The *Arabidopsis* SeedGenes Project

Tzafrir I[1], Rogers R[1], Sweeney C[1], Hutchens S[1], Dickerman A[2], McElver J[3], Patton D[3], Meinke D[1]

[1] *Department of Botany, Oklahoma State University, Stillwater, OK 74078*
[2] *Virginia Bioinformatics Institute, Virginia Tech University, Blacksburg, VA 24061*
[3] *Syngenta, Research Triangle Park, NC 27709*

The purpose of the *Arabidopsis* SeedGenes Project is to coordinate the collection, analysis, and presentation of information on essential, non-redundant genes that give a seed phenotype when disrupted by mutation. *Arabidopsis* appears to contain ~750 such *EMB* genes required for seed development. Our long-term goal is to help the community establish a complete collection of *Arabidopsis* genes with a mutant seed phenotype. This information is needed to focus attention on genes with important cellular functions and to assess from a genetic perspective the extent of functional redundancy in the *Arabidopsis* genome.

The SeedGenes Project database (www.seedgenes.org) is organized into two major sections. One section deals with genes and the other with mutant alleles. The database can be queried for detailed information on a single gene to create a SeedGenes Profile. Queries can also generate lists of genes or mutants that fit specified criteria. The third database release (June 2003) contains information on ~220 essential genes and ~310 mutants. Included are many genes identified at Syngenta and all known genes described in the literature. The goal over the next two years is to present information on a total of 500 mutants defective in 300 *EMB* genes. Future enhancements include the addition of Nomarski images of mutant seed phenotypes, a tutorial on screening for seed defects, and information on genes expressed in young seeds.

Forward genetics will not enable the identification of every *EMB* gene through random mutagenesis. Small genes in particular will be difficult to detect. We are therefore using a bioinformatics-based approach to identify promising candidates for missing *EMB* genes based on sequence comparisons with essential genes in other model organisms. Candidate genes with evidence of seed expression and limited functional redundancy will then be subjected to reverse genetics using available insertion lines and screened for seed phenotypes.

Funding provided by the NSF 2010 Program and the SR Noble Foundation.