

Metabolic Pathways of Type 2 Diabetes
Intersection of Genetics, Transcriptomics, and Metabolite Profiling

by

Christine Therese Ferrara

Department of Pharmacology and Cancer Biology
Duke University

Date: _____

Approved:

Christopher Newgard, Supervisor

Donald McDonnell

Deborah Muoio

Svati Shah

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Pharmacology and Cancer Biology in the Graduate School
of Duke University

2008

ABSTRACT

Metabolic Pathways of Type 2 Diabetes
Intersection of Genetics, Transcriptomics, and Metabolite Profiling

by

Christine Therese Ferrara

Department of Pharmacology and Cancer Biology
Duke University

Date: _____

Approved:

Christopher Newgard, Supervisor

Donald McDonnell

Deborah Muoio

Svati Shah

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Pharmacology and Cancer Biology in the Graduate School
of Duke University

2008

Copyright by
Christine Therese Ferrara
2008

Abstract

Type 2 diabetes is characterized by insufficient insulin secretion to maintain euglycemia in the setting of peripheral insulin resistance. The majority of insulin-resistant diabetics are obese, yet not all insulin-resistant obese individuals develop diabetes. This obesity/diabetes dichotomy suggests that genetic factors play a pivotal role in disease pathogenesis.

Gene mapping has identified genetic quantitative trait loci (QTL) influencing disease-related phenotypes. To uncover molecular pathways leading from genotype to clinical trait, we classify phenotypes in greater depth and identify QTL that influence combinations of physiological traits, mRNA levels, and metabolite abundance. A major challenge then becomes deciphering the causal interrelationships among correlated phenotypes.

In this dissertation, we develop methods for building causal direction into an undirected network by including QTLs for each phenotype. We then apply and validate these methods in an F2 intercross between the diabetes-resistant C57BL/6 *leptin*^{ob/ob} (B6^{ob/ob}) and the diabetes-susceptible BTBR *leptin*^{ob/ob} (BTBR^{ob/ob}) mouse strains. We show that genomic analysis can be integrated with liver transcriptional and metabolite profiling data to construct causal networks for specific metabolic processes in liver. This causal network construction led to the discovery of a pathway by which glutamine induces *Phosphoenolpyruvate carboxykinase* (*Pck1*) expression.

To investigate glutamine induction of *Pck1* in the context of diabetes, we perform mRNA expression analysis and metabolic profiling in liver of the parental strains. We find glutamine is decreased with obesity in both strains; in the diabetes-resistant B6 strain, liver *Pck1* expression parallels glutamine abundance, but in the diabetes-susceptible BTBR strain, *Pck1* is elevated with obesity. Follow-up *in vitro* studies indicate that α -ketoglutarate, which is elevated nearly two fold in the livers of BTBR relative to B6 mice *in vivo*, may mediate the glutamine effect. We hypothesize that hepatic *Pck1* is regulated by glutamine abundance in the liver of B6 animals, but in the presence of high α -ketoglutarate, *Pck1* becomes uncoupled from glutamine regulation in the livers of diabetes-susceptible BTBR mice.

Our method of causal network construction led to the discovery of glutamine induction of a key hepatic gluconeogenic enzyme, a pathway potentially disrupted in the diabetes-susceptible BTBR mouse. Future studies will include identifying hepatic mediators of the glutamine effect, and applying QTL-directed networks to multiple organs to ultimately define causal relationships between tissues involved in diabetes progression.

Contents

Abstract.....	iv
List of Tables	ix
List of Figures.....	x
List of Abbreviations	xiii
Acknowledgements.....	xviii
1. Introduction.....	1
1.1 Diabetes Pathogenesis.....	1
1.2 Hepatic Gluconeogenesis.....	3
1.3 Metabolites in the Context of Diabetes.....	7
Metabolic Dysregulation.....	7
Amino Acids as Signaling Molecules.....	9
1.4 Genetic Contribution to Diabetes Development.....	11
Gene Mapping and Expression QTL (eQTL).....	12
Metabolic QTL (mQTL).....	14
Genetic Networks of Causality	16
1.5 Interorgan Communication	17
1.6 Project Goals.....	20
2. Inferring Causal Phenotypes from Segregating Populations	36
2.1 Introduction.....	36
2.2 Network Construction.....	38
2.3 Results.....	44

Hundred Phenotypes Network	44
Cyclic Networks.....	45
Metabolite and Gene Expression Network	46
2.4 Discussion.....	47
3. Genetic Networks of Liver Metabolism Revealed by Integration of Metabolic and Transcriptional Profiling.....	58
3.1 Introduction.....	58
3.2 Materials and Methods.....	61
3.3 Results.....	69
Individuals of the F2 population segregate by physiological phenotype.....	69
Metabolites of similar function are highly correlated across the F2 population.....	70
Expected and novel correlations between transcripts and metabolites	72
Correlations and co-mapping of transcripts and metabolites produce causal network models.....	74
3.4 Discussion.....	77
4. Glutamine Regulation of Hepatic Transcription: Mechanism and Implications in Diabetes Phenotypes	95
4.1 Introduction.....	95
4.2 Materials and Methods.....	97
4.3 Results.....	102
Glutamine treatment elevates a unique subset of hepatic metabolites.....	102
Glutamine metabolism and glutamine-derived metabolites play a role in altered Pck1 expression	104
Glutamine regulation of Pck1 in the context of diabetic states	105

4.4 Discussion.....	107
5. The Role of Metabolites in Inter-tissue Communication.....	118
5.1 Introduction.....	118
5.2 Materials and Methods.....	121
5.3 Results.....	126
BTBR mice have specific defects in liver mitochondrial lipid oxidation pathways	126
The BTBR strain has increased hepatic microsomal and peroxisomal lipid oxidation	127
Dicarboxylate fatty acids and dicarboxylate derivatives alter insulin secretion.....	128
Dicarboxylic acids may act as signaling molecules in insulin secretion	130
5.4 Discussion.....	131
6. Future Directions and Concluding Remarks.....	144
6.1 Intrahepatic Studies.....	146
In vitro Mechanistic Studies	146
In vivo Significance of Hepatocyte Discoveries.....	151
6.2 Interorgan Studies	154
Broad Application of Causal Network Construction	154
In vivo Investigation of Regulatory Genetic Regions.....	159
6.3 Summary	162
References.....	169
Biography.....	195

List of Tables

Table 1. Metabolite Codes (Acyl-carnitines).....	83
Table 2. Metabolite Codes (Amino Acids and Organic Acids).....	84
Table 3. PCR Primer Sequences.....	85
Table 4. Amino acid correlation matrix.....	88
Table 5. Glx network correlations and mapping.....	91
Table 6. Glx network correlations of amino acid biochemical group.....	92

List of Figures

Figure 1. Hierarchy of systemic metabolic changes in obesity-related overnutrition.	23
Figure 2. Substrate cycles in the glycolytic/gluconeogenic pathways.....	24
Figure 3. Phosphoenolpyruvate carboxykinase, cytosolic (Pck1) gene promoter.....	25
Figure 4. Metabolic changes during catabolic states.	26
Figure 5. The Cori cycle and glucose/alanine cycle represent interorgan communication with the liver.	27
Figure 6. Glutamate integrates carbohydrate and amino acid metabolism.....	28
Figure 7. Strategy for QTL mapping.	29
Figure 8. C57BL/6 and BTBR mouse strain phenotypes.	30
Figure 9. Genome mapping reveals clinical QTLs (cQTLs).	31
Figure 10. Cis- and trans-regulation of transcript abundance.....	32
Figure 11. Possible relationships between QTLs, metabolites, and mRNA.....	33
Figure 12. Hepatic fatty acid oxidation pathways.	34
Figure 13. Cross-talk between fatty acid oxidation systems in times of metabolic overload.....	35
Figure 14. Spurious edges and partial correlations.....	51
Figure 15. Distinguishing direct and indirect effects of a common QTL.....	52
Figure 16. Directed acyclic graph (DAG).....	53
Figure 17. QDG performance in inferring correct direction of recovered edges.....	54
Figure 18. QDG performance of inferring direction in various patterns.....	55

Figure 19. Cyclic networks.....	56
Figure 20. Causal network from a segregating F2 sample.....	57
Figure 21. Clinical phenotypes of the leptin-deficient F2 population.	86
Figure 22. Heat map of correlations between liver metabolites.	87
Figure 23. Linkage hot spots for metabolic quantitative trait loci (mQTL).	89
Figure 24. Heat map of correlations between liver metabolites and select liver transcripts.	90
Figure 25. Glx network.....	93
Figure 26. Glutamine changes hepatic gene expression.	94
Figure 27. Glutamine alters abundance of liver metabolites.	113
Figure 28. Relationship between glutamine metabolism and Pck1 expression.....	114
Figure 29. Glutamine related metabolites alter hepatic gene expression.....	115
Figure 30. Induction of Pck1 expression by glutamine and α -ketoglutarate is dose dependent.....	116
Figure 31. Liver Pck1 mRNA, glx and α -ketoglutarate show strain-specific differences.	117
Figure 32. Liver monocarboxylic fatty acid oxidation differs by strain.....	137
Figure 33. Strain differences in hepatic microsomal and peroxisomal oxidation.....	138
Figure 34. Adipic acid alters insulin secretion in INS-1 cell line.....	139
Figure 35. Diethyl-3,4-pyridinedicarboxylate (DEPDC) enhances glucose stimulated insulin secretion.	140
Figure 36. DEPDC enhances insulin secretion from non-glucose secretagogues.	141
Figure 37. DEPDC and calcium-induced insulin secretion	142

Figure 38. DEPDC alters calcium flux dynamics in INS-1 cells.....	143
Figure 39. Expression QTL (eQTL) of urea cycle enzymes in liver tissue.	164
Figure 40. Liver mRNA for urea cycle enzymes in C/EBP α knockout mouse.	165
Figure 41. Gene expression network model shows strain-specific differences between B6 and BTBR mice.....	166
Figure 42. Modules can be correlated with physiological phenotypes.	167
Figure 43. Hypothetical co-mapping QTL strategy.....	168

List of Abbreviations

α KG, α -ketoglutarate**

AARE, amino acid response element

Acot8, acyl-CoA thioesterase 8

Actb, β -actin

Agxt, alanine:glyoxylate aminotransferase

AMPK, AMP activated protein kinase

Arg1, arginase, liver isoform

Asl, argininosuccinate lyase

Ass1, argininosuccinate synthetase, 1

Asx, asparagine+aspartate

B6, C57BL/6

BCH, 2-aminobicyclo[2.2.1]heptane-2-carboxylic acid

BF, Bayes factor

BMI, body mass index

C/EBP, CAAT enhancer binding protein

C3, propionyl carnitine

C5:1, tiglyl carnitine

C5's, isovaleryl, 3-methylbutyryl, or 2-methylbutyryl carnitine

C5-DC, glutaryl carnitine

C6:1-DC, hexenedioyl carnitine

C6-DC, adipoyl carnitine

C8:1-DC, octenedioyl carnitine

cADPR, cyclic ADP ribose

cAMP, cyclic AMP

CDKAL1, CDK5 regulatory subunit associated protein 1-like 1

ChIP-on-chip, chromatin immunoprecipitation-on-chip

CHOP, CAAT/enhancer binding protein homologous transcription factor
ChREBP, carbohydrate responsive element-binding protein
Cps1, carbamoyl-phosphate synthetase 1
cM, centimorgan
CPT-1, carnitine palmitoyltransferase 1
cQTL, clinical QTL
CRE, cAMP response element
CREB, cAMP responsive element binding protein
Cyp4a10, cytochrome P450, family 4, subfamily a, polypeptide 10
Cyp4a14, cytochrome P450, family 4, subfamily a, polypeptide 14
DAG, directed acyclic graph
db/db, leptin receptor gene mutation
DEPDC, diethyl-3,4-pyridinedicarboxylic acid
DG, directed graph
DMM, dimethyl malate
eIF-2 α , eukaryotic translation initiation factor 2A
EGF, epidermal growth factor
eQTL, expression QTL
EST, expressed sequence tag
Fbp1, fructose biphosphatase 1
FDR, false discovery rate
FGF, fibroblast growth factor
FXR, farnesoid X receptor
FTO, fat mass and obesity associated
G6pc, glucose 6 phosphatase, catalytic
GC/MS, gas chromatography/mass spectrometry
Gck, glucokinase
GGM, Gaussian graphical model

Gls2, glutaminase 2 (liver, mitochondria)
Glud1, glutamate dehydrogenase 1
Glx, glutamine+glutamate
GO term, gene ontology term
Got1, glutamate oxaloacetate transaminase 1, soluble
Got2, glutamate oxaloacetate transaminase 2, mitochondrial
Gpt1, glutamic pyruvic transaminase 1, soluble
Gpt2, glutamic pyruvate transaminase 2
GRU, glucocorticoid response unit
HLA-DR4, human leukocyte antigen DR4
Hlcs, holocarboxylase synthetase
HHEX-IDE, hematopoietically expressed homeobox/insulin degrading enzyme
HNF-1, hepatocyte nuclear factor 1
IL6, interleukin 6
INF γ , interferon gamma
IRS-1, insulin receptor substrate 1
Ivd, isovaleryl-CoA dehydrogenase
K_{ATP}, ATP-sensitive potassium channel
KIC, α -ketoisocaproate
LOD, logarithm of the odds
LXR, nuclear receptor subfamily 1, group H, member 3
Mb, megabase
Mccc1, methylcrotonoyl-Coenzyme A carboxylase 1 (alpha)
MODY, maturity onset diabetes of the young
MPP, maximum posterior probability
mQTL, metabolic QTL
MS/MS, tandem mass spectrometry
mTOR, mammalian target of rapamycin

NAADP, nicotinic acid adenine dinucleotide phosphate
NAG, N-acetylglutamate
NF-1, nuclear factor 1
NSRE, nutrient sensing response element
Oat, ornithine aminotransferase
ob/ob, leptin deficiency
ODG, ordered dependency graph
Otc, ornithine transcarbamylase
PC1, first principal component
Pck1, phosphoenolpyruvate carboxykinase, cytosolic
Pck2, phosphoenolpyruvate carboxykinase, mitochondrial
Pcx, phosphoenolpyruvate carboxylase
PEP, phosphoenolpyruvate
PFK2/FBPase2, 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2
Pfk1, phosphofructokinase, liver, B-type
PKA, cAMP-dependent protein kinase
Pklr, pyruvate kinase, liver, red blood cell
PNDM, permanent neonatal diabetes mellitus
PPAR, peroxisome proliferator activated receptor
Ppara, peroxisome proliferator activated receptor alpha
Ppargc1a, peroxisome proliferative activated receptor, gamma, coactivator 1 alpha
QDG, QTL-directed dependency graph
QTL, quantitative trait loci
RMA, robust multichip average
SEM, structural equation models
Slc1a2, solute carrier family 1 (glial high affinity glutamate transporter), member 2
Slc38a3, solute carrier family 38, member 3
STEDMAN, the Study of the Effects of Diet on Metabolism and Nutrition

TCA, tricarboxylic acid cycle

TCF7L2, transcription factor 7-like 2

UDG, undirected dependency graph

UTR, untranslated region

**Complete metabolite codes are provided in Table 1 (acyl-carnitines) and Table 2 (amino acids and organic acids)

Acknowledgements

I have had the fortunate and unique opportunity to work with two amazing research advisors. Despite their different backgrounds and expertise, Christopher Newgard and Alan Attie complement each other as scientists and I am fortunate to have been mentored by both. I cannot thank Chris enough for his guidance over the past several years. Chris has supported me both inside and outside the lab to pursue my talents and passions. He has been there in times of celebration, and has encouraged me in times of frustration. None of this would have been possible had Alan Attie not taken me under his wing. He has been a wonderful mentor and provided the never ending daily enthusiasm and awe for science. At our monthly advising conference calls between Madison and Durham, we would often joke more than one advisor was required to get me through graduate school, but I honestly cannot thank the two of them enough for providing me with this collaborative experience.

With a graduate school career involving two institutions, I don't know how to begin to thank all of the people who have touched my life these past several years. I would like to thank Hans Hohmeier and Mette Jensen, who showed unending patience and were always there to answer my questions. Mark Keller was likewise there in Wisconsin when I arrived, giving me research advice and support when I was discouraged. I must thank Daniel Blasiolo who taught me everything I know about hepatocyte isolations; without his help, I would have been greatly handicapped. I am also

grateful to have worked with statisticians Elias Chaibub and Ping Wang, two individuals who are still brilliant at 1:00 am in responses to last minute e-mails. James Bain, Brett Wenner, Robert Stevens, and Olga Ilkayeva also deserve recognition for their generous help with the metabolic profiling.

There are too many names to list those who have made each day enjoyable. I want to thank Jeremy Lavine who has seen me through the ups and downs of both research and life events. The cave (and graduate school) would not have been the same without him. I also want to thank my best friend Nicole Mazzola who, in spite of being physically in California, was never anywhere other than right beside me.

Lastly, I thank my parents. Mom and Dad, I truly do not know how to thank you for the amazing opportunities you have given me. I thank you for pushing me to my full potential and at the same time accepting me for who I am. Robert and Becky, I am so blessed to have you two in my life and I could not wish for more supportive and loving step-parents. I am so fortunate to have so many people standing behind me, and I thank you all.

1. Introduction

1.1 Diabetes Pathogenesis

Type 2 diabetes mellitus affects over 5% of the US population, making it the most common of all metabolic disorders [1]. The pathogenesis of type 2 diabetes often involves an initial peripheral tissue insulin resistance, which increases the amount of insulin required to maintain euglycemia. The pancreatic β -cell acutely increases insulin secretion to meet this demand; however, if the insulin requirement is chronically elevated, islet function and mass are decreased, eventually leading to insulinopenia [2,3]. This underlying β -cell decompensation manifests clinically as elevated fasting and postprandial blood glucose levels, diagnostic criteria for diabetes [4,5].

In humans, diabetes is often correlated with obesity, leading to a long-standing hypothesis that insulin resistance is a consequence of overnutrition and elevated dietary fatty acids [6]. Chronic metabolic overload has a detrimental effect on whole body metabolism, and there is increasing evidence that the liver and adipose play a “causal” role to drive this metabolic disequilibrium (Figure 1). The adipose acts as a storage depot for excess fatty acids and also serves as an endocrine organ, releasing adipokines and inflammatory signaling molecules. The liver plays a pivotal role in whole body energy homeostasis by handling all fuel types: the liver is one of two organs capable of glucose production through gluconeogenesis; it manages fat load by oxidizing fatty acids for energy or storing excess triglycerides as fuel; and the liver is involved in overall nitrogen balance, utilizing tissue-derived amino acids as gluconeogenic precursors and eliminating

ammonia as urea[7]. Because the liver coordinates fuel usage and is thus in a position to drive metabolic dysregulation, we have focused on the liver as our organ of study.

The ability of the liver to organize whole body fuel metabolism is altered significantly by excessive nutrients, especially by elevated fatty acids. It has been shown that liver metabolism is altered at the transcriptional level, indicating that lipid metabolites can affect gene expression. Fatty acids can regulate transcription via signaling pathways that converge on second messengers, such as protein kinase C, or by activating transcription factors including, but not limited to PPARs, LXR, and FXR [8-14]. Through transcriptional changes in key metabolic pathways, chronic overfeeding impairs the liver's ability to effectively oxidize fatty acids and increases triglyceride production, thus elevating both intrahepatic and circulating levels of fatty acid derivatives[6]. These accumulating intrahepatic lipid-derived metabolites, including diacylglycerols and ceramides, not only exacerbate altered liver gene expression and metabolic pathways, but also contribute to hepatic insulin resistance[15-22]. The loss of insulin sensitivity further prevents the liver from sensing and managing nutrient overload.

The “reactive” tissues, importantly the skeletal muscle and pancreatic islets, are now exposed to lipid-derived metabolites and inflammatory mediators produced in the liver and adipose not only from nutritional overload, but also tissue malfunction. In the skeletal muscle, the mechanism of fatty acid induced insulin resistance is still emerging [6]. Elevated circulating plasma fatty acids can alter glucose transport into skeletal muscles, and intramuscular accumulation of lipid-derived metabolites, including cytosolic and mitochondrial metabolites, has been correlated with decreased insulin

sensitivity [6,7,23]. While the effects on the skeletal muscle are important, clinical type 2 diabetes is defined by hyperglycemia, which only develops when the pancreatic islets no longer can compensate for increased insulin demand. In the short-term, free fatty acids actually stimulate insulin secretion, but after prolonged elevation, these metabolites decrease β -cell function and survival [6,7,23]. It is this β -cell decompensation that leads to progression from metabolic syndrome, initiated by driving organs such as the liver and adipose, to hyperglycemia and type 2 diabetes.

1.2 Hepatic Gluconeogenesis

Elevated fasting and postprandial blood glucose levels are considered to reflect β -cell decompensation, and are thus diagnostic of type 2 diabetes. Islet insulin secretion is required to maintain euglycemia following a meal, but growing evidence suggests that increased hepatic glucose output is the main cause of fasting hyperglycemia [24]. To comprehensively understand pathways of diabetes pathology, it is necessary to identify the factors that cause increases in glucose production.

Gluconeogenesis is defined as the production of new glucose, and occurs in the liver and the kidney. Synthesis of glucose from three and four carbon precursors is essentially a reversal of the reactions of glycolysis; however, there are three enzymatic steps of glycolysis that proceed with a large negative free energy, and are in effect “irreversible.” These three glycolytic steps are carried out by: 1) Pyruvate kinase (Pklr), which catalyzes the conversion of phosphoenolpyruvate (PEP) to pyruvate and ATP; 2)

Phosphofructokinase-1 (Pfk1), converting fructose-6-phosphate to fructose-1,6-bisphosphate and ADP; and 3) Hexokinase/Glucokinase (Gck) responsible for the first step of glycolysis in which glucose is phosphorylated by ATP to produce glucose-6-phosphate.

Unique enzymes have evolved to allow gluconeogenesis to bypass the irreversible reactions of glycolysis (Figure 2). The first bypass converts glucose-6-phosphate to glucose through the action of Glucose-6-phosphatase (G6pc), reversing the first step of glycolysis catalyzed by Hexokinase/Glucokinase (glucose cycle). Secondly, fructose-1,6-bisphosphate is converted to fructose-6-phosphate by Fructose-1,6-bisphosphatase (Fbp1), thereby reversing the Phosphofructokinase reaction, a rate limiting step of glycolysis (fructose cycle). Finally, the conversion of pyruvate to PEP, the reversal of the Pyruvate kinase reaction, requires the action of two mitochondrial enzymes (pyruvate/PEP cycle). The first enzyme is Pyruvate carboxylase (Pcx), which utilizes ATP and bicarbonate to convert pyruvate to oxaloacetate. Since oxaloacetate cannot be transported to the cytosol, it is converted to PEP by mitochondrial Phosphoenolpyruvate carboxykinase (Pck2) or to aspartate or malate which can exit the mitochondria. Cytosolic aspartate and malate are converted back to oxaloacetate via transamination or oxidation, respectively, prior to conversion to cytosolic PEP by cytosolic Phosphoenolpyruvate carboxykinase (Pck1).

Short and long-term regulation of these three key gluconeogenic steps are required to maintain euglycemia[25]. The short-term regulation of the fructose cycle has been well characterized. Glucagon, acting through cAMP and PKA, alters

phosphorylation of the bifunctional enzyme PFK2/FBPase2 which serves as both a kinase and a phosphatase. PKA-phosphorylated PFK2/FBPase2 promotes gluconeogenesis by dephosphorylating fructose-2-6-bisphosphate, a negative allosteric regulator of Fbp1. Insulin activates protein phosphatases to dephosphorylate PFK2/FBPase2; PFK2/FBPase2 then acts as a kinase and increases abundance of fructose-2-6-bisphosphate, thus promoting glycolysis. Fructose-1,6-bisphosphatase is also regulated by the ATP:ADP ratio. When this is high, gluconeogenesis can proceed maximally [26-28].

The pyruvate/PEP equilibrium is also regulated by glucagon which, via PKA, phosphorylates Pckr to inhibit endogenous activity and prevent glycolysis. In the phosphorylated state, Pckr is also more susceptible to allosteric inhibition by alanine and ATP. Just as the fructose cycle favors glycolysis when the ATP:ADP ratio falls, a reduction in energy levels inhibits Pck and Pck activities. Flux through the pyruvate/PEP cycle and the fructose cycle are further interconnected by fructose-1,6-bisphosphate, a product of glycolysis, which activates Pckr [26,27,29,30].

Long-term regulation of glycolysis and gluconeogenesis occurs via transcriptional changes of pathway enzymes. In general, glucagon (via cAMP) and insulin act reciprocally on expression; cAMP reduces expression of glycolytic enzymes *Gck*, *Pfkf*, and *Pckr*, while insulin increases transcript levels. The reverse is true for the gluconeogenic enzymes, where cAMP increases transcription and insulin inhibits gene expression [26,27].

Under physiologic conditions, the rate of gluconeogenesis is determined by flux through the pyruvate/PEP cycle. The cytosolic form of Phosphoenolpyruvate carboxykinase (Pck1) is considered to be the pace-setting enzyme and has such a pronounced effect on gluconeogenic flux that it has become a marker gene for hepatic glucose output[26,31]. *Pck1* transcription is under complex regulation by hormonal, environmental, and nutritional inputs, which act at specific regions of the Pck1 promoter (Figure 3) [32]. Region one contains the NF-1 site to maintain basal transcription, and the CRE which mediates cAMP-stimulated transcription by binding transcription factors including CREB and CAAT enhancer binding proteins [33]. The second region of the Pck1 promoter binds proteins that interact with transcription factors binding to region one. Importantly, this region contains a binding site for HNF-1, essential for basal gene expression, and the C/EBP binding site, required for the full response to cAMP. The glucocorticoid response element (GRU) in region three coordinates hormonal regulation by glucocorticoids and insulin. The GRU also interacts with the cAMP response unit to mediate the cAMP control of gene transcription. [32,34-50].

Transcriptional regulation of Pck1 is complex, and the resultant mRNA abundance reflects inputs from nutritional, hormonal, and environmental factors. Investigation of nutrient control of *Pck1* expression, however, has largely led to the conclusion that nutrient-induced hormonal (insulin/glucagon) signals are responsible for altered transcription [31,33,51]. While this may address how *Pck1* expression is altered in a diabetic state where insulin levels are dysregulated, diet-induced metabolic perturbations precede diabetes onset and alter gene expression in primary tissues such as

the liver. The islet is downstream of these events, and thus insulin secretion and ultimate β -cell decompensation marking type 2 diabetes are reactive to previous gene changes in other organs [6,7]. It is therefore important to isolate the metabolite-driven mechanism of transcriptional regulation in the liver.

1.3 Metabolites in the Context of Diabetes

Metabolic Dysregulation

Metabolic dysregulation in type 2 diabetes involves all three major fuel types. As peripheral insulin resistance prevents major organs from sensing nutrient load appropriately, the body perceives that it is undergoing starvation. In the fasted state, the majority of organs can utilize stored glycogen (liver and muscle) and lipid (adipose) for energy; the central nervous system, however, relies heavily on blood glucose. Peripheral organs therefore possess several regulatory mechanisms to use alternate fuel sources in times of need, thus “sparing” glucose for the brain (Figure 4).

In starvation states, the liver increases hepatic glucose output and depends on peripheral tissues for gluconeogenic precursors. Insulin resistant skeletal muscle, unable to adequately take up circulating glucose, instead relies on stored glycogen for energy. Intramuscular glucose is metabolized through glycolysis, but rather than being oxidized completely via the TCA cycle, pyruvate is converted into alanine or lactate and released. The muscle also breaks down its own proteins and releases amino acids, particularly alanine and glutamine. These skeletal muscle protein-derived substrates are transported to

the liver where they can be used for gluconeogenesis and glucose production (Figure 5) [52,53].

Alanine delivered to the liver is transaminated by alanine aminotransferase, converting alanine and α -ketoglutarate into pyruvate and glutamate. The carbon skeleton, now in the form of pyruvate, is converted into glucose through the multi-step gluconeogenic pathway while the amine nitrogen is eliminated via the urea cycle[54]. The urea cycle and gluconeogenesis are highly integrated, and glutamate is at the center of these metabolic events. Glutamate serves as substrate for the synthesis of N-acetylglutamate, an essential allosteric activator of Carbamoyl phosphate synthetase 1 (Cps1) in the urea cycle. Liver glutamate is essential for disposal of amino acids even in the fed state, as many of the dietary complement of amino acids are converted to glutamate via transamination prior to their utilization [55-58] (Figure 6).

Over the past century, the Western diet has evolved such that nearly 40% of caloric intake is in the form of fat, due in large part to fast food and high-fat, processed meat. The Study of the Effects of Diet on Metabolism and Nutrition (STEDMAN) weight loss project showed, not surprisingly, that plasma free fatty acids were elevated in obese humans (BMI:30-50) compared to lean participants (BMI:18.5-24.9). Interestingly, a set of eight plasma amino acids, were also significantly elevated in the obese subjects relative to the lean group, including all three branched-chain amino acids (leucine, isoleucine, valine)[59]. In a study comparing the Zucker *db/db* rat to their lean counterparts, branched-chain amino acids were also elevated both in the liver and plasma[60].

This dissertation will describe the metabolic profile of a mouse model of diabetes. As is described in Chapters 4 and 5, the hepatic metabolite profiles of the C57BL/6 *leptin*^{ob/ob} (B6^{ob/ob}, diabetes-resistant) mouse and BTBR *leptin*^{ob/ob} (BTBR^{ob/ob}, diabetic) mouse parallels that of the STEDMAN human study. The hepatic metabolites of the BTBR^{ob/ob} mouse mimic the obese human profile with elevated amino acid species (notably the branched-chain amino acids), as well as increased fatty acid derivatives, compared to the B6^{ob/ob}. Since both the BTBR^{ob/ob} and the B6^{ob/ob} are insulin resistant, the differences in hepatic metabolite profile and susceptibility to hyperglycemia depend on the genetic differences between the strains. These mouse strains therefore comprise a model in which to study the genetic determinants of key metabolic changes that contribute to diabetes pathogenesis, as well as the downstream effects of these metabolic changes.

Amino Acids as Signaling Molecules

Branched-chain amino acids and alanine are of particular importance in catabolic states such as diabetes because they are released from muscle during protein breakdown. Plasma glutamine is the most abundant amino acid in the plasma, and its levels are altered by starvation states [54,61-66]. Since plasma amino acid composition can reflect either pathological or nutritional status, it is important to consider the signaling role that these metabolites have on target tissues. In recent years, it has been shown that amino acid availability alters expression of many genes, including amino acid transporters and transcription factors[67,68]. Amino acids regulate multiple steps of transcription and translation, from chromatin remodeling, to stabilization of spliced RNA, to protein

translation[69], yet there is limited information on molecular signaling pathways of such amino acid regulation [11,12,70-73].

The majority of studies of amino acid regulation of gene expression focus on transcriptional responses to amino acid deprivation. Several genes that are altered by this depletion contain amino acid response elements (AARE) or nutrient sensing response elements (NSRE) in their promoter regions. The CHOP transcription factor is one such gene that responds to amino acid deficit (specifically leucine scarcity). Amino acid limitation increases uncharged tRNA, which activates Gcn2p. This kinase phosphorylates eIF-2 α of the translational machinery, making eIF-2 α inactive. Although the majority of protein translation is halted, there is an increase in Gcn4 protein which activates a CHOP transcription factor ATF4[69]. Another characterized pathway of amino acid sensing is the mTOR cascade, which regulates protein translation through 4EBPs and 6sk1 in response to increased nutrients[74,75].

The aforementioned studies reveal limited information about the mechanism by which amino acid excess can alter gene expression. Glutamine is an example of such an amino acid that regulates transcription of hepatic enzymes when it is supplemented. Importantly, glutamine increases expression of *Argininosuccinate synthetase (Ass1)* and *Arginase (Arg1)*, members of the urea cycle [56-58,76]. In some cases the ability of glutamine to regulate transcription may be dependent on the presence of insulin or other metabolites, such as additional amino acids or glucose[77,78]. Glutamine is not only at the crossroads of amino acid and glucose metabolism, but it is also the most abundant amino acid in plasma. Given that glutamine has both transcriptional and allosteric

regulatory roles, and that plasma levels reflect nutrient status, we have focused our studies on this amino acid (glx: glutamine+glutamate). Revealing the mechanism of glutamine-induced transcription will enhance our understanding of glucose and amino acid metabolism, two pathways that are often disrupted in states of insulin resistance and diabetes.

1.4 Genetic Contribution to Diabetes Development

Lifestyle-induced hyperlipidemia and weight gain are experienced by many individuals as a function of aging. There is increasing evidence that obesity-related overnutrition leads to insulin resistance, but the link between metabolic overload and type 2 diabetes is less direct. Although the majority of insulin resistant type 2 diabetics are obese, only a subset of insulin-resistant and obese subjects develop diabetes[79]. This obesity/diabetes dichotomy indicates that while environmental factors do play a causal role in diabetes, genetic differences are also very important[80].

The genetic element of type 2 diabetes has been extensively studied. Maturity onset diabetes of the young (MODY) and permanent neonatal diabetes mellitus (PNDM) are among the very few cases of diabetes in which a single gene disorder is responsible for disease development. Through genome wide association scans, there have been 18 genetic loci linked to diabetes development [81]. Of the risk alleles identified, follow up studies have confirmed that CDKAL1 and HHEX-IDE alter insulin secretion, while FTO is predictive of increased adiposity[82]. The strongest genetic contributor, TCF7L2, was

discovered by the DeCode group. This locus encodes a transcription factor that is thought to be involved in the Wnt signaling pathway and in the production of incretin GLP-1[81,83].

While there have been advances in unraveling the genetic aspect of diabetes, the genes identified account for less than 5% of type 2 diabetes. The molecular pathways leading from the identified genetic loci to diabetes pathology remain largely unknown, and hypotheses mainly include altered pathways of insulin secretion. There is a remarkable scarcity of insulin resistance genes that have been recovered by genome wide association scans, likely due to the strong environmental factors that contribute to such a state [84]. Insulin resistance is less heritable than insulin secretion; a further complication arises in uncovering genetic aspects of type 2 diabetes is the fact that this disease is polygenic, and individual genes show only a modest effect sizes [81,85-90].

Gene Mapping and Expression QTL (eQTL)

A powerful way to establish a causal link between genetic loci and disease phenotypes is gene mapping [91,92]. Gene mapping requires an F2 population, a scrambled genetic mix of the two parental organisms resulting from random recombination events. As a result of this recombination, no two individuals of an F2 population are identical, and therefore segregate by genotype and phenotype. Genotypes are measured at markers throughout the genome to determine which parental strain contributed a specific genetic region in an F2 individual[93]. We then correlate genotype at these markers with phenotype, at each marker asking, “If the genotype was changed from A to B, would the phenotype be significantly different?” (Figure 7). A logarithm of

the odds score (LOD score) represents the strength of this correlation, and one can decipher distinct regions (quantitative trait loci, QTL) throughout the genome influencing the phenotype considered. A LOD score above a determined threshold shows significant linkage between the phenotype and that genetic marker; the “peak marker” is defined as the genetic region with the most significant correlation to the phenotype (highest LOD score).

We have the tools to explore the genetic contribution to the diabetes/obesity dichotomy. At 10 weeks of age, the leptin-deficient BTBR^{ob/ob} mouse has severe diabetes. Compared to the diabetes-resistant B6^{ob/ob} mouse, the BTBR^{ob/ob} mouse has elevated body weight, fasting blood glucose, and plasma triglycerides. The BTBR^{ob/ob} males at 10 weeks of age have insufficient plasma insulin compared to the male B6^{ob/ob} (Figure 8). To investigate the genetic contribution underlying the parental phenotypic differences, we can generate an F2 population that will segregate genetically and also contain individuals with a range of diabetes clinical phenotypes (body weight, plasma insulin, fasting blood glucose). By studying this population through gene mapping, we will be able to explore the genetic contributions to the diabetes clinical traits seen in the BTBR and absent in the B6 strains[94].

We have shown that we can map clinical traits (fasting glucose, plasma insulin, body weight) to genetic loci (clinical QTLs, cQTLs) in this F2 population, which confirms that disease phenotypes are heritable and influenced by genetic factors (Figure 9) [94-96]. The central dogma dictates that DNA alters RNA which has downstream effects, and it seems logical to ask if transcripts can also be mapped to genetic regions.

Gene expression can be mapped to genetic loci (expression QTLs, eQTLs), indicating that the definition of “phenotype” need not be limited to clinical or physiological traits. When eQTLs co-localize with cQTLs, one can hypothesize that they share a common regulator and propose potential pathways from genotype to phenotype through specific transcripts [97].

Expression QTLs are complicated by the fact that transcripts can map in “*trans*” or in “*cis*” fashion (Figure 10). A transcript is said to map in “*trans*” if the eQTL is a region of the genome distinct from the physical location of the gene itself. For example, consider that chromosome 2 encodes a transcription factor regulating gene n on chromosome 5. The expression of gene n would therefore be influenced by chromosome 2, and expression levels of gene n would map to an eQTL on chromosome 2, in spite of gene n being physically located on chromosome 5. It is conceivable that several transcripts will map in *trans* to the same eQTL if a transcription factor alters the expression of multiple genes. If, however, expression of gene n mapped in “*cis*” to where it is physically located on chromosome 5, that would indicate that gene n itself has changed (e.g., mutation in the promoter region) leading to a change in gene n expression.

Metabolic QTL (mQTL)

Transcript mapping does provide one avenue for understanding genotype to phenotype relationships, but gene expression does not always parallel physiological outcome. Clinical phenotypes such as body weight and blood glucose levels are far downstream of many biological inputs, and gene expression is often not predictive of biological change. There are many ways that gene expression could be uncoupled from

the phenotype, for example by failure of mRNA translation into a functional protein[98-100]. In a disease such as type 2 diabetes, we know *a priori* that environmental factors such as diet and inactivity interact with genetic factors, pushing a subset of individuals to develop hyperglycemia while others remain unaffected[2].

The potential for gene expression to be dissociated from physiological phenotype makes it necessary to incorporate additional intermediates to fully uncover the molecular pathway from genotype to disease state. Metabolites are considered a response to genetic and environmental factors, much like a clinical phenotype [101]. Mouse models of neurological disease, muscular dystrophy, and diabetes are characterized by distinct metabolic signatures reflecting pathology [102]. Of equal importance is the fact that a metabolic profile may help to define a population at risk for disease development. A human study of familial early-onset cardiovascular disease by Shah, *et al.* showed specific heritable groups of plasma metabolites not only in diseased individuals, but also in unaffected family members with a genetic propensity for this disease[103]. Identifying specific “predictive” metabolic signatures will allow for pharmacological or lifestyle intervention before disease onset.

Metabolite levels are particularly relevant for type 2 diabetes, characterized by altered energy metabolism and often preceded by a “metabolic syndrome”. Metabolites therefore serve as important predictive intermediates that link gene expression and clinical measurements. The recent advent of technologies for quantitative measurement of large clusters of metabolites (sometimes termed “metabolomics”) allow metabolites to

be mapped to metabolic QTLs (mQTLs) to aid in understanding of pathways leading from genotype to phenotype[104].

Genetic Networks of Causality

When considering correlations and co-mapping of clinical traits, metabolite levels, and gene expression, the limitation remains that the only one-way arrow of causality that can be confirmed is from genotype to phenotype, whether that phenotype be mRNA, metabolite levels, or clinical traits. There are three ways in which a co-mapping and correlated transcript and metabolite/clinical trait could be related[97] (Figure 11). In the reactive model, the metabolite is downstream of both the genetic locus and mRNA, for example an increase in *Glucokinase* expression results in increased glucose-6-phosphate. The causal model suggests that the metabolite/clinical trait is upstream of the RNA, which has been seen for example in glucose regulation of gene expression via the ChREBP transcription factor[105,106]. Finally it is possible that the correlated and co-mapping phenotypes are independent of one another, yet still dependent on the genetic locus. An example of this would be a genetic locus encoded a transcription factor that regulated both lipid oxidation and gluconeogenesis expression. An acyl-carnitine (product of lipid oxidation) and *Pyruvate kinase* mRNA would still be correlated and co-map, yet would not be causal or reactive to each other.

In past years, metabolite/transcript connections could only be described as a correlation depicted in association networks. However, with the advancements in transcriptomics and metabolomics, we now have defined eQTL and mQTL and identified genetic regions as drivers for downstream phenotypes. By using QTL of metabolite and

transcript nodes in the association networks, it is possible to infer directionality between two phenotypes. In a *causal* network, one can generate data-driven hypotheses about the entire pathway in which a genetic region can influence a biological response.

The causal versus reactive role of metabolites is important both for therapy and diagnosis of type 2 diabetes. As the signaling role of metabolites is realized, their position as “downstream” of gene expression is being re-evaluated, and it is important to integrate this information. If a metabolite is found to be causal for gene expression change or a clinical trait, it can serve as a therapeutic target for pharmacological intervention. If a metabolite is reactive to a diabetic phenotype such as insulin resistance, it can be used as a biomarker for diagnosis, disease development, or responsiveness to treatment.

1.5 Interorgan Communication

The pathogenesis of type 2 diabetes is clearly not limited to one organ. Insulin resistance of the liver and other peripheral tissues is both causal and reactive to metabolic disturbances; demands for increased insulin alter β -cell function; lifestyle and dietary habits influence multiple aspects of disease progression. The ultimate progression from insulin resistance to diabetes diagnosis requires β -cell decompensation, such that secreted insulin is insufficient to meet the increasing demand. While the focus of our studies has been on the understanding the relationships between intrahepatic metabolites and hepatic gene transcription in diabetic and non-diabetic states, the liver and the islet are in close communication. It is therefore important to consider metabolic analytes secreted from the

liver in response to metabolic overload and how these signals could influence insulin output [6,7]

Insulin secretion is regulated in a biphasic manner by nutritional and hormonal signals, but the primary regulator is glucose. Glucose metabolism in the normal β -cell increases the ATP:ADP ratio, which closes the K_{ATP} channels. This closure causes membrane depolarization, which then activates voltage-gated calcium channels. Calcium-mediated signaling pathways ultimately cause the exocytosis of insulin granules. There is also a K_{ATP} independent component of insulin secretion which may play a role in the second sustained phase of insulin secretion. Other secretagogues such as free fatty acids and amino acids have a small effect on basal insulin secretion, but become significant potentiators of insulin secretion as glucose concentrations rise above a threshold of 6-8mM in the blood in the postprandial state.

Metabolism of glucose and other fuel secretagogues produces many downstream byproducts that may not necessarily change the ATP:ADP ratio. As such, these metabolite products may contribute to the less well characterized K_{ATP} independent pathway of sustained insulin release. The liver integrates metabolism of carbohydrates, proteins, and lipids, and to understand the liver/islet communication in diabetes pathogenesis, we have therefore explored the effects of select non-glucose metabolites on insulin secretion [6,7,107,108].

In the short-term fatty acids enhance glucose-stimulated insulin secretion, while chronic exposure suppresses islet function[109,110]. Long-chain fatty acids make up the bulk of lipid content in the Western diet, but in the liver, these are metabolized into their

medium and short-chain acyl-carnitine derivatives. Under normal conditions, the liver metabolizes fatty acids via mitochondrial β -oxidation, producing monocarboxylic acid derivatives that are used for fuel. When mitochondrial oxidation capacity is exceeded fatty acids are diverted to microsomal ω -oxidation, producing dicarboxylic acids that are subsequently degraded in the peroxisome [111-117] (Figure 12). This ω -oxidation is dramatically increased in times of elevated mobilization of free fatty acids, such as during starvation and insulin resistance, or when consuming a high-fat diet [118]. Thus in overnutrition, the liver metabolizes excess fatty acids both by conventional mitochondrial β -oxidation, and “overspill” pathways, producing short and medium-chain acyl-carnitines and dicarboxylates. Importantly, dicarboxylate products have a feedback effect on hepatic oxidation pathways and can alter secretion of liver-derived products (Figure 13) [113,119,120].

The STEDMAN study certainly found that there was a difference in circulating plasma free fatty acids in obese versus lean subjects, but also found that obese subjects had significantly elevated levels of short and medium-chain acyl-carnitines[59]. In a study comparing control versus diabetic rat livers, ω -oxidation was drastically increased in the diabetic cohort[121]. The B6 diabetes-resistant and BTBR diabetic predisposed mice also show a strain-dependent difference in their hepatic fatty acid profiles, specifically in the shorter monocarboxylic and dicarboxylic acyl-carnitines (Chapter 5). Several of these specific monocarboxylic and dicarboxylic fatty acids have been implicated as serving a role in insulin secretion[122-124], placing these metabolites as potential candidate interorgan “messengers” between the liver and islet.

A dicarboxylic acid is an organic compound with two carboxylic acid functional groups. As the chain length of these dicarboxylates decrease, the analytes begin to resemble members of the TCA cycle: the five-carbon dicarboxylate (glutarate) is structurally similar to α -ketoglutarate; the four carbon dicarboxylate is succinate itself. Much recent work has been done on understanding the contribution of anaplerosis (generation of new TCA intermediates) in fuel-regulated insulin secretion. These newly formed TCA intermediates are cycled between the mitochondria and cytosol, and disruption in this cycling has been implicated in diabetes pathogenesis[125]. The fact that these intermediates are exported from the mitochondria suggests a signaling role independent of their ability to generate ATP [108,125-133]. Short and medium-chain dicarboxylates resemble these TCA intermediates, and we therefore consider their role in insulin secretion via anaplerosis and as extramitochondrial signaling molecules.

1.6 Project Goals

At the time I began my dissertation, our laboratories had established a B6^{ob/ob}/BTBR^{ob/ob} derived leptin-deficient F2 population that segregated by diabetes phenotypes (blood glucose, plasma insulin, body weight, and liver triglycerides). We had demonstrated strong correlations between these clinical traits and liver transcripts and hepatic metabolites. While QTL had been identified for both liver mRNA (eQTL) and these clinical traits (cQTL), thus revealing broad genomic regions where potential master

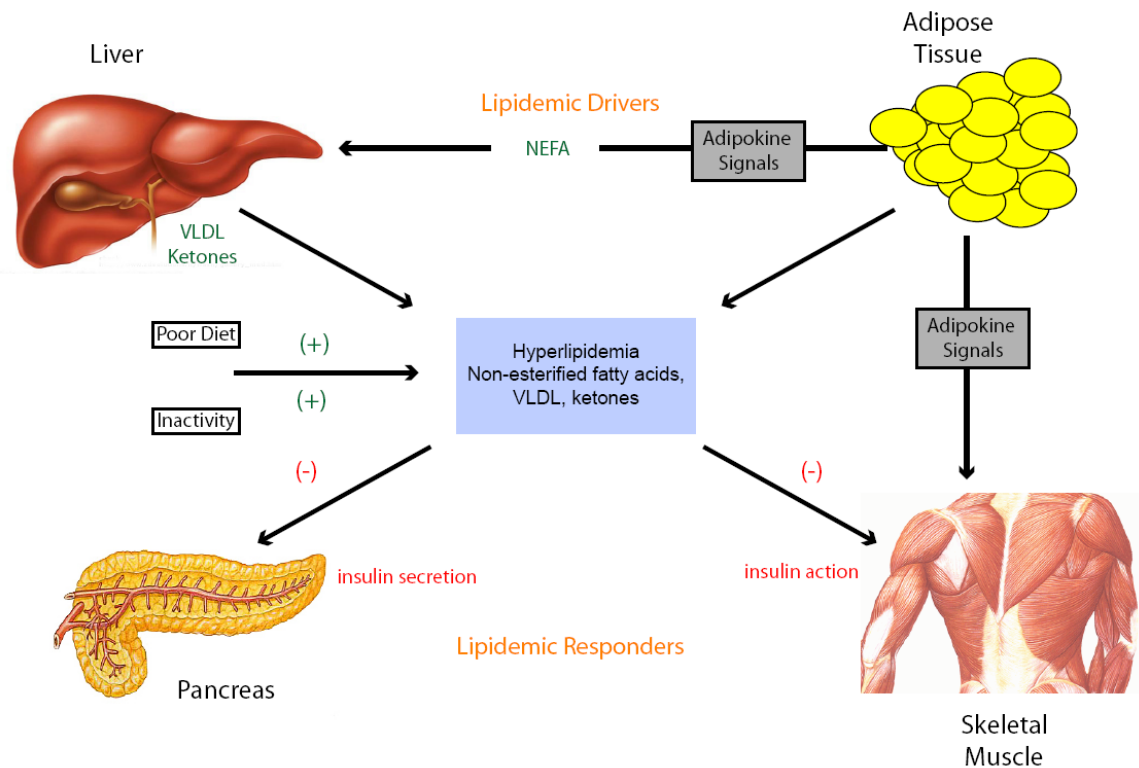
regulator genes may reside, the molecular pathway from genotype to physiological output was largely unknown.

One of the first challenges was to advance association studies linking phenotypes (clinical traits, expression traits) to causal directed networks and establish relationships between these correlated inputs. Secondly, we recognized that clinical trait was a downstream product of molecular and transcriptional happenings, so we asked whether hepatic metabolites map to genetic regions and thus complement the genomic and transcriptomic data. I hypothesized that genetic regions (mQTL) would alter abundance of hepatic metabolites, and that these analytes may not only be key intermediates between gene expression and clinical trait, but may actually be upstream and causal for gene expression changes.

To address these questions, I worked closely with the statistics department and we developed a method to incorporate QTL data into an association network, using the gene mapping information to infer causal versus reactive direction between two related phenotypes. Once we established the methodology for developing a QTL-directed dependency graph (QDG), we applied this to our segregating F2^{ob/ob} population and incorporated the genomic information with hepatic metabolite and transcript data. We perform studies on hepatocytes from the parental strains to validate network construction in the liver. Finally, we address how tissues communicate by studying how potential liver-derived metabolites alter insulin secretion in INS-1 cells and isolated rat islets. These studies involved collaboration between the Sarah W. Stedman Nutrition and Metabolism Center at Duke University and the Biochemistry and the Biostatistics

Departments at the University of Wisconsin, Madison. Quantitative, analytical measurements of organic acids and acyl-carnitines by GC/MS and MS/MS, respectively, were conducted at the Stedman Center's Metabolomics/Biomarker core laboratory and microarray analysis was performed by Dr. Attie's collaborator, Rosetta, Inc (a subsidiary of Merck).

Ultimately, we would like to characterize all interconnections between causal tissues, such as the liver and adipose, and reactive organs, the islet and the muscle. The methods developed in the course of this dissertation for integration of information about genetic variance, gene expression, and comprehensive metabolic profiling is widely applicable to other tissues where gene mapping of an F2 population has been applied. Just as we used genetic regions to identify a causal relationship between two liver phenotypes (e.g., transcript and metabolite), we can also define causality between two correlated nodes of different tissues (e.g., hepatic metabolite and islet mRNA abundance). A causal network relating organ systems will again provide data-driven, testable hypotheses about master regulators of diabetes pathology, and ultimately help us focus on crucial targets for therapeutic intervention.



Muoio and Newgard (2006) Annu Rev Biochem. 75:367-401.

Figure 1. Hierarchy of systemic metabolic changes in obesity-related overnutrition.

In states of metabolic overload, multiple organs are impacted. The adipose inadequately sequesters excess fuel and secretes adipokines; the liver fails to oxidize fatty acids and releases their excess fuels as VLDL and ketones. Adipokines, free fatty acids, and triglycerides derived from “driving organ” dysfunction have downstream effects on the islet and skeletal muscle. Ultimately the skeletal muscle fails to efficiently take up glucose while the islet undergoes β -cell decompensation, secreting insulin levels inadequate to meet the increased demand. VLDL: very-low density lipoproteins; NEFA: non-esterified fatty acids.

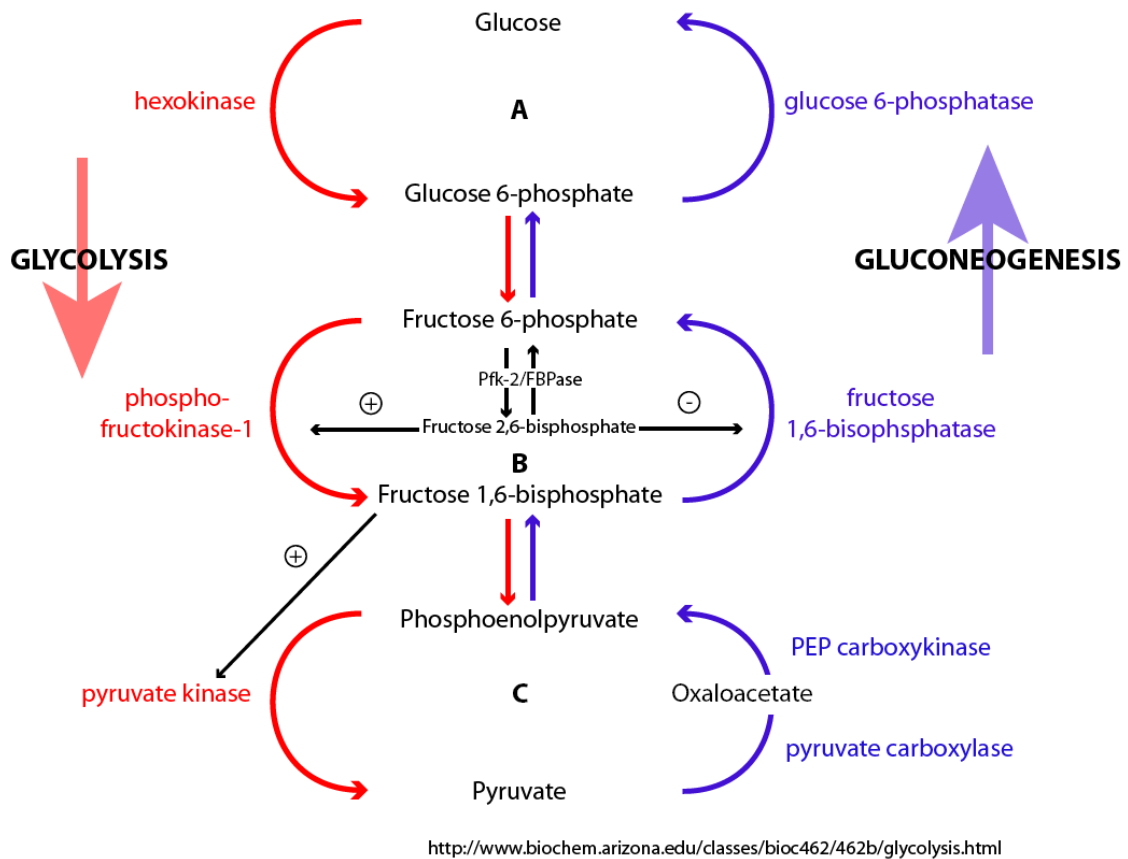
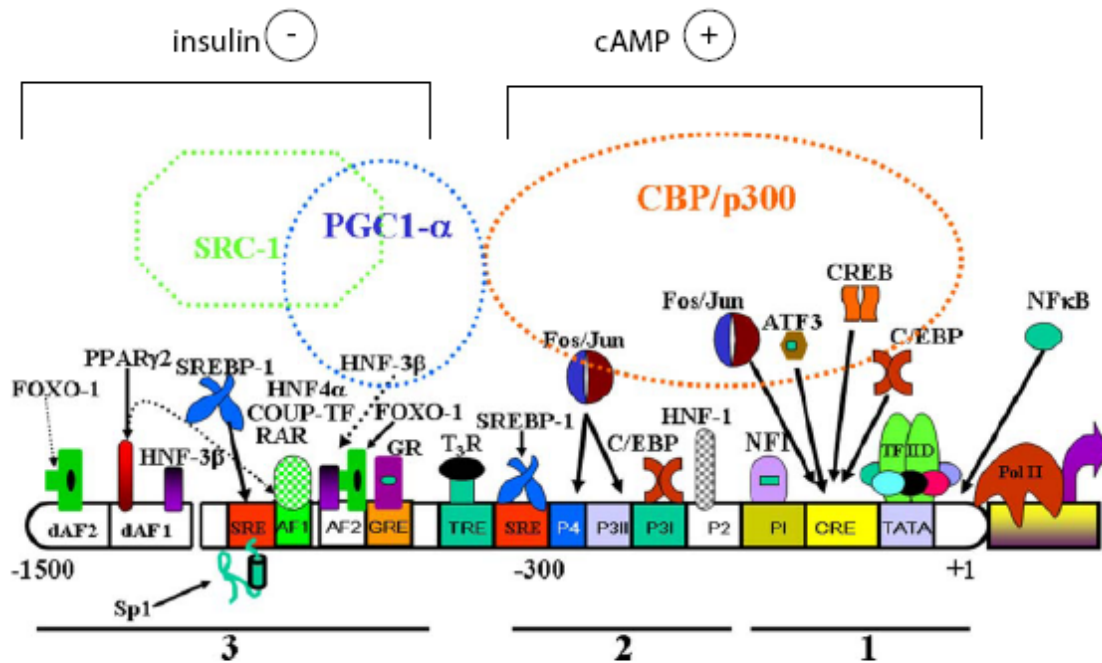


Figure 2. Substrate cycles in the glycolytic/gluconeogenic pathways.

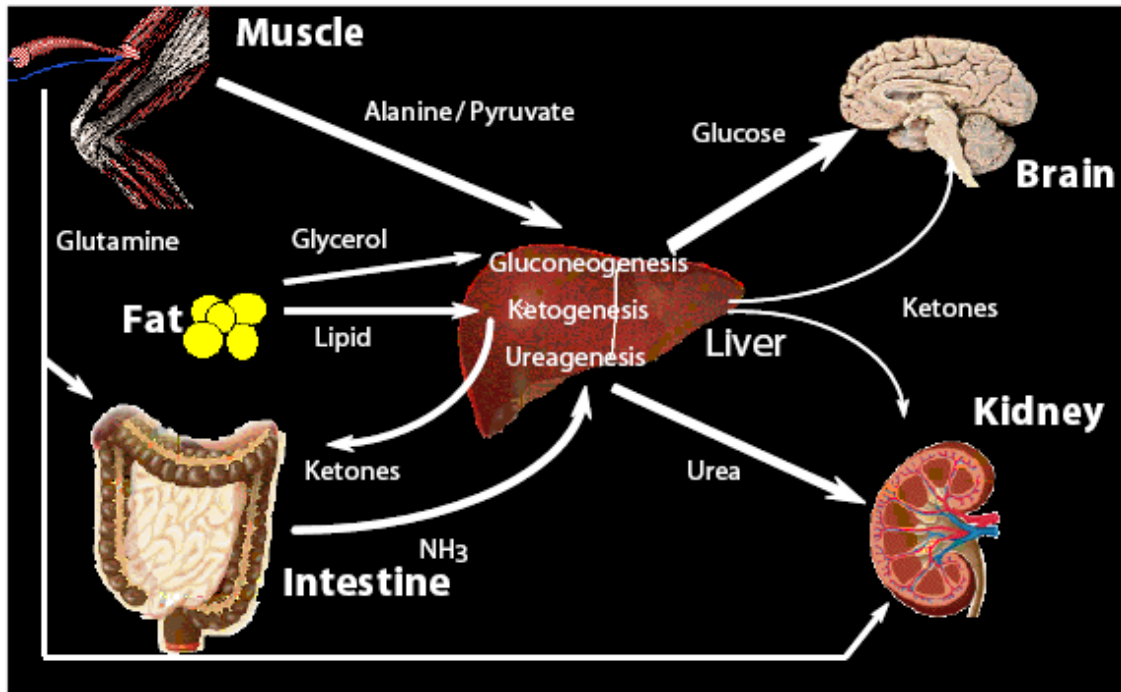
The three steps of glycolysis that are irreversible are the conversion of glucose to glucose-6-phosphate, fructose-6-phosphate to fructose-1,6-bisphosphate, and pyruvate to phosphoenolpyruvate. Gluconeogenesis therefore has unique enzymes to catalyze the reverse reactions, establishing three glycolysis/gluconeogenesis “cycles”: the glucose cycle (A), fructose cycle (B), and pyruvate/PEP cycle. (C) The enzymes of the hepatic substrate cycles are subject to short and long-term regulation by hormones, energy status, and metabolites. PEP: phosphoenolpyruvate; Pfk2/FBPase: 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2.



Chakravarty, et. al (2005) Crit Rev Biochem Mol Bio. 40(3):129-154

Figure 3. Phosphoenolpyruvate carboxykinase, cytosolic (Pck1) gene promoter.

CREB: cAMP regulatory binding protein; ATF3: activating transcription factor 3; NFκB: nuclear factor κB; COUP-TF: chicken ovalbumin upstream promoter transcription factor (NR2F); PPARα2: peroxisome proliferator-activated receptor α2; SREBP-1: sterol regulatory element binding protein; HNF: hepatocyte nuclear factor; C/EBP: CAAT/enhancer binding protein; NF1: nuclear factor 1; CREM: cAMP regulatory element modifier; CBP/p300: CREB binding protein/p300; PGC-1γ: peroxisome proliferator activated receptor coactivator 1γ; SRC-1: sterol receptor coactivator 1; RAR: retinoic acid receptor; AF1: accessory factor 1; AF2: accessory factor 2; dAF1: distal accessory factor 1; dAF2: distal accessory factor 2; TRE: thyroid hormone regulatory element; Pol II: RNA polymerase II; GRE: glucocorticoid regulatory element; CRE: cAMP regulatory element; P1, P2, P3(I), P3(II), and P4: protein binding sites.



www.e971.com/caep/brightmoon/tnt/4.ppt

Figure 4. Metabolic changes during catabolic states.

In times of starvation (fasting, perceived starvation such as diabetes), peripheral tissues spare glucose for the central nervous system and catabolize alternate fuel substances for energy. As a result of this tissue breakdown, the liver receives tissue-derived amino acids and fatty acids and, via gluconeogenesis and ketogenesis, respectively, converts them into glucose and ketones to be used for energy by peripheral tissues and the central nervous system.

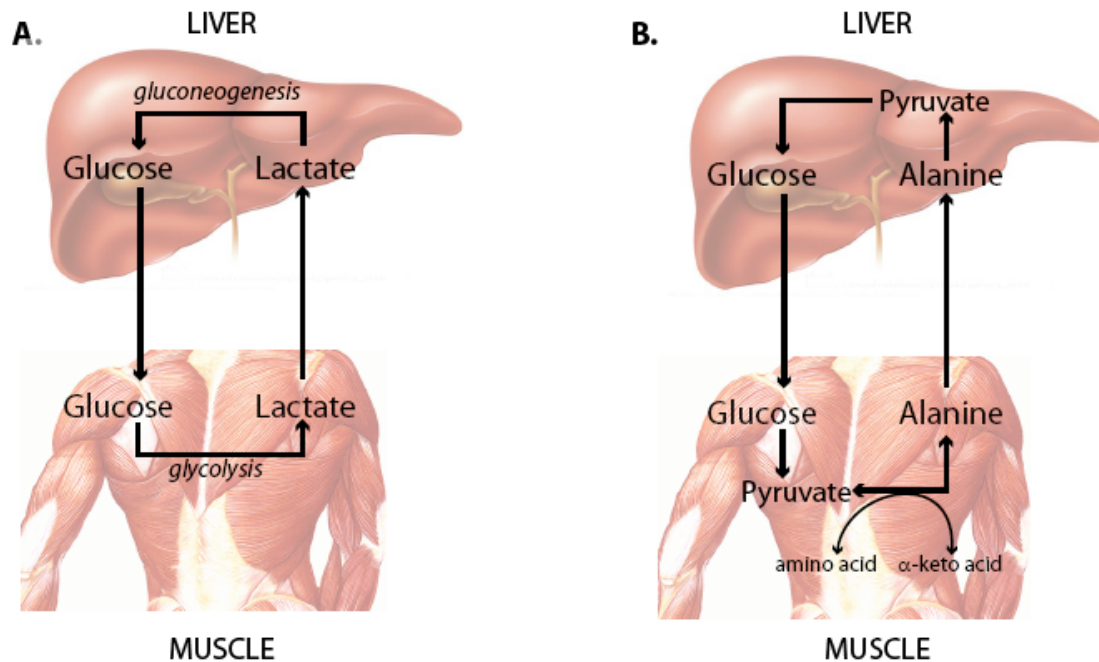


Figure 5. The Cori cycle and glucose/alanine cycle represent interorgan communication with the liver.

A. The Cori cycle involves the utilization of lactate, produced by glycolysis in non-hepatic tissues, as a carbon source for hepatic gluconeogenesis. In this way the liver can convert the anaerobic byproduct of glycolysis, lactate, back into more glucose for peripheral tissues. B. The glucose-alanine cycle is increased during periods of fasting as skeletal muscle protein is degraded for the energy value of the amino acid carbons. Within the liver alanine is converted back to pyruvate which is then a source of carbon atoms for gluconeogenesis; the amino group is converted to urea in the urea cycle and excreted.

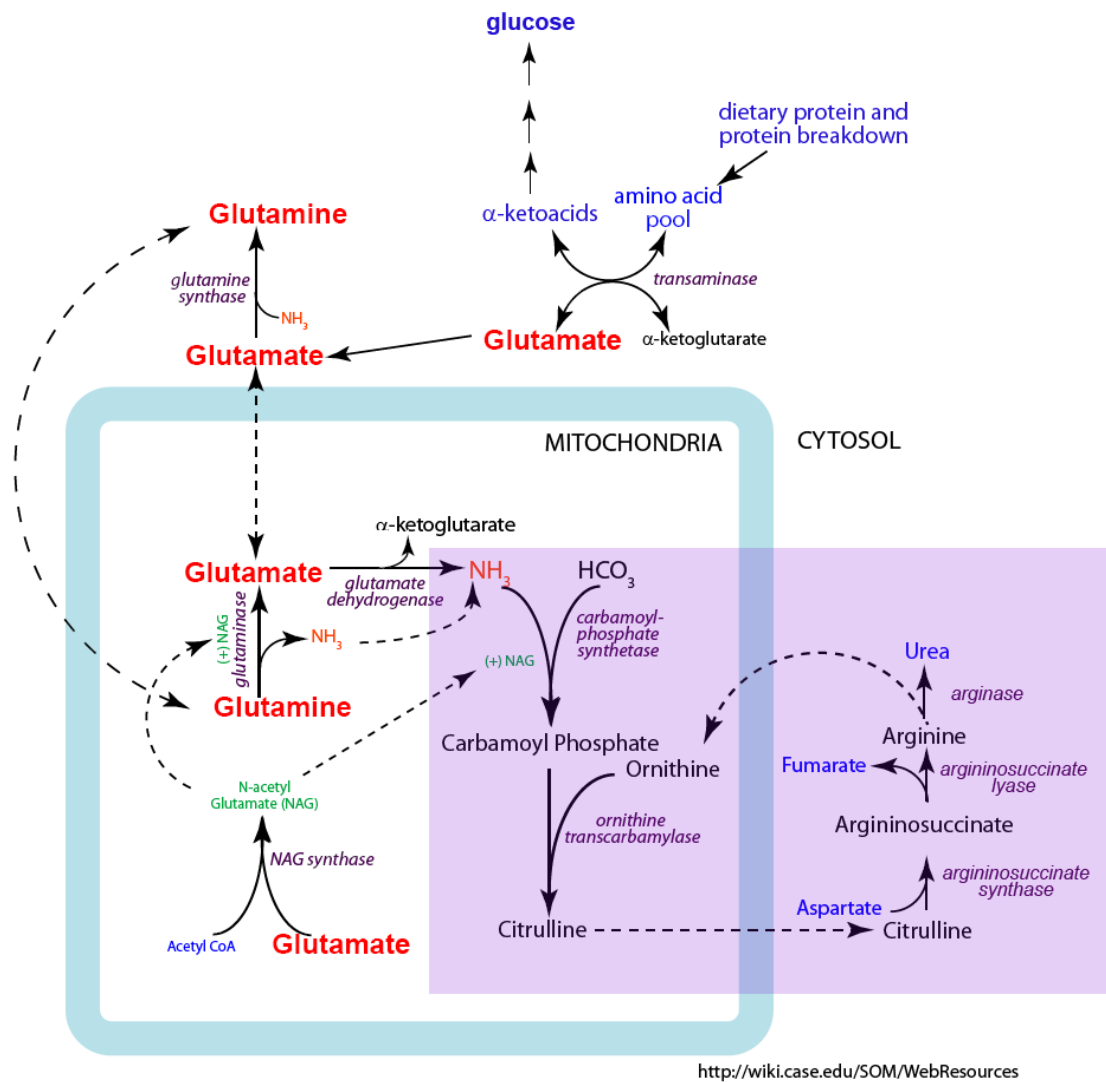
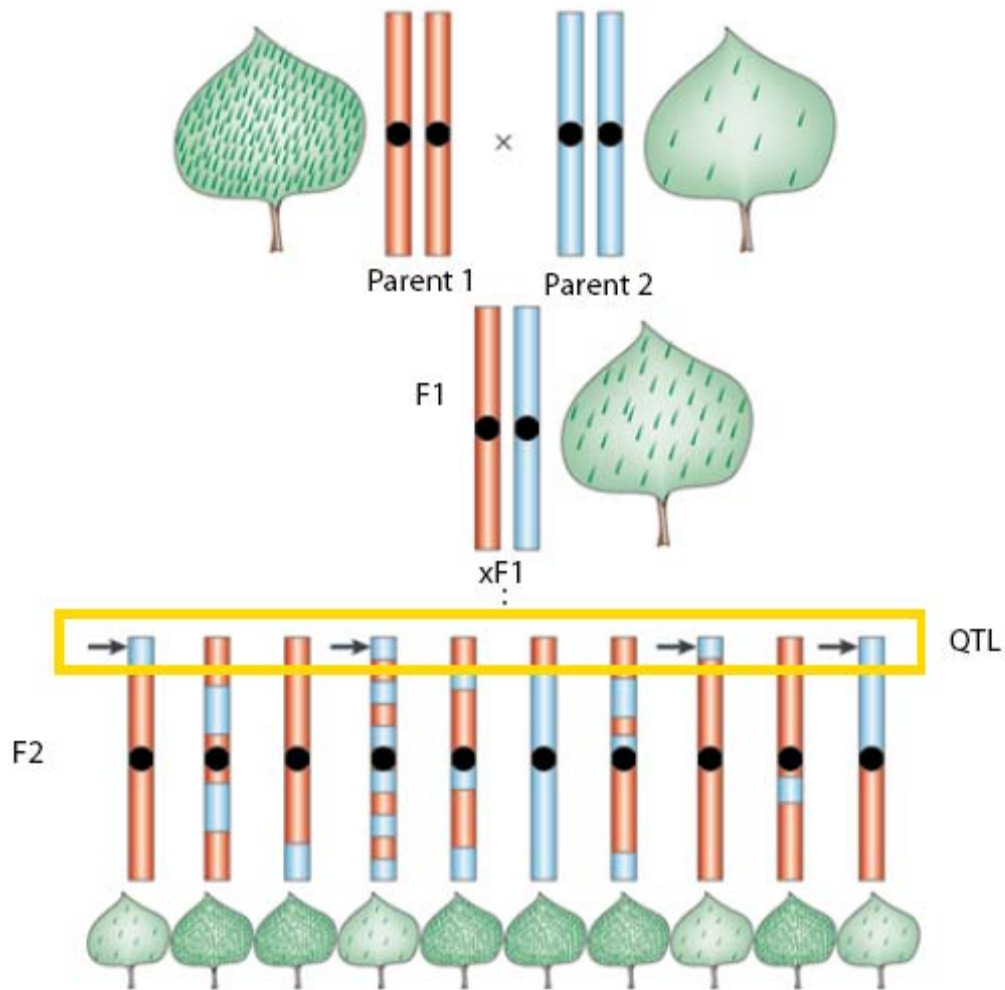


Figure 6. Glutamate integrates carbohydrate and amino acid metabolism.

Amino acids from peripheral tissue degradation and the diet are transaminated in the liver to glutamate and an α -keto acid. The α -keto acid carries the carbon chain to gluconeogenesis while glutamate carries the nitrogen group to the urea cycle (purple box). Glutamate not only delivers ammonia to the urea cycle, but is also the precursor for N-acetylglutamate (NAG), a potent allosteric activator of carbamoyl phosphate synthetase which catalyzes the first step of the urea cycle.



Mauricio (2001) Nat Rev Genet. 2:370-381.

Figure 7. Strategy for QTL mapping.

The phenotype mapped here is the density of hairs on a plant leaf. Inbred parental strains that differ in the density of hairs are crossed to form an F1 population. This F1 population has individuals with an identical genetic and phenotypic make-up and an intermediate level of hairs. F1 individuals are crossed among themselves to produce an F2 population. The F2 individuals are scored for several genetic markers, as well as for the hair density phenotype. The arrow and boxed region indicates the section of the chromosome that drives the density of hairs, the quantitative trait locus (QTL).

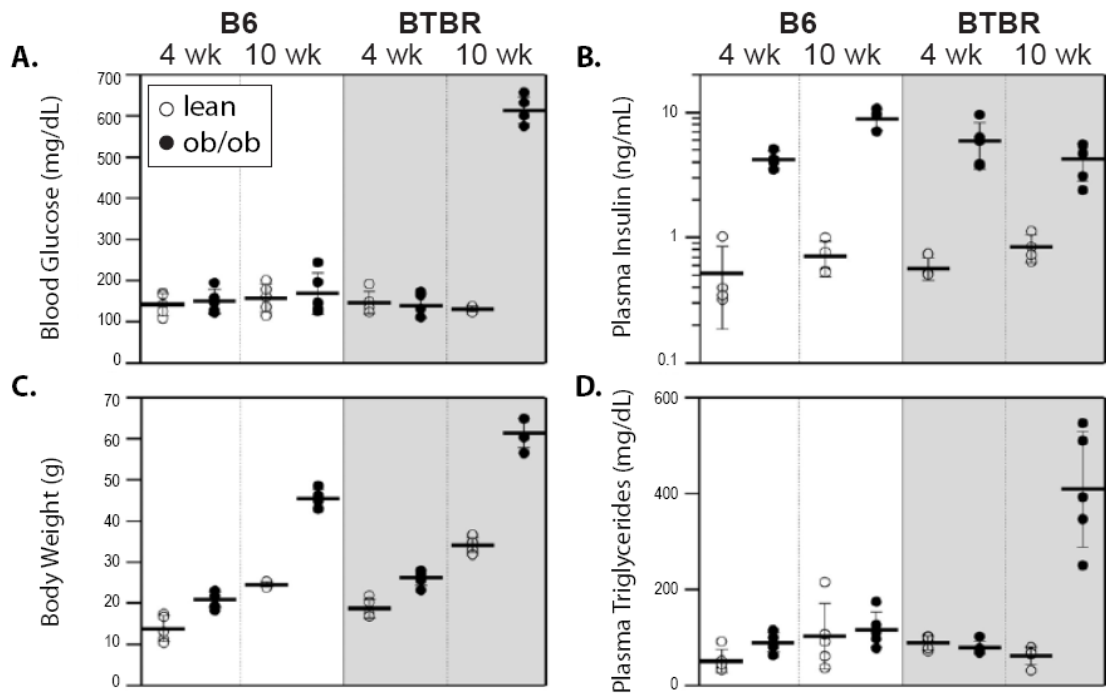
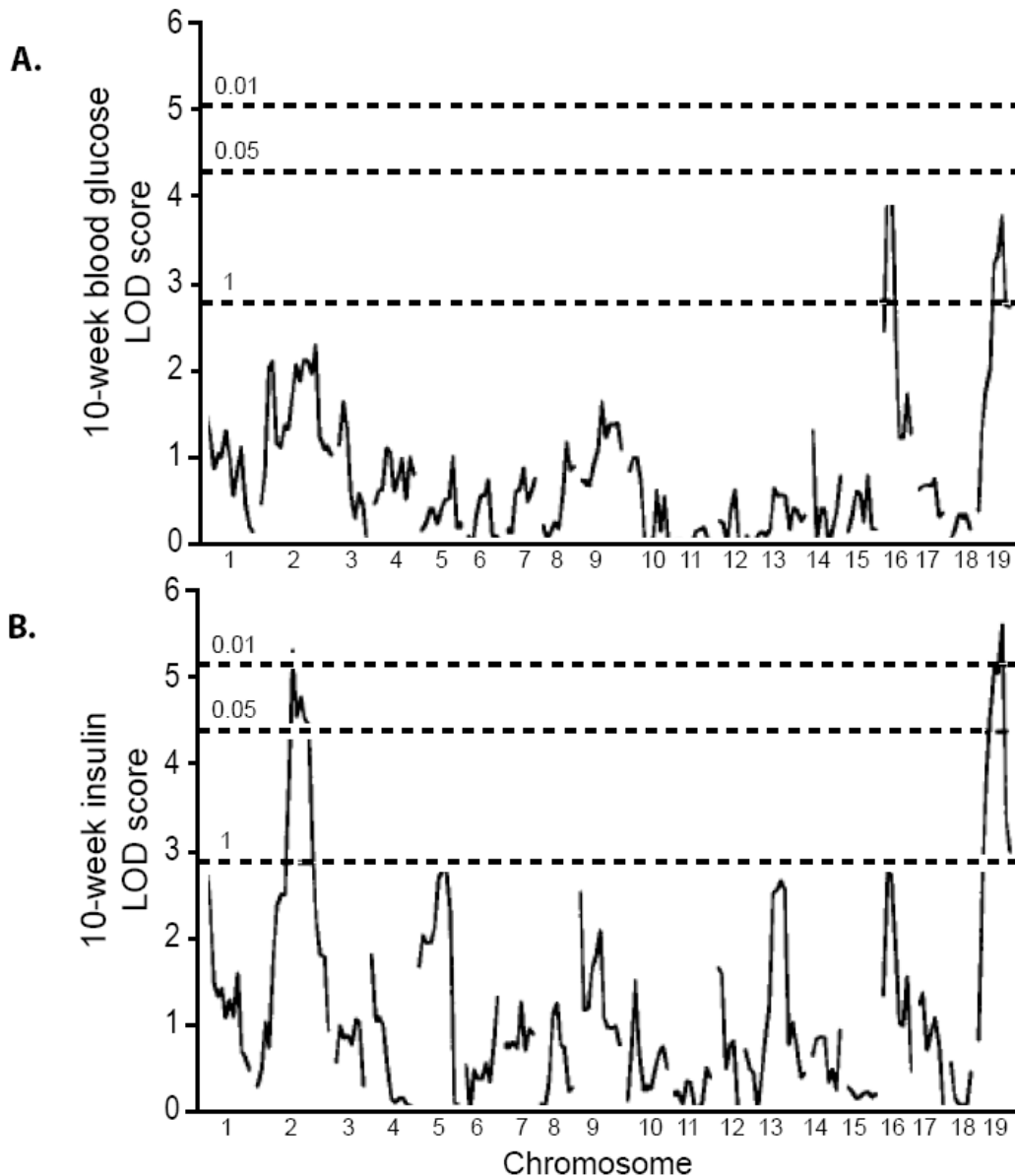


Figure 8. C57BL/6 and BTBR mouse strain phenotypes.

Clinical phenotypes blood glucose (A), plasma insulin (B), body weight (C), and plasma triglycerides (D) are shown for five animals for each of the eight groups of mice used in the study. Open (lean) and closed (*ob/ob*) circles represent individual mice. Horizontal bars show mean values for each group.



Stoehr, et al (2000) Diabetes. 49(11): 1946-1954.

Figure 9. Genome mapping reveals clinical QTLs (cQTLs).

Genome interval mapping indicates cQTL for (A) 10-week fasting plasma glucose on chromosomes 16 and 19 and (B) 10-week plasma insulin on chromosomes 2 and 19 for an F2 population derived from B6^{ob/ob} and BTBR^{ob/ob} parental strains. Dashed lines indicate genome-wide significance guidelines. Maximum likelihood interval mapping techniques were used to generate LOD scores.

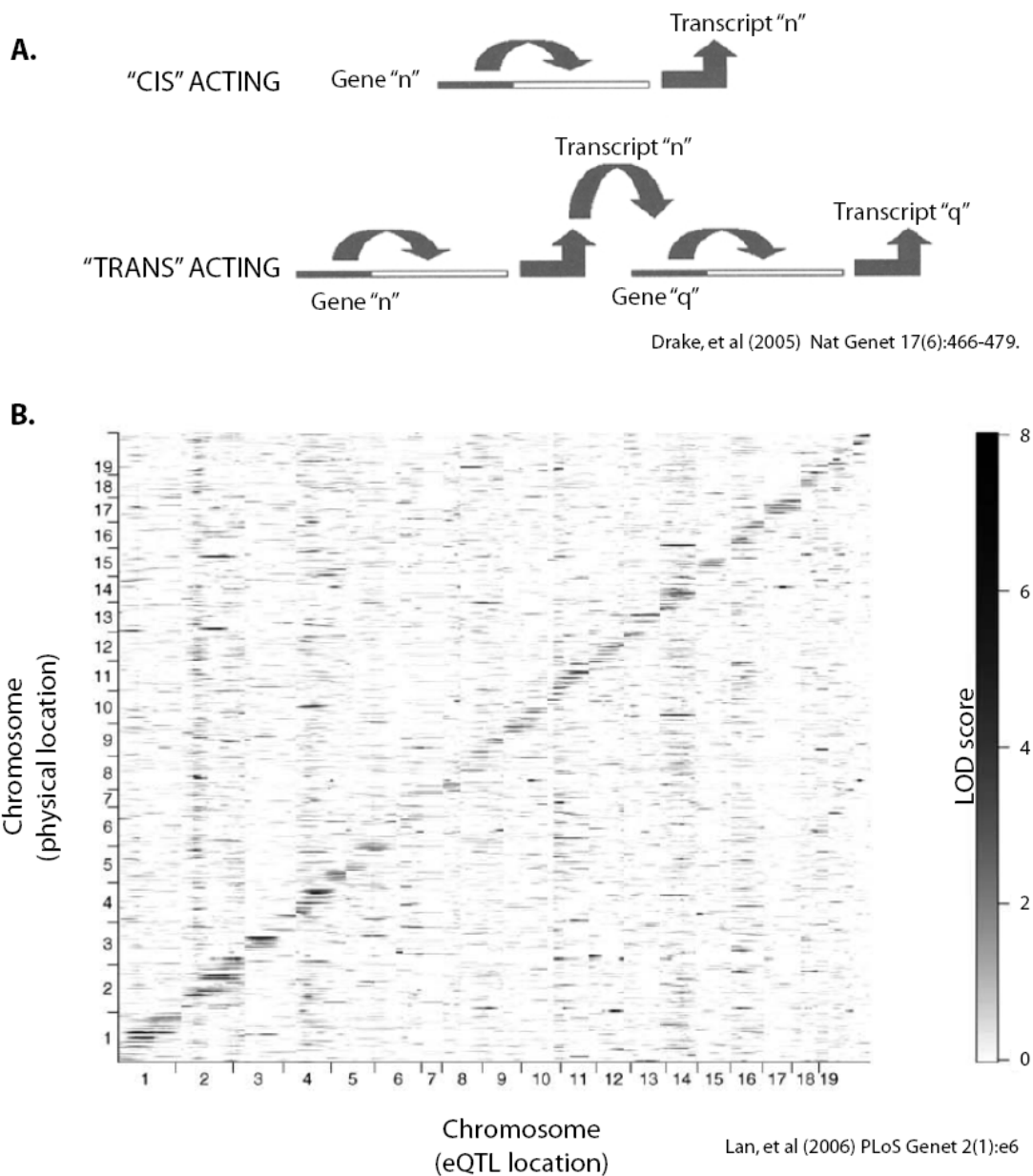


Figure 10. Cis- and trans-regulation of transcript abundance.

A. A transcript maps in *cis* if the eQTL is located at the gene encoding the transcript. In the case of *trans*-regulation, the eQTL of a given transcript is at a location distinct from the locus encoding the transcript. B. Gene expression from liver microarray show that the majority of transcripts map in *cis* to where they are physically encoded (diagonal). The LOD score represents the strength of the relationship between the transcript and its eQTL.

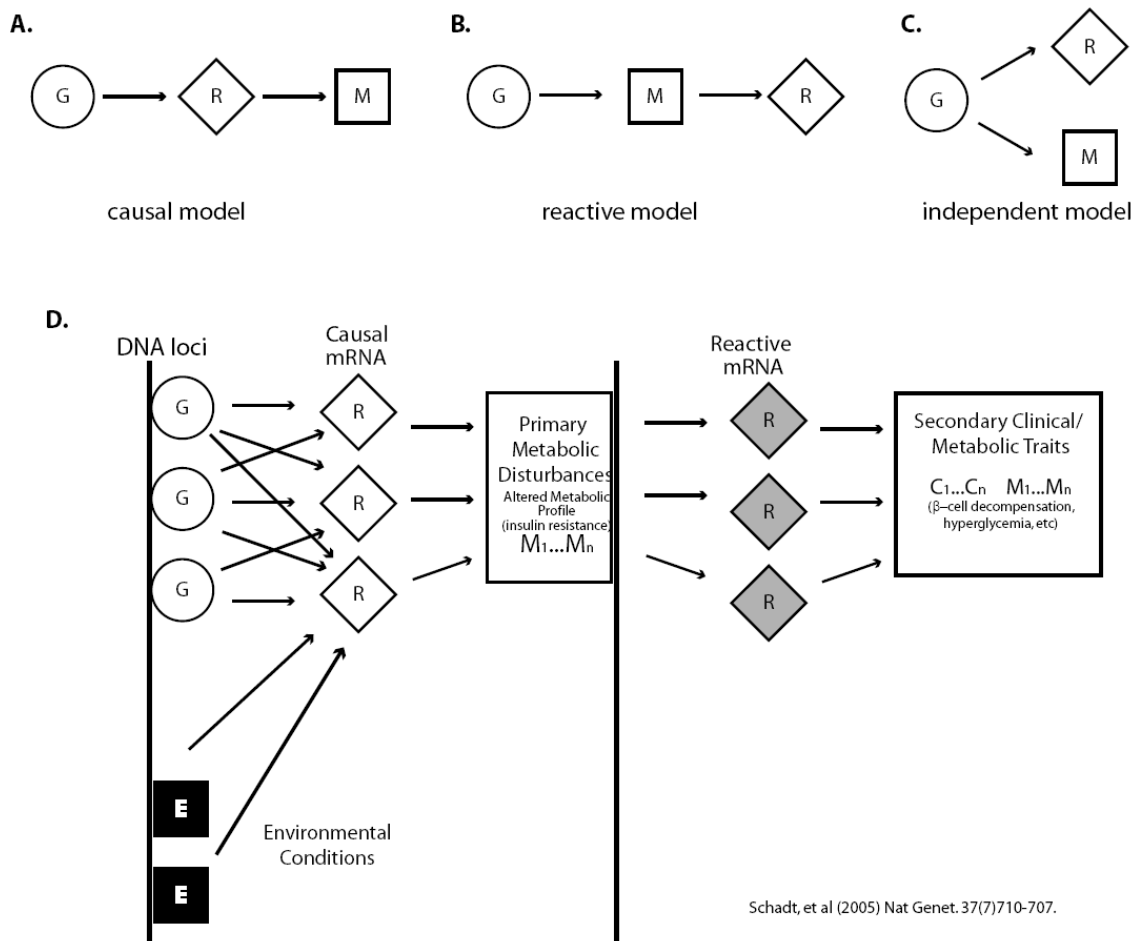
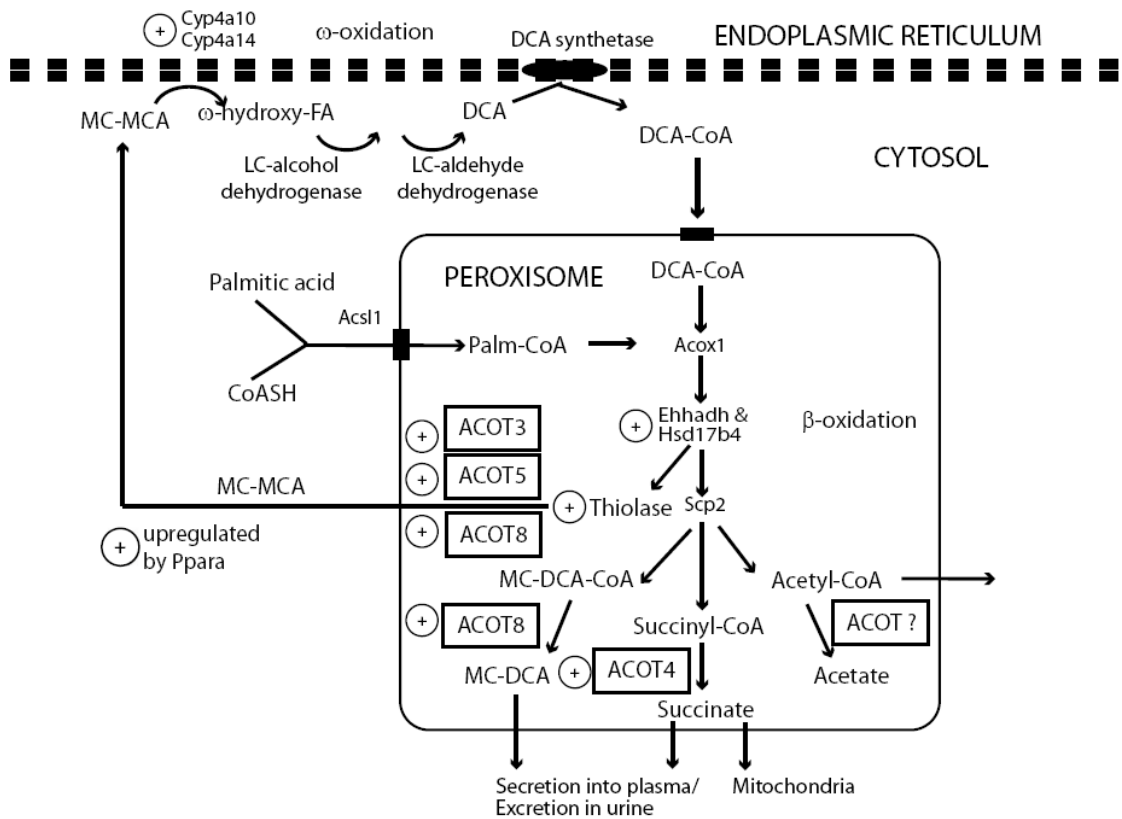


Figure 11. Possible relationships between QTLs, metabolites, and mRNA.

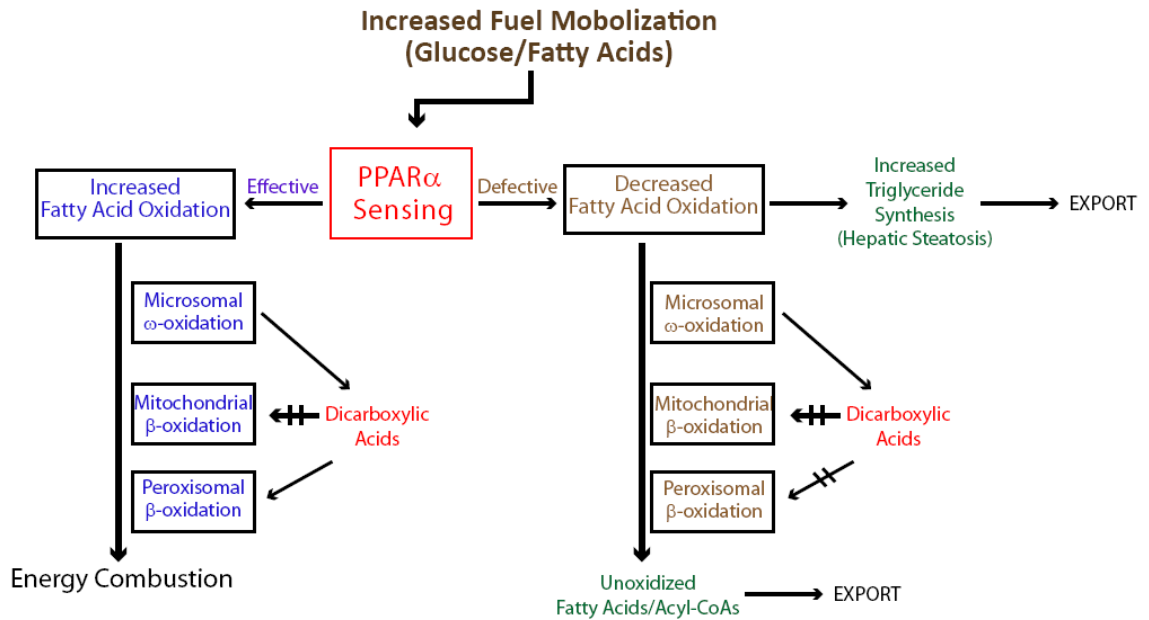
There are three ways in which a correlating transcript (R) and metabolite (M) that have eQTL and mQTL at the same locus (G) can be related. In the causal model (A), the locus alters gene expression which has a downstream effect on metabolite levels. The reactive model (B) suggests that the mRNA is modulated by metabolites. The metabolite and mRNA may also be independently related (C), where the QTL acts on these traits but the metabolite and mRNA do not influence one another. These models do not work in isolation, and the complexity can generate secondary and tertiary metabolite and gene expression changes (D).



Westin, et al (2005) J Biol Chem. 280(46):38125-32.

Figure 12. Hepatic fatty acid oxidation pathways.

In the liver, mitochondrial, peroxisomal, and microsomal fatty acid oxidation systems are regulated by PPAR α -mediated gene expression. Under normal conditions, the liver can accommodate fatty acids via mitochondrial β -oxidation. In times of overnutrition which exceed mitochondrial capacity, fatty acids are metabolized via ω -oxidation to dicarboxylic acids, which are subsequently broken down in the peroxisome.



Reddy and Rao (2006) Am J Physiol Gastrointest Liver Physiol. 290(5) G852-G858.

Figure 13. Cross-talk between fatty acid oxidation systems in times of metabolic overload.

Elevated fatty acids increase PPAR α signaling and gene expression changes in microsomal ω -oxidation and mitochondrial and peroxisomal β -oxidation. If mitochondrial oxidation capacity is exceeded or defective, microsomal ω -oxidation engages, producing dicarboxylic acids, which can be used for energy or secreted. At elevated levels, however, the dicarboxylates inhibit the already obstructed oxidation pathways, and un-metabolized fatty acid derivatives are released, further elevating circulating plasma levels of these metabolites.

2. Inferring Causal Phenotypes from Segregating Populations

2.1 Introduction

Network models derived from microarray experiments have shed light on the manner and extent of connectedness among expressed genes. However, these networks mostly summarize association, connecting phenotypes without causal direction. Genetical genomic studies with microarray data in a segregating population have shown evidence for both *cis*-acting and *trans*-acting genomic regions [134,135]. These studies yield expression quantitative trait loci (eQTL) that provide a causal experimental system where genotype at that loci drives phenotype[136]. Directed graphs inferred from such a system can yield causal gene networks that predict biochemical pathways, generating new hypotheses about functional relationships among expressed genes[137-141].

In the past, generation of networks has required a network node that consists of a gene physically located within the QTL support interval of one or more transcripts; this gene thus becomes a candidate regulator for the other transcripts[141]. Unfortunately, phenotypes such as metabolites and physiological traits cannot serve as causal nodes in these networks. Schadt, *et al.* was the first to consider causal clinical traits as network nodes. By using pairs of phenotypes that share multiple QTLs, they defined directionality between a clinical trait and transcript through likelihood-based causality model selection tests, using conditional correlation measures to determine which trait/trait relationship is best supported by the data [135]. Li, *et al.* proposed the application of structural equation models (SEM) to genetic analysis of multiple traits. SEM allows for characterization of

pleiotropic and heterogeneous genetic effects of multiple loci on multiple traits as well as the physiological interactions among traits [142].

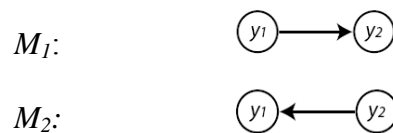
Several of these current methods require network nodes consisting of genes physically located within the QTL support interval of other transcripts [137-141] or require common QTLs [135], but despite these constraints have laid the foundation for construction of more complex networks. We have derived a novel causal network reconstruction method for genetical genomics experiments termed QTL-directed dependency graph (QDG) which drops these QTL restrictions, permitting construction of causal networks among phenotypes that may not have common QTLs. This also allows for the use of phenotypes (e.g., body weight and blood glucose) other than gene expression and considers covariates such as sex and age.

The QDG approach is widely applicable to human studies and outbred populations. It also can accommodate feedback loops which are common in biology. We present a series of simulation studies where we evaluate the recovery rate of undirected graphs and compare the performance of the QDG and conventional PC-algorithm for causal orientation in a network containing 100 phenotypes. We assess the accuracy of the QDG method for the interference of small acyclic graphs, and show that the integrated QDG approach overcomes serious limitations of both genetical genomics network methods and the PC-algorithm.

2.2 Network Construction

The QDG method first builds an undirected graph that infers associations among phenotypes using an undirected dependency graph (UDG) or PC-skeletons [143,144]. These two methods are better suited to causal inference than other networks such as Gaussian graphical models (GGM), which include edges between any pair of nodes that has significant partial correlation [145]. However, partial correlations can exist even when two nodes are uncorrelated, leading to spurious edges (Figure 14). In directed acyclic graphs, UDGs and PC-skeletons avoid this problem by first removing edges where there is no significant correlation ($\alpha=0.05$). To avoid edges between two nodes that are not directly connected, but are connected via an intermediate node, we test the partial correlation between the nodes of interest, conditional on the intermediate. If it is not significantly different from zero, we remove the edge between the two nodes, leaving only connections to the intermediate node. In directed cyclic graphs, it may not be possible to remove all spurious edges. In evaluating the two methods for edge recovery, the PC-skeleton algorithm performs fewer computations than the UDG algorithm for generating sparse graphs, but it is less accurate in small sample sizes [144].

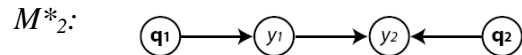
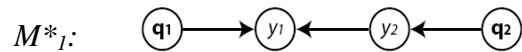
The second step of the QDG algorithm is to causally orient every edge connecting a pair of phenotypes in an undirected network by asking, “Does phenotype y_1 drive phenotype y_2 , or *vice versa*?”



Model selection procedures cannot distinguish between M_1 and M_2 because they are distribution or likelihood equivalent:

$$f(y_1)f(y_2/y_1)=f(y_1,y_2)=f(y_2)f(y_1/y_2)$$

where $f()$ is the predictive density of the argument conditions. Including QTL with genotypes \mathbf{q}_1 affecting y_1 and \mathbf{q}_2 affecting y_2 , we can resolve direction because new models



are *not* likelihood equivalent since the predictive densities disagree,

$$f(\mathbf{q}_1)f(y_1|\mathbf{q}_1)f(y_2/y_1,\mathbf{q}_2)f(\mathbf{q}_2)\neq f(\mathbf{q}_2)f(y_2/\mathbf{q}_2)f(y_1/y_2,\mathbf{q}_1)f(\mathbf{q}_1).$$

Therefore, we distinguish between models M^*_1 and M^*_2 using a LOD score that conditions on genotypes at multiple QTL (derived from earlier gene mapping of phenotypes), and can now create a directed graph (DG).

It is possible that a strong QTL affecting an upstream trait may also be incorrectly detected as a QTL for a downstream phenotype, so we must first distinguish QTLs with direct and indirect effects. We employ the following to differentiate these direct and indirect QTLs: 1) identify all pairs of connected phenotypes that share a common QTL, and 2) apply a generalization of the method proposed by Schadt, *et al.* to each of these pairs in order to determine if the common QTL directly affects both traits or if it has an indirect effect in one of the phenotypes. This method includes comparing the “causal”, “reactive”, and “independent” models for the genetic locus and the two co-mapping traits (Figure 15). The log likelihood is calculated for each model and the number of

parameters is taken into consideration to generate a model score. Based on each model score, the relationship between the two traits is determined. Schadt used this method to determine the causal direction between a pair of phenotypes, and we have applied this to test direct versus indirect effect of a common QTL on two phenotypes[135]. Whenever a pair of connected phenotypes have a common QTL, we test the null hypothesis that no partial correlation between the phenotypes conditional on the common QTL. If the null is accepted, the edge between the phenotypes is dropped and the common QTL is considered to have pleiotropic effects.

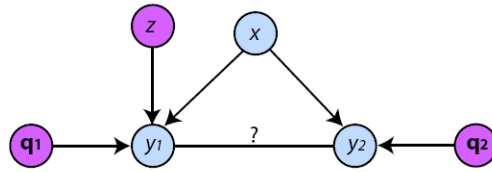
The next step is to orient causal edges between all pairs of connected phenotypes in the undirected network using associated multiple QTLs to break likelihood equivalence. Edge orientation among pairs of phenotypes is based on a linear regression model with phenotypes regressed on QTL genotypes and on additive or interacting covariates such as sex, age, and other phenotypes. We orient each edge, conditional on all other nodes (phenotypic, genotypic, or covariate) that affect the two nodes of the specific edge under investigation.

For each determined edge, we now evaluate a LOD score comparing the two possible orientations. We orient the edge in favor of the direction with the higher likelihood in the ratio, thus creating a QTL-directed graph (QDG). For the simple network presented above, this ratio is a contrast of LODs,

$$LOD=(LOD_1+LOD_{2|1})-(LOD_2+LOD_{1|2})$$

Where $LOD_1+LOD_{2|1}$ represents the likelihood that \mathbf{q}_1 affects y_1 and that y_1 is causal for y_2 and $LOD_2+LOD_{1|2}$ represents the likelihood that \mathbf{q}_2 affects y_2 and that y_2 is causal for y_1 .

The p-values for the LOD scores are computed using 10,000 permutations. To generate the null distribution for the direction, we permute blocks of data. Consider the network:



where we are interested in computing the p-value for the arrow $y_1 \rightarrow y_2$ conditional on covariates x and z . Here x represents a covariate on both phenotypes (x could be another phenotype, age, sex, etc); z is a covariate of y_1 only. \mathbf{q}_1 and \mathbf{q}_2 represent two sets of QTLs affecting y_1 and y_2 , respectively. In order to break the connections that are directly causal to the direction of the edge, we permute the corresponding pairs of nodes (and their common covariates) as a block. We permute y_1, y_2, x as a block (blue) by breaking the connections with z, \mathbf{q}_1 , and \mathbf{q}_2 (which constitute another block, purple). The data is permuted 10,000 times to generate the null distribution around a LOD of zero.

We then calculate the LOD scores of each 10,000 individual permuted blocks of data; while we have broken connections that affect the direction of the edge, we are still utilizing a QTL and covariate block z, \mathbf{q}_1 , and \mathbf{q}_2 in each permutation, but these permuted data sets do not correspond to the “true” y_1, y_2, x block to which they belong in the actual individual. These LOD scores will be approximately, but not exactly, equal

to zero and these values are then plotted on a histogram. The p-value is calculated from this histogram, allowing for $\alpha=0.05$.

In more complex networks, we orient edges in two steps: (1) build an initial directed network orienting each edge as above; (2) recomputed the LOD score for each edge connecting a pair of phenotype nodes by condition on all other phenotypes causative to either or both nodes. We repeat the second step for all edges until no edge switches direction which involves some iteration to find the best orientation across the entire graph.

For example, consider a directed network with four phenotypes in which a preliminary directed network may look like $1 \rightarrow 2 \rightarrow 3 \leftarrow 4$. The second step recomputes the LOD score for the edges between nodes 2 and 3 conditioning respectively on nodes 1 and 4. [Note the edge 1-2 is not tested at first because neither node 1 or 2 is downstream of, and therefore reactive to, any other node in the network considered.] Similarly, we recomputed the LOD score for the 3-4 edge, conditioning on node 2 for phenotype 3. If the direction of the edge 2-3 switches, the graph becomes $1 \rightarrow 2 \leftarrow 3 \leftarrow 4$. We then recomputed the LOD scores for the 1-2 edge using node 3 as a covariate for 2, and again for the 2-3 edge conditioning respectively on 1 and 4. We repeat this step until no more edges change direction.

In complex networks, this algorithm may find more than one solution. That is, starting the algorithm from a different edge orientation may yield a different graph. If the graph oscillates between different directions without converging to a solution in 30 steps, we include all distinct graphs as solutions. In order to get the best solution we: 1) re-run

this algorithm using different initial edges to get all possible solutions 1000 times; 2) score each solution using a likelihood-based overall measure of fit; 3) select the graph with the overall highest score. Although we start the QDG method as a pairwise algorithm considering the relationship between two nodes and their upstream effectors, the last steps represent a multivariate algorithm, where we perform computations of direction on subsets of the graph.

In summary, the QDG algorithm is as follows:

1. Construct an undirected dependency graph (PC-skeleton or association graph);
UDG
2. After removing indirect QTL, use QTL genotypes to direct each edge in the
UDG to create a directed graph; DG
3. Randomly choose an ordering of all edges in the DG to make an ordered
directed graph; ODG
4. For an edge in ODG, recompute the direction LOD score using all other
causative phenotypes to either or both nodes, according to the DG, in addition
to QTL genotypes. If the direction changes, update the DG and move to the
next edge in ODG.
5. Repeat step (4) until no edges change direction, providing one solution. If
graphs do not converge within 30 steps, these distinct graphs are distinct
solutions
6. Repeat steps (3), (4), and (5) 1,000 times and store all different solutions

7. Score all solutions in (6) using a likelihood-based measure of fit of the whole graph; that with the highest score is the solution; QDG

2.3 Results

In all simulation studies, we used the PC-skeleton to recover edges, if one existed, between phenotypes. We then assessed the performance of the PC-skeleton to recover a true edge by calculating the average (across all edges in the respective graphs) of the percentage of times the algorithm recovered each of the true edges. Edge recovery rates were strong, increasing with sample size. We then assessed QDG for power to determine the correct direction for these previously PC-skeleton-derived recovered edges. Finally we evaluate the QDG power to infer causality by contrasting it with the PC-algorithm[146], which can orient some but not all edges in an undirected network.

Hundred Phenotypes Network

A complicated network with 100 phenotypes and 107 edges was used to assess our methodology (Figure 16a). Each phenotype node was affected by two or three QTL (not shown in figure), and we allowed for only additive genetic effects. The QTL for each phenotype were randomly selected among 200 markers, with 10 markers unevenly distributed on each of 20 autosomes. We allowed different phenotypes to potentially share common QTL. Our 100 phenotype network results illustrate the percent of edges *recovered* using the PC-skeleton increases with sample size (Figure 16b).

When comparing the PC-algorithm and QDG methods to determine causal direction, we see that as sample size increased, both methods showed an increase in the percentage of recovered edges whose direction was correctly inferred and a decrease in the percent of incorrect directions. In all cases, however, the QDG method resulted in more correct directions than the PC-algorithm and when sample size was greater than 60, fewer incorrect directions (Figure 17).

The network in Figure 16a demonstrates many features that can be inferred with varying degrees of difficulty by the PC-algorithm. Nodes organized in an unshielded collider pattern, such as $x \rightarrow z \leftarrow y$ are easier to detect than nodes organized in a bifurcation or line pattern such as $x \leftarrow z \rightarrow y$ or $x \rightarrow z \rightarrow y$. The PC-algorithm performed better in the unshielded collider pattern relative to other patterns, both in its ability to recover edges and also to identify correct causal direction to these recovered edges. The QDG algorithm, however, was able to infer correct direction equally well for all patterns and at a greater rate than PC-algorithm (Figure 18).

The QDG method is robust with respect to the number of QTL per phenotype. If only one QTL per phenotype is chosen at random among the total QTL affecting the phenotype, there is a little loss in percent correct directions.

Cyclic Networks

We investigated the properties of three cyclic graphs simulated according to Chaibub, *et al* (Figure 19)[147]. As sample size increased from 100 to 200, the average percentage of true recovered edges using the PC-skeletons improved to nearly 100% for graphs (a) and (b). In graph (c) with a direct reciprocal interaction between nodes 2 and 5,

allows nodes 1 and 5 (and 4 and 2) to be connected in two manners. Node 1 can be related to 5 via node 2 (4,2,5, respectively); on the other hand, node 1 and 5 are independent of one another and both affect node 2 (4,2,5). Given the former relationship it is possible a spurious edge may be detected between 1 and 5 (4, 2); however spurious edges were only recovered two out of 100 realizations.

Just as increasing sample size allowed for the PC-skeleton to improve the number of *recovered* edges detected, increasing sample size allowed QDG to correctly *direct* a greater percentage of these recovered edges in cyclic networks. Although the QDG method cannot detect direct reciprocal interactions (such as $2 \leftrightarrow 5$), it can infer the stronger direction corresponding to a higher regression coefficient.

Metabolite and Gene Expression Network

We inferred a causal network among expression and metabolite traits and validated the network with *in vitro* experiments[104]. Here we simulate from a connected subset of this inferred network to study its statistical properties.

The causal network was constructed using *in vivo* gene expression and metabolite data from a sample of 58 mice from an F2 leptin-deficient *ob/ob* population generated from the C57BL/6 *leptin^{ob/ob}* (B6^{*ob/ob*}) and BTBR *leptin^{ob/ob}* (BTBR^{*ob/ob*}) founder strains[95]. The B6 and BTBR strains, when made obese, differ in diabetes susceptibility; The B6^{*ob/ob*} are diabetes resistant where as the BTBR^{*ob/ob*} develop severe diabetes. The included transcripts had a high positive correlation with liver glx (glutamine+glutamate) and/or had co-mapping eQTL and mQTL[104].

In the construction and simulations, age was used as an additive covariate and we allowed for sex-by-genotype interactions. Estimates from the real data were used as the parameter values in the simulations and we adopted $\alpha=0.05$ for the UDG algorithm to recover edges. Our QDG algorithm with random edge orderings converged on two solutions (Figure 20). We simulated 1,000 realizations from the network, recovering a true edge 75% of the time, and correctly oriented those recovered edges on average 83%. False edges were detected overall at rates below 2%.

Our network predicted six edge directions which were tested by isolating hepatocytes from the B6 and BTBR parental strains. Real-time PCR was used to measure gene expression changes in response to 10mM glutamine supplement *in vitro*[104]. We confirmed causal network predictions through this subsequent laboratory work as described in Chapter 3 and published [104], and thus showed our QDG method is valid for hypothesis generation.

2.4 Discussion

In genetical genomics, or eQTL studies, we are increasingly interested in inferring causal networks for sets of phenotypes that map to coincident genomic regions. We typically believe that there is a “master regulator” and that most co-mapping is due to indirect effects. Thus our objective is to untangle the direct and indirect effects of QTLs and phenotypes. We provide a novel method that can first infer what phenotypes have

proximal relationships (undirected graphs), and then orient these edges to infer causal relationships and produce unified causal networks.

In recent years, there has been a large effort to model phenotypic networks. Approaches began by constructing networks based on associations and correlations [95,104,148,149] and since this early work, the goal has been to infer causality within these networks [137,139-141,145,150]. These preliminary methods have developed the foundations of causal network reconstruction from which we can now move onto inferring causality in complex networks that more closely resemble biology.

We have proposed a simple approach for causal network construction that extends existing causal network methods by accommodating phenotypes that do not share common QTLs. We have shown how to infer causal direction in an association network of phenotypes by conditioning on causal QTL, even when graphs include cycles. We begin with feature selection to reduce from thousands of phenotypes to a manageable number [95,149], followed by a focus on a particular functional group. We then use QTL mapping methods to identify major QTL of phenotypes in this functional group [151].

This proposed method is advantageous in that it can handle multiple QTLs and covariates (additive or interactive), with computing time proportional to the number of edges. This can be readily extended to include discrete driving factors other than genotypes, such as sex or treatment group. Furthermore, this method can be applied to a wide variety of phenotypes. The incorporation of genotypes (or discrete driving factors) allows for the reconstruction of a *single* directed network of phenotypes. This has advanced previous work where reconstruction using Bayesian networks was limited to

the inference of equivalence classes, that is a set of networks that have the same likelihood[152]. Another advantage of QDG is that it does not require sequence information, and can therefore be useful in the discovery process of a broader range of genetical genomics experiments[139].

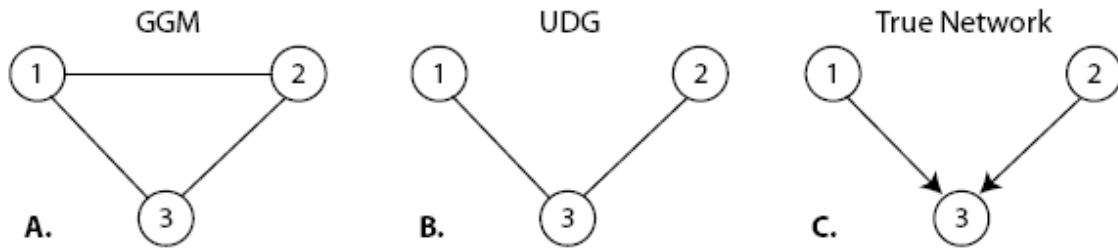
Our QDG method shows superior performance over the PC-algorithm to infer pleiotropic effect of one gene expression phenotype on two other phenotypes, or more broadly, the directions of edges that bifurcate from any node[146]. Both the PC-algorithm and QDG are effective in detecting unshielded collider patterns, which correspond to the epistatic effect of two gene expression phenotypes on a third phenotype. However, the enhanced ability of the QDG algorithm to infer both pleiotropy and epistasis will make QDG superior as eQTL and mQTL studies become more prevalent in experimental crosses.

In this work, we assume that a set of multiple QTLs affecting each of the phenotypes had been previously determined. With sample sizes in the order of 300 to 500, we expect to have enough power to detect QTLs for a *set* of functionally related phenotypes, rather than individual phenotypes, for our causal network methods. This will become important as this method is applied to biological systems where inter-tissue communication occurs. The QDG also can infer directionality in cyclic graphs containing three or more phenotypes, representative of feedback loops.

A key feature of the QDG is that phenotypes with no detectible QTLs need not be excluded in our analysis, and on a similar note, we can orient an edge between a pair of phenotypes with common QTLs. The only requirement to determine directionality is that

there be at least one QTL or causal phenotype uniquely affecting one of the pair of nodes under consideration. Even in a situation where two connected y_1 and y_2 contain only one common QTL \mathbf{q} and causal models are thus likelihood equivalent (e.g., $f(\mathbf{q})f(y_1/\mathbf{q})f(y_2/y_1, \mathbf{q})=f(\mathbf{q})f(y_2/\mathbf{q})f(y_1/y_2, \mathbf{q})$), our method may determine directionality. We can modify the network by breaking genotype-phenotype edges to consider models $\mathbf{q} \rightarrow y_1 \rightarrow y_2$ and $\mathbf{q} \rightarrow y_2 \rightarrow y_1$, in addition to the pleiotropic model where by \mathbf{q} affects both phenotypes[135]. Providing the pleiotropic model is not supported, we can infer a direction between the two phenotypes.

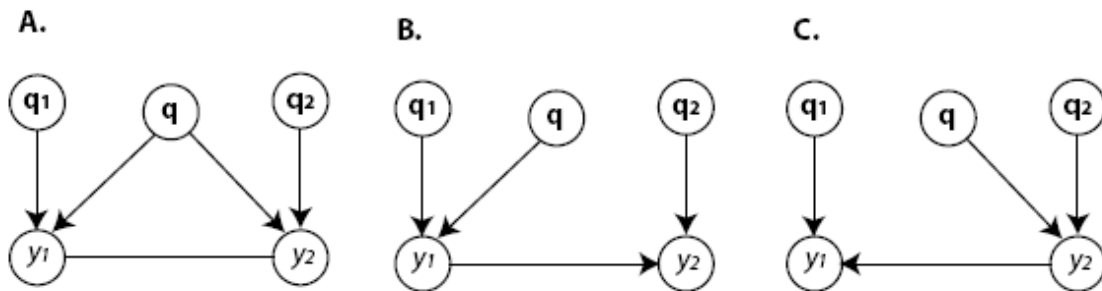
In conclusion, we have presented a novel, efficient causal network construction method QDG with four main advantages: 1) we screen out edges associated with partial correlation that are not causative in directed acyclic graphs; 2) we have overcome the previous QTL constraints, dropping requirements that a phenotype have a QTL, that two phenotypes share common QTLs, and that an expression trait be physically located within the QTL; 3) we can infer feedback loops, common in genetic, metabolic, and biochemical networks; and 4) we can infer a single network instead of an equivalence class of networks. The QDG network method developed in the course of this dissertation research is appropriate for a range of population structures, including but not limited to inbred crosses, outbred crosses, and natural populations. The causal networks that emerge form the basis for hypotheses about critical pathways that can be individually tested and aid in our understanding of phenotype relationships.



Chaibub, et. al (2008) Genetics. 179:1089-1100

Figure 14. Spurious edges and partial correlations.

A. The Gaussian graphical model (GGM) will recover an edge between any two nodes where there is a correlation, and is unable to distinguish partial correlation relationships. B. The undirected dependency graph (UDG) and PC-skeleton methods indicate that nodes 1 and 2 are independent, and removes edge 1-2 which is a partial correlation, given 3. UDG or PC skeletons eliminate spurious edges and are therefore better suited to build the framework of the true network (C).



Chaibub, et al (2008) Genetics. 179:1089-1100.

Figure 15. Distinguishing direct and indirect effects of a common QTL.

Suppose that QTL mapping of phenotype y_1 detected QTLs q and q_1 , and mapping of phenotype y_2 detected the common QTL q plus QTL q_2 . A strong QTL directly affecting an upstream trait may also be (incorrectly) detected as a QTL for a downstream phenotype. In order to resolve this situation, we apply a generalization of Schadt, *et al.* 2005, allowing for multiple QTLs q_1 and q_2 and use these to infer direction; we then score the three models above. Model A supports both traits being directly affected by the common QTL q . Model B implies that q directly affects y_1 but should not be included as a QTL of phenotype y_2 . Model C supports the reverse situation where y_1 is downstream of y_2 which is directly impacted by q .

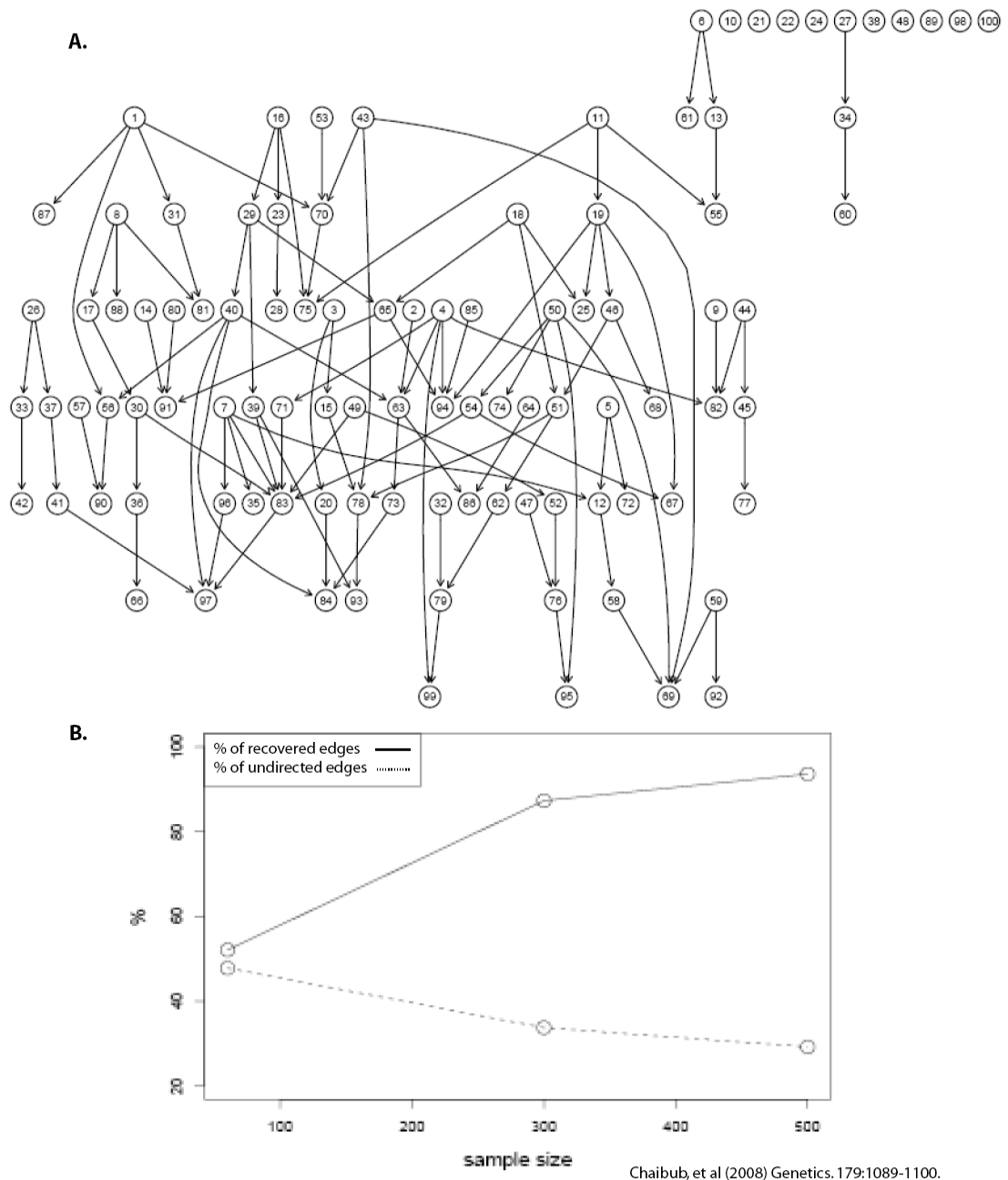
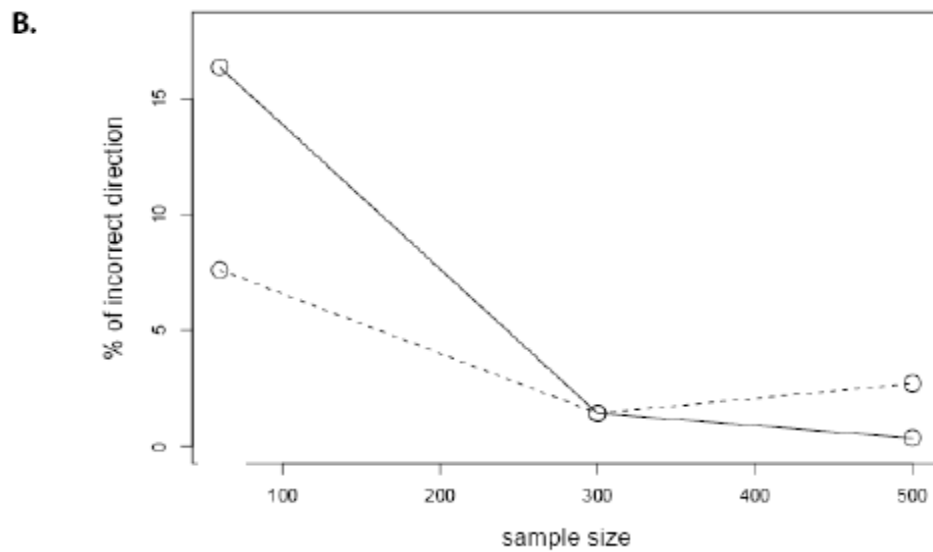
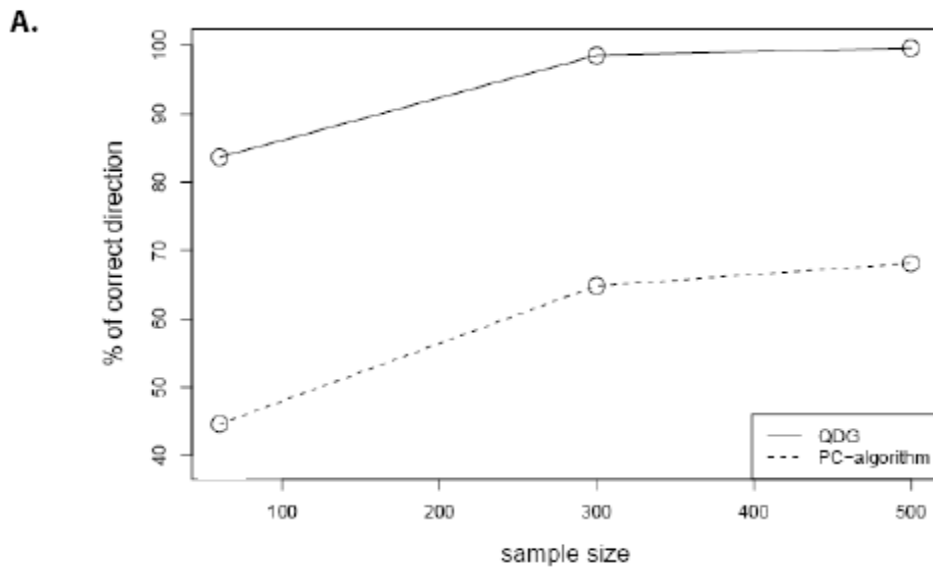


Figure 16. Directed acyclic graph (DAG).

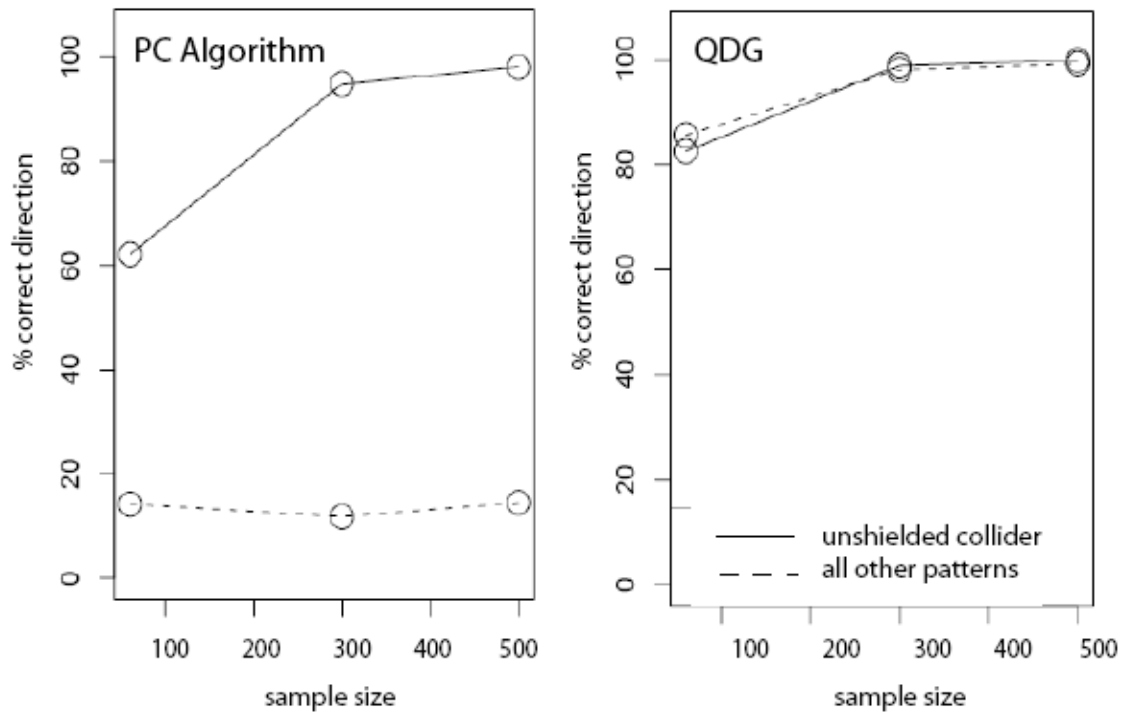
A. This graph was composed by 100 nodes (phenotypes) connected by 107 edges using the random DAG function in the pcalg R package B. Average percentage of recovered edges and average percent of edges that could not be detected by the PC-algorithm relative to the total number of true edges existing in the network.



Chaibub, et al (2008) Genetics. 179:1089-1100.

Figure 17. QDG performance in inferring correct direction of recovered edges.

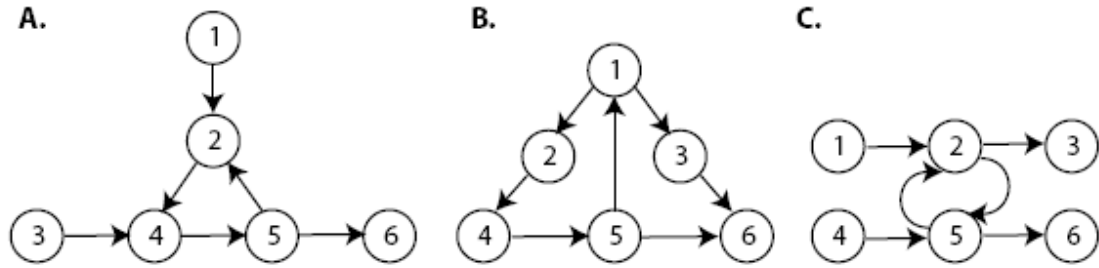
Comparison of the average percent of correctly inferred directions (A) and incorrectly inferred directions (B) by QDG and PC-algorithm for DAG of Figure 16a.



Chaibub, et. al. (2008) Genetics. 179:1089-1100.

Figure 18. QDG performance of inferring direction in various patterns.

Comparisons of PC-algorithm (A) versus QDG (B) success in detecting unshielded collider pattern $1 \rightarrow 2 \leftarrow 3$ (solid line) and other patterns (dotted line) of the network in Figure 16a.



Chaibub, et. al. (2008) *Genetics*.179:1089-1100.

Figure 19. Cyclic networks.

Graphs A and B show simple cyclic graphs. Graph C contains a reciprocal interaction between nodes 2 and 5.

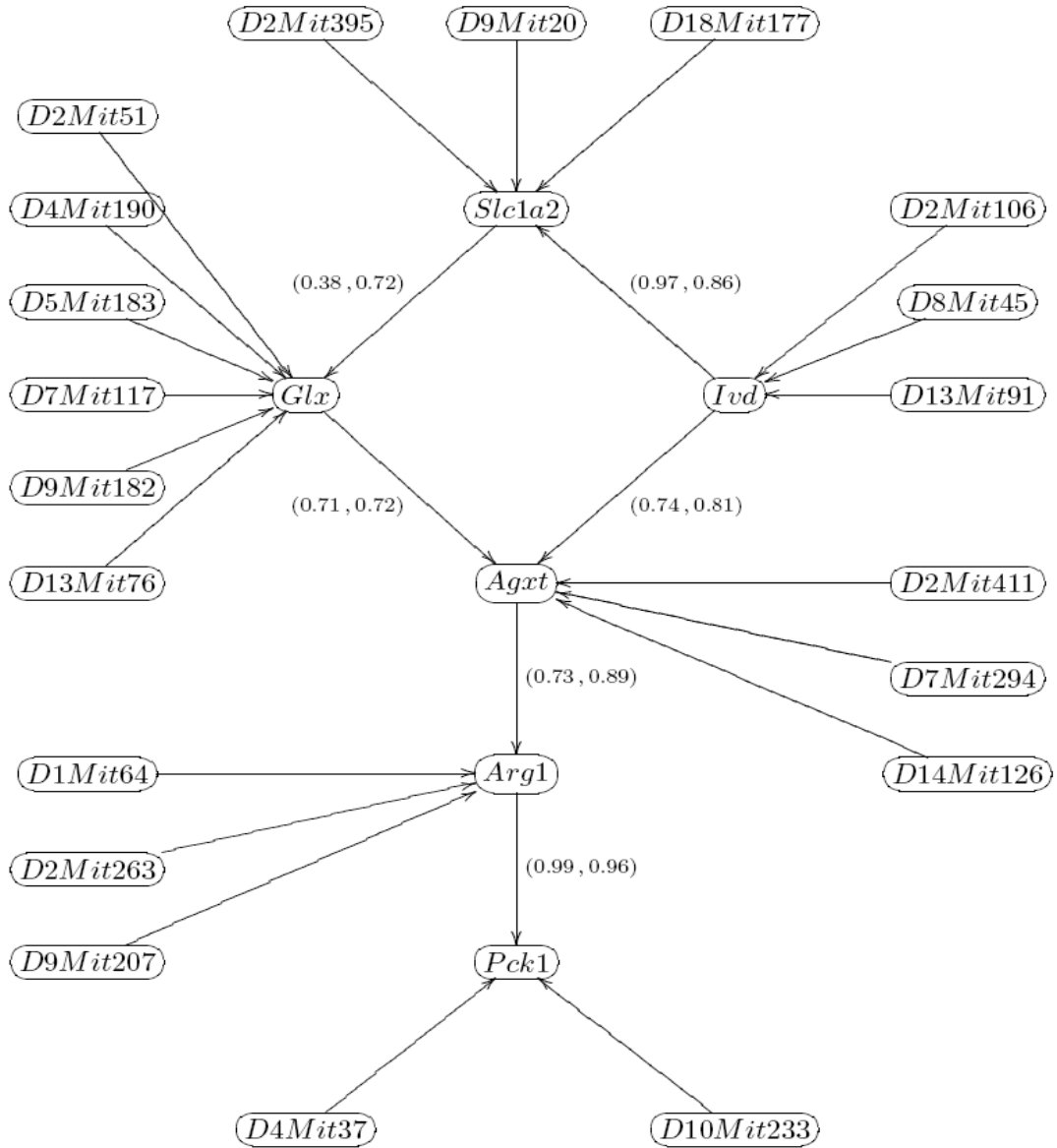


Figure 20. Causal network from a segregating F2 sample.

The F2 population from the B6^{ob/ob} and BTBR^{ob/ob} parental strains segregate by genotype and phenotype as published in Ferrara, *et al.* 2008. Values represent the proportion of times the edge was detected, and the proportion of times this edge was given the correct direction (e.g., (0.99,0.96) from *Arg1* to *Pck1* indicates a 1% chance that this is not truly and edge and that 96% of the time, the edge arrowhead is *Arg*→*Pck1* as indicated).

3. Genetic Networks of Liver Metabolism Revealed by Integration of Metabolic and Transcriptional Profiling

3.1 Introduction

Genetic linkage and association studies have the power to establish a causal link between gene loci and physiological traits. These studies can make novel connections between biological processes that would not otherwise be predictable based on current knowledge. The pace of gene discovery has greatly accelerated in recent years, and numerous quantitative trait loci (QTL) influencing disease-related phenotypes have been identified through gene mapping and positional cloning. While it has become relatively straightforward to map a phenotype to a broad genomic region, identification of the individual gene(s) driving phenotype changes remains difficult. Consequently, only a few percent of the many QTL that have been mapped have had their underlying gene(s) identified [153-159]. Another limitation of traditional QTL mapping is that it is based on association with a physiological phenotype, but often does not reveal the molecular pathways leading to that phenotype.

One way to uncover molecular mechanisms of disease states is to broadly expand the types of phenotypes analyzed in genetic screens. For example, with microarray technology, one can measure the abundance of virtually all mRNAs in a segregating sample. Importantly, mRNA abundance shows sufficient heritability in outbred populations and experimental crosses to allow mapping of gene loci that control gene expression, termed expression QTL (eQTL) [95,135]. When eQTL co-localize with a

physiological QTL, one can hypothesize a shared regulator and offer a potential pathway leading to the physiological trait [135,160].

The pathway between a QTL and a physiological trait often involves changes in the steady-state levels of metabolic intermediates, in addition to changes in mRNA abundance. These metabolites can correlate with the genetic, transcriptional, translational, post-translational, and environmental influences on phenotype [101,159]. Moreover, metabolites are intermediates in signaling pathways that can regulate gene expression. For example, fatty acids act as ligands for several of the PPAR nuclear hormone receptors, bile acids activate FXR in liver, and diacylglycerol regulates protein kinase C [161-163]. Metabolite abundance reflects a biological response to exogenous and endogenous inputs, and when investigating pathways from genotype to phenotype, metabolites can provide a powerful complement to gene expression data and give novel insights into disease pathogenesis mechanisms [59,94,96,101,108,132,159,164-169].

Our laboratories have begun to apply targeted metabolic profiling to study mechanisms underlying obesity-induced diabetes [59,108,132,164-166], but have not yet attempted to integrate these methods with genotyping and transcriptional profiling. This has included the application of gas chromatography/mass spectrometry (GC/MS) and tandem mass spectrometry (MS/MS) for measurements of acyl-carnitine, organic acid, amino acid, free fatty acid, and long and medium-chain acyl-CoA metabolites in tissue extracts and bodily fluids. Herein, we have applied these methods to measure various metabolites in liver samples from mouse strains that differ in susceptibility to obesity-induced diabetes.

C57BL/6 *leptin*^{ob/ob} (B6^{ob/ob}) mice are obese but essentially resistant to diabetes, whereas BTBR *leptin*^{ob/ob} (BTBR^{ob/ob}) mice are severely diabetic [96]. In an F2 cohort derived from these parental strains, we have shown that the range of blood glucose, insulin levels, and body weight exceeds that of either the B6^{ob/ob} or BTBR^{ob/ob} parental strains. We went on to identify several diabetes-related QTL in this F2 sample [94,96]. In the current study, we focused on a subset of 60 F2 mice that have previously been evaluated in detail with regard to liver gene expression profiles [168] to ask if the abundances of hepatic metabolic intermediates would show sufficient heritability to enable us to map metabolic QTL (mQTL). Because we previously performed mRNA expression profiling on liver samples from this F2 sample, we were also able to investigate the potential for integrative analysis of the expression profiling and metabolite data sets.

We show that liver metabolites do map to distinct genetic regions, thereby demonstrating that tissue metabolite profiles are heritable. In addition, we show that mQTL co-localize with eQTL, suggesting common genetic regulators. Finally, as a proof of principle of the practical significance of this multi-disciplinary approach, we illustrate the construction of a specific causal network that links gene expression and metabolic changes, and demonstrate its validity by targeted gene expression analysis.

3.2 Materials and Methods

Animals

BTBR, B6, and B6^{ob/+} mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and bred at the University of Wisconsin. The lineage and characteristics of the BTBR strain have been reviewed by Ranheim, *et al.* Mice were housed in an environmentally controlled facility (12-hour light and dark cycles) and were weaned at 3 weeks of age onto a 6% fat diet (Purina; #5008). Mice had ad libitum access to food and water, except for 4-hour fasting periods before blood draws and sacrifice (by CO₂ asphyxiation). Plasma glucose levels were measured using a commercially available kit (994-90902; Wako Chemicals). Plasma insulin levels were measured by radioimmunoassay (RI-13K; Linco Research).

The facilities and research protocols were approved by the University of Wisconsin Institutional Animal Care and Use Committee.

Genotyping

Sixty F2 *leptin*^{ob/ob} mice ranging in age from 13 to 26 weeks were genotyped as previously described [96]. Mapmaker/EXP was used to compile genotype data into framework map.

RNA Collection and Microarray

Liver RNA was arrayed as described in Lan, *et al* [95]. Ten to 12-week-old male and female F2 *leptin*^{ob/ob} mice were sacrificed by CO₂ asphyxiation after a 4-hour fast.

Total RNA from sixty F2 mice using RNazol reagent (Tel-Test) and was further purified using an RNeasy kit (Qiagen). The sample labeling, microarray hybridization, washing, and scanning were performed according to the manufacturer's protocols (Affymetrix). Labeled cRNA was prepared and hybridization assay procedures including preparation of solutions were carried out as described in the Affymetrix GeneChip Expression Analysis Technical Manual. A total of 60 MOE430A and MOE430B arrays were used to monitor the expression levels of approximately 45,000 genes or ESTs. The distribution of fluorescent material on the array was obtained using G2500A GeneArray Scanner (Affymetrix). Microarray Suite (MAS) version 5.0 and GeneChip Operating Software (GCOS) supplied by Affymetrix was used to perform gene expression analysis. Expression levels of all the transcripts were estimated using the RMA algorithm [170].

Liver Metabolite Quantification

Amino acids, acyl-carnitines and organic acids were measured using stable isotope dilution techniques [15, 18, 56]. Amino acids and acyl-carnitine species were measured using flow injection tandem mass spectrometry and sample preparation methods described previously [15, 56]. Briefly, samples were equilibrated with a cocktail of internal standards, de-proteinated by precipitation with methanol, aliquoted supernatants were dried, and then esterified with hot, acidic methanol (acyl-carnitines) or *n*-butanol (amino acids). The data were acquired using a Micromass Quattro micro TM system equipped with a model 2777 autosampler, a model 1525 μ HPLC solvent delivery system and a data system controlled by MassLynx 4.0 operating system (Waters, Milford,

MA) [15, 56]. Organic acids were quantified using a previously described method that utilizes Trace GC Ultra coupled to a Trace DSQ MS operating under Excalibur 1.4 (Thermo Fisher Scientific, Austin, TX) [132].

Sixty-seven liver metabolites were measured, comprised of 15 amino acids and urea cycle intermediates, 45 acyl-carnitine derivatives, and 7 organic acids (TCA cycle intermediates and related analytes). The specific metabolites are listed in Table 1 and Table 2. All MS analyses employed stable-isotope-dilution. The standards serve both to help identify each of the analyte peaks and provide the reference for quantifying their levels. Quantification was facilitated by addition of mixtures of known quantities of stable-isotope internal standards from Isotec (St. Louis, MO), Cambridge Isotope Laboratories (Andover, MA), and CDN Isotopes (Pointe-Claire, Quebec, CN) to samples, as follows: Acyl-carnitine assays--D₃-acetyl, D₃-propionyl, D₃-butyryl, D₉-isovaleryl, D₃-octanoyl, and D₃-palmitoyl carnitines; Amino acid assays--¹⁵N₁, ¹³C₁-glycine, D₄-alanine, D₈-valine, D₇-proline, D₃-serine, D₃-leucine, D₃-methionine, D₅-phenylalanine, D₄-tyrosine, D₃-aspartate, D₃-glutamate, D₂-ornithine, D₂-citrulline, and D₅-arginine; Organic acid assays--D₃-lactate, D₃-pyruvate, ¹³C₄-succinate, D₂-fumarate, D₄-glutarate, ¹³C₁-malate, D₆-α-ketoglutarate, and D₃-citrate. In addition to mass, analytes are identified on the basis of the particular MS/MS transitions that we monitor for each class of metabolites. For example, all acyl-carnitine methyl esters produce a fragment m/z 99. We make the assumption that all even mass precursors ions of m/z 99 are acyl-carnitines to which we assign plausible molecular structures. We differentiate isobaric structures e.g., dicarboxylic and hydroxylated acyl-carnitines, by comparing of MS/MS spectra for

precursors of m/z 85 butylated acyl-carnitine species. We can infer whether the original compound had one or two carboxyl groups on the basis of the mass change from methyl to butyl esters.

Given our sample size, we initially analyzed metabolite abundance by hierarchical clustering using the distance function 1-correlation [40, 57-60]. Pairwise Spearman correlation coefficients of $r > 0.254$ and $r > 0.330$ reflected p-values $p < 0.05$ and $p < 0.01$, respectively. To test whether the 15 amino acids are significantly correlated as a group, groups of 15 metabolites were permuted 1,000 times and the percentage of pairwise correlations exceeding 0.5 was recorded for each group. The 15 amino acids cluster significantly as a group based on 1,000 permutations ($p < 0.001$).

QTL Analysis

Detection and mapping of QTL was performed as previously described [9, 21]. Briefly, genotypes of 512 F2 mice at 293 markers were assembled using MAPMAKER/EXP [171]. A previously established subset of 60 mice with transcript data was used for expression QTL analysis [168]. Interval mapping methods adjusted for sex as implemented in R/qtl were used to compute linkage to the traits of interest and to investigate mode of inheritance [172]. The traits included the 45,265 probe sets surveyed by microarray analysis, and the 67 liver metabolites assayed by MS methods.

We used standard interval mapping implemented in R/qtl to map each of the transcripts and liver metabolites at 1-cM resolution with age as additive covariates and sex as both additive and interactive covariates [172]. A LOD threshold of 5.0 is required

to reach a level of $p < 0.05$ in this data set with sample size 60 based on 10,000 permutations. We used threshold of 3.0 in order to highlight genetic regions to which groups of metabolites map. To visualize regions of mQTL co-localization in highly correlated metabolites (Figure 23), we constructed heat maps where metabolites are ordered as in hierarchical clustering using 1-correlation, as in Figure 22. When mice with the B6 allele at a marker have greater levels of metabolites on average than mice with the BTBR allele at that marker, the LOD score at that marker is multiplied by -1. This adjustment allows us to visualize whether the B6 or BTBR allele results in elevated metabolite abundance.

Networks

Causal networks were constructed using the methods described in Chapter 2 and in Chaibub, *et al* [147]. Although the network has the ability to accommodate 100 or more transcripts, we chose a limited number of transcripts passing several selection filters. The transcripts for the glx network were derived from the top 250 most correlated transcripts ($p < 0.002$) according to the WebQTL software (www.genenetwork.org). A hypergeometric test was performed and identified the GO term category “metabolism” as one of the two processes significantly enriched by these correlates ($p < 0.004$). Transcripts were chosen from this category, with an additional requirement being that they have at least one eQTL overlapping with the glx mQTL (Table 5). QTL in the genetic region encompassing a 1.5 LOD support interval around LOD peaks that are at least 3.0 are also included [40, 41]. Based on 10,000 permutations for each of the transcripts, the LOD

threshold is significantly higher to reach significance ($\text{LOD} > 5.0$ is required for $p < 0.05$), but the 3.0 threshold was used include major and minor putative QTL [9, 23].

If more than one probe set was used to identify a transcript of interest, only probe sets with a grade A annotation by Affymetrix were considered. For these probe sets, only those with all eleven oligonucleotides aligning (via BLAST) to their appropriate target sequence provided by the National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) were considered acceptable. If more than one primer set still identified the transcript, an average of the probe sets in the network.

We built an undirected dependency graph (UDG) of order 6 with *glx* and these transcripts as nodes with a two-tailed significance level of 0.05 [173]. We remove edges that are based on spurious or partial correlations, and then orient causal edges between all pairs of connected phenotypes using associated multiple QTLs to break likelihood equivalence. Quantitative trait loci for *glx* and the selected transcripts were identified with R/qtl [172] using a 3.0 LOD cutoff; the marker closest to each peak provided key information for inferring causal direction. We oriented phenotype edges using our QTL-directed dependency graph (QDG) approach. For any two phenotypes connected by an edge, the direction LOD score was computed by regressing these phenotypes on each other and on their respective multiple QTLs, adjusting for age and for QTL-sex interactions, and by other phenotypes that might be directly connected to either phenotype by an UDG edge. Our QDG algorithm used random starts to converge to possible solutions. The best solution was determined by an approximate Bayes factor (BF)[65, 66].

We estimated network parameters from the true data and simulated synthetic data according to the causal network in Figure 25. We simulated 1,000 realizations from the causal network and for each edge, we recorded the percentage of undirected edges recovered by the UDG algorithm and the percentage correctly inferred direction by the orientation steps of the QDG algorithm. Overall, the average percentage of true recovered edges was 75% and the average percentage of correctly inferred direction was 83%. False edges were detected at a rate below 2%. To calculate the false discovery rate for the network, we simulated 1,000 data sets from the true network. For each data set, the UDG algorithm was used to infer the network topology, and computed the fraction of false edges (those detected that do not exist in the true network) relative to the total number edges detected by the UDG algorithm. The FDR for the network topology, computed as the average fraction for these 1,000 simulations, is 0.014.

In Vitro Hepatocyte Experiments

Hepatocytes from 10-week-old lean male B6 and BTBR parental strain mice (n=5 for each genotype) were isolated by liver perfusion [174]. Hepatocytes were seeded at subconfluency (3.5×10^6 cells/6 well plate) in low glucose DMEM (Gibco) supplemented with FBS (10% vol/vol; Gibco), pen/strep antibiotic (1%, Gibco), glutamine (2mM; Gibco), and pyruvate (1mM; Gibco). Cells were left to attach for 3 hours in an incubator at 37°C, 5% CO₂. After washing with PBS, the cells were treated with unsupplemented DMEM (Sigma) with 1g/L glucose, pen/strep (1%), +/- 10mM glutamine. Cells were treated for 24 hours.

RNA was extracted from hepatocytes using RNeasy kits (Qiagen) after treatment described above. Hepatocytes in 6-well plates were homogenized in 0.35 mL RLT buffer and stored at -80 C. RNA was purified using RNeasy-mini columns (Qiagen) according to the manufacturer's directions. The ratio of the optical densities from RNA samples measured at 260 and 280 nm was used to evaluate nucleic acid purity and total RNA concentrations were determined by the absorbance at 260 nm. The quality of total RNA was estimated based on the integrity of 28S and 18S rRNA separated using 1% agarose gel electrophoresis.

Gene expression was measured using a 7500 fast real-time PCR system (Applied Biosystems). cDNA was synthesized from 1ug of total RNA using the SuperScriptIII first-strand cDNA synthesis kit (Invitrogen) primed with a mixture of oligo-dT and random hexamers. Primers were obtained from Integrated DNA Technologies and MWG Biotechnology. The SYBR Green PCR core reagent kit (Applied Biosystems) was used to determine relative expression. The housekeeping gene *Actb* was used as a normalization control. Primer sequences for transcripts of the glx network are provided in Table 3.

Gene Expression Analysis

The fold changes relative to the untreated hepatocytes for each animal were calculated. An overall ANOVA analysis was performed with gene transcripts nested within subject; interest focused on gene transcript effects and possible gene transcript differences between strains. This analysis showed that glutamine-induced expression

change differed by gene ($p < 0.0001$). Significant overall gene transcript effects allows separate transcript-specific paired t-tests between the difference in delta CT values of untreated and glutamine-induced gene expression (relative to *Actb*) in each strain separately. Statistics on these data were analyzed with Prism software version 4.02 (Graph Pad Software) and the aov command in R (www.r-project.org).

3.3 Results

Individuals of the F2 population segregate by physiological phenotype

At 10 weeks of age, the BTBR^{ob/ob} mouse has severe diabetes compared to the diabetes-resistant B6^{ob/ob} mouse, manifesting fasting hyperglycemia and insulin resistance. To investigate the genetic contribution underlying these phenotypic differences in the parental strains, we performed breeding studies to generate a segregating F2 *leptin*^{ob/ob} population. Individuals in the F2 population exhibited a wide range of “clinical phenotypes,” which we defined by traditional markers such as blood glucose, plasma insulin levels, plasma triglycerides, and body weight. In fact, the range in these markers in the sixty individuals of the F2 population was greater than the range represented by the parental strains (Figure 21). This suggests that we have animals with extreme phenotypes due to additive effects of alleles at distinct loci [94,96].

Metabolites of similar function are highly correlated across the F2 population

We determined the concentration of 67 liver metabolites, comprised of 15 amino acids and urea cycle intermediates, 45 acyl-carnitines, and 7 organic acids (TCA cycle intermediates and related metabolites) in the F2 sample.

We created a correlation matrix of all pairwise comparisons among individual metabolites. Unsupervised hierarchical clustering revealed “hot spots” of highly correlated metabolites (Figure 22). It is striking that several hot spots correspond to the biochemical pathway to which the metabolites belong. For example, 12 of the 15 amino acids cluster in this matrix. Moreover, when we consider pairwise correlations between all amino acids, 75% had absolute correlation coefficients greater than 0.5 ($p < 0.01$) (Table 4). Permutation analysis of these pairwise correlations confirm that the 15 amino acids correlate as a functional group ($p < 0.001$). Several specific acyl-carnitine derivatives are also clustered, such as hexadecadienoyl carnitine (C16:2), 3-hydroxy-tetradecanoyl carnitine or dodecenedioyl carnitine (C14:1-OH/C12:1-DC), and 3-hydroxy-palmitoleoyl carnitine or *cis*-5-tetradecenedioyl carnitine (C16:1-OH/C14:1-DC). The fact that metabolites of a common functional group are highly correlated suggests that there are potential regulators of these biochemical pathways segregating in this F2 sample.

In another cluster, pyruvate correlates most highly with alanine ($r = 0.53$, $p < 0.01$), and also with lactate and tiglyl carnitine (C5:1) ($p < 0.01$). Alanine and short-chain acyl-carnitines are products of peripheral protein and fatty acid catabolism, respectively, and are delivered to the liver. The liver uses alanine, along with pyruvate and lactate, as gluconeogenic substrates and rapidly interconverts these metabolites through

transamination and oxidation/reduction. The clustering of these metabolites based on their relative concentration in F2 animals suggests that static metabolic profiling can be used as a marker for changes in flux through certain metabolic pathways.

It has been demonstrated that mRNA abundance, as determined with microarray technology, is sufficiently heritable to map QTL [95,150,159,160,167-169,175]. Lan, *et al.* showed that using expression mapping, specifically in this F2 intercross, can uncover mechanisms that explain correlations between specific transcripts [95]. We therefore sought to determine if metabolite abundance, as measured in F2 liver samples by mass spectrometry, was similarly heritable. If so, resulting metabolic QTL (mQTL) could be integrated with expression QTL (eQTL) to form network models of gene expression that might ultimately help to explain diabetes susceptibility and resistance in the BTBR^{ob/ob} and B6^{ob/ob} strains, respectively [176,177].

We found that individual metabolites mapped to specific regions of the genome. By permutation analysis, 21% of the metabolites map significantly to genomic regions (LOD>5.0, p<0.05), indicating those genomic regions could potentially influence (either directly or indirectly) the abundance of these metabolites. We used LOD threshold of 3.0 to investigate both major and minor putative mQTL where groups of metabolites map. Figure 23 displays a heat map, with metabolites organized by hierarchical clustering as in Figure 22. The 12 amino acids that clustered based on correlation (citrulline, tyrosine, and alanine are the exceptions) map to common mQTL, e.g., an overlapping region of chromosome 9. Amino acids that act together in specific pathways show additional common mQTL. For example glx (glutamine+glutamate) and urea cycle intermediates

arginine, asx (asparagine+aspartate), and ornithine, map to a common region of chromosome 7. The gluconeogenic substrates alanine and pyruvate have a mapping profile distinct from the majority of amino acids in that they lack the prominent mQTL on chromosome 9 (Figure 23). This unique alanine/pyruvate mQTL may explain why alanine clusters with pyruvate rather than the amino acids in the correlation matrix (Figure 22).

Expected and novel correlations between transcripts and metabolites

The foregoing results demonstrate that metabolites of a functional class often are correlated with one another and have common mQTL. To better understand how gene expression and metabolites are related, we adopted the approach used by Carrari, *et al.* [178] and created a correlation matrix between liver metabolites and selected liver transcripts of our 60 F2 mice. Three categories of transcripts were chosen, based on gene ontology terms relating to the biological process in which they play a role: 1) carbohydrate metabolism (glucose metabolism, gluconeogenesis, glycolysis, carbohydrate biosynthesis, TCA cycle, glucose transport, and glycogen metabolism); 2) lipid metabolism (fatty acid biosynthesis, fatty acid oxidation, steroid metabolism, cholesterol metabolism and biosynthesis, and lipid biosynthesis); and 3) protein metabolism (urea cycle, amino acid biosynthesis, protein catabolism, and amino acid transport). We organized the metabolites into functional classes to reveal whether biochemical groups of metabolites correlated in a specific pattern with transcripts of a particular pathway (Figure 24).

We found evidence for correlations among functionally similar metabolites and transcripts when organized by biological process. For example, several long-chain acyl-carnitine species show a positive correlation with groups of transcripts involved in glycolysis, fatty acid biosynthesis, steroid metabolism, cholesterol metabolism, and lipid biosynthesis. In contrast, a subset of medium-chain acyl-carnitines and short-chain acyl-carnitines exhibit a negative correlation to these same individual transcripts. These findings are consistent with recent studies from our laboratories showing that long-chain acyl-carnitines accumulate in muscle of animals with diet-induced obesity at the expense of short-chain acyl-carnitines, and that this abnormality is resolved when obese animals are exercised [165].

The 15 amino acids displayed a common correlation pattern with mRNA transcripts in pathways of protein metabolism, as well as glycolysis, the TCA cycle, and several lipid metabolism transcripts. These amino acids are very tightly correlated with one another, leading us to investigate the role played by individual transcripts in control of amino acid abundance. Our data show that two very highly correlated metabolites often correlate with the same set of individual transcripts. However, we also see that within this metabolite group, subsets of amino acids will have a unique transcript correlation pattern. For example, thirteen of fifteen amino acids correlate ($r > 0.35$, $p < 0.01$) with *Slc38a3*, a sodium-dependent transporter that mediates entry of a select group of amino acids across the plasma membrane. There are pathways by which the few known *Slc38a3* amino acid substrates (alanine, asparagine, histidine, and glutamine) could serve as precursors for biosynthesis of non-substrate amino acids that also correlate

with this transporter [179,180]. In contrast, only valine and leucine+isoleucine correlate as highly ($r>0.35$, $p<0.01$) with *Ppargc1a* mRNA, and could represent a unique metabolic pathway involving the branched-chain amino acids.

Correlations and co-mapping of transcripts and metabolites produce causal network models

One hypothesis that follows from our results is that unique genetic regulators could affect the abundance of clusters of metabolites. Unlike mRNA transcripts, metabolites can be interconverted with other metabolites, generating a cluster to which the precursor metabolite will be highly correlated [181]. The downstream product metabolites will also be correlated with the regulatory transcript and co-map with the eQTL of the regulatory transcript [138,159].

Glutamate is a substrate and product in amino acid catabolic and biosynthetic pathways. Glutamate can act either as an ammonium donor or acceptor in transamination reactions (via α -ketoglutarate) and the glutamate dehydrogenase reaction, and can also be rapidly synthesized from glutamine via glutaminase, thus providing precursor metabolites for the generation of other organic acids and amino acids. Glutamine can also act as a signaling molecule to alter expression of urea cycle and gluconeogenic enzymes [55-58,76]. Given that glutamine and glutamate (glx) can generate a network of related metabolites and can also change gene expression, we focused on glx as the start-point for building a proof-of-principle causal network from the F2 liver expression and metabolite profiling data sets. We generated a network featuring glx and a limited number of transcripts that passed multiple, stringent selection filters (see materials and methods).

This provided a testable network that would enable us to gain insights into metabolite-transcript relationships.

Transcript nodes of the network are highly correlated to glx ($p < 0.05$ by 10,000 permutations) as well as other amino acids (Table 5, Table 6). Table 5 depicts the overlap of the glx mQTL interval and the physical location of the transcripts or their eQTL encompassing a 1.5 LOD support interval around LOD peaks that are at least 3.0 [182,183]. We note that glx is correlated with mRNA of two transporters: sodium-dependent amino acid transporter *Slc38a3* and glutamate transporter *Slc1a2*, whose genes are located on chromosomes 9 (102.5Mb) and 2 (107.5Mb), respectively. Additionally, the glx mQTL on chromosome 9 spans a region containing *Slc38a3* and the mQTL on chromosome 2 and 9 overlaps with the eQTL of *Slc1a2* (Table 5). We hypothesize that both *Slc1a2* and *Slc38a3* could mediate the entry of glx into liver cells, but that *Slc1a2* may also have expression regulated by glx abundance.

Table 5 also shows that glx is significantly correlated to *Argininosuccinate synthetase 1 (Ass1)*, *Arginase 1 (Arg1)*, *Phosphoenolpyruvate carboxykinase 1 (Pck1)*, *Isovaleryl coenzyme A dehydrogenase (Ivd)* and *Alanine:glyoxylate aminotransferase (Agxt)* mRNAs. The physical location and/or mapping location of these transcripts with respect to the glx mQTL indicates that the metabolite-transcript relationship may go beyond correlation. For example, on chromosome 2, we see that the glx mQTL co-maps with the eQTL for *Agxt*, *Arg1*, *Ass1*, and *Ivd* [183]. This is consistent with network models in which the QTL regulates glx, which then regulates gene expression or

conversely, the QTL regulates mRNA abundance of the four transcripts, which then regulate glx [135].

Using the method described by Chaibub, *et al.*[147], we generated a causal network consisting of glx and these highly correlated transcripts (FDR=0.014), incorporating mQTL and eQTL to determine directionality between the nodes (Figure 25). This network model predicts that modulation of glutamine and/or glutamate levels should lead to a change in the expression of *Agxt*, *Arg1*, and *Pck1*. To test this prediction, we isolated hepatocytes from lean B6 and BTBR parental strains and measured changes in gene expression as a result of addition of 10 mM glutamine to the cultured cells. Glutamine exposure changed transcript abundance, and no transcript-specific strain differences in glutamine effect on gene expression were found (p=0.53) (Figure 26). Glutamine significantly increased expression of *Agxt*, *Arg1*, *Pck1*, and *Ass1* in both strains (p<0.05 for both strains); the increases in *Pck1* and *Ass1* confirm prior studies [56-58,76].

Given its role as a glutamate transporter, it is not surprising that *Slc1a2* is upstream of glx in the best proposed causal network (BF=163) (Figure 25, solid lines). However, glutamine exposure in vitro reduced *Slc1a2* expression in isolated hepatocytes from either mouse strain, supporting the second-best causal network solution (Figure 25, dotted lines). Glutamine also reduced *Ivd* expression in the B6 strain but showed no effect in the BTBR strain, despite *Ivd* being upstream of glx in our best causal network. Our causal network predicts *Slc38a3* should be unchanged by glutamine treatment. Our hepatocyte experiments confirm this prediction (Figure 26). *Argininosuccinate lyase*

(*Asl*), which is neither correlated nor co-maps with *glx*, served as a negative control and indeed was not altered by glutamine treatment.

3.4 Discussion

Genomics, transcriptomics, proteomics, and metabolomics have delivered large arrays of data, allowing one to correlate physiological states with patterns of gene expression, protein levels, and metabolite abundance. A major challenge in the analysis and interpretation of this data is delivering models of causation from correlations [97,135]. Mouse models of diabetes provide a unique method for exploring correlation structure since metabolic dysregulation creates a window for simultaneous application of multiple “omic” technologies.

We have previously shown that diabetes traits show strong heritability in an F2 intercross between the diabetes-resistant B6^{*ob/ob*} and the diabetes-susceptible BTBR^{*ob/ob*} mouse strains. We assume that the disease phenotype is brought about by a complex pattern of gene expression changes in key tissues [94,96]. However, we also recognize the complexity inherent in discriminating the gene expression changes that cause diabetes from those that occur as a consequence of the disease. For example, many genes are known to be responsive to elevated blood glucose levels [184]. Through correlation alone, it is difficult to distinguish these “reactive” genes from ones that are “causal” for the disease.

We have taken advantage of the high heritability of mRNA abundance phenotypes, and via microarray technology, have mapped gene loci controlling gene expression at the genome-wide level [95]. This establishes at least one node in a network simply because genetic variation leads to changes in gene expression and not *vice versa*. However, it does not establish whether the link between a locus and a phenotype is direct or via multiple steps and pathways [150,185].

The purpose of the current study was to explore the possibility that the levels of metabolites in tissues are sufficiently heritable in an F2 intercross to provide significant linkage signals, leading to metabolic QTL. Given that many pathways converge upon common metabolites and that these pathways have multiple controllers, any one genetic locus may not alter metabolite levels significantly, and therefore may not be identified as a metabolite QTL. Nonetheless, in our F2 sample, we found significant linkage signals, including some that are quite strong (*e.g.* tyrosine: LOD>7, $p < 0.005$; chromosome 2).

Our results reveal that metabolites can be mapped to distinct genetic regions, much like mRNA transcripts. Although QTL mapping in an F2 sample does not provide sufficient resolution to identify individual genes with high certainty, it can yield novel information about regulatory networks. Phenotypes mapping to the same locus can be hypothesized to be co-regulated by that locus. With our definition of “phenotype” now including transcripts, metabolites, and physiological traits, we can begin to devise relationships between these phenotypes and genetic regions.

This F2 study provides evidence of co-regulation of biologically related pathways. An example is the correlations we found between amino acids and short-chain

acyl-carnitine derivatives. These findings are consistent with our understanding of metabolic physiology. In a catabolic, “glucose starved” state, muscle degrades proteins and delivers amino acids to the liver for glucose production. The liver transaminates amino acids to corresponding α -keto acid gluconeogenic substrates. Alpha-ketoglutarate is often the α -keto acid acceptor for these transaminase reactions, generating glutamate as a product. Glutamate, which can also be generated from glutamine in the glutaminase reaction, is then deaminated to produce ammonia by glutamate dehydrogenase, to be fixed through the urea cycle. Additionally, hepatic fatty acid oxidation and amino acid catabolism yield even and odd-numbered short-chain acyl-CoAs, which can be used for fuel and for production of ketone bodies. These short-chain acyl-CoA species are readily converted to the cognate carnitine esters, which we have profiled by MS/MS in this study.

The amino acid metabolites provide the most striking evidence of functional clustering. We see in both the correlation matrix (Figure 22) and the genetic linkage data (Figure 23) that the majority of amino acids group together. However, a subset of the amino acids, asx, glx, arginine, and ornithine uniquely map to chromosome 7. Our data predict that these metabolites are driven by different genetic regulators, leading to a unique mapping signature, even within a group of highly correlated metabolites. The C/EBP transcription factors have been shown to alter expression of enzymes acting in the urea cycle and gluconeogenic pathway [170,186-191], and the C/EBP α isoform is encoded on chromosome 7. Although we cannot determine that metabolites are mapping

to the same individual genes, we can identify genetic regions that coordinate groups of metabolites and transcripts and contain plausible candidate genes.

The relationship between mRNA transcripts and metabolites, however, can be bi-directional. Our network identifies a specific metabolite, glx that regulates gene expression. This is consistent with previous studies where glutamine alone increases hepatic expression of *Argininosuccinate synthetase* and *Phosphoenolpyruvate carboxykinase*, but when combined with other essential amino acids, alters additional transcripts of urea cycle and gluconeogenic pathways [56,57,76,77]. Our work extends these prior observations by showing that glutamine also changes expression of *Agxt*, *Arg1*, *Ivd*, and *Slc1a2*, but does not alter *Slc38a3*, despite the positive correlation with this transcript. The combination of pathway construction based on transcriptional and metabolic profiling and direct model testing in living cells provides evidence for a new pathway by which glx can regulate a key gluconeogenic enzyme. Future studies will be needed to investigate if this pathway is perturbed in development of diabetes.

The glutamine-induced reduction in *Slc1a2* expression was unexpected given that this glutamate transporter is upstream of glx in the best-proposed causal network (Figure 25, solid lines). *Slc1a2* mRNA abundance, however, maps in *trans* (to a locus distinct from the physical location of the gene) to chromosome 9, its eQTL overlapping with the glx mQTL. It is therefore possible that glutamine could regulate *Slc1a2*, as indicated by the second causal network (Figure 25, dotted lines). Several studies have shown that *Slc1a2* expression in astrocytes is reduced by increased ammonia [186-188,191-194]. Despite the positive correlation between *Slc1a2* and glx *in vivo*, the glutamine-treated

hepatocytes produce ammonia via glutaminase, and could decrease expression of hepatic *Slc1a2 in vitro*. We also did not predict altered expression of *Ivd*, an enzyme of leucine oxidation. It is interesting to note that *Ivd* is a case where a gene maps both in *cis* (to the locus containing the *Ivd* gene) and in *trans*, here overlapping with the glx mQTL on chromosomes 2 and 13. Studies have shown that glutamine has an inverse relationship with leucine oxidation, and this could be mediated by glutamine-induced decreased *Ivd* expression [189,190].

We show that the combined use of eQTL and mQTL, with correlations allows one to derive a network and establish data-driven hypotheses about metabolite and gene expression relationships. For example, glycine and serine are the two amino acids most highly correlated with glx, and the transcript most highly correlated with glx is *Agxt* (Table 4, Table 5). Indeed, in our experiments, *Agxt* was upregulated by glutamine. We hypothesize that the upregulation by glx of *Agxt* is one mechanism by which glx is correlated with glycine and serine since *Agxt* catalyzes the transamination of glyoxalate to form glycine, which can then be converted to serine. In further support of this hypothesis, in the F2 sample, serine and glycine correlate ($r > 0.5$, $p < 0.01$) to *Agxt*.

The concurrent use of transcriptomics and metabolomics is not limited to one biochemical pathway. For example, the correlation between amino acids and transcripts of carbohydrate and lipid metabolism might reflect a broader signaling function of amino acids beyond pathways of protein metabolism. Furthermore, this correlation, co-mapping, and causal network analysis can uncover roles for transcripts of unknown function. We note Riken clones and ESTs are among the transcripts highly correlated to individual

metabolites. By incorporating these transcripts of unknown function as nodes into causal networks, along with transcripts from known pathways, we may infer the functions of these previously unidentified mRNA species.

In conclusion, this study shows that metabolites, in addition to transcripts and physiological traits, can be mapped to genetic regions, providing a powerful tool to establish connections between genetic loci and physiological traits. The groups of metabolites and transcripts that are correlated or co-map to physiological traits in our F2 sample may offer insight into metabolic pathways that are causal or reactive to diabetes pathology.

Table 1. Metabolite Codes (Acyl-carnitines)

Metabolite Symbol	Trivial name(s)
C2	Acetyl carnitine
C3	Propionyl carnitine
C4-DC/C4-DC	Methylmalonyl carnitine or Succinyl carnitine
C4/Ci4	Butyryl carnitine or Isobutyryl carnitine
C4-OH	3-Hydroxy-butyryl carnitine, b-hydroxy butyryl carnitine
C5:1	Tiglyl carnitine
C5's	Isovaleryl carnitine, 3-methylbutyryl carnitine or 2-Methylbutyryl carnitine
C5-DC	Glutaryl carnitine
C5-OH/C3-DC	3-Hydroxy-isovaleryl carnitine or Malonyl carnitine
C6	Hexanoyl carnitine
C6-DC	Adipoyl carnitine
C7-DC	Pimeloyl carnitine, heptanedioyl carnitine
C8	Octanoyl carnitine
C8:1	Octenoyl carnitine
C8:1-DC	Octenedioyl carnitine
C8:1-OH/C6:1-DC	3-Hydroxy- <i>cis</i> -5-octenoyl carnitine or Hexenedioyl carnitine
C10:3	Decatrienoyl carnitine
C10:2	Decadienoyl carnitine
C10:1	Decenoyl carnitine
C10	Decanoyl carnitine
C10-OH/C8-DC	3-Hydroxy-decanoyl carnitine or Suberoyl carnitine
C12:1	Dodecenoyl carnitine
C12	Lauroyl carnitine
C12-OH/C10-DC	3-Hydroxy-dodecanoyl carnitine or Sebacoyl carnitine
C14:2	Tetradecadienoyl carnitine
C14:1	Tetradecenoyl carnitine
C14	Myristoyl carnitine
C14:1-OH	3-Hydroxy-tetradecenoyl carnitine
C14-OH/C12-DC	3-Hydroxy-tetradecanoyl carnitine or Dodecanedioyl carnitine
C16	Palmitoyl carnitine
C16-OH/C14-DC	3-Hydroxy-hexadecanoyl carnitine or Tetradecanedioyl carnitine
C16:2	Hexadecadienoyl carnitine
C16:1	Palmitoleoyl carnitine
C16:1-OH/C14:1-DC	3-Hydroxy-palmitoleoyl carnitine or <i>cis</i> -5-Tetradecenedioyl carnitine
C18:2-OH	3-Hydroxy-linoleyl carnitine
C18:2	Linoleyl carnitine
C18:1	Oleoyl carnitine
C18	Stearoyl carnitine
C18:1-OH	3-Hydroxy-octadecenoyl carnitine
C18-OH/C16-DC	3-Hydroxy-octadecanoyl carnitine or Hexadecanedioyl carnitine, thapsoyl carnitine
C18:1-DC	Octadecenedioyl carnitine
C20	Arachidoyl carnitine, eicosanoyl carnitine
C20-OH/C18-DC	3-Hydroxy-eicosanoyl carnitine or Octadecanedioyl carnitine
C22	Behenoyl carnitine, docosanoyl carnitine
C20:4	Arachidonoyl carnitine

Table 2. Metabolite Codes (Amino Acids and Organic Acids)

Metabolite Symbol	Trivial name(s)
Ala	Alanine
Arg	Arginine
Asx	Asparagine+Aspartate
Cit	Citrulline
Citrate	Citrate
Fumarate	Fumarate
Glutarate	Alpha-Ketoglutarate
Glx	Glutamine+Glutamate
Gly	Glycine
Lactate	Lactate
Leu.Ile	Leucine+Isoleucine
Malate	Malate
Met	Methionine
Ornithine	Ornithine
Phe	Phenylalanine
Pro	Proline
Pyruvate	Pyruvate
Ser	Serine
Succinate	Succinate
Tyr	Tyrosine
Val	Valine

Table 3. PCR Primer Sequences.

Gene Symbol	Gene Accession	Forward	Reverse
Actb	NM_007393	ACTATTGGCAACGAGCGGTT	ATGGATGCCACAGGATTCCA
Agxt	NM_016702	CCCCTGAGGAGAGGGTGCT	GCCACACGGTCCACATTCTC
Arg1	NM_007482	CATGGGCAACCTGTGTCCTT	CAGGGAGTCACCCAGGAGAA
Asl	NM_133768	TGATGTGAGTCACGTGTGGGA	TTCGACACTGGATTTGCTGT
Ass1	NM_007494	CAACAGCCCAGATGTCCTGA	CGGGTTGTGCCATCTTTGAT
Gls2	XM_125928	CAGGAGCCATCCTCCTGGTT	GGCTGTTCCCCAGCTTGTCT
Glud1	NM_008133	TTGGGCTCCCTGTTTGGAGT	AAGCCTTGAAGGGGACACA
Got1	NM_010324	GCCCTGGGTTTTGCCAGTAG	TGGGCAGGTACTCGTGGTTG
Got2	NM_010325	AGCGGCTGACCAAGGAGTTC	CCCACATTGCCAGAGGTGAC
Gpt1	NM_182805	GGTCCTGGCCCTCTGTGTCT	TGCGTTCTGCCCTTCTCTTG
Gpt2	NM_173866	CAGGTTTGCCTGAGCCTTGA	CCTCCATTTTCAGCCCTTGC
Ivd	NM_019826	CATGTGAGGGAAGCCTTTGG	CATGAGGCGGGTGTACATGTC
Oat	NM_172478	GAGGAGGAGCGTGAGGTGGT	ACCGAAAAGAGCGGAACTC
Pck1	NM_011044	CAACCTGAGCTGACCCTGA	TTCCCACCATATCCGCTTCC
Slc1a2	NM_001077514	CCTTTCCAAGTCTGAGCTGGA	AATGGACTGCGTCTTGGTCA
Slc38a3	NM_023805	GCCAATACGGGCATCATCCT	TGAGGAGCAGGTGGATGGAA

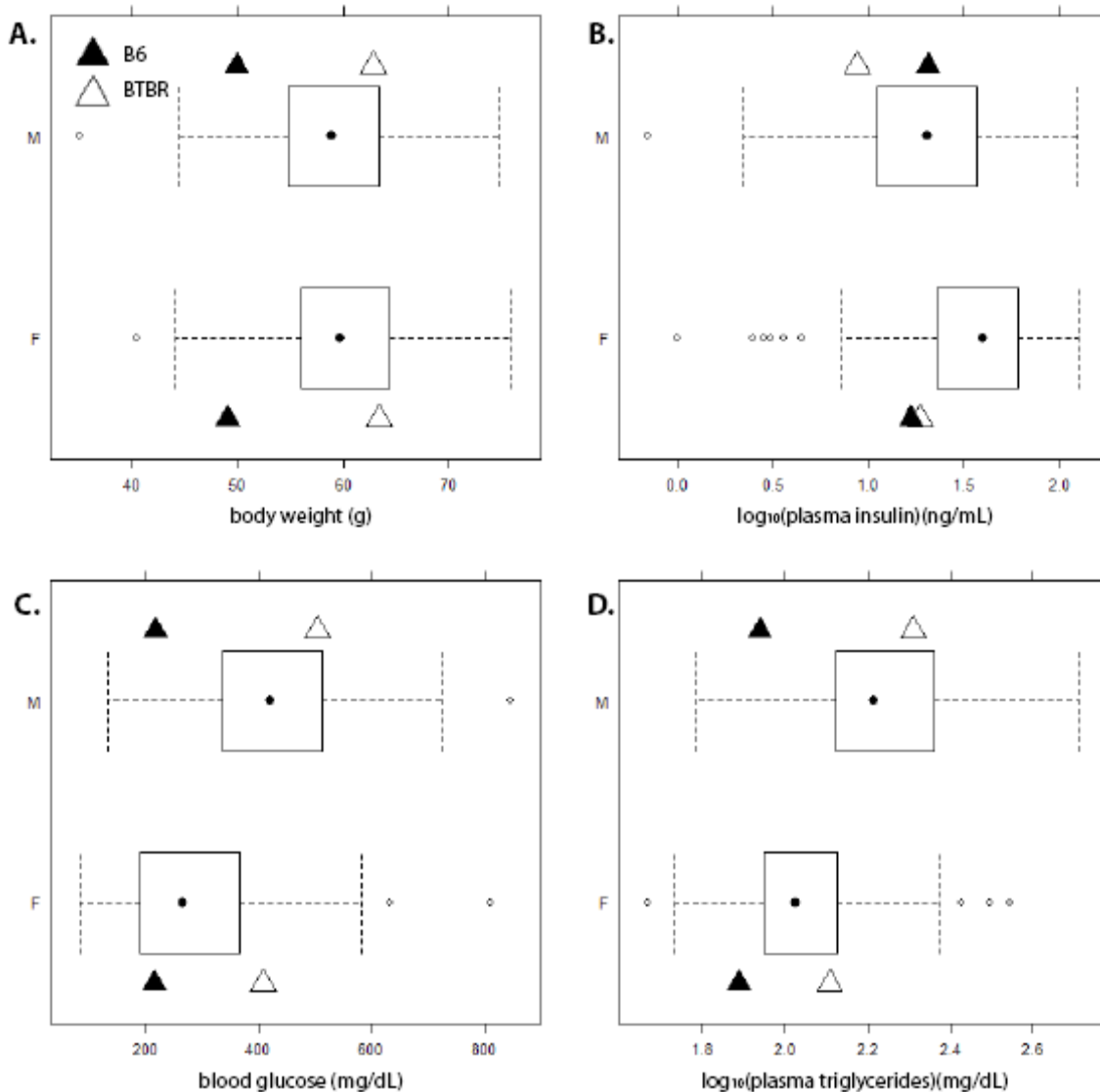


Figure 21. Clinical phenotypes of the leptin-deficient F2 population.

The F2^{ob/ob} population shows a range of clinical phenotypes commonly associated with diabetes: body weight (A), plasma insulin (B), fasting blood glucose (C), and plasma triglycerides (D). Within this population are individuals with phenotypes more extreme than either the C57BL/6^{ob/ob} (black triangle) or BTBR^{ob/ob} (white triangle) parental strains. Graphically depicted are boxplots for each phenotype; the boxes represent 50% of the population and bars are ± 1.5 standard deviations.

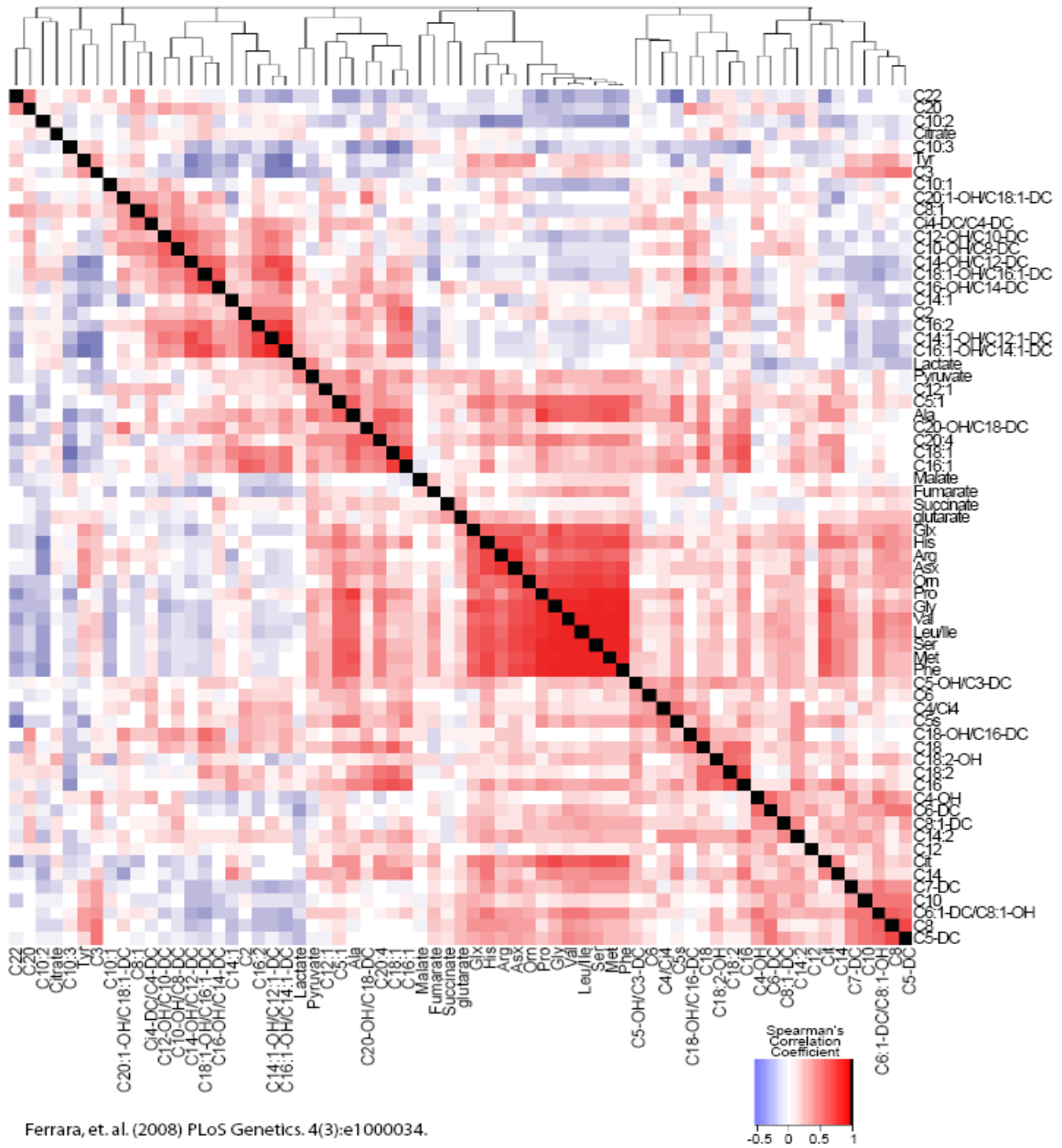


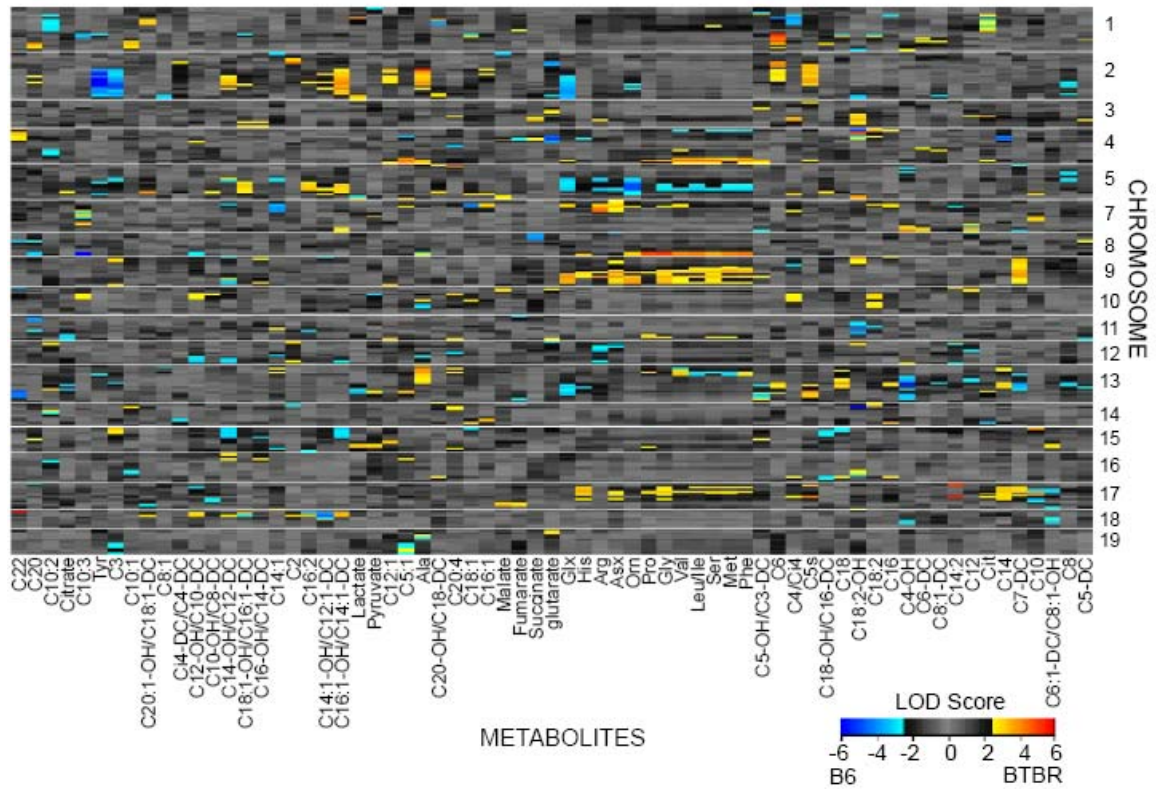
Figure 22. Heat map of correlations between liver metabolites.

Each square represents the Spearman's correlation coefficient between the metabolite of the column with that of the row ($|r| > 0.254$, $p < 0.05$; $|r| > 0.330$, $p < 0.01$). Metabolite order is determined as in hierarchical clustering using the distance function 1-correlation. Self-self correlations are identified in black. Metabolite codes are listed in Table 1 and Table 2. Individual correlation coefficients between the amino acids can be found in Table 4.

Table 4. Amino acid correlation matrix.

Pairwise correlations between the 15 amino acids were calculated and recorded as the Spearman's correlation coefficient ($|r| > 0.254, p < 0.05$; $|r| > 0.330, p < 0.01$). Metabolites are ordered according to the strength of their correlation as in Figure 22.

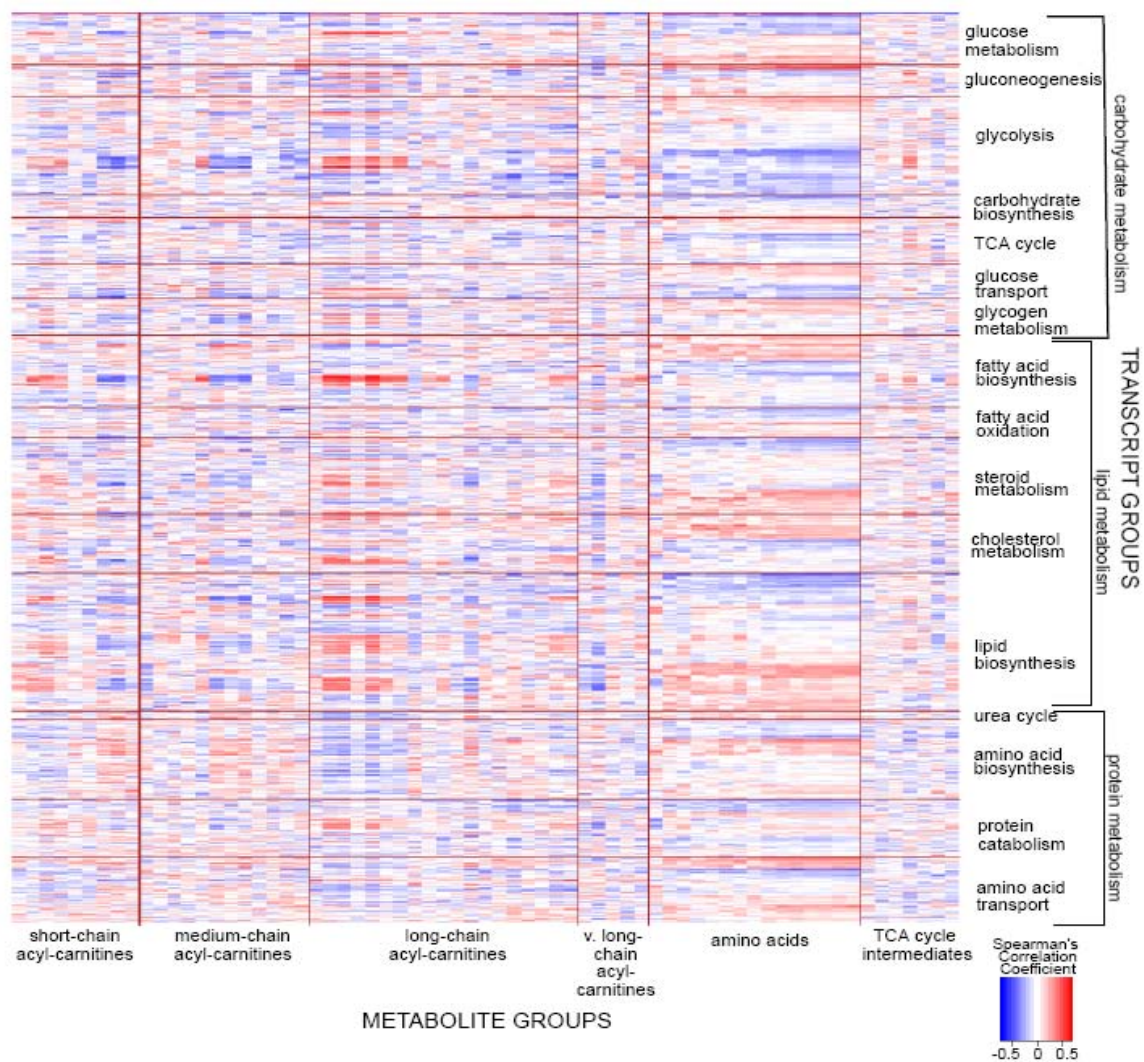
Spearman's Correlation Coefficient (Metabolite with Metabolite)		METABOLITE														
		Metabolite Symbol	Tyr	Ala	Glx	His	Arg	Asx	Orn	Pro	Gly	Val	Leu.Ile	Ser	Met	Phe
METABOLITE	Tyr	1.00	-0.24	0.37	0.25	0.42	0.38	0.43	0.05	0.15	0.15	0.26	0.20	0.18	0.21	-0.12
	Ala	-0.24	1.00	0.25	0.42	0.22	0.38	0.30	0.77	0.61	0.56	0.55	0.61	0.65	0.65	0.41
	Glx	0.37	0.25	1.00	0.62	0.58	0.56	0.61	0.48	0.71	0.59	0.65	0.65	0.58	0.61	0.46
	His	0.25	0.42	0.62	1.00	0.70	0.73	0.52	0.58	0.68	0.64	0.67	0.69	0.71	0.71	0.52
	Arg	0.42	0.22	0.58	0.70	1.00	0.85	0.56	0.38	0.56	0.53	0.60	0.61	0.57	0.59	0.30
	Asx	0.38	0.38	0.56	0.73	0.85	1.00	0.69	0.56	0.67	0.66	0.72	0.73	0.69	0.72	0.39
	Orn	0.43	0.30	0.61	0.52	0.56	0.69	1.00	0.70	0.72	0.74	0.79	0.76	0.74	0.77	0.46
	Pro	0.05	0.77	0.48	0.58	0.38	0.56	0.70	1.00	0.85	0.83	0.84	0.85	0.89	0.89	0.64
	Gly	0.15	0.61	0.71	0.68	0.56	0.67	0.72	0.85	1.00	0.88	0.90	0.92	0.90	0.90	0.74
	Val	0.15	0.56	0.59	0.64	0.53	0.66	0.74	0.83	0.88	1.00	0.96	0.92	0.91	0.93	0.61
	Leu.Ile	0.26	0.55	0.65	0.67	0.60	0.72	0.79	0.84	0.90	0.96	1.00	0.95	0.95	0.97	0.61
	Ser	0.20	0.61	0.65	0.69	0.61	0.73	0.76	0.85	0.92	0.92	0.95	1.00	0.96	0.95	0.66
	Met	0.18	0.65	0.58	0.71	0.57	0.69	0.74	0.89	0.90	0.91	0.95	0.96	1.00	0.99	0.67
	Phe	0.21	0.65	0.61	0.71	0.59	0.72	0.77	0.89	0.90	0.93	0.97	0.95	0.99	1.00	0.63
	Cit	-0.12	0.41	0.46	0.52	0.30	0.39	0.46	0.64	0.74	0.61	0.61	0.66	0.67	0.63	1.00



Ferrara, et. al. (2008). PLoS Genet. 4(3)e1000034.

Figure 23. Linkage hot spots for metabolic quantitative trait loci (mQTL).

Each row represents a marker; each column represents a metabolite. Metabolites are ordered as in hierarchical clustering using the distance function 1-correlation (as in Figure 22). The LOD color scale is indicated, showing blue (red) when the B6 (BTBR) allele at that marker results in an elevated level of metabolite. Metabolite codes are listed in Table 1 and Table 2.



Ferrara, et. al. (2008). PLoS Genet. 4(3)e1000034.

Figure 24. Heat map of correlations between liver metabolites and select liver transcripts.

Each square represents the Spearman's correlation coefficient between the metabolite of the column with the transcript of the row ($|r| > 0.254$, $p < 0.05$; $|r| > 0.330$, $p < 0.01$). Metabolites are organized into their biochemical class; transcripts are selected based on gene ontology terms relating to biological processes in which they play a role. Correlation coefficients between individual amino acids with select transcripts are found in Table 6.

Table 5. Glx network correlations and mapping.

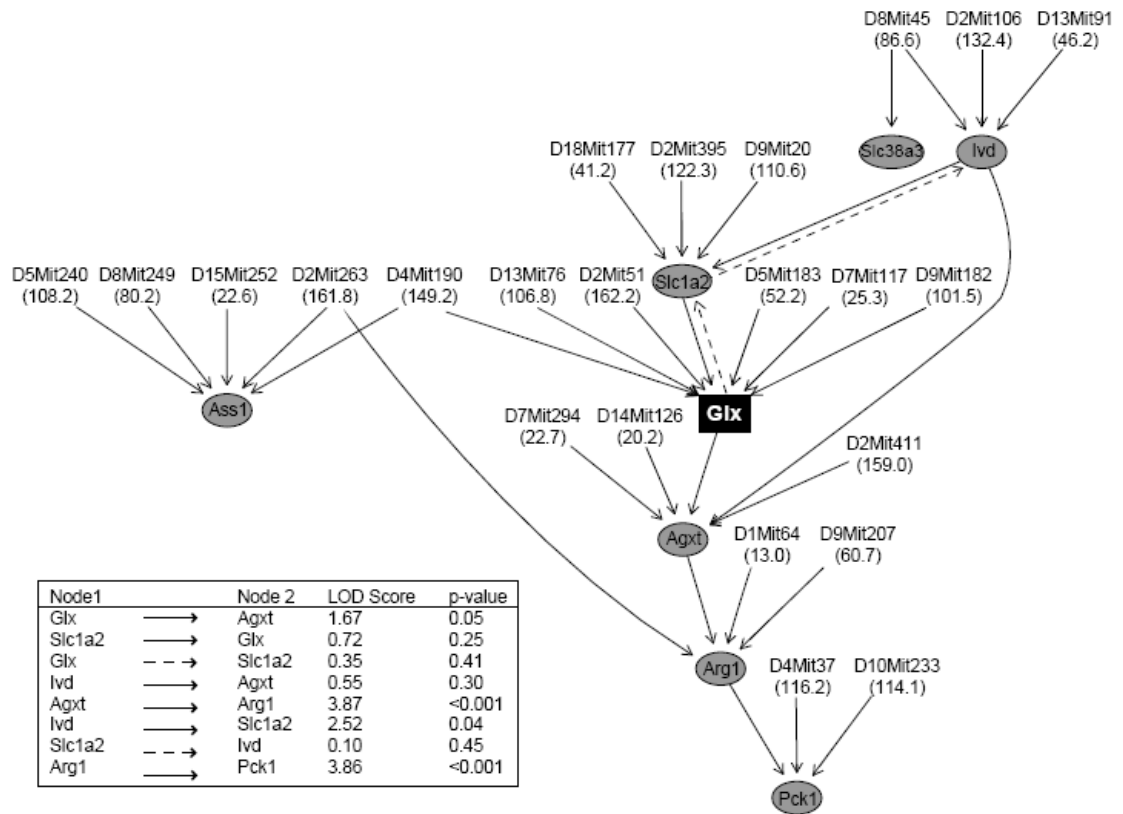
The Spearman's correlation coefficient of glx with the transcripts of the network is recorded. The physical location (Mb) of each transcript is noted (boxed and in bold), and eQTL or mQTL LOD>3.0 ±1.5 LOD interval are given (Mb).

Gene/Metabolite Symbol	GLX	Agxt	Arg1	Asl	Ass1	Ivd	Pck1	Pck1	Slc1a2	Slc38a3	
Spearman's Correlation to Glx	1.00	0.55	0.35	0.23	0.40	0.33	0.36	0.39	0.49	0.42	
mQTL/eQTL (LOD>3 ± 1.5 LOD) (Mb)	Chromosome	1	95.0	13.0-171.5	13.0-156.6						
		2	114.0-178.5	68.2-178.5	145.3-171.9		31.3;149.7-178.4	118.6;68.9-151.9	172.8	172.8	102.6;105.2-141.8
		3									
		4	3.5-150.1				22.6-150.1		40.8-121.5	3.6-144.6	
		5	39.3-117.7			130.3	47.7-129.2				
		6									
		7	13.9-35.1	13.9-28.7							
		8					48.5-119.5	46.0-170.7			36.6-100.0
		9	85.1-112.0		54.2-74.5					90.2-117.6	107.5
		10			24.6				107.2-120.4		
		11									
		12									
		13	81.4-112.0					33.1-86.4			
		14		19.2-23.1					21.1-98.3		
		15				3.3-91.8	3.3-76.0				
		16									
		17									
		18							15.5-85.7	30.7-48.3	
		19									

Table 6. Glx network correlations of amino acid biochemical group.

The Spearman's correlation coefficient of the 15 individual amino acids with the transcripts of the glx network is recorded ($|r|>0.254$, $p<0.05$; $|r|>0.330$, $p<0.01$). Amino acids are ordered as in Figure 24.

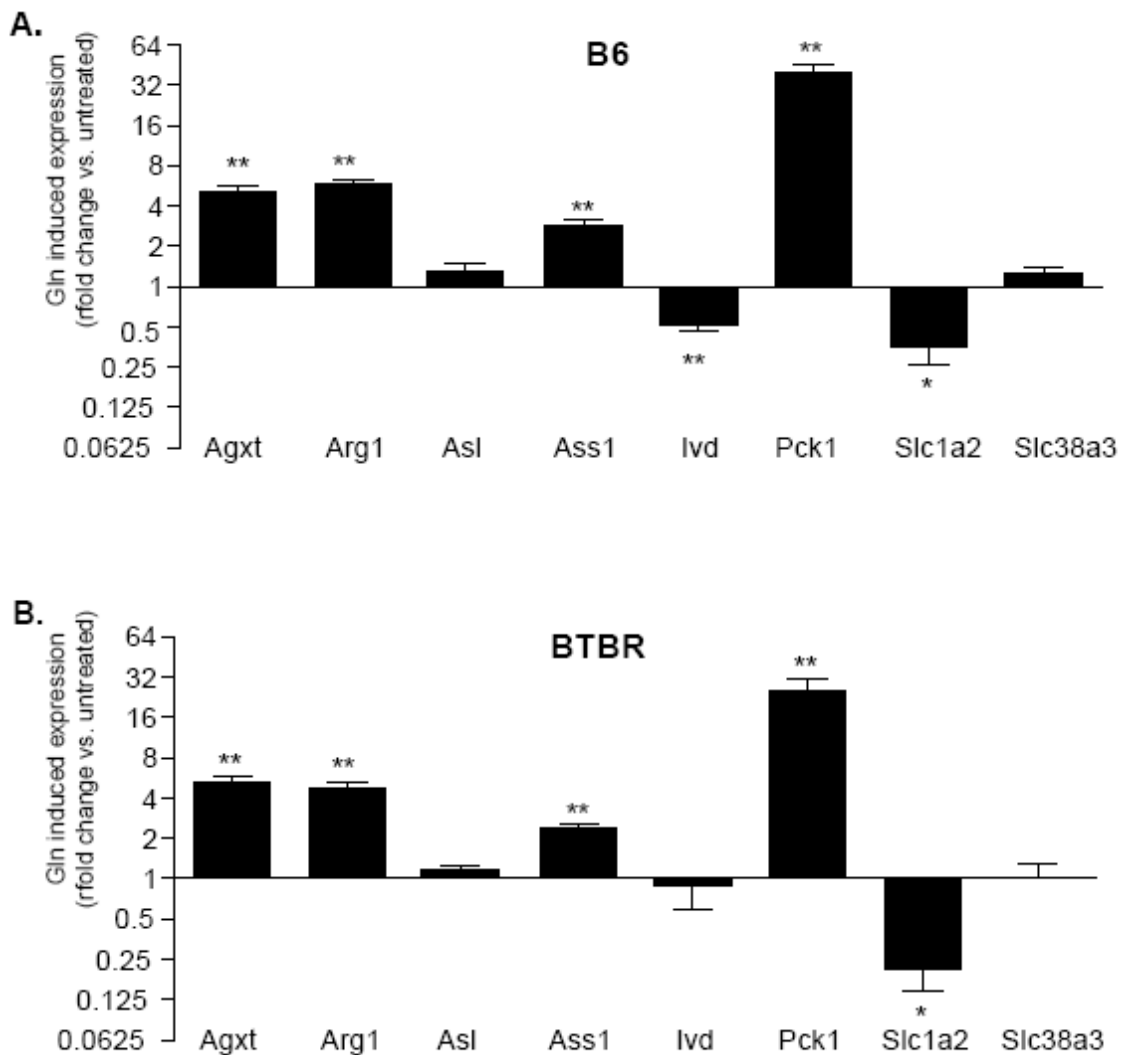
Spearman's Correlation Coefficient (Transcript with Metabolite)	TRANSCRIPT									
	Gene Symbol	Agxt	Arg1	Asl	Ass1	Ivd	Pck1	Pck1	Slc1a2	Slc38a3
	Gene Location chr (Mb)	1 (95.0)	10 (24.6)	5 (130.3)	2 (31.3)	2 (118.6)	2 (172.8)	2 (172.8)	2 (102.6)	9 (107.5)
AMINO ACID	Tyr	0.33	0.20	0.22	0.16	0.35	0.04	-0.02	0.44	-0.06
	Cit	0.33	0.25	-0.01	0.14	0.00	0.44	0.36	0.22	0.52
	Ala	0.19	0.09	0.00	0.12	-0.18	0.24	0.21	-0.16	0.33
	Arg	0.49	0.43	0.20	0.27	0.20	0.30	0.25	0.24	0.35
	Asx	0.47	0.37	0.14	0.22	0.09	0.31	0.20	0.15	0.37
	Glx	0.55	0.35	0.23	0.40	0.33	0.36	0.39	0.49	0.42
	His	0.51	0.42	0.16	0.31	0.11	0.46	0.40	0.17	0.46
	Orn	0.40	0.28	0.01	0.15	0.12	0.36	0.23	0.26	0.38
	Pro	0.37	0.22	-0.05	0.12	-0.10	0.45	0.32	0.04	0.46
	Gly	0.51	0.33	0.14	0.35	0.18	0.56	0.45	0.36	0.54
	Ser	0.54	0.40	0.17	0.36	0.18	0.54	0.41	0.33	0.53
	Phe	0.50	0.39	0.11	0.34	0.08	0.57	0.42	0.26	0.60
	Met	0.48	0.39	0.06	0.30	0.07	0.56	0.43	0.25	0.58
	Leu.Ile	0.52	0.42	0.20	0.40	0.16	0.60	0.41	0.32	0.55
	Val	0.53	0.44	0.23	0.42	0.19	0.64	0.46	0.32	0.55



Ferrara, et. al. (2008). PLoS Genet. 4(3)e1000034.

Figure 25. Glx network.

This network consists of a select number of transcripts (grey circles) among the 250 mRNA that are most correlated to glx (black rectangle) ($p < 0.002$). The microsatellite marker (Mb) for peak eQTL or mQTL altering levels of transcripts and metabolites, respectively, are given. For any two phenotypes connected by an edge, the direction LOD score and p-value are indicated (insert). The best solution was determined by an approximate Bayes factor (BF) and indicated in solid lines, the second best solution in dotted lines



Ferrara, et. al. (2008) PLoS Genetics. 4(3):e100034

Figure 26. Glutamine changes hepatic gene expression.

Hepatocytes from 10-week-old lean B6 (A) and BTBR (B) were treated overnight +/- 10mM glutamine (n=5 per strain). Transcripts were measured by RT-PCR and expression was normalized to *Actb* control. Significance calculated based on the difference of delta CT value of each transcript between the untreated and glutamine-treated hepatocytes for each individual animal (*p<0.05, **p<0.005)

4. Glutamine Regulation of Hepatic Transcription: Mechanism and Implications in Diabetes Phenotypes

4.1 Introduction

The liver organizes whole body energy homeostasis under fasting and fed conditions. Dietary and hormonal signaling pathways can modify hepatic metabolic processes either by acute changes in enzyme activity, or by long-term transcriptional regulation of enzyme abundance. In the context of obesity-induced insulin resistance, there are multiple exogenous and endogenous inputs affecting hepatic function [6,7]. Overnutrition is a main feature of obesity-associated diabetes, and it is therefore important not only to identify the nutrient signaling pathways involved in hepatic gene expression, but also to distinguish these pathways from hormonal influences.

Several studies have investigated the effects of obesity and insulin resistance on plasma and tissue metabolite levels. In the STEDMAN study, human subjects with obesity-related insulin resistance had anticipated increases in plasma medium and long-chain fatty acids, as well as elevation in plasma amino acids (notably the branched-chain amino acids, aromatic amino acids, and “glutamine family” of amino acids) [59]. A study analyzing amino acid metabolism in Zucker diabetic fatty rats in both the pre-diabetic and diabetic stages found that plasma and hepatic gluconeogenic amino acids were reduced, suggestive of increased utilization for glucose production [60]

Studies on nutrient regulation of liver metabolism have identified nuclear receptors and transcription factors that mediate fatty acid induced gene expression

changes [195-199]. Amino acid abundance also impacts liver function, affecting virtually every step from gene transcription initiation to protein translation. Unlike fatty acids, however, the key mediators of such amino acid induced effects are not well characterized[56,57,76,77,200]. The mammalian target of rapamycin (mTOR) pathway detects elevated amino acids and integrates this input with insulin signals, altering protein translation initiation and gene transcription through downstream effectors [201-204]. Amino acid deprivation is also known to regulate gene expression and translation in a *cis*-regulated fashion, with effector proteins binding to amino acid response elements (AARE) and nutrient sensing response elements (NSRE) within promoter regions [67,69,73,74,201,205-215].

Although recent work has identified several molecular mechanisms of amino acid-dependent transcription, the critical signaling intermediates are largely unknown. Studies summarized in Chapter 3 and in a published study outlined a causal network whereby glutamine acts to increase *Phosphoenolpyruvate carboxykinase (Pck1)* expression, which we tested in isolated hepatocytes from lean C57BL/6 (B6) and BTBR mice[104]. In order to further investigate the mechanism of glutamine regulation of *Pck1* expression, and to separate the effects of glutamine from the effects of hormones and obesity, we have performed additional studies in hepatocytes from the lean B6 mouse strain [26,31]. Our data indicate that α -ketoglutarate may be a key glutamine-derived mediator of *Pck1* induction. Phosphoenolpyruvate carboxykinase is implicated in diabetes development, and we find that the B6 and BTBR strain show a very different

hepatic expression profile. The leptin-deficient B6^{ob/ob} mouse has decreased levels of *Pck1* while the BTBR^{ob/ob} mouse has increased expression compared to their respective lean counterparts, despite the fact that both strains show an obesity-dependent decrease in glx levels. In addition, we find that liver α -ketoglutarate is nearly two fold higher in the BTBR strain than in the B6 mouse, and we postulate that this elevation of liver α -ketoglutarate is responsible for uncoupling glutamine abundance from *Pck1* expression.

4.2 Materials and Methods

Animals

BTBR, B6, and B6^{ob/+} mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and bred at the University of Wisconsin. The lineage and characteristics of the BTBR strain have been reviewed by Ranheim, *et al.* Mice were housed in an environmentally controlled facility (12-hour light and dark cycles) and were weaned at 3 weeks of age onto a 6% fat diet (Purina; #5008). Mice had ad libitum access to food and water, except for 4-hour fasting periods before blood draws and sacrifice (by CO₂ asphyxiation). The facilities and research protocols were approved by the University of Wisconsin Institutional Animal Care and Use Committee.

Plasma measurements

Glucose was measured by the glucose oxidase method using commercially available kits (Sigma-Aldrich, St. Louis MO). For lean mice, insulin was measured by radioimmunoassay. For *ob/ob* mice, insulin was measured by an in-house developed ELISA using a pair of insulin/proinsulin antibodies (clones D6C4 and D3E7-BT; Research Diagnostics, Concord, MA) according to procedures previously described[216].

RNA Collection and Microarray

Only male mice at the 4 and 10-week time point were considered and five mice of each strain and condition were used (five B6 lean 4-week-old, five B6 ^{*ob/ob*} 4-week-old, five B6 lean 10-week-old, five B6 ^{*ob/ob*} 10-week-old; similar for BTBR). RNA preparations and all hybridizations were performed at Rosetta Inpharmatics (Merck & Co., Seattle, WA). The procedure for gene expression profiling, data normalization, and GO term enrichment analysis is described in Keller *et al.*, 2008 [216]. Custom ink-jet microarrays used in this study (Agilent Technologies, Palo Alto, CA) consisted of 4,732 control probes and 39,558 non-control oligonucleotides extracted from mouse Unigene clusters and combined with RefSeq sequences and RIKEN full-length cDNA clones. Mouse tissues were homogenized and total RNA extracted using Trizol reagent (Invitrogen, CA) according to manufacturer's protocol. Total RNA was reverse transcribed and labeled with either Cy3 or Cy5 fluoro-chrome. For a given strain, labeled complementary RNA (cRNA) from each animal of that strain was hybridized against a

pool of labeled cRNAs constructed from equal aliquots of RNA from all the animals for that strain (over both time points). All hybridizations were performed in fluor-reversal for 48 hours in a hybridization chamber, washed, and scanned using a confocal laser scanner. Arrays were quantified on the basis of spot intensity relative to background, adjusted for experimental variation between arrays using average intensity over multiple channels, and fitted to a previously described error model to determine significance [217]. Gene expression measurements are reported as intensity values or re-ratio values. The re-ratio values are derived from a ratio splitter pairwise ratio builder function in Resolver 6.0 to account for strain-specific reference pools (Rosetta Biosoftware, Seattle, WA). The ratio splitting operation generates intensity profiles, including error modeling of the channels of the ratio scan group normalization, forward transformation of intensities, group de-trending, and inverse transformations. Experiments are then rebuilt by rationing each sample to a new baseline value which is a super-pool average of all array hybridizations in the experiment.

Liver Metabolite Quantification

Amino acids, acyl-carnitines and organic acids were measured using stable isotope dilution techniques [132,164,218] as described in Chapter 3.

In Vitro Hepatocyte Experiments

Hepatocytes from lean 10-week-old male B6 parental strain mice were isolated by liver perfusion [174]. Hepatocytes were seeded at subconfluency (3.5×10^6 cells/6 well plate) in low glucose DMEM (Gibco) supplemented with FBS (10% vol/vol; Gibco), pen/strep antibiotic (1%, Gibco), glutamine (2mM; Gibco), and pyruvate (1mM; Gibco). Cells were left to attach for 3 hours in an incubator at 37°C, 5% CO₂. After washing with PBS, the cells were treated with unsupplemented DMEM (Sigma) with 1g/L glucose, pen/strep (1%), +/- appropriate concentration of glutamine, alanine, dimethyl-glutamate, dimethyl-2-oxoglutarate, or 2-Aminobicyclo[2.2.1]heptane-2-carboxylic acid (BCH) (Sigma). Cells were treated for 24 hours.

RNA was extracted from hepatocytes using RNeasy kits (Qiagen) after treatment described above. RNA isolation and gene expression measurements by real time quantitative PCR was conducted as described in Chapter 3. Primer sequences and gene accession codes for transcripts measured are listed in Table 3. For amino acid measurements, hepatocytes were plated as above and lysed in 500µL H₂O. Organic acid measurements were made on hepatocytes plated at subconfluency on a 100mm dish (4-8

x 10⁶ cells/dish), lysed with 750µL 0.1M HCl. Sample metabolites were profiled as previously described [132,164,218] in Chapter 2.

Statistical Analysis

To analyze *in vivo* data, we classified liver genes and hepatic metabolites into differential expression or abundance patterns, respectively, according to methods described in Keller, *et al*[216]. Briefly, we used an empirical Bayes hierarchical modeling approach called EBarrays [136,168,219], which is implemented in R, a publicly available statistical analysis environment and available through Bioconductor (www.bioconductor.org). EBarrays describes the probability distribution of a set of expression measurements, accounting for differences among values in their true expression/abundance levels, measurement fluctuations, and distinct expression/abundance patterns for a given gene/metabolite among different conditions.

Given that there are four conditions within each strain (4-week-old vs. 10-week-old; lean or obese), 15 expression patterns are possible. These patterns include changes seen in obesity only, age only, and strain only, as well as combinations of these major patterns. Each gene or metabolite was assigned to the expression pattern with the maximum posterior probability (MPP); significance was determined at MPP <0.7, suggesting a 30% chance that it could be one of 14 other patterns ($\alpha \approx 0.02$).

The *in vitro* data was analyzed with Prism software version 4.02 (Graph Pad Software). The number of biological replicates for each experiment is indicated in the

figure legend. Analysis of amino acid and organic acid measurements in response to glutamine used a paired t-test between the untreated and glutamine-treated hepatocytes. For gene expression data, the fold change of mRNA relative to the untreated hepatocytes for each animal is depicted graphically. Expression of each gene was normalized to *Actb*. Statistical analysis was performed on the difference in delta CT values of metabolite-induced gene expression versus gene expression in the untreated (or glutamine-treated) hepatocytes, as indicated in figure legends. The transcriptional effects of glutamine, α -ketoglutarate, glutamate and alanine were analyzed by paired t-test to the untreated and glutamine-treated hepatocytes. The glutamine effect on *Gls2*, *Glud1*, *Got1*, *Got2*, *Gpt1*, *Gpt2*, *Oat*, and *Pck1* was analyzed by one-way ANOVA, with a Dunnet's post test comparing values to the untreated hepatocytes. The dose response curves for glutamine and α -ketoglutarate utilized a sigmoidal (variable slope) equation; significance was determined by a one-way ANOVA with a Dunnet's post test relative to untreated hepatocytes.

4.3 Results

Glutamine treatment elevates a unique subset of hepatic metabolites

Our previous work has defined a causal network by which glx (glutamine+glutamate) alters gene expression in an F2 population generated from the B6^{ob/ob} and BTBR^{ob/ob} parental strains. The component of this causal network for which

we have the most confidence is glx altering *Phosphoenolpyruvate carboxykinase (Pck1)* expression through transcript nodes *Alanine:glyoxylate aminotransferase (Agxt)* and *Arginase (Arg1)* (inferred direction significance, $p < 0.001$). Importantly, the function of this network was confirmed in our published *in vitro* hepatocyte study described in Chapter 3[104,147].

Our F2 metabolite correlation analysis shows that liver glx groups with many other amino acids, but is also at the interface between the amino acids and α -ketoglutarate. This correlation data makes biological sense given that glutamate is both a substrate and product of transamination reactions in amino acid biosynthesis and catabolism, respectively. In order to uncover glutamine-derived metabolites that could mediate the glutamine effect, we profiled metabolites in isolated hepatocytes exposed to 10mM glutamine.

In response to glutamine, hepatocyte alanine, asx (asparagine+aspartate) and ornithine levels were significantly elevated (Figure 27a). Alanine is of great interest because it is an important gluconeogenic precursor which plays a role in the glucose sparing effect. Alanine is also a substrate and product of *Agxt*, the transcript to which glx is most highly correlated. As a group, alanine, asx, and ornithine are significant because their biosynthesis from glutamate each involves a specific transaminase enzyme which catalyzes the production of α -ketoglutarate.

Organic acid profiling of glutamine-treated hepatocytes is consistent with glutamine production of α -ketoglutarate. There is a slight elevation in α -ketoglutarate and

a significant elevation in TCA cycle intermediates (succinate, fumarate, malate) downstream of the α -ketoglutarate entry site (Figure 27b). Citrate is significantly decreased whereas lactate is increased, which suggests shunting of pyruvate away from pathways of glucose oxidation.

Glutamine metabolism and glutamine-derived metabolites play a role in altered *Pck1* expression

In attempts to identify the mechanism of glutamine action, we first tested whether glutamine alters transcription of specific transaminase enzymes required to produce the aforementioned amino acids. We find no change in *Glutaminase (Gls2)*, *Glutamic oxaloacetate aminotransferase (Got1, Got2)*, *Glutamic pyruvate aminotransferase (Gpt1)*, *Ornithine aminotransferase (Oat)*, or *Glutamate dehydrogenase (Glud1)* mRNA in response to 10mM glutamine, indicating that glutamine treatment does not act at the transcript level to stimulate important genes of exogenous and endogenous glutamine metabolism (Figure 28a). We next tested whether enhancing glutamine metabolism by stimulating enzyme activity would alter the glutamine effect. BCH is an analogue of leucine that activates glutamine metabolism via allosteric stimulation of glutamine dehydrogenase activity and may increase glutaminase activity [220,221]. Preliminary data reveals that 10mM BCH not only potentiated the glutamine effect, but that BCH alone acting on endogenous glutamine stores could elevate *Pck1* expression (Figure 28b).

Our data thus far suggests that glutamine-derived metabolites may contribute to the induction of *Pck1* expression, and we therefore investigated candidate metabolites, α -

ketoglutarate, glutamate, and alanine. We find that 10mM α -ketoglutarate not only recapitulates the glutamine effect to stimulate *Pck1* transcription, but like glutamine, also causes upregulation of *Arg1* and *Agxt* mRNA levels. Furthermore, 10mM dimethyl-glutamate elicits a similar pattern of gene induction as that observed with 10mM glutamine. In contrast, alanine treatment did not significantly induced *Pck1* and had no effect on *Arg1* and *Agxt* expression (Figure 29).

In the B6 strain, we illustrate that both the glutamine and α -ketoglutarate induction of *Pck1* is dose-dependent (Figure 30). At physiological plasma concentrations (~1-2mM), the glutamine response is one half of the response observed at 10mM [78,222,223]. The α -ketoglutarate dose response curve is shifted relative to glutamine. Glutamine has a significant inductive effect above 3mM, with the EC₅₀ of 3.1mM; α -ketoglutarate induces *Pck1* significantly at 7mM, and has an EC₅₀ of 6.1mM.

Glutamine regulation of Pck1 in the context of diabetic states

Increased gluconeogenesis is the primary cause for increased hepatic glucose output and fasting hyperglycemia in type 2 diabetes[24,224,225], making *Pck1* a candidate driver for diabetes phenotypes. In order to define the causal role of *Pck1* in development of diabetes, we considered a time course of liver *Pck1* expression and related hepatic metabolites in both the BB6 and BTBR strains. We analyzed the obese “pre-diabetic” state at 4 weeks, prior to elevated blood glucose, and the obese “diabetic” state at 10 weeks, where both strains are insulin resistant but only the BTBR display

hyperglycemia. For comparison, we studied the lean counterparts of both strains at the two time points.

Our microarray data shows that hepatic *Pck1* expression has a drastically different expression profile across the eight subsets of animals (Figure 31a). In the B6 animal, at both 4 and 10-weeks of age, there is an obesity-dependent *decrease* in hepatic *Pck1* expression, yet in the BTBR animal, there is an obesity-dependent *increase* in liver *Pck1* mRNA. Importantly, at 10 weeks, the BTBR^{ob/ob} diabetic animal has elevated levels of *Pck1* mRNA compared to the B6^{ob/ob} but this difference is not seen in the lean counterpart.

We have shown previously *in vitro* that there is no strain difference in induction of *Pck1* expression in response to a supraphysiological 10mM dose of glutamine[104], but the data in Figure 30 shows that glutamine and α -ketoglutarate induction of *Pck1* are dose dependent. To relate this to the differences in *Pck1* profiles in our time course study, we measured endogenous hepatic metabolite levels in the two strains at different ages by MS-based methods. Our *in vivo* data shows that in both strains there is an obesity-dependent decrease in hepatic glx at both the 4-week and 10-week time points (Figure 31b). While both strains decrease liver glx levels in response to obesity, hepatic α -ketoglutarate is elevated nearly two-fold in the BTBR strain relative to the B6 at all time points and in both lean and obese physiological states (Figure 31c). This data reveals that at all times the BTBR strain preserves at least one metabolite capable of inducing *Pck1* at

(lean:glx+ α KG; *ob/ob*: α KG), whereas the B6 strain is without either inducer in the obese state.

We had shown previously *in vivo* that α -ketoglutarate was able to recapitulate the glutamine-induction of *Agxt* and *Arg1* mRNA, the two transcript nodes between glx and *Pck1* in our proposed network[104]. If elevated hepatic α -ketoglutarate *in vivo* is sufficient to uncouple glx regulation of *Pck1* expression in the livers of the BTBR strain, we predict that *Agxt* and *Arg1* should show different *in vivo* profiles between the strains. We in fact do find that liver *Agxt* expression profile in the B6 animal parallels the obesity dependent decrease in hepatic glx; however, *Agxt* mRNA is unchanged in the BTBR^{*ob/ob*} animals (Figure 31d). Liver *Arg1* shows an obesity dependent increase in the BTBR, very similar to the *Pck1* mRNA profile, but *Arg1* mRNA is not elevated in the B6 obese state (Figure 31e). While we cannot rule out that other glutamine-derived metabolites are involved in *Pck1* mRNA induction, this data does suggest that α -ketoglutarate may be able to disrupt the specific network pathway from glutamine (through *Agxt* and *Arg1*) to *Pck1*.

4.4 Discussion

In patients with type 2 diabetes, the degree of fasting hyperglycemia is highly dependent on the rate of hepatic glucose production. Gluconeogenesis has been shown to be exaggerated in patients with non-insulin-dependent diabetes, resulting from an

increase in availability of endogenous glucose precursors, diminished secretion of insulin, or hepatic insulin resistance [24,225]. The relationship between insulin secretion and hepatic glucose production is reciprocal such that the degree of fasting hyperglycemia is closely related to the degree of impaired islet responsiveness to secretagogues[226]. Hepatic gluconeogenesis clearly is an integral part of diabetes pathology, such that it has become a target for therapeutic intervention in diabetes treatment[227].

In those with type 2 diabetes, there is an increased production of gluconeogenic precursors which fuels the increased gluconeogenesis. In both human and rodent models of obesity and insulin resistance, the “pre-diabetic” state is characterized by altered levels of tissue and plasma fatty acids, amino acids, and organic acids[59,60]. In diabetic patients, however, there is an additional intrahepatic mechanism to increase conversion of these precursors into glucose; this is not present in normal individuals where substrate-induced stimulation of gluconeogenesis fails to increase total glucose production[225,228]. In this study, we investigate the role of metabolites as transcriptional regulators contributing to this intrahepatic mechanism and compare metabolite sensing in two mouse strains which differ in diabetes susceptibility.

Phosphoenolpyruvate carboxykinase catalyzes one of the irreversible steps of gluconeogenesis, and is regulated primarily at the transcriptional level. Our previous work shows that *Pck1* expression is increased in isolated hepatocytes in response to glutamine treatment[104], thus making *Pck1* a candidate gene involved in the metabolite driven hepatic glucose production. We explored the mechanism of glutamine-induced

Pck1 expression in the B6 lean animals to isolate metabolite driven effects from hormonal or obesity- induced changes.

We measured metabolites in isolated hepatocytes from the B6 strain in response to 10mM glutamine to characterize glutamine-derived intermediates that are instrumental in induction of *Pck1*. We find that glutamine elevates alanine, asx, and ornithine, all non-essential amino acids that use glutamate as an ammonia donor in transamination reactions. Glutamine also elevated TCA cycle intermediates distal to α -ketoglutarate entry. Citrate, the first intermediate of the TCA cycle after acetyl-CoA entry, was decreased, suggesting that pyruvate is being shunted away from oxidation pathways. Lactate is also increased, further supporting our theory that gluconeogenesis is proceeding while glucose oxidation pathways are shut down.

While glutamine did not induce expression of the specific transaminase enzymes required to produce our specific elevated amino acids, preliminary data suggests that accelerating the metabolism of glutamine via BCH enhances the glutamine effect. We therefore considered glutamine-derived metabolites that could mediate the induction of *Pck1* and show that at 10mM, α -ketoglutarate, glutamine, and glutamate are equally effective at inducing *Pck1*. Moreover, α -ketoglutarate recapitulates the glutamine-induction of *Alanine:glyoxylate aminotransferase* and *Arginase* expression, the two network transcript nodes between glx and *Pck1*.

When treated with supraphysiologic concentrations of glutamine (10mM), approaching saturating concentrations for transport[78], the B6 and BTBR hepatocytes

have a similar induction of *Pck1*[104], but to give physiological credence to these findings, we tested the dose-dependency of the glutamine and α -ketoglutarate effect. In fed rats, plasma glutamine is approximately 0.6mM, while the estimated concentrations of this metabolite in the cytosol and mitochondria are 7 and 20mM respectively[53,229]. The α -ketoglutarate concentrations are estimated to be much lower, with levels of approximately 10 μ M and 1-1.5mM in the plasma and cytosol, respectively[230,231]. Our data shows a dose dependency of both glutamine and α -ketoglutarate with the α -ketoglutarate curve slightly right shifted. The EC₅₀ for glutamine and α -ketoglutarate are 3.1mM and 6.6mM, respectively. We also find that at concentrations of either metabolite resembling physiological levels in plasma, neither metabolite alters *Pck1* expression. While this data does indicate that *in vitro* glutamine is more a potent activator of *Pck1* transcription than is α -ketoglutarate, no hormones were added to hepatocyte cultures, and glucagon or insulin may alter the dose-dependence of the metabolites *in vivo*.

We next investigated changes in these metabolites relative to changes in gene expression in livers from the B6 diabetes-resistant and BTBR diabetes-susceptible strains. The leptin-deficient B6 and BTBR mice are both insulin resistant, but only the 10-week-old BTBR^{ob/ob} mouse develops diabetes. Our results show that in the B6^{ob/ob} strain, despite being insulin resistant, has suppressed liver *Pck1* mRNA compared to the B6 lean counterpart at both 4 and 10 weeks of age. The B6 hepatic *Pck1* expression profile parallels glx metabolite abundance; at 4 and 10 weeks, the B6^{ob/ob} animals have both decreased glx and *Pck1* mRNA. The BTBR strain shows a similar obesity-

dependent decrease in liver glx abundance, yet the BTBR^{ob/ob} mice have elevated *Pck1* mRNA at both time points. Interestingly, liver α -ketoglutarate in the BTBR strain is nearly twice that of the B6 strain. Based on these data, we hypothesize that in the presence of elevated α -ketoglutarate, liver *Pck1* becomes uncoupled from glutamine regulation in the BTBR strain, contributing to elevated *Pck1* expression and increased hepatic gluconeogenesis. In support of this hypothesis we find that *Agxt* also shows an obesity-dependent decrease in the B6 strain that parallels glx abundance, yet is uncoupled from glx regulation in the BTBR strain. Additionally, the lack of an obesity-dependent increase in B6 hepatic *Arg1* expression could be due to the absence of inducing concentrations of liver α -ketoglutarate in this strain.

While further studies are needed to fully characterize glutamine induction of *Pck1* mRNA, this study illustrates one molecular pathway that is potentially dysregulated in the diabetes-susceptible strain. Our results suggest that elevated liver α -ketoglutarate could be a means by which regulation of hepatic gluconeogenesis is disrupted in the BTBR strain. The implications of hepatic α -ketoglutarate as a potential contributor to development of diabetes is underscored by previous work in our lab showing that the BTBR is more responsive to the secretagogue α -ketoisocaproate (KIC), which acts via conversion to α -ketoglutarate and leucine[232]. Future studies on the BTBR strain will determine whether a similar phenomenon occurs in the liver with hepatic glutamine being more efficiently converted to α -ketoglutarate. If so, this may provide a unifying

mechanism for development of dysfunction in two key organs of metabolic fuel homeostasis.

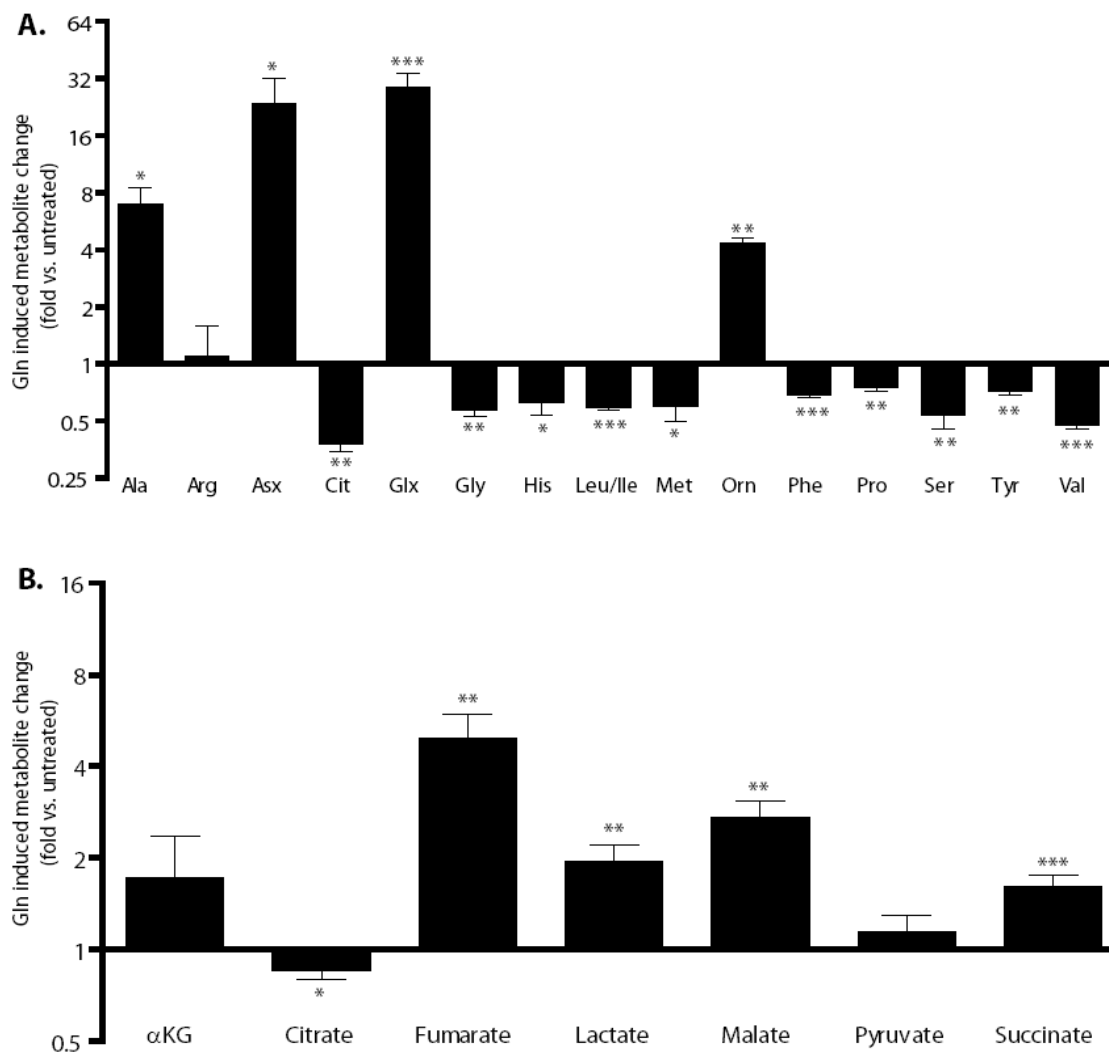


Figure 27. Glutamine alters abundance of liver metabolites.

Amino acids (A) (n=5) and organic acids (B) (n=8) were profiled in isolated hepatocytes from 10-week-old lean male B6 mice treated +/- 10mM glutamine. For each metabolite, statistical significance was determined by a paired t-test between the glutamine-treated and untreated hepatocytes (*p<0.05, **p<0.01, ***p<0.001). Metabolite codes are provided in Table 1 and Table 2.

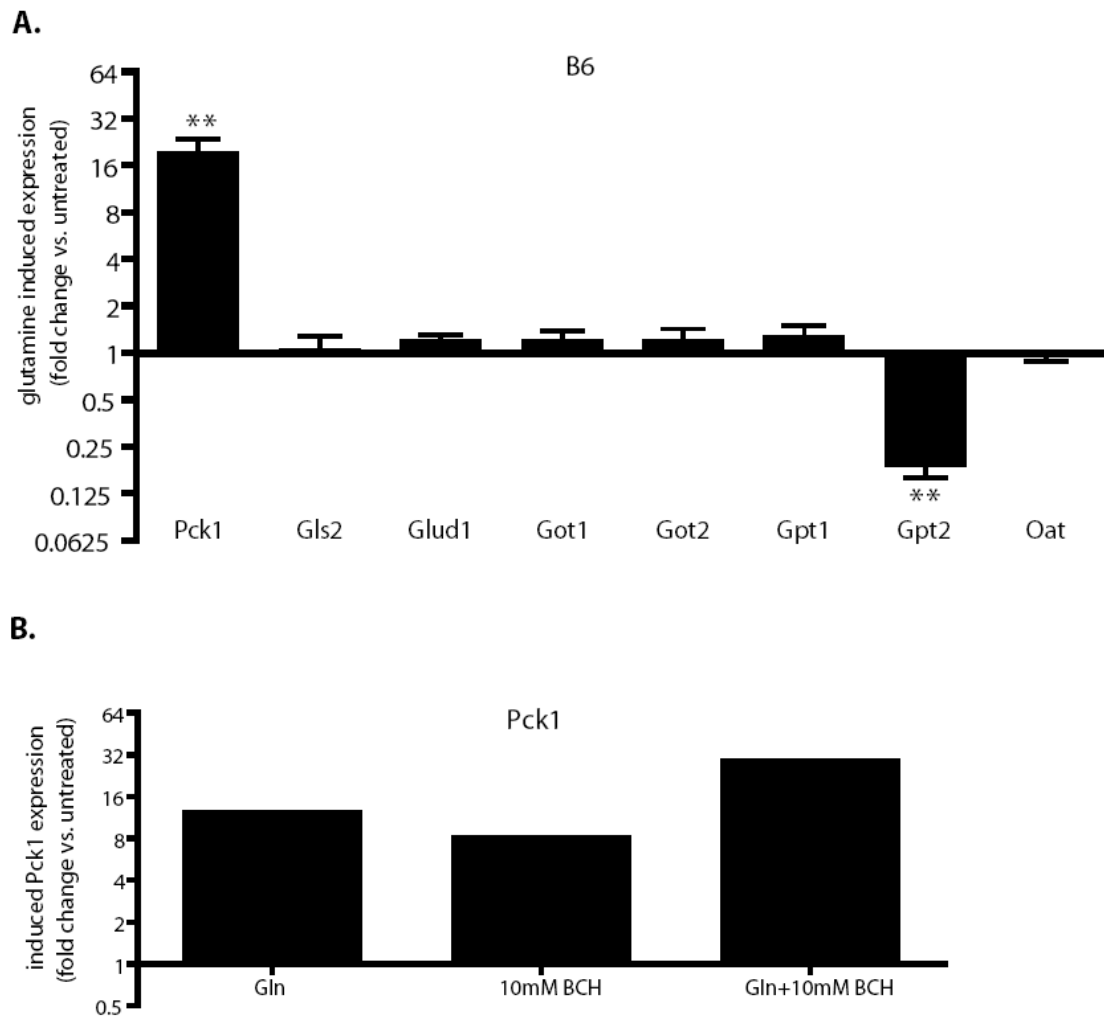


Figure 28. Relationship between glutamine metabolism and Pck1 expression.

Hepatocytes from 10-week-old male B6 mice were isolated and treated with 10mM glutamine (n=6) (A), or 10mM glutamine +/- 10mM BCH (B). Transcripts were measured by RT-PCR and expression was normalized to *Actb*. Statistics are based on the difference of delta CT of each transcript between the untreated and treated hepatocytes (*p<0.05, **p<0.01). Preliminary data in (B) was not analyzed.

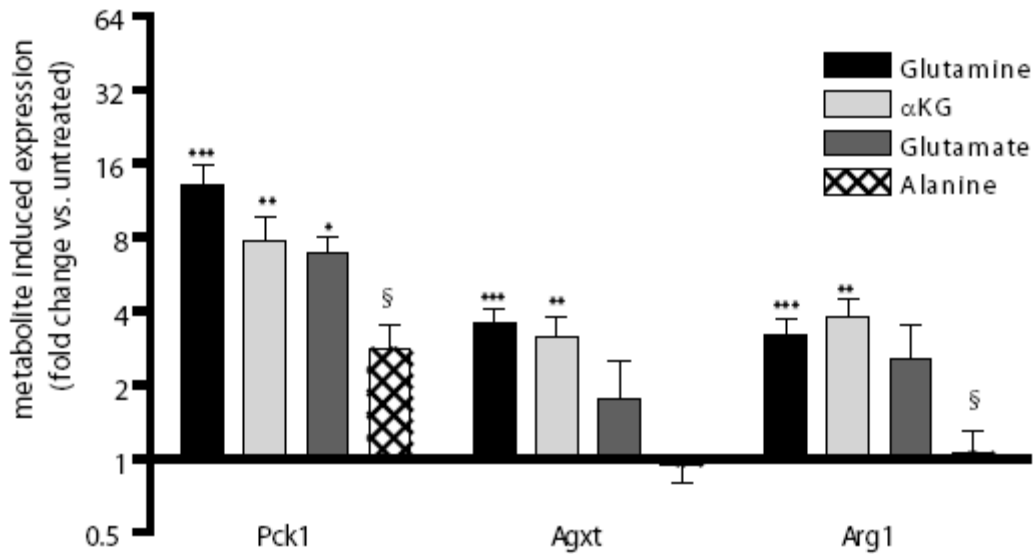


Figure 29. Glutamine related metabolites alter hepatic gene expression.

Hepatocytes from 10-week-old male B6 mice were isolated and treated with 10mM of the indicated metabolite (glutamine, black n=9; α -ketoglutarate, light grey n=7; glutamate, dark grey n=2; alanine, hatched n=2). Transcripts were measured by RT-PCR and expression was normalized to *Actb*. Statistics are based on the difference of delta CT of each transcript between the metabolite-treated compared to untreated hepatocytes (*p<0.05, **p<0.01, ***p<0.001) or compared to the corresponding gene treated with 10mM glutamine (§p<0.05).

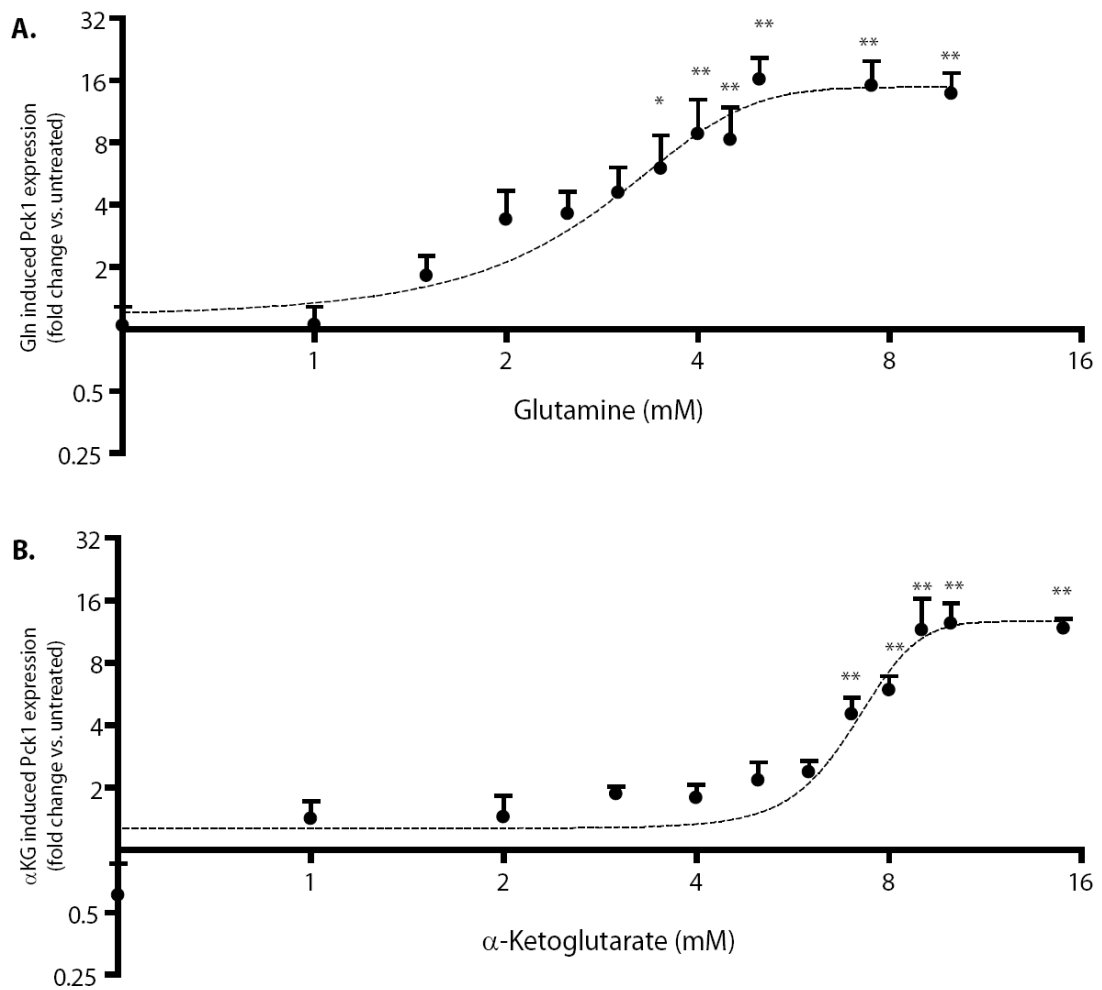


Figure 30. Induction of Pck1 expression by glutamine and α -ketoglutarate is dose dependent.

Hepatocytes from 10-week-old male B6 mice were isolated and treated indicated dose of metabolite (n=5-6). Transcripts were measured by RT-PCR and *Pck1* expression was normalized to *Actb*. Statistics are based on the difference of delta CT between the untreated and metabolite-treated hepatocytes (*p<0.05, **p<0.01).

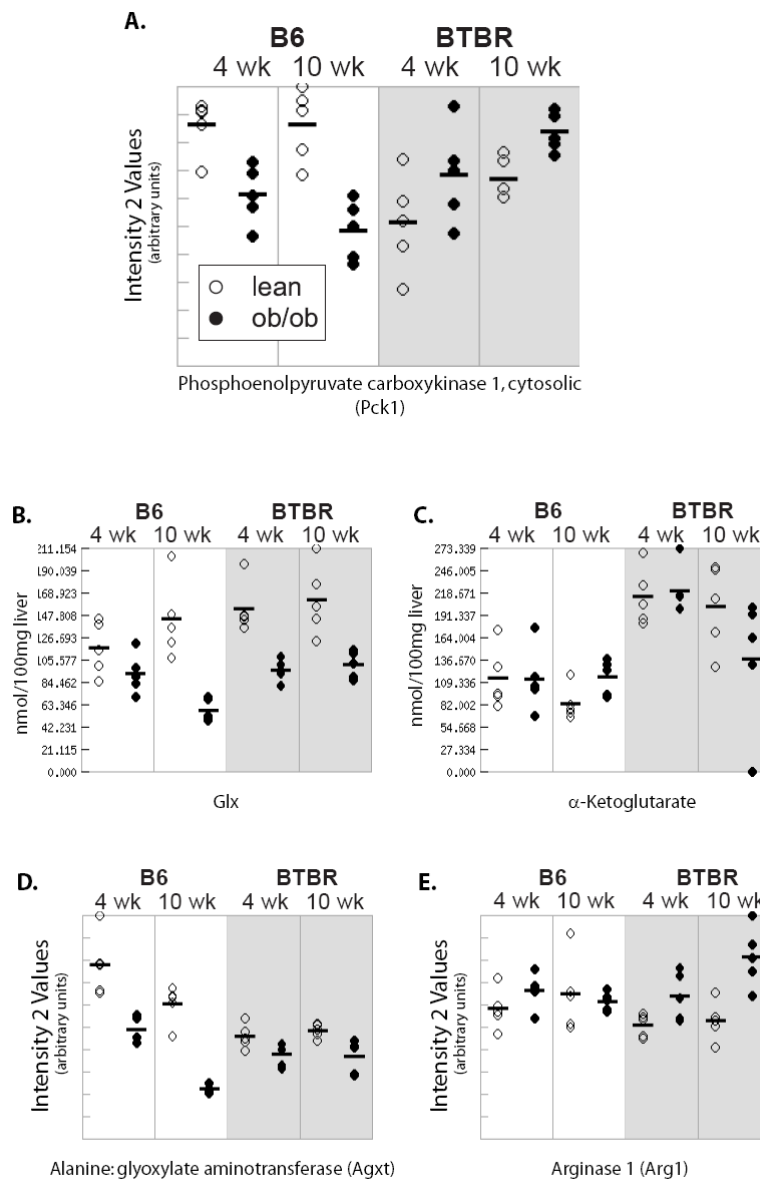


Figure 31. Liver Pck1 mRNA, glx and α -ketoglutarate show strain-specific differences.

Liver *Pck1* expression (A), glx (glutamine+glutamate) (B) and α -ketoglutarate (C) profiles (nmol/100mg liver tissue) as well as intermediate glx network nodes *Agxt* (D) and *Arg1* (E) are shown for the five animals in each of eight subsets. Open (lean) and closed (*ob/ob*) circles represent individual mice. Horizontal bars show mean values for each group.

5. The Role of Metabolites in Inter-tissue Communication

5.1 Introduction

Type 2 diabetes is characterized by fasting hyperglycemia associated with defects in the pancreatic islet, the liver, and the peripheral tissues. Identification of “causal” elements that serve to trigger progression to such broad tissue failure remains a major challenge in developing effective treatment plans[233]. There is increasing evidence that the liver and adipose act as “driving” organs while defects in the muscle and islet occur in response to adipokines and other factors that are secreted by the former tissues[6,7,234]. Just as adipose-derived signaling molecules affect downstream tissues, liver-derived metabolites may contribute to the β -cell decompensation that marks diabetes.

Hepatic insulin resistance often results from obesity-related metabolic overload. Excess circulating free fatty acids from the diet or delivered from insulin-resistant adipocytes impair liver insulin signaling and alter hepatic lipid metabolism, resulting in decreased lipid oxidation and increased triglyceride synthesis[235]. Hepatic lipid oxidation occurs primarily in the mitochondria, but in times of increased lipid mobilization and mitochondrial overload, the oxidation takes place in alternate organelles. In the endoplasmic reticulum the cytochrome P450 enzymes catalyze the first step to convert monocarboxylic fatty acids to dicarboxylates in a process referred to as ω -oxidation; these dicarboxylic acids undergo chain shortening via β -oxidation in the peroxisome. In states of hepatic insulin resistance, generation of malonyl-CoA diverts

fatty acids away from CPT-1 and mitochondrial fatty acid oxidation towards esterification pathways, further enabling these alternate routes for oxidation [113,115,121,123]. Accumulation of lipid-derived metabolites not only alters hepatic metabolism, but can lead to release of these substances into the plasma where they can affect other peripheral tissues.

The potential for fatty acids to stimulate insulin secretion has been extensively studied, revealing fatty acids acutely enhance secretion, but that chronic exposure alters islet function and contributes to β -cell decompensation [6,122,125,236,237]. Moreover, the “glucolipotoxicity” concept proposes that free fatty acids and hyperglycemia are synergistic in their detrimental effects on islet function[6]. Recent work has also shown that fatty acyl derivatives, including dicarboxylic acids, have the potential to stimulate insulin secretion through both metabolism and signaling pathways[238].

Our mouse model of type 2 diabetes shows several strain differences in fatty acid oxidation between the diabetes-resistant C57BL/6 *leptin*^{ob/ob} (B6^{ob/ob}) and diabetes-susceptible BTBR *leptin*^{ob/ob} (BTBR^{ob/ob}). The BTBR^{ob/ob} mice have elevated hepatic transcripts acting in mitochondrial β -oxidation, but display decreased expression of enzymes necessary for complete usage of these fatty acids. Our data also reveals an obesity-dependent increase in expression of transcripts of the ω -oxidation and peroxisomal β -oxidation pathways in both strains, but with more pronounced upregulation in the BTBR strain. Moreover, the BTBR^{ob/ob} strain also exhibits upregulation of ω -oxidation transcripts that redirect dicarboxylic acids from pathways of

complete oxidation within the peroxisome, and therefore makes these derivatives available to other tissues. Metabolic profiling reveals that strain-specific differences in mRNA levels are accompanied by strain-specific differences in select liver monocarboxylic and dicarboxylic acyl-carnitine derivatives; the diabetes-resistant B6 strain shows an obesity-dependent decrease in the concentrations of these analytes, while in the BTBR strain these metabolites are either unchanged or increased in response to obesity.

We hypothesize that release of these metabolites and delivery to the islet could modulate β -cell function, and in this study we explore the role of fatty acid-derived dicarboxylate metabolites as insulin secretagogues. We reveal that adipic acid, a six-carbon dicarboxylic acid, enhances glucose-stimulated insulin secretion. Fuel secretagogues can stimulate insulin secretion via TCA anaplerosis and extramitochondrial signaling[6,108,125,132,133,239-241], and there is increasing evidence that fatty acids serve as mediators of the K_{ATP} independent pathway[238]. To investigate the signaling potential of fatty acid derivatives, we explore the insulin-secreting mechanism of non-metabolizable dicarboxylic acid, and suggest that the site of action may include processes distal to fuel metabolism.

5.2 Materials and Methods

Animals

BTBR, B6, and B6^{ob/+} mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and bred at the University of Wisconsin. The lineage and characteristics of the BTBR strain have been reviewed by Ranheim, *et al.* Mice were housed in an environmentally controlled facility (12-hour light and dark cycles) and were weaned at 3 weeks of age onto a 6% fat diet (Purina; #5008). Mice had ad libitum access to food and water, except for 4-hour fasting periods before blood draws and sacrifice (by CO₂ asphyxiation). The facilities and research protocols were approved by the University of Wisconsin Institutional Animal Care and Use Committee.

RNA Collection and Microarray and Metabolite Profiling

Only male mice at the 4 and 10-week time point were considered and five mice of each strain and condition were used (five B6 lean 4-week-old, five B6^{ob/ob} 4-week-old, five B6 lean 10-week-old, five B6^{ob/ob} 10-week-old; similar for BTBR). RNA collection, gene expression analysis, and liver metabolite profiling was conducted as in Chapter 4.

Islet Cell Lines

All reagents for islet related experiments were purchased from Sigma (St. Louis, MO), unless otherwise indicated. 832/3 cells were derived from parental INS-1 cells by a transfection-selection strategy and exhibit a robust insulin secretion response[242]. Cells

were cultured in RPMI 1640 medium containing 11.1mM glucose and supplemented with 10% FBS, 10mM HEPES, 2mM glutamine, 1mM sodium pyruvate, and 50 μ M β -mercaptoethanol at 37° C in a humidified atmosphere containing 5% CO₂.

Insulin Secretion Assay

Insulin secretion from cell lines was measured by preincubating cells in glucose-free secretion buffer (114mM NaCl, 4.7mM KCl, 1.2mM KH₂PO₄, 1.16mM MgSO₄, 20mM HEPES, 2.5mM CaCl₂, 0.2% bovine serum albumin, pH 7.2) for two hours. Cells were then incubated for in secretion buffer containing 2.5mM or 15mM glucose, +/- appropriate amount of indicated test compound. After two hours of cell incubation, insulin accumulation in the assay buffer was measured by radioimmunoassay (Coat-A-Count kit, DPC, Los Angeles, CA; RI-13K, Linco Research, St. Charles, MO) or ELISA (Linco Research, St. Charles, MO), as described previously[216].

In vitro islet experiments

Islets were harvested from adult Spague-Dawley rats weighing approximately 250g as previously described [243,244]. Islets were initially incubated in RPMI 1640 medium containing 8mM glucose and supplemented with 10% FBS, 10mM HEPES, 2mM glutamine, 1mM sodium pyruvate, 50 μ M β -mercaptoethanol, 20U/mL penicillin, 20ug/mL streptomycin, 0.05ug/mL amphotericin B (Gibco, Carlsbad, CA) for 24 hours prior to insulin secretion assay. Islets were washed and incubated for one hour in

secretion buffer containing 114mM NaCl, 4.7mM KCl, 1.2mM KH₂PO₄, 1.16mM MgSO₄, 20mM HEPES, 2.5mM CaCl₂, 25.5mM NaHCO₃, 0.2% bovine serum albumin, pH 7.2, supplemented with 2.5mM glucose. Islets were then divided into groups and incubated in 500μL secretion buffer containing 2.5mM or 16.7mM glucose for two hours. Insulin secretion from islets was measured by RIA (Coat-a-Count).

Insulin Secretion Measurements

All insulin measurements were normalized to protein concentration. Islets protein was measured via the BCA protein assay (Pierce, Rockford, IL), while 832/3 protein content was measured by either the BCA protein assay or Bradford assay (Sigma). Cell samples were homogenized on ice in lysis buffer containing 0.1% Triton X-100. The BCA assay consisted of adding 200μL bicinchoninic acid and copper sulfate (50:1) to protein lysates, incubating at 37° C for 30 minutes, and calculating protein concentration by measuring absorbance at 340nm. The Bradford assay consisted of adding 200μL of Bradford reagent to protein lysates, incubating at room temperature for 15 minutes, and calculating protein concentration was by measuring absorbance at 540nm.

Calcium Imaging

Calcium influx was estimated using calcium-sensitive fluorphore fura 2 as previously described[245]. INS-1 cells were plated onto 25-mm glass coverslips (Fisher, Pittsburgh, PA) in RPMI 1640 media four days prior to imaging. Confluent slides were

washed and incubated in solution containing 2.8mM glucose, 125mM NaCl, 5.9mM KCl, 1.28mM CaCl₂, 1.2mM MgCl₂, 17mM HEPES, pH 7.4 containing 2.5mM glucose for one hour. Cells were incubated with fura 2-AM for 30 min at 37°C in a loading solution that contained the measurement solution plus 5 μM fura 2-AM and 5μM pluronic acid. After 20 minutes, the loading solution was replaced with measurement solution containing 2.8mM glucose, 85mM NaCl, 50mM KCl, 2mM CaCl₂, 1mM MgCl₂, 15mM HEPES, pH 7.4. Calcium flux was measured via an inverted microscope in conjunction with a 175-W Xenon arc lamp Metafluor Imaging System (Nikon Instrument, Lewisville, TX) as described [245]. Treatments were 25mM glucose+0.4%DMSO vehicle or 25mM glucose+0.5mM DEPDC. Reactions were quenched by 30mM KCl.

Statistical Analysis

The *in vivo* gene and metabolite analysis was conducted as described in Chapter 4 and in Keller, *et al* [216]. For a metabolite to be “causal” it must meet several requirements. At the 4-week time point (the pre-diabetic stage), it must show a change with obesity in the B6 that is different than the change seen in the BTBR. We assume here that a change with obesity in the B6 is an “obesity only” metabolite; in the BTBR strain however, this metabolite could be a “diabetes risk” metabolite or an “obesity only” metabolite. By eliminating those metabolites that change in the same direction in both strains, we eliminate the “obesity only” metabolites. We repeat the same procedure for the 10-week time point, at which stage the BTBR^{ob/ob} has diabetes. Here we assume again

that a metabolite changing in the B6 strain with obesity is an “obesity only” metabolite, while that changing in the BTBR strain with obesity is a “diabetes responsive” metabolite, a “diabetes risk” metabolite, or an “obesity only” metabolite. Again, by eliminating those metabolites that change in the same direction at 10 weeks, we eliminate the “obesity only” metabolites. Finally, we compare the 4-week metabolites (causal “diabetes risk”), with the 10-week metabolites (causal “diabetes risk” + reactive “diabetes responsive”), and those common are determined to be the causal metabolites.

The *in vitro* insulin secretion data was analyzed with Prism software version 4.02 (Graph Pad Software). For dose response curves to DEPDC in INS-1 cells and islets, the basal insulin secretion induced by DEPDC and stimulated insulin secretion induced by DEPDC and adipic acid were analyzed separately to determine significant changes as a result of the dicarboxylates. These data were analyzed by ANOVA and Dunnett’s post test compared to untreated conditions at corresponding glucose concentrations. To analyze the effects of known secretagogues (α -ketoglutarate, KIC, dimethyl malate, and KCl), we conducted ANOVA and Dunnett’s post test compared to untreated conditions at corresponding glucose concentration. To determine the effect of DEPDC, an ANOVA with Bonferroni correction comparing 1mM DEPDC to no DEPDC at otherwise identical secretagogue and glucose concentrations.

5.3 Results

BTBR mice have specific defects in liver mitochondrial lipid oxidation pathways

In previous chapters, the B6 and BTBR mouse strains were shown to differ in their susceptibility to diabetes development. The leptin-deficient B6^{ob/ob} and BTBR^{ob/ob} mice are both insulin resistant at 10 weeks, but only the 10-week-old BTBR^{ob/ob} mice develop hyperglycemia indicative of β -cell decompensation and type 2 diabetes [95,216].

Our microarray data shows an obesity-dependent increase in liver transcripts encoding enzymes of mitochondrial fatty acid oxidation in both strains (e.g., medium and long-chain acyl-CoA dehydrogenase; enoyl-CoA hydratase). This suggests that in both the B6 and BTBR mice, obesity causes an increase in fatty acid oxidation via conventional mitochondrial β -oxidation at the level of liver mRNA. The BTBR^{ob/ob} mouse, however, has specific defects in the ability to completely metabolize certain acyl-carnitine species. Compared to the B6 strain, the BTBR strain has lower levels of *Methylcrotonoyl-CoA carboxylase, 1 (Mccc1)* and *Isovaleryl-CoA dehydrogenase (Ivd)* mRNA levels in the liver. At both 4 and 10 weeks of age, *Holocarboxylase synthetase, alpha (Hlcs)* is increased in the B6 strain in the obese state, but remains unchanged in the BTBR mouse (Figure 32a).

The aforementioned enzymes are specifically required for the breakdown of odd-chain acyl-carnitines which are produced during branched-chain amino acid catabolism. We profiled liver metabolites and find strain-specific differences in short and medium-

chain liver acyl-carnitines which are consistent with the defects in mitochondrial β -oxidation seen at the transcript level. Relative to the B6 strain, the BTBR has increased levels of the five-carbon acyl-carnitine species, which reflects the relative decrease in *Mccc1* and *Ivd* that act in the catabolic pathways of C5:1 (tiglyl carnitine) and the C5's (isovaleryl, 3-methylbutyryl, or 2-methylbutyryl carnitine), respectively.

Holocarboxylase synthetase is involved in metabolism of biotin, an essential cofactor for carboxylase reactions. The B6 strain shows an obesity-dependent decrease in propionyl carnitine (C3) at both 4 and 10 weeks, suggestive of decreased activity of biotin-requiring propionyl-CoA carboxylase. The BTBR strain, which does not show an obesity-dependent increase in *Hlcs* transcription also does not exhibit the decrease in propionyl carnitine with obesity as seen in the B6 livers; rather the BTBR has an increase in this metabolite with obesity consistent with reduced catabolism (Figure 32b).

The BTBR strain has increased hepatic microsomal and peroxisomal lipid oxidation

Fatty acids are primarily oxidized in the mitochondria under normal conditions. However, during times of metabolic overload and increased fatty acid mobilization, or defective mitochondrial oxidation, lipids can be converted to dicarboxylic acids by microsomal ω -oxidation and are subsequently broken down in the peroxisome [113]. We find a strain difference in hepatic expression of *Peroxisome proliferator receptor, alpha* (*Ppara*), a transcription factor responsible for regulating both ω -oxidation and

peroxisomal β -oxidation. *Ppara* shows an obesity-dependent decrease in the B6 mouse, but an obesity-dependent increase in the BTBR strain. Several *Ppara* target genes show increased hepatic expression only in the BTBR strain. Cytochrome P450 members *Cyp4a10* and *Cyp4a14*, important enzymes of ω -oxidation, are only increased in the BTBR strain in response to obesity. *Acyl-CoA thioesterase, isoform 8 (Acot8)*, which alters metabolism of dicarboxylates and promotes their release from the peroxisome prior to complete oxidation[113], also shows a more pronounced obesity-dependent expression elevation in the BTBR strain (Figure 33a)

Several medium-chain dicarboxylates show strain-specific differences consistent with the gene expression data for pathways of ω -oxidation and peroxisomal β -oxidation. Glutaryl carnitine (C5-DC), adipoyl carnitine (C6-DC), hexenedioyl carnitine (C6:1-DC) and octenedioyl carnitine (C8:1-DC) show an obesity-dependent decrease in the B6 and are either increased or unchanged in response to obesity in the BTBR (Figure 33b). Importantly, glutaryl carnitine, adipoyl carnitine, and hexenedioyl carnitine are considered “causal” metabolites, that is at both the 4 and 10-week time point, these liver metabolites showed a different response to obesity in the B6 versus BTBR strain (e.g., in response to obesity C6:1-DC at 4 weeks is decreased in the B6 but unchanged in the BTBR; at 10 weeks, is decreased in the B6 but increased in the BTBR; see methods).

Dicarboxylate fatty acids and dicarboxylate derivatives alter insulin secretion

Studies have shown that liver-derived acyl-carnitine species and dicarboxylic acids have the potential to be released into the blood stream [59,113], and we therefore

considered that specific members of these metabolites could modulate islet function. We were particularly interested in the medium-chain dicarboxylic acids which our statistical analysis identified as causal metabolites. Given that even-numbered medium-chain dicarboxylic acids are the primary urinary dicarboxylate species of patients with mitochondrial β -oxidation defects[113], we tested adipic (C6) acid, in our 832/3 INS-1 cell line. Preliminary data shows that adipic acid potentiates glucose-stimulated insulin secretion in INS-1 cells, with a trend towards increased basal insulin secretion (Figure 34).

Liver-derived dicarboxylic acids can enter the mitochondria via a CPT-independent pathway, and if delivered to the islet these metabolites could therefore potentially enter the mitochondria and contribute to anaplerosis [111]. Insulin secretagogues are thought to act via conversion to TCA cycle intermediates, which alter the ATP:ADP ratio through the oxidative pathway. However, there is emerging evidence that extramitochondrial TCA intermediates, which are generated by pyruvate cycling pathways, also have a key signaling role [6,107,108,129,132,239,240] To characterize the signaling role of dicarboxylic fatty acids derivatives we investigated the secretagogue potential of a non-metabolizable dicarboxylic acid, diethyl-3,4-pyridinedicarboxylate (DEPDC) (Figure 35a).

DEPDC is unable to enhance basal insulin secretion, but potentiates glucose-stimulated insulin secretion in a dose-dependent manner in both INS-1 cells and primary rat islets. We show that 1mM DEPDC enhances insulin secretion by nearly five-fold

compared at 15mM glucose (Figure 35b,c). We subsequently tested whether DEPDC was dependent on stimulatory glucose for its effect, or if it could potentiate the effects of other metabolic fuels. Our data reveals that DEPDC can potentiate insulin secretion initiated by α -ketoglutarate, and α -ketoisocaproate (KIC) at basal glucose (Figure 36a,b), but not that of dimethyl-malate, which has no secretory capacity at basal glucose (Figure 36c). DEPDC potentiates insulin secretion initiated by depolarizing KCl, even in the absence of glucose (Figure 37a,b), as well as insulin secretion induced by calcium ionophore ionomycin (not shown). We find also that DEPDC requires extracellular calcium entry, as blocking L-type calcium channels with verapamil, abolished DEPDC action even in the presence of high KCl (Figure 37c).

Dicarboxylic acids may act as signaling molecules in insulin secretion

Our data thus far indicates that DEPDC may act downstream of calcium entry to amplify insulin secretion initiated by other secretagogues. Extracellular calcium entry is required for DEPDC secretory action, and we therefore considered calcium-dependent processes on which this compound could act. There are several intracellular calcium stores in the β -cell including the acidic vesicles and the ER, and these stores have been shown to alter insulin secretion. Release of calcium from these stores is a calcium-dependent process requiring second messengers cyclic ADP-ribose (cADPR) [246-249] and nicotinic acid adenine dinucleotide phosphate (NAADP) [250-254]. The

structural similarity of DEPDC and these second messengers led us to consider that DEPDC enhances calcium mobilization in the islet.

Preliminary data shows that DEPDC enhances calcium mobilization in INS-1 cells compared to a stimulatory concentration of glucose by altering the kinetics, but not the magnitude of the response. Thus, we observe a much more immediate influx of calcium when DEPDC is applied, but at later time points, cells treated with glucose alone attain the same amplitude of calcium flux as those treated with glucose and DEPDC (Figure 38).

5.4 Discussion

The causal factors driving progression from metabolic dysregulation to frank diabetes are numerous. No one tissue is responsible for diabetes pathogenesis; rather it is communication between organs that contributes to disease progression. Obesity and diet-induced changes in plasma nutrients cause hepatic insulin resistance and transcriptional changes which disturb key metabolic processes. Since the liver is responsible for regulating whole body energy homeostasis, such alterations can have drastic effects on downstream tissues[226]. The clinical hallmark of type 2 diabetes is elevated fasting and postprandial plasma glucose, indicating both increased hepatic glucose output and inadequate insulin secretion from the β -cell. To effectively treat diabetes, a better

understanding of pathways that facilitate communication between such organs will be required.

While we have evidence that metabolic pathways are altered in primary tissues at the transcriptional level, messenger RNA can only mediate its effects within the tissue where it is expressed. It is the products of such altered biochemical pathways that have the potential to link changes in one tissue with another. A well-characterized example of such mediators are adipokines, which signal changes in the adipose energy status to other organs that control whole body metabolism[7]. Tissue metabolites also have the potential to serve as such interorgan mediators, and represent the biological output of environmental, genetic, nutrient, and hormonal influences on a tissue. We therefore focused our attention on changes in specific liver metabolites occurring in response to obesity and development of diabetes in mice, and tested the idea that these metabolites could alter islet function.

We performed microarray analysis of mRNA expression profiles and comprehensive MS-based metabolic analysis in livers from two mouse strains that differ in their susceptibility to disease. In our time-course analysis summarized in this chapter, we were able to assess transcripts and metabolites from 4-week “pre-diabetic” and 10-week “diabetic” lean and obese mice from both strains. Our data reveals that the BTBR strain has decreased hepatic expression of enzymes required to catabolize certain acyl-carnitines (*Mccc1*, *Ivd*, *Hlcs*). These defects result in elevated concentrations of several

odd-numbered, short-chain acyl-carnitines (C5s, C5:1, and C3) specifically in the BTBR^{ob/ob} compared to the B6^{ob/ob} at 10 weeks of age.

Our data also indicates that at the transcriptional level, the BTBR strain experiences hepatic fatty acid overload that exceeds the mitochondrial capacity for oxidation, possibly causing diversion of excess lipids towards ω -oxidation and subsequent peroxisomal β -oxidation pathways. The key transcriptional mediator of these pathways, *Ppara*, is downregulated with obesity in the B6 strain at 4 and 10 weeks; yet in the BTBR strain, this transcript is either unchanged or increases in the obese state. Expression of several Ppar target genes are upregulated in the BTBR strain, including *Cyp4a10*, *Cyp4a14*, and *Acot8*. These cytochrome P450 isoforms catalyze the first step of ω -oxidation, and the increase seen in the BTBR diabetes-susceptible strain is consistent with other studies finding an induction of the CYP4A family in diabetic states[255]. *Acot8* redirects dicarboxylic acids away from oxidation to succinate, and thus medium-chain dicarboxylates are released from the peroxisome prior to complete catabolism. The strain-specific differences in the medium-chain dicarboxylic acids glutaryl carnitine, adipoyl carnitine, hexenedioyl carnitine, and octenedioyl carnitine correlate with anticipated changes in metabolic byproducts. Importantly, our analysis reveals that the five-carbon and six-carbon dicarboxylate acyl-carnitines may be considered “causal” for diabetes development.

To characterize the liver/islet communication, we show that adipic acid (a six-carbon dicarboxylic acid) potentiates glucose-stimulated insulin secretion in a dose

dependent manner in 832/3 INS-1 cells, which is consistent with the potentiating effect of non-esterified fatty acids on insulin secretion[256]. Our data also indicates that dicarboxylic acids have a trend to elevate basal insulin secretion suggests that they can be metabolized by the islet and participate in anaplerosis. This has been seen for other nutrient secretagogues such as α -ketoisocaproate, which alters basal insulin secretion via conversion to α -ketoglutarate [232]. It is known that in the hepatic peroxisome medium-chain dicarboxylates may be completely metabolized to succinate [113], and it is plausible that dicarboxylic acids are metabolized in the islet to contribute the pool of TCA cycle intermediates, pyruvate cycling, and secretory signals.

Non-glucose fuel secretagogues are thought to alter insulin secretion in part by anaplerosis, but there is evidence that these intermediates may also serve as signaling molecules to promote insulin secretion independent of their metabolism[257]. Export of TCA cycle intermediates and derivatives from the mitochondria provides the cytosol with dicarboxylic acids (citrate, malate, isocitrate, glutamate) [129,239,257,258]. It has been shown that as signaling molecules, these metabolites affect the amplifying pathway of insulin secretion independent of ATP by acting via unidentified extramitochondrial pathways that may alter exocytosis machinery[239,258]. To characterize the signaling role of liver derived medium-chain dicarboxylates, we studied the secretory potential of a non-metabolizable dicarboxylic acid, diethyl-3,4-pyridinedicarboxylate.

DEPDC enhances glucose-stimulated insulin secretion in 832/3 INS-1 insulinoma cells by five-fold compared to stimulatory glucose alone; this strong potentiating effect is

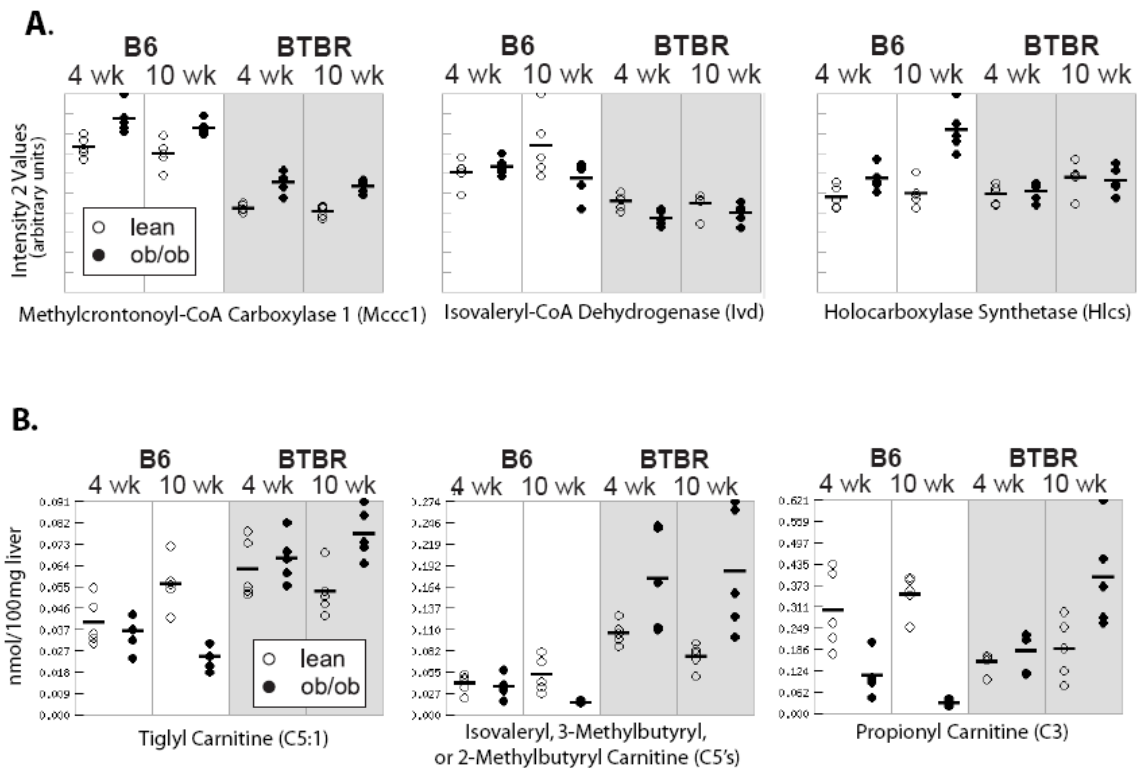
also seen in primary rat islets. We find that while DEPDC is unable to stimulate basal insulin secretion, it can potentiate insulin secretion from non-glucose secretagogues including α -ketoglutarate and α -ketoisocaproate, and also non-fuel secretagogues KCl (a depolarizing agent) and ionomycin (cell membrane calcium ionophore). This data indicates that the site of action of DEPDC may be downstream of the triggering pathway of insulin secretion and is dependent on extracellular calcium entry.

DEPDC is structurally similar to nicotinic acid, a moiety found on second messengers NAADP and cADPR which act in calcium-dependent signaling pathways. NAADP and cADPR are involved in the calcium-dependent calcium release from intracellular stores in the lysosome and ER, respectively [259,260], and we investigated whether DEPDC could alter calcium flux in INS-1 cells. We show in two independent experiments that, while DEPDC does not increase the amplitude of calcium flux in response to glucose, it does hasten the intracellular calcium increase. Several studies have shown that the intracellular calcium stores play a role in β -cell insulin secretion. Mechanisms by which the increased cytoplasmic calcium contributes to secretion include contributing to membrane depolarization and thus enhancing extracellular calcium entry, or by acting directly on the exocytosis machinery[261].

While the direct target of DEPDC remains to be characterized, our results show that dicarboxylic acids can serve as signaling molecules and alter islet insulin secretion. We hypothesize that upon release from the liver, these dicarboxylates serve to stimulate islet insulin secretion to handle acute nutrient excess. However, similar to the effects of

chronic elevated non-esterified fatty acids, it is possible that these dicarboxylates cause overstimulation of islets and eventual development of dysfunction, causing β -cell decompensation and resulting hyperglycemia.

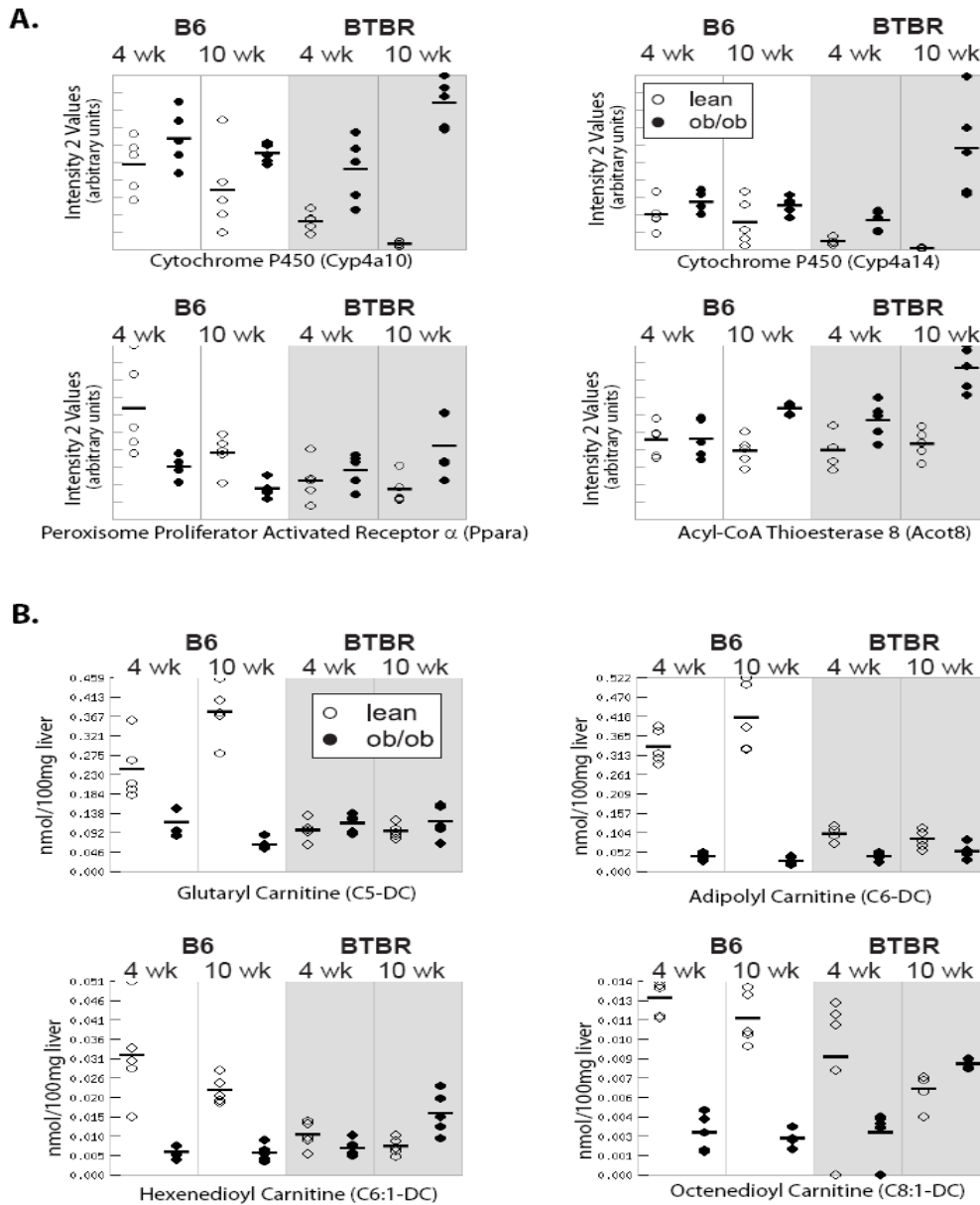
The glucolipotoxicity model of β -cell dysfunction states that islet functional impairment results from free fatty acid excess in the presence of stimulatory levels of glucose[262,263]; this concept is important in our mouse model system where only the 10-week-old BTBR^{ob/ob} mouse is hyperglycemic. Our *in vivo* data indicates that in the obese state, the B6 animal has a decrease in liver medium-chain dicarboxylate production. In contrast, the BTBR strain lacks this downregulation, and in some cases has elevated hepatic dicarboxylate levels in response to obesity. In a model consistent with the glucolipotoxicity concept, we propose that the presence of medium-chain dicarboxylates in the hyperglycemic 10-week-old BTBR^{ob/ob} mouse contributes to islet dysfunction and diabetes progression, whereas the B6 mouse decreases production of hepatic dicarboxylates to preserve an appropriate glucose-stimulated insulin secretion responses.



www.diabetes.wisc.edu

Figure 32. Liver monocarboxylic fatty acid oxidation differs by strain.

A. Intensity 2 values (arbitrary units) of specific liver transcripts of mitochondrial β -oxidation as measured by microarray. B. Liver monocarboxylic acyl-carnitine derivatives (nmol/100mg liver tissue). Open (lean) and closed (*ob/ob*) circles represent each of the five individual male mice in each subset. Horizontal bars show mean values for each group.



www.diabetes.wisc.edu

Figure 33. Strain differences in hepatic microsomal and peroxisomal oxidation

A. Intensity 2 values (arbitrary units) of liver transcripts of ω -oxidation and peroxisomal β -oxidation, measured by microarray. B. Liver dicarboxylic acyl-carnitine derivatives (nmol/100mg liver tissue). Open (lean) and closed (*ob/ob*) circles represent each of the five individual male mice. Horizontal bars show mean values for each group.

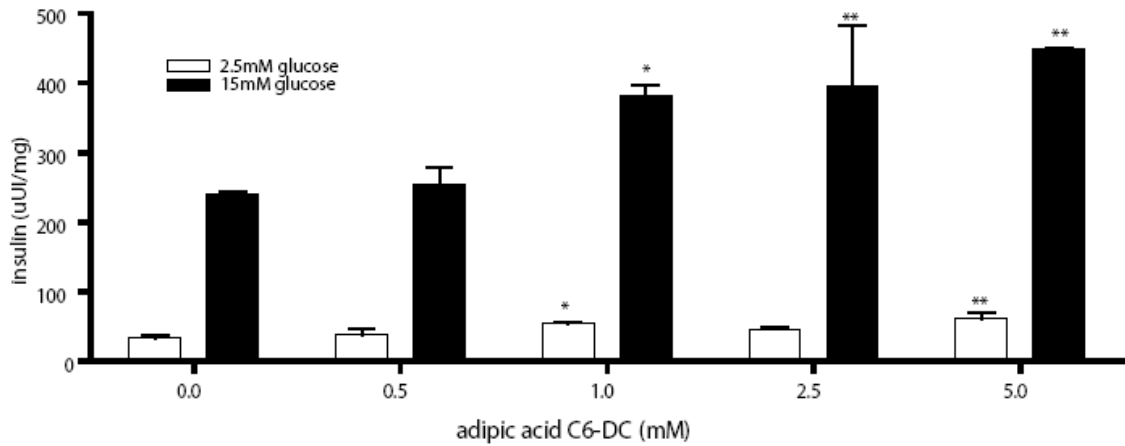


Figure 34. Adipic acid alters insulin secretion in INS-1 cell line.

The 832/3 INS-1 insulinoma cell line was treated with indicated amounts of adipic acid at basal glucose (2.5mM, white bars), or stimulatory glucose (15mM, black bars). Analysis compared the effect of adipic acid relative to no treatment at the corresponding glucose concentration (* $p < 0.05$, ** $p < 0.01$) (n=3).

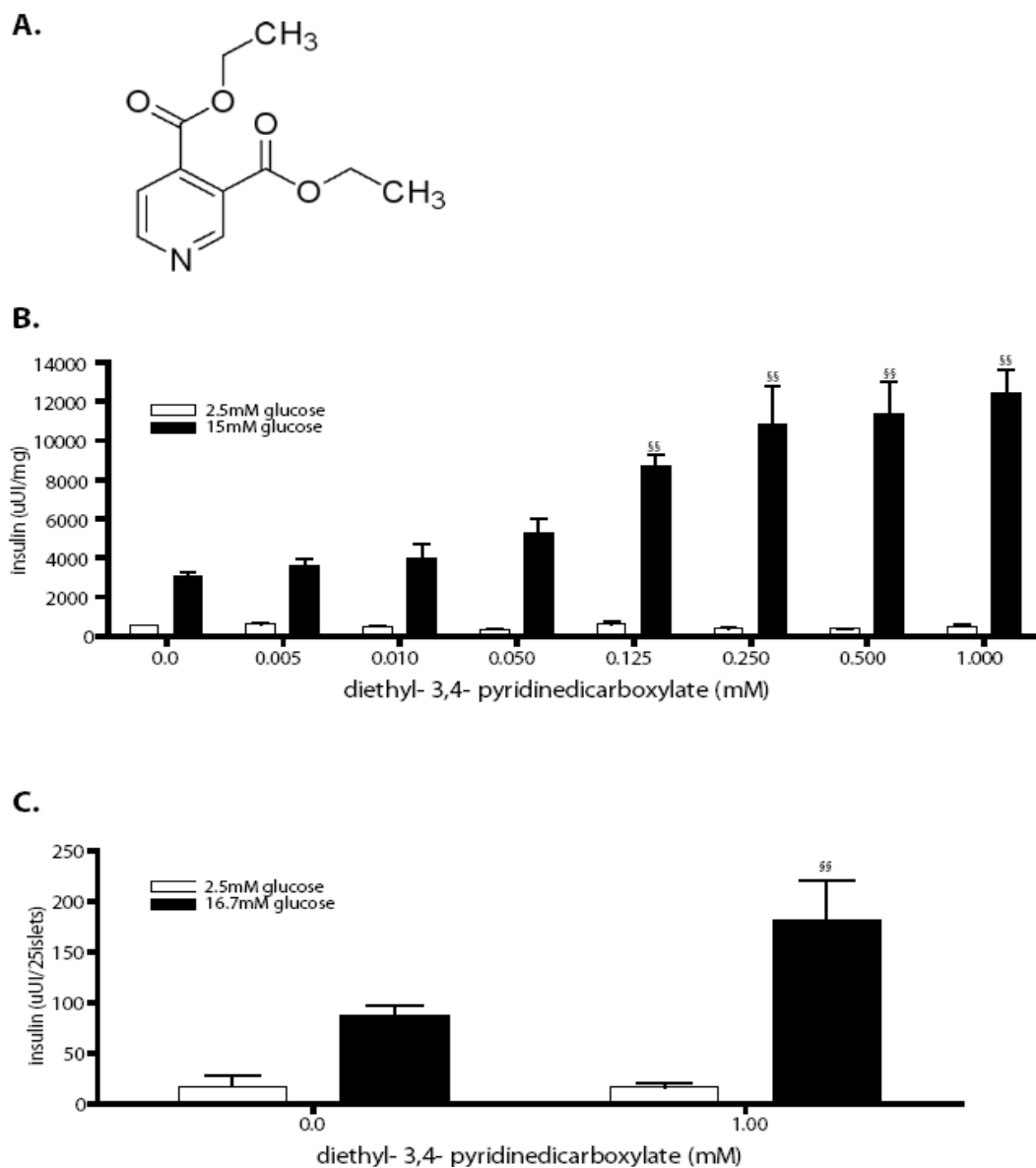


Figure 35. Diethyl-3,4-pyridinedicarboxylate (DEPDC) enhances glucose stimulated insulin secretion.

DEPDC (A) was added at indicated concentration to either 832/3 INS-1 insulinoma cell lines (B) (n=3) or primary rat islets (C) (n=3) at basal (white bars) or stimulatory (black bars) glucose concentrations. Statistical analysis compared the effect of adipic acid relative to no treatment at the corresponding glucose concentration (*p<0.05, **p<0.01).

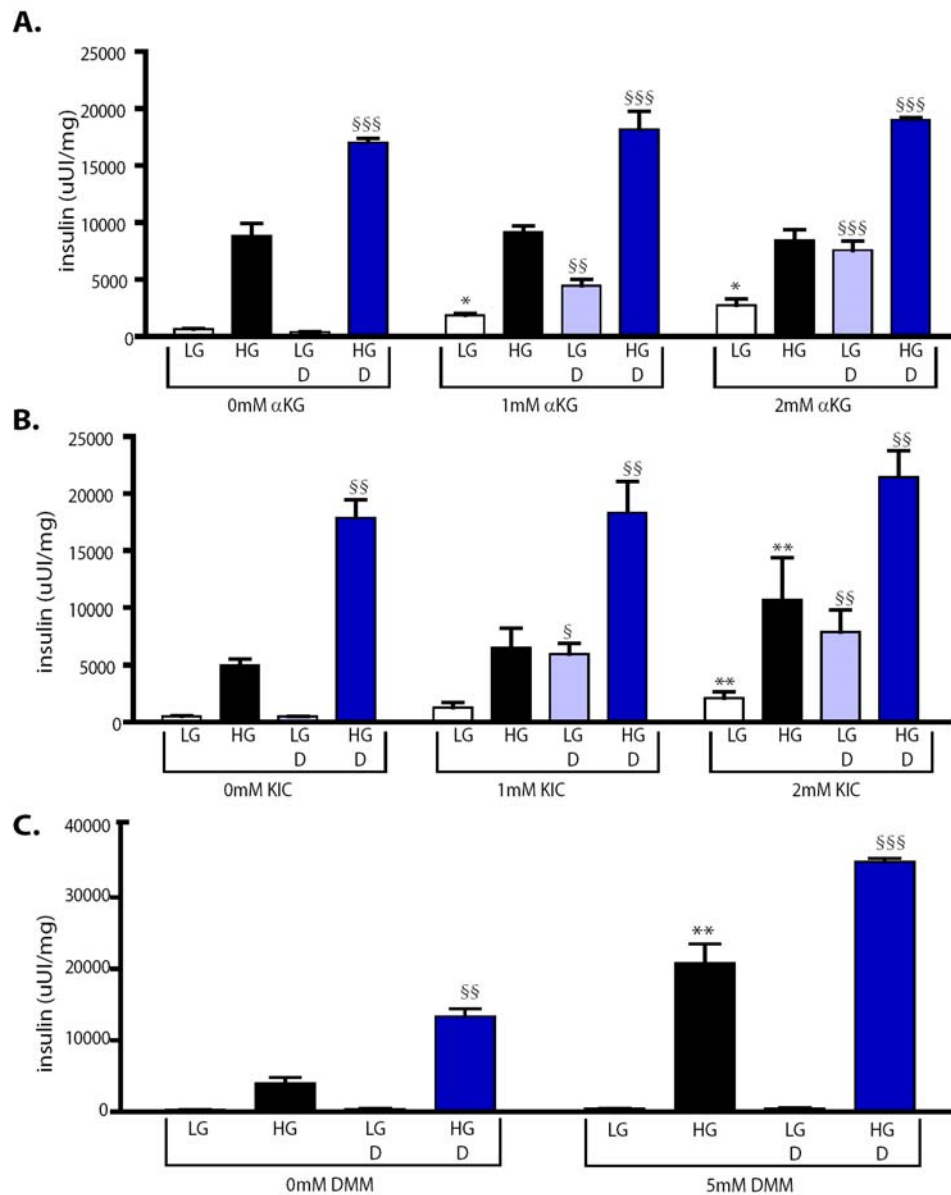


Figure 36. DEPDC enhances insulin secretion from non-glucose secretagogues.

832/3 INS-1 cells treated with α -ketoglutarate (A), α -ketoisocaproate (B), and dimethylmalate (C) +/- 1mM DEPDC in 2.5mM or 15mM glucose. Effects of α KG, KIC, DMM are relative to the untreated cells (* p <0.05, ** p <0.01) at corresponding glucose level. DEPDC effect (blue bars) determined vs. identical conditions without DEPDC (§ p <0.05, §§ p <0.01, §§§ p <0.001). D: DEPDC, LG: 2.5mM glucose, HG: 15mM glucose (n=3).

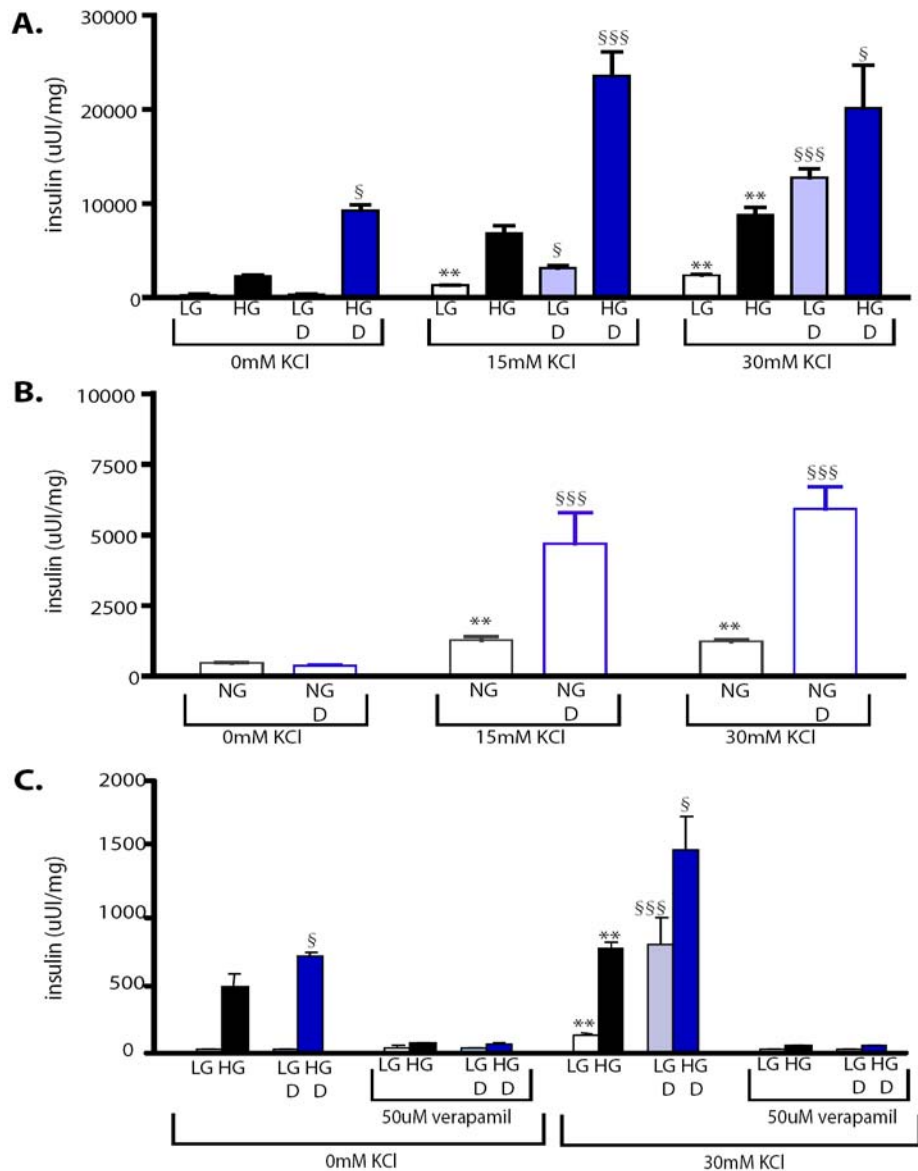


Figure 37. DEPDC and calcium-induced insulin secretion

832/3 INS-1 cells treated with KCl +/- 1mM DEPDC in 2.5mM or 15mM glucose (A) or 0mM glucose (B), and +/- 50uM verapamil (C). Analysis compared the effects of KCl relative to untreated (*p<0.05, **p<0.01 vs. corresponding glucose level). DEPDC effect (blue bars) determined vs. identical conditions without DEPDC (§p<0.05, §§p<0.01, §§§p<0.001) D: DEPDC, LG: low glucose, HG: high glucose, NG: no glucose. (n=3)

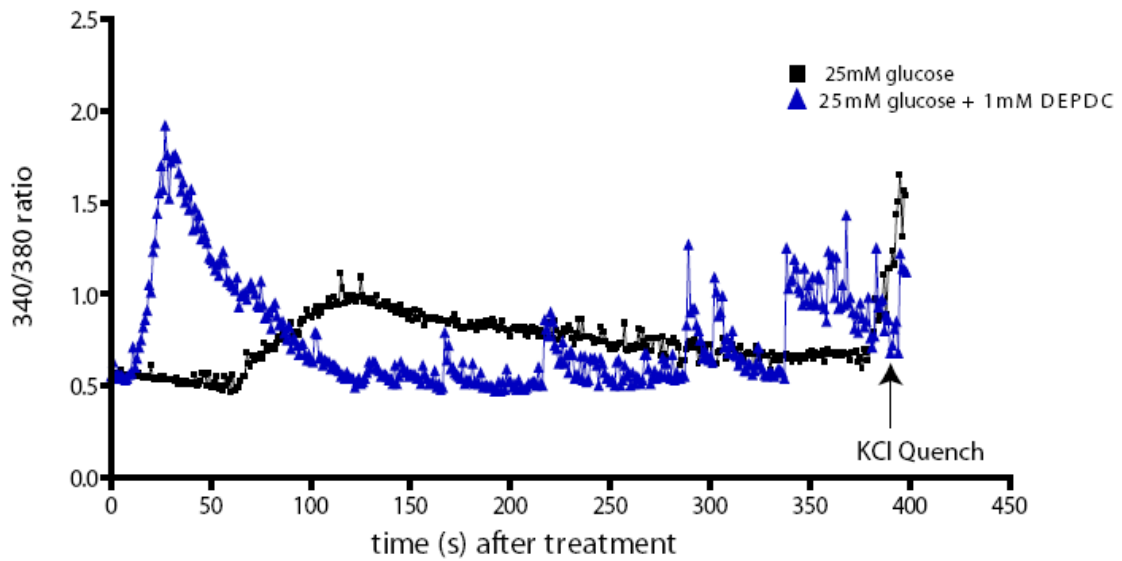


Figure 38. DEPDC alters calcium flux dynamics in INS-1 cells.

832/3 INS-1 cells were treated with 25mM glucose (black) or 25mM glucose with 1mM DEPDC (blue) (n=2). Depicted is a representative image of calcium mobilization in response to treatments. The x-axis shows the time (in seconds) relative to the administration of the treatment. The reaction was quenched with 30mM KCl as a positive control.

6. Future Directions and Concluding Remarks

Type 2 diabetes now affects more than 200 million people worldwide, and nearly one third of these individuals are undiagnosed. Type 2 diabetes is often preceded by a “metabolic syndrome,” which is characterized by insulin resistance and impaired glucose tolerance, yet is a potentially reversible metabolic state. Although the metabolic syndrome is clinically silent, it is associated with increased risk for cardiovascular disease and by the time of diabetes diagnosis, a significant fraction of patients exhibit both microvascular and macrovascular complications [264]. As metabolic syndrome progresses to type 2 diabetes, hyperglycemia exacerbates the already compromised vascular system, increasing the potential for life-threatening cardiovascular events. Given that diabetes comorbidities are often established during the preceding metabolic syndrome, therapeutic intervention prior to islet dysfunction and hyperglycemia is essential [265].

Current management of hyperglycemia and insulin resistance involves a stepwise approach of diet and exercise modifications followed by pharmaceutical intervention. There are several classes of medications designed to reduce fasting blood glucose and HBA1c, and each drug class targets a specific dysregulated process [264]. The most common initial drug of choice is metformin which activates AMPK and inhibits hepatic glucose output. Additional drug classes include the sulfonylurea compounds which increase insulin secretion, the thiazolidinediones (PPAR activators) that increase insulin

sensitivity, and α -glucosidase inhibitors which block the digestion and absorption of starch and other glucose-containing polymers in food[266,267]. While these medications have been shown to improve glycemic index nearly two times more effectively than diet and exercise alone, the side effects are numerous. Metformin, often the initial drug of choice, can cause fatal lactic acidosis in patients with kidney disease[268]. The sulfonylurea compounds can cause hypoglycemia, which is not only dangerous, but also increases appetite and weight gain, counterproductive in a disease where obesity is causal. In 2007, the FDA issued a safety alert on GlaxoSmithKline's Avandia (rosiglitazone of the thiazolidinediones) after a study showing a nearly seven-fold increase in heart failure in patients treated for three years [264,268-272].

One of the major challenges in treating type 2 diabetes is that side effects of hyperglycemia range from kidney disease to gallstones to sleep apnea, and monotherapy is often ineffective [268,269]. The likelihood of single-therapy failure is proportional to elevated baseline obesity (BMI, waist circumference), elevated initial fasting plasma glucose and plasma triglyceride values, and younger age at time of diabetes diagnosis [266]. In isolation, metformin, sulfonylurea compounds, and thiazolidinediones have a failure rate of 20-30% at five years, and at nine years, less than 25% of patients achieve their target fasting blood glucose or HbA1c [266,267,270]. Insulin injections have been shown to improve glycemic index when added in combination with oral drugs. Incretin mimetics (DPP4 inhibitor sitagliptin; GLP1 analog exenatide; amylin analog pramlintide) have recently gained interest as diabetes therapeutics. These newer peptide analogs do

not cause hypoglycemia and have the potential to aid in weight loss, however they are very costly[266,267,270].

Despite the shortcomings of current pharmacological intervention, the need for timely management of blood sugar in diabetic patients is clear. Hyperglycemia has predictable toxic effects on multiple organs, and the duration of unmanaged blood glucose is positively correlated with the risk for vascular-related adverse events [273,274]. As diabetes progresses, the number of organ systems disturbed by hyperglycemia increases, as does the number medications required to treat all symptoms of disease. However, if we can establish how metabolic syndrome progresses to diabetes, this may provide opportunities to identify the early pathogenic processes that affect multiple tissues. In identifying upstream drug targets in tissues causal for whole body metabolic dysregulation, we will aid in development of monotherapy that may be effective in a broader patient population, and can intervene prior to frank diabetes.

6.1 Intrahepatic Studies

In vitro Mechanistic Studies

In states of hepatic insulin resistance caused by obesity-related overnutrition, liver gluconeogenic and lipid oxidation pathways are dysregulated, leading to increased plasma glucose and exacerbation of already elevated plasma free fatty acids. In the insulin resistant state, the liver is unable to perform its role in organizing whole body fuel

metabolism. The liver is a critical driver for metabolic dysregulation in diabetes development[7,226,234], and many drugs (including metformin and PPAR agonists) have been designed to correct hepatic processes that are dysfunctional [264,269,271,275-279]. These drug classes, however, are not effective in all diabetic individuals, which could indicate the presence of unidentified pathways that contribute to liver malfunction.

Hepatic glucose output is not only a major contributor to overall hyperglycemia, but is also used as a marker of disease progression and severity[24,224,228]. Our F2 study revealed a pathway by which glutamine can regulate expression of *Phosphoenolpyruvate carboxykinase (Pck1)*, a key gluconeogenic enzyme, in isolated hepatocytes[104]. While we have gained insights into what glutamine-derived metabolites may be mediating the *Pck1* response, we would like to extend these studies to understand how several metabolic pathways are coordinated by individual, or groups of, regulatory analytes.

Our *in vitro* hepatocyte experiments reveal that glutamine not only alters gluconeogenic pathways via *Pck1*, but also induces expression of *Arg1* and *Ass1* acting in the urea cycle. We also show that glutamine increases hepatic levels of ornithine and decreases hepatic citrulline, consistent with the changes in *Arg1* and *Ass1* mRNA, respectively (Chapters 3 and 4). It is interesting to note that amino acids involved in the urea cycle (asx, glx, ornithine, arginine) are among the subset of the highly correlated amino acid group that uniquely map to a region of chromosome 7 [104], colocalizing with the eQTL of *Pck1* and urea cycle enzymes *Argininosuccinate lyase (Asl)*, *Arginase 1*

(*Arg1*), *Carbamoyl-phosphate synthetase 1 (Cps1)*, and *Argininosuccinate synthetase 1 (Ass1)* mRNA (Figure 39). Of potential interest, also located in this broad region of chromosome 7 is transcription factor CCAAT enhancer binding protein α (C/EBP α). The *Cebpa*^{-/-} mouse shows reduced transcription of all enzymes of the urea cycle except for *Ornithine transcarbamylase (Otc)*, which does not map to this region in our F2 population (Figure 40). Additionally, this mouse displays hypoglycemia, presumably from reduced *Pck1* and *G6pc* expression[280]. While gene mapping does not provide the level of resolution necessary to identify individual genes, it does provide us with a list of candidate genes, such as C/EBP α , for further testing.

We have begun to explore our C/EBP α hypothesis and find that glutamine does not alter expression of *Cebpa* mRNA (data not shown). However, given that transcription factors are altered via post-translational modifications, glutamine could modify C/EBP α activity to mediate these changes in gene expression. One way to test this is via a C/EBP reporter assay in which a C/EBP responsive luciferase construct is transfected into isolated hepatocytes receiving control media or 10mM glutamine, and the reporter plasmid containing the C/EBP binding site will provide luciferase readouts. One of the drawbacks of this type of assay is that the glutamine effect on *Pck1* expression is time limited; after 24 hours, the hepatocytes no longer show the glutamine-induced elevation. It will therefore be critical to transfect the plasmid in the presence of the glutamine treatment, and to interpret results anticipating submaximal luciferase readouts.

The C/EBP α hypothesis just described illustrates one way to use gene mapping (eQTL and mQTL) to identify transcription factors that may orchestrate entire pathways driven by a single metabolite. We would also like to identify novel transcription factors and regulatory mechanisms that mediate metabolite-induced gene expression. One means by which this can be done is chromatin immunoprecipitation-on-chip (ChIP-on-chip) assays, which allows us to investigate group of genes and identify DNA sequences occupied by specific binding proteins in response to treatment. The majority of these DNA binding proteins will be those involved in chromatin remodeling such as transcription factors, replication machinery, and histone modifying proteins. The ultimate goal of the ChIP-on-chip method is to localize protein binding sites in order to identify functional elements in the genome or a group of genes [170,281].

To apply this technology to our system, we first design a chip containing genes most highly correlated to and sharing QTL with a metabolite of interest. We can then compare metabolite treatment induced specific common modifications in this group of genes [281]. We can ask whether a transcription factor (C/EBP α in the case of glutamine) is involved, but this analysis is not restricted to a pre-identified candidate protein such a known transcription factor. This technique allows us to test for the presence of RNA polymerase II binding (transcription initiation complex), phosphorylated RNA polymerase II (active transcription), and epigenetic factors; we are thus not limited to utilizing antibodies directed at known transcription factors such as C/EBP α . We can assess whether two genes that are altered by a metabolite also have similar modifications

in their promoter regions. From this data we can hypothesize and test likely effectors are responsible for the intermediate step between metabolite and transcription. For example, our F2 data indicates that *Arg1*, *Pck1*, and *Ass1* co-map and are transcriptionally induced by glutamine. If we show by ChIP-on-chip analysis that glutamine-treatment alters the acetylation status of the promoter region, we can hypothesize that a the glutamine mechanism may involve histone deacetylase or related enzyme.

A major advantage of ChIP-on-chip analysis is that we can develop the molecular pathway upstream of gene transcription to identify therapeutic targets in steps prior to protein production and enzymatic reactions. Many of the current drugs that target hepatic glucose production are aimed at reducing the activities of gluconeogenic enzymes including Glucose-6-phosphatase and Fructose 1,6-bisphosphatase[275,279,282], yet targeting specific enzymes may be too narrow of an approach to treat dysregulated glucose metabolism. With this ChIP-on-chip technology, we can examine entire metabolic pathways that are altered transcriptionally by a given metabolite(s), and can include genes of different pathways such as lipid oxidation and gluconeogenesis, to see if distinct metabolic processes are coordinated by a common upstream process. We can reveal not only cooperative regulation between pathways, but can also uncover the mechanism by which these pathways are co-regulated. This allows for identification of therapeutic targets very upstream of individual enzymes, and development of such drugs will limit the need for combination therapy.

Another benefit of ChIP-on-chip technology is that our analysis is not limited to the proximal promoter region of the genes of interest. For decades, traditional assays for measuring transcriptional activity included generating plasmids containing the proximal promoter of a gene of interest and tagging it with a reporter gene. The drawback of such a method is that it only allowed for a few kilobases of DNA to be assessed due to plasmid construction and subsequent transfection. It is becoming increasingly apparent that there are regulatory regions both significantly upstream and downstream of the promoter start site [282]. The chip analysis allows us to look at entire genes to define unique sequences in exons, UTRs, and distal promoters that alter regulation. Identifying these regions in multiple genes that control relevant metabolic pathways could provide insights into new therapeutic targets.

In vivo Significance of Hepatocyte Discoveries

Glutamine regulation of gluconeogenesis *in vitro* is a novel and exciting mechanism by which amino acids could alter hepatic metabolism. Glutamine is the most abundant amino acid in the plasma, and in states of catabolism, including starvation and diabetes, muscle protein breakdown delivers amino acids in the form of glutamine and alanine to the liver for glucose production. Hence, plasma glutamine is similar to the free fatty acids in that both are metabolites that reflect of altered nutritional status and have the ability to transcriptionally regulate liver metabolism.

Our hepatocyte experiments have only started to uncover the molecular mechanism of metabolite-induced transcription. A longer-term goal will be to characterize the biological effects of glutamine *in vivo* and to describe how this single amino acid may contribute to diabetes clinical traits (in this case, hepatic glucose output manifesting in hyperglycemia). We show in the time course data that often the stress of obesity is required to reveal strain-specific differences between the B6 and the BTBR metabolite and gene expression profiles. It is therefore more probable that obesity aggravates the genetic predisposition in the BTBR and it is only when environmental and genetic factors combine that we see hyperglycemia.

The extension of our glutamine studies must consider both the lean and obese states of both strains. The time course data shows an obesity-dependent decrease in glutamine and glutamate in both strains at 10 weeks, but that α -ketoglutarate is increased in the livers of the BTBR. We hypothesize that this increased level of α -ketoglutarate prevents the BTBR from detecting relative glx depletion the obese state, and thus *Pck1* expression is not downregulated in the BTBR^{ob/ob} mice. We did not see a strain-specific difference in the lean hepatocyte *Pck1* response to 10mM glutamine and this could be because it was not confounded by obesity; for example, counter-regulatory mechanisms to prevent *Pck1* may remain intact in the B6 animal when made obese but are dysfunctional in the BTBR^{ob/ob} mice.

We can administer metabolites to lean and obese mice via tail vein injection, intraperitoneal injection, or oral gavage and measure hepatic gene expression in the

context of other physiological changes that come with obesity (hypertriglyceridemia, hyperglycemia in the BTBR, etc). In these animals receiving metabolites, we can also measure hepatic glucose output and hepatic insulin sensitivity via hyperinsulinemic euglycemic clamp to see if individual analytes can change liver physiology as a whole.

In these *in vivo* studies, we would like to not only verify that both 10-week-old lean B6 and BTBR mice recapitulate our *in vitro* glutamine induction of *Pck1*, but also test our hypothesis that increased levels of hepatic glutamine and α -ketoglutarate in an obese background promotes gluconeogenesis and hepatic glucose output. To determine whether glutamine is sufficient to regulate *Pck1*, we can administer glutamine to the obese B6 and BTBR mice at 4 and 10 weeks of age to elevate hepatic levels of glx to that of their lean counterparts. If glutamine is sufficient, then maintaining elevated glx should maintain *Pck1* expression at the level seen in the lean animals. To determine whether increased α -ketoglutarate is sufficient override glx regulation and prevent the obesity-dependent decrease in *Pck1* mRNA, we can administer α -ketoglutarate to the diabetes-resistant B6^{ob/ob} animals at 4 and 10 weeks of age.

Our *in vitro* data suggests that enhancing glutamine metabolism by BCH potentiates the glutamine-induction of *Pck1*, and that α -ketoglutarate is a key mediator. We would like to test whether hyperglycemia in the 10-week-old BTBR^{ob/ob} mice can be prevented by knockdown of *Glutamate dehydrogenase* either by adenovirus-mediated delivery of *Glud1*shRNA or by conditional gene knock-out in transgenic mice. It would also be instructive to determine if progression to hyperglycemia occurs in the 4-week-old

BTBR^{ob/ob} mice if we elevate certain hepatic metabolites or express a constitutively active (GTP-inhibition insensitive) Glud1 enzyme.

This glutamine study has only scratched the surface of identifying metabolite regulators of transcription. Our laboratory and others have shown that in the “pre-diabetic” state, obese subjects display unique hepatic and plasma metabolic profiles. We show in our time course study that at 4 weeks, the BTBR^{ob/ob} animal, which will become diabetic, has an elevation in several hepatic amino acids (alanine), organic acids (α -ketoglutarate and pyruvate) and acyl-carnitine derivatives (isovaleryl, 3-methylbutyryl, 2-methylbutyryl, and tiglyl carnitine) relative to the 4-week-old B6^{ob/ob} animal. The STEDMAN study, involving comprehensive metabolic profiling of obese and insulin resistant, but non-diabetic humans revealed significant increases in eight of 15 amino acids surveyed, as well as select short-and medium-chain acyl-carnitines compared to lean controls [59]. The aforementioned techniques can easily be applied to other metabolites or groups of metabolites to identify novel mediators of transcriptional and physiological changes.

6.2 Interorgan Studies

Broad Application of Causal Network Construction

Studying the F2 liver in isolation has been enlightening, but we recognize that liver reacts to peripheral tissues and has effects on other organs, including the pancreatic

islets. We showed in Chapter 3 and 4 that various metabolites affect hepatic transcription, yet we understand that the liver is part of a larger, whole-body system. We attempt to address this by showing that dicarboxylates (potentially liver-derived) can alter insulin secretion, both by metabolic and more distal mechanisms (Chapter 5).

In Chapters 4 and 5, we discussed a time course study and focused on liver metabolite levels and hepatic expression data. In that study, the islet, gastrocnemius, soleus, hypothalamus, and adipose were analyzed by microarray as well, generating unique profiles for each gene depending on the strain (B6 vs. BTBR), age (4 week vs. 10 week), and physiological state (lean vs. obese). Our laboratory has attempted to characterize the interorgan communication of diabetes by defining modules of transcripts with “like expression profiles” across all six tissues [216]. This analysis initially looked at tissues independently and categorized the transcript into one of 15 patterns based on strain-dependent, age-dependent, and/or obesity-dependent changes in expression. This approach allowed for identification of primary and secondary drivers of changes in gene expression in the various tissues profiled. For example, a gene module that shows an elevation in the 4-week-old BTBR^{ob/ob} but not the 4-week-old B6^{ob/ob} is considered to be causal since it was differentially expressed in the “pre-diabetic” state; a module that only is elevated in the diseased 10 week BTBR^{ob/ob} is considered to be a result of diabetes. We also identified modules that show a strict age-effect or obesity-effect.

Transcripts of similar pattern were grouped into modules for each tissue and a first principal component was identified for each module. The first principal component

of each module was then correlated to that of the other modules not only within the same tissue, but also between tissues (Figure 41). These modules gain physiological relevance when showing a correlation to physiological traits such as plasma insulin or fasting glucose (Figure 42).

The above approach has demonstrated that there is high correlation between transcript groups of different tissues, suggesting interorgan communication. It has not, however, allowed us to infer causality between modules within or between tissues. As we described in Chapter 2, to define a causal relationship between two correlated phenotypes, we need a variable uniquely affecting one phenotype to infer direction. Our method for building a QTL-directed dependency graph (QDG) describes using a genetic locus as this variable, but this variable could be other covariates as well. We validated this methodology in a small F2 population in Chapter 3, and have now expanded on our initial F2 study. We currently have gene expression data from a much larger F2 population from tissues that were analyzed in the time course study (liver, islet, hypothalamus, adipose). By using the methods developed in Chapter 2 and 3, we are currently generating more complex tissue mega-networks by including up to a thousand transcripts. One organ will have several of these mega-networks, each mega-network consisting of several thousand transcripts.

These mega-networks within a single organ will be characterized by a principal component, similar to what was done in the time course study. Now, however, because we have QTL data, we can define causality between correlated networks. The causal

relationship between mega-networks is not confined to one organ, and therefore one can demonstrate the causal role that a mega-network in the liver has on a mega-network in the islet. When characterizing interorgan communication, it is essential to include tissue metabolite measurements with expression data, since it is the chemical changes induced by the transcriptional changes, not the mRNA itself that physically bridges the gap between the tissues. Metabolites generated in an individual organ can be transported in the plasma, and thus serve as the mobile mediator of communication between the tissues. We have demonstrated in Chapter 3 that metabolites are a powerful complement to gene expression data when investigating the pathway from genotype to phenotype in a single organ. In our *in vitro* experiments, we have shown that amino acids can alter gene expression, thus adding this class of analytes to a growing list of metabolite signaling molecules [161-163]. Just as was done for the liver in Chapter 3, one can use mQTL and eQTL of individual tissues to construct causal networks within each tissue.

To truly understand interorgan communication, it is essential to construct a plasma network. Using QTL data, we will then be able to draw a causal arrow from a primary tissue network to the plasma network to a reactive tissue network. Plasma metabolite data will be of great importance to elucidate interorgan communication, but other secreted factors should be considered in building this soluble network. We have measured approximately 50 plasma proteins including growth factors (FGF, EGF, growth hormone) and inflammatory mediators (IL6, INF γ). Cytokines, whether derived from adipose tissue or peripheral blood cells, act through inflammatory signaling pathways and

alter intermediates of the insulin signaling cascade. Additionally the cytokines are causal for the cardiovascular consequences of obesity and insulin resistance[283-285]. Adipose tissue also secretes several adipokines that alter function of peripheral tissues: adiponectin increases insulin sensitivity, reduces hepatic glucose output, and is inversely related to diabetes risk; leptin decreases appetite and increases energy expenditure[286-288]. We also can include blood glucose, insulin, and plasma triglycerides as phenotypes within this plasma network, just as they were included in correlation analysis in the time course study.

In identifying the role of plasma intermediates in whole body networks we not only can understand interorgan communication in a systemic disease, but we also may identify specific, accessible biomarkers of diabetes development and improve current screening procedures. By defining plasma metabolites/factors that cause dysfunction in “primary” organs (liver or adipose), we can identify those at risk for developing metabolic syndrome. Plasma free fatty acids, the majority of which are derived from the diet, are an example of this class of initial causal metabolites, and have led to lifestyle and pharmacologic interventions in individuals with elevated cholesterol levels. Additionally, plasma factors may be reflective of organ pathology. By defining agents that serve as an intermediate between the liver and islet, we can identify those individuals that have hepatic insulin resistance before β -cell decompensation occurs. Finally, by defining plasma metabolites that ultimately result from diabetes, we can use this

biomarker in place of fasting plasma glucose to monitor disease progression and response to treatment.

In vivo Investigation of Regulatory Genetic Regions

One of the challenges in the mega-network experiment is identifying “master regulators” that significantly alter downstream events, whether these regulators are individual transcripts, metabolites, or entire tissue/plasma networks. In a disease as complex as diabetes, it is likely that there are several master regulators that act in concert to cause one to manifest clinical diabetes. It is therefore unrealistic to attempt single gene/metabolite manipulations in attempts to see whole body physiological changes. What we can hope to find, however, are genetic regions that control entire networks in one or several tissues. We can imagine metabolites and transcripts of a given mega-network having overlapping mQTL and eQTL such that the entire mega-network maps to a common genetic region(s) (Figure 43a).

We can also imagine that a mega-network in one tissue maps to a similar QTL as downstream/upstream mega-network in another tissue or in the plasma. From this data, we can identify causal regions on multiple chromosomes and assess whether interaction between these various regions is required. If we were to map only gene expression or only metabolite abundance, we would be left with perhaps hundreds of candidate regions. While our analysis may still suggest that several regions of the genome containing causal driving genes, we will be left with fewer false positive regions by virtue of mapping

entire networks (containing metabolite and expression data) from multiple tissues to these loci, and thus have more confidence to follow up on these specific QTL.

To explore the contribution of one particular genetic region to these phenotypes—remembering again that phenotype includes gene expression, metabolite abundance, clinical traits, and perhaps entire networks—we must analyze this genetic locus in isolation. Since a congenic mouse has only one region of the genome that differs genetically from a parent strain, it is a model for such a confirmation study. Consider a congenic strain containing genetic region derived from B in a genomic background of strain A. If this congenic strain (with an introgressed region from strain B) recapitulates the phenotype seen in the homozygous parental strain B, then it suggests that this specific genetic region contains genes contributing to the phenotype appearance.

While production of congenic strains takes several generations, our laboratory has already produced congenic strains for chromosomes 2 and 16 with BTBR chromosomal regions introgressed into the B6 background. The chromosome 2 and 16 congenic strains contain peak markers for plasma insulin and glucose, respectively, but the regions to which these diabetes phenotypes map remains broad. With the addition of tissue network mapping we can gain insight into which smaller genetic regions to focus on for construction of sub-congenics (Figure 43b). We can measure tissue metabolites and gene expression and also assess the clinical phenotypes of these congenic strains. Plasma insulin and glucose levels can be measured, glucose-stimulated insulin secretion assays can be performed, and insulin sensitivity can be determined via hyperinsulinemic

euglycemic clamp. To investigate smaller regions of the genetic loci, these congenic mice are further back-crossed to B6^{ob/ob} mice heterozygous to produce sub-congenic strains.

Congenic experiments will most likely confirm that isolated genetic regions do have an influence on the transcripts, metabolites, clinical traits, and perhaps whole networks that map to these loci. The difficulty will come in interpreting these results since it is likely that one genetic region will be necessary but not sufficient to completely alter the aforementioned measurements. When comparing the phenotypes of any single congenic strain to the parents, we will probably achieve intermediate values between the parental strains, indicating that other genetic regions or environmental factors contribute. A further complexity arises from the fact that the only one-way arrow of causality that can be drawn from genotype to phenotype (clinical traits, expression, metabolites). While the co-mapping does generate hypotheses about how genotype could affect pathways related to clinical traits or how a genetic region regulates two tissue networks, it does not indicate causality between the phenotypes that map. Additional *in vivo* or *in vitro* experiments are required to test the causal relationship between these two phenotypes as was done in the glutamine experiment described in Chapters 3 and 4.

While congenic studies do leave questions unanswered, they provide invaluable information. Congenic strains do not provide the level of resolution to investigate individual genes within that region, but the technique of overlaying cQTLs, eQTLs, and mQTLs or entire network QTL from several tissues and plasma gives a region that can affect multiple aspects of diabetes pathology, and thus contains key targets for therapy.

The genetic regions that affect phenotype measurements in numerous tissues will not only be fewer in number across the genome, they should also be narrow with fewer individual genes to screen. This will hasten identification of targets as this smaller list of genes is systematically analyzed and also provide candidate regions that affect multiple organs of diabetes pathophysiology.

6.3 Summary

In conclusion, we began by using metabolic and microarray data in an F2^{ob/ob} population segregating for genotype and diabetes phenotypes to define metabolites and transcripts that are correlated and share QTL. From this data we constructed a causal network, providing a data-driven hypothesis about how a metabolite and select transcripts could be related. We validated this network in an *in vitro* system, and this served as a pilot study for generation of more complex hypotheses about groups of metabolites and entire metabolic pathways. We focused on glutamine and glutamine-derived metabolites (α -ketoglutarate, glutamate, and alanine) and how these regulate a single transcript, *Pck1*.

Still to be determined is the more precise mechanism by which a single metabolite can regulate groups of genes in a single or multiple related pathways and the *in vivo* significance of such regulation. New technologies, including, but not limited to the CHIP-on-chip analysis, will allow us to fill in the gaps of the molecular mechanisms in which metabolites alter expression of biochemical pathways of an entire organ. We are currently

expanding our initial liver studies to the whole body by constructing multi-tissue networks in attempts to understand interorgan communication *in vivo*. We have subsequently tested the *in vitro* insulin secretory capacity of select liver-derived metabolites to explore the role metabolites play in this organ cross-talk. By identifying upstream dysregulated metabolic pathways and causal metabolite mediators we may pinpoint ways to block the progression of clinically silent metabolic syndrome to frank type 2 diabetes.

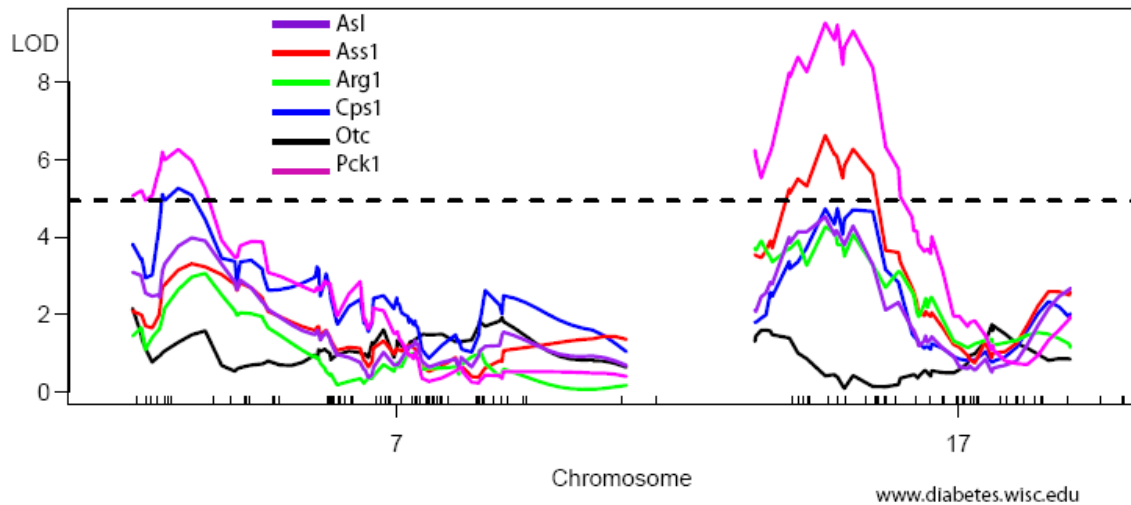
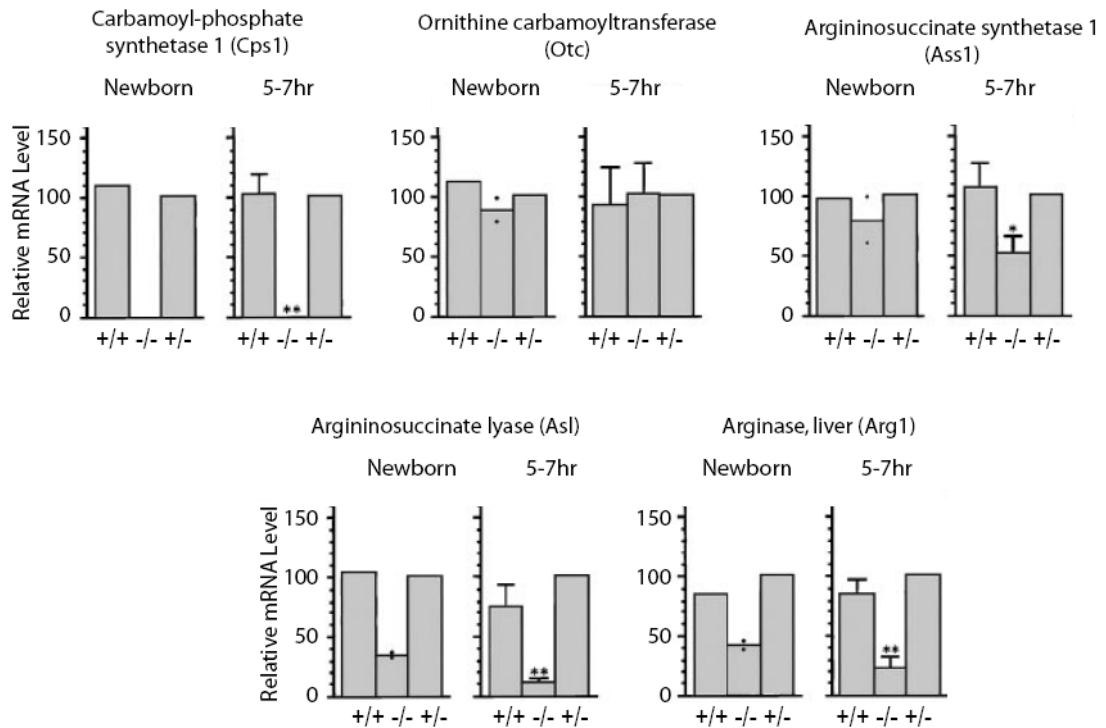


Figure 39. Expression QTL (eQTL) of urea cycle enzymes in liver tissue.

Liver samples from 500 leptin-deficient F2 mice were analyzed by microarray. Transcripts of the urea cycle enzymes *Arginase 1* (*Arg1*, green), *Argininosuccinate lyase* (*Asl*, purple), *Argininosuccinate synthetase 1* (*Ass1*, red), *Carbamoyl-phosphate synthetase 1* (*Cps1*, blue) have putative eQTL common with *Phosphoenolpyruvate carboxykinase, cytosolic* (*Pck1*, pink) mRNA on chromosomes 7 and 17. *Ornithine transcarbamylase* (*Otc*, black) does not map significantly to these regions



Kimura, et al. (1998) J Biol Chem. 273(42):27505-27510.

Figure 40. Liver mRNA for urea cycle enzymes in C/EBP α knockout mouse.

Total liver RNAs from wild type (+/+), homozygous (-/-), and heterozygous (+/-) littermates within 2 hours (newborn) and 5-7 hours after birth were examined for urea cycle enzymes. Expression data was normalized to *Actb*. Significant difference from +/- animals was determined by Student's t-test (* $p < 0.05$, ** $p < 0.01$).

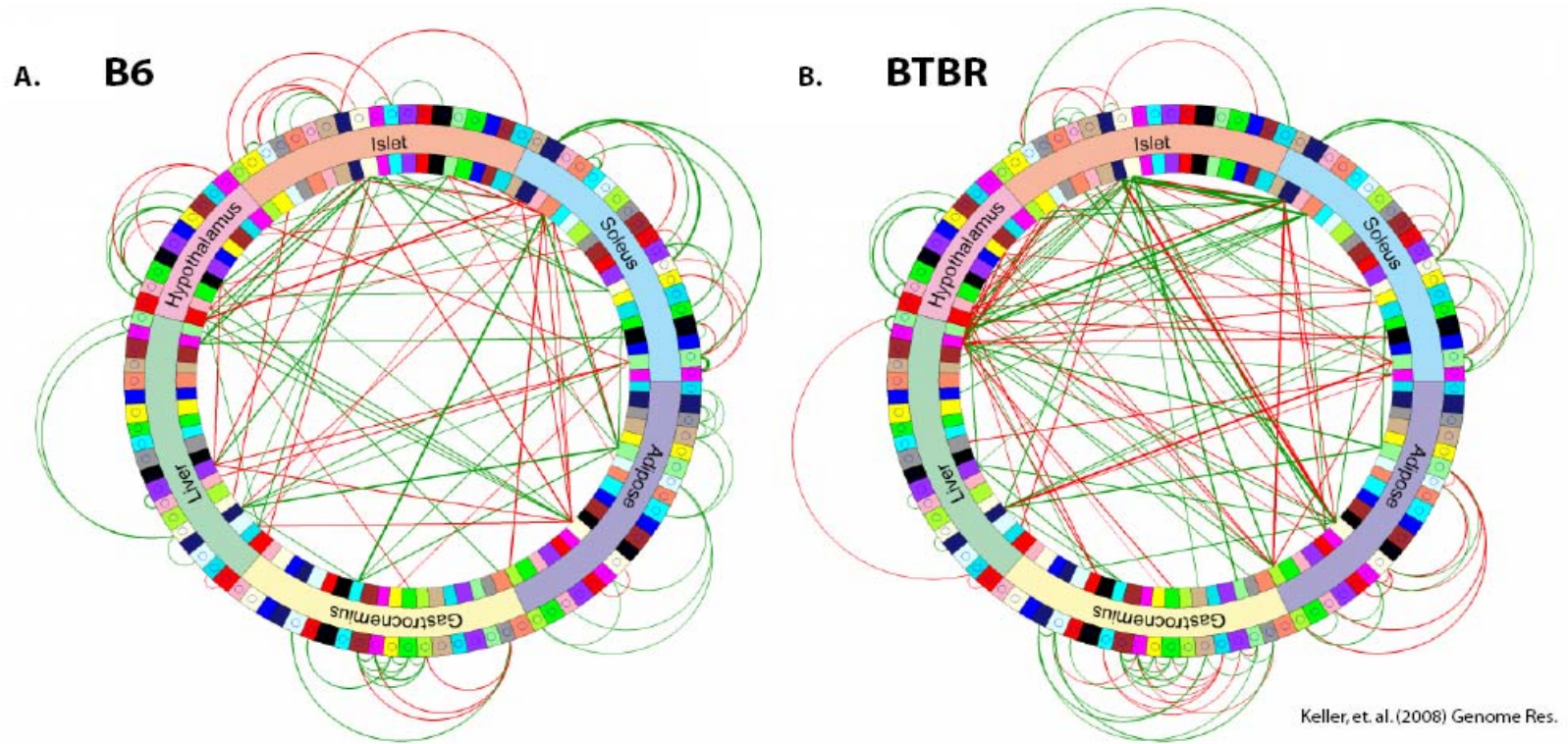
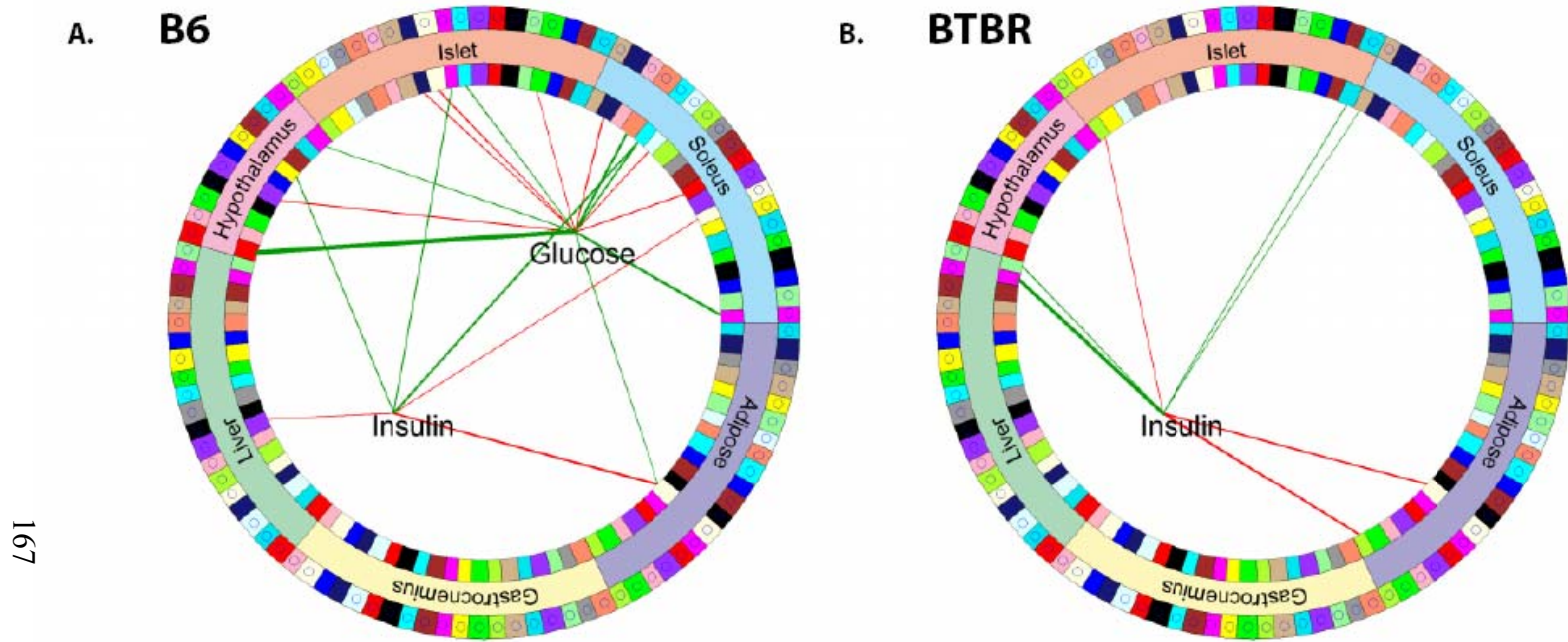


Figure 41. Gene expression network model shows strain-specific differences between B6 and BTBR mice.

Within each of the six tissues, transcripts displaying “like patterns” of gene expression across the eight subsets of mice were organized into a modules (colored bricks). A gene-gene network was constructed based on the partial correlation between the PC1 of each module. This was done both within and between the six tissues profiled. Inter-tissue edges within the network are shown as lines connecting the inside modules; intra-tissue edges are depicted as arcs connecting outside modules. Green indicates a positive partial correlation, red is negative.



Keller, et. al. (2008) Genome Res.

Figure 42. Modules can be correlated with physiological phenotypes.

Within each of the six tissues, transcripts displaying “like patterns” of gene expression across the eight subsets of mice were organized into a modules (colored bricks). A gene to clinical trait network was constructed for each strain based on the partial correlation between the module PC1 and either glucose or insulin. Line thickness is proportional to the magnitude of the partial correlation coefficient. Green indicates positive partial correlation, red is negative.

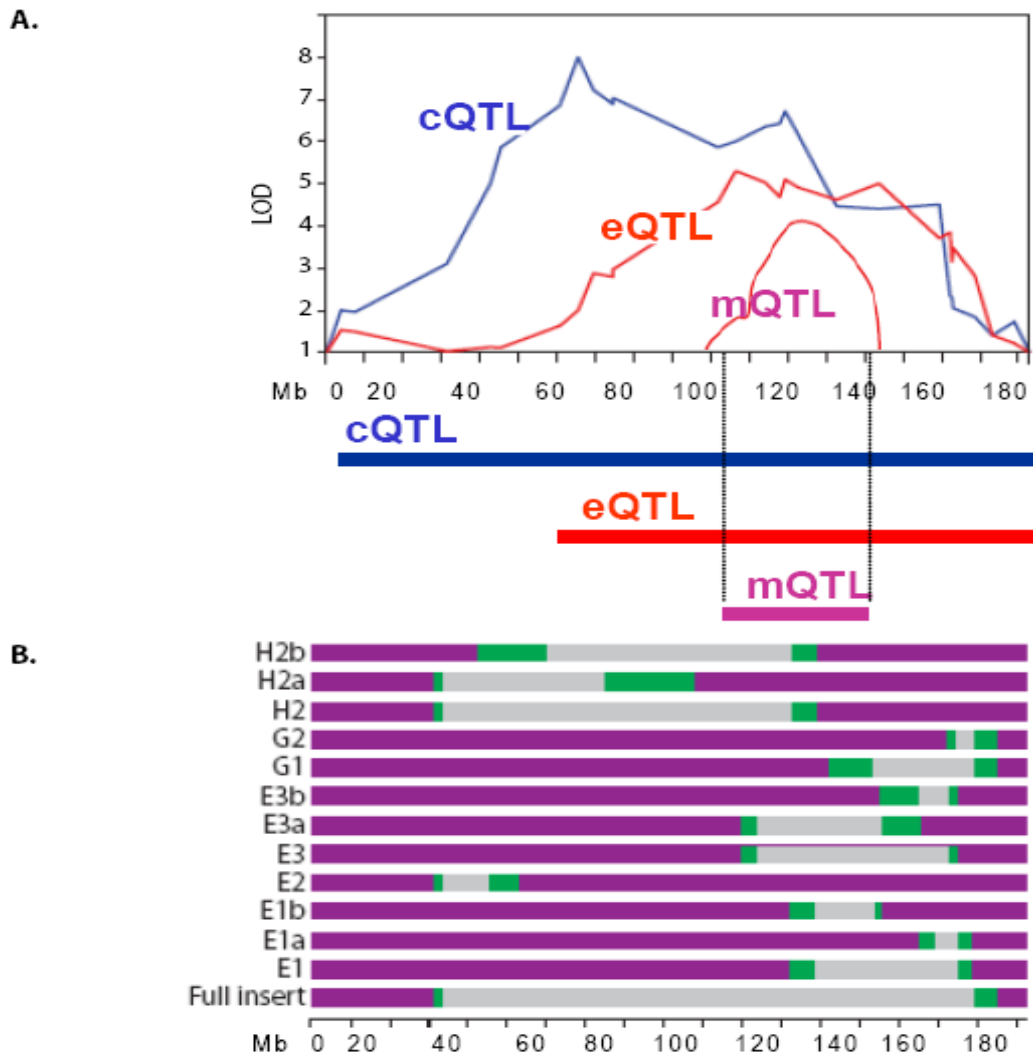


Figure 43. Hypothetical co-mapping QTL strategy.

A. Shown is a hypothetical illustration of how cQTL (for example insulin), eQTL and mQTL for individual or groups of transcripts and metabolites, respectively, could co-map. B. Schematic of the chromosome 2 congenic strains currently in the Attie lab (purple=B6; grey=BTBR; green=unknown genotype). The genetic region containing any one QTL is broad, but the region common to all traits helps to narrow the focus. In this example, subcongenic strains E3a, E3, E1b, and E1 likely contain a driver candidate gene affecting gene expression, metabolite abundance, and a clinical phenotype.

References

1. Mayfield J (1998) Diagnosis and classification of diabetes mellitus: new criteria. *Am Fam Physician* 58: 1355-1362, 1369-1370.
2. Kahn CR, Vicent D, Doria A (1996) Genetics of non-insulin-dependent (type-II) diabetes mellitus. *Annu Rev Med* 47: 509-531.
3. Weir GC, Laybutt DR, Kaneto H, Bonner-Weir S, Sharma A (2001) Beta-cell adaptation and decompensation during the progression of diabetes. *Diabetes* 50 Suppl 1: S154-159.
4. Gavin JR, 3rd (1998) New classification and diagnostic criteria for diabetes mellitus. *Clin Cornerstone* 1: 1-12.
5. Kuzuya T, Nakagawa S, Satoh J, Kanazawa Y, Iwamoto Y, et al. (2002) Report of the Committee on the classification and diagnostic criteria of diabetes mellitus. *Diabetes Res Clin Pract* 55: 65-85.
6. Muoio DM, Newgard CB (2008) Mechanisms of disease: molecular and metabolic mechanisms of insulin resistance and beta-cell failure in type 2 diabetes. *Nat Rev Mol Cell Biol* 9: 193-205.
7. Muoio DM, Newgard CB (2006) Obesity-related derangements in metabolic regulation. *Annu Rev Biochem* 75: 367-401.
8. Sung M, Kim I, Park M, Whang Y, Lee M (2004) Differential effects of dietary fatty acids on the regulation of CYP2E1 and protein kinase C in human hepatoma HepG2 cells. *J Med Food* 7: 197-203.
9. Latruffe N, Cherkaoui Malki M, Nicolas-Frances V, Jannin B, Clemencet MC, et al. (2001) Peroxisome-proliferator-activated receptors as physiological sensors of fatty acid metabolism: molecular regulation in peroxisomes. *Biochem Soc Trans* 29: 305-309.
10. Latruffe N, Vamecq J (1997) Peroxisome proliferators and peroxisome proliferator activated receptors (PPARs) as regulators of lipid metabolism. *Biochimie* 79: 81-94.
11. Steineger HH, Arntsen BM, Spydevold O, Sorensen HN (1997) Retinoid X receptor (RXR alpha) gene expression is regulated by fatty acids and dexamethasone in hepatic cells. *Biochimie* 79: 107-110.

12. Steineger HH, Arntsen BM, Spydevold O, Sorensen HN (1998) Gene transcription of the retinoid X receptor alpha (RXRalpha) is regulated by fatty acids and hormones in rat hepatic cells. *J Lipid Res* 39: 744-754.
13. Ide T, Shimano H, Yoshikawa T, Yahagi N, Amemiya-Kudo M, et al. (2003) Cross-talk between peroxisome proliferator-activated receptor (PPAR) alpha and liver X receptor (LXR) in nutritional regulation of fatty acid metabolism. II. LXRs suppress lipid degradation gene promoters through inhibition of PPAR signaling. *Mol Endocrinol* 17: 1255-1267.
14. Pawar A, Xu J, Jerks E, Mangelsdorf DJ, Jump DB (2002) Fatty acid regulation of liver X receptors (LXR) and peroxisome proliferator-activated receptor alpha (PPARalpha) in HEK293 cells. *J Biol Chem* 277: 39243-39250.
15. Choi CS, Savage DB, Kulkarni A, Yu XX, Liu ZX, et al. (2007) Suppression of diacylglycerol acyltransferase-2 (DGAT2), but not DGAT1, with antisense oligonucleotides reverses diet-induced hepatic steatosis and insulin resistance. *J Biol Chem* 282: 22678-22688.
16. Samuel VT, Choi CS, Phillips TG, Romanelli AJ, Geisler JG, et al. (2006) Targeting foxo1 in mice using antisense oligonucleotide improves hepatic and peripheral insulin action. *Diabetes* 55: 2042-2050.
17. Yu C, Chen Y, Cline GW, Zhang D, Zong H, et al. (2002) Mechanism by which fatty acids inhibit insulin activation of insulin receptor substrate-1 (IRS-1)-associated phosphatidylinositol 3-kinase activity in muscle. *J Biol Chem* 277: 50230-50236.
18. Zhang D, Liu ZX, Choi CS, Tian L, Kibbey R, et al. (2007) Mitochondrial dysfunction due to long-chain Acyl-CoA dehydrogenase deficiency causes hepatic steatosis and hepatic insulin resistance. *Proc Natl Acad Sci U S A* 104: 17075-17080.
19. Holland WL, Knotts TA, Chavez JA, Wang LP, Hoehn KL, et al. (2007) Lipid mediators of insulin resistance. *Nutr Rev* 65: S39-46.
20. Stratford S, Hoehn KL, Liu F, Summers SA (2004) Regulation of insulin action by ceramide: dual mechanisms linking ceramide accumulation to the inhibition of Akt/protein kinase B. *J Biol Chem* 279: 36608-36615.
21. Summers SA (2006) Ceramides in insulin resistance and lipotoxicity. *Prog Lipid Res* 45: 42-72.

22. Summers SA, Nelson DH (2005) A role for sphingolipids in producing the common features of type 2 diabetes, metabolic syndrome X, and Cushing's syndrome. *Diabetes* 54: 591-602.
23. Bergman RN, Ader M (2000) Free fatty acids and pathogenesis of type 2 diabetes mellitus. *Trends Endocrinol Metab* 11: 351-356.
24. Consoli A, Nurjhan N (1990) Contribution of gluconeogenesis to overall glucose output in diabetic and nondiabetic men. *Ann Med* 22: 191-195.
25. Fleig WE, Noether-Fleig G, Roeben H, Ditschuneit H (1984) Hormonal regulation of key gluconeogenic enzymes and glucose release in cultured hepatocytes: effects of dexamethasone and gastrointestinal hormones on glucagon action. *Arch Biochem Biophys* 229: 368-378.
26. Pilkis SJ, Claus TH (1991) Hepatic gluconeogenesis/glycolysis: regulation and structure/function relationships of substrate cycle enzymes. *Annu Rev Nutr* 11: 465-515.
27. Pilkis SJ, Granner DK (1992) Molecular physiology of the regulation of hepatic gluconeogenesis and glycolysis. *Annu Rev Physiol* 54: 885-909.
28. Pilkis SJ, el-Maghrabi MR, Claus TH (1990) Fructose-2,6-bisphosphate in control of hepatic gluconeogenesis. From metabolites to molecular genetics. *Diabetes Care* 13: 582-599.
29. Klover PJ, Mooney RA (2004) Hepatocytes: critical for glucose homeostasis. *Int J Biochem Cell Biol* 36: 753-758.
30. Kletzien RF, Weber CA, Stumpo DJ (1981) Coordinate regulation of gluconeogenesis by the glucocorticoids and glucagon: evidence for acute and chronic regulation by glucagon. *J Cell Physiol* 109: 83-90.
31. Chakravarty K, Cassuto H, Reshef L, Hanson RW (2005) Factors that control the tissue-specific transcription of the gene for phosphoenolpyruvate carboxykinase-C. *Crit Rev Biochem Mol Biol* 40: 129-154.
32. Olswang Y, Cohen H, Papo O, Cassuto H, Croniger CM, et al. (2002) A mutation in the peroxisome proliferator-activated receptor gamma-binding site in the gene for the cytosolic form of phosphoenolpyruvate carboxykinase reduces adipose tissue size and fat content in mice. *Proc Natl Acad Sci U S A* 99: 625-630.

33. Gurney AL, Park EA, Liu J, Giralt M, McGrane MM, et al. (1994) Metabolic regulation of gene transcription. *J Nutr* 124: 1533S-1539S.
34. Cheong J, Coligan JE, Shuman JD (1998) Activating transcription factor-2 regulates phosphoenolpyruvate carboxykinase transcription through a stress-inducible mitogen-activated protein kinase pathway. *J Biol Chem* 273: 22714-22718.
35. Crawford DR, Leahy P, Hu CY, Chaudhry A, Gronostajski R, et al. (1998) Nuclear factor I regulates expression of the gene for phosphoenolpyruvate carboxykinase (GTP). *J Biol Chem* 273: 13387-13390.
36. Faber S, O'Brien RM, Imai E, Granner DK, Chalkley R (1993) Dynamic aspects of DNA/protein interactions in the transcriptional initiation complex and the hormone-responsive domains of the phosphoenolpyruvate carboxykinase promoter in vivo. *J Biol Chem* 268: 24976-24985.
37. Giralt M, Park EA, Gurney AL, Liu JS, Hakimi P, et al. (1991) Identification of a thyroid hormone response element in the phosphoenolpyruvate carboxykinase (GTP) gene. Evidence for synergistic interaction between thyroid hormone and cAMP cis-regulatory elements. *J Biol Chem* 266: 21991-21996.
38. Hall RK, Scott DK, Noisin EL, Lucas PC, Granner DK (1992) Activation of the phosphoenolpyruvate carboxykinase gene retinoic acid response element is dependent on a retinoic acid receptor/coregulator complex. *Mol Cell Biol* 12: 5527-5535.
39. Hall RK, Sladek FM, Granner DK (1995) The orphan receptors COUP-TF and HNF-4 serve as accessory factors required for induction of phosphoenolpyruvate carboxykinase gene transcription by glucocorticoids. *Proc Natl Acad Sci U S A* 92: 412-416.
40. Imai E, Miner JN, Mitchell JA, Yamamoto KR, Granner DK (1993) Glucocorticoid receptor-cAMP response element-binding protein interaction and the response of the phosphoenolpyruvate carboxykinase gene to glucocorticoids. *J Biol Chem* 268: 5353-5356.
41. Imai E, Stromstedt PE, Quinn PG, Carlstedt-Duke J, Gustafsson JA, et al. (1990) Characterization of a complex glucocorticoid response unit in the phosphoenolpyruvate carboxykinase gene. *Mol Cell Biol* 10: 4712-4719.
42. Leahy P, Crawford DR, Grossman G, Gronostajski RM, Hanson RW (1999) CREB binding protein coordinates the function of multiple transcription factors

- including nuclear factor I to regulate phosphoenolpyruvate carboxykinase (GTP) gene transcription. *J Biol Chem* 274: 8813-8822.
43. Liu JS, Park EA, Gurney AL, Roesler WJ, Hanson RW (1991) Cyclic AMP induction of phosphoenolpyruvate carboxykinase (GTP) gene transcription is mediated by multiple promoter elements. *J Biol Chem* 266: 19095-19102.
 44. Park EA, Gurney AL, Nizielski SE, Hakimi P, Cao Z, et al. (1993) Relative roles of CCAAT/enhancer-binding protein beta and cAMP regulatory element-binding protein in controlling transcription of the gene for phosphoenolpyruvate carboxykinase (GTP). *J Biol Chem* 268: 613-619.
 45. Roesler WJ, McFie PJ, Dauvin C (1992) The liver-enriched transcription factor D-site-binding protein activates the promoter of the phosphoenolpyruvate carboxykinase gene in hepatoma cells. *J Biol Chem* 267: 21235-21243.
 46. Scott DK, Hall RK, Granner DK (1995) Retinoid receptors cause distortion of the retinoic acid response element in the phosphoenolpyruvate carboxykinase gene promoter. *Biochem J* 310 (Pt 2): 483-490.
 47. Scott DK, Mitchell JA, Granner DK (1996) The orphan receptor COUP-TF binds to a third glucocorticoid accessory factor element within the phosphoenolpyruvate carboxykinase gene promoter. *J Biol Chem* 271: 31909-31914.
 48. Scott DK, Mitchell JA, Granner DK (1996) Identification and characterization of the second retinoic acid response element in the phosphoenolpyruvate carboxykinase gene promoter. *J Biol Chem* 271: 6260-6264.
 49. Tebbey PW, Hall RK, Granner DK (1995) The role of the TATA box in the hormonal regulation of phosphoenolpyruvate carboxykinase gene expression. *Biochem Biophys Res Commun* 215: 1006-1013.
 50. Xing L, Quinn PG (1993) Involvement of 3',5'-cyclic adenosine monophosphate regulatory element binding protein (CREB) in both basal and hormone-mediated expression of the phosphoenolpyruvate carboxykinase (PEPCK) gene. *Mol Endocrinol* 7: 1484-1494.
 51. Marshall S (2006) Role of insulin, adipocyte hormones, and nutrient-sensing pathways in regulating fuel metabolism and energy homeostasis: a nutritional perspective of diabetes, obesity, and cancer. *Sci STKE* 2006: re7.

52. Pozefsky T, Tancredi RG, Moxley RT, Dupre J, Tobin JD (1976) Effects of brief starvation on muscle amino acid metabolism in nonobese man. *J Clin Invest* 57: 444-449.
53. Jacob R, Rosenthal N, Barrett EJ (1986) Characterization of glutamine transport by liver plasma membrane vesicles. *Am J Physiol* 251: E509-514.
54. Felig P (1973) The glucose-alanine cycle. *Metabolism* 22: 179-207.
55. Brosnan JT (2000) Glutamate, at the interface between amino acid and carbohydrate metabolism. *J Nutr* 130: 988S-990S.
56. Curi R, Newsholme P, Procopio J, Lagranha C, Gorjao R, et al. (2007) Glutamine, gene expression, and cell function. *Front Biosci* 12: 344-357.
57. Curi R, Lagranha CJ, Doi SQ, Sellitti DF, Procopio J, et al. (2005) Molecular mechanisms of glutamine action. *J Cell Physiol* 204: 392-401.
58. Curthoys NP, Watford M (1995) Regulation of glutaminase activity and glutamine metabolism. *Annu Rev Nutr* 15: 133-159.
59. Haqq AM, Lien LF, Boan J, Arlotto M, Slentz CA, et al. (2005) The Study of the Effects of Diet on Metabolism and Nutrition (STEDMAN) weight loss project: Rationale and design. *Contemp Clin Trials* 26: 616-625.
60. Wijekoon EP, Skinner C, Brosnan ME, Brosnan JT (2004) Amino acid metabolism in the Zucker diabetic fatty rat: effects of insulin resistance and of type 2 diabetes. *Can J Physiol Pharmacol* 82: 506-514.
61. Marliss EB, Aoki TT, Pozefsky T, Most AS, Cahill GF, Jr. (1971) Muscle and splanchnic glutamine and glutamate metabolism in postabsorptive and starved man. *J Clin Invest* 50: 814-817.
62. Russell DM, Walker PM, Leiter LA, Sima AA, Tanner WK, et al. (1984) Metabolic and structural changes in skeletal muscle during hypocaloric dieting. *Am J Clin Nutr* 39: 503-513.
63. Aoki TT, Muller WA, Cahill GF, Jr. (1972) Hormonal regulation of glutamine metabolism in fasting man. *Adv Enzyme Regul* 10: 145-151.
64. Aoki TT, Toews CJ, Rossini AA, Ruderman NB, Cahill GF, Jr. (1975) Glucogenic substrate levels in fasting man. *Adv Enzyme Regul* 13: 329-336.

65. Felig P, Wahren J, Sherwin R, Palaiologos G (1977) Amino acid and protein metabolism in diabetes mellitus. *Arch Intern Med* 137: 507-513.
66. Felig P, Wahren J, Karl I, Cerasi E, Luft R, et al. (1973) Glutamine and glutamate metabolism in normal and diabetic subjects. *Diabetes* 22: 573-576.
67. Fafournoux P, Bruhat A, Jousse C (2000) Amino acid regulation of gene expression. *Biochem J* 351: 1-12.
68. Averous J, Maurin AC, Bruhat A, Jousse C, Arliguie C, et al. (2005) Induction of IGFBP-1 expression by amino acid deprivation of HepG2 human hepatoma cells involves both a transcriptional activation and an mRNA stabilization due to its 3'UTR. *FEBS Lett* 579: 2609-2614.
69. Kilberg MS, Pan YX, Chen H, Leung-Pineda V (2005) Nutritional control of gene expression: how mammalian cells respond to amino acid limitation. *Annu Rev Nutr* 25: 59-85.
70. Tollet P, Stromstedt M, Froyland L, Berge RK, Gustafsson JA (1994) Pretranslational regulation of cytochrome P450A1 by free fatty acids in primary cultures of rat hepatocytes. *J Lipid Res* 35: 248-254.
71. Collins QF, Xiong Y, Lupo EG, Jr., Liu HY, Cao W (2006) p38 Mitogen-activated protein kinase mediates free fatty acid-induced gluconeogenesis in hepatocytes. *J Biol Chem* 281: 24336-24344.
72. Seo T, Oelkers PM, Giattina MR, Worgall TS, Sturley SL, et al. (2001) Differential modulation of ACAT1 and ACAT2 transcription and activity by long chain free fatty acids in cultured cells. *Biochemistry* 40: 4756-4762.
73. Jousse C, Bruhat A, Ferrara M, Fafournoux P (2000) Evidence for multiple signaling pathways in the regulation of gene expression by amino acids in human cell lines. *J Nutr* 130: 1555-1560.
74. Kimball SR, Jefferson LS (2006) New functions for amino acids: effects on gene transcription and translation. *Am J Clin Nutr* 83: 500S-507S.
75. Matsumura T, Morinaga Y, Fujitani S, Takehana K, Nishitani S, et al. (2005) Oral administration of branched-chain amino acids activates the mTOR signal in cirrhotic rat liver. *Hepatol Res* 33: 27-32.

76. Curi R, Lagranha CJ, Doi SQ, Sellitti DF, Procopio J, et al. (2005) Glutamine-dependent changes in gene expression and protein activity. *Cell Biochem Funct* 23: 77-84.
77. Li Z, Yarmush ML, Chan C (2004) Insulin concentration during preconditioning mediates the regulation of urea synthesis during exposure to amino acid-supplemented plasma. *Tissue Eng* 10: 1737-1746.
78. Lavoigne A, Husson A, Quillard M, Chedeville A, Fairand A (1996) Glutamine inhibits the lowering effect of glucose on the level of phosphoenolpyruvate carboxykinase mRNA in isolated rat hepatocytes. *Eur J Biochem* 242: 537-543.
79. Gerich JE (1998) The genetic basis of type 2 diabetes mellitus: impaired insulin secretion versus impaired insulin sensitivity. *Endocr Rev* 19: 491-503.
80. Radha V, Vimalaswaran KS, Deepa R, Mohan V (2003) The genetics of diabetes mellitus. *Indian J Med Res* 117: 225-238.
81. Perry JR, Frayling TM (2008) New gene variants alter type 2 diabetes risk predominantly through reduced beta-cell function. *Curr Opin Clin Nutr Metab Care* 11: 371-377.
82. Pascoe L, Tura A, Patel SK, Ibrahim IM, Ferrannini E, et al. (2007) Common variants of the novel type 2 diabetes genes CDKAL1 and HHEX/IDE are associated with decreased pancreatic beta-cell function. *Diabetes* 56: 3101-3104.
83. Florez JC (2007) The new type 2 diabetes gene TCF7L2. *Curr Opin Clin Nutr Metab Care* 10: 391-396.
84. Florez JC (2008) Newly identified loci highlight beta cell dysfunction as a key cause of type 2 diabetes: Where are the insulin resistance genes? *Diabetologia* 51: 1100-1110.
85. Froguel P, Velho G, Passa P, Cohen D (1993) Genetic determinants of type 2 diabetes mellitus: lessons learned from family studies. *Diabete Metab* 19: 1-10.
86. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, et al. (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316: 889-894.
87. Mitchell SM, Weedon MN, Owen KR, Shields B, Wilkins-Wall B, et al. (2003) Genetic variation in the small heterodimer partner gene and young-onset type 2 diabetes, obesity, and birth weight in U.K. subjects. *Diabetes* 52: 1276-1279.

88. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, et al. (2008) Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet.*
89. Kimber CH, Doney AS, Pearson ER, McCarthy MI, Hattersley AT, et al. (2007) TCF7L2 in the Go-DARTS study: evidence for a gene dose effect on both diabetes susceptibility and control of glucose levels. *Diabetologia* 50: 1186-1191.
90. Grarup N, Andersen G (2007) Gene-environment interactions in the pathogenesis of type 2 diabetes and metabolism. *Curr Opin Clin Nutr Metab Care* 10: 420-426.
91. Nevins JR, Huang ES, Dressman H, Pittman J, Huang AT, et al. (2003) Towards integrated clinico-genomic models for personalized medicine: combining gene expression signatures and clinical factors in breast cancer outcomes prediction. *Hum Mol Genet* 12 Spec No 2: R153-157.
92. Huang ES, Nevins JR, West M, Kuo PC (2004) An overview of genomic data analysis. *Surgery* 136: 497-499.
93. Mauricio R (2001) Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat Rev Genet* 2: 370-381.
94. Stoehr JP, Byers JE, Clee SM, Lan H, Boronenkov IV, et al. (2004) Identification of major quantitative trait loci controlling body weight variation in ob/ob mice. *Diabetes* 53: 245-249.
95. Lan H, Chen M, Flowers JB, Yandell BS, Stapleton DS, et al. (2006) Combined expression trait correlations and expression quantitative trait locus mapping. *PLoS Genet* 2: e6.
96. Stoehr JP, Nadler ST, Schueler KL, Rabaglia ME, Yandell BS, et al. (2000) Genetic obesity unmasks nonlinear interactions between murine type 2 diabetes susceptibility loci. *Diabetes* 49: 1946-1954.
97. Drake TA, Schadt EE, Lusis AJ (2006) Integrating genetic and gene expression data: application to cardiovascular and metabolic traits in mice. *Mamm Genome* 17: 466-479.
98. Cervino AC, Li G, Edwards S, Zhu J, Laurie C, et al. (2005) Integrating QTL and high-density SNP analyses in mice to identify *Insig2* as a susceptibility gene for plasma cholesterol levels. *Genomics* 86: 505-517.
99. Xu X, Wang L, Ding D (2004) Learning module networks from genome-wide location and expression data. *FEBS Lett* 578: 297-304.

100. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, et al. (2006) Genomic and transcriptional aberrations linked to breast cancer pathophysiology. *Cancer Cell* 10: 529-541.
101. Fiehn O (2002) Metabolomics--the link between genotypes and phenotypes. *Plant Mol Biol* 48: 155-171.
102. Griffin JL (2006) Understanding mouse models of disease through metabolomics. *Curr Opin Chem Biol* 10: 309-315.
103. Svati H, Shah ERH, James R. Bain, Michael J. Muehlbauer, Robert D. Stevens, Brett R. Wenner, Michael Sketch, Elaine Dowdy, Christopher B. Granger, Geoffrey S. Ginsburg, Christopher B. Newgard, William E. Kraus (ahead of print) High Heritability of Metabolomic Profiles in Families Burdened with Early-onset Cardiovascular Disease.
104. Ferrara CT, Wang P, Neto EC, Stevens RD, Bain JR, et al. (2008) Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. *PLoS Genet* 4: e1000034.
105. Ma L, Robinson LN, Towle HC (2006) ChREBP**Mlx* is the principal mediator of glucose-induced gene expression in the liver. *J Biol Chem* 281: 28721-28730.
106. Ma L, Tsatsos NG, Towle HC (2005) Direct role of ChREBP.*Mlx* in regulating hepatic glucose-responsive genes. *J Biol Chem* 280: 12019-12027.
107. Newgard CB, Lu D, Jensen MV, Schissler J, Boucher A, et al. (2002) Stimulus/secretion coupling factors in glucose-stimulated insulin secretion: insights gained from a multidisciplinary approach. *Diabetes* 51 Suppl 3: S389-393.
108. Ronnebaum SM, Ilkayeva O, Burgess SC, Joseph JW, Lu D, et al. (2006) A pyruvate cycling pathway involving cytosolic NADP-dependent isocitrate dehydrogenase regulates glucose-stimulated insulin secretion. *J Biol Chem* 281: 30593-30602.
109. Las G, Mayorek N, Dickstein K, Bar-Tana J (2006) Modulation of insulin secretion by fatty acyl analogs. *Diabetes* 55: 3478-3485.
110. Olofsson CS, Collins S, Bengtsson M, Eliasson L, Salehi A, et al. (2007) Long-term exposure to glucose and lipids inhibits glucose-induced insulin secretion downstream of granule fusion with plasma membrane. *Diabetes* 56: 1888-1897.

111. Mingrone G, Castagneto M (2006) Medium-chain, even-numbered dicarboxylic acids as novel energy substrates: an update. *Nutr Rev* 64: 449-456.
112. Westin MA, Alexson SE, Hunt MC (2004) Molecular cloning and characterization of two mouse peroxisome proliferator-activated receptor alpha (PPARalpha)-regulated peroxisomal acyl-CoA thioesterases. *J Biol Chem* 279: 21841-21848.
113. Westin MA, Hunt MC, Alexson SE (2005) The identification of a succinyl-CoA thioesterase suggests a novel pathway for succinate production in peroxisomes. *J Biol Chem* 280: 38125-38132.
114. Westin MA, Hunt MC, Alexson SE (2007) Peroxisomes contain a specific phytanoyl-CoA/pristanoyl-CoA thioesterase acting as a novel auxiliary enzyme in alpha- and beta-oxidation of methyl-branched fatty acids in mouse. *J Biol Chem* 282: 26707-26716.
115. Westin MA, Hunt MC, Alexson SE (2008) Short- and medium-chain carnitine acyltransferases and acyl-CoA thioesterases in mouse provide complementary systems for transport of beta-oxidation products out of peroxisomes. *Cell Mol Life Sci*.
116. Niot I, Gresti J, Boichot J, Sempore G, Durand G, et al. (1994) Effect of dietary n-3 and n-6 polyunsaturated fatty acids on lipid-metabolizing enzymes in obese rat liver. *Lipids* 29: 481-489.
117. Ferdinandusse S, Denis S, Van Roermund CW, Wanders RJ, Dacremont G (2004) Identification of the peroxisomal beta-oxidation enzymes involved in the degradation of long-chain dicarboxylic acids. *J Lipid Res* 45: 1104-1111.
118. Hardwick JP (2008) Cytochrome P450 omega hydroxylase (CYP4) function in fatty acid metabolism and metabolic diseases. *Biochem Pharmacol* 75: 2263-2275.
119. Reddy JK, Rao MS (2006) Lipid metabolism and liver inflammation. II. Fatty liver disease and fatty acid oxidation. *Am J Physiol Gastrointest Liver Physiol* 290: G852-858.
120. Reddy JK, Hashimoto T (2001) Peroxisomal beta-oxidation and peroxisome proliferator-activated receptor alpha: an adaptive metabolic system. *Annu Rev Nutr* 21: 193-230.
121. Engels W, van Bilsen M, Wolffenbuttel BH, van der Vusse GJ, Glatz JF (1999) Cytochrome P450, peroxisome proliferation, and cytoplasmic fatty acid-binding

- protein content in liver, heart and kidney of the diabetic rat. *Mol Cell Biochem* 192: 53-61.
122. Montague W, Taylor KW (1968) Regulation of insulin secretion by short chain fatty acids. *Nature* 217: 853.
123. Kaikaus RM, Sui Z, Lysenko N, Wu NY, Ortiz de Montellano PR, et al. (1993) Regulation of pathways of extramitochondrial fatty acid oxidation and liver fatty acid-binding protein by long-chain monocarboxylic fatty acids in hepatocytes. Effect of inhibition of carnitine palmitoyltransferase I. *J Biol Chem* 268: 26866-26871.
124. Malaisse WJ, Greco AV, Mingrone G (2000) Effects of aliphatic dioic acids and glycerol-1,2,3-tris(dodecanedioate) on D-glucose-stimulated insulin release in rat pancreatic islets. *Br J Nutr* 84: 733-736.
125. Boucher A, Lu D, Burgess SC, Telemaque-Potts S, Jensen MV, et al. (2004) Biochemical mechanism of lipid-induced impairment of glucose-stimulated insulin secretion and reversal with a malate analogue. *J Biol Chem* 279: 27263-27271.
126. Farfari S, Schulz V, Corkey B, Prentki M (2000) Glucose-regulated anaplerosis and cataplerosis in pancreatic beta-cells: possible implication of a pyruvate/citrate shuttle in insulin secretion. *Diabetes* 49: 718-726.
127. Gunawardana SC, Liu YJ, Macdonald MJ, Straub SG, Sharp GW (2004) Anaplerotic input is sufficient to induce time-dependent potentiation of insulin release in rat pancreatic islets. *Am J Physiol Endocrinol Metab* 287: E828-833.
128. MacDonald MJ (2002) Differences between mouse and rat pancreatic islets: succinate responsiveness, malic enzyme, and anaplerosis. *Am J Physiol Endocrinol Metab* 283: E302-310.
129. Macdonald MJ (2003) Export of metabolites from pancreatic islet mitochondria as a means to study anaplerosis in insulin secretion. *Metabolism* 52: 993-998.
130. Prentki M (1996) New insights into pancreatic beta-cell metabolic signaling in insulin secretion. *Eur J Endocrinol* 134: 272-286.
131. Schuit F, De Vos A, Farfari S, Moens K, Pipeleers D, et al. (1997) Metabolic fate of glucose in purified islet cells. Glucose-regulated anaplerosis in beta cells. *J Biol Chem* 272: 18572-18579.

132. Jensen MV, Joseph JW, Ilkayeva O, Burgess S, Lu D, et al. (2006) Compensatory responses to pyruvate carboxylase suppression in islet beta-cells. Preservation of glucose-stimulated insulin secretion. *J Biol Chem* 281: 22342-22351.
133. Lu D, Mulder H, Zhao P, Burgess SC, Jensen MV, et al. (2002) ¹³C NMR isotopomer analysis reveals a connection between pyruvate cycling and glucose-stimulated insulin secretion (GSIS). *Proc Natl Acad Sci U S A* 99: 2708-2713.
134. Brem RB, Yvert G, Clinton R, Kruglyak L (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science* 296: 752-755.
135. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, et al. (2005) An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* 37: 710-717.
136. Kendzioriski C, Wang P (2006) A review of statistical methods for expression quantitative trait loci mapping. *Mamm Genome* 17: 509-517.
137. Bing N, Hoeschele I (2005) Genetical genomics analysis of a yeast segregant population for transcription network inference. *Genetics* 170: 533-542.
138. Jansen RC, Nap JP (2001) Genetical genomics: the added value from segregation. *Trends Genet* 17: 388-391.
139. Keurentjes JJ, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, et al. (2007) Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci U S A* 104: 1708-1713.
140. Zhu J, Lum PY, Lamb J, GuhaThakurta D, Edwards SW, et al. (2004) An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet Genome Res* 105: 363-374.
141. Kulp DC, Jagalur M (2006) Causal inference of regulator-target pairs by gene mapping of expression phenotypes. *BMC Genomics* 7: 125.
142. Li R, Tsaih SW, Shockley K, Stylianou IM, Wergedal J, et al. (2006) Structural model analysis of multiple quantitative traits. *PLoS Genet* 2: e114.
143. de la Fuente A, Bing N, Hoeschele I, Mendes P (2004) Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics* 20: 3565-3574.

144. Shipley B (2000) *Cause and Correlation in Biology*. New York: Cambridge University Press.
145. Schafer J, Strimmer K (2005) An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics* 21: 754-764.
146. Spirtes P, Glymour C, Scheines R (1993) *Causation, Prediction, and Search*: MIT Press.
147. Chaibub Neto E, Ferrara CT, Attie AD, Yandell BS (2008) Inferring Causal Phenotype Networks from Segregating Populations. *Genetics*.
148. Schadt EE, Lum PY (2006) Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Reverse engineering gene networks to identify key drivers of complex disease phenotypes. *J Lipid Res* 47: 2601-2613.
149. Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4: Article17.
150. Li H, Chen H, Bao L, Manly KF, Chesler EJ, et al. (2006) Integrative genetic analysis of transcription modules: towards filling the gap between genetic loci and inherited traits. *Hum Mol Genet* 15: 481-492.
151. Yi N, Yandell BS, Churchill GA, Allison DB, Eisen EJ, et al. (2005) Bayesian model selection for genome-wide epistatic quantitative trait loci analysis. *Genetics* 170: 1333-1344.
152. Friedman N, Linial M, Nachman I, Pe'er D (2000) Using Bayesian networks to analyze expression data. *J Comput Biol* 7: 601-620.
153. DiPetrillo K, Wang X, Stylianou IM, Paigen B (2005) Bioinformatics toolbox for narrowing rodent quantitative trait loci. *Trends Genet* 21: 683-692.
154. Glazier AM, Nadeau JH, Aitman TJ (2002) Finding genes that underlie complex traits. *Science* 298: 2345-2349.
155. Guo SW, Lange K (2000) Genetic mapping of complex traits: promises, problems, and prospects. *Theor Popul Biol* 57: 1-11.
156. Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6: 95-108.

157. Nadeau JH, Frankel WN (2000) The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. *Nat Genet* 25: 381-384.
158. Petretto E, Liu ET, Aitman TJ (2007) A gene harvest revealing the archeology and complexity of human disease. *Nat Genet* 39: 1299-1301.
159. Wentzell AM, Rowe HC, Hansen BG, Ticconi C, Halkier BA, et al. (2007) Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways. *PLoS Genet* 3: 1687-1701.
160. Kliebenstein DJ, West MA, van Leeuwen H, Loudet O, Doerge RW, et al. (2006) Identification of QTLs controlling gene expression networks defined a priori. *BMC Bioinformatics* 7: 308.
161. Graber R, Sumida C, Nunez EA (1994) Fatty acids and cell signal transduction. *J Lipid Mediat Cell Signal* 9: 91-116.
162. Quest AF, Ghosh S, Xie WQ, Bell RM (1997) DAG second messengers: molecular switches and growth control. *Adv Exp Med Biol* 400A: 297-303.
163. Wang H, Chen J, Hollister K, Sowers LC, Forman BM (1999) Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Mol Cell* 3: 543-553.
164. An J, Muoio DM, Shiota M, Fujimoto Y, Cline GW, et al. (2004) Hepatic expression of malonyl-CoA decarboxylase reverses muscle, liver and whole-animal insulin resistance. *Nat Med* 10: 268-274.
165. Koves TR, Li P, An J, Akimoto T, Slentz D, et al. (2005) Peroxisome proliferator-activated receptor-gamma co-activator 1alpha-mediated metabolic remodeling of skeletal myocytes mimics exercise training and reverses lipid-induced mitochondrial inefficiency. *J Biol Chem* 280: 33588-33598.
166. Monetti M, Levin MC, Watt MJ, Sajan MP, Marmor S, et al. (2007) Dissociation of hepatic steatosis and insulin resistance in mice overexpressing DGAT in the liver. *Cell Metab* 6: 69-78.
167. Deutsch S, Lyle R, Dermitzakis ET, Attar H, Subrahmanyam L, et al. (2005) Gene expression variation and expression quantitative trait mapping of human chromosome 21 genes. *Hum Mol Genet* 14: 3741-3749.
168. Kendzioriski CM, Chen M, Yuan M, Lan H, Attie AD (2006) Statistical methods for expression quantitative trait loci (eQTL) mapping. *Biometrics* 62: 19-27.

169. Lum PY, Chen Y, Zhu J, Lamb J, Melmed S, et al. (2006) Elucidating the murine brain transcriptional network in a segregating mouse population to identify core functional modules for obesity and diabetes. *J Neurochem*.
170. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249-264.
171. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, et al. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1: 174-181.
172. Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19: 889-890.
173. Shipley B (2000) *Cause and Correlation in Biology*. New York: Cambridge University Press.
174. Princen HM, Huijsmans CM, Kuipers F, Vonk RJ, Kempen HJ (1986) Ketoconazole blocks bile acid synthesis in hepatocyte monolayer cultures and in vivo in rat by inhibiting cholesterol 7 alpha-hydroxylase. *J Clin Invest* 78: 1064-1071.
175. Bao L, Wei L, Peirce JL, Homayouni R, Li H, et al. (2006) Combining gene expression QTL mapping and phenotypic spectrum analysis to uncover gene regulatory relationships. *Mamm Genome* 17: 575-583.
176. Chesler EJ, Lu L, Shou S, Qu Y, Gu J, et al. (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat Genet* 37: 233-242.
177. Schauer N, Semel Y, Roessner U, Gur A, Balbo I, et al. (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat Biotechnol* 24: 447-454.
178. Carrari F, Baxter C, Usadel B, Urbanczyk-Wochniak E, Zanon MI, et al. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol* 142: 1380-1396.
179. Gu S, Langlais P, Liu F, Jiang JX (2003) Mouse system-N amino acid transporter, mNAT3, expressed in hepatocytes and regulated by insulin-activated and phosphoinositide 3-kinase-dependent signalling. *Biochem J* 371: 721-731.

180. Gu S, Villegas CJ, Jiang JX (2005) Differential regulation of amino acid transporter SNAT3 by insulin in hepatocytes. *J Biol Chem* 280: 26055-26062.
181. Steuer R (2006) Review: on the analysis and interpretation of correlations in metabolomic data. *Brief Bioinform* 7: 151-158.
182. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95: 14863-14868.
183. Manichaikul A, Dupuis J, Sen S, Broman KW (2006) Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174: 481-489.
184. Lan MS, Wang HW, Chong J, Breslin MB (2007) Coupling of glucose response element from L-type pyruvate kinase and G6Pase promoter enhances glucose responsive activity in hepatoma cells. *Mol Cell Biochem* 300: 191-196.
185. Hall R, Beale M, Fiehn O, Hardy N, Sumner L, et al. (2002) Plant metabolomics: the missing link in functional genomics strategies. *Plant Cell* 14: 1437-1440.
186. Desjardins P, Belanger M, Butterworth RF (2001) Alterations in expression of genes coding for key astrocytic proteins in acute liver failure. *J Neurosci Res* 66: 967-971.
187. Hazell AS, Butterworth RF (1999) Hepatic encephalopathy: An update of pathophysiologic mechanisms. *Proc Soc Exp Biol Med* 222: 99-112.
188. Ye ZC, Sontheimer H (2002) Modulation of glial glutamate transport through cell interactions with the extracellular matrix. *Int J Dev Neurosci* 20: 209-217.
189. Hankard RG, Haymond MW, Darmaun D (1996) Effect of glutamine on leucine metabolism in humans. *Am J Physiol* 271: E748-754.
190. Le Bacquer O, Mauras N, Welch S, Haymond M, Darmaun D (2007) Acute depletion of plasma glutamine increases leucine oxidation in prednisone-treated humans. *Clin Nutr* 26: 231-238.
191. Zhou BG, Norenberg MD (1999) Ammonia downregulates GLAST mRNA glutamate transporter in rat astrocyte cultures. *Neurosci Lett* 276: 145-148.
192. Albrecht J, Jones EA (1999) Hepatic encephalopathy: molecular mechanisms underlying the clinical syndrome. *J Neurol Sci* 170: 138-146.

193. Butterworth RF (2001) Neurotransmitter dysfunction in hepatic encephalopathy: new approaches and new findings. *Metab Brain Dis* 16: 55-65.
194. Butterworth RF (2002) Glutamate transporters in hyperammonemia. *Neurochem Int* 41: 81-85.
195. Jump DB, Thelen A, Ren B, Mater M (1999) Multiple mechanisms for polyunsaturated fatty acid regulation of hepatic gene transcription. *Prostaglandins Leukot Essent Fatty Acids* 60: 345-349.
196. Carlsson L, Linden D, Jalouli M, Oscarsson J (2001) Effects of fatty acids and growth hormone on liver fatty acid binding protein and PPARalpha in rat liver. *Am J Physiol Endocrinol Metab* 281: E772-781.
197. Gnerre C, Schuster GU, Roth A, Handschin C, Johansson L, et al. (2005) LXR deficiency and cholesterol feeding affect the expression and phenobarbital-mediated induction of cytochromes P450 in mouse liver. *J Lipid Res* 46: 1633-1642.
198. Ory DS (2004) Nuclear receptor signaling in the control of cholesterol homeostasis: have the orphans found a home? *Circ Res* 95: 660-670.
199. Sanguino E, Roglans N, Alegret M, Sanchez RM, Vazquez-Carrera M, et al. (2005) Atorvastatin reverses age-related reduction in rat hepatic PPARalpha and HNF-4. *Br J Pharmacol* 145: 853-861.
200. Claeysens S, Gangneux C, Brasse-Lagnel C, Ruminy P, Aki T, et al. (2003) Amino acid control of the human glyceraldehyde 3-phosphate dehydrogenase gene transcription in hepatocyte. *Am J Physiol Gastrointest Liver Physiol* 285: G840-849.
201. Jefferson LS, Kimball SR (2003) Amino acids as regulators of gene expression at the level of mRNA translation. *J Nutr* 133: 2046S-2051S.
202. Proud CG (2004) mTOR-mediated regulation of translation factors by amino acids. *Biochem Biophys Res Commun* 313: 429-436.
203. Nair KS, Short KR (2005) Hormonal and signaling role of branched-chain amino acids. *J Nutr* 135: 1547S-1552S.
204. Choi JH, Park MJ, Kim KW, Choi YH, Park SH, et al. (2005) Molecular mechanism of hypoxia-mediated hepatic gluconeogenesis by transcriptional regulation. *FEBS Lett* 579: 2795-2801.

205. Kimball SR, Jefferson LS (2004) Amino acids as regulators of gene expression. *Nutr Metab (Lond)* 1: 3.
206. Bruhat A, Jousse C, Fafournoux P (1999) Amino acid limitation regulates gene expression. *Proc Nutr Soc* 58: 625-632.
207. Jousse C, Bruhat A, Fafournoux P (1999) Amino acid regulation of gene expression. *Curr Opin Clin Nutr Metab Care* 2: 297-301.
208. Jousse C, Bruhat A, Harding HP, Ferrara M, Ron D, et al. (1999) Amino acid limitation regulates CHOP expression through a specific pathway independent of the unfolded protein response. *FEBS Lett* 448: 211-216.
209. Jousse C, Bruhat A, Ferrara M, Fafournoux P (1998) Physiological concentration of amino acids regulates insulin-like-growth-factor-binding protein 1 expression. *Biochem J* 334 (Pt 1): 147-153.
210. Thiaville MM, Dudenhausen EE, Zhong C, Pan YX, Kilberg MS (2008) Deprivation of protein or amino acid induces C/EBPbeta synthesis and binding to amino acid response elements, but its action is not an absolute requirement for enhanced transcription. *Biochem J* 410: 473-484.
211. Bruhat A, Fafournoux P (2001) Recent advances on molecular mechanisms involved in amino acid control of gene expression. *Curr Opin Clin Nutr Metab Care* 4: 439-443.
212. Bruhat A, Jousse C, Carraro V, Reimold AM, Ferrara M, et al. (2000) Amino acids control mammalian gene transcription: activating transcription factor 2 is essential for the amino acid responsiveness of the CHOP promoter. *Mol Cell Biol* 20: 7192-7204.
213. Averous J, Bruhat A, Jousse C, Carraro V, Thiel G, et al. (2004) Induction of CHOP expression by amino acid limitation requires both ATF4 expression and ATF2 phosphorylation. *J Biol Chem* 279: 5288-5297.
214. Averous J, Bruhat A, Mordier S, Fafournoux P (2003) Recent advances in the understanding of amino acid regulation of gene expression. *J Nutr* 133: 2040S-2045S.
215. Jousse C, Averous J, Bruhat A, Carraro V, Mordier S, et al. (2004) Amino acids as regulators of gene expression: molecular mechanisms. *Biochem Biophys Res Commun* 313: 447-452.

216. Keller MP, Choi Y, Wang P, Davis DB, Rabaglia ME, et al. (2008) A Gene Expression Network Model of Type 2 Diabetes Links Cell Cycle Regulation in Islets with Diabetes Susceptibility. *Genome Res.*
217. He YD, Dai H, Schadt EE, Cavet G, Edwards SW, et al. (2003) Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. *Bioinformatics* 19: 956-965.
218. Wu JY, Kao HJ, Li SC, Stevens R, Hillman S, et al. (2004) ENU mutagenesis identifies mice with mitochondrial branched-chain aminotransferase deficiency resembling human maple syrup urine disease. *J Clin Invest* 113: 434-440.
219. Newton MA, Kendzierski CM, Richmond CS, Blattner FR, Tsui KW (2001) On differential variability of expression ratios: improving statistical inference about gene expression changes from microarray data. *J Comput Biol* 8: 37-52.
220. Zaleski J, Wilson DF, Erecinska M (1986) beta-2-Aminobicyclo-(2.2.1)-heptane-2-carboxylic acid. A new activator of glutaminase in intact rat liver mitochondria. *J Biol Chem* 261: 14091-14094.
221. Zaleski J, Wilson DF, Erecinska M (1986) Glutamine metabolism in rat hepatocytes. Stimulation by a nonmetabolizable analog of leucine. *J Biol Chem* 261: 14082-14090.
222. Lavoigne A, Meisse D, Quillard M, Husson A, Renouf S, et al. (1998) Glutamine and regulation of gene expression in rat hepatocytes: the role of cell swelling. *Biochimie* 80: 807-811.
223. Quillard M, Renouf S, Husson A, Meisse D, Lavoigne A (1997) Glutamine and regulation of gene expression in mammalian cells. Special reference to phosphoenolpyruvate carboxykinase (PEPCK). *Biochimie* 79: 125-128.
224. Consoli A, Nurjhan N, Capani F, Gerich J (1989) Predominant role of gluconeogenesis in increased hepatic glucose production in NIDDM. *Diabetes* 38: 550-557.
225. Tappy L (1995) Regulation of hepatic glucose production in healthy subjects and patients with non-insulin-dependent diabetes mellitus. *Diabete Metab* 21: 233-240.
226. Halter JB, Ward WK, Porte D, Jr., Best JD, Pfeifer MA (1985) Glucose regulation in non-insulin-dependent diabetes mellitus. Interaction between pancreatic islets and the liver. *Am J Med* 79: 6-12.

227. Wiernsperger NF (1999) Membrane physiology as a basis for the cellular effects of metformin in insulin resistance and diabetes. *Diabetes Metab* 25: 110-127.
228. Consoli A (1992) Role of liver in pathophysiology of NIDDM. *Diabetes Care* 15: 430-441.
229. Haussinger D, Soboll S, Meijer AJ, Gerok W, Tager JM, et al. (1985) Role of plasma membrane transport in hepatic glutamine metabolism. *Eur J Biochem* 152: 597-603.
230. Martin M, Ferrier B, Baverel G (1989) Transport and utilization of alpha-ketoglutarate by the rat kidney in vivo. *Pflugers Arch* 413: 217-224.
231. Titheradge MA, Picking RA, Haynes RC, Jr. (1992) Physiological concentrations of 2-oxoglutarate regulate the activity of phosphoenolpyruvate carboxykinase in liver. *Biochem J* 285 (Pt 3): 767-771.
232. Rabaglia ME, Gray-Keller MP, Frey BL, Shortreed MR, Smith LM, et al. (2005) Alpha-Ketoisocaproate-induced hypersecretion of insulin by islets from diabetes-susceptible mice. *Am J Physiol Endocrinol Metab* 289: E218-224.
233. Kahn SE, Porte D, Jr. (1988) Islet dysfunction in non-insulin-dependent diabetes mellitus. *Am J Med* 85: 4-8.
234. Petruzzelli M, Lo Sasso G, Portincasa P, Palasciano G, Moschetta A (2007) Targeting the liver in the metabolic syndrome: evidence from animal models. *Curr Pharm Des* 13: 2199-2207.
235. Arner P (2002) Insulin resistance in type 2 diabetes: role of fatty acids. *Diabetes Metab Res Rev* 18 Suppl 2: S5-9.
236. Opara EC, Hubbard VS, Burch WM, Akwari OE (1992) Characterization of the insulinotropic potency of polyunsaturated fatty acids. *Endocrinology* 130: 657-662.
237. Kashyap S, Belfort R, Gastaldelli A, Pratipanawatr T, Berria R, et al. (2003) A sustained increase in plasma free fatty acids impairs insulin secretion in nondiabetic subjects genetically predisposed to develop type 2 diabetes. *Diabetes* 52: 2461-2474.
238. Komatsu M, Yajima H, Yamada S, Kaneko T, Sato Y, et al. (1999) Augmentation of Ca²⁺-stimulated insulin release by glucose and long-chain fatty acids in rat

pancreatic islets: free fatty acids mimic ATP-sensitive K⁺ channel-independent insulinotropic action of glucose. *Diabetes* 48: 1543-1549.

239. MacDonald MJ, Fahien LA, Brown LJ, Hasan NM, Buss JD, et al. (2005) Perspective: emerging evidence for signaling roles of mitochondrial anaplerotic products in insulin secretion. *Am J Physiol Endocrinol Metab* 288: E1-15.
240. Ronnebaum SM, Joseph JW, Ilkayeva O, Burgess SC, Lu D, et al. (2008) Chronic Suppression of Acetyl-CoA Carboxylase 1 in β -Cells Impairs Insulin Secretion via Inhibition of Glucose Rather Than Lipid Metabolism. *J Biol Chem* 283: 14248-14256.
241. Joseph JW, Odegaard ML, Ronnebaum SM, Burgess SC, Muehlbauer J, et al. (2007) Normal flux through ATP-citrate lyase or fatty acid synthase is not required for glucose-stimulated insulin secretion. *J Biol Chem* 282: 31592-31600.
242. Hohmeier HE, Mulder H, Chen G, Henkel-Rieger R, Prentki M, et al. (2000) Isolation of INS-1-derived cell lines with robust ATP-sensitive K⁺ channel-dependent and -independent glucose-stimulated insulin secretion. *Diabetes* 49: 424-430.
243. Naber SP, McDonald JM, Jarett L, McDaniel ML, Ludvigsen CW, et al. (1980) Preliminary characterization of calcium binding in islet-cell plasma membranes. *Diabetologia* 19: 439-444.
244. Milburn JL, Jr., Hirose H, Lee YH, Nagasawa Y, Ogawa A, et al. (1995) Pancreatic beta-cells in obesity. Evidence for induction of functional, morphologic, and metabolic abnormalities by increased long chain fatty acids. *J Biol Chem* 270: 1295-1299.
245. Taylor JT, Huang L, Keyser BM, Zhuang H, Clarkson CW, et al. (2005) Role of high-voltage-activated calcium channels in glucose-regulated beta-cell calcium homeostasis and insulin release. *Am J Physiol Endocrinol Metab* 289: E900-908.
246. Liang Y, Matschinsky FM (1994) Mechanisms of action of nonglucose insulin secretagogues. *Annu Rev Nutr* 14: 59-81.
247. Takasawa S, Nata K, Yonekura H, Okamoto H (1993) Cyclic ADP-ribose in insulin secretion from pancreatic beta cells. *Science* 259: 370-373.
248. Takasawa S, Ishida A, Nata K, Nakagawa K, Noguchi N, et al. (1995) Requirement of calmodulin-dependent protein kinase II in cyclic ADP-ribose-mediated intracellular Ca²⁺ mobilization. *J Biol Chem* 270: 30257-30259.

249. Okamoto H (1999) The CD38-cyclic ADP-ribose signaling system in insulin secretion. *Mol Cell Biochem* 193: 115-118.
250. Johnson JD, Mislser S (2002) Nicotinic acid-adenine dinucleotide phosphate-sensitive calcium stores initiate insulin signaling in human beta cells. *Proc Natl Acad Sci U S A* 99: 14566-14571.
251. Mitchell KJ, Lai FA, Rutter GA (2003) Ryanodine receptor type I and nicotinic acid adenine dinucleotide phosphate receptors mediate Ca²⁺ release from insulin-containing vesicles in living pancreatic beta-cells (MIN6). *J Biol Chem* 278: 11057-11064.
252. Masgrau R, Churchill GC, Morgan AJ, Ashcroft SJ, Galione A (2003) NAADP: a new second messenger for glucose-induced Ca²⁺ responses in clonal pancreatic beta cells. *Curr Biol* 13: 247-251.
253. Patel S (2003) NAADP on the up in pancreatic beta cells-a sweet message? *Bioessays* 25: 430-433.
254. Heart E, Yaney GC, Corkey RF, Schultz V, Luc E, et al. (2007) Ca²⁺, NAD(P)H and membrane potential changes in pancreatic beta-cells by methyl succinate: comparison with glucose. *Biochem J* 403: 197-205.
255. Yoshinari K, Takagi S, Sugatani J, Miwa M (2006) Changes in the expression of cytochromes P450 and nuclear receptors in the liver of genetically diabetic db/db mice. *Biol Pharm Bull* 29: 1634-1638.
256. Yaney GC, Corkey BE (2003) Fatty acid metabolism and insulin secretion in pancreatic beta cells. *Diabetologia* 46: 1297-1312.
257. Guay C, Madiraju SR, Aumais A, Joly E, Prentki M (2007) A role for ATP-citrate lyase, malic enzyme, and pyruvate/citrate cycling in glucose-induced insulin secretion. *J Biol Chem* 282: 35657-35665.
258. Maechler P, Wollheim CB (1999) Mitochondrial glutamate acts as a messenger in glucose-induced insulin exocytosis. *Nature* 402: 685-689.
259. Evans AM, Wyatt CN, Kinnear NP, Clark JH, Blanco EA (2005) Pyridine nucleotides and calcium signalling in arterial smooth muscle: from cell physiology to pharmacology. *Pharmacol Ther* 107: 286-313.
260. Mitchell KJ, Pinton P, Varadi A, Tacchetti C, Ainscow EK, et al. (2001) Dense core secretory vesicles revealed as a dynamic Ca⁽²⁺⁾ store in neuroendocrine cells

with a vesicle-associated membrane protein aequorin chimaera. *J Cell Biol* 155: 41-51.

261. Lemmens R, Larsson O, Berggren PO, Islam MS (2001) Ca²⁺-induced Ca²⁺ release from the endoplasmic reticulum amplifies the Ca²⁺ signal mediated by activation of voltage-gated L-type Ca²⁺ channels in pancreatic beta-cells. *J Biol Chem* 276: 9971-9977.
262. Poitout V, Robertson RP (2008) Glucolipotoxicity: fuel excess and beta-cell dysfunction. *Endocr Rev* 29: 351-366.
263. Kohnke R, Mei J, Park M, York DA, Erlanson-Albertsson C (2007) Fatty acids and glucose in high concentration down-regulates ATP synthase beta-subunit protein expression in INS-1 cells. *Nutr Neurosci* 10: 273-278.
264. Moneva MH, Dagogo-Jack S (2002) Multiple drug targets in the management of type 2 diabetes. *Curr Drug Targets* 3: 203-221.
265. Joron GE, Laryea E, Jaeger D, Macdonald L (1986) Cause of death in 1144 patients with diabetes mellitus: an autopsy study. *Cmaj* 134: 759-764.
266. Turner RC, Cull CA, Frighi V, Holman RR (1999) Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus: progressive requirement for multiple therapies (UKPDS 49). UK Prospective Diabetes Study (UKPDS) Group. *Jama* 281: 2005-2012.
267. Vandekoppel S, Choe HM, Sweet BV (2008) Managed care perspective on three new agents for type 2 diabetes. *J Manag Care Pharm* 14: 363-380.
268. Fonseca VA, Kulkarni KD (2008) Management of type 2 diabetes: oral agents, insulin, and injectables. *J Am Diet Assoc* 108: S29-33.
269. McDonnell ME (2007) Combination therapy with new targets in Type 2 diabetes: a review of available agents with a focus on pre-exercise adjustment. *J Cardiopulm Rehabil Prev* 27: 193-201.
270. Kahn SE, Haffner SM, Heise MA, Herman WH, Holman RR, et al. (2006) Glycemic durability of rosiglitazone, metformin, or glyburide monotherapy. *N Engl J Med* 355: 2427-2443.
271. Rondinone CM (2005) Diabetes: the latest developments in inhibitors, insulin sensitisers, new drug targets and novel approaches. October 18-19, 2004, The Hatton, London, UK. *Expert Opin Ther Targets* 9: 415-418.

272. Nissen SE (2007) Perspective: effect of rosiglitazone on cardiovascular outcomes. *Curr Cardiol Rep* 9: 343-344.
273. Vasudevan AR, Burns A, Fonseca VA (2006) The effectiveness of intensive glycemic control for the prevention of vascular complications in diabetes mellitus. *Treat Endocrinol* 5: 273-286.
274. Valensi P, Paries J, Attali JR (2003) Cardiac autonomic neuropathy in diabetic patients: influence of diabetes duration, obesity, and microangiopathic complications--the French multicenter study. *Metabolism* 52: 815-820.
275. Radziuk J, Bailey CJ, Wiernsperger NF, Yudkin JS (2003) Metformin and its liver targets in the treatment of type 2 diabetes. *Curr Drug Targets Immune Endocr Metabol Disord* 3: 151-169.
276. Massi-Benedetti M (2006) Changing targets in the treatment of type 2 diabetes. *Curr Med Res Opin* 22 Suppl 2: S5-13.
277. Rayasam GV, Tulasi VK, Davis JA, Bansal VS (2007) Fatty acid receptors as new therapeutic targets for diabetes. *Expert Opin Ther Targets* 11: 661-671.
278. Seda O, Sedova L (2007) Peroxisome proliferator-activated receptors as molecular targets in relation to obesity and type 2 diabetes. *Pharmacogenomics* 8: 587-596.
279. Agius L (2007) New hepatic targets for glycaemic control in diabetes. *Best Pract Res Clin Endocrinol Metab* 21: 587-605.
280. Kimura T, Christoffels VM, Chowdhury S, Iwase K, Matsuzaki H, et al. (1998) Hypoglycemia-associated hyperammonemia caused by impaired expression of ornithine cycle enzyme genes in C/EBPalpha knockout mice. *J Biol Chem* 273: 27505-27510.
281. Kim S, Yamazaki M, Zella LA, Meyer MB, Fretz JA, et al. (2007) Multiple enhancer regions located at significant distances upstream of the transcriptional start site mediate RANKL gene expression in response to 1,25-dihydroxyvitamin D3. *J Steroid Biochem Mol Biol* 103: 430-434.
282. Wiernsperger NF, Bailey CJ (1999) The antihyperglycaemic effect of metformin: therapeutic and cellular mechanisms. *Drugs* 58 Suppl 1: 31-39; discussion 75-82.
283. Reilly MP, Rohatgi A, McMahon K, Wolfe ML, Pinto SC, et al. (2007) Plasma cytokines, metabolic syndrome, and atherosclerosis in humans. *J Investig Med* 55: 26-35.

284. Guest CB, Park MJ, Johnson DR, Freund GG (2008) The implication of proinflammatory cytokines in type 2 diabetes. *Front Biosci* 13: 5187-5194.
285. Bastard JP, Maachi M, Lagathu C, Kim MJ, Caron M, et al. (2006) Recent advances in the relationship between obesity, inflammation, and insulin resistance. *Eur Cytokine Netw* 17: 4-12.
286. Arner P (2005) Insulin resistance in type 2 diabetes -- role of the adipokines. *Curr Mol Med* 5: 333-339.
287. Muoio DM, Newgard CB (2005) Metabolism: A is for adipokine. *Nature* 436: 337-338.
288. Xydakis AM, Case CC, Jones PH, Hoogeveen RC, Liu MY, et al. (2004) Adiponectin, inflammation, and the expression of the metabolic syndrome in obese individuals: the impact of rapid weight loss through caloric restriction. *J Clin Endocrinol Metab* 89: 2697-2703.

Biography

Born

September 20, 1980, San Diego, CA

Education

Duke University Medical Scientist Training Program; Durham, NC 2003-2009
Duke University Graduate School PhD Pharmacology 2005-2008
Advisor: Dr. Christopher Newgard
In collaboration with Dr. Alan Attie, University of Wisconsin, Madison

University of Notre Dame; South Bend, IN BS Biology '98-99,'00-03

Peer-Reviewed Publications

Genetic Networks of Liver Metabolism Revealed by Integration of Metabolic and Transcriptional Profiling.

PLoS Genet. 2008 Mar 14; 4(3):e1000034.

Ferrara CT, Wang P, Chaibub E, Stevens RD, Bain JR, Wenner BR, Ilkayeva OR, Keller MP, Blasiolo DA, Kendzioriski C, Yandell BS, Newgard CB, Attie AD.

Inferring Causal Phenotype Networks From Segregating Populations.

Genetics 2008 May 27; 179:1089-1100

Chaibub E, Ferrara CT, Attie AD, Yandell BS.

Glutamine Regulation of Liver Transcription: Mechanism and Relevance to Diabetes (in progress)

Ferrara CT, Stevens RD, Bain JR, Wenner BR, Ilkayeva OR, Attie AD, Newgard CB.

Honors and Awards

Duke University School of Medicine, Durham, NC

Alpha Omega Alpha Medical Honors Society Nominee-Spring 2007

University of Notre Dame, Notre Dame, IN

Summa Cum Laude, College of Science 2003

Dean's List- Fall 1998, 2000, 2001, 2002 & Spring 1999, 2001, 2002

Academic Honor Committee for College of Science-Fall 2001-Fall 2002

Alpha Epsilon Delta Pre-Medical Honor Society