Pest Management Science SCI

The Power and Potential of Genomics in Weed Biology and Management

Journal:	*Pest Management Science*
Manuscript ID	Draft
Wiley - Manuscript type:	Perspective
Date Submitted by the Author:	n/a
Complete List of Authors:	Ravet, Karl; Colorado State University College of Agricultural Sciences, Bioagricultural Sciences and Pest Management Patterson, Eric; Colorado State University, Bioagricultural Sciences and Pest Management Kraehmer, Hansjoerg ; Bayer CropScience, Weed Control; Hamouzova, Katerina; Czech University of Life Sciences Prague, Department of Agroecology and Biometeorology Fan, Longjiang; Zhejiang Chinese Medical University, Institute of Bioinformatics Jasieniuk, Marie ; University of California - Davis, Plant Sciences Lawton-Rauh, Amy; Clemson University College of Science, Genetics and Biochemistry; Clemson University College of Science, Clemson University, Dept Genetics & Biochemistry Malone, Jenna; University of Adelaide, School of Agriculture, Food & Wine McElory, Scott; The University of Tennessee, Knoxville, Plant Sciences Department Merotto, Aldo; Universidade Federal do Rio Grande do Sul, Department of Crop Science Westra, Philip ; Colorado State University, Bioagricultural Sciences and Pest Management Preston, Christopher; The University of Adelaide, School of Agriculture, Food & Wine Vila-Aiub, Martin; IFEVA-University of Buenos Aires, Faculty of Agronomy Busi, Roberto; University of Western Australia, School of Plant Biology Tranel, Patrick; University of Illinois, Crop Sciences Reinhardt, Charlie; University of Pretoria, Department of Plant Production and Soil Science Saski, Christopher; Clemson University, Institute for Translational Genomics Beffa, Roland; Bayer CropScience AG, Weed Control Neve, Paul; University of Warwick, Warwick HRI Gaines, Todd; Colorado State University, Bioagricultural Sciences and Pest Management
Key Words:	genomics, weed biology, weed management, weed evolution, herbicide resistance, weedy species

The Power and Potential of Genomics in Weed Biology and Management

Running title: Power and Potential of Weed Genomics

Karl Ravet[1], Eric L. Patterson[1], Hansjörg Krähmer[2], Kateřina Hamouzová[3], Longjiang Fan[4], Marie Jasieniuk[5], Amy Lawton-Rauh[6], Jenna M. Malone[7], J. Scott McElroy[8], Aldo Merotto Jr[9], Philip Westra[1], Christopher Preston[7], Martin M. Vila-Aiub[10], Roberto Busi[11], Patrick J. Tranel[12], Carl Reinhardt[13], Christopher Saski[14], Roland Beffa[2], Paul Neve[15,*], and Todd A. Gaines[1,*]

[1] Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO 80523, USA.

[2] Bayer AG, Industriepark Hoechst, D-65926 Frankfurt am Main, Germany.

[3] Department of Agroecology and Biometeorology, Czech University of Life Sciences Prague, Prague, Czech Republic.

[4] Institute of Crop Science & Institute of Bioinformatics, Zhejiang University, Hangzhou 310058, China.

[5] Department of Plant Sciences, University of California - Davis, Davis, CA 95616, USA.

[6] Department of Genetics and Biochemistry, 316 Biosystems Research Complex, Clemson University, Clemson, SC 29634-0381.

[7] School of Agriculture, Food & Wine, University of Adelaide, Glen Osmond, SA 5064, Australia.

[8] Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, AL 36849, USA.

[9] Department of Crop Sciences, Agricultural School, Federal University of Rio Grande do Sul, Porto Alegre, RS. Brazil.

[10] Facultad de Agronomía, Departamento de Ecología, IFEVA-CONICET, Universidad de Buenos Aires, Buenos Aires 1417, Argentina

[11] Australian Herbicide Resistance Initiative, School of Agriculture and Environment, University of Western Australia, Crawley, WA 6009, Australia.

[12] Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA

[13] Department of Plant and Soil Sciences, University of Pretoria, Hatfield, Pretoria 0028, South Africa.

[14] Clemson University Genomics and Computational Biology Lab, Clemson University, Clemson, SC 29634-0323.

[15] Biointeractions & Crop Protection Department, Rothamsted Research, West Common, Harpenden, Hertfordshire, AL5 2JQ, UK.

*Corresponding authors: Todd Gaines, todd.gaines@colostate.edu; Paul Neve, paul.neve@rothamsted.ac.uk.

ABSTRACT

There have been previous calls for, and efforts focused on, realizing the power and potential of weed genomics. Sustained advances in genome sequencing and assembly technologies now make it possible for individual research groups to generate reference genomes for multiple weed species at reasonable costs. Here, we present the outcomes from several meetings, discussions, and workshops focused on establishing an International Weed Genomics Consortium (IWGC) for a coordinated international effort in weed genomics. We review the 'state of the art' in genomics and weed genomics, including technologies, applications, and on-going weed genome projects. We also report the outcomes from a workshop and a global survey of the weed science community to identify priority species, key biological questions, and weed management applications that can be addressed through greater access to genomic resources. Major focus areas include the evolution of herbicide resistance and weedy traits, the development of molecular diagnostics, and the identification of novel targets and approaches for weed management. There is increasing interest in, and need for, weed genomics, and the establishment of the IWGC will provide the necessary global platform for communication and coordination of weed genomics research.

Keywords: genomics, weed biology, weed management, weed evolution, herbicide resistance, weedy species

1. INTRODUCTION

In their 2017 State of the World's Plants report, Kew Gardens estimated that whole genome sequences are now available for 225 plant species.[1] Of sequenced vascular plants, 58% were crop species, 18% were crop wild relatives and 22% were model species and their wild relatives. Clearly, the commercial and societal value of plants for providing food, materials, fiber, energy, and medicinal products has been a major motivating factor in plant genomics efforts. However, it is notable that the weeds that compete with these crops, resulting in an average 30% annual yield loss across several crops,[2] are currently under-represented.

The power and potential of weed genomics to provide biological insight into the discovery of new herbicide targets and new approaches in weed management has long been recognized. A symposium addressing the potential impacts of biotechnology and genomics in weed science was held at the Weed Science Society of America (WSSA) symposium in Toronto in 2000.[3] Here, various authors considered the potential of genomics for discovering new herbicide targets,[4] providing insights into weed diversity[5] and for generating wider traction in evolutionary ecology of weeds and management research, including studies of weedy traits, invasiveness, seed dormancy, allelopathy, biological control, gene flow, and introgression.[6]

At the WSSA meeting in 2007, an Emerging Technologies Symposium[7] considered how advances in genomics, transcriptomics, proteomics and bioinformatics might be applied to studies of weed biology. Many others have considered the potential of weed molecular biology[8] and weed genomics[9-11] to contribute to advances in basic and applied weed science and called for the pooling of resources towards a community effort to generate genomic data and resources for major weed species.[12] Sustained advances in sequencing technologies exponentially increase the rate at which genome sequence data in non-model organisms can be generated at rapidly diminishing costs. Considering these advances, it is timely to review the aspirations and potential for an international, community-based effort by weed scientists to sequence the genomes of important, global weed species.

Here, we review the global status of weed genomics research, the impacts that new sequencing technologies have on the availability of genomic data from weeds and present the outcomes of several recent international workshops, discussions and surveys that have attempted to visualize a global effort in weed genomics, through the auspices of an International Weed Genomics Consortium (IWGC). Whilst these efforts remain preliminary, we feel it is timely to present our first efforts to determine global priorities for weed genomics, including major species to be sequenced, biological questions and models to be addressed and the novel weed management tools, resources, targets, and approaches that may arise from such an effort.

2. THE STATE OF THE ART IN GENOMICS AND WEED GENOMICS

Discovery in genomics is primarily being driven by advances in sequencing technologies. When second (next) generation sequencing was first developed, the cost per base pair plummeted and opened genomics and transcriptomics[13,14] to non-model organisms. Third generation sequencing technologies that provide the ability to sequence long DNA molecules (> 5kb) are highly advantageous for weed genome sequencing efforts.[15,16] The production of longer individual sequences imparts more complete and accurate recapitulation of complex regions of the genome. Reads from third generation sequencing technologies are long enough to span most repetitive regions [simple sequence repeats (SSRs), tandem DNA arrays, long stretches of homo-polymers, low complexity sequence, repetitive DNA elements, telomeres, etc.] and can therefore resolve these areas and accurately scaffold large contiguous DNA sections (contigs). Previously, short-read-only assemblies (so called “shotgun” assemblies) could not resolve complex regions and assemblies remained highly fragmented, regardless of the amount of short-read data used in their generation.[17,18] Highly repetitive, gene-poor regions such as centromeres and telomeres remain difficult to assemble, even with long-read data, and still require other strategies such as Hi-C sequencing, mate-pair sequencing, linkage mapping, and optical mapping to construct sequences for entire chromosomes. Computational assembly algorithms have also advanced to handle hybrid data sets (2nd and 3rd generation), and can also manage higher amounts of heterozygous sites and the ability to split haplotypes, which is crucial for outcrossing species.[19,20] Once complete, reference genomes become valuable tools for studying structural variation, DNA rearrangements, and polyploid genome evolution.[21] These structural rearrangements are essential for understanding evolution, domestication, phylogeny, reproduction, invasiveness, and herbicide resistance in weedy species. In short, technological advancements have drastically enabled genome reconstruction efforts to deliver more complete, contiguous, and accurate genome assemblies for more complex species.

Recently, the sequencing of complex crop genomes has been completed for multiple species, e.g. sugar beet, quinoa, grain amaranth, millet, and sorghum.[22-26] These results greatly aid efforts to sequence weed genomes as some of these crops have weedy relatives (Figure 1; grain amaranth, foxtail millet, quinoa, perennial ryegrass, etc.). One of the most ambitious recent genomes to be published is the hexaploid wheat genome (15.3 Gb).[27] The wheat sequencing effort makes it clear that almost no genome is out of reach for a dedicated research group.

Despite the increasing ease with which long-read sequence data can be generated and assembled, few economically important weed genomes have been fully sequenced. Numerous *de novo* transcriptomes have been completed in weed species for herbicide resistance studies[28] and to study weed biology and evolution.[29,30] Weed genome assemblies have been published for *Conyza canadensis*,[31] *Echinocloa crus-galli*,[32] *Thlaspi arvense*,[33] *Raphanus raphanistrum*,[34] and *Lolium perenne*;[35] however, these genomes remain fragmented and are not presented as chromosome-scale pseudomolecules, making some types of analysis impossible. Importantly, these three projects are far from representing the broad diversity of weed species and they remain relatively isolated efforts. Currently, several projects are underway to sequence additional weed species, including *Amaranthus palmeri* (Palmer amaranth), *A. tuberculatus* (waterhemp), *Echinochloa colona* (junglerice), and *Kochia scoparia* (kochia). These assemblies have not yet been completed nor made publicly available. There likely are additional weed genome sequencing efforts underway of which we are not aware. Furthermore, it is likely that multiple groups are independently sequencing the genomes of the same species, which is an inefficient use of resources and could be better coordinated via the IWGC platform.

Looking forward, several new weed genome sequencing projects will be initiated and the pace at which sequence data becomes available to weed researchers will accelerate dramatically. Indeed, it may soon be possible for researchers to move beyond sequencing a single individual to embark on pangenomics projects[36] that aim to elucidate gene and genome diversity across a species range, and provide opportunities for studies into weed population genomics and eco-evolutionary dynamics.

3. BUILDING A GLOBAL WEED GENOMICS COMMUNITY

We believe the weed science community is now ready to successfully engage in a community-based approach to weed genomics through the IWGC. Firstly, advances in sequencing technology have reduced the cost of *de novo* genome assembly such that multiple weed genomes may be successfully completed. Secondly, a critical mass of interested and motivated scientists from academia and industry has been prompted to both drive the genomics effort and to utilize the resulting genomics resources. Here, we report on recent efforts to launch the IWGC.

Initially, a "Grass and weed genomics workshop" held in Prague at the 7th International Weed Science Congress in 2016 (IWSC; http://iwss.info) brought together over 30 weed scientists from academia and industry. This workshop aimed to determine global priorities for weed biology and management and highlighted the importance of developing weed genomics databases and skills to address those priorities in key weed species. A fundamental output of the workshop was the establishment of an international working group to develop a concerted effort towards developing weed genomics.

Following the IWSC 2016, this working group, which included members from North America, South America, Europe, Africa, Asia and Australia, met at the first official IWGC workshop at Rothamsted Research (UK) in March 2017. The aim of this meeting was to better define the overall effort including the organization, structure, objectives and financing of the consortium. The IWGC concept was then presented to the weed science community during a keynote session at the Global Herbicide Resistance Challenge in Denver in May 2017.[37] A workshop that addressed the question 'how do we bridge the gap between weed genomics and weed management' was also held at the conference[38] with the intention of identifying key applications of weed genomics and activities of the IWGC (see below). These meetings were followed by an online survey of the weed science community to assess the level of interest in the IWGC concept (see below).

The main objectives of the IWGC are to obtain reference genomes for the most important weedy species globally, to provide open-access to the data, and to offer user-friendly genome analytical tools and training. The expectation is that the IWGC and its associated website will become a central resource not only for a broad array of scientists with diverse areas of expertise and interests around the globe, but will also represent a key platform for stimulating 1) data sharing, 2) partnerships between academia and industry, 3) collaborations between international research groups, 4) education of the next generation of weed scientists, 5) the transfer of knowledge and experience to developing countries and 6) an open forum for discussion.

4. FOCAL SPECIES FOR A GLOBAL WEED GENOMICS INITIATIVE

Initial feedback on the IWGC proposal highlighted the importance of careful selection of 'model' weed species for genome sequencing. A survey circulated amongst the weed science community following the GHRC addressed this point through two questions. The survey consisted of 245 participants that represented a range of geographic regions and research interests (Supporting Information Figures 1 and 2). Participants were asked to choose three priority species from amongst a pre-selected list of 10 species that could be targets for future sequencing efforts, based on the output of previous workshops and discussions, and for which genome assembly projects were not completed or known to be in progress as of March 2017. For example, *Amaranthus palmeri* and *Conyza canadensis* were not included in the pre-selected list since their genomes were sequenced or in progress at the time of the survey. As this list of species may not have reflected the needs of the entire community, a second question offered the possibility of nominating two additional unlisted weed species.

In summary, responses to these two questions revealed a consensus towards sequencing weed models within the genera of *Conyza*, *Sorghum*, *Poa*, *Lolium, Amaranthus*, *Echinochloa,* and to a lesser extent *Alopecurus*, *Eleusine* and *Digitaria* (Figures 2, 3). The analysis of preferences based on the 10 listed species (Figure 2) revealed two species that were globally important: *Lolium rigidum* (rigid ryegrass, 143 votes) and *Conyza bonariensis* (hairy fleabane, 110 votes). A cluster of four species, namely, *Amaranthus tuberculatus* (waterhemp, 86 votes), *Sorghum halepense* (Johnsongrass, 84 votes), *Alopecurus myosuroides* (blackgrass, 62 votes) and *Digitaria sanguinalis* (hairy crabgrass, 55 votes) were viewed as next in importance. Respondents also identified several additional genera, consisting of multiple target weed species (Figure 3). Several species appeared in the second list (Figure 3A) for which genomes are in progress or complete, such as *Amaranthus palmeri* and *Conyza canadensis*, indicating their importance to the weed science community and a need for improved communication about the status of genome sequencing projects in weeds.

Fortunately, several of the species identified as high priority have closely related crop relatives whose genomes and genome annotations (Figure 1) will be highly useful for annotating new weed genome assemblies. Additionally, genomes of several of the identified priority species or related species within the same genus are in progress and/or published, including the published *C. canadensis*[31] and *E. crus-galli*[32] genomes and the sequencing of BACs from *A. palmeri*,[39] with a whole genome assembly in progress.

5. WEED GENOMICS PROVIDES NOVEL INSIGHT IN WEED BIOLOGY

As mentioned above, a workshop was also held at the GHRC to examine and discuss application domains (weed biology and weed management) for the IWGC.[38] Prior to the workshop, participants were asked to submit up to five biological questions and/or weed management applications where they thought weed genomics could or should make a significant contribution. In total, 92 questions were submitted. These questions were analyzed to identify major emerging themes and areas where potential new insights and advances could be made, given access to weed genomics resources. From this analysis, several focal areas for weed biology research and for weed management application were identified (Figure 4). These areas are discussed below. Broadly, three areas in which weed genomics can make significant advances were identified: (i) understanding of the fundamental molecular, physiological, genetic, ecological and evolutionary processes that underlie weed adaptation (*basic plant biology*), (ii) insights into new targets and new approaches for weed management (*translational plant science*) and (iii) management strategies that make weed adaptation (*applied evolutionary biology*) more difficult, or slower, to evolve.

5.1.Herbicide resistance

To date, our understanding of the molecular basis of herbicide resistance has been largely informed by single-gene sequencing and identification of single-point mutations causing target-site resistance (TSR). More recently, second-generation sequencing technologies have also enabled transcriptomic approaches (e.g., RNA-Seq) to identify candidate genes for more complex non-target-site resistance (NTSR mechanisms, such as changes in herbicide metabolism and translocation). Genomics offers the promise to go beyond transcriptomes to provide further novel insights into the genomic basis of complex resistance traits. Moreover, as NTSR genes, such as cytochrome P450 genes,[40] are identified, functional genomics will enable researchers to address why differential regulation of TSR and NTSR genes occurs (regulatory mechanisms), what types of mutations produce this adaptive molecular variation, and other basic questions about the processes involved in resistance evolution.

Gene expression can be controlled by a variety of mechanisms; however, mutations in *cis* and *trans* regulatory elements and in the transcription factors that bind them can only be understood with genomic tools. Additionally, gene copy number variation (CNV) and the resulting changes in gene expression have been shown repeatedly to be important herbicide resistance mechanisms in multiple species for different resistance traits.[41,42] Genomic and computational resources are essential to answer questions related to CNVs. Thinking even further ahead, epigenetic mechanisms may be playing a yet unknown role in herbicide resistance[43] and adaptive evolution in weeds, yet without genomic resources, all work on epigenetics is currently restricted to model species with reference genomes, such as *Arabidopsis*. Research into CNVs and epigenetics is expected to generate new hypotheses of whether herbicide resistance mechanisms are linked to abiotic stress responses, such as tolerance to flooding, drought, heat, cold, or xenobiotics. Having multiple weed genomes will also enable functional validation of candidate genes.

Some important research needs in weed biology and herbicide resistance evolution identified in the GHRC 2017 workshop include 1) identifying the original source and amount of novel molecular variation in NTSR genes; 2) determining whether NTSR mechanisms are linked to stress responses, such as flooding, drought, heat, or cold tolerance; 3) discovering whether NTSR genes are genetically linked and/or co-regulated with such stress response pathways; 4) asking if pre-adaptation for NTSR occurs in weed populations that have adaptations to stressful environments and 5) determining why enhanced metabolic resistance appears to be relatively common in grass weeds, but relatively less common in dicot weeds.

5.2. Weed evolution

Over 60 years ago, Harper (1956)[44] surmised that 'arable weeds constitute an ecological group …… that have been selected by the very practices that were originally designed to suppress them'. The relative importance of 'general-purpose' genotypes[45], phenotypic plasticity and rapid, ongoing adaptation for explaining the prominence of some plant species as agricultural weeds has been the subject of ongoing debate,[45-47] leading to a recognition that weed species may be ideal models for studying adaptation in plants.[48] Given that the evolution of herbicide resistance provides evidence of the importance of, and potential for, rapid weed adaptation, it is likely that ongoing selection for other weedy traits is a pervasive force that impacts all weed management efforts.

Many weedy traits (including non-target site herbicide resistance) have complex genetic architecture, and understanding the potential for evolution of those traits in the face of novel management and environmental challenges requires knowledge of the additive genetic variation that underlies traits and the resulting trait heritability. Approaches based on quantitative genetics and population genomics, including genome-wide association studies (GWAS) and whole genome diversity scans, can deliver increased power to unravel the genetic basis of complex traits in weeds, their phylogenetic histories, and the demographic and population genetic processes that mediate responses to environment- and management-based selection pressures. These approaches will be enabled by greater access to genomic resources for weed species.

As the costs decline and accessibility increases for genomic data in non-model organisms, many commentators envisage that we are rapidly moving towards the age of pangenomics,[36] where genome sequencing efforts are focused on multiple individuals and populations to capture the full range of genetic diversity within a species. This may be particularly important for species where understanding adaptation is a major focus (such as weeds), as there is an increasing realization of the importance of gene duplication, genomic rearrangements, and neo-functionalization in rapid plant adaptation to environmental stress.[41,49-51] Whilst these pangenomics approaches may be a longer-term aspiration for the IWGC, they can only be possible by initial access to reference genomes for key weed species.

5.3.Weedy traits and stress tolerance

Through natural and human-mediated selection, weeds have been and continue to be selected for optimal fitness in agricultural environments. Since weeds are not constrained by conscious breeding efforts that may deliberately select for only a few specific traits, especially related to yield and pest resistance, natural selection acts on a host of weedy traits that make weeds well-adapted to compete with crops. Important weedy traits include prolific seed production (high rates of population increase), extended seed dormancy, large dispersion rates, adaptive germination traits, increased abiotic stress tolerance, and high competitiveness. Additional weedy traits of high interest for weed management and crop improvement are allelopathy and seed shattering. These life history and resistance traits have been shaped by evolution in response to human and environmental pressures and may have been fixed in populations. Identifying and understanding the genetic bases of these traits will be facilitated by access to the genome of the weed species.

In general, weeds tend to demonstrate high phenotypic plasticity, especially when subjected to biotic or abiotic stress.[52] Weeds are notoriously hardy when they are subjected to abiotic stresses such as drought and flooding, cold and heat. Thus, owing to their high genetic diversity, weeds likely outcompete crops in changing environments over space and time. Identifying the gene(s) responsible for these adaptive and stress resistance traits would open new avenues for crop improvement to breed crops that may, for example, be more resilient to climate change. In theory, genes responsible for weedy traits may be incorporated into crop varieties (via introgression if the crop and weed are sexually compatible, or via transgenic approaches where they are not) or the corresponding crop gene may be modified to mimic the function of the weed gene (via genome editing). While functional genomics in weeds remains at a preliminary stage, genomic resources for weed species developed now will pay large dividends in the years to come.

5.4. Weed taxonomy and identification

Where closely related weed species coexist, and where definitive identification via morphological traits is not possible, genomic resources may play a role in assigning species identity. This may be particularly important in instances where control options and efficacies differ between closely-related weed species and where hybrid complexes between co-occurring species have been reported. For example, two related aquatic plant species of the *Myriophyllum* genus, and their hybrid, were distinguished using three Kompetitive Allele specific PCR (KASP) markers.[53] Molecular diagnostics may also enable confirmation of species identity where novel invasions of weed species beyond their normal range are suspected. A recent example used species diagnostic KASP markers to confirm the recent invasion into Brazil of populations of *Amaranthus palmeri*, distinguishing this species from the complex of other weedy *Amaranthus* species previously documented in Brazil.[54]

Increasing access to genomic resources and sequence data for closely-related weed species will continue to enable and inform these efforts.

5.5. Weed dispersal and gene flow

Elucidating the spatial dynamics of weed dispersal (via seed, fruits, and propagules) and gene flow (via seed and pollen) is important for understanding the evolution and spread of weeds and weedy traits. This knowledge can, in turn, inform the most appropriate spatial scales for the implementation of weed management strategies. Where dispersal is limited, weed management interventions can be planned and implemented at field to farm scales, whereas for highly mobile species, a concerted, regional or landscape-based approach may be warranted. The degree to which the evolution of herbicide resistance at large-scale is determined by multiple independent evolutionary events versus rarer, isolated events with subsequent spread remains incompletely understood, and likely varies for different resistance traits and species.[55] Studies to quantify the dispersal of herbicide resistance alleles have used a variety of methods. Manipulative field experiments[56,57] and field observations of gene flow between herbicide resistant and susceptible crop varieties have been conducted.[58] Population genetics analyses have used the sequence and/or frequency of herbicide target genes,[59] AFLP markers[60] and microsatellites/simple sequence repeats.[61,62] With access to less expensive sequencing technologies it becomes possible to generate orders of magnitude more data (tens of thousands of markers) for genotyping-by-sequencing (GBS) and population genomics studies.[63] These approaches will significantly increase power to determine genetic structure, and associated gene flow and dispersal processes in weed populations. Associated with this, the greater genome coverage achieved provides extra power to determine the areas of the genome that are under selection at landscape scales in weed populations, not just for resistance to herbicides, but for weediness traits in general.

A related issue that has elicited much discussion within the weed science community has been the potential for 'transgene escape' via introgression of transgenic crops and their weedy relatives. Several mitigation strategies have been proposed, which link crop protection traits with other traits that will lower the fitness of weedy populations should introgression occur.[64] Increased access to weed genomes will facilitate efforts to identify candidate 'fitness-reducing' traits that can be coupled with crop protection traits in tandem constructs.

6. WEED GENOMICS FOR NOVEL AND IMPROVED WEED MANAGEMENT

6.1. Resistance diagnostics

Understanding the underlying genetics of herbicide resistance mechanisms and the development of diagnostic methods for those genetic traits is one immediate practical application of weed genomics. Information on the presence and frequency of herbicide resistance is most valuable when available prior to planning and making herbicide applications. Most current diagnostic procedures use either PCR-based assays to genotype for known TSR mutations, and/or directly measure herbicide metabolism using analytical procedures.[65] These diagnostics can be made faster and less expensive once the molecular variation underlying NTSR is known, i.e., specific mutations in regulatory regions or other molecular variation that can be detected using PCR. PCR-based methods for resistance diagnostics would preferably be DNA-based, as DNA is less expensive to extract and manipulate for diagnostics than RNA. Without a full understanding of intron/exon and promoter structure of a gene, these DNA based methods rely on inferences from closely related species.

Some resistance mechanisms may be amenable for protein-based detection methods using antibodies (TSR or NTSR protein overexpression). Such methods would work for mechanisms in which the abundance of a given protein (e.g., a cytochrome P450) is much higher in a resistant plant than in a susceptible plant. Antibody-based detection methods can be adapted to field applications, as a leaf can be crushed in buffer and the extraction applied to a strip or column containing the antibody for rapid detection and visualization.

It must be emphasized that DNA and protein based diagnostic methods are necessarily specific to known mutations and mechanisms. If a weed population carries an unidentified mechanism that is not tested for in the assay, the diagnosis would produce a false result of herbicide sensitivity. Biological diagnostic assays, which are independent of mechanism, are better at avoiding false results but typically require more time. The RISQ assay[66] is a current leading example of a cost-effective test that can be easily employed for a result within approximately two weeks, but necessary seeds or seedlings are not available at all times of year. A major challenge for resistance diagnostics in the years ahead is to develop simple, inexpensive, and robust molecular diagnostics that encompass all known mechanisms while somehow addressing the potential for as yet unknown mechanisms.

6.2. New targets for weed control

New herbicides are being discovered at a slowing pace and no new modes of action have been marketed in more than 30 years.[67] When new candidate molecules are found, they may not be brought to market for a range of reasons such as insufficient efficacy, narrow range of usefulness, non-selectivity, inappropriate residual activity, and/or mammalian toxicity.[68] Weed genomics can contribute to new herbicide discovery by helping to 1) identify the mode of action of new compounds with unknown targets using sequence-based approaches, 2) discover the target proteins of existing modes of action for which all molecular targets are not yet known (e.g., synthetic auxins, cellulose biosynthesis inhibitors), and 3) design the chemical structure of candidate inhibitors based on potential new molecular targets discovered in the genome. It should be noted that having genomic sequence is no guarantee of finding new herbicide targets. Gene knockout approaches for chemical discovery have been attempted using *Arabidopsis* by the chemical discovery industry, but no commercial products with new target sites have reached the market.[67] We anticipate that having the gene sequences of economically important weeds may aid these efforts. If a gene of known importance (knock-out lethal) is discovered, molecules to inhibit the specific enzyme in weeds could be developed and tested. Novel molecular targets may be discovered through computational approaches using the sequences of all expressed genes available from transcriptomes and genomes. Such genomic and computational approaches have the potential to produce novel targets more quickly and require less investment when compared to the chemical library approach currently used in industry. Additionally, candidate molecules from other crop protection or medical sectors with known targets could be evaluated against plant targets using complete proteomes available from weed genomes, as has been conducted for malarial drugs as candidate herbicides using *Arabidopsis* as a model.[69,70]

Potentially disruptive technologies such as RNAi and gene drive may provide new tools for weed management with facilitation by weed genomics. RNAi targets could function as herbicide synergists, and/or as stand-alone herbicides, depending on the efficiency of transcript silencing that can be achieved.[71] Gene drive technology[72] could be employed to reverse herbicide resistance or to spread deleterious mutations through weed populations that impact reproductive success, competitiveness or other fitness-related traits. The more complete understanding of herbicide resistance mechanisms expected to result from weed genomics could also be utilized to discover and develop chemical synergists to reverse and/or down-regulate resistance mechanisms. Finally, the discovery of novel herbicide resistance mechanisms that have evolved in weeds can be used to develop new herbicide resistance traits in crops through targeted gene editing.

6.3. Proactive resistance management

The potential for weed genomics to provide additional insight and understanding for resistance management featured heavily in responses and discussions amongst delegates at the 2017 GHRC conference. In Figure 4 (the word cloud), we distinguish participant responses between a general resistance management category and a more specific recognition of the potential for weed genomic resources to enable more predictive or pre-emptive approaches. In general, resistance management would be facilitated by, for example, access to resistance diagnostics, and through the identification of novel targets for weed control, which would enable more diverse weed management strategies and moderate selection pressures by reducing reliance on current weed management tools. As these aspects have been discussed above, here we focus on application for pre-emptive resistance management.

Access to comprehensive genome sequences for major global weeds will result in major efforts to understand the herbicide- and stress-responsive pathways that are implicated in the evolution of non-target site resistance. Based on the premise that these mechanisms evolve via selection and recombination of standing genetic variation within weed populations, it should be possible to gain a better understanding of relative resistance risks prior to, or during the early stages of selection for novel resistance traits. These insights will catalyze pre-emptive studies that assess risks of resistance for new herbicides, and even new modes of action. They may also help to address longstanding and recalcitrant questions about why some weed species are more prone to the rapid evolution of resistance than others. Increased knowledge of the underlying molecular genetic mechanisms of NTSR can inform questions about the molecular basis of cross-resistance patterns, the repeatability and genetic convergence of evolutionary outcomes and the adaptive landscape for the evolution of resistance traits.[55] This understanding will inform the rational design of herbicide mixture and rotation strategies to ensure that these do not promote the use of herbicides with cross-resistance mechanisms, as well as the desirability and practicability of current approaches that aim to 'stack' transgenic resistance traits in crop plants, often with little understanding or consideration for the potential selection of cross-resistance mechanisms. Armed with this basic knowledge, it will be possible to improve models that integrate molecular genetics, population genetic, ecology, evolutionary biology and agronomy towards the design, testing and implementation of proactive resistance management strategies.

6.4. New traits for crop improvement and crop/weed comparative genomics

Despite the continuous progress made in crop breeding, projection models for crop production hint to a progressive decline in yield for most major crops, which threatens food security globally.[51] There are two main reasons for these predictions. First, breeders have long capitalized on optimizing the combination of beneficial genes from various origins within a given species, which led to so called "gene erosion".[73] Breeders now consider that the introgression of genes from other species would greatly benefit the breeding effort. Second, modern varieties have been selected based on their performance in a specific climate, thus they are generally geographically adapted. Fluctuations in the environment can result in dramatic crop yield loss, and thus the predicted global climate change will significantly impact crop production.

Agricultural weeds and crop wild-relatives are emerging as a promising source of genetic diversity for crop improvement.[51] As previously mentioned, many traits responsible for weed competitiveness are unknown, and most of the physiological characteristics associated with weed fitness have an unknown genetic basis. Understanding the genetic mechanisms underlying the physiological processes that make weedy species so competitive in agriculture settings will provide new resources for developing new crop germplasm that can out-compete weeds and remain efficient in various climatic scenarios. Among the most interesting traits are developmental traits improving plant biomass production and seed yield, and stress tolerance traits enhancing plant adaptation to environmental challenges.

As an example, the cultivated rice/weedy rice system represents an excellent model for crop/weed comparative genomics because they belong to the same genus, a high quality and fully annotated reference genome is available, and many cultivated rice cultivars and wild relatives have been sequenced. Re-sequencing of several rice and weedy rice accessions from China demonstrated that Chinese weedy rice was de-domesticated independently from cultivated rice through exoferality and experienced a strong genetic bottleneck.[74] Following a similar approach, 38 weedy rice accessions from the United States were re-sequenced and compared to over 100 previously sequenced rice genomes, including weedy rice accessions from China.[75] Such work provides insights into evolutionary processes responsible for weediness traits in rice and identified genomic regions that could be used for crop improvement.[76] Importantly, with the more robust genomic tools in rice, we can now detect the difference between exoferal (derived from crop-wild relative hybridization) and endoferal (escaped crop genotypes) weeds.[77,78] Through comparative genomics analyses of exo- and endoferal weedy rice, we can now discover quantitative genetic differences in weedy versus domesticated traits and how population structure may modulate these differences in the field.

7. DELIVERING AN INTERNATIONAL WEED GENOMICS CONSORTIUM

To reach its full potential, weed genomics must connect communities from diverse disciplines of biology such as weed science, plant genetics, molecular physiology, evolution, and ecology. The IWGC is moving forward on these objectives to initiate the envisaged genome sequencing, website, and training initiatives. Annotated genomes at chromosome-scale assembly will be released in a user-friendly database environment. To provide a platform for community engagement with the IWGC, we have established a discussion forum at Plantae (www.plantae.org), a free online resource for the plant science community. Interested readers are invited to register with Plantae and join the conversation at https://community.plantae.org/discussion/4896069111202710923/international-weed-genomics-consortium. The weed genomics website developed by Dr. Scott McElroy, www.weedgenomics.org, may be further developed into the online weed genomics resource for the IWGC, based in part on other successful genome initiatives such as www.rosaceae.org. The weed genomics website will provide genome browsers, searching capability, comparative and diversity genomics tools, and visualization of gene expression and genotyping data sets. The website will also provide a platform for IWGC training in bioinformatics.

ACKNOWLEDGMENTS

This work was partially supported by the Biotechnology and Biological Sciences Research Council (BBSRC), UK (grant number BB/N022319/1 to P.N. and T.G. and grant number BB/L001489/1 to P.N.). The work at Rothamsted forms part of the Smart Crop Protection (SCP) strategic programme (BBS/OS/CP/000001) funded through the BBSRC's Industrial Strategy Challenge Fund.

REFERENCES

1. Kew, State of the World's Plants. Available at: https://stateoftheworldsplants.com. (2017).

2. Oerke EC, Crop losses to pests. *The Journal of Agricultural Science* **144**: 31-43 (2006).

3. Shaner DL, Introduction to the effect of biotechnology and genomics on weed science. *Weed Sci* **49**: 248-248 (2001).

4. Hess FD, Anderson RJ and Reagan JD, High throughput synthesis and screening: the partner of genomics for discovery of new chemicals for agriculture. *Weed Sci* **49**: 249-256 (2001).

5. Jasieniuk M and Maxwell BD, Plant diversity: New insights from molecular biology and genomics technologies. *Weed Sci* **49**: 257-265 (2001).

6. Weller SC, Bressan RA, Goldsbrough PB, Fredenburg TB and Hasegawa PM, The effect of genomics on weed management in the 21st century. *Weed Sci* **49**: 282-289 (2001).

7. Anderson JV, Emerging technologies: an opportunity for weed biology research. *Weed Sci* **56**: 281-282 (2008).

8. Gressel J, Molecular Biology of Weed Control (Vol. 1). CRC Press. (2002).

9. Stewart Jr CN, Weedy and Invasive Plant Genomics, in ed. by John Wiley & Sons, 2009).

10. Tranel PJ and Horvath DP, Molecular biology and genomics: new tools for weed science. *Bioscience* **59**: 207-215 (2009).

11. Basu C, Halfhill MD, Mueller TC and Stewart Jr CN, Weed genomics: new tools to understand weed biology. *Trends Plant Sci* **9**: 391-398 (2004).

12. Stewart CN, Tranel PJ, Horvath DP, Anderson JV, Rieseberg LH, Westwood JH, Mallory-Smith CA, Zapiola ML and Dlugosch KM, Evolution of weediness and invasiveness: Charting the course for weed genomics. *Weed Sci* **57**: 451-462 (2009).

13. Shendure J and Ji H, Next-generation DNA sequencing. *Nature Biotech* **26**: 1135-1145 (2008).

14. Ansorge WJ, Next-generation DNA sequencing techniques. *New Biotechnol* **25**: 195-203 (2009).

15. Schadt EE, Turner S and Kasarskis A, A window into third-generation sequencing. *Human Mol Genet* **19**: R227-R240 (2010).

16. Stapley J, Reger J, Feulner PG, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP and Slate J, Adaptation genomics: the next generation. *Trends Ecol Evol* **25**: 705-712 (2010).

17. Alkan C, Sajjadian S and Eichler EE, Limitations of next-generation genome sequence assembly. *Nat Methods* **8**: 61-65 (2011).

18. Treangen TJ and Salzberg SL, Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* **13**: 36-46 (2012).

19. Xiao C-L, Chen Y, Xie S-Q, Chen K-N, Wang Y, Han Y, Luo F and Xie Z, MECAT: fast mapping, error correction, and *de novo* assembly for single-molecule sequencing reads. *Nat Methods* **14**: 1072 (2017).

20. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH and Phillippy AM, Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* **27**: 722-736 (2017).

21. Chaney L, Sharp AR, Evans CR and Udall JA, Genome mapping in plant comparative genomics. *Trends Plant Sci* **21**: 770-780 (2016).

22. Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJ, Ohyanagi H, Mineta K, Michell CT and Saber N, The genome of *Chenopodium quinoa*. *Nature* **542**: 307-312 (2017).

23. Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, Rupp O, Sörensen TR, Stracke R and Reinhardt R, The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature* **505**: 546-549 (2014).

24. Lightfoot D, Jarvis DE, Ramaraj T, Lee R, Jellen E and Maughan P, Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (*Amaranthus hypochondriacus*) chromosomes provide insights into genome evolution. *BMC Biology* **15**: 74 (2017).

25. Varshney RK, Shi C, Thudi M, Mariac C, Wallace J, Qi P, Zhang H, Zhao Y, Wang X and Rathore A, Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotech* **35**: 969 (2017).

26. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob ur R, Ware D, Westhoff P, Mayer KFX, Messing J and Rokhsar DS, The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**: 551-556 (2009).

27. Zimin AV, Puiu D, Hall R, Kingan S, Clavijo BJ and Salzberg SL, The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *GigaScience* **6**: 1-7 (2017).

28. Giacomini DA, Gaines T, Beffa R and Tranel PJ, Optimizing RNA-seq studies to investigate herbicide resistance. *Pest Manag Sci*: in press (2018).

29. Babineau M, Mahmood K, Mathiassen SK, Kudsk P and Kristensen M, *De novo* transcriptome assembly analysis of weed *Apera spica-venti* from seven tissues and growth stages. *BMC Gen* **18**: 128 (2017).

30. Jayasena AS, Fisher MF, Panero JL, Secco D, Bernath-Levin K, Berkowitz O, Taylor NL, Schilling EE, Whelan J and Mylne JS, Stepwise evolution of a buried inhibitor peptide over 45 My. *Mol Biol Evol* **34**: 1505-1516 (2017).

31. Peng Y, Lai Z, Lane T, Nageswara-Rao M, Okada M, Jasieniuk M, O'Geen H, Kim RW, Sammons RD, Rieseberg LH and Stewart Jr CN, *De novo* genome assembly of the

economically important weed horseweed using integrated data from multiple sequencing platforms. *Plant Physiol* **166**: 1241-1254 (2014).

32. Guo L, Qiu J, Ye C, Jin G, Mao L, Zhang H, Yang X, Peng Q, Wang Y, Jia L, Lin Z, Li G, Fu F, Liu C, Chen L, Shen E, Wang W, Chu Q, Wu D, Wu S, Xia C, Zhang Y, Zhou X, Wang L, Wu L, Song W, Wang Y, Shu Q, Aoki D, Yumoto E, Yokota T, Miyamoto K, Okada K, Kim D-S, Cai D, Zhang C, Lou Y, Qian Q, Yamaguchi H, Yamane H, Kong C-H, Timko MP, Bai L and Fan L, *Echinochloa crus-galli* genome analysis provides insight into its adaptation and invasiveness as a weed. *Nature Comm* **8**: 1031 (2017).

33. Dorn KM, Fankhauser JD, Wyse DL and Marks MD, A draft genome of field pennycress (*Thlaspi arvense*) provides tools for the domestication of a new winter biofuel crop. *DNA Research* **22**: 121-131 (2015).

34. Moghe GD, Hufnagel DE, Tang H, Xiao Y, Dworkin I, Town CD, Conner JK and Shiu S-H, Consequences of whole-genome triplication as revealed by comparative genomic analyses of the wild radish (*Raphanus raphanistrum*) and three other Brassicaceae species. *Plant Cell* **26**: 1925-1937 (2014).

35. Byrne SL, Nagy I, Pfeifer M, Armstead I, Swain S, Studer B, Mayer K, Campbell JD, Czaban A, Hentrup S, Panitz F, Bendixen C, Hedegaard J, Caccamo M and Asp T, A synteny-based draft genome sequence of the forage grass *Lolium perenne*. *Plant J* **84**: 816-826 (2015).

36. Golicz AA, Batley J and Edwards D, Towards plant pangenomics. *Plant Biotechnol J* **14**: 1099-1105 (2016).

37. Ravet K, International Weed Genomics Consortium (IWGC): an initiative to coordinate and enable research in weed genomics. *Global Herbicide Resistance Challenge Proceedings* **2**: 51 (2017).

38. Neve P, Bridging the gap between weed geomics and herbicide resistance management. *Global Herbicide Resistance Challenge Proceedings* **2**: 25 (2017).

39. Molin WT, Wright AA, Lawton-Rauh A and Saski CA, The unique genomic landscape surrounding the *EPSPS* gene in glyphosate resistant *Amaranthus palmeri*: a repetitive path to resistance. *BMC Gen* **18**: 91 (2017).

40. Iwakami S, Endo M, Saika H, Okuno J, Nakamura N, Yokoyama M, Watanabe H, Toki S, Uchino A and Inamura T, Cytochrome P450 CYP81A12 and CYP81A21 are associated with resistance to two acetolactate synthase inhibitors in *Echinochloa phyllopogon*. *Plant Physiol* **165**: 618-629 (2014).

41. Patterson EL, Pettinga DJ, Ravet K, Neve P and Gaines TA, Glyphosate resistance and *EPSPS* gene duplication: Convergent evolution in multiple plant species. *J Hered*: (2017).

42. Laforest M, Soufiane B, Simard M-J, Obeid K, Page E and Nurse RE, Acetyl-CoA carboxylase overexpression in herbicide resistant large crabgrass (*Digitaria sanguinalis*). *Pest Manag Sci* **73**: 2227–2235 (2017).

43. Markus C, Pecinka A, Karan R, Barney JN and Merotto A, Epigenetic regulation– Contribution to herbicide resistance in weeds? *Pest Manag Sci* **74**: 275-281 (2018).

44. Harper J. The evolution of weeds in relation to resistance to herbicides; 1956. pp. 179-188.

45. Baker HG (1965) Characteristics and modes of origin of weeds. Characteristics and Modes of Origin of Weeds. pp. 147-172.

46. Clements DR, DiTommaso A, Jordan N, Booth BD, Cardina J, Doohan D, Mohler CL, Murphy SD and Swanton CJ, Adaptability of plants invading North American cropland. *Agr Ecosyst Environ* **104**: 379-398 (2004).

47. Neve P, Vila-Aiub M and Roux F, Evolutionary thinking in agricultural weed management. *New Phytol* **184**: 783-793 (2009).

48. Vigueira CC, Olsen KM and Caicedo AL, The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity* **110**: 303-311 (2013).

49. DeBolt S, Copy number variation shapes genome diversity in *Arabidopsis* over immediate family generational scales. *Genome Biol Evol* **2**: 441-453 (2010).

50. Boyko A and Kovalchuk I, Genome instability and epigenetic modification—heritable responses to environmental stress? *Curr Opin Plant Biol* **14**: 260-266 (2011).

51. Mickelbart MV, Hasegawa PM and Bailey-Serres J, Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nat Rev Genet* **16**: 237-251 (2015).

52. Kuester A, Conner JK, Culley T and Baucom RS, How weeds emerge: a taxonomic and trait□based examination using United States data. *New Phytol* **202**: 1055-1068 (2014).

53. Patterson E, Fleming M, Kessler K, Nissen S and Gaines T, A KASP genotyping method to identify northern watermilfoil, Eurasian watermilfoil, and their interspecific hybrids. *Front Plant Sci* **8**: 752 (2017).

54. Küpper A, Borgato EA, Patterson EL, Netto AG, Nicolai M, Carvalho SJd, Nissen SJ, Gaines TA and Christoffoleti PJ, Multiple resistance to glyphosate and acetolactate synthase inhibitors in Palmer amaranth (*Amaranthus palmeri*) identified in Brazil. *Weed Sci* **65**: 317-326 (2017).

55. Neve P, Busi R, Renton M and Vila□Aiub MM, Expanding the eco□evolutionary context of herbicide resistance research. *Pest Manag Sci* **70**: 1385-1393 (2014).

56. Busi R, Yu Q, Barrett-Lennard R and Powles S, Long distance pollen-mediated flow of herbicide resistance genes in *Lolium rigidum*. *Theor Appl Genet* **117**: 1281 (2008).

57. Sarangi D, Tyre AJ, Patterson EL, Gaines TA, Irmak S, Knezevic SZ, Lindquist JL and Jhala AJ, Pollen-mediated gene flow from glyphosate-resistant common waterhemp (*Amaranthus rudis* Sauer): consequences for the dispersal of resistance genes. *Sci Rep* **7**: 44913 (2017).

58. Rieger MA, Lamond M, Preston C, Powles SB and Roush RT, Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* **296**: 2386-2388 (2002).

59. Menchari Y, Camilleri C, Michel S, Brunel D, Dessaint F, Le Corre V and Délye C, Weed response to herbicides: regional□scale distribution of herbicide resistance alleles in the grass weed *Alopecurus myosuroides*. *New Phytol* **171**: 861-874 (2006).

60. Délye C, Michel S, Bérard A, Chauvel B, Brunel D, Guillemin J-P, Dessaint F and Le Corre V, Geographical variation in resistance to acetyl-coenzyme A carboxylase-inhibiting herbicides across the range of the arable weed *Alopecurus myosuroides* (black-grass). *New Phytol* **186**: 1005-1017 (2010).

61. Okada M, Hanson BD, Hembree KJ, Peng Y, Shrestha A, Stewart CN, Jr., Wright SD and Jasieniuk M, Evolution and spread of glyphosate resistance in *Conyza canadensis* in California. *Evol Appl* **6**: 761-777 (2013).

62. Fernández L, Haro LA, Distefano AJ, Carolina Martínez M, Lía V, Papa JC, Olea I, Tosto D and Esteban Hopp H, Population genetics structure of glyphosate□resistant Johnsongrass (*Sorghum halepense* L. Pers) does not support a single origin of the resistance. *Ecol Evol* **3**: 3388-3400 (2013).

63. Küpper A, Manmathan HK, Giacomini D, Patterson EL, McCloskey WB and Gaines TA, Population genetic structure in glyphosate-resistant and-susceptible Palmer amaranth (*Amaranthus palmeri*) populations using genotyping-by-sequencing (GBS). *Front Plant Sci* **9**: 29 (2018).

64. Gressel J, Dealing with transgene flow of crop protection traits from crops to their relatives. *Pest Manag Sci* **71**: 658-667 (2015).

65. Beffa R, Figge A, Lorentz L, Hess M, Laber B and Ruiz-Santaella JP, Weed resistance diagnostic technologies to detect herbicide resistance in cereal growing areas. A review. *Julius-Kühn-Archiv* **434**: 75 (2012).

66. Kaundun SS, Hutchings SJ, Dale RP, Bailly GC and Glanfield P, Syngenta 'RISQ' test: a novel in-season method for detecting resistance to post-emergence ACCase and ALS inhibitor herbicides in grass weeds. *Weed Res* **51**: 284-293 (2011).

67. Duke SO, Why have no new herbicide modes of action appeared in recent years? *Pest Manag Sci* **68**: 505-512 (2012).

68. Peters B and Strek H, Herbicide discovery in light of rapidly spreading resistance and ever increasing regulatory hurdles. *Pest Manag Sci*: In press (2018).

69. Corral MG, Leroux J, Tresch S, Newton T, Stubbs KA and Mylne JS, Exploiting the evolutionary relationship between malarial parasites and plants to develop new herbicides. *Angew Chem Int Edit* **56**: 9881-9885 (2017).

70. Corral MG, Leroux J, Stubbs KA and Mylne JS, Herbicidal properties of antimalarial drugs. *Sci Rep* **7**: 45871 (2017).

71. Zotti M, dos Santos EA, Cagliari D, Christiaens O, Taning CNT and Smagghe G, RNAi technology in crop protection against arthropod pests, pathogens and nematodes. *Pest Manag Sci*: in press (2018).

72. Esvelt KM, Smidler AL, Catteruccia F and Church GM, Concerning RNA-guided gene drives for the alteration of wild populations. *Elife* **3**: e03401 (2014).

73. Van de Wouw M, Kik C, van Hintum T, van Treuren R and Visser B, Genetic erosion in crops: concept, research results and challenges. *Plant Genet Res* **8**: 1-15 (2010).

74. Qiu J, Zhou Y, Mao L, Ye C, Wang W, Zhang J, Yu Y, Fu F, Wang Y, Qian F, Qi T, Wu S, Sultana MH, Cao Y-N, Wang Y, Timko MP, Ge S, Fan L and Lu Y, Genomic variation associated with local adaptation of weedy rice during de-domestication. *Nature Comm* **8**: 15323 (2017).

75. Li L-F, Li Y-L, Jia Y, Caicedo AL and Olsen KM, Signatures of adaptation in the weedy rice genome. *Nat Genet* **49**: 811-814 (2017).

76. Stewart Jr CN, Becoming weeds. *Nat Genet* **49**: 654 (2017).

77. Kanapeckas KL, Tseng TM, Vigueira CC, Ortiz A, Bridges WC, Burgos NR, Fischer AJ and Lawton□Rauh A, Contrasting patterns of variation in weedy traits and unique crop features in divergent populations of US weedy rice (*Oryza sativa* sp.) in Arkansas and California. *Pest Manag Sci* **DOI**: 10.1002/ps.4820 (2018).

78. Kanapeckas KL, Vigueira CC, Ortiz A, Gettler KA, Burgos NR, Fischer AJ and Lawton-Rauh AL, Escape to ferality: the endoferal origin of weedy rice from crop rice through de-domestication. *PLoS One* **11**: e0162676 (2016).

FIGURE LEGENDS

Figure 1. Simplified phylogenetic relationship between selected weed species (bold font) and cultivated crops (underlined). Selected weeds that are (A) eudicots along with important crops, pseudo-cereals, and vegetables; and (B) grass weeds along with cultivated cereal crops.

Figure 2. Weeds selected as high-priority species among the pre-selected species. A list of 10 species was proposed, pre-selected based on the output of previous workshops and discussions, and for which genome assembly projects were not completed or known to be in progress as of March 2017. For example, *Amaranthus palmeri* and *Conyza canadensis* were not included in the pre-selected list since their genomes were sequenced or in progress at the time of the survey. Participants were asked to select up to 3 species. Total number of participants: 245.

Figure 3. Additional weed species selected as high-priority species. Participants were asked to choose up to two additional species that were not pre-selected in Figure 2. A) Highest ranked species and B) highest ranked species by genus. Several species appear in this list for which genomes are in progress or complete, such as *Amaranthus palmeri* and *Conyza canadensis*, indicating their importance to the weed science community and a need for improved communication about the status of genome sequencing projects in weeds. Species and genera receiving less than 5 votes were not presented. Total number of participants: 245.

Figure 4. A word map depicting the frequency that key words were mentioned when delegates at the weed genomics workshop (Denver, 2017) were asked to define priority areas

for future research that addresses weed management questions through increased access to weed genomic resources.

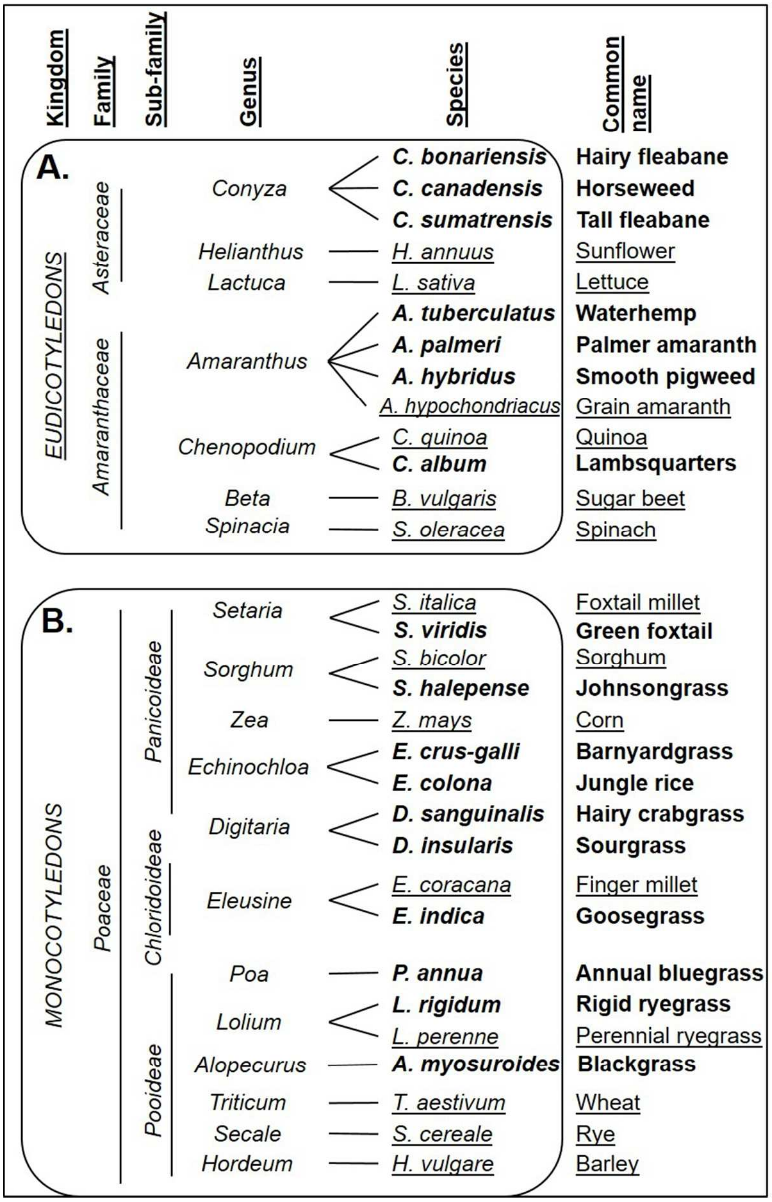


Figure 1. Simplified phylogenetic relationship between selected weed species (bold font) and cultivated crops (underlined). Selected weeds that are (A) eudicots along with important crops, pseudo-cereals, and vegetables; and (B) grass weeds along with cultivated cereal crops.

146x226mm (150 x 150 DPI)

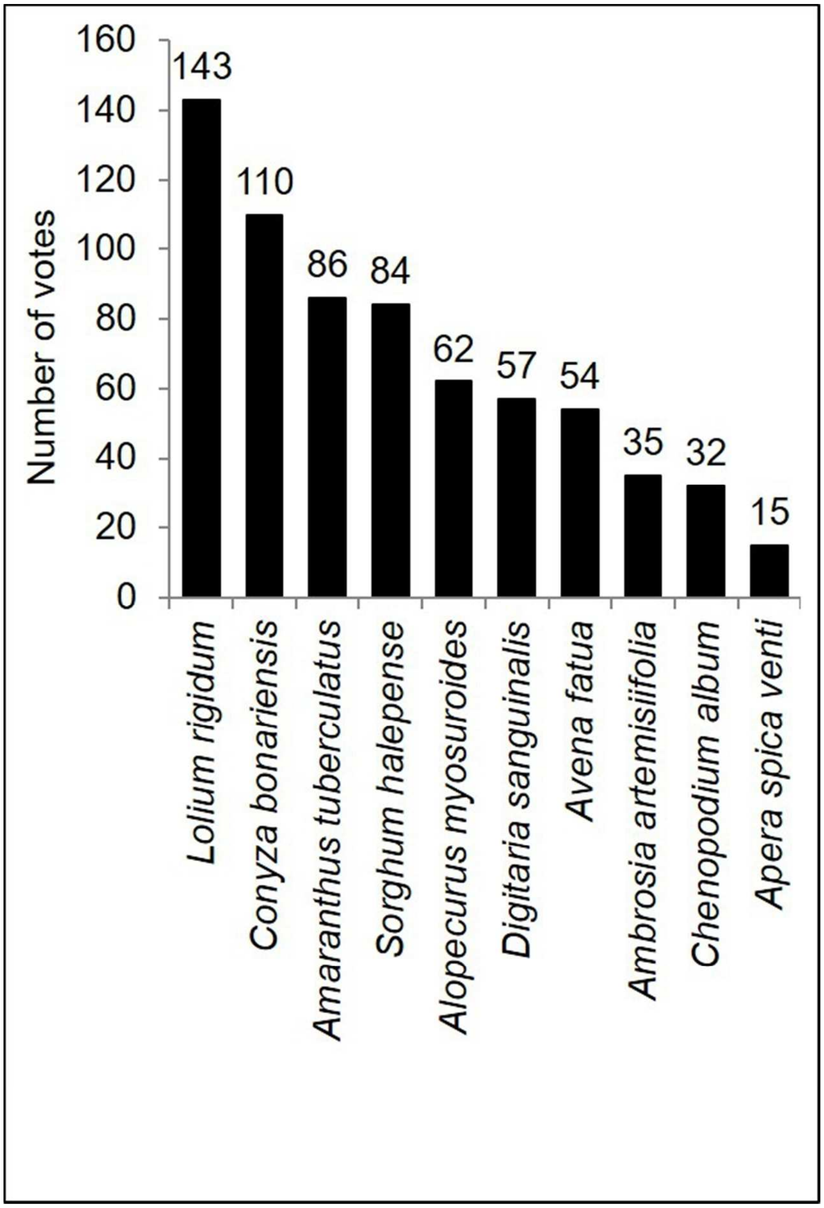


Figure 2. Weeds selected as high-priority species among the pre-selected species. A list of 10 species was proposed, pre-selected based on the output of previous workshops and discussions, and for which genome assembly projects were not completed or known to be in progress as of March 2017. For example, Amaranthus palmeri and Conyza canadensis were not included in the pre-selected list since their genomes were sequenced or in progress at the time of the survey. Participants were asked to select up to 3 species. Total number of participants: 245.

106x156mm (150 x 150 DPI)

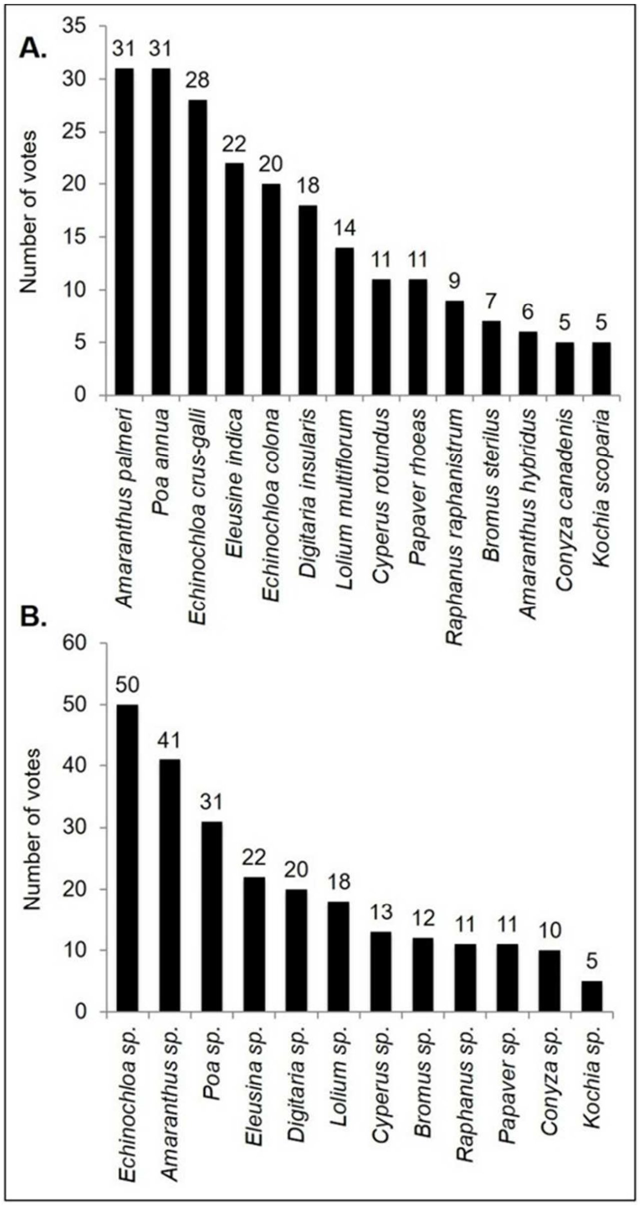


Figure 3. Additional weed species selected as high-priority species. Participants were asked to choose up to two additional species that were not pre-selected in Figure 2. A) Highest ranked species and B) highest ranked species by genus. Several species appear in this list for which genomes are in progress or complete, such as Amaranthus palmeri and Conyza canadensis, indicating their importance to the weed science community and a need for improved communication about the status of genome sequencing projects in weeds. Species and genera receiving less than 5 votes were not presented. Total number of participants: 245.

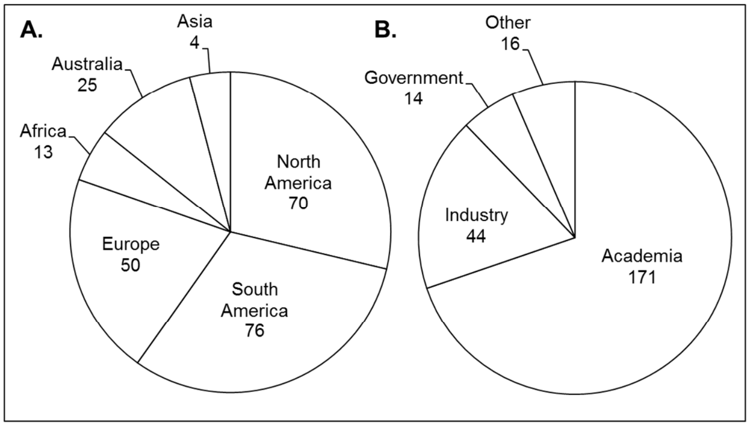
93x175mm (150 x 150 DPI)

Figure 4. A word map depicting the frequency that key words were mentioned when delegates at the weed genomics workshop (Denver, 2017) were asked to define priority areas for future research that addresses weed management questions through increased access to weed genomic resources.

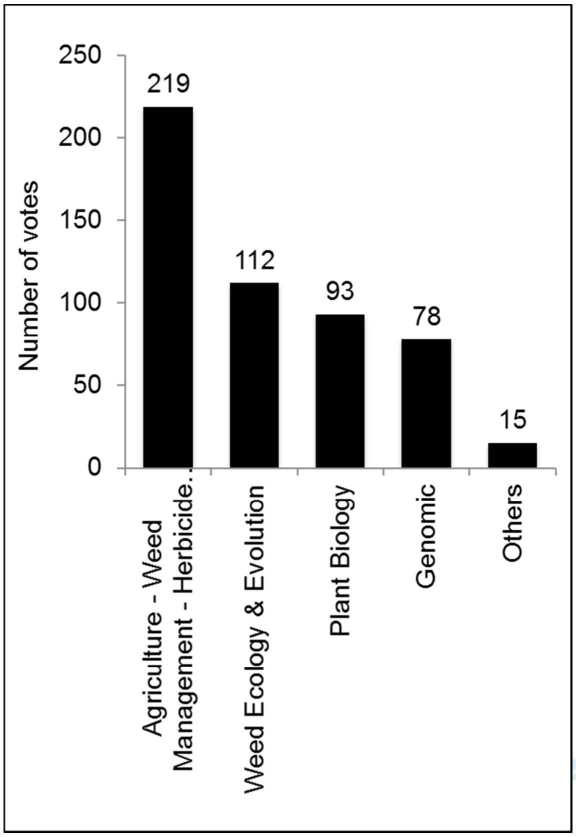
246x149mm (96 x 96 DPI)

The Power and Potential of Genomics in Weed Biology and Management

Supporting Information



Supporting Information Figure 1. Distribution of the survey's participants. The distribution of participants was analyzed based (A) on their geographic origins and (B) on their type of professional appointments. Total number of participants: 245.



Supporting Information Figure 2. Primary scientific interests of participants. Participants were asked about their main scientific/research interests. They could select among four categories (genomics, plant biology, weed ecology and evolution, and agriculture/weed management/herbicide resistance). Additionally, they could choose “other” and give details about their interests. Participants could check several interests. Total number of participants: 245.

Response to Reviewers

PM-17-0847.R1

We thank the editor and the two reviewers for the constructive comments. We have conducted a major revision of the manuscript and are submitting it as a new submission with additional co-authors. Responses to specific comments are indicated below.

Editor's comments :
I like the suggestion by reviewer 2 to put in place a platform to allow reader engagement. Doing so would also at least partially address some of the concerns of reviewer 1.

We have started a discussion board at Plantae.org, referenced in the final paragraph of the paper.

Reviewer's comments to the Author:
Referee: 1

COMMENTS FOR AUTHORS

As a concept I want to tell you that I think the idea of a consortium is a good idea and I applaud the efforts of the authors to take such action. I have a tremendous respect for the authors on this manuscript and I am sure such efforts will be successful.
However, rather than hiding my next comments in a comment to the editor I want to be direct with the authors. I think you are very premature with this manuscript. I think you need to take a step back and actually begin to accomplish something before you publish a paper on what you WANT to do. When I read this paper you have some great conceptual ideas but you have not fleshed out any of these ideas for how you are going to accomplish them. You have conducted a survey and people think this is an important effort but there is no information on how to make these things happen. How are you going to foster and create collaboration for instance? How are you going to make whole genomes actually useful to people? Who is going to develop user friendly genome analytical tools? What type of tools? How are you going to make the genomes open access? Who is going to host them? Are they going to be searchable? How will they be searchable? How are people going to share data? You have outlined a great set of objectives in L76-84 but from what I can see nothing has been accomplished except for a survey as of yet.

In the major revision, we have provided more details on the goals of the IWGC. Some of the specific requested by the reviewer have not yet been defined, but we have an organizational structure in place to make decisions regarding genome access and genomics tools.

L30 – Delete "High-quality". What is high quality? How do you define high quality? Cannot. It is a vague term.

Agreed. We have deleted references to high quality.

L236-240 – Delete this paragraph. I have looked up www.weedgenome.org and the page is not active. None of the things you say are being developed have been developed. It is disingenuous to discuss this site when nothing is working or active. It is easy for someone to say that they want to do something or they plan to do something but nothing has actually been done with this site.

We have made arrangements with www.weedgenomics.org to be the host site for the IWGC. The weedgenome.org page is also now live, although we have removed the specific reference to it from the paper.

What really needs to be done in the weed genomics area is to actually grow the number of people who can work in and understand genomics/transcriptomics/computational biology/bioinformatics and want to apply those skills to weed species. To me that is where this consortium should be focusing. Trying to develop workshops, computer science trainings (and bootcamps) to get people in weed science moving in this direction.

We agree that training efforts will be a major impact of the IWGC, and we discuss this point in the manuscript.

Referee: 2

COMMENTS FOR AUTHORS
This manuscript is part meeting report from the Global Herbicide Resistance Challenge conference, part white paper calling for the establishment of an International Weeds Genomic Consortium, and part report on the data collected from a survey sent to meeting participants (and whoever they passed them on to). The paper highlights the advantages of such a consortium and some of the considerations for the responsibilities and requirements of such a consortium. The picture is painted with a very broad brush, and few specifics are noted or even suggested. This is likely because no mechanisms have yet been set in place to develop a platform for discussion of such specifics. As it is written, it is a valuable starting point for discussion and a potentially excellent advertisements for any developing nucleus of researchers interested in this endeavor (I certainly found it as such). However, there really needs to be some platform in place for interested parties to engage in this project. Merely stating the results of the (admittedly biased) survey and calling for action seems insufficient and too likely to die on the vine without some sort of a platform to allow readers to connect and engage. I believe without such a platform (perhaps on free science-based platforms designed for such a purpose such as Trellis or Plantae) this manuscript would have minimal impact. However, with such a platform in place, interested readers are likely to engage and help spur this well-argued call for action.

We appreciate the suggestion for a discussion platform, and we have started a discussion group on Plantae for the IWGC.

Comments in PDF:

Delete redundant period on line 50 (1st paragraph of introduction)

Change made.

Page 11, line 218: Given the likelihood that epigenetics plays in plasticity of weed, I am surprised that methalomics (or other genomic scale analyses of epigenetic marker were not included)

We discuss epigenetics and the possible role in evolution of weedy traits on lines 238-244.

Page 11, line 218: I am also surprised by the lack of mention of Evo-devo type questions such as how weeds have eveolved resistances under particular selection regimes or how they are (or perhaps are not) undergoing allelic selection at the extremes of their range to help identify factors limiting the spread of some invasives - that could be answered by development and population-based use of high density snp that could (should) be identified by sequencing efforts.

Section 5 "WEED GENOMICS PROVIDES NOVEL INSIGHT IN WEED BIOLOGY" now goes into detail on these types questions and how genomics can assist in this type of research.

Page 11, line 220: Only if efforts include some sort of scheme to identify specific differences associated with resistance. Such analyses would be arguably easier with a well annotated reference genome, but would still require significant resequencing efforts.

Agreed, but we would say that resequencing is only affordable and feasible when a well annotated reference genome is available.

Page 11, line 221: Markers for detection of resistance? I suppose if the test is cheap enough, a grower may collect weeds samples in his field and get them analyzed for resistance to help inform his choice of herbicides, but this seems like a stretch. Likewise, surveying growers or extension agents should more easily allow mapping of the spread of resistance.

Section 6a is now expanded to go into more detail on this topic.

Page 12, line 237: This web site just shows a picture and has the note"Under construction. Any comments or suggestions please contact with Dr. Fan (fanlj@zju.edu.cn)"

We apologize for this oversight. We have now selected www.weedgenomics.org as the home site for the IWGC, and updates to the website are in progress.

Page 13, line 256: It would be good to develop some sort of a mechanism for seeking input to answer these questions and note it in this publication.

We have started the Plantae discussion board.

Page 13, line 259: It will be difficult to compete with NCBI for data open access to and storage of data, or google for identifying resources or collaborators. However, some sort of database for funding or the development of proposal generating/screening process could be quite valuable.

Agreed, no specific changes made in response.

Page 13, line 269: Again, it is nice that this is recognized, but this manuscript would benefit from development of and access to some sort of mechanism or platform for discussing these ideas: I believe there are two platforms that might be utilized to initiate such a platform (Trellis (through AAAS) and Plantae (through the ASPB which is completely free and looking for such partnerships).

We have started the Plantae discussion board.