
SEMINAR ANNOUNCEMENT

DATE: Thursday, 10/16/2003

TIME: 4:00PM - 5:00PM

PLACE: 112 Lefevre Hall

TITLE: Full-Length cDNAs and Characterization of
Genes in *Arabidopsis thaliana*

SPEAKER: Dr. Nickolai Alexandrov
Manager for Bioinformatics, Ceres Inc.

Abstract:

We have characterized a large collection of full-length cDNAs from *Arabidopsis thaliana*, together with another set of full-length cDNAs from Salk/Riken. We provide a statistical characterization of Arabidopsis transcripts in terms of their size, UTR lengths, 3' end cleavage sites, splicing variants, and coding potential. A statistical analysis of gene structure is also presented including intron and exon lengths, intergenic distances, features of promoters, variant 5' ends and splicing patterns of mRNAs transcribed from the same locus. Non-canonical splicing sites constitute less than 1% of all splicing sites. Putative alternative transcription start sites were observed in 30% of highly expressed genes and in more than 50% of the genes with low expression. The intergenic distance varies significantly depending on the orientation of the neighboring genes. Transcription start sites correlate remarkably with a GC skewed peak in the DNA sequence. The conclusions drawn not only provide a better understanding of the Arabidopsis genome but should also allow better predictions to be made for, as yet, poorly defined genes.