Accelerated Discovery of Cynodon *Genes and DNA Markers by cDNA Sequencing*

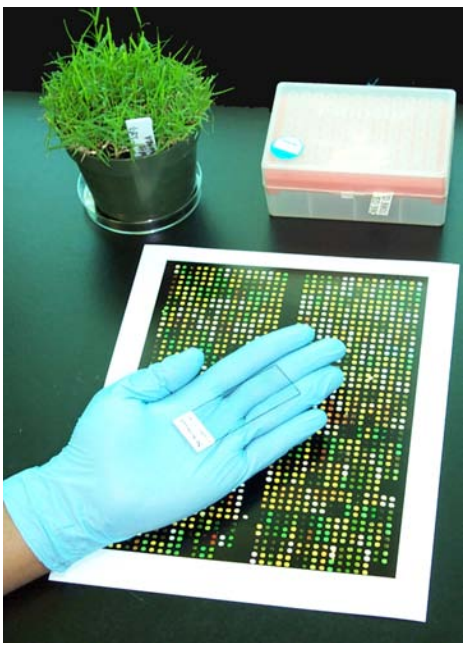
Andrew H. Paterson
University of Georgia

Objectives:

1. We will more than double the Internet-accessible 'gene encyclopedia' for *Cynodon*, by sequencing 12,000 *Cynodon* expressed sequence tags (ESTs).
2. To nurture application of EST resources to many questions in *Cynodon* biology and improvement, we will develop a freely-available online resource of "conserved intron scanning primers".

Start Date: 2006
Project Duration: three years
Total Funding: $90,000

While most genes have been identified for some crops such as maize, rice, and sorghum, *Cynodon* (bermudagrass) lags far behind. Our recently completed USGA project yielded sequences for portions of nearly 5,000 *Cynodon* genes, and a partner project (A. Guenzi) has yielded nearly 4,000 more for a total of about 9,000. While seemingly large, these numbers are tiny in comparison to 868,456 known for wheat, 472,163 for barley, and similarly high numbers for many other crops. Studies in other organisms suggest that EST sequencing is the most cost-effective gene discovery method up to at least 100,000 sequences-- in other words, there is much more to learn about *Cynodon* genes from this efficient approach.



University of Georgia scientists are identifying genes responsible for natural adaptation of different Cynodon *taxa to diverse environments, especially genes associated with drought resistance.*

Building on our development of the first 'genetic map' of the *Cynodon* genome, we will further expand knowledge of the 'sequences' (spelling) of large numbers of *Cynodon* genes, identifying the *Cynodon* versions of many genes for which functions are known or suspected from research in other taxa. Our analytical approach will thus leverage a large body of existing data for other plants to identify and determine the functions of many *Cynodon* genes.

We will also leverage NSF grants to the PI of $4 million for grass genomics, and $600,000 for informatics, in applying bioinformatic methods to *Cynodon* ESTs, and making the resulting information available to researchers. Finally, this project will benefit greatly from the recent sequencing by the US Department of Energy of the entire genome of sorghum, an important model for *Cynodon* and other tropical grasses.

EST sequencing is the most efficient approach to gene discovery in higher organisms, and remains cost-effective up to about 100,000 ESTs. Our current USGA project together with a companion project has determined partial sequences for about 9,000 *Cynodon* ESTs, as compared to 596,799 known for wheat, 425,127 for barley, and similarly high numbers for many other crops. This proposed project will increase the number of *Cynodon* gene sequences to about 20,000, more than doubling our knowledge of its genes and their functions.

Working from the established community resource of 36,864 cDNA clones individually archived in 384-well plates during our prior funding period, during this first year of funding we completed the sequencing of the additional 12,000 clones proposed, and have begun sequence data processing.

Graduate student Changsoo Kim is near completion of this work. ESTs from *Cynodon* have been trimmed to remove vector and low-quality sequence, and aligned to best-matching sites of the most current rice sequence (at www.tigr.org) using BLASTN($E=1x10^{-10}$).

Redundant hits have been removed and PCR primers designed from highly conserved (0-2 mismatches) alignments considering intergeneric sequence conservation. Primer design criteria include implied intron size (200-1,500 bp), oligonucleotide melting temperature (58-620^{o} C), size range (18-22bp), GC content (50%), and primer-dimer formation potential (minimized).

Design criteria are focused on compatibility with common sets of PCR conditions (specifically, target annealing temperatures of 55/60^{0} C). EC_oligos software is used to design all possible intron-spanning conserved PCR primer sets. An in house Perl script is used to parse CISP sets from the output file. These CISP sets are BLAST aligned (E=0.002) to the rice genome and the results loaded into an ACCESS (Microsoft) database to ascertain uniqueness and determine exact locations in the rice genome.

Summary Points

- Identification of corresponding rice/sorghum genes will permit scientists to deduce the probable functions of many *Cynodon* genes, and also reveal features that are present in the genomic DNA that surrounds *Cynodon* genes but are absent from the ESTs, such as the 'promoters' (on/off switches) that regulate their expression.
- While a complete sequence of the *Cynodon* genome remains far in the future, our work will build the framework for efficient progress by linking this project to the first detailed *Cynodon* genetic map.