Reviewer's report

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

Version: 1 Date: 23 August 2010

Reviewer: Antonio Camargo

Reviewer's report:

The manuscript titled “Gene expression profiling in whole blood identifies distinct biological pathways associated with obesity” by Ghosh et al addresses the differential gene expression in whole blood between obese and lean people. The approach is relevant and the work is well designed and carried out. However, there is a series of issues that should be considered to strengthen the quality and value of the manuscript before being accepted.

1. Is the question posed by the authors well defined? Yes
2. Are the methods appropriate and well described? Yes
3. Are the data sound?
4. Does the manuscript adhere to the relevant standards for reporting and data deposition? Yes
5. Are the discussion and conclusions well balanced and adequately supported by the data? No
6. Are limitations of the work clearly stated? No
7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
8. Do the title and abstract accurately convey what has been found? No
9. Is the writing acceptable? Yes

- Major Compulsory Revisions

1. Table 1 shows the characteristics of the subjects included in the current study (20 obese and 20 lean), but the data from 3 samples from each group were not used due to poor hybridization. Table should show the characteristic of the subject included in the data analysis (17 obese and 17 lean), where the results come from.

2. The study was carried out in men and women. A new analysis showing significant changes for each sex would be worth and would provide an accurate analysis regarding gender dimorphism.

3. The authors have confirmed gene expression changes by RT-PCR. What is
the rational of choosing genes? Are they representative of several levels of gene expression/pathways?

4. Results are not clearly described. Results section is actually a Results and Discussion section where the results are discussed in a long and speculative way. Further, Discussion section (the second discussion) focuses in the advantages of the use of whole blood instead to others strategies such as isolation of PBMCs. It should be taking into consideration that, as the authors comment, “compared to whole blood, several cell types including neutrophils, basophils, eosinophils, platelets, reticulocytes and erythrocytes are depleted in PBMCs which lead to loss of important transcription information”, the use of a very heterogeneous cell population could introduce noise and could mask gene expression differences in specific cells. Different amount of each one of the types cell between obese and lean could be also responsible of the different gene expression.

5. The manuscript would improve if the main issues are clearly splitted along the text, the gene expression changes found between groups (obese and lean) and the advantages and desvantages of the use of whole blood versus cell types isolation.

6. Although authors comment that “transcriptional differences are causal or caused”, sometimes along the manuscript is confused the relationship between gene expression and genetic variation. It should be clarified the fact that genetic variation could affect the expression of the genes, but gene expression can be modified because obesity independently of the genetic background.

Page 2. Abstract: Conclusion: “This study represents a novel approach for the elucidation of the genetic bases of obesity”

Page 4. “Second, as differences in gene expression are often driven by sequence variants in gene regulatory regions, our study provides a mechanism for the selection of obesity-associated candidate genes for the determination of possible regulatory sequence variants”.

Page 10: “Accordingly, we examined whether biological pathways implicated from gene-set enrichment analysis of the current study could provide a set of mechanism-based gene predictors that would be capable of predicting obese and lean subjects with high accuracy”

What predicting value have differentially expressed genes because the obesity? Would be these genes differentially expressed in prone people to develop obesity?

7. In order to clarify the differential gene expression in the 11 predictor genes associated with the phenotype, it would be worth look for sequence variants in these gene regulatory regions (e.g. promoter regions).

- Minor Essential Revisions
1. In results section, it is not mentioned the number of transcript detected in whole blood, and how many are up-regulated and down-regulated.

2. Quality of the RNA samples was not shown, have the authors assessed the RNA quality before the microarrays performing? Which method did they use?

- Discretionary Revisions
Nothing.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests