

FOR IMMEDIATE RELEASE

Contact:

Dr. Ben Goertzel

Biomind LLC

1405 Bernerd Place

Rockville MD 20851

Phone: 240-505-6518

<http://www.biomind.com>

[ben@biomind.com](mailto:ben@biomind.com)

**New Website [go.biomind.com](http://go.biomind.com), Powered by Machine Learning Technology, Identifies Functions of Hundreds of Previously Unclassified Genes**

Rockville, MD -- April 25, 2005 – The mapping of the human genome is complete, but the understanding of gene function much less so. The Gene Ontology database ([www.geneontology.org](http://www.geneontology.org)) classifies thousands of human genes regarding the biological processes, molecular functions and cellular components to which they pertain – but thousands more of the 25000 or so human genes remain unclassified. Completing the Gene Ontology is a slow and painstaking task, because at the present time the database is being constructed primarily by hand, with biologists continually creating new entries based on their knowledge and discoveries.

Thanks to advanced bioinformatics technology, there is now at least a partial solution to this problem. Maryland bioinformatics firm Biomind LLC has applied advanced machine learning algorithms to guess the appropriate Gene Ontology categories for hundreds of previously unclassified human genes. The Biomind software makes its guesses by combining information about gene sequences with data mined from a large corpus of microarray data, incorporating data from a variety of different experiments. The results of this analysis are available at [go.biomind.com](http://go.biomind.com).

The basic concept is that if a gene “behaves and looks similar” to the genes in some Gene Ontology category (say, “DNA repair” or “glycolysis”) then perhaps that gene belongs in that category. But, as Biomind CEO Dr. Ben Goertzel puts it, “getting from this basic concept to a working classification system with high accuracy involves a lot of computer science.” Statistical testing suggests that the system’s prediction will be roughly 80% accurate – not perfect, to be sure, but good enough to pay a lot of attention to. And this accuracy may be improved in the near future, without changing the algorithms at all – simply by utilizing a broader collection of microarray data.

According to Dr. Goertzel, the information on gene function available at [go.biomind.com](http://go.biomind.com) “should be interesting to anyone doing genomics research. It suggests probable functions for a whole lot of genes that were previously not understood at all, some that were never mentioned in the research literature, because the biologists haven’t gotten around to studying them yet. For instance,

there are dozens of genes that our software says are associated with programmed cell death – a key cause of aging – and that weren't previously known to have anything to do with anything. Biologists can look at this and then run traditional low-throughput experiments to understand exactly how why these genes are associated with aging. The microarray data and machine learning technology are being used to point scientific thinking and expensive wet lab work in the right direction.”

The task of understanding how genomes lead to the building of cells and organisms is far from complete. But we gather more and more relevant data each year – so much data that no human mind can fully comprehend it all. The data was collected by machines like microarrays, and it takes machine algorithms like Biomind's software to guide the human mind toward the proper interpretations.

**For additional information, including preprints describing detailed technical work related to [go.biomind.com](http://go.biomind.com), contact: Dr. Ben Goertzel --**  
Phone: 240-505-6518 -- [ben@biomind.com](mailto:ben@biomind.com)