

## Molecular Mechanisms of Herbicide Resistance

Christophe Délye, Arnaud Duhoux, Fanny Pernin, Chance W. Riggins, and Patrick J. Tranel\*

**Key words:** Allele-specific PCR, dCAPS, DNA sequencing, gene expression, mutation, PCR, qPCR.

Resistance to herbicides occurs in weeds as the result of evolutionary adaptation (Jasieniuk et al. 1996). Basically, two types of mechanisms are involved in resistance (Beckie and Tardif 2012; Délye 2013). Target-site resistance (TSR) is caused by changes in the tridimensional structure of the herbicide target protein that decrease herbicide binding, or by increased activity (e.g., due to increased expression or increased intrinsic activity) of the target protein. Nontarget-site resistance (NTSR) is endowed by any mechanism not belonging to TSR, e.g., reduction in herbicide uptake or translocation in the plant, or enhanced herbicide detoxification (reviewed in Délye 2013; Yuan et al. 2007).

Mutations endowing herbicide resistance can be classified into two types. The first type is structural changes in a DNA sequence encoding a protein, i.e., structural mutations. Structural mutations endowing herbicide resistance are expected to cause a structural modification in the tridimensional structure of a protein that will lead to a decrease in the efficacy of an herbicide. For example, mutations conferring an amino acid substitution at the herbicide-binding site of a target protein can decrease the affinity of the herbicide for the target protein (TSR). Alternatively, mutations at the active site of a metabolic enzyme or a transporter protein can improve the activity of these proteins in herbicide degradation or compartmentation away from its site of action, respectively (NTSR). In the case of structural changes in DNA sequence, seeking the cause for resistance means identifying and being able to detect the relevant structural mutations in the DNA of resistant plants.

The second type of mutations associated with herbicide resistance results in a difference in the

expression of one or several genes in resistant plants compared to sensitive plants, i.e., regulatory mutations (Délye 2013; Yuan et al. 2007). These mutations are changes in a DNA sequence that can cause an increase in the expression of the herbicide target protein that compensates for the herbicide inhibitory action (TSR), or a variation in the expression of herbicide-metabolizing enzyme(s) or of transporter proteins that will lead to an increase in herbicide degradation or compartmentation away from its site of action, respectively (NTSR). Identifying regulatory mutation(s) responsible for changes in gene expression is not straightforward, because these mutations can be of diverse nature. Examples include whole-gene amplification (e.g., Gaines et al. 2010), structural changes in the promoter sequence of the gene encoding the protein showing a variation in expression, or even structural changes in the promoter sequence or in the coding sequence of a gene encoding a protein that regulates the expression of the protein showing a variation in expression (Délye 2013). Epigenetic processes (e.g., DNA methylation) can also be involved in the regulation of gene expression (Délye 2013). Thus, in the case of regulatory mutations or of epigenetic regulation, seeking and detecting herbicide resistance is most easily achieved by identifying and being able to detect significant differences in expression of genes between resistant and sensitive plants.

A variety of approaches are available for confirming and evaluating herbicide resistance in weeds (Burgos et al. 2013). The aim of this paper is to provide the inexperienced researcher with information to investigate herbicide resistance at the DNA level. Protocols and guidelines are provided for investigating both structural changes in DNA sequence and changes in gene expression.

DOI: 10.1614/WS-D-13-00096.1

\* First, second, and third authors: Researcher, PhD Student, and Research Technical Assistant, INRA, UMR1347 Agroécologie, F-21000 Dijon, France; fourth and fifth authors: Postdoctoral Research Assistant and Professor, Department of Crop Sciences, University of Illinois, Urbana, IL 61801. Corresponding author's E-mail: tranel@illinois.edu

### DNA and RNA Basics

Most DNA-based assays for herbicide resistance rely on the polymerase chain reaction (PCR) to selectively amplify a DNA sequence of interest from the milieu of DNA that is not of interest. Thus, we

begin with a discussion of extracting DNA for use in PCR, performing the PCR, and analyzing the PCR products.

Most standard “genomic” DNA extraction procedures yield DNA from the nuclear, chloroplastic, and mitochondrial genomes, and thus are suitable for a wide range of downstream molecular analyses, including PCR. Although DNA can be extracted from all types of plant material (e.g., leaves, roots, stems, seeds, preserved tissue), young, newly emerged leaves are often best for extraction because they are easier to grind up than mature tissue and typically contain lower amounts of polysaccharides, polyphenolics, and other secondary metabolites that can interfere with DNA isolation. Attention should be given to selecting healthy, clean tissue (e.g., without visible signs of infection, fungi, insects, soil) and proper handling to prevent plant-to-plant contamination. In the absence of fresh tissue, high-quality DNA can also be extracted from preserved material. For example, if immediate processing of fresh material is not possible (such as under field conditions), then the plant tissue can be kept on ice for several hours or preserved for future workup by drying or freezing. Leaves should be dried relatively quickly and stored in a moisture-free environment to prevent rotting. Leaf tissue can be pressed between paper towels and air-dried for several days at room temperature or placed in a zip-lock bag with silica gel to absorb moisture. Alternatively, leaf tissue can be frozen by first rinsing and blotting dry, and then storing at  $-20\text{ C}$  or indefinitely at  $-80\text{ C}$ . Frozen tissue should not be allowed to thaw before processing because this increases the risk of DNA degradation by endogenous nucleases.

Following are two common and relatively simple protocols that we use routinely to obtain DNA suitable for PCR and other molecular assays. These protocols are only two of numerous DNA extraction methods that have been published in the literature or are available online, and there is also a variety of commercial DNA extraction kits available, such as the DNeasy Plant kits (Qiagen, Valencia, CA) and E.Z.N.A. Plant DNA kits (Omega Bio-Tek, Norcross, GA). These and other commercial kits typically yield DNA of higher purity, but one must factor in the cost of the kits with the time, expertise, scale, and goals of the project.

**Tip:** Use molecular biology-grade reagents for all of the following protocols.

**Easy Way: Brutus DNA Extraction.** Brutus DNA extraction basically consists of grinding a plant fragment in a salty buffer, then boiling it to release DNA from plant tissues (Délye et al. 2002). The quality and quantity of the DNA solution is sufficient to provide DNA templates suitable for a range of PCR-based techniques, ranging from basic PCRs to sophisticated techniques such as TaqMan (e.g., Délye et al. 2010). This easy DNA extraction protocol is fast and costs almost nothing. It can be performed from different plant tissues (leaf, root, green stem), either fresh or dried (Délye et al. 2011); however, Brutus extraction from ripe seeds is often not suitable for PCR.

Materials needed: Extraction buffer (to make 500 ml, mix 50 ml 1 M Tris-HCl, pH = 9.5, 10 ml 0.5 M EDTA pH = 8, and 37.275 g KCl in 200 ml deionized water, bring to 500 ml final volume with deionized water and sterilize in an autoclave); a means to grind the tissue, e.g., a bead mill or disposable plastic pestles; and a tube-heating device at  $95\text{ C}$  (e.g., a water bath).

**Tip:** Any type of grinding device can be used, provided it efficiently shreds plant tissues. For processing a small number of samples, tissue can be ground with pestles designed for use in 1.5 ml microcentrifuge tubes (e.g., Fisher Scientific 12-141-364; Pittsburgh, PA). For a large number of samples, tissue can be ground in 1.5 ml microcentrifuge tubes with a 3-mm glass bead or in 0.5 ml or 0.2 ml microcentrifuge tubes with a 2-mm glass bead, using a bead mill (e.g., Retsch MM400; Hann, Germany).

General steps are as follows:

1. Collect 10 to 20 mm<sup>2</sup> section of plant tissue in a microcentrifuge tube.
2. Add 100  $\mu\text{l}$  extraction buffer per sample.
3. Grind tissue. If using a bead mill, use two rounds of 90 s at 30 shakes  $\text{s}^{-1}$ .
4. Incubate tubes at  $95\text{ C}$  for 5 min.
5. Place tubes in ice for 10 min.
6. Centrifuge tubes (at least  $3,000$  to  $4,000 \times g$  for 1 min, so that tissue fragments lie at the bottom of the tube and are not pipetted to reaction mixes in downstream experiments).
7. Store DNA samples at  $-20\text{ C}$ , or use immediately.

**Tip:** Brutus DNA samples are very stable, even under repeated freeze/thaw cycles. Furthermore, tubes containing the mixture of boiled buffer and tissues can be reused when the DNA solution has run out: simply add 50 to 100  $\mu\text{l}$  fresh extraction buffer to the ground plant tissue remaining in the tube, and repeat the extraction procedure.

**CTAB DNA Extraction.** One of the most popular plant DNA extraction buffers is CTAB (cetyltrimethyl ammonium bromide; also called hexadecyltrimethyl ammonium bromide). The most commonly used CTAB extraction protocol is that of Doyle and Doyle (1990), and it has been successfully employed for a wide variety of plant species. Numerous modifications of this basic method are available. The following protocol works for several species, and would be a good option to try if the Brutus method is not successful. The CTAB procedure includes chloroform extraction and DNA precipitation steps, which result in cleaner DNA than is obtained with the Brutus method.

Material needed: CTAB extraction buffer (to make 100 ml, mix 10 ml 1M Tris-HCl pH 8, 28 ml 5M NaCl, 4 ml 0.5M EDTA, 2 g CTAB, and 40 ml deionized water; heat to dissolve; bring to 100 ml final volume with deionized water; and filter sterilize with e.g., Nalgene Rapid-Flow Tissue Culture Filter Unit, 0.22  $\mu\text{m}$ ; (Rochester, NY); pestles for grinding samples; microcentrifuge capable of spinning 1.5 ml polypropylene tubes at  $13,500 \times g$  at room temperature; water bath.

**Tip:** The protocol provided uses a pestle for grinding samples, but could be modified to use a bead mill for higher sample throughput.

The extraction steps are as follows:

1. Collect 10 to 20  $\text{mm}^2$  tissue in a 1.5 ml tube and grind with pestle.
2. Add 600  $\mu\text{l}$  CTAB buffer to ground tissue with pestles still in tubes. Several samples can be processed at once. Additional grinding with pestles might be necessary to fully pulverize tissue before proceeding to the next step.
3. Incubate homogenate at 65 C for 20 min to 1 h in water bath. Occasionally mix by inversion to avoid aggregation of homogenate.
4. Add 400  $\mu\text{l}$  chloroform and mix by inversion.

**Tip:** Chloroform is often used as a solvent in DNA extraction protocols; however, dichloromethane is less toxic and can be substituted for chloroform (Chaves et al. 1995).

5. Centrifuge 5 min at  $\geq 13,500 \times g$  in a microcentrifuge at room temperature.
6. Transfer upper aqueous layer (approximately 500  $\mu\text{l}$ ) into a new 1.5 ml tube and add an equal volume of isopropanol. Mix by inversion.
7. Centrifuge 10 min at  $\geq 13,500 \times g$  at room temperature.
8. Discard supernatant and add 250  $\mu\text{l}$  of 80% ethanol. Centrifuge for 2 min as before.

**Tip:** The nucleic acid pellet is usually visible at this stage and attention must be given not to lose the pellet when decanting.

9. Discard supernatant and add 250  $\mu\text{l}$  of 95% ethanol. Centrifuge for 2 min as before.
10. Discard supernatant and dry pellet completely.

**Tip:** Residual ethanol should be removed before dissolving the DNA pellet in sterile distilled water or TE buffer because it can inhibit downstream enzymatic reactions. The pellet can be air-dried for 30 min, or a Savant SpeedVac (Thermo Scientific, Waltham, MA) with medium heat can also be used to dry the pellet. Do not dry the pellets too long in the SpeedVac (5 min is usually long enough) or they will be difficult to resuspend.

11. Resuspend dry DNA pellet in 100  $\mu\text{l}$  water or TE buffer. Store at  $-20 \text{ C}$ .

**Tip:** DNA is more stable in TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) than in water, so TE buffer is recommended for long-term storage. However, EDTA can interfere with PCR, so we often store DNA in a 0.1 $\times$  EDTA TE buffer (10 mM Tris-HCl pH 8, 0.1 mM EDTA).

**RNA Extraction and cDNA Synthesis.** DNA can also be synthesized from messenger RNA (mRNA) using a reverse-transcriptase enzyme. This enzyme synthesizes DNA complementary to RNA (cDNA) from the 3' end of a primer (a small, single-strand DNA molecule) hybridized on the RNA strand, using the RNA strand as a template. cDNA is of particular interest when working on genes with complex intron–exon structure, because, like mRNAs, cDNAs do not contain introns. For “simple” genes such as the gene encoding acetolactate synthase (ALS) (ca. 2,000 base pairs [bp] with no introns in most plants), genomic DNA and cDNA are identical (within transcribed regions). For genes such as the gene encoding chloroplastic acetyl CoA carboxylase (ACCase) in grasses (family Poaceae), the cDNA sequence is ca. 7,000 bp long. The corresponding genomic DNA is ca. 10,000 to 12,000 bp long, and contains 32 introns (Délye 2005; Huang et al. 2002).

Traditional protocols for RNA extraction and purification and cDNA synthesis are time consuming, and often involve either dangerous reagents, and/or complex purification procedures. However, RNA extraction and cDNA synthesis are now most easily done using combinations of commercial kits, which can be used without particular training. An example of a combination of widely distributed kits yielding good quality cDNA is RNeasy Plant Mini Kit (Qiagen) plus RNase Free DNase Set for RNA extraction and genomic DNA removal (Qiagen), followed by 5PRIME Masterscript Kit (Fisher Scientific) for reverse-transcription reactions.

**Tip:** This is only one example. Other kits can be combined to produce cDNA.

**Tip:** RNA can be extracted from all living tissues. Young, actively growing tissues typically will provide the highest yields of RNA. However, when extracting RNA (unlike when extracting DNA), one must select tissue in which the gene of interest is likely to be expressed (e.g., aboveground plant part for foliar herbicides, roots for soil-applied herbicides).

RNA must be treated with care. RNA is a single-stranded molecule and, due to its chemistry, and to the presence of RNAses everywhere, including in

your plant samples and on your skin, RNA is more fragile than double-stranded DNA. It is essential that RNAses are absent or neutralized when handling RNA (efficient RNase-neutralizing solutions are commercially available). The extraction steps are as follows:

1. Proceed to the extraction of RNA from plant tissue samples immediately after collection. Because collection of plant tissue can induce wound stress response or RNA degradation due to RNase release, it is best to freeze samples in liquid nitrogen immediately after collection. If plant tissue sample are not to be processed immediately, store at  $-80\text{ C}$ .

**Tip:** Never let a frozen sample thaw, because this causes RNA degradation.

2. Use RNase-free plasticware and reagents.
3. Perform reverse-transcription reactions as soon as possible after RNA extraction.
4. Check RNA quantity and quality using a UV spectrophotometer. RNA quality is assessed using the A260/A280 and A260/A230 absorption ratios. RNA samples with A260/A280 and A260/A230 ratios between 1.8 and 2.2 are generally considered suitable.
5. Do not store RNA samples more than a few days at  $-20\text{ C}$ . Longer-term storage must be at  $-80\text{ C}$ . When storing RNA samples for several months, check RNA quantity and quality again prior to any experiment to estimate RNA degradation.

**Tip:** cDNA is much more stable than RNA and, depending on the downstream applications, might be a better option for long-term storage. RNA degradation can occur within a few weeks even if stored at  $-80\text{ C}$ .

6. Do not let RNA pellets dry. Rapidly dissolve in RNase-free water or buffer.
7. Do not subject RNA samples to frequent freeze-and-thaw cycles. Let RNA samples gently thaw on ice. Work on ice as much as possible.

After RNA is extracted, cDNA synthesis using reverse-transcription mixes can be conducted with several types of primers. If no specific gene is

targeted, or the sequence of gene(s) of interest is not known in the species considered, then random hexamer primers (i.e., a mixture of DNA molecules consisting of six randomly chosen nucleotides) or poly-dT primers (i.e., a primer including a poly-T sequence that is complementary to the poly-A tail of mRNAs), or both, can be used for reverse transcription. Random hexamer primers are expected to generate short cDNA fragments by hybridizing randomly on their complementary sequences on mRNAs. Poly-dT primers are expected to generate cDNA fragments corresponding to the 3' end of mRNAs. If a specific, known gene or region of a gene is targeted, then gene-specific primers can be used for the reverse transcription.

**PCR.** The polymerase chain reaction (PCR) can massively replicate a given DNA region (amplicon) from small or minute amounts of DNA. PCR enables easy and rapid gene sequencing. PCR is also the tool of choice for DNA-based mutation diagnosis. A PCR reaction mix typically consists of a buffer containing template DNA or cDNA, primers (see below), a thermostable DNA polymerase (e.g., *Taq* polymerase), and the four DNA nucleotides (dNTPs). PCRs are run as a succession of cycles, each with three steps, which are carried out in a thermocycler. At each cycle, the quantity of amplicon present in the reaction mix is theoretically doubled, until the dNTP stock is exhausted. The three steps in each cycle are:

1. Denaturation: the reaction mix is heated at 93 to 95 C to dissociate double helix DNA into single stranded DNA.
2. Annealing: the temperature is lowered to allow primer hybridization to their target DNA sequences.
3. Extension: the temperature is increased to the reaction temperature optimal for DNA polymerase activity (68 to 72 C). The DNA polymerase synthesizes a new DNA strand by adding dNTPs to form a sequence complementary to that of the DNA template strand, starting from the 3' end of the primers.

*Primer Design.* A primer is a short (10 to 30 bp) single strand of nucleic acid (oligonucleotide) that serves as a starting point for DNA synthesis. A pair of primers is used in PCR to amplify the DNA region flanked by the primers. During the PCR, the primers hybridize with their target sequence on

DNA. Subsequent DNA synthesis occurs from the 3' end of a primer using the DNA strand the primer hybridized with as a template.

For a pair of primers, one primer ("forward" primer) has the same sequence as a short region on the coding strand of template DNA, and the other ("reverse" primer) has the same sequence as a short region on the complementary strand of template DNA. Forward primers hybridize to their complementary strand, which is the noncoding strand of template DNA, whereas reverse primers hybridize to the coding strand of template DNA. The degree of specificity of a primer is directly linked to how closely its sequence matches its target sequence. The last five bases at the primer 3' end are particularly crucial for specificity.

**Tip:** The longer the primers, the more stable the hybridization to its target sequence. Optimal primer length for basic PCR is generally considered to be 20 to 25 nucleotides. The pH of the PCR mix and the temperature used for primer hybridization in PCRs can influence the specificity of primers.

If the sequence of the targeted gene is known in the species of interest, primer design can be directly performed using software such as Primer3 (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). Software programs such as these design and rank primers based on their stability and specificity, taking into account the PCR conditions and input data, such as expected primer size range, expected amplicon size, and possibility for the primers to hybridize together or outside of the targeted region.

If the sequence of the targeted gene is not known in the species of interest, then primer design can be done based on the alignment of sequences of the homologous gene from closely related species (if available) as follows:

1. Obtain nucleotide sequences for homologues of the gene of interest that are available in the GenBank/EMBL database (<http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>).
2. Align these sequences. Several tools are available for this purpose, including online tools (e.g., <http://multalin.toulouse.inra.fr/multalin/>).
3. Search for DNA regions containing fully or very highly conserved sequences among species.
4. Design primers targeting these conserved regions.

**Tip:** If DNA sequence alignments do not enable easy identification of conserved regions, try aligning protein sequences. Protein sequences are often more conserved than DNA sequences; alignment of protein sequences allows you to focus in on areas of the gene that are most conserved.

**Tip:** The nucleotides at the 3' end of the primers are crucial for primer specificity and should be positioned on nucleotide positions very highly conserved in the alignment. The last 3' nucleotide of the primer should not be positioned on the third nucleotide of a codon, because this nucleotide can vary among individuals in a species.

If few DNA regions are highly conserved among the sequences aligned, there are several options regarding the variable nucleotides: (1) use the nucleotide most often present at a given position; (2) use inosine at variable positions (inosine forms hydrogen bonds with all four natural DNA bases) (Ohtsuka et al. 1985); or (3) use a mixture of oligonucleotides that vary at one or a small number of nucleotides (often called degenerate primers).

Once the primers are designed, they can be custom synthesized for you by any one of several companies (e.g., Integrated DNA Technologies, Coralville, IA; Life Technologies, Carlsbad, CA; Eurofins MWG Operon, Huntsville, AL).

**Tip:** Generally, ordering oligonucleotides using the default or least-expensive options in terms of quantity and purity is sufficient for most PCR applications.

**Tip:** Custom oligonucleotides are now very inexpensive (if ordering just a pair of primers, the shipping cost might be more expensive than the oligonucleotides themselves). For this reason, and because primers predicted to work sometimes do not, it often is advantageous to design and order a few different primers and then determine which combinations work best.

*Optimizing PCR.* Several parameters can be adjusted to optimize PCR. Mostly, PCR is optimized by adjusting the composition of the reaction mix (including primer and DNA concentration) and the program entered in the thermocycler. In this paper, we do not detail PCR optimization because this has been done elsewhere (e.g., Roux 1995, 2009). Our aim is to provide guidelines on how to set up a robust PCR assay.

*PCR mix.* There is a broad variety of commercial and in-lab mixes available for PCR. Optimizing a PCR mix can be a tedious issue. It is generally a time-saving approach to use a PCR mix that has proved effective for a purpose similar to yours, and to optimize PCR cycling programs. The PCR mix provided in Table 1 has proved robust and efficient for PCR using DNA from different extraction procedures and a variety of thermostable polymerases. (Délye et al. 2002)

**Tip:** PCR efficacy is dependent on the primer concentration. It is often a good idea to test a range of primer concentrations when optimizing a PCR reaction.

**Tip:** Commercial kits exist that are designed to optimize PCR efficacy and/or specificity.

*PCR programs.* In many instances, short durations can be successfully used for denaturation, annealing and extension steps. This saves time and can spare thermocyclers. A typical PCR program is as follows:

1. Initial denaturation step (optional): 5 min, 95 C.
2. Three-step cycles consisting of: denaturation (5 s at 95 C), annealing (10 to 30 s at annealing temperature), and extension (15 to 120 s at 72 C); repeated 30 to 40 times.
3. Final extension step (optional): 5 to 10 min at 72 C.

The initial denaturation step can be used to ensure complete DNA denaturation for difficult DNA targets (i.e., high guanine and cytosine [GC] contents) or for crude DNA extracts (e.g., Brutus extracts). Also, this step is generally mandatory if you are using an antibody-bound DNA polymerase (e.g., a "Hot Start" DNA polymerase; available from numerous sources). For the following (cycled)

Table 1. Example mix for polymerase chain reaction (PCR), and stock solutions for preparing the mix.<sup>a,b</sup>

|                              |   |
|------------------------------|---|
| 10× buffer (for 10 ml)       | 1.66 ml (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> 1 M [1.66 mM]<br>6.7 ml Tris-HCl 1 M, pH = 8.8 [670 mM]<br>200 µl MgCl <sub>2</sub> 1 M [20 mM]<br>Bring to 10 ml final volume with distilled water<br>Sterilize with autoclave<br>Add 70 µl β-mercaptoethanol 14.3 M [100 mM]<br>Store at -20 C  |
| 2× stock solution (for 1 ml) | 200 µl 10× buffer<br>100 µl Brij 58 at 1% wt/vol (Sigma-Aldrich, P5884) <sup>c</sup> [0.1%]<br>14 µl dNTP mixture (dATP, dCTP, dTTP, and dGTP each at 25 mM) [350 µM each dNTP] <sup>d</sup><br>20 µl bovine serum albumin 20 mg ml <sup>-1</sup> [400 µg ml <sup>-1</sup> ]<br>Bring to 1 ml final volume with distilled water<br>Store at -20 C |
| PCR mix (20 µl)              | 10 µl 2× stock solution<br>5 µl primer solution 0.4 to 1.6 µM [0.1 to 0.4 µM] <sup>e</sup><br>5 µl DNA<br>0.2 to 0.5 units polymerase   |

<sup>a</sup> Use molecular biology-grade reagents.

<sup>b</sup> Final concentration in the solution is given between brackets.

<sup>c</sup> Source: Sigma-Aldrich, St. Louis, MO.

<sup>d</sup> Abbreviations: dNTP, deoxynucleoside triphosphate; dATP, deoxyadenosine triphosphate; dCTP, deoxycytidine triphosphate, dTTP, deoxythymidine triphosphate; dGTP, deoxyguanosine triphosphate.

<sup>e</sup> Range of primer concentrations that most often works.

denaturation steps, 5 s is generally sufficient. Short denaturation steps also preserve the activity of the polymerase. For annealing, 10 to 30 s is generally sufficient. Try several annealing temperatures to optimize the PCR. Generally, testing the annealing temperature computed for the primers and temperatures 5 C above and below this temperature will be sufficient. Alternatively, test 55, 60, and 65 C.

**Tip:** Annealing temperature is an important determinant for PCR success. Decreasing the annealing temperature might enhance PCR efficiency but increases the risk for nonspecific PCR amplification. Conversely, increasing the annealing temperature can decrease the PCR efficiency, but decreases the risk for nonspecific PCR amplification. The trick is to select a temperature that is a compromise between specificity and efficacy of PCR amplification.

**Tip:** Annealing temperatures cannot exceed the temperature used for the extension step (generally 72 C). If your primers have an annealing temperature close to the extension temperature, try two-step cycles: 5 s denaturation, directly followed by extension.

For the extension step, a general rule is to use 1 min extension time per 1,000 bp of the targeted amplicon. For amplicons longer than 2,000 to 2,500 bases, dedicated types of thermostable DNA polymerases or mixtures of polymerases dedicated to “long range” PCRs are commercially available.

**Tip:** Many papers describing PCR methods use a final extension step at the end of the cycling program, typically consisting of 10 min at 72 C. In many cases, this is not necessary.

**Tip:** Most thermocyclers provide the option of holding the samples at 4 C after the PCR program. This is not necessary, because PCR mixes can safely remain at room temperature for several hours. Not using this option can also spare thermocyclers.

*Amount of DNA.* The PCR method is extremely sensitive, requiring only a few DNA molecules to be present in the reaction mix to yield successful amplification.

**Tip:** Brutus DNA solutions are not quantified. Test a range of 5-fold or 10-fold dilutions to identify the best working dilution in your lab. Generally, 25-fold to 100-fold dilutions are adequate. If you used a similar amount of plant tissues for extraction for all your samples, the dilution test only needs to be performed on a few samples. Then, use the same dilution for all samples.

*Designing a Proper PCR Experiment.* The PCR method is extremely efficient in amplifying DNA. Thus extreme care must be taken to reduce the likelihood for sample-to-sample contamination. A typical PCR experiment should include: (1) the untested samples to be amplified; (2) positive controls, i.e., samples for which successful amplification had been obtained consistently (to check for the efficacy of PCR); and (3) negative controls, i.e., samples without DNA (no amplification should be obtained for these samples, which check for DNA contamination of the PCR reagents).

**Tip:** In case of DNA contamination in your PCR assay, it is often most effective to use new batches of all reagents (PCR mixes, polymerase, primers and distilled water) than to try to identify the source of contamination. Also, we recommend using pipette tips with aerosol barriers to reduce the potential for sample contamination.

**Gel Electrophoresis of PCR Products.** Horizontal agarose gel electrophoresis is a simple method for the separation of PCR products or amplicons based on their size and charge. An electric field is applied to the gel and induces the negatively charged DNA molecules to migrate through an agarose matrix towards the anode. Short, lightweight molecules migrate faster through the agarose matrix of the gel, and are therefore separated from longer, heavier molecules. At the end of the electrophoresis, DNA molecules are revealed using one of various DNA staining reagents. The following materials are needed:

- Electrophoresis buffer: Tris Borate EDTA (TBE) buffer 0.5× (obtained by dilution of a 10× stock solution with deionized water. To make 1 L 10× stock solution, dilute in 500 ml deionized water: 108 g Tris base, 55 g boric acid, and 20 ml 1 M EDTA; bring to 1 L final

volume with deionized water and sterilize in an autoclave).

- Loading buffer. To make 100 ml, mix 50 ml 1 M EDTA pH = 8, 1.5 g ficoll, 24 g urea, 0.05 g bromophenol blue, and 0.05 g xylene-cyanol, then bring to 100 ml final volume with autoclaved deionized water.
- Agarose, molecular biology grade.
- Microwave oven.
- Horizontal electrophoresis system with power source.
- UV transilluminator and camera.
- Ethidium bromide (can be purchased as a 10 mg ml<sup>-1</sup> solution) or other DNA staining agent.

**Tip:** Ethidium bromide is considered a mutagen and must be handled with care. Dedicated, clearly identified areas must be set up for manipulations involving ethidium bromide. There are alternatives to ethidium bromide that are advertised as being safer (e.g., SYBRgreen-based dyes available from several sources), but ethidium bromide is still most broadly used.

General steps for gel electrophoresis are as follows:

1. Weigh the agarose. See Table 2 to determine the amount of agarose. Place it in an Erlenmeyer flask with half the volume of 0.5× TBE buffer required for the gel. Do not close the flask.
2. Melt the agarose gel in the buffer by microwaving up to ebullition, then gently agitating the flask (careful, it is hot!) until obtaining a homogenous, translucent solution (no solid agarose grains must be seen).
3. Add the rest of the volume of 0.5× TBE buffer required for the gel. Homogenize by gently agitating.
4. Pour the gel mixture on its support with a comb (to form the loading wells). Wait for complete solidification.
5. Place the gel in the electrophoresis device. Cover with 0.5× TBE buffer.
6. Add 1 µl loading buffer per 20 µl PCR product to each sample. Centrifuge briefly.

**Tip:** After the PCR is complete, the samples are ready for electrophoretic analysis. Samples can be analysed immediately after PCR, or stored at 4 C for later analysis (−20 C for long-term storage).

Table 2. Concentration of agarose to be used for gel electrophoresis according to the expected size in base pairs (bp) of your amplicon of interest.

| Amplicon size   | Agarose concentration<br>(g agarose 100 ml <sup>-1</sup> 0.5× TBE<br>[tris borate EDTA] buffer) |
|-----------------|---|
| Above 1,000 bp  | 0.8   |
| 600 to 1,000 bp | 1   |
| 200 to 600 bp   | 2   |
| < 200 bp        | 3 to 3.5  |

**Tip:** To avoid transferring amplicons from one PCR reaction to another, the loading buffer can be deposited at the top of the tube. It will blend with the PCR mix during the centrifugation step. If this is done carefully, there is no need to change pipette tips between samples. Alternately, there are commercial PCR mixes incorporating a loading buffer, which allows for sample loading on electrophoresis gels after PCR completion without further workup.

7. Load 1 to 15 µl of each sample on the gel. Loading a commercially available DNA ladder on one gel lane is generally useful for amplicon size estimation.
8. Run electrophoresis at 100V for 20 min to 45 min, depending on the size of the gel and the degree of separation expected (the longer the migration, the better the separation).

Amplicons are generally visualized with a UV transilluminator after dipping the gel into an ethidium bromide solution (0.5 µg ml<sup>-1</sup>). Alternatively, ethidium bromide can be added to the gel (0.5 µg ml<sup>-1</sup>) just prior to casting it. Ethidium bromide renders DNA fluorescent under UV light by intercalating between the DNA bases. Other DNA-staining reagents can also be used. Use a camera to save a photograph of the gel.

**Tip:** Be sure to check the negative controls, which should not show any trace of an amplicon. If they do, all the PCR experiment results are invalidated. Do not be ashamed: PCR is a very sensitive technique, and contamination can happen even with experienced molecular biologists.

## DNA Sequencing

The easiest way to obtain sequence data for your gene of interest is to directly use the PCR amplicon

as a template for Sanger sequencing (Sanger et al. 1977). Sanger sequencing consists of two steps: (1) sequencing reaction using a DNA fragment as a template, and special nucleotides that terminate and label the amplified molecules; and (2) visualization of the labelled molecules by capillary electrophoresis. This second step requires expensive equipment, and so is carried out by, for example, a sequencing company or a centralized campus facility. The first step can be contracted along with the second step, or can be done in-lab using for example, the Applied Biosystems BigDye Terminator Cycle Sequencing Kit (Life Technologies).

**Tip:** Classical thermostable DNA polymerases do not have a proofreading activity. Thus, erroneous nucleotides are incorporated in some of the DNA molecules generated in the PCR mixes (Tindall and Kunkel 1988). It is important to realize that the amplicon used for sequencing is actually a population of individual DNA molecules. When an amplicon is directly sequenced, the sequence obtained is a consensus of the sequences of the DNA molecules generated by PCR and present in the PCR mix. If misincorporation occurs during the first PCR cycles, erroneous DNA molecules will represent a substantial part of the DNA molecules in the amplicon. Thus, erroneous nucleotides will show in the consensus sequence. One option to lower the risk of sequencing errors, is to perform sequencing from at least three to five different PCR mixes (i.e., perform three to five identical PCR reactions in separate tubes, pool them, and sequence the pool). Alternatively, use a thermostable polymerase with proofreading activity.

It is generally advisable to first purify the PCR amplicon that will be used as the template prior to the sequencing reaction. There is a range of kits available to purify PCR amplicons of different sizes. Alternatively, sequencing companies propose amplicon purification as part of their sequencing service. If amplicons other than the one of expected size are observed during gel electrophoresis (even if minor in abundance), then the amplicon of interest should be isolated by gel purification, because the sequence of all amplicons in the PCR mix will superimpose to generate the consensus sequence observed. Sequencing different amplicons in the same Sanger reaction generally yields illegible sequences. To “gel-purify” the amplicon:

1. Pour an agarose gel with wide combs and leave at least one empty lane between samples.
2. Load the entire PCR contents into the wells and run the gel for 30 to 45 min (allow enough time to fully separate the bands).
3. Stain the gel with ethidium bromide.
4. Place the gel on a UV transilluminator (wear eye and skin protection to prevent direct exposure to UV light), turn on the UV lamp, and carefully excise the band with a razor or bistoury blade. The amplicon can then be purified with a commercial gel isolation kit (e.g., QIAquick Gel Extraction Kit, Qiagen) and can subsequently serve as the template for sequencing.

For the sequencing reaction, the same PCR primers used to generate the amplicon can be used. Only one primer is used in each sequencing reaction. It is generally recommended to perform two sequencing reactions, one with each primer. Sequencing from the forward primer will yield the sequence of the coding DNA strand, and sequencing from the reverse primer will yield the sequence of the complementary DNA strand. Thus, the amplicon will be sequenced on both strands. The forward and reverse sequences are then compared to detect and eliminate sequencing errors. The forward sequence of the amplicon can be obtained from the reverse sequence by reverse complementing. Reverse complementing and alignment of the sequences of both strands can be performed using sequencing software (see below). Currently, Sanger sequencing platforms yield highly accurate sequence data of up to about 1,000 nucleotides in length (Metzker 2005). To sequence larger amplicons, additional sequencing reactions are performed using primers spaced 500 to 800 nucleotides apart (the partial sequences must overlap to allow assembly into the full sequence of the amplicon).

There are many programs (freeware and non-freeware) available for analysis of sequence data, so it is really up to the user's personal preference as to which one to use. A simple and free general-purpose program that can be routinely used to manipulate, edit, align, and assemble sequences is BioEdit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). An extensive compilation of software programs for a variety of evolutionary and molecular studies can be found at Joe Felsenstein's website (<http://evolution.genetics.washington.edu/phylip/software.html>).

The sequences of the DNA molecules in the PCR product can differ due to biological reasons (i.e., nucleotide polymorphism). For example, polyploid

or heterozygous plants can contain different gene copies, and/or different alleles of a gene, which exhibit different nucleotide sequences. Sanger sequencing yields a single sequence resulting from the superimposition of the sequences of the different genes or alleles. This sequence thus displays multiple nucleotides at the variable nucleotide positions. If there is an interest in the precise sequence of individual genes or alleles, it is necessary to separate individual DNA molecules from the amplicon mixture by cloning, and then individually sequence a few randomly selected ones. In this case, each sequencing reaction will yield a sequence derived from a single DNA molecule.

**Tip:** Cloning involves a substantial amount of work. Before embarking into it, be sure it is really necessary for your purpose.

**Tip:** Any individual DNA molecule obtained in a PCR may contain one or more errors introduced by PCR. Thus, when sequencing cloned products, several clones should be individually sequenced. Sequencing three to five clones per gene or allele is usually sufficient, but keep in mind that you cannot identify which molecule derives from which gene or allele prior to sequencing.

**Tip:** A proofreading polymerase can be used for PCR when downstream applications call for cloning, to reduce sequencing errors.

Cloning of PCR products involves ligating DNA fragments into plasmid vectors carrying an antibiotic resistance gene, and then transferring these recombinant vectors into bacteria cells (generally *Escherichia coli* cells). Electroporation (Dower et al. 1988) is most effective for this purpose, but chemical or heat shock-based transfer procedures that do not require an electroporation device are currently implemented in commercial kits. Transformed cells are selectively identified from non-transformed cells on a solid agar medium containing the appropriate antibiotic. Individual colonies of

transformed cells are then grown in liquid medium, whereupon the recombinant plasmid vector is amplified. The plasmid is then purified from the cells and used for DNA sequencing or other downstream applications. Several kits are available for cloning PCR products, including the TOPO TA Cloning Kit (Life Technologies), the pGEM-T Easy Vector System (Promega, Madison, WI), or the pDrive system (Qiagen).

The approaches discussed above generally are used to obtain sequence information from one or a few genes. For large-scale gene sequencing projects, or to sequence one or a few amplicons in very large numbers of individual plants, using one of several next-generation sequencing (NGS) techniques is advisable. Current NGS technologies include Roche/454 GSFLX pyrosequencing (454 Life Sciences, Branford, CT), Illumina (San Diego, CA), and SOLiD (Life Technologies) (Thudi et al. 2012). NGS technology can generate enormous amounts of sequence data, but up-front costs of sample preparation, sequencing, and data processing are high. NGS technologies are thus of interest if sequencing a few genes or amplicons in a massive number of plants or of weed populations is required. Another use of NGS techniques is to gain access to the sequence of genes of interest in a weed species where no or few genomic data are available (Vigueira et al. 2013). This can be achieved by sequencing and assembling the transcriptome or a draft of the genome of the species of interest (e.g., Riggins et al. 2010). Sequencing and assembly can be carried out by a sequencing company or centralized campus facilities.

The recent rise of the NGS technologies has also opened the possibility to study weed adaptive traits with a complex genetic determinism involving both structural and regulatory mutations (Vigueira et al. 2013), such as NTSR (Délye 2013; Yuan et al. 2007). An approach of choice for this purpose is quantitative transcriptome sequencing, also known as RNA sequencing or RNAseq. It implies RNA samples are converted to cDNA. Millions of short sequences are then generated using NGS. Currently, the Illumina NGS technology is most adequate for quantitative transcriptome sequencing, because of the massive number of short reads generated that enable to maximize transcriptome coverage even for rare transcripts (Ness et al. 2011). Illumina short reads thus simultaneously contain transcriptome sequence information and a measure of gene expression (Ozsolak and Milos 2010). Yet, Illumina reads are short (currently 100 to 200 nucleotides),

which renders *de novo* transcriptome assembly tricky (Ness et al. 2011). The short sequence reads can be mapped against a reference genome sequence or assembled *de novo* to produce transcriptome resources, including the structure and abundance of each transcript (see Martin and Wang 2011 and Ward et al. 2012 for reviews). Guidelines for transcriptome-based analysis of NTSR in weeds have been proposed (Délye 2013). Yet, this approach is still in its infancy (Délye et al. 2013; Vigueira et al. 2013), and only a few attempts have been made to date (e.g., Gaines et al. 2013; Gardin et al. 2013).

## Detecting Mutations

Amplification of a gene of interest, followed by DNA sequencing, can be used to detect DNA polymorphisms that can confer herbicide resistance. However, when a particular DNA polymorphism (mutation) is known to cause resistance, the researcher might want to screen plants only for the presence of this mutation (mutation genotyping). For this purpose, methods are available that are simpler, faster, and cheaper than DNA sequencing.

**Easy, Robust Method: dCAPS, CAPS.** The most robust, in-lab PCR-based techniques that can be used to detect DNA mutations are based on the differential cleaving by restriction enzymes between amplicons carrying and not carrying the targeted mutation(s). Restriction enzymes recognise a specific DNA sequence (recognition site) and hydrolyze (“cut”) the DNA molecule at, or close to, their recognition sites.

The basic approach involves first performing a PCR to amplify a gene region of interest, incubating (digesting) the amplicon with the appropriate restriction enzyme, and analyzing the PCR products by gel electrophoresis. Observing two small molecular-weight DNA fragments in place of one larger molecular-weight fragment indicates that the fragment has been digested.

The Cleaved Amplified Polymorphic Sequence (CAPS) technique uses a restriction site that is naturally present in one type of amplicon (mutant or wild-type) and absent in the other due to the resistance-endowing mutation. Thus, CAPS can only be implemented if the specific mutation sought either creates or abolishes a restriction site. Because there are hundreds of commercially available restriction enzymes with unique recognition sites,

the CAPS approach might work in some cases. However, the mutation of interest most often will not create a restriction site polymorphism; therefore, the derived Cleaved Amplified Polymorphic Sequence (dCAPS) technique has been developed. dCAPS uses PCR to create a restriction enzyme recognition site in a sequence where none exists. The dCAPS primer is located close to the mutation of interest, and contains one or more mismatched nucleotides so as to create a restriction enzyme recognition site encompassing the mutation of interest (Neff et al. 1998). The (d)CAPS technique has been successfully used to detect mutations conferring herbicide resistance, including in the ACCase and ALS genes (e.g., Délye and Boucansaud 2008; Délye et al. 2011; Kaundun and Windass 2006).

Obviously, to use the dCAPS method, the nucleotide sequences of wild-type and mutant alleles should be known. Ideally, development and testing of the assay also will require DNA samples from both wild-type and mutant plants. Heterozygous plants should be available to verify the reliability of the assay.

**Tip:** If heterozygous plants are not available, one can mix equal amounts of DNA from wild-type and mutant plants to obtain a “mock” heterozygote.

*Primer Design for CAPS.* Primers must flank the mutation (and thus the restriction site) and yield an amplicon that can be easily discriminated from the resulting digestion products by gel electrophoresis if the enzyme recognition site is present. For example, you can design primers to amplify a 300 bp fragment with the restriction site near the middle.

**Tip:** Keep in mind that the same restriction site might be present elsewhere in the gene. If it is present more than once in your amplicon, it can confound interpretation of the digestion productions.

**Tip:** Design the primers so that they amplify a relatively small fragment (< 400 bp).

**Tip:** If the target gene has introns and genomic DNA is used as a PCR template (this is generally more convenient than using cDNA), design the primers to hybridize within a single exon or be sure to use the genomic DNA sequence when designing primers.

*Primer Design for dCAPS.* Here, primer pairs consist of one “dCAPS” primer that creates the restriction site in the amplicon, and of one “conventional” PCR primer. dCAPS primer sequences targeting the codon of interest can be generated using the free software dCAPS Finder 2.0 (<http://helix.wustl.edu/dcaps/dcaps.html>) (Neff et al. 2002). This software provides a range of dCAPS primer and restriction enzyme combinations to be used in dCAPS assays. dCAPS primers are then selected on the basis of: (1) restriction enzyme availability (and cost), and (2) absence of mismatch at the last 3' nucleotide position in the primer sequence, which introduces a risk for primer inefficiency.

**Tip:** Setting dCAPS Finder “mismatch” option to 1 might not yield satisfactory primers. The option can be increased to 3 if dCAPS primers containing no more than one mismatch in the last three to five 3' nucleotides are selected.

The following rules should be followed when designing a dCAPS assay:

1. Design assays so that only amplicons containing wild-type codons are cut (provides a control for restriction enzyme activity).
2. Use long dCAPS primers (ca. 40 nucleotides) to enable easy discrimination of undigested and digested (about 40 bp removed) amplicons by standard agarose gel electrophoresis. For this reason, design the “conventional” primer so that the amplicon generated is 150 to 350 nucleotides long.
3. Ensure that all nucleotides in the restriction enzyme recognition site are exclusively located in the dCAPS primer sequence and in the part of the targeted codon where any variation would cause amino acid substitution, so as to detect only nucleotide polymorphisms endowing resistance (to avoid false positive detection).

4. When several restriction enzymes can be used, select one with no other recognition site in the amplicon. Several freeware programs are available to identify restriction sites in DNA sequence (e.g., Webcutter, <http://rna.lundberg.gu.se/cutter2/>). Alternatively, position the non-dCAPS primer to exclude any other enzyme recognition site in the amplicon.

When the primers are designed, PCR optimization is performed as described above.

**Tip:** When dCAPS primers are designed following these rules, it is not necessary that mutant plants be available. For genes where resistance-endowing mutations are well known (e.g., ACCase, ALS, EPSPS, and psbA; Burgos et al. 2013), one can design assays to detect these mutations before resistance has evolved in a weed species.

*Restriction Digests.* The composition of PCR mixes is usually compatible with reaction mixes for restriction enzymes. Thus, no purification of the PCR product is necessary prior to digestion by restriction enzymes. Digestions are performed in 10 µl volumes using the following general steps:

1. Put 1 to 5 µl PCR mix in tubes. The amount of PCR mix used for digestion depends on the intensity of the amplicons observed.

**Tip:** dCAPS assays are based on the absence of digestion of amplicons carrying the mutation(s) targeted. An excess of amplicon must be avoided, because it can create false positive detection due to partial digestion of excessively abundant amplicons.

2. Add 5 units of restriction enzyme.
3. Add 0.5 µl of 10× enzyme reaction buffer (supplied with the enzyme).
4. Fit the total volume to 10 µl with distilled water.
5. Incubate digestion mixes 3 h at the temperature optimal for the activity of the restriction enzyme selected.

**Tip:** Duration of digestion can be cut down to 15 min using “fast digestion” mixes, if available for the enzyme used.

6. Visualize digestion patterns by electrophoresis (see above).

**Tip:** Run samples of undigested PCR product on the gel as a negative control for digestion (allows one to more easily discriminate digested and undigested fragments).

7. Interpret the results. See Figure 1 for an example of dCAPS patterns.

### Another Simple Method: Allele-Specific PCR.

Allele-specific PCR utilizes a pair of primers in which one primer (allele-specific primer) selectively binds to only one allele (Sommer et al. 1992). Allele specificity of the PCR is due to the presence of a nucleotide exactly matching one of the alleles but not the other at the 3' end of the allele-specific primer. At a specific annealing temperature, a 3' mismatch does not prime in a PCR (Sommer et al. 1992). The presence or absence of one given

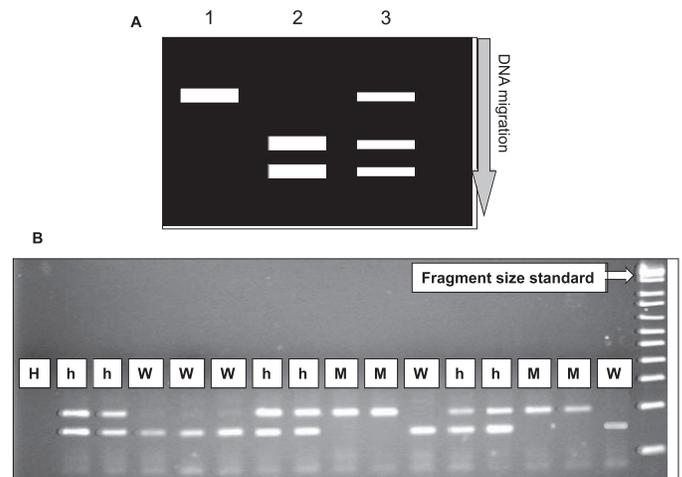


Figure 1. Band patterns that can be observed after agarose gel electrophoresis of the digestions of PCR mixes using a (d)CAPS assay digesting wild-type amplicons. Interpretation is as follows: (A) expected patterns for a theoretical dCAPS assay. Lane 1, undigested amplicon: homozygous mutant; Lane 2, fully digested amplicon: homozygous wild-type; Lane 3, mixture of digested (two bottom fragments) and undigested (top fragment) amplicons with similar intensity on the gel: heterozygous mutant. Note that, although not depicted here, a faint trace of undigested amplicon can be observed in homozygous wild-type samples if amplicon material is very abundant (restriction enzyme can be saturated by its substrate). (B) example of dCAPS patterns obtained with a dCAPS assay targeting ALS codon 197 in *Lolium* spp. Lanes h, heterozygous mutants; lanes W, homozygous wild-types; lanes M, homozygous mutants. Lane H, water control (no DNA added to the PCR mix).

nucleotide at the targeted position will thus result in the presence or absence of an amplicon, respectively.

Standard allele-specific PCR using two primers results in the detection of one allele by yielding a single amplicon when this allele is present, and no amplicon when it is absent (e.g., Wagner et al. 2002). The lack of an amplicon can also occur due to PCR failure (e.g., due to poor-quality DNA template) and, therefore, standard allele-specific PCR is prone to false negatives.

An improvement of the technique is bidirectional allele-specific PCR that uses four primers. Bidirectional allele-specific PCR enables concurrent detection of the presence or absence of two distinct alleles in a single PCR (e.g., Délye et al. 2002; Kaundun et al. 2011). Two primers are classical forward and reverse primers flanking the mutation site. The other two are allele-specific primers. In a given assay, one forward allele-specific primer targets one allele, and one reverse allele-specific primer targets the other allele. Using bidirectional allele-specific PCR, three amplicons sizes are obtained: one is specific for one allele (e.g., wild-type), the second is specific for the other allele (e.g., mutant), and the third one is amplified in all samples (positive internal control). As for (d)CAPS markers, sequences of the alleles of interest must be known for primer design. Also, to set up the assay, it is mandatory to have biological samples with homozygous wild-type and mutant genotypes, and heterozygous mutants.

The following rules should be followed when designing a bidirectional allele-specific PCR assay:

1. One allele-specific primer is designed on the coding DNA strand, the second, on the noncoding strand.
2. The sizes of the amplicons specific to each allele should easily be separated on agarose gels.
3. Allele-specific primers to be used in the same assay should have melting temperature as close as possible.
4. The last nucleotide in the sequence of each allele-specific PCR primers must be the nucleotide specific to the allele targeted.

**Tip:** The specificity of an allele-specific primer can be improved by creating an additional mismatch at the penultimate 3' nucleotide position in the primer. Kwok et al. (1994) and Pettersson et al. (2003) provide guidelines for designing allele-specific primers.

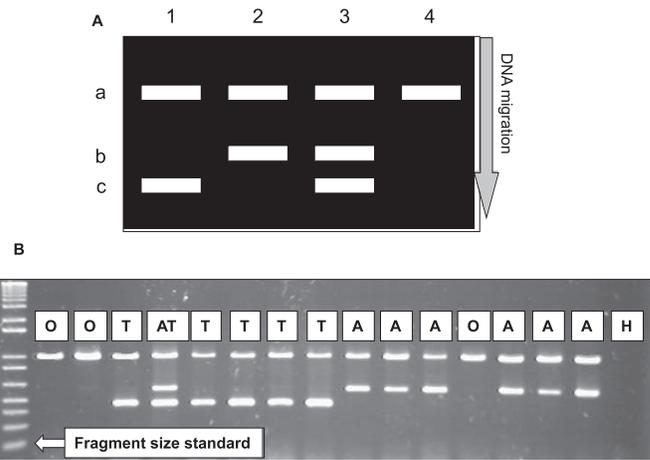


Figure 2. Bidirectional allele-specific PCR patterns that can be obtained after agarose gel electrophoresis for an assay designed to generate a smaller allele-specific amplicon from the mutant allele. (A) expected patterns for a theoretical bidirectional allele-specific PCR assay. Up to three amplicons are generated: “a,” positive internal control (should always be observed); “b,” wild-type allele; “c,” mutant allele. Interpretation is as follows: Lane 1, occurrence of the mutant allele; Lane 2, occurrence of the wild-type allele; Lane 3, heterozygous mutant. If only two different nucleotides are known to be present at the targeted position, lanes 1 and 2 can be interpreted as homozygous mutant and wild-type, respectively. If nucleotide(s) different from the two nucleotides targeted by the assay is (are) present on all gene copies at the targeted position, none of the allele-specific primers will hybridize, and pattern 4 will be obtained. In this case, sequencing fragment “a” will allow identification of the new allele. (B) example of patterns obtained with a bidirectional allele-specific PCR assay targeting the first nucleotide at ACCase codon 1781 in *Alopecurus myosuroides*. There are four possible nucleotide variants at this position: A (if the codon encodes an isoleucine residue; wild-type allele), C or T (if the codon encodes a leucine residue; mutant alleles), and G (if the codon encodes a valine residue; mutant allele). The allele-specific assay is designed to detect the A and T variants. Lanes T, occurrence of the T allele; lanes A, occurrence of the A allele; lane AT, heterozygous mutant; lane O, plant not carrying A or T alleles (homozygous mutant for the C allele). Lane H, water control (no DNA added to the PCR mix).

5. Test a range of annealing temperatures until one temperature is found that allows specific amplification.

Interpretation of bidirectional allele-specific PCR patterns is illustrated in Figure 2.

Bidirectional allele-specific PCR is sensitive to the annealing temperature used and to the composition of the PCR mix (especially to the pH). Because no digestion step is involved, bidirectional allele-specific PCR is faster than (d)CAPS, but can require a fair amount of troubleshooting to develop a robust assay. Also, because the diagnostic step occurs during the PCR, variations among template samples (DNA quantity, quality, and purity) can

decrease the robustness of the assay. In contrast, the diagnostic step of the (d)CAPS assay is the restriction enzyme digest, which is a much less finicky reaction than PCR. Bidirectional allele-specific PCR allows for detection of only two alleles per assay, but you know precisely which allele(s) is (are) detected. The (d)CAPS assay can be used to detect several mutations at the same nucleotide or adjacent nucleotides (on the basis of the digestion of the wild-type amplicons vs. absence of digestion of amplicons carrying any mutation disrupting the restriction enzyme recognition site). In cases where several mutations are possible at the targeted position, as observed for instance in ALS or ACCase (Beckie and Tardif 2012; Tranel and Wright 2002), (d)CAPS does not enable identification of the specific mutation disrupting the restriction enzyme recognition site. Both allele-specific PCR and (d)CAPS assays (as well as other assays not discussed here) can and have been used to detect mutations conferring herbicide resistance. Choosing between dCAPS and bidirectional allele-specific PCR therefore depends on the planned study. If several mutations are present at the nucleotide positions of interest, or if the robustness of a bidirectional allele-specific PCR becomes a concern, then (d)CAPS is the assay of choice.

**“High-Tech” Methods.** Other PCR-based techniques can be used to detect DNA mutations. They require expensive equipment and reagents, and are best suited for very high-throughput analyses (hundreds or thousands of samples). Three examples of high-tech methods are TaqMan, Multiplex SNaPshot (both from Life Technologies) and Scorpions (Sigma-Aldrich, St. Louis, MO).

TaqMan uses a pair of PCR primers flanking the mutation site, and a short DNA probe carrying a fluorophore (reporter dye) at the 5' end and a quencher at the 3' end. During the PCR reaction, the PCR primers and the TaqMan probe anneal simultaneously on the DNA sequence. During the PCR elongation step, exonuclease activity of the thermostable DNA polymerase degrades the probe, which releases the reporter dye from the quencher, thereby producing fluorescence. The TaqMan technology can combine two probes with two different fluorochromes, making it possible to detect two different alleles. TaqMan assays have been developed to genotype ALS (e.g., Warwick et al. 2008) or ACCase (e.g., Délye et al. 2010).

Scorpions uses a probe directly carrying a fluorophore that has a specific target sequence for

PCR. This system has the advantage over systems such as TaqMan in that no separate probe carrying a fluorophore is required to bind to the amplified target, making detection both faster and more efficient. Assays combining Scorpions, allele-specific PCR, and quantitative PCR technology have been developed to quantify mutant ACCase alleles in bulk samplings of *Lolium* spp. (Kaundun et al. 2006).

Multiplex SNaPshot can detect mutations at up to 10 sites in a single assay. It is a combination of PCR and single-base sequencing. The DNA regions carrying a mutation are amplified by PCR. DNA oligonucleotide probes anneal to a target sequence immediately adjacent to the variable nucleotide positions. A DNA polymerase then extends the probe by incorporating one dye-labelled nucleotide corresponding to the nucleotide present at the variable position. Thus, the size of the DNA fragment obtained for each mutation site targeted is the size of the probe plus one fluorescent base, for which the color depends on the nucleotide incorporated. The mixture of probes with incorporated dye-labelled nucleotides is then separated on a sequencer, and the nucleotide present at each position investigated is detected based on the size of the corresponding fragment and the wavelength of the incorporated dye-labelled nucleotide. SNaPshot assays have been developed to genotype ACCase (Alarcón-Reverte et al. 2013).

## Detecting Changes in Gene Expression

Quantitative PCR (qPCR) is a versatile analytical tool in which a fluorescent dye is used to monitor the amount of PCR product produced in real time during each PCR cycle. There are two basic categories of detection dye chemistries: nonspecific dyes that intercalate with any double-stranded DNA fragment, and target-specific dyes that utilize fluorescent probes and/or primers (Nolan et al. 2006). Nonspecific dyes, such as Sybr Green and EvaGreen, are typically less costly and simpler to use than target-specific dyes, but they do have one important drawback: any form of double-stranded DNA, including nontemplate fragments and primer-dimers, will be detected by nonspecific dyes. However, these unwanted products can readily be detected using a melting curve analysis and eliminated by assay optimization and/or redesigning the primers.

Herbicide resistance caused by an increase in the amount of a protein can be due to a change in the

regulation of the corresponding gene or to an increase in the number of genomic copies of this gene. Both origins of resistance can be detected by techniques based on qPCR. Although qPCR is predominately used for gene expression analyses with cDNA as the template (discussed in the next section), it is also compatible with genomic DNA. Thus, we first focus on the use of qPCR for detecting gene amplification, which was associated with glyphosate resistance initially in Palmer amaranth (*Amaranthus palmeri* S. Wats.) (Gaines et al. 2010).

### Testing for Gene Amplification Using qPCR.

The basic qPCR experiment for testing gene amplification involves calculating the relative quantity of a target gene based on an endogenous control or reference gene. Ideally, the reference gene should be single-copy in all individuals. Example reference genes that have been used in studies of glyphosate-resistant weeds include ALS (Gaines et al. 2010) and carbamoylphosphate synthetase (Tranel et al. 2011) for *Amaranthus* spp., and cinnamoyl-CoA reductase for Italian ryegrass [*Lolium perenne* L.ssp. *multiflorum* (Lam.) Husnot] (Salas et al. 2012).

**Tip:** Several candidate reference genes should be evaluated experimentally—rather than just picking one from the literature—to confirm suitability in a particular plant species. General guidelines for selecting a reference gene are provided in a subsequent section.

*Primer Design.* The same basic principles of primer design discussed for regular PCR also apply to qPCR, with only slight amendments. Both forward and reverse qPCR primers should be designed to anneal at 60 to 64 °C and within 2 °C of each other. The amplicon should also be 70 to 300 bp in length, which is optimum for the shortened cycling conditions of a typical qPCR experiment. Another consideration when designing qPCR primers for a gene amplification analysis is the region of the gene in which the primers are anchored. Ideally, the primers should target a region of the gene showing little nucleotide polymorphism to maximize assay sensitivity and reliability. Exon–intron junctions, introns, and 5′ untranslated regions tend to be more variable across genes and species, so these regions should be avoided for a qPCR assays. In any case,

several different primer sets should be evaluated experimentally before large-scale screening. Integrated DNA Technologies ([www.idtdna.com](http://www.idtdna.com)) offers a free online qPCR primer design tool along with a downloadable qPCR Application Guide with additional background information and useful tips on primer design and troubleshooting.

A qPCR experiment for gene amplification should include positive and negative controls (samples with and without gene amplification) and no-template controls. Three technical replicates (i.e., replicates of the same biological sample) should be run and the resultant quantification cycle (C<sub>q</sub>) values averaged. Plate-to-plate variation can be assessed by including interplate control samples (i.e., including one or more common samples on each plate). Individuals from multiple populations could also be included to evaluate background gene levels in the species.

**Tip:** The C<sub>q</sub> value refers to the PCR cycle during which fluorescence increases above some threshold level. It also is referred to as a C<sub>t</sub> (threshold cycle) value.

*Absolute vs. Relative (Comparative C<sub>q</sub> Method) Quantitation.* Before beginning the qPCR experiment, one must decide how the data are to be analyzed. The results of your assay can be analyzed either by absolute or relative quantitation methods. Absolute quantitation involves determining quantities of unknown samples from a standard curve. This method can be useful if one wants to know the exact copy number of a particular gene of interest in a particular individual. Although this method results in a high level of accuracy, a standard curve of known quantities must be performed for each experimental run, which requires more reagents and takes up space on the plate. Alternatively, relative quantitation using the Comparative C<sub>q</sub> Method (Pfaffl 2001) does not require standard curves for each plate as long as the PCR efficiencies of the target and control genes are essentially equivalent. To use this method, a validation experiment (Livak and Schmittgen 2001) is first performed by establishing standard curves for the target and endogenous control genes from the same biological sample. The standard curves are done using five to six serial dilutions of genomic DNA (gDNA) for each gene with three replicates for each dilution point. Resultant C<sub>q</sub> values are averaged for each input amount and  $\Delta C_q$ s calculated (C<sub>q</sub> target –

Cq control). The log input of the gDNA is then plotted against the  $\Delta Cq$ s. To pass the test, the absolute value of the slope must be  $< 0.1$ .

*qPCR Protocol for Detection of Gene Amplification.*

This protocol was originally designed to detect EPSPS gene amplification in *Amaranthus* spp. with the 7900-HT detection system (Life Technologies), which uses 384-well plates. It can easily be adapted for 96-well plates and other detectors, and for other genes/species.

1. Dilute all genomic DNA samples to  $10 \text{ ng } \mu\text{l}^{-1}$ . Include positive and negative control samples (resistant and sensitive plants) and no-template controls (NTC) for each gene.
2. Prepare a dilution series to evaluate primer efficiency and determine the dynamic range of the assay. A 5-fold serial dilution, such as  $1\times$ ,  $0.2\times$ ,  $0.04\times$ ,  $0.008\times$ ,  $0.0016\times$ , is sufficient and should be prepared for the target and control genes using one sample template. The proper range must be experimentally determined for each assay.

**Tip:** The starting DNA concentration is not critical, as long as the Cq values of the samples are obtained within the dilution series range. For example, most of our DNA yields from CTAB average  $200$  to  $400 \text{ ng } \mu\text{l}^{-1}$ , so our dilution series starts with this as the upper range (basically undiluted stock) and is serially diluted five times ( $200 \text{ ng } \mu\text{l}^{-1}$  ( $1\times$ ),  $40 \text{ ng}$ ,  $8 \text{ ng}$ ,  $1.6 \text{ ng}$ ,  $0.32 \text{ ng}$ ) to produce a standard curve.

**Tip:** PCR efficiency is satisfactory if the slopes of the dilution curves are close to  $-3.3$  ( $\pm 10\%$ ). Slopes higher or lower than this range indicate the need for assay optimization. Look for an  $R^2$  value  $> 98\%$ . Percentage efficiency (E) is calculated by  $E = [10^{-1/\text{slope}} - 1] * 100$  (Radstrom et al. 2003).

3. Prepare a master mix for each gene. Account for all samples and replicates in calculating the master mix plus some extra (10%) for pipetting variation. Our calculations are based on a  $10 \mu\text{l}$  final volume ( $9 \mu\text{l}$  mix +  $1 \mu\text{l}$  template) for a 384-well plate. For example, a master mix for one gene for 10 samples would be calculated as shown in Table 3.

Table 3. Preparation of master mix for each gene in a quantitative polymerase chain reaction (qPCR) assay. Volumes are based on a  $10\text{-}\mu\text{l}$  reaction using  $1 \mu\text{l}$  of template DNA.

| Reagent                                    | Amount for three          |   |
|--|---------------------------|---|
|  | Amount for one reaction   | triplicates each of 10 samples <sup>a</sup> |
|  | ----- $\mu\text{l}$ ----- |   |
| EvaGreen                                   | 5.0                       | 165.0                                       |
| forward primer (0.2 to 0.4 $\mu\text{M}$ ) | 0.3                       | 9.9   |
| reverse primer (0.2 to 0.4 $\mu\text{M}$ ) | 0.3                       | 9.9   |
| water                                      | 3.4                       | 112.2                                       |
| Total                                      | 9.0                       | 297.0                                       |

<sup>a</sup> Includes 10% extra to allow for pipetting error.

**Tip:** Thaw all samples and reagents on ice. EvaGreen is light-sensitive, so handle accordingly. Master mix reagents should be mixed thoroughly but not vortexed. Accurate pipetting is crucial with such small volumes (an electronic pipette is helpful, but not necessary).

4. Pipette  $9 \mu\text{l}$  mix in each well, then add  $1 \mu\text{l}$  DNA individually.
5. Gently tap the plate on the benchtop to settle the mix in the bottom of the wells and overlay with ABI Prism optical adhesive cover (#4311971). Spin the plate in a centrifuge for 2 min at room temperature before loading into the 7900 HT.

*Data Analysis.* Following the run, the data are stored in a Sequence Detection System (SDS) file (format for the 7900-HT). First, open the file and verify that the baseline and threshold were generated correctly. These values will likely be set automatically by the software (default settings) and no adjustments might be necessary. Otherwise, the threshold can be adjusted manually and should be within the geometric phase of the PCR amplification curve and above the baseline. Second, visually inspect each well and remove or exclude those with poor amplification. Replicates of the samples should have similar Cq values, otherwise remove outliers before analysis. The dissociation curves (for SYBR Green assays only) will help identify wells that should be removed, including empty wells and wells with nonspecific amplification products. Primer-dimer peaks are common in the no-template controls, but should disappear in the standard curve and unknown samples. If you ran a standard curve, make sure the slope is around  $-3.3$  (indicating that your assay is 100% efficient) and the correlation coefficient ( $R^2$ ) is  $> 0.98$ . Variation in the data can be

caused by a variety of factors, but the more common sources are imprecise pipetting, incomplete mixing of reagents, poor quality of DNA/RNA template, improper threshold or baseline setting, or improper handling of the qPCR plate. Low precision also can be caused by poorly designed primers or inherent biological properties of the genes under investigation. After quality checking the results, the data can then be exported to one of a variety of qPCR programs available online or to Microsoft Excel for analysis.

Assuming good standard curves and a validation experiment that demonstrated equal amplification efficiencies of the target and reference genes, the Comparative C<sub>q</sub> Method can then be used to test for gene amplification. The following steps can easily be done in Excel:

1. Calculate the mean and standard deviation values of the replicates for both genes of each biological sample. Sample C<sub>q</sub> values should fall within the limits of the standard curve and standard deviations should be  $\leq 0.3$  in order to distinguish between a one- to twofold difference in copy number (Bubner and Baldwin 2004).
2. Calculate the  $\Delta C_q$  value for each sample by subtracting the mean C<sub>q</sub> of the reference gene from the mean C<sub>q</sub> of the target gene.
3. Calculate the standard deviation of the  $\Delta C_q$  value. This is calculated from the standard deviations (*s*) of the target and reference gene values by the following formula:  $s = (s_1^2 + s_2^2)^{1/2}$ .
4. Calculate the  $\Delta\Delta C_q$  value by subtracting the  $\Delta C_q$  of the reference sample (plant without gene amplification) from the  $\Delta C_q$  of the unknown samples (plants in which gene amplification is suspected).
5. Calculate the fold-difference and range by incorporating the  $\Delta C_q$  standard deviations using the formula:  $2^{-\Delta\Delta C_q}$ , with  $\Delta\Delta C_q \pm$  the standard deviation of the  $\Delta C_q$  value.

**Tip:** The fold-difference for the reference individual and any other individual lacking gene amplification should be approximately equal to 1, assuming the reference gene is single copy and stable.

**Measuring Gene Expression Using Reverse-Transcription-qPCR (RT-qPCR).** RT-qPCR is currently the technique of choice for gene expression quantification. However, a range of factors can

influence the reliability of direct gene expression measurement (e.g., experimental error due to variations in RNA or cDNA concentration among samples, operator error, and variation in the efficacy of the reverse-transcription reaction). Thus, expression data is generally analyzed as relative expression ratios using a normalization strategy: raw gene expression data is normalized using expression data of one (or preferably several) reference gene(s).

*Reference Genes.* A reference gene must have a constitutive and constantly stable expression in all experimental conditions studied. For herbicide-resistance studies, reference gene sets should consist of genes for which expression has been proven stable before and after herbicide application, and among resistant and sensitive plants. To date, reference genes with a stable expression under herbicide action have only been validated for two grass weeds, blackgrass (*Alopecurus myosuroides* Huds.) (Petit et al. 2012) and *Lolium* spp. (Duhoux and Délye 2013).

**Tip:** There are no “universal” reference genes identified to date. Reference genes suitable for one herbicide and one weed species might not be suitable for a different herbicide/species pair. It is a prerequisite to any study aiming at comparing gene expression among samples that the stability of the reference genes used has been adequately validated in the experimental system studied.

Three validated reference genes are generally considered to be adequate for accurate normalization (Vandesompele et al. 2002). To obtain three validated reference genes, it is advisable to test the stability of at least six candidate reference genes (Bustin et al. 2009, 2010). Candidate genes should be involved in different metabolic pathways, so that they are under independent regulations of expression. This reduces the risk for a similar effect of herbicide application on all candidate reference genes. A list of potentially useful candidate genes (that is, genes that have frequently been used as validated reference genes in the literature) is given in Table 4. Another possibility to identify genes with a stable expression that can be used as candidate reference genes is to use data from transcriptome-wide studies, when available (e.g., Czechowski et al. 2005: the genes identified in *Arabidopsis thaliana* (L.) Heynh. in this work are candidate reference genes of interest for studies addressing species in the family Brassicaceae).

Table 4. Examples of candidate reference genes for use in reverse-transcription quantitative polymerase chain reaction (RT-qPCR).

| Metabolic pathway (Gene ontology) | Candidate reference gene                                |
|-----------------------------------|---|
| Microtubule-based process         | Beta-tubulin <sup>a</sup>                               |
| Oxidation-reduction process       | Glyceraldehyde-3-phosphate dehydrogenase <sup>a,b</sup> |
| Translational initiation          | Cap binding protein <sup>b</sup>                        |
| Protein catabolic process         | Ubiquitin <sup>a,b</sup>                                |
| Translational elongation          | Eukaryotic elongation factor 1 alpha                    |
| Translational elongation          | Eukaryotic elongation factor 4 alpha                    |
| Transmembrane transport           | Sucrose proton symporter                                |
| Carbon fixation                   | Ribulose 1-5-bisphosphate                               |
| Microtubule-based process         | Actin   |
| Protein folding                   | Cyclophilin   |
| Protein catabolic process         | Ubiquitin conjugating enzyme E2                         |
| DNA binding                       | Histone 3   |
| Vacuolar fusion                   | SAND protein  |

<sup>a</sup> Reference genes validated in the grass weed *Alopecurus myosuroides* with herbicides inhibiting ACCase (Petit et al. 2012).

<sup>b</sup> Reference genes validated in the grass weed *Lolium* sp. with herbicides inhibiting ALS (Duhoux and Délye 2013).

**Tip:** It is advisable not to use ribosomal genes as reference genes, because of their high expression level compared to most other genes.

*Experimental Design for the Validation of Reference Genes.* The stability of reference gene expression must be checked across all experimental conditions studied. cDNA samples used to test candidate reference genes must therefore cover the full range of modalities intended in the study: (1) plant material (e.g., resistant and sensitive phenotypes, different geographical origins); (2) herbicide application modalities (e.g., herbicide dose, time before and after herbicide application); and (3) plant tissue (select the tissues where herbicides are applied where herbicide target is most abundant or active; this generally means the youngest, most actively growing plant parts, such as leaf and meristems for foliar herbicides, and root tips for soil-applied herbicides).

**Tip:** A crucial point here is to collect plant material rapidly and to neutralize plant metabolism without delay to avoid modifications of the gene expression patterns induced by response to wounding. The most widespread and effective method for this purpose is snap freezing of plant material in liquid nitrogen, followed by storage at  $-80$  C. Frozen plant material must not be allowed to thaw, and must be processed as quickly as possible to avoid RNA degradation (see RNA extraction above).

*Obtaining Expression Data for Candidate Reference Genes.* This process is as follows:

1. Choose a set of candidate reference genes.

**Tip:** At least six candidate genes should be tested for stability of expression using RT-qPCR. It is thus wise to start the procedure with more than six potential candidate genes, because this means that at least six candidate genes must pass steps 2 to 13 below.

- Design primer pairs for each candidate reference gene (see section “primer design” above). The expected amplicon should be 70 to 300 bp long for an optimal efficiency of PCR.
- Test the primer pairs in classical PCR followed by agarose gel electrophoresis. This step can be conducted on genomic DNA or cDNA. If conducted on genomic DNA, consider the possibility for introns to be present in the targeted region. A single, clear amplicon must be obtained.
- Remove the candidate reference genes not passing this step from your list. Alternatively, design new primers for the reference gene and go back one step.

**Tip:** Designing primers flanking intron-containing regions can be useful to check genomic DNA contamination in cDNA samples.

5. Check the sequence of the amplicon.
6. Remove the candidate reference genes not passing this step from your list.
7. Extract RNA from your samples. Check RNA quality and perform cDNA synthesis from a similar amount of starting RNA for all samples.

**Tip:** It is crucial to use several biological replicates to adequately test the stability of reference genes, taking into account biological variation (e.g., different individual plants used in a same experimental modality). Technical replicates should also be used for each biological replicate to account for experimental variation (e.g., independent reverse-transcription reactions).

8. Check efficiency and specificity of the primer pairs in qPCR using dilution series as described in the previous section.
9. Remove the candidate reference genes that do not show a single-peak melting curve (see Figure 3).
10. Remove the candidate reference genes with PCR efficiency values outside of the 90% to 110% range from your list.

**Tip:** If the efficiency value is not satisfactory for a given gene, designing new primers is generally easier, faster, and cheaper than trying to optimize qPCR.

11. Chose one working dilution for all your cDNA samples.

**Tip:** If the dilution series used enabled satisfactory assessment of qPCR efficiency, the midpoint dilution of the series can generally be used as the working dilution for all cDNA samples. This requires that a similar amount of RNA is used for all samples when performing the reverse-transcription reaction.

12. Perform qPCR for every remaining candidate gene on the diluted cDNA samples. Include the dilution series in each qPCR run as a control for the efficiency of the reaction.

13. Extract Cq values for each cDNA sample and for each gene after positioning the threshold.

**Tip:** The threshold for fluorescence detection can be set automatically or manually. When setting the threshold manually, use the logarithmic amplification plot to position the threshold so that it is above the background fluorescence, below the linear region, and at the beginning of the region of exponential amplification.

*Selecting a Set of Validated Reference Genes for Expression Data Analysis.* Three software packages are commonly used to estimate gene stability and identify the more stable genes. They are available at: <http://www.gene-quantification.de/>. A brief description is provided below for the three software packages:

**Tip:** The three software packages are each based on a specific algorithm. They provide complementary results, and should be used together to identify the most stable genes to be used as reference genes.

1. BestKeeper (Pfaffl et al. 2004):

Input:

- qPCR efficiency values for each candidate reference gene.
- Cq values for each gene and each sample.

Output:

- Samples that should be removed from the analysis due to experimental errors for good stability analysis.
- Variation in Cq values and its standard deviation (SD) for each gene. Genes with  $SD < 1.00$  and P value for the Pearson correlation coefficient below 0.01 are considered stable. Most stable genes are genes with lowest standard deviation.

**Tip:** it is possible to analyse the stability of the genes within a subgroup of sample by only entering the Cq values of samples in this subgroup.

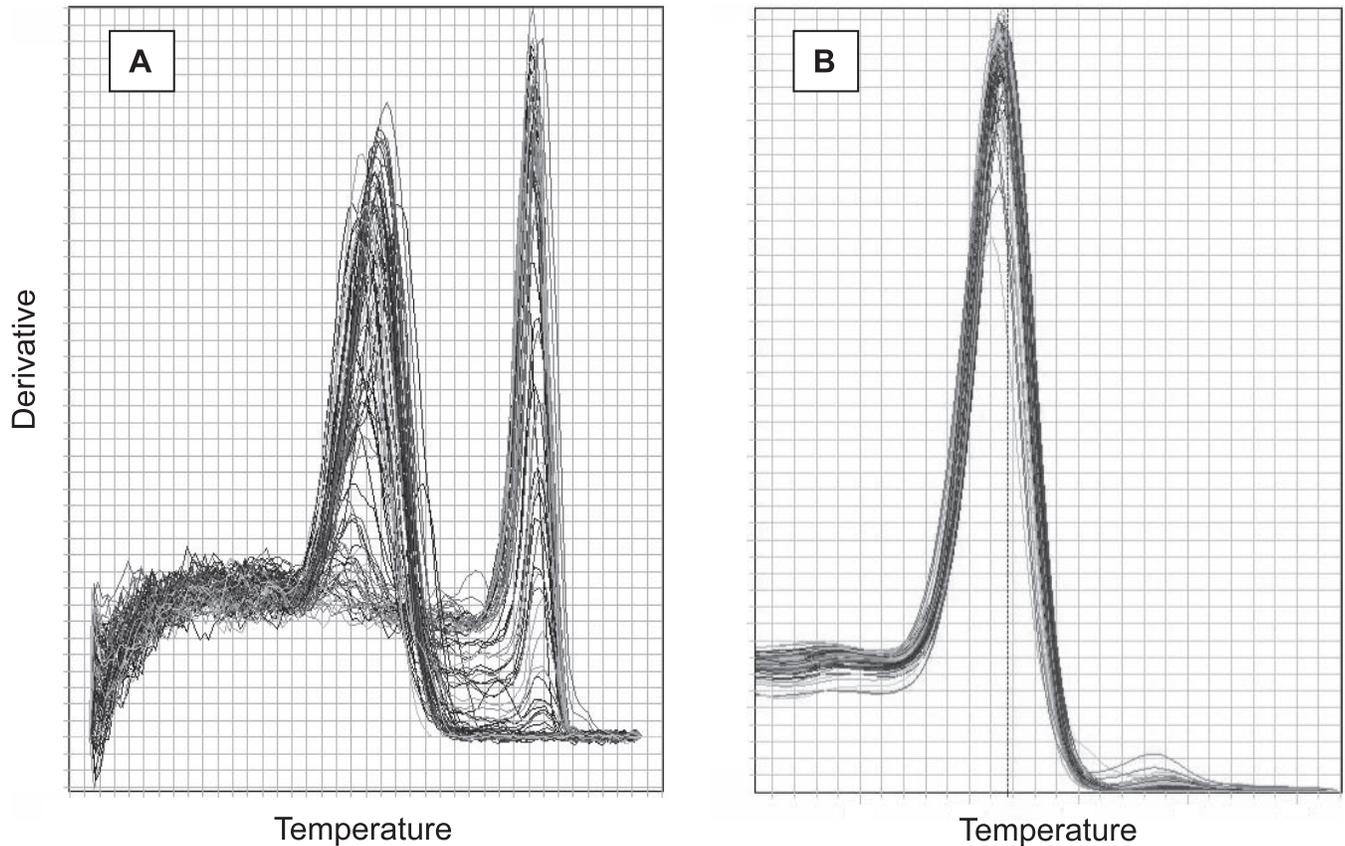


Figure 3. Primer specificity test for RT-qPCR. (A) melting curves generated for one primer pair that does not yield a single-peak melting curve because of primer dimer formation or nonspecific amplification. This primer pair is thus not suitable for RT-qPCR. (B) melting curves generated for one primer pair that yield a single-peak melting curve. This primer pair can be proceeded to the next step (step # 10) of the candidate reference gene validation process.

## 2. NormFinder (Andersen et al. 2004):

Input:

- Cq data transformed using the  $2^{-\Delta\Delta Cq}$  method (Livak and Schmittgen 2001). This method uses a calibrator defined as the lowest Cq value obtained for each gene (i.e., the Cq value from the sample with the highest expression level for the gene considered). Input data for the sample  $i$  and the gene  $j$  will thus be:  $I_{ij} = 2^{-(C_{qij} - C_{qminj})}$ .

Output:

- Stability value (SV) for each gene. The most stable genes are the genes with the lowest SVs.

**Tip:** Stability should be assessed without considering subgroups in a first step (overall stability). NormFinder allows assigning subgroups to each sample in the sampling (e.g., sensitive or resistant, treated or untreated). In a second step, stability should be analysed within subgroups. Subgroups should include a minimum of eight samples.

## 3. geNorm (Vandesompele et al. 2002):

Input:

- qPCR efficiency for each candidate reference gene.
- Cq data transformed using the  $2^{-\Delta\Delta Cq}$  method.

Output:

- Expression stability value (M) for each gene. Genes with  $M < 1.5$  are considered stable.
- Ranking of the genes from the two most stable down to the least stable.
- Optimal number of reference genes for adequate normalization.

**Tip:** The 0.15 cutoff threshold for the pairwise variation between consecutive normalization factors that is proposed to identify the number of reference genes most adequate for normalization is not to be considered as an absolute threshold value. If pairwise variation values are always above 0.15, using the number of reference

genes yielding the lowest normalization factor value is also appropriate.

**Tip:** It is possible to assess the stability of the genes within a subgroup of samples by only entering the transformed Cq values of samples in the subgroup.

*How Many Reference Genes?* Never use a single reference gene for normalization. On the other hand, keep in mind that it will be necessary to measure the expression of all reference genes in all samples analysed to normalize the expression of the target genes investigated. Thus, the more reference genes used, the more labor-intensive the experiments. In most cases, using the three most stable genes allows adequate normalization of gene expression (provided all genes have been found stable using the three software packages) (Vandesompele et al. 2002).

**Tip:** Remember it is preferable to use genes involved in different metabolic pathways.

*Quantification and Comparison of Relative Expression Levels of Genes of Interest. Primer design for genes of interest.* Follow steps 2 to 10 in section *Obtaining Expression Data for Candidate Reference Genes* using the sequence for your gene(s) of interest.

Comparison of the expression data generated for

**Tip:** qPCR efficiency values can be largely above or below 100% when working with genes of interest showing important variation in their expression among samples. This is acceptable, provided that the associated correlation coefficient that indicates the linearity of the expression is  $> 98\%$ . Caution: this tip does not apply to reference genes.

genes of interest among samples is performed on the basis of the respective relative expression levels. Relative expression levels are computed based on the PCR efficiency of the target gene and of the reference genes, and on the Cp deviation of the target and the reference genes in the analysed

sample compared to a reference sample (Pfaffl 2001). Relative expression level quantification software implement this approach and allow the user to compare the expression of a target gene among different samples. A “reference sample” or group of samples must be defined beforehand.

A useful software for this purpose is REST-mcs (Relative Expression Software Tool—multiple condition solver, Pfaffl et al. 2002) that is available at <http://www.gene-quantification.de/>.

Input:

- qPCR efficiency for each target gene and reference genes.
- Cq values for each sample and each gene (reference and target genes).

Output:

- Relative expression ratio in the “sample group” using the “reference group” as a baseline.
- Significance of the difference observed among samples (P value  $< 0.05$ ).

Another option for data analysis is the qPCR library developed for the R software (<http://www.dr-spiess.de/qpcR.html>)

**Tip:** In experiments involving response to herbicides, obvious reference samples would be cDNA from untreated, herbicide-sensitive plants.

**Tip:** In some samples, no expression of a target gene can be detected, and thus no Cq value can be generated, but expression of the reference genes is as expected (e.g., in untreated sensitive plants when studying a gene potentially involved in resistance). In such cases, a possibility is to use more concentrated dilutions of these samples in qPCR so as to be able to generate a Cq value by starting with more copies of the target gene. It might be that, even so, expression of the target gene still is not detected. This indicates extremely low, or even absent expression of the target gene in the samples in question. For the purpose of gene expression comparison, a possibility is to arbitrarily attribute a Cq value equal to the maximum number of cycles in the qPCR for the target gene in the samples in question.

## Dealing with Polyploid Species

The approaches discussed above apply straightforwardly to diploid species, i.e., species with a single genome that contains single-copy genes of interest. Yet, there are quite a few major weed species that are polyploid, i.e., they contain several genomes [e.g., *Avena* spp., *Chenopodium* spp., *Echinochloa* spp., ricefield bulrush [*Schoenoplectus mucronatus* (L.) Palla]]. The ploidy level can vary among species within a given genus, or even among plants in a given species. For instance, common lambsquarters (*Chenopodium album* L.) plants can be diploid (one genome), tetraploid (two genomes), or hexaploid (three genomes) (Bhargava et al. 2006). Studying polyploid species means facing two types of additional challenges.

First, because polyploid plants have several genomes, they are expected to contain several copies of a gene of interest. The genomic organisation of such species is not simple (Scarabel et al. 2010), and the number of copies of a given gene is not necessarily the same in the different genomes. Thus, a species with X number of genomes might contain more than X number of copies of a given gene; for example, a recent study conducted on the tetraploid species rice barnyardgrass [*Echinochloa phyllopogon* (Stapf) Koso-Pol.] identified two ALS copies, as expected, but also four chloroplastic ACCase copies (Iwakami et al. 2012). As every copy of a gene encoding a protein involved in herbicide resistance might carry mutations endowing herbicide resistance (e.g., Yu et al. 2013), it is important to obtain and sequence all copies of this gene present in a given species. Specific techniques such as Southern hybridization (Southern 1975 [not reviewed in this paper]) can be used to assess the number of copies of the gene of interest in a polyploid species.

The second additional challenge resides in the possibility that some of the gene copies present across the different genomes of polyploid species are pseudogenes or silenced genes (e.g., Huang et al. 2002). Thus, it is important to check that all genes analysed are readily expressed. It is therefore essential to know the ploidy level of a species of interest beforehand, and to be aware that studying polyploid weeds presents additional challenges.

### Mistakes to Avoid after a Genetic Change is Identified

It must be kept in mind that having identified a genetic change (either a structural mutation or gene

expression differences reflecting occurrence of regulatory mutation[s]) between resistant and sensitive plants does not necessarily mean that your work is done. Ultimately, you will need evidence that the identified genetic change is actually responsible for the observed resistance phenotype. Such evidence can be obtained using genetic and biochemical approaches. For example, one could clone genes that are identical except for the implicated mutation, produce the encoded enzymes using an *E. coli* expression system, and then compare herbicide sensitivities of the enzymes using an in vitro biochemical assay (Dayan et al. 2014). Alternatively, one could compare herbicide sensitivities of whole organisms (*E. coli*, yeast, or plant) bearing transgenes that differ only by the implicated mutation. If multiple mutations are found within a candidate resistance gene, systematic experimentation is required to determine which specific mutation, or combination of mutations, confers resistance. If an identified mutation is identical to one that has been previously demonstrated to confer resistance in a different species, “guilt by homology” provides reasonable—but not complete—certainty of the mutation’s involvement in resistance. Nevertheless, this “guilt by homology” says nothing about the potential involvement of an additional resistance mechanism within your particular population. Genetic analysis (Mallory-Smith et al. 2014) can be used to determine if multiple resistance mechanisms are present within a population, and to test for cosegregation of the candidate genetic difference. Cosegregation analysis also can be a useful approach to obtain evidence that a change in gene expression confers resistance.

As a final tip with regard to making inferences based on identified genetic changes, one needs to keep in mind that weed populations typically exist as collections of genetically diverse individuals and that resistance can be endowed by a range of genes (reviewed in Délye et al. 2013). Identification of a particular genetic change conferring resistance in one or a few individuals of a population certainly is no guarantee that other genetic changes conferring resistance are not also present within the population from which the plants were obtained, or even within the plants in which the particular genetic change was identified.

### Literature Cited

Alarcón-Reverte R, Hanley S, Kaundun SS, Karp A, Moss SR (2013) A SNaPshot assay for the rapid and simple detection of

- known point mutations conferring resistance to ACCase-inhibiting herbicides in *Lolium* spp. *Weed Res* 53:12–20
- Andersen L, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–5250
- Beckie HJ, Tardif FJ (2012) Herbicide cross-resistance in weeds. *Crop Prot* 35:15–28
- Bhargava A, Shukla S, Ohri D (2006) Karyotypic studies on some cultivated and wild species of *Chenopodium* (Chenopodiaceae). *Genet Resour Crop Evol* 53:1309–1320
- Bubner B, Baldwin IT (2004) Use of real-time PCR for determining copy number and zygosity in transgenic plants. *Plant Cell Rep* 23:263–271
- Burgos NR, Tranel PJ, Streibig JC, Davis VM, Shaner D, Norsworthy JK, Ritz C (2013) Review: confirmation of resistance to herbicides and evaluation of resistance levels. *Weed Sci* 61:4–20
- Bustin SA, Beaulieu JF, Huggett J, Jaggi R, Kibenge FSB, Olsvik PA, Penning LC, Toeggel S (2010) MIQE precis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol* 11:74
- Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55:611–622
- Chaves AL, Vergara CE, Mayer JE (1995) Dichloromethane as an economic alternative to chloroform in the extraction of DNA from plant tissues. *Plant Mol Biol Rep* 13:18–25
- Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis thaliana*. *Plant Physiol* 139:5–17
- Dayan FE, Owens DK, Corniani N, Silva FML, Watson SB, Howell J, Shaner DL (2015) Biochemical markers and enzyme assays for herbicide mode of action and resistance studies. *Weed Sci* 63:23–63
- Délye C (2005) Weed resistance to acetyl coenzyme A carboxylase inhibitors: an update. *Weed Sci* 53:728–746
- Délye C (2013) Unravelling the genetic bases of non-target-site-based resistance (NTSR) to herbicides: a major challenge for weed science in the forthcoming decade. *Pest Manag Sci* 69:176–187
- Délye C, Boucansaud K (2008) A molecular assay for the proactive detection of target site-based resistance to herbicides inhibiting acetolactate synthase in *Alopecurus myosuroides*. *Weed Res* 48:97–101
- Délye C, Jasieniuk M, Le Corre V (2013) Deciphering the evolution of herbicide resistance in weeds. *Trends Genet* 29:649–658
- Délye C, Matějček A, Gasquez J (2002) PCR-based detection of resistance to acetyl-CoA carboxylase-inhibiting herbicides in black-grass (*Alopecurus myosuroides* Huds.) and ryegrass (*Lolium rigidum* Gaud.). *Pest Manag Sci* 58:474–478
- Délye C, Michel S, Bérard A, Chauvel B, Brunel D, Guillemain JP, Dessaint F, Le Corre V (2010) Geographical variation in resistance to acetyl-coenzyme A carboxylase-inhibiting herbicides across the range of the arable weed *Alopecurus myosuroides* (blackgrass). *New Phytol* 186:1005–1017
- Délye C, Pernin F, Michel S (2011) ‘Universal’ PCR assays detecting mutations in acetyl-coenzyme A carboxylase or acetolactate-synthase that endow herbicide resistance in grass weeds. *Weed Res* 51:353–362
- Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* 16:6127–6145
- Doyle JJ, Doyle JL (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus* 12:13–15
- Duhoux A, Délye C (2013) Reference genes to study herbicide stress response in *Lolium* sp.: up-regulation of P450 genes in plants resistant to acetolactate-synthase inhibitors. *PLOS ONE* 8:10.1371/journal.pone.0063576
- Gaines T, Lorentz L, Maiwald F, Ott MC, Perret P, Yu Q, Han H, Busi R, Streck H, Beffa R, Powles S (2013) Identifying differential expression in non-target-site herbicide resistance genes in *Lolium rigidum* using high-throughput sequencing. Page 69 in *Global Herbicide Resistance Challenge*. Perth, Australia: Australian Herbicide Resistance Initiative [Abstract]
- Gaines TA, Zhang W, Wang D, Bukun B, Chisholm ST, Shaner DL, Nissen SJ, Patzoldt WL, Tranel PJ, Culpepper AS, Grey TL, Webster TM, Vencill WK, Sammons RD, Jiang J, Preston C, Leach JE, Westra P (2010) Gene amplification confers glyphosate resistance in *Amaranthus palmeri*. *Proc Natl Acad Sci U S A* 107:1029–1034
- Gardin J, Beffa R, Gouzy J, Carrère S, Délye C (2013) A transcriptomics-based approach enables the first identification of candidate genes involved in non-target-site-based resistance to herbicides in a grass weed (*Alopecurus myosuroides*). Page 68 in *Global Herbicide Resistance Challenge*. Perth, Australia: Australian Herbicide Resistance Initiative [Abstract]
- Huang S, Sirikhachornkit A, Faris JD, Su X, Gill BS, Haselkorn R, Gornicki P (2002) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. *Plant Mol Biol* 48:805–820
- Iwakami S, Uchino A, Watanabe H, Yamasue Y, Inamura T (2012) Isolation and expression of genes for acetolactate synthase and acetyl-CoA carboxylase in *Echinochloa phyllopogon*, a polyploid weed species. *Pest Manag Sci* 68:1098–1106
- Jasieniuk M, Brûlé-Babel AL, Morrison IN (1996) The evolution and genetics of herbicide resistance in weeds. *Weed Sci* 44:176–193
- Kaundun SS, Cleere SM, Stanger CP, Burbidge JM, Windass JD (2006) Real-time quantitative PCR assays for quantification of L1781 ACCase inhibitor resistance allele in leaf and seed pools of *Lolium* populations. *Pest Manag Sci* 62:1082–1091
- Kaundun SS, Dale RP, Zelaya IA, Dinelli G, Marotti I, McIndoe E, Cairns A (2011) A novel P106L mutation in EPSPS and an unknown mechanism(s) act additively to confer resistance to glyphosate in a South African *Lolium rigidum* population. *J Agric Food Chem* 59:3227–3233
- Kaundun SS, Windass JD (2006) Derived cleaved amplified polymorphic sequence, a simple method to detect a key point mutation conferring acetyl CoA carboxylase inhibitor herbicide resistance in grass weeds. *Weed Res* 46:34–39
- Kwok S, Chang SY, Sninsky JJ, Wang A (1994) A guide to the design and use of mismatched and degenerate primers. *PCR Methods Appl* 3:S39–S47
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta C_t$  Method. *Methods* 25:402–408
- Mallory-Smith C, Hall L, Burgos N (2015) Experimental methods to study gene flow. *Weed Sci* 63:12–22
- Martin JA, Wang Z (2011) Next-generation transcriptome assembly. *Nat Rev Genet* 12:671–682

- Metzker ML (2005) Emerging technologies in DNA sequencing. *Genome Res* 15:1767–1776
- Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *Plant J* 14:387–392
- Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* 18:613–615
- Ness RW, Siol M, Barrett SCH (2011) De novo sequence assembly and characterization of the floral transcriptome in cross- and self-fertilizing plants. *BMC Genomics* 12:298
- Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. *Nat Protoc* 1:1559–1582
- Ohtsuka E, Matsuki S, Ikehara M, Takahashi Y, Matsubara K (1985) An alternative approach to deoxyligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. *J Biol Chem* 260:2605–2608
- Ozsolak F, Milos PM (2010) RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet* 12:87–98
- Petit C, Pernin F, Heydel JM, Délye C (2012) Validation of a set of reference genes to study response to herbicide stress in grasses. *BMC Res Notes* 5:18
- Pettersson M, Bylund M, Alderborn A (2003) Molecular haplotype determination using allele-specific PCR and pyrosequencing technology. *Genomics* 82:390–396
- Pfaffl MW (2001) A new mathematical model relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45
- Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30:e36
- Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper-Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26:509–515
- Radstrom P, Lofstrom C, Lovenklev M (2003) Strategies for overcoming PCR inhibition. Pages 149–161 in Dieffenbach CW, Dveksler GS, eds. *PCR Primer: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press
- Riggins CW, Peng Y, Stewart CN, Jr, Tranel PJ (2010) Characterization of de novo transcriptome for waterhemp (*Amaranthus tuberculatus*) using GS-FLX 454 pyrosequencing and its application for studies of herbicide target-site genes. *Pest Manag Sci* 66:1042–1052
- Roux KH (1995) Optimization and troubleshooting in PCR. *Genome Res* 4:S185–S194
- Roux KH (2009) Optimization and troubleshooting in PCR. *Cold Spring Harbor Protoc* 1–6, doi:10.1101/pdb.ip66. <http://cshprotocols.cshlp.org/content/2009/4/pdb.ip66>. Accessed June 18, 2013
- Salas RA, Dayan FE, Pan Z, Watson SB, Dickson JW, Scott RC, Burgos NR (2012) EPSPS gene amplification in glyphosate-resistant Italian ryegrass (*Lolium perenne* ssp. *multiflorum*) from Arkansas. *Pest Manag Sci* 68:1223–1230
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–5467
- Scarabel L, Locascio A, Furini A, Sattin M, Varotto S (2010) Characterisation of ALS genes in the polyploid species *Schoenoplectus mucronatus* and implications for resistance management. *Pest Manag Sci* 66:337–344
- Sommer SS, Groszbar AR, Bottema CDK (1992) PCR amplification of specific alleles (PASA) is a general method for rapidly detecting known single base-pair changes. *Biotechniques* 12:82–87
- Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503–517
- Thudi M, Li Y, Jackson SA, May GD, Varshney RK (2012) Current state-of-art of sequencing technologies for plant genomics research. *Brief Funct Genomics* 2:3–11
- Tindall KR, Kunkel TA (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27:6008–6013
- Tranel PJ, Riggins CW, Bell MS, Hager AG (2011) Herbicide resistances in *Amaranthus tuberculatus*: a call for new options. *J Agric Food Chem* 59:5808–5812
- Tranel PJ, Wright TR (2002) Resistance of weeds to ALS-inhibiting herbicides: what have we learned. *Weed Sci* 50:700–712
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:10.1186/gb-2002-3-7-research0034
- Vigueira CC, Olsen KM, Caicedo AL (2013) The red queen in the corn: agricultural weeds as models of rapid adaptive evolution. *Heredity* 110:303–311
- Wagner J, Haas HU, Hurlle K (2002) Identification of ALS inhibitor-resistant *Amaranthus* biotypes using polymerase chain reaction amplification of specific alleles. *Weed Res* 42:280–286
- Ward JA, Ponnola L, Weber CA (2012) Strategies for transcriptome analysis in nonmodel plants. *Am J Bot* 99:267–276
- Warwick SI, Xu R, Sauder C, Beckie HJ (2008) Acetolactate synthase target-site mutations and single nucleotide polymorphism genotyping in ALS-resistant kochia (*Kochia scoparia*). *Weed Sci* 56:797–806
- Yu Q, Ahmad-Hamdani MS, Han H, Christoffers MJ, Powles SB (2013) Herbicide resistance-endowing ACCase gene mutations in hexaploid wild oat (*Avena fatua*): insights into resistance evolution in a hexaploid species. *Heredity* 110:220–231
- Yuan JS, Tranel PJ, Stewart CN, Jr (2007) Non-target-site herbicide resistance: a family business. *Trends Plant Sci* 12:6–13

Received June 18, 2013, and approved November 19, 2013.