mechanisms underlying herbicide resistance in agricultural weeds (Délye *et al.*, 2011; Ghanizadeh & Harrington, 2017; Jugulam & Shyam, 2019; Gaines *et al.*, 2020), there are strikingly few comprehensive whole-genome assays of resistant weeds (Kreiner *et al.*, 2021b; Van Etten *et al.*, 2020; Cai *et al.*, 2022) suggesting that many of the genes responsible for the NTSR response are rarely captured. Our study using a sequenced and assembled genome, whole-genome resequencing of natural populations, and a gene expression survey offers a unique and a more comprehensive opportunity to identify loci associated with NTSR compared with previous assays and to further investigate the evolutionary forces that underlie the maintenance of resistance alleles in natural populations.

Our results provide strong evidence that altered translocation and detoxification underlie resistance in *I. purpurea*. While altered translocation is likely due to a change in phosphate transporters (Preston & Wakelin, 2008), detoxification is hypothesized to be enriched in cytochrome P450s and glycosyltransferases (chemical modification), and ABC and sugar transporters (transport to vacuoles; Yuan *et al.*, 2007; Gaines *et al.*, 2020). We found evidence of selection on genes involved



**Fig. 5** Loci associated with the cost of glyphosate resistance identified by the (a) whole-genome selection scan and differential expression analysis in the (b) absence and (c) presence of herbicide. Top panel of (a) represents  $G_{ST}$  between the resistant and the susceptible populations, mid-panel is the Md-rank-P-value, and the bottom panel represents the allele frequency. Salmon vertical lines represent NFD6 and NAC25, in that order. Red and blue represent resistant and susceptible populations, respectively. Black horizontal dotted lines represent 95 percentile values while vertical lines represent regions with  $G_{ST}$  above 0.6. The differentially expressed cost genes shown here were chosen based on their functional annotation and had FDR < 0.05 and  $P$ -value < 0.00005. Color key represents log counts-per-million values.

in this pathway. We also found evidence of selection (and in some cases, differential expression) of genes involved in plant signaling and environmental stress. Our results thus expand what we currently know about the detoxification NTSR mechanism in this species to include plant signaling and stress responses, both of which are either hypothesized (Délye, 2013) or shown to be involved in herbicide resistance (Radwan, 2012; Duhoux & Délye, 2013; Dyer, 2018; Vega *et al.*, 2020). While we do not currently have functional genomics resources for this species, our study using a cytochrome P450 inhibitor supports the hypothesis that resistant *I. purpurea* individuals have the ability to detoxify the herbicide. The next step in understanding polygenic resistance in *I. purpurea* involves determining the contribution of each of the candidate loci under selection (and/or showing differential regulation) to both resistance and its associated cost. With future development of genome editing protocols for *I. purpurea*, we will be able to experimentally test the function of loci hypothesized to be contributing to glyphosate resistance.

Due to the involvement of multiple genes from the herbicide detoxification pathway, and evidence for selection on regions of the genome found on separate chromosomes, we hypothesized that multiple loci would show evidence of ILD, following (Yeaman *et al.*, 2016), who have shown the presence of statistical associations among physically unlinked loci driven by selection. Our results support this hypothesis. We found high ILD values between regulatory regions and resistance alleles, and between intervals harboring genes involved in the same molecular pathways (e.g. detoxification, and stress signaling and response), in contrast to low background LD and ILD. Local regions of long-distance linkage disequilibrium and ILD within species might be generated by demographic processes like population structure

(Nei & Li, 1973; Wilson & Goldstein, 2000), genetic drift (Schaper *et al.*, 2012), or could be due to other processes like selection (Hohenlohe *et al.*, 2012; Hench *et al.*, 2019). Furthermore, polygenic local adaptation among loci wherein adaptive alleles at two independent loci would be transmitted together (either due to additivity or epistasis) can generate linkage disequilibrium (Hohenlohe *et al.*, 2012; Yeaman *et al.*, 2016; Behrouzi & Wit, 2017; Sohail *et al.*, 2017; Hench *et al.*, 2019). Further work with a more directed functional and genetic approach will be necessary to disentangle the underlying mechanistic basis (e.g. demography vs selection), and understand the evolutionary consequences (e.g. the degree to which LD would constrain/facilitate adaptation) of this LD among the resistance loci.

A factor that should counteract the continued evolution of resistance is the potential for fitness costs of resistance, either due to the negative fitness effects of loci that are linked to resistance loci or due to the pleiotropic effects of resistance alleles themselves. While costs are central to theories of resistance evolution (Simms & Rausher, 1987; Stahl *et al.*, 1999; Baucom & Mauricio, 2004; Vila-Aiub *et al.*, 2009), there are currently no examples, to our knowledge, in which the loci underlying fitness costs of nontarget-site resistance have been identified. Our results suggest putative candidate loci associated with the previously identified cost of glyphosate resistance. Specifically, we found a highly differentiated region on Chr6 that exhibited alternate alleles in resistant and susceptible populations, and found this region to contain loci required for normal seed development and maturation (*NAC25*, *NFD6*). One of these genes, *NFD6*, was differentially regulated in the resistant individuals, further supporting its role in the low seed quality, and thus fitness cost, that we have previously described (Van Etten *et al.*, 2016). Further functional

studies are needed to validate the role of these genes in incurring fitness cost of resistance.

Additionally, our results strongly suggest genetic hitchhiking may act to maintain the cost in this species. Both *NAC25* and *NFD6* are physically linked on chromosome 6 to the regions under selection containing serine–threonine kinase genes and a regulatory region that is itself exhibiting ILD to the *CYP76A2* gene on chromosome 10. Although recombination should decouple cost alleles that are physically linked to resistance alleles, these loci would not completely decouple if the recombination rate ( $c$ ) is much lower than the selection coefficient ( $s$ ) of the resistance locus (i.e.  $c \ll s$ , Stephan *et al.*, 1992). The requirement that  $c \ll s$  is not improbable given the close proximity of cost and resistance loci (< 85 kb) and the strong ongoing selection for herbicide resistance. Furthermore, if the ratio  $d/s < 10^{-4}$ , the hitchhiking would almost be complete and the cost alleles could become fixed in the populations (Fay & Wu, 2000). Alternatively, it is possible that new compensatory mutations arising in the population could increase in frequency over time due to selection, decoupling the cost and resistance alleles and thus reducing fitness cost associated with glyphosate resistance (Vogwill *et al.*, 2016; Lenormand *et al.*, 2018).

Overall, our work identified the potential genes associated with NTSR glyphosate resistance in *I. purpurea* – our whole-genome and transcriptome assays strongly support the role of detoxification conferring glyphosate resistance in this species, and we additionally identified a role for plant sensing and stress. Interestingly, we show that NTSR glyphosate resistance in *I. purpurea* likely involves multiple loci that exhibit linkage disequilibrium and ILD. We also provide strong evidence to support the idea that fitness costs may be due to loci in strong linkage with resistance loci. Our work highlights the importance of multilevel, multipopulation study in identifying the genetic mechanisms underlying polygenic defense traits, and for understanding the complex genetic interplay between defense and cost.

## Acknowledgements

We thank M. Rausher for supplying germplasm used in the *I. purpurea* genome assembly, and T. Newsum, S. Paranjape, and K. Johnson for growth room phenotyping, and M. Palmer and MBGNA for growth room support. We likewise thank J. Opp in the University of Michigan sequencing core for sequencing support, Amanda Cummings for high molecular weight DNA preparation, and the Georgia Genomics and Bioinformatics Core, which provided the PacBio Sequel sequencing service. We also thank E. B. Josephs, R. L. Rogers, and the Ross-Ibarra laboratory for providing feedback on the manuscript. Funding for this work was provided by the University of Michigan and USDA (United States Department of Agriculture) NIFA awards 24892 & 28497 to RSB, NSF (National Science Foundation, USA) DEB 1442199 and IOS 1444567 to JLM, and NSF IOS 1611853 to AH.




## Competing interests

None declared.

## Author contributions

RSB and SG designed the project. RSB, MLVE and SG designed the experiments. AH and JLM assembled and annotated the genome. AS performed the growth room experiment. MLVE performed the WGS experiment. SG analyzed whole-genome resequencing and RNA-Seq data and performed linkage analysis. RSB, SG, AH and JLM wrote the paper. SG and AH contributed equally to this work.

## ORCID

Regina S. Baucom  <https://orcid.org/0000-0001-7960-498X>  
Sonal Gupta  <https://orcid.org/0000-0002-4419-2345>  
Alex Harkess  <https://orcid.org/0000-0002-2035-0871>

## Data availability

The datasets generated during and/or analyzed during the current study have been deposited in SRA database under the project PRJNA922500. The *I. purpurea* genome assembly and annotation are available in the CoGe platform – <https://genomevolution.org/coge/GenomeInfo.pl?gid=58735>. Biomass data are available via the Dryad data repository: [10.5061/dryad.7wm37pvxm](https://doi.org/10.5061/dryad.7wm37pvxm).

## References

- Alvarado-Serrano DF, Van Etten ML, Chang S-M, Baucom RS. 2019. The relative contribution of natural landscapes and human-mediated factors on the connectivity of a noxious invasive weed. *Heredity* 122: 29–40.
- Amrhein N, Martinoia E. 2021. An ABC transporter of the ABCC subfamily localized at the plasma membrane confers glyphosate resistance. *Proceedings of the National Academy of Sciences, USA* 118: e2100136118.
- Anders S, Pyl PT, Huber W. 2015. HTSeq—a PYTHON framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166–169.
- Baduel P, Hunter B, Yeola S, Bomblies K. 2018. Genetic basis and evolution of rapid cycling in railway populations of tetraploid *Arabidopsis arenosa*. *PLoS Genetics* 14: e1007510.
- Baucom RS. 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.
- Baucom RS, Mauricio R. 2004. Fitness costs and benefits of novel herbicide tolerance in a noxious weed. *Proceedings of the National Academy of Sciences, USA* 101: 13386–13390.
- Beckie HJ. 2020. Herbicide resistance in plants. *Plants* 9: 435.
- Behrouzi P, Wit EC. 2017. Detecting epistatic selection with partially observed genotype data using copula graphical models. *arXiv [Stat.AP]*. doi: [10.48550/arXiv.1710.00894](https://doi.org/10.48550/arXiv.1710.00894).
- Bergelson J, Purrington CB. 1996. Surveying patterns in the cost of resistance in plants. *The American Naturalist* 148: 536–558.
- Browning SR, Browning BL. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics* 81: 1084–1097.
- Busi R, Neve P, Powles S. 2013. Evolved polygenic herbicide resistance in *Lolium rigidum* by low-dose herbicide selection within standing genetic variation. *Evolutionary Applications* 6: 231–242.
- Cai L, Comont D, MacGregor D, Lowe C, Beffa R, Neve P, Saski C. 2022. The blackgrass genome reveals patterns of non-parallel evolution of polygenic herbicide resistance. *New Phytologist* 237: 1891–1907.



- Christopher JT, Preston C, Powles SB. 1994. Malathion antagonizes metabolism-based chlorsulfuron resistance in *Lolium rigidum*. *Pesticide Biochemistry and Physiology* 49: 172–182.
- Cummins I, Wortley DJ, Sabbadin F, He Z, Coxon CR, Straker HE, Sellars JD, Knight K, Edwards L, Hughes D *et al.* 2013. Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. *Proceedings of the National Academy of Sciences, USA* 110: 5812–5817.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.
- Délye C. 2013. Unravelling the genetic bases of non-target-site-based resistance (NTSR) to herbicides: a major challenge for weed science in the forthcoming decade. *Pest Management Science* 69: 176–187.
- Délye C, Gardin JAC, Boucansaud K, Chauvel B, Petit C. 2011. Non-target-site-based resistance should be the centre of attention for herbicide resistance research: *Alopecurus myosuroides* as an illustration: why we need more research on NTSR. *Weed Research* 51: 433–437.
- Délye C, Jasieniuk M, Le Corre V. 2013. Deciphering the evolution of herbicide resistance in weeds. *Trends in Genetics* 29: 649–658.
- Dimaano NG, Iwakami S. 2021. Cytochrome P450-mediated herbicide metabolism in plants: current understanding and prospects. *Pest Management Science* 77: 22–32.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21.
- Dobzhansky T. 1971. *Genetics of the evolutionary process*. New York, NY, USA: Columbia University Press.
- Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 39–40.
- Duhoux A, Délye C. 2013. Reference genes to study herbicide stress response in *Lolium* sp.: up-regulation of P450 genes in plants resistant to acetolactate-synthase inhibitors. *PLoS ONE* 8: e63576.
- Dyer WE. 2018. Stress-induced evolution of herbicide resistance and related pleiotropic effects. *Pest Management Science* 74: 1759–1768.
- Fay JC, Wu CI. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155: 1405–1413.
- Fay JC, Wu C-I. 2003. Sequence divergence, functional constraint, and selection in protein evolution. *Annual Review of Genomics and Human Genetics* 4: 213–235.
- Fox J, Weisberg S. 2018. *An R companion to applied regression*. Los Angeles, CA, USA: SAGE Publications.
- Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A, Dayan FE. 2020. Mechanisms of evolved herbicide resistance. *The Journal of Biological Chemistry* 295: 10307–10330.
- Gaines TA, Patterson EL, Neve P. 2019. Molecular mechanisms of adaptive evolution revealed by global selection for glyphosate resistance. *New Phytologist* 223: 1770–1775.
- Garud NR, Messer PW, Buzbas EO, Petrov DA. 2015. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genetics* 11: e1005004.
- Ghanizadeh H, Harrington KC. 2017. Non-target site mechanisms of resistance to herbicides. *Critical Reviews in Plant Sciences* 36: 24–34.
- Günther T, Coop G. 2013. Robust identification of local adaptation from allele frequencies. *Genetics* 195: 205–220.
- Hawkins NJ, Bass C, Dixon A, Neve P. 2018. The evolutionary origins of pesticide resistance. *Biological Reviews of the Cambridge Philosophical Society* 94: 135–155.
- Hench K, Vargas M, Höppner MP, Owen McMillan W, Puebla O. 2019. Inter-chromosomal coupling between vision and pigmentation genes during genomic divergence. *Nature Ecology & Evolution* 3: 657–667.
- Hohenlohe PA, Bassham S, Currey M, Cresko WA. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 367: 395–408.
- Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12: 491.
- Hoshino A, Jayakumar V, Nitasaka E, Toyoda A, Noguchi H, Itoh T, Shin-I T, Minakuchi Y, Koda Y, Nagano AJ *et al.* 2016. Genome sequence and analysis of the Japanese morning glory *Ipomoea nil*. *Nature Communications* 7: 13295.
- Huang X-X, Zhao S-M, Zhang Y-Y, Li Y-J, Shen H-N, Li X, Hou B-K. 2021. A novel UDP-glycosyltransferase 91C1 confers specific herbicide resistance through detoxification reaction in Arabidopsis. *Plant Physiology and Biochemistry* 159: 226–233.
- Jiang J, Ma S, Ye N, Jiang M, Cao J, Zhang J. 2017. WRKY transcription factors in plant responses to stresses. *Journal of Integrative Plant Biology* 59: 86–101.
- Josephs EB, Van Etten ML, Harkess A, Platts A, Baucom RS. 2021. Adaptive and maladaptive expression plasticity underlying herbicide resistance in an agricultural weed. *Evolution Letters* 5: 432–440.
- Jugulam M, Shyam C. 2019. Non-target-site resistance to herbicides: recent developments. *Plants* 8: 417.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* 37: 540–546.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. CANU: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research* 27: 722–736.
- Kreiner JM, Sandler G, Stern AJ, Tranel PJ, Weigel D, Stinchcombe JR, Wright SI. 2021a. Repeated origins, gene flow, and allelic interactions of herbicide resistance mutations in a widespread agricultural weed. *eLife*. doi: 10.7554/eLife.70242.
- Kreiner JM, Stinchcombe JR, Wright SI. 2018. Population genomics of herbicide resistance: adaptation via evolutionary rescue. *Annual Review of Plant Biology* 69: 611–635.
- Kreiner JM, Tranel PJ, Weigel D, Stinchcombe JR, Wright SI. 2021b. The genetic architecture and population genomic signatures of glyphosate resistance in *Amaranthus tuberculatus*. *Molecular Ecology* 30: 5373–5389.
- Kuester A, Chang S-M, Baucom RS. 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.
- Lai Z, Vinod K, Zheng Z, Fan B, Chen Z. 2008. Roles of Arabidopsis WRKY3 and WRKY4 transcription factors in plant responses to pathogens. *BMC Plant Biology* 8: 68.
- Lee J-H, Kim WT. 2011. Regulation of abiotic stress signal transduction by E3 ubiquitin ligases in Arabidopsis. *Molecules and Cells* 31: 201–208.
- Lenormand T, Harmand N, Gallet R. 2018. Cost of resistance: an unreasonably expensive concept. *Rethinking Ecology* 3: 51–70.
- Lenth RV. 2016. Least-squares means: the R package LSMEANS. *Journal of Statistical Software, Articles* 69: 1–33.
- Leon RG, Dunne JC, Gould F. 2021. The role of population and quantitative genetics and modern sequencing technologies to understand evolved herbicide resistance and weed fitness. *Pest Management Science* 77: 12–21.
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27: 2987–2993.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25: 1754–1760.
- Liu N. 2015. Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. *Annual Review of Entomology* 60: 537–559.
- Liu W, Bai S, Zhao N, Jia S, Li W, Zhang L, Wang J. 2018. Non-target site-based resistance to tribenuron-methyl and essential involved genes in *Myosoton aquaticum* (L.). *BMC Plant Biology* 18: 225.
- Liu Y, Geyer R, van Zanten M, Carles A, Li Y, Hörold A, van Nocker S, Soppe WJJ. 2011. Identification of the Arabidopsis REDUCED DORMANCY 2 gene uncovers a role for the polymerase associated factor 1 complex in seed dormancy. *PLoS ONE* 6: e22241.
- Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, Zhang Q, Vilhjálmsson BJ, Korte A, Nizhynska V *et al.* 2013. Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. *Nature Genetics* 45: 884–890.
- Lotterhos KE, Card DC, Schaal SM, Wang L, Collins C, Verity B. 2017. Composite measures of selection can improve the signal-to-noise ratio in genome scans. *Methods in Ecology and Evolution* 8: 717–727.



- Lowry DB, Hoban S, Kelley JL, Lotterhos KE, Reed LK, Antolin MF, Storfer A. 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources* 17: 142–152.
- Lyons EH. 2008. *CoGe, a new kind of comparative genomics platform: insights into the evolution of plant genomes*. Berkeley, CA, USA: University of California.
- Mangiafico SS. 2015. *An R companion for the handbook of biological statistics*. [WWW document] URL [rcompanion.org/documents/RCompanionBioStatistics.pdf](http://rcompanion.org/documents/RCompanionBioStatistics.pdf) [accessed 1 May 2020].
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* 17: 10–12.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M *et al.* 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–1303.
- Mithila J, Godar AS. 2013. Understanding genetics of herbicide resistance in weeds: implications for weed management. *Advances in Crop Science and Technology* 1: 1–3.
- Murphy BP, Tranel PJ. 2019. Target-site mutations conferring herbicide resistance. *Plants* 8: 382.
- Nei M, Li WH. 1973. Linkage disequilibrium in subdivided populations. *Genetics* 75: 213–219.
- Pan L, Yu Q, Han H, Mao L, Nyporko A, Fan L, Bai L, Powles S. 2019. Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in *Echinochloa colona*. *Plant Physiology* 181: 1519–1534.
- Pan L, Yu Q, Wang J, Han H, Mao L, Nyporko A, Maguza A, Fan L, Bai L, Powles S. 2021. An ABCC-type transporter endowing glyphosate resistance in plants. *Proceedings of the National Academy of Sciences, USA* 118: e2100136118.
- Pandian BA, Sathishraj R, Prasad PVV, Jugulam M. 2021. A single gene inherited trait confers metabolic resistance to chlorsulfuron in grain sorghum (*Sorghum bicolor*). *Planta* 253: 48.
- Petkov PM, Graber JH, Churchill GA, DiPetrillo K, King BL, Paigen K. 2005. Evidence of a large-scale functional organization of mammalian chromosomes. *PLoS Genetics* 1: e33.
- Piasecki C, Yang Y, Benemann DP, Kremer FS, Galli V, Millwood RJ, Cechin J, Agostinetto D, Maia LC, Vargas L *et al.* 2019. Transcriptomic analysis identifies new non-target site glyphosate-resistance genes in *Conyza bonariensis*. *Plants* 8: 157.
- Portereiko MF, Sandaklie-Nikolova L, Lloyd A, Dever CA, Otsuga D, Drews GN. 2006. NUCLEAR FUSION DEFECTIVE1 encodes the Arabidopsis RPL21M protein and is required for karyogamy during female gametophyte development and fertilization. *Plant Physiology* 141: 957–965.
- Preston C, Tardif FJ, Christopher JT. 1996. Multiple resistance to dissimilar herbicide chemistries in a biotype of *lilium rigidum* due to enhanced activity of several herbicide degrading enzymes. *Pesticide Biochemistry and Physiology* 54: 123–134.
- Preston C, Wakelin AM. 2008. Resistance to glyphosate from altered herbicide translocation patterns. *Pest Management Science* 64: 372–376.
- Privé F, Aschard H, Ziyatdinov A, Blum MGB. 2018. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinformatics* 34: 2781–2787.
- Qian C, Yan X, Shi Y, Yin H, Chang Y, Chen J, Ingvarsson PK, Nevo E, Ma X-F. 2020. Adaptive signals of flowering time pathways in wild barley from Israel over 28 generations. *Heredity* 124: 62–76.
- Radwan DEM. 2012. Salicylic acid induced alleviation of oxidative stress caused by clethodim in maize (*Zea mays* L.) leaves. *Pesticide Biochemistry and Physiology* 102: 182–188.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. EDGER: a BIOCONDUCTOR package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–140.
- Scarabel L, Pernin F, Délye C. 2015. Occurrence, genetic control and evolution of non-target-site based resistance to herbicides inhibiting acetolactate synthase (ALS) in the dicot weed *Papaver rhoeas*. *Plant Science* 238: 158–169.
- Schaper E, Eriksson A, Rafajlovic M, Sagitov S, Mehlig B. 2012. Linkage disequilibrium under recurrent bottlenecks. *Genetics* 190: 217–229.
- Schluter D. 2000. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press.
- Simms EL, Rausher MD. 1987. Costs and benefits of plant resistance to herbivory. *The American Naturalist* 130: 570–581.
- Sohail M, Vakhrusheva OA, Sul JH, Pulit SL, Francioli LC, Genome of the Netherlands Consortium, Alzheimer's Disease Neuroimaging Initiative, van den Berg LH, Veldink JH, de Bakker PIW *et al.* 2017. Negative selection in humans and fruit flies involves synergistic epistasis. *Science* 356: 539–542.
- Song J, Xing Y, Munir S, Yu C, Song L, Li H, Wang T, Ye Z. 2016. An ATL78-like ring-h2 finger protein confers abiotic stress tolerance through interacting with RAV2 and CSN5B in Tomato. *Frontiers in Plant Science* 7: 1305.
- Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. 1999. Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. *Nature* 400: 667–671.
- Steffen JG, Kang I-H, Portereiko MF, Lloyd A, Drews GN. 2008. AGL61 interacts with AGL80 and is required for central cell development in Arabidopsis. *Plant Physiology* 148: 259–268.
- Stephan W, Wiehe THE, Lenz MW. 1992. The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theoretical Population Biology* 41: 237–254.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Van Etten M, Lee KM, Chang S-M, Baucom RS. 2020. Parallel and nonparallel genomic responses contribute to herbicide resistance in *Ipomoea purpurea*, a common agricultural weed. *PLoS Genetics* 16: e1008593.
- Van Etten ML, Kuester A, Chang S-M, Baucom RS. 2016. Fitness costs of herbicide resistance across natural populations of the common morning glory, *Ipomoea purpurea*. *Evolution* 70: 2199–2210.
- Van Etten ML, Soble A, Baucom RS. 2021. Variable inbreeding depression may explain associations between the mating system and herbicide resistance in the common morning glory. *Molecular Ecology* 30: 5422–5437.
- Vega T, Gil M, Martin G, Moschen S, Picardi L, Nestares G. 2020. Stress response and detoxification mechanisms involved in non-target-site herbicide resistance in sunflower. *Crop Science* 60: 1809–1822.
- Vemanna RS, Vennapusa AR, Easwaran M, Chandrashekar BK, Rao H, Ghanti K, Sudhakar C, Mysore KS, Makarla U. 2017. Aldo-keto reductase enzymes detoxify glyphosate and improve herbicide resistance in plants. *Plant Biotechnology Journal* 15: 794–804.
- Vila-Aiub MM, Neve P, Powles SB. 2009. Fitness costs associated with evolved herbicide resistance alleles in plants. *New Phytologist* 184: 751–767.
- Vogwill T, Kojadinovic M, MacLean RC. 2016. Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of *Pseudomonas*. *Proceedings Biological Sciences* 283: 1830.
- Wallace B. 1953. On coadaptation in *Drosophila*. *The American Naturalist* 87: 343–358.
- Weir BS. 1996. *Genetic data analysis II*. Sunderland, MA, USA: Sinauer Associates.
- Wilson JF, Goldstein DB. 2000. Consistent long-range linkage disequilibrium generated by admixture in a bantu-semitic hybrid population. *American Journal of Human Genetics* 67: 926–935.
- Xiong F, Ren J-J, Yu Q, Wang Y-Y, Kong L-J, Otegui MS, Wang X-L. 2019. AtBUD13 affects pre-mRNA splicing and is essential for embryo development in Arabidopsis. *The Plant Journal* 98: 714–726.
- Yannicari M, Gígón R, Larsen A. 2020. Cytochrome P450 herbicide metabolism as the main mechanism of cross-resistance to ACCase- and ALS-inhibitors in *Lolium* spp. populations from Argentina: a molecular approach in characterization and detection. *Frontiers in Plant Science* 11: 600301.
- Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. 2018. Quantifying how constraints limit the diversity of viable routes to adaptation. *PLoS Genetics* 14: e1007717.
- Yeaman S, Hodgins KA, Lotterhos KE, Suren H, Nadeau S, Degner JC, Nurkowski KA, Smets P, Wang T, Gray LK *et al.* 2016. Convergent local adaptation to climate in distantly related conifers. *Science* 353: 1431–1433.
- Yu Q, Powles S. 2014. Metabolism-based herbicide resistance and cross-resistance in crop weeds: a threat to herbicide sustainability and global crop production. *Plant Physiology* 166: 1106–1118.

- Yuan JS, Tranel PJ, Stewart CN Jr. 2007. Non-target-site herbicide resistance: a family business. *Trends in Plant Science* 12: 6–13.
- Yurchenko AA, Daetwyler HD, Yudin N, Schnabel RD, Vander Jagt CJ, Soloshenko V, Lhasaranov B, Popov R, Taylor JF, Larkin DM. 2018. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and acclimation. *Scientific Reports* 8: 1–16.
- Zhao X-Y, Qi C-H, Jiang H, You C-X, Guan Q-M, Ma F-W, Li Y-Y, Hao Y-J. 2019. The MdWRKY31 transcription factor binds to the MdRAV1 promoter to mediate ABA sensitivity. *Horticulture Research* 6: 66.
- Zhong D, Pai A, Yan G. 2005. Costly resistance to parasitism. *Genetics* 169: 2127–2135.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Fig. S1** PCA of the resistant and susceptible populations sampled.

**Fig. S2** Admixture analysis of the resistant and susceptible populations sampled.

**Fig. S3** Chromosome scaffolding and renaming.

**Fig. S4** Synteny of the *Ipomoea purpurea* genome against related Convolvulaceae species, including *Ipomoea nil*, *Ipomoea trifida*, and *Ipomoea triloba*.

**Fig. S5** Signs of selection across conserved haplotype of multiple glycosyltransferases for each individual on chromosome 10.

**Fig. S6** Region of chromosome 11 showing signs of selection.

**Fig. S7** Region of chromosome 13 showing signs of selection.

**Fig. S8** Differentially expressed transcripts (with FDR  $P$ -value  $< 0.05$ ) of interest under treatment (herbicide exposed) that are unique to resistance populations DW and WG as compared to susceptible populations (RB and HA).

**Methods S1** Variant calling for whole-genome resequencing data.

**Table S1** Population information for each population used in the study.

**Table S2** SNP outliers (in the top 5% Bayes factor and 5% Rho) identified via BAYENV2 for R vs S (with BI).

**Table S3** Functional annotation of genes present within  $\pm 5$  kb of BAYENV2 SNP outliers (with BI).

**Table S4** Genome selection summary statistics (with BI) for windows with Gst above 95 percentile value (window size of 300 SNPs; step size of 50 SNPs).

**Table S5** Functional annotations of genes under potential selection (with BI) identified via md-rank-P approach.

**Table S6** Functional annotations of genes under potential selection (with BI) identified via BAYENV2 and md-rank-P approach.

**Table S7** RNA-Seq sample information used in the study.

**Table S8** RNA-Seq gene counts estimated for all samples using HTSeq.

**Table S9** Differential expression analysis design for treatment.

**Table S10** Sample information used for the malathion assay.

**Table S11** SNP outliers (in the top 5% Bayes factor and 5% Rho) identified via BAYENV2 for R vs S (with BI removed).

**Table S12** Functional annotation of genes present within  $\pm 5$  kb of BAYENV2 SNP outliers (without BI).

**Table S13** Genome selection summary statistic (without BI) for windows with Gst above 95 percentile value (window size of 300 SNPs; step size of 50 SNPs).

**Table S14** Functional annotations of genes under potential selection (without BI) identified via md-rank-P approach.

**Table S15** Functional annotations of genes under potential selection (without BI) identified via BAYENV2 and md-rank-P approach.

**Table S16** Mutations and their effects present within the genes of interest.

**Table S17** List of differentially expressed genes between treated (herbicide sprayed) resistant vs susceptible individuals for WG population.

**Table S18** List of differentially expressed genes between treated (herbicide sprayed) resistant vs susceptible individuals for DW population.

**Table S19** Population-wise counts of differentially expressed transcripts under treatment (herbicide exposed).

**Table S20** List of differentially expressed genes between treated (herbicide sprayed) resistant vs susceptible individuals (common genes among all populations).

**Table S21** Differentially expressed transcripts of interest under treatment (herbicide exposed) that are unique to resistance populations DW and WG as compared to susceptible populations (RB and HA).

**Table S22** List of differentially expressed genes between control (nonherbicide sprayed) resistant vs susceptible individuals.

**Table S23** Pairwise contrast statistics for normalized above-ground biomass between the four treatment conditions.

**Table S24** Interchromosomal linkage disequilibrium summary statistics (99<sup>th</sup> percentile value and max  $r^2$ ) for the five regions under selection that exhibited  $GST > 0.39$ .

**Table S25** Individual interchromosomal linkage disequilibrium interactions, above the 99 percentile cutoff  $r^2$  value, for SNPs within the region under selection.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.