

EMB Genes of Arabidopsis with Unknown Cellular Functions

Rosanna Pena- Muralla, Rebecca Rogers, and David Meinke

Department of Botany, Oklahoma State University, Stillwater, OK 74078 USA

Many genes of *Arabidopsis thaliana* are annotated to encode proteins with unknown functions. Determining what functions these proteins perform is a long-term objective of genomics efforts worldwide. We describe here a collection of *Arabidopsis* genes with unknown functions required for normal embryo development. These genes represent a valuable subset of the *Arabidopsis* unknowns because they are known to be essential. Included in this collection are proteins with defined motifs but uncertain cellular functions and proteins with uncertain functions based on marginal BLASTP matches. From an initial collection of 56 candidate unknowns with a knockout seed phenotype (www.seedgenes.org), we have confirmed 30 gene identities through the recovery of duplicate alleles derived from a combination of forward genetics (46 Syngenta mutants) and reverse genetics (25 Salk mutants). Another four genes have been confirmed through molecular complementation. Approximately half of the confirmed genes have no paralogs in *Arabidopsis* and most do not appear to have counterparts outside of plants. RT-PCR analysis confirmed that gene expression is for the most part not embryo-specific, consistent with general functions throughout the life cycle. We conclude that *EMB* genes represent a valuable resource for identifying novel proteins associated with important plant processes.

Funded by the NSF 2010 Program and the S.R. Noble Foundation.