Rapid weed adaptation and range expansion in response to agriculture over the past two centuries

Julia M. Kreiner1,2,*, Sergio M. Latore2,3, Hernán A. Burbano3,4, John R. Stinchcombe5, Sarah P. Otto2,6, Detlef Weigel4, Stephen I. Wright3

North America has experienced a massive increase in cropland use since 1800, accompanied more recently by the intensification of agricultural practices. Through genome analysis of present-day and historical samples spanning environments over the past two centuries, we studied the effect of these changes in farming on the extent and tempo of evolution across the native range of the common waterhemp (*Amaranthus tuberculatus*), a now pervasive agricultural weed. Modern agriculture has imposed strengths of selection rarely observed in the wild, with notable shifts in allele frequency trajectories since agricultural intensification in the 1960s. An evolutionary response to this extreme selection was facilitated by a concurrent human-mediated range shift. By reshaping genome-wide diversity across the landscape, agriculture has driven the success of this weed in the 21st century.

Agricultural practices across North America have rapidly intensified over the past two centuries through cropland expansion (1), monoculture plantings (2, 3), and increased chemical inputs (4, 5). Since the beginning of the 1800s, cropland usage has expanded from 8 million to 200 million hectares in Canada and the United States alone (1). Since the mid-1900s, development of new crop varieties—including high-yield and herbicide-resistant wheat, corn, and soy (6, 7)—has greatly improved the efficiency of food production in all farming sectors. Combined with increased reliance on pesticides, fertilizers, irrigation, and large-scale mechanization, this global transformation is often referred to as the agricultural Green Revolution (8–10). Pesticide effectiveness, however, has been limited by the evolution of resistance across numerous pest species (11–14). Although technological innovation for efficient food production has risen with increasing global food demands, the concomitant landscape conversion has become one of the foremost drivers of global biodiversity loss (15).

Species that have managed to survive, and even thrive, in the face of such extreme environmental change provide notable examples of rapid adaptation on contemporary time scales and illustrate the evolutionary consequences of anthropogenic change. One such species is the common waterhemp (*Amaranthus tuberculatus*), an agricultural weed that is native to North America and persists in large part in natural, riparian habitats (16, 17), which provides an opportunity to investigate the time scale and extent of contemporary agricultural adaptation. The genetic changes underlying weediness are particularly important to understand in *A. tuberculatus* because it has recently become one of the most problematic agricultural weeds in North America as a result of its widespread adaptation to herbicides, its persistence in fields across seasons, and its strong ability to compete with both soy and corn (18, 19). Determining the roles of newly arisen mutations, genetic variants predating the onset of environmental change (20, 21), migration across the range (22), and their interactions (23, 24) will inform the temporal and spatial scales at which contemporary adaptation occurs and management strategies should be used.

To understand how changing agricultural practices have shaped the success of a ubiquitous weed, we analyze genomic data from contemporary paired natural and agricultural populations alongside historical herbarium samples collected from 1828 until 2011 (Fig. 1). With this design, we identify candidate agriculturally adaptive alleles (i.e., those that occur at consistently higher frequencies in agricultural compared with nearby natural sites), track their frequencies across nearly two centuries, and link the tempo of weed adaptation to demographic changes and key cultural shifts in modern agriculture.

The genome-wide signatures of agricultural adaptation

To find alleles favored under current farming practices, we looked for those that were consistently overrepresented in extant populations collected in agricultural habitats compared with neighboring riparian (natural) habitats (25) using Cochrane–Mantel–Haenszel (CMH) tests (Fig. 2A). Alleles associated with agricultural environments [the 0.1% of single-nucleotide polymorphisms (SNPs) with the lowest CMH *P* values; *n* = 7264] are significantly enriched for 29 gene ontology (GO) biological process terms related to growth and development; reproduction; cellular metabolic processes; and responses to abiotic, endogenous, and external stimuli, including responses to chemicals (table S1). The importance of chemical inputs in shaping weed agricultural adaptation is clear in that the most significant agriculturally associated SNP (raw *P* value = 8.6 × 10−11; [false discovery rate (FDR)-corrected] *Q* value = 0.00062) falls just 80 kb outside the gene protoporphyrinogen oxidase (PPO)—the target of PPO-inhibiting herbicides (Fig. 2B). PPO-inhibiting herbicides were widely used in the 1990s and have had a recent resurgence to control and slow the spread of glyphosate-resistant weeds (26, 27). Other genes with the strongest agricultural associations include *ACO1*, which has been shown to confer oxidative stress tolerance (28); *HB13*, involved in pollen viability (29) as well as drought and salt tolerance (30); *PME3*, involved in growth through germination timing (31); *CAM1*, a regulator of senescence in response to stress (32, 33); and both *CRY2* and *CPD*, two key regulators of photomorphogenesis and flowering through brassinosteroid signaling (34–37) (table S2). These changes are consistent with agricultural selection to grow in high-stress and high-disturbance environments among fast-growing crops. Natural-versus-agricultural allele frequency differentiation (*F*ST) is highly correlated with the CMH test statistic (Pearson’s correlation coefficient *r* = 0.987), with 78% (98%) of CMH focal SNPs overlapping with the top 0.01% (0.1%) of *F*ST hits (fig. S1). Despite negligible genome-wide differentiation among environments suggesting widespread gene flow (*F*ST = 0.0008; with even lower mean *F*ST between paired sites = −0.0029; Fig. 2C), our results suggest that strong antagonistic selection acts to maintain spatial differentiation for particular alleles, with 403 SNPs showing a CMH *Q* < 0.10 (251 after aggregating linked SNPs).

To further investigate the extent to which herbicides shape adaptation to agriculture, we assayed patterns of environmental differentiation by known herbicide-resistance variants. Eight such alleles were present in contemporary samples, only six of which were common (table S3): a deletion of codon 210 within *PPO* (38), a copy number amplification and a nonsynonymous mutation within 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*) conferring resistance to glyphosate herbicides (39), and 3 separate nonsynonymous mutations within acetolactate synthase (ALS) conferring resistance to ALS-inhibiting herbicides (19). Although these resistance alleles were at intermediate frequencies in agricultural populations, ranging from 0.08 to 0.35, they tended to be rarer but still frequent in natural populations, ranging from 0.04 to 0.22 (Fig. 2C). Four of six common resistance alleles
show significant allele frequency differences across environments (EPSPSamp: $F = 9.02$, $P = 0.008$; PPO210: $F = 53.39$, $P = 1.04 \times 10^{-11}$; ALS374: $F = 4.95$, $P = 0.028$; ALS376: $F = 4.37$, $P = 0.038$), two of which are among the strongest signals of differentiation genome-wide. Natural-versus-agricultural $F_{ST}$ at the PPO210 deletion, 0.21, is higher than anywhere else in the genome and is even stronger when calculated within population pairs ($F_{ST} = 0.27$) (Fig. 2C). Similarly, the EPSPS amplification is ranked 20th among genome-wide biallelic $F_{ST}$ values at 0.14 (within-pair $F_{ST} = 0.22$), in support of herbicides as a foremost driver of agricultural adaptation (Fig. 2D).

To infer the importance of selective trade-offs in adaptation across natural and agricultural environments, we implemented a Wright-Fisher allele frequency–based migration-selection balance model for these four differentiated resistance alleles as well as the top 30 independent CMH outliers. Assuming that these alleles are at a steady state between migration and selection, we inferred that the costs of resistance per migrant that has arrived into natural environments are consistently higher than the benefits of resistance per migrant that has arrived into agricultural environments (permigrant cost-to-benefit ratio ranges from 1.39 for EPSPSamp to 5.03 for the PPO210 deletion; Fig. 2D and table S3). Thus, the spread of these four common herbicide-resistance alleles appears to be constrained either by more consistent selection against resistance in herbicide-free, natural environments or by particularly high rates of migration of susceptible alleles from natural into agricultural environments. In comparison, for the top 30 independent CMH outliers, the costs per migrant that has arrived in natural environments were about equally likely to be stronger or weaker (12/28, 42%) than the benefits per migrant in agricultural environments (fig. S2). This population genetic inference provides a previously unused and sensitive alternative to experimental studies of fitness costs that vary greatly depending on context (40), highlighting the potentially important role of resistance costs across a diverse set of individuals within complex agricultural and natural environments. In these field settings, further work is necessary to understand the contributions of temporal and spatial heterogeneity in both migration and selection for and against resistance across the landscape.

**Agriculturally adaptive alleles change rapidly**

With the genome-wide set of 251 modern agriculture-associated alleles, we searched for signatures of temporal evolution using newly collected whole-genome sequence data from a set of historical herbarium samples ($n = 108$) dating back to 1828. These samples provide snapshots of the genetic changes that have occurred over this time period and across environment types, with collections from natural and weedy (agricultural and disturbed) habitats (Fig. 1). Of the 165 loci for which we had sufficient information in the historical SNP set (sequenced to 10× coverage, on average), 151 were segregating with the same reference-alternate allele combination (i.e., 11 were dropped because of multiallelism), and only three were invariant. To model allele frequency change through time at these alleles, we implemented logistic regressions of genotypes (within-individual allele frequencies) at each locus by collection year, where twice the slope of the logit-transform is equivalent to the strength of selection ($s$) in a diploid model of selection (where $s$ is the fitness difference between homozygotes, assuming additivity; see materials and methods for model and simulations (41)).

Consistent with the rapid change in land use and farming practices in the recent past, the frequency of these 154 contemporary agricultural alleles has increased substantially over
the past two centuries. Whereas in natural environments agriculturally associated alleles have increased by 6% on average since 1870—the earliest time point at which we have collections across environment types—these same alleles have increased by 22% in disturbed and agricultural environments (Fig. 3A). This observed change greatly exceeds the expected change over this time period, based on genomewide patterns that reflect drift, migration, selection, and demographic change [null 95% interquantile range for allele frequency change in natural sites = (−2.7, 2.0%); for change in agricultural and disturbed sites = (3.3, 7.9%)]. We generated these null expectations by randomly sampling a set of 154 loci with the same distribution of contemporary allele frequencies (fig. S4) and calculating their frequency change through time across herbarium samples, separately in each environment, 1000 times [see materials and methods (4)]. That the observed change in natural environments is also more extreme than what is expected is consistent with ongoing migration of agriculturally selected alleles and subsequent costs in natural environments.

The considerable increase in frequency of these alleles across environments corresponds to notably strong selection, even when estimated over century-long time periods. The 154 agriculture-associated alleles collectively exhibit a selective strength of $s = 0.022$ since the 1870s in agricultural and disturbed habitats. However, these alleles exhibit much weaker selection, $s = 0.0056$, in natural habitats [agricultural and disturbed null interquantile range = (0.0026, 0.0068); natural null interquantile range = (−0.0018, 0.0018)]. An open question in evolutionary biology is what distribution of selection coefficients underlie adaptation (42). We estimate that selection on agricultural-associated loci varies between −0.196 and 0.150 in natural habitats and between −0.090 and 0.372 in agricultural and disturbed habitats, reflective of left- and right-skewed distributions, respectively (Fig. 3B and fig. S5). The top 15 agriculture-associated alleles that we infer have experienced the strongest selection over the past ~150 years include SNPs that map near $PPO$, $ACO1$, $CCB2$, $WRKY13$, $BPL3$, and $ATPD$ (table S4). We find that both the total frequency change of agriculture-associated alleles and the estimated strength of selection in agricultural and disturbed environments are positively correlated with the extent of contemporary linkage disequilibrium around these loci (the number of SNPs within 1 Mb with $r^2 > 0.25$) (frequency change: $F = 5.16$, $P = 0.024$, $r = 0.12$; strength of selection: $F = 3.99$, $P = 0.048$, $r = 0.058$; fig. S6), consistent with theoretical expectations for the genomic signatures of recent positive selection (43, 44).

We next investigated how well the trajectory of modern agricultural alleles reflects the rise of industrialized agricultural regimes over the past century. When we split our samples into those that predate versus those that come after the intensification of agriculture during the Green Revolution, we find that the increase in frequency of agricultural alleles was negligible in agricultural and disturbed environments before the 1960s (predicted 1870-to-1960 change = 0.005). By contrast, change subsequent to 1960 nearly completely accounts for the observed rise in frequency of modern agricultural alleles (predicted 1960-to-2018 change = 0.219 versus total 1870-to-2018 change = 0.221) (Fig. 3C). Corresponding estimates of selection by logistic regression using only data from before 1960 show no evidence of selection on these loci in disturbed and agricultural habitats [$s = 0.0008$, null interquantile range = (−0.0044, 0.0020)] or in natural habitats [$s = 0.0006$, null interquantile range = (−0.004, 0.004)]. However,
samples collected after 1960 reflect a marked shift in selection—a collective $s = 0.054$ in disturbed and agricultural environments and a collective $s = 0.028$ in natural environments [agricultural and disturbed null interquantile range $= (0.0064, 0.0020)$; natural null interquantile range $= (-0.0056, 0.0054)$] (Fig. 3C and fig. S8). Together, these results suggest that although most contemporary agricultural alleles were present in historical populations, these alleles only became associated with agricultural and human-managed sites over the past century, on time scales and at rates consistent with the rapid uptake and intensification of agrochemicals, controlled irrigation, and mechanization in agriculture.

The historical trajectory of known herbicide-resistance alleles epitomizes extreme selection over the past 50 years (Fig. 3D). Five of seven known biallelic herbicide-resistance alleles present in our contemporary, paired-environment collections are absent from our historical samples, consistent with the suggested importance of resistance adaptation from de novo mutation (13, 45) and a particularly recent increase in their frequency. Only 3 of 108 historical samples show variation for herbicide resistance—two samples homozygous for resistance at ALS574 and one heterozygous for resistance at ALS122—all of which were sampled after the onset of herbicide applications in the 1960s (Fig. 3D). Resolving the very low historical and much higher contemporary frequencies of resistance, we estimate that since the approximate onset of herbicide use in 1960, these seven resistance alleles have collectively experienced a selective strength of $s = 0.198$ (logistic $Z$ value $= 2.11, P = 0.035$) per year across environment types. Maximum likelihood–based estimates of selective strengths for each resistance allele are significant for five of the seven and are strongest for PPO210 ($s > 0.194$), EPSPS106 ($s > 0.106$), and ALS574 ($s > 0.088$) (Fig. 3D and table S3).

**Concurrent temporal shifts in ancestry underlie agricultural adaptation**

Finally, we explored whether historical demographic change over the past two centuries has played a role in agricultural adaptation. Early taxonomy described two different *A. tuberculatus* varieties as separate species, with few distinguishing characteristics [seed dehiscence and tepal length (16)]. Sauer’s 1955 revision of the genus, which used herbarium specimens to gauge the distribution and migration of congeners over the past two centuries (46), led him to describe an expansion of the southwestern var. *rudis* type [at the time, *A. tamariscinus* (Sauer)] northeastward into the territory of var. *tuberculatus* (*A. tuberculatus* (Sauer)) sometime between 1856 to 1905 and 1906 to 1955. Our sequencing of >100 herbarium
samples dating back to 1828, combined with 349 contemporary sequences (25, 47), allowed us to directly observe the change in the distribution of these two ancestral types, adding further temporal resolution to Sauer’s morphological observations of the species’ range shifts, and to assess the role of agriculturally adaptive standing genetic variation across varieties.

Range-wide, we see clear shifts in the distribution of var. *rudis* ancestry based on fastSTRUCTURE (48) inference at \( K \) (number of subpopulations) = 2 (fig. S9) across three time spans—1830 to 1920, 1920 to 1980, and 1980 to 2018 (time span: \( F = 5.47, P = 0.0045 \))—and particularly so in the east (time span × longitude: \( F = 5.49, P = 0.0045 \)), consistent with a recent expansion of var. *rudis* ancestry (Fig. 4A). Furthermore, we see strong state- and province-specific shifts in ancestry through time in our historical sequences (time span by state interaction: \( F = 4.22, P = 7 \times 10^{-5} \)), highlighting not only the shift of var. *rudis* eastward (with increases through time in Ontario, Ohio, Illinois, and Missouri) but also the recent introduction of var. *tuberculatus* ancestry into the most western part of the range in Kansas (Fig. 4B). *A. tuberculatus* demography thus appears to have been markedly influenced by human-mediated landscape change over the past two centuries, consistent with the massive recent expansion of effective population size that we had previously inferred from contemporary samples over this same time frame (45). That this shift has been most notable over the past 40 years is further consistent with the

Fig. 4. Temporal shifts in the distribution of var. *rudis* ancestry have facilitated polygenic agricultural adaptation. (A) Longitudinal clines in var. *rudis* ancestry over three time spans, illustrating the expansion of var. *rudis* ancestry eastward over the past two centuries. In (A) to (C), dots represent individual-level ancestry estimates. (B) The distribution of individual-level var. *rudis* ancestry by state and through time, illustrating state-specific changes in ancestry. Horizontal lines within each distribution represent first and third quartiles of ancestry. (C) Increasing sorting of individual-level var. *rudis* ancestry into agricultural environments (Env.) on contemporary time scales. (D) Environment-specific metrics of selection (CMH P value and cross-population extended haplotype homozygosity (XPEHH)) across the genome in 100-kb windows positively correlate with var. *rudis* ancestry in agricultural but not natural habitats (XPEHH by environment: \( F = 10.97, P = 9.3 \times 10^{-4} \); CMH by environment: \( F = 108.51, P < 10^{-16} \)).
time scale of agricultural intensification, shifts toward conservation tillage, and rampant herbicide-resistance evolution within the species (19, 45, 49, 50), which suggests that selection may facilitate the colonization of var. rudis ancestry outside its historical range. Along these lines, we find that this contemporary range expansion has facilitated the sorting of var. rudis ancestry across environments (a longitude by time span by environment interaction: \(F = 5.13, P = 4 \times 10^{-2}; \) Fig. 4C), with increasing overrepresentation of var. rudis ancestry in agricultural and disturbed environments in the eastern portion of the range through time, as has been previously suggested (25).

To investigate whether agricultural adaptation has drawn disproportionately from var. rudis ancestry, we reconstructed fine-scale ancestry across the genome. On the basis of analyses in 100-kb windows, we find a least squares mean of 5.5% (95% confidence interval (CI) = (5.0, 5.9%)) more var. rudis ancestry genome-wide in agricultural environments compared with the adjacent natural habitat (fig. S10). The environment-specific proportion of var. rudis ancestry is not only positively correlated with recombination rate (\(F = 18.85, P = 1.4 \times 10^{-16}; \) fig. S10) and gene density (\(F = 8.53, P = 0.004, r = 0.050 \)) but also with SNP and haplotype-based evidence of environment-specific selection. Agricultural but not natural populations have an excess of cross-population haplotype homozygosity (agricultural versus natural XPEHH) and within-environment differentiation (CMH P value) in genomic regions of high var. rudis ancestry (XPEHH) by environment: \(F = 10.97, P = 9.3 \times 10^{-16}; \) CMH by environment: \(F = 108.51, P < 10^{-16}; \) Fig. 4D), which implies that ancestry composition genome-wide in large part determines the extent of polygenic agricultural adaptation. These findings suggest that the expansion of var. rudis ancestry across the range, particularly in the past 40 years, has facilitated waterhemp’s success in agricultural habitats by providing access to preadapted, standing genetic variation.

**Discussion**

Agricultural adaptation in *A. tuberculatus*, a weed plant in North America, has occurred over extremely short time scales, facilitated by range shifts in response to the agriculturalization of its native habitat. The human-mediated expansion of the southwestern lineage of the species northeastward since the latter half of the 20th century has introduced new genetic variation across the range, on which selection in agricultural settings could act. Negligible genome-wide differentiation across habitats in this species refutes the idea of agricultural populations existing as separate to natural ecosystems (51). Despite substantial gene flow, the prevalence of agricultural alleles has increased rapidly since the intensification of agriculture over the past 60 years—in agricultural environments by a selection coefficient of nearly 6% per year and even in natural habitats by >2% per year. The selective intensity of industrial agriculture is on par with the intensities that *Arabidopsis* populations in extreme hot and dry environments are predicted to face by 2070 under the high-emissions scenario of climate change (32). The effects of agricultural herbicides are even more extreme—range-wide, evolved resistance mutations have experienced selective strengths of 20% on average per year since herbicides were first introduced—permeating even into natural habitats.

Although modern, industrial agriculture imposes strengths of selection rarely observed in the wild, *A. tuberculatus* has in turn escalated the weed management-evolution arms race through a multitude of interdependent mechanisms: range expansion, polygenic adaptation from standing genetic variation, and large-effect herbicide-resistance mutations. Together, these results highlight that anthropogenic change not only leads to the formation of new habitats but also provides an opportunity for range expansion that may facilitate and interact with local adaptation, reshaping genetic variation for fitness within native species.

**REFERENCES AND NOTES**

16. J. D. Sauer, Madroño 13, 5–16 (1956).
interests. Data and materials availability: All new sequence data have been archived at the NCBI Sequence Read Archive (BioProject ID PRJNA878842). Scripts and accompanying metadata have been archived on Github (www.github.com/jkreinz/TemporalAdaptation) and on Zenodo (53). License information: Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo7293

Materials and Methods
Figs. S1 to S13
Tables S1 to S4

References (54–68)
Data S1 and S2

Submitted 25 February 2022; resubmitted 24 June 2022
Accepted 17 October 2022
10.1126/science.abo7293
Rapid weed adaptation and range expansion in response to agriculture over the past two centuries

Julia M. Kreiner, Sergio M. Latorre, Hernán A. Burbano, John R. Stinchcombe, Sarah P. Otto, Detlef Weigel, Stephen I. Wright

Science, 378 (6624), • DOI: 10.1126/science.abo7293

Selection by modern agriculture

Intensive agriculture creates extreme environmental changes, including soil disturbance, water and nutrient addition, and application of chemical pesticides. Kreiner et al. documented how these extreme environments have driven changes to a native North American agricultural weed, common waterhemp (Amaranthus tuberculatus), using genomic data from paired agricultural and natural settings, as well as historic herbarium samples from both habitat types (see the Perspective by Waselkov and Olsen). The authors found that eastward range expansion of a distinct southwestern variety introduced new genetic diversity and facilitated rapid adaptation to agricultural environments in genes related to growth, environmental tolerance, and herbicide resistance. —BEL

View the article online
https://www.science.org/doi/10.1126/science.abo7293

Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service

Science (ISSN ) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title Science is a registered trademark of AAAS.
Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works