Reviewer's report

Title: Gene expression profiling in whole blood identifies distinct biological pathways associated with obesity

Version: 1 Date: 18 August 2010

Reviewer: robert koza

Reviewer's report:

The manuscript submitted by Ghosh et al describes a comprehensive microarray analyses of human blood to identify gene expression patterns associated with obesity. These types of analyses may be able to identify novel pathways that are associated with obesity in humans. Although highly descriptive, this manuscript demonstrates that analyses of whole blood, a relatively non-invasive approach compared to tissue biopsies, could provide some insight into the etiology of obesity. The experimental design and statistical analyses described in this manuscript are fundamentally sound, however, several issues regarding the interpretation of the data need to be addressed:

Major compulsory Revisions:

1. The interpretation of the gene expression data is questionable because the contribution of RNA from the red blood cell component and remaining cellular components of whole blood appears to differ between obese and lean individuals. This is indicated by the author’s own observation that the majority of genes that were upregulated in the obese subjects are most highly expressed in erythrocytes and reticulocytes. Several references to increased hemocrit in obese individuals are noted in the literature as well (i.e., Wysocki M et al Atherosclerosis 1991, 88(1):21-28). Although the reasons noted in the manuscript as to why PAXgene whole blood RNA was used instead of fractionated blood components are reasonable, is there a way to adjust gene expression data for hemocrit or for erythrocyte cell specific mRNA expression?

2. Why is there a difference in the number of individuals used for phenotypic characterization of subjects (i.e 12 per group; results, page 5 line 8/9 and Table 1) different from the number of individuals analyzed via Affymetrix microarray (i.e 17 per group; page 6 line 1; Figure 3)?

3. The data in Figure 1 are confusing. The X-axis indicates log2 ratio differences between the obese and lean cohorts; however, the bars are stacked and are no longer associated with the X-axis in a meaningful way. Would it be possible to re-design this figure to show a more direct correlation between the TaqMan and microarray analysis.

Level of interest: An article of importance in its field
Quality of written English: Acceptable

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

Declaration of competing interests:
I declare that I have no competing interests