Use of Marker Assisted Selection in a Product Development Breeding Program

Donna J. Cahill 1 and Daria H. Schmidt 2

1 Research Scientist, Pioneer Hi-Bred International, A DuPont Company donna.cahill@pioneer.com
2 Research Director, Pioneer Hi-Bred International, A DuPont Company daria.schmidt@pioneer.com

Abstract
Marker Assisted Selection (MAS) has held promise for impacting, perhaps revolutionizing, plant breeding disciplines. There are examples of commercial breeding programs making use of MAS in product development as a tool in breeding schemes for many agriculturally important crops, including grain, oilseeds, vegetables, ornamentals, and tree crops. In most cases, levels of efficiency in selection at early generations and characterization in later generations are the tangible deliverables from MAS. Building the infrastructure of a high throughput MAS program is a costly and time consuming process, with vision and patience required. As a case study, markers have proven to be a useful tool in the development of disease-resistant soybean (Glycine max) varieties. The efficiency they bring to the breeding process is measured by the success in accumulating desired traits, particularly soybean cyst nematode resistance, in a larger percentage of candidate experimental lines throughout the breeding pipeline. This technology has been sufficiently streamlined to allow for high throughput data generation in a timely manner that supports "traditional" cultivar development. The improvements for sample turn-around and total data point capacity are continuing to make MAS a cost-effective tool in the breeding of high performing soybean cyst nematode (SCN) resistant varieties. MAS for soybean product development at Pioneer Hi-Bred has been used for almost 10 years and has significantly impacted early generation single plant selections, resulting in numerous successful commercial soybean varieties.

Media summary
Marker assisted selection (MAS) has been successfully used in soybean (Glycine max) product development in early generation single plant selection for key pest resistance.

Key Words
Soybean, Glycine max, Soybean Cyst Nematode, Single Nucleotide Polymorphism, Phytophthora Sojae, Marker Assisted Selection

Introduction
Examples of marker-trait associations can be readily found in plant science literature and are too numerous to review in this document. To demonstrate the influence and vast scope that marker-trait associations have on every-day consumer products, Table 1 includes examples of these associations in species affecting a single dining experience. The hops in a glass of beer, the olive and pepper hors d'oeuvre, lentil soup, slice of melon, potato side dish, broccoli, baked pears with sugar, and finally, a cup of coffee all have marker-trait associations allowing selection for traits of interest.
Table 1: Selected Snapshot from mid-year, 2003 literature search referencing marker-trait associations for plant species.

<table>
<thead>
<tr>
<th>Scientific Name</th>
<th>Trait</th>
<th>Common Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avena sativa L.</td>
<td>Oat crown rust</td>
<td>Oat (Zhu and Kaeppler 2003)</td>
</tr>
<tr>
<td>Brassica napus</td>
<td>Oil composition</td>
<td>Canola (Burns et al. 2003)</td>
</tr>
<tr>
<td>Brassica spp.</td>
<td>Quinine reductase</td>
<td>Broccoli (Mithen et al. 2003)</td>
</tr>
<tr>
<td>Capsicum spp.</td>
<td>Capsaicinoids</td>
<td>Pepper (Blum et al. 2002)</td>
</tr>
<tr>
<td>Cocos nucifera</td>
<td>Lethal yellowing</td>
<td>Coconut palm (Cardena et al. 2003)</td>
</tr>
<tr>
<td>Coffee arabica</td>
<td>Root-knot nematode</td>
<td>Coffee (Noir et al. 2003)</td>
</tr>
<tr>
<td>Cucumis melo</td>
<td>Fusarium wilt</td>
<td>Melon (Burger et al. 2003)</td>
</tr>
<tr>
<td>Dioscorea alata</td>
<td>Anthracnose Resistance</td>
<td>Water yam (Mignouna et al. 2003)</td>
</tr>
<tr>
<td>Elaegnus</td>
<td>Oil quality</td>
<td>Olive (Rosa et al. 2002)</td>
</tr>
<tr>
<td>Glycine max</td>
<td>Lepidopteran pests</td>
<td>Soybean (Walker et al. 2002)</td>
</tr>
<tr>
<td>Gossypium spp.</td>
<td>Fiber strength</td>
<td>Cotton (Zhang et al. 2003)</td>
</tr>
<tr>
<td>Heliotus annuus</td>
<td>Male fertility restorer</td>
<td>Sunflower (Horn et al. 2003)</td>
</tr>
<tr>
<td>Hordeum vulgare</td>
<td>Leaf rust resistance</td>
<td>Barley (Mammadov et al. 2003)</td>
</tr>
<tr>
<td>Humulus lupulus</td>
<td>Plant gender</td>
<td>Hops (Patzak et al. 2002)</td>
</tr>
<tr>
<td>Larix spp.</td>
<td>Wood characteristics</td>
<td>Larches (Arcade et al. 2002)</td>
</tr>
<tr>
<td>Lathyrus odoratus</td>
<td>Tendril morphology</td>
<td>Sweet pea (Hanada et al. 2003)</td>
</tr>
<tr>
<td>Lens culinaris</td>
<td>Anthracnose</td>
<td>Lentils (Tullu et al. 2003)</td>
</tr>
<tr>
<td>Linum usitatissimum</td>
<td>Rust resistance</td>
<td>Flax (Bo et al. 2002)</td>
</tr>
<tr>
<td>Lolium perenne</td>
<td>Temperature response</td>
<td>Loliun forage grass (Skot et al. 2002)</td>
</tr>
<tr>
<td>Lupinus angustifolius</td>
<td>Diaportha resistance</td>
<td>Lupins (Yang et al. 2003)</td>
</tr>
<tr>
<td>Manihot esculenta</td>
<td>Early root bulking</td>
<td>Cassava (Okogbenin et al. 2002)</td>
</tr>
<tr>
<td>Micanthus sinensis</td>
<td>Biomass</td>
<td>Silvergrass (Atienza et al. 2003)</td>
</tr>
<tr>
<td>Nicotiana tabacum</td>
<td>Black shank disease</td>
<td>Tobacco (Johnson et al. 2002)</td>
</tr>
<tr>
<td>Oriza sativa L.</td>
<td>Bacterial blight</td>
<td>Rice (Cao et al. 2003)</td>
</tr>
<tr>
<td>Pennisetum glaucum</td>
<td>Stover yield</td>
<td>Pearl Millet (Yadav et al. 2003)</td>
</tr>
<tr>
<td>Phaseolus vulgaris</td>
<td>Rust</td>
<td>Common bean (Kelly et al. 2003)</td>
</tr>
<tr>
<td>Prunus persica</td>
<td>Powdery mildew</td>
<td>Peach (Scorza et al. 2002)</td>
</tr>
<tr>
<td>Pyrus spp.</td>
<td>Black spot and pear scab</td>
<td>Pear (Yamamota et al. 2002)</td>
</tr>
<tr>
<td>Saccharaum</td>
<td>Sugar yields</td>
<td>Sugarcane (Ming et al. 2002)</td>
</tr>
<tr>
<td>Secale cereale</td>
<td>Male fertility restorer</td>
<td>Rye (Stracke et al. 2003)</td>
</tr>
<tr>
<td>Sesame indicum</td>
<td>Capsule morphology</td>
<td>Sesame (Uzun et al. 2003)</td>
</tr>
<tr>
<td>Solanum tuberosum</td>
<td>Late blight</td>
<td>Potato (Simko 2002)</td>
</tr>
<tr>
<td>Sorghum bicolor</td>
<td>Midge resistance</td>
<td>Sorghum (Tao et al. 2003)</td>
</tr>
<tr>
<td>Zea mays</td>
<td>Protein quality</td>
<td>Maize (Dreher et al. 2003)</td>
</tr>
</tbody>
</table>

Examples of Molecular Markers for MAS being used in commercial seed businesses are more difficult to track. However, more and more companies are coming forward with statements acknowledging their use of markers in the development of their products, and there are several examples of the use of DNA diagnostics in commercial companies. Harris Moran Seed Company mentions the use directly on their October, 2003 website. http://www.harrismoran.com/products/biotechstatement.htm

“Today Harris Moran employs many advanced techniques to assist and improve the classic plant breeding work that is the core of our business. One example is the use of Molecular Markers to allow the identification of specific genes and plant traits. Our breeders use markers to assist in selection of new plant lines with important characteristics, our plant pathologist use markers to identify plant diseases and our quality assurance folks use markers to help insure the seed you buy is true to type - truly versatile and valuable technology”

Syngenta is involved world wide in research and product development in the areas of vegetables, flowers, and agronomic crops. Syngenta indicated they utilize MAS in one or more of their product lines, as noted on their October, 2003 website. http://www.syngenta.com/en/about_syngenta/research_tech_how.aspx#research

“While traditional breeding approaches have worked well, they can be very time consuming and sometimes inefficient. Each plant has tens of thousands of genes, so crossing two plants results in a multitude of combinations. With advances in technology, scientists can now precisely identify through
Marker technology allows some of the individual genes responsible for producing a particular characteristic. Traits such as fruit color or resistance to a particular pest, make breeding new varieties quicker and more precise. New technologies have opened doors to certain improvements that were not possible before. Marker technology helps indicate the presence of specific characteristics in plants. In marker-assisted breeding, lab analysis and identification of the "marker" linked to a gene supplements the time-consuming process of crossing plants, growing them and selecting the best in the field or greenhouse.

Zeraim Gederas, a vegetable seed company based in Israel, references the use of molecular markers in their company profile document found on their October, 2003 website. [http://zeraimgedera.com/](http://zeraimgedera.com/).

"The target of Zeraim Gederas R&D is the creation and improvement of new varieties combining classic breeding for quality traits such as taste, uniformity and high yield, with new safe molecular methods such as assisted selection with genetic markers on the DNA level for disease resistance and certain quality traits."

Application of markers to commercial product development is referenced in the area of forestry as well. Forest Research, based in New Zealand, referenced the use of allele specific hybridization for wood density parameters on their October, 2003 website. [http://www.forestresearch.co.nz/](http://www.forestresearch.co.nz/)

**"ASSOCIATION STUDIES"**

When markers are used to select for desirable traits the statistical correlations between marker alleles and phenotypic extremes of the trait are often established by analysing segregation data from controlled crosses. Such correlations are valid for a particular cross, but may not be valid for crosses between different individuals because linkage relationships among marker and trait alleles vary among individuals in an outcrossing population. It is possible to identify markers and gene sequences that are associated with trait variation in a population by using association tests to determine gametic or linkage disequilibrium. We have isolated DNA and taken increment cores from an association population of about 2000 radiata pine trees. Initially we will be looking for association between wood density parameters and SSR or candidate gene alleles SNPs. Our current preferred method of SNP genotyping is to use allele-specific primers.

Case study of SNP marker application in Soybean Product Development at Pioneer Hi-Bred

Justification for development of a production system for SNP Markers

Molecular markers lend themselves nicely in several areas of applied plant breeding towards product development in many crops, including soybean. At Pioneer, we employ molecular technologies to progeny screen for individual plants with resistance to key high-value traits, which include Soybean Cyst Nematode (*Heterodera glycines*), SCN, Phytophthora root rot (*Phytophthora sojae*), PRR, and Brown Stem Rot (*Phialophora gregata*), BSR.

Soybean cyst nematode has been long recognized as a globally significant pest in nearly all soybean regions, including North America, Canada, Indonesia, Japan, Korea, China, and South America (Riggs and Niblack 1999). Soybean production loss due to SCN is significant and may be greater than three million metric tons (Wrather, et al. 1997) globally. Infection causes significant root damage and can often result in plant stunting and yellowing. Nematode cysts can remain in the soil for more than ten years. Resistance to SCN is quantitative in nature (Mansur, et al.1993), making it difficult to assay in a field setting. Bioassays can be challenging and costly in greenhouse and growth chamber systems as well.

Phytophthora Root Rot is a significant fungal pest of soybeans throughout the world, primarily in Argentina, Australia, Brazil, Canada, China, Hungry, Italy, Japan, and the United States (Schmitthenner 1999). The fungus can infect at any time during the soybean growth cycle; seedling infection may result in stand reduction requiring replant, and infection in older plants can result in significant yield loss. Developing lines with race-specific resistance has been very successful to date and researchers rely on race specific screening bioassays to detect resistance.
Brow Stem Rot occurs world-wide including regions of Argentina, Brazil, Canada, Egypt, Japan, Mexico, United States on several legume host species (Gray and Grau 1999). Economic losses are most often observed in seasons wherein cool, wet weather occurs during pod fill, followed by hot dry weather. First symptoms usually appear in the leaf at the R4 stage, although the fungus can be isolated from vascular tissue before these symptoms appear. Rbs resistance genes have been characterized in soybean breeding germplasm, and breeders have used traditional breeding methodology to develop resistant cultivars. The disease is somewhat problematic to screen in the field, although certain environments have been developed for consistent screening locations.

The significance of these traits from an economic standpoint combined with the challenges and cost in screening and characterizing lines based on phenotype were key items considered when making the decision to pursue a marker selection strategy. The soybean plant breeders were seeking a marker assisted selection tool that complemented the breeding process currently in place, with the goal of enhancing traditional variety development methods. The best-case scenario for a product development breeder would be to yield test lines that were fixed for the resistance traits required to meet customer needs. An extremely efficient use of yield testing resources in product development is a significant outcome of fixing key traits early in the line development process.

Development of a high throughput system took several years and required the coming together of new approaches in both the lab and field. Evolution of robotics for DNA processing and handling in mass numbers has been a critical development in the lab. Managing of samples through a sample tracking system from tissue arrival to data export was and continues to be a second key piece. Communication of result and sample selection in the field is a third step in the process that came together to move sample numbers per year from the tens of thousands to hundreds of thousands.

**Technical overview of the current MAS SNP production process**

Detection of DNA samples for their marker genes is facilitated through allele specific hybridization, (ASH). This method utilizes a single nucleotide polymorphism (SNP) in the soybean genome as the base for marker technology (Figure 1). Marker development for the current set of production markers used in the lab involved significant research efforts. For example, to determine the trait/marker associations established for SCN, a mapping project was undertaken (Webb 2000; Webb et al. 1995). The significant QTL markers were cloned and sequenced. Primers were developed to amplify the identified region, and SNPs between resistant and susceptible marker loci were discovered. Probes to identify the SNPs were then developed and tested across hundreds of pedigrees before the markers were moved into production. The MAS lab uses polymerase chain reaction (PCR) to amplify the regions of interest and hybridizes the probes to the sample DNA fixed onto a nitrocellulose membrane. Labeled probes detect which of the marker alleles are present in each sample for each locus.

![Probe 1](image.png)

### Allele 1

- **Target 1**: 3' ... CTAGAT GGTAC ... 5'
- **Probe 1**: 5' CCATGаCAATGT 3'

### Allele 2

- **Target 2**: 3' ... CTAGAT GGTAC ... 5'
- **Probe 2**: 5' CCATGаCAATGT 3'

**Figure 1.** A schematic where in targets 1 and 2 are the sample DNA PCR product of a single locus. The alleles differ by a single base pair, noted in red and green. Labeled probes specify which allele(s) are present in the sample via hybridization.

The MAS lab routinely screens for QTL markers for races 1,2,3,5, and 14 of SCN. They also routinely screen for several PMG loci, and for BSR. Prior to screening a segregating population, the parents of the cross are fingerprinted to determine parent marker profiles and marker/trait association. Parents of a cross must be segregating for the marker QTL in order to be able to detect resistance QTL in subsequent progeny. Routinely, breeders will request marker information for multiple disease traits that are...
detectable through the marker lab. The lab utilizes a multiplexing PCR process, which makes optimum use of the lab resources and allows for selection for several resistance traits in one sample.

Current throughput for population development has grown from approximately 20,000 data points in 1995 to approximately 2 million data points in 2003. This growth has come about by improvements to in-lab procedures, gains in sample efficiency by multiplexing (detecting a greater number of loci per sample), tissue sampling improvement, and information management programs. Routine traits selected for screening via MAS are SCN (several of the key loci), PMG (Rps1, Rps3, Rps4, Rps6 loci), BSR (Rbs3). MAS markers for several other traits and for several other species are currently under development. The MAS lab services the North American and South American Soybean breeding programs. Material and information flow is very dependent on a strong breeder/lab interface. There are several MAS "seasons" throughout a calendar year, which coincide with breeding cycles in North America, South America, and the tropical nursery locations.

Steps in the current process

Plant samples are collected into bullet tubes in the field using a leaf punch collection device and placed in the appropriate position in a 96-well plate (Figure 2). Plates are kept cool during the collection process and depending upon location, samples may be lyophilized (freeze-dried) prior to shipment to the lab. A location may collect as few as 10,000 plant samples, or as many as 100,000 for any given season.

The lab checks in plates as they arrive and lyophilizes them, if needed. The next process involves DNA extraction and transfer (Figure 3).

After extraction, DNA is transferred into plates for the PCR process (Figure 4). Plates are sorted and organized into groups of 15, based on the marker loci requested. The 16th plate contains checks and controls. PCR is performed to amplify the regions of the genome that are to be detected. The post-PCR product is then transferred to a nitrocellulose membrane.
Membranes are hybridized with the appropriate labeled allele specific oligonucleotide probes. Images are developed on X-ray film and scanned into a computer for scoring (Figure 5). A solid black dot on a single image would indicate that the plant sampled is homozygous for that particular marker allele; a gray to dark dot on both images would indicate that the plant sampled is heterozygous for the marker allele; no dot (no hybridization) would indicate that the plant sampled does not carry the marker allele.

The images are analyzed (Figure 6, left) and results are converted into a tabular form for upload into a database. Sample results are electronically communicated to the breeders and are used to select individual plants for advancement to the next level of testing. Integration of sample bar codes with the lab results insures that the correct selection is advanced, and the entry list for the next generation is directly generated from the selection file (Figure 6, right).

Current Implications of MAS in Soybean Product Development
Marker assisted selection has made a significant contribution in our efforts to efficiently screen for pest resistance early in the breeding process. Once data on individual plants from a segregating population is returned to the breeder at the research center, the breeder has the choice of selecting individuals that are
fixed for all or a majority of the loci that were screened. The trait quality of the breeders' single plant yield trials in the subsequent season has been greatly improved by use of MAS. Downstream plot resources are better utilized for experimental lines that have been fixed for traits, as lines susceptible are discarded before yield tests costs are incurred. Over the past ten years wherein the Pioneer soybean breeders have been utilizing MAS for product development as a significant tool, the percentage of experimental lines confirmed for traits has more than doubled at each pre-commercial stage in the pipeline.

The MAS assay allows for sampling of more than one trait per tissue sample in a very short timeframe. Breeders can screen for a number of soybean fungal diseases in conjunction with SCN. Bioassays often require separate tests and have resource costs that quickly add up. Costs are tied up in packaging seed, planting, inoculating, scoring, and data processing for each assay required in a one assay/one experiment traditional screening approach. Additionally, costs are incurred in land rental for field plots, or space management in greenhouses or growth chambers. Marker technology allows for detection for a number of traits in one experiment. MAS assays have an additional benefit in that they provide consistent results that do not rely on a pest/plant interaction. MAS provides validation information for various bioassay processes, which results in increased confidence in both bioassay based approaches and a DNA detection approaches. MAS lends itself well in the purification process and increases the confidence in quality of breeder seed delivered to pre-production facilities.

Trait selection via MAS has succeeded in favorably impacting the rate of genetic gain and variety performance in market segments where the traits under selection are critical for our soybean customer’s success.

References


