

Reviewer's report

Title: Uncovering Mechanisms of Transcriptional Regulations by Systematic Mining of Cis Regulatory Elements with Gene Expression Profiles

Version: 1 **Date:** 7 February 2008

Reviewer: Brett McKinney

Reviewer's report:

The authors make modifications to MotifADE, a method for identifying cis elements associated with differential gene expression, and they apply it to mouse cell data to find transcription factors (TFs) associated with regulation of genes involved in myogenesis. The authors modify MotifADE to test statistical significance of binding motifs in up-regulated and down-regulated genes, separately. This avoids the potential cancelling effect of up and down-regulated genes when they are tested together.

Major Essential Revisions:

Finding sequence motifs that modulate gene expression is an important problem, and the authors have made useful changes to MotifADE. I believe this manuscript would be improved if there were a better discussion of MotifADE in the Background. For example, make it clear to the reader that statistical tests for the difference in t-statistic means (as opposed to expression level means) are being performed, and explain why this is necessary. A reasonable description of MotifADE is given in the Methods, but I think it would help the reader to provide a more organized description and motivation near the beginning of the paper.

The authors also look for two-way synergistic effects between binding motifs in both up and down-regulated genes. The authors define a synergistic effect between TFs A and B using two significance criteria for genes in the AB intersection: that the expression of intersection genes shows a significant difference between (1) the A complement B set of genes and (2) the B complement A set of genes. This is an interesting extension but not much appears to be found. Is there any prior biological information that would suggest plausibility for the synergy found between TFs for down-regulated genes and that this synergy is more than a false positive? How might this be tested experimentally?

The authors perform some experimental validation, but the paper would be strengthened if predictions by other methods were compared. Which TFs would be missed if MotifADE were used without modification?

Minor Essential Revisions:

In the tables showing significant TFs, regarding the Occurrence column, it would be helpful to state in the caption out of how many total genes occurrence's were

found.

What is the availability of software used in this paper?

What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.