Modelling Metabolic Pathways Involved in the Pathogenesis of Non-Alcoholic Fatty Liver Disease

Adviti Naik*, Aleš Belič

Faculty of Electrical Engineering, University of Ljubljana, Tržaška Cesta 25, 1000 Ljubljana *adviti.naik@fe.uni-lj.si

Simulation Notes Europe SNE 23(2), 2013, 71 - 76
DOI: 10.11128/sne.23.tn.10183
Received: Jan. 10, 2013 (Selected MATHMOD 2012 Postconf. Publ.); Revised Accepted: June 10, 2013;

Abstract. Abstract: Non-alcoholic fatty liver disease (NAFLD) is a poorly understood complex disorder that occurs at a high frequency in Western populations with unhealthy dietary and lifestyle habits. Systems biology tools may provide further insight into the multidimensional nature of NAFLD. A dynamic semiquantitative model of the metabolic and signaling pathways suggested to be important in the pathogenesis of NAFLD has been generated using an object-oriented library of components based on differential equations. Initial model validation procedures to draw a comparison between experimental data and model simulations has suggested a close correlation between experimental models and the in silico network. Furthermore, the stiffness of multisubstrate reactions indicates the presence of a balance between pathway fluxes within the network. The model aims to identify novel mechanisms of NAFLD pathogenesis and regulatory components, thus providing a basis for experimental hypotheses.

Introduction

Non-Alcoholic fatty Liver Disease (NAFLD) is the most common chronic liver disease in Western countries affecting approximately 20-30% of the general population and 70-80% of obese populations. It encompasses a wide disease spectrum ranging from benign steatosis, non-alcoholic steatohepatitis (NASH) that is characterised by inflammation, hepatocellular carcinoma (HCC) and ultimately, cirrhosis ([2]). The accumulation of triglycerides in hepatocytes is a hallmark of hepatic steatosis. The major contributor of hepatic triglycerides is de novo lipogenesis and the plasma triglyceride pool derived from the adipose tissues ([3]).

Aberrant perturbations resulting in imbalances between lipid transport, oxidation, synthesis and storage contribute to the pathogenesis of NAFLD ([9]).

NAFLD is the hepatic manifestation of the metabolic syndrome as it encompasses several features of the metabolic syndrome such as dyslipidemia, obesity, insulin resistance, hyperglycaemia and hypertension. Although several molecular mediators and genetic polymorphisms have been implicated in the initiation and progression of NAFLD, this disease remains poorly understood due to inter-individual variations in disease pathogenesis, lack of representative animal models and non-invasive diagnostic methods ([1]). Polymorphisms in various genes, such as PNPLA3 that encodes adiponutrin ([14],[15]), have been identified in association with NAFLD, however a majority of these remain to be replicated in diverse and significantly powered populations and/or do not have a significant contributing effect. Western lifestyles, including insufficient physical activity and the consumption of high-fat and high-sugar diets have a major impact on disease pathogenesis. Hence, the complexity of NAFLD arising from the close interaction between genetic and environmental factors has proved to be difficult to dissect to fully understand its manifestations.

Biological systems can be considered as fundamentally nonlinear systems that operate at nominal equilibrium states, which on perturbation results in the activation of mechanisms to restore the equilibrium state or to reach a different state of equilibrium. Systems biology provides an integrated tool to study non-linear biological physiology using a mathematical modelling approach. It provides a rapid, cheap and repeatable analysis methodology in conjunction with experimentation in order to understand molecular biology. Systems biology has been increasingly utilized to study metabolic networks, especially in prokaryotes ([4]) due to their unicellular nature and the absence of complex metabolic and signalling pathways .

International collaborations are working towards ambitious projects such as the Physiome Project, which aims to provide a framework for modelling biochemistry, biophysics and anatomy aspects of the human body ([7]). However, the complexity of metabolic pathways in eukaryotes renders this task more challenging and as a result only a few reconstructions have been built for multicellular organisms, which are either static networks ([10]) or models focussing on a single cell type ([5]), hence not accounting for tissue-tissue interactions, which is an important deregulated mechanism in NAFLD. This article aims at describing the generation of a dynamic in silico model integrating the NAFLD metabolic and signalling pathways, whilst focussing at various hierarchical levels of metabolites and proteins regulated at the level of gene expression and post-translation.

1 Generation of the Metabolic Network

1.1 Modelling approach

An object-oriented modelling and simulation programme, Dymola 7.4, was utilised to curate a majority of the metabolic reactions implicated in the pathogenesis of NAFLD (Appendix Figure5). The metabolic network was generated utilising a systems biology library of components, wherein each component functions as a biological component on the basis of differential equations assigned to it.

1.2 Curation of reactions

The metabolic network includes enzyme-catalysed reactions, which are regulated by various transcriptional factors and molecular mediators at the transcriptional and post-translational levels. Evidence for the incorporation of these reactions within the metabolic network was derived from literature searches (n = 470), the Kyoto encyclopaedia of genes and genomes (KEGG, http://www.genome.jp/kegg/) and the Reactome (www.reactome.org) databases. The reversibility and stoichiometry of the reactions were maintained only in the presence of evidence from the literature.

The major compartments modelled into the network include the liver, adipose tissue, peripheral tissue, pancreas and macrophages. Within each of these compartments, pathways involved in the metabolism of glucose and lipids, such as glycolysis, gluconeogenesis, citric acid cycle, pentose phosphate pathway, de novo lipogenesis, β-oxidation, lipolysis, amino acid metabolism and ketone body synthesis have been included (Appendix Figure6). Moreover, transport and excretory reactions between tissues via the blood have also been portrayed, thus ensuring a near- physiological distribution of metabolites. A total of 160 reactions involving 150 metabolites are depicted in the complete model.

1.3 Assigning flux distributions

At branch points in the network, approximate values of the metabolite flux distribution into each of the pathway branches were assigned based on a mass isotopomers study and flux analysis in a human hepatocellular carcinoma cell line, HepG2, using labelled 13-C glucose ([6],[11])). Reactions with undetermined fluxes are assigned nominal flux values which allow stable simulations of the model.

2 Model Validation

The completed model was put through validation procedures, whereby model simulations of published experimental situations were compared to the experimental data to ensure correspondence. The value of all model variables at steady state has been defined as 1. A disagreement between model simulations and experimental data indicates errors or omitted components, which then need to be identified through a close evaluation of the network and further literature mining.

2.1 Simulation of fasting conditions

Fasting results in decreased glucose levels in the blood and hence, decreased insulin and increase in glucagon secretion. Changes in these blood parameters signals an upregulation of gluconeogenesis (synthesis of glucose), a decrease in glycogen stores due to conversion into glucose, decrease in de novo lipogenesis due to unavailability of glucose substrate, an increase in β-oxidation to serve as an energy source in the absence of sufficient glucose and an increase in the level of fatty acids in the blood ([8]).

In the in silico metabolic model if the component defining the source of mass flow of glucose in reduced by 10-fold, the levels of glucagon increases by 2-fold and insulin decreases by 4.5-fold (Figure1). The simulation of other variables using these parameters depicting fasting results in a close correlation between experimental data and model simulations. A 1.2-fold increase in glucose-6-phosphate levels, a crucial enzyme in gluconeogenesis, a 12.5-fold decrease in glycogen stores,

a 1.25-fold decrease in fatty acid synthase (FAS) levels, a rate-limiting enzyme in de novo lipogenesis, a 1.15-fold increase in carnitine acyl transferase 1 (CPT-1) levels, the first enzyme catalysing β-oxidation and a 1.2-fold increase in fatty acid levels in blood were observed (Figure2).

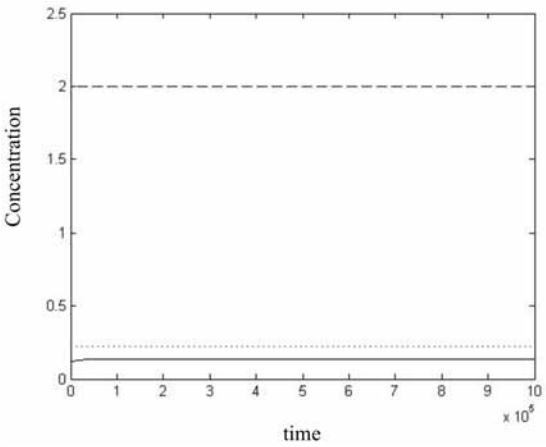

Figure 1: Modelling fasting conditions: Solid line- serum glucose levels, dotted line- serum insulin levels, dashed line- serum glucagon levels.

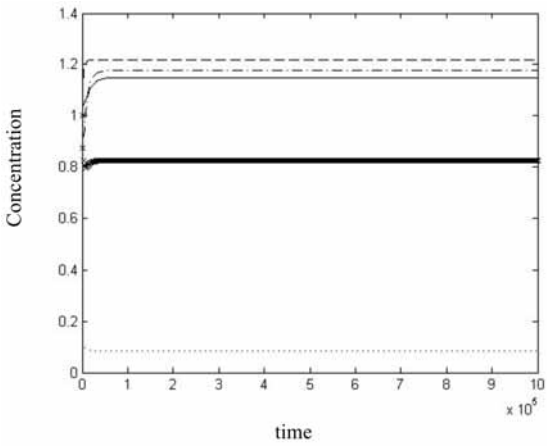

Figure 2: Simulation of fasting conditions: Solid line- Carnitine acyl-transferase (CPT-1), dotted line- hepatic glycogen stores, dashed line- glucose-6-phophatase, dash dotted line- serum fatty acid levels, thick line- Fatty acid synthase (FAS).

2.2 Simulation of the absence of Stearoyl CoA Desaturase-1 (SCD-1)

Stearoyl CoA desaturase (SCD-1) is a crucial lipogenic enzyme that converts saturated fatty acids into monounsaturated fatty acids, which is the major substrate for triglyceride and cholesterol ester biosynthesis ([12]). Data from an experimental mouse model with null expression of the SCD-1 isoform due to targeted disruption of the gene on a high-fat diet, indicated a downregulation of lipogenic enzymes, decreased accumulation of hepatic triglycerides and plasma leptin and an upregulation of β-oxidation and plasma ketone bodies ([13]).

The equation describing enzyme levels in the metabolic model is as follows:

$$f = C.fi - k \cdot Q \quad (1)$$

where, f is the real value describing the utilization of the enzyme in the reaction, $C.fi$ is the initial utilization of the enzyme at steady state, k is the real value describing the rate of enzyme degradation and Q is the concentration of the enzyme.

The disruption of SCD-1 in the in silico model was simulated by increasing the value of k from its initial value of 0.01 to 0.9999. The simulation other variables indicated a 3.3-fold decrease in hepatic triglycerides, 1-fold decrease in plasma leptin and 1.1-fold downregulation of FAS. Furthermore, a 1.47-fold decrease in sterol regulatory element-binding protein 1-c (SREBP1-c), which is a crucial transcription factor regulating de novo lipogenesis, was observed. B-oxidation and plasma ketone bodies were upregulated by 1.7-fold and 1.2-fold, respectively, as observed on simulating the CPT-1 enzyme and β-hydroxybutyrate and acetoacetate plasma concentrations (Figure3). Thus, the model simulations were in close correlation with experimental data. It is important to note that, as the physiological concentration of the metabolites and enzymes is not incorporated into the model, the model simulations provide only relative and not absolute values of variable changes.

2.3 Stiffness of fluxes through reactions with $\geq$2 substrates

Model simulations testing varying values of fluxes through network branch points indicates that only a limited set of flux values can be defined for branch point reactions that involve ≥ 2 metabolite substrates. This limited tolerance can be explained for reactions whose substrates are generated via separate network branches as the reaction depends on reaching equilibrium between different pathway branch flux values.

The example of glycerol-3-phosphate acyl-transferase has been described here to demonstrate this phenomenon. Glycerol-3-phosphate acyl-transferase

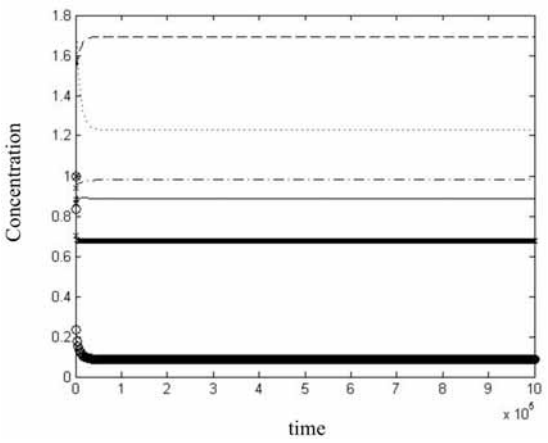

Figure 3: Simulation of Stearoyl CoA Desaturase-1 null conditions: Solid line- Fatty acid synthase (FAS), dotted line-serum ketone body levels, dashed line- Carnitine acyl-transferase (CPT-1), dash dotted line- serum leptin, thick line- SREBP1c, double thickness line- hepatic triglycerides.

catalyzes the biosynthesis of lysophosphatidic acid, the first intermediate in the formation of triglycerides. The generation of lysophosphatidic acid depends on the concentration of 2 substrates, glycerol-3-phosphate and fatty acyl CoAs (Figure4). Hence, the fluxes from the glycolysis pathway as well as the fatty acid synthesis pathway govern the formation of lysophosphatidic acid. The flux through this reaction in the model has been assigned a percentage of 5% of the fatty acyl CoA pool.

On varying the reaction flux percentage to any value >10%, the model reached instability, thus suggesting an intolerant focal point within the network. This observation is true for the majority of reactions involving $\geq$2 substrates. Interestingly, these reactions are catalyzed be highly-regulated enzymes, thus providing an explanation for the evolutionary development of stringent regulatory mechanisms to ensure the maintenance of the flux within narrow limits, thus preventing disruptions of rate-limiting metabolic reactions. Hence, enzymes catalysing these reactions may signify novel molecular mediators that may play a crucial role in the pathogenesis of NAFLD.

3 Discussion

The multi-dimensional nature and complexity of NAFLD has resulted in the poor understanding of the disease pathogenesis. As the progression of NAFLD

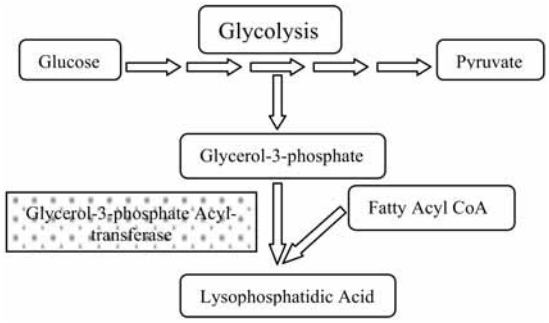

Figure 4: Diagrammatic scheme of the reaction catalysed by glycerol-3-phosphate acyl-transferase.

appears to be a result of multiple perturbations or predisposition factors, the application of systems biology may provide an additional dimension to the dissection of the disease. The generation of the metabolic network, focussing on tissue-tissue interaction as well as intra-tissue signalling, aims to provide a framework to test various hypotheses of NAFLD initiation and progression by in silico simulations and thus prove as a basis for further experimental analysis. The object-oriented differential equation -based model provides an easy to manipulate, dynamic and visual network, which may be made available to the wider scientific community for utilization as a systems biology tool. Due to the interconnected nature of the metabolic and signalling pathways that have been identified in relation to NAFLD, the model size and complexity depicts the intricacy of the disease network.

The future aim of this project is to substantially simplify the network to its core components for further transparency, whilst still maintaining the precision and full functionality of the model. An initial step in model simplification is to reduce linear pathways into single components and clump similarly regulated reactions or reactions for which regulatory mechanism have not yet been identified. In spite of the drawback that the model generation is based solely on evidence from the literature, model simulations that do not correspond to experimental data indicate the presence of unidentified components in the system. Hence, permutations and combinations of various regulatory components for specific reactions may provide a basis for experimental hypotheses to identify novel regulatory mechanisms. This method will also allow further model validation in order to generate a robust model of NAFLD pathogenesis.

In order to enable the generation of more physiological and quantitative simulations, flux data is required

Appendix

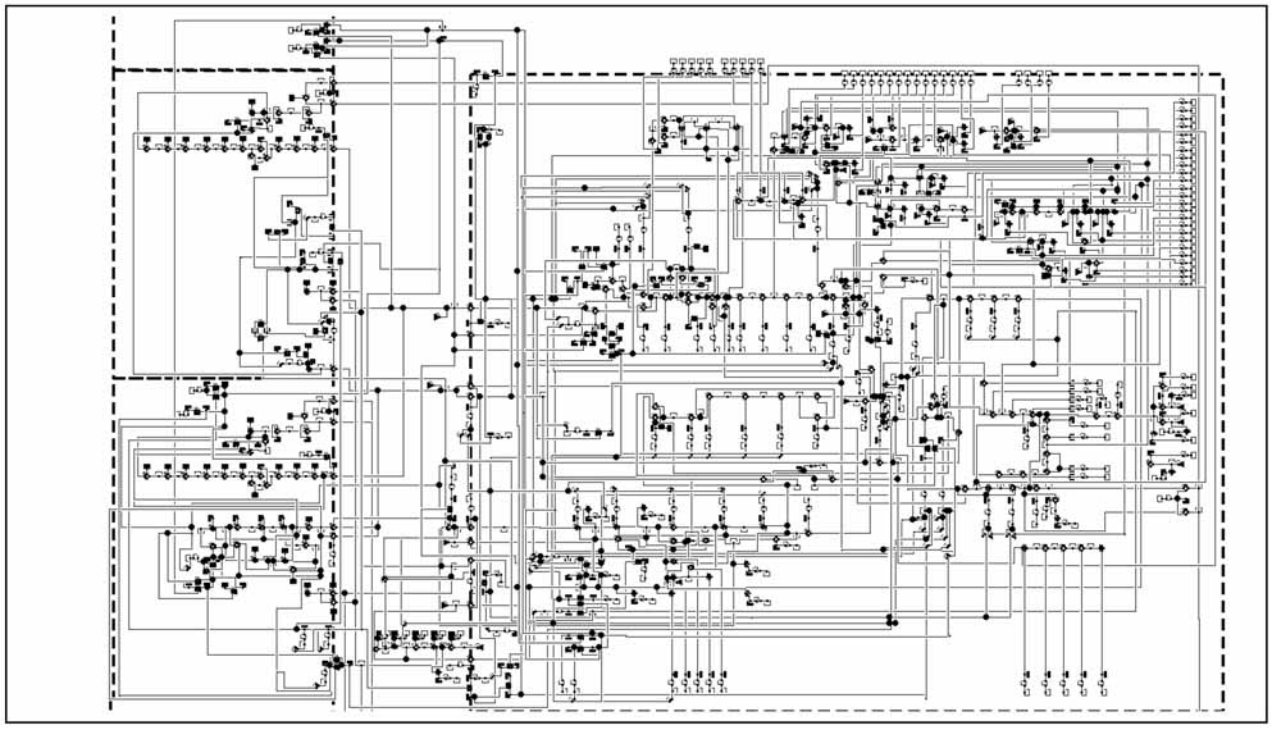

Figure 5: Dymola model of NAFLD metabolism. Dashed lines- tissue compartments, straight lines- regulatory signalling pathways and metabolic reactions, black and white shapes- systems biology components (metabolites, proteins, regulatory blocks etc.).

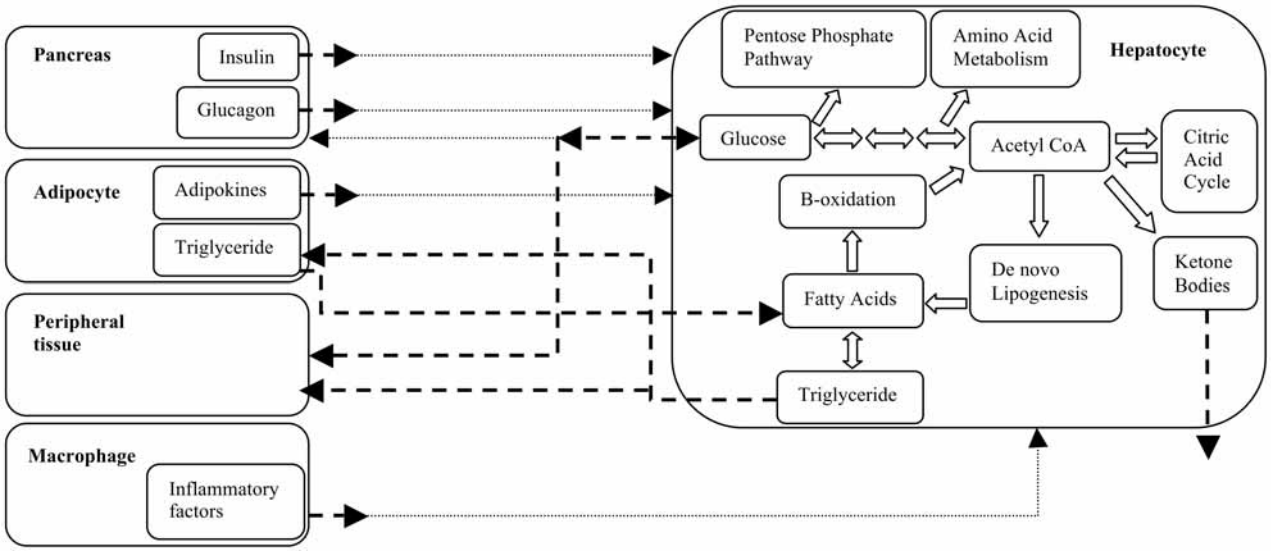

Figure 6: Schematic diagram of NAFLD metabolic model. Dashed lines- transport pathways, dotted lines- regulatory signalling pathways.

to be incorporated within the model. The observation of the stiffness of reactions involving ≥ 2 substrates indicates the significance of determining the fluxes in metabolic pathways. Moreover, it also suggests that the balance between fluxes in various interconnected pathways within the network is stably maintained. Hence, it may be hypothesised that this balance is disrupted in NAFLD due to various dietary perturbations or genetic predispositions, thus resulting in the accumulation of lipid intermediates. The simulation of probable molecular disruptions or dietary effects within the model may suggest novel mechanisms of lipid accumulation and inflammatory responses, which are characteristic of NAFLD. Furthermore, metabolic flux analysis comparing normal and steatotic experimental models will aid in the identification of pathways with disturbed fluxes in NAFLD pathogenesis.

The diversity in the manifestation of the NAFLD disease spectrum has proved as a major hindrance in effective treatment strategies. Liver biopsy, an invasive and expensive procedure, is the golden standard for diagnosing the progression of NAFLD in patients. The in silico NAFLD model also aims at identifying non-invasive serum markers that correspond to intrinsic perturbations and hence, can predict the transition of NAFLD from relatively benign to more detrimental forms.

References

[1] Cohen JC, Horton JD, Hobbs, HH. Human Fatty Liver Disease: Old Questions and New Insights. *Science*. 2011; 332: 1519-1523.

[2] De Alwis NM, Day CP. Non-alcoholic fatty liver disease: the mist gradually clears. *Journal of Hepatology*. 2008; 48(1): 104-112.

[3] Donnelly KL, Smith CI, Schwarzenberg SJ, Jessurun J, Boldt MD, Parks EJ. Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *Journal of Clinical Investigation*. 2005; 115: 1343-1351.

[4] Feist AM, Palsson BO. The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nature Biotechnology*. 2008; 26: 659-667.

[5] Gille C, Bolling C, Hoppe A, Bulik S, Hoffmann S, Hubner K, Karlstadt A, Ganeshan R, Konig M, Rother K, WEidlich M, Behre J, Holzhutter HG. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Molecular Systems Biology*. 2010; 6: 411.

[6] Hofmann U, Maier K, Niebel A, Vacun G, Reuss M, Mauch K. Identification of metabolic fluxes in hepatic cells from transient 13C-labeling experiments: Part I Experimental observations. *Biotechnology and Bioengineering*. 2008; 100: 344-354.

[7] Hunter P, Robbins P, Noble D. The IUPS human physiome project. *Pflugers Archiv-European Journal of Physiology*. 2002; 445: 1-9.

[8] Kersten S, Seydoux J, Peters JM, Gonzalez FJ, Desvergne B, Wahli W. Peroxisome proliferator-activated receptor alpha mediates the adaptive response to fasting. *Journal of Clinical Investigation*. 1999; 103: 1489-1498.

[9] Kohjima, M, Enjoji M, Higuchi N, Kato M, Kotoh K, Yoshimoto T, Fujino T, Yada M, Yada R, Harada N, Takayanagi R, Nakamuta M. Re-evaluation of fatty acid metabolism-related gene expression in nonalcoholic fatty liver disease. *International Journal of Molecular Medicine*. 2007; 20: 351-358.

[10] Ma H, Sorokin A, Mazein A, Selkov A, Selkov E, Demin O, Goryanin I. The Edinburgh human metabolic network reconstruction and its functional analysis. *Molecular Systems Biology*. 2007; 3: 135.

[11] Maier K, Hofmann U, Reuss M, MauchK. Identification of metabolic fluxes in hepatic cells from transient 13C-labeling experiments: Part II Flux estimation. *Biotechnology and Bioengineering*. 2008; 100: 355-370.

[12] Miyazaki M, Kim YC, Gray-Keller MP, Attie AD, Ntambi JM. The biosynthesis of hepatic cholesterol esters and triglycerides is impaired in mice with a disruption of the gene for stearoyl-CoA desaturase 1. *Journal of Bioliogical Chemistry*. 2000; 275: 30132-30138.

[13] Ntambi JM, Miyazaki M, Stoehr JP, Lan H, Kendziorski CM, Yandell BS, Song Y, Cohen P, Friedman JM, Attie AD. Loss of stearoyl-CoA desaturase-1 function protects mice against adiposity. *JProceedings of National Academy of Science USA*. 2002; 99: 11482-11486.

[14] Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, Boerwinkle E, Cohen JC, Hobbs HH. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nature Genetics*. 2008; 40: 1461-1465.

[15] Speliotes EK, Yerges-Armstrong LM, Wu J, Hernaez R, Kim LJ, Palmer CD, Gudnason V *et al.* Genome-Wide Association Analysis Identifies Variants Associated with Nonalcoholic Fatty Liver Disease That Have Distinct Effects on Metabolic Traits. *PLoS Genet* 2011; 7(3): e1001324. doi:10.1371/journal.pgen.1001324