The *Arabidopsis* SeedGenes Project: Approaching Saturation for Essential Genes

Rosanna Muralla[1], Colleen Sweeney[1], Johnny Lloyd[1], Allan Dickerman[2], and David Meinke[1]

[1] Department of Botany, Oklahoma State University, Stillwater, OK 74078
[2] Virginia Bioinformatics Institute, Blacksburg, VA 24061

The purpose of the SeedGenes Project (www.seedgenes.org) is to present detailed information on essential genes in *Arabidopsis*. Emphasis is placed on genes required for seed development. The December 2007 database release includes more than 350 total genes and 600 mutant alleles. When pending additions found through the SeedGenes Access page are included, over 350 *EMB* genes alone have been identified to date. Approximately 65% of these gene identities have been confirmed through molecular complementation or the analysis of duplicate mutant alleles. Another 150 to 200 genes required for gametophyte development have been reported in the literature, although many of these identities remain to be confirmed. At least 50 genes required for seedling survival have also been described.

Lethals represent a critical and informative part of the *Arabidopsis* mutant collection. A robust dataset of essential genes is needed to define the minimal gene set required for plant growth and development and to complement ongoing research with other model organisms. We estimate that 30-50% of the *EMB* genes in *Arabidopsis* have been identified to date. Because forward genetic screens for lethals become less efficient with progress towards saturation, most of the remaining essential genes will likely be identified through reverse genetics. We describe here several different strategies that we have pursued to identify additional *EMB* genes by focusing on candidates that: (A) represent orthologs of essential genes in other organisms; (B) function in the same metabolic pathway or cellular process as known *EMBs*; (C) share protein interactors with a known *EMB* gene product; and (D) are associated with insertion mutants that fail to generate knockout homozygotes.

Research supported by the NSF 2010 Program. Individuals working on mutant alleles of genes first described in SeedGenes are requested to include the original *EMB* gene symbol in future publications.