molecular systems biology

Integration of clinical data with a genome-scale metabolic model of the human adipocyte

Adil Mardinoglu, Rasmus Agren, Caroline Kampf, Anna Asplund, Intawat Nookaew, Peter Jacobson, Andrew J. Walley, Philippe Froguel, Lena M. Carlsson, Mathias Uhlen and Jens Nielsen

Corresponding author: Jens Nielsen, Chalmers University of Technology

Review timeline:

Submission date:	24 October 2012
Editorial Decision:	19 December 2012
Revision received:	07 January 2013
Accepted:	11 February 2013

Editor: Thomas Lemberger

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 19 December 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the two referees who agreed to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. They raise, however, several concerns on your work, which should be convincingly addressed in a revision of the work. The recommendations provided by the reviewers are very clear in this regard. In particular, a clear demonstration that the presented iAdipocytes1809 model outperforms previous human or adipocyte models (Recon1, iAB586) is required.

Please include a "Data availability" section in Materials & Methods section that specifies the links to the datasets and the model presented in this study.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favorable.

© European Molecular Biology Organization

Referee reports:

Reviewer #1 (Remarks to the Author):

Mardinoglu et al. construct a human adipocyte genome-scale metabolic model (GEM) by integrating an experimentally validated dataset of adipocyte expressed proteins generated from immunohistochemistry of human tissue samples with previously published GEMs and public databases on metabolism. The human adipocyte GEM was then used to create predictive models of metabolic differences between lean and obese subjects utilizing defined data sets from gene expression and clinical serum chemistry. The study results give rise to interesting new predictions on adipocyte metabolism and specific pathway relationships with clinical and therapeutic implications.

Reviewer #2 (Remarks to the Author):

In this work Mardinoglu et al. have manually reconstructed a comprehensive and, apparently functional, genome-scale metabolic model (GEM) of adipocyte metabolism. Although not explicitly stated, this seems to be an improvement on the previously reconstructed adipocyte-specific GEM (iAB586).
In order to more rigorously define protein abundance in their adipocyte model the authors carry out an ambitious immunohistochemisty-based proteomic assessment of human adipocytes, where they identify the presence or absence of proteins associated with 14,337 genes. In addition, the authors appear to have improved on the previous reconstruction in at least two more aspects. First, in this reconstruction the authors assigned proteins to cellular compartments based on enzyme localization data from Uniprot and the Human protein atlas. Second, the authors included 59 individual fatty acids as metabolites rather than relying on generic pools.
The authors then use this newly reconstructed adipocyte model to predict how known mRNA abundance differences in lean vs obese subjects from the SOS Sib Pair Study will impact the abundance of various metabolites (i.e. Reporter Metabolites). From this analysis, they identified three potentially interesting metabolites (androsterone, ganglioside GM2 and Heparan sulfate proteoglycans) that are predicted to have differential abundance in lean vs obese individuals. Interestingly, separate lines of investigation from other researchers have implicated these metabolites in obesity related adipose metabolism.
This work appears to be a very substantial contribution to the growing field of genome-scale metabolic modeling, however several limitations of the study design and manuscript organization make it difficult to assess the true merit of this work.

Comments/Questions:
1. Is this newly reconstructed model any better than the previously reconstructed adipose model (iAB586)?
A. In the results section where clinical data from (McQuaid et al. 2011) is used to "qualitatively" predict lipid droplet formation over 24 hours. First, I could not find the data that demonstrated that "iAdipocytes1809 successfully predicted the amount of LDs qualitatively over a 24 hour period". What was the metric of success? How does this compare to the Recon 1 GEM or iAB586?
B. Would Recon 1 or iAB586 have predicted the same Reporter Metabolites given the gene expression inputs from the SOS Sib Pair Study?
C. Can either iAdipocytes 1809 or iAB586 predict known metabolite flux differences in lean vs obese humans given expression inputs from the SOS Sib Pair Study?

2. The "Results" section could be condensed/rewritten for easier reading.
A. Parts of the results section read like a Methods section.
Examples: i. "...Expression console software from Affymetrix and the quality assessment was carried out using R..."
ii. "Annotations of high-resolution images were performed by certified pathologists..."
iii. "Localization information of reactions was inferred from the resources used for the model reconstruction..."
B. The last part of the last paragraph on page 13 starting with "It is known that mitochondrial acetyl-CoA plays a central role in different pathways in the mitochondria..."may be better in the Discussion

© European Molecular Biology Organization

section.

C. Third section of results, paragraphs starting with: "In order to outline...", & "Our analysis demonstrated...". It seems like this is a validation of the model, but the significance of the results are not easy to discern. Why was this analysis done? How does it move the story forward?

3. To generate the Reporter Metabolites for Male and Female, lean and obese, it appears that gene expression data was used from the SOS Sib Pair Study. Was the whole transcriptome pattern used in the model for each case or were only certain genes with differential expression used to analyze the model? In general this section of the results could be written a bit more clearly.

4. The authors state, "Global protein profiling of adipocytes found in the breast and soft tissue samples showed some differences". It is not clear how this impacted the reconstruction of the model.

1st Revision - authors' response 07 January 2013

Detailed description of the changes made in response to the referees

Reviewer #1 (Remarks to the Author):

Mardinoglu et al. construct a human adipocyte genome-scale metabolic model (GEM) by integrating an experimentally validated dataset of adipocyte expressed proteins generated from immunohistochemistry of human tissue samples with previously published GEMs and public databases on metabolism. The human adipocyte GEM was then used to create predictive models of metabolic differences between lean and obese subjects utilizing defined data sets from gene expression and clinical serum chemistry. The study results give rise to interesting new predictions on adipocyte metabolism and specific pathway relationships with clinical and therapeutic implications.

We wish to thank Reviewer #1 for detailed reading of our manuscript and finding the topic of our study of potential interest.

Reviewer #2 (Remarks to the Author):

In this work Mardinoglu et al. have manually reconstructed a comprehensive and, apparently functional, genome-scale metabolic model (GEM) of adipocyte metabolism. Although not explicitly stated, this seems to be an improvement on the previously reconstructed adipocyte-specific GEM (iAB586).
In order to more rigorously define protein abundance in their adipocyte model the authors carry out an ambitious immunohistochemisty-based proteomic assessment of human adipocytes, where they identify the presence or absence of proteins associated with 14,337 genes. In addition, the authors appear to have improved on the previous reconstruction in at least two more aspects. First, in this reconstruction the authors assigned proteins to cellular compartments based on enzyme localization data from Uniprot and the Human protein atlas. Second, the authors included 59 individual fatty acids as metabolites rather than relying on generic pools.
The authors then use this newly reconstructed adipocyte model to predict how known mRNA abundance differences in lean vs obese subjects from the SOS Sib Pair Study will impact the abundance of various metabolites (i.e. Reporter Metabolites). From this analysis, they identified three potentially interesting metabolites (androsterone, ganglioside GM2 and Heparan sulfate proteoglycans) that are predicted to have differential abundance in lean vs obese individuals. Interestingly, separate lines of investigation from other researchers have implicated these metabolites in obesity related adipose metabolism.

© European Molecular Biology Organization

This work appears to be a very substantial contribution to the growing field of genome-scale metabolic modelling, however several limitations of the study design and manuscript organization make it difficult to assess the true merit of this work.

We wish to thank Reviewer #2 for detailed reading of our manuscript and providing constructive comments. Please find below the answers to the points raised.

Comments/Questions:

1. Is this newly reconstructed model any better than the previously reconstructed adipose model (iAB586)?

The newly reconstructed *iAdipocytes1809* is more comprehensive and detailed than the previously reconstructed adipose model *iAB586* and it has several properties which speak strongly in its favour. Firstly, it is significantly larger in scope not only in terms of number of reactions/genes/metabolites but also in which parts of metabolism that is included. Secondly, *iAdipocytes1809* was validated for its ability to perform 250 metabolic functions. This, together with a very thorough quality control process, ensures that it is a connected and functional model. Thirdly, we have based the model on very extensive direct protein-level evidence generated within this project. This in itself represents a considerable improvement over existing models. Fourthly, *iAdipocytes1809* is formulated using 59 individual fatty acids rather than relying on generic pools. This enables mapping and integration of lipidomics data in a way that is not possible using *iAB586*. Finally, productions of all metabolites in the model were checked with minimum input to the model to ensure the connectivity. This is now clarified in the discussion part of the paper.

A. In the results section where clinical data from (McQuaid et al. 2011) is used to "qualitatively" predict lipid droplet formation over 24 hours. First, I could not find the data that demonstrated that "iAdipocytes1809 successfully predicted the amount of LDs qualitatively over a 24 hour period". What was the metric of success? How does this compare to the Recon 1 GEM or iAB586?

One of the improvements in *iAdipocytes1809* over the previously published adipocyte GEM (iAB586) is that is allows for modelling of the formation of lipid droplets (LDs). Simulation of LD formation could not be compared with *Recon 1* or *iAB586* since those models do not contain that feature.
Our metric of success was that the model simulations were able to capture the known behaviour of decreased LD dynamics in obese patients. We have now clarified this in the manuscript by inserting the following text:
"Based on measurements of the uptake of glucose and TAG and the release of NEFAs over a 24 hour period (Figure 5a) we simulated the change in LD size (Figure 5b, Dataset 9a). We found from our simulations that lean subjects have large dynamic changes in LD formation compared with obese subjects, which is in agreement with experimental data (Arner et al., 2011). Furthermore, we predicted a lower acetyl-CoA production in obese subjects, as shown in Figure 5c (Dataset 9b)."

B. Would Recon 1 or iAB586 have predicted the same Reporter Metabolites given the gene expression inputs from the SOS Sib Pair Study?

We have included a comparison of Reporter Metabolites calculated using *iAB586* with those calculated using *iAdipocytes1809*. In short, the results are similar in the parts of metabolism that are covered by *iAB586*, but the larger scope of *iAdipocytes1809* enabled identification of several additional Reporter Metabolites which were highly interesting for our study. We chose to compare to the published adipocyte model rather than to Recon1, since tissue-specific models are better suited, because they filter out the effects of having multiple cell types in the transcription data. A supplementary figure (Figure S8), which includes the analysis of reporter metabolites using *iAB586,* as well as the following additional text has been included in the manuscript.
"In order to illustrate the improvement of *iAdipocytes1809* over the published *iAB586,* the Reporter Metabolites were also calculated for male and female obese subjects by using *iAB586* (Figures S8). Reporter Metabolites involved in the mitochondrial dysfunction as well as different amino acids were identified to be similar to the Reporter Metabolite analysis using *iAdipocytes1809*. However, *iAB586* could not detect several of the most significant and in our view most interesting Reporter Metabolites identified when using *iAdipocytes1809* due to the increase in number of reactions,

© European Molecular Biology Organization

metabolites and genes (see Discussion). Thus, the Reporter Metabolite analysis with *iAdipocytes1809 and iAB586* provides an unbiased confirmation that *iAdipocytes1809* represent a significant advancement of the adipocyte metabolic network compared with *iAB586*."

C. Can either iAdipocytes 1809 or iAB586 predict known metabolite flux differences in lean vs obese humans given expression inputs from the SOS Sib Pair Study?

We are not aware of any measured internal fluxes in adipocytes but as discussed in the paper we did use uptake and secretion rates to calculate internal fluxes using a random sampling algorithm. We believe that the quality of iAdipocytes1809 will allow to calculate the fluxes from transcription data e.g. using the algorithm of Shlomi et. al. (2008), Nature Biotechnology, 26, 1003 - 1010.

2. The "Results" section could be condensed/rewritten for easier reading.
A. Parts of the results section read like a Methods section.
Examples: i. "...Expression console software from Affymetrix and the quality assessment was carried out using R..."

This part has moved to Materials and methods as:

"The normalization of the microarrays was carried out using the Expression Console software from Affymetrix and the quality assessment was carried out using R Statistical and Computing language and the Bioconductor software (Gentleman et al., 2004)."

ii. "Annotations of high-resolution images were performed by certified pathologists..."

This part has moved to Materials and methods as:

"Annotation of high-resolution images was manually performed by certified pathologists. Relative expression was indicated with four different color codes ranging from strong (red), moderate (orange), weak (yellow), and no expression (white) (Kampf et al., 2004)."
iii. "Localization information of reactions was inferred from the resources used for the model reconstruction..."

This part has moved to Materials and methods as:

" Localization information of proteins was inferred from manually curated Uniprot data (Apweiler et al., 2011) and recently generated HPA data on intracellular localization of proteins (Lundberg and Uhlen, 2010)."

B. The last part of the last paragraph on page 13 starting with "It is known that mitochondrial acetyl-CoA plays a central role in different pathways in the mitochondria..."may be better in the Discussion section.

This part has moved to Discussion as:

"Mitochondrial acetyl-CoA plays a central role in different pathways in the mitochondria and it reacts with oxaloacetate to form citrate, which can be transported from the mitochondria to the cytosol where it is participating in FA synthesis (Dean et al., 2009). Acetyl-CoA derived through other principal sources, including degradation of amino acid and ketone bodies and fatty acid oxidation processes are insufficient for FA synthesis. Increasing the acetyl-CoA concentration and eventually FA synthesis in adipose tissue of obese subjects results in whole body regulation of metabolism, including stimulation of muscle insulin action and suppression of hepatosteatosis, as reported by Cao et al. (2008). "

C. Third section of results, paragraphs starting with: "In order to outline...", & "Our analysis demonstrated...". It seems like this is a validation of the model, but the significance of the results are not easy to discern. Why was this analysis done? How does it move the story forward?

We agree with the comment of the reviewer and it has been clarified in the manuscript as:
"Although the generated proteome data for adipocytes covers the entire set of cellular processes (e.g. signaling, metabolism, cell cycle), GEMs are applicable only for the study of metabolism. To get a general overview of the global changes between obese, overweight and lean subjects the enrichment of differentially expressed genes was calculated for KEGG pathways (Figures S1 and S2) and for

© European Molecular Biology Organization

biological process Gene Ontology (BP:GO) terms (Figures S3 and S4). This was done using DAVID and for male and female subjects (Huang et al., 2009). In order to check the correlation of the genome-wide transcription data of SAT with the proteome data, the enrichment of differentially expressed genes in male and female obese subjects was calculated for the most significant KEGG pathways from the analysis of the proteome data (Figure 2c). This was done as the transcriptome data for SAT represent not only adipocytes but also other cell types; including immune cells and preadipocytes linked with different BMIs.
The analysis demonstrates that some of the metabolic and signaling pathways found to be enriched in adipocytes based on the proteome data also show significant changes in gene expression between lean and obese subjects, both in males and females. Similarly, we find that BP:GO terms that are enriched based on the proteome data (assessed with DAVID (Huang et al., 2009)) also show enrichment based on the transcriptome data (Dataset 3). Thus, enriched BP:GO terms such as post-translational protein modification, cellular protein metabolic process, lipid metabolic process, cellular lipid metabolic process and FA metabolic process are found both from the adipocyte-specific proteome data and from comparison of expression data for lean and obese subjects."

3. To generate the Reporter Metabolites for Male and Female, lean and obese, it appears that gene expression data was used from the SOS Sib Pair Study. Was the whole transcriptome pattern used in the model for each case or were only certain genes with differential expression used to analyze the model? In general this section of the results could be written a bit more clearly.

Reporter Metabolites performs a statistical test to see if there is a significant change in expression in the genes encoding enzymes "associated" with the metabolite, i.e. the enzymes use the metabolite as a substrate or product. The analysis is therefore performed using all genes present in the model. This has been clarified in the manuscript.

"Modeling using *iAdipocytes1809* can be applied to predict metabolic states under various perturbations, study regulation of the adipocytes, identify potential therapeutic targets and discover novel biomarkers for the development of more effective therapies. Here, we used the model to identify Reporter Metabolites (Patil and Nielsen, 2005) of male and female obese subjects compared to lean subjects using gene expression data obtained from the SOS Sib Pair Study. Reporter Metabolites are metabolite nodes in the metabolic network around which there are significant transcriptional changes. Here, 20 statistically significant Reporter Metabolites are presented for up and down regulated genes in male and female obese subjects through the employment of *iAdipocytes1809* (Figure 6). The most significant results from our Reporter Metabolites analysis for up and down regulated genes are correlated with the KEGG pathways enrichment results of significantly expressed genes in the obese subject groups (Figures S1 and S2)."

4. The authors state, "Global protein profiling of adipocytes found in the breast and soft tissue samples showed some differences". It is not clear how this impacted the reconstruction of the model.

We wanted the model to be a general representation of adipocyte metabolism, and as such it includes all the proteins expressed in any of the three tissues. This has been clarified in the manuscript.
"Since adipocytes obtained from breast and two different soft tissues were used for the protein profiling, there was some variation between the samples. In order to estimate the effect of these variations on the functionality of the adipocytes we used the functional annotation tool DAVID to calculate the enrichment in KEGG pathways (Huang et al., 2009). The results are presented in Figure 2c for breast, the two types of soft tissues, as well as for all proteomics data used for the GEM reconstruction. The analysis demonstrated that for all the adipocyte-specific proteome data there is enrichment in terms of metabolic pathways including FA metabolism, elongation and biosynthesis as well as major signaling pathways including adipocytokine, insulin, neurotrophin and PPAR signaling pathways. The figure shows that the samples from different tissues exhibit some differences, but that the overall pattern is similar. We therefore decided to reconstruct a general GEM for adipocytes by incorporating all proteins that were expressed in any of the three tissues."

© European Molecular Biology Organization