

**SYSTEMS BIOLOGY OF MICROBIOTA METABOLITES AND ADIPOCYTE
TRANSCRIPTION FACTOR NETWORK**

A Dissertation

by

KYUNGOH CHOI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Arul Jayaraman
Committee Members,	Zhilei Chen
	Katy Kao
	Robert Alaniz
Head of Department,	M. Nazmul Karim

August 2013

Major Subject: Chemical Engineering

Copyright 2013 Kyungoh Choi

ABSTRACT

The overall goal of this research is to understand roles of gut microbiota metabolites and adipocyte transcription factor (TF) network in health and disease by developing systematic analysis methods. As microbiota can perform diverse biotransformation reactions, the spectrum of metabolites present in the gastrointestinal (GI) tract is extremely complex but only a handful of bioactive microbiota metabolites have been identified. We developed a metabolomics workflow that integrates *in silico* discovery with targeted mass spectrometry. A computational pathway analysis where microbiota metabolisms are modeled as a single metabolic network is utilized to predict a focused set of targets for multiple reaction monitoring (MRM) analysis. We validated our methodology by predicting, quantifying in murine cecum and feces and characterizing tryptophan (TRP)-derived metabolites as ligands for the aryl hydrocarbon receptor.

The adipocyte process of lipid droplet accumulation and differentiation is regulated by multiple TFs that function together in a network. Although individual TF activation is previously reported, construction of an integrated network has been limited due to different measurement conditions. We developed an integrated network model of key TFs - PPAR γ , C/EBP β , CREB, NFAT, FoxO1, and SREBP-1c - underlying adipocyte differentiation. A hypothetic model was determined based on literature, and stochastic simulation algorithm (SSA) was applied to simulate TF dynamics. TF activation profiles at different stages of differentiation were measured using 3T3-L1

reporter cell lines where binding of a TF to its DNA binding element drives expression of the *Gaussia* luciferase gene. Reaction trajectories calculated by SSA showed good agreement with experimental measurement. The TF model was further validated by perturbing dynamics of CREB using forskolin, and comparing the predicted response with experimental data.

We studied the molecular recognition mechanism underlying anti-inflammatory function of a bacterial metabolite, indole in DC2.4 cells. The indole treatment attenuated the fraction of cells that were producing the pro-inflammatory cytokine, TNF α and knockdown of nuclear receptor related 1 (Nurr1; NR4A2) resulted in less indole-derived suppression of TNF α production. The first discovery of NR4A2 as a molecular mediator of the endogenous metabolite, indole is expected to provide a new strategy for treatment of inflammatory disorders.

DEDICATION

To my husband, Sanghyun

my daughter, Susie

my brother, Minseong

and my parents, Hansu and Younghee

ACKNOWLEDGEMENTS

I would like to sincerely thank my committee chair, Dr. Arul Jayaraman, for his guidance and support throughout the course of my Ph.D. I have learned how to develop and verify my ideas for new research. He has encouraged my work and given many advices for the future career. It has been a great pleasure to discuss science with him.

I also thank my committee members, Dr. Zhilei Chen, Dr. Katy Kao and Dr. Robert Alaniz for their kind encouragement which helps me accomplish my research. Dr. Zhilei Chen and Dr. Katy Kao were friendly mentors I would follow the example of young scientists. Especially, Dr. Robert Alaniz provided essential help for flow cytometry analysis and *in vivo* mice experiments.

I want to extend my appreciation to Dr. Kyongbum Lee for the simulation studies of the adipocyte transcription factor network and prediction of microbiota metabolites and to Gautham Sridharan for mass spectrometry analysis. It would be difficult to complete my research without their contribution.

Special thanks go to Dr. Jayaraman lab members, Dr. Jeongyun Kim, Daenyun Kim, Manjunath Hegde, Billy Newton, Colby Moya, Shreya Maiti, Nandita Kohli, Sasi Kiran Pasupuleti, Dmitriy Verkhovturov, Yufang Ding, Cory Klemashevich, Sneha Jani, Michelle Olson, Daniel Howsmon, Darshan Prabhakaran, Eric Yang, Selena Hernandez, David Chaar and Pallavi Nair. I also acknowledge Dr. Alaniz lab and Dr. Kyongbum Lee members, Jane Miller, Madhu Katepalli, James K. Sims and Sara Manteiga.

Thanks also go to my friends, the department faculty and staff for making my time at Texas A&M University harmonious experiences. Finally, deeply thanks to my brother and parents for their encouragement and to the most precious husband and daughter for their love and support.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vii
LIST OF FIGURES	ix
LIST OF TABLES	xi
CHAPTER I INTRODUCTION	1
1.1 Background	1
1.2 Motivation	3
1.3 Research objectives, importance, and novelty	4
CHAPTER II LITERATURE REVIEW	8
2.1 Gut microbiota.....	8
2.2 Metabolomic analysis of GI tract.....	13
2.3 AhR signaling in GI tract	15
2.4 Transcriptional control of adipocyte differentiation	18
2.5 Inflammatory signaling in adipose tissue.....	21
2.6 NR4A regulation in inflammation.....	24
CHAPTER III PREDICTION, IDENTIFICATION, AND QUANTIFICATION OF MICROBIOTA METABOLITES IN MURINE GUT USING IN SILICO TARGETED METABOLOMICS	28
3.1 Introduction	28
3.2 Materials and methods	32
3.3 Results	42
3.4 Discussion	54
CHAPTER IV ANALYSIS OF TRANSCRIPTION FACTOR NETWORK UNDERLYING 3T3-L1 ADIPOCYTE DIFFERENTIATION	63

4.1 Introduction	63
4.2 Materials and methods	65
4.3 Results	76
4.4 Discussion	85
CHAPTER V MECHANISMS UNDERLYING RECOGNITION OF TRYPTOPHAN DERIVATIVES IN HOST CELLS	93
5.1 Introduction	93
5.2 Materials and methods	96
5.3 Results	99
5.4 Discussion	103
CHAPTER VI CONCLUSIONS AND RECOMMENDATIONS	107
6.1 Conclusions	107
6.2 Recommendations	109
REFERENCES	113

LIST OF FIGURES

FIGURE	Page
3.1 Probabilistic pathway construction with a tryptophan source.	43
3.2 Representative tryptophan biotransformation routes and intermediates predicted by the pathway construction algorithm.	45
3.3 LC-MS/MS chromatogram for indole	48
3.4 LC-MS/MS chromatogram for all target metabolites	49
3.5 Comparison of metabolite concentrations between young (6 weeks of age) and old (15 weeks of age) mice	50
3.6 AhR activation by TRP-derived metabolites.....	52
3.7 Microbiota metabolites screen for AhR activation in the H4IIE reporter cells...	53
3.8 Dose dependent activation of AhR upon exposure to indole-3-acetate and tryptamine	55
3.9 Dose dependent activation of AhR upon exposure to indole 3-pyruvate	56
3.10 Dose-dependent AhR activation in the H4IIE reporter cells.....	57
4.1 Schematic of TF network model	71
4.2 The six TF reporter constructs properly present activation of each target TF	77
4.3 3T3-L1 PPAR γ reporter cells after 19 days of induction for differentiation.....	78
4.4 3T3-L1 TF reporter cells after 19 days of induction for differentiation.....	80
4.5 TF activity profiles represented as fold of increase rates of <i>Gaussia</i> luciferase (Gluc) activities normalized with relative fluorescence units (RFU)	81
4.6 Simulated profiles of TF dynamics	84
4.7 Perturbation of CREB activity profiles by forskolin treatments	86
4.8 Simulations of TF activity profiles representing the forskolin treatment.....	87

5.1	Expression of NR4A receptors in DC2.4 cells.	101
5.2	The NR4A2 knockdown attenuates indole-derived suppression of TNF α production in DC2.4 cells	102

LIST OF TABLES

TABLE	Page
2.1 Interactions between the transcription factors to regulate adipogenesis	22
3.1 Optimized MRM-MS parameters for detection of TRP metabolites	36
3.2 Solvent gradient method for metabolite separation	38
4.1 Mass action kinetic model of TF interactions	73

CHAPTER I

INTRODUCTION

1.1 BACKGROUND

The human gastrointestinal tract is colonized by $\sim 10^{14}$ bacteria belonging to $\sim 1,000$ species that are collectively termed the gut microbiota (1). The gut microbiota have been classically known to be important in two main capacities - the digestion of dietary molecules and the generation of energy from fermentation of indigestible carbohydrates to short-chain fatty acids (2). Recent studies have shown that the intestinal microbiota is a “super organ” that mediates a wide range of functions including innate and adaptive immunity, inflammation, cell proliferation, and communication with the gut epithelium and distant organs (3). Alterations in the microbiota composition and/or function (dysbiosis) are increasingly correlated to several disorders, including several inflammatory diseases of the gut such as inflammatory bowel disease (IBD) and colitis (4, 5), as well as obesity, insulin resistance and type 2 diabetes (T2D) (6, 7). The effects of dysbiosis are not localized to the GI tract alone but are seen systemically as well as studies have associated several allergies (8, 9) and autism (10) to intestinal microbiota dysbiosis. Increasing evidence shows that the functional outputs of the microbiota (i.e., metabolites) are important modulators of function in the host. For example, conversion of choline into methylamines and trimethylamine by the microbiota causes non-alcoholic fatty liver disease (NAFLD) (11) and cardiovascular diseases (12).

Obesity is one of the most common diseases that have been strongly associated with microbiota dysbiosis. Obesity correlates with complications of several metabolic disorders (e.g., insulin resistance and T2D) and increases risk for many other diseases such as NAFLD, atherosclerosis, and certain cancers. Development of the obesity and metabolic disorders is a complex process and associated with pathways which connect metabolism and immune systems (13). Obesity is characterized by an increase in body fat mass, specifically in that of white adipose tissue (WAT) (14). Expansion of the body fat mass can occur either due to an increase in the number of adipocytes and/or an increase in size of adipocytes. The increase of adipocyte cell size is an important component of the increase in adipose tissue mass (15). Adipocyte enlargement arises as a result of genetic or environmental conditions that progressively increase lipid loading. The enzymes mediating adipocyte lipid accumulation and utilization are regulated through the temporally coordinated action of several transcription factors (16, 17). Therefore, it is important to understand the dynamics and interaction between different members of the regulatory network for modulating complex phenotypes such as lipid balance in adipocytes. Phylogenetic (16S rRNA) sequencing and metabolomics analysis have shown that the composition and metabolite profile of the intestinal microbiota are significantly altered in obese subjects (18, 19). The microbiota in obese subjects or animals are able to extract more energy from diet (20), and interact with epithelial host cells to control energy storage and expenditure (21).

1.2 MOTIVATION

In the United States, IBD is one of the five most prevalent gastrointestinal disease burdens and an overall care cost is more than \$1.7 billion. Each year, there are more than 700,000 physician visits, 100,000 hospitalizations, and disability in 119,000 patients due to IBD (<http://www.cdc.gov/ibd/>). Several studies have established a strong correlation between intestinal inflammatory disease such as Crohn's disease and dysbiosis of the intestinal microbiota (22). For example, a clinical metabolomics study suggested that poor utilization of tryptophan and an imbalance of tryptophan-derived microbiota metabolites correlated strongly with ileal crohn's disease (ICD) (23). Although there is a high level of interest in identifying microbiota metabolites as well as characterizing molecular mechanisms for their immunomodulatory functions, there has been little success due to the large number of metabolites present and the complex network of biosynthetic reactions carried out by the microbiota (24).

The prevalence of obesity and related metabolic disorders has vastly increased throughout the world. Medical care costs of obesity in the United States are about \$147 billion in 2008 (25). Approximately, 21 million people suffer from type 2 diabetes and an annual total cost for treatment is over \$80 billion in the United States (26). Several therapies such as gastric bypass and suppression of appetite have been developed for control of obesity but have had little success. Therefore, there is a need to develop new approaches against obesity. Since adipocyte differentiation and enlargement is an orchestrated process by the sequential induction of key transcription factors,

understanding an entire network between the transcription factors will provide a framework for manipulating obesity.

Recently, it is demonstrated that the bacterial metabolite indole attenuates indicators of inflammation and improves tight junction properties in intestinal epithelial cells (27). This study demonstrated for the first time that microbiota metabolites are not simply metabolic by-products but are active molecular determinants of cellular behavior and responses *in vivo*. Moreover, development of an integrated transcription factor network underlying adipocyte differentiation and lipid accumulation has offered a new insight into understanding of obesity related disorders. Accordingly, a study for microbiota metabolites-triggered change in the transcription factor network is expected to provide a clue for treatment of obesity and metabolic disorders.

1.3 RESEARCH OBJECTIVES, IMPORTANCE, AND NOVELTY

The spectrum of metabolites present in the GI tract is extremely complex, as the microbiota can carry out a diverse range of biosynthetic reactions (24), including those that are not present in the mammalian host (28). Classical approaches such as isolating and culturing individual bacteria and identifying metabolites produced by them has not yielded much success, as many bacterial species in the GI tract cannot be cultured under standard laboratory conditions. Although untargeted metabolomics for profiling and identifying metabolites produced in the GI tract have used in recent several studies (29-31), the untargeted approach suffers from two limitations. First, the simultaneous identification and quantification of a large number of metabolites remains challenging,

because of the variations in the dynamic ranges of different metabolites (32). Second, it is extremely difficult to distinguish between metabolites produced by the microbiota from those produced by host metabolism, as the microbiota and host share significant similarity in the biochemical reactions that they can carry out. Therefore, the systematic prediction, identification and quantification of microbiota metabolites are a significant and novel contribution of this work.

To our knowledge, this is the first study to present a novel methodology which integrates bioinformatic prediction of microbiota metabolites with targeted metabolomics. The methodology predicts a panel of target metabolites that can be uniquely ascribed to the microbiota biochemical reactions. Targeted metabolomics improves dynamic range of detection and provides absolute quantitation of metabolites. In future, extension of the study with different *in vivo* models that alter gut homeostasis and ecology will enable us to determine the effect of perturbations on the function of the microbiota and its effect on host function. The identification of a mechanism for the immunomodulatory role of indole on immune cells (i.e., inter-domain signal recognition) is also a novel aspect of this study.

During the cellular transformation from preadipocytes into mature adipocytes, cells engage a network of multiple transcription factors. While many studies have identified roles for different transcription factors, data from these studies provide limited information on both expression dynamics and interaction between regulatory network molecules. In fact, simple inhibition of an individual regulatory molecule often has adverse effects; for example, the inhibition of peroxisome proliferator activated receptor-

γ (PPAR γ) *in vivo* reduces adipogenesis, but also increases insulin resistance, one of the chief complications of type 2 diabetes mellitus (33). In order to understand entire regulatory events underlying differentiation and enlargement of adipocytes, integration for dynamic interactions between transcription factors is required.

The integrated model of the transcription factor network underlying 3T3-L1 adipocyte differentiation is developed by combinatorial approach of experiment and simulation in this study. Furthermore, the model can be applied for prediction of overall change of transcription factor activity profiles triggered by a specific microbiota metabolite and its influence on development of obesity and metabolic disorders.

The specific objectives were to:

- Predict metabolites that can be derived from the dietary amino acid tryptophan (TRP) by the microbiota, quantify the levels of the predicted TRP derivatives in murine cecum and feces using multiple reaction monitoring (MRM) mass spectrometry and demonstrate the biological activity of the metabolites at their physiological concentrations measuring aryl hydrocarbon receptor (AhR) activation
- Elucidate the molecular mechanism underlying recognition of TRP derivatives in eukaryotic cells
- Develop an integrated network model of six adipocyte transcription factors (PPAR γ , SREBP-1c, NFAT, CREB, C/EBP β and FoxO1) based on published connections between the different molecules, validate the model by measuring of activation dynamics during adipocyte differentiation and by perturbing the

activation levels of CREB to determine its effects on the other transcription factors

CHAPTER II

LITERATURE REVIEW

2.1 GUT MICROBIOTA

The human gastrointestinal tract (GI) contains a variety of microorganisms which are collectively termed gut microbiota. This community of $\sim 10^{14}$ bacteria is composed of at least 1000 different species and dominated by anaerobic bacteria (34). Firmicutes, Actinobacteria and Bacteroidetes are the main bacterial phyla in human gut microbiota. The largest bacterial phylum, Firmicutes contains more than 200 genera, including *Bacillus*, *Lactobacillus*, *Clostridium* and *Mycoplasma* (35). Genomes of gut microbiota contain more than 100 times the number of genes as compared to our genome and are being increasingly described as a super organ playing important roles in immune system, energy balance, metabolism and host defense (36).

2.1.1 GUT MICROBIOTA AND DISEASES

Alterations of gut microbiota have been associated with several disorders and metabolic diseases that occur not only in the GI tract but in distant sites. Inflammatory Bowel Disease (IBD) is one of the most commonly reported diseases that is associated with dramatic changes in gut microbiota. Studies have shown that the number of mucosal adherent bacteria belonging to species such as *Proteobacteria* and *Enterobacteriaceae* is increased in patients with IBD (37). The population of *Bifidobacterium* was also found to be significantly reduced in rectal biopsies from IBD

patients (38). In addition, a decreased distribution of *Firmicutes* and butyrate producers such as *Clostridium* was noted in IBD patients (39). Roediger and co-workers demonstrated that overgrowth of sulphate-reducing bacteria (SRB) is associated with acute ulcerative colitis (40). A metabolite of SRB, hydrogen sulfide was shown to have genotoxic properties and produce mucosal inflammation (41).

Obesity is the most prevalent metabolic disease worldwide and is also associated with reduced bacterial diversity and alterations in metabolic pathways and bacterial genes in the GI tract (42). Ley and co-workers found obese mice have 50% more *Firmicutes* and less *Bacteroidetes* as compared to lean mice (18). This was associated with an increase of cecal short chain fatty acids (SCFA), implying increased fermentation by gut microbiota and energy harvest from diet (20). Colonization of GI tract in germ-free mice with microbiota from conventionally raised mice showed an increase of insulin resistance and body fat. In addition, the colonization study found that the increase of body fat correlated with downregulation of fasting induced adipocyte factor (FIAF) in gut epithelium. Since the FIAF is a secreted lipoprotein lipase (LPL) inhibitor, downregulation of FIAF leads to higher LPL activity and triglyceride (TG) storage in adipose tissue (21). In addition, gut microbiota modulate plasma LPS level that triggers chronic low grade inflammation leading to diabetes (7). Non-alcoholic fatty liver disease (NAFLD) is another disease related to obesity. The gut microbiota contribute to development of NAFLD through inflammasomes that shape metabolic pathways of lipid accumulation (43). Inflammasome deficiency caused changes in configuration of gut microbiota and exacerbated hepatic steatosis and inflammation

through influx of Toll-like receptor 4 (TLR4) and TLR9 agonists, leading to enhanced tumour-necrosis factor (TNF)- α expression.

In addition to its effect on metabolic diseases, alterations in the gut microbiota have a significant impact on immune responses in the host. For example, a systemic autoimmune disease, Rheumatoid Arthritis (RA) is associated with changes in gut microbiota composition and function. In serum samples of RA patients, antibodies to *P. gingivalis* were reported to be more frequent than controls and C-reactive protein concentration and a titer of RA-related autoantibodies were higher in individuals infected with *P. gingivalis* (44). Moreover, RA is closely related to periodontal disease. In subjects of RA and periodontitis, serum antibodies against disease producing periodontal bacteria were more frequently detected (45).

Immune system development is also associated with commensal gut microbiota. Following birth, breastfeeding confers beneficial effects in configuration of gut microbiota, including increased colonization of *Bifidobacteria* and reduced prevalence of an asthma-associated pathogen, *C. difficile*. Antibiotics ingested by infants suppress colonization of commensal bacteria and cause emergence of *C. difficile* (46). Penders and co-workers found that infants colonized with *C. difficile* and *E. coli* had higher risk of developing recurrent wheeze and eczema (47). The presence of *C. difficile* and *E. coli* could be associated with a decrease of beneficial bacteria, and results in less induction of Treg cells by the beneficial bacteria leading to immune dysregulation.

2.1.2 GUT MICROBIOTA-DERIVED METABOLITES

Conversion of proteins, carbohydrates and nonnutritive compounds by gut microbiota leads to formation of a variety of metabolites that have adverse or beneficial effects on human health. The breakdown of complex carbohydrates in the colon results in oligo- and monomeric compounds which are further broken down to short chain fatty acids (SCFA) such as butyrate and propionate (48). The formation of SCFA in the colon significantly impacts host responses and function. The SCFA is a significant source of energy, contributing 10% of daily calorie requirement for health (49). Butyrate is a main energy source of colonic epithelial cells and colonocytes metabolize 90% of butyrate (50). It functions as a main regulator of growth by promoting proliferation of epithelial cells and enhancing differentiation. In intracolonic milieu, the SCFA lowers pH, which inhibits growth of pH sensitive pathogenic bacteria and prevents degradation of primary bile acids to carcinogenic secondary bile acids (51). It has been proposed butyrate lowers risk of colon cancer by inhibiting genotoxic activity of hydrogen peroxide and nitrosamides in colon cells (52). It has also been reported that butyrate plays a chief role in chronic inflammation of intestinal mucosa. In a clinical study of Crohn's disease, butyrate was shown to decrease expression of proinflammatory cytokines such as interleukin-1 β (IL-1 β), IL-6 and tumor necrosis factor- α (TNF α) by inhibiting activation of nuclear factor κ B (NF- κ B) (53).

While bacterial fermentation of carbohydrates primarily occurs in a proximal colon, fermentation of proteins mainly occurs in a distal colon (54). Some of products resulting from the amino acid fermentation are associated with diseases. For example,

H₂S (present as HS⁻ at neutral pH) which is formed through bacterial degradation of cysteine and methionine was shown to cause DNA hypomethylation, leading to abnormalities of cell proliferation in crypt epithelium in ulcerative colitis (55). In addition, sulfate-reducing bacteria (SRB) generate sulfide from dietary sulfate by coupling with oxidation of lactate, succinate, ethanol and H₂ (56). The anaerobic fermentation of aromatic amino acids by gut microbiota gives rise to several indole and phenol-containing compounds. Phenolic metabolites such as *p*-cresol were proposed to act as a procarcinogen in the etiology of colon cancer (57). In a rat model of depression, plasma concentrations of indoxylsulfate and indole-3-acetate were decreased (58). These altered compounds are metabolic products of amino acids by gut microbiota (59). The indoxylsulfate is generated by gut microbiota containing tryptophanase metabolism and indole-3-acetate derives from transamination on tryptophan by gut microbiota.

Bile acids facilitate lipid absorption and participate in regulation of cholesterol, triglyceride and energy homeostasis, and the gut microbiota have been reported to modulate bile acid profiles (i.e., the spectrum of molecules produced) through dehydroxylation, oxidation, sulfation and deconjugation, resulting in generation of secondary or tertiary bile acids (60). For example, deconjugation of glycocholate by bacteria generates cholate, which is transformed to deoxycholate by the bacterial enzyme 7- α -dehydroxylase. Suhre and co-workers detected that deoxycholate levels are higher in plasma from diabetic patients (61). Deoxycholate is the secondary bile acid and constitutes a major part (35%) of circulating bile acid pool in human. The secondary bile acids can passively enter the bile acid pool or be excreted to feces. The loss of bile acids

in feces is compensated by *de novo* biosynthesis of bile acids in the liver. Therefore, the higher level of deoxycholate in diabetic patients indicates alterations in the composition of bile acid pool and related biosynthetic pathways including conversion of primary to the secondary bile acids by gut microbiota. This observation is supported by the fact that composition of gut microbiota is between control and diabetic individuals (62).

2.2 METABOLOMIC ANALYSIS OF GI TRACT

To date, a majority of the observations on the alterations in gut microbiota in different disease models have been derived from high throughput sequencing studies. The metagenome data derived from these sequencing studies provide information only on alterations in the microbiota composition (i.e., population changes and/or enriched gene content) but do not provide information on functional changes arising from these alterations (63). Recent studies have used mass spectrometry (MS) methods to profile alterations in the metabolite levels (i.e., the products of metabolic reactions) from fecal material and biofluids to determine functional changes arising due to alterations in the metagenome. Analytical techniques such as high resolution nuclear magnetic resonance (NMR) spectroscopy, FTICR (Fourier transform ion cyclotron resonance)-MS, and chromatography have been coupled with MS enable comprehensive profiling of metabolites that are correlated to microbiota in the GI tract.

Zheng and co-workers (29) applied an untargeted LC-MS (ultra performance liquid chromatography coupled to a tandem quadrupole-time-of-flight mass spectrometry) and GC-MS (gas chromatography coupled with a time-of-flight mass

spectrometry) to profile metabolite changes in a rodent model of antibiotic exposure. In this study, panels of 223 fecal and 202 urinary metabolites were significantly altered, showing extensive gut microbiota modulation of host metabolism involving tryptophan, tyrosine and short-chain fatty acid metabolism. In another untargeted MS study, metabolites in plasma of germ-free (GF) mice were compared to samples of conventional mice (64) to demonstrate that a large number of compounds in circulation can be attributed to the presence of microbiota in the GI tract. Specifically, phenyl group-containing organic acids considerably increased in presence of gut microbiota. In addition, the levels of indole-containing metabolites derived from tryptophan such as indoxylsulfate and indole-3-propionic acid (IPA) were decreased in GF mice as compared to conventional mice. When GF mice were injected with IPA, the high serum level of IPA detected after 1 h decreased more than 90% within 5 h, indicating that the presence of IPA in serum of conventional mice results from continuous production from gut microbiota. Indoxylsulfate, a known nephrotoxin that accumulates in blood of patients suffering from chronic kidney failure (65), is generated by transformation of a bacterial metabolite indole in liver. This finding illustrates bacterial metabolites can reach other organs and be modified by mammalian chemical processes, resulting broad effects on the entire body.

Antunes and co-workers (66) used Fourier transform ion cyclotron resonance MS with direct infusion (DI-FT-ICR-MS) to obtain the chemical composition of the GI tract upon antibiotic treatment. They observed that antibiotic treatment disrupted homeostasis of hormone such as steroids which have been implicated in immunological responses to

infection and affected levels of 87% of detected 2230 metabolite features. Predictive mapping of altered metabolites to corresponding metabolic pathways identified various host metabolic functions such as steroid, bile acid, fatty acid and eicosanoid metabolism as being altered. Marchesi and co-workers (67) investigated fecal extracts of Crohn's disease (CD) and ulcerative colitis (UC) patients by employing high resolution ^1H NMR spectroscopy (68). In both fecal extracts of CD and UC patients, levels of methylamine, trimethylamine, acetate and butyrate were reduced, implying a change in gut microbial community. In a CD patient study of 15 twin pairs, individuals with ileal CD had increased abundance of *E. coli* and a lower abundance of *F. prausnitzii* as compared to healthy co-twins (69). This study also showed that increased fecal levels (i.e., poor utilization) of tryptophan and phenylalanine correlated with ileal CD.

2.3 AHR SIGNALING IN GI TRACT

The aryl hydrocarbon receptor (AhR) is a ligand-activated transcription factor that has been shown to play a significant role in the intestinal immune system (70, 71). AhR is a member of the basic helix-loop-helix / Per-Arnt-Sim homology superfamily which plays a central role in sensing environmental factors such as light and oxygen. The AhR is localized in the cytoplasm and it translocates to the nucleus upon binding ligands such as polycyclic aromatic hydrocarbons. In the nucleus, ligand bound AhR forms heterodimers with the AhR nuclear translocator (Arnt) protein. The AhR-Arnt complex activates expression of genes containing the dioxin response element (DRE; GCGTG). Well known AhR target genes that are regulated by the AhR include

xenobiotic metabolizing enzymes such as Cyp1A1 of the cytochrome P450 family member (72).

While the AhR was identified as a part of xenobiotic metabolism, it has demonstrated function in immune response as well. AhR is expressed in cells involved in the innate immune response such as innate lymphoid cells (ILC) and in CD4⁺ T cells which are crucial for the adaptive immunity (70). Li and co-workers investigated AhR expression in a broad range of immune cells isolated from various tissues. Intraepithelial lymphocytes (IEL) (30) from skin and intestine demonstrated a high level of AhR expression. AhR deficient mice were short of IELs, which resulted in reduction of anti-microbial peptides, an altered microbiota composition, a low level of epithelial turnover, and increased susceptibility for intestinal inflammation (71). Similarly, Kiss and co-workers showed AhR deficient mice had decreased immune response and were highly susceptible to infection with *Citrobacter rodentium* (70). On the other hand, activation of AhR signaling inhibited colitis and inflammation in GI tract of mice (68). In addition to control of the immune response, it has been shown that AhR functions as a tumor suppressor. Fan and co-workers examined a role of AhR in the liver tumorigenesis induced by a genotoxic chemical diethylnitrosamine (DEN) (73). DEN-exposed *Ahr*^{-/-} mice exhibited elevation of tumor incidence in livers as compared to that in wild type littermates. They also found that expression levels of proinflammatory cytokines such as IL-6 and TNF α were higher in *Ahr*^{-/-} mice than *Ahr*^{+/+} counterpart. Hall and co-workers reported that activation of AhR inhibits a metastatic process in breast cancer cell lines

(74). Treatment of exogenous AhR agonists inhibited motility and cell invasiveness in a Boyden chamber assay.

Ligands for AhR include various toxic halogenated aromatic hydrocarbons, heterocyclic amines, polycyclic aromatic hydrocarbons and a variety of indole containing compounds (75-77). Although many dietary compounds were initially suggested as natural ligands for AhR, several reports have indicated that endogenous physiological ligands for AhR exist. For example, in absence of added exogenous ligands, expression of AhR dependent genes was induced *in vitro* (78, 79) and AhR knockout mice showed developmental defects (80). To date, many studies have demonstrated that a plant specific tryptophan derivative, indole-3-carbinol and its acidic condensation product such as indolo[3,2-*b*]carbazole are AhR ligands (72, 81, 82). In addition, endogenous metabolites of tryptophan such as tryptamine and indole acetic acid were shown to bind to and activate AhR (83). Several photooxidation products of tryptophan were identified to bind to AhR with a high affinity and induce AhR dependent gene expression (84, 85). Wincent and co-workers demonstrated that 6-formylindolo[3,2-*b*]carbazole (FICZ) is formed by exposure of tryptophan to visible light and FICZ acts a substrate for CYP1A1, CYP1A2 and CYP1B1 (86). In addition, they found sulfo conjugates of phenolic metabolites of FICZ are present in human urine. These findings indicate that FICZ is naturally occurring activators of AhR signaling pathway.

2.4 TRANSCRIPTIONAL CONTROL OF ADIPOCYTE DIFFERENTIATION

Adiposity can arise through either an increase in the number of adipocytes (hyperplasia) and/or an increase in the size of adipocytes (hypertrophy). Hyperplasia involves recruitment, proliferation and differentiation of preadipocytes while hypertrophy occurs due to an increase of lipid content in adipocytes (15).

The morphological changes and accumulation of lipid during adipogenesis are caused by shifts in the set of genes that are expressed at different stages of differentiation (17). Adipogenesis is known to follow a well characterized sequence of events that are regulated by the activity of different transcription factors. Initial growth arrest of preadipocytes is induced upon hormonal stimulation and then one or two additional rounds of cell division (clonal expansion) follows. During this period, there is a temporal increase in the expression of CCAAT/enhancer binding protein β (C/EBP β) and C/EBP δ (87). It was demonstrated ectopic expression of C/EBP β and C/EBP δ in 3T3-L1 preadipocytes induces adipogenesis without extracellular hormones (87, 88). Subsequently, expression of peroxisome proliferator activated receptor- γ (PPAR γ) and C/EBP α is stimulated (89). Several studies showed a direct link between PPAR γ and C/EBPs. Ectopic expression of C/EBP β and/or C/EBP δ in nonadipogenic NIH 3T3 fibroblasts induced PPAR γ , and exposure to PPAR γ ligands prompted differentiation into adipocytes (90). Moreover, functional regulatory elements of C/EBP were identified in the PPAR γ promoter (91). Recently, it was shown that the ectopic expression of C/EBP β in fibroblasts was incapable of inducing C/EBP α without PPAR γ ligands. In addition, retroviral expression of C/EBP β in ppar $\gamma^{-/-}$ mouse embryonic fibroblasts did not

stimulate expression of C/EBP α (92). On the other hand, C/EBP α -deficient embryonic fibroblasts did not express PPAR γ and restoration of C/EBP α restored PPAR γ expression levels (93). Therefore, the principal pathway of adipogenesis appears to involve 3 steps: i) expression of PPAR γ induced by C/EBP β and C/EBP δ , ii) activation of C/EBP α by PPAR γ , and iii) positive feedback loop between PPAR γ and C/EBP α .

It is well established that PPAR γ and C/EBP α function as master regulators for differentiation of adipocytes (94). Coexpression of PPAR γ and C/EBP α could transdifferentiate myoblasts into adipocytes (95). Loss of function studies demonstrated a dominant role of PPAR γ for fat cell formation and differentiation. PPAR γ deficiency in mice caused embryonic lethality due to a placental defect. A mutant mouse that was rescued by using wild type tetraploid cells showed severe lipodystrophy (96). In another study, embryonic stem cells with homozygous deletion of PPAR γ gene could not differentiate into adipocytes and did not participate in adipose tissue formation (97). Similar with PPAR γ studies, C/EBP α knockout mice died soon after birth due to inability of glucose production, as C/EBP α is required for gluconeogenesis in liver (98). Ablation of the *c/ebpa* gene in tissues except liver showed that C/EBP α is required for WAT formation. In addition, deficiency of C/EBP α expression resulted in insulin resistance in cell culture models (93).

Sterol regulatory element binding protein-1c (SREBP-1c) was identified as an E-box binding transcription factor involved in adipogenesis (99). SREBP-1c promotes adipogenesis and activates expression of lipoprotein lipase and fatty acid synthase although it cannot initiate adipogenesis by itself (100). In adipose tissue, expression of

SREBP-1c increases in response to feeding and insulin stimulates expression of SREBP-1c in 3T3-L1 cells (101). It was shown that SREBP-1c could activate PPAR γ expression by binding to the PPAR γ promoter (102). Dominant negative SREBP-1c expression inhibited adipogenesis but the inhibition could be overcome by treatment with a PPAR γ ligand, thiazolidinedione (TZD). This result suggests a possible role of SREBP-1c in the production of endogenous PPAR γ ligands (103).

Cyclic AMP (43) response element binding protein (CREB) is another primary regulator involved in initiating adipocyte differentiation. Stimulation of CREB activity is induced by differentiation inducing reagents such as insulin and dexamethasone. A study by Zhang and co-workers provides convincing evidence that CREB is activated at a very early time point in adipogenesis and participates in induction of C/EBP β expression (104). This result is consistent with an earlier study which mentioned a role of cAMP signaling for C/EBP β expression (87). CREB exerts its function by binding to promoters of C/EBP β and/or C/EBP δ and activate several adipocyte specific genes such as phosphoenolpyruvate carboxykinase (PEPCK) and adipocyte specific fatty acid binding protein (aP2). In 3T3-L1 cells, expression of dominant negative CREB blocks adipogenesis while expression of constitutively active CREB induces adipocyte differentiation (16).

Nuclear factor of activated T cell (NFAT) is a member of the transcription factor family that regulates early immune response and cytokine production. It was shown that blocking the nuclear translocation of NFAT inhibits adipogenesis and NFAT is able to bind to aP2 promoter (105). NFAT has also been proposed as a regulator of PPAR γ

activity since NFAT forms a composite enhancer complex with C/EBP and associates physically with PPAR γ (106).

Forkhead transcription factor FoxO1 is induced in an early stage of adipocyte differentiation but activation of FoxO1 is delayed until the end of clonal expansion phase. FoxO1 has been shown to attenuate PPAR γ activation in 3T3-L1 adipocytes (107). Constitutively active FoxO1 prevented preadipocyte differentiation while dominant negative FoxO1 restored adipocyte differentiation of fibroblasts from insulin receptor deficient mice.

As described above, the expression and activation of transcription factors were orchestrated in a temporally controlled manner to regulate adipogenesis. Based on the interactions between the transcription factors in the literature, Table 2.1 summarizes interactions between transcription factors and external stimuli that regulate their activity.

2.5 INFLAMMATORY SIGNALING IN ADIPOSE TISSUE

Obesity is characterized by chronic low-grade inflammation and progressive infiltration of macrophages during obesity development. In a hypertrophic state, adipocytes secrete TNF α , which stimulates preadipocytes to produce monocyte chemoattractant protein-1 (MCP-1) (108). In addition, endothelial cells within the adipose tissue secrete MCP-1 in response to the TNF α . The overproduction of MCP-1 induces macrophage infiltration into the adipose tissue (109, 110). Increased secretion of leptin and decreased production of adiponectin in adipocytes also stimulate transport of macrophages to the adipose tissue and contribute to accumulation of macrophages (111).

Table 2.1. Interactions between the transcription factors to regulate adipogenesis

Effector	Target	Type of interaction	Reference
CREB	C/EBP β	Activation	(87), (104)
CREB	C/EBP δ	Activation	(87)
NFAT	PPAR γ	Activation	(106)
FoxO1	PPAR γ	Repression	(107)
C/EBP β	PPAR γ	Activation	(90), (91)
C/EBP δ	PPAR γ	Activation	(90), (91)
PPAR γ	C/EBP α	Activation	(92)
C/EBP α	PPAR γ	Activation	(93)
SREBP-1c	PPAR γ	Activation	(102)

Physical damage to endothelium, caused by crowding or oxidative damage due to a lipolytic condition may also contribute to the macrophage recruitment. Once macrophages are actively present, they perpetuate a cycle of macrophage recruitment, production of the inflammatory cytokines and impairment of adipocyte function (112).

TNF α is one of crucial mediators in the adipose tissue inflammation. Increased levels of TNF α correlate with development of insulin resistance (113) whereas TNF α deficient obese mice remain insulin sensitive (114). Macrophage-derived TNF α which acts through TNF α receptors of adipocytes induces production of proinflammatory cytokines and lipolysis while suppressing production of anti-inflammatory cytokines in adipocytes. Saturated fatty acids (SFAs) are released during macrophage induced adipocyte lipolysis, and several studies have shown SFAs are endogenous ligands for Toll-like receptor 4 (TLR4) which is necessary for recognition of lipopolysaccharide (LPS) (115). The TLR4 activation by SFAs induces expression of inflammatory genes through nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) dependent mechanism (116). An *in vitro* coculture study with macrophages and adipocytes showed that SFAs upregulate expression of TNF α in macrophages through a mitogen-activated protein kinase (MAPK) activity (117). Adipocyte derived T helper cell (Th2) produced cytokines such as interleukin-4 (IL-4) and IL-13 regulate macrophage polarization (proinflammatory M1 state or anti-inflammatory M2 state) which in turn influences production of proinflammatory cytokines in adipocytes. IL-13 treatment increased expression of PPAR δ in macrophages, and adipocytes cocultured

with PPAR δ null macrophages produced more TNF α , MCP-1 and IL-6 as compared to wild type macrophages (118).

Adiponectin which is secreted from adipocytes exerts a protective role against inflammation by promoting macrophage polarization toward an anti-inflammatory M2 phenotype (119). The adiponectin expression is stimulated by PPAR γ ligands such as thiazolidinedione (TZD) (120). In addition to PPAR γ , other transcription factors such as SREBP-1c and C/EBP α have been demonstrated to positively regulate transcription of the adiponectin gene (121, 122). In human primary macrophages, PPAR γ , PPAR α and Liver X receptor (LXR) agonists were shown to increase expression of adiponectin receptors, implying that these transcription factors synergistically drive toward an anti-inflammatory state (123). On the other hand, adiponectin expression is negatively regulated by TNF α and IL-6 (124, 125). As mentioned above, obesity-derived adipose tissue inflammation associates with comprehensive molecular mechanisms.

2.6 NR4A REGULATION IN INFLAMMATION

The nuclear receptor superfamily has emerged as a therapeutic target for the treatment of several diseases such as atherosclerosis, diabetes and obesity (126). Studies on the mechanisms used by the nuclear receptor family to sense environmental cues and translate the signals into specific gene expression have expanded our understanding of these diseases. In addition to ligand activated transcription factors such as PPAR γ , retinoid X receptor (RXR) and liver X receptor (LXR) (127), many orphan receptors whose ligands remain unknown are included in the NR superfamily. Among the orphan

receptors, the NR4A subgroup consists of neuron-derived clone 77 (Nur77; NR4A1), nuclear receptor related 1 (Nurr1; NR4A2) and neuron derived orphan receptor 1 (NOR1; NR4A3).

Members of the NR4A subgroup are operated in a ligand independent manner (128). Structural studies have revealed that NR4A members do not possess a conventional ligand-binding pocket due to bulky side chains of hydrophobic amino acid residues (129-131). However, several small molecules have been recently described as agonists that may interact in unconventional ways. These molecules can work either by modulating C-terminal LBD (131) or N-terminal AF-1 transactivation domain (132). Expression of NR4A members is known to be rapidly induced by various stimuli in a wide range of cultured cells and tissues. Such stimuli includes agonists of G protein-coupled receptors (133, 134), activators of protein kinase and cAMP signaling (135, 136), UV light (137) and mechanical stress (138). Once expression is induced, NR4A members bind to nerve growth factor-induced protein B responsive element (NBRE; AAAGGTCA) as monomers (139). As homodimers or heterodimers with the other NR4A member, they bind to Nur-responsive element, which consists of a repeat of AAAT(G/A)(C/T)CA sequence. In addition, NR4A1 and NR4A2 (but not NR4A3) can form heterodimers with the RXR and bind to a combination of the NBRE and retinoic acid response elements (140).

Recent functional studies have pointed a critical role of NR4A members in control of inflammation. Three NR4A receptors are induced in macrophages during inflammation. The expression of NR4As in macrophages depends on the activation of

NF- κ B signaling (141). Interestingly, functional studies indicated that NR4A receptors both repress and activate inflammatory genes in macrophages. For example, inducible kinase inhibitor of nuclear factor kappa-B kinase (IKK) μ /IKK ϵ which is a NF- κ B activating kinase was identified as a direct NR4A1 target gene and expression of NR4A1 in macrophages potentiated inflammatory gene expression in response to LPS (142). On the other hand, Bonta and co-workers showed that overexpression of each NR4A receptor reduces expression of MCP-1, IL-6 and macrophage inflammatory protein 1 α (MIP1 α) (143). In addition, Saijo and co-workers demonstrated that NR4A2 represses transcription of TNF α and IL-1 β in macrophages and microglia. This transrepression is mediated by recruitment of corepressor for element-1 silencing transcription factor (CoREST) complex to a target promoter and clearance of NF- κ B (144).

Several natural and synthetic compounds have been recently identified as agonists for NR4A receptors, but endogenous ligands have not been identified. The first agonist 6-mercaptopurine was identified through high throughput screening for all NR4A members and some related thiopurine compounds were demonstrated to activate NR4A3. Activation of NR4A2 and NR4A3 by 6-mercaptopurine was independent of the ligand-binding domain (LBD) and involved modulation of activation function-1 (AF-1) within the amino-terminal regulatory domain (129, 145, 146). Several compounds based on 1,1-bis(3'-indolyl)-1-(*p*-substituted phenyl)methanes (C-DIMs) have been identified as agonists of NR4A1. In deletion studies, C-DIM activity was dependent on the LBD of NR4A1. The C-DIMs have been demonstrated to induce proapoptotic pathways in cancer cells, supporting a proapoptotic function of NR4A1 (132, 147). A series of compounds

based on a benzimidazole scaffold were discovered by *in silico* screening methods. These compounds were shown to be potent agonists of NR4A2 *in vitro* (148). Zhan and co-workers identified cytosporone B from *Dothiorella* sp. HTF3 as an agonist of NR4A1, which binds to the LBD and induces apoptosis in cancer cells *in vitro* (149). Thus, manipulation and exploitation of NR4A by these small molecule agonists from the preliminary studies may provide therapeutic modalities against inflammatory diseases.

CHAPTER III

PREDICTION, IDENTIFICATION, AND QUANTIFICATION OF MICROBIOTA METABOLITES IN MURINE GUT USING IN SILICO TARGETED METABOLOMICS

3.1 INTRODUCTION

The human gastrointestinal (GI) tract is colonized by $\sim 10^{14}$ bacteria belonging to $\sim 1,000$ species that are collectively termed the microbiota. Disruptions in the microbiota composition (dysbiosis) are increasingly correlated to not only gut diseases such as inflammatory bowel disease (IBD) and colitis (4), but also obesity, insulin resistance and type 2 diabetes (6, 42). There is increasing evidence that the functional outputs of the microbiota, specifically the metabolites they produce, are important modulators of host physiology. For example, recent work from our laboratory demonstrated that the tryptophan (TRP)-derived bacterial metabolite indole attenuates indicators of inflammation and improves tight junction properties in intestinal epithelial cells (27). An increase in the conversion of dietary choline into trimethylamine by the gut microbiota has also been correlated to non-alcoholic fatty liver disease (11) and cardiovascular diseases (12).

Despite a high level of interest, only a handful of bioactive microbiota metabolites in the GI tract have been identified. One major challenge is that the spectrum of metabolites present in the GI tract is extremely complex, as the microbiota can carry out a diverse range of biotransformation reactions, including those that are not present in

the mammalian host (28). Classical approaches such as isolating and culturing individual bacteria and identifying metabolites produced in these cultures has not yielded much success, as many bacterial species in the GI tract cannot be cultured under standard laboratory conditions. Moreover, this approach also does not account for community-level interactions between the bacteria as metabolites produced by one bacterial species can be utilized or modified by other species resident in the local microenvironment. Another challenge is that it is also extremely difficult to determine whether a metabolite is the product of microbiota or host metabolism, as most metabolite classes are present in both bacteria and mammals due to the high degree of conservation of metabolic pathways across organisms.

In addressing these challenges, metabolomics of fecal or bodily fluid samples has emerged as an attractive approach to explore the metabolite profiles of the GI tract, and to compare these profiles under different conditions. Mass spectrometry (MS)-based approaches have been especially useful in high-throughput identification of a broad spectrum of metabolites. For example, a recent study by Zheng and co-workers (29) used chromatographic separation coupled with MS to characterize the impact of antibiotic treatment on the metabolome of rat fecal and urine samples, and observed that the levels of more than 200 metabolites were significantly altered. Interestingly, TRP-derived compounds such as indole and tryptamine were among the significantly altered metabolites in both fecal and urine samples. In a related study, Antunes and co-workers (66) used Fourier transform ion cyclotron resonance MS to detect more than 2,000 metabolite features in murine fecal samples, and found that a single high dose of

streptomycin causes significant changes in 88% of these features. In addition to MS, highresolution NMR spectroscopy has also been used to broadly profile the metabolites whose levels in feces, bodily fluids or host tissues are significantly altered by interventions (such as bariatric surgery) that are expected to perturb the gut microbiota (150). In general, MS offers greater sensitivity compared to NMR, and has become the dominant analytical platform for metabolomics (151).

To date, a majority of MS studies on profiling GI tract metabolites have utilized an untargeted approach to achieve high throughput. While this approach offers the benefit of potential for discovery, it also has several drawbacks. Due to the complexity of the mass spectra obtained, especially when full scan tandem mass spectrometry (MS/MS) is employed, metabolite identification can be difficult, as it is difficult to discriminate between ions and ion fragments having the same mass-to-charge ratio (m/z). High-resolution time-of-flight (TOF) mass spectrometers can somewhat alleviate this problem (151). However, quantification remains a challenge, because the dynamic range of different metabolites in a biological sample can span up to nine orders of magnitude (32), and an untargeted approach precludes tailoring of MS parameters for ionization and fragmentation of specific, low-abundance metabolites. In contrast, targeted approaches, where the analytes are determined a priori, can use quantitative methods such as multiple reaction monitoring (MRM) that afford custom optimization of MS parameters for individual metabolites to enable sensitive detection. On the other hand, a targeted approach is limited in its discovery potential, as only a focused set of metabolites is analyzed simultaneously. Expanding the number of MRM transitions, and thus detecting

more analytes in a single run, requires a reduction in the dwell time to preserve peak shape, which can compromise sensitivity (151).

In this work, we present a novel targeted metabolomics methodology that addresses the inherent discovery limitation by integrating an *in silico* prediction step into the MS workflow to identify bioactive microbiota metabolites. To date, bioinformatics tools have been utilized in metabolomics generally for post-hoc analysis to process raw data (152), or perform statistical comparisons (153). Recently, Greenblum and co-workers presented an elegant metagenomic study that places obesity or IBD-associated variations in human gut microbiota gene abundances in the context of a microbial community-level metabolic network reconstructed *in silico* (154). Here, we similarly model microbiota metabolism using a metabolic network representing a single “microbial metabolic organ” to computationally explore the products of microbiota metabolism. We thus exploit efficient computational algorithms for network analysis and the growing catalogue of annotated microbial genomes to conduct *in silico* discovery experiments. As a result, the analytical effort can focus on a relatively small number of metabolites to support facile quantitation. To validate our methodology, we use computational pathway analysis to predict bacterial products of tryptophan (TRP) metabolism, and utilize MRM coupled with liquid chromatography (LC) to quantify the levels of the predicted metabolites in murine cecum and feces. To determine bioactivity of the metabolites, we use a *Gaussia luciferase* (Gluc) reporter system measuring aryl hydrocarbon receptor (AhR) activation. Our results have the potential to facilitate mechanistic studies on host-bacteria interactions in the intestinal tract.

3.2 MATERIALS AND METHODS

3.2.1 MATERIALS

All chemicals including HPLC-grade solvents and high-purity metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO) unless noted otherwise.

3.2.2 PREDICTION OF MICROBIOTA METABOLITES

We used a probabilistic pathway construction algorithm to predict possible metabolites that can be derived from a given source metabolite through the metabolic reactions of the gut microbiota. Previously, the algorithm was used to explore novel synthesis pathways for both native and nonnative metabolites in a particular microbial host (*Escherichia coli*) (155). In this study, the algorithm was modified to construct bacterial biotransformation pathways that are nonnative to the mammalian host organism, i.e. mouse, such that the intermediates of the constructed pathways could be unambiguously sourced to bacterial metabolism. A list of metabolic reactions for the mouse was compiled from published genome-scale models (156, 157) and manually crossreferenced with the KEGG database to identify the corresponding reaction and compound identification numbers. The algorithm recursively constructs a tree, starting from a source metabolite as a root of the tree. A single reaction is randomly selected from a list of candidate bacterial reactions in the KEGG enzyme database that involve the source metabolite as a main reactant. This list, obtained through KEGG's FTP site, comprised all bacterial enzymes catalogued in the database as specified by the genes

entry. An enzyme was determined to be bacterial if at least one bacterial species encoded the enzyme in its genome. The selected reaction is then added to the tree and represented by an edge. This edge expands the tree by attaching new nodes representing the product metabolites and cofactors of the selected reaction. The construction thus proceeds in a depth-first fashion. Each of these nodes is a new root for the recursion, unless the corresponding metabolite or cofactor is already present in the host organism or was previously added to the tree. To achieve reasonable runtimes (on the order of tens of minutes for a run of 1,000 – 2,000 iterations), the size of the search space was constrained by placing an upper limit on the number of reactions that can be used to construct a pathway. In this study, the upper limit was varied from 20 to 50, which had no observable impact on the distribution of the predicted metabolites. When the addition of a reaction to the tree violates the upper limit, the algorithm backtracks and proceeds by adding to the tree another reaction that has not been previously explored, effectively identifying an alternative pathway. If none of these alternative routes satisfy the pathway length limit, the algorithm further backtracks and continues from there. The algorithm finishes when all permitted-length branches of the tree terminate in a metabolite that is native to the host organism. Due to the probabilistic nature of selecting the reactions, the completed tree does not exhaustively enumerate all possible pathways. Rather, each tree represents a single pathway from the source metabolite to one or more product metabolites that are native to the host organism. Therefore, the search is iterated many times to explore a diverse number of possible pathways. In our previous work, we found that the probabilistic search matches an exhaustive search in terms of sampling diversity,

and dramatically outperforms the exhaustive search in terms of computational efficiency (155).

3.2.3 SAMPLE COLLECTION

Female C57BL/6 mice at 5 weeks and 14 weeks of ages were purchased from Jackson Laboratories (Bar Harbor, Maine) and allowed to orient to colony for 1 week. All mice were maintained in a pathogen free animal facility located at Texas A&M Health Science Center. The animals were handled in accordance with the Institutional Animal Care and Use Committee guidelines under an approved animal use protocol. The young and old mice ($n = 7$) were sacrificed at 6 and 15 weeks of age, respectively. The entire cecum (tissue with luminal contents) and fecal pellets were collected from each animal. The samples were weighed, flash frozen, and stored at -80°C before processing for extraction.

3.2.4 METABOLITE EXTRACTION

Metabolites were extracted from cecum or fecal pellets luminal contents and fecal samples using a solvent-based method (158) with minor modifications. Briefly, 1.5 ml of ice-cold methanol/chloroform (2:1, v/v) was added to a sample tube containing a pre-weighed luminal content or fecal sample. After homogenization on ice, the sample tube was centrifuged under refrigeration (4°C) at $15,000 \times g$ for 10 min. The supernatant was then transferred to a new sample tube through a (70- μm) cell strainer. After adding 0.6 mL of ice-cold water, the sample tube was vortexed vigorously and centrifuged

under refrigeration (4 °C) at $15,000 \times g$ for 5 min to obtain phase separation. The upper and lower phases were separately collected into fresh sample tubes with a syringe, taking care not to disturb the interface. To improve signal intensity for MS, 500 μ L of the polar phase was concentrated by solvent evaporation in a Savant speedvac concentrator (Thermo Scientific, Asheville, NC), and then reconstituted in 50 μ L of methanol/water (1:1, v/v). Extracted metabolites were stored at -80°C until analysis.

3.2.5 METABOLITE ANALYSIS

Prior to sample analysis, MS parameters were optimized for each target metabolite to identify the MRM transition (precursor/product fragment ion pair) with the highest intensity under direction injection at 10 μ L/min. The following parameters were optimized operating in positive mode: declustering potential (DP), entrance potential (EP), collision energy (CE), and collision cell exit potential (CXP). The optimized parameter values for the target metabolites analyzed in this study are shown in Table 3.1. The target metabolites in samples were detected and quantified on a triple quadrupole linear ion trap mass spectrometer (3200 QTRAP, AB SCIEX, Foster City, CA) coupled to a binary pump HPLC (1200 Series, Agilent, Santa Clara, CA). Peak identification and integration were performed using Analyst software (version 5, Agilent, Foster City, CA). Samples were maintained at 4°C on an autosampler prior to injection. Chromatographic separation was achieved on a hydrophilic interaction column (Luna 5 μ m NH₂ 100 Å 250 mm \times 2 mm, Phenomenex, Torrance, CA) using a solvent gradient method (159).

Table 3.1. Optimized MRM-MS parameters for detection of TRP metabolites

Compound	Precursor (Da)	Product (Da)	DP (V)	EP (V)	CE (V)	CXP (V)
Indole	118.0	91.0	41.0	10.0	27.0	2.5
Indole 3-acetate	176.0	130.0	31.0	9.0	19.0	4.0
Indole 3-acetamide	175.0	130.0	26.0	10.0	19.0	4.0
Tryptamine	161.0	144.2	11.0	4.0	15.0	4.0
Hydroxyindole	134.1	77.1	26.0	10.0	37.0	2.5
Tryptophan	205.0	188.0	21.0	10.0	17.0	4.0

Solvent A was an ammonium acetate (20 mM) solution in water with 5 % acetonitrile (v/v). The pH of solvent A was adjusted to 9.5 immediately prior to analysis using ammonium hydroxide. Solvent B was pure acetonitrile. Injection volume was 10 μ L. The gradient method, including solvent flow rate and composition, is shown in Table 3.2.

3.2.6 PARTITION OF METABOLITES IN EXTRACTION SOLVENT

Since only the upper (polar) phase of the biphasic metabolite extract was used for analysis, the partition of all five metabolites between the two phases was determined for absolute quantification. Briefly, 5 μ L of the metabolite's stock solution (1 mg/mL) was dispensed into a 1.5 mL microfuge tube. To this pure stock solution, 750 μ L of 2:1 methanol:chloroform was added, followed by 300 μ L of water. Following centrifugation and phase separation, the upper phase was collected and analyzed for the target metabolite using the corresponding optimized MRM parameters. The difference between measured and theoretical amounts of the metabolite was used to determine the partition between phases. For example, the theoretical amount of tryptophan in the stock solution-solvent mixture is 25 nmol. The measured concentration in the upper phase was 34 μ M, which corresponds to an estimated total amount of ~27 nmol in the upper phase (volume 800 μ L), suggesting that all of the tryptophan is recovered within a margin of (e.g. pipetting) error. Similarly, all other metabolites analyzed in this study showed full partition into the upper phase, except indole. For indole, of the 42 nmol dispensed, only

Table 3.2. Solvent gradient method for metabolite separation

Time (min)	Flow rate ($\mu\text{L}/\text{min}$)	A (%)	B (%)
0	300	15	85
15	300	100	0
28	300	100	0
30	300	15	85
50	300	15	85

17 nmol was recovered in the upper phase, indicating that the actual amount of indole extracted from the tissue is ~2.5 times greater than the measured amount.

3.2.7 STATISTICAL ANALYSIS

Comparisons of medians between the metabolite levels of young and old mice were performed with the non-parametric two-sided Mann-Whitney U-test. The null hypothesis that the two medians are the same was rejected for $p < 0.05$. Comparisons of means for the reporter experiments were performed using the Student's t-test. The null hypothesis that the two means are the same was rejected for $p < 0.05$.

3.2.8 CELL CULTURE

A MCF-7 human breast cancer cell line was obtained from ATCC (Manassas, VA). MCF-7 cells were cultured at 37 °C with 5 % CO₂ in RPMI 1640 medium (MP Biomedicals, Solon, OH) supplemented with 10% (v/v) fetal bovine serum (FBS), glucose (2.5 g/L), HEPES (10 mM), sodium pyruvate (1 mM), sodium bicarbonate (2 g/L), penicillin (100 U/ml) and streptomycin (100 µg/ml). A stably transfected rat hepatoma (H4IIE) cell line containing an AhR-responsive enhanced green fluorescent protein (EGFP) reporter was kindly provided by Prof. Michael Denison (University of California, Davis, CA) (160). The H4IIE cells were cultured at 37 °C with 5 % CO₂ in alpha minimum essential medium (Mediatech, Manassas, VA) supplemented with 10 % (v/v) FBS, glucose (3.5 g/L), sodium bicarbonate (2.2 g/L), penicillin (100 U/ml) and

streptomycin (100 µg/ml). Serum-free medium was used for assays on AhR activity in the H4IIE reporter cells.

3.2.9 CONSTRUCTION OF GAUSSIA LUCIFERASE REPORTER PLASMID

A lentiviral reporter plasmid for monitoring activation of AhR was constructed as described below. AhR response elements (RE) in target promoter were identified using the TRANSFAC database 7.0 Public. An oligonucleotide containing three repeats of the binding sequence (CTGAGGCTAGCGTGCGT) separated by 4 - 6 bases (spacer sequence) was chemically synthesized with two restriction enzyme (EcoRI and AfeI) cleavage sites at the ends. The RE oligonucleotide was cloned into a lentiviral vector (161) in which expression of the *Gaussia* luciferase (Gluc) is under the control of a minimal CMV promoter and red fluorescent protein (RFP) is constitutively expressed. Expression of Gluc is induced when ligand-activated AhR binds to its RE. Clones containing the correct RE were identified by multiple restriction enzyme digests and verified by sequencing.

3.2.10 GENERATION OF A STABLE MCF-7 REPORTER CELL LINE

A stable MCF-7 AhR reporter cell line was generated by lentiviral transduction. To produce lentiviral particles, AhR reporter plasmid and helper plasmids psPAX (plasmid 12260, Addgene, MA) and pMD2.G (plasmid 12259, Addgene) were co-transfected into 293T/17 cells using the calcium phosphate transfection method (162). After 24 h following the transfection, the medium was replenished and 5 mM of sodium

butyrate was added. After an additional 24 h of incubation, culture supernatants containing viral particles were collected, pooled, filtered with 0.45 μm filters, and centrifuged for 2 h at 4°C at 48000 x g. The viral titer was measured using a Lenti-X qRT-PCR titration kit (Clontech, Palo Alto, CA). To transduce MCF-7 cells, a concentrated aliquot of virus particles ($\sim 1 \times 10^8$ IFU) was added to the cells in presence of Polybrene (hexadimethrine bromide). After 4 h of incubation with the virus particles, the medium was replenished.

3.2.11 AHR ACTIVATION STUDIES

The MCF-7 and H4IIE reporter cells were seeded in 24-well tissue culture plates and grown to 70% confluence. Cells were treated with indicated concentrations of target metabolites (i.e., indole, tryptamine, indole-3-acetate, indole-3-pyruvate and indole-3-acetamide). The negative and positive controls were 0.1% (v/v) N, N-dimethylformamide (DMF) and 20 nM 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), respectively. For assays using the MCF-7 reporter cells, 20 μL of culture supernatant was collected at 24 h and 48 h post-treatment. The MCF-7 supernatant samples were stored at -20 °C until the secreted luciferase activity was measured. The luciferase activity (relative light units; RLU) was used to calculate the rate of Gluc production (RLU divided by the time over which Gluc was secreted). To account for differences in cell density between different experiments, the Gluc production rate was normalized by the intensity (relative fluorescence units; RFU) of the constitutively expressed RFP measured at 550/600 nm excitation/emission. The AhR activity reported by H4IIE cells

was quantified using fluorescence optical microscopy based on the green fluorescence intensity of EGFP.

3.3 RESULTS

3.3.1 TRYPTOPHAN DERIVATIVES

Starting with tryptophan as the source metabolite, biochemical transformation pathways were computationally constructed from candidate bacterial reactions catalogued in the KEGG database (163). As the goal was to predict metabolic derivatives that could be formed by the gut microbiota, but not the murine host, pathway construction was terminated when a metabolite was encountered that could also be formed through host (murine) metabolism. Due to the probabilistic nature of the pathway construction algorithm, the results can vary with the number of iterations. Therefore, we performed a series of simulations with varying iteration numbers, and examined the diversity of predicted metabolic derivatives. No further increase in the number of unique metabolic derivatives was observed beyond iteration number 1,500 (Fig. 3.1A). Moreover, the ordering of the metabolic derivatives in terms of their relative selection frequency also remained constant. Holding the iteration number constant at 1,500, we next performed 50 separate runs of the pathway construction algorithm to calculate average frequencies for the selected metabolic derivatives (Fig. 3.1B). Of the 26 predicted metabolic derivatives, 12 derivatives with the highest selection frequency accounted for > 90% of all metabolite selections as pathway intermediates. Interestingly,

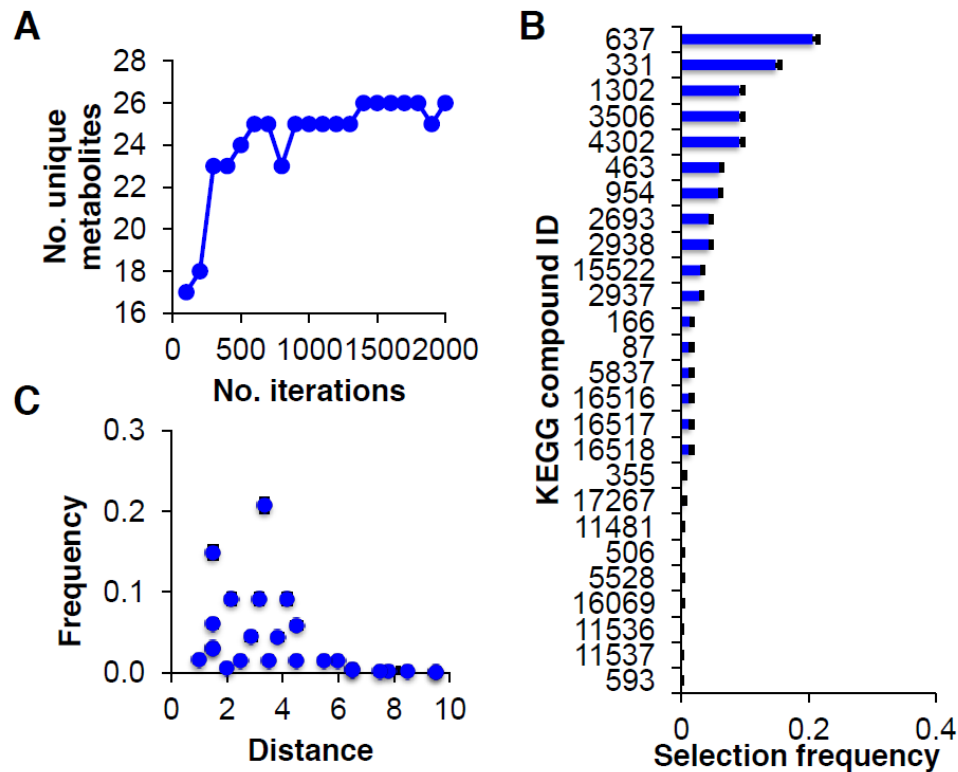


Fig. 3.1. Probabilistic pathway construction with a tryptophan source. (A) Number of unique metabolites selected using probabilistic pathway construction with TRP as the source metabolite. (B) Average selection frequencies of TRP pathway metabolites at iteration number 1,500 (50 repeated runs). (C) Correlation of metabolite selection frequency and distance from the source metabolite. In (B) and (C), error bars indicate standard deviations.

a significant correlation between selection frequency and reaction distance (number of enzymatic steps) from the source metabolite TRP was not observed for these 12 most frequently selected metabolites. However, there was a significant negative correlation between selection frequency and reaction distance when all 26 metabolites were included in the analysis (Fig. 3.1C). The metabolite farthest away from the source metabolite was acetyl-phosphate, with an average reaction distance of 10.5, which is well below the pathway length limit of 50 we set for the algorithm.

Representative tryptophan biotransformation routes and intermediates predicted by the pathway construction algorithm are shown in Fig. 3.2. To explore potential biotransformation of TRP products that could be present in the intestinal mucosa, we also analyzed the metabolic derivatives of two neighboring metabolites, tryptamine and indole, both of which can be formed from TRP in a single enzymatic step. Tryptamine can be formed via an aromatic-L-amino-acid decarboxylase (EC 4.1.1.28), which is encoded by the dopa decarboxylase gene (*ddc*) in the mouse (164), although the expression of this enzyme has not been reported in intestinal tissue. Indole can be formed via tryptophanase (*tnaA*), which is expressed in a number of species of the murine gut microbiome. For tryptamine, the number of unique derivatives identified by our algorithm was 54 at an iteration number of 1,500. Similar to tryptophan, the selection frequencies were highly uneven, with 10 metabolites accounting for >95% of all metabolite selections. Every one of these 10 metabolites was also predicted as TRP derivatives. For indole, the maximal number of unique derivatives identified by our algorithm was only three: 1-(2-carboxyphenylamino)-1-deoxy-D- ribulose 5-phosphate,

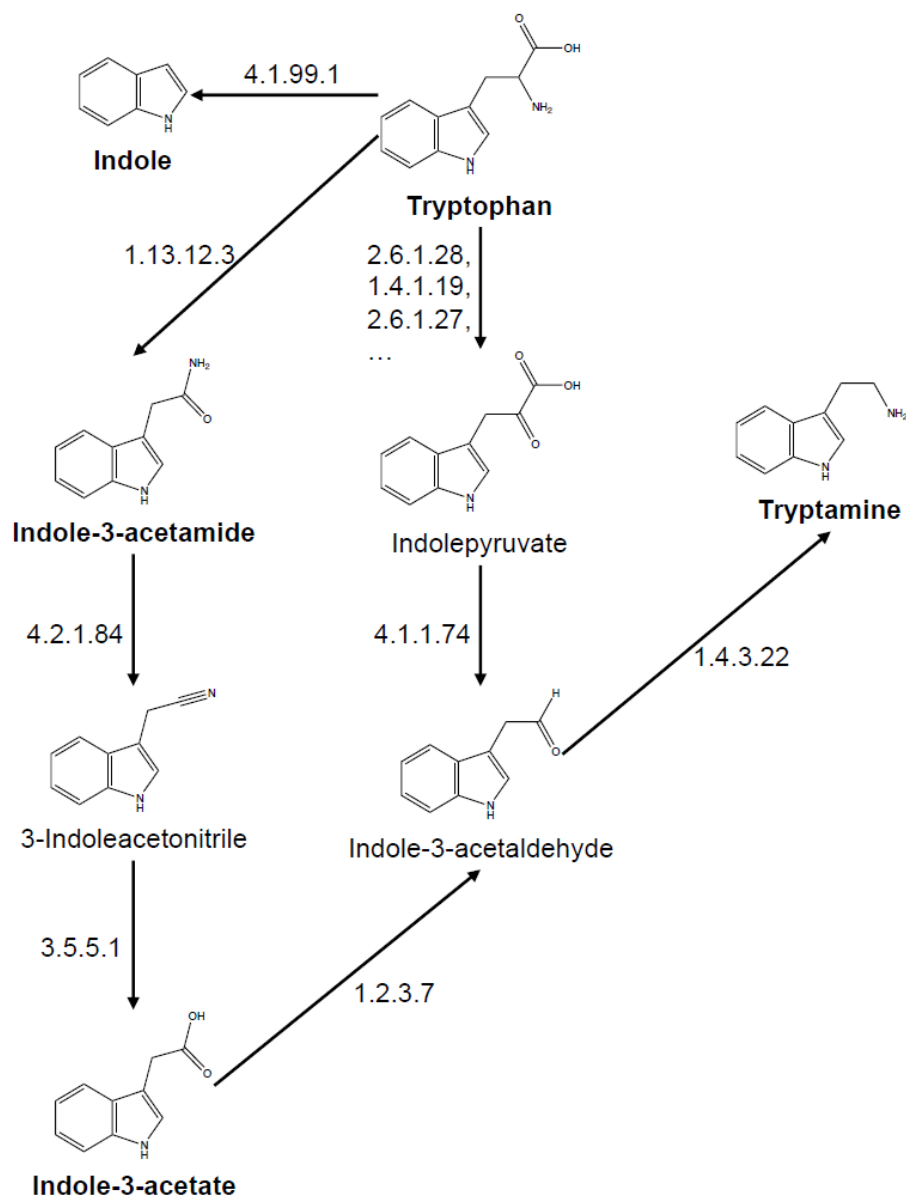


Fig. 3.2. Representative tryptophan biotransformation routes and intermediates predicted by the pathway construction algorithm. Metabolites quantified in this study using MRM MS are indicated in bold. Arrows representing enzymatic reactions are labeled by the corresponding enzyme commission number (EC#). Some reactions (e.g. conversion of tryptophan into indolepyruvate) can be catalyzed by multiple enzymes. Similarly, some metabolic derivatives (e.g. indole-3-acetaldehyde) can be produced through multiple routes.

indoleglycerol phosphate, and N-(5-phospho-D-ribosyl)anthranilate. The iteration number needed to obtain a stable set of relative selection frequencies for the predicted derivatives was also lower than TRP or tryptamine (only 200), presumably due to the lower connectivity of reactions involving indole. Similar to the high-frequency tryptamine derivatives, every indole derivative was a subset of the predicted TRP derivatives.

3.3.2 QUANTIFICATION OF METABOLITES USING MULTIPLE REACTION MONITORING (MRM)

Of the 12 high-frequency metabolites predicted by the pathway algorithm to be derived from TRP by bacteria, a subset was selected for MRM analysis based on availability of pure standards and ease of ionization and fragmentation. For example, indole 3-acetaldehyde was excluded, because a high-purity was not available to generate a standard curve. Indole 3-pyruvate was excluded, because it ionized poorly, and thus had a high limit of detection (LOD) in comparison with the other metabolites targeted for analysis. The final panel of metabolites targeted for analysis comprised indole, hydroxyindole, indole 3-acetate, indole 3-acetamide, tryptamine and TRP. Hydroxyindole was added to the panel based on recent reports suggesting that it is derived from indole through bacterial reactions (20). All six metabolites were simultaneously identified and quantified in murine fecal and cecum extracts using MRM MS following separation by hydrophilic interaction chromatography (HILIC). Metabolite identification was performed based on both chromatographic retention time

and mass signatures, as we found that even an optimized MRM transition (precursor-product ion pair) did not always uniquely identify a metabolite in a complex biological sample. For indole, the optimal ion pair transition based on MS signal intensity was 118/91 m/z. The elution time of this transition ranged from 2.2 and 2.9 min for standards containing the corresponding high-purity chemical (Fig. 3.3A). The linear response of signal area under curve (AUC) with respect to concentration is shown for concentrations up to 85 μ M (Fig. 3.3A; $R^2 = 0.998$, $p < 10^{-4}$), with an estimated LOD (defined as the peak height of 10-times the baseline signal) of 4.6 μ M. For tissue extracts, we observed multiple peaks corresponding to the signal-optimized indole transition, highlighting the importance of matching the retention time to unambiguously identify a metabolite (Fig. 3.3B, Fig. 3.4 for chromatograms of all metabolites). To determine whether the MRM method could be used to detect physiological differences in metabolite levels, we compared the fecal pellet and cecum concentrations of TRP and its derivatives in young (age 6 weeks) and old (age 15 weeks) female C56BL/6 mice fed standard chow (Fig. 3.5). All six metabolites were detected in both fecal pellet and cecum extracts from young and old mice. For cecum extracts, statistically significant differences ($p < 0.05$) were found between young and old mice for all metabolites except tryptamine. For fecal extracts, statistically significant differences were found only for indole 3-acetate. Unlike the fecal extracts, metabolite levels in the cecum extracts could be meaningfully expressed as tissue concentrations by normalizing the absolute molar quantities with respect to the weight of the source tissues and approximating tissue density with that of

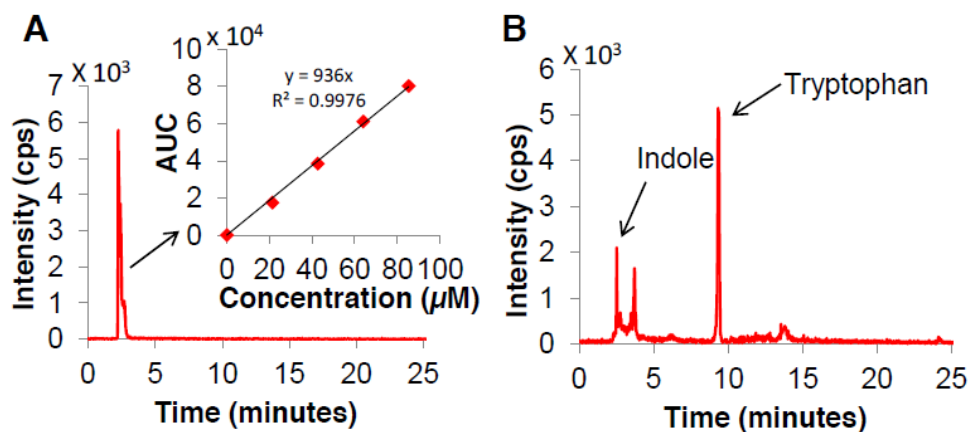


Fig. 3.3. LC-MS/MS chromatogram for indole. (A) indole standard (107 ng/L). Peaks correspond to 118 > 91 (m/z) MRM transition. Retention time of indole lies between 2.2 and 2.9 minutes. Signal Area Under Curve (AUC) is also plotted as a function of indole concentration ($R^2 = 0.998$, $p < 10^{-4}$). (B) Representative sample with peaks corresponding to the 118 \rightarrow 91 (m/z) MRM transition. Sample shown was extracted from the cecum of a young mouse (6 weeks of age).

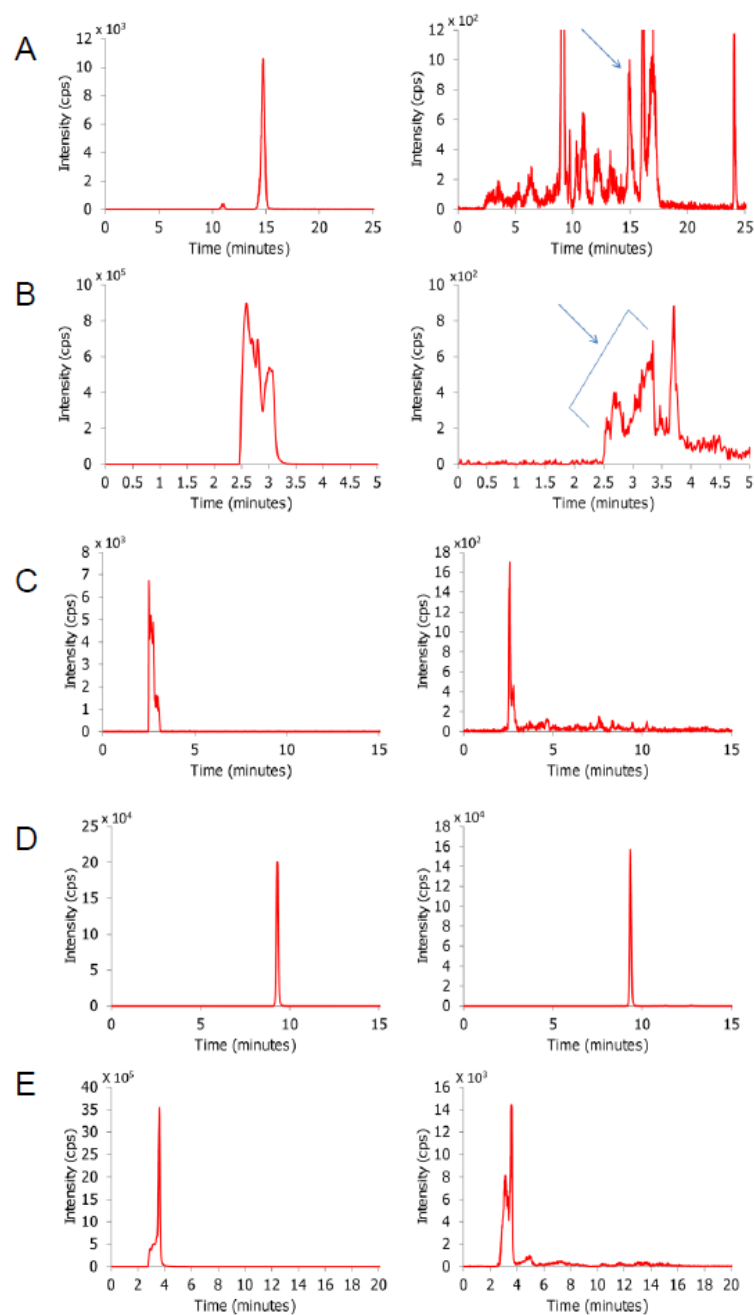


Fig. 3.4. LC-MS/MS chromatogram for all target metabolites. (A) Indole 3-Acetate, MRM: 176>130 (B) Indole 3-Acetamide, MRM: 175>130 (C) Hydroxyindole MRM: 134.1>77.1 (D) Tryptophan MRM: 205>188 (E) Tryptamine, MRM: 161>144.2, standards (left) and metabolite extracts of week 5 cecum samples (right).

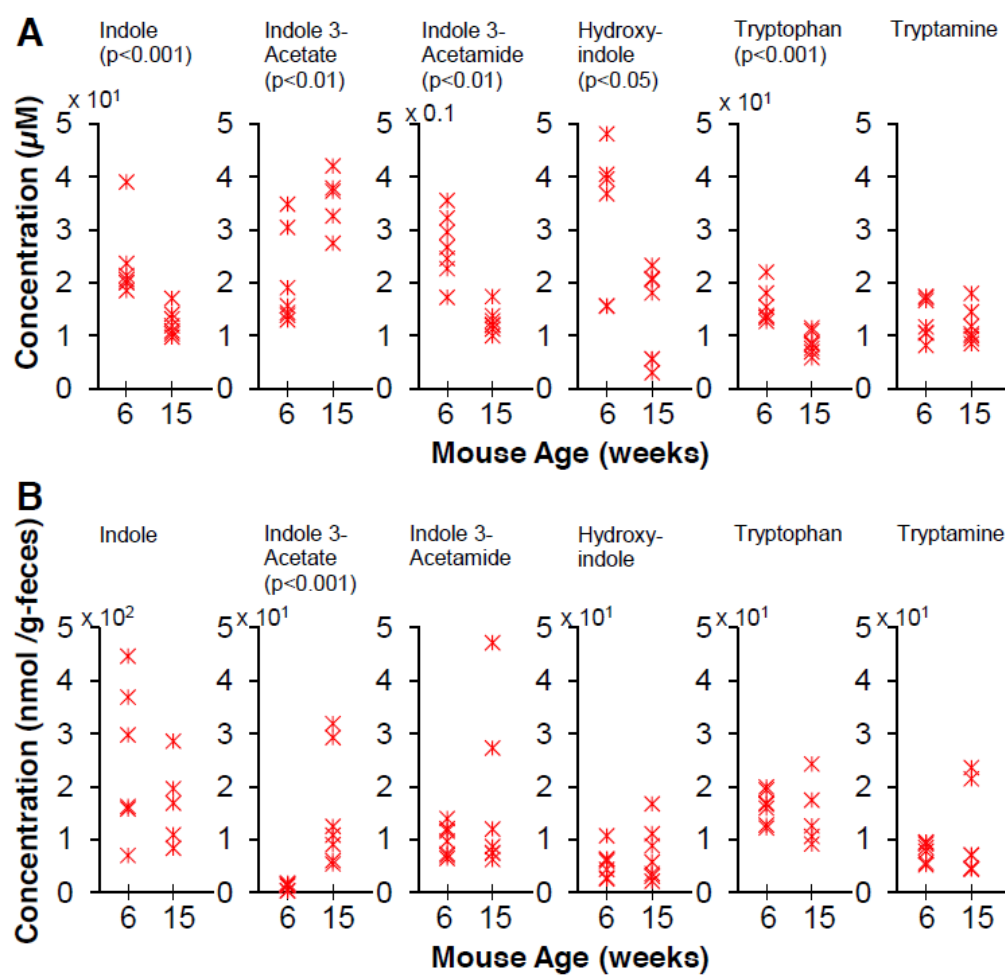


Fig. 3.5. Comparison of metabolite concentrations between young (6 weeks of age) and old (15 weeks of age) mice. Metabolite concentrations were normalized to the corresponding mass of fecal pellet. (A) cecum extracts (B) fecal extracts.

water (1 g/ml). For example, after accounting for the partition of indole between the polar and nonpolar phases of the metabolite extraction solvent, the normalized tissue concentrations for indole ranged from 9.7 to 40 μ M. These are conservative estimates, as it is expected that the efficiency for the extraction of metabolites from tissue into solvent is less than 100 %.

3.3.3 ACTIVATION OF AHR BY MICROBIOTA METABOLITES

We investigated if the identified microbiota metabolites could be putative ligands in host cells and activate eukaryotic signaling pathways. Specifically, we investigated the ability of identified microbiota metabolites to act as agonists for the AhR, as previous studies (83) have shown that endogenous TRP-derived metabolites can activate AhR signaling. MCF-7 cells with a stably integrated *Gaussia luciferase* (Gluc) reporter plasmid for AhR binding activity were incubated with 100 μ M microbiota metabolites or 20 nM TCDD (positive control) for 48 h, and luciferase activity in culture supernatants was measured. Exposure to TCDD, a high-affinity AhR agonist, resulted in a 2-fold increase in the rate of AhR-driven luciferase activity (RLU/h/RFU) as compared to the solvent control (Fig. 3.6). Exposure to indole-3-acetate, tryptamine and indole-3-pyruvate also resulted in similar induction of AhR activity. Indole-3-acetamide resulted in a 1.3-fold statistically significant ($p < 0.02$) increase in AhR binding activity at 24 h but not at 48 h. Interestingly, indole did not induce activation of AhR, suggesting that the core indole moiety is by itself not sufficient to elicit a significant response. Similar results were also observed in the H4IIE rat liver reporter cell line using GFP as the

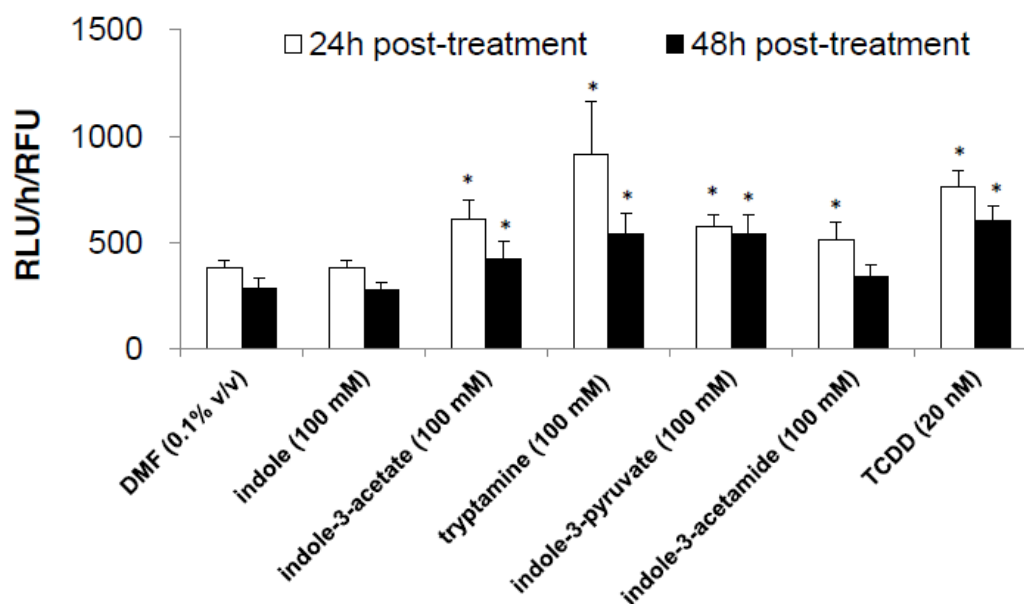


Fig. 3.6. AhR activation by TRP-derived metabolites. Activity assays were performed at a dose of 100 μ M for each metabolite. AhR activity is reported as the rate of luciferase activity normalized to red fluorescence of constitutively expressed RFP (RLU/h/RFU). Positive and negative controls were 20 nM TCDD and 0.1% (v/v) DMF, respectively.

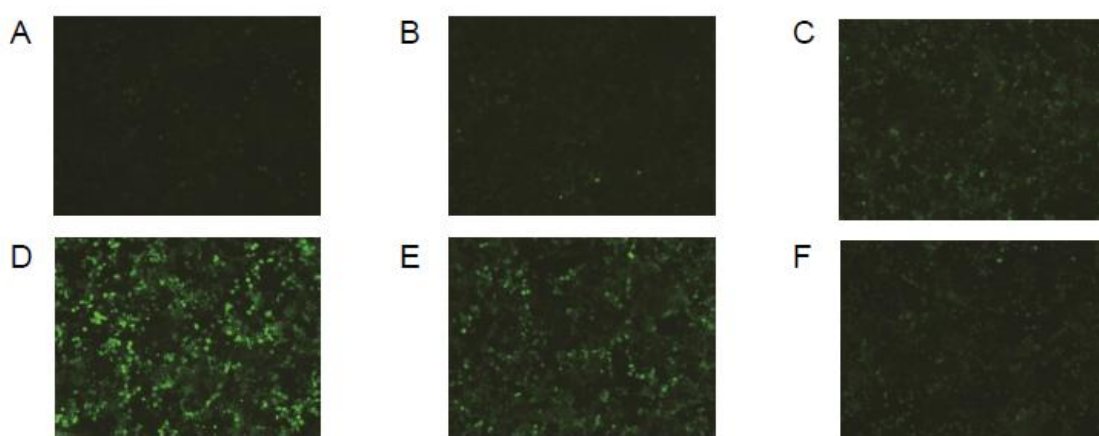


Fig. 3.7. Microbiota metabolites screen for AhR activation in the H4IIE reporter cells. The H4IIE cells containing an AhR-responsive enhanced green fluorescent protein (EGFP) reporter were seeded in 24-well tissue culture plates and cultured to 70% confluence. Cells were treated with 100 μ M of target metabolites or 0.1 % (v/v) of DMF. Green fluorescent images were taken at 48h post-treatment. (A) DMF, (B) indole, (C) indole-3-acetate, (D) tryptamine, (E) indole-3-pyruvate, (F) indole-3-acetamide

readout (Fig. 3.7). For metabolites that induced AhR binding activity at 100 μ M, we performed dose response experiments to determine concentration effects on induction of AhR activity. Exposing MCF-7 reporter cells to increasing concentrations of indole-3-acetate and tryptamine resulted in a dose-dependent increase in AhR activity (Fig. 3.8). Exposure to indole 3-pyruvate, one of the *in silico* predicted metabolites, but not quantified due to poor ionization, also showed a dose dependent increase (Fig. 3.9). The dose-dependent increase in AhR-dependent reporter activity was also confirmed using H4IIE reporter cells (Fig. 3.10).

3.4 DISCUSSION

In this study, we present a methodology for the prediction, identification and quantification of gut microbiota metabolites that integrates computational pathway analysis into a targeted metabolomics workflow. The effects of physiological or pathological perturbations on the microbiota have been investigated using metagenomic analyses characterizing the composition of the gut microbial community, the enrichment (or depletion) of bacterial genes, or the expression levels of genes for specific metabolic pathways (20, 165, 166). A limitation of these analyses is that they do not provide direct information on the products of bacterial metabolism such as which molecules are formed from bacterial biotransformation reactions and the concentrations at which these metabolic products are present. Thus, the ability to unambiguously identify bacterial metabolites and quantify their levels in the GI tract is expected to have a significant impact on the study of human gut microbiome function.

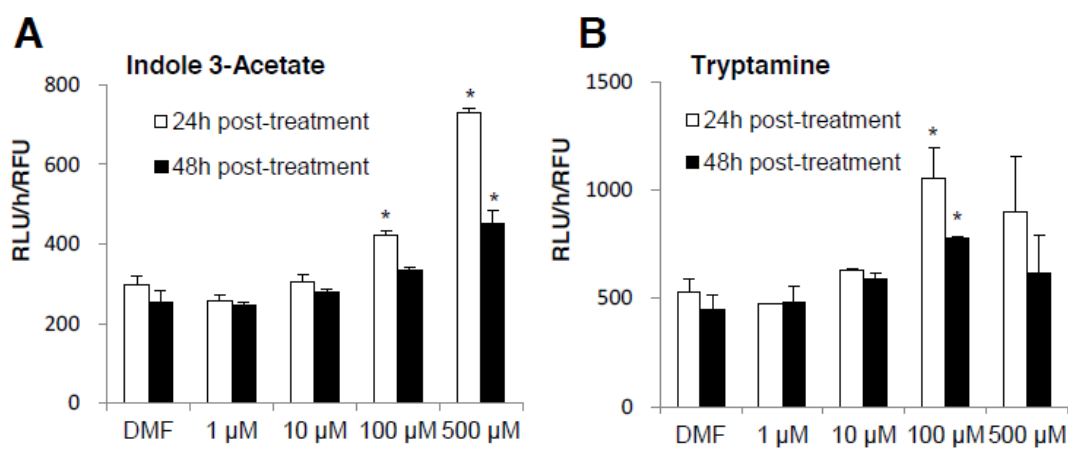


Fig. 3.8. Dose dependent activation of AhR upon exposure to indole-3-acetate and tryptamine. Data shown are mean \pm standard deviation from four replicate experiments. Asterisk (*) indicates statistical significance at $p < 0.05$.

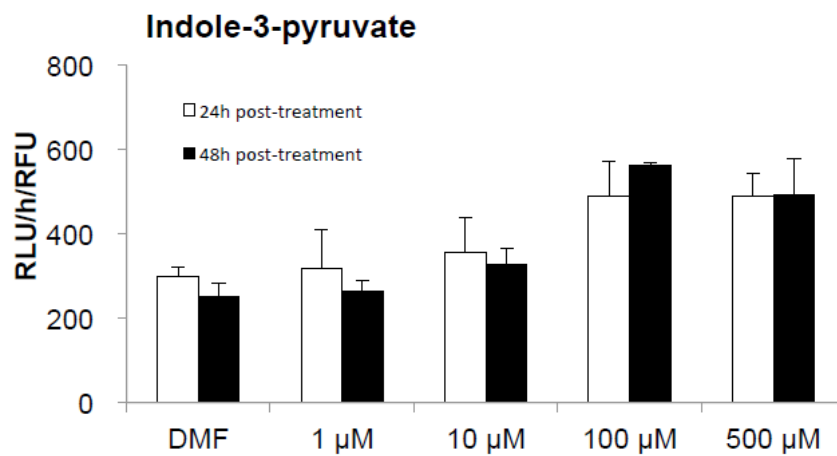


Fig. 3.9. Dose dependent activation of AhR upon exposure to indole 3-pyruvate. Data shown are mean \pm standard deviation from four replicate experiments. Asterisk (*) indicates statistical significance at $p < 0.05$.

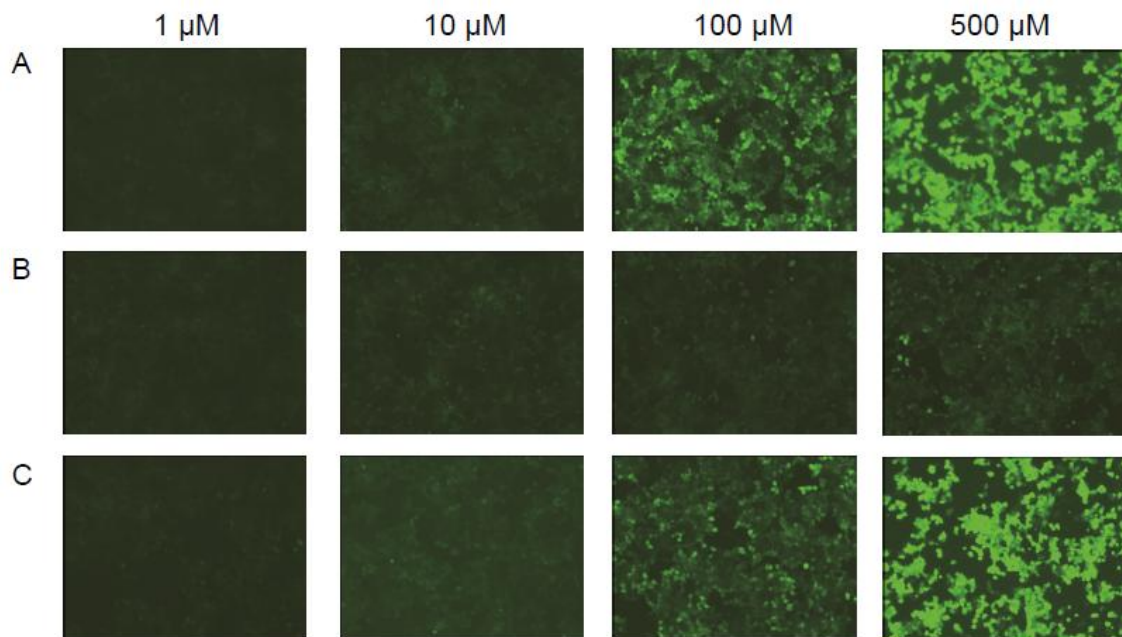


Fig. 3.10. Dose-dependent AhR activation in the H4IIE reporter cells. (A) Tryptamine, (B) Indole 3-Acetate, and (C) Indole 3-Pyruvate. The H4IIE cells containing an AhR-responsive enhanced green fluorescent protein (EGFP) reporter were seeded in 24-well tissue culture plates and cultured to 70% confluence. Cells were treated with indicated concentrations of target metabolites. Green fluorescent images were taken at 48h post-treatment.

One limitation of the *in silico* metabolite prediction step is that its reliability is predicated on the accuracy and completeness of the KEGG pathway database. While the KEGG database is continually and frequently updated, it is certainly possible that our predictions lag behind newly published discoveries regarding microbiota metabolism. For example, Wikoff and coworkers recently identified indole 3-propionate as a TRP-derived metabolite that is produced by the gut microbiota (64). However, at the time of completion of this work, this metabolite was not listed in the KEGG compound database, and thus could not be identified by our algorithm. Similarly, the discrimination of bacterial metabolites from metabolites that could also be produced by host cells depends on an accurate model of host metabolism. The most relevant host genomes, e.g., mouse and human, have been sequenced and largely annotated, and several published models are available that exhaustively catalogue the metabolic reactions. However, prior studies with genome-scale model reconstruction suggest that several iterations may be required until the published models can be considered stable, consensus reconstructions (167).

Another limitation of the prediction algorithm is that it does not differentiate between reactions based on their gene abundance or expression level when constructing a possible biotransformation route. Consequently, each candidate reaction that can be connected to the source metabolite has an equal likelihood of selection, which unlikely reflects the true engagements of metabolic reactions in the gut microbiota. In principle, metabolites produced by reactions encoded by genes that are highly abundant and/or highly expressed should more likely be present at quantifiable levels compared to the products of depleted or minimally expressed pathways. In this regard, both of the present

limitations of the *in silico* prediction step could be addressed by incorporating metagenomic data on as they become available. Our pathway construction already accepts user-specified selection weights, and could be extended in a straightforward manner to explore a microbiota metabolic network weighted by relative gene abundances or expression levels.

A key advantage of MRM as a MS technique for targeted metabolomics is that the detection of precursor-product ion pairs offers greater specificity compared to full scan MS, while allowing absolute quantitation. This targeted approach also allows for enhanced sensitivity as well as improved LOD, because instrument-specific parameters can be tailored and optimized for the detection of each individual MRM transition, whereas full scan methods are restricted to one fixed set of parameters for all analytes, which may be suboptimal for the ionization and fragmentation of certain metabolites. In practice, even optimized MRM transitions may not represent a unique mass signature for a metabolite, as there are cases where multiple analytes present in a biological sample share the same transition. For example, indole 3-acetamide shows a strong signal for the $175 \rightarrow 130$ transition, as does arginine, a highly abundant amino acid. This overlap in MRM transitions by different compounds underscores the critical importance of chromatographic separation in identifying metabolites. Another common problem that potentially compromises specificity is the tendency for metabolites possessing thermally labile bonds to decompose due to heated electrospray ionization (HESI). For example, a sample containing only high-purity TRP showed a strong signal for the indole transition

(118 → 91) at the same retention time as TRP (Fig. 3.6B), presumably due to partial decomposition into indole at the ion source.

To our knowledge, this is the first study to use MRM to quantify physiological concentrations of microbiota-produced metabolites present in intestinal tissue extracts. While the literature on absolute concentrations of microbiota metabolites is relatively sparse, we found good agreement between our results and previously reported values. In an early study, Whitt and coworkers used an enzymatic assay to determine an indole concentration of ~40 nmol/g tissue in murine cecum (168), which is comparable to our results (16 - 31 nmol/g tissue in cecum of young mice). The earlier study also reported that indole was absent in the cecum of germ-free mice, which is consistent with our pathway analysis algorithm's prediction that indole is a product of bacterial metabolism. A comparison of samples from young and old mice suggests that the age of the host impacts the levels of microbiota metabolites in the GI tract as reflected in cecum and fecal concentrations of bacterial TRP derivatives. This observation is consistent with several recent studies showing that age may influence levels of other microbiota-associated metabolites. For example, Vaahtovuori and coworkers reported differences in bacteria-derived cellular fatty acids in stool samples between 5 - 7 and 15 - 19 weeks old mice (169).

The AhR is a ligand-activated transcription factor that plays an important role in the mucosal immune system (71), and several TRP-derived chemicals have been identified as AhR ligands (81). In our study, we observed that three of the predicted TRP derivatives (indole 3-acetate, tryptamine and indole 3-pyruvate) were able to activate

AhR in a dose-dependent manner (Fig. 3.8, Fig. 3.9). This result is consistent with a previous study by Heath-Pagliuso *et al.*, who showed that tryptamine and indole-3-acetate function as AhR agonists (83). However, this previous study, while assuming that these two metabolites could be derived from TRP through enzymatic reactions, did not confirm that they are actually present in the GI tract in quantifiable amounts. In this work, we show that putative AhR agonists like tryptamine and indole-3-acetate are present in the cecum at intracellular concentrations (i.e., the levels present in the cecal bacteria and the surrounding tissue) ranging from 1 to 5 μ M, and that they activate the AhR *in vitro* at concentrations of \sim 100 μ M. This apparent discrepancy between the measured concentration of metabolites and the concentrations required to activate the AhR *in vitro* could be due to two reasons. First, the measured concentration of these metabolites is a conservative estimate, as it does not account for the metabolite extraction efficiency. Second, the concentrations needed to activate the AhR *in vitro* are extracellular concentrations (i.e., concentrations in the culture medium), and since the AhR is an intracellular (cytosolic) receptor, the extracellular medium concentrations may not equate to intracellular levels, depending on the rates of uptake and metabolism inside the cells. Further analysis is needed to model the kinetics of ligand uptake and processing that lead to the activation of the AhR.

An obvious extension of this work is to predict and identify molecules that can be derived by the intestinal microbiota from different source metabolites. One such source metabolite could be the amino acid phenylalanine that is reduced in the serum of germ-free mice, and therefore, a source of putative bioactive molecules. The

methodology described here can also be applied to identifying bioactive derivatives of environmental contaminants such as bisphenol A that can be generated *in vivo* by the GI tract microbiota and either have increased activity or have potentially different spectrum of activity (170).

CHAPTER IV

ANALYSIS OF TRANSCRIPTION FACTOR NETWORK UNDERLYING 3T3-L1 ADIPOCYTE DIFFERENTIATION

4.1 INTRODUCTION

Obesity is characterized by an increase in body fat mass, specifically in that of white adipose tissue (WAT) (14). Expansion of the body fat mass can occur either due to an increase in the number of adipocytes and/or an increase in size of adipocytes. The increase of adipocyte cell size is an important component of the increase in adipose tissue mass (15). Adipocyte enlargement arises as a result of genetic or environmental conditions that progressively increase lipid loading and therefore, cell size. The lipid content in adipocytes at any point is the result of a balance between formation and breakdown of lipid droplets in adipocytes. This lipid balance is extremely important as both high and low lipid levels have deleterious consequences. Several *in vitro* and *in vivo* studies (171-173) have shown that increased lipid loading and enlarged adipocytes correlate strongly with complications of obesity (e.g., recruitment of macrophages and inflammation, insulin resistance). Similarly, inadequate lipid levels in adipocytes are also not desirable as it leads to metabolic disorders and imbalance in supply of energy substrates to other organs such as the liver as well as insulin resistance (174). Therefore, control and regulation of the lipid balance in adipocytes (e.g., the ability to control lipid accumulation without significantly altering the extent of adipogenesis) is extremely important.

The enzymes mediating adipocyte lipid accumulation and utilization are regulated through the temporally coordinated action of several transcription factors (175, 176). *In vitro* studies have shown that during the clonal expansion of growth-arrested preadipocytes the expression of CCAAT/enhancer binding protein β (C/EBP β) and δ (C/EBP δ) is increased (87). Subsequently, expression of peroxisome proliferator activated receptor- γ (PPAR γ) and C/EBP α are stimulated (177). PPAR γ is a master switch for adipocyte differentiation (176, 178), and along with sterol regulatory element binding protein-1c (SREBP-1c), directly controls the expression of several genes in lipogenesis. C/EBP α and C/EBP β also sequentially regulate the expression of diacylglycerol acyltransferase 2 (DGAT2), which catalyzes the final step in triacylglycerol (TG) synthesis and lipogenesis, while PPAR γ also regulates the expression of lipases (adipose triacylglycerol lipase, ATGL) that functions in lipid breakdown. In addition, SREBP-1c promotes adipogenesis. Although SREBP-1c cannot initiate adipogenesis by itself (179), it can activate PPAR γ expression (180) which leads to an increase in adipogenesis. Cyclic AMP response element binding protein (CREB) is another early-transcriptional regulator as compounds that upregulate CREB activity (e.g., insulin, dexamethasone) also induce differentiation (175). Nuclear factor of activated T cells (NFAT) was demonstrated to form a composite enhancer complex with C/EBP and potentiate PPAR γ expression (181), while the forkhead transcription factor (FoxO1) has been shown to counter PPAR γ activation in 3T3-L1 adipocytes (182). While these studies have identified roles for different transcription factors, data from these studies provide limited information on both expression dynamics and interaction

between regulatory network molecules generated in these studies. In fact, simple inhibition of individual regulatory molecules often has adverse effects; for example, the inhibition of PPAR γ *in vivo* reduces adipogenesis, but also increases insulin resistance, one of the chief complications of type-2 diabetes mellitus (183). Therefore, it is important to consider the dynamics and interaction between different members of the regulatory network for modulating complex phenotypes such as lipid balance in adipocytes.

In this study, we developed an integrated network model of adipocyte transcription factors based on published connections between the different molecules, and validated it using measurements of activation dynamics during adipocyte differentiation and lipid loading. The model was verified by perturbing the activation levels of CREB and determining its effects on the other transcription factors. The transcription factor network is expected to provide a framework for manipulating adipocyte differentiation and lipid accumulation during obesity.

4.2 MATERIALS AND METHODS

4.2.1 MATERIALS

3T3-L1 cells were kindly provided by Prof. Barbara Corkey (Boston University School of Medicine, MA). Tissue culture reagents including Dulbecco's Modified Eagle's Medium (DMEM), calf serum (CS), fetal bovine serum (FBS), human insulin,

and penicillin/streptomycin were purchased from Invitrogen (Carlsbad, CA). Unless otherwise noted, all other chemicals were purchased from Sigma (St. Louis, MO).

4.2.2 CELL CULTURE AND DIFFERENTIATION

3T3-L1 cells were seeded in 6-well tissue culture plates in preadipocyte growth medium consisting of DMEM supplemented with CS (10 % v/v), penicillin (200 U/ml) and streptomycin (200 µg/ml). Medium was replenished every other day. Two days post-confluence, differentiation was induced using a standard adipogenic cocktail (1 µg/ml insulin, 0.5 mM isobutylmethylxanthine, and 1 µM dexamethasone) added to a basal medium (DMEM with 10 % FBS and penicillin/streptomycin). After 48 h, the first differentiation medium was replaced with a second differentiation medium consisting of the basal adipocyte medium supplemented with only 1 µg/ml of insulin. After another 48 h, the second medium was replaced with the adipocyte basal medium and replenished every other day.

4.2.3 CONSTRUCTION OF GAUSSIA LUCIFERASE REPORTER PLASMIDS

Lentiviral reporter plasmids for monitoring activation of six transcription factors (TFs) - PPAR γ , SREBP-1c, NFAT, CREB, C/EBP β and FoxO1 - during adipocyte differentiation and lipid loading were constructed as described below. For each TF, consensus binding sites (response elements; RE) in target gene promoter regions were identified using TRANSFAC database 7.0 Public (AGGACAAAGGTCA for PPAR γ , CATGTG for SREBP-1c, GGAAAATTTGAGTCA for NFAT, TGACGTCA for CREB,

ATTGCGCAAT for C/EBP β and AGTTGGACGCGAC for FoxO1). Response element (RE) oligonucleotides containing the binding sequence for each TF were chemically synthesized. Each RE oligonucleotide consists of three consensus binding sequences separated by 4 - 6 bases (spacer sequence) and a unique restriction enzyme (*EcoRI* and *AfeI*) cleavage sites at the ends. The RE oligonucleotides were cloned into Gluc-DRE2-viral vector (pCS-sMAR8-pA1-DRE2-hPGK-cHS4-tACTB-SPA-Gluc-CMVmin) (184) in which expression of the *Gaussia* luciferase (Gluc) is under the control of a minimal promoter. Expression of Gluc is induced only when a TF binds to its consensus binding site. Clones containing the correct RE were identified by multiple restriction enzyme digests and sequenced to verify fidelity.

4.2.4 GENERATION OF STABLE REPORTER CELL LINES

Stable reporter cell lines for each TF were generated by lentiviral transduction. To produce lentiviral particles, each TF reporter plasmid and two helper plasmids (psPAX; Addgene plasmid 12260 and pMD2.G; Addgene plasmid 12259, Dr. Trono, Lausanne, Switzerland) were co-transfected into 293T/17 cells using the calcium phosphate transfection method (185). After 24 h following the transfection, the medium was replenished and 5 mM of sodium butyrate was added, and incubated for an additional 24 h. Supernatants containing viral particles were collected, pooled, filtered with 0.45 μ m filters, and centrifuged for 2 h at 4°C at 48000 xg. The viral titer was measured using a Lenti-X qRT-PCR titration kit (Clontech, Palo Alto, CA). To transduce 3T3-L1 preadipocytes, concentrated virus particles ($\sim 2 \times 10^8$ IFU) were added

to the cells in presence of Polybrene (hexadimethrine bromide). The cells were incubated with the virus particles for 15 h, and the medium was replenished the next day. The efficiency of transduction was assessed by microscopic analysis.

4.2.5 VALIDATION OF REPORTER PLASMID FUNCTION

Plasmids for constitutive expression of the TFs were purchased from Addgene (pCMV5-FLAG-FoxO1, pSV Sport SREBP-1c, pEGFP-C1 NFAT3, pcDNA flag PPAR gamma) or Invitrogen (pCMV-Sport6-CREB). 293T/17 cells were seeded in 6-well tissue culture plates, and ~1 µg of each expression plasmid was cotransfected along with the corresponding reporter plasmid (~ 1 µg) using the calcium phosphate transfection. For control experiments, the same amount of pEYFP-N1 plasmid (constitutive expression of yellow fluorescent protein) was transfected. At 48 h post-transfection, supernatants were collected and luciferase activity measured using the BioLux *Gaussia* Luciferase Flex assay kit (New England Biolabs, Ipswich, MA). Additionally, the PPAR γ reporter plasmid was validated using thiazolidinedione (TZD) as the agonist. PPAR γ was overexpressed from plasmid (pcDNA flag PPAR gamma) in 3T3-L1 PPAR γ reporter cells and 25µM TZD was used to activate PPAR γ for 24 h. The C/EBP β reporter cell line was validated by up-regulating C/EBP β with 100 ng/mL oncostatin M (OSM) for 12 h (186). Luciferase activity in the supernatant was determined as described above.

4.2.6 MEASUREMENT OF TRANSCRIPTION FACTOR ACTIVITY PROFILES

For profiling TF activation in each reporter cell line, 3T3-L1 preadipocyte reporter cells for each TF were seeded in 6-well tissue culture plates and differentiated into adipocytes as described above. At different stages post-differentiation, 30 μ L supernatant samples were collected at 24 h post-medium change from day 0 (induction of differentiation) until day 17. Samples were stored at -20°C prior to assessing luciferase activity. The luciferase activity (Relative Light Units; RLU) measured was used to calculate the rate of Gluc production (RLU divided by the time over which Gluc was secreted). To account for differences in cell density between different experiments, the RLU rate was normalized with the red fluorescence intensity (Relative fluorescence units; RFU) measured at 550 nm (excitation) and 600 nm (emission). The fold-increase in TF activity was determined by normalizing the (RLU/h)/RFU at each time point to the corresponding value at the start of differentiation (i.e., day 0).

4.2.7 PERTURBATION OF THE TRANSCRIPTION FACTOR NETWORK BY FORSKOLIN

CREB, C/EBP β and PPAR γ reporter cell lines were seeded in 6-well tissue culture plates and differentiated into adipocytes as described above. Cells were treated with 10 μ M of forskolin or 0.1% DMSO at day 0 for 48 h. TF-driven luciferase activity in the supernatants was determined and the TF activation profiles were determined as described above.

4.2.8 TRANSCRIPTION FACTOR NETWORK

An interaction network model of the TFs analyzed in this study was assembled from the published literature (87, 90-93, 104, 180, 182, 187). A schematic of the model is shown in Fig. 4.1. The model included only four of the six measured TFs, i.e. CREB, C/EBP, PPAR γ , and SREBP-1c, as interactions involving the other two TFs, NFAT and FoxO1 were not consistently documented in the literature. The guiding principles in assembling the model were simplicity and identifiability. While it is likely that some of the TFs interact with each other through intermediaries and co-activators, these molecules were not experimentally monitored in this study, and thus it was not attempted to model these details. Including additional details could have improved the fit of the model to the data by increasing the degree of freedom, but the “improved” fit would have occurred at the expense of diminished identifiability. For these reasons, the activation of a TF by another factor was assumed to occur directly, and modeled using simple mass action kinetics. The exception was the activation of PPAR γ by SREBP-1c, which was assumed to involve the induction of a metabolic enzyme leading to the production of an activating ligand. The identity of this ligand remains to be definitively established, and a number of candidates have been proposed such as 13-hydroxyoctadecadienoic acid, 15-hydroxyeicosatetraenoic acid and 1-O-hexadecyl-2-Azelaoyl-sn-glycero-3-phosphocholine (188, 189); however, there is reasonable consensus that SREBP-1c activation of PPAR γ involves the endogenous production of a metabolite ligand. Including the putative PPAR γ ligand, the TF network model

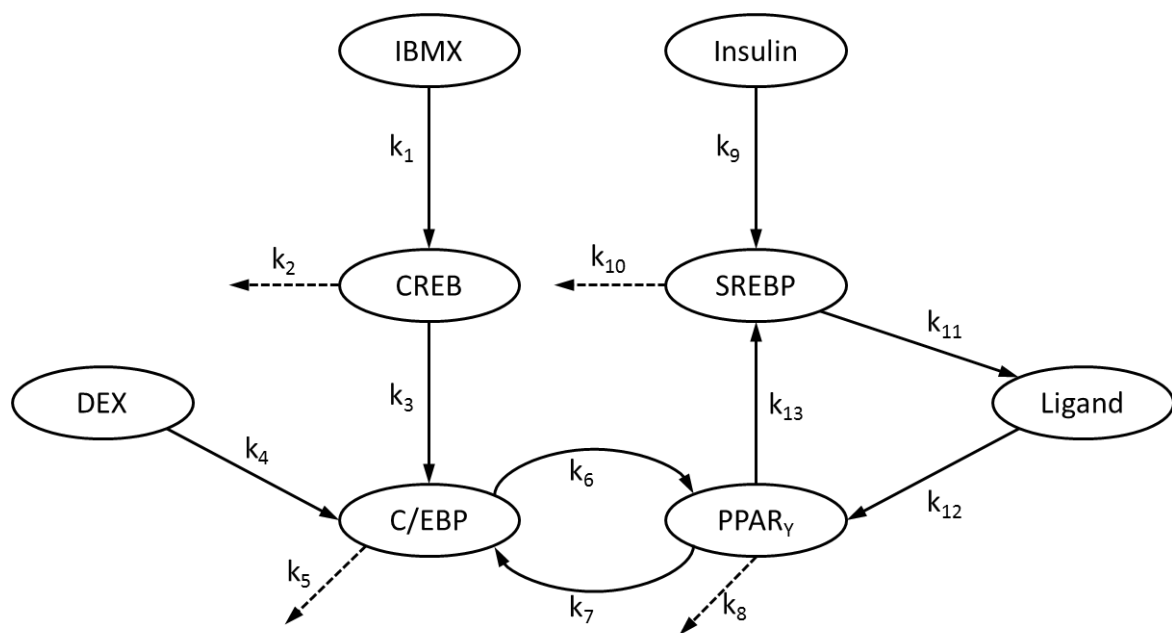


Fig. 4.1. Schematic of TF network model. Arrows indicate direction of interaction. Model parameters (k_2 , k_5 , k_8 , and k_{10}) labeling dotted arrows represent first-order decay rate constants for the TFs. See text for abbreviations.

comprises 5 species whose levels varied continuously with time. In addition, the model included three input species representing the differentiation inducing hormones insulin, dexamethasone (DEX), and isobutylmethylxanthine (IBMX). The levels of these input species were set to one (arbitrary units) or zero to reflect the addition and withdrawal of the hormones at different stages of a differentiation experiment. For the initial induction period from time zero to 48 h, the levels of all three input species were set to one. For the second induction period from 48 h to 96 h, DEX and IBMX were set to zero, while insulin was kept at one. For the maintenance period from 96 h to end of the experiment, all input species were set to zero.

The rates of change of the 5 time variant species were described using ordinary differential equations (ODEs). Partial differentiation equations or stochastic equations were not considered, because the data collected reflected the pooled averages of many ($> 10^6$) cells in a well-mixed environment. The general structure of the ODEs describing the TF dynamics was as follows.

$$\frac{dy_i}{dt} = \sum_j k_{j>i} y_j - k_d y_i \quad (1)$$

In equation (1), y_i is the amount of active TF i , y_j is the amount of a TF j that activates i , and $k_{j>i}$ is the rate constant for first-order activation of i by j . Every TF was assumed to decay at a rate proportional to its amount with a first-order degradation rate constant k_d . The rate of production and degradation of the metabolite ligand was modeled similarly using first-order rate expressions. The full set of ODEs describing the dynamics of the TFs and ligand are listed in Table 4.1.

Table 4.1. Mass action kinetic model of TF interactions

TF (or Ligand)	Differential equation	k	min	max	median
CREB	$\frac{dCREB}{dt} = k_1 IBMX - k_2 CREB$	k ₁	1.04E+01	1.50E+01	1.26E+01
		k ₂	8.65E-01	1.18E+00	9.93E-01
C/EBP β	$\frac{dCEBP}{dt} = k_4 DEX + k_3 CREB + k_7 PPAR - k_5 CEBP$	k ₃	5.99E-07	2.25E+01	2.42E+00
		k ₄	2.17E-09	3.89E+00	1.65E+00
		k ₅	5.00E-01	8.79E+01	1.05E+01
		k ₇	9.12E-02	2.38E+01	2.96E+00
PPAR γ	$\frac{dPPAR}{dt} = k_6 CEBP + k_{12} Ligand - k_8 PPAR$	k ₆	3.43E-01	3.85E+01	1.50E+01
		k ₈	5.00E-01	5.35E+01	2.35E+01
		k ₁₂	2.27E-01	4.32E-01	2.82E-01
SREBP-1c	$\frac{dSREBP}{dt} = k_9 Insulin + k_{13} PPAR - k_{10} SREBP$	k ₉	2.72E-01	1.56E+00	4.22E-01
		k ₁₀	8.61E-02	1.40E+00	2.52E-01
		k ₁₃	1.29E-08	4.74E-01	5.62E-02
Ligand	$\frac{dLigand}{dt} = k_{11} SREBP - k_{12} Ligand$	k ₁₁	8.37E-01	1.00E+02	4.89E+01

4.2.9 PARAMETER ESTIMATION

The model parameters, i.e. rate constants, were estimated from the experimentally obtained TF time course data (control condition without forskolin treatment) using a nonlinear constrained optimization procedure. The objective function was to minimize the sum of squared residuals (SSR) measuring the discrepancy between the measured and calculated TF levels for all time points.

$$\min \sum_t \sum_i (y_{i,t}^{exp} - y_{i,t}^{model})^2 \quad (2)$$

In the above expression, $y_{i,t}$ refers to the i th TF activity level (in RLU/h/RFU) at time point t , with the superscript denoting experimentally measured or model calculated value. The optimization variables were the rate constants k_1 through k_{14} describing activation (production) or degradation of the 5 model species (Table 4.1). The experimental data were averages of two independent experiments, where each experiment included two biological replicates for each TF. Parameter optimization was iterated 50 times. Each time, randomly generated noise drawn from a normal distribution was added to the experimental data. A noise level of 5 % standard deviation was used based on average variances in the measured TF activities across replicate experiments. For each of the 50 iterations, the optimization was repeated 5 times using a randomly generated set of initial parameter values. The upper bounds were set to 100 to ensure that the first-order rate parameters are within one order of magnitude from the observed rates of increase or decrease. The lower bounds on the TF degradation rate constants (k_2 , k_5 , k_8 and k_{10}) were set to 0.5, to ensure that there is a finite degradation rate for each of the

factors. We chose a minimal rate constant value that is equivalent to a half-life of ca. 36 h, which is 5- to 10-fold longer than the reported half-lives of TF modeled in this study. The lower bounds on all other parameters were set to zero. The initial values of the TFs were set to one ('1') to reflect the normalization of the experimental data. The initial value of the metabolite ligand was set to zero, as it was assumed that the ligand is not present in preadipocytes prior to induction. The optimization was performed using the LSQNONLIN function (trust-region-reflective method) of the Optimization Toolbox in MATLAB (Natick, MA).

4.2.10 SIMULATION OF THE TRANSCRIPTION FACTOR NETWORK

The TF network described above was evaluated by comparing the simulated TF time profiles against experimental data from a perturbed condition, i.e. forskolin exposure, which was not used in estimating the parameters. Model simulations were performed using all 50 sets of parameters, leading to 50 different sets of TF time profiles. Addition of forskolin, which rapidly increases intracellular cAMP (190), was modeled as step increase in IBMX during the first induction period (time 0 to 48 h) from a value of 1 to 1.7. This step increase in IBMX was set based on the measured profile of CREB, which is the direct target of IBMX in the model. All other model parameters and initial values were the same as the parameter estimation problem.

4.3 RESULTS

4.3.1 GENERATION OF GAUSSIA LUCIFERASE REPORTER CELL LINES TRACKING TRANSCRIPTION FACTOR ACTIVITY

A panel of six transcription factors (TFs) (PPAR γ , SREBP-1c, NFAT, CREB, C/EBP β and FoxO1) was chosen to develop a regulatory network underlying adipocyte differentiation and lipid loading. Reporter plasmids were developed as described in Materials and Methods and validated by determining if activation of a transcription factor led to expression of the *Gaussia* luciferase gene from the reporter plasmid. For the transcription factors CREB, SREBP-1c, NFAT, and FoxO1, plasmids containing the full length gene for each TF were over-expressed in 293T/17 cells by transfection along with the corresponding reporter plasmid. Figure 4.2A shows 5-30 fold increase in luciferase activity when transcription factors were overexpressed for 48 h. The 3T3-L1 PPAR γ reporter cell line was validated by over-expressing PPAR γ and activating it with a chemical agonist, thiazolidinedione (TZD). Figure 4.2B shows that overexpression of PPAR γ increases PPAR γ -driven luciferase activity by 1.3-fold, and additional stimulation with 25 μ M TZD resulted in a 1.5-fold increase in luciferase activity. The C/EBP β reporter was validated by stimulating 3T3-L1 preadipocyte C/EBP β reporter cells with the cytokine oncostatin M (OSM), which resulted in a 1.7 fold increase in C/EBP β -driven luciferase activity (Fig. 4.2C). Together, these results confirmed the ability of the developed reporter cell lines to report activation of the different TFs.

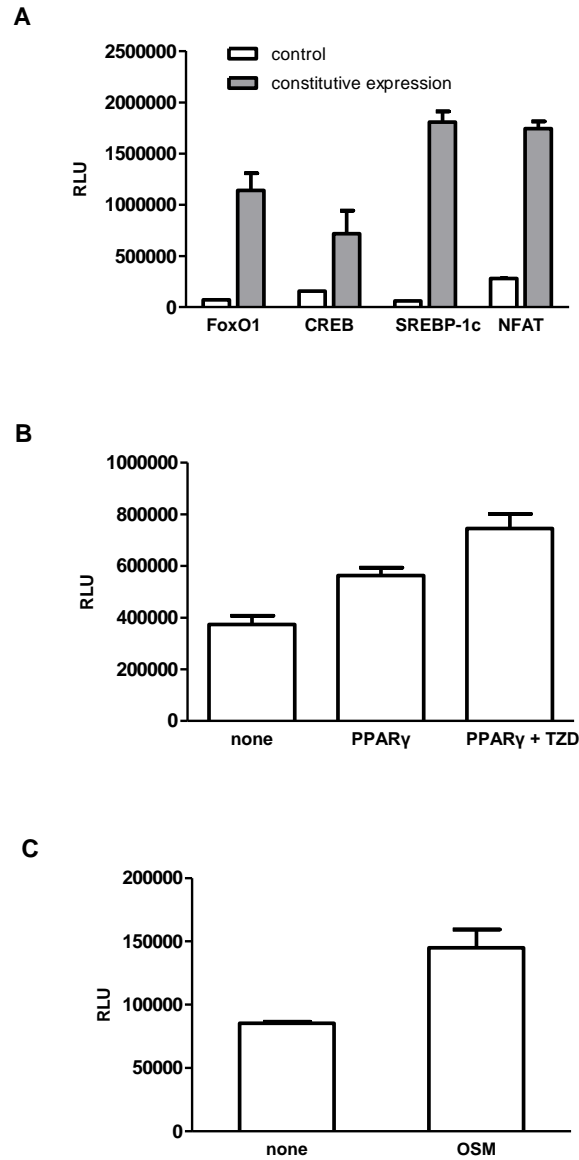


Fig. 4.2. The six TF reporter constructs properly present activation of each target TF. (A) When plasmids for constitutive expression of TFs were cotransfected with corresponding reporter plasmids into 293T/17 cells, Gluc activities were induced. (B) When a plasmid for constitutive expression of PPAR γ was transfected into 3T3-L1 PPAR γ reporter cells, there was 1.5 fold increase of Gluc activity. Treatment of a PPAR γ agonist (TZD) contributed to 1.3 fold increase. (C) Treatment of a C/EBP β agonist (OSM) on 3T3-L1 C/EBP β reporter cells caused 1.7 fold increase. Data represent mean \pm SD.

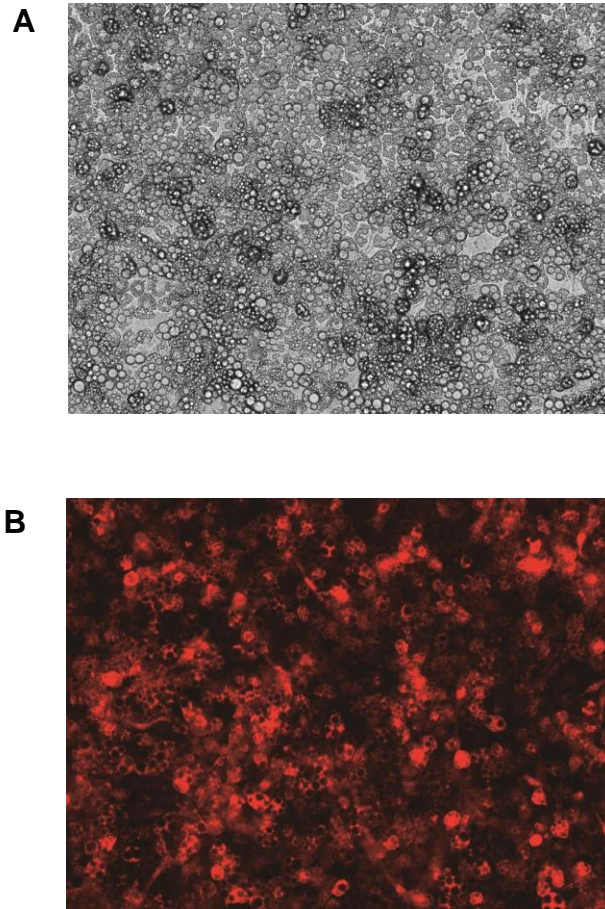


Fig. 4.3. 3T3-L1 PPAR γ reporter cells after 19 days of induction for differentiation (A) transmitted light (B) red fluorescence microscopy images.

Lentivirus-mediated integration of the reporter plasmids into 3T3-L1 preadipocytes did not affect the ability of the different reporter cells to differentiate into adipocytes. Figure 4.3 shows representative transmitted light and red fluorescence images of the PPAR γ reporter cell line after 19 days in culture, and clearly demonstrate that adipocyte differentiation was not affected (i.e., reporter cells show comparable accumulation of lipid droplets and cell morphology to 3T3-L1 adipocytes without the reporter plasmid). Moreover, no differences in differentiation and lipid loading were observed between the different reporter cell lines (Fig. 4.4). In addition, more than 90% of reporter cells demonstrated RFP expression, and suggest proper integration of the reporter DNA.

4.3.2 PROFILES OF TRANSCRIPTION FACTOR ACTIVITY DURING ADIPOCYTE DIFFERENTIATION AND ENLARGEMENT

3T3-L1 preadipocyte reporter cells were differentiated into adipocytes and the activity of each TF was monitored throughout the differentiation and maturation process. The different TFs demonstrated markedly different activation dynamics. As shown in Fig. 4.5, CREB reporter cells demonstrated a pronounced 13-fold increase in normalized luciferase activity (RLU/h/RFU) on day1 after addition of differentiation media. However, this increase was transient as the normalized luciferase activity rapidly decreased and returned to day 0 levels by day 7. On the other hand, PPAR γ reporter cells demonstrated a 3-fold increase in luciferase activity in the first 24 h after induction of differentiation. However, unlike CREB, this increase was sustained, as nearly 2-fold

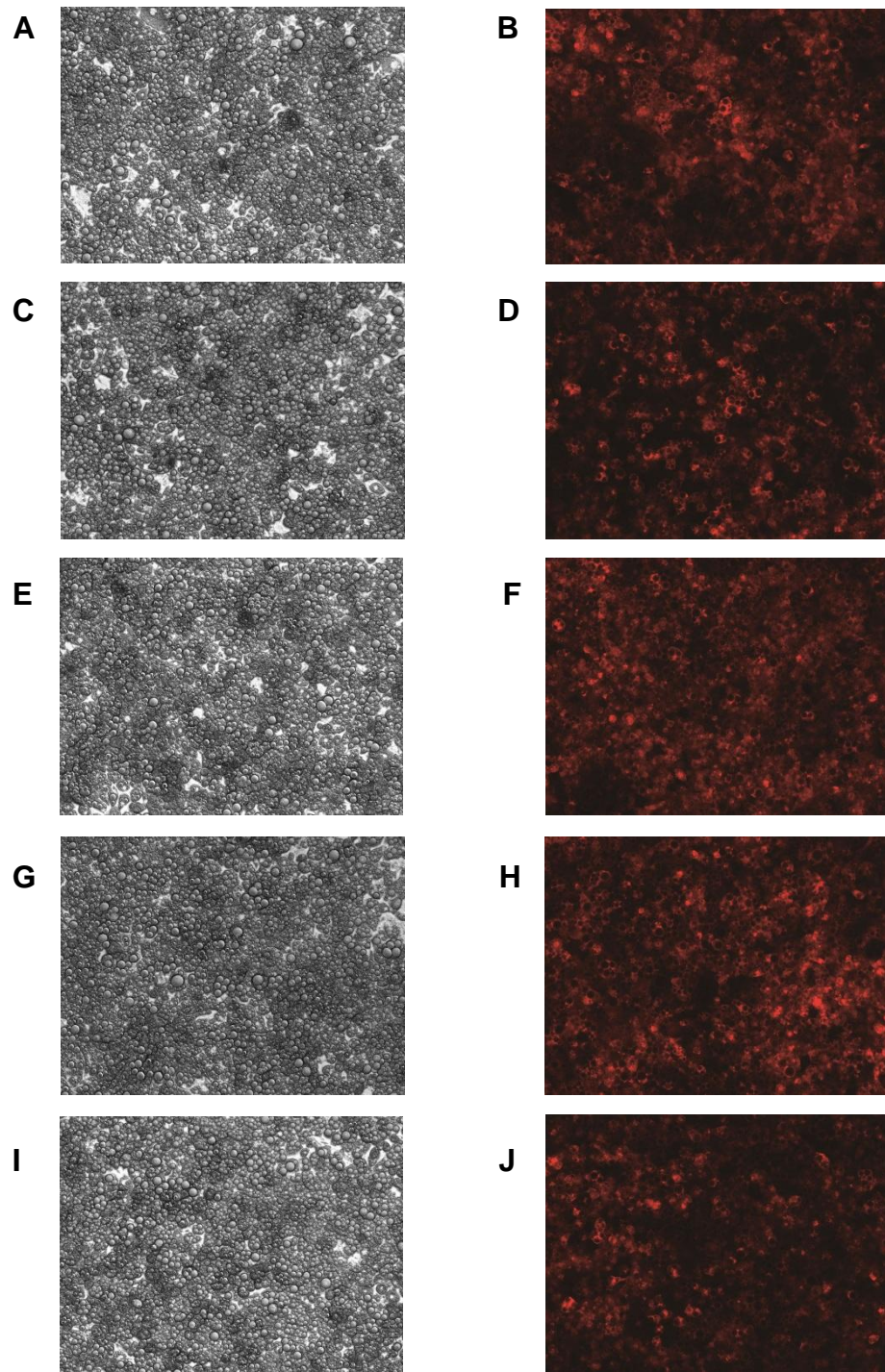


Fig. 4.4. 3T3-L1 TF reporter cells after 19 days of induction for differentiation. Transmitted light images of (A) FoxO1, (C) CREB, (E) NFAT, (G) SREBP-1c and (I) C/EBP β and red fluorescence images of (B) FoxO1, (D) CREB, (F) NFAT, (H) SREBP-1c and (J) C/EBP β .

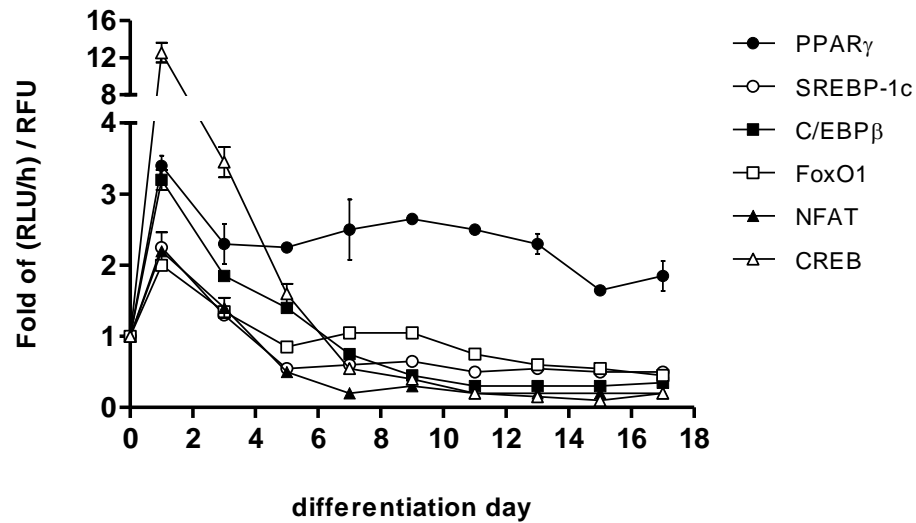


Fig. 4.5. TF activity profiles represented as fold of increase rates of *Gaussia* luciferase (Gluc) activities normalized with relative fluorescence units (RFU). The Gluc activities which reflect a degree of each TF binding onto response elements were measured every 24 h post-medium change from a time point of differentiation induction (day 0) to day 17. The increase rate of Gluc activity (RLU/h) was normalized with relative fluorescence units (RFU). The fold of the increase rate of Gluc activity was determined by normalizing the RLU/h/RFU at each time point to the corresponding value at the start of differentiation (i.e., day 0). Data represent mean \pm SD.

higher activation of PPAR γ was observed until day 17 post-differentiation compared to that in preadipocytes. This trend was especially noticeable beyond day 7 post-differentiation. C/EBP β , SREBP-1c, NFAT and FoxO1 all demonstrated similar activation dynamics with maximum activation observed at day 1 until day 3 and gradually decreasing to near-baseline values after day 5.

4.3.3 SIMULATION OF TRANSCRIPTION FACTOR ACTIVITY PROFILES

Based on literature reports, a simple network model was assembled describing the interactions between TFs as well as their stimulation by adipogenic hormones (Fig. 4.1). Only a subset of the TFs (CREB, C/EBP β , PPAR γ , and SREBP-1c) was considered, as the interactions involving NFAT and FoxO1 were less clearly documented. Activation and decay of TFs were modeled using mass action kinetics (Table 4.1). The model's rate parameters were fitted to the normalized activity profiles shown in Fig. 4.5 by minimizing the sum of squared residuals (SSR) between the measured and model-calculated TF activity profiles.

Across 50 iterations of parameter optimization, the SSR varied from 4.82 to 7.22, with a mean value of 5.99. The mean values of the optimized parameters ranged from 10^0 to 10^1 , in terms of order of magnitude, except for the rate constant describing activation of SREBP-1c by PPAR γ (k_{13}). The estimated value of this rate constant ranged from 0 to 0.47 across the iterations, with a mean value of 0.06 and standard deviation of 0.12, indicating that this parameter is not statistically significantly greater than 0 (one-sided t -test, $p > 0.05$). A majority of the parameters were reasonably well

constrained, as assessed by the ranges of the values obtained across the iterations of parameter optimization (Table 4.1). The parameters with the greatest uncertainty were associated with the degradation rate of C/EBP β (k_5) and the production rate of the putative ligand for PPAR γ (k_{11}), as their estimated values ranged from 0.5 to 87.9 and 0.8 to 100, respectively. Given the uncertainties in the estimated parameters, model simulations were performed using all of the 50 sets of optimized parameters to obtain 50 sets of profiles from which means and standard deviations were calculated (Fig. 4.6).

Overall, the simulated profiles were largely consistent with the measured profiles, and correctly ordered the peak activities for the TFs. Importantly, the simulations showed a sustained elevation in PPAR γ activity that extends to day 11, in good agreement with the measured profile. On the other hand, there were noticeable differences between the shapes of the measured and simulated profiles. The rise and fall of the simulated profiles were generally somewhat sharper, presumably due to the simplifying assumptions regarding the concentrations of the hormones, which were modeled as step inputs. Specifically, the simulated profile for C/EBP β showed an initially sharp rise until day 2, followed by a rapid decline, whereas the measured profile showed a broader peak extending from day 1 to day 5. In addition to the assumption regarding step change in the hormonal input levels, it is also possible that the discrepancy could be due to the fact that the C/EBP β reporter cannot discriminate between other isoforms of C/EBP.

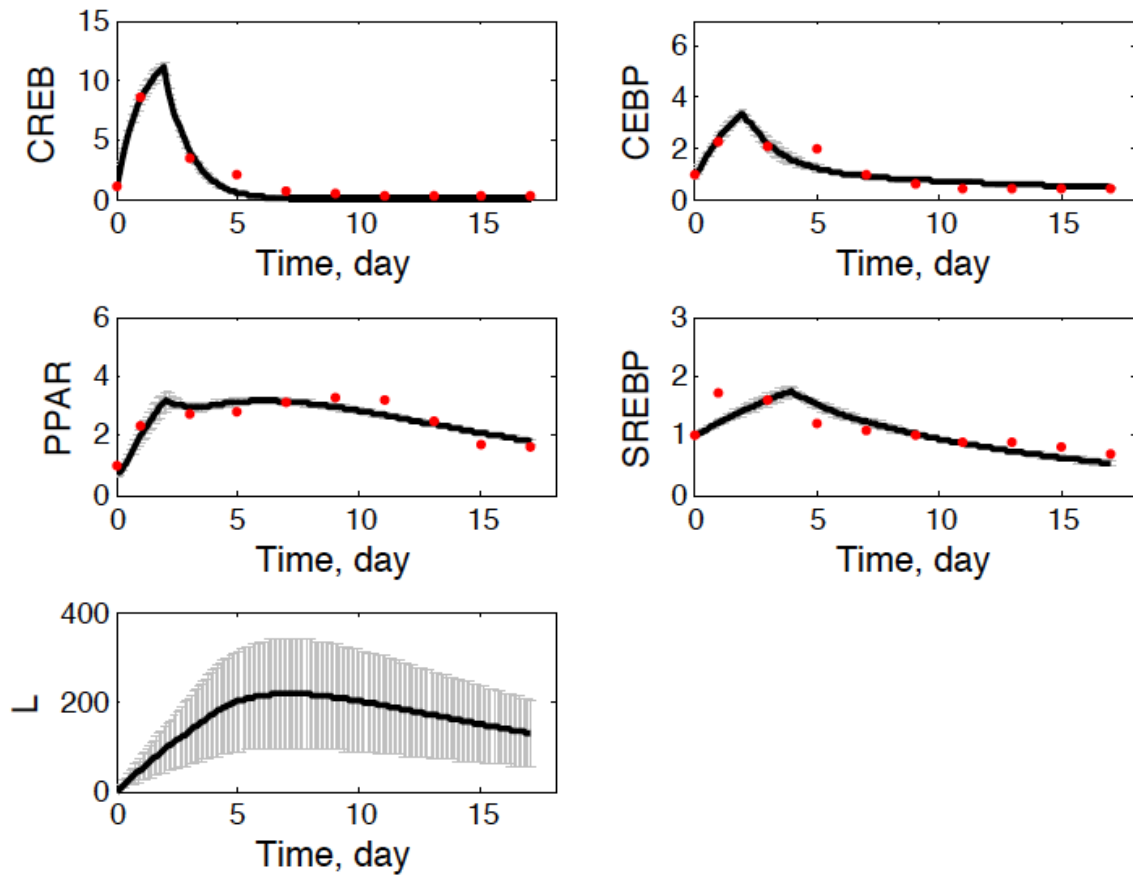


Fig. 4.6. Simulated profiles of TF dynamics. Simulations were performed with each of the 50 sets of parameters estimated from the measurement profiles. The average profiles are shown in bold solid lines, with error bars representing one standard deviation across the 50 simulations. Red circles indicate measured TF activities.

4.3.4 PERTURBATION OF THE TRANSCRIPTION FACTOR NETWORK BY FORSKOLIN TREATMENTS

In order to evaluate the TF network model, we perturbed the activity of CREB using forskolin, and investigated its effects on the activation of C/EBP β and PPAR γ . In the simulations, this perturbation was modeled as an increase in the concentration of IBMX relative to the unperturbed condition. As shown in Fig. 4.7, forskolin treatment on day 0 resulted in 1.7-fold increase in the maximum activation of CREB as compared to the control, while retaining the same activation dynamics. The maximum activation of C/EBP β and PPAR γ increased 1.3-fold and 1.5-fold, respectively, after forskolin treatment. These results are largely consistent with the model simulations, which also showed a positive, but attenuated fold-change in the C/EBP β and PPAR γ activities relative to the unperturbed condition (Fig. 4.8). When the average maximum activity of CREB (on day 1) was increased to 1.8-fold in the simulation (relative to the unperturbed simulation) by adjusting the initial IBMX input level, the corresponding increases in the C/EBP β and PPAR γ maximum activities (based on the averages of the simulated profiles) were 1.3-fold and 1.2-fold, respectively.

4.4 DISCUSSION

In this study, we developed an integrated transcription factor (TF) network model underlying adipocyte differentiation and enlargement and validated it using dynamic activation profiles of different TFs during differentiation of preadipocytes into adipocytes and lipid accumulation. Dynamic activation profiles of the different TFs were

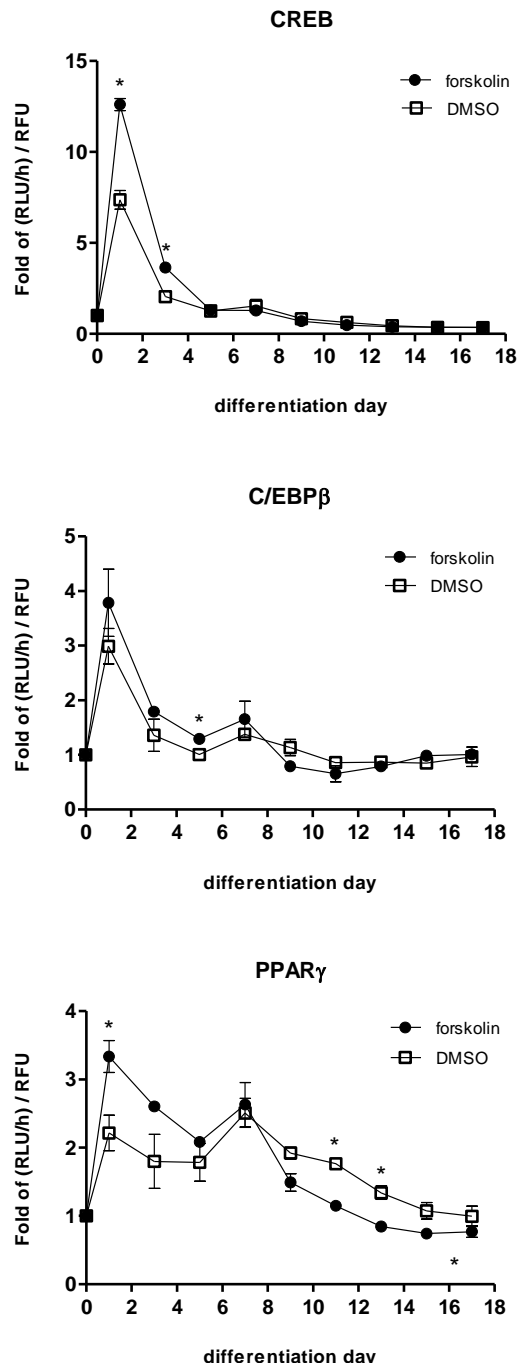


Fig. 4.7. Perturbation of CREB activity profiles by forskolin treatments. Cells were treated with 10 μ M of forskolin or 0.1% DMSO at day0 for 48 h. For each time point, fold of RLU/h/RFU was calculated by dividing with a value of day 0. Data represent mean \pm SD. *: p < 0.05

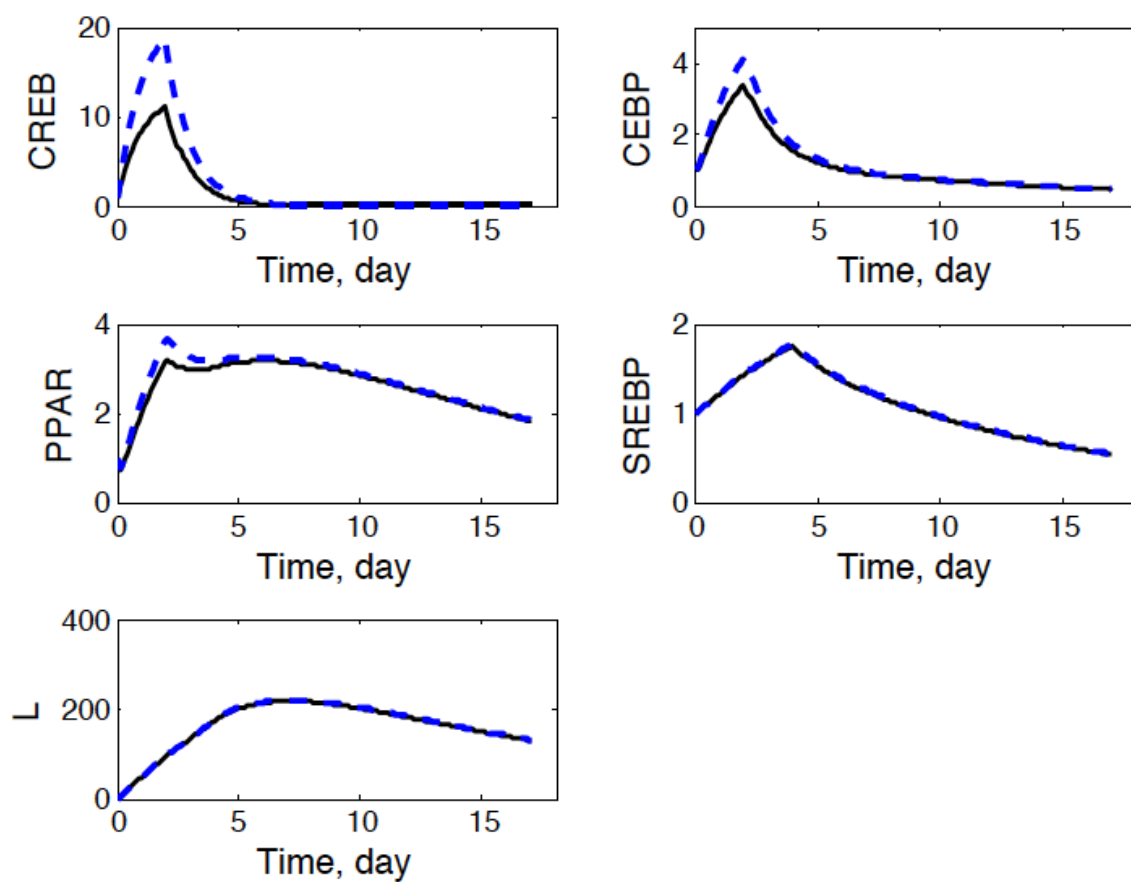


Fig. 4.8. Simulations of TF activity profiles representing the forskolin treatment (dashed blue line) and vehicle control (solid black line) conditions.

obtained using 3T3-L1 reporter cell lines where binding of a transcription factor to its recognition sequence leads to expression of a *Gaussia* luciferase reporter gene. Since the luciferase gene has a secretion leader sequence, non-invasive dynamic profiling of TF activity is possible by monitoring the culture supernatant for luciferase activity. This approach enabled us to directly monitor TF function (i.e., binding to its recognition sequence) as opposed to inferring activity from mRNA and protein data. The non-invasive measurement technique also allows for frequent sampling, which, in turn, facilitates data-based model construction. Lastly, this approach also allows for TF activation profiles to be derived from the same population of cells over the differentiation and maturation.

The differentiation of 3T3-L1 preadipocytes into mature adipocytes is characterized by accumulation of lipid droplets and alteration of cell shape which are mediated by multiple enzymes and regulatory TFs. While previous data on the activation of different TFs underlying adipocyte differentiation and enlargement are available in the literature, differences in the experimental conditions and readout methods limit their utility for developing an integrated network model. In this study, we address this by developing a TF model using time-course data from the same cell population.

Our results show that different TFs demonstrate distinct activation profiles during adipocyte differentiation and enlargement, which are consistent with published reports. The temporal correlation between CREB and C/EBP β is one such example (Fig. 4.5) where activation of C/EBP β follows the activation of CREB until day5. Zhang and coworkers (104) have demonstrated that CREB induces expression of C/EBP β at 8 h

after induction of the adipogenic program. Similarly, SREBP-1c has been shown to promote adipogenesis by generation of a PPAR γ ligand (179), and it has been proposed that SREBP-1c could activate expression of PPAR γ by binding to a PPAR γ promoter (180). NFAT has also been proposed as a regulator of PPAR γ activity since NFAT associates physically with PPAR γ (187). Our data show that both SREBP-1c and NFAT show comparable activation profiles as C/EBP β , which is known to precede and activate expression of PPAR γ (191), during the initial differentiation phase(days 1-3). Thus, our data provide corroborating evidence for SREBP-1c and NFAT in regulating PPAR γ activity. PPAR γ is a key TF which plays critical roles for induction and maintenance of fully differentiated adipocytes. Our data also show that PPAR γ activity increased gradually from the induction of differentiation and stayed elevated between days 7 and 11. While the activation levels decreased after day 11, they were still significantly higher than baseline (preadipocyte) levels for the duration of the experiment. Interestingly, only PPAR γ demonstrated sustained activation as the activity of other TFs decreased to pre-induction levels after adipocyte maturation (days 7-10). This result is consistent with the primary role for PPAR γ in adipocyte differentiation and enlargement. Interestingly, FoxO1 did not show significant activation over the time course of the experiment. A recent study has shown that a direct protein-protein interaction leads to repression of PPAR γ activity by FoxO1 (192). However, since FoxO1 interferes with PPAR γ by recruitment to the PPAR γ response element, measurement of FoxO1 binding to its own response element may not show any correlation to alterations in PPAR γ activity.

To more quantitatively characterize the interactions between TFs regulating the adipocyte differentiation program, we assembled a mass action model of TF activation and decay based on connections documented in the published literature. For four of the TFs monitored in this study, namely CREB, C/EBP β , PPAR γ , and SREBP-1c, there was a general consensus regarding the activation cascade. For the sake of parsimony, we modeled the cascade as a sequence of direct activation events between TFs. The sole exception was the activation of PPAR γ by SREBP-1c, as there was strong evidence that SREBP-1c activation induces the production of a metabolite ligand for PPAR γ (103). As published data on adipocyte TF network activation rates were not available, the rate parameters were estimated from the measured profiles collected in this study. Outside of the general assumption that the rate constants have to be nonnegative, the only other constraint applied to the parameters was to set a lower bound for the turnover rates of the TFs. The estimated decay rate constants were significantly higher than the lower bound, ranging from $\sim 1 \text{ day}^{-1}$ to $\sim 23 \text{ day}^{-1}$. These decay rate constants correspond to half-life ranges of $\sim 17 \text{ h}$ to $\sim 40 \text{ min}$, which are comparable to published turnover rates of TFs (193-198). As alluded to earlier, there is substantial uncertainty regarding the estimated decay rate constant for C/EBP β . One possible explanation is that the reporter construct measured not only C/EBP β , but also the other isoforms of the C/EBP family. In particular, the measured profile could reflect the activation of C/EBP α , which follows the earlier activation of the β and δ isoforms during the initial stages of adipogenesis. Conflating the α , β and δ isoforms could lead to an apparent broadening of the peak TF activity period, as the decline in the β and δ isoforms could be masked by the rise in the

α isoform. It is thus possible that two distinct sets of activation and decay rates, and correspondingly discriminatory measurements, may be needed to more accurately model the dynamics of C/EBP.

The other large uncertainty in parameter estimation involved the production rate of the PPAR γ ligand upon induction of SREBP-1c. Similar to the C/EBP decay rate parameter, the simplest explanation is the lack of direct and specific experimental observations on the dynamics of the putative ligand. Moreover, it is likely that multiple ligands are produced in differentiating adipocytes (199), although a dominant endogenous ligand has not yet been conclusively identified. Another, related limitation regarding the interpretation of the simulated ligand time profile is that it is not possible to determine an absolute concentration range for the ligand, because normalized TF time profiles were used in fitting the model.

Despite these limitations, the mathematical model revealed several noteworthy features of the adipocyte TF network. First, model simulations of forskolin stimulation showed that an increase in CREB activity is attenuated as the stimulus propagates to C/EBP and PPAR γ , indicating correct ordering of upstream and downstream TFs. This prediction was in good agreement with measured data, which also showed an attenuation of TF stimulation downstream from CREB. Second, the estimated values for rate parameters suggest that feedback activation of SREBP-1c by PPAR γ is negligible in comparison to activation of SREBP-1c by insulin, corroborating the downstream placement of SREBP-1c relative to PPAR γ in recently published studies (200, 201). Third, the estimated parameters suggest that the strength of the positive activation loop between C/EBP and

PPAR γ is comparable to the strengths of the activation cascades connecting the TFs to hormonal inputs. This finding is consistent with previous experiments showing that direct (e.g. genetic) manipulation of either C/EBP or PPAR γ could drive the adipogenesis program to similar extents as hormonal stimulation (93, 202).

The integrated model of the TF network developed in this study provides a powerful tool for direct control of the entire regulatory network and further adipocyte metabolism. For example, we can predict overall change of activity profiles for key TFs triggered by a specific perturbation (i.e. agonist or antagonist) and its influence on expression levels of target genes which are regulated by the TFs. We expect the systematic understanding for the TF network will be a key basis to develop a new treatment against obesity. In addition, the combinatorial approach of experiment and simulation reported in this study will be applicable to understand complex networks in other cellular regulatory mechanisms.

CHAPTER V

MECHANISMS UNDERLYING RECOGNITION OF TRYPTOPHAN DERIVATIVES IN HOST CELLS

5.1 INTRODUCTION

The human gastrointestinal (GI) tract consists of $\sim 10^{14}$ bacteria belonging to $\sim 1,000$ species which are collectively termed gut microbiota. Recently, the microbiota metabolites have been considered as important modulators of host physiology instead of metabolic byproducts. For example, our laboratory demonstrated that tryptophan (TRP)-derived bacterial metabolite indole attenuates several inflammatory indicators in intestinal epithelial cells (IECs) (27). In addition, non-alcoholic fatty liver disease and cardiovascular diseases correlates with conversion of choline into trimethylamine by the gut microbiota (11, 12).

In our prior work (Chapter III), we predicted bacterial products of TRP metabolism by computational pathway analysis, and quantified physiological levels of the predicted metabolites in murine cecum and feces by utilizing multiple reaction monitoring (MRM) mass spectrometry. Further, the TRP derivatives were characterized as bioactive molecules by demonstrating agonist activity for aryl hydrocarbon receptor (AhR). However, the mechanisms underlying recognition of the TRP derivatives (i.e., the receptor(s) involved, the signaling pathways utilized) have not been elucidated yet. Identification of molecular signaling targets for the TRP derivatives is necessary to comprehensively understand the influence of the TRP derivatives on host cell function.

The AhR is a ligand-activated transcription factor that has been shown to play an important role in the intestinal immune system (70, 71). In the absence of ligand, AhR is localized in the cytoplasm. Upon ligand binding, AhR translocates into the nucleus where it heterodimerizes with the AhR nuclear translocator (Arnt) protein. The AhR-Arnt complex specifically interacts with the dioxin response element (DRE) in promoter regions of various target genes and modulates their transcription. In addition, the AhR modulates target gene expression through interaction with nuclear factor- κ B (NF- κ B) p65 (203). The AhR interacts with a wide spectrum of diverse natural or synthetic compounds such as halogenated aromatic hydrocarbons, heterocyclic amines, polycyclic aromatic hydrocarbons and several indole containing compounds (75-77). Although many dietary compounds were initially suggested as natural ligands for AhR, several reports have indicated that endogenous physiological ligands for AhR exist (78-80). For example, endogenous metabolites of tryptophan such as tryptamine and indole acetic acid were shown to bind to and activate AhR (83).

Several studies showed that synthetic 3,3'-diindolylmethane (DIM) derivatives (i.e., DIM-C-pPhCF₃, DIM-C-pPhOCH₃, DIM-C-Ph) activates NR4A1, a member of the nerve growth factor-induced B (NGFI-B) subfamily of orphan nuclear receptors (132, 147). The DIM is generated through heterodimerization of indole-3-carbinol in the acid environment of the gut (81) and exhibits anticancer activities through activation of the AhR (204). The NR4A subgroup consists of neuron-derived clone 77 (Nur77; NR4A1), nuclear receptor related 1 (Nurr1; NR4A2) and neuron derived orphan receptor 1 (NOR1; NR4A3). Recent functional studies have pointed a critical role of NR4A members in

control of inflammation. For example, NR4A2 represses transcription of TNF α and IL-1 β through recruitment of corepressor for element-1 silencing transcription factor (CoREST) complex to a target promoter and clearance of NF- κ B (144). To date, endogenous ligands for NR4A receptors have not been identified.

Dendritic cells (DCs) are specialized antigen presenting cells (APCs) that possess unique properties for adaptive immune responses. DCs migrate from peripheral tissues to lymphoid tissues in which they present antigens to T cells (205). DCs exist in functionally distinct immature and mature cells. The immature DCs in the peripheral tissues are mainly phagocytic cells while mature DCs in lymphoid tissues are characterized by high surface expression of co-stimulatory molecules for T cell activation (205). In the gut, DCs distribute in the lamina propria of the small and large intestine, the Peyer's patch and mesenteric lymph nodes (206, 207). Intestinal DCs have been shown to drive development of regulatory T cells (T_{reg}) which are essential to counteract inflammatory responses (208). In addition, intestinal DCs impart gut-homing properties to lymphocytes and induce B cells to acquire expression of IgA which supports several mucosal functions (206, 209). Although the exact role of DCs in intestinal disease is unknown, several observations in models of inflammatory diseases suggest that aberrant DCs drive dysregulated inflammation (210).

The goal of this study is to elucidate the molecular recognition mechanism(s) underlying sensing on TRP derivatives and in anti-inflammatory function. In previous studies (70, 71, 141-144), both AhR and NR4A receptors have shown to play important roles in modulation of inflammation. Moreover, endogenous metabolites of tryptophan

have shown to activate AhR, and synthetic derivatives of DIM retaining AhR activation have been demonstrated as NR4A agonists (83, 147). Therefore, we hypothesized that TRP derivatives are recognized by AhR and/or NR4A receptors and the sum of these interactions determines the extent of anti-inflammatory activity. Our data indicate that knockdown of NR4A2 attenuates, but not completely abolishes, the anti-inflammatory activity of indole to suppress expression of pro-inflammatory cytokine TNF α in dendritic cells. Our results suggest that NR4A2 is an anti-inflammatory signaling target of indole. In a future knockdown study of AhR, it will be clarified whether AhR also functions as a receptor for indole signaling.

5.2 MATERIALS AND METHODS

5.2.1 CELL CULTURE

The DC2.4 murine dendritic cell line was obtained from ATCC (Manassas, VA). DC2.4 cells were cultured at 37 °C with 5 % CO₂ in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5% (v/v) fetal bovine serum (FBS), glucose (3.5 g/L), sodium bicarbonate (2 g/L), nonessential amino acids (0.1 mM), HEPES (10 mM), penicillin (100 U/ml) and streptomycin (100 μ g/ml).

5.2.2 WESTERN BLOT ANALYSIS

Whole cell lysates were extracted using RIPA buffer. The lysates were heated at 100 °C for 5 min, separated on 4% to 20% SDS-PAGE (BIO-RAD, Hercules, CA) and

transferred to a nitrocellulose membrane (Invitrogen, Carlsbad, CA). The membrane was blocked with 5 % (v/v) nonfat dry milk in Tris-buffered saline overnight and incubated with a 1:500 dilution of anti-NR4A2 antibody (clone N-20) (Santa Cruz Biotechnology, Santa Cruz, CA) for 1 h at room temperature. After washing for 15 min, a 1:5000 dilution of a goat anti-rabbit IgG antibody (Santa Cruz Biotechnology, Santa Cruz, CA) was added, and the membrane was incubated for 45 min. After washing for 20 min, the membrane was incubated with a SuperSignal Kit (Pierce Chemical, Rockford, IL) for 5 min and imaged using Bio-rad Model 3000 VersaDoc Imaging System.

5.2.3 LENTIVIRUS PRODUCTION AND INFECTION OF DC2.4 CELLS

293T/17 cells were seeded at 8×10^5 cells per 6 well tissue culture plate in 2 ml DMEM supplemented with heat inactivated 10% (v/v) FBS. Next day, cells were transfected with pLKO.1-puro NR4A2 shRNA or non-target control plasmids (Sigma, St. Louis, MO) and two helper plasmids (psPAX; Addgene plasmid 12260 and pMD2.G; Addgene plasmid 12259, Dr. Trono, Lausanne, Switzerland) using a TransIT-293 Transfection Reagent (Mirus Bio, Madison, WI) following the protocol recommended by the manufacturer. After 6 h, the medium was removed and replaced with fresh 2 ml OptiMEM (Invitrogen, Carlsbad, CA) and 10 % (v/v) FBS. At 48 h post-transfection, supernatants containing viral particles were collected and filtered with a 0.45 μ m PES membrane (VWR, Radnor, PA). To transduce DC2.4 cells, the virus particles were 10-fold diluted in the culture medium and added to the cells in presence of Polybrene

(hexadimethrine bromide). After 4 h, the virus particles were removed and the cells were cultured in replenished medium until further analysis.

5.2.4 QUANTITATIVE RT-PCR

At 3 days and 4 days post-transduction, total RNA was isolated using a RNA Miniprep kit (Qiagen, Valencia, CA) according to instructions of the manufacturer. The primers for analysis of expression levels of NR4A2 gene were designed by using PrimerQuest online software (5'-GCACGTCAAAGAACTGGAAAG-3' and 5'-TGCCTGCAGGTTAGGAAATAG-3'). qRT-PCR was performed by using a SYBR Green iScript one-step RT-PCR kit in a MyiQ single-color real-time PCR detection system (BIO-RAD, Hercules, CA). The templates were analyzed in triplicates in 25 μ l reactions, and 100 ng of total RNA was used for each reaction with the final forward and reverse primer concentrations at 0.1 μ M each. The average threshold cycles (C_t) calculated by the MyiQ optical system software were normalized to average C_t values of 18S rRNA, and relative changes between the samples were determined.

5.2.5 FLOW CYTOMETRY AND INTRACELLULAR CYTOKINE STAINING

Changes in TNF α production in response to DMF (solvent control) or indole were determined by intracellular cytokine staining (ICS). DC2.4 cells infected with NR4A2 shRNA lentiviral particles were transferred into a 96 well U-bottom tissue culture plate at 2 days post- transduction, and exposed to indole or DMF overnight (~19 h). Next day, 5 μ g/ml lipopolysaccharide (LPS) (Invivo Gen, San Diego, CA) and 1,000-

fold diluted GolgiPlug (BD Biosciences, San Jose, CA) were added to the cells for 4 h. After two washes in PBSA (PBS supplemented with 0.5% BSA), cells were fixed with 4% paraformaldehyde on ice, permeabilized with Perm/Wash buffer (BD Biosciences, San Jose, CA) and stained with Alexa Fluor 700 TNF α antibody (BD Biosciences, San Jose, CA) for 1 h. The cells were resuspended in PBSA and stored in the dark at 4°C until analysis. Unstained or unstimulated cells were used as negative controls. Data were acquired on a BD FACSAria II cell sorter system (BD Biosciences, San Jose, CA) and analyzed with FlowJo software (Tree Star, Ashland, OR).

5.3 RESULTS

5.3.1 EXPRESSION OF NR4A RECEPTORS IN DC2.4 CELLS

In an effort to identify a NR4A member involved in recognition and inflammatory signaling of indole, we investigated expression levels of NR4A receptors in DC2.4 cells. Western blot analysis using an antibody that recognizes NR4A1 or NR4A2 receptor showed that expression of NR4A2 was much higher than NR4A1 (Fig. 5.1). Based on this result, NR4A2 was firstly targeted for a functional study of knockdown.

5.3.2 KNOCKDOWN OF NR4A2 BY SHRNA LENTIVIRAL TRANSDUCTION

After transient lentiviral transduction of NR4A2 shRNA, expression levels of NR4A2 mRNA were measured by qRT-PCR to determine a time point for the most

efficient knockdown. At 3 days post-transduction, the NR4A2 mRNA level was reduced to 47% of the control. However, the NR4A2 mRNA level increased back up to 88% of the control at 4 days post-transduction. Accordingly, all experiments were conducted at 3 days post-transduction to investigate the impact of NR4A2 knockdown on TNF α production.

5.3.3 A ROLE OF NR4A2 ON ANTI-INFLAMMATORY SIGNALING OF INDOLE IN DC2.4 CELLS

Previously, our lab reported that indole exerts anti-inflammatory effects on intestinal epithelial cells (IECs) (27). To test if indole exposure alters DC2.4 function, we exposed DC2.4 cells to 0.5 mM and 1 mM indole overnight (~ 19 h) before stimulation with LPS (1 μ g/ml) for 4 h followed by intracellular cytokine staining (ICS) and flow cytometry. Similar to the effects on IECs, the indole treatment attenuated the fraction of cells that were producing the pro-inflammatory cytokine TNF α (i.e., 13 % decrease with 0.5 mM and 29 % decrease with 1 mM indole) (Fig. 5.2A). Control DC2.4 cells were treated with DMF solvent control. These data demonstrate that indole attenuates indicators of inflammation in DC2.4 cells. To explore an involvement of the NR4A2 in anti-inflammatory signaling of indole, DC2.4 cells were transduced with NR4A2 shRNA lentiviral particles and its effect on TNF α production was investigated at 3-days post-transduction. As shown in Fig. 5.2B, ~ 53% knockdown of NR4A2 resulted in less suppression of TNF α production than the untransduced control (i.e., 3 % decrease

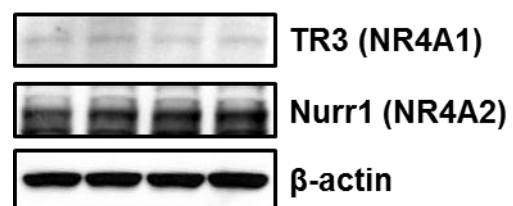


Fig. 5.1. Expression of NR4A receptors in DC2.4 cells. The western blot analysis was performed in quadruplicate.

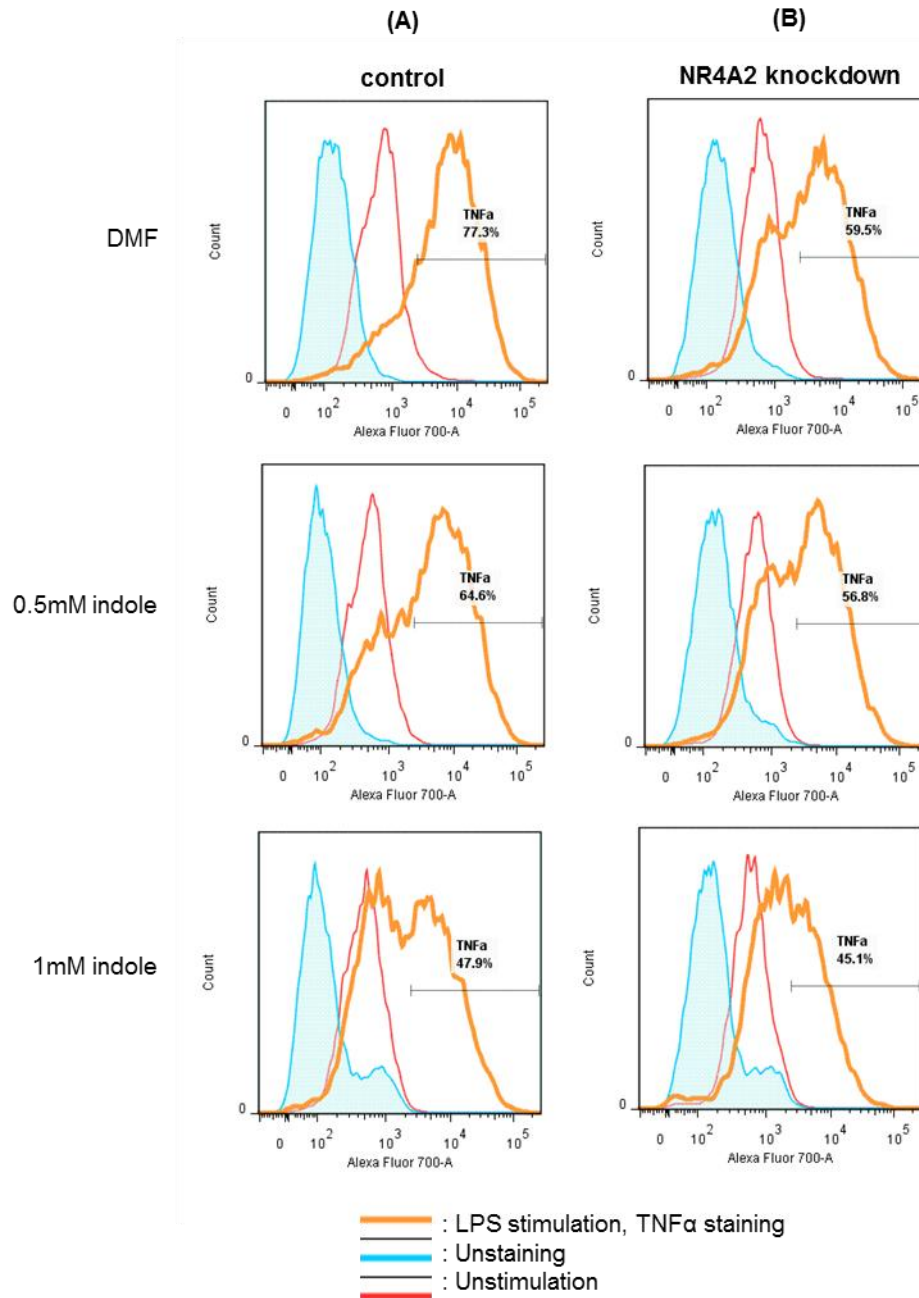


Fig. 5.2. The NR4A2 knockdown attenuates indole-derived suppression of TNFα production in DC2.4 cells. ICS and flow cytometry of TNFα were performed after exposure to indole overnight (~ 19 h) followed by stimulation with LPS (1 μg/ml) for 4 h. (A) Untransduced cells (B) NR4A2 shRNA lentivirus transduced cells.

with 0.5 mM and 14 % decrease at 1 mM indole). From these data, it can be concluded that the NR4A2 is important in mediating the anti-inflammatory effects of indole in DC2.4 cells.

5.4 DISCUSSION

Increasing evidence suggests that the metabolites produced by the gastrointestinal (GI) tract microbiota are important modulators of human health and disease (48, 211). For example, the bacterial metabolite indole has been recently shown to strengthen host cell-barrier properties and modulate inflammation by increasing anti-inflammatory IL-10 and decreasing pro-inflammatory IL-12 and IL-8 cytokines (27). Although several microbiota-derived metabolites have been identified and studied for their function, in-depth research is required to understand molecular mechanisms underlying sensing and signaling of these metabolites.

Dendritic cells (DCs) play an important role in initiation and maintenance of immune responses by helping naïve T cells to develop into effector T-helper (Th) cells. Depending on microenvironment and stimuli, DCs produce various cytokines (IL-12, IL-6, TNF α , IL-10, TGF β , etc.) that determine type and function of Th cells and thus DCs play a critical role in shaping overall immunity (212-214). In addition to dietary and microbial signals, intestinal DCs take microbiota metabolites from intestinal lumen and epithelial cell layer to orchestrate function of gut-associated lymphoid tissue (210, 215). Due to the diffusible nature of indole (64), we hypothesized that DCs in the Peyer patches are exposed to microbiota metabolites such as indole by M-cell transcytosis (206)

and these metabolites likely alter production of cytokines by DC's, thereby, leading to modulation of mucosal inflammation. Since indole was previously shown to alter production of several inflammatory cytokines in IECs (27), we first tested indole among several TRP derivatives to monitor a change of DCs' production of a pro-inflammatory cytokine (i.e., TNF α). As shown in Fig. 5.2A, the indole treatment resulted in a significant decrease of in cell population that were expressing TNF α . This result demonstrates that indole is recognized by DC2.4 cells and exerts the anti-inflammatory effect.

To elucidate the signaling pathway(s) through which the anti-inflammatory effects of indole are mediated in DC2.4 cells, we carried out knockdown of the NR4A2 nuclear receptor and investigated the resultant effect on DC 2.4 function. NR4A2 is a member of the NR4A subfamily of orphan nuclear receptors (132, 147). NR4A receptors act as transcription factors, regulating expression of downstream genes in apoptosis, proliferation, cell migration, angiogenesis and inflammation (216). Saijo and co-workers reported that NR4A2 represses transcription of TNF α and IL-1 β in macrophages and microglia, and this trans-repression is mediated by recruitment of co-repressor for element-1 silencing transcription factor (CoREST) complex to a target promoter and clearance of NF- κ B (144). Although the NR4A subfamily is known to be operated in a ligand independent manner (128), several compounds based on 1,1-bis(3'-indolyl)-1-(p-substituted phenyl)methanes (C-DIMs) have been recently identified as agonists for NR4A (132, 147). DIM is generated through heterodimerization of indole-3-carbinol in the acid environment of the gut (81). Since indole-3-carbinol can be structurally

considered as a derivative of the indole, we hypothesized that indole attenuates production of TNF α in part through activation of NR4A2. As shown in Fig. 5.2B, NR4A2 knockdown resulted in a smaller down-regulation of TNF α upon indole exposure as compared to the control. This result demonstrates that activation of NR4A2 is required for down-regulation of TNF α production by indole in DC2.4 cells. Similar results were also obtained in primary DC's (not shown).

While knockdown of NR4A2 resulted in a decrease in the indole-mediated suppression of TNF α , complete abolishment of TNF α expression was not observed. This could be due to low knockdown efficiency (~53%) or due to the presence of additional receptors that can bind indole and mediate its effect on inflammatory cytokine production.

An example of a second receptor for indole is the aryl hydrocarbon receptor (AhR). The AhR is a ligand-activated transcription factor that plays an important role in the intestinal immune system (70, 71). Immunomodulation by AhR activation in DCs has been reported in several studies. For example, the AhR agonist benzo(α)pyrene was shown to inhibit TNF α secretion by LPS-stimulated mouse bone marrow-derived dendritic cells (217). The decrease of TNF α production was likely due to AhR activation because the AhR antagonist alpha-naphthoflavone counteracted the effect of benzo(α)pyrene. In another study, the well-known AhR agonist TCDD suppressed binding of NF- κ B/Rel which has been described to be involved in pro-inflammatory gene regulation such as IL-6 to its response element and blocked translocation into the nucleus in TNF α treated DC2.4 cells (218, 219). Current work in our laboratory is

focused on evaluating the effects of AhR knockdown on the indole-mediated down-regulation of TNF α expression in DC 2.4 cells.

In summary, we have identified NR4A2 as a nuclear receptor to which indole binds to induce changes in TNF α production in DC 2.4 cells. Activation of NR4A2 is essential for maximal anti-inflammatory activity of indole in DC2.4 cells. Although exogenous and synthetic ligands for NR4A2 are previously reported, this is the first report of a putative endogenous ligand for NR4A2 and provides a new tool for modulation of inflammatory disorders.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1 CONCLUSIONS

In this work, we have successfully presented a methodology for the systematic prediction, identification and quantification of metabolites produced by the intestinal microbiota (Chapter III). We predicted tryptophan (TRP) derived metabolites by constructing biochemical transformation pathways from candidate bacterial reactions in the KEGG database and excluding metabolites which could be formed through host metabolism. Among predicted TRP derivatives, the final panel of metabolites (i.e., indole, hydroxyindole, indole 3-acetate, indole 3-acetamide, tryptamine, TRP) was quantified in murine fecal and cecum extracts using MRM mass spectrometry. To our knowledge, this is the first study to use MRM mass spectrometry for absolute quantification of microbiota metabolites in murine tissue samples. Additionally, some of the identified metabolites (i.e., indole 3-acetate, tryptamine, indole 3-pyruvate) were shown to activate AhR signaling, thus suggesting putative biological roles for these metabolites. The strength of our methodology is the ability to unambiguously identify bacterial metabolites by discriminating host-derived metabolites and quantify their physiological levels in the GI tract. This systematic approach is expected to have a significant impact on the study of human gut microbiome function and lead to the discovery of derivatives of natural compounds generated in the body with biological significance that have not previously been investigated.

We developed an integrated network model of dynamics of key transcription factors (i.e., PPAR γ , C/EBP β , CREB, NFAT, FoxO1, SREBP1c) underlying adipocyte differentiation and lipid loading (Chapter IV). A hypothetical network model was determined based on published connections, and stochastic simulation algorithm was applied to simulate random fluctuations of the transcription factors. Transcription factor activation profiles were experimentally measured using 3T3-L1 reporter cell lines where the binding of a transcription factor to its DNA binding element drives expression of the *Gaussia* luciferase gene. The simulated profiles were consistent with the measured profiles, and correctly ordered peak activities of the transcription factors. The model was further verified by perturbing the activation level of CREB and determining its effects on CREB-downstream transcription factors (i.e., PPAR γ , C/EBP β). The integrated model of the transcription factor network developed in this study is expected to provide a powerful tool for direct control of the entire regulatory network and a framework for manipulating adipocyte metabolism against obesity.

We investigated the molecular recognition mechanism of indole and its anti-inflammatory signaling in DC2.4 cells (Chapter V). It was shown that indole suppresses the production of pro-inflammatory cytokine, TNF α , upon LPS stimulation. To understand how indole exerts the anti-inflammatory effect in DC2.4 cells, a functional study of NR4A2 knockdown was performed. The NR4A2 knockdown resulted in less suppression of indole on TNF α production. This result indicates that NR4A2 is a molecular mediator of indole for the suppression of TNF α production. The finding of

indole as a NR4A endogeneous ligand is expected to provide a new strategy for treatment of inflammatory disorders.

6.2 RECOMMENDATIONS

An extension of the methodology described in Chapter III is to predict and identify microbiota metabolites which can be derived from different source compounds. One such source compound could be phenylalanine of which derivatives have been shown to be associated with several diseases such as diabetes and autism. For example, phenylacetylglutamine whose precursor, phenylacetic acid is produced by gut bacterial metabolism was found in higher concentrations in the plasma of diabetic patients (61) and lower levels in urine of autistic children as compared to control (220). The methodology can also be applied to identifying bioactive derivatives of environmental contaminants such as bisphenol A which was demonstrated to modulate glucose transport (221) and increase lipid accumulation in adipocytes (222).

In the studies of Chapter III and Chapter V, we only focused on regulation of bacterial metabolites on host function. In the future work, an interplay model in which microbiota metabolites are converted to other compounds by host metabolism can be explored to understand comprehensive contribution of microbiota metabolites on human diseases. For example, hippurate whose expression levels decrease in urine of autistic individuals is formed by hepatic glycine conjugation of dietary and gut microbial-derived benzoate (223). In liver, gut bacterial metabolite indole is transformed to indoxyl by cytochrome P-450, and the indoxyl is further converted to indoxyl sulfate by

sulfotransferase (64). The indoxyl sulfate is known to accumulate in the blood of patients with chronic kidney failure (65). In this way, co-metabolites generated by host enzymes from metabolites produced by the microbiota, could play an important role in health and disease. It should be noted that the algorithm presented in Chapter III can be modified to predict these co-metabolites by starting with a specific microbiota metabolite and constructing host-derived transformation pathways shown in Chapter III.

Gut microbiota are important factors in obesity and metabolic disorders, but little is known about direct influence of microbiota metabolites on adipocyte differentiation and inflammation. The integrated transcription factor (TF) network we developed in Chapter IV can be used as a platform to investigate how microbiota metabolites modulate adipocyte differentiation, lipid accumulation, and inflammation. A strategy for this purpose is as follows. First, the TRP derivatives identified in Chapter III are used to treat 3T3-L1 TF reporter cells showing the earliest activation (i.e., CREB) or persistent activation at a later stage of differentiation process (i.e., PPAR γ). After getting TRP metabolites-triggered perturbation of CREB or PPAR γ profiles, a change in the activation of other TFs and of the entire TF network can be determined, as any perturbation to the TF network is expected to lead to a change in the expression level of target genes which are regulated by these TFs.

Obesity is characterized by low-grade chronic inflammation and macrophage infiltration. To mimic the environment of macrophage infiltration, the spent medium of LPS-stimulated macrophage J774A.1 cells can be introduced into the 3T3-L1 adipocyte culture medium. In a preliminary study, it was demonstrated that the spent medium of

LPS-stimulated J774A.1 cells induced activation of NFAT (data not shown). In addition, the binding activity of NFAT was shown to be significantly increased in white adipose tissues of obese mice as compared to control (224). Therefore, an increase of the NFAT activity by the spent medium of LPS-stimulated J774A.1 cells can be linked to adipocyte inflammation during obesity development. A study to investigate whether microbiota metabolites are involved in modulation of adipocyte inflammation can be carried out in two ways, as TRP metabolites can modulate secretion of pro-inflammatory molecules in J774A.1 cells or interfere with pro-inflammatory signaling in 3T3-L1 cells. In the first approach, TRP-derived metabolites can be added to the J774A.1 culture medium during LPS stimulation to modulate composition of the spent medium. Then, the spent medium with altered composition can be used to stimulate 3T3-L1 NFAT reporter cells to monitor a change of NFAT activity. In the second approach, 3T3-L1 NFAT reporter cells can be exposed to TRP metabolites after introducing spent medium from LPS-stimulated J774A.1 cells.

In the functional study of NR4A2 knockdown (Chapter V), it was demonstrated that NR4A2 is a molecular mediator of anti-inflammatory signaling of indole. The AhR is another attractive target for the knockdown study because several TRP-derived metabolites were shown to have ability of AhR activation in our and other reports. If knockdown of AhR also attenuates indole-derived suppression of TNF α production, we can hypothesize both NR4A2 and AhR participate in sensing of indole for anti-inflammatory activity. Further, our data in which the NR4A2 knockdown did not completely abolish indole-derived suppression of TNF α production support a possibility

that the anti-inflammatory activity is an output of the sum of several mechanisms. Simultaneous knockdown of NR4A2 and AhR is expected to provide information on the extent to which the different pathways are involved in indole-mediated suppression of cytokine production in DC 2.4 cells. For other TRP derivatives which may show indole-like behavior, a similar approach can be applied to identify mechanisms involved in mediating their anti-inflammatory functions.

Together, these strategies and experiments can provide fundamental understanding of the interaction between intestinal microbiota and mucosal host interactions cells, which can in turn lead to the development of novel treatment strategies and modalities against obesity and gastrointestinal inflammation.

REFERENCES

1. Tiihonen K, Ouwehand AC, & Rautonen N (2010) Human intestinal microbiota and healthy ageing. *Ageing research reviews* 9(2):107-116.
2. Arora T & Sharma R (2011) Fermentation potential of the gut microbiome: implications for energy homeostasis and weight management. *Nutrition reviews* 69(2):99-106.
3. Holmes E, Li JV, Athanasiou T, Ashrafi H, & Nicholson JK (2011) Understanding the role of gut microbiome–host metabolic signal disruption in health and disease. *Trends in microbiology* 19(7):349-359.
4. Chassaing B & Darfeuille-Michaud A (2011) The commensal microbiota and enteropathogens in the pathogenesis of inflammatory bowel diseases. *Gastroenterology* 140:1720-1728.
5. Man SM, Kaakoush NO, & Mitchell HM (2011) The role of bacteria and pattern-recognition receptors in Crohn's disease. *Nature reviews gastroenterology and hepatology* 8:152-168.
6. Burcelin R, Serino M, Chabo C, Blasco-Baque V, & Amar J (2011) Gut microbiota and diabetes: from pathogenesis to therapeutic perspective. *Acta diabetologica* 48(4):257-273.
7. Cani PD, *et al.* (2007) Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes* 56(7):1761-1772.

8. Arnold IC, *et al.* (2011) Helicobacter pylori infection prevents allergic asthma in mouse models through the induction of regulatory T cells. *Journal of clinical investigation* 121(8):3088-3093.
9. Round JL & Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nature reviews immunology* 9(5):313-323.
10. Robinson CJ, Bohannan BJM, & Young VB (2010) From Structure to Function: the Ecology of Host-Associated Microbial Communities. *Microbiology and molecular biology reviews* 74(3):453-476.
11. Dumas M-E, *et al.* (2006) Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proceedings of the national academy of sciences* 103(33):12511-12516.
12. Wang Z, *et al.* (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472(7341):57-63.
13. Tilg H & Kaser A (2011) Gut microbiome, obesity, and metabolic dysfunction. *Journal of clinical investigation* 121(6):2126-2132.
14. Wickelgren I (1998) Obesity: How Big a Problem? *Science* 280(5368):1364-1367.
15. Boone C, Mouro J, Gregoire F, & Remacle C (2000) The adipose conversion process: regulation by extracellular and intracellular factors. *Reproduction nutrition development* 40:325-328.
16. Reusch JEB, Colton LA, & Klemm DJ (2000) CREB activation induces adipogenesis in 3T3-L1 cells. *Molecular and cellular biology* 20(3):1008-1020.

17. Rosen ED & Spiegelman BM (2000) Molecular regulation of adipogenesis. *Annual review of cell and developmental biology* 16(1):145-171.
18. Ley RE, *et al.* (2005) Obesity alters gut microbial ecology. *Proceedings of the national academy of sciences* 102(31):11070-11075.
19. Waldram A, *et al.* (2009) Top-down systems biology modeling of host metabotype–microbiome associations in obese rodents. *Journal of proteome research* 8(5):2361-2375.
20. Turnbaugh PJ, *et al.* (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444(7122):1027-1131.
21. Bäckhed F, *et al.* (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the national academy of sciences* 101(44):15718-15723.
22. Sartor RB (2006) Mechanisms of disease: pathogenesis of crohn's disease and ulcerative colitis. *Nature clinical practice gastroenterology and hepatology* 3(7):390-407.
23. Willing B, *et al.* (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. *Inflammatory bowel diseases* 15(5):653-660.
24. Zilber-Rosenberg I & Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS microbiology reviews* 32(5):723-735.

25. Finkelstein EA, Trogon JG, Cohen JW, & Dietz W (2009) Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health affairs* 28(5):w822-w831.
26. Blonde L (2007) State of diabetes care in the United States. *American journal of managed care* 13 Suppl 2:S36-40.
27. Bansal T, Alaniz RC, Wood TK, & Jayaraman A (2010) The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proceedings of the national academy of sciences* 107(1):228-233.
28. Van Duynhoven J, *et al.* (2011) Metabolic fate of polyphenols in the human superorganism. *Proceedings of the national academy of sciences* 108(Suppl 1):4531-4538.
29. Zheng X, *et al.* (2011) The footprints of gut microbial-mammalian co-metabolism. *Journal of proteome research* 10:5512-5522.
30. Arumugam M, *et al.* (2011) Enterotypes of the human gut microbiome. *Nature* 473:174-180.
31. Antunes LC, Davies JE, & Finlay BB (2011) Chemical signaling in the gastrointestinal tract. *F1000 biology report* 3:4.
32. Want EJ, Cravatt BF, & Siuzdak G (2005) The expanding role of mass spectrometry in metabolite profiling and characterization. *Chembiochem* 6:1941-1951.

33. Savage DB (2005) PPAR[gamma] as a metabolic regulator: insights from genomics and pharmacology. *Expert reviews in molecular medicine* 7(01):1-16.
34. Xu J & Gordon JI (2003) Honor thy symbionts. *Proceedings of the national academy of sciences* 100(18):10452-10459.
35. Zoetendal EG, Vaughan EE, & De Vos WM (2006) A microbial world within us. *Molecular microbiology* 59(6):1639-1650.
36. Vrieze A, *et al.* (2010) The environment within: how gut microbiota may influence metabolism and body composition. *Diabetologia* 53(4):606-613.
37. Nagalingam NA & Lynch SV (2012) Role of the microbiota in inflammatory bowel diseases. *Inflammatory bowel diseases* 18(5):968-984.
38. Macfarlane GT, Blackett KL, Nakayama T, Steed H, & Macfarlane S (2009) The gut microbiota in inflammatory bowel disease. *Current pharmaceutical design* 15(13):1528-1536.
39. Fava F & Danese S (2011) Intestinal microbiota in inflammatory bowel disease: friend of foe? *World journal of gastroenterology* 17(5):557-566.
40. Roediger WE, Duncan A, Kapaniris O, & Millard S (1993) Reducing sulfur compounds of the colon impair colonocyte nutrition: implications for ulcerative colitis. *Gastroenterology* 104(3):802-809.
41. Attene-Ramos MS, Wagner ED, Gaskins HR, & Plewa MJ (2007) Hydrogen sulfide induces direct radical-associated DNA damage. *Molecular cancer research* 5(5):455-459.

42. Turnbaugh PJ, *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature* 457(7228):480-484.
43. Henao-Mejia J, *et al.* (2012) Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. *Nature* 482(7384):179-185.
44. Mikuls TR, *et al.* (2009) Antibody responses to *Porphyromonas gingivalis* (P. *gingivalis*) in subjects with rheumatoid arthritis and periodontitis. *International immunopharmacology* 9(1):38-42.
45. Moen K, *et al.* (2006) Synovial inflammation in active rheumatoid arthritis and psoriatic arthritis facilitates trapping of a variety of oral bacterial DNAs. *Clinical and experimental rheumatology* 24(6):656-663.
46. Fallani M, *et al.* (2010) Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics. *Journal of pediatric gastroenterology and nutrition* 51(1):77-84.
47. Penders J, *et al.* (2007) Gut microbiota composition and development of atopic manifestations in infancy: the koala birth cohort study. *Gut* 56(5):661-667.
48. Blaut M & Clavel T (2007) Metabolic diversity of the intestinal microbiota: implications for health and disease. *Journal of nutrition* 137(3):751S-755S.
49. Nordgaard I & Mortensen PB (1995) Digestive processes in the human colon. *Nutrition* 11(1):37-45.
50. Roediger WE (1982) Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* 83(2):424-429.

51. Grubben MJ, *et al.* (2001) Effect of resistant starch on potential biomarkers for colonic cancer risk in patients with colonic adenomas: a controlled trial. *Digestive diseases and sciences* 46(4):750-756.
52. Wollowski I, Rechkemmer G, & Pool-Zobel BL (2001) Protective role of probiotics and prebiotics in colon cancer. *American journal of clinical nutrition* 73(2 Suppl):451S-455S.
53. Segain JP, *et al.* (2000) Butyrate inhibits inflammatory responses through NFkappaB inhibition: implications for Crohn's disease. *Gut* 47(3):397-403.
54. Macfarlane GT & Macfarlane S (2012) Bacteria, colonic fermentation, and gastrointestinal health. *Journal of AOAC international* 95(1):50-60.
55. Pitcher MC & Cummings JH (1996) Hydrogen sulphide: a bacterial toxin in ulcerative colitis? *Gut* 39(1):1-4.
56. Florin T, Neale G, Gibson GR, Christl SU, & Cummings JH (1991) Metabolism of dietary sulphate: absorption and excretion in humans. *Gut* 32(7):766-773.
57. Bone E, Tamm A, & Hill M (1976) The production of urinary phenols by gut bacteria and their possible role in the causation of large bowel cancer. *American journal of clinical nutrition* 29(12):1448-1454.
58. Zheng S, *et al.* (2010) Urinary metabonomic study on biochemical changes in chronic unpredictable mild stress model of depression. *Clinica chimica acta* 411(3-4):204-209.
59. Chung KT, Anderson GM, & Fulk GE (1975) Formation of indoleacetic acid by intestinal anaerobes. *Journal of bacteriology* 124(1):573-575.

60. Russell DW & Setchell KD (1992) Bile acid biosynthesis. *Biochemistry* 31(20):4737-4749.
61. Suhre K, *et al.* (2010) Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *Plos one* 5(11):e13953.
62. Cani PD & Delzenne NM (2009) The role of the gut microbiota in energy metabolism and metabolic disease. *Current pharmaceutical design* 15(13):1546-1558.
63. Frank DN & Pace NR (2008) Gastrointestinal microbiology enters the metagenomics era. *Current opinion in gastroenterology* 24(1):4-10.
64. Wikoff WR, *et al.* (2009) Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the national academy of sciences* 106(10):3698-3703.
65. Deguchi T, *et al.* (2002) Major role of organic anion transporter 3 in the transport of indoxyl sulfate in the kidney. *Kidney international* 61(5):1760-1768.
66. Antunes LCM, *et al.* (2011) Effect of antibiotic treatment on the intestinal metabolome. *Antimicrobial agents and chemotherapy* 55(4):1494-1503.
67. Marchesi JR, *et al.* (2007) Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *Journal of proteomic research* 6(2):546-551.
68. Monteleone I, *et al.* (2011) Aryl hydrocarbon receptor-induced signals up-regulate IL-22 production and inhibit inflammation in the gastrointestinal tract. *Gastroenterology* 141(1):237-248.

69. Willing B, *et al.* (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal crohn's disease. *Inflammatory bowel diseases* 15(5):653-660.
70. Kiss EA, *et al.* (2011) Natural aryl hydrocarbon receptor ligands control organogenesis of intestinal lymphoid follicles. *Science* 334(6062):1561-1565.
71. Li Y, *et al.* (2011) Exogenous stimuli maintain intraepithelial lymphocytes via aryl hydrocarbon receptor activation. *Cell* 147(3):629-640.
72. Ito S, Chen C, Satoh J, Yim S, & Gonzalez FJ (2007) Dietary phytochemicals regulate whole-body CYP1A1 expression through an arylhydrocarbon receptor nuclear translocator-dependent system in gut. *Journal of clinical investigation* 117(7):1940-1950.
73. Fan Y, *et al.* (2010) The aryl hydrocarbon receptor functions as a tumor suppressor of liver carcinogenesis. *Cancer research* 70(1):212-220.
74. Hall JM, *et al.* (2010) Activation of the aryl-hydrocarbon receptor inhibits invasive and metastatic features of human breast cancer cells and promotes breast cancer cell differentiation. *Molecular endocrinology* 24(2):359-369.
75. Safe SH (1994) Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment. *Critical reviews in toxicology* 24(2):87-149.
76. Kafafi SA, Afeefy HY, Said HK, & Kafafi AG (1993) Relationship between aryl hydrocarbon receptor binding, induction of aryl hydrocarbon hydroxylase and 7-ethoxyresorufin O-deethylase enzymes, and toxic activities of aromatic

- xenobiotics in animals. A new model. *Chemical research in toxicology* 6(3):328-334.
77. Waller CL & McKinney JD (1995) Three-dimensional quantitative structure-activity relationships of dioxins and dioxin-like compounds: model validation and Ah receptor characterization. *Chemical research in toxicology* 8(6):847-858.
 78. Ma Q & Whitlock JP, Jr. (1996) The aromatic hydrocarbon receptor modulates the Hepa 1c1c7 cell cycle and differentiated state independently of dioxin. *Molecular and cellular biology* 16(5):2144-2150.
 79. Sinal CJ & Bend JR (1997) Aryl hydrocarbon receptor-dependent induction of cyp1a1 by bilirubin in mouse hepatoma hepa 1c1c7 cells. *Molecular pharmacology* 52(4):590-599.
 80. Fernandez-Salguero P, *et al.* (1995) Immune system impairment and hepatic fibrosis in mice lacking the dioxin-binding Ah receptor. *Science* 268(5211):722-726.
 81. Bjeldanes LF, Kim JY, Grose KR, Bartholomew JC, & Bradfield CA (1991) Aromatic hydrocarbon responsiveness-receptor agonists generated from indole-3-carbinol in vitro and in vivo: comparisons with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Proceedings of the national academy of sciences* 88(21):9543-9547.
 82. Gillner M, *et al.* (1993) Interactions of indolo[3,2-b]carbazoles and related polycyclic aromatic hydrocarbons with specific binding sites for 2,3,7,8-tetrachlorodibenzo-p-dioxin in rat liver. *Molecular pharmacology* 44(2):336-345.

83. Heath-Pagliuso S, *et al.* (1998) Activation of the Ah Receptor by Tryptophan and Tryptophan Metabolites. *Biochemistry* 37(33):11508-11515.
84. Rannug A, *et al.* (1987) Certain photooxidized derivatives of tryptophan bind with very high affinity to the Ah receptor and are likely to be endogenous signal substances. *Journal of biological chemistry* 262(32):15422-15427.
85. Wei YD, Rannug U, & Rannug A (1999) UV-induced CYP1A1 gene expression in human cells is mediated by tryptophan. *Chemico-biological interactions* 118(2):127-140.
86. Wincent E, *et al.* (2009) The suggested physiologic aryl hydrocarbon receptor activator and cytochrome P4501 substrate 6-formylindolo[3,2-b]carbazole is present in humans. *Journal of biological chemistry* 284(5):2690-2696.
87. Cao Z, Umek RM, & McKnight SL (1991) Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. *Genes and development* 5(9):1538-1552.
88. Yeh WC, Cao Z, Classon M, & McKnight SL (1995) Cascade regulation of terminal adipocyte differentiation by three members of the C/EBP family of leucine zipper proteins. *Genes and development* 9(2):168-181.
89. Tontonoz P, Hu E, Graves RA, Budavari AI, & Spiegelman BM (1994) mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. *Genes and development* 8(10):1224-1234.

90. Wu Z, Xie Y, Bucher NL, & Farmer SR (1995) Conditional ectopic expression of C/EBP beta in NIH-3T3 cells induces PPAR gamma and stimulates adipogenesis. *Genes and development* 9(19):2350-2363.
91. Clarke SL, Robinson CE, & Gimble JM (1997) CAAT/enhancer binding proteins directly modulate transcription from the peroxisome proliferator-activated receptor gamma 2 promoter. *Biochemical and biophysical research communications* 240(1):99-103.
92. Zuo Y, Qiang L, & Farmer SR (2006) Activation of CCAAT/enhancer-binding protein (C/EBP) alpha expression by C/EBP beta during adipogenesis requires a peroxisome proliferator-activated receptor-gamma-associated repression of HDAC1 at the C/ebp alpha gene promoter. *Journal of biological chemistry* 281(12):7960-7967.
93. Wu Z, *et al.* (1999) Cross-regulation of C/EBP alpha and PPAR gamma controls the transcriptional pathway of adipogenesis and insulin sensitivity. *Molecular cell* 3(2):151-158.
94. Payne VA, *et al.* (2007) Sequential regulation of diacylglycerol acyltransferase 2 expression by CAAT/enhancer-binding protein beta (C/EBPbeta) and C/EBPalpha during adipogenesis. *Journal of biological chemistry* 282(29):21005-21014.
95. Hu E, Tontonoz P, & Spiegelman BM (1995) Transdifferentiation of myoblasts by the adipogenic transcription factors PPAR gamma and C/EBP alpha. *Proceedings of the national academy of sciences* 92(21):9856-9860.

96. Barak Y, *et al.* (1999) PPAR gamma is required for placental, cardiac, and adipose tissue development. *Molecular cell* 4(4):585-595.
97. Rosen ED, *et al.* (1999) PPAR gamma is required for the differentiation of adipose tissue in vivo and in vitro. *Molecular cell* 4(4):611-617.
98. Wang ND, *et al.* (1995) Impaired energy homeostasis in C/EBP alpha knockout mice. *Science* 269(5227):1108-1112.
99. Tontonoz P, Kim JB, Graves RA, & Spiegelman BM (1993) ADD1: a novel helix-loop-helix transcription factor associated with adipocyte determination and differentiation. *Molecular and cellular biology* 13(8):4753-4759.
100. Kim JB & Spiegelman BM (1996) ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. *Genes and development* 10(9):1096-1107.
101. Kim JB, *et al.* (1998) Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. *Journal of clinical investigation* 101(1):1-9.
102. Fajas L, *et al.* (1999) Regulation of peroxisome proliferator-activated receptor gamma expression by adipocyte differentiation and determination factor 1/sterol regulatory element binding protein 1: implications for adipocyte differentiation and metabolism. *Molecular and cellular biology* 19(8):5495-5503.
103. Kim JB, Wright HM, Wright M, & Spiegelman BM (1998) ADD1/SREBP1 activates PPARgamma through the production of endogenous ligand. *Proceedings of the national academy of sciences* 95(8):4333-4337.

104. Zhang J-W, Klemm DJ, Vinson C, & Lane MD (2004) Role of CREB in transcriptional regulation of CCAAT/enhancer-binding protein β gene during adipogenesis. *Journal of biological chemistry* 279(6):4471-4478.
105. Ho I-C, Kim JH-J, Rooney JW, Spiegelman BM, & Glimcher LH (1998) A potential role for the nuclear factor of activated T cells family of transcriptional regulatory proteins in adipogenesis. *Proceedings of the national academy of sciences* 95(26):15537-15541.
106. Yang XY, *et al.* (2000) Activation of human T lymphocytes is inhibited by peroxisome proliferator-activated receptor γ (PPAR γ) agonists: PPAR γ co-association with transcription factor NFAT. *Journal of biological chemistry* 275(7):4541-4544.
107. Nakae J, *et al.* (2003) The forkhead transcription factor Foxo1 regulates adipocyte differentiation. *Developmental cell* 4(1):119-129.
108. Xu H, *et al.* (2003) Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. *Journal of clinical investigation* 112(12):1821-1830.
109. Kanda H, *et al.* (2006) MCP-1 contributes to macrophage infiltration into adipose tissue, insulin resistance, and hepatic steatosis in obesity. *Journal of clinical investigation* 116(6):1494-1505.
110. Kamei N, *et al.* (2006) Overexpression of monocyte chemoattractant protein-1 in adipose tissues causes macrophage recruitment and insulin resistance. *Journal of biological chemistry* 281(36):26602-26614.

111. Sierra-Honigmann MR, *et al.* (1998) Biological action of leptin as an angiogenic factor. *Science* 281(5383):1683-1686.
112. Wellen KE & Hotamisligil GS (2003) Obesity-induced inflammatory changes in adipose tissue. *Journal of clinical investigation* 112(12):1785-1788.
113. Stenholm S, *et al.* (2010) Adipocytokines and the metabolic syndrome among older persons with and without obesity: the InCHIANTI study. *Clinical endocrinology* 73(1):55-65.
114. Uysal KT, Wiesbrock SM, Marino MW, & Hotamisligil GS (1997) Protection from obesity-induced insulin resistance in mice lacking TNF- α function. *Nature* 389(6651):610-614.
115. Shi H, *et al.* (2006) TLR4 links innate immunity and fatty acid-induced insulin resistance. *Journal of clinical investigation* 116(11):3015-3025.
116. Suganami T, *et al.* (2007) Role of the Toll-like receptor 4/NF- κ B pathway in saturated fatty acid-induced inflammatory changes in the interaction between adipocytes and macrophages. *Arteriosclerosis, thrombosis, and vascular biology* 27(1):84-91.
117. Suganami T, Nishida J, & Ogawa Y (2005) A paracrine loop between adipocytes and macrophages aggravates inflammatory changes: role of free fatty acids and tumor necrosis factor α . *Arteriosclerosis, thrombosis, and vascular biology* 25(10):2062-2068.

118. Kang K, *et al.* (2008) Adipocyte-derived Th2 cytokines and myeloid PPARdelta regulate macrophage polarization and insulin sensitivity. *Cell metabolism* 7(6):485-495.
119. Ohashi K, *et al.* (2010) Adiponectin promotes macrophage polarization toward an anti-inflammatory phenotype. *Journal of biological chemistry* 285(9):6153-6160.
120. Iwaki M, *et al.* (2003) Induction of adiponectin, a fat-derived antidiabetic and antiatherogenic factor, by nuclear receptors. *Diabetes* 52(7):1655-1663.
121. Seo JB, *et al.* (2004) Adipocyte determination- and differentiation-dependent factor 1/sterol regulatory element-binding protein 1c regulates mouse adiponectin expression. *Journal of biological chemistry* 279(21):22108-22117.
122. Qiao L, *et al.* (2005) C/EBPalpha regulates human adiponectin gene transcription through an intronic enhancer. *Diabetes* 54(6):1744-1754.
123. Chinetti G, Zawadzki C, Fruchart JC, & Staels B (2004) Expression of adiponectin receptors in human macrophages and regulation by agonists of the nuclear receptors PPARalpha, PPARgamma, and LXR. *Biochemical and biophysical research communications* 314(1):151-158.
124. Kim KY, *et al.* (2005) c-Jun N-terminal kinase is involved in the suppression of adiponectin expression by TNF-alpha in 3T3-L1 adipocytes. *Biochemical and biophysical research communications* 327(2):460-467.

125. Fasshauer M, *et al.* (2003) Adiponectin gene expression and secretion is inhibited by interleukin-6 in 3T3-L1 adipocytes. *Biochemical and biophysical research communications* 301(4):1045-1050.
126. Olefsky JM (2001) Nuclear receptor minireview series. *Journal of biological chemistry* 276(40):36863-36864.
127. Chawla A, Repa JJ, Evans RM, & Mangelsdorf DJ (2001) Nuclear receptors and lipid physiology: opening the X-files. *Science* 294(5548):1866-1870.
128. Wilson TE, Fahrner TJ, Johnston M, & Milbrandt J (1991) Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* 252(5010):1296-1300.
129. Wansa KD, Harris JM, Yan G, Ordentlich P, & Muscat GE (2003) The AF-1 domain of the orphan nuclear receptor NOR-1 mediates trans-activation, coactivator recruitment, and activation by the purine anti-metabolite 6-mercaptopurine. *Journal of biological chemistry* 278(27):24776-24790.
130. Wang Z, *et al.* (2003) Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature* 423(6939):555-560.
131. Wansa KD, Harris JM, & Muscat GE (2002) The activation function-1 domain of Nur77/NR4A1 mediates trans-activation, cell specificity, and coactivator recruitment. *Journal of biological chemistry* 277(36):33001-33011.
132. Chintharlapalli S, *et al.* (2005) Activation of Nur77 by selected 1,1-Bis(3'-indolyl)-1-(p-substituted phenyl)methanes induces apoptosis through nuclear pathways. *Journal of biological chemistry* 280(26):24903-24914.

133. Pearen MA, *et al.* (2006) The orphan nuclear receptor, NOR-1, is a target of beta-adrenergic signaling in skeletal muscle. *Endocrinology* 147(11):5217-5227.
134. Fernandez PM, *et al.* (2000) Nuclear receptors Nor1 and NGFI-B/Nur77 play similar, albeit distinct, roles in the hypothalamo-pituitary-adrenal axis. *Endocrinology* 141(7):2392-2400.
135. Davis IJ & Lau LF (1994) Endocrine and neurogenic regulation of the orphan nuclear receptors Nur77 and Nurr-1 in the adrenal glands. *Molecular and cellular biology* 14(5):3469-3483.
136. Lim RW, Zhu CY, & Stringer B (1995) Differential regulation of primary response gene expression in skeletal muscle cells through multiple signal transduction pathways. *Biochimica et biophysica acta* 1266(1):91-100.
137. Miyakoshi J, *et al.* (1998) Enhanced NOR-1 gene expression by exposure of Chinese hamster cells to high-density 50 Hz magnetic fields. *Molecular and cellular biochemistry* 181(1-2):191-195.
138. Bandoh S, Tsukada T, Maruyama K, Ohkura N, & Yamaguchi K (1997) Mechanical agitation induces gene expression of NOR-1 and its closely related orphan nuclear receptors in leukemic cell lines. *Leukemia* 11(9):1453-1458.
139. Wilson TE, Fahrner TJ, & Milbrandt J (1993) The orphan receptors NGFI-B and steroidogenic factor 1 establish monomer binding as a third paradigm of nuclear receptor-DNA interaction. *Molecular and cellular biology* 13(9):5794-5804.

140. Perlmann T & Jansson L (1995) A novel pathway for vitamin A signaling mediated by RXR heterodimerization with NGFI-B and NURR1. *Genes and development* 9(7):769-782.
141. Pei L, Castrillo A, Chen M, Hoffmann A, & Tontonoz P (2005) Induction of NR4A orphan nuclear receptor expression in macrophages in response to inflammatory stimuli. *Journal of biological chemistry* 280(32):29256-29262.
142. Pei L, Castrillo A, & Tontonoz P (2006) Regulation of macrophage inflammatory gene expression by the orphan nuclear receptor Nur77. *Molecular endocrinology* 20(4):786-794.
143. Bonta PI, *et al.* (2006) Nuclear receptors Nur77, Nurr1, and NOR-1 expressed in atherosclerotic lesion macrophages reduce lipid loading and inflammatory responses. *Arteriosclerosis, thrombosis, and vascular biology* 26(10):2288-2294.
144. Saijo K, *et al.* (2009) A Nurr1/CoREST pathway in microglia and astrocytes protects dopaminergic neurons from inflammation-induced death. *Cell* 137(1):47-59.
145. Ordentlich P, Yan Y, Zhou S, & Heyman RA (2003) Identification of the antineoplastic agent 6-mercaptopurine as an activator of the orphan nuclear hormone receptor Nurr1. *Journal of biological chemistry* 278(27):24791-24799.
146. Wansa KD & Muscat GE (2005) TRAP220 is modulated by the antineoplastic agent 6-mercaptopurine, and mediates the activation of the NR4A subgroup of nuclear receptors. *Journal of molecular endocrinology* 34(3):835-848.

147. Cho SD, *et al.* (2007) Nur77 agonists induce proapoptotic genes and responses in colon cancer cells through nuclear receptor-dependent and nuclear receptor-independent pathways. *Cancer research* 67(2):674-683.
148. Dubois C, Hengeler B, & Mattes H (2006) Identification of a potent agonist of the orphan nuclear receptor Nurr1. *Chemmedchem* 1(9):955-958.
149. Zhan Y, *et al.* (2008) Cytosporone B is an agonist for nuclear orphan receptor Nur77. *Nature chemical biology* 4(9):548-556.
150. Li JV, *et al.* (2011) Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut* 60(9):1214-1223.
151. Lu W, Bennett BD, & Rabinowitz JD (2008) Analytical strategies for LC-MS-based targeted metabolomics. *Journal of chromatography. b, analytical technologies in the biomedical and life sciences* 871(2):236-242.
152. Brown M, *et al.* (2011) Automated workflows for accurate mass-based putative metabolite identification in LC/MS-derived metabolomic datasets. *Bioinformatics* 27(8):1108-1112.
153. Martin FP, *et al.* (2010) Dietary modulation of gut functional ecology studied by fecal metabonomics. *Journal of proteomic research* 9(10):5284-5295.
154. Greenblum S, Turnbaugh PJ, & Borenstein E (2012) Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proceedings of the national academy of sciences* 109(2):594-599.

155. Yousofshahi M, Lee K, & Hassoun S (2011) Probabilistic pathway construction. *Metabolic engineering* 13(4):435-444.
156. Selvarasu S, Karimi IA, Ghim GH, & Lee DY (2010) Genome-scale modeling and in silico analysis of mouse cell metabolic network. *Molecular biosystems* 6(1):152-161.
157. Sigurdsson MI, Jamshidi N, Steingrimsso E, Thiele I, & Palsson BO (2010) A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. *BMC systems biology* 4:140.
158. Sellick C, *et al.* (2010) Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: matching extraction approaches to cell type and metabolite targets. *Metabolomics* 6(3):427-438.
159. Bajad SU, *et al.* (2006) Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *Journal of chromatography. a* 1125(1):76-88.
160. Nagy SR, Sanborn JR, Hammock BD, & Denison MS (2002) Development of a green fluorescent protein-based cell bioassay for the rapid and inexpensive detection and characterization of ah receptor agonists. *Toxicological sciences* 65(2):200-210.
161. Tian J, Alimperti S, Lei P, & Andreadis ST (2010) Lentiviral microarrays for real-time monitoring of gene expression dynamics. *Lab on a chip* 10(15):1967-1975.

162. Carlotti F, *et al.* (2004) Lentiviral vectors efficiently transduce quiescent mature 3T3-L1 adipocytes. *Molecular therapy* 9(2):209-217.
163. Kanehisa M, Goto S, Furumichi M, Tanabe M, & Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic acids research* 38(Database issue):D355-360.
164. Juorio AV & Durden DA (1984) The distribution and turnover of tryptamine in the brain and spinal cord. *Neurochemical research* 9(9):1283-1293.
165. Jones BV, Begley M, Hill C, Gahan CG, & Marchesi JR (2008) Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proceedings of the national academy of sciences* 105(36):13580-13585.
166. Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiology and molecular biology reviews* 68(4):669-685.
167. Aziz RK, *et al.* (2012) SEED servers: high-performance access to the SEED genomes, annotations, and metabolic models. *Plos one* 7(10):e48053.
168. Whitt DD & Demoss RD (1975) Effect of microflora on the free amino acid distribution in various regions of the mouse gastrointestinal tract. *Applied microbiology* 30(4):609-615.
169. Vahtovuo J, Toivanen P, & Eerola E (2001) Study of murine faecal microflora by cellular fatty acid analysis; effect of age and mouse strain. *Antonie van leeuwenhoek* 80(1):35-42.

170. Van de Wiele T, *et al.* (2005) Human colon microbiota transform polycyclic aromatic hydrocarbons to estrogenic metabolites. *Environmental health perspectives* 113(1):6-10.
171. Olefsky JM (1977) Mechanisms of decreased insulin responsiveness of large adipocytes. *Endocrinology* 100:1169-1177.
172. Friedman JM, *et al.* (2000) Obesity in the new millenium: effect of fat cell size on its sensitivy to insulin measured by a new method. adipose expression of tumor necrosis factor-alpha: direct role in obesity linked insulin resistance. *Nature* 404:632-634.
173. Greenberg AS & Obin MS (2006) Obesity and the role of adipose tissue in inflammation and metabolism. *American journal of clinical nutrition* 83:461S-465S.
174. Chehab FF (2008) Obesity and lipodystrophy-where do the circles intersect? *Endocrinology* 149:925-934.
175. Reusch JE, Colton LA, & Klemm DJ (2000) CREB activation induces adipogenesis in 3T3-L1 cells. *Molecular and cellular biology* 20:1008-1020.
176. Rosen ED & Spiegelman BM (2000) Molecular regulation of adipogenesis. *Annual review of cell and developmental biology* 16:145-171.
177. Tontonoz P, Hu E, Graves RA, Budavari AI, & Spiegelman BM (1994) mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. *Genes and development* 8(10):1224-1234.

178. Rangwala SM & Lazar MA (2000) Transcriptional control of adipogenesis. *Annual review of nutrition* 20:535-559.
179. Kim JB & Spiegelman BM (1996) ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. *Genes and development* 10(9):1096-1107.
180. Fajas L, *et al.* (1999) Regulation of peroxisome proliferator-activated receptor γ expression by adipocyte differentiation and determination factor 1/sterol regulatory element binding protein 1: implications for adipocyte differentiation and metabolism. *Molecular and cellular biology* 19(8):5495-5503.
181. Yang TTC & Chow C-W (2003) Transcription cooperation by NFAT·C/EBP composite enhancer complex. *Journal of biological chemistry* 278(18):15874-15885.
182. Nakae J, *et al.* (2003) The forkhead transcription factor Foxo1 regulates adipocyte differentiation. *Developmental cell* 4(1):119-129.
183. Savage DB (2005) PPAR gamma as a metabolic regulator: insights from genomics and pharmacology. *Expert reviews in molecular medicine* 7:1-16.
184. Tian J & Andreadis ST (2009) Independent and high-level dual-gene expression in adult stem-progenitor cells from a single lentiviral vector. *Gene therapy* 16(7):874-884.
185. Carlotti F, *et al.* (2004) Lentiviral vectors efficiently transduce quiescent mature 3T3-L1 adipocytes. *Molecular therapy* 9(2):209-217.

186. Walker EC, *et al.* (2010) Oncostatin M promotes bone formation independently of resorption when signaling through leukemia inhibitory factor receptor in mice. *Journal of clinical investigation* 120(2):582-592.
187. Yang XY, *et al.* (2000) Activation of human T lymphocytes is inhibited by peroxisome proliferator-activated receptor γ (PPAR γ) agonists. *Journal of biological chemistry* 275(7):4541-4544.
188. Davies SS, *et al.* (2001) Oxidized alkyl phospholipids are specific, high affinity peroxisome proliferator-activated receptor gamma ligands and agonists. *Journal of biological chemistry* 276(19):16015-16023.
189. Limor R, *et al.* (2008) Lipoxygenase-derived metabolites are regulators of peroxisome proliferator-activated receptor gamma-2 expression in human vascular smooth muscle cells. *American journal of hypertension* 21(2):219-223.
190. Shoback DM & Brown EM (1984) Forskolin increases cellular cyclic adenosine monophosphate content and parathyroid hormone release in dispersed bovine parathyroid cells. *Metabolism: clinical and experimental* 33(6):509-514.
191. Zhu Y, *et al.* (1995) Structural organization of mouse peroxisome proliferator-activated receptor gamma (mPPAR gamma) gene: alternative promoter use and different splicing yield two mPPAR gamma isoforms. *Proceedings of the national academy of sciences* 92(17):7921-7925.
192. Fan W, *et al.* (2009) FOXO1 transrepresses peroxisome proliferator-activated receptor γ transactivation, coordinating an insulin-induced feed-forward response in adipocytes. *Journal of biological chemistry* 284(18):12188-12197.

193. Fu M, *et al.* (2001) Platelet-derived growth factor promotes the expression of peroxisome proliferator-activated receptor gamma in vascular smooth muscle cells by a phosphatidylinositol 3-kinase/Akt signaling pathway. *Circulation research* 89(11):1058-1064.
194. Waite KJ, Floyd ZE, Arbour-Reily P, & Stephens JM (2001) Interferon-gamma-induced regulation of peroxisome proliferator-activated receptor gamma and STATs in adipocytes. *Journal of biological chemistry* 276(10):7062-7068.
195. Ron D, Brasier AR, McGehee RE, Jr., & Habener JF (1992) Tumor necrosis factor-induced reversal of adipocytic phenotype of 3T3-L1 cells is preceded by a loss of nuclear CCAAT/enhancer binding protein (C/EBP). *Journal of clinical investigation* 89(1):223-233.
196. Bergalet J, *et al.* (2011) HuR-mediated control of C/EBPbeta mRNA stability and translation in ALK-positive anaplastic large cell lymphomas. *Molecular cancer research* 9(4):485-496.
197. Chan CP, Mak TY, Chin KT, Ng IO, & Jin DY (2010) N-linked glycosylation is required for optimal proteolytic activation of membrane-bound transcription factor CREB-H. *Journal of cell science* 123(Pt 9):1438-1448.
198. Giandomenico V, Simonsson M, Gronroos E, & Ericsson J (2003) Coactivator-dependent acetylation stabilizes members of the SREBP family of transcription factors. *Molecular and cellular biology* 23(7):2587-2599.
199. Penumetcha M & Santanam N (2012) Nutraceuticals as ligands of PPARgamma. *PPAR research* 2012:858352.

200. Yellaturu CR, *et al.* (2009) Insulin enhances post-translational processing of nascent SREBP-1c by promoting its phosphorylation and association with COPII vesicles. *Journal of biological chemistry* 284(12):7518-7532.
201. Yellaturu CR, Deng X, Park EA, Raghow R, & Elam MB (2009) Insulin enhances the biogenesis of nuclear sterol regulatory element-binding protein (SREBP)-1c by posttranscriptional down-regulation of Insig-2A and its dissociation from SREBP cleavage-activating protein (SCAP).SREBP-1c complex. *Journal of biological chemistry* 284(46):31726-31734.
202. Tontonoz P, Hu E, & Spiegelman BM (1994) Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 79(7):1147-1156.
203. Carlson DB & Perdew GH (2002) A dynamic role for the Ah receptor in cell signaling? Insights from a diverse group of Ah receptor interacting proteins. *Journal of biochemical and molecular toxicology* 16(6):317-325.
204. Chen I, McDougal A, Wang F, & Safe S (1998) Aryl hydrocarbon receptor-mediated antiestrogenic and antitumorigenic activity of diindolylmethane. *Carcinogenesis* 19(9):1631-1639.
205. Banchereau J & Steinman RM (1998) Dendritic cells and the control of immunity. *Nature* 392(6673):245-252.
206. Coombes JL & Powrie F (2008) Dendritic cells in intestinal immune regulation. *Nature reviews immunology* 8(6):435-446.

207. Kelsall B (2008) Recent progress in understanding the phenotype and function of intestinal dendritic cells and macrophages. *Mucosal immunology* 1(6):460-469.
208. Powrie F, Read S, Mottet C, Uhlig H, & Maloy K (2003) Control of immune pathology by regulatory T cells. *Novartis foundation symposium* 252:92-98; discussion 98-105, 106-114.
209. Cerutti A (2008) The regulation of IgA class switching. *Nature reviews immunology* 8(6):421-434.
210. Iwasaki A & Kelsall BL (1999) Mucosal immunity and inflammation. I. Mucosal dendritic cells: their specialized role in initiating T cell responses. *American journal of physiology* 276(5 Pt 1):G1074-1078.
211. Musso G, Gambino R, & Cassader M (2011) Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. *Annual review of medicine* 62:361-380.
212. Abbas AK, Murphy KM, & Sher A (1996) Functional diversity of helper T lymphocytes. *Nature* 383(6603):787-793.
213. Mosmann TR & Coffman RL (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annual review of immunology* 7:145-173.
214. Bettelli E, Oukka M, & Kuchroo VK (2007) T(H)-17 cells in the circle of immunity and autoimmunity. *Nature immunology* 8(4):345-350.
215. Strober W (2009) The multifaceted influence of the mucosal microflora on mucosal dendritic cell responses. *Immunity* 31(3):377-388.

216. Mohan HM, *et al.* (2012) Molecular pathways: the role of NR4A orphan nuclear receptors in cancer. *Clinical cancer research* 18(12):3223-3228.
217. Hwang JA, Lee JA, Cheong SW, Youn HJ, & Park JH (2007) Benzo(a)pyrene inhibits growth and functional differentiation of mouse bone marrow-derived dendritic cells. Downregulation of RelB and eIF3 p170 by benzo(a)pyrene. *Toxicology letters* 169(1):82-90.
218. Ruby CE, Leid M, & Kerkvliet NI (2002) 2,3,7,8-Tetrachlorodibenzo-p-dioxin suppresses tumor necrosis factor-alpha and anti-CD40-induced activation of NF-kappaB/Rel in dendritic cells: p50 homodimer activation is not affected. *Molecular pharmacology* 62(3):722-728.
219. Mann J, Oakley F, Johnson PW, & Mann DA (2002) CD40 induces interleukin-6 gene transcription in dendritic cells: regulation by TRAF2, AP-1, NF-kappa B, AND CBF1. *Journal of biological chemistry* 277(19):17125-17138.
220. Yap IK, *et al.* (2010) Urinary metabolic phenotyping differentiates children with autism from their unaffected siblings and age-matched controls. *Journal of proteomic research* 9(6):2996-3004.
221. Sakurai K, *et al.* (2004) Bisphenol A affects glucose transport in mouse 3T3-F442A adipocytes. *British journal of pharmacology* 141(2):209-214.
222. Sargis RM, Johnson DN, Choudhury RA, & Brady MJ (2010) Environmental endocrine disruptors promote adipogenesis in the 3T3-L1 cell line through glucocorticoid receptor activation. *Obesity* 18(7):1283-1288.

223. Schwab AJ, *et al.* (2001) Hepatic uptake and metabolism of benzoate: a multiple indicator dilution, perfused rat liver study. *American journal of physiology. gastrointestinal and liver physiology* 280(6):G1124-1136.
224. Kim HB, *et al.* (2006) NFATc4 and ATF3 negatively regulate adiponectin gene expression in 3T3-L1 adipocytes. *Diabetes* 55(5):1342-1352.