Omics in Weed Science: A Perspective from Genomics, Transcriptomics, and Metabolomics Approaches

Authors: Amith S. Maroli, Todd A. Gaines, Michael E. Foley, Stephen O. Duke, Münevver Doğramacı, et. al.

Source: Weed Science,  66(6) : 681-695

Published By: Weed Science Society of America

URL: https://doi.org/10.1017/wsc.2018.33
Omics in Weed Science: A Perspective from Genomics, Transcriptomics, and Metabolomics Approaches

Amit S. Maroli1, Todd A. Gaines2, Michael E. Foley3, Stephen O. Duke4, Münevver Doğramaci5, James V. Anderson6, David P. Horvath7, Wun S. Chao8 and Nishanth Tharayil9

1Postdoctoral Fellow (ORCID: 0000-0002-0647-0316), Department of Plant and Environmental Sciences, Clemson University, Anderson, SC, USA; current: Department of Environmental Engineering and Earth Sciences, Clemson University, Anderson, SC, USA; 2Assistant Professor (ORCID: 0000-0003-1485-7665), Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO, USA; 3Plant Physiologist, Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit, USDA-Agricultural Research Service, Fargo, ND, USA; 4Research Leader (ORCID: 0000-0001-7210-5168), Natural Products Utilization Research Unit, National Center for Natural Products Research, USDA-Agricultural Research Service, University, MS, USA; 5Research Molecular Biologist (ORCID: 0000-0002-1412-8174), Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit, USDA-Agricultural Research Service, Fargo, ND, USA; 6Research Chemist (ORCID: 0000-0002-1801-5767), Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit, USDA-Agricultural Research Service, Fargo, ND, USA; 7Research Plant Physiologist (ORCID: 0000-0002-8458-7692), Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit, USDA-Agricultural Research Service, Fargo, ND, USA; 8Research Molecular Geneticist, Red River Valley Agricultural Research Center, Sunflower and Plant Biology Research Unit, USDA-Agricultural Research Service, Fargo, ND, USA and 9Associate Professor (ORCID: 0000-0001-6866-0804), Department of Plant and Environmental Sciences, Clemson University, Clemson, SC, USA

Abstract

Modern high-throughput molecular and analytical tools offer exciting opportunities to gain a mechanistic understanding of unique traits of weeds. During the past decade, tremendous progress has been made within the weed science discipline using genomic techniques to gain deeper insights into weedy traits such as invasiveness, hybridization, and herbicide resistance. Though the adoption of newer “omics” techniques such as proteomics, metabolomics, and physionomics has been slow, applications of these omics platforms to study plants, especially agriculturally important crops and weeds, have been increasing over the years. In weed science, these platforms are now used more frequently to understand mechanisms of herbicide resistance, weed resistance evolution, and crop–weed interactions. Use of these techniques could help weed scientists to further reduce the knowledge gaps in understanding weedy traits. Although these techniques can provide robust insights about the molecular functioning of plants, employing a single omics platform can rarely elucidate the gene-level regulation and the associated real-time expression of weedy traits due to the complex and overlapping nature of biological interactions. Therefore, it is desirable to integrate the different omics technologies to give a better understanding of molecular functioning of biological systems. This multidimensional integrated approach can therefore offer new avenues for better understanding of questions of interest to weed scientists. This review offers a retrospective and prospective examination of omics platforms employed to investigate weed physiology and novel approaches and new technologies that can provide holistic and knowledge-based weed management strategies for future.

Introduction

The identities of all organisms are embedded in their genes, which are often influenced by developmental and environmental cues. Sequential and temporal decoding of these genes confers physiological distinctiveness to each individual (Anderson 2008). In the last two decades, the term “omics” has been suffixed with several fields of study in biology (Brunetti et al. 2018). Recent advances in high-throughput functional omics technologies (Table 1) have facilitated an understanding of the various molecular–environmental interactions that regulate biological systems (Kitano 2002). The use of omics techniques to study various biological aspects would provide greater opportunities to dissect the molecular and physiological
mechanisms in developing resilient phenotypes. Among the various omics platforms, functional genomics has seen rapid progress, resulting in a growing number of sequenced plant genomes. This has facilitated the development of plants selected for specific agronomic traits and biological processes (Kantar et al. 2017; Nelson et al. 2018). The traditional giants of omics platforms encompass genomics, transcriptomics, and proteomics (Palsson 2002; Rochfort 2005). While genomics aims to understand how the genome functions, transcriptomics and proteomics perform systematic qualitative and quantitative analysis of the transcriptome and proteome content, respectively, in a tissue, cell, or subcellular compartment. Other recent omics techniques such as metabolomics, phenomics, and lipidomics complement the traditional techniques to depict a precise picture of the entire cellular process.

Omics approaches in weeds science have been gaining momentum over the past decade. As with other domains, the number of studies using genomic approaches to investigate weed biology and physiology has increased over the years (Basu et al. 2004; Bevan and Walsh 2005; Kreiner et al. 2018). DNA-based molecular studies using simple sequence repeats (SSRs), microsatellites, amplified fragment length polymorphisms (AFLPs), and inter simple sequence repeats (ISSRs) have provided tremendous opportunities to study weedy characteristics such as resilience, dormancy, and...
invasiveness, as well as weed genetic diversity and hybridization among related weed species (Corbett and Tardif 2006; Horvath 2010). Excellent reviews on weed genomics and DNA-based herbicide-resistance techniques have been produced by Basu et al. (2004), Corbett and Tardif (2006), Stewart (2009), and Tranel and Horvath (2009). Recently, the weed science community has initiated the International Weed Genomics Consortium to facilitate genomics for weed science (Ravet et al. 2018). However, applications of other omics for studying agronomically important weeds are at a nascent stage, as seen by the limited number of published studies (Table 2).

In addition to genomics, other omics techniques have also been used to investigate areas critical to weed science, including stress response, weediness/invasiveness, herbicide resistance, and genetic diversity (Délye 2013; Grossmann et al. 2010; Keith et al. 2017; Stewart 2009; Stewart et al. 2009, 2010; Zhang and Reichers 2008). However, due to the complexity of the molecular and environmental interactions, no single omics analysis can independently explain the intricacies of fundamental physiology (Fukushima et al. 2009; Hirai et al. 2004; Liberman et al. 2012). Hence, an integrated systems biology approach is needed to provide precise information about the molecular, biochemical, and physiological status of the target organism (Figure 1). An integrated systems biology approach can help not only in annotating unknown genes, but also in identifying their regulatory networks and the metabolic pathways they would influence (Pérez-Alonso et al. 2018). This would aid in understanding the genotype–phenotype relationship and consequently help to improve the existing weed management strategies in agricultural fields. Although there are several omics platforms, the present review will strive to highlight omics approaches used to study physiological aspects of agriculturally important weeds that have not been previously touched upon, such as elucidating physiology of bud dormancy, deciphering the mechanisms of herbicide resistance, and identifying potential herbicidal phytochemicals using omics approaches.

### Transcriptomics to Investigate Herbicide Resistance

Compared with the availability of genome sequence information and genetic resources for model plants such as mouse-ear cress [Arabidopsis thaliana (L.) Heynh.] (Arabidopsis Genome Initiative 2000), barbelclover (Medicago truncatula Gaertn.) (Bell et al. 2001), and purple false brome [Brachypodium distachyon (L.) P. Beauv.] (Vogel et al. 2010) and the genome sequences of several other dicot and monocot crops that are either sequenced or soon will be, to date only four draft genome assemblies have been completed for agronomic weed species (Table 3). Next-generation sequencing (NGS) techniques such as RNA-Seq have enabled accurate and powerful transcriptome analysis approaches for non-model species such as weeds, without requiring a fully assembled genome. A review of 15 RNA-Seq studies conducted in

### Table 2. Examples of omics papers on phytotoxins, including herbicides.

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Phytotoxin/herbicide</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcriptomics</td>
<td>Cinidon-ethyl, tribenuron-methyl, and 2,4-D</td>
<td>Pasquer et al. 2006</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>2,4-D</td>
<td>Raghavan et al. 2006</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Flufenacet</td>
<td>Lechelt-Kunze et al. 2003</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Atrazine and bentazon</td>
<td>Zhu et al. 2009</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>L-DOPA</td>
<td>Golisz et al. 2011</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Fagamine and phenolics</td>
<td>Golisz et al. 2008</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Diclofop-methyl</td>
<td>Gaines et al. 2014</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Pyroxasulam</td>
<td>Duhoux et al. 2015</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Juglone</td>
<td>Chi et al. 2011</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Benzoazoxolin-2(3H)-one</td>
<td>Baerson et al. 2005</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Imazapyr, primisulfuron-methyl, glyphosate, clorsulam-methyl, and sulfometuron-methyl</td>
<td>Das et al. 2010</td>
</tr>
<tr>
<td>Transcriptomics</td>
<td>Imazapyr</td>
<td>Manabe et al. 2007</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Cantharidin</td>
<td>Bajsa et al. 2015</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Diquox, paraquat, and norflurazon</td>
<td>Nestler et al. 2012</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Paraquat and glyphosate</td>
<td>Ahsan et al. 2008</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Dicamba and clopyralid</td>
<td>Kelley et al. 2006</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Butachlor</td>
<td>Kumari et al. 2009</td>
</tr>
<tr>
<td>Proteomics</td>
<td>Amiprophos-methyl</td>
<td>Wang et al. 2011</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Eremophilanes</td>
<td>Cantrell et al. 2007</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Ascaulitoxin</td>
<td>Duke et al. 2011</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Glyphosate</td>
<td>Maroli et al. 2015, 2017</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Saflufenacil</td>
<td>Grossmann et al. 2010</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>Cinmethylin</td>
<td>Grossmann et al. 2012b</td>
</tr>
</tbody>
</table>
weeds to find candidate genes for herbicide resistance and abiotic stress tolerance identified that increased replicate number and controlling genetic background were important factors to increase detection power and minimize the false-discovery rate (Giacomini et al. 2018). The first weed species transcriptomes released were horseweed (Erigeron canadensis L.) (Peng et al. 2010) and waterhemp [Amaranthus tuberculatus (Moq.) J. D. Sauer] (Riggins et al. 2010), with at least 22 weed transcriptomes sequenced and assembled to date, including weeds of agronomic crops, turfgrass, and invasive weeds (Gaines et al. 2017; McElroy 2018). Compared with transcriptomics methods such as microarray, which provide relative quantification, NGS-based transcriptome approaches produce absolute quantification of transcript expression as well as the sequence of all transcripts in a given sample. All expressed genes can be studied for changes in regulation (for example, upregulation of cytochrome P450s to increase herbicide metabolism), and several genes can be examined for candidate nonsynonymous mutations that could confer resistance. The identification of transcripts with differential regulation and/or mutations generates a hypothesis to be tested with subsequent validation. For discussion about transcriptomics in weeds before the introduction of NGS, the reader is referred to reviews by Lee and Tranel (2008) and Horvath (2010).

RNA-Seq measures the transcriptome abundance at a given time from a genome. The data can be used for various analyses, such as identification of differentially expressed transcripts

Table 3. Draft genome assemblies of agronomic weed species sequenced using next-generation sequencing technologies.

<table>
<thead>
<tr>
<th>Agronomic weed</th>
<th>Plant line</th>
<th>Estimated genome size</th>
<th>Sequence coverage</th>
<th>Genome coverage</th>
<th>Sequencing platforms</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild radish</td>
<td>5th generation inbred</td>
<td>515 Mb</td>
<td>NA</td>
<td>49.3%</td>
<td>Illumina® Genome Analyzer II</td>
<td>Moghe et al. 2014</td>
</tr>
<tr>
<td>(Raphanus raphanistrum L.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horseweed</td>
<td>Tennessee glyphosate-resistant biotype (TN-R)</td>
<td>335 Mb</td>
<td>350x</td>
<td>92.3%</td>
<td>Roche 454 GS-FLX, Illumina® HiSeq 2000, and PacBio® RS</td>
<td>Peng et al. 2014</td>
</tr>
<tr>
<td>(Erigeron canadensis L.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field pennycress</td>
<td>MN106</td>
<td>539 Mb</td>
<td>87x</td>
<td>76.5%</td>
<td>Illumina® HiSeq 2000, Illumina® MxSeq, and PacBio® SMRT</td>
<td>Dorn et al. 2015</td>
</tr>
<tr>
<td>(Thlaspi arvense L.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barnyardgrass</td>
<td>ST808</td>
<td>1.41Gb</td>
<td>171x</td>
<td>90.7%</td>
<td>Illumina® HiSeq 2000 and PacBio® RS II</td>
<td>Guo et al. 2017</td>
</tr>
<tr>
<td>[Echinochloa crus-galli (L.) P. Beauv.]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
between treatments, analysis of sequence variants, or characterization of alternative splicing. Due to its digital nature, RNA-Seq has a linear-detection dynamic range over five orders of magnitude, enabling quantification of even transcripts with very low expression. A typical RNA-Seq experiment consists of the steps outlined in Figure 2. Numerous downstream transcriptome data analyses can also be used to help interpret data, such as identification of enriched pathways with differentially expressed transcripts. RNA-Seq is also advantageous for studying complex gene families, such as those involved in enhanced metabolic resistance (for example, cytochrome P450s, glutathione-S-transferases, glucosyl transferases, ABC transporters). The results of the RNA-Seq experiments alone are not sufficient to prove causation for a candidate mechanism. RNA-Seq should be considered an experimental approach to generate robust hypotheses for candidate gene function. Subsequent forward genetics validation experiments are essential to prove function, such as testing cosegregation of a molecular marker (increased gene expression and/or a mutation) with resistance, testing for the presence of the molecular marker in unrelated populations of the same species, and preferably expression or knockout in a heterologous system such as *Arabidopsis* or yeast (for example, Cummins et al. 2013; LeClere et al. 2018). Differential expression can then be measured on validation samples using qRT-PCR on cDNA.

In weed science, several transcriptomic studies have focused on herbicide-resistance traits, including target-site resistance mechanisms (Riggins et al. 2010; Wiersma et al. 2015; Yang et al. 2013) and non-target site resistance (NTSR) mechanisms (An et al. 2014; Gaines et al. 2014; Gardin et al. 2015; Leslie and Baucom 2014; Peng et al. 2010; Riggins et al. 2010; Yang et al. 2013). Studying the entire transcriptome is especially useful for NTSR mechanisms, because NTSR generally involves multiple genes and gene families (Délye 2013). Examples of RNA-Seq studies on NTSR in grasses include acetyl-CoA carboxylase (ACCCase)-inhibitor resistance in rigid ryegrass (*Lolium rigidum* Gaudin) (Gaines et al. 2014) and *Brachypodium hybridum* Catalán, Joch. Müll., L.A. Mur & T. Langdon (Matzrafi et al. 2017) and acetylalanate synthase (ALS)-inhibitor resistance in *L. rigidum* (Duhoux et al. 2015) and blackgrass (*Alopecurus myosuroides* Huds.) (Gardin et al. 2015). Both NTSR and target-site mechanisms were evaluated in barnyardgrass [*Echinochloa crus-galli* (L.) P. Beauv.], using NGS to identify candidate genes involved in ALS-inhibitor and synthetic auxin (quinclorac) resistance (Yang et al. 2013). Responses to glyphosate and glyphosate resistance have also been studied using RNA-Seq in tall morningglory [*Ipomoea purpurea* (L.) Roth] (Leslie and Baucom 2014) and kochia [*Bassia scoparia* (L.) A. J. Scott] (Wiersma et al. 2015). Employing differential expression (DE) analysis using RNA-Seq, the study by Leslie and Baucom (2014) found a range of candidate genes that may explain differences in glyphosate response between populations, including metabolism, signaling, and defense-related genes with differential expression. Similarly, in glyphosate-resistant *B. scoparia*, RNA-Seq was used to confirm overexpression of 5-enolpyruvylshikimate-3-phosphate synthase (*EPSPS*) due to gene amplification, to determine that no other genes in the shikimate pathway besides *EPSPS* were differentially expressed between resistant and susceptible populations, and to establish that no candidate resistance-conferring mutations were present in the *EPSPS* sequence from the resistant population (Wiersma et al. 2015). This mutation analysis is referred to as deep sequencing, and it is used to identify mutations that may be expressed at a low level and not detected by traditional sequencing approaches. Recently, a mutation for dicamba resistance in *B. scoparia* was identified using transcriptomics and subsequently functionally validated using forward genetics and expression in heterologous systems (LeClere et al. 2018).

**Figure 2.** Workflow of transcript analyses by RNA-Seq and qRT-PCR.

**Transcriptomics to Investigate Bud Dormancy and Vegetative Growth**

Early studies on bud dormancy employed traditional or accessible molecular approaches (Anderson et al. 2005; Horvath and Anderson 2002; Horvath et al. 2002). However, accessibility of the genome sequence for *A. thaliana* (Arabidopsis Genome Initiative 2000) and availability and adoption of cDNA microarray technology for plant genes (Schena et al. 1995) allowed researchers to examine transcriptome profiles for a variety of tissues and treatments and enabled the development of cross-species...
adoption of Arabidopsis platforms (Horvath et al. 2003). Omics approaches were employed for elucidating signals, pathways, and mechanisms governing dormancy in underground adventitious buds (UABs) of leafy spurge (Euphorbia esula L.) (Anderson and Horvath 2001; Anderson et al. 2004, 2007; Foley et al. 2013; Horvath et al. 2006, 2008). Euphorbia esula is a noxious and perennial rangeland weed that can reproduce and spread vegetatively from an abundance of UABs (Anderson et al. 2005). As dormancy in these buds often contributes to escape from control measures, it is essential to understand the seasonal dormancy cycles (paradormancy, endodormancy, and ecodormancy) for UABs. Horvath et al. (2006, 2008) studied the transcriptome of E. esula UABs during transitions in these well-defined phases of dormancy under greenhouse and field conditions using high-density microarrays constructed from an E. esula expressed sequence tag database (Anderson et al. 2007). This work helped to identify transcripts encoded by a gene with similarity to DORMANCY ASSOCIATED MADS-BOX, which has since been strongly implicated in dormancy processes of several perennial plant systems (Horvath 2015). Meta-analysis of microarray-based transcriptome data also identified transcripts similar to Arabidopsis COP1, HY5, MAF3-like, RD22, and RVE1 as potential molecular markers for endodormancy in E. esula UABs (Doğramaci et al. 2015).

Studies have also been done to determine the impact of growth regulators on dormancy by examining changes in transcriptome profiles of E. esula UABs in response to foliar glyphosate treatment (Doğramaci et al. 2014, 2015, 2016). Although glyphosate is widely used as a broad-spectrum herbicide (Duke and Powles 2008), it is also known to have hormeric activity (Belz and Duke 2014; Velini et al. 2008). When applied at sublethal concentrations, it can cause tillering in some plant species due to axillary and root-bud growth. Maxwell et al. (1987) reported that glyphosate application at higher rates (2 to 6 kg ae ha$^{-1}$) to E. esula under field conditions caused an increase in the number of stems per square meter as a result of shoot growth from UABs, a phenomenon referred to as "witches' broom." Discovery or development of a growth regulator that could induce or inhibit shoot growth from UABs would be a significant step toward long-term control of other perennial weeds such as Canada thistle (Cirsium arvense (L.) Scop.), field bindweed (Convolvulus arvensis L.), and hedge bindweed (Calystegia sepium (L.) R. Br.). Initial studies conducted using qRT-PCR, indicated that glyphosate had the most significant impact on abundance of ENT-COPALYL DIPHOSPHATE SYNTHETASE 1, which is involved in a committed step for gibberellic acid (GA) biosynthesis, and auxin transporters, including PINs, PIN-LIKES, and ABC TRANSPORTERS. Foliar glyphosate treatment also reduced the abundance of transcripts involved in cell cycle processes, which was consistent with altered growth patterns (Doğramaci et al. 2014). RNA-Seq identified nearly 13,000 differentially expressed transcripts in UABs in response to foliar glyphosate treatment (Doğramaci et al. 2015). Of these transcripts, 6,239 had significant changes ≥ 2-fold in either direction, which included transcripts associated with many processes involving shoot apical meristem maintenance and stem growth. The foliar glyphosate treatment increased shikimate abundance in UABs before decapitation of aboveground shoots, indicating that EPSPS, the target site of glyphosate, was inhibited. Interestingly, the abundance of shikimate in new aerial shoots (6 wk after growth-inducing decapitation) derived from UABs of foliar glyphosate-treated plants was similar to controls. The abundance of transcripts (i.e., EPSPS, EMB1144, SKI1) involved in various stages of chorismate/shikimate biosynthesis had little change in amplitude, indicating glyphosate was not directly affecting transcription for components of the pathway in these tissues. Hormone analyses indicated that auxins, gibberellins (precursors and catabolites of bioactive gibberellins), and cytokinins (precursors and bioactive cytokinins) were more abundant in the aboveground shoots derived from UABs of glyphosate-treated plants versus the control. Based on the accumulation of transcriptome and metabolite data, it was proposed that the classic stunted and bushy phenotypes resulting from vegetative reproduction of E. esula UABs following foliar glyphosate treatment involve complex interactions, including shoot apical meristem maintenance, hormone biosynthesis and signaling (auxin, cytokinins, gibberellins, and strigolactones), cellular transport, and detoxification mechanisms (Doğramaci et al. 2015).

An expanded investigation into glyphosate-induced witches' brooming under field conditions was accomplished (Doğramaci et al. 2016). Field plots treated with high rates (3.3 and 6.7 kg ae ha$^{-1}$) of glyphosate had increased UAB-derived shoots displaying the stunted and bushy phenotype characteristics. qRT-PCR analysis to quantify the abundance of a selected set of transcripts in UABs of nontreated versus treated plants (0 vs. 6.7 kg ae ha$^{-1}$) further supported the impact that glyphosate has on molecular processes involved in biosynthesis or signaling of tryptophan or auxin, GA, ethylene, and cytokinins, as well as cell cycle processes. Moreover, these glyphosate-induced effects on vegetative growth and transcript abundance persisted in the field for at least 2 yr. Transcriptome studies have now progressed to a point where testable hypothesis-driven studies could be initiated as a step toward next-generation approaches for weed management. Though foliar application of glyphosate to E. esula causes effects that impact molecular processes in UABs, this broad-spectrum herbicide would not be ideal for manipulation of bud growth in rangeland perennial weeds due to its effect on non–target plant species. Nevertheless, this proof of concept project sets the stage to screen commercially available libraries of compounds, growth regulators, natural products, and other bioactive molecules that could be applied to perturb bud growth and shoot development.

Although transcriptome and metabolite analysis can identify potentially important signals, pathways, and molecular mechanisms involved in dormancy and glyphosate-induced witches’ brooming, it is important to remember that these changes in transcript abundance do not reflect a direct association with activity occurring at the posttranscriptional levels (Beckwith and Yanovsky 2014). Moreover, as with many weedy species, the genome for E. esula has not been completely sequenced or annotated. Therefore, research employing molecular, genomics, and genetics approaches must rely on the annotated genomes of model species such as Arabidopsis.

**Metabolomics and Fluxomics to Understand Weed Physiology**

The realization that genes, transcripts, and proteins alone cannot completely explain several physiological responses has triggered a marked increase in employing approaches that can relate gene expression to the final phenotypic outcome. Metabolomics is one such approach that comprehensively identifies and quantitates low-molecular-weight metabolites (metabolome), thus offering a powerful approach for molecular phenotyping (Fiehn 2002).
A common workflow for metabolomics experiments involves metabolite extraction, chromatographic separation, detection, data processing, metabolite identification, and statistical validation (Figure 3). Most often in plant metabolomics, metabolite separation is carried out by either liquid chromatography or gas chromatography followed by mass spectrometer detection (De Vos et al. 2007; Haggarty and Burgess, 2017; Maroli et al. 2015, 2017).

Metabolomics has been used in the past decade to study the mechanisms of action (MOAs) of synthetic and natural herbicides using several model plant species, such as maize (Zea mays L.) (Aranibar et al. 2001), sterile oat (Avena sterilis L.) (Aliferis and Chrysyn-Tokousbalides 2006), and Arabidopsis (Jaini et al. 2017; Sumner et al. 2015; Wu et al. 2018). However, limited studies have employed metabolomics to characterize weed physiology in response to herbicide applications (Aliferis and Chrysyn-Tokousbalides 2011; Miyagi et al. 2010), herbicide-resistance mechanisms (Aliferis and Jabaji 2011; Maroli et al. 2015, 2017; Serra et al. 2015; Vivancos et al. 2011), and non-target site herbicide-resistance mechanisms such as detoxification and metabolism (Wang et al. 2017). As reviewed earlier, application of genomics and transcriptomics has helped to identify herbicide-resistance mechanisms in some weeds (Chen et al. 2017; Délye 2013; Gaines et al. 2010; Nandula et al. 2012; Wright et al. 2018a, 2018b). Apart from this, metabolomics approaches have been recently adopted to understand effect of chemical stresses on perennial ryegrass (Lolium perenne L.) (Serra et al. 2015), to identify complementary glyphosate resistance mechanisms in Palmer amaranth (Amaranthus palmeri S. Watson) (Maroli et al. 2015), to determine glyphosate-induced global physiological perturbations in glyphosate-resistant (Fernández-Escalada et al. 2016, 2017) and glyphosate-tolerant (Maroli et al. 2017) weeds, and to examine herbicide metabolism in herbicide-resistant weeds (Wang et al. 2017). For determination of physiological perturbations, both Fernández-Escalada et al. (2016) and Maroli et al. (2017) investigated the metabolic changes induced in the weeds following exposure to nonlethal doses of glyphosate.

Although MOAs of most herbicides have been well identified, in many cases the sequence of phytotoxic events that result in plant death is unclear, particularly for slow-acting herbicides, which exhibit a significant time lag between herbicide application and plant death. Using genetics and biochemical and metabolic analyses, Fernández-Escalada et al. (2016) studied the physiological and metabolic perturbations of glyphosate-resistant (Fernández-Escalada et al. 2016, 2017) and glyphosate-tolerant (Maroli et al. 2017) weeds, and to examine herbicide metabolism in herbicide-resistant weeds (Wang et al. 2017). For determination of physiological perturbations, both Fernández-Escalada et al. (2016) and Maroli et al. (2017) investigated the metabolic changes induced in the weeds following exposure to nonlethal doses of glyphosate. Although MOAs of most herbicides have been well identified, in many cases the sequence of phytotoxic events that result in plant death is unclear, particularly for slow-acting herbicides, which exhibit a significant time lag between herbicide application and plant death. Using genetics and biochemical and metabolic analyses, Fernández-Escalada et al. (2016) studied the physiological and metabolic perturbations of glyphosate-resistant (Fernández-Escalada et al. 2016, 2017) and glyphosate-tolerant (Maroli et al. 2017) weeds, and to examine herbicide metabolism in herbicide-resistant weeds (Wang et al. 2017). For determination of physiological perturbations, both Fernández-Escalada et al. (2016) and Maroli et al. (2017) investigated the metabolic changes induced in the weeds following exposure to nonlethal doses of glyphosate. Although MOAs of most herbicides have been well identified, in many cases the sequence of phytotoxic events that result in plant death is unclear, particularly for slow-acting herbicides, which exhibit a significant time lag between herbicide application and plant death. Using genetics and biochemical and metabolic analyses, Fernández-Escalada et al. (2016) studied the physiological and metabolic perturbations of glyphosate-resistant (Fernández-Escalada et al. 2016, 2017) and glyphosate-tolerant (Maroli et al. 2017) weeds, and to examine herbicide metabolism in herbicide-resistant weeds (Wang et al. 2017). For determination of physiological perturbations, both Fernández-Escalada et al. (2016) and Maroli et al. (2017) investigated the metabolic changes induced in the weeds following exposure to nonlethal doses of glyphosate.
Figure 3. Designing a metabolomics study. (A) The various approaches for performing a metabolomics experimental study. GC-MS, gas chromatography–mass spectrometry; HILIC-LC-MS/MS, hydrophilic interaction chromatography for liquid chromatography–tandem mass spectrometry; LC-MS/MS, liquid chromatography–tandem mass spectrometry. (B) The general metabolomics workflow. It involves formulating a biological question, setting up an experimental design to test the hypothesis, sample treatment and harvest, metabolite extraction, clean-up, chromatographic separation, identification, statistical validation, and functional interpretation.
amino acids following glyphosate treatment is due to proteolysis. However, using SIRM analysis, Maroli et al. (2016) essentially described the underlying cause of higher amino acid accumulation in the susceptible biotype. It was shown that glyphosate-induced amino acid accumulation in susceptible *A. palmeri* biotypes is a consequence of proteolysis (catabolism) coupled with de novo synthesis of certain amino acids. In contrast, amino acid concentrations in the glyphosate-resistant biotype were predominantly due to de novo synthesis (anabolism). Thus, it can be seen from this study that the use of modern omics platforms has helped to establish the connection between metabolome and metabolic pool dynamics to elucidate the link between the glyphosate MOA and de novo amino acid synthesis.

**Integrated Omics Approaches to Understanding Phytotoxin MOA**

Herbicides with new MOAs are desperately needed to combat evolved and evolving herbicide resistance (Duke and Heap 2017), and no new commercial herbicides with a clearly new MOA have been commercialized since the 4-hydroxyphenylpyruvate dioxygenase inhibitors in the 1980s (Duke 2012). Thus, discovery of herbicides with new MOAs is of prime importance in herbicide discovery efforts. Evidence from the natural phytotoxin literature suggests that there are many more viable MOAs than the current 20 MOAs of commercial herbicides (Dayan and Duke 2014). However, determination of the MOA of phytotoxins is not a trivial pursuit, because what we observe after herbicide treatment of a plant is the manifestation of many secondary and tertiary effects resulting from an effect on the primary target site. The literature is full of papers confusing secondary and tertiary effects with primary effects. Many of the older herbicides were commercialized before their MOAs were known, partly because the target sites were not easy to determine, due to the difficulty in working back from physiological effects to a molecular target site.

With the advent of omics technologies, new strategies for MOA determination have been devised (Duke et al. 2013; Grossmann et al. 2012a). Omics-based MOA discovery consists of building a database of any one of the different omics responses to herbicides with known MOAs and then comparing the response profile of a phytotoxin with an unknown MOA to profiles generated by phytotoxins with known MOAs. To our knowledge, this has been done in industry with only one omics method—metabolomics. This approach can be highly effective if the new compound happens to have an MOA that is in the database of omics responses to compounds with known MOAs. If not, the approach will indicate that the compound has a new MOA not represented in the database. Most companies involved in herbicide discovery have tried this approach, but only BASF has published a detailed description of how it has used omics methods to identify MOAs (Grossman et al. 2012a, 2012b). In that case, it combined both metabolomic and physionomic methods to build extensive databases of omics responses to phytotoxins with known MOAs against which to evaluate data from compounds of unknown MOAs. While a growing number of omics technologies are available to choose from, some of them being quite narrow (for example, lipidomics and glycomics), the scope of this segment will be limited to transcriptomics, proteomics, metabolomics, physionomics, and combined approaches.

Most MOA transcriptomics has been done with *Arabidopsis*. Transcription responses to several herbicides with known MOAs have been published (Table 2). However, a major problem with this method of determination is that at doses of the toxicant that have even a subletal effect on the plant (for example, the dose that reduces growth by 50%), expression of many genes is affected within a short time after treatment. Many of the affected genes are those involved in stress responses and metabolic detoxification and other means of dealing with xenobiotics. For example, Baerson et al. (2005) found that the phytotoxic allelochemical benzoxazolin-2(3H)-one (BOA) upregulated many *Arabidopsis* genes involved in metabolism of xenobiotics and cell rescue and defense within 24 h after treatment. An extreme example is that of cantharidin, a potent natural phytotoxin that significantly affected gene expression of more than 6% of the genes of *Arabidopsis* within 2 h of treatment with a dose that reduced growth by 30% (Baja et al. 2011a, 2011b). Eventually, 10% of the genome was affected. This is not a surprise, as cantharidin and the herbicide chemical analogue endothall both inhibit all of the serine/threonine protein phosphatases (*Arabidopsis* has more than 20) in plants (Baja et al. 2011a, 2012). These enzymes are heavily involved in signaling pathways and gene expression.

Proteomics has been used considerably less than transcriptomics to probe the MOAs or mechanisms of resistance to herbicides. Zhang and Reichers (2008) reviewed the use of proteomics in weed science research. The effects of paraquat, diuron, and norflurazon on *Chlamydomonas reinhardtii* were studied with proteomics (Nestler et al. 2012). Although the abundance of the target protein of norflurazon, pythoene desaturase, was unaffected, the amounts of other enzymes of the plastidic terpene pathway were affected. Diuron increased the amount of its target, the D1 protein of photosystem II, whereas some other proteins involved in photosynthetic electron transport decreased. The effects of the auxinic herbicides dicamba and clopyralid on the proteome of soybean [*Glycine max* (L.) Merr.] were examined by Kelley et al. (2006). They found four proteins to be strongly affected, and one of them was the product of the GH3 gene, a gene that they found to be strongly upregulated at the transcriptional level. Kumari et al. (2009) found that butachlor reduced levels of proteins involved in photosynthesis and respiration of the alga *Aulosira fertilissima*. Because the MOA of butachlor is inhibition of very long chain lipid synthesis, these effects are secondary or tertiary. Likewise, amiprophos-methyl, a herbicide that affects microtubule function, had effects on proteins associated with diverse physiological and biochemical processes but not directly associated with tubulin (Wang et al. 2011). More recently, the natural phytotoxin α-terthienyl was found to affect 16 proteins associated with energy transduction, of which the transketolase protein was greatly reduced (Zhao et al. 2018). A transketolase-altered mutant was less sensitive to the phytotoxin, and the enzyme from the mutant was less inhibited by the compound. But the weak effect of the toxin on the enzyme is not what one would expect for a primary target site.

Studies using natural phytotoxins with unknown target sites have revealed distinct metabolic effects but no clear indication of a molecular target (Cantrell et al. 2007; Duke et al. 2011). Other metabolomic studies of phytotoxin MOAs are discussed in Duke et al. (2013). One of the more complete studies of this type was that of Trenkamp et al. (2009), who examined the effects of glufosinate, glyphosate, sulcotrione, foramsulfuron, benfuresate, and an experimental herbicide on the metabolome of *Arabi- dopsis*. Results matched the MOA for some but not all of the phytotoxins. More systematic approaches that rely on metabolic profiles of an array of phytotoxin MOAs have been more
successful (Grossmann et al. 2010, 2012a, 2012b). Perhaps the only new phytotoxin MOAs discovered by omics methods are the determination that cinmethylin’s target site is tyrosine amino-transferase (Grossmann et al. 2012b) and that of a phenylalanine analogue (PHE1) is IAA synthesis (Grossmann et al. 2012a), although in the latter case the specific enzyme target to the IAA synthesis pathway was not determined. In both cases, physionomic and metabolomic databases were used to narrow the search for the target sites. Verification of the omics indications were followed by physiological and biochemical studies.

Limitations, Conclusions, and Future Directions

Relying on orthologous genomes to annotate related genomes of weedy species has pitfalls associated with proposing biological interactions and processes based on spurious assumptions that homologous genes have conserved functionality across species (Doğramacı et al. 2015). Similarly, assembling a quality de novo reference transcriptome for weeds can be computationally difficult, due to complex gene families and high levels of heterozygosity that often occur in weeds. Polyploidy further complicates reference transcriptome assembly, although the assembly can be completed and yield insight into evolution of polyploidy in weeds (Chen et al. 2016). This becomes more critical when studying herbicide resistance, because RNA-Seq will detect all differences in gene expression and sequence; therefore, using highly unrelated resistant and susceptible populations will result in a large number of false positives (genes with significant DE that are completely unrelated to herbicide resistance). Hence, not all resistance mechanisms can be detected using RNA-Seq. Therefore, for a successful RNA-Seq experiment in non-model species such as weeds, ideally, a high-quality reference transcriptome is desired for identifying and quantifying DE genes or sequence variations. Additionally, the quality of the de novo reference transcriptome is also important, as genes will only be identified for DE and/or sequence variation if they are present in the reference assembly. In contrast, the major challenge for developing an effective high-throughput metabolomics platform lies in the chemical complexity, heterogeneity, and dynamic range of the metabolites and the challenges in developing a single extraction procedure for all metabolites. Plant extracts have a complicated biochemical composition and require extensive extraction and separation procedures to achieve reproducible results. Furthermore, very few of these metabolites can act as distinct biomarkers for a particular herbicide or phytotoxin MOA. Exceptions are EPSPS, PPO, and ceramide synthase, which cause dramatic increases in the pools of shikimic acid (Duke et al. 2003), protoporphyrin IX (Dayan and Duke 2003), and sphingoid bases (Abbas et al. 2002), respectively. Unfortunately, most other metabolites or phytotoxins do not have such dramatic effects.

As no single omics method is likely to reveal the MOA of a herbicide or natural phytotoxin, omics approaches to probe MOAs, though powerful, have to be used with caution. Several factors can influence the outcome of an omics experiment. First, physiological effects and responses are dose dependent, such that at high doses, results might be confounded by secondary targets, while at low doses, a plant might compensate too rapidly to observe anything meaningful. Second, the results are highly dependent on exposure time, wherein responses can be rapid, gradual, or delayed or sometimes may even reverse over time. Third, omics responses can vary between plant tissues and cell types, such that important effects in some cells could be masked when the entire tissue or organ is extracted. Finally, metabolic-pool sizes can be deceiving, as the pool size is determined by both input and output of the pool. In many cases, changes in pool fluxes would be much more informative about the effect of a herbicide than the pool size. Moreover, even when omics methods suggest a molecular target site, it must be verified by physiological and biochemical methods. For example, histone deacetylase was found to be the target site of a phytotoxic metabolic product of BOA, and its MOA was further probed by transcriptome analysis (Venturelli et al. 2015). It is evident that the transcriptome data would have been very unlikely to reveal the molecular target site.

Another caveat involving omics studies is that all herbicides and phytotoxins cause stress, including oxidative stress, so omics methods can be misleading to the naïve researcher. For example, Ahsan et al. (2008) found both paraquat and glyphosate to enhance the amount of proteins involved in defense against oxidative stress in leaves of glyphosate-susceptible rice (Oryza sativa L.). They concluded that this was an “alternative” effect, rather than a secondary or tertiary effect of herbicide-induced stress. Clearly, the approximately 50-fold level of resistance of crops made resistant to glyphosate by means of a glyphosate-resistant EPSPS (Nandula et al. 2007) is proof that there is no alternate primary effect of glyphosate. Comparing results with different omics approaches is rare, but quite different effects have been reported with different omics approaches. For example, in the same experiment in which cantharidin’s effects on the transcriptome were determined (Bajsa et al. 2011a, 2011b), samples were taken for proteome studies (Bajsa et al. 2015). A remarkable lack of correlation between transcriptome and proteome results was observed, although the lack of correspondence between transcriptome and proteome data could be due to multiple factors (Duke et al. 2013; Narayanan and Van de Ven 2014; Payne 2015). Similarly, Zhao et al. (2018) found decreases in the transketolase protein of Arabidopsis treated with α-terthienyl, but the gene for this enzyme was upregulated by the same treatment. They hypothesized that the decrease in protein was due to direct interaction with α-terthienyl, which resulted in upregulation of the gene to compensate.

The disconnect between comparing individual omics platforms to understand weed genetics, diversity, heterozygosity, and importantly, evolution of herbicide resistance in weeds, especially non-target site resistance, highlights the need to develop an integrated omics platform. As an example of developing blueprints for constructing low-cost genomic assemblies in weed species, Horvath et al. (2018) have sequenced gene space and transcriptome assemblies of E. esula that were used to identify promoter sequences, high-quality markers, and repetitive elements. Based on this framework, a reliable sequence for >90% of the expressed E. esula protein-coding genes was made available. Compared with conventional screening techniques, developing herbicides with new MOAs and chemistries or evaluating natural products for use as bioherbicides can be achieved at a much faster rate using next-generation omics. Despite requiring a cautionary approach, integrated systems biology can revolutionize weed management practices by providing hitherto unknown biological information (Han et al. 2017; Kraehmer 2012). A holistic line of action with multidisciplinary integrated approaches and collaboration between weed scientist, extension specialist, and farmers is required to allow for the development of long-term, weed management strategies. Though information on candidate genes is lacking for most weed species, global gene expression profiling
techniques, such as microarrays, can serve as effective tools for understanding NTSR mechanisms (Peng et al. 2010), while RNA-Seq and whole-metabolome profiling can identify genes and metabolites involved in regulating biochemical processes in a weed. An outcome of a systems biology approach is the ambitious RNAi technology (BioDirect™) developed by Monsanto to exploit precise RNA segments coding for EPSPS protein in reversing glyphosate resistance in weeds (Holommon 2012; Shaner and Beckie 2014). In conclusion, it can be said that genomics, transcriptomics, and other methods for high-throughput screening can yield promising results for elucidating basic weed biology concepts as well as insights into the response of weeds to biotic and abiotic stresses and crop–weed competition. Thus, with the aid of these omics platforms, improved knowledge of weed biology, genetics, and physiology can be gained quickly, paving the way for the development of long-term, sustainable weed management practices.

Acknowledgements. The authors would like to thank the WSSA for providing the funding and platform to organize this symposium at the 55th Annual Meeting of the Weed Science Society of America, Lexington, KY. The authors would also like to acknowledge Amy Lawton-Rauh and Patrick Tranel for their contributions to the symposium. No conflicts of interest have been declared.

References


Chen S, McElroy JS, Dane F, Goertzen LR (2016) Transcriptome assembly and comparison of an allotetraploid weed species, annual bluegrass, with its two diploid progenitor species, Poa supina Schrad and Poa annua K. Kunth. Plant Genome 9, 10.3835/plantgenome2015.06.0050


Chen S, McElroy JS, Dane F, Goertzen LR (2016) Transcriptome assembly and comparison of an allotetraploid weed species, annual bluegrass, with its two diploid progenitor species, Poa supina Schrad and Poa infirma Kunth. Plant Genome 9, 10.3835/plantgenome2015.06.0050


Manabe Y, Tinker N, Colberta V (2005) CSR1, the sole target of glyphosate in Arabidopsis thaliana. Proc Natl Acad Sci USA, 10.1073/pnas.1712372115
species. Accessed: April 12, 2018
species. Accessed: April 12, 2018
species. Accessed: April 12, 2018
Mondula VK, Reddy KN, Koger CH, Poston DH, Rimando AM, Duke SO, Bond JA, Ribeiro DN (2012) Multiple resistance to glyphosate and
pyrithiobac in Palmer amaranth (*Amaranthus palmeri*) from Mississippi and response to flumiclorac. Weed Sci 60:179–188


Raghavan V, Ong EK, Dalling MJ, Stevenson TW (2006) Regulation of genes associated with auxin, ethylene and ABA pathways by 2,4-dichlorophenoxy-acetic acids in *Arabidopsis*. Funct Integr Genom 6:60–70


Zhao L, Hu Q, Huang Y, Keller AA (2017b) Response at genetic, metabolic, and physiological levels of maize (Zea mays) exposed to a Cu(OH)2 nanopesticide. ACS Sustainable Chem Eng 5:8294–8301

