

ABSTRACT

ADDITIONAL INFORMATION:

INTRODUCTION

Crown eukaryotes (metazoa, fungi and plants) have some mechanisms that are conserved in cell morphogenesis, such as membrane vesicle transport, while others, like extracellular structures or the roles of different Rho-like GTPases, do not. Several other cellular processes, such as cytokinesis, frequently employ conserved proteins to perform functionally distinct tasks. The molecular mechanisms that ensure the communication between the cytoskeleton and cell surface are subject to this, as evidenced by the fact that the data available from various genome projects has recently increased, making it possible to search for homologous components of known signaling and morphogenetic pathways. Such searches can yield experimentally testable hypotheses and general conclusions about the evolution of morphogenetic processes. Formins, also known as formin homology (FH) proteins, are involved in cellular and organismal morphogenesis of both metazoa and fungi. They play an important role in the establishment and maintenance of cell/tissue polarity, cytokinesis, and positioning of the mitotic spindle on cells. They interact with actin, profilin (which regulates insulin secretion), Rho-like GTPases, yeast Spa2 protein, septins, proteins containing SH3 or WW domains as well as dynein and microtubules. The localization of BNI1, a yeast formin homolog encoded by the cell periphery, involves positioning cortical actin patches towards different regions of the plasma membrane. This may necessitate some type of contact with the bloodstream and/or plasmalemma, although there is no evidence to support such an interaction. Furthermore, metazoan formins are considered to be cytoplasmic or nuclear proteins. The extent of formin function in plants is uncertain, although two Arabidopsis thaliana proteins with the formin-homology 2 (FH2) domain have been identified in recent times. Given that all known formins belong to a well-defined family, this class of proteins may be suitable for systematic genome sequencing. I am presenting here the results of an approach that has led to identification of putative plant formin genes, and also found evidence that recent eukaryotes may use the evolutionarily old formin domain in various contexts and ways ('modules' as defined by Hartwell & al.

CONCLUSION

The tree harvest procedure is a useful and efficient method of supervised learning from gene expression data. Its primary objective is to identify the additive and interaction structure of clusters of genes, in relation to an outcome measure. However, this technique requires significant samples for successful analysis and any other procedures with similar objectives. This method had some flaws as there were not enough samples for the actual data, so we decided to try tree harvesting on ever-larger gene expression datasets when they are available; this time, we proceeded with a stepwise approach whereby we used the sum and products of the average gene Expression of chosen clusters—then selecting models that are interpretable and biologically plausible. The gene expression of clusters' average could be used to construct alternative models, such as tree-based models or boosting methods (as demonstrated by Friedman et al.).

Additional data: Additional information about clustering from the lymphoma data-driven harvest model and other available data can be found in the online version of this article.