Spring 2014

Gene expression in the kidney of rats affected by metabolic syndrome: A meta-analysis

Brittany Leigh Hatchett
James Madison University

Gene Expression in the Kidney of Rats Affected by Metabolic Syndrome: A Meta-Analysis

_____

A Project Presented to

the Faculty of the Undergraduate

College of Science and Mathematics

James Madison University

_____

in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Science

_____

by Brittany Leigh Hatchett

May 2014

Accepted by the faculty of the Department of Biology, James Madison University, in partial fulfillment of the requirements for the Degree of Bachelor of Science.

FACULTY COMMITTEE: HONORS PROGRAM APPROVAL:

Project Advisor: Terrie Rife, Ph.D.,   Barry Falk, Ph.D.,
Associate Professor, Biology   Director, Honors Program

Reader: Janet Daniel, Ph. D.,
Associate Professor, Biology

Reader: Timothy Bloss, Ph.D.,
Associate Professor, Biology

Reader: Nusrat Jahan, Ph.D.,
Associate Professor, Mathematics and Statistics

Dedication

I would like to dedicate this thesis to my brother, Devin Hatchett, who constantly motivates and

encourages me to work hard to achieve my goals, and to always do the best work possible.

Table of Contents

List of Figures

Figures

Tables

Acknowledgments

I would like to acknowledge and thank my advisor, Dr. Terrie K. Rife, for her constant support and guidance throughout this project. Without her continuing advice and encouragement, I would not have been able to complete this project. She has given me the lab experience that will be integral in my later career pursuits.

I would also like to thank one of my readers, Dr. Timothy Bloss, for dedicating his time serving on my committee. He has been an influential member in bettering and expanding this project as it progressed.

I would also like to thank my other reader, Dr. Janet Daniel, for her investment in this project and service on my committee. Her knowledge of metabolism was integral to developing this project. She has inspired my continued desire to work in the field of endocrinology through her Human Metabolism class.

I would like to thank Dr. Nusrat Jahan for her extensive help with the statistical methodology of this thesis. Her guidance and knowledge of the program R were integral in completing the statistical analysis required for this project. The UBM group of Lauren van Reesema, Keith Zirkle, and Jonah Williams were also helpful in providing an excellent example to guide my project.

Lastly, I am grateful for the support and encouragement of my friends and family for motivating me and helping me to complete this project.

5

Abstract

Metabolic syndrome is a clustering of risk factors that make a person more susceptible to cardiovascular disease and diabetes, and subsequently kidney disease. The overall metabolic health of American society is decreasing at an alarming rate. Metabolic syndrome is difficult to study due to its multi-factorial nature, which can vary from study to study. This work utilizes a meta-analysis to examine the trends in changes of gene expression that occur in rat kidneys from three different models of metabolic syndrome. Microarray studies used for this analysis were GSE4800, and GSE7193 obtained from the Gene Expression Omnibus, and E-MEXP-1695 obtained from the European Bioinformatics Institute. Following processing, 88 genes were found to be significant, including 73 upregulated genes, and 15 downregulated genes. Approximately 42.0% of the significant genes had already been associated with metabolic syndrome. Five novel genes were chosen to examine in further detail using real time PCR. These genes include downregulated genes RGD1309350, and Klk1c9 and upregulated genes Stk32c, Ampd3, and Mgmt. Preliminary qPCR results verified the trends found through meta-analysis for Klk1c9, and Mgmt, but not for the remaining three genes. Future work involves continuing to verify the gene expression with qPCR.

Background

In recent years, the overall metabolic health of the general population has been steadily

decreasing contributing to an exponential increase in disease and health care costs. This

decrease in metabolic health has been attributed to poor diets high in fat and sugar along with an

increase in sedentary lifestyles (Hong et al., 2014). It is thought that there is a genetic

component to obesity and subsequent metabolic dysfunction as well. As a result of this shift in

diet and activity level, many more people are becoming obese, and developing health

complications. Due to this increased prevalence of metabolic disease the term metabolic

syndrome (MetS) was generated to describe a clustering of risk factors for health complications

such as kidney disease, cardiovascular disease (CVD), and stroke. It is estimated that 20-25% of

the world's adult population have metabolic syndrome (Alberti, Zimmet, Shaw, & Grundy,

2006). Insulin resistance is also thought to underlie MetS, and in conjunction with other risk

factors, can significantly increase a patient's risk for type 2 diabetes mellitus (T2DM) (Grundy et

al., 2004). Patients with metabolic syndrome have a 5-fold greater risk in developing T2DM,

and up to 80% of patients with T2DM will develop CVD (Grundy, 2007). Patients with

metabolic syndrome are at a high risk for renal complications such as microalbuminuria and/or

chronic kidney disease (Locatelli, Pozzoni, & Del Vecchio, 2006). Kidney damage (as signaled

by microalbuminuria) is also an integral factor in cardiovascular disease risk assessment because

many patients with microalbuminuria (especially diabetic patients) have increased

atherosclerosis (Gobal, Deshmukh, Shah, & Mehta, 2011). Patients with metabolic syndrome

have a 2-fold higher chance of developing CVD (Grundy, 2007). It is recommended that weight

reduction and an increase in physical activity be used to combat the growing prevalence of such

disease (Grundy, 2012). These factors contribute to the necessity of expanding this area of research.

Metabolic Syndrome

The term metabolic syndrome is a relatively new description for a group of risk factors that, when present together, can increase a person's risk for coronary artery disease, stroke, type 2 diabetes, and kidney disease. Although this condition is rapidly increasing in prevalence, it is not yet clearly defined in the literature. There are a variety of definitions of metabolic syndrome, coming from different organizations, which are summarized in the following discussion and in Table 1.

The National Cholesterol Education Program's Adult Treatment Panel III (ATP III) classifies metabolic syndrome using the following criteria: waist circumference > 102 cm (>40 in) for men and > 88 cm (>35 in) for women, triglycerides (TG) ≥ 150 mg/dL, HDL cholesterol < 40 mg/dL for men and < 50 mg/dL for women, blood pressure ≥ 130/85 mmHg and fasting blood glucose level ≥ 110 mg/dL. Metabolic syndrome can be diagnosed when 3 of the previously listed risk factors are present. Although many patients meeting this criterion for MetS do have insulin resistance, it is not required in order to be diagnosed with MetS (Grundy et al., 2004).

According to the World Heath Organization, insulin resistance (as indicated by T2DM, impaired fasting glucose, or impaired glucose tolerance) is required in addition to two other MetS risk factors. The risk factors include: high blood pressure (≥140/90 mmHg), plasma triglycerides ≥ 150 mg/dL (≥1.7 mmol/L), HDL cholesterol < 35 mg/dL (0.9 mmol/L) in men or < 39 mg/dL (1.0 mmol/L) in women, BMI > 30 kg/m^2 and/or waist:hip ratio > 0.9 in men and >

0.85 in women, and urinary albumin excretion rate ≥ 20 μg/min or albumin:creatine ratio ≥ 30 mg/g (Grundy et al., 2004).

The European Group for the Study of Insulin Resistance (EGIR) requires insulin resistance in conjunction with two or more of the following: central obesity (waist circumference ≥ 94 cm for males and ≥ 80 cm for females), dyslipidemia (triglycerides ≥ 2.0 mmol/L), hypertension (blood pressure ≥ 140/90 mm Hg, and a fasting plasma glucose ≥ 6.1 mmol/L (Bloomgarden, 2004).

The International Diabetes Federation (IDF) defines metabolic syndrome as having central obesity (waist circumference defined by ethnicity-specific values) and two of the following: raised triglycerides (>150 mg/dL), reduced HDL cholesterol (<40 mg/dL for males, <50 mg/dL for females), raised blood pressure (systolic >130 mm Hg, diastolic >85 mm Hg), and raised fasting plasma glucose (>100 mg/dL or previously diagnosed T2DM) (Alberti et al., 2006).

This inconsistency in definitions further exemplifies the necessity of more research in this area of metabolism, and shows why its clinical determination is often complicated. However, most epidemiological studies have used the ATP III definition based on its easy utility in the clinical setting (Singh & Kari, 2013). Given the presence of such a wide array of components and criteria, a decision had to be made as to what criteria and definition would be used in this thesis to diagnose MetS. The definition of MetS used in this study was made stringently so that it would be an inclusive mixture similar to most of the definitions: where two or more risk factors are occurring in conjunction with insulin resistance (in our case, each study has 3 risk factors in conjunction with insulin resistance).

Table 1. Criteria for diagnosing MetS based on differing criteria of various organizations.

Organization	Components of MetS	Requirements for Diagnosis
The National Cholesterol Education Program's Adult Treatment Panel III	Waist circumference ≥ 102 cm or 40 in (male), ≥ 88 cm or 35 in (female) TG ≥ 1.7 mmol/L (150 mg/dl) HDL-C < 40 mg/dL (male), < 50 mg/dL (female) Blood Pressure ≥ 130/85 mmHg Fasting plasma glucose ≥ 6.1 mmol/L (110 mg/dL)	Must have 3 or more components
World Health Organization	Waist:hip ratio > 0.90 (male), > 0.85 (female) or BMI > 30 kg/m^2 TG ≥ 2.0 mmol/L and HDL-C ≤ 0.9 mmol/L (male), ≤ 1.0 mmol/L (female) Blood Pressure ≥ 140/90 mmHg Microalbuminuria: Urinary albumin excretion ratio ≥ 20 µg/min or albumin:creatine ratio ≥ 30 mg/g	Insulin resistance indicated by the presence of diabetes mellitus, impaired fasting glucose, or impaired glucose tolerance and 2 or more components
European Group for the Study of Insulin Resistance	Waist circumference ≥ 94 cm or 37 in (male), ≥ 80 cm or 31.5 in (female) TG ≥ 2.0 mmol/L and/or HDL-C < 1.0 mmol/L Blood Pressure ≥ 140/90 mmHg Fasting plasma glucose ≥ 6.1 mmol/L	Insulin resistance as the top 25% of the fasting insulin values among non-diabetic individuals and 2 or more components
International Diabetes Federation	TG ≥ 150 mg/dL (1.7 mmom/L) HDL < 40 mg/dL (1.03 mmol/L) in males, < 50 mg/dL (1.29 mmol/L) in females Blood Pressure > 130/85 Fasting Plasma Glucose > 100 mg/dL (5.6 mmol/L)	Central obesity and 2 or more components

Insulin Resistance

Although each organization has a slightly different definition of MetS, insulin resistance is a

recurring factor in a majority of the definitions. This suggests that insulin resistance underlies

the pathology metabolic syndrome by some mechanism, and thus it is important to define insulin

resistance and explain its role in the pathology of MetS.

Insulin resistance is another term for reduced insulin sensitivity in tissues that are capable

of responding to insulin. When the tissues are insulin resistant, increased glucose transport does

not occur after insulin stimulation, and thus glucose remains in the blood. This often leads to a

state of hyperinsulinemia where more and more insulin is secreted in response to hyperglycemia.

Often, the overproduction of insulin tires out the pancreatic beta cells (cells that produce insulin),

and when they are no longer able to produce insulin, the patients becomes hyperglycemic

(pertaining to high blood glucose) (Alberti et al., 2006).

The presence of insulin resistance is often signaled by hyperglycemia and glucose

intolerance, which are two diagnostic indices for T2DM. Upon assessment of acute and chronic

T2DM, it itself classifies as a risk factor for CVD. In addition, insulin resistance independently

has been linked to a risk of CVD, as well as a rise in blood pressure. Insulin resistance has also

been linked to obesity, where a rise in fat content correlates with increasing insulin resistance

(Grundy et al., 2004). Insulin resistance mediates many of the risk factors for MetS, which

shows the prevalent role insulin resistance is playing in its diagnosis.

Patients with metabolic syndrome have a 5-fold increase in diabetes risk compared to

healthy individuals, which contributes to an association of metabolic syndrome with a pre-

diabetic state (Stern, Williams, Gonzalez-Villalpando, Hunt, & Haffner, 2004). Since metabolic

syndrome is becoming an increasingly prevalent precursor to heart disease, diabetes, and kidney

disease, the goal of this project is to examine genes that are likely to be influential in the

pathology of metabolic syndrome in the kidney. This could give valuable insight into possible

causes of metabolic syndrome, which could aid in studying the various conditions that often stem

from metabolic dysfunction.

Functions of the Kidney

The main function of the kidneys is to recondition the blood to remove toxins, and retain

nutrients producing urine as the waste product. Urine is composed of metabolic wastes that the

body needs to move along with water. Water is removed in specific quantities to regulate the

osmolarity of bodily fluids. Blood flows through a series of glomeruli, which are tangles of

blood vessels surrounded by a cup-shaped structure called the Bowman's capsule. The kidneys

are efficient at filtering the blood, due to the high pressure on the glomerular capillaries. The

wastes that are filtered from the blood form urine. Before the urine is fully formed, many solutes

and water are reabsorbed into the blood. The urine is collected in the renal tubule and later joins

the collecting ducts to travel to the bladder for excretion (Frayn, 2010). Figure 1 shows the

general anatomy of the kidney, and the functional unit of the kidney, the nephron.
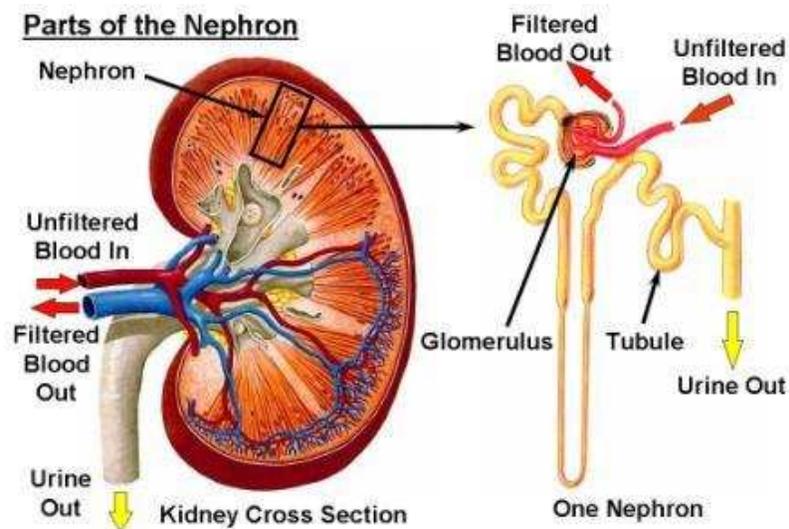
Figure 1. General Kidney Anatomy. Retrieved from http://www.unckidneycenter.org/images/glomerulus.jpg on 4/07/14.

Renal Complications of Metabolic Syndrome

Patients with metabolic syndrome are at a high risk for renal complications such as

microalbuminuria and/or chronic kidney disease (Locatelli et al., 2006). Metabolic syndrome

patients have a 2.5-fold higher risk of CKD and a 2-fold higher risk of microalbuminuria as

compared to a healthy individual (Singh & Kari, 2013). Microalbuminuria occurs when small

amounts of protein (albumin) are lost via the urine due to pathology in glomerular filtration,

when it would normally be returned to the bloodstream when the kidneys are functioning

correctly. The kidneys are at risk for damage from chronic kidney disease (CKD) in patients

with metabolic syndrome as a result of atherosclerotic vascular damage (Okada et al., 2013).

The pathologic effects of metabolic syndrome on the kidney can be manifested in the following

ways: tubular atrophy, interstitial fibrosis, and global and segmental sclerosis. It is hypothesized

that the combination of insulin resistance, hypertension, dyslipidemia and inflammation result in

increased expression of adipocytokines, angiotensin, inflammatory cytokines such as interleukin

(IL)-6 and tumor necrosis factor α (TNFα), which results in renal fibrosis (Singh & Kari, 2013).

The high blood glucose (hyperglycemia) noted in many cases of metabolic syndrome may also

be contributing to glomerular damage. Microalbuminuria has also been noted in individuals in

this pre-diabetic state, which is of concern because it may indicate CKD. A pre-diabetic state is

one is which blood glucose levels are elevated, but not enough to be diagnosed as diabetes

(Grundy, 2012). Since a pre-diabetic state can lead to T2DM, the kidney is a valuable organ to

study in relation to metabolic syndrome because diabetes can lead to diabetic nephropathy,

which is a common microvascular complication of T2DM (Frayn, 2010). These conditions can

result in renal failure, making it necessary for the patient to receive dialysis to filter toxins and

metabolic wastes out of the blood.

Goals of this Thesis

Based on the widespread pathologic effects of metabolic syndrome, this thesis aims to:

1. Further understand the pathologic effects of metabolic syndrome on the kidney by examining

    changes in gene expression caused by metabolic syndrome in the kidney throughout the rat

    genome to determine which genes should be examined in further detail.

2. Verify selected results using real-time polymerase chain reaction.

Microarray Analysis

Microarray analysis is used in this project to study changes in gene expression caused by metabolic syndrome that may lead to kidney disease. Microarrays are a research tool aimed at indentifying patterns of gene expression among an entire genome. Each microarray slide is printed with the entire genome of the organism of interest, using small fragments of complementary DNA (cDNA) for each gene. The cDNA is arranged in a grid-like pattern allowing for each gene to have a precise location on the microarray slide (Figure 2). There are two distinct types of microarray analysis: single-channel arrays and two-channel arrays. Microarray data from 3 different studies will be combined and analyzed in this thesis, from single channel arrays.
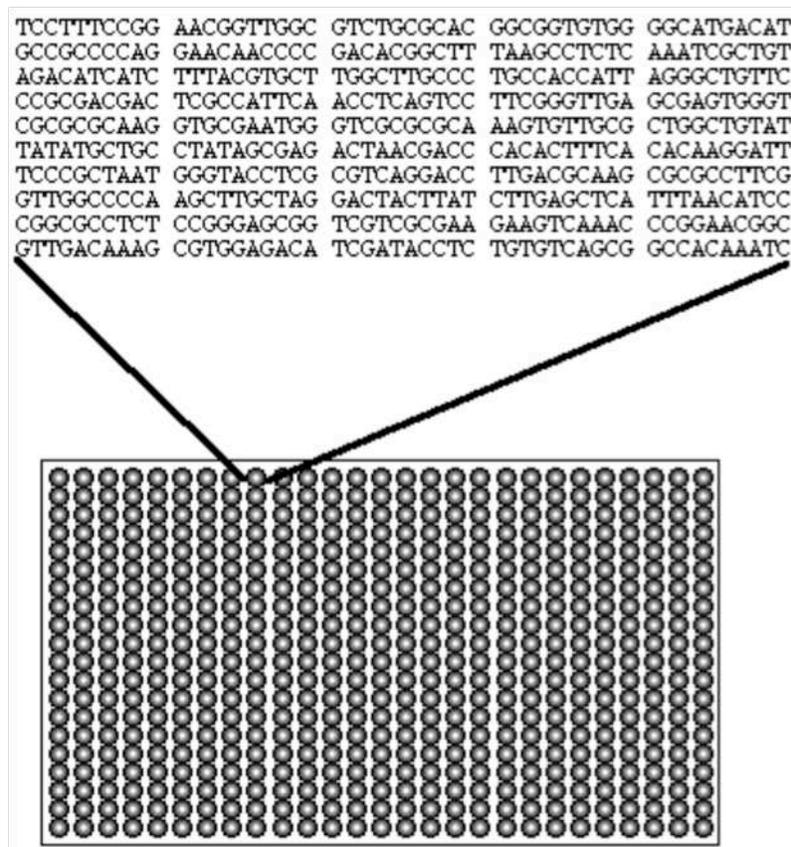
Figure 2. Example of a microarray slide, where each bead represents a distinct gene in the rat genome. This figure shows more nucleotides per individual gene than the arrays used in this thesis. Image source: Dr. Terrie Rife, James Madison University.

14

Messenger RNA (mRNA) is isolated and extracted from kidney tissues of both the wild-type and experimental groups. The mRNA is reverse transcribed into cDNA and is hybridized to the microarray slide. The wild-type and experimental samples are each hybridized to different microarray slides in addition to rat genomic cDNA. In a single channel array, the wild-type and experimental cDNA samples are tagged with a red florescent dye, while rat genomic cDNA is labeled with a green fluorescent dye. The slides are then scanned to determine the ratio of red to green pixels for each gene. The green genomic cDNA is a universal standard, so the ratios are treated as absolute levels of gene expression and are compared between the wild-type and experimental group to determine if a gene is upregulated or downregulated.

Meta-Analysis

A meta-analysis is a technique used to combine data sets from multiple different sources. Since metabolic syndrome includes many factors contributing to its clinical diagnosis, a meta-analysis can be used to combine these factors and thus gain a wider understanding of the possible genes involved in metabolic syndrome. There are now a large variety of studies related to metabolic syndrome or the components of it; many of which employ microarray technology to investigate genes related to distinct components of metabolic syndrome.

However, based on the costly nature of using microarrays, these studies are often narrow in scope (investigate only one condition or aspect of a condition) and have small data sets. Using a meta-analysis, the data for specific conditions as a whole can be combined and be directly compared although there are likely to be differences in experimental protocols. This is an inexpensive way to investigate multivariate conditions such as MetS using data from studies more concentrated to a single factor of the condition. Using a meta-analysis increases the sample

size and subsequently the statistical power of the results (Choi, Yu, Kim, & Yoo, 2003). Figure 3 demonstrates the technique of using a meta-analysis as applied to this thesis.

Three different studies, each investigating a specific aspect of MetS, were used to analyze the common genes between the three studies, and determine changes in gene expression caused by risk factors for MetS. The meta-analysis is used to determine which changes in gene expressions overlap between the studies. Since all three risk factors can be present together and denote a MetS condition, genes that overlap in all three studies may be influential in the regulation and pathology of MetS. For example, in Figure 3 below, while Gene B and Gene C may be significant in two of the three studies, only Gene A is significant in all three studies. The conclusion drawn from this hypothetical data would be that Gene A could be playing a role in MetS, and thus should be further investigated.
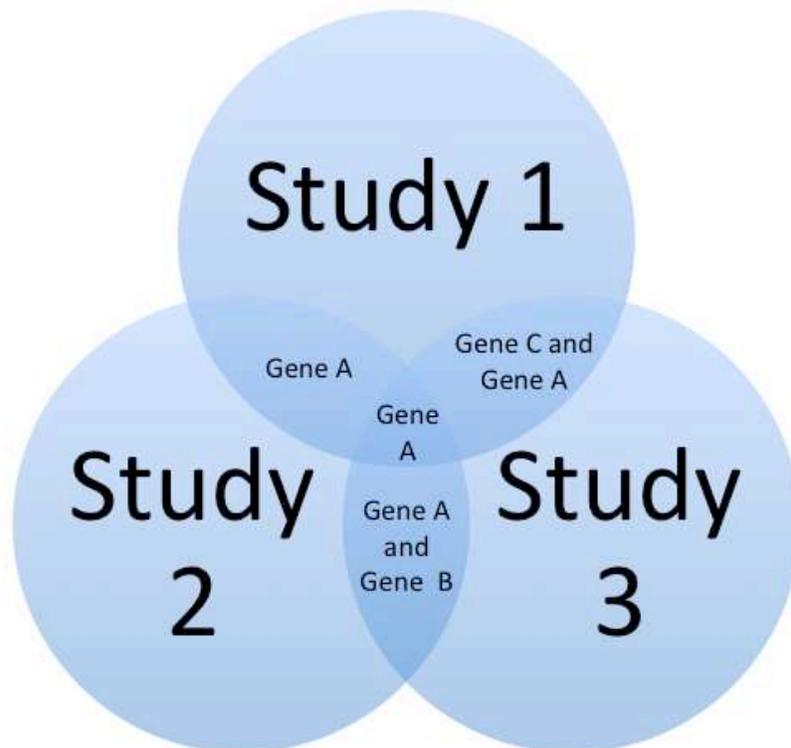
Figure 3. Venn Diagram showing the meta-analysis technique. Each study has insulin resistance and the characteristics identified above. The intersections show which genes were significant in the overlapping studies. Gene A was significant in all three studies and thus should be examined further.

Purpose

Due to the increasing prevalence of MetS and its complications, metabolic disease has gained

increased awareness and attention for the general public and scientists. Because of its

multivariate nature and multi-factorial effects, scientific literature has had significant difficulty in

determining the exact mechanism underlying MetS and its pathogenesis. Three studies, which

are detailed in the Methods section of this thesis, were chosen to represent MetS. For all of the

studies, single channel microarray data for the kidney was available, and a revised statistical

protocol from a previous study was used to isolate upregulated and downregulated genes. This

thesis aims to determine what genes are upregulated and which genes are downregulated in

kidneys of rodent models with metabolic syndrome. Once identified, further real-time PCR

analysis will be completed to verify the accuracy of the meta-analysis data. The answers to these

questions will be used to infer as to which genes could be playing a central role in the pathology

of MetS.

Methods

Data Collection

In order to collect microarray data related to MetS, existing studies concerning various characteristics of MetS were obtained via the National Center for Biotechnology (NCBI) Gene Expression Omnibus (GEO) database and the European Bioinformatics Institute (EBI) ArrayExpress database. Two studies from NCBI GEO database and one study from EBI ArrayExpress were chosen for analysis based on the following characteristics:

1. Study used kidney tissue samples.

2. Insulin resistance and at least two other symptoms of metabolic syndrome were evident in all rat models at the time of data collection.

3. All studies were single channel Affymetrix microarrays so that they could be easily compared to investigate the entire rat genome.

4. Raw data was available online.

The raw microarray data was obtained via NCBI's GEO2R or ArrayExpress and extracted into a comma-separated values (.csv) file. This format of the data was compatible for statistical analysis using the program R, and Microsoft Excel.

Studies Chosen for Analysis

After a thorough search for studies meeting the above criteria was performed, three studies were selected for use in the meta-analysis. All three studies used an Affymetrix Rat Genome 230 2.0 One-Channel Array. All three studies also investigated a total of 31099 genes via microarrays.

Table 2. Descriptions of the studies chosen for meta-analysis.

Study	Access Number	Length of Treatment	Rats Used	Criteria for Relation To Metabolic Syndrome (in addition to insulin resistance	Number of Arrays Used	Link to Microarray Data
1	GSE4800	Started at 5 wks, fed for 8 wks	13 wk old male DS rats and Lewis rats	High triglyceride levels, high blood pressure, and a high fasting blood glucose level	6 (3 Lewis, 3 DS)	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4800
2	GSE7193	12 wks	12 wk old male ZO rats and ZL rats	Large waistline, high blood pressure, and a high fasting blood glucose level	6 (3 ZL control diet, 3 ZO control diet)	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7193
3	E-MEXP-1695	6 months	BN rats and Goto-Kakizaki rats	High triglyceride levels, high blood pressure, and a high fasting blood glucose level	11 (5 normal BN control, 6 Goto-Kakizaki insulin resistant)	http://www.ebi.ac.uk/arrayexpress/experiments/E-MEXP-1695

Study One: Dahl Salt-Sensitive Rats in Hypertension Study

Study one performed by (Yasui, Kajimoto, Sumiya, Okuda, & Iwai, 2007), used 13-week old

male Lewis rats and 13-week old male Dahl Salt-Sensitive (DS) fed a high salt diet (8% NaCl) in

order to study salt-sensitive hypertension. Male DS rats and Lewis rats were fed the high salt

diet starting at 5 weeks of age, and the treatment continued for 8 weeks. The DS rats were fed a

high salt diet to induce salt-sensitive hypertension. Their goal was to narrow down the genes

that may be involved in salt-sensitive hypertension in the kidney in the chromosome 10

quantitative trait loci region. Their control group in the microarray analysis included the Lewis

rats on a high salt diet, compared to the treatment rats (DS rats on a high salt diet). The

treatment rats developed the following three metabolic syndrome criteria: high triglyceride level,

high blood pressure, and a high fasting blood glucose level (Table 2). They did two other

microarray experiments; the data of which is available via GSE 4800. They selected genes that

modulated expression more than 2-fold or less than 0.5 fold in all three of their studies. They determined that 8 genes were differentially expressed by their criteria.

After reverse transcription, PCR verification, and Northern blot analysis, it was determined that the expression levels of Ccl2 mRNA were 10-fold higher in the DS rats fed a high salt diet than in the Lewis rats fed a high salt diet. This was the only gene whose trend was verified via PCR, showing that it was differentially expressed between DS and Lewis rats after a high salt diet, concluding that the gene is upregulated in the DS rats in response to a high salt diet. Ccl2 is known to be involved in renal injury, and higher expression of Ccl2 could aggravate macrophage infiltration, which can ultimately result in an exacerbation of salt-sensitive hypertension. They concluded that this gene may be playing an influential role in hypertension in the kidneys.

Why is this study applicable to this meta-analysis? For our study, the DS rats on a high salt diet were used as the treatment and the Lewis rats fed a high salt diet were the control. A total of 6 arrays were used from this study (3 DS arrays, and 3 Lewis arrays). This study uses Dahl-Salt sensitive rats, which by definition exhibit hypertension, insulin resistance, hyperinsulinemia, and hypertriglyceridemia (Animal models of disease: Metabolic, renal, and cardiovascular.2011). Although this study focuses on studying hypertension, the characteristics these rats exhibit fit our criteria for a diagnosis of MetS, making it an applicable study to include in the meta-analysis.

Study Two: Zucker Obese Rat Study

The study by (Song, Liu, Ressom, Tiwari, & Ecelbarger, 2008) used 12-week old male Zucker Lean (ZL) rats as the control rat and 12-week-old male Zucker Obese (ZO) rats as the treatment rat in an experiment studying obesity as it related to the development of type 2 diabetes. Zucker

Obese rats have a deletion mutation in the leptin receptor, which causes misregulation of appetite and hunger. These rats develop a large waistline, high blood pressure, and a high fasting blood glucose level (Table 2). In addition to studying the difference in gene expression between ZO and ZL rats, the study also examines the effects of different compounds used as potential diabetes drugs, such as rosiglitazone and thiazolidinediones.

Using the microarray data, 903 probe sets were determined to be significantly altered with at least a 1.5-fold change between the control and treatment groups. There were 300 probe sets that were increased and 244 probe sets that were decreased in obese rats compared to lean rats. Increased genes included the B-subunit of the epithelial sodium channel (ENaC), the thiazide-sensitive Na-Cl cotransporter, and aquaporin 3. Decreased genes included angiotensin-converting enzyme, type 1 (ACE1) (Song et al., 2008).

Why is this study applicable to this meta-analysis? For our study, 3 arrays from the ZO rats and 3 arrays from the ZL rats were used as the treatment and control respectively. Zucker Obese rats were used to study obesity and T2DM. Zucker rats by definition have the following characteristics: obesity, insulin resistance, hyperinsulinemia, and hypertriglyceridemia (Harlan Laboratories, 2014). Although this study focused on the effects of obesity and diabetes, it is useful for this meta-analysis since the Zucker Diabetic Fatty (ZDF) rats used for PCR verification are also obese and diabetic.

Study Three: Insulin Resistance in a Transgenic Rat Strain

In a study by (Wallis et al., 2008) the cardio-metabolic syndrome (similar to our definition of metabolic syndrome) was examined using Brown Norway (BN) rats as the control rats and Goto-Kakizaki (BN.GK-D1Wox18/D1Got254, a transgenic insulin resistant strain) rats as the treatment rats. This study utilizes this type of rat to examine genes related to diabetes and

obesity, because GK rats carry a mutation in a G-protein receptor that is primarily associated

with obesity. These rats exhibit high triglyceride levels, high blood pressure, and a high fasting

blood glucose level (Table 2).

At the 12 and 24 week stage the congenic rats showed elevated blood glucose levels and

hyperinsulinemia. Six-month-old rats were used for microarray analysis. According to the

microarray data, a significant difference in transcription was observed for 478 distinct genes, and

the trends were verified for 38 of the genes. The results of this study identified a genetic basis of

risk factors for the cardio-metabolic syndrome.

Why is this study applicable to this meta-analysis? This study uses Goto-Kakizaki rats to

investigate the cardio-metabolic syndrome. Goto-Kakizaki rats naturally develop T2DM early in

life, but they are not obese. These rats are commonly used in studies investigating metabolic

phenomena, and again meet our criteria for having metabolic syndrome. Since T2DM is a

common result of MetS, this study adds to our meta-analysis with regards to the later stages of

MetS.

Pre-Processing of Data

The data was processed by a previous group of students involved in the Undergraduate Research

in Biology and Mathematics (UBM) program (Lauren van Reesema, Keith Zirkle, and Jonah

Williams) and Dr. Nusrat Jahan. The raw data in its entirety was extracted from either the NCBI

GEO Omnibus database or EBI ArrayExpress and converted into .csv files for later use with

Microsoft Excel and the program R for statistical analysis. For each of the three studies, separate

.csv files were generated for the specified wild type and treatment data.

The resulting data was then quantile normalized before it was individually analyzed via

the program R. Quantile normalization is performed so that all of the arrays have the same

distribution, shape and similar spread. Next, for each gene, a t-test was performed to determine individual significant differences in gene expression. This analysis resulted in a test statistic to account for study variation and a p-value for each gene.

The statistical program R is used for analysis of microarray data. Before any statistical analysis can be performed, the data were pre-processed. The pre-processing included averaging gene duplicate values, log-2 base transformation and quantile normalization of the data. The original data are right skewed, but all statistical tests require data to have at least approximately normal distribution. Log transformation is a technique that makes right skewed distribution more symmetric by minimizing the scale of differences between values. With this technique, if the gene signal intensity is very similar between the treatment and the control, the treatment to control ratio will be close to one, and the $\log_2(1)$ is equal to 0. If the treatment signal intensity is larger than the control, the ratio will be greater than one and log transformation will produce a positive number, signaling an upregulated gene. On the other hand, if treatment signal intensity is smaller than the control intensity, a negative number will result signaling a downregulated gene.

Mean log(gene expression) = Mean log(treatment expression) – Mean log(wild-type expression)

Meta-Analysis

The meta-analysis performed in this project was performed by Dr. Jahan, using the program R and a code presented in Appendix 1. An effect size (u) was calculated for each gene using the following formula, where y_m is the mean treatment gene expression of the gene m, and x_m is the mean gene expression of the wild-type:

$$u_m = \frac{\bar{y}_m - \bar{x}_m}{s_i^{pool}}$$

The effect size represents the standardized mean difference in gene expression for the treatment and wild-type rats. The mean gene expression of the control is subtracted from the mean treatment gene expression. This is then divided by the pooled standard deviation and a correction factor (s_i^{pool}). Using a pooled standard deviation accounts for the variability in the three studies and standardizes the sample sizes of the studies. The correction factor is used to ensure that the denominator is large enough to not artificially inflate the meta-Z value. The correction factor used is the 90th percentile point.

The meta-Z statistic is calculated for each gene using the mean and variance. The mean is divided by the square root of the variance. The purpose of the meta-Z value is to determine the threshold value for significant genes corresponding to a specified quantile or cut off point (Hu, Greenwood, & Beyene, 2005). The meta-Z values were then used to determine the statistical significance of each gene.

Determination of Significant Genes within the Meta-Analysis

Genes were considered to be significantly upregulated if the meta-Z score was greater than or equal to +1.8888 and were considered to be downregulated if the meta-Z score was less than or equal to -2.054. These meta-Z cut off values were assigned stringently in order to minimize the presence of false positives and negatives. In order to determine the meta-Z cut offs, the difference between the meta-Z and the perm-Zs (permuted meta-Zs) were taken into account. Meta-Z cutoff values of 1.8888 and -2.054 were found using a difference of 2 between the true meta-Z and the average of the permuted meta-Z. The perm-Zs are calculated in the same way

that the meta-Zs are calculated, but used the random permutation of the original data. Permutation is described further in the next paragraph with regards to the false discovery rate. If a gene is not different between the treatment and control, then the random permutation of the data will produce perm-Z values that would be very different from the meta-Z values (Hu et al., 2005). If a gene is significant based on permutation, it can be concluded that the gene is not significantly different between the treatment and wild-type samples and its significance in the permuted data is purely due to random chance.

The false discovery rate (FDR) method can be used to control the amount of false positive genes in the data set. The FDR is determined by dividing the number of expected false positive genes by the number of truly significant genes. The number of truly significant genes must be estimated since it is unknown. This means that the FDR is also estimated using a permutation method. This method is described by (Tusher, Tibshirani, & Chu, 2001). Permuting the data rearranges it to account for every possible scenario of gene expression. A weighted average, weighted variance, and a meta-Z statistic are generated for each gene in each permutation. Since the data is permuted β times, there are β permuted meta-Zs for each gene. The permuted meta-Z values are arranged from smallest to largest within each permutation and using the same threshold quantile that is used to identify significant genes from the meta-Zs, significant genes can be determined within each permutation. These significant genes represent the false positives, since the permuted data should not have differential effects.

In cases where gene expression is inaccurately reported as significantly different, a type 1 (false positive) error occurs (Storey & Tibshirani, 2003). Although in most statistical tests a significance level of 0.05 is used to determine the error rate, in microarray studies an alternate method must be used due to the large amount of genes being tested at once. Since approximately

30,000 genes are being tested simultaneously, the error rate of each gene is combined producing an overall large error rate. For example, when testing 30,000 genes each with an error rate of 0.05, there would be 1,500 expected false positives. This error rate is way too large for the study to have any statistical significance and relevance.

Quality Check of Significant Genes

In order for genes to truly be considered significantly upregulated or significantly downregulated, some additional parameters were put in place to ensure that these significant genes made biological sense. The first criterion was that the average fold change for all three studies was at least a 1.5 fold difference between the treatment animals and the wild-type animals. This was used in order to determine whether or not the genes were biologically significant, using the following equation:

$$\log_2(1.5) = 0.58496$$

This means that the average fold change of significant genes must be ≥ 0.58496 or ≤ -0.58496. The second criterion was that for a gene to be deemed upregulated, it must have a fold change $\geq$ 0.58496 in at least two of the three studies, and must have a positive fold change in the third study. For a gene to be deemed downregulated, it must have a fold change ≤ -0.58496 in at least two of the three studies, and must have a negative fold change in the third study. This ensures that the average fold change is not artificially inflated or deflated based on a particular study. There are many possible scenarios that can be considered in regards to the previously described criteria. For example, a gene could be really highly upregulated in one study, but not quite upregulated enough in the other two studies to be statistically significant. Since the first study is inflating the average expression of the gene, this gene would not be retained in the list of genes to further investigate since it does not meet our criteria that a gene must be significant in at least

26

2 of the three studies and follow the general trend for gene expression in the third study. Another case could be that a gene is highly upregulated in two of the studies, but downregulated in the third study. This gene would also not be included in the list of genes to analyze further since it does not meet the specified criteria.

For example, the gene titled Syt10 (synaptotagmin X) meets our statistical cutoff value for the meta-z statistic to be called statistically significant and has an overall gene expression of 1.5 fold for all three studies, as seen by its average fold change of 0.808. The fold change for study one is 0.877, and the fold change for study two is 1.309. These positive values (>0.58) support the conclusion that this gene is upregulated. However, the fold change for study three is 0.239, which still shows a slight upregulation, but not enough to meet our criteria of 0.58 (0.239 < 0.58). Since all three studies agree at this basic biologic level (upregulation) even though the third does not, it will be retained in the list of genes for further examination.

To carry out this screening, any gene with a fold change < 0.58496 or > -0.58496 was flagged in study one. Then, any gene with a fold change < 0.58496 or > -0.58496 and was flagged in study two. If it had been flagged in both studies one and two, it was entirely ruled out because it did not meet our criteria (since it would be out in two of the three studies). The same criteria were used for study three. Again, if it was < 0.58496 or > -0.58496 it was flagged, and if it had been previously flagged in either study one or study two (or both studies) it was ruled out overall.

Next, some genes were thrown off simply because we were unable to determine their biological function. Lastly, genes that had already been associated with metabolic syndrome were excluded from analysis, since this thesis aims to find novel genes that could be influential in the pathology of metabolic syndrome.

Tissue Extraction

In order to verify the gene expression changes in these novel genes, tissues were obtained from Zucker Diabetic Fatty rats fed a 66% fructose diet or a standard diet for 10 weeks while monitoring their blood glucose and insulin levels. Zucker Diabetic Fatty rats exhibit type 2-diabetes, hyperlipidemia, glucose intolerance, obesity, and hyperinsulinemia (Animal models of disease: Metabolic, renal, and cardiovascular.2011). The rats were characterized as diabetic at the time of tissue extraction by their marked hyperglycemia and hyperinsulinemia. Kidney tissue was harvested by Divya Bansal, a former graduate student in the lab. These tissues were surgically removed and stored in 3-4 mL of RNA later® solution to inhibit RNase activity. The tissues were stored in this solution for two days at 4°C, then transferred to a new tube and stored at -80°C until use.

RNA Isolation

To study gene expression, RNA was isolated from the kidney tissue by Divya Bansal using the TRIzol method. Four basic steps were used in the process of RNA isolation: breaking open the kidney cells, preparing the RNA, determining the concentration of the prepared RNA, and checking the RNA for degradation using gel electrophoresis. Throughout this process, care was taken to avoid RNase activity by using RNAZap (Ambion) detergent wipes to chemically treat everything that could have potentially come in contact with the sample.

To homogenize the kidney tissue, approximately 500 μL of TRIzol reagent per 50-100 mg of tissue was added to a dounce homogenizer in order to lyse kidney cells. The kidney sample was dissolved in the TRIzol reagent as quickly as possible to prevent degradation of the sample by RNases.

After homogenization, the sample was moved to a RNase free 15 mL tube and incubated at room temperature for 5 minutes. Approximately 200 mL of chloroform per one mL of TRIzol reagent was added to the tube and shaken vigorously for 20 seconds. This mixture was incubated at room temperature for an additional 5 minutes. Centrifugation at 12,000 × G at 4°C for 15 minutes was used to pellet the cells by dividing the hydrophilic RNA from the hydrophobic proteins and bulkier DNA. The upper aqueous clear layer contains the RNA and was placed in a new tube with 500 mL of isopropanol per one mL of TRIzol reagent used to precipitate the RNA. This tube was vortexed and incubated at room temperature for 5 minutes, then centrifuged at 12,000 × G at 4°C for 8 minutes. The supernatant was poured off, and one mL of 75% ethanol per one mL of TRIzol reagent was added to the pellet. The sample was then centrifuged at 7,500 × G at 4°C, and the ethanol was poured off. To remove the remaining ethanol, the sample was centrifuged again for one minute and the remaining ethanol was pipetted off. The sample was then dried in the fume hood for approximately 5 minutes to allow the remaining ethanol to evaporate. Once the samples were dried the RNA was resuspended in 50 μL of RNA Storage Solution and stored at -80°C.

Examining the Quality of the RNA

To quantify the isolated RNA, the Nano-vue spectrophotometer was used to measure the absorbance of the sample at 260 nm (A260) and 280 nm (A280), and determine the concentration of the isolated RNA. The quality of the RNA was determined using a ratio (A260/A280), which represents the ratio of nucleotide absorbance and protein absorbance. Pure RNA should have a A260/A280 of 2.0-2.2, but any ratio over 1.9 was considered acceptable RNA quality.

To ensure the RNA quality, the RNA was run on a 1.2% agarose gel. Two distinct bands are shown for good quality RNA: a 28S band and a 18S band. These represent the large and small subunits of ribosomal RNA. The 28s band appears twice as bright as the 18S band.

Reverse Transcription of RNA

Once the RNA was deemed to have adequate quality, it was reversed transcribed from RNA into cDNA using a Superscript II Invitrogen kit. A mixture of 5 μg of RNA, 2 μL of random hexamers, 1 μL of dNTPs, and nuclease free water was added together to make a total volume of 10 μL. The sample was then incubated at 70 °C for five minutes, and placed on ice for one minute to denature the RNA and primers. Next, a mixture of 2 μL of 10x RT buffer, 2 μL of 0.1 M DTT, 2 μL of 50 mM $MgCl_2$, and 1 μL of RNase out inhibitor was added and the sample was incubated for two minutes. Next, 0.5 μL of 50 units/μL Superscript II RT was added to the sample, which was then incubated at 42°C for 50 minutes. Lastly, 0.5 μL of 100 units/μL RNase H was added to the sample, which was then incubated at 37°C for 30 minutes. The samples were then stored at -20 °C until quantitative real-time PCR could be performed.

Primer Design

In order to verify the gene expression of the chosen novel genes, primers specific to each gene had to be designed. The NCBI Primer Blast program was used to accomplish this (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). The mRNA sequence was located and entered into the program. The following parameters were changed from the defaults: PCR product size (70-300 bp) and melting temperature (57° C - 63°

C, with optimal melting temperature at 60° C). The returned forward and reverse primer sequences were checked to make sure that they spanned at least one intron, and that the sequences did not match anything else in the genome besides the gene being amplified.

Table 3. Primer sets used in qPCR.

Product Name	Sequence	Product Length
B-actin Fwd B-actin Rev	5'-AAC CCT AAG GCC AAC CGT GAA AAG-3' 5'-TCA TGA GGT AGT CTG TCA GGT-3'	241 bp
Mgmt Fwd Mgmt Rev	5'-GTG AGC GAG GCG TGC ATG GG-3' 5'- CCCACCGAGCACCTCAGGAC-3'	95 bp
Ampd3 Fwd Ampd3 Rev	5'- ACCAACGCTTGCTGGTCGGT-3' 5'- TCTCTGCCAACAGCCGGACA-3'	105 bp
Klk1c9 Fwd Klk1c9 Rev	5'- CCTGTTCCTCGCCCTGTCCCT-3' 5'- ATCAGGACACCCCCGCAGAAAG-3'	144 bp
Stk32c Fwd Stk32c Rev	5'- CCGCGGGCTCCGATGTGTC-3' 5'- GCTTCTGCACGATGCACACCT-3'	264 bp
RGD 1309350 Fwd RGD 1309350 Rev	5'-GCCTACGTTTGTGCCGCCTG-3' 5'-GTAGCGCTCCGTGTCGAAGG-3'	152 bp

Figure 4 shows a DNA gel of the primers sepcified above. A one kilobase ladder was used, and 12 µL (including 6X loading dye) of each primer solution was loaded in each well.
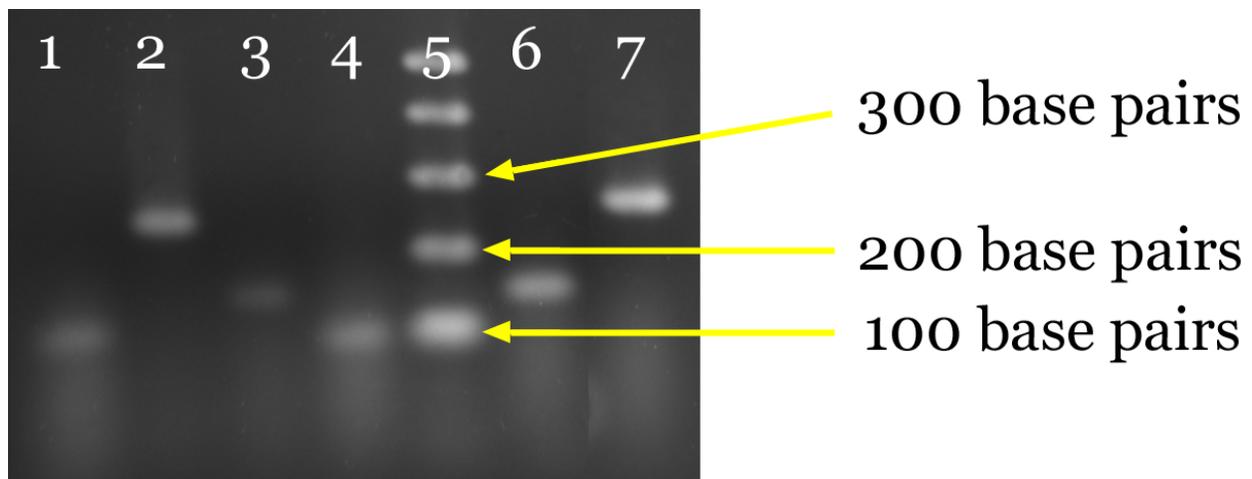
Figure 4. Gel showing the 5 primers and the housekeeping gene (B-actin). In lane one, Ampd3 was run. Lane 2 contains B-actin. Lane 3 contains Klk1c9. Lane 4 contains Mgmt, Lane 5 contains the molecular weight marker. Lane 6 contains RGD 1309350. Lane 7 contains Stk32c.

Quantitative Real-Time Polymerase Chain Reaction

To amplify DNA containing the genes of interest and B-actin, quantitative real-time polymerase chain reaction (qPCR) was used. A fluorescent dye called SYBR green is used to identify double-stranded DNA, thus inducing a brighter fluorescence as amplification continues.

To perform qPCR on the various samples, a master mix containing 1X SYBR green (enzymes and a dye that fits between double stranded DNA), and 0.2 μM forward and reverse primers for B-actin, Mgmt, Ampd3, Klk1c9, RGD1309350, or Stk32c was made. The various primer sets used are presented in Table 3. For each reaction, components were added in the following proportions: 10 μL of SYBR green master mix, 2 μL of 2mM primer, 2 uL of cDNA and 6 uL of water totaling 20 uL for each reaction. Each gene was amplified with its specific primer for each of the three control samples, and each of the 4 treatment samples. The reactions were then run on a Bio Rad Opticon Monitor CFX96 Touch™ Real-Time PCR Detection System (185-5196) with the following cycling conditions: 50°C for 2 minutes, 90°C for 2 minutes to denature cDNA and activate enzymes, followed by 50 cycles of: 95°C for 15 seconds, 60°C for 30 seconds, with a plate read at 60°C, which was used to analyze the data in regression mode. After the 50 cycles, the temperature was lowered to 60°C for 10 minutes. The machine then constructed a melting curve starting at 76°C, taking measurements every 0.2°C up to 98°C. After the amplification was complete, the products were run on a 2% agarose gel to confirm that the correct cDNA fragments were amplified (Figure 4).

Data Analysis For Determining Gene Expression

Data was analyzed using the $2^{-\Delta\Delta C_T}$ method, or the Livak method. This method assumes that both the target gene and the housekeeping gene are amplified with efficiencies near 100% and within

5% of each other. First, normalization of the C_T of the target gene and the housekeeping gene was performed.

$$\Delta C_{T(test)}=C_{T(target,test)}-C_{T(ref,test)}$$

$$\Delta C_{T(calibrator)}=C_{T(target,calibrator)}-C_{T(ref,calibrator)}$$

Next the ΔC_T of the test sample was normalized to the ΔC_T of the calibrator:

$$\Delta\Delta C_T=\Delta C_{T(test)}- \Delta C_{T(calibrator)}$$

Last, the expression ratio was determined using the following equation:

$$2^{-\Delta\Delta C_T}=\text{Normalized expression ratio}$$

Once the normalized expression ratio is obtained, it is converted into a percentage. This percentage is indicative of the percent of the control expression that the treatment is expressing. For example, if the normalized expression ratio is 0.66, it can be concluded that the gene is expressed at 66% of the control expression. This gene would be considered downregulated because it is expressed 34% less in the treatment than it is in the control.

Results

To determine what genes may be influential in the pathology of metabolic syndrome, a meta-analysis of 3 microarray studies was performed. After running the R code presented in Appendix 1 and before screening for the significant genes, there were 12,328 genes in common in this meta-analysis. Based on the meta-Z cutoffs of 1.8888 and -2.054, 597 genes were found to be significant: 551 of these genes were upregulated, and 46 of the genes were downregulated.

After screening the genes to determine which ones had an average fold change for all three studies that was at least a 1.5 fold difference between the treatment animals and the wild-type animals, 158 genes remained (this list of genes is available in the Appendix 2). Of the 158 genes, 135 were upregulated, and 23 were downregulated (Table 4). After screening the gene list according to the fold change in study one, 18 of the 158 genes were flagged because they did not have a fold change that was changed by 1.5 fold in study one. After screening the gene list according to the fold change in study 2, 51 of the 158 genes were not changed by 1.5 fold in study two (2 of which had already been flagged in study one). After screening the gene list according to the fold change in study three, only 90 genes remained that were not "off" in two of the studies.

Table 4. Genes meeting the criteria for an average fold change of > 0.58496 or < -0.58496.

Genes With a Fold Change of 1.5-fold or Greater	
Upregulated	Downregulated
Havcr1, Scg2, RGD1562844, Eps8l3, Pthlh, C4bpb, Socs2, Clcf1, Thsd7b, Atpbd3, Kntc1, Fezf2, Gkn2, Clu, Hfe2, RGD1565709, Ckap2, Traf4af1, Prl3d1, LOC500118, Tmsb10, Mfap5, Abra, Adamts1, Dscc1, Ctsw, Timp1, Fcrls, Ttc25, Dfnb59, Dlgap5, Tnfrsf12a, Mgmt, RT1-Db1, Fetub, LOC290595, Ltbp2, Ttk, Rbm11, Cox6a2, Tekt1, LOC691979, Mbnl3, Clec2d, Pemt, Ccna2, Xkr4, Syt10, Ptgs2, C3, Prc1, Wtap, Mfap3, RGD1561849, Kcne1, Krt25, Cd8a, Tekt5, Gpx7, Fes, Ccl2, Casp4, Apoa1, Mpp3, RGD1563091, Arhgap11a, RGD1310862, Cox15, Mfap4, Lum, LOC378467, Nkg7, P2rx4, Col8a1, Pqlc3, Ptprz1, Spp1, RGD1559980, Sass6, Asns, Lsp1, Nipsnap3a, Eml2, S100b, Ubash3b, Slc18a1, Pmf1, Mical1, Tap1, Trpv1, Slc15a3, LOC497995, Depdc1, Fgb, Rftn2, Rrm2, Gda, LOC654482, Psmb9, Dclk1, Samsn1, Tdrd1, RGD1311558, Rab19, Kcnip3, Bin2a, Pcp4l1, RGD1305713, Cds2, Grem1, Pycard, Itih1, Bub1, Jrk, Col3a1, Athl1, Ccl20, Cmtm3, Slc7a7, LOC503175, RGD1565844, Cebpa, RGD1307621, C1qb, Ampd3, Actn3, Ccdc80, Cxcl10, Obfc2a, Fgr, Agap2, Lgals3bp, Tmed6, Hck, Cks2	Abcb11, Rnf40, F2, Snap25, Bmp15, F5, Apoc1, Nell1, Mlc1, Serpinc1, Cep78, Grhl3, Krt14, Rp1h, Rbp4, Mrgprg, Scgb1c1, Slc22a13, RGD1309350, Hsd3b6, Klks3, Stk32c, Cyp2c

After removing genes that we were unable to determine their biological function, there were a total of 88 genes (73 were upregulated and 15 were downregulated) (Table 5). Lastly, after genes that had already been associated with metabolic syndrome were excluded from analysis 51 genes remained. A total of 37 out of the total 88 genes used for analysis had already been associated with MetS in the literature (Table 5). This means that 42.0% of the significant genes had already been associated with MetS. The 51 novel genes can be seen in Table 5.

Table 5. Genes already associated with MetS, and remaining novel genes. Remaining novel genes refers to their novelty to MetS association, not novelty in function.

Genes Already Associated with MetS	Remaining Novel Genes
Hfe2, Havcr1, Spp1, Athl1, Rrm2, Kcnip3, Ccl2, C1qb, Pthlh, Tnfrsf12a, RGD1563091, Ccl20, Socs2, C4bpb, RT1-Db1, Timp1, Fes, Tap1, Serpinc1, Hsd3b6, Snap25, Mlc1, Mfap5, Psmb9, Ltbp2, Col8a1, Lum, Grem1, Ptgs2, Clcf1, Ccna2, Clu, Abcb11, Rbp4, Cyp2c, Rp1h, Pemt	Abra, Ampd3, Arhgap11a, Atpbd3, Casp4, Ccdc80, Cep78, Ckap2, Cmtm3, Dclk1, Dscc1, Eml2, Eps8l3, Fetub, Gkn2, Grhl3, Jrk, Klk1c9, Kntc1, Krt25, Lgals3bp, LOC500118, LOC654482, Mgmt, Mrgprg, Nell1, Nipsnap3a, Pcp4l1, Pmf1, Pqlc3, Prc1, Prl3d1, Ptprz1, Rab19, Rbm11, Rftn2, RGD1307621, RGD1309350, RGD1311558, RGD1562844, RGD1565709, Sass6, Scg2, Scgb1c1, Slc15a3, Stk32c, Syt10, Tmsb10, Traf4af1, Ttk, Xkr4

Five genes were chosen as novel genes in relation to MetS that will be investigated further. The five chosen genes are: Ampd3, Klk1c9, Mgmt, RGD1309350, Stk32c. These genes were chosen based on their role in metabolism. A majority of the genes can be classified as having to do with an amino acid's role in metabolism or DNA packaging. Some genes were simply chosen for their statistical characteristics. RGD1309350, Klk1c9, and Stk32c are expected to be downregulated and Ampd3, and Mgmt are expected to be upregulated (Table 8).

Preliminary PCR data and subsequent analysis using the Livak method described in the Methods is shown in Table 6. The values in the table represent the percentage of the treatment compared to the control. For example, a value of 0.66128 means that the treatment's gene expression is 66.128% of the controls gene expression, and thus will be considered downregulated.

Table 6. Gene expression of treatment samples depicted as a percentage of the control.

Primer	Ampd3	Klk1c9	Mgmt	RGD1309350	Stk32c
Treatment 1	66.128 %	122.0171 %	217.8472 %	57.4349 %	464.8719 %
Treatment 2	74.9154 %	103.6761 %	317.8472 %	2208.508 %	516.7023 %
Treatment 3	86.9545 %	67.6932 %	710.2532 %	46.6516 %	36.7122 %
Treatment 4	0.006746 %	2.0152 %	38.1124 %	1.8136 %	2.0751 %

If the four treatments are averaged, the average gene expression can be obtained and used to determine whether the gene is upregulated or downregulated. The average gene expressions are as follows: Ampd3 is 0.571681, Klk1c9 is 0.738504, Mgmt is 3.210156, RGD1309350 is 5.78602, and Stk32c is 2.550904 (Table 7).

Table 7. Average gene expression ± a standard deviation.

Primer	Ampd3	Klk1c9	Mgmt	RGD1309350	Stk32c
Average	57.168 ± 38.6 %	73.850 ± 52.9 %	321.016 ± 284.1 %	578.602 ± 1086.9 %	255.090 ± 273.3 %

Table 8. Information used for the five chosen genes.

Gene	Expected Gene Expression	Average Fold Change	Fold Changes For Individual Studies (1,2,3)	Meta-Z	PCR verification
Ampd3	upregulated	0.608	1.12, 0.61, 0.091	2.78	57.17 %
Klk1c9	downregulated	-1.436	-1.93, -2.16, -0.22	-3.48	73.85 %
Mgmt	upregulated	0.880	0.87, 0.60, 1.17	3.31	321.02 %
RGD1309350	downregulated	-1.241	-1.44, -1.68, -0.60	-4.16	578.60 %
Stk32c	downregulated	-1.948	-3.85, -0.072, -1.92	-2.96	255.09 %

Discussion

The complex nature of metabolic syndrome is multi-faceted, which has been influential in increasing the amount of research and attention being devoted to this disorder. This thesis aims to elucidate some of the factors that may be playing a role in its pathology by studying gene expression. Overall, the purpose of this thesis is to determine what genes may be influential in the development and progression of MetS in the kidney.

The purpose of using a meta-analysis to study the effect metabolic syndrome has on gene expression is to gain statistical power. Since metabolic syndrome is a multi-factorial disease, a meta-analysis is used to make a stronger and broader conclusion about the disease by examining studies with different components of the disease. By combining the results of these studies, the statistical power is increased which leads to a stronger conclusion about the regulation of gene expression by metabolic syndrome. Due to limited amounts of available microarray data, not all aspects of metabolic syndrome could be addressed in this thesis. However, since 37 out of 88 (42.0%) genes were already associated with metabolic syndrome, it seems that the methodology used in this meta-analysis is valid.

When analyzing the data by fold change to ensure that the genes were actually biologically significantly upregulated, the genes had to have significantly upregulated or significantly downregulated fold changes in 2 of the 3 studies and follow the same trend in the third study. When performing this step, many of the genes had significant fold changes in studies one and two, but were thrown off in the third study. This could be due to the length of diet treatment of the rats in the studies, with the third study treating the rats the longest. This

38

could mean that the rats in that study were farther along in their progression of MetS, which could differentially affect gene expression.

Although there are many advantages to using meta-analyses, there are also some disadvantages. Due to the fact that different laboratories performed the different studies used in the meta-analysis, it is possible that there is some variation in the protocols used for microarray analysis. Other factors that could influence the accuracy of the meta-analysis include different technologies, different techniques and/or accuracy of microarray hybridization, and differences in the treatment of animals. Also, since the technique of a meta-analysis examines so many genes at once (12,328 in this case), there is a possibility that genes are upregulated or downregulated by random chance. Although we were stringent in the statistical analysis of this data, the random chance present in this study could be influencing gene expression and thus could be a source of error. Human error is of course a possibility in any experiment, and could be contributing to the discontinuity of the data. These factors could account for some of the inconsistencies observed in the data.

The trends in gene expression were verified in 2 of the 5 chosen genes via real-time PCR. We used a fourth animal model for PCR verification. As mentioned in the Methods section, each study used a different type of rat to examine gene expression. Our lab used Zucker Diabetic Fatty rats to verify the gene expression observed in the meta-analysis, so the genetic differences in the rats could be influential in their gene expression patterns.

Another topic worth noting is the role that hypertension could be playing in metabolic syndrome. In this study, it is not considered as a separate factor that can occur in MetS, but hypertension may actually have its own independent effects on kidney function, separate from the effects triggered by metabolic syndrome. All of the rat models used in the three chosen

39

studies exhibited high blood pressure, which show their link to hypertension. Since the rats used for PCR verification were not hypertensive, the effects of gene expression in the kidney can be soley attributed to metabolic syndrome pathogenesis, whereas in the studies used to gather the microarray data this distinction cannot necessarily be made. The microarray studies could show a bias towards the effects of hypertension on the kidneys since the rats used in all three studies exhibited high blood pressure, whereas the Zucker Diabetic Fatty rats used for PCR verification do not exhibit high blood pressure. Although the ZDF model was able to verify two genes from the meta-analysis, there could be some false negatives in the PCR data resulting from the effects of hypertension independently. Perhaps using a hypertensive diabetic rat model would ensure that all influential genes are PCR verified.

The two genes that were verified are Klk1c9 and Mgmt. Klk1c9, Kallikrein 1-related peptidase C9, encodes a protein that exhibits serine-type endopeptidase activity and is involved in the positive regulation of vasoconstriction. Klk1c9 was significantly downregulated in studies one and two, and followed the trend for downregulation in study three. Thus, this gene was expected to be downregulated, which agrees with the PCR results that Klk1c9 was expressed at 73.850% of the control.

In a study by Yamamuro, rats exposed to a tickling stimulus had increased expression of Klk1c9 and reduced amylase production (Yamamuro et al., 2013). The reduction of amylase by submandibular glands could have implications for the digestion of carbohydrates, which links this gene to metabolic dysfunction. According to the Rat Genome Database, this gene has also been associated with cardiovascular disease. Since CVD is a common complication of MetS and T2DM, this gene could be influential in the end-stage of MetS.

Another study identified a broad chromosome 1 region around Klk1 that was influential in regulating blood pressure levels. Since high blood pressure is one of the factors used to denote MetS in our project, it is not surprising that this gene was verified as being downregulated in the treatment rats compared to the control rats, which could mean that it is less active in controlling blood pressure in rats with metabolic syndrome (Yasui et al., 2007).

O-6-methylguanine-DNA transferase (Mgmt) was significantly upregulated in all three studies, which agrees with the preliminary PCR data that it is expressed 321.016 % of the control in the treatment. Mgmt, which encodes a protein that exhibits calcium ion binding and methylated-DNA-[protein]-cysteine S-methyltransferase activity. It is also involved in DNA dealkylation and DNA repair (Laulederkind et al., 2013). DNA repair is an important step in DNA replication, where mutations can be the cause of disease.

Mgmt could be influential in the pathology of MetS since obesity and metabolic dysfunction have a founded genetic component. The Rat Genome Database also cites that Mgmt is associated with neurologic disease. In many cases of MetS (especially those that evolve into T2DM) patients experience peripheral neuropathy, which is loss of sensation caused by damage to neurons. This links Mgmt to MetS since it is involved in DNA repair and neurologic disease, which can occur in the pathology of MetS.

Although the remaining genes trends were not verified in PCR, some of them could possibly still be contributing to metabolic syndrome pathogenesis. The general functions of the unverified genes will be briefly discussed, to give some context as to why they were chosen for this project.

Adenosine monophosphate deaminase 3 (Ampd3) was significantly upregulated in studies one and two, and followed the general trend for upregulation in the third study, which

was not supported by the PCR percentage of **+/-** 57.168 %. Ampd3 is a gene that encodes a protein that exhibits AMP deaminase activity, and is involved in AMP catabolism. It also participates in the purine metabolic pathway and has been previously associated with heart disease and stroke (Laulederkind et al., 2013).

RGD1309350 (similar to transthyretin (4L369)) was significantly downregulated in all three studies, but this was not supported by the PCR percentage of **+/-** 578.602 %. RGD1309350 encodes a protein that exhibits hydroxyisourate hydrolase activity and is involved in the purine nucleobase metabolic process (Laulederkind et al., 2013). This gene was chosen for its possible role in epigenetics.

Stk32c was significantly downregulated in studies one and three, and followed the trend for downregulation in study two, which did not agree with the PCR percentage of **+/-** 255.090 %. Stk32c encodes a protein that exhibits ATP binding and protein serine/threonine kinase activity (Laulederkind et al., 2013). The phosphorylation step could be influential in signaling cascades related to MetS. Stk32c was the most inconclusive gene with regards to the PCR data (gene expression was greatly varied among the 4 treatments), suggesting a possible error with the primer used for its amplification. Stk32c was expected to be downregulated, which disagrees with the average PCR data, but further PCR analysis in necessary to confirm this finding.

Based on the discontinuity of the PCR data with the statistical expectations, future work includes continuing to collect PCR data to verify the upregulation or downregulation trends. This project could be expanded is by changing the definition used to define and diagnose metabolic syndrome. Since the literature is so inconclusive as to the exact criteria needed to diagnose a person with MetS, there may be some value in tweaking the criteria used to include

studies in the meta-analysis. This could potentially result in more applicable studies that when included in the meta-analysis, may uncover more novel genes to investigate.

Since the gene expression data is still preliminary, a much larger pool of data may be necessary to verify the trends already observed through the PCR data. As future work reveals more consistent PCR data, a more conclusive statement about the individual role of these genes in the pathology of metabolic syndrome may provide influential insight into the etiology and pathology of this increasingly prevalent disorder.

Overall, the importance of this study was to find novel genes that relate to the pathologic effects of metabolic syndrome. Although preliminary PCR data was only able to verify the trends in gene expression for two of the genes, these genes could be investigated even further and eventually become the target of a new therapy or drug for metabolic syndrome or one of its debilitating complications. This study serves as preliminary research into possible genes involved in the pathogenesis of metabolic syndrome that can be continued in future work.

Appendices

Appendix 1

R Code used in the meta-analysis:
```
#study1
study1<-read.csv("N:/UserG-L/jahannx/Diabetes/UBM
data/TransformedandMergedStudy1.csv",header=F,na.strings="NA")
head(study1)
tail(study1)

study2<-read.csv("N:/UserG-L/jahannx/Diabetes/UBM
data/TransformedandMergedStudy2.csv",header=F,na.strings="NA")
head(study2)
tail(study2)

study5<-read.csv("N:/UserG-L/jahannx/Diabetes/UBM
data/TransformedandMergedStudy5.csv",header=F,na.strings="NA")
head(study5)
tail(study5)

n<-length(study1[,1])
n

Gene.ID<-study1[,1]
Gene.Symbol<-study1[,2]

#this should match for IDs and symbols in all 3 merged studies (since they're sorted)

#log transformation
col3<-study1[,3];z1<-log(col3,2)
col4<-study1[,4];z2<-log(col4,2)
col5<-study1[,5];z3<-log(col5,2)
col6<-study1[,6];z4<-log(col6,2)
col7<-study1[,7];z5<-log(col7,2)
col8<-study1[,8];z6<-log(col8,2)

qnorm_w <-(1/3)*(sort(z1)+sort(z2)+sort(z3))
qw1<- qnorm_w[rank(z1)]; qw2<- qnorm_w[rank(z2)]; qw3<- qnorm_w[rank(z3)]
qnorm_m<-(1/3)*(sort(z4)+sort(z5)+sort(z6))
qm1<- qnorm_m[rank(z4)]; qm2<- qnorm_m[rank(z5)]; qm3<- qnorm_m[rank(z6)]

###***Quantile**Normalization**code***********
```

```
qw <- cbind(z1, z2,z3);qm<-cbind(z4,z5,z6)
qz1<-apply(apply(qw,2,sort),1,mean);qz2<-apply(apply(qm,2,sort),1,mean)
n<-length(qw[,1])
qzw<-matrix(rep(0,3*n),n,3)
for(i in 1:3){
        qzw[,i]<-qz1[rank(qw[,i])]
}

qzm<-matrix(rep(0,3*n),n,3)
for(i in 1:3){
        qzm[,i]<-qz2[rank(qm[,i])]
}
zzn1<-data.frame(qzw,qzm)

###****produces **same **normalized**data**as**above****
zzn1<- cbind(qw1,qw2,qw3,qm1,qm2,qm3)
head(zzn1)
###***************************************************

# To use std dev correction proposed in SAM
se1<- sqrt(apply(zzn1,1,var)/6)
head(se1)
se11<- sort(se1)
s01<- quantile(se1,.90)
# To verify that 90th percentile is computed correctly
length(which (se11 > s01))

#m<-qzm
#w<-qzw

w<-zzn1[,1:3]; m<-zzn1[,4:6]

#get mean for each gene in mutant arrays
mean.m<-apply(m,1,mean)
#get mean for each gene in wild-type arrays
mean.w<-apply(w,1,mean)

#find difference between means for each gene
diff1<-mean.m-mean.w

#define how many mutant arrays
n.m=3
#define how many wild-type arrays
n.w=3

#get std. dev. for each gene in mutant arrays
```

```
s.m<-apply(m,1,sd)
#get std. dev. for each gene in wild-type arrays
s.w<-apply(w,1,sd)

a<-(1/n.m + 1/n.w)/(n.m+n.w-2)

#gene specific scatter is computed
si1<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

#gene relative difference between group means is computed (SAM method)
di1 <- diff1/(si1+s01)

#Weight for effect size meta-analysis
s1<-(1/n.m + 1/n.w + (di1)^2/(2*(n.m+n.w)))


#indiv1<-data.frame(Gene.Symbol, Gene.ID,qzw[,1],qzw[,2],qzw[,3],qzm[,1],qzm[,2],qzm[,3],
diff,di1, si1, s1)

indiv1<-data.frame(Gene.Symbol, Gene.ID,qw1,qw2,qw3,qm1,qm2,qm3, diff1,di1, si1, s1)
head(indiv1)


################################

#study 2
study2<-read.csv("N:/UserG-L/jahannx/Diabetes/UBM
data/TransformedandMergedStudy2.csv",header=F,na.strings="NA")
head(study2)

Gene.ID<-study2[,1]
Gene.Symbol<-study2[,2]

#log transformation
col3<-study2[,3];z1<-log(col3,2)
col4<-study2[,4];z2<-log(col4,2)
col5<-study2[,5];z3<-log(col5,2)
col6<-study2[,6];z4<-log(col6,2)
col7<-study2[,7];z5<-log(col7,2)
col8<-study2[,8];z6<-log(col8,2)

qnorm_w <-(1/3)*(sort(z1)+sort(z2)+sort(z3))
qw1<- qnorm_w[rank(z1)]; qw2<- qnorm_w[rank(z2)]; qw3<- qnorm_w[rank(z3)]
qnorm_m<-(1/3)*(sort(z4)+sort(z5)+sort(z6))
qm1<- qnorm_m[rank(z4)]; qm2<- qnorm_m[rank(z5)]; qm3<- qnorm_m[rank(z6)]
#zzn2<- data.frame(qw1,qw2,qw3,qm1,qm2,qm3)
```

```
zzn2<- cbind(qw1,qw2,qw3,qm1,qm2,qm3)

w<-zzn2[,1:3]; m<-zzn2[,4:6]
mean.m<-apply(m,1,mean)
mean.w<-apply(w,1,mean)

s.m<-apply(m,1,sd)
s.w<-apply(w,1,sd)

n.m<-3
n.w<-3
diff2<-mean.m-mean.w

a<-(1/n.m + 1/n.w)/(n.m+n.w-2)

#gene specific scatter is computed
si2<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

# To use std dev correction proposed in SAM
se1<- sqrt(apply(zzn2,1,var)/6)
head(se1)
se11<- sort(se1)
s02<- quantile(se11,.90)
# To verify that 90th percentile is computed correctly
length(which (se11 > s02))


#gene relative difference between group means is computed (SAM method)
di2 <- diff2/(si2+s02)

#Weight for effect size meta-analysis
s2<-(1/n.m + 1/n.w + (di2)^2/(2*(n.m+n.w)))

indiv2<-data.frame(Gene.Symbol, Gene.ID,qw1,qw2,qw3,qm1,qm2,qm3, diff2,di2, si2, s2)
head(indiv2)

#################################

study5<-read.csv("N:/UserG-L/jahannx/Diabetes/UBM
data/TransformedandMergedStudy5.csv",header=F,na.strings="NA")
head(study5)

Gene.ID<-study5[,1]
Gene.Symbol<-study5[,2]

#log transformation was not needed
```

```
col3<-study5[,3]
col4<-study5[,4]
col5<-study5[,5]
col6<-study5[,6]
col7<-study5[,7]
col8<-study5[,8]
col9<-study5[,9]
col10<-study5[,10]
col11<-study5[,11]
col12<-study5[,12]

zzn5<-cbind(col3,col4,col5,col6,col7,col8,col9,col10,col11,col12)

#specify groups
w<-zzn5[,1:4]; m<-zzn5[,5:10]

#get mean for each gene in mutant arrays
mean.m<-apply(m,1,mean)
#get mean for each gene in wild-type arrays
mean.w<-apply(w,1,mean)

#find difference between means for each gene
diff5<-mean.m-mean.w

#get std. dev. for each gene in mutant arrays
s.m<-apply(m,1,sd)
#get std. dev. for each gene in wild-type arrays
s.w<-apply(w,1,sd)

#define how many mutant arrays
n.m=6
#define how many wild-type arrays
n.w=4

# To use std dev correction proposed in SAM
se1<- sqrt(apply(zzn5,1,var)/(n.w+n.m))
head(se1)
se11<- sort(se1)
s05<- quantile(se11,.90)
# To verify that 90th percentile is computed correctly
length(which (se11 > s05))

a<-(1/n.m + 1/n.w)/(n.m+n.w-2)

#gene specific scatter is computed
```

```
si5<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

#gene relative difference between group menas is computed (SAM method)
di5 <- diff5/(si5+s05)

p=s05+si5[2]
diff5[2]/p

#Weight for effect size meta-analysis
s5<-(1/n.m + 1/n.w + (di5)^2/(2*(n.m+n.w)))

indiv5<-data.frame(Gene.Symbol, Gene.ID,col3,col4,col5,col6,col7,col8,col9,col10,col11,col12,
diff5,di5, si5, s5)
head(indiv5)




##############################################################################
all.new<-data.frame(Gene.Symbol,Gene.ID,diff1,di1,s1,diff2,di2,s2,diff5,di5,s5)


#write this data file into an Excel file for easy viewing
#write.csv(all.new,"N:/UserG-L/jahannx/Diabetes/UBM data/", row.names=T)

write.csv(all.new, file="allNew.csv")
getwd()

#Computing an average combined study effect size (meta-analysis effect size)
mu <- (di1*(1/s1) + di2*(1/s2)+di5*(1/s5))/(1/s1+1/s2+1/s5)
#Computing variance for the average combined study effect size
vmu <- (1/s1+1/s2+1/s5)/(1/s1+1/s2+1/s5)**2
#Computing meta z score
z0 <- mu/sqrt(vmu)

#Creating a file with meta-z's
meta.all<- data.frame(Gene.Symbol,Gene.ID,diff1,di1,diff2,di2,diff5,di5,z0)

write.csv(meta.all, file="metaAll.csv")
getwd()

meta.final<- data.frame(Gene.Symbol,Gene.ID,diff1,di1,diff2,di2,diff5,di5,z0, zdiff, zdiffabs)
write.csv(meta.final, file="meta.final.csv")
getwd()


plot(z0)
```

```
plot(x <- sort(z0), type="s", title="Distribution of Meta-Zs")

####################################################
#Different effect size standard error
####################################################

s11 = (si1+s01)**2
s22 = (si2+s02)**2
s55 = (si5+s05)**2
#Computing an average combined study effect size (meta-analysis effect size)
mu.n <- (di1*(1/s11) + di2*(1/s22)+di5*(1/s55))/(1/s11+1/s22+1/s55)
#Computing variance for the average combined study effect size
vmu.n <- (1/s11+1/s22+1/s55)/(1/s11+1/s22+1/s55)**2
#Computing meta z score
z0.n <- mu.n/sqrt(vmu.n)

#Creating a file with meta-z's
meta.n<- data.frame(Gene.Symbol,Gene.ID,diff1,di1,s11,diff2,di2,s22,diff5,di5,s55,z0.n)

write.csv(meta.n, file="metaN.csv")
getwd()
##########################################################
##########################################################


#Starting permutation
w1<-1:n
muF<-1:n
muVar<-1:n
b<-1:n
mB<-1:n
B <- 100


for(j in 1:B)
{
x <- zzn1[,sample (ncol(zzn1))]
w <- x[, 1:3]
m <- x[, 4:6]
n.m=3
n.w=3

mean.w<-apply(w,1,mean)
s.w<-apply(w,1,sd)
mean.m<-apply(m,1,mean)
s.m<-apply(m,1,sd)
```

```
a<-(1/n.m + 1/n.w)/(n.m+n.w-2)
#gene specific scatter is computed
sep<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

se<- sqrt(apply(x,1,var)/6)
se0<- sort(se)
s01<- quantile(se0,.90)

size1 <- ((mean.m-mean.w)/(sep+s01))
s1<- (1/n.m + 1/n.w + size1^2/(2*n.m+2*n.w))


###############################################################
x <- zzn2[,sample (ncol(zzn2))]
w <- x[, 1:3]
m <- x[, 4:6]
n.m=3
n.w=3
mean.w<-apply(w,1,mean)
s.w<-apply(w,1,sd)
mean.m<-apply(m,1,mean)
s.m<-apply(m,1,sd)

a<-(1/n.m + 1/n.w)/(n.m+n.w-2)
#gene specific scatter is computed
sep<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

se<- sqrt(apply(x,1,var)/6)
se0<- sort(se)
s0<- quantile(se0,.90)

size2 <- ((mean.m-mean.w)/(sep+s0))
s2<- (1/n.m + 1/n.w + size2^2/(2*n.m+2*n.w))

###############################################################
x <- zzn5[,sample (ncol(zzn5))]
w <- x[, 1:4]
m <- x[, 5:10]
n.w=4
n.m=6
mean.w<-apply(w,1,mean)
s.w<-apply(w,1,sd)
mean.m<-apply(m,1,mean)
s.m<-apply(m,1,sd)

a<-(1/n.m + 1/n.w)/(n.m+n.w-2)
```

```
#gene specific scatter is computed
sep<- sqrt( a*((n.m-1)*s.m^2+(n.w-1)*s.w^2))

se<- sqrt(apply(x,1,var)/10)
se0<- sort(se)
s0<- quantile(se0,.90)

size5 <- ((mean.m-mean.w)/(sep+s0))
s5<- (1/n.m + 1/n.w + size2^2/(2*n.m+2*n.w))

###############################################################################

all <- data.frame(Gene.ID, Gene.Symbol, size1,s1, size2, s2,size5, s5)

for(i in 1:n)
{
    w1[i] <- sum(1/all[i,4]+ 1/all[i,6] + 1/all[i,8] )
    muF[i]<- (sum(1/all[i,4]*all[i,3]+ 1/all[i,6]*all[i,5] + 1/all[i,8]*all[i,7] ))/w1[i]
    muVar[i] <- w1[i]/(w1[i])^2
        b[i]<-muF[i]/sqrt(muVar[i])
}
mB <- cbind(mB,b)
}

n1<-B+1
permZ<-data.frame(Gene.Symbol,Gene.ID,mB[,2:n1])

write.csv(permZ, file="permZ.csv")
getwd()


#write.csv(permZ,"/Users/keithwzirkle/Documents/UBM
DATA/100PermutedZ.csv",row.names=T)

#permZ<-read.csv("/Users/keithwzirkle/Documents/UBM DATA/100PermutedZ.csv")

zsum<-1:n
zbar<-1:n
zdiff<-1:n
zdiffabs<-1:n
for(i in 1:n)
{
        zsum[i]<-sum(permZ[i,4:102],na.rm=T)
        zbar[i]<-(1/B)*zsum[i]
        zdiff[i]<-z0[i]-zbar[i]
        zdiffabs[i]<-abs(zdiff[i])
```

```
}

quantile(zdiff,.99,na.rm=T)
quantile(zdiff,.01,na.rm=T)



perm.Z<-permZ[,3:102]

count<- 1:n
for(i in 1:n){
count[i]<-0
for(j in 1:B){
if (abs(z0[i]) > abs(perm.Z[i,j]))
count[i] = count[i]+1
}
}
countRate = count/B
nm<-length(which(countRate[i]<=0.05))

count<-0
for(j in 1:B)
{
for(i in 1:n)
{
nm1<-length(which(perm.Z[i,]<=t2))
nm2<-length(which(perm.Z[i,]>=t1))
count<-count+nm1+nm2
}
}
n2<-B+2



nm<-count/B

fdr<-nm/dn

meta.final<- data.frame(Gene.Symbol,Gene.ID,diff1,di1,diff2,di2,diff5,di5,z0, zdiff, zdiffabs)
write.csv(meta.final, file="meta.final.csv")
getwd()

meta.B<- data.frame(Gene.Symbol,Gene.ID,z0,zbar,countRate)
write.csv(meta.B, file="meta.B.csv")
getwd()
```

```
####################################################

#metaZs<-read.csv("/Users/keithwzirkle/Documents/UBM
DATA/MetaZs.csv",header=F,na.strings="NA")
#metaZ<-metaZs[,3]

metaZs<-read.csv("/Users/keithwzirkle/Documents/UBM DATA/MetaZs.csv")
 #metaZ<-metaZs[,4]

#length(metaZ)
[1] 12329

#metaZ<-Z.ID[,3]

#quantile(metaZ,.999,na.rm=T)
#quantile(metaZ,.01)

t1<-(5.048028)
t2<-(-6.151412)

#t1<-quantile(z.diff,.995,na.rm=T);t2<-quantile(z.diff,.015,na.rm=T)

dn1<-length(which(metaZ<=t2))
dn2<-length(which(metaZ>=t1))
dn<-dn1+dn2
```

Gene > Symbol	Gene > Name	Gene > Description	Average Fold Change	Meta-Z Statistic	Already Associated with MetS?
Abcb11			-0.597781208	-2.710864478	Yes
Abra	actin-binding Rho activating protein	ENCODES a protein that exhibits actin binding (ortholog) AND transcription coactivator activity (ortholog) AND INVOLVED IN positive regulation of Rho protein signal transduction (ortholog) AND positive regulation of sequence-specific DNA binding transcription factor activity (ortholog) AND positive regulation of transcription from RNA polymerase II promoter (ortholog) AND FOUND IN sarcomere AND actin cytoskeleton (ortholog) AND myofibril (ortholog) AND INTERACTS WITH C60 fullerene AND all-trans-retinoic acid AND ammonium chloride	0.980237889	1.89020657	
Ampd3	adenosine monophosphate deaminase 3	ENCODES a protein that exhibits AMP deaminase activity AND INVOLVED IN AMP catabolic process (inferred) AND IMP biosynthetic process (inferred) AND IMP salvage (inferred) AND PARTICIPATES IN purine metabolic pathway AND ASSOCIATED WITH Heart Diseases (ortholog) AND Muscular Diseases (ortholog) AND Stroke (ortholog) AND INTERACTS WITH DDE AND N-nitrosodiethylamine AND ammonium acetate	0.608729665	2.781628136	
Apoc1	apolipoprotein C-I	ENCODES a protein that exhibits fatty acid binding (ortholog) AND lipase inhibitor activity (ortholog) AND phospholipase inhibitor activity (ortholog) AND INVOLVED IN regulation of lipid transport AND cholesterol efflux (ortholog) AND cholesterol metabolic process (ortholog) AND PARTICIPATES IN altered lipoprotein metabolic pathway AND lipoprotein metabolic pathway AND ASSOCIATED WITH Alzheimer Disease (ortholog) AND Carcinoma Pancreatic Ductal (ortholog) AND Diabetes Mellitus Type 2 (ortholog) AND FOUND IN high-density lipoprotein particle AND very-low-density lipoprotein particle AND endoplasmic reticulum (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND 3-methylcholanthrene AND alpha-hexachlorocyclohexane	-0.675625342	-2.420277079	Yes
Arhgap 11a	Rho GTPase activating protein 11A	INVOLVED IN signal transduction (inferred) AND FOUND IN intracellular (inferred) AND INTERACTS WITH sodium dichromate AND (-)-epigallocatechin 3-gallate (ortholog) AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog)	0.731745946	2.227572701	
Athl1	ATH1 acid	ENCODES a protein that exhibits catalytic activity (inferred) AND INVOLVED IN carbohydrate metabolic	0.624919691	3.390794405	Yes

	trehalase-like 1 (yeast)	process (inferred) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND paracetamol (ortholog) AND propiconazole (ortholog)			
C1qb	complement component 1 q subcomponent B chain	ENCODES a protein that exhibits protein homodimerization activity (ortholog) AND INVOLVED IN aging AND brain development AND response to glucocorticoid stimulus AND PARTICIPATES IN Staphylococcus aureus infection pathway AND Trypanosoma cruzi infection pathway AND coagulation cascade pathway AND ASSOCIATED WITH Ocular Hypertension AND Retrograde Degeneration AND Spinal Cord Injuries AND FOUND IN complement component C1 complex AND INTERACTS WITH 1 3-dinitrobenzene AND 17alpha-ethynylestradiol AND 2 3 7 8-tetrachlorodibenzodioxine	0.610797	2.6019769	Yes
C4bpb	complement component 4 binding protein beta	INVOLVED IN complement activation classical pathway (inferred) AND innate immune response (inferred) AND PARTICIPATES IN forkhead class A signaling pathway AND coagulation cascade pathway AND complement system pathway AND FOUND IN extracellular region (inferred) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 3H-1 2-dithiole-3-thione AND N-nitrosodiethylamine	1.257901022	2.172353776	Yes
Casp4	caspase 4 apoptosis-related cysteine peptidase	ENCODES a protein that exhibits scaffold protein binding AND INVOLVED IN germ cell programmed cell death (ortholog) AND FOUND IN neuron projection AND neuronal cell body AND protein complex AND INTERACTS WITH 1 1 1-trichloro-2 2-bis(4-hydroxyphenyl)ethane AND 1 2 4-trimethylbenzene AND 2 3 7 8-tetrachlorodibenzodioxine	0.745985144	3.58812773	
Ccdc80	coiled-coil domain containing 80	ENCODES a protein that exhibits fibronectin binding (ortholog) AND glycosaminoglycan binding (ortholog) AND heparin binding (ortholog) AND INVOLVED IN extracellular matrix organization (ortholog) AND positive regulation of cell-substrate adhesion (ortholog) AND FOUND IN basement membrane (ortholog) AND extracellular matrix (ortholog) AND interstitial matrix (ortholog) AND INTERACTS WITH diuron AND 17alpha-ethynylestradiol (ortholog) AND 17beta-estradiol (ortholog)	0.600693747	2.32826001	
Ccl2	chemokine (C-C motif) ligand 2	ENCODES a protein that exhibits CCR2 chemokine receptor binding AND chemokine activity AND heparin binding AND INVOLVED IN aging AND cellular calcium ion homeostasis AND chemokine-mediated signaling pathway AND PARTICIPATES IN angiotensin II signaling pathway AND vascular endothelial growth factor signaling pathway AND granulocyte-macrophage colony-stimulating factor signaling pathway AND ASSOCIATED WITH Acute	0.75498848	2.4415923	Yes

		Lung Injury AND Anti-Glomerular Basement Membrane Disease AND Aortic Aneurysm Abdominal AND FOUND IN cytoplasm AND extracellular space AND neuronal cell body AND INTERACTS WITH ( )-pilocarpine AND (3S 5R)-fluvastatin AND (S)-colchicine			
Ccl20	chemokine (C-C motif) ligand 20	ENCODES a protein that exhibits chemokine activity AND cytokine activity (ortholog) AND INVOLVED IN chemokinesis AND PARTICIPATES IN chemokine mediated signaling pathway AND cytokine mediated signaling pathway AND rheumatoid arthritis pathway AND FOUND IN extracellular space (inferred) AND INTERACTS WITH 1 2 4-trimethylbenzene AND 2-amino-2-deoxy-D-glucopyranose AND aldehydo-D-glucosamine	0.617613332	3.002862362	Yes
Ccna2	cyclin A2	ENCODES a protein that exhibits protein kinase binding (ortholog) AND INVOLVED IN organ regeneration AND positive regulation of fibroblast proliferation AND response to estradiol stimulus AND PARTICIPATES IN p53 signaling pathway AND cell cycle pathway mitotic AND ASSOCIATED WITH Carcinoma Embryonal (ortholog) AND Heart Failure (ortholog) AND Mammary Neoplasms Experimental (ortholog) AND FOUND IN cytoplasm (ortholog) AND female pronucleus (ortholog) AND male pronucleus (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 2-acetamidofluorene AND 3 3' 4 4' 5-pentachlorobiphenyl	0.811236924	1.995412127	Yes
Cep78	centrosomal protein 78	FOUND IN centrosome (ortholog) AND INTERACTS WITH 2 4-dinitrotoluene AND 2 6-dinitrotoluene AND ammonium chloride	-0.829944001	-3.734428671	
Ckap2	cytoskeleton associated protein 2	INVOLVED IN mitotic cytokinesis (ortholog) AND negative regulation of microtubule depolymerization (ortholog) AND positive regulation of transcription from RNA polymerase II promoter (ortholog) AND FOUND IN centrosome (ortholog) AND cytoplasm (ortholog) AND cytoplasmic microtubule (ortholog) AND INTERACTS WITH 2-amino-2-deoxy-D-glucopyranose AND aldehydo-D-glucosamine AND beta-D-glucosamine	1.038361252	2.405706833	
Clcf1	cardiotrophin-like cytokine factor 1	ENCODES a protein that exhibits ciliary neurotrophic factor receptor binding (ortholog) AND cytokine activity (ortholog) AND protein heterodimerization activity (ortholog) AND INVOLVED IN B cell differentiation (ortholog) AND JAK-STAT cascade (ortholog) AND cell surface receptor signaling pathway (ortholog) AND PARTICIPATES IN Jak-Stat signaling pathway AND cytokine mediated signaling pathway AND ASSOCIATED WITH Cold-Induced Sweating Syndrome 1 (ortholog) AND FOUND IN CNTFR-	1.205824058	2.371990249	Yes

		CLCF1 complex (ortholog) AND CRLF-CLCF1 complex (ortholog) AND extracellular region (ortholog) AND INTERACTS WITH 2 4-dinitrotoluene AND 2 6-dinitrotoluene AND C60 fullerene			
Clu	clusterin	ENCODES a protein that exhibits ATPase activity (ortholog) AND misfolded protein binding (ortholog) AND ubiquitin protein ligase binding (ortholog) AND INVOLVED IN cellular response to growth factor stimulus AND endocrine pancreas development AND negative regulation of apoptotic process AND ASSOCIATED WITH Brain Ischemia AND Ischemia AND Arteriosclerosis (ortholog) AND FOUND IN aggresome AND growth cone AND neuron projection AND INTERACTS WITH 1-naphthyl isothiocyanate AND 17alpha-ethynylestradiol AND 17beta-estradiol	1.084572253	2.102581467	Yes
Cmtm3	CKLF-like MARVEL transmembrane domain containing 3	FOUND IN integral to membrane (inferred) AND INTERACTS WITH cisplatin AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND benzo[a]pyrene (ortholog)	0.615966112	2.093043498	
Col8a1	collagen type VIII alpha 1	INVOLVED IN camera-type eye morphogenesis (ortholog) AND epithelial cell proliferation (ortholog) AND positive regulation of cell-substrate adhesion (ortholog) AND PARTICIPATES IN syndecan signaling pathway AND FOUND IN extracellular matrix (ortholog) AND INTERACTS WITH 2-amino-2-deoxy-D-glucopyranose AND C60 fullerene AND aldehydo-D-glucosamine	0.712786297	2.384292096	Yes
Ctu1	cytosolic thiouridylase subunit 1	ENCODES a protein that exhibits tRNA binding (ortholog) AND INVOLVED IN tRNA thio-modification (ortholog) AND tRNA wobble uridine modification (ortholog) AND FOUND IN cytosol (ortholog) AND mitochondrion (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND benzo[a]pyrene (ortholog) AND copper(2) sulfate (ortholog)	1.175427262	2.495972515	
Cxcl10	chemokine (C-X-C motif) ligand 10	ENCODES a protein that exhibits chemokine activity AND INVOLVED IN cellular response to heat AND immune response AND negative regulation of angiogenesis AND PARTICIPATES IN Retinoic acid-inducible gene (RIG) I-like receptor signaling pathway AND Toll-like receptor signaling pathway AND chemokine mediated signaling pathway AND ASSOCIATED WITH Brain Ischemia AND Bronchiolitis Viral AND Diabetes Mellitus Type 1 AND FOUND IN extracellular space AND external side	0.599933141	2.408571619	Yes

		of plasma membrane (ortholog) AND extracellular region (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 2 4 6-trinitrotoluene AND 2 4-dinitrotoluene			
Cyp2c1 1	cytochro me P450 subfamily 2 polypepti de 11	ENCODES a protein that exhibits testosterone 16-alpha-hydroxylase activity AND caffeine oxidase activity (ortholog) AND drug binding (ortholog) AND INVOLVED IN cellular amide metabolic process (ortholog) AND drug catabolic process (ortholog) AND drug metabolic process (ortholog) AND PARTICIPATES IN arachidonic acid metabolic pathway AND linoleic acid metabolic pathway AND phase I biotransformation pathway via cytochrome P450 AND ASSOCIATED WITH Depressive Disorder AND Hypertension AND Kidney Failure Chronic AND FOUND IN intracellular membrane-bounded organelle (ortholog) AND INTERACTS WITH 1 1-dichloroethene AND 1 2-dichloroethene AND 17alpha-ethynylestradiol	-2.145074413	-5.115602983	Yes
Dclk1	doublecor tin-like kinase 1	ENCODES a protein that exhibits protein kinase activity AND INVOLVED IN protein phosphorylation AND INTERACTS WITH C60 fullerene AND N-nitrosodiethylamine AND ammonium chloride	0.656868266	2.538316659	
Dscc1	DNA replicatio n and sister chromatid cohesion 1	INVOLVED IN maintenance of mitotic sister chromatid cohesion (ortholog) AND post-translational protein acetylation (ortholog) AND regulation of DNA replication (ortholog) AND FOUND IN chromatin (ortholog) AND chromosome centromeric region (ortholog) AND nucleoplasm (ortholog) AND INTERACTS WITH cefaloridine AND cisplatin AND coumarin	0.973403399	2.41296766	
Eml2	echinoder m microtub ule associate d protein like 2	ENCODES a protein that exhibits protein C-terminus binding AND receptor binding AND FOUND IN intracellular membrane-bounded organelle (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 2 4-dinitrotoluene AND ammonium chloride	0.688631322	3.047660641	
Eps8l3	EPS8-like 3	INTERACTS WITH 17alpha-ethynylestradiol AND bis(2-ethylhexyl) phthalate AND bisphenol A	1.61228436	2.874181705	
Fes	feline sarcoma oncogene	ENCODES a protein that exhibits immunoglobulin receptor binding (ortholog) AND non-membrane spanning protein tyrosine kinase activity (ortholog) AND phosphatidylinositol binding (ortholog) AND INVOLVED IN peptidyl-tyrosine phosphorylation (ortholog) AND positive regulation of actin cytoskeleton reorganization (ortholog) AND positive regulation of microtubule polymerization (ortholog) AND PARTICIPATES IN interleukin-4 signaling pathway AND vascular endothelial growth factor signaling pathway AND FOUND IN cytoplasm (ortholog) AND	0.764013198	1.901805644	Yes

		cytoplasmic vesicle (ortholog) AND extrinsic to internal side of plasma membrane (ortholog) AND INTERACTS WITH C60 fullerene AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND N-methyl-N'-nitro-N-nitrosoguanidine (ortholog)			
Fetub	fetuin B	ENCODES a protein that exhibits metalloendopeptidase inhibitor activity (ortholog) AND INVOLVED IN binding of sperm to zona pellucida (ortholog) AND negative regulation of endopeptidase activity (ortholog) AND single fertilization (ortholog) AND FOUND IN extracellular space (inferred) AND INTERACTS WITH ammonium chloride AND cadmium dichloride AND clofibrate	0.87169286	4.57425494	
Gkn2	gastrokine 2	FOUND IN extracellular region (inferred) AND INTERACTS WITH N-acetyl-L-cysteine AND N-nitrosodiethylamine AND diuron	1.085770069	2.104517196	
Grem1	gremlin 1	ENCODES a protein that exhibits protein tyrosine kinase activator activity (ortholog) AND receptor agonist activity (ortholog) AND vascular endothelial growth factor receptor 2 binding (ortholog) AND INVOLVED IN apoptotic process AND negative regulation of cell growth AND negative regulation of leukocyte chemotaxis AND PARTICIPATES IN Bone morphogenetic proteins signaling pathway AND FOUND IN cell surface AND extracellular space AND INTERACTS WITH 1 1 1-trichloro-2 2-bis(4-hydroxyphenyl)ethane AND 1 2 4-trimethylbenzene AND 2-amino-2-deoxy-D-glucopyranose	0.634975734	2.801365652	Yes
Grhl3	grainyhead-like 3 (Drosophila)	ENCODES a protein that exhibits DNA binding (ortholog) AND sequence-specific DNA binding RNA polymerase II transcription factor activity (ortholog) AND INVOLVED IN central nervous system development (ortholog) AND cochlea morphogenesis (ortholog) AND ectoderm development (ortholog) AND FOUND IN nucleus (ortholog) AND INTERACTS WITH diuron AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND 5-fluorouracil (ortholog)	-0.834607587	-2.074108985	
Havcr1	hepatitis A virus cellular receptor 1	INVOLVED IN phagocytosis engulfment AND response to antibiotic AND response to drug AND ASSOCIATED WITH Acute Kidney Injury AND Cardio-Renal Syndrome AND Colitis AND FOUND IN apical plasma membrane AND cell surface AND cytoplasm AND INTERACTS WITH 1-naphthyl isothiocyanate AND 11-deoxycorticosterone AND 4 4'-diaminodiphenylmethane	2.11505311	2.284800851	Yes
Hfe2	hemochromatosis type 2 (juvenile)	ENCODES a protein that exhibits coreceptor activity (ortholog) AND INVOLVED IN BMP signaling pathway (ortholog) AND iron ion homeostasis (ortholog) AND positive regulation of transcription from RNA polymerase II promoter (ortholog) AND	1.063908487	3.71900411	Yes

		PARTICIPATES IN Bone morphogenetic proteins signaling pathway AND ASSOCIATED WITH Hemochromatosis (ortholog) AND FOUND IN basolateral plasma membrane AND cell surface (ortholog) AND extracellular space (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND C60 fullerene AND cefaloridine			
Hsd3b6	hydroxy-delta-5-steroid dehydrog enase, 3 beta- and steroid delta-isomerase 1	ENCODES a protein that exhibits 3-beta-hydroxy-delta5-steroid dehydrogenase activity; steroid delta-isomerase activity; INVOLVED IN androgen biosynthetic process; C21-steroid hormone biosynthetic process; hippocampus development; PARTICIPATES IN cortisol biosynthetic pathway; estradiol biosynthetic pathway; steroid hormone biosynthetic pathway; ASSOCIATED WITH Hyperprolactinemia; Hypertension; FOUND IN mitochondrial inner membrane (ortholog); mitochondrial intermembrane space (ortholog); INTERACTS WITH (-)-citrinin; 1,1,1-trichloro-2,2-bis(4-hydroxyphenyl)ethane; 17beta-estradiol	-1.39076549	-2.088623506	Yes
Jrk	jerky homolog (mouse)	ENCODES a protein that exhibits mRNA binding (ortholog) AND FOUND IN cytoplasm (ortholog) AND nucleus (ortholog) AND ribonucleoprotein complex (ortholog) AND INTERACTS WITH L-methionine (ortholog) AND bisphenol A (ortholog) AND choline (ortholog)	0.630346942	2.54592418	
Kcnip3	Kv channel interactin g protein 3 calsenilin	ENCODES a protein that exhibits ion channel binding AND potassium channel regulator activity AND sequence-specific DNA binding AND INVOLVED IN intracellular protein transport AND behavioral response to pain (ortholog) AND negative regulation of transcription from RNA polymerase II promoter (ortholog) AND PARTICIPATES IN calcium/calcium-mediated signaling pathway AND FOUND IN axon AND axon terminus AND dendrite AND INTERACTS WITH ammonium chloride AND tetrachloromethane AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog)	0.643793895	2.721307663	Yes
Klk1c9	kallikrein 1-related peptidase C9	ENCODES a protein that exhibits serine-type endopeptidase activity (inferred) AND INVOLVED IN positive regulation of vasoconstriction AND INTERACTS WITH cyclosporin A AND sirolimus AND tacrine	-1.435727195	-3.475117881	
Knstrn	kinetocho re-localized astrin/SP AG5 binding protein	INVOLVED IN chromosome segregation (ortholog) AND mitotic anaphase (ortholog) AND mitotic sister chromatid segregation (ortholog) AND FOUND IN kinetochore (ortholog) AND microtubule plus end (ortholog) AND mitotic spindle (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND N-nitrosodiethylamine AND cefaloridine	1.037466529	2.451820757	
Kntc1	kinetocho re	INVOLVED IN mitotic cell cycle checkpoint (ortholog) AND FOUND IN kinetochore microtubule (ortholog)	1.11972911	2.774124549	

	associated 1	AND spindle pole (ortholog) AND INTERACTS WITH cefaloridine AND (-)-epigallocatechin 3-gallate (ortholog) AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog)			
Krt25	keratin 25	ENCODES a protein that exhibits structural molecule activity (inferred) AND INVOLVED IN hair follicle morphogenesis (ortholog) AND intermediate filament organization (ortholog) AND FOUND IN cytoplasm (inferred) AND intermediate filament (inferred) AND INTERACTS WITH L-methionine (ortholog) AND choline (ortholog) AND dimethylarsinic acid (ortholog)	0.772193878	2.108727634	
Lgals3bp	lectin galactoside-binding soluble 3 binding protein	ENCODES a protein that exhibits scavenger receptor activity (inferred) AND INVOLVED IN cell adhesion (inferred) AND FOUND IN extracellular matrix (ortholog) AND extracellular vesicular exosome (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND 17beta-estradiol AND 2 3 7 8-tetrachlorodibenzodioxine	0.590089574	2.192901147	
LOC290595	hypothetical gene supported by AF152002	null	0.868616408	2.938463796	
LOC378467			0.720173027	4.09776505	
LOC500118	family with sequence similarity 221, member A	INTERACTS WITH fonofos (ortholog); parathion (ortholog); terbufos (ortholog)	1.014978436	2.472776597	
LOC654482	hypothetical protein LOC654482	INTERACTS WITH arsenic atom (ortholog) AND hydralazine (ortholog) AND valproic acid (ortholog)	0.658975121	2.321353547	
Loc691979			0.822319504	2.797529061	
Ltbp2	latent transforming growth factor beta binding protein 2	ENCODES a protein that exhibits calcium ion binding (inferred) AND growth factor binding (inferred) AND heparin binding (inferred) AND ASSOCIATED WITH Glaucoma 3 Primary Congenital D (ortholog) AND Tooth Agenesis Selective 6 (ortholog) AND FOUND IN extracellular matrix (ortholog) AND extracellular space (ortholog) AND extracellular vesicular exosome (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND N-nitrosodiethylamine AND ammonium chloride	0.867851363	2.42845486	Yes

Lum	lumican	ENCODES a protein that exhibits collagen binding (ortholog) AND INVOLVED IN cartilage development AND response to growth factor stimulus AND response to organic cyclic compound AND ASSOCIATED WITH Fibrosis AND Myocardial Reperfusion Injury AND Spinal Cord Injuries AND FOUND IN proteinaceous extracellular matrix AND extracellular matrix (ortholog) AND extracellular space (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND N-ethyl-N-nitrosourea AND ammonium chloride	0.72153317	2.192195487	Yes
Mfap5	microfibrillar associated protein 5	INVOLVED IN extracellular fibril organization (ortholog) AND PARTICIPATES IN Notch signaling pathway AND FOUND IN extracellular matrix (ortholog) AND INTERACTS WITH dexamethasone AND diuron AND perfluorooctane-1-sulfonic acid	0.991854272	2.682661968	Yes
Mgmt	O-6-methylguanine-DNA methyltransferase	ENCODES a protein that exhibits calcium ion binding AND methylated-DNA-[protein]-cysteine S-methyltransferase activity AND methyltransferase activity (ortholog) AND INVOLVED IN DNA dealkylation involved in DNA repair AND cellular response to ionizing radiation AND cellular response to organic cyclic compound AND ASSOCIATED WITH Colonic Neoplasms AND Melanoma AND Bile Duct Neoplasms (ortholog) AND FOUND IN nucleus AND INTERACTS WITH ( )-pilocarpine AND 2 3 7 8-tetrachlorodibenzodioxine AND 2 4-dinitrotoluene	0.879915656	3.310127963	
Mlc1	megalencephalic leukoencephalopathy with subcortical cysts 1	ENCODES a protein that exhibits protein transporter activity AND protein complex binding (ortholog) AND INVOLVED IN caveolin-mediated endocytosis AND cellular response to cholesterol AND protein transport AND ASSOCIATED WITH Megalencephalic leukoencephalopathy with subcortical cysts (ortholog) AND FOUND IN caveola AND clathrin-coated vesicle AND cytoplasmic vesicle AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 2 4-dinitrotoluene AND C60 fullerene	-0.741371495	-3.638487407	Yes
Mrgprg	MAS-related GPR member G	ENCODES a protein that exhibits G-protein coupled receptor activity (inferred) AND FOUND IN integral to membrane (inferred) AND plasma membrane (inferred) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine (ortholog)	-1.074920272	-2.062276727	
Nell1	NEL-like 1 (chicken)	ENCODES a protein that exhibits heparin binding AND identical protein binding AND protein kinase C binding AND INVOLVED IN induction of apoptosis AND positive regulation of ossification AND protein homotrimerization AND ASSOCIATED WITH abnormal craniofacial bone morphology AND ASSOCIATED WITH Craniosynostoses AND FOUND IN cytoplasm AND extracellular space AND nuclear envelope (ortholog) AND INTERACTS WITH 17alpha-	-0.706297723	-2.199570686	

		ethynylestradiol AND ammonium chloride AND dichlorine			
Nipsnap3a	nipsnap homolog 3A (C. elegans)	FOUND IN mitochondrion (ortholog) AND INTERACTS WITH 17alpha-ethynylestradiol AND 2-methoxyethanol AND N-nitrosodiethylamine	0.691604782	2.930527011	
Pcp4l1	Purkinje cell protein 4-like 1	INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND N-nitrosodiethylamine AND benzo[a]pyrene	0.640118534	2.134991422	
Pemt			0.816829673	2.843496686	Yes
Pmf1	polyamine-modulated factor 1	ENCODES a protein that exhibits leucine zipper domain binding (ortholog) AND transcription coactivator activity (ortholog) AND INVOLVED IN chromosome segregation (ortholog) AND FOUND IN MIS12/MIND type complex (ortholog) AND INTERACTS WITH cisplatin AND 2-methylcholine (ortholog) AND 4 4'-diaminodiphenylmethane (ortholog)	0.679702364	2.589740252	
Pqlc3	PQ loop repeat containing 3	INTERACTS WITH N-nitrosodiethylamine AND all-trans-retinoic acid AND diuron	0.709869718	2.307354938	
Prc1	protein regulator of cytokinesis 1	ENCODES a protein that exhibits identical protein binding (ortholog) AND protein kinase binding (ortholog) AND INVOLVED IN cytokinesis (ortholog) AND FOUND IN cytoplasm (ortholog) AND microtubule cytoskeleton (ortholog) AND nucleus (ortholog) AND INTERACTS WITH 2-acetamidofluorene AND 2-amino-2-deoxy-D-glucopyranose AND aldehydo-D-glucosamine	0.799037232	3.684119167	
Prl3d1	Prolactin family 3 subfamily d member 1	ENCODES a protein that exhibits hormone activity (inferred) AND FOUND IN extracellular region (inferred) AND INTERACTS WITH ammonium chloride AND bisphenol A	1.027358136	2.033813612	
Psmb9	proteasome (prosome macropain) subunit beta type 9	ENCODES a protein that exhibits threonine-type endopeptidase activity (inferred) AND INVOLVED IN antigen processing and presentation (ortholog) AND PARTICIPATES IN ubiquitin/proteasome degradation pathway AND ASSOCIATED WITH Colorectal Neoplasms (ortholog) AND Diabetes Mellitus Type 1 (ortholog) AND Esophageal Neoplasms (ortholog) AND FOUND IN proteasome core complex (ortholog) AND spermatoproteasome complex (ortholog) AND INTERACTS WITH 1 3-dinitrobenzene AND 17alpha-ethynylestradiol AND 2 3 7 8-tetrachlorodibenzodioxine	0.658897547	2.277446527	Yes
Ptgs2	prostaglandin-endopero	ENCODES a protein that exhibits lipid binding AND oxidoreductase activity acting on single donors with incorporation of molecular oxygen incorporation of two	0.805086205	2.312262692	Yes

	xide synthase 2	atoms of oxygen AND enzyme binding (ortholog) AND INVOLVED IN angiogenesis AND bone mineralization AND cellular response to ATP AND PARTICIPATES IN sphingosine 1-phosphate signaling pathway AND Leishmania infection pathway AND arachidonic acid metabolic pathway AND ASSOCIATED WITH Acute Kidney Injury AND Acute Lung Injury AND Adenocarcinoma AND FOUND IN caveola AND cytoplasm AND protein complex AND INTERACTS WITH (-)-epigallocatechin 3-gallate AND (R)-lipoic acid AND (R)-noradrenaline			
Pthlh	parathyroid hormone-like hormone	ENCODES a protein that exhibits hormone activity AND peptide hormone receptor binding (ortholog) AND INVOLVED IN adenylate cyclase-activating G-protein coupled receptor signaling pathway AND endochondral ossification (ortholog) AND endoderm development (ortholog) AND PARTICIPATES IN Hedgehog signaling pathway AND ASSOCIATED WITH BRACHYDACTYLY TYPE E2 (ortholog) AND Hypercalcemia (ortholog) AND FOUND IN Golgi apparatus (ortholog) AND cytoplasm (ortholog) AND intracellular (ortholog) AND INTERACTS WITH L-ascorbic acid AND all-trans-retinoic acid AND candesartan	1.4156472	3.142766628	Yes
Ptprz1	protein tyrosine phosphatase receptor-type Z polypeptide 1	ENCODES a protein that exhibits protein tyrosine phosphatase activity (inferred) AND INVOLVED IN axonogenesis (ortholog) AND ASSOCIATED WITH Helicobacter Infections (ortholog) AND FOUND IN perineuronal net (ortholog) AND proteinaceous extracellular matrix (ortholog) AND INTERACTS WITH ( )-pilocarpine AND ammonium chloride AND cisplatin	0.709466161	3.012278667	
Rab19	RAB19 member RAS oncogene family	ENCODES a protein that exhibits GTP binding (inferred) AND INVOLVED IN protein transport (inferred) AND small GTPase mediated signal transduction (inferred) AND FOUND IN plasma membrane (inferred) AND INTERACTS WITH C60 fullerene AND rosiglitazone AND thioacetamide	0.643823425	2.360917141	
Rbm11	RNA binding motif protein 11	ENCODES a protein that exhibits nucleic acid binding (inferred) AND nucleotide binding (inferred)	0.848448999	2.360917141	
Rbp4	retinol binding protein 4 plasma	ENCODES a protein that exhibits protein heterodimerization activity AND retinol binding AND retinol transporter activity AND INVOLVED IN response to ethanol AND retinol metabolic process AND retinol transport AND PARTICIPATES IN retinoic acid metabolic pathway AND ASSOCIATED WITH Diabetes Mellitus Type 2 AND Obesity AND Avitaminosis (ortholog) AND FOUND IN extracellular	-1.066787632	-3.21934866	Yes

		space AND protein complex AND INTERACTS WITH 17alpha-ethynylestradiol AND 2 3 7 8-tetrachlorodibenzodioxine AND N-methyl-N-nitrosurea			
Rftn2	raftlin family member 2	INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 3H-1 2-dithiole-3-thione AND benzo[a]pyrene (ortholog)	0.663521893	1.927202011	
RGD13 07621	hypotheti cal LOC3141 68	INTERACTS WITH dibutyl phthalate AND furan AND copper atom (ortholog)	0.611118691	2.394147272	
RGD13 09350	similar to transthyre tin (4L369)	ENCODES a protein that exhibits hydroxyisourate hydrolase activity (inferred) AND INVOLVED IN purine nucleobase metabolic process (inferred) AND transport (inferred) AND PARTICIPATES IN purine metabolic pathway AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND mercaptopurine AND purine-6-thiol	-1.240507347	-4.159261603	
RGD13 11558	similar to 4930506 M07Rik protein	ENCODES a protein that exhibits kinase binding (inferred) AND INVOLVED IN axonogenesis AND FOUND IN axon (inferred) AND cytosol (inferred) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND (-)-epigallocatechin 3-gallate (ortholog) AND benzo[a]pyrene (ortholog)	0.643903769	2.319702918	
RGD15 62844	similar to serine (or cysteine) proteinas e inhibitor clade B member 9	INTERACTS WITH cisplatin	1.668952038	3.066907682	
RGD15 63091	similar to OEF2	ASSOCIATED WITH Tumoral Calcinosis Normophosphatemic Familial (ortholog) AND FOUND IN cytoplasm (ortholog) AND intracellular membrane-bounded organelle (ortholog) AND INTERACTS WITH chloroprene AND diethylstilbestrol AND (-)-epigallocatechin 3-gallate (ortholog)	0.740667222	2.387430322	Yes
RGD15 65709	similar to ovostatin-2	ENCODES a protein that exhibits endopeptidase inhibitor activity (inferred) AND FOUND IN extracellular space (inferred) AND INTERACTS WITH indole-3-methanol AND sodium dichromate	1.0581821	2.101246609	
Rnf40	ring finger protein 40 E3 ubiquitin protein ligase	ENCODES a protein that exhibits protein complex binding AND syntaxin-1 binding AND ubiquitin-protein ligase activity AND INVOLVED IN ubiquitin-dependent protein catabolic process AND histone H2B ubiquitination (ortholog) AND histone monoubiquitination (ortholog) AND FOUND IN neuron projection AND protein complex AND HULC complex (ortholog) AND INTERACTS WITH 2 3 7 8-tetrachlorodibenzodioxine AND 2 4-dinitrotoluene AND	-0.605638319	-2.577830214	

		ammonium chloride			
Rp1h	cylindro matosis (turban tumor syndrome)	ENCODES a protein that exhibits Lys63-specific deubiquitinase activity (ortholog); proline-rich region binding (ortholog); protein kinase binding (ortholog); INVOLVED IN necroptotic process (ortholog); negative regulation of canonical Wnt signaling pathway (ortholog); negative regulation of NF-kappaB import into nucleus (ortholog); PARTICIPATES IN nuclear factor kappa B signaling pathway; tumor necrosis factor mediated signaling pathway; Retinoic acid-inducible gene (RIG) I-like receptor signaling pathway; ASSOCIATED WITH Familial cylindromatosis (ortholog); Skin Neoplasms (ortholog); Trichoepithelioma multiple familial (ortholog); FOUND IN centrosome (ortholog); cytoplasmic microtubule (ortholog); cytosol (ortholog); INTERACTS WITH furan; methapyrilene; valproic acid	-1.032160361	-2.06854147	Yes
Rrm2	ribonucle otide reductase M2	ENCODES a protein that exhibits ribonucleoside-diphosphate reductase activity thioredoxin disulfide as acceptor (ortholog) AND INVOLVED IN mitotic cell cycle AND positive regulation of cell proliferation AND pyrimidine nucleobase metabolic process AND PARTICIPATES IN glutathione metabolic pathway AND p53 signaling pathway AND purine metabolic pathway AND ASSOCIATED WITH Endometrial Neoplasms AND Urinary Bladder Neoplasms AND Breast Neoplasms (ortholog) AND FOUND IN nuclear envelope AND INTERACTS WITH 17alpha-ethynylestradiol AND N-nitrosodiethylamine AND cefaloridine	0.662952702	3.042839946	Yes
RT1-Db1	RT1 class II locus Db1	ENCODES a protein that exhibits peptide antigen binding (ortholog) AND INVOLVED IN T-helper 1 type immune response (ortholog) AND antigen processing and presentation of exogenous peptide antigen via MHC class II (ortholog) AND detection of bacterium (ortholog) AND PARTICIPATES IN interleukin-12 signaling pathway AND Leishmania infection pathway AND Mycobacterium tuberculosis infection pathway AND ASSOCIATED WITH Addison Disease (ortholog) AND Alzheimer Disease (ortholog) AND Anti-Glomerular Basement Membrane Disease (ortholog) AND FOUND IN external side of plasma membrane (ortholog) AND late endosome membrane (ortholog) AND lysosomal membrane (ortholog) AND INTERACTS WITH 3H-1 2-dithiole-3-thione AND N-nitrosodiethylamine AND amitriptyline	0.879445692	2.467619932	Yes
Sass6	spindle assembly 6 homolog	INVOLVED IN centriole replication (ortholog) AND centrosome duplication (ortholog) AND FOUND IN centriole (ortholog) AND centrosome (ortholog) AND INTERACTS WITH (-)-epigallocatechin 3-gallate	0.694572707	2.243878548	

	(C. elegans)	(ortholog) AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog) AND 4-hydroxynon-2-enal (ortholog)			
Scg2	secretogr anin II	ENCODES a protein that exhibits chemoattractant activity (ortholog) AND cytokine activity (ortholog) AND INVOLVED IN MAPK cascade (ortholog) AND angiogenesis (ortholog) AND cellular component movement (ortholog) AND FOUND IN secretory granule AND extracellular space (ortholog) AND INTERACTS WITH (S)-nicotine AND 1 3-dinitrobenzene AND Delta(9)-tetrahydrocannabinol	1.855872057	2.096114031	
Scgb1c 1	secretogl obin family 1C member 1	FOUND IN extracellular region (inferred) AND INTERACTS WITH 17alpha-ethynylestradiol AND bisphenol A AND genistein	-1.096292096	-3.920606626	
Serpinc 1	serpin peptidase inhibitor clade C (antithro mbin) member 1	ENCODES a protein that exhibits heparin binding AND serine-type endopeptidase inhibitor activity AND protease binding (ortholog) AND INVOLVED IN negative regulation of inflammatory response AND response to nutrient AND PARTICIPATES IN glypican signaling pathway AND coagulation cascade pathway AND complement system pathway AND ASSOCIATED WITH Endotoxemia AND Inflammation AND Jaundice AND FOUND IN extracellular space AND INTERACTS WITH 2 4-dinitrotoluene AND 2 6-dinitrotoluene AND paracetamol	-0.824421779	-2.318811156	Yes
Slc15a3	solute carrier family 15 member 3	ENCODES a protein that exhibits proton-dependent oligopeptide secondary active transmembrane transporter activity (inferred) AND symporter activity (inferred) AND INVOLVED IN protein transport (inferred) AND FOUND IN lysosomal membrane AND INTERACTS WITH ammonium chloride AND cefaloridine AND chloroprene	0.671338592	2.790907733	
Snap25	synaptoso mal-associate d protein 25	ENCODES a protein that exhibits SNARE binding AND myosin binding AND protein N-terminus binding AND INVOLVED IN axonogenesis AND calcium ion-dependent exocytosis of neurotransmitter AND endosomal transport AND PARTICIPATES IN insulin secretion pathway AND ASSOCIATED WITH Down Syndrome (ortholog) AND FOUND IN SNARE complex AND actin cytoskeleton AND axon AND INTERACTS WITH C60 fullerene AND ammonium chloride AND arsenite(3-)	-0.649618657	-2.337333133	Yes
Socs2	suppresso r of cytokine signaling 2	ENCODES a protein that exhibits growth hormone receptor binding (ortholog) AND insulin-like growth factor receptor binding (ortholog) AND INVOLVED IN aging AND response to peptide hormone stimulus AND cellular response to hormone stimulus (ortholog) AND PARTICIPATES IN Jak-Stat signaling pathway AND interleukin-2 signaling pathwayinsulin signaling pathway AND ASSOCIATED WITH Sepsis AND Spinal Cord	1.2336994	2.352739818	Yes

		Injuries AND Breast Neoplasms (ortholog) AND INTERACTS WITH 1 3-dinitrobenzene AND 3H-1 2-dithiole-3-thione AND ammonium chloride			
Spp1			0.703696953	2.403035612	Yes
Stk32c	serine/threonine kinase 32C	ENCODES a protein that exhibits ATP binding (inferred) AND protein serine/threonine kinase activity (inferred) AND INTERACTS WITH 17beta-estradiol AND 17alpha-ethynylestradiol (ortholog) AND 2 3 7 8-tetrachlorodibenzodioxine (ortholog)	-1.947577257	-2.964186019	
Syt10	synaptotagmin X	ENCODES a protein that exhibits identical protein binding (ortholog) AND protein heterodimerization activity (ortholog) AND protein homodimerization activity (ortholog) AND INVOLVED IN response to stress AND FOUND IN cell junction (inferred) AND integral to membrane (inferred) AND synaptic vesicle membrane (inferred) AND INTERACTS WITH ammonium chloride AND bis(2-ethylhexyl) phthalate AND bisphenol A	0.808265442	2.623735972	
Tap1	transporter 1 ATP-binding cassette sub-family B (MDR/TAP)	ENCODES a protein that exhibits ADP binding AND ATP binding AND ATPase activity AND INVOLVED IN ATP catabolic process AND peptide transport AND protection from natural killer cell mediated cytotoxicity AND PARTICIPATES IN antigen processing and presentation pathway AND phagocytosis pathway AND primary immunodeficiency disease pathway AND ASSOCIATED WITH Alveolitis Extrinsic Allergic (ortholog) AND Arthritis Reactive (ortholog) AND Asthma (ortholog) AND FOUND IN TAP complex AND intracellular membrane-bounded organelle AND integral to endoplasmic reticulum membrane (ortholog) AND INTERACTS WITH 2 6-dinitrotoluene AND 2-nitrofluorene AND 4-(N-nitrosomethylamino)-1-(3-pyridyl)butan-1-one	0.674822925	2.019832933	Yes
Timp1	TIMP metallopeptidase inhibitor 1	ENCODES a protein that exhibits protease binding AND metalloendopeptidase inhibitor activity (ortholog) AND INVOLVED IN aging AND cartilage development AND cell activation AND PARTICIPATES IN interleukin-6 signaling pathway AND ASSOCIATED WITH Brain Injuries AND Cholestasis AND Diabetes Mellitus Experimental AND FOUND IN extracellular matrix AND basement membrane (ortholog) AND proteinaceous extracellular matrix (ortholog) AND INTERACTS WITH (S)-colchicine AND 1 2 4-trimethylbenzene AND 1-naphthyl isothiocyanate	0.960824309	1.967044627	Yes
Tmsb10	thymosin beta 10	ENCODES a protein that exhibits actin monomer binding (inferred) AND INVOLVED IN actin cytoskeleton organization (inferred) AND sequestering of actin monomers (inferred) AND spermatid development (inferred) AND FOUND IN cytoplasm (inferred) AND	1.003384925	3.127600457	

		cytoskeleton (inferred) AND INTERACTS WITH 17alpha-ethynylestradiol AND 2 3 7 8-tetrachlorodibenzodioxine AND 2-nitrofluorene			
Tnfrsf12a	tumor necrosis factor receptor superfamily member 12a	INVOLVED IN cell adhesion (ortholog) AND cell death (ortholog) AND positive regulation of axon extension (ortholog) AND PARTICIPATES IN cytokine mediated signaling pathway AND FOUND IN cell surface (ortholog) AND plasma membrane (ortholog) AND ruffle (ortholog) AND INTERACTS WITH 1 3-dinitrobenzene AND 2 3 7 8-tetrachlorodibenzodioxine AND 2 4-dinitrotoluene	0.897563874	3.039993412	Yes
Ttk	Ttk protein kinase	ENCODES a protein that exhibits identical protein binding (ortholog) AND PARTICIPATES IN cell cycle pathway mitotic AND INTERACTS WITH cefaloridine AND cisplatin AND lithium atom	0.858999329	2.488305361	
Xkr4	XK Kell blood group complex subunit-related family member 4	FOUND IN integral to membrane (inferred) AND INTERACTS WITH C60 fullerene AND cadmium dichloride	0.808544465	2.092256329	

References

Alberti, G., Zimmet, P., Shaw, J. & Grundy, S. M. (2006). The IDF consensus worldwide definition of the metabolic syndrome. Retrieved March 18, 2014, from http://www.idf.org/webdata/docs/IDF_Meta_def_final.pdf

Animal models of disease: Metabolic, renal, and cardiovascular. (2011). Retrieved March 20, 2014, from http://www.criver.com/files/pdfs/rms/rm_rm_d_animal_models_of_disease.aspx

Bloomgarden, Z. T. (2004). Definitions of the insulin resistance syndrome: The 1st world congress on the insulin resistance syndrome. *Diabetes Care, 27*(3), 824-830.

Choi, J. K., Yu, U., Kim, S., & Yoo, O. J. (2003). Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics (Oxford, England), 19 Suppl 1*, i84-90.

Frayn, K. N. (2010). *Metabolic regulation* (3rd ed.). Oxford, United Kingdom: Wiley-Blackwell.

Gobal, F., Deshmukh, A., Shah, S., & Mehta, J. L. (2011). Triad of metabolic syndrome, chronic kidney disease, and coronary heart disease with a focus on microalbuminuria death by overeating. *Journal of the American College of Cardiology, 57*(23), 2303-2308. doi:10.1016/j.jacc.2011.02.027; 10.1016/j.jacc.2011.02.027

Grundy, S. M. (2007). Metabolic syndrome: A multiplex cardiovascular risk factor. *The Journal of Clinical Endocrinology and Metabolism, 92*(2), 399-404. doi:10.1210/jc.2006-0513

Grundy, S. M. (2012). Pre-diabetes, metabolic syndrome, and cardiovascular risk. *Journal of the American College of Cardiology, 59*(7), 635-643. doi:10.1016/j.jacc.2011.08.080; 10.1016/j.jacc.2011.08.080

Grundy, S. M., Brewer, H. B.,Jr, Cleeman, J. I., Smith, S. C.,Jr, Lenfant, C., National Heart, Lung, and Blood Institute, & American Heart Association. (2004). Definition of metabolic syndrome: Report of the national heart, lung, and blood institute/american heart association conference on scientific issues related to definition. *Arteriosclerosis, Thrombosis, and Vascular Biology, 24*(2), e13-8. doi:10.1161/01.ATV.0000111245.75752.C6

Harlan Laboratories. (2014). Zucker obese rats. Retrieved 4/16, 2014, from http://www.harlan.com/products_and_services/research_models_and_services/research_models/zucker_obese_rats.hl

Hong, N., Oh, J., Lee, Y. -., Youn, J. C., Park, S., Lee, S. -., . . . Kang, S. -. (2014). Comparison of association of glomerular filtration rate with metabolic syndrome in a community-based population using the CKD-EPI and MDRD study equations. *Clinica Chimica Acta, 429*, 157-162. doi:10.1016/j.cca.2013.12.008

Hu, P., Greenwood, C. M., & Beyene, J. (2005). Integrative analysis of multiple gene expression profiles with quality-adjusted effect size models. *BMC Bioinformatics, 6*, 128. doi:10.1186/1471-2105-6-128

Laulederkind, S. J., Hayman, G. T., Wang, S. J., Smith, J. R., Lowry, T. F., Nigam, R., . . . Shimoyama, M. (2013). Rat genome database. *Brief Bioinform,* doi:10.1093/bib/bbt007

Locatelli, F., Pozzoni, P., & Del Vecchio, L. (2006). Renal manifestations in the metabolic syndrome. *Journal of the American Society of Nephrology : JASN, 17*(4 Suppl 2), S81-5. doi:10.1681/ASN.2005121332

Okada, R., Yasuda, Y., Tsushita, K., Wakai, K., Hamajima, N., & Matsuo, S. (2013). The number of metabolic syndrome components is a good risk indicator for both early- and late-stage kidney damage. *Nutrition, Metabolism, and Cardiovascular Diseases : NMCD,* doi:10.1016/j.numecd.2013.08.004; 10.1016/j.numecd.2013.08.004

Singh, A. K., & Kari, J. A. (2013). Metabolic syndrome and chronic kidney disease. *Current Opinion in Nephrology and Hypertension, 22*(2), 198-203. doi:10.1097/MNH.0b013e32835dda78; 10.1097/MNH.0b013e32835dda78

Song, J., Liu, H., Ressom, H. W., Tiwari, S., & Ecelbarger, C. M. (2008). Chronic rosiglitazone therapy normalizes expression of ACE1, SCD1 and other genes in the kidney of obese zucker rats as determined by microarray analysis. *Experimental and Clinical Endocrinology & Diabetes : Official Journal, German Society of Endocrinology [and] German Diabetes Association, 116*(6), 315-325.

Stern, M. P., Williams, K., Gonzalez-Villalpando, C., Hunt, K. J., & Haffner, S. M. (2004). Does the metabolic syndrome improve identification of individuals at risk of type 2 diabetes and/or cardiovascular disease? *Diabetes Care, 27*(11), 2676-2681.

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America, 100*(16), 9440-9445. doi:10.1073/pnas.1530509100

Tusher, V. G., Tibshirani, R., & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America, 98*(9), 5116-5121. doi:10.1073/pnas.091062498

Wallis, R. H., Collins, S. C., Kaisaki, P. J., Argoud, K., Wilder, S. P., Wallace, K. J., . . . Gauguier, D. (2008). Pathophysiological, genetic and gene expression features of a novel rodent model of the cardio-metabolic syndrome. *PloS One, 3*(8), e2962. doi:10.1371/journal.pone.0002962; 10.1371/journal.pone.0002962

Yamamuro, T., Hori, M., Nakagawa, Y., Hayashi, T., Sakamoto, S., Ohnishi, J., . . . Urayama, O. (2013). Tickling stimulation causes the up-regulation of the kallikrein family in the

submandibular gland of the rat. *Behavioural Brain Research, 236*(1), 236-243. doi:10.1016/j.bbr.2012.09.001; 10.1016/j.bbr.2012.09.001

Yasui, N., Kajimoto, K., Sumiya, T., Okuda, T., & Iwai, N. (2007). The monocyte chemotactic protein-1 gene may contribute to hypertension in dahl salt-sensitive rats. *Hypertension Research : Official Journal of the Japanese Society of Hypertension, 30*(2), 185-193. doi:10.1291/hypres.30.185