Use of Pyrosequencing Technology to Genotype Imidazolinone-Tolerant Wheat

Mackenzie T. Ellison, Frank Fischinger, Robert S. Zemetra, and Joseph C. Kuhl*

M.T. Ellison and J.C. Kuhl, Dep. of Plant, Soil, and Entomological Sciences, University of Idaho, Moscow, ID 83844;

F. Fischinger, Bio-Rad Laboratories, Hercules, CA 94547;

R.S. Zemetra, Dep. of Crop and Soil Science, Oregon State University, Corvallis, OR 97331.

Received __________ *Corresponding author (jkuhl@uidaho.edu).

Abbreviations: ALS (acetolactate synthase), AHAS (acetohydroxyacid synthase), AHAS LSU (acetohydroxyacid synthase large sub-unit), PCR (polymerase chain reaction), SNP (single nucleotide polymorphism)
ABSTRACT

Cultivars of several cereal crops have been developed with acetohydroxyacid synthase (AHAS) insensitivity to imidazolinone herbicides and are now an important tool for weed management. Options for screening for imidazolinone resistant lines include direct herbicide application, biochemical assays for AHAS activity and DNA-based methods. Herbicide and biochemical assays for AHAS activity provide limited information as to mutation copy number and provide no information as to the genome on which the mutation is located without extensive test crossing. Wheat (*Triticum aestivum* L.) can have between one and six copies of the resistant acetohydroxyacid synthase on any of three genomes. A novel DNA-based screening protocol is described here in which pyrosequencing is used to screen for the S653N imidazolinone tolerant mutation in wheat. One assay is shown to successfully detect zero to four copies of the S653N mutation, while additional assays can detect the presence of S653N in individual wheat genomes. All of these assays are based on a single 298-bp PCR fragment and can be easily scaled up or down depending on the number lines that need to be screened. Potential applications include detection of mutant copy number in segregating populations, and the selection of parental lines with genome specific mutant composition.
The enzyme acetolactate synthase (ALS, EC 2.2.1.6) also known as acetohydroxyacid synthase (AHAS) is a critical enzyme in the metabolic pathway for the synthesis of the branched chain amino acids valine, leucine and isoleucine. This enzyme is the site of action of many herbicide families, including imidazolinones, sulfonylureas, triazolopyrimidines and pynmidoxybenzoates (Shaner et al., 1984; Chaleff and Mauvais, 1984; Subramanian et al., 1990). Imidazolinone herbicides control a wide spectrum of weeds, are effective at low rates of application, and have low mammalian toxicity and environmental effects (Shaner and O'Connor, 1991).

Grassy weeds similar to wheat (*Triticum aestivum* L.) with respect to growth habit and life cycle, such as jointed goatgrass (*Aegilops cylindrica* L. AEGCY), feral rye (*Secale cereal* L. SECCE), and downy brome (*Bromus tectorum* L. BROTE) are difficult to control in cultivated wheat using cultural and mechanical methods and have limited chemical controls available (Ball et al., 1999; Geier et al., 2004; Pester et al., 2001). In response to the absence of selective chemical controls, cultivars have been developed with AHAS that are insensitive to imidazolinones (Ball et al., 1999), a strategy which is now an important tool for the management of weeds in wheat (Geier et al., 2004; Kniss et al., 2008).

There are several mutations that confer tolerance to AHAS-inhibiting herbicides occurring in the large sub-unit (LSU) of an AHAS polypeptide. The most common mutations providing tolerance to one or more herbicides in the imidazolinone class are at the positions of Ala122, Pro197, Ala205, Trp574 and Ser653 (numbers indicate the amino acid position in *Arabidopsis thaliana* AHAS) of the AHAS LSU (Tan et al., 2005). The mutation at Ser653 confers tolerance to imidazolinones but not to other AHAS inhibitors. Imidazolinone tolerance is conferred by a single point mutation from G to A of the coding sequence corresponding to a
change from serine to asparagine at residue 653 (Sathasivan et al., 1991). Wheat is an
allohexaploid with three genomes (A, B, and D), and a chromosome number of 2n=6x=42
(Poehlman and Sleper, 1995). Each genome has a homologous \textit{AHAS} gene located on the long
arm of chromosome 6 (6A, 6B and 6D). Imidazolinone resistance is additive with higher levels
of resistance observed with each additional resistant allele (Pozniak, et al., 2004). Up to six
resistant copies possible with two in each genome.

In order to produce new wheat cultivars containing mutant AHAS that confer tolerance to
imidazolinones, sufficient screening methods need to be in place to detect the mutation. Options
for screening are direct herbicide application, biochemical assays for AHAS activity and DNA-
based methods. Testing by application of herbicide and biochemical assays for AHAS activity
can provide limited information as to mutation copy number; this is especially true in trying to
discriminate between 3 and 4 copies of the allele. These methods of screening also provide no
information as to the genome on which the mutation is located without extensive test crossing
(Willenborg et al., 2008; Newhouse et al., 1992). DNA-based methods such as polymerase chain
reaction (PCR) targeting DNA from the \textit{AHAS} LSU followed by traditional Sanger sequencing
can be used to identify the mutant allele and the particular genome on which it is located;
however such a method is not practical for more than a modest number of samples. PCR-based
assays (without sequencing) have been recently demonstrated in barley (\textit{Hordeum vulgare} L.),
however, as a diploid, barley only has one genome and as such is a simpler case than wheat (Lee
et al., 2011). PCR assays for wheat lose their simplicity because instead of a single PCR
reaction, at the minimum, one is required for each genome, and indeed, the protocol developed
by BASF Corporation (Florham Park, New Jersy) for varietal development uses seven different
PCR reactions for a complete picture (Zhao et al., 2005).
An alternative DNA-based screening protocol is described here in which pyrosequencing is used to screen for the S653N imidazolinone tolerant mutation(s) in all three wheat genomes. Pyrosequencing is a sequencing-by-synthesis, primer-directed polymerase extension method that relies on the detection of pyrophosphate which is released upon nucleotide incorporation and was first described by Ronaghi et al., 1998. Pyrosequencing is well suited to single-nucleotide polymorphism (SNP) analysis allowing for rapid genotyping as well as having the ability to differentiate between homozygous and heterozygous loci (Ahmadian et al., 2000) as a percentage, which is very useful in the case of wheat which has three AHAS genes.
MATERIALS AND METHODS

Plant material

Wheat plants used for establishing zero through four copies of the S653N mutation included Lambert (Zemetra et al., 1995) as wild-type, e.g. zero copies. Segregating populations included Dune05 by FS4 (source of the D mutation) and Dune05 by Teal 11A (source of the B mutation) (Hughes and Hucl, 1993). Multiple backcrosses to Dune05 providing a single heterozygous mutation on the D genome and a three-copy mutant, homozygous on the B genome and heterozygous on the D genome. UICF Brundage (Zemetra et al., 1998), a two-copy mutant, is homozygous for the D genome mutation. ORI2101835, an advanced breeding line from Oregon State University, is a four-copy mutant, homozygous for the A and D genome mutations. Genome specific S653N mutations are denoted by lower case genome letters, such that one mutant copy in the B genome is denoted as AABbDD.

Unknowns used for assay validation were segregating crosses and doubled haploids from crosses with FS4 and Teal 11A generate by the University of Idaho wheat breeding program. An independently created four-copy homozygous B and D genome mutant was produced by Northwest Plant Breeding Co. in Pullman, WA for Oregon State University.

PCR Amplification of AHAS DNA

Genomic DNA was recovered by collecting 100mg of fresh leaf tissue from 2 week old wheat seedlings and DNA was isolated using the DNeasy Plant Mini kit according to the manufacturers’ protocol (Qiagen, Valencia, CA). PCR amplification of a 298-bp region of the AHAS gene was performed using a forward primer (5’-TACCTTGGCAACCCAGAAAAT-3’) and biotinylated reverse primer (5’/5Biosg/CTGATTGCGCATGTCACACTT-3’) both at 0.5 μM final concentration. All PCR and sequencing primers were designed using PyroMark Assay
Design 2.0 software (Qiagen, Valencia, CA). Reactions were performed in a total volume of 20 μL with a 1x PCR buffer (minus Mg), 3 mM MgCl₂, 2 units Platinum Taq (Invitrogen, Grand Island, NY), 0.5 μM dNTPs and 25 ng genomic wheat DNA on a PTC-220 DNA Engine Dyad® (BioRad, Hercules, CA) with the following PCR parameters: 94°C for 2 min followed by 10 cycles of 94°C for 30 sec, 69 to 66°C (decreasing by 0.3°C every cycle) and 72°C for 18 seconds followed by an additional 35 cycles of 94°C for 30 sec, 66°C for 30 sec, 72°C for 18 sec, then a final extension at 72°C for 5 min. Forty-five cycles ensure that all PCR components were exhausted, in particular the biotinylated primer. Amplified DNA quality was checked using one μl of the PCR reaction on a 1.2% agarose gel.

**Pyrosequencing Assays**

Total copy number of the S653N imidazolinone tolerant mutation was detected using a genome non-specific sequencing primer 5’-TGCCTATGATCCCAA-3’. The target region for pyrosequencing is underlined in Figure 1. Genome specific sequencing primers for A, B, and D genomes were 5’-GCCATACTTGTGGATATCATC-3’, 5’-GCCATACTTGTGGATATCATCATT-3’, and 5’-GCCATACTTGTGGATATCATCATCATA-3’, respectively, with a single base-pair change at the 3’ end of the primers corresponding to a genome specific SNP. Genome specific primers utilized a target region from the primers to the S653N SNP. An experiment to determine PCR bias was performed using a genome nonspecific primer identical to the genome specific primers except it was missing the 3’ genome specific nucleotide, 5’-GCCATACTTGTGGATATCATC-3’. All pyrosequencing assays utilized the PCR product described above. Unless otherwise specified, all pyrosequencing assays involved the genome non-specific sequencing primer. The sequence alignment with primers and mutations is presented in Figure 1. Fifteen μl of PCR product was used in pyrosequencing assays.
on a PyroMark Q24 instrument (Qiagen, Valencia, CA) according to the manufacturers’
protocol. All assays included a “no PCR” control in position A1 to detect residual reagents in
the PyroMark cartridge.

Plants with zero through four copies of the S653N mutation were initially determined by
the BASF protocol (Zhao et al., 2005). The same DNA was used for pyrosequencing assays to
determine the percentage of the mutant SNP. Four replications of each category were performed
and a 95 and 99% confidence interval was determined based on pyrosequencing output. After
establishment of confidence intervals, new confirmed data points were folded into the confidence
interval determination, and disagreements between the two assays were settled using the genome
specific assay to determine copy number for each genome.
RESULTS

A four replicate trial of each of the five mutation levels available (wild type, one, two, three and four S653N mutant copies out of six possible), as determined by the BASF protocol, was used to create a confidence interval correlating pyrosequencing output to S653N copy number. Pyrosequencing results (percent mutant allele) for wild type (zero copies), the punitive 95% confidence interval was 2.3 – 4.2, for one-copy mutants it was 17.4 – 20.6, for two-copy mutants it was 33.4 – 37.1, for three-copy mutants it was 54.8 – 56.7 and for four-copy mutants it was 62.4 – 64.6 (Figure 2, Table 1). Standard deviation was greatest at four copies, however two copies had greater standard deviation than three copies. An example of pyrosequencing output for each copy level is presented in Figure 3. Once the pyrosequencing assay was finalized, repeated runs were made with the following controls; template only, sequencing primer alone, negative PCR sample, and negative PCR sample minus sequencing primer. All controls indicated an artifact-free assay. Confidence intervals were tested against 76 unknowns from segregating populations with five levels of mutation copy available and the results were confirmed by the BASF protocol with no discrepancies (Supplemental Table 1).

Polymerase chain reaction bias was performed using 5 lines of varying mutation levels using a genome nonspecific primer identical to the genome specific primers except it was missing the 3’ genome specific nucleotide. In this way pyrosequencing detected the proportion of each genome specific nucleotide, C, T, and A, corresponding to A, B, and D genomes, respectively (Figure 1). Each line was run four times. The genomic bias was calculated by averaging the results of each genome specific primer (for the A, B, and D genomes). The mean PCR amplification rate, which should theoretically be 33.3% for each of the three genomes, was 27.4, 35.2 and 37.3 for the D, A and B genomes, respectively (Table 2). The D genome
consistently amplified at a lower rate than A or B genomes. The B genome amplified at the
highest rate in all lines except for the four-copy mutant ORI 2101835 where the A genome
amplified to the greatest extent.

Genome specific primers were evaluated for genome specificity and mutant percentage
using the five mutation levels mentioned above. Results were similar to those expected (Table
3), although 1- and 2-copy mutant levels tended to be low (data not shown), i.e. less than the
expected 50% and 100%. Mutant detection was specific to the genome of the sequencing
primer, although B genome specific primers showed low levels of the mutant SNP with no
copies of the mutation present. A and D genome specific sequencing primers both showed 0%
mutant SNP in the appropriate lines. One versus two copies per genome were generally easily
distinguishable.
DISCUSSION

A single PCR reaction was combined with pyrosequencing to detect copy number of the S653N mutation in wheat. Zero to four copies were reliably detected with this assay (Figure 2, Table 1). When evaluating zero-copy lines, a low level artifact S653N SNP was detected, rather than the expected zero percent, an average of 3.13 percent was observed. This could be caused by a low level of truncated sequencing primer, missing an A on the 3’ end. All copy numbers were easily determined with no overlap between S653N frequencies as indicated by the 95% and 99% confidence intervals. Our data shows increasing variation as copy number increases.

Previously, Wasson et al. (2002) observed a $R^2$ statistic of 0.9963 relating peak height to allele frequency at 5% intervals from 1 to 99%, although this was based on hand mixed samples to generate the various SNP frequencies. Although not tested, our assay should also be effective for five and six copies, however increasing variance may become an issue.

Pyrosequencing values represented as percent S653N SNP of zero to four copies are between two and five percentage points higher than expected values. This is a minor inconvenience considering that all copy number levels are easily distinguishable. However, it is conceivable that these values could be adjusted with an algorithm to more closely match expected values. Since this was not a goal of this research no suggestions will be made here, however such a strategy appears feasible. The pyrosequencing instrument, PyroMark Q24, is capable of making such adjustments in real time.

The assay described here utilizes a single PCR reaction, a great simplification from the seven PCR reactions required by the BASF protocol (Zhao et al., 2005). After PCR a relatively simple procedure is required to prepare the PCR product for pyrosequencing. Pyrosequencing is very rapid, requiring as little as fifteen minutes per run. Pyrosequencing results provide clear
confirmation that the assay targeted the correct region. This assay is easily scaled up to medium throughput due to ease of use and short processing times. Higher throughput instruments are available that handle 96 samples at a time compared to 24 samples of the instrument used here.

Variation in PCR efficiency may distort the amplicon abundances causing distortion of pyrosquencing results (Sikaroodi and Gilevet, 2012; Wagner et al., 1994). Therefore, the relative abundance of each genome was determined for the amplicon used in these experiments. Polymerase chain reaction bias was detected such that the D genome amplified at a lower frequency, 27.4%, than A and B genomes, 35.2% and 37.3%, respectively (Table 2). Assuming no other bias in the assay, this would generate distortion in the mutation frequency.

Unfortunately the lines used here do not provide the best combination of genome mutants to make a direct comparison, e.g. single mutant copy in each of the three genomes. However, it is possible to compare four-copy mutants, seventeen AAbbdd lines and eight aaBBdd lines, which averaged 72.8% and 62% S653N, respectively (Supplemental Table 1). The higher S365N percentage of AAbbdd lines compared to aaBBdd lines is expected considering the B genome is amplified at a higher level, 37.3%, than the A genome, 35.3%, although the difference is greater than expected. It would appear that PCR bias relative to the genome does contribute to skewing of S653N detection in this case.

Although one assay will reliably detect total S653N copy number, in certain circumstances the detection of copy number within each of wheat’s three genomes is desired. This was accomplished through utilization of the genome specific SNP 38 base pairs upstream of the S653N polymorphism (Figure 1). Three assays, one for each genome, are required to completely evaluate copy number. This could be reduced if one or two genomes are known to carry the mutation, thereby eliminating unnecessary assays. The B genome specific sequencing
primer detected low levels of mutant SNP when no mutation was present, which is similar to results using the genome non-specific sequencing primer. B and D genome specific assays detected less than the 50% and 100% expected for the 1 and 2 mutant copies, respectively. This may reflect loss of signal as the distance from the sequencing primer increased.

Results presented here demonstrate a novel strategy for determining copy number of the S653N mutation in hexaploid wheat. These assays require significantly fewer steps than current assays. Wheat breeders would benefit from knowing both the total S653N copy number and genome specific values. This has obvious applications in terms of determining those lines for use in a crossing block, but is also of value to determine seed purity. However, additional work is needed to define sensitivity of this assay for S653N detection. It might be possible to measure allele frequencies in a group of individuals using far fewer reactions (Sham et al., 2002). Such an approach would aid in identification of S653N contaminated seed, e.g. purity of wild type wheat, as well as testing seed lot purity of AHAS resistant lines. In general, pyrosequencing assays could be developed for a wide range of genes of interest in wheat with similar applications, successful development of such assays depends on the polymorphism present and the nature of the surrounding sequence.

ACKNOWLEDGMENTS

This work was partially funded by a gift from BASF and start-up money from College of Agriculture and Life Sciences, University of Idaho.
References


Figure captions

Figure 1. Alignment of acetohydroxyacid synthase (AHAS) sequences representing A, B and D genomes wild type and related mutant alleles. The first arrow identifies the SNP associated with genome identification. The second arrow identifies the SNP associated with the S653N mutation. Four primers are shaded. Forward and reverse PCR primers are at the beginning and end. Genome specific sequencing primers are located immediately before the first arrow, while the genome non-specific sequencing primer for detecting total mutants is before the second arrow. Underlined is the target region for pyrosequencing with the genome non-specific sequencing primer.

Figure 2. Distribution of S653N mutation at 0-4 copies as expressed as percent SNP. The number of lines tested at 0-4 copies were 11, 9, 16, 15, and 25, respectively. Crosses indicate theoretical percentage at each copy number.

Figure 3. Pyrosequencing histogram showing 0 – 4 copies of the AHAS mutant allele. Percent of the A SNP in the shaded region reflects the proportion of S653N mutation out of six possible copies, theoretical values are 0%, 16.7%, 33.3%, 50.0% and 66.7%. Zero copy is represented by Brundage. One to four S653N mutant copies are represented by following mutant genotypes BbDDAA, BbDdAA, bbDdAA, and bbddAA, respectively.
Table 1. Confidence interval of 0-4 S653N mutation SNPs expressed in percent mutation from pyrosequencing, n=4.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Copy#</th>
<th>Mean</th>
<th>SD</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>99% CI Lower</th>
<th>99% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lambert</td>
<td>0</td>
<td>3.13</td>
<td>0.53</td>
<td>2.08</td>
<td>4.17</td>
<td>1.75</td>
<td>4.50</td>
</tr>
<tr>
<td>07-678†</td>
<td>1</td>
<td>19.00</td>
<td>0.82</td>
<td>17.40</td>
<td>20.60</td>
<td>16.90</td>
<td>21.10</td>
</tr>
<tr>
<td>UICF Brundage</td>
<td>2</td>
<td>35.35</td>
<td>1.32</td>
<td>32.77</td>
<td>37.94</td>
<td>31.95</td>
<td>38.75</td>
</tr>
<tr>
<td>07-678†</td>
<td>3</td>
<td>55.75</td>
<td>0.50</td>
<td>54.77</td>
<td>56.73</td>
<td>54.46</td>
<td>57.04</td>
</tr>
<tr>
<td>ORI2101835</td>
<td>4</td>
<td>70.80</td>
<td>4.43</td>
<td>62.11</td>
<td>79.49</td>
<td>59.39</td>
<td>82.21</td>
</tr>
</tbody>
</table>

† 07-678 is a segregating population derived from a cross between a line containing B and D homozygous mutations (4 total copies) and a wild type line, each 07-678 genotype is an individual plant.
Table 2. PCR amplification bias as determined by pyrosequencing of the genome specific SNP 38bp upstream of S653N, numbers reflect mean percentage of each genome detected in the amplified product, n=4.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Genome</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>D</td>
</tr>
<tr>
<td>Lambert</td>
<td>34.5</td>
<td>37.8</td>
<td>27.3</td>
</tr>
<tr>
<td>07-678†</td>
<td>36.0</td>
<td>36.3</td>
<td>27.3</td>
</tr>
<tr>
<td>UICF Brundage</td>
<td>33.5</td>
<td>39.5</td>
<td>27.3</td>
</tr>
<tr>
<td>07-678†</td>
<td>34.3</td>
<td>39.0</td>
<td>27.0</td>
</tr>
<tr>
<td>ORI2101835</td>
<td>37.8</td>
<td>34.0</td>
<td>28.3</td>
</tr>
<tr>
<td>Mean</td>
<td>35.2</td>
<td>37.3</td>
<td>27.4</td>
</tr>
<tr>
<td>SD</td>
<td>1.69</td>
<td>2.23</td>
<td>0.49</td>
</tr>
</tbody>
</table>

† 07-678 is a segregating population derived from a cross between a line containing B and D homozygous mutations (4 total copies) and a wild type line, each 07-678 genotype is an individual plant.
Table 3. A partial table indicating percent mutant *AHAS* allele, indicated by lower case genome letter, expected to be detected by genome specific sequencing primers. PCR primers are the same as those used for total mutant allele detection.

<table>
<thead>
<tr>
<th>Sequencing Primer</th>
<th>AABBDD</th>
<th>AAbbDD</th>
<th>AABBDd</th>
<th>AABbDd</th>
<th>AAbbDD</th>
<th>AABBdd</th>
</tr>
</thead>
<tbody>
<tr>
<td>A specific</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>B specific</td>
<td>0%</td>
<td>50%</td>
<td>0%</td>
<td>50%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>D specific</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>50%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>Total mutant %†</td>
<td>0%</td>
<td>17%</td>
<td>17%</td>
<td>33%</td>
<td>33%</td>
<td>33%</td>
</tr>
</tbody>
</table>

† Percentage total mutants out of 6 possible copies