Generalized Logical Networks of Metabolomic Interactions in Alfalfa and *Medicago truncatula*

Samuel D. Palmer[1], Curtis Luce[2], Omar Holguin[3], Champa Sengupta-Gopalan[4] , and Joe Song[5]

***Short Abstract* – A metabolomic approach was used to create a model on the effects of methionine synthesis and accumulation in legumes. Comparative metabolic profiles of soluble primary and secondary metabolites of alfalfa and *Medicago truncatula* were quantized into discrete variables prior to modeling the interactions. Using a *K*-th order generalized logical network modeling software, significant interactions among metabolites in alfalfa and *M. truncatula* were identified. These interactions are used to understand further the accumulation of Met-rich □-zein proteins.**

I. Purpose

Alfalfa (*Medicago sativa*) is an important forage legume providing quality protein and is low in methionine (Met), an important amino acid. A genetic engineering approach to increase the Met content of alfalfa is to express genes encoding for high Met protein. Seed storage proteins of corn, the □- (15kD) and □- (10 and 18kD) zeins, are very high in Met and are ideal candidates for introducing into alfalfa. The □-zein gene engineered behind the CaMV 35S promoter was introduced in alfalfa and *Medicago truncatula*, a model legume. Our analysis of these transformants shows a ~10 fold higher level of accumulation of □ -zein protein and transcript in the leaves of *M. truncatula* plants when compared to alfalfa □ -zein expressors. Our hypothesis is that the two *Medicago* species differ with regards to the amino acid composition and in the rate of synthesis of Met rich proteins. A metabolomic approach was used to create a model on the effects of methionine synthesis and accumulation in legumes. Comparative metabolic profiling of soluble primary and secondary polar metabolites in *M. truncatula* and alfalfa was carried out by GC-MS, UPLC and LC-MS.

II. Quantization of continuous random variables

When the quantization of continuous random variables into discrete random variables is performed, it is important to preserve interactions between variables and avoid creating interactions that are not present in the original data. We attempt to maintain the data's original distribution by maximizing the log likelihood of the quantization grid. This approach allows us to quantize variables based on all dimensions unlike a single dimensional quantization using *k*-means or the mutual information pairwise approach presented by Hartemink [1].

III. Modeling interactions among metabolites in alfalfa and medicago truncatula

The *K*-th order generalized logical network modeling software determines significant interactions among metabolites in alfalfa and *M. truncatula*. Using inference of the *K*-th order the temporal information in the data sets provides a basis to examine interactions with metabolites using the generalized logical network model. In a generalized logical network, a logical function, associated with each metabolite as a node, describes its behavior dictated by some other influential microbes. The optimal logics at each metabolite node in the network will be searched so that they best explain the observed data. Determination of an optimal logic will involve parent node selection and truth-table generation. The maximum number of parents is set to a given number. If the current node shows consistent behavior during transition from one state to another given the parent nodes, then the parent nodes will be kept. The actual goodness of the transition will be calculated using the chi-square test.

IV. Conclusion

These analyses revealed a number of metabolic characteristics which are discretely associated with the fate of Met between species. Particular interest was focused on those metabolites associated with methionine biosynthetic pathway. An understanding of the basis for the differences between *M. sativa* and the model legume, *M. truncatula*, with regards to the accumulation of the Met-rich □ -zein protein will allow us to increase the Met-containing proteins in *M. sativa* using genetic engineering approaches.

References

[1] Alexander J. Hartemink. Principled computational methods for the validation and discovery of genetic regulatory networks. PhD thesis, Massachusetts Institute of Technology, 2001.

Acknowledgements: NSF CREST Center for Bioinformatics and Computational Biology at NMSU.

[1,2,5]Department of Computer Science, New Mexico State University, E-mail:[1]sdpalme@nmsu.edu,[2]cluce@cs.nmsu.edu, [5] joemsong@cs.nmsu.edu
[3,4]Department of Plant & Environmental Sciences, New Mexico State University, E-mail: [3]frholgui@nmsu.edu, [4]csgopala@nmsu.edu