

A Reference Database for *Medicago Truncatula* Genes, Proteins, and Metabolites

Dianjing Guo, Xingjing Li, Aejaaz Kamal, Olga Brazhnik, and Pedro Mendes

*Virginia Bioinformatics Institute, Virginia Tech
1880 Pratt Dr., Blacksburg, VA 24061*

Abstract

As a model plant for legumes as well as a rich source of natural products (such as flavonoids, isoflavonoids and triterpenes), *Medicago truncatula* (M.t.) is one of the subjects of current major US genomics initiatives. Nevertheless, data sources of gene, protein, and metabolite in relation to M.t. are very limited in publicly available biological databases.

Information about genes, proteins, and metabolites is usually distributed among multiple databases. Retrieval and organization of this information can be a laborious task. We present a relational database, B-Net, that is intended to gather information from multiple sources representing genes, proteins, metabolites, and biochemical reactions of M.t. This database represents known facts about the

biochemistry of M.t., classified according to the Gene Ontology. We anticipate this new resource to be particularly useful as a reference data set but also a qualitative proteome and metabolome database.

B-net consists of four main sections: GENE, PROTEIN, COMPOUND, and REACTION. All entries contain cross-reference information. Original metabolite and reaction data was imported from KEGG but filtered through a well-defined set of criteria. Information on genes and proteins from *Medicago truncatula* were gathered from public resources such as the published literature, and EST databases. The database schema supports adding data from any other species, and we foresee reusing it to create references for other organisms (the informatics system will be made available).