The efficacy of GeneThresher® methylation filtering technology in the plant kingdom

U. Warek[1], J.A. Bedell[1], M.A. Budiman[1], A.N. Nunberg[1], R.W. Citek[1], D. Robbins[1], N. Lakey[1] and P.D. Rabinowicz[2]
[1]*Orion Genomics LLC, 4041 Forest Park Avenue, St. Louis, MO 63108, USA.,Email: uwarek@oriongenomics.com*
[2]*The Institute for Genomic Research, Rockville, MD 20850, USA*

Keywords: gene discovery, sequencing, SSRs, synteny

Introduction The genomes of many plants are known to be composed of a large fraction of repetitive DNA, while a small portion is dedicated to genes. The bulk of the repetitive DNA constitutes transposable elements and is heavily methylated. GeneThresher technology has been developed to take advantage of these differential methylation patterns by filtering genomic shotgun libraries to exclude methylated sequences (Rabinowicz *et al*., 1999; Palmer *et al*., 2003; Martienssen *et al*., 2004). The result is a gene-enriched genomic shotgun library. Random shotgun sequencing of plant gene space, enabled by GeneThresher technology, is a rapid and cost-effective strategy for comprehensive gene discovery in agriculturally important crops.

Materials and methods We have applied GeneThresher to a number of plants that span the major branches of the plant kingdom. We have tested species representing monocots, dicots, gymnosperms, and non-vascular plants with a last common ancestor estimated at 500 million years ago.

Results Gene enrichment was achieved in all plants tested suggesting that GeneThresher will be effective across the whole plant kingdom. Genes discovered in the filtered sequences appear to be a random, unbiased representation of the gene set and represent the 5', internal, and 3' portions of genes with equal frequency. GeneThresher subclone libraries contain virtually all of the genes in a plant genome, and preferentially represent exons and introns, promoters, non-coding RNAs, and simple sequence repeats, while minimizing the representation of interspersed repeats. DNA sequence obtained from GeneThresher libraries provides a robust view of the functional parts of the genome, and enables the design of DNA microarrays bearing the complete gene set of a plant. A combination of GeneThresher sequencing with low coverage BAC sequencing recovers rare methylated genes and anchors the sequence to a physical map. This is a plant-specific modification of the strategy used to sequence the rat genome (Gibbs *et al*., 2004).

After only a 1.1x coverage of the ~250 Mb gene space of the Sorghum bicolor genome, we have tagged more than three quarters of the Sorghum gene set using GeneThresher methylation-filtering technology (Bedell *et al*., 2005). The Sorghum genome, as with other plants, is composed of a large percentage of methylated repetitive elements that have expanded the genome several-fold. A comparison of the Sorghum gene set to the completed Rice and *Arabidopsis* genomes enhances our understanding of the gene family structure of Sorghum and adds to our knowledge of the syntenic relationships among the grasses. In addition to allowing the rapid cataloguing of the gene sets of large plant genomes, GeneThresher sequences can enhance and complement other genome discovery efforts.

References

Bedell, J. A., M. A. Budiman, A. Nunberg, *et al*. (2005) Sorghum genome sequencing by methylation filtration. PLoS Biol 3(1): e13.

Gibbs, R. A., G. M. Weinstock, M. L. Metzker, *et al*. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428(6982): 493-521.

Martienssen, R. A., P. D. Rabinowicz, A. O'Shaughnessy and W. R. McCombie (2004) Sequencing the maize genome. Curr Opin Plant Biol 7(2): 102-7.

Palmer, L. E., P. D. Rabinowicz, A. L. O'Shaughnessy, V. S. Balija, L. U. Nascimento, S. Dike, M. de la Bastide, R. A. Martienssen and W. R. McCombie (2003) Maize genome sequencing by methylation filtration. Science 302(5653): 2115-7.

Rabinowicz, P. D., K. Schutz, N. Dedhia, C. Yordan, L. D. Parnell, L. Stein, W. R. McCombie and R. A. Martienssen (1999) Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nat Genet 23(3): 305-8.