THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# Heterogeneity of Human Metabolism in Health and Disease: a Modeling Perspective

Pouyan Ghaffari Nouran


Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Heterogeneity of Human Metabolism in Health and Disease: a Modeling Perspective
Pouyan Ghaffari Nouran
ISBN 978-91-7597-586-3



Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie Nr 4267
ISSN 0346-718X

Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: Illustration of the systems biology approach from genes to community.

Printed by Chalmers Reproservice
Gothenburg, Sweden 2017

# Heterogeneity of human metabolism in health and disease: a modeling perspective

Pouyan G. Nouran

Department of Biology and Biological Engineering
Chalmers University of Technology

## Abstract

Metabolism is broadly defined as the sum of biochemical reactions within cells that are involved in maintaining the living state of the organism. Profound importance of metabolism comes from the fact that it is the sole source of energy that allows life to resist to be degraded into entropy. Human metabolism is a complex interactive network consisting of highly regulated functional pathways, impacting or been impacted by many other cellular process. Internal or external perturbations may cause dysfunction of some of these functional or regulatory pathways and may lead to the rise of abnormal phenotypes. Many human diseases associated with irregular metabolic transformations that perturb normal physiology and lead to phenotype dysfunction. Discovering how biological systems reorganize their activities to force specific phenotypic transformation, e.g., normal to cancer/diabetes/obesity, is a main challenge in life science. This thesis is dedicated to investigating genome-scale metabolic transformations from health to disease states, with specific focus on non-symmetric reprogramming in cancer metabolism.

The human gut microbiome has been associated with a variety of human diseases, but to go beyond association studies and elucidate causalities is a major challenge. We developed a comprehensive computational platform, CASINO (Community and Systems-level Interactive Optimization), for simulation of the microbial communities using genome-scale metabolic modeling. We demonstrated the power of the toolbox in predicting metabolic interactions between gut microbiota and host, through a diet-intervention study of obese and overweight individuals. Our modeling platform could provide a quantitative description of the altered plasma and fecal amino acid levels in response to dietary intervention.

Next, we proceed to investigate heterogeneity of cancer related metabolic transformations at the genome-scale. First, we reconstructed genome scale metabolic models (GEMs) for eleven human cancer cell lines based on RNA-Seq data. We used the generated models to investigate inter-cell line heterogeneity of metabolic reprogramming and also to identify potential anti-growth factors. This was followed by two consecutive studies on two main subtypes of the non-small cell lung cancer, lung adenocarcinoma (LAC) and lung squamous cell carcinoma (SCC), by generating RNA sequencing (RNAseq) data for cancer biopsies and for normal tissue samples. We followed a systemic approach to investigate the heterogeneity and direction of the metabolic transformation in lung cancer at three levels of biochemical organization: global metabolic network level, individual biochemical pathways level and at the level of specific enzymatic reactions. We observed large heterogeneity in the expression of enzymes involved in the majority of the metabolic pathways, and identified significant association between some of these variations and patient prognosis. Our findings provide mechanistic insights into complex metabolic behavior of tumors and may be used to develop more effective diagnostic and prognostic methods.

**Keywords:** cancer metabolism; genome-scale metabolic models; heterogeneity; community modeling; gut microbiome; lung cancer; amino acids metabolism; fatty acid metabolism

# List of publications

This thesis is based on the work contained in the following publications:

I. Shoaie, S., **Ghaffari, P**., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot … & Nielsen J. (2015). Quantifying Diet-Induced Metabolic Changes of the Human Gut Microbiome. Cell Metabolism, 22(2), 320-331.

II. **Ghaffari, P**., Mardinoglu, A., Asplund, A., Shoaie, S., Kampf, C., Uhlen, M., & Nielsen, J. (2015). Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. Scientific reports, 5, 8183.

III. **Ghaffari, P**., Djureinovic. D., Mattsson. J., Babaei, P., Fredrik Pontén, Mardinoglu, A., Micke, P., Uhlen, M., & Nielsen, J. Heterogeneity of amino acids metabolism affects lung cancer prognosis. Submitted for publication.

IV. **Ghaffari, P**., Mattsson, J., Mardinoglu, A., Micke, P., Uhlen, M., & Nielsen, J. Stratification of lung cancer patients based on heterogeneous expression of FABP5. Submitted for publication

V. **Ghaffari, P.**, Mardinoglu, A., & Nielsen, J. (2015). Cancer metabolism: a modeling perspective. Frontiers in physiology 6 (2015): 382

Additional publications not included in this thesis:

VI. Mardinoglu, A., Shoaie, S., Bergentall M., **Ghaffari, P**., Zhang C., Bäckhed F., Nielsen J., (2015) The gut microbiota modulates host amino acid and glutathione metabolism in mice. Molecular systems biology, 11.10 (2015): 834.

VII. **P. Ghaffari**, A. Mardinoglu, and J. Nielsen, Comparative analysis of metabolic reprograming across human cancer cell lines at genome-scale. Manuscript in preparation (2017).

# Contribution Summary

I. Formulated, implemented and developed software. Participated in analysis of data and writing the paper.

II. Reconstructed the models, performed the analysis, prepared and wrote the paper.

III. Designed the study, reconstructed the models, performed the analysis, prepared and wrote the paper.

IV. Designed the study, performed the analysis, prepared and wrote the paper.

V. Carried out the literature review, prepared and wrote the paper.

VI. Assisted in analysis of data and preparation and writing of the paper.

VII. Designed the study, reconstructed the models, performed the analysis, and wrote the paper.

# Preface

This dissertation is submitted for the partial fulfillment of the degree of doctor of philosophy. It is based on work carried out between 2013 and 2017 in the Systems and Synthetic Biology group, Department of Biology and Biological Engineering, Chalmers University of Technology under the supervision of Professor Jens Nielsen. The research was funded by the Knut and Alice Wallenberg Foundation, the Chalmers Foundation, and the Novo Nordisk Foundation.

Pouyan G. Nouran
May 2017

# Table of contents

# Abbreviations

ATP: Adenosine triphosphate
BCAA: Branched-chain amino acid
CASINO: Community and Systems-level Interactive Optimization
CBM: Constraint-based modeling
CL-GEM: Cell line specific GEM
Corr-Corr: Correlation-Correlation
DNA: Deoxyribonucleic acid
ER: Endoplasmic reticulum
ETC: Electron transport chain
FADH: Flavin adenine dinucleotide
FBA: Flux balance analysis
FDG-PET: Fluorodeoxyglucose positron emission tomography
GEM: Genome scale metabolic models
HGC: High gene count
HMR: Human Metabolic Reaction
HPA: Human protein atlas
KEGG: Kyoto Encyclopedia of Genes and Genomes
LAC: lung adenocarcinoma
LGC: Low gene count
MVG: Maximum variant genes
NADP: Nicotinamide adenine dinucleotide phosphate
NADPH: Nicotinamide adenine dinucleotide phosphate
NGS: Next generation sequencing
ODE: Ordinary differential equation
OXPHOS: Oxidative phosphorylation
PPI: Protein-protein interaction
PPP: Pentose phosphate pathway
QCD: Quartile coefficient of dispersion
RNA: Ribonucleic acid
RNAseq: RNA sequencing
ROS: Reactive oxygen species
RQCD: Relative quartile coefficient of dispersion
SCC: lung squamous carcinoma
SCFA: Short-chain fatty acid
TCA cycle: tricarboxylic cycle
TCGA: The cancer genome atlas
TNBC: Triple negative breast cancer

To Dilara,

For being by my side through it all

# Introduction

The world comprises diverse types of complex systems, such as biological, ecological and socio-economic systems. Complex systems are described by several common properties, including being thermodynamically open, containing diverse types of components with nonlinear interactions, exhibiting high degree of spatiotemporal heterogeneity and having feedback loops. These common characteristics come together to bring up two new properties, emergence and self-organization, which discriminate complex systems from non-complex ones. There are different types of the complexity with diverse implications across many fields of science and technology. Structural complexity refers to configurational and compositional sophistication of a system, functional complexity denotes nonlinear and heterogeneous dynamics of a system, and self-organizing complexity described based on emergence property of a system evolving from internal and external interactions at different special and temporal scales. Systems with self-organizing complexity have often been called complex adaptive systems. Majority of biological, ecological and social systems can be considered as complex adaptive systems with different degree of self-organizing complexity (Freeman et al., 2001). Herbert A. Simon in his seminal book 'The Sciences of the Artificial' (Simon, 1980) identified three waves of studies interested in science of complexity. The post-World War I wave, which was characterized by creative evolution and holism. The post-World War II wave which was shaped by information theory, cybernetics and general systems theory, and mainly focused on understanding parameters impacting system stability. The post- 80s wave which signified understanding causes and mechanisms underlying complex behaviors of natural systems. The diversity of the approaches in this period can be identified by concepts such as catastrophe, chaos, fractals, neural networks, genetic networks, cellular automata and self-organizations. It seems that, development of the theoretical and practical modeling of biological and ecological systems follows this historic line of approaches to describe the complexity.

Complex biological systems do not exist in stable conditions, but in temporal and spatial varying environments. They are always subjected to and respond to internal and external stimulus, such as chemical or physical perturbations or changes in composition of their environment. In order to maintain homeostasis, these systems activate several feedback loops to sense the environment and adapt their behavior fast enough to new structures of the environment. To understand biological systems, even at the basic level of cells and tissues, it is not enough to characterize individual molecules, and obtaining a broad view of the interactions between constituent molecules and compartments of the cell is required. This is more important when we want to investigate mechanisms underlying complex disease such as cancer, Alzheimer and diabetes. Systems biology has been developed to understand the complexity of biological systems at all levels of organization (from ecosystems to cells and molecules) in normal and disturbed conditions. It collects, integrates and analyse the mass of multidimensional and multi-omics data generated by high-throughput technologies, in order to develop descriptive and predictive models of biological systems. Systems biology synthesis the bottom-up (data-driven) and the top-down (model-driven) approaches into a hypothesis-driven approach in quest for basic biological principles and mechanisms (Auffray et al., 2009; Wang et al., 2015; Werner et al., 2014).

In the first line of studies we focused on modeling metabolic interactions between species in the human gut ecosystem. The human gut ecosystem contains a diversity of species outnumbering the number of human cells by a factor of 10, and the composition of this community is host-specific which dynamically evolve through life, subject to both endogenous and exogenous alterations (Huttenhower et al., 2012). The human gut microbiota provides essential health benefits to its host by contributing to digestion and secretion of nutrients, by providing resistance to pathogens, and especially by regulating immune homeostasis. The overall structural balance in the gut ecosystem, especially the presence or absence of key species, is important in insuring the normal host physiological states. However, the underlying mechanisms of gut microbiota influences on health through a complex network of interactions between host, microbiota and diets remain largely undefined (Nicholson et al., 2012). Integrating multi-dimensional high-throughput omics data using proper mathematical modeling techniques can help in this hurdle. Here, we developed a comprehensive computational platform, CASINO (Community And Systems-level Interactive Optimization) toolbox, to model interactions between species within microbial communities through genome-scale metabolic modeling. After validating the predictive power of the toolbox, we used it to simulate the diet-induced changes in composition and metabolism of the human gut microbiota.

In the second line of studies, we concentrated on understanding heterogeneity of cancer-related metabolic reprogramming at genome-scale. Metabolic reprogramming in tumors is closely related to malignant transformation by supporting growth and survival of the cancer cells through elevated supply of the energy and biosynthetic precursors, activation/deactivation of oncogenes/growth suppressors, and maintaining proliferative signaling (Hanahan and Weinberg, 2011). Genetic diversity which partly define the effect of tissue-of-origin, inconstant patterns of clonal architecture and environmental factors, such as gradients of nutrients and oxygen availability and tumor-host interactions, can give rise to a heterogeneous metabolic signature of cancer cells (Hensley et al., 2016b; Meacham and Morrison, 2013; Yuneva et al., 2012). In this line of study, we used mRNA expression profiles of the protein coding genes, generated from human cancer cell lines and biopsy samples of human lung cancer sub-types, to assess cancer-related metabolic transformations. We followed a systemic approach to investigate the scale and direction of heterogeneity of metabolic reprogramming in cancer cell lines and lung cancer patients, at three levels of: metabolic network, biochemical pathways and individual enzymatic reactions.

Generally speaking, this thesis describes application of systems biology approach to advance our understanding on different patterns of metabolic transformations in health and disease. Findings in this thesis, can gain insight into underlying mechanisms of cancer plasticity and support the development of future personalized therapeutic strategies.

# Background

## Complex systems science

If we pay attention to most of the events that occur around us, from those affecting our own health condition to those changing the environment where we live in and even those influencing our social status, we realize that they are the consequence of the evolving dynamical organization of systems that, on their turn, consist of layers of subsystems and basic components with complex patterns of interaction. One of the main efforts of modern science, defined as complex systems science, is to provide proper and descriptive images of theses observed phenomena to understand rules that are lying behind, and to manipulate and control them in desired direction. The clear recognition of complex systems concepts and principles and its widespread application in science, which provides new ways of understanding mechanisms in physical, biological, engineering and social systems, is a very recent phenomenon. In a simplest way, we can define a complex system as an ensemble of the multiple constituents which are interacting in a disordered way to form and maintain a new organization. Main properties of the complex systems that establish our natural world have been manifested as: (Kwapien and Drozdz, 2012; Lansing, 2003; Segel, 1986).

- Nonlinearity of interactions. An interaction is linear in some variable x, if the interaction force f vary proportionately to x (linearity: $f(kx) = kf(x)$, where k is a constant). Linear systems are characterized by additivity and nonlinear systems characterized by non-additivity.
- Sensitivity to initial conditions. Complex systems response to perturbations are mainly difficult to predict, as their nonlinear nature permits small change in initial state of the system to be intensified into large differences in following system trajectories. This property of complex systems highlights the importance of errors in system analysis and how our interpretations could be biased under effect of errors.
- Symmetry breaking. In general term symmetry means invariance under an operation such as scaling, reflection, relocating and rotation. For example, molecular structure of a crystal is invariant under spatial transformations. Complexity of the systems increase by each symmetry breaking.
- Self-organization. Self-organization happens when a system bifurcates, through local interactions between constituents and under its own dynamics, to a form exhibiting new overall order. The process is spontaneous without any external regulation. Chemical oscillation, protein folding, formation of the lipid membranes, crystallization, neural networks and animals swarming are examples of self-organization.
- Emergence. We can define emergence in simplest form as: the rising of a new perceivable coherent macro property or structure during the dynamic process of the self-organization within complex systems. The emergent property cannot be reduced to sum of the micro properties of the constituents. For example, life as an emergent property cannot be reduced

to underlying laws of chemistry and physics. In fact, one can claim where emergence appears, causation declines.

Figure 1 demonstrates a fundamental difference between simple and complex systems in sensing of environment and hierarchy of internal interactions to produce response.

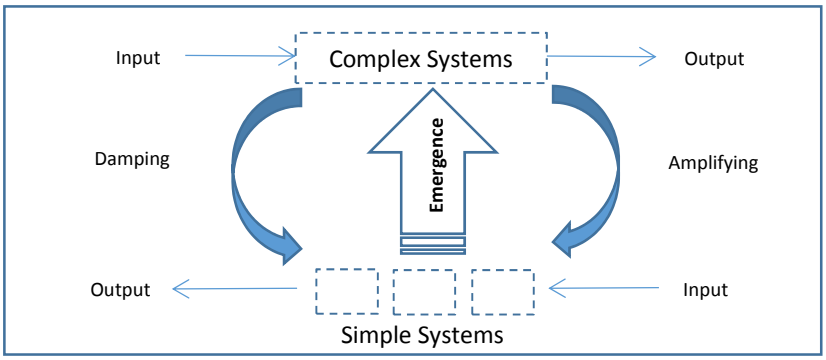


**Figure 1 complex systems vs simple systems**

## Systems biology: a tool to study complex biological systems

The divide and conquer approach (reductionism) has dominated biological research, similar to other scientific disciplines, for many years. This strategy made great progresses in understanding basic mechanisms shaping living cells and in identifying many essential components accounting for specific cellular phenotypes in health and disease. However, the complexity of biological systems makes it a challenging effort to analyze and understand structure and dynamics of cellular and organismal functions, and many critical questions cannot be answered just by focusing on individual molecular components in isolation. Repeatedly, critics have emerged to challenge the reductionist approach in biology, but opponents lacked applicable research tools and techniques and just indicate the idea to escape reductionism and mechanistic explanation. The recent advancements of complex systems science and implementation of mathematical tools to investigate structure and dynamics of complex biological networks, has begun to take root as an important sub-discipline of biological science defined as systems biology and computational biology. Kitano, in his seminal paper in Science 2002, recommended four key properties to be considered to derive a system level understanding of the biological systems: system structures, system dynamics, the control method and the design method. He also proposed an idealized cycle of systems biology research in a hope for gradual transformation of the biological research to fit this cycle towards systematic hypothesis-driven life science (Kitano, 2002).

In medical science, the ultimate goal of systems biology is understanding factors that influence human health and cause diseases across the multiple levels of organization, from molecular interactions to networks of biochemical pathways, from the cell–cell interactions to tissue level, and from organs to the physiological functioning of the whole organism. Systems biology body of science covers the experimental methods for generation of high-throughput datasets (omics data),

and development and application of the computationally derived methods and network approaches to integrated and analyze generated multi-dimensional datasets. High-throughput datasets derived from transcriptomics (mRNA transcripts), proteomics (protein levels), metabolomics (small molecule metabolites) and fluxomics (rates of metabolic reactions) can provide global information of the temporal and special dynamics in the cell and tissue levels. Integration of the omics datasets with each other and other data types have been used to reconstruct models of cell metabolism, cell signaling and disease networks to understand the mechanisms of interactions and to identify new therapeutic targets (Berg, 2014; Wang et al., 2015).

## Network analysis: a tool to represent complex biological systems

Large high-throughput omics datasets highlight the challenge of how to reduce the complexity of data, integrate multiple sources of data and connect findings to external information to develop robust models and testable hypothesis from this huge amount of data. Biological systems, similar to most complex systems, do not function in isolation, rather dynamic interactions of the elements determines the function and behavior of these systems. These interactions shape different types of biological networks with characteristic topology, such as protein-protein interaction (PPI) networks, gene regulatory networks, signal transduction networks and metabolic networks (Barabasi et al., 2011). Network representation of multi-dimensional omics data provides a simplified portrait of the complex biological systems by reducing the dimension of the data, and enables using tools from graph theory and network science to analyze data. At an abstract level, the system components can be reduced to nodes that are connected to each other by edges that represent the interactions between components. The nodes and specific pattern of edges together form a characteristic network topology. Networks can be undirected or directed depending on the nature of the interactions. When interactions between nodes have well-defined direction, it is called a directed network, and otherwise called undirected network. Metabolic networks with defined direction of material flow from substrate to product are an example of directed network and PPI networks are example of undirected network. Despite remarkable diversity of the natural complex networks, realizing that their topology are shaped by couple of simple common principles was probably the most important discoveries of graph theory (Barabasi and Oltvai, 2004). One of these simple principles is that topological properties of the real networks cannot be explained by random network model and they mainly characterized by power-law degree distributions. It means that, for many real networks the number of the nodes with a given degree follows power law distribution not Poisson degree distribution, and the probability of connectivity for each node follows:

$$P(k) \sim k^{-\gamma}$$

where k is the degree and $\gamma$ is the degree exponent with value between 2 and 3 for most networks (Barabasi and Albert, 1999). Networks with power degree distribution are called scale-free or scale-rich networks. It has been shown that cellular metabolism follows a scale-free topology by analyzing metabolic networks of 43 different organisms (Jeong et al., 2000). Another common feature of the complex networks is called small-world effect, which means any two nodes within these networks are connected with a small number of the nodes. Small-world effect is stronger in

scale-free networks, making them ultra-small and consequently any local perturbation propagates very quickly in these networks. Investigations increasingly revealed that functions of the biological networks are closely related to their topological structure, meaning that global and local structural features of these networks can disclose essential properties of the biological systems. Differential network analysis, which compares the topological alterations of complex networks under different conditions, may help to discover disease biomarkers (Wang et al., 2015).

## Genome-scale metabolic modeling

One of the fundamental objective of systems biology is to reconstruct a predictive model of a living cell encompassing all its components in a quest for manipulating biological systems in desired direction. But, insufficient broad knowledge on biological processes and related kinetics is an important hurdle. However, cell metabolism is the only domain we have been able to make first steps towards realizing this visionary model, thanks to hundred years of biochemistry research and the continued progression of high-throughput omics techniques and next generation sequencing (NGS) methods in the past decade, which transformed biology into a data-rich scientific discipline by simultaneous measurement of a large number of cellular components. By now, metabolism is most probably the best studied cellular process and recent advances in the computational study of cellular metabolism have accelerated its computational investigation on a genome scale. Despite this significant progress, we have been able to measure reaction flux rates in eukaryote cells just for limited number of reactions in central metabolism (Niklas et al., 2010). Reconstructing and analyzing genome-scale metabolic models (GEMs) can bridge this gap and characterize the large scale metabolic state of the cells/tissues/organisms by facilitating the prediction of the inner and outer flux rates of the metabolic networks. Furthermore, GEMs are capable of providing context-specific metabolic descriptions by integrating multi-layer omics data, and simulation of the metabolic interactions within multi-species ecosystem under dynamic environmental perturbations (Mardinoglul and Nielsen, 2015; Nielsen, 2017; Robinson and Nielsen, 2016).

GEMs are mathematical portrait of the complex biochemical transformations occurring within a specific cell or tissue through a compartmented assembling of the metabolites, reactions and enzymes into the complex metabolic network (Figure 2). The constructed stoichiometry matrix incorporates stoichiometric coefficients of the metabolic reactions complemented by detailed mapping of the enzymes into their associated reactions. In general, metabolic networks are modeled under quasi-steady state assumption and constraint-based modeling (CBM) approach is used to analyze reconstructed GEMs. CBM imposes physico-chemical constraints such as mass balance, thermodynamics and min/max flux capacity boundaries to shape the feasible solution space, and the reconstructed model usually is under-determined with range of potential alternative flux distributions satisfying the constraints. Flux balance analysis (FBA) is mainly used to select optimum flux distribution by optimizing the formulated model for pre-defined objective function such as biomass yield, substrate consumption, ATP production and redox potential (Figure 2).

Since the first GEM reconstructed for *Haemophilus influenza* in 1999, increasing number of the metabolic models have been generated for diverse range of species spanning from bacteria to

humans. Human models, in contrast to GEMs generated for microorganisms, are reconstructed as generic forms including all metabolites and reactions have been documented to be present in human metabolism integrated with known catalyzing enzymes. First generic human metabolic models, Edinburgh Human Metabolic Network (EHMN) and Recon 1, were manually reconstructed from accumulated high quality genetic and biochemical data, followed by more comprehensive Recon2 and Human Metabolic Reaction 2.0 (HMR2) models encompassing current advances in human biochemical knowledge (Ghaffari Nouran et al., 2015; Mardinoglu et al., 2013; Robinson and Nielsen, 2016). In recent years, accumulation of the disease related high-throughput omics data made it possible to project this data into the generic human GEMs and to reconstruct context-specific models of human metabolism in health and disease. An extensive body of work has showed successful applications of the generated GEM to investigate metabolic foundations of the non-healthy phenotypes and to identify new selective biomarkers and therapeutic targets (Ghaffari et al., 2015; Jerby and Ruppin, 2012; Nielsen, 2017; Oberhardt et al., 2009; Shoaie et al., 2015).

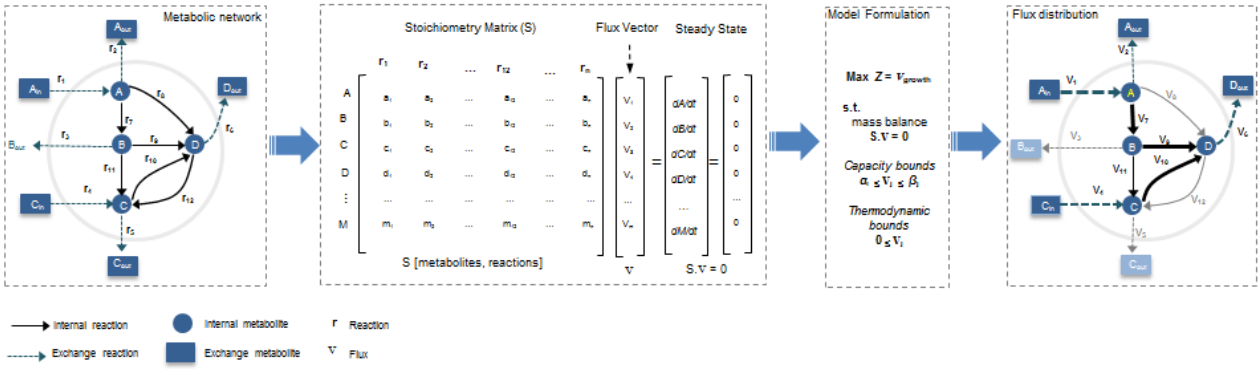


**Figure 2 constraint-based modeling (CBM) and flux balance analysis (FBA).** Conceptual illustration of modeling a simple metabolic network. The stoichiometric matrix is reconstructed under steady-state condition and model is formulated by introducing physico-chemical constraints and defining a relevant objective function. Formulated model is solved through FBA providing optimal and feasible flux distribution.

# Human gut microbiota

Since the first observations related to the presence of microbes in the human mouth and faces, it has been discovered that our body contains a massive number of microorganisms, including bacteria, eukaryotes and viruses, colonizing in different sites of the body from skin to the gastrointestinal tract. These microorganism, collectively is called microbiota and their collective genome is referred to as metagenome (Huttenhower et al., 2012). Human microbiota outnumbers the total number somatic cells in the body by a factor of 10, and the human gut metagenome is twofold greater than the human genome. The gastrointestinal tract hosts a large portion of the human microbiota with an overall weight of 1.5 kg, comparable to the weight of the adult human brain (Zhao, 2013). The density of cells changes in different sites of the gut, from $10^3$ to $10^8$ cells

per gram of luminal content in the small intestine to $10^{11}$-$10^{12}$ cells per gram of luminal content in the distal colon, with varying diversity and composition of the gut microbiota along the gastrointestinal tract. In humans, the gut microbiota shows high inter-subject variation of diversity, reflecting the maternal hand-over of the species at birth, and subsequent evolution of the microbial landscape through life driven by the complex and dynamic interplay between factors such as geography, age, life-style, diet, diseases and drug usage (Backhed et al., 2005; Claesson et al., 2012; Dethlefsen and Relman, 2011). Metabolic phenotype of the host, and consequently biochemistry and susceptibility to disease, is highly influenced and modulated by evolving trajectory of the gut microbiota. Despite high inter-subject variation of the gut microbiota, some recent studies have shown similarity of changes in the composition of the microbiota between two time points across subjects (Faith et al., 2013). However, there are biologically and metabolically flexible components that represent dynamic behavior in response to dietary perturbations or environmental stresses, by altering the diversity and composition of the species (Clemente et al., 2012).

The cross-talk between gut microbiota and host immune system, shapes the composition of the microbiota as well as the development of the immune system, through vast array of signaling and metabolic pathways. Immune-mediated signaling processes act upon several organs such as brain, gut, muscle and liver, together with complex biochemical interactions between host and microbiota involving chains of microbe-host multidirectional metabolic axes. Within these metabolic axes, host and microbiota co-metabolize large array of dietary and xenobiotic compounds, including short-chain fatty acids (SCFAs), bile acids, lipids and choline, many of which are essential for host health and microbial symbiosis. For each region of the human gastrointestinal tract there exists distinctive composition of the microbiota with specific inter-species and host-species metabolic dialogues (Li et al., 2008).

# Modeling ecosystems

Models are mathematical abstraction of the reality, comprising essential and necessary features of the real system, to make predictions of dynamics and behavior of the system which is under study. Ecological systems, composed of large number of biotic and abiotic components that interact with each other through chemical, physical and biological processes going on in multiple layers, are of overwhelming complexity. Due to this complexity, ecosystem models are mainly simplified to limited number of the well-understood parameters relevant to solve the problem in the intended direction and in accepted level of difficulty. The simplification process usually reduces the dynamics of complex ecosystem to a reasonable number of parameters, called state variables, and mathematical functions which describe the interaction between these variables. Along with parameters and interactions, most of the models developed for ecosystems have to incorporate spatial and temporal dynamics as inevitable part of the problem (Auger and Roussarie, 1994; Costanza et al., 1993). One of the most well-known ecosystems models is Lotka-Volterra predator and prey model which originally developed to explain observed fluctuations in fish populations.

This model uses pairs of ordinary differential equations (ODEs), representing predator and prey species.

$$\frac{dY}{dt} = \gamma.\beta.X.Y - \delta.Y$$

$$\frac{dX}{dt} = \alpha.X - \beta.X.Y$$

Where, Y is the number/concentration of the predator species, X is the number/concentration of the prey species, $\alpha$ is the prey species' growth rate, $\beta$ is the predation rate of Y upon X, $\gamma$ is the assimilation efficiency of Y and $\delta$ is the mortality rate of the predator species. Although, the model illustrates some features of the modelled biological ecosystem like growth and mortality, later studies revealed that real interactions in nature is far from Lotka-Volterra's model (Peterson, 2013).

Microbial communities comprise nearly 50% of the earth's biomass and their functions and interactions play a key role in the ecosystem of the planet and human life by driving most of the biogeochemical cycles. They also have been used widely in food fermentations, for example in the production of yogurt, cheeses, vinegar and soy sauce (Hanemaaijer et al., 2015; Kallmeyer et al., 2012; Lin et al., 2000). Despite major impact of the microbial communities on human health and sustainability, we still have little insight into the design principles that regulate their ecosystems evolution, functionality, robustness, and consequently limited opportunities to control and optimize the performance of these communities in desired direction.

## Genome scale modeling of ecosystems

Recent advances in high-throughput technologies enabled us to study metagenomics, meta-transcriptomics and metabolomics of the microbial communities. This high resolution omics data made it possible to identify specific molecules such as DNA, mRNA, proteins and metabolites, which provided information on the identities of the occurring species, their relative abundances and their metabolic potentials (Segata et al., 2013). However, information acquired from meta-omics studies, including 16S rRNA gene sequencing, only provide indirect view into the biological activities of the microbes within the ecosystem, and almost no insight into how genetic features and the interaction between species and with their environment give rise to the community structure and community state, i.e., the concentrations and fluxes of nutrients within ecosystem. Inference of the community structure and state, requires mathematical modeling methodologies that can integrate heterogeneous physiochemical, physiological information with omics and meta-omics data in context of a unified model with optimum level of information content and predictive power. Constrain-based genome-scale metabolic modeling of the community metabolism, capable of integrating high-throughput multi-dimensional data, is promising for this purpose.

One way to use GEMs for studying ecosystems is network-based analysis, which uses topology of the network to gain insight into the structure and state of the community, regardless of stoichiometric parameters. This approach has been applied to develop a computational framework, by integrating GEMs with metagenomics data, to predict levels of the complementarity and

competition between the 154 species of the gut microbiota and to study the rules behind the community-level assembly of microorganism within ecosystem (Levy and Borenstein, 2013). Following successful applications of the CBM and FBA to predict metabolic fluxes and the phenotype of the individual species, taking into account the stoichiometry information and constraining for the metabolic fluxes, couple of researchers have been interested in using the analogue path to study communities. Stolyar et al. reported the first implementation of the multispecies CBM to model interactions between sulphate-reducing bacteria and methanogenic archaea, assuming each model as a discreet compartment and using one extra compartment for metabolic exchanges between species (Stolyar et al., 2007). Freilich et al. used stoichiometry-based metabolic models of 118 species to study the cooperative and competitive potential between bacterial pairs, and identified correlations between bacterial interactions and their coexistence pattern (Freilich et al., 2011). Zomorrodi et al. proposed a bi-level optimization approach to model community of species, including lower-level objective functions for species and an upper-level objective function for community to simultaneously optimize for species-level and community level (Zomorrodi and Maranas, 2012).

# Cancer metabolism: the fundamentals

According to the World Health Organization (Lunt et al.) cancer is a major public health problem worldwide, with more than 14 million new cases and more than 8 million deaths in 2012 (Ferlay et al., 2015). One of the oldest areas of research in cancer biology, dating back to early 20 century, is cancer metabolism. This is founded on the principle that metabolic activities are transformed in tumor cells in a direction that supports the acquisition and maintenance of malignant properties. The past decade has experienced a fast increase in investigations on mechanisms underlying cancer-related metabolic alterations, resulting in accumulated evidences on association between malignant transformation and several pathways in human metabolism (Heiden et al., 2009; Schulze and Harris, 2013). These studies revealed that some altered metabolic properties are happening quite frequently during neoplastic transformation across different types of the cancer cells, which resulted in considering deregulated cellular metabolism as one of the hallmarks of cancer (Hanahan and Weinberg, 2011). Accordingly, several selective metabolic targets have been identified and started to enter clinical studies (Galluzzi et al., 2013). Cancer-mediated altered metabolic activity supports catabolic cells survival during nutrient limitation, anabolic cells proliferation during nutrient-replete periods, and maintained redox homeostasis under stressed conditions (Boroughs and DeBerardinis, 2015). Characterization of these reprogrammed metabolic activities may provide opportunities to diagnose cancer in early stages by imaging malignant tissues noninvasively and predicting tumor behavior, and to inhibit cancer progression by blocking critical pathways for tumor growth and survival (Ward and Thompson, 2012).

Aerobic glycolysis or the Warburg effect is the classical example of metabolic reprogramming in cancer cells. Normal cells employ glycolysis as a physiological response to hypoxic condition, but in early 1920s Otto Warburg observed that tumor cells increase glucose uptake regardless of oxygen availability, a phenomena that has been observed in many tumor types (Koppenol et al.,

2011). Glycolysis produces energy faster but far less efficient than oxidative phosphorylation, making cancer cell dependent on increased influx of the glucose. Elevated glycolytic flux also provides glycolytic intermediates required by the subsidiary metabolic pathways to support cell proliferation (Lunt and Vander Heiden, 2011). Like glycolysis, intermediates derived from tricarboxylic cycle (TCA cycle) are also use as precursors for biosynthetic pathways. To maintain intermediate pools, carbon needs to be resupplied into the TCA cycle by anaplerotic pathways at sites other than acetyl-CoA. Anaplerotic fluxes in tumor cells can be provided by glutaminolysis, pyruvate carboxylation, and in some tissues oxidation of the branched-chain amino acids (BCAAs), valine and isoleucine (Owen et al., 2002). The general induction of some metabolic pathways by malignancy to support core functions like catabolism, anabolism and redox homeostasis may reflect their regulation by a finite set of signaling pathways. Tumor cells frequently activate phosphatidylinositol 3-kinase (PI3K) and its downstream pathways AKT and mammalian target of rapamycin (mTOR) to promote anabolic growth through elevated glycolytic flux and fatty acid synthesis, with minimum dependence on external stimulation by growth factors (Yuan and Cantley, 2008). Another commonly detected deregulated signaling pathway in cancer cells is gain of function by MYC that overexpresses several effective genes in anabolic growth, including enzymes involved in glycolysis, glutaminolysis, serine metabolism, fatty acid biosynthesis and mitochondrial metabolism (Stine et al., 2015). The tumor suppressor gene p53 transcription factor is deleted or mutated in nearly half of the human cancers. Some recent studies indicated that tumor-suppressive function of the p53 might be independent of its canonical activities, such as cell cycle arrest, DNA repair and apoptosis, but rather dependent on regulation of the metabolism under oxidative stress (Kruiswijk et al., 2015). Tumor cells usually proliferate faster than the rate of new blood vessel formation and reside in hypoxic environment, ranging from 0 to 2% of oxygen concentration. Hypoxia-inducible factor–1 (HIF-1), which induces upregulation of the enzymes involved in elevated glycolytic flux, coordinates metabolic adaptation to hypoxia (Semenza, 2012).

## Tumor bioenergetics

Otto Warburg's hypothesis that impaired mitochondrial function, makes tumor cells depend on glycolysis as major source of energy, even in the presence of oxygen, was widely accepted for decades. Warburg's observation that cancer cells consume glucose in much higher rate than normal cell has been validated extensively in many tumor types and has been employed in clinic for cancer diagnosis by fluorodeoxyglucose positron emission tomography (FDG-PET) technique, which uses 2-deoxy-2-[fluorine-18] fluoro-D-glucose, a radioactive analogue of glucose, to visualize glucose uptake in cancer cells and adjacent normal tissue (Kelloff et al., 2005). However, recent studies reveled that most of the tumor cells generate the majority of the required energy through mitochondrial metabolism, except the ones experience mutations in genes involved in mitochondrial respiration, such as Succinate Dehydrogenase (SDH) and Fumarate Hydratase (FH) (Zu and Guppy, 2004). Yet, cancer cells bearing mutations in SDH or FH still rewire their metabolism to replenish necessary intermediates for TCA cycle and for proliferation (Cardaci et al., 2015; Mullen et al., 2012). Todays, there is an overall consensus around the idea that Warburg's effect is mainly caused by loss of tumor suppressor genes, activation of the oncogenes and

overexpression of the PI3K signaling pathway, and one of the key advantages of the upregulated glycolysis in cancer cells is to provide biosynthetic precursors for proliferating cells (Pavlova and Thompson, 2016).

In addition to glucose, amino acids and fatty acids can replenish the TCA cycle and maintain mitochondrial ATP production in cancer cells. Fatty acids oxidation in the mitochondria produces acetyl-CoA which enters the TCA cycle, and NADH and FADH which are used by electron transport chain (ETC) to generate ATP (Carracedo et al., 2013). Glutaminolysis generates α-ketoglutarate from glutamine to fuel TCA cycle (Hensley et al., 2013). Amino acids leucine, isoleucine and valine, which are found in elevated amounts in plasma of pancreatic cancer patients, can be converted to acetyl-CoA and enter the TCA cycle (Mayers et al., 2014). In conditions of nutrient deprivation and metabolic stress, some cancer cells scavenge external fatty acids to fuel mitochondrial oxidation of fatty acids and to produce ATP (Nieman et al., 2011). This metabolic flexibility of the TCA cycle increases tumor cells plasticity and makes them resistant to metabolic perturbations caused by fluctuating fuels availability during the evolution of the cancer.

## Macromolecules biosynthesis and redox balance

Tumor are characterized by uncontrolled high rate of cellular proliferation and consequently biosynthetic pathways are of critical importance for cancer cells to maintain macromolecules production rate required for cell division and tumor growth. Nucleic acids, proteins and lipids, comprising up to 80% of the dry mass of mammalian cells, are the macromolecules that contribute to tumor cell metabolism in large extent. Generally speaking, cellular biosynthesis involves three main phases: acquiring simple nutrients such as hydrocarbons, fatty acids and amino acids from extracellular matrix, employing central metabolism like glycolysis, TCA cycle, amino acid biosynthesis and pentose phosphate pathway (PPP) to convert imported compounds into the biosynthetic intermediates, and finally assembling large and complex macromolecules through energy consuming processes (DeBerardinis and Chandel, 2016).

Protein synthesis is a highly regulated process that requires access to a full set of essential and nonessential amino acids to function. Proliferating cancer cells overexpress surface transporters to import amino acids from the environment, convert glutamine to glutamate through deamidation and transamidation reactions, and activate transamination reactions to synthesis most of the nonessential amino acids from glutamate (Hensley et al., 2013). Under conditions of nutrient depletion, cells can activate some catabolic pathways to degrade macromolecules from internal and external sources and replenish pool of required intracellular intermediates. Autophagy is an important survival pathway capable of recycling intracellular proteins and organelles through a highly regulated process and delivering them to the lysosome (Galluzzi et al., 2014). When nutrients are scarce, cells can internalize other components from the extracellular milieu through macropinocytosis, and supply the central metabolism with necessary flux of carbon and nitrogen (Commisso et al., 2013). Some evidences indicate that intracellular and extracellular protein degradation can be suppressed by activity of mTORC1 signaling which may help cells to maximize net rate of protein synthesis when amino acids are available (Palm et al., 2015).

Cancer cells need extra sources of acetyl-CoA and cytosolic NADPH as reducing power to elevate fatty acids synthesis required for membrane biosynthesis and cellular signaling. Glycolysis is the prominent source of the acetyl-CoA for fatty acid synthesis, along with glutamine and acetate as alternative carbon sources when access to glycolysis-derived acetyl-CoA is diminished by mitochondrial dysfunction or hypoxia (Metallo et al., 2012; Schug et al., 2015). Membrane biosynthesis can also be supplied by acquiring fatty acids from the extracellular space, in particular when ability to satisfy biosynthetic demands of highly proliferative cancer cells is compromised by metabolic stress under hypoxia (Nieman et al., 2011).

As mentioned earlier, the majority of cancer cells employ mitochondrial oxidation of glucose, lipids and glutamine coupled with ETC and oxygen as the final electron acceptor, to satisfy an important portion of their energy demand. Mitochondrial oxidative metabolism is connected to the production of reaction oxygen species (ROS), which at high levels, can damage nucleotide and consequently impair cell viability (Sabharwal and Schumacker, 2014). ROS are chemically reactive intracellular species that contain oxygen and include the superoxides ($O_2^-$), hydrogen peroxide ($H_2O_2$), hydroxyl radicals ($OH\cdot$) and singlet oxygen (Murphy, 2009). The reduction of the oxygen in the mitochondria and through cytosolic NADPH oxidases (NOXs) produces superoxide, and enzymatic activity of superoxide dismutase 1 or 2 converts superoxide to hydrogen peroxide, followed by partial reduction of the hydrogen peroxide to hydroxyl radical or its full detoxification to water (Brand, 2010). ROS regulate key cellular processes, such as metabolism, apoptosis and cancer cells invasiveness. Cancer cells mainly use reduced glutathione together with high ratio of $NADPH/NADP^+$ to detoxify ROS. GHS is synthesized from glycine, glutamate and cysteine, and its oxidation by GSH peroxidase is coupled with hydrogen peroxide turnover rate. NADPH is an electron carrier cofactor and its constant regeneration from $NADP^+$ is necessary to maintain cellular redox homeostasis. In the mitochondria, NADPH production is partially controlled by one-carbon metabolism and isocitrate dehydrogenase 2 (IDH2), and cytosol uses multiple source to produce NADPH including PPP, one-carbon metabolism, malic enzyme 1 and isocitrate dehydrogenase 1 (IDH1) (Martinez-Outschoorn et al., 2017).

Redox balance is required during tumorigenesis and metastasis. It seems that moderate levels of ROS can support cancer progression, but when ROS concentration exceeds antioxidant capacity of the cells, a lethal oxidative stress can occurs (Gorrini et al., 2013). A rising model of cancer redox homeostasis is that during tumor initiation, which is characterized by high proliferation rate and increased metabolic activity of cancer cells and subsequently increased level of ROS, tumor cells increase their antioxidant capacity to detoxify ROS. Under hypoxic and nutrient-deprived conditions, which characterized by limited fluxes through glycolysis and PPP and consequently reduced NADPH levels, tumor cells activate adenosine5-monophosphate kinase (AMPK) to stimulate PPP dependent NADPH production and downregulate anabolic pathways with high demand of NADPH, such as lipid biosynthesis. When cancer cells detached from the matrix, they encounter high concentrations of ROS and their ability to upregulate NADPH production and antioxidant capacity is required for distant metastasis (Jiang et al., 2016; Piskounova et al., 2015).

## Therapeutic potentials of cancer metabolism

Cancers reprogram their metabolism to meet the biosynthetic, bioenergetic, and redox requirements of malignant cells during tumorigenesis and metastasis. Cancer-related metabolic alterations supports cancer cells proliferation and survival, but at the same time make them vulnerable to perturbations. However, these metabolic transformations are not homogeneous and shows remarkable intra-cancer and inter-cancer plasticity. Understanding how and why metabolic alterations occurs and the mechanism behind, progressively opens insights into the development of new therapeutic strategies for human cancer. Antimetabolites, such as antifolates, that target specific part of the cancer cell metabolism, were among the first developed cancer therapies and still remain as important group of chemotherapy agents in the clinic. Glycolysis was an early attractive therapeutic target following experimental and clinical observations that revealed majority of tumors experience a remarkable elevation in glucose uptake (Vander Heiden, 2011). Lactate dehydrogenase A (Felson et al.), that catalysis conversion of the pyruvate to lactate, was the first target identified in glycolysis pathway. Hexokinase 2 (HK2), which phosphorylate glucose to glucose-6-phosphate, experience overexpression in many type of the cancer and preclinical mouse studies demonstrated that inhibition of the HK2 suppresses lung and breast cancer cells progression (Patra et al., 2013). Other potential glycolysis-related targets are D-3-phosphoglycerate dehydrogenase (PHGDH) and Serine Hydroxymethyltransferase 2 (SHMT2), enzymes involved in de novo serine and glycine synthesis pathway. Currently, it is not known that targeting enzymes within one-carbon metabolism, such as PHGDH and SHMT2, would be effective in suppressing tumor growth without causing systemic toxicity (Nilsson et al., 2014; Possemato et al., 2011).

Emerging evidences have begun to show that mitochondrial metabolism is potentially a key target for cancer therapy, partially, due to identification of the anticancer effect of metformin, an antidiabetic drug to control patients' blood glucose levels, through several epidemiological studies (Weinberg et al., 2015). Laboratory based studies also indicated that metformin inhibits mitochondrial ETC complex I, hampers mitochondrial ATP production and constraints the biosynthetic capacity of the mitochondria in cancer cells (Griss et al., 2015). Mitochondrial metabolism potentially can be targeted by inhibiting autophagy or glutaminase. Autophagy refuel TCA cycle by supply amino acids and its short term inhibition has decreased tumor progress in mouse models of lung cancer without inducing systemic toxicity (Karsli-Uzunbas et al., 2014). Glutamine is an important component to support TCA cycle metabolism in several types of the cancer, even in the absence of autophagy, and consequently inhibition of glutaminase can diminish tumorigenesis (Xiang et al., 2015). During metabolic stress, tumors can use acetate to support cancer cells proliferation and survival which makes acetate metabolism an alternative therapeutic target (Comerford et al., 2014).

Cancer cells elevate their antioxidant capacity to maintain redox balance during growth and metastasis, thus selectively targeting of tumors redox metabolism is a potential therapeutic approach (Gorrini et al., 2013). The reducing equivalent HADPH plays a key role in cancer cells antioxidant defense system but most of the cytosolic and mitochondrial NADPH-generating pathways and enzymes, such as PPP, one-carbon metabolism, malic enzyme1, IDH1 and IDH2,

are critical for normal cell function, which make them difficult to target. But, there are two potential strategies to target NADPH-generating systems with minimal side effects, namely inhibiting glucose-6-phosphate dehydrogenase (G6PDH) and administering high doses of vitamin C. G6PDH is a catabolic enzyme that participates in oxidative PPP and catalyzes the conversion of $NADP^+$ to NADPH. Approximately 400 million people worldwide are estimated to be G6PDH deficient, however some cancer types are dependent on this pathway for NADPH generation and maintain redox homeostasis, making it potential therapeutic target. In colorectal cancer cells that harbor KRAS mutations, administering high doses of vitamin C can induce cancer cell death by depleting NADPH and glutathione pools and elevating ROS levels (Yun et al., 2015).

Most of the metabolic inhibitors are unlikely to be an efficient therapeutic target as a single agent, so it would be more effective approach to consider combination of agents to target a malignant function through multiple pathways. For example, to impair ATP sources within cancer cells, metformin, which inhibits mitochondrial ATP production, can be combined with current clinical PI3K inhibitors, which reduces glycolysis. It is necessary to consider that normal proliferating cells, such as stem cells and immune cells, mainly reprogram their metabolism similar to cancer cells, and adaptive immune system is very sensitive to metabolic perturbations, so inhibition of some metabolic processes and associated enzymes may cause systemic toxicity (Erez and DeBerardinis, 2015; Ito and Suda, 2014; Pearce et al., 2013). Feasibility of metabolic targets for cancer therapy depends on whether they can be tolerated by normal tissues as well as whole body metabolism.

## Heterogeneity of cancer metabolism

Cancer is a complex disease that emerges through an iterative process of genetic alteration, clonal extension and selection within adaptive and dynamic landscape of host tissue ecosystem, over a changing time frame from 1 to 50 years. The intricate cross-talk between cancer cells and environment may induce selective pressure in favor of clones that are more capable to adapt for and to survive under temporally and specially changing environmental conditions, regardless of growth promoting mutations (Greaves and Maley, 2012). Despite tumors general tendency towards uniformity of metabolic transformations to support growth stimulating adaptations, heterogeneous patterns of cancer mediated metabolic reprograming occurs due to genetic diversity, diverse patterns of clonal architecture, dynamic tumor microenvironment and patient-specific whole body metabolism (Meacham and Morrison, 2013). Apparently, a common metabolic reprograming model cannot describe diverse metabolic alterations happen during tumorigenesis. Cancers like colorectal, liver, pancreas, and leukemia rely on glycolysis, while glioblastoma, melanoma and lymphoma have been categorized as oxidative cancers (Lehuede et al., 2016; Obre and Rossignol, 2015). Molecular profiling of 176 diffuse large B-cell lymphoma uncovered the existence of robust subtypes of cancer cells characterized by expression pattern of enzymes associated with OXPHOS (Monti et al., 2005). Later studies revealed that expression level of TCA cycle and ETC enzymes stratify lymphoma into two subtypes with active or non-active OXPHOS metabolism. Analyzing expression profiles of metabolic genes from 22 different cancer types, Hu et al. showed general similarity between expression profiles of tumors and corresponding normal tissues (Hu et al.,

2013). Tumors demonstrate distinctive metabolic features characterized by tissue of origin, e.g., Myc-driven live and lung cancer cells display different phenotypes related to glutamine metabolism (Yuneva et al., 2012). Cancer cell subpopulations depending on spatiotemporal factors such as nutrient availability, oxygenation, pH, growth factors and other cellular populations, may adapt different metabolic profiles to maintain tumorigenesis. For example estrogen receptor positive breast cancer may prefer oxidative phosphorylation (OXPHOS) for energy production but triple negative breast cancer (TNBC) usually shows a classical Warburg effect (Choi et al., 2013). Prostate cancer uses both OXPHOS and glycolysis for energy production and exhibits switches in energy metabolic during different stages of tumor progression (Elia and Fendt, 2016). Also, association between proliferating and non-proliferating sub populations of breast cancer and different activity states of pyruvate kinase suggested impact of glucose metabolism in tumorigenesis (Israelsen et al., 2013). Using intraoperative $^{13}$C-glucose infusions to investigate metabolism in lung cancer patients, Hensley et al. identified metabolically heterogeneous areas within and between tumors, also showed evidence for contributions of non-glucose nutrients in different regions, including lactate as alternative source of carbon (Hensley et al., 2016a). Metabolic cooperation of the intra-tumor sub-populations can help cancer cells to tolerate spatial and temporal fluctuations of the tumor microenvironment. A two-compartment metabolic cross-talk has been proposed to model the relationship between cells present in hypoxic zone with those exist in oxygenated microenvironment. In this model, oxidative cancer cell subpopulations uptake lactate released by glycolytic cells within hypoxic area, as extra fuel to elevate mitochondrial oxidative metabolism. Hypoxic subpopulations that secretes lactate are mainly characterized by high expression of the lactate exporter enzyme monocarboxylate transporters 4 (MCT4) and aerobic subpopulations that uptake lactate are characterized by high expression of the lactate importer enzyme monocarboxylate transporters 1 (MCT1) (Baenke et al., 2015; Yoshida, 2015). Cancer cells are closely interacting with the tumor microenvironment, including endothelial cells, macrophages, T cells and fibroblasts, and this interaction has an important role in shaping the characteristics of tumor cells. Tumor-microenvironment interactions can influence cancer cells metabolic reprogramming to a large extent and induce positive or negative effect on tumor growth and survival. Cancer cells may attempt to actively manipulate the microenvironment, by secreting diverse repertoire of soluble and vesicle-associated factors, to support tumor progression and metastasis, and to resist to therapeutic interventions (Quail and Joyce, 2013).

# Results

The result section can be divided into two main sub-sections: modeling metabolic interactions between species in human gut ecosystem and modeling cancer metabolism at genome-scale.

## Modeling metabolic interactions between species in human gut ecosystem

### CASINO toolbox: multi-objective and multi-dimensional optimization

A collaborative multi-dimensional optimization approach has been employed to design CASINO toolbox. GEMs are based on single-objective linear optimization model, but presence of multiple GEMs capable of interacting within the network of the community systems results in a multi-objective problem with a nonlinear community objective function. Along with nonlinearity of objective function, emerged topology of community network enhances the complexity of the model. The linearity of variables and the convexity of the solution space are conserved by defining three classes of variables: input, output and connecting, distributed into two independent levels of optimization: species-level and community level. The initial topology profile of the community and interaction between species identified through initialization process in species-level, and followed by an iterative multi-dimensional optimization to find final optimum solution. At both levels, maximum biomass production has been used as objective function. At species level, each model tries to independently maximize its biomass yield, while community tries to synchronize competition between individuals to maintain a balanced optimum state (Figure 3).

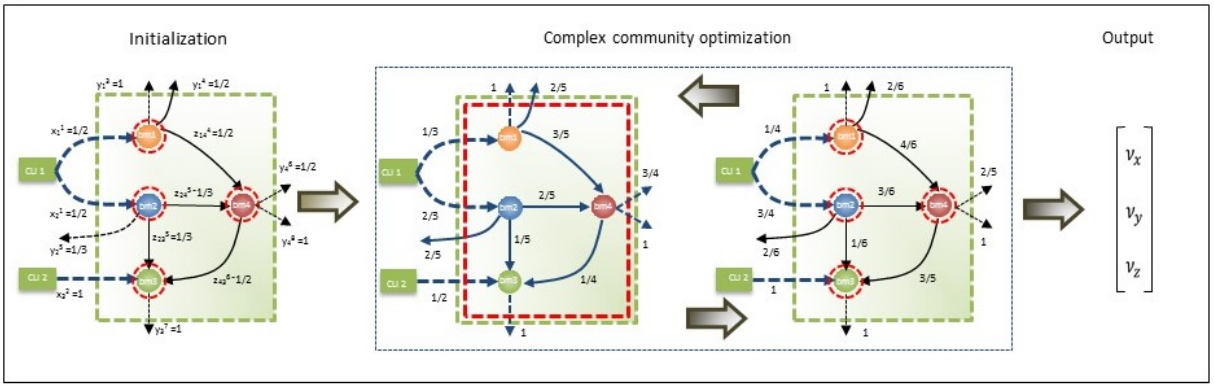


**Figure 3 CASINO toolbox.** The problem solving starts with initialization process to find an initial feasible solution and topology of community network. Using outputs from initialization, optimization process enters an iterative and multi-dimensional loop of community-level and species-level optimization till converging to an optimal solution.

Initialization process defines a network structure for the community and tries to reconstruct a feasible profile of the network topology and interactions through a cascade of activation. The activation cascade starts with classifying species into two classes: primary species which can grow independently using available sources, and non-primary species which are partially dependent on compounds secreted by other species to grow. Initially identified primary species within the community are activated by supplying them available resource to grow. The community resource pool is updated by adding the compounds produced initially by activated species, and the community is rescreened to find subset of the non-primary species that can grow using the updated resource pool. The updated set of primary species are activated and produce compounds are added to the community resource pool. This cascade of activations continues until the whole network has been activated. The initialization process finalizes by reconstructing the topology of the community, identification of the interactions between species, defining the community object function and generating community constraint matrix. At this stage, the defined model is locally optimum (species grow on maximum biomass yield rate using resource allocated to them), but feasible and non-optimum at community-level (resources distributed between species under feasibility conditions but do not satisfy community-level optimality conditions).

Community-level optimization, is a multi-level iterative optimization process which is based on solution space, topology of the community network, constraints and objective function that defined at the end of initialization process. In each intermediate stage, community-level optimization finds the optimum distribution of resources and distributes them to individual models. At species-level, individual models consume allocated resources and secrete compounds at optimum condition, and reshape the boundaries of the intermediate feasible solution space to find a new set of community-level optimal solutions. This iterative process continues till converging to an optimum solution, satisfying a defined threshold.

Objective function is defined at two levels: at upper level maximizing community-level biomass together with connection between species, and at species-level maximizing biomass yield of each individual model. These two objective functions induce two forces, community force and species force, and the summation of these forces defines the direction of optimization within feasible solution space. Power centrality degrees of species calculated based on topology of the reconstructed community network and used to tune the community force.

$$\text{Maximize} \quad S = \quad [\varphi \times \theta'].C_E.\omega_\mu.\delta$$

Subject to

$$s = \alpha.X + \beta.Z^{in} + \gamma.Y + \delta.Z^{out}$$

$$\omega_\mu = \frac{\mu(\mathrm{i})}{\min(\mu_{\mathrm{PL}})}$$

$$\varphi = [\alpha, \beta, \gamma, \delta] \quad : Coefficients\ matrix$$

$$\theta = [X, Y, Z] \quad : Inputs, outputs\ and\ connecting\ parameters$$

$$PL \quad : \quad Primary\ species\ list$$

$\delta$ is a binary vector that activates certain parameter in the objective equation. The relative centrality scores of species ($C_E$) is used to adjust binary vector. μ(i) is the biomass yield of individuals and $\mu_{PL}$ is biomass yield of species belong to PL.

## Analyzing diet-induced changes in composition and metabolism of the gut microbiota

To evaluate the predictive power of the toolbox, we used CASINO to predict abundances of species and the interactions between them in two *in-silico* microbial communities including FBBR (*Faecalibacterium prausnitzii, Bifidobacterium adolescentis, Bacteroides thetaiotaomicron and Ruminococcus bromii*) and EBBR (*Eubacterium rectale, Bifidobacterium adolescentis, Bacteroides thetaiotaomicron and Ruminococcus bromii*). The model was able to predict the contribution of individual species to overall phenotype of each community as well as consumption and release profile of the metabolites within communities. Next, we simulated the effect of diet intervention on 45 overweight and obese subjects that have been stratified into two groups based on their gut microbial gene richness, high gene count (HGC; n = 27) and low gene count (LGC; n = 18), with threshold of 480,000 genes (Le Chatelier et al., 2013). Previously, it was shown that microbial gene richness of subjects influences the phenotypic response to dietary interventions (Cotillard et al., 2013). Individuals in this study subjected to 6 weeks of the energy restricted high protein diet, and we used reconstructed GEMs for abundant bacteria *B. adolescentis, B. thetaiotaomicron, E. rectale, F. prausnitzi* and *Lactobacillus reuteri* to model the changes in human gut microbiota at baseline and after dietary intervention. To translate foods to metabolites, the dietary macronutrients of diets were calculated and used as inputs for GEMs and CASINO toolbox. Our simulations predicted the profiles of the 14 amino acids and three short chain fatty acids (SCFAs) secreted by gut ecosystems and contribution of the each species to the overall phenotype, before and after dietary intervention. We observed a general reduction in levels of amino acids and SCFAs after 6 weeks of energy restricted diet, with higher rate of decrease within LGC group compared to HGC group. (Figure 4).

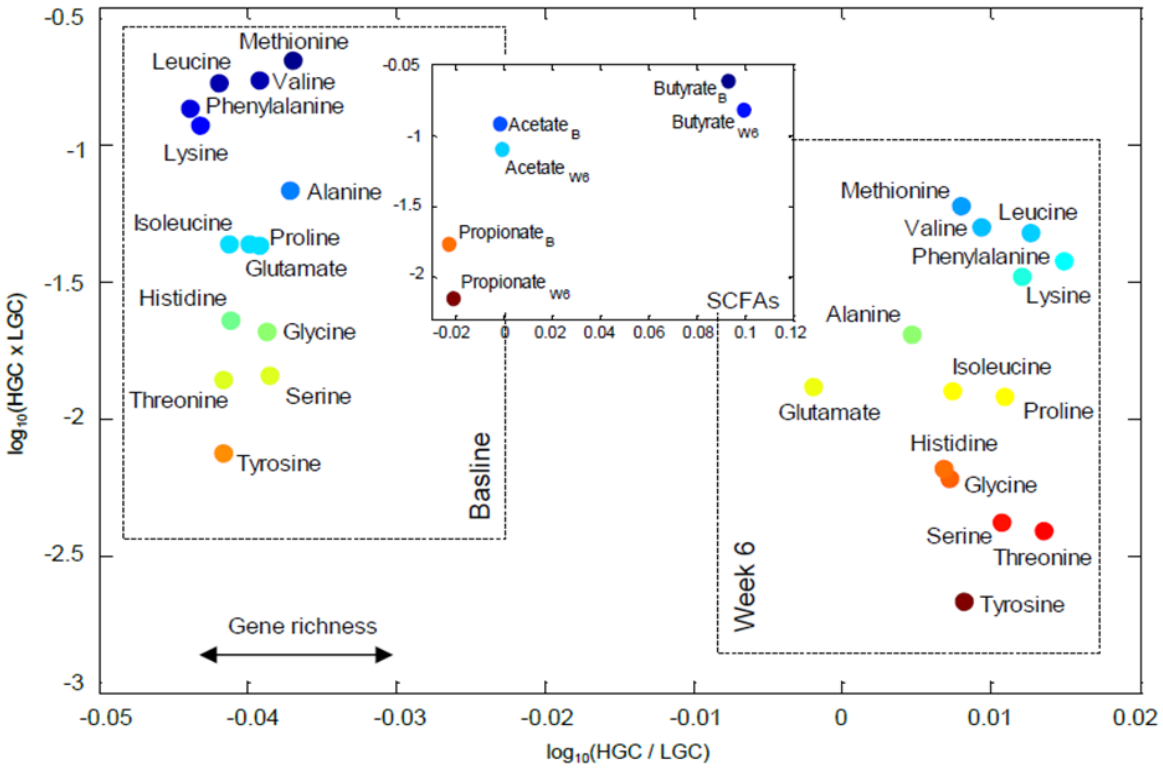


**Figure 4. Predicting the effect of the dietary intervention on fecal metabolomics of LGC and HGC individuals.** Summary of average phenotypic predictions for baseline and after 6 weeks. Metabolites clustered at top left shows prediction at baseline and the group gathered at bottom right represents predictions after six weeks of dietary intervention. The x axis shows the ratio of predicted level of metabolites between HGC and LGC, and the y axis shows the sum of predicted level of metabolites in the two groups. The colors show the metabolites' distance from zero on the y axis (from dark blue at the top to dark red at the bottom).

Next, we performed simulations to design a diet required to change gut metabolism for LGC individual, assuming as non-optimal phenotype, to improved gut phenotype of HGC individuals. Using the abundance of five species *F. prausnitzi, E. rectale, thetaiotaomicron, B. adolescentis and L.reuteri* , the relative consumption of the eight essential amino acids were predicted for LGC subjects at baseline and for HGC subjects after 6 weeks of dietary intervention. We observed that improved phenotypic state is associated with increased consumption of essential amino acids (Figure 5A). The emerged pattern of amino acids consumption rates was correlated with amino acids content of different food types, and the direction of the correlation was used to find the contribution of dietary consumptions to improve phenotype. The results showed that intake of vegetables, dairy products and white meat have positive effect on improving the phenotype, while rice, bread and pastries have inverse effect (Figure 5B).

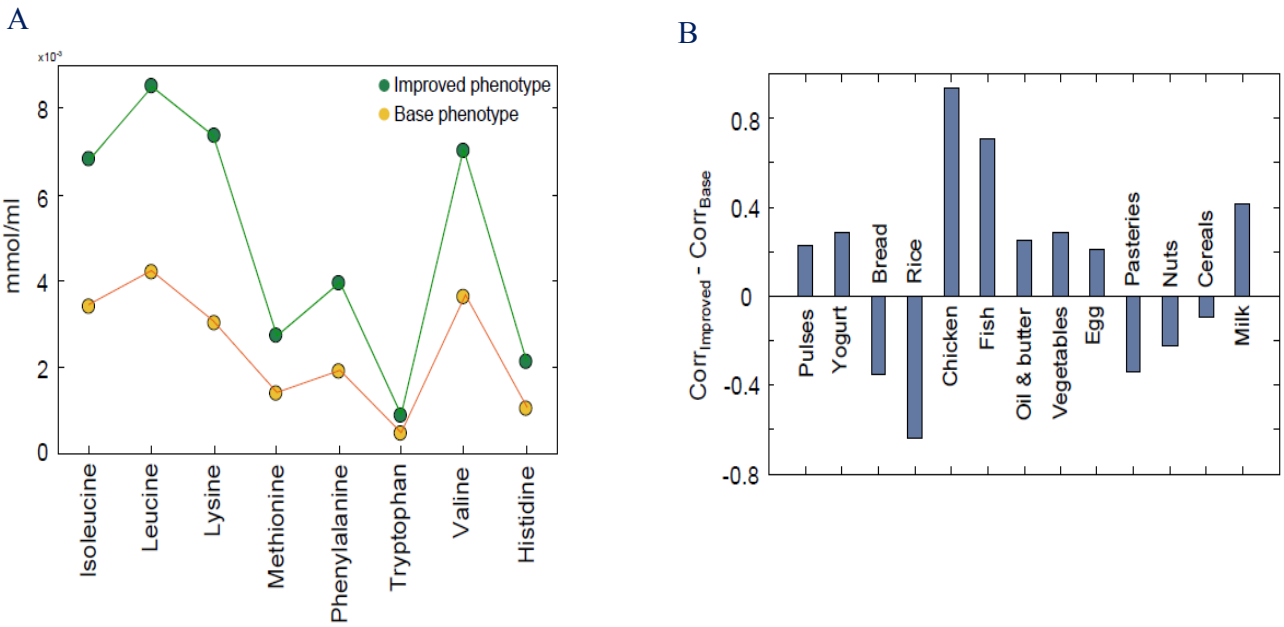


**Figure 5. Modeling dietary composition to improve phenotype.** A. Simulated consumption of eight essential amino acids by gut microbiome of LGC individuals at baseline (yellow line) and by HGC at week six (green line). B. Correlation between amino acids contents of different food categories and baseline/improved phenotypes. The direction of $Corr_{Improved}$ - $Corr_{Base}$ points to positive or negative effect of each food to improve phenotype.

# Modeling cancer metabolism at genome-scale

## Identifying anti-growth factors for human cancer cell lines

Established human cancer cell lines are routinely used as model systems to study human cancers in simplified laboratory conditions. Cell lines derived from original human tumors of different tissue of origin, at one time in the history of tumorigenesis, and adapted for *in vitro* growth in extended periods of time. Human cell lines are preferred because they are pure, easily proliferate and can be genetically manipulated, when same experimental protocols are used provide reproducible results, and identified mechanisms by analyzing their perturbation-response patterns can often be extrapolated to behavior of human tumors *in vivo*. The acquired knowledge from *in vitro* analysis of human cancer cell lines can be used to identify biomarkers, drug targets and therapeutic agents' mechanisms of actions (Jain et al., 2012; Moghaddas Gholami et al., 2013).

In this study, we used mRNA expression profiles of the 20,314 protein coding genes from eleven human cancer cell lines to reconstruct functional cell line specific GEMs (CL-GEMs). Human Metabolic Reaction database 2 (HMR2) and 56 metabolic functions are known to occur in all human cell types have been used to reconstruct the models. The functionality of the generated models have been validated by using consumption and release (CORE) profiles the cancer cell lines (Jain et al., 2012). The resulting CL-GEMs contained 4,209 to 4,432 metabolites and 5,297 to 5,584

reactions associated with 2,193 to 2,328 genes. Pair wise comparison of the models showed each cell line has an average of 272 metabolites, 353 genes and 517 reactions different from other cell lines, and the glioblastoma cell line (U-251 MG) and hepatocellular carcinoma cell line (Hep-G2) are the ones with highest degree of difference based on reactions and genes, respectively (Figure 6).

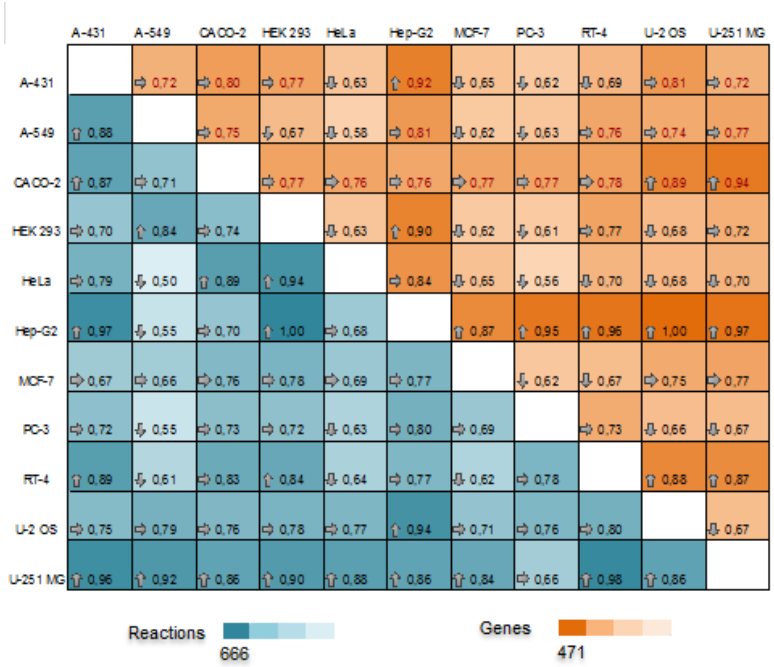


**Figure 6. Pair wise comparison of CL-GEMs.** Ratios of pair wise difference between CL-GEMS compared to maximum observed difference across all models has been presented for reactions on the lower triangular part of matrix and for genes on the upper triangular part of matrix.

We defined a heterogeneity degree based on average hamming distance of the constituent parameters in reconstructed models and used it to investigate the divergence between CL-GEMs as well as to compare CL-GEMs with previously reconstructed GEMs for healthy cell types. For each model, heterogeneity degree formulated as:

$$d^{h}_{ij} = \frac{\log_2(\bar{h}_{ij})}{\log_2(H_{ij})} \qquad i = 1\ldots n,\ j = 1\ldots m$$

Where $d^{h}_{ij}$ is the heterogeneity degree of model i based on parameter j, $\bar{h}_{ij}$ is the average Hamming distance of model i to all other models based on parameter j, $H_{ij}$ is maximum Hamming distance of model i based on parameter j in comparison to integrated vector of parameter j ($Iv_j$). In this formulation, n is the number of models, m is the number of the constituent parameters (here parameter are genes, reaction and metabolites, m = 3), and $Iv_j$ is a unique union of corresponding parameter across all models. For example,$I.v_1$ represent a union vector of all genes present at least in one of the models.

Cell lines experienced a fall in divergence of metabolic networks comparing to healthy cell types based on reactions and genes, revealing a tendency towards uniformity of metabolism. However, the observed loss of heterogeneity was not significant, and interestingly models preserved their heterogeneity based on metabolites indicating capability of cancer cells to metabolize diverse types of the metabolites (Figure 6). This analysis highlight the importance of developing cancer specific models rather than a generic cancer model, to investigate cancer related metabolic alterations and the mechanism behind.

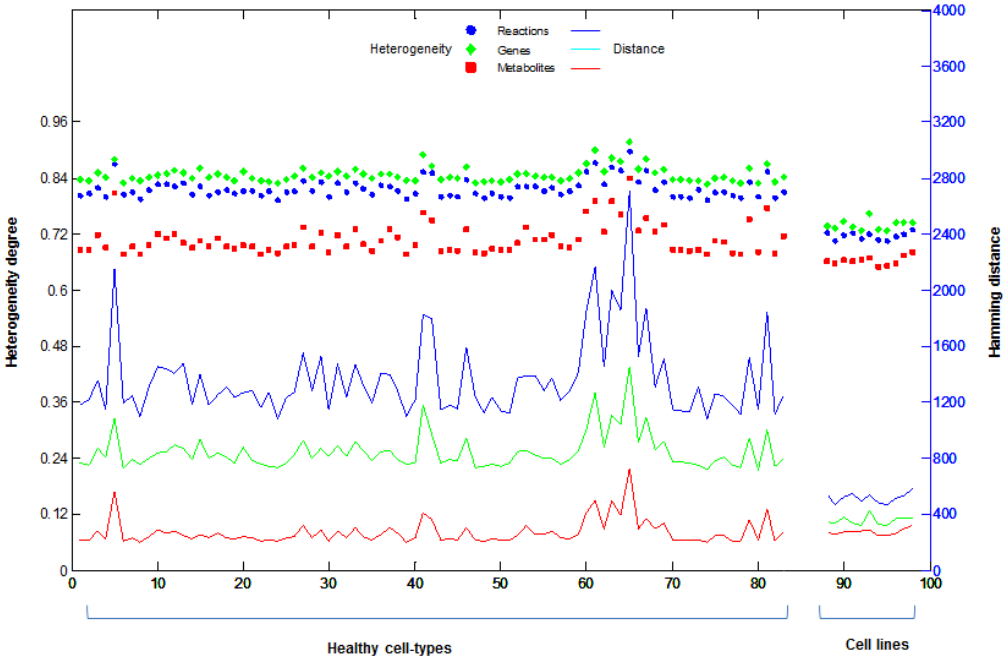


**Figure 7. The heterogeneity of CL-GEMs and healthy cell-types.** Heterogeneity degrees of CL-GEMs and GEMs reconstructed for 83 healthy cell-types are presented by dots and projected on the left hand side axis, and the hamming distance between models are demonstrated by lines and projected on the right hand side axis. There is a slight tendency towards loss of heterogeneity in CL-GEMs comparing to normal tissue models, based on genes and reactions, but models preserve their heterogeneity degree based on metabolites.

We used the concept of antimetabolites to identify antigrowth factors, by inducing growth inhibiting perturbations into the reconstructed CL-GEMs. Antimetabolites are structural analogue of endogenous metabolites that can disrupt the cellular process by tricking and inhibiting all associated enzymes. Antimetabolites, such as antipyrimidines (e.g. Cytarabine, 5-Fluorouracil), antifolates (e.g. Methotrexate) and antipurines (e.g. 6-Mercaptopurine), are among the earliest and most commonly used chemotherapeutic agents since their discovery (Hebar et al., 2013; Lazar and Birnbaum, 2012). Perturbing metabolic network by testing essentiality of all metabolites, we identified 138 antimetabolites that can inhibit the growth of any of eleven CL-GEMs. Next, we performed *in silico* toxicity test by checking the essentiality of detected antimetabolites for energy process in GEMs reconstructed for healthy cell types, and ended up with 85 potential antigrowth

factors. From 85 identified antimetabolites, 60 were effective on all CL-GEMs and remaining 25 were effective one maximum two cell lines.

Analogue of L-carnitine, one of the 60 antimetabolites predicted to be effective on all eleven cell lines, was selected for further valuation. L-carnitine plays an essential role in fatty acids biosynthesis and mitochondrial energy metabolism. Carnitine-shuttle system transfers long chain fatty acids across the mitochondrial membrane to be oxidized through β-oxidation which results in cyclical shortening of fatty acids and production of acetyl CoA, NADH and FADH (Carracedo et al., 2013). Two cell lines with distinct phenotypic origins, prostate carcinoma cell line (*PC-3*) and epidermoid carcinoma cell line (*A-431*), were selected for *in vitro* evaluation using perhexiline malate salt (perhexiline) to mimic the effect of L-carnitine analogue. Perhexiline inhibits carnitine palmitoyltransferase 1 (*CPT1*) which form conjugated fatty acids-carnitine to be translocated from cytosol to the mitochondria (Figure 8).

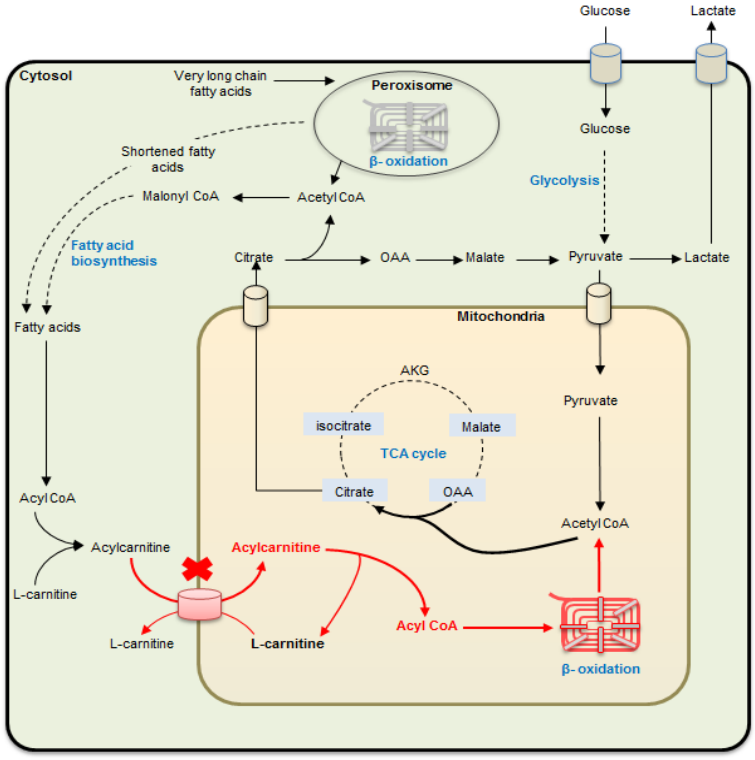


**Figure 8. The predicted mechanism of action of an L-carnitine analogue.**

L-carnitine was predicted as an essential metabolite, and the use of its analogue was proposed for inhibiting the growth in all eleven human cancer cell lines. L-carnitine antimetabolite may inhibit mitochondrial β-oxidation of fatty acids and reduce *de novo* fatty acids biosynthesis, which are required for energy production and synthesis of the cell membrane.

We evaluated growth inhibitory effect of perhexiline on cell lines by implementing four concentrations (2, 4, 8 and 24 μM) at two time points, 24 and 48 hours, with eight replicates. Both cell lines experienced significant decrease (t-test, p-value = 0.05) in cells viability in presence of perhexiline with concentrations more than 4 μM, validating the relevancy of predicted antimetabolite (Figure 9).

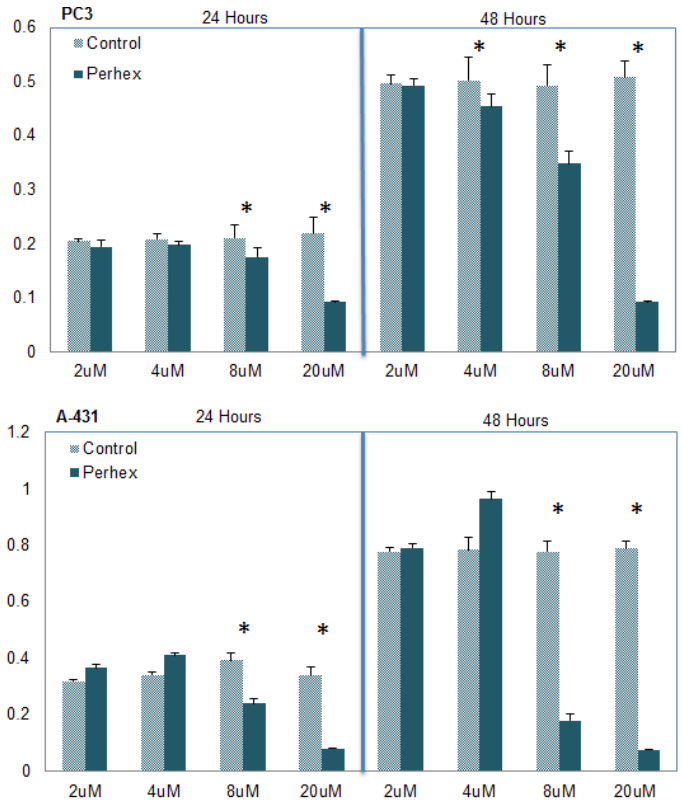


**Figure 9. Inhibitory effect of Perhexiline on PC3 and A-431 cell lines.**

Inhibitory Effect of Perhexiline on viability of epidermoid carcinoma cell line, A-431 and prostate carcinoma cell line, PC-3, were evaluated by 2, 4, 8 and 20 µM of Perhexiline at 24 and 48 hours. Results for analyzing eight replicates of all concentrations and corresponding controls were demonstrated by bar plots, mean ± standard deviation. Significance of difference indicated by star symbol * (Student's t-test, p-value = 0.05).

## Heterogeneity of amino acids metabolism affects lung adenocarcinoma prognosis

In this study, we assessed cancer-related metabolic transformations by analyzing mRNA expression profile of 108 lung adenocarcinoma (LAC) tumors and 10 matched control samples. Using generated RNAseq data, we reconstructed one generic metabolic model, two gender-specific models (male and female), four stage based models (stage1 to 4) and 3 smoking-based models (smokers, ex-smokers and never smokers) for LAC, and one generic model for normal tissue samples (Figure 10A). Reconstructed GEMs have been used to identify potential growth inhibiting factors, and after performing *in silico* toxicity test, we predicted 58 potential non-toxic antimetabolites for lung cancer and classified them based on presence in associated metabolic pathways (Figure 10B).

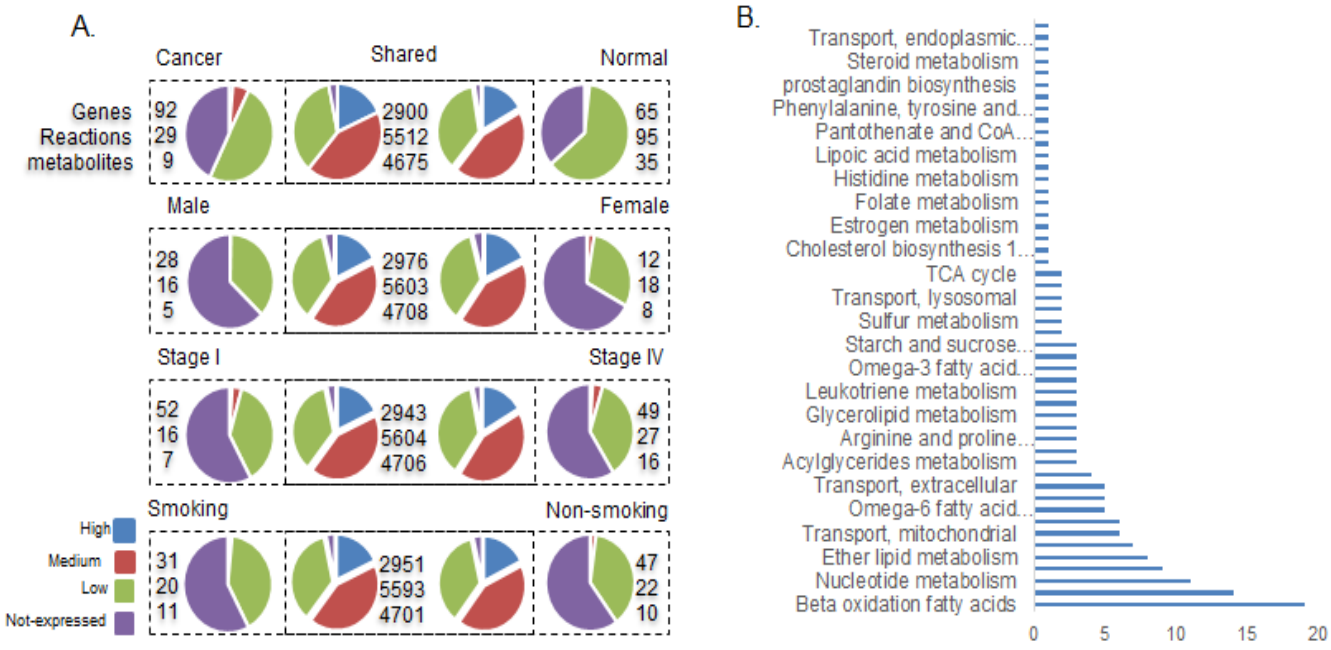


**Figure 10. Reconstructed GEMs for lung cancer and predicted antimetabolites.**

**A**. Distribution of genes, reactions and metabolites in generic models reconstructed for cancer and normal samples, as well as, GEMs reconstructed based on different clinical parameters. Cancer and normal GEMs are compared based on similarity and difference of genes, reactions and metabolites, and expression level of genes associated to reactions classified in four levels: not-expressed (FPKM < 1), low (1=< FPKM<10), medium (10=<FPKM<40) and high (40=<FPKM). **B**. distribution of predicted 58 antimetabolites based on their presence in metabolic pathways. X-axis represents number of anti-metabolites and y-axis specifies corresponding metabolic pathway.

We followed a systemic approach to investigate the scale and direction of heterogeneity of metabolic reprogramming in LAC patients, at three levels of: metabolic network, biochemical pathways and individual enzymatic reactions. At the level of metabolic networks, we used Kruskal-Wallis one-way analysis of variance to compare expression profiles of the genes included in the models, and calculated p-value = 0.0048 rejected the hypothesis of the similarity of GEMs at significance level of $\alpha = 0.01$. Performing pairwise Ad Hoc analysis using Fisher's LSD test with confidence level of 0.05, we observed that GEMs reconstructed for matched control samples and stage IV samples show highest rate of significance difference (Figure 11).

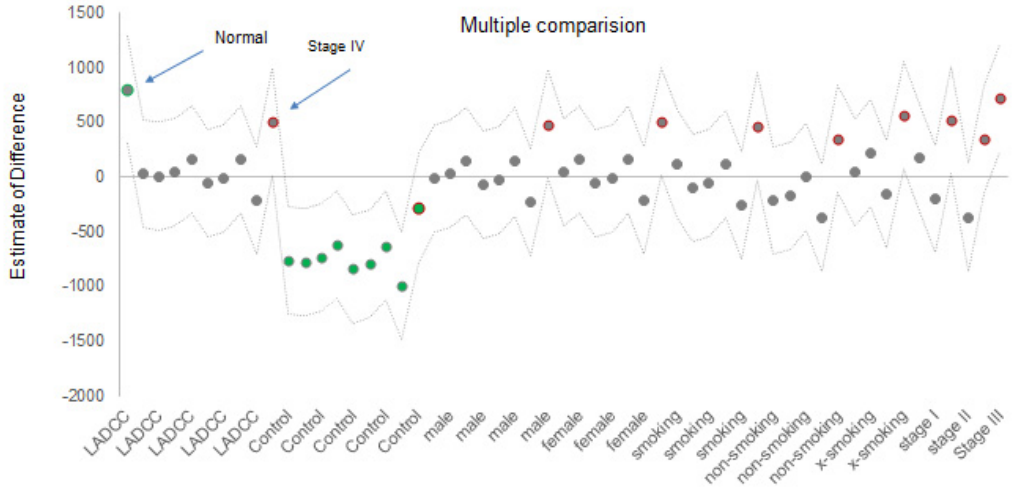


**Figure 11. Global shifts in mRNA expression patterns of reconstructed GEMs for lung cancer.** Markers represent pairwise estimate of difference resulted from multiple comparison analysis and lines show upper and lower bounds. Green markers show normal GEM and red borderline around markers shows stage IV model.

At biochemical pathways level, we calculated the divergence in expression patterns of 126 metabolic pathways define in HMR2 database as:

$$\Delta E_i^j = \log_2 X_i^j - Average\,(\log_2 Y_i) \qquad i = 1\ldots n,\ j = 1\ldots m$$

Where $X_i^j$ is expression vector of genes associated to pathway i in tumor sample j, and Yi is expression vector of genes associated to pathway i in normal control samples. n is number of pathways and m is the number of samples. Wilcoxon signed ranked test with p-value = 1e-03 has been used to determine significant up/downregulation of metabolic pathways using over calculated $\Delta E$. Fraction of significantly upregulated (H) and downregulated (L) samples were calculated and the results were transferred to (H+L) and (H-L) coordinates, to observe the patterns of up/downregulation within samples. As expected pathways necessary for cell division and proliferation, such as nucleotide metabolism, purine and pyrimidine metabolism, were frequently upregulated. On the other hand, fatty acids metabolism showed frequent downregulation. Genes associated with amino acids metabolism demonstrated heterogeneous patterns of expression with slight tendency towards upregulation, except for arginine and proline metabolism which was one of the most frequently upregulated metabolic pathways, placing within the upper 96 percentile of all 126 metabolic pathways studied here (Figure 12).

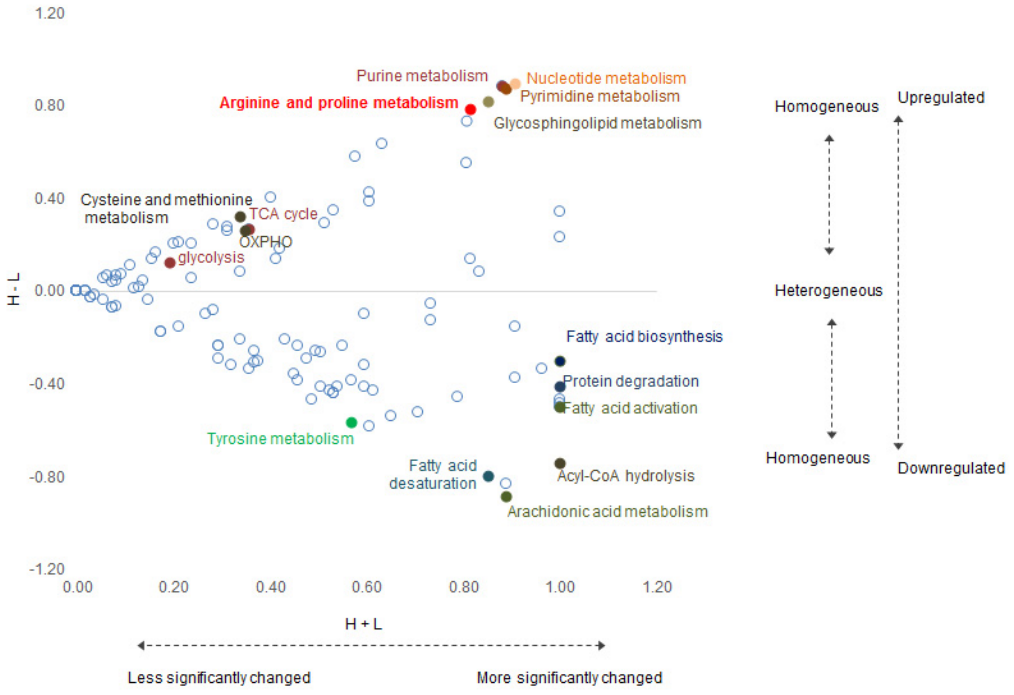


**Figure 12. Cancer related transformations in expressional profiles of metabolic pathways.** Genome scale heterogeneity of metabolic pathways across LAC patients are shown in H+L and H-L coordinates. H is fraction of patients in which a pathway is significantly upregulated and L is fraction of patients in which a pathway is significantly downregulated. Wilcoxon signed ranked test adjusted for multiple hypothesis testing. P-value<0.01 has been used to determine the significance of changes.

To get further insight into arginine and proline metabolic pathway, we used the quartile coefficient of dispersion (Bonett, 2006) to analyze variation of expression of protein coding genes associated with each biochemical reaction within pathway. We found that in contrast to general pattern of significant upregulation of pathway in LAC samples, argininosuccinate synthase 1 (*ASS1*) which catalysis the conversion of citrulline to argininosuccinate in arginine biosynthesis, shows high variation of expression across the LAC patients. Therefore, we stratified patients based on *ASS1* expression level into two groups, upper and lower 15 percentile groups. Previously, loss of *ASS1* activity has been reported in prostate cancer, malignant melanoma, hepatocellular carcinoma and mesothelioma (Long et al., 2013; Patil et al., 2016; Phillips et al., 2013; Rabinovich et al., 2015). Interestingly, here we found that *ASS1* shows a dual activity states in LAC patients, samples in lower percentile group shows loss of activity compared to control samples but upper percentile group shows average of 2.5 fold upregulations compared to control samples (Figure 13A). Next, we did survival time analysis using Kaplan-Meier estimator, to investigate association between *ASS1* variation of expression and life expectancy of LAC patients. From this, we found significant decrease (Mann–Whitney U test, p-value = 0.05) in survival time patients placed in lower percentile

group compared with upper percentile group, going down from median survival time of 1950 to 1270 days (figure 3C). We confirmed this findings analyzing RNAseq data form 490 LAC samples and 50 matched control samples available in TCGA database.

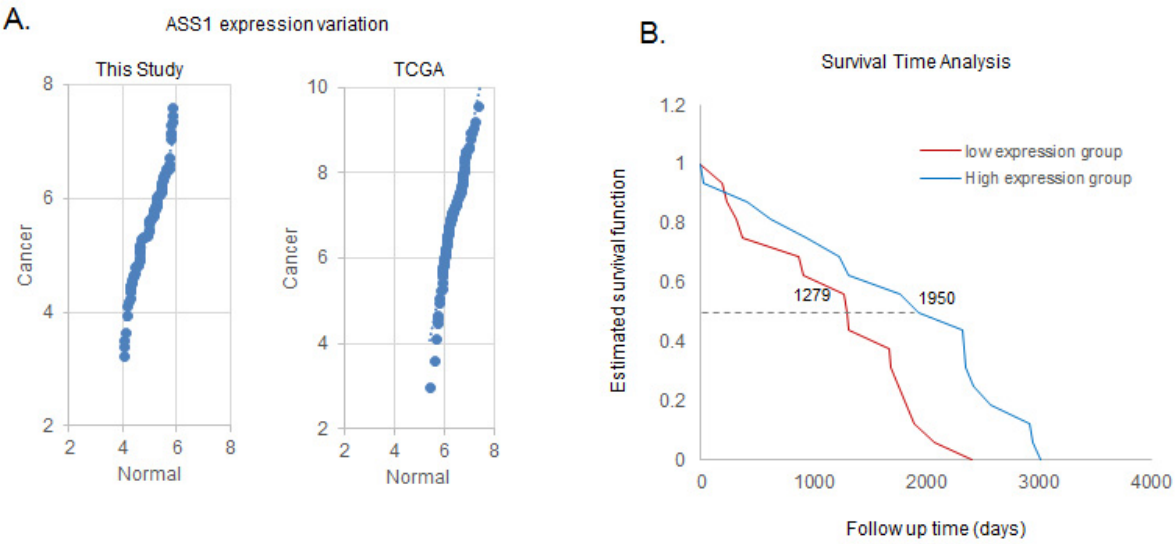


**Figure 13. Expression of *ASS1* in LAC and control samples.** A. q-q- plot shows range variation in expression of *ASS1* within cancer and normal samples, analyzed in this study and from TCGA database. Cancer samples shows 2.5 fold and 3.5-fold higher expression variation in our study and TCGA data, respectively. B. survival time analysis. Survival time of patients stratified by expression of *ASS1* in two groups: upper and lower percentile groups. Dash line demonstrated median survival time of each group in days

Following these observations, we continued with analyzing differences in metabolism of patients stratified by expression of *ASS1* into upper and lower percentile groups. We found significant upregulation of carnitine shuttle, β-oxidation of fatty acids, acyl-CoA hydrolysis and TCA cycle in upper percentile group compared with lower percentile group. This indicates more active mitochondria and elevated oxidative energy metabolism in upper percentile group. Analyzing expression pattern of the biochemical reactions within arginine and proline metabolic pathway, we found downregulation of the arginine catabolizing enzymes arginase and nitric oxide synthase, upregulation of the creatine kinase, which catalysis the reversible conversion of the creatine to phosphocreatine, in upper percentile group (Figure 14A). Phosphocreatine is a high energy molecule that can diffuse from mitochondria to cytosol and be converted to creatine resulting in ATP production. Phosphocreatine can serve as rapidly releasable reserve of energy during short term and high energy demanding periods (Schlattner et al., 2006). This findings indicates a possible mechanism of energy fluctuations tolerance in tumor cells within upper percentile group, where

they activate phosphocreatine shuttle to maintain a buffer for ATP production under effect of hypoxic condition (Figure 14B).

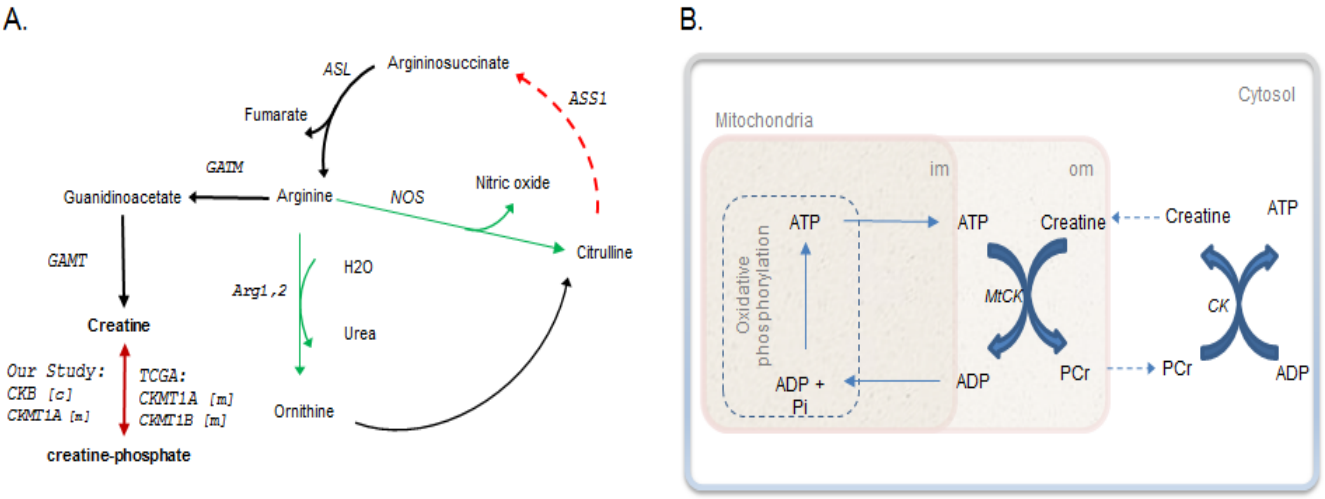


**Figure 14. Differences in metabolism of patients stratified by *ASS1*.** A. in upper percentile group of patients stratified by expression level of ASS1, only the path from arginine to creatine-phosphate was upregulated. Figure shows significant upregulation (Mann–Whitney U test, p-value = 0.05) of creatine phosphate shuttle enzymes: mitochondrial creatine kinases 1A and cytosolic creatine kinase B (*CKB*) in our study and mitochondrial creatine kinases 1A and 1B (*CKMT1A* and *CKMT1B*) in the TCGA database. B. Proposed energy buffer system which activated by creatine-phosphate shuttle to tolerate fluctuation in energy supply. Pcr: phosphocreatine; om: outer membrane; im: inner membrane.

We proceed with investigating the association between *ASS1* expression at mRNA and protein levels, using immunohistochemistry to analyze protein expression pattern of the same tissue samples from LAC patients. Consistent with RNAseq data, we found significant upregulation (Mann–Whitney U test, p-value = 1e-2) of protein level in upper percentile group (Figure 15A). Next, we stratified patients based on *ASS1* protein expression level into upper and lower 25 percentile groups, and found significant (Mann–Whitney U test, p-value = 0.05) increase in survival time of the patients placed in upper percentile group compared to lower percentile group, rising from 1324 to 1970 days (Figure 15B).

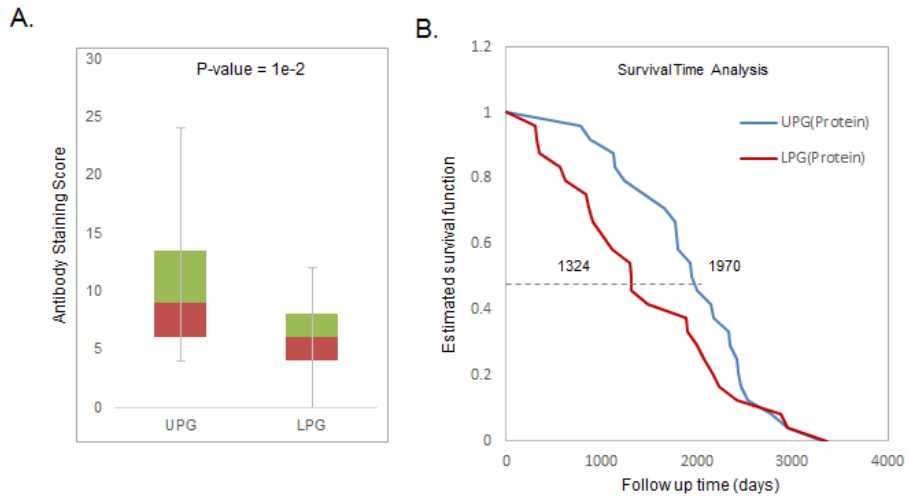


**Figure 15. Protein expression profile of *ASS1* in LAC samples.** A. Change in protein level related to variation of ASS1 transcript level. LAC samples classified into upper and lower 25 percentile groups based on RNA expression data to assess the change in protein level between two groups. Protein level was significantly elevated in upper percentile compared to lower percentile group (Mann–Whitney U test, p-value = 1e-3). B. Survival time analysis. LAC samples stratified into upper and lower 25 percentile groups based on antibody staining score, and Kaplan-Meier estimator was used to compare survival time of the two groups. Blue line represents upper percentile group and red line represents lower percentile group. Median survival time of each group is depicted in the figure by dash line and corresponding value in days. UPG: upper percentile group; LPG: lower percentile group.

## Stratification of lung cancer patients based on heterogeneous expression of *FABP5*

Cancer progresses by a repeated process of genetic diversification and clonal expansion within the adaptive microenvironment of tissue-of-origin. Irregular and divers patterns of tumors initiation, progress and metastasis result in different responses to similar treatments within pathologically defined same subtypes of cancer (Greaves and Maley, 2012). Using distinctive molecular profiles of the cancer cells, generated by high-throughput omics technologies, to classify heterogeneous populations of cancer cells into meaningful biological and clinical subtypes, is one of the main goals of the computational and systems biology. Recent attempts in stratification of tumors based on RNAseq data have resulted in discovery of the new subtypes in breast cancer, ovarian cancer and glioblastoma (Hofree et al., 2013).

In this study, we followed a systemic approach to investigate the association between prognosis of lung squamous cell carcinoma (SCC) patients and heterogeneous patterns of genes mRNA expression, using RNAseq data generated for 67 SCC samples and 7 normal lung tissue samples. Three consecutive steps of filtrations have been employed to filter genes based on heterogeneity of expression, correlation between heterogeneity and survival time of patients, and consistency of

mRNA expression and protein expression levels. We defined relative quartile coefficient of dispersion (RQCD) to measure comparative variation of expression of protein coding genes. Comparative expression dispersion for each gene is calculated as:

$$RQCD^i = Log_2(\frac{{q_c}^i}{{q_n}^i})$$

Where ${q_c}^i$ is quartile coefficient of dispersion (QCD) of gene i through cancer samples and ${q_n}^i$ is QCD of gene i across normal samples. For each gene, QCD calculated as:

$$q^i = \frac{{Q_3}^i - {Q_1}^i}{{Q_3}^i + {Q_1}^i}$$

Where ${Q_3}^i$ is the third quartile and ${Q_1}^i$ is the first quartile of expression of gene i across samples. To find maximum variant genes (MVG), a cut-off threshold of more than 95 percentile of RQCDs has been used. Next, we stratified patients into two groups, upper and lower 25 percentile groups, based on expression level of MVGs, and used Kaplan-Meier estimator to assess the effect of this stratification on survival time of SCC patients. We used the same procedure to investigate RNAseq data available in TCGA database for 490 SCC and 50 normal control samples, intersected the results and found 26 genes with same properties in both studies. Next, we implement the third filtration using reliability score defined in HPA database, and found 14 out 26 genes with consistency between mRNA and proteins expression levels. These 14 genes show high heterogeneity of expression in cancer samples but relatively homogeneous expression pattern in normal samples, their heterogeneity of expression is associated with significant change in survival time of SCC patients, and have high reliability score. We classified these 14 genes into anti-cancer and pro-cancer groups, based on negative or positive effect of their variation of expression on survival time of patients, which resulted in 6 anti-cancer and 8 pro-cancer genes. Next, we investigate the effect of pairing genes in each group on survival time of SCC patients and observed tendency towards synergetic effect of pairing in pro-cancer group without recognizable pattern in anti-cancer group (Figure 16).

We selected fatty acid binding protein 5 (*FABP5*), the only metabolic gene within the group of pro-cancer genes which showed completely different expression pattern compared with other members of the lipid binding protein family (Figure 17 ), and stratified SCC patients to upper and lower 25 percentile groups based on expression of *FABP5*. We used Mann–Whitney U test (p-value = 0.01) combined with minimum fold change = 3 to investigate global expression changes consistent with *FABP5* variation of expression, and detected pattern of 6 significantly upregulated keratins in upper percentile group (Figure 18A). Considering very low expression level of these keratins in normal samples, and keeping in mind that keratins have widely been used as diagnostic and prognostic markers in tumor pathology, this distinctive pattern of expression has potential to be evaluated as prognostic marker for SCC patients stratified by high expression of *FABP5*.

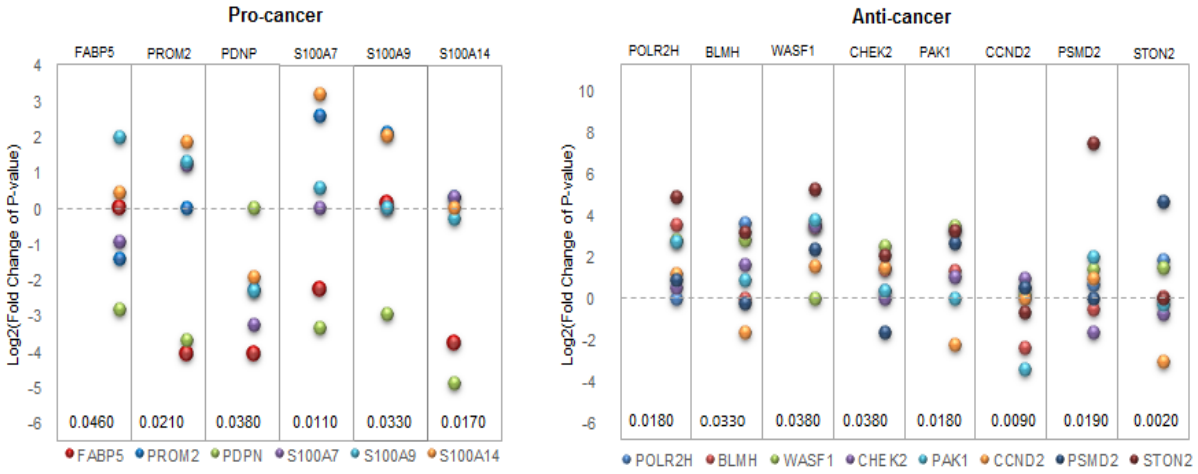


**Figure 16. Association between survival time of patients and stratifying by pairs of genes.** Positive or negative effect of the pairing genes in each group on survival time of the SCC patients has been investigated. Genes placed in pro-cancer or anti-cancer groups paired with other members of the same group, SCC patients stratified by these pairs, and p-value for the significance of difference between survival times of the stratified patients has been calculated. Rows show genes names with p-values before pairing, and columns represent fold change of p-values after pairing in log2 scale. Positive values show more significant stratification of patients after pairing genes and negative values represent invers effect.

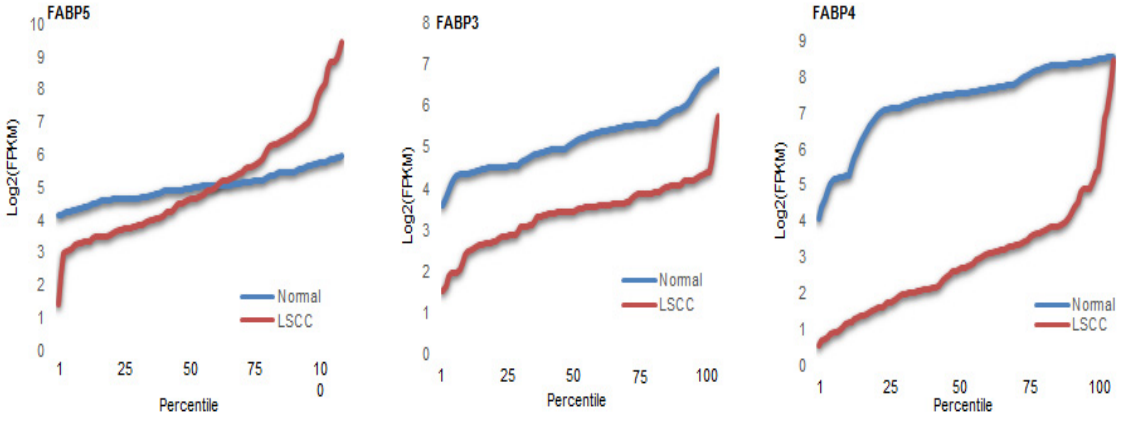


**Figure 17. Expression patterns of members of the lipid binding protein family across SCC and normal control samples.** Positive or negative effect of the pairing genes in each group on survival time of the SCC patients has been investigated. Genes placed in pro-cancer or.

Next we analyzed co-expression pattern of 382 genes, related to cancer pathway in KEGG database, with FABP5, using distance and direction of the change in correlation of expression of each gene

with FABP5 from normal to cancer. We defined correlative distance for gene i in cancer samples from to normal samples as:

$R_i = r_{ci}^b - r_{ni}^b \qquad i = 1 \ldots n$

Where $r_{ci}^b$ is the Pearson correlation of expression of gene i with reference gene b in cancer samples, $r_{ni}^b$ is the Pearson correlation of expression of gene i with reference gene b in normal samples, $R_i$ is the correlation distance of gene i, and n is the number of genes analyzed. Positive values of Ri, defined here as gain of correlation, shows increased correlation of expression with reference gene moving from normal state to cancer, and negative values of Ri, define here as loss of correlation, shows decrease correlation of expression with reference gene moving from normal state to cancer. Projecting calculated correlative distances on correlation-correlation (Corr-Corr) plot revealed that hypoxia inducible factor 1 alpha subunit (*HIF1α*), placental growth factor (*PGF*) and fibroblast growth factor 11(*FGF11*) are among the group with highest gain of correlation, and epidermal growth factor (*EGF*) and growth factor receptor bound protein 2(*GRB2*) belong to the group with highest loss of correlation (Figure 18B).

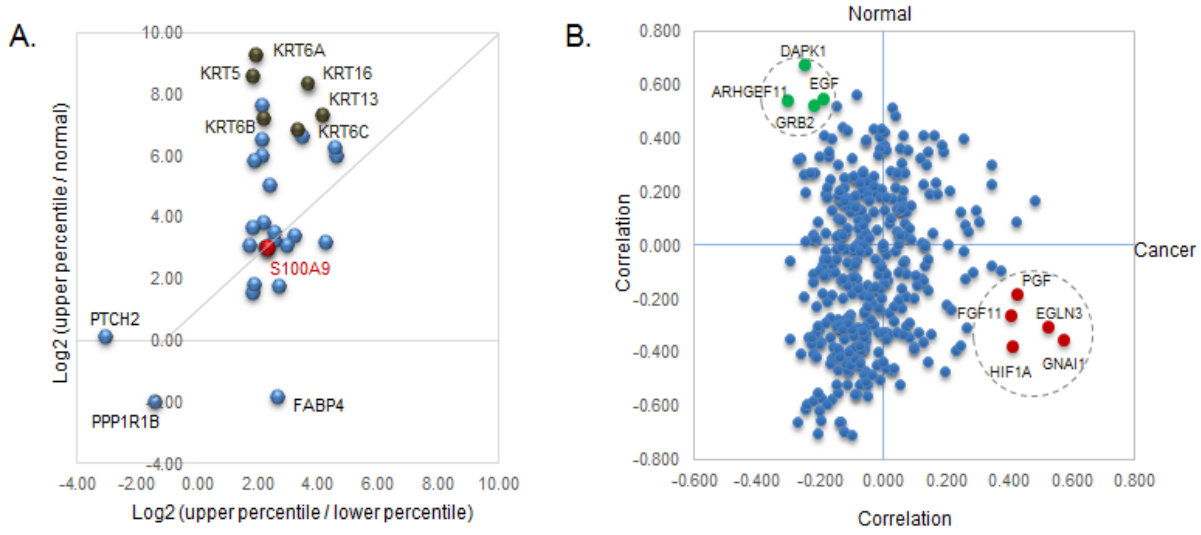


**Figure 18. Patterns of expression change consistent with *FABP5*.** A. Global expression change consistent with *FABP5* upper and lower percentile groups. 27 genes were detected to be significantly upregulated and 2 genes were detected to be significantly downregulated in upper percentile group, using Mann–Whitney U test with p-value = 0.01 and applying minimum 3 fold change cut-off. Vertical axis shows fold change of expression in upper percentile group compared with normal samples and horizontal axis shows fold change of expression in upper percentile group compared with lower percentile group, in Log2 scale. B. Co-expression pattern of cancer related genes with *FABP5*. Changes in expressional correlation of cancer related genes with *FABP5,* from normal samples to cancer samples, has been demonstrated in the plot. Genes with highest gain of correlation in SCC samples clustered in lower-left corner of plot, and genes with highest loss of correlation in SCC samples placed in upper-left corner of plot. Red markers: highest gain of correlation; green marker: highest loss of correlation.

We continued with investigating differences in metabolism of patients stratified by *FABP5* into two sub-groups. At the level of biochemical pathways, we found upregulation of the glycolysis, and fatty acids activation in endoplasmic reticulum (ER), and downregulation of the fatty acid biosynthesis and mitochondrial carnitine shuttle in upper percentile group. At the level of biochemical reactions and within the central metabolism, Glucose transporter 1 (*GLUT1*), monocarboxylate transporter1 (*MCT1*) which transports lactate, and the fatty acid translocase FAT/CD36 were significantly overexpressed. On the other hand, mitochondrial pyruvate carrier 1 (*MPC1*) and Pyruvate Dehydrogenase Kinase 1 (*PDK1*) were significantly downregulated in upper percentile group of patients. Overexpression of *GLUT1* and *MACT1*, and downregulation of the *MPC1* indicates high glycolytic activity. Consistent with this change, pyruvate Dehydrogenase Kinase 1 (*PDK1*) which regulates metabolic flux through TCA cycle by inhibiting formation of acetyl-CoA from pyruvate, and *HIF1α* which regulates glycolysis flux, were significantly overexpressed in upper percentile group. Also, we observed overexpression of Acyl-CoA Synthetase 3 (*ACSL3*) which activates acyl-CoA in ER and diacylglycerol o-acyltransferase 2 (*DGAT2*) which catalysis the rate limiting reaction in synthesis of triglycerides, and downregulation of monoacylglycerol lipase (*MGLL*) which regulates intracellular level of fatty acids by mobilizing fatty acid from lipid droplets in upper percentile group.

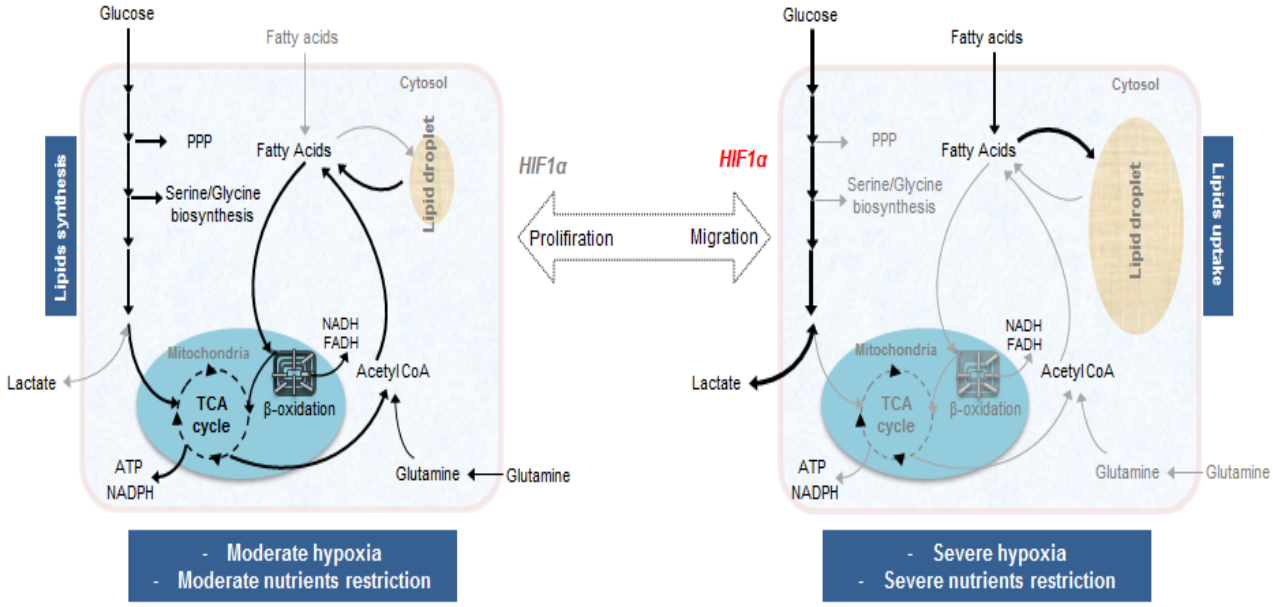


**Figure 19. Metabolic switch in SCC patients stratified by *FABP5*.** Figure depicts main differences in central metabolism of SCC patients classified in two upper and lower percentile groups by expression level of *FABP5*.

All these observations pointed to a bi-state model of fatty acid metabolism in SCC, proliferation state and migration state, which is detectable by expression of *FABP5*. Proliferation state is characterized by fatty acid oxidation and *de novo* biosynthesis under moderate hypoxia. Migration

state is characterized by elevated fatty acids uptake and storage under sever hypoxia (Figure 19). Possibly, recurrent periods of hypoxia and re-oxygenation induce selective pressure in favor of tumor cells to downregulated energy-intensive fatty acid synthesis and accumulate fatty acids in lipid droplets that can be used as energy source upon re-oxygenation. Our findings, provide mechanistic insights into how different directions of metabolic reprogramming within SCC patients can be characterized by heterogeneous expression patterns of particular gene. This property has important implications for personalized treatment, and can be used for stratification of new sub-classes of cancers and identification of potential new drug targets.

# Conclusion

This thesis aimed to use computational and systems biology approach to get new insights into how complex biological systems reprogram their structure and activities to adapt for internal and external perturbations.

In **Paper I**, we described how genome scale metabolic modeling can be used to explore metabolic interactions between the predominant subpopulation of human gut microbiota and its host. We developed CASINO toolbox, a computational platform based on multi-objective and multi-dimensional optimization approach, to simulate gut microbiota composition and interactions through a dietary intervention. We identified association between consumption of 8 essential amino acids and healthier phenotype of gut microbiota. We also showed that CASINO can be used to predict dietary change compatible with specific phenotypic transformation in gut microbiota, e.g. from non-healthy to healthy phenotype. CASINO can be improved and extended towards a more comprehensive platform capable of designing personalized dietary interventions and/or predicting beneficial probiotic supplements for treatment of the specific metabolic disorders which are related to gut microbiota.

In **Paper II**, we demonstrated potential of genome scale metabolic modeling on predicting effective growth inhibiting factors for human cancer cell lines. We used mRNA expression data and reconstructed cell line specific genome scale metabolic models (CL-GEMs) for eleven human cancer cell lines with different phenotypic origin. We used generated CL-GEMs to predict potential antimetabolites, and experimentally validated growth inhibitory effect of the L-carnitine analogue, one of the identified targets. The study of human cancer cell lines metabolism using CL-GEMs has revealed promising performance of the GEMs in identification of new therapeutic targets, as well as gaining more mechanistic insight into cancer-related metabolic reprogramming. This approach can be used and be extended to study the effects of combining established and new therapeutic agents to design more effective cancer-specific therapeutic strategies.

In **Paper III**, we used context-specific genome scale models, reconstructed using mRNA expression profiles of lung adenocarcinoma (LAC) and normal lung tissues samples, to study genome-scale heterogeneity of the metabolic reprogramming in lung cancer. Investigating direction of changes in cancer-related metabolic reprogramming at three levels of whole metabolic network, biochemical pathways and reactions, revealed homogeneous behavior of pathways necessary for cell division and proliferation, and heterogeneous behavior of glycolysis and majority of amino acids metabolism across lung cancer patients. We identified divergent regions within arginine and proline metabolism, one of the pathways with high homogeneous pattern of expression, and found significant difference between survival times of patients stratified by expression level of the argininosuccinate synthetase 1 (*ASS1*), one of the enzymes with extreme heterogeneity of expression. We observed a difference in energy metabolism of patients stratified by ASS1 expression which can explain significant change in survival rate of patients.

Following observing association between heterogeneous expression of an enzyme and survival time of the patients in previous study, in **Paper IV** we proceed with defining a systemic approach to investigate correlation between heterogeneity of expression and cancer prognosis. For this study, we used generated RNA sequencing (RNAseq) data from lung squamous cell carcinoma (SCC) biopsies and control samples and implemented three consecutive filters that resulted in 14 genes with desired property. These gene were classified into pro-cancer and anti-cancer groups based on association between genes variation of expression and survival time of SCC patients. We found significant difference in survival time of the patients stratified in two groups by expression level of fatty acid binding protein 5 (*FABP5*), and observed distinct behavior of fatty acid metabolism in these groups which was characterized by expression of *FABP5*. Our analysis and findings in papers II and IV, provide mechanistic insights into how different directions of metabolic reprogramming within same sub-types of a specific cancer, can be projected in and be characterized by heterogeneous expression patterns of particular genes. This property has important implications for personalized treatment, and can be used for stratification of new sub-classes of cancers and identification of potential new drug targets.

# Perspective

Considerable progress has been made in the past two decades to understand the mechanisms underlying metabolic reprogramming in cancers and biological and medical consequences and liabilities of this transformation. A common theme that has arisen from studies on cancer biology is essentiality of metabolic reprogramming for biology of tumors, especially to ensure cancer cells ability to proliferate and survive by manipulating metabolic pathways to supply energy, produce biosynthetic precursors, and maintain redox homeostasis. Furthermore, comprehensive molecular profiling of cancer cells using high-throughput technologies and wide-range studies on cultured tumor cells have revealed a remarkable intra-cancer and inter-cancer heterogeneity of metabolic transformations. We have passed the era when cancer metabolism was simply considered to be the Warburg effect.

Our knowledge on molecular mechanism of cancer related metabolic reprogramming mainly gained by in vitro studies using immortalized cancer cell lines rather than intact tumors. Perhaps, modeling an exact tumor microenvironment in culture is impossible, but we need more advanced culture conditions to cover the in vivo conditions as much as possible. Developing technologies for direct in vivo analysis of metabolic fluxes, along with computational and mathematical modeling platforms to understand the distinct metabolic phenotypes of different cell types (cancer cells, fibroblasts, lymphocytes, macrophages, and endothelial cells) within solid tumors in context of whole microenvironment, will play an important role in detecting metabolic targets and deploying them in clinical trials. This can help us to understand metabolic interactions between different subpopulations of cancer cells within a tumor, and between tumor and stroma, providing new therapeutic opportunities. Studies on cancer metabolism mainly have been focused on metabolic transformations that support cancer cell proliferation and we have learned a lot, but much less is known about the metabolism of non-proliferating malignant cells with high potential for survival under severe tumor microenvironment conditions, and to develop drug resistance. It has been hypothesized that tumors metabolic plasticity can contribute to drug resistance and cancer metastasis, so developing novel experimental methods and computational approaches can help to get more insight into mechanism of plasticity-induced resistance and survival of cancer cells, and to design more effective therapeutic strategies. Our approach on characterizing heterogeneity of metabolic transformations within the same sub-types of lung cancer can be extended by including more molecular profiles and pathophysiological parameters. Developing mathematical models to capture and integrate different facets of cancer heterogeneity can provide important clues about how cancer cells respond to treatment and develop drug resistance. Now, it is clear that cancer risk increases in patients with disease such as obesity and diabetes, but we don’t have enough insight into how these diseases interact with cancer and how to break the links. We believe that approach we followed in simulating metabolic interactions between host and gut microbiota, has high potential to be extended for modeling metabolic interactions between tumor and host in quest for more effective personalized treatment strategies.

In brief, why and how metabolism becomes transformed from normal to neoplastic? Which functions are activated or deactivated by this transformation? What are the metabolic effects of the oncogenes and tumor suppressor genes? What are the main metabolic interactions between cancer cell and microenvironment? How to use detected metabolic reprogramming for therapeutic interventions? And, what is the impact of whole-body metabolism on cancer progression and drug response are among key research-driving questions in the field. These are all key questions going forward and I am confident that systems biology approach, by synthesizing model-driven and data-driven approaches, will assist in answering these questions in the future.

# Acknowledgments

I would like to express my honest appreciation to my supervisor, Jens Nielsen for the opportunity he provided form me to do my PhD, for his support and motivation, and immense knowledge. I met so many people and learned a lot during PhD, thanks for involving me in such a fantastic multi-national and multi-skill group

Special thanks to my co-supervisor, Adil Mardinoglu. You have been present and supported me during most difficult times of my studies. Thanks for all the discussions, guidance, tips and ideas.

My sincere thanks also goes to Saeed Shoaie, with whom I experienced my first project in systems biology. We had a long and fluctuating pattern of cooperation during your studies and later during my studies, but you always have been a good friend and always supported me when needed.

I would like to thank Amir Feizi and Antonio Maras for interesting conversations that we had about science, art and politics. I learned a lot from both of you.

I would like to express my sincere gratitude to Bernhard Mehlig, who introduced me to the world of complex systems and reshaped my thoughts and opinions about physics and mathematics.

This work would not have been possible without the contributions of my co-authors and collaborators, to whom I am very grateful.

I would like to acknowledge and mention my thanks for administrative support that I received from Erica Dahlin, Martina Butorac, Josefine Jäwert and Anna Bolling.

Finally, I would like to thank my family and my old friends, you are always there and I am seeing it.

# References


Auffray, C., Chen, Z., and Hood, L. (2009). Systems medicine: the future of medical genomics and healthcare. Genome Med *1*.

Auger, P.M., and Roussarie, R. (1994). Complex Ecological Models with Simple Dynamics - from Individuals to Populations. Acta Biotheor *42*, 111-136.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. Science *307*, 1915-1920.

Baenke, F., Dubuis, S., Brault, C., Weigelt, B., Dankworth, B., Griffiths, B., Jiang, M., Mackay, A., Saunders, B., Spencer-Dene, B., et al. (2015). Functional screening identifies MCT4 as a key regulator of breast cancer cell metabolism and survival. Journal of Pathology *237*, 152-165.

Barabasi, A.L., and Albert, R. (1999). Emergence of scaling in random networks. Science *286*, 509-512.

Barabasi, A.L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. Nat Rev Genet *12*, 56-68.

Barabasi, A.L., and Oltvai, Z.N. (2004). Network biology: Understanding the cell's functional organization. Nat Rev Genet *5*, 101-U115.

Berg, E.L. (2014). Systems biology in drug discovery and development. Drug Discov Today *19*, 113-125.

Bonett, D.G. (2006). Confidence interval for a coefficient of quartile variation. Comput Stat Data An *50*, 2953-2957.

Boroughs, L.K., and DeBerardinis, R.J. (2015). Metabolic pathways promoting cancer cell survival and growth. Nature Cell Biology *17*, 351-359.

Brand, M.D. (2010). The sites and topology of mitochondrial superoxide production. Exp Gerontol *45*, 466-472.

Cardaci, S., Zheng, L., MacKay, G., Van den Broek, N.J.F., MacKenzie, E.D., Nixon, C., Stevenson, D., Tumanov, S., Bulusu, V., Kamphorst, J.J., et al. (2015). Pyruvate carboxylation enables growth of SDH-deficient cells by supporting aspartate biosynthesis. Nature Cell Biology *17*, 1317-+.

Carracedo, A., Cantley, L.C., and Pandolfi, P.P. (2013). Cancer metabolism: fatty acid oxidation in the limelight. Nature Reviews Cancer *13*, 227-232.

Choi, J., Kim, D.H., Jung, W.H., and Koo, J.S. (2013). Metabolic interaction between cancer cells and stromal cells according to breast cancer molecular subtype. Breast Cancer Res *15*.

Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., Harris, H.M.B., Coakley, M., Lakshminarayanan, B., O'Sullivan, O., et al. (2012). Gut microbiota composition correlates with diet and health in the elderly. Nature *488*, 178-+.

Clemente, J.C., Ursell, L.K., Parfrey, L.W., and Knight, R. (2012). The Impact of the Gut Microbiota on Human Health: An Integrative View. Cell *148*, 1258-1270.

Comerford, S.A., Huang, Z., Du, X., Wang, Y., Cai, L., Witkiewicz, A.K., Walters, H., Tantawy, M.N., Fu, A., Manning, H.C., et al. (2014). Acetate Dependence of Tumors. Cell *159*, 1591-1602.

Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., Parker, S.J., Kamphorst, J.J., Hackett, S., Grabocka, E., Nofal, M., Drebin, J.A., Thompson, C.B., et al. (2013). Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. Nature *497*, 633-+.

Costanza, R., Wainger, L., Folke, C., and Maler, K.G. (1993). Modeling Complex Ecological Economic-Systems - toward an Evolutionary, Dynamic Understanding of People and Nature. Bioscience *43*, 545-555.

Cotillard, A., Kennedy, S.P., Kong, L.C., Prifti, E., Pons, N., Le Chatelier, E., Almeida, M., Quinquis, B., Levenez, F., Galleron, N., et al. (2013). Dietary intervention impact on gut microbial gene richness. Nature *500*, 585-+.

DeBerardinis, R.J., and Chandel, N.S. (2016). Fundamentals of cancer metabolism. Sci Adv *2*.

Dethlefsen, L., and Relman, D.A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proceedings of the National Academy of Sciences of the United States of America *108*, 4554-4561.
Elia, I., and Fendt, S.M. (2016). In vivo cancer metabolism is defined by the nutrient microenvironment. Transl Cancer Res *5*, S1284-S1287.
Erez, A., and DeBerardinis, R.J. (2015). Metabolic dysregulation in monogenic disorders and cancer - finding method in madness. Nature reviews. Cancer *15*, 440-448.
Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., et al. (2013). The Long-Term Stability of the Human Gut Microbiota. Science *341*, 44-+.
Felson, D.T., Callaghan, M.J., Parkes, M., Marjanovic, E.J., Lunt, M., Oldham, J.A., Gait, A., and Hutchinson, C.E. (2011). Bone Marrow Lesions in Knee Osteoarthritis Change in 6 to 12 Weeks. Osteoarthr Cartilage *19*, S10-S10.
Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer *136*, E359-386.
Freeman, W.J., Kozma, R., and Werbos, P.J. (2001). Biocomplexity: adaptive behavior in complex stochastic dynamical systems. Biosystems *59*, 109-123.
Freilich, S., Zarecki, R., Eilam, O., Segal, E.S., Henry, C.S., Kupiec, M., Gophna, U., Sharan, R., and Ruppin, E. (2011). Competitive and cooperative metabolic interactions in bacterial communities. Nature communications *2*.
Galluzzi, L., Kepp, O., Vander Heiden, M.G., and Kroemer, G. (2013). Metabolic targets for cancer therapy (vol 12, pg 829, 2013). Nat Rev Drug Discov *12*, 965-965.
Galluzzi, L., Pietrocola, F., Levine, B., and Kroemer, G. (2014). Metabolic Control of Autophagy. Cell *159*, 1263-1276.
Ghaffari Nouran, P., Mardinoglu, A., and Nielsen, J. (2015). Cancer metabolism: a modeling perspective. Frontiers in Physiology *6*.
Ghaffari, P., Mardinoglu, A., Asplund, A., Shoaie, S., Kampf, C., Uhlen, M., and Nielsen, J. (2015). Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. Scientific reports *5*, 8183.
Gorrini, C., Harris, I.S., and Mak, T.W. (2013). Modulation of oxidative stress as an anticancer strategy. Nat Rev Drug Discov *12*, 931-947.
Greaves, M., and Maley, C.C. (2012). Clonal evolution in cancer. Nature *481*, 306-313.
Griss, T., Vincent, E.E., Egnatchik, R., Chen, J., Ma, E.H., Faubert, B., Viollet, B., DeBerardinis, R.J., and Jones, R.G. (2015). Metformin Antagonizes Cancer Cell Proliferation by Suppressing Mitochondrial-Dependent Biosynthesis. Plos Biol *13*.
Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646-674.
Hanemaaijer, M., Roling, W.F.M., Olivier, B.G., Khandelwal, R.A., Teusink, B., and Bruggeman, F.J. (2015). Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure. Front Microbiol *6*.
Hebar, A., Valent, P., and Selzer, E. (2013). The impact of molecular targets in cancer drug development: major hurdles and future strategies. Expert review of clinical pharmacology *6*, 23-34.
Heiden, M.G.V., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. Science *324*, 1029-1033.
Hensley, C.T., Faubert, B., Yuan, Q., Lev-Cohain, N., Jin, E., Kim, J., Jiang, L., Ko, B., Skelton, R., Loudat, L., et al. (2016a). Metabolic Heterogeneity in Human Lung Tumors. Cell *164*, 681-694.
Hensley, C.T., Faubert, B., Yuan, Q., Lev-Cohain, N., Jin, E., Kim, J., Jiang, L., Ko, B., Skelton, R., Loudat, L., et al. (2016b). Metabolic Heterogeneity in Human Lung Tumors. Cell *164*, 681-694.

Hensley, C.T., Wasti, A.T., and DeBerardinis, R.J. (2013). Glutamine and cancer: cell biology, physiology, and clinical opportunities. Journal of Clinical Investigation *123*, 3678-3684.
Hofree, M., Shen, J.P., Carter, H., Gross, A., and Ideker, T. (2013). Network-based stratification of tumor mutations. Nat Methods *10*, 1108-1115.
Hu, J., Locasale, J.W., Bielas, J.H., O'Sullivan, J., Sheahan, K., Cantley, L.C., Vander Heiden, M.G., and Vitkup, D. (2013). Heterogeneity of tumor-induced gene expression changes in the human metabolic network. Nature biotechnology *31*, 522-529.
Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., Fulton, R.S., et al. (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.
Israelsen, W.J., Dayton, T.L., Davidson, S.M., Fiske, B.P., Hosios, A.M., Bellinger, G., Li, J., Yu, Y.M., Sasaki, M., Horner, J.W., et al. (2013). PKM2 Isoform-Specific Deletion Reveals a Differential Requirement for Pyruvate Kinase in Tumor Cells. Cell *155*, 397-409.
Ito, K., and Suda, T. (2014). Metabolic requirements for the maintenance of self-renewing stem cells. Nat Rev Mol Cell Biol *15*, 243-256.
Jain, M., Nilsson, R., Sharma, S., Madhusudhan, N., Kitami, T., Souza, A.L., Kafri, R., Kirschner, M.W., Clish, C.B., and Mootha, V.K. (2012). Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. Science *336*, 1040-1044.
Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabasi, A.L. (2000). The large-scale organization of metabolic networks. Nature *407*, 651-654.
Jerby, L., and Ruppin, E. (2012). Predicting Drug Targets and Biomarkers of Cancer via Genome-Scale Metabolic Modeling. Clinical Cancer Research *18*, 5572-5584.
Jiang, H., Hegde, S., Knolhoff, B.L., Zhu, Y., Herndon, J.M., Meyer, M.A., Nywening, T.M., Hawkins, W.G., Shapiro, I.M., Weaver, D.T., et al. (2016). Targeting focal adhesion kinase renders pancreatic cancers responsive to checkpoint immunotherapy. Nat Med *22*, 851-+.
Kallmeyer, J., Pockalny, R., Adhikari, R.R., Smith, D.C., and D'Hondt, S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. Proceedings of the National Academy of Sciences of the United States of America *109*, 16213-16216.
Karsli-Uzunbas, G., Guo, J.Y., Price, S., Teng, X., Laddha, S.V., Khor, S., Kalaany, N.Y., Jacks, T., Chan, C.S., Rabinowitz, J.D., et al. (2014). Autophagy Is Required for Glucose Homeostasis and Lung Tumor Maintenance. Cancer discovery *4*, 914-927.
Kelloff, G., Hoffman, J.M., Johnson, B., Scher, H.I., Siegel, B.A., Cheng, E.Y., Cheson, B.D., O'Shaughnessy, J., Guyton, K.Z., Mankoff, D.A., et al. (2005). Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development. Clinical Cancer Research *11*, 2785-2808.
Kitano, H. (2002). Systems biology: A brief overview. Science *295*, 1662-1664.
Koppenol, W.H., Bounds, P.L., and Dang, C.V. (2011). Otto Warburg's contributions to current concepts of cancer metabolism (vol 11, pg 325, 2011). Nature Reviews Cancer *11*, 618-618.
Kruiswijk, F., Labuschagne, C.F., and Vousden, K.H. (2015). p53 in survival, death and metabolic health: a lifeguard with a licence to kill. Nat Rev Mol Cell Bio *16*, 393-405.
Kwapien, J., and Drozdz, S. (2012). Physical approach to complex systems. Phys Rep *515*, 115-226.
Lansing, J.S. (2003). Complex adaptive systems. Annu Rev Anthropol *32*, 183-204.
Lazar, M.A., and Birnbaum, M.J. (2012). Physiology. De-meaning of metabolism. Science *336*, 1651-1652.
Le Chatelier, E., Nielsen, T., Qin, J.J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.M., Kennedy, S., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541-+.
Lehuede, C., Dupuy, F., Rabinovitch, R., Jones, R.G., and Siegel, P.M. (2016). Metabolic Plasticity as a Determinant of Tumor Growth and Metastasis. Cancer Research *76*, 5201-5208.

Levy, R., and Borenstein, E. (2013). Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. Proceedings of the National Academy of Sciences of the United States of America *110*, 12804-12809.
Li, M., Wang, B.H., Zhang, M.H., Rantalainen, M., Wang, S.Y., Zhou, H.K., Zhang, Y., Shen, J., Pang, X.Y., Zhang, M.L., et al. (2008). Symbiotic gut microbes modulate human metabolic phenotypes. Proceedings of the National Academy of Sciences of the United States of America *105*, 2117-2122.
Lin, B.L., Sakoda, A., Shibasaki, R., Goto, N., and Suzuki, M. (2000). Modelling a global biogeochemical nitrogen cycle in terrestrial ecosystems. Ecol Model *135*, 89-110.
Long, Y., Tsai, W.B., Wangpaichitr, M., Tsukamoto, T., Savaraj, N., Feun, L.G., and Kuo, M.T. (2013). Arginine Deiminase Resistance in Melanoma Cells Is Associated with Metabolic Reprogramming, Glucose Dependence, and Glutamine Addiction. Mol Cancer Ther *12*, 2581-2590.
Lunt, S.Y., Muralidhar, V., Hosios, A.M., Israelsen, W.J., Gui, D.Y., Newhouse, L., Ogrodzinski, M., Hecht, V., Xu, K., Acevedo, P.N.M., et al. (2015). Pyruvate Kinase Isoform Expression Alters Nucleotide Synthesis to Impact Cell Proliferation. Molecular Cell *57*, 95-107.
Lunt, S.Y., and Vander Heiden, M.G. (2011). Aerobic Glycolysis: Meeting the Metabolic Requirements of Cell Proliferation. Annu Rev Cell Dev Bi *27*, 441-464.
Mardinoglu, A., Gatto, F., and Nielsen, J. (2013). Genome-scale modeling of human metabolism a systems biology approach. Biotechnology journal *8*, 985-+.
Mardinoglul, A., and Nielsen, J. (2015). New paradigms for metabolic modeling of human cells. Curr Opin Biotech *34*, 91-97.
Martinez-Outschoorn, U.E., Peiris-Pages, M., Pestell, R.G., Sotgia, F., and Lisanti, M.P. (2017). Cancer metabolism: a therapeutic perspective (vol 14, pg 11, 2017). Nat Rev Clin Oncol *14*, 113-113.
Mayers, J.R., Wu, C., Clish, C.B., Kraft, P., Torrence, M.E., Fiske, B.P., Yuan, C., Bao, Y., Townsend, M.K., Tworoger, S.S., et al. (2014). Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. Nat Med *20*, 1193-1198.
Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. Nature *501*, 328-337.
Metallo, C.M., Gameiro, P.A., Bell, E.L., Mattaini, K.R., Yang, J.J., Hiller, K., Jewell, C.M., Johnson, Z.R., Irvine, D.J., Guarente, L., et al. (2012). Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. Nature *481*, 380-U166.
Moghaddas Gholami, A., Hahne, H., Wu, Z., Auer, F.J., Meng, C., Wilhelm, M., and Kuster, B. (2013). Global proteome analysis of the NCI-60 cell line panel. Cell reports *4*, 609-620.
Monti, S., Savage, K.J., Kutok, J.L., Feuerhake, F., Kurtin, P., Mihm, M., Wu, B.Y., Pasqualucci, L., Neuberg, D., Aguiar, R.C.T., et al. (2005). Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. Blood *105*, 1851-1861.
Mullen, A.R., Wheaton, W.W., Jin, E.S., Chen, P.H., Sullivan, L.B., Cheng, T., Yang, Y.F., Linehan, W.M., Chandel, N.S., and DeBerardinis, R.J. (2012). Reductive carboxylation supports growth in tumour cells with defective mitochondria. Nature *481*, 385-U171.
Murphy, M.P. (2009). How mitochondria produce reactive oxygen species. Biochemical Journal *417*, 1-13.
Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-Gut Microbiota Metabolic Interactions. Science *336*, 1262-1267.
Nielsen, J. (2017). Systems Biology of Metabolism: A Driver for Developing Personalized and Precision Medicine. Cell Metabolism *25*, 572-579.
Nieman, K.M., Kenny, H.A., Penicka, C.V., Ladanyi, A., Buell-Gutbrod, R., Zillhardt, M.R., Romero, I.L., Carey, M.S., Mills, G.B., Hotamisligil, G.S., et al. (2011). Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. Nat Med *17*, 1498-U1207.
Niklas, J., Schneider, K., and Heinzle, E. (2010). Metabolic flux analysis in eukaryotes. Curr Opin Biotech *21*, 63-69.

Nilsson, R., Jain, M., Madhusudhan, N., Sheppard, N.G., Strittmatter, L., Kampf, C., Huang, J., Asplund, A., and Mootha, V.K. (2014). Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. Nature communications *5*.
Oberhardt, M.A., Palsson, B.O., and Papin, J.A. (2009). Applications of genome-scale metabolic reconstructions. Molecular systems biology *5*.
Obre, E., and Rossignol, R. (2015). Emerging concepts in bioenergetics and cancer research: Metabolic flexibility, coupling, symbiosis, switch, oxidative tumors, metabolic remodeling, signaling and bioenergetic therapy. Int J Biochem Cell B *59*, 167-181.
Owen, O.E., Kalhan, S.C., and Hanson, R.W. (2002). The key role of anaplerosis and cataplerosis for citric acid cycle function. Journal of Biological Chemistry *277*, 30409-30412.
Palm, W., Park, Y., Wright, K., Pavlova, N.N., Tuveson, D.A., and Thompson, C.B. (2015). The Utilization of Extracellular Proteins as Nutrients Is Suppressed by mTORC1. Cell *162*, 259-270.
Patil, M.D., Bhaumik, J., Babykutty, S., Banerjee, U.C., and Fukumura, D. (2016). Arginine dependence of tumor cells: targeting a chink in cancer's armor. Oncogene *35*, 4957-4972.
Patra, K.C., Wang, Q., Bhaskar, P.T., Miller, L., Wang, Z.B., Wheaton, W., Chandel, N., Laakso, M., Muller, W.J., Allen, E.L., et al. (2013). Hexokinase 2 Is Required for Tumor Initiation and Maintenance and Its Systemic Deletion Is Therapeutic in Mouse Models of Cancer (vol 24, pg 213, 2013). Cancer cell *24*, 399-399.
Pavlova, N.N., and Thompson, C.B. (2016). The Emerging Hallmarks of Cancer Metabolism. Cell Metabolism *23*, 27-47.
Pearce, E.L., Poffenberger, M.C., Chang, C.H., and Jones, R.G. (2013). Fueling immunity: insights into metabolism and lymphocyte function. Science *342*, 1242454.
Peterson, R.O. (2013). How Species Interact Altering the Standard View on Trophic Ecology. Science *339*, 142-143.
Phillips, M.M., Sheaff, M.T., and Szlosarek, P.W. (2013). Targeting Arginine-Dependent Cancers with Arginine-Degrading Enzymes: Opportunities and Challenges. Cancer Res Treat *45*, 251-262.
Piskounova, E., Agathocleous, M., Murphy, M.M., Hu, Z.P., Huddlestun, S.E., Zhao, Z.Y., Leitch, A.M., Johnson, T.M., DeBerardinis, R.J., and Morrison, S.J. (2015). Oxidative stress inhibits distant metastasis by human melanoma cells. Nature *527*, 186-+.
Possemato, R., Marks, K.M., Shaul, Y.D., Pacold, M.E., Kim, D., Birsoy, K., Sethumadhavan, S., Woo, H.K., Jang, H.G., Jha, A.K., et al. (2011). Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature *476*, 346-U119.
Quail, D.F., and Joyce, J.A. (2013). Microenvironmental regulation of tumor progression and metastasis. Nat Med *19*, 1423-1437.
Rabinovich, S., Adler, L., Yizhak, K., Sarver, A., Silberman, A., Agron, S., Stettner, N., Sun, Q., Brandis, A., Helbling, D., et al. (2015). Diversion of aspartate in ASS1-deficient tumours fosters de novo pyrimidine synthesis. Nature *527*, 379-+.
Robinson, J.L., and Nielsen, J. (2016). Integrative analysis of human omics data using biomolecular networks. Molecular bioSystems *12*, 2953-2964.
Sabharwal, S.S., and Schumacker, P.T. (2014). Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? Nature Reviews Cancer *14*, 709-721.
Schlattner, U., Tokarska-Schlattner, M., and Wallimann, T. (2006). Mitochondrial creatine kinase in human health and disease. Biochim Biophys Acta *1762*, 164-180.
Schug, Z.T., Peck, B., Jones, D.T., Zhang, Q.F., Grosskurth, S., Alam, I.S., Goodwin, L.M., Smethurst, E., Mason, S., Blyth, K., et al. (2015). Acetyl-CoA Synthetase 2 Promotes Acetate Utilization and Maintains Cancer Cell Growth under Metabolic Stress. Cancer cell *27*, 57-71.
Schulze, A., and Harris, A.L. (2013). How cancer metabolism is tuned for proliferation and vulnerable to disruption (vol 491, pg 364, 2012). Nature *494*, 130-130.

Segata, N., Boernigen, D., Tickle, T.L., Morgan, X.C., Garrett, W.S., and Huttenhower, C. (2013). Computational meta'omics for microbial community studies. Molecular systems biology *9*.
Segel, L.A. (1986). Theoretical Biology and Complexity - 3 Essays on the Natural-Philosophy of Complex-Systems - Rosen,R. Nature *319*, 457-457.
Semenza, G.L. (2012). Hypoxia-Inducible Factors in Physiology and Medicine. Cell *148*, 399-408.
Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E., de Wouters, T., Juste, C., Rizkalla, S., Chilloux, J., et al. (2015). Quantifying Diet-Induced Metabolic Changes of the Human Gut Microbiome. Cell Metab *22*, 320-331.
Simon, H.A. (1980). Cognitive Science - the Newest Science of the Artificial. Cognitive Sci *4*, 33-46.
Stine, Z.E., Walton, Z.E., Altman, B.J., Hsieh, A.L., and Dang, C.V. (2015). MYC, Metabolism, and Cancer. Cancer discovery *5*, 1024-1039.
Stolyar, S., Van Dien, S., Hillesland, K.L., Pinel, N., Lie, T.J., Leigh, J.A., and Stahl, D.A. (2007). Metabolic modeling of a mutualistic microbial community. Molecular systems biology *3*.
Vander Heiden, M.G. (2011). Targeting cancer metabolism: a therapeutic window opens. Nat Rev Drug Discov *10*, 671-684.
Wang, R.S., Maron, B.A., and Loscalzo, J. (2015). Systems medicine: evolution of systems biology from bench to bedside. Wires Syst Biol Med *7*, 141-161.
Ward, P.S., and Thompson, C.B. (2012). Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. Cancer cell *21*, 297-308.
Weinberg, S.E., Sena, L.A., and Chandel, N.S. (2015). Mitochondria in the Regulation of Innate and Adaptive Immunity. Immunity *42*, 406-417.
Werner, H.M.J., Mills, G.B., and Ram, P.T. (2014). Cancer Systems Biology: a peek into the future of patient care? Nat Rev Clin Oncol *11*, 167-176.
Xiang, Y., Stine, Z.E., Xia, J.S., Lu, Y.Q., O'Connor, R.S., Altman, B.J., Hsieh, A.L., Gouw, A.M., Thomas, A.G., Gao, P., et al. (2015). Targeted inhibition of tumor-specific glutaminase diminishes cell-autonomous tumorigenesis. Journal of Clinical Investigation *125*, 2293-2306.
Yoshida, G.J. (2015). Metabolic reprogramming: the emerging concept and associated therapeutic strategies. J Exp Clin Canc Res *34*.
Yuan, T.L., and Cantley, L.C. (2008). PI3K pathway alterations in cancer: variations on a theme. Oncogene *27*, 5497-5510.
Yun, J., Mullarky, E., Lu, C.Y., Bosch, K.N., Kavalier, A., Rivera, K., Roper, J., Chio, I.I.C., Giannopoulou, E.G., Rago, C., et al. (2015). Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. Science *350*, 1391-1396.
Yuneva, M.O., Fan, T.W.M., Allen, T.D., Higashi, R.M., Ferraris, D.V., Tsukamoto, T., Mates, J.M., Alonso, F.J., Wang, C.M., Seo, Y., et al. (2012). The Metabolic Profile of Tumors Depends on Both the Responsible Genetic Lesion and Tissue Type. Cell Metabolism *15*, 157-170.
Zhao, L.P. (2013). The gut microbiota and obesity: from correlation to causality. Nat Rev Microbiol *11*, 639-647.
Zomorrodi, A.R., and Maranas, C.D. (2012). OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. PLoS computational biology *8*.
Zu, X.L., and Guppy, M. (2004). Cancer metabolism: facts, fantasy, and fiction. Biochem Bioph Res Co *313*, 459-465.