A systems genetics analysis of metastatic mammary cancer development in mice fed varying levels of dietary fat

Ryan R. Gordon

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Nutrition, School of Public Health.

Chapel Hill
2009

Approved By:

Advisor: Daniel Pomp

Reader: Melinda Beck

Reader: Kent Hunter

Reader: David Threadgill

Reader: Steven Zeisel

ABSTRACT

Ryan R. Gordon

A systems genetics analysis of metastatic mammary cancer development in mice fed varying levels of dietary fat

(Under the direction of Daniel Pomp)

High dietary fat intake and/or obesity may increase the risk of susceptibility to certain forms of cancer. To study the interactions of dietary fat, obesity, and metastatic mammary cancer, a population of F_2 mice cosegregating obesity quantitative trait loci (QTL) and the MMTV-PyMT transgene was created. The mice were fed either a very high-fat or a matched-control-fat diet, and evaluated for growth, body composition, age at mammary tumor onset, tumor progression, and pulmonary metastases development. Single nucleotide polymorphism (SNP) genotyping across the genome facilitated analyses of QTL and QTL by diet interaction effects. To further investigate the complex genetic architecture that modifies mammary cancer and metastasis, expression profiles of axillary tumors were characterized with the Illumina Mouse-6 whole genome sentrix arrays. Using a systems-based analysis pipeline developed in R, we conducted a genome-wide expression QTL (eQTL) analysis was conducted. In addition, network and pathway QTL analyses for mammary tumors that have

developed in the presence of varying degrees of obesity, and during exposure to high or normal fat diets.

Results demonstrated that mice fed a high-fat diet are not only more likely to experience decreased mammary cancer latency but they also have increased tumor growth and occurrence of pulmonary metastases over an equivalent time. We identified 25 modifier loci for mammary cancer and pulmonary metastasis, likely representing 13 unique loci after accounting for pleiotropy, as well as novel QTL x diet interactions at a majority of these loci. Transciptome mapping revealed several candidate genes potentially underlying both tumor and metastasis QTL. These candidates were subsequently prioritized using multiple analytic approaches, including but not limited too causality testing, copy number variation analysis and database evaluations.

To my wife, Gretja, for her unending support, encouragement and passion towards life.

To my parents, Jim and Kathie, for believing in me, even when others did not.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMD	Average metastasis density
Chr	Chromosome
cM	Centimorgan
CNV	Copy number variation
DEXA	Dual-energy X-ray absorptiometry
eQTL	Expression quantitative trait loci
FDR	False discovery rate
HFD	High fat diet
IPA	Ingenuity Database
Mb	Mega base
MCD	Matched control diet
MET	Metastasis
MMTV	Mouse mammary tumor virus
PyMT	Polyoma middle T
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism
TA	Axillary tumor
TC	Tumor count
TI	Inguinal tumor
TOID	Tumor onset in days
TTW	Total tumor weight

INTRODUCTION

Significance

According to National Institute of Health statistics for the US, breast cancer is the most common cancer type and accounts for the second leading cause of cancer-related deaths in women, excluding skin cancers (National-Cancer-Institute 2005). Although breast cancer is thought to be a disease that primarily affects women, about 1% of all breast cancer cases diagnosed occur in men. It has been estimated that in 2009, 192,370 women and 1,990 men in the United States will be diagnosed with some form of breast cancer and that over 40,170 women and 450 men will die of this disease (American-Cancer-Society 2009), typically from secondary metastatic disease (Sporn 1996). The treatment and research of this disease costs the United States on average $8.1 billion a year (National Cancer Institute 2007). Additionally, this disease has a significant global impact, as indicated by the estimate that each year breast cancer causes 502,000 deaths worldwide (World-Health-Organiziation 2006).

Breast cancer is an extremely complex disease contributed to by a combination of environmental and genetic pressures. A substantial amount of effort has been expended in the attempt to identify the associated risk factors of breast cancer; however, the majority of underlying mechanisms that result in this altered disease state still remain unclear. While some successful attempts to understand the genetic predisposition to mammary cancer have been achieved (i.e., identifying small-to-low effect familial risk factors), support for

interactions between environmental components, such as dietary fat, and cancer have been less fruitful. The purpose of this work was to test the hypothesis that mice predisposed to mammary tumor development and challenged with either high or normal levels of dietary fat will experience variability in the pathogenesis of mammary cancer as a result of variations in genetic predisposition, gene expression, and somatic mutations.

Breast cancer and metastasis paradigm

The majority of breast cancer cases begin as aberrant cellular growths in the ducts or lobules of the mammary glands (American-Cancer-Society 2009). Multiple factors that alter breast cancer risk exist including, but are not limited to, diet (Murtaugh et al. 2008), breast feeding (Huo et al. 2008), age at first pregnancy, parity, and family history (Zografos et al. 2004). Differential susceptibility to breast cancer is thought to be mediated by three different classes of cancer modifiers: rare high-penetrance cancer-associated alleles, rare genomic mutations that confer intermediate risk, or common low penetrance polygenes (Turnbull and Rahman 2008). While mutations in genes like *BRCA1, BRAC2, TRP53* and *PTEN* can result in increased breast cancer risk (King et al. 2003; Rohan et al. 2006; Song et al. 2006; Walsh et al. 2006), these inherited alleles contribute to only about 15-20% of all breast cancers (Balmain et al. 2003), suggesting that a polygenic etiology is responsible for the majority of breast cancer cases. In most cancer cases the primary tumor is considered nonfatal and can be readily treated by its surgical removal. However, in many instances as tumor growth progresses, abnormal cells invade the lymphatic system or other vasculature and spread (metastasize) to distant sites in the body, such as the brain, bones, and lungs. These

secondary cancers are far more difficult to treat and typically result in mortality (Murphy 2001).

The ability of a tumor to metastasize is an inefficient process (Chambers et al. 2002) mediated by a series of steps which ultimately allow the carcinogenic cells to escape and survive beyond the site of the primary tumor. As a tumor grows its energy requirements increase along with its need for additional vasculature (Lunt et al. 2009). The process of angiogenesis provides the tumor access to oxygen, nutrients and an outlet to eliminate waste products (Carmeliet and Jain 2000). This newly formed vasculature in turn, provides a route for metastatic cells to escape the primary tumor and pass into the general circulation (Carmeliet and Jain 2000). Additionally, the disruption of the cell-to-cell adhesion properties of the tissue surrounding the tumor, via secreted proteases, may allow the cancer cells to intravasate into the lymphatic or circulatory systems (Bogenrieder and Herlyn 2003; Kroemer and Pouyssegur 2008). Circulation is a harsh environment and most cancer cells undergo anoikis once they become detached from the extracellular matrix (Steeg 2006). Therefore, it is critical for metastasizing cells to protect themselves against programmed cell death until they are able to enter an arrested state. Alterations in gene expression of the metastatic cells may contribute to this process (Douma et al. 2004; Geiger and Peeper 2005; Howard et al. 2008; Zhan et al. 2004); however, the complete anti-anoikis mechanisms are not yet known.

Multiple mechanisms exist by which metastatic cells can escape from circulation. One possibility is for the cells to arrest themselves by becoming lodged in capillaries that are too small to pass through (Steeg 2006). These cells may also non-specifically bind to coagulation factors such as fibrinogen and fibrin, which increases the likelihood that they

will become lodged (Chambers et al. 2002). Additionally, the carcinogenic cells may take an active role in their arrest by secreting tumor-derived endothelial-cell adhesion molecules, which allow the cells to bind to the interior surface of the vessels (Glinskii et al. 2005). Once arrested the cells can extravasate the endothelium and begin their attempts to colonize at new locations (Miles et al. 2008). The ability to survive in a new tissue/organ environment requires that the surrounding conditions are optimal for growth, the cells are able to protect themselves from apoptosis, and that the cells are able to stimulate angiogenesis (Chambers et al. 2002).

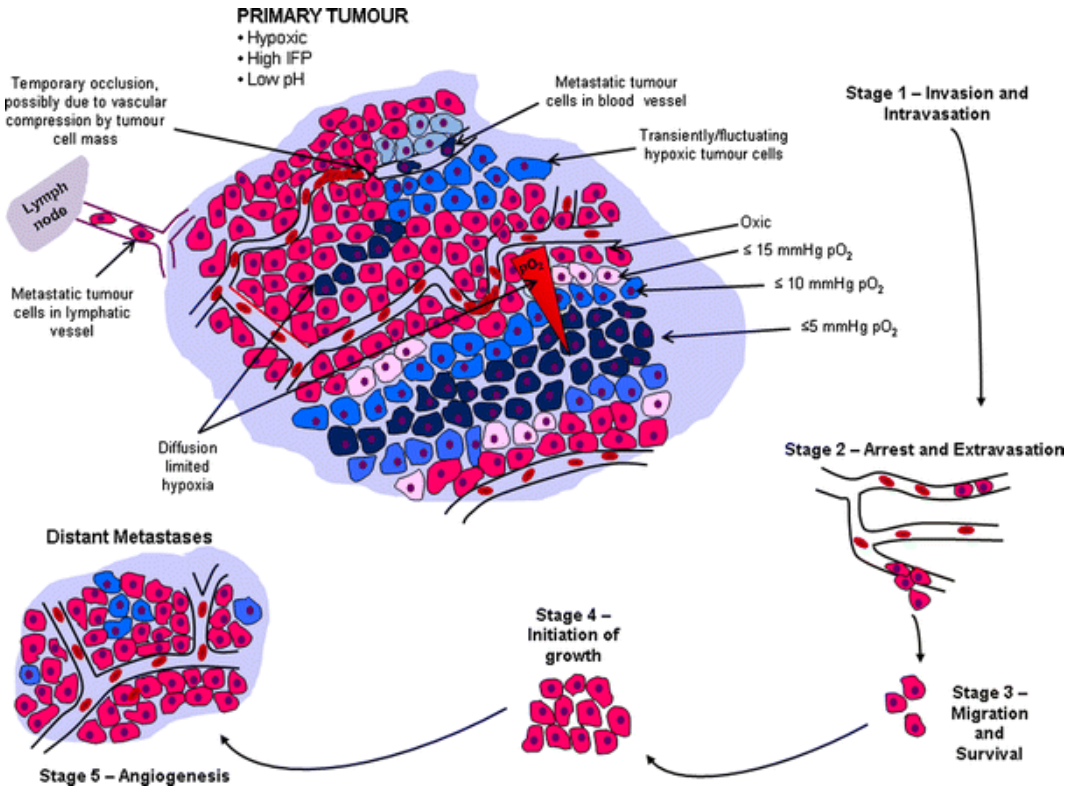


Figure 1.1: Schemata of the stages of the metastatic process (Lunt et al. 2009)

From a genetic standpoint, three contrasting paradigms for metastatic cancer currently exist: progression, initiation, and the predisposition models. The progression model suggests that the cancer's ability to metastasize is a rare occurrence during the cancer's progression in a population of cells due to random mutations and epigenetic alterations (Nowell 1976). The initiation model, which was derived from expression profiling, follows the belief that mutations that occur during early development of the primary tumor are responsible for metastatic potential (Ramaswamy et al. 2003). However, existing cancer data cannot be explained by either of these two models alone, and as a result a third model (predisposition) may be more appropriate (Threadgill 2005). The predisposition model explains that an individual's susceptibility to metastatic cancer is a direct result of his/her genetic makeup. In layman's terms, the cancer's ability to spread to distant sites in the body is facilitated by inherited genes. While multiple genes that contribute to metastatic processes have been characterized (Nguyen and Massague 2007), these most likely only represent a small fraction of those involved in the metastatic cancer process. Additionally, an estimated 60-70% of patients have progressed to metastatic disease by the time of their diagnosis (Eccles et al. 1994). Further elucidation of the genetic underpinnings influencing metastatic cancer is essential for decreasing cancer mortality. Therefore, in this dissertation further characterization of the genetic architecture altering the metastatic breast cancer processes will be performed.

Obesity and Mammary Cancer

Obesity costs the U.S. ~75 billion dollars annually in associated health costs (Finkelstein et al. 2004). Obesity occurs when energy intake exceeds energy expenditure over a long period of time resulting in the storage of reserved energy as excessive adipose. According to the American Obesity Association (2005), 64.5 percent (about 127 million) of the adult American population is overweight or obese, and each year obesity contributes to an excess of 300,000 deaths in the US. Many of these deaths are through comorbidities, including those related to increased risk of cardiovascular disease, type II diabetes, and certain forms of cancer. Though many monogenic models of obesity exist in mice and in humans (see (Rankinen et al. 2006)), obesity is primarily a complex disease. As such, many factors may influence an individual's risk of becoming obese, ranging from lifestyle (e.g. diet) to genetic predisposition, as well as interaction effects between genotype and environment.

In the past few decades an increasing amount of attention has been spent trying to dissect the complex interaction occurring between obesity and cancer etiologies. Evidence suggests that a strong correlation exists between increasing levels of adiposity and increased risk for the development of breast cancer among postmenopausal women (Agnoli et al. 2009; Honda et al. 1998; Lahmann et al. 2004; Reeves et al. 2007). The evidence supporting a link between obesity and premenopausal breast cancer is less clear-cut. However, given the observation of Friedenreich et al. (2002) that postmenopausal breast cancer risk was highly associated with excess weight gain over a 20-yr period, excess body fat throughout one's life appears to have negative breast health implications (Santen et al. 2007). This association appears to be further supported by findings that a link not only exists between obesity and

breast cancer in postmenopausal women, but also between weight gain over an extended period of time and the increased incidence of breast cancer as well (Feigleson et al. 2004).

Changes in the concentrations of serum estrogen have been established to alter the risk for developing breast cancer (Hankinson et al. 1998; Kaaks et al. 2005), with higher levels typically associated with increased disease risk (Eliassen et al. 2006; Key et al. 2003). Links have been reported between increased adiposity and higher levels of estrogenic compounds among postmenopausal women (Lukanova et al. 2004; McTiernan et al. 2006). This increased production of estrogenic compounds is thought to be the result of aromatase conversion of androstenedione, produced in the adrenal gland, to estrone (Bray 2002). Given that this rate of estrone production is directly related to the size of the adipose deposits, it is potentially a significant source of estrogenic compounds, particularly in postmenopausal women (Bray 2002). This association between obesity and circulating estrogens has also been confirmed amongst postmenopausal breast cancer survivors as well, in which obesity was linked to higher levels of estrone and estradiol and, ultimately, to increased risk for a recurrence of cancer (Calle and Thun 2004; McTiernan et al. 2003).

Not only has obesity been implicated in increasing one's risk for breast cancer, excessive fat accumulation is also believed to lead to a poorer prognosis for survival. Individuals who are obese are more likely to have cancer that has progressed to more aggressive later stages at the time of diagnosis (Cui et al. 2002). This advanced stage of cancer is directly correlated to a decreased chance of survival (National-Cancer-Institute 2005). Several large prospective studies have confirmed the link between a higher BMI, or higher waist-to-hip ratio, and increased breast cancer mortality (Borugian et al. 2003; Calle et al. 2003; Petrelli et al. 2002; Whiteman et al. 2005). These findings most likely occur,

partially because of the difficulty associated with and delay in detecting a tumor in obese women compared to women with less body fat (Oestreicher et al. 2002). This late diagnosis may lead to proliferation and progression of more advanced tumors and an overall greater tumor burden at the time of detection (Berclaz et al. 2004; Demirkan et al. 2007). Additionally, it has been observed that obese patients have a higher incidence of lymph node involvement as compared to patients of normal weight (Daniell 1988), reinforcing this link between obesity and metastatic cancer development.

A mechanistic insight into the obesity and metastatic cancer development cancer paradigm may come directly from the adipocyte microenvironment. Inflammatory proteins, such as IL-6, produced in response to the inflammation associated with obesity help promote the angiogenic properties of a tumor by stimulating the production of vascular endothelial growth factor (VEGF). VEGF is an adipokine that is a well established promoter of the vascularizton of mammary tumors (Rega et al. 2007). Leptin, another adipokine positively correlated with increased adipose stores appears to be a potent mediator of angiogenesis as well (Vona-Davis and Rose 2009). Whereas a reduction in the circulating levels of adiponectin, an adipokine negatively correlated with increasing adiposity (Weyer et al. 2001), can have a negative impact of mammary tissue health. This negative impact is accomplished by promoting both tumor onset and tumor growth (Lam et al. 2009). Additionally, excess adipose tissue can result in an increase in the circulating levels of matrix metalloproteinase 2 and 9 (Bouloumie et al. 2001), both of which can contribute to the disruption of the cell-to-cell adhesion properties of the tissue surrounding a tumor. This disruption, as pointed out earlier, can allow the cancer cells to intravasate into the lymphatic or circulatory systems (Bogenrieder and Herlyn 2003; Kroemer and Pouyssegur 2008).

Ultimately, each one of these processes could individually or synergistically contribute to the increased mortality associated with the obese state.

Dietary fat and breast cancer

A complex relationship between diet and genetics underlies breast cancer susceptibility. Whereas genetic predisposition to mammary cancer has been confirmed by identification of multiple, small-to-low effect familial risk factors, the evidence connecting dietary components to cancer susceptibility has been limited and inconsistent. Links between diet and incidence of breast cancer have gained increasing attention (Key et al. 2004) and a broad spectrum of dietary components ranging from alcohol consumption (Lew et al. 2009; Terry et al. 2006) to fiber intake (Cade et al. 2007) have been investigated. To date the association between total dietary fat and breast cancer risk has produced conflicting results. While some studies have shown a positive association between fat intake and breast cancer (Cho et al. 2003; Lee et al. 2005; Thiebaut et al. 2007), others have failed to find such an association (Kim et al. 2006; Wakai et al. 2005). In mouse models this relationship seems to be much clearer, as observed in the original studies in this field by Tannenbaum (1942). Tannenbaum found that mice fed a high-fat diet experienced a higher frequency of mammary cancer as compared to the low-fat fed controls. This relationship has been subsequently confirmed many times in mice (e.g. (Cleary et al. 2004)), but the underlying mechanisms remain largely unknown. Furthermore, given that it was recently estimated that approximately 34% of the total energy in the American diet is derived from fat (Kerver et al. 2006), a strong need to clarify this relationship in an experimental model exists. Therefore,

using the unique mouse population that we developed, further investigation of how dietary fat modulates cancer susceptibility and progression is proposed.

Mouse models of breast cancer

Mouse models have been used to gain a better understanding of the inherited genetic and environmental determinants for cancer risk in the human population (Callahan and Smith 2000; Cavanna et al. 2007). The use of mouse models is mainly due to the striking similarities between biological and genetic processes related to cancer development and progression in humans and mice (e.g. (Lin et al. 2003)). Unlike in the human population, rigid control of all the environmental factors allowing for the systematic removal of external influences is possible (Balmain 2002). Additionally, evaluation of the environmental components of the breast cancer paradigm (i.e. diet) through selective inbreeding, which can provide us with a genetically homogenous research population, is possible. By utilizing inbred populations of mice, many other investigators have made substantial findings in the fields of breast cancer (Hennighausen 2000; Park et al. 2003; Thomas et al. 1996) and other polygenic diseases (Bower et al. 2006; Cheverud et al. 2004; Korstanje and DiPetrillo 2004; Liu et al. 2007a).

Whereas mouse models can provide us with valuable information in regards to cancer processes, they have their limitations as well. A few essential differences between mouse models and humans are as follows; carcinogenic risk factors often differ among mice and humans, mouse tumors are commonly mesodermal sarcomas while tumors in humans are typically epithelial carcinomas, and the sites of spontaneous tumor formation can vary

(Anisimov et al. 2005). However, in the case of breast cancer mice appear to develop cancer in the same general locations with a remarkably similar histopathological course as that observed in humans (Anisimov et al. 2005; Balmain and Harris 2000). Regardless of the limitations associated with using mice to investigate human diseases, mice remain an incredibly valuable resource in complex trait dissection.

Quantitative trait analysis

In complex genetic disorders, phenotypes alone do not allow us to dissect the underlying genetic problems, since typically phenotypes represent end-points that can potentially be reached in many ways (Gusella and MacDonald 2002). For essentially any disorder, phenotypes differ either qualitatively or quantitatively depending on the underlying genetic architecture (Gusella and MacDonald 2002). One specific technique for determining the genetic underpinnings for any given trait is the mapping of quantitative trait loci (QTL). QTL mapping has been described as the first step toward the identification of genes and casual polymorphisms for traits of importance in agriculture and human medicine (Seaton et al. 2002). This procedure utilizes an approach with no prior assumptions about the potential importance of specific genes or genetic regions. Instead, the results of the scan are used in an unbiased manner to identify chromosomal segments which are highly correlated with a particular phenotype. These regions, in turn, become the focus of more intensive follow-up analyses to uncover the underlying genes (Comuzzie and Allison 1998). The simplest and most efficient way to detect QTL is by using inbred-line crosses because the limited genetic variation within strain. Therefore, by crossing two inbred parental lines, the resulting

population will exhibit a fixed difference between every marker and trait locus. Thus, all linked loci in the F_1 generation used to create the mapping population are in linkage disequilibrium (Slate 2005). In the subsequent F_2 generation, QTL then represent genetic variation between the founder lines which are fixed for alternative alleles at the QTL (Seaton et al. 2002).

The method of QTL mapping has proven to be an efficient platform for aiding in the process of understanding complex traits, including obesity and a variety of diet-related cancers. The usefulness of this method is particularly evident in the review by Dragani (2003), in which he describes the detection of over 100 QTL for a wide spectrum of cancer phenotypes. While QTL analysis is a beneficial tool for the dissection of complex traits, like any technology, associated limitations also exist. The successful completion of a QTL analysis requires large crosses consisting of hundreds and possibly thousands of animals, the tedious collection of relevant phenotypes, and the detection of informative loci spanning the entire genome (Singer et al. 2004). Yet, the most recognized limitation is the long and arduous processes involved in making the transition from a QTL to the underlying gene or genes of interest (Flint et al. 2005; Miles and Wayne 2008; Moore and Nagle 2000). Regardless of the limitations, when used as a starting point, this method still remains very useful for dissecting complex traits. Furthermore, with the advent of multiple promising analytic techniques, many of the aforementioned roadblocks may be circumvented. Therefore, we are proposing to utilize this QTL method as the foundation for our pursuit of the genetic architecture of metastatic mammary cancer.

Gene expression and eQTL

Given that QTL analysis are compounded by a substantial roadblock in identifying the genes that underlie disease phenotypes, alternative approaches have been taken. Other methods such as *in vivo selection*, gene expression analysis and clinical verification have proven to be successful. In a particular experiment, a combination of these techniques was used in a mouse model to identify gene sets which predicted if mammary tumors would metastasize to the lung (Minn et al. 2005). The gene sets were then evaluated in humans and found to be clinically correlated with the development of lung metastasis when expressed in the primary tumor (Minn et al. 2005). Ultimately revealing multiple lung metastasis signature genes which appear to enhance metastatic growth within breast and the lung, along with a subset of metastatic genes that were rarely expressed in the primary tumors but were strongly selected for once they reached the lungs (Minn et al. 2005). Successful attempts have been made using these methods, but they tend to be labor-intensive and costly relative to their ultimate yield. Additionally, While QTL analysis and differential expression can reveal pathways, candidate regions and genes potentially linked to disease phenotypes, these methods are both independent assessments of the paradigm. As such, it is difficult to infer a relationship between the results of the two methods.

However, A relatively new approach has been developed with the potential to bridge this disconnect by treating the expression of each transcript identified through microarray analysis as a quantitative trait. The traits can then be tested for associations with genotypic data to develop what is known as an expression QTL (eQTL) (Jansen and Nap 2001; Schadt et al. 2003a). In general, these eQTL are identified in an identical manner as traditional QTL. However, because the traits tested are actual transcripts and not endpoint phenotypes

(e.g. body weight, body fat, tumor size), the difficult task of assigning an actual gene to the QTL is less of an issue. When running an eQTL analysis two distinct classes of loci are detected. The first class is loci that map within close proximity to the actual physical location of the expressed gene (*cis*-acting), and the second class is loci which map independently of the expressed genes physical location (*trans*-acting) (Pomp et al. 2008). While *trans*-acting eQTL represent loci controlled by unknown regulators, *cis*-acting eQTL exhibit self-regulation (Alberts et al. 2005). Therefore, by comparing or overlaying these *cis*-acting eQTL with locations of traditional phenotypic QTL detected in the same mapping population, the potential exists to significantly enrich for candidate genes that are both positional and functional in nature (Wang et al. 2007b). The field of research on metastatic breast cancer has produced only a few experiments utilizing this multifaceted approach (Crawford et al. 2008). A multifaceted approach which combines the traditional technique of QTL mapping and the emerging method of eQTL mapping to increase understanding of the dietary fat metastatic breast cancer paradigm was utilized in our studies (see following chapters).

Copy number variation (CNV) and Chromosomal aberrations

Recently a new source of genetic variation that can potentially impact on disease processes has been identified, known as copy number variants (CNV). CNV is described as segments of DNA that are over or under represented due to insertions/deletions occurring naturally over time or acutely due to tissue-specific somatic mutations (Feuk et al. 2006). Approximately 12% of the human genome has been estimated to be affected by CNV

(Beckmann et al. 2007) and this over/under representation of chromosomal segments can have profound influences on the expression of the gene within these afflicted regions. Multiple diseases such as Crohn's disease (Fellermann et al. 2006), lupus (Yang et al. 2007), and HIV (Gonzalez et al. 2005) have already been linked in part to CNV. It is plausible that CNV may be linked to many common complex diseases such as cancers (Shelling and Ferguson 2007), yet currently our knowledge of this paradigm remains limited.

Taking into account CNV is especially important when analyzing tumor tissue, given the substantial amount of evidence linking the accumulation of CNV to cancer pathogenesis (de Tayrac et al. 2009; Fridlyand et al. 2006; Reis-Filho et al. 2005; van Beers and Nederlof 2006; Yusenko et al. 2009). CNV can be classified into two categories, those that are inherited (germ-line) and those that are acquired during the replication of cells (somatic). In humans, germ-line CNVs are detected across all tissues in both healthy and diseased individuals (Shlien and Malkin 2009). The presence of CNV in genomic regions encoding cancer modifiers can lead to increased risk for the development of cancer (Albertson et al. 2003). Somatic CNV are acquired during DNA replication and are not found uniformly through out all tissue types, and their presence can impart a growth advantage to cells harboring them, resulting in disease (Greenman et al. 2007). Tsafrir et al. (2006) research suggests that as cancer progresses it is possible for tumors to continue acquiring somatic CNV, which could potentially alter their metastatic tendencies. To date few, if any, large association studies have used this combined approach in the characterization of complex diseases. Therefore, CNV will be examined within our population and its association with mammary cancer susceptibility and progression will be evaluated.

Model Review

As previously suggested, the use of animal models to dissect the genetic architecture of complex diseases has yielded valuable information. The proposed cross involves M16i and FVB, two phenotypically divergent mouse models which have been extensively characterized. The following section includes a brief review of the M16 model of obesity and a mammary cancer-focused review of the FVB transgenic cancer model.

M16i an inbred obesity line

One particular model of interest is the M16, which was derived from the ICR strain through 27 generations of selection for weight gain from 3 to 6 weeks of age (Hanrahan et al. 1973). When fed the same diet, M16 mice will gain weight faster and reach a greater end-weight than its control line (ICR). M16 mice also appear to be hyperglycemic when compared to its control (Allan et al. 2004). The inbred version of the M16, the M16i has been used in multiple QTL mapping projects. One project in particular was carried out by creating a large F_2 intercross between the M16i and its control line (ICR), and this ultimately resulted in a large number of QTL being detected (Rocha et al. 2004a, b). Another large intercross study utilizing these M16 and ICR strains (Allan et al. 2005) yielded 95 QTL. Of the 95, 39 QTL had effects on body weight and growth traits, 36 on traits influencing body composition, 12 related to energy intake, four linked to feed efficiency, and eight were associated with serum leptin, insulin and blood glucose. This strain of mice has also been investigated for the impact of environmental components on body size. For example, dietary fat was found to significantly increase body weight and adiposity (Allan et al. 2004). Recently, this mouse model was investigated for the environmental effect of exercise, which

was found to significantly reduce bodyweight after only 6 days of free access to a running wheel (Nehrenberg et al. 2009).

FVB/NJ-TgN(MMTV-PyMT)^{634Mul}

A transgenic mouse model that has been a widely used to study breast cancer is the Polyoma Middle T Oncoprotein (PyMT) mouse. In this model the PyMT antigen is under the control of mouse mammary tumor virus LTR (MMTV) which restricts it to the mammary epithelium (Guy et al. 1992). Multifocal tumors develop in all mammary glands of females and are detectable by palpation at 60 d of age, and by 100 d of age, 85% of the animals develop pulmonary metastasis. Furthermore, male carriers of PyMT experience delayed tumor latency and overall decreased severity (both decreased tumor size and propensity to metastasize) in comparison to females (Lifsted et al. 1998). To verify that the PyMT mouse was an appropriate model for breast cancer investigation, given that PyMT does not naturally occur in humans, the progression patterns and morphology of their tumors were compared to those in humans and were found to be strikingly similar (Lin et al. 2003). Additionally, it has been observed that gene expression patterns detected in PyMT-induced tumors share common characteristics with those in humans that are associated with poor of survival (Lin et al. 2003).

The FVB/NJ-PyMT is a mouse model that has played an important role in the cancer field for increasing the understanding of mammary cancer and the metastatic processes involved. Le Voyer et al. (2000) attempted to assess the basis for the tumor acceleration they observed in the FVB/NJ-PyMT I/LnJ F_2 population. First they ruled out the possibility that the observed change in tumor latency was a result of alterations in the expression of the

transgene using western blots. Then a QTL analysis of this backcross revealed two loci, *Ampt1* and *Ampt2*, responsible for the tumor acceleration, located on Chromosomes 15 and, respectively. The researchers were also able to identify a suggestive locus on Chromosome 7.

Le Voyer et al. (2001) evaluated the mammary tumors arising in an F_1 hybrid background (I/LnJ x FVB/N-TgN) that resulted in earlier onset but reduced total tumor mass than in the parental strain using quantitative and molecular approaches. Their experiments ultimately led to the identification of three loci, designated *Mmtg1-3*, which are associated with tumor growth modification. Both *Mmtg1* and 2 mapped to Chromosome 4 which was a region that had been previously associated with mammary tumorigenesis. *Mmtg3* mapped to the proximal portion of Chromosome 7; this was the same region in which this group had previously mapped a potential latency modifier gene.

Work by Cozma et al. (2002) utilized an approach that combined genetics, genomics and bioinformatics to identify interesting candidate genes for the *Ampt* mammary tumor latency modifiers. By utilizing public databases they were able to identify a large number of papers that found genes within the regions similar to those of *Ampt1* and 2, Chromosomes 15 and 9. The literature search was then further reduced to identify gene pairs that were known to interact or to be in a common pathway relating to breast cancer. This literature search led to the identification of one particular gene pair of interest, *c-Myc* and *Cdc25A*, which was tested for tumor response by creating a double-transgenic (PyVT/Myc) mouse model. The double-transgenic model displayed early and more aggressive tumor growth, and the data collected suggested that *c-Myc* and *Cdc25A* are indeed *Ampt1* and *Ampt2*, respectively.

Hunter et al. (2001) set out to identify the approximate genomic locations of metastasis efficiency genes as a preliminary step for high-resolution mapping and positional cloning of genes of interest. They attempted to identify these phenotypically relevant loci by analyzing four different genetic mapping experiments utilizing three inbred strains that altered only the metastatic phenotype of the mammary tumor, and one backcross that altered tumor growth rate, tumor latency and metastatic efficiency. These analyses lead them to identify a significant locus of a metastasis suppressor on the proximal end of Chromosome 19, designated *Mtes1*. Other suggestive loci were found on Chromosomes 6, 9, 13 and 17. Upon further investigation (Hunter et al. 2001), *Mets1* was suggested to co-localize with the mouse orthologue of the human breast cancer metastasis suppressor gene, *Brms1*.

Park et al. (2005) began to further investigate the underlying genetic components that are responsible for metastatic progression, more specifically the *Mets1* locus. They indentified these genetic factors this by utilizing various methods. The first approached used was to determine candidate regions on Chromosome 19 through an evaluation of a multiple cross and mapping study previously cited (Hunter et al. 2001). The evaluation resulted in identification of five regions of interest in which the researchers were able to seek candidates genes, one of which was *Sipa1*, a gene associated with metastatic progression of human prostate cancer. By using bioinformatics and molecular techniques (such as RT-PCR and western blot analysis), the investigators were able to eliminate several candidates. However, biochemical analysis revealed a polymorphism of interest in *Sipa1* and, when further evaluated, *Sipa1* was considered a strong candidate for one of the genetic polymorphisms underlying the *Mtes1* locus. When *Sipa1* was evaluated in a human cohort, germline

polymorphisms of this gene were implicated in modifying the aggressiveness of breast cancer (Crawford et al. 2006).

Recently the primary tumors collected from multiple recombinant inbred mouse strains, including those derived from FVB-PyMT, were used to generate expression profiles for metastatic signature genes (Yang et al. 2005). These expression profiles were subsequently used to identify relevant metastatic eQTL mapping to the three regions of interest on Chromosomes 7, 17 and 18 (Crawford et al. 2007). The locus on Chromosome 17 was further investigated because it encompassed ribosomal RNA processing 1 homolog B (Rrp1b), a factor known to interact with *Sipa1* and shown to be highly correlated with metastasis-predictive gene expression. Additional analyses revealed that *Rrp1b* may function as a modifier of tumor progression and as a metastasis and metastatic predictive marker for humans (Crawford et al. 2007). By combining the aforementioned gene expression results with additional analyses, Crawford et al. (2008) were able to generate a transcriptional network, the Diasporin Pathway, which may be able to predict the metastatic potential of tumor in both humans and mice.

Summary

In summary, breast cancer is a complex disease resulting from a combination of, and interaction between, environmental and genetic factors. It is thought that differential genetic susceptibility to breast cancer is mediated by three different types of cancer modifiers: rare high-penetrance cancer-associated alleles, rare intermediate penetrance mutations, or common low penetrance polygenes. While mutations in genes like BRCA1, BRCA2, TRP53

and PTEN can result in increased breast cancer risk, these inherited alleles contribute to only 15-20% of all breast cancers, suggesting that a polygenic etiology is responsible for the other 80-85% of breast cancer cases. However, the identity of most of these polygenes has yet to be revealed.

While genetic factors clearly influence breast cancer susceptibility, environmental factors an equally important role as well. Links between diet and incidence of breast cancer have been reported for a broad spectrum of dietary components. In particular, dietary fat has received significant attention, although reports have reached conflicting results. Whereas some human studies have shown a positive association between fat intake and breast cancer, others have failed to replicate these findings. Given this lack of congruity and the large number of individuals afflicted with this disease, a need exists to clarify further the relationships between diet and cancer. Furthermore, whereas numerous studies have investigated genetic or dietary factors linked to breast cancer, few have focused on the gene x diet interactions that are likely to be major contributors to differential risk. Understanding how diet might influence expression and effects of these cancer predisposition loci is increasingly needed as such genes are identified and used in human diagnostics paradigms. With the current analytical techniques available, i.e. QTL/eQTL mapping and CNV analysis, it is possible to address these gaps in our knowledge of the genetic underpinnings of metastatic breast cancer. Additionally, given their extensive use in the field of complex trait analysis, mouse models utilized in the subsequent research papers are aptly suited for investigating the etiology of metastatic breast cancer.

Finally, in this dissertation the hypothesis that mice predisposed to mammary tumor development and challenged with either high or normal levels of dietary fat will experience

variability in the pathogenesis of mammary cancer as a result of variations in genetic predisposition, gene expression, and copy number variation, will be tested. The overall goals of the dissertation are to characterize the effects of dietary fat on mammary cancer phenotypes and elucidate the genetic architecture underlying development, severity and metastatic potential of mammary cancer. In addition, the question of how dietary fat interacts with susceptibility genes will be assessed.

CHAPTER II

Genotype X Diet Interactions in Mice Predisposed to Mammary Cancer: I.

Body Weight and Fat

Ryan Gordon, Kent W. Hunter, Peter Sørensen, Daniel Pomp

Mamm Genome. (2008) March;19(3):163-78

Abstract:

High dietary fat intake and obesity may increase susceptibility to certain forms of cancer. To study the interactions of dietary fat, obesity and metastatic mammary cancer, we created a population of F_2 mice cosegregating obesity QTL and the MMTV-PyMT transgene, we fed the F_2 mice either a very high fat or a matched control fat diet, and we measured growth, body composition, age at mammary tumor onset, tumor number and severity, and formation of pulmonary metastases. SNP genotyping across the genome facilitated analyses of QTL and QTL x diet interaction effects. Here we describe development of the F_2 population (n=615) which resulted from a cross between the polygenic obesity model M16i and FVB/NJ-TgN (MMTV-PyMT)^{634Mul}, effects of diet on growth and body composition, and QTL and QTL x Diet and/or gender interaction effects for growth and obesity-related phenotypes. We identified 38 QTL for body composition traits that were significant at the genome-wide 0.05 level, likely representing 9 distinct loci after accounting for pleiotropic

effects. QTL x diet and/or gender interactions were present at 15 of these QTL, indicating that such interactions play a significant role in defining the genetic architecture of complex traits such as body weight and obesity.

Introduction:

Obesity costs the U.S. ~75 billion dollars annually in associated health costs (Finkelstein et al. 2004). It occurs when energy intake exceeds energy expenditure over a long period of time resulting in the storage of reserved energy as excessive adipose. According to the American Obesity Association (2005), 64.5 percent (about 127 million) of the adult American population is overweight or obese, and each year obesity causes in excess of 300,000 deaths in the US. Many of these deaths are through comorbidities, including those related to increased risk of cardiovascular disease, type II diabetes, and certain forms of cancer. Though many monogenic models of obesity exist in mice and in humans (see (Rankinen et al. 2006)), obesity is primarily a complex and polygenic trait. As such, there are many factors that can influence an individual's risk of becoming obese, ranging from lifestyle (e.g. diet) to genetic predisposition, as well as interaction effects between genotype and environment.

Substantial progress has been made in understanding the genetic underpinnings of obesity in mice, and well over 200 quantitative trait loci (QTL) have been mapped for various obesity-related phenotypes (Pomp 1997). Furthermore, genetic variation in response to high-fat diet has been well characterized (e.g. (Taylor and Phillips 1997; West et al. 1994a; West et al. 1994b)). However, while many QTL studies have focused on finding chromosomal regions controlling obesity in either normal or high fat diets, much less focus

has been placed on examination of QTL x diet interactions by feeding both normal and high levels of fat within the same segregating population (e.g. (Cheverud et al. 2004)).

Links between dietary fat and breast cancer risk have been shown in a meta-analysis of epidemiologic research and migration studies from countries of low to high risk for breast cancer (Forman 2007). High fat diet has been studied as a risk factor for mammary tumors in mice (e.g. (Cleary et al. 2004; Tannenbaum 1942)), and several genetic modifiers of mammary tumor development have been identified in various transgenic mouse models (Connelly et al. 2007; Le Voyer et al. 2000; Le Voyer et al. 2001; Liao et al. 2007; Rajkumar et al. 2007). However, no studies have reported on diet x QTL interaction effects on mammary tumor development and associated metastasis. Understanding how diet might influence expression and effects of cancer predisposition loci is increasingly important as such genes are identified and used in human diagnostics paradigms (Bild et al. 2006).

Many patients afflicted with cancer experience cachexia, the loss of both adipose and muscle tissue (Tisdale 1997), and as the disease progresses so does the associated wasting (Dewys et al. 1980). This in turn can have serious implications on cancer survivability, and some estimates suggest that cachexia is responsible for ~20% of cancer related deaths (Tisdale 2002). While many mediators of cachexia have been characterized (for a detailed review see (Baracos 2006)) the exact mechanisms of action, which are critical for developing treatments, are still not fully understood (Gordon et al. 2005).

We developed a large F_2 population originating from a cross between the polygenic obesity model M16i (Allan et al. 2005) and FVB/NJ-TgN(MMTV-PyMT)^{634Mul}, a transgenic line that generates aggressive mammary tumors with subsequent pulmonary metastases (Guy

et al. 1992). Half the F_2 mice harbored PyMT while half did not, and the F_2 mice were fed either high fat or matched control diets. Overall, our objectives were to evaluate the phenotypic effects of dietary fat on growth, body composition and cancer traits and to map, in a coordinated fashion, chromosomal positions of predisposition genes (QTL) for obesity and cancer and their associated diet x QTL interactions. In this paper we describe the population, effects of diet on growth and body composition, and QTL and diet x QTL interaction effects for growth and obesity-related phenotypes. In a companion paper we focus on cancer phenotypes.

Materials and Methods:

Resource Populations: An F_2 population (n=615) was generated by crossing M16i, a polygenic obesity line (Allan et al. 2005), and FVB/NJ-TgN(MMTV-PyMT)^{634Mul} (PyMT), a line transgenic for the Polyoma Middle T Oncoprotein, leading to the development of mammary tumors and subsequent pulmonary metastasis (Guy et al. 1992). Since PyMT females have impaired reproduction, the initial cross was between M16i females and FVB males hemizygous for PyMT. In the F_1, males hemizygous for PyMT were crossed to females without PyMT to generate the F_2 population in four consecutive replicate breeding cycles (repeated mating of F_1 pairs). F_2 litters were standardized at birth to 10 pups, maintaining as many females as possible. Individual mice were identified by toe-clipping at postnatal day (d) 12, and toe-clips were used for PCR analysis (Pomp 1991) to identify mice hemizygous for the PyMT transgene with the following primers: forward 5'-AACGGCGGAGCGAGGAACTG-3': reverse 5'-ATCGGGCTCAGCAACACAAG-3'.

F_2 mice were weaned at 3 wk of age and randomly assigned, using a random number generator, within litter, sex and genotype (PyMT or no PyMT) to receive one of two synthetic purified diets at 4 wk of age. Mice designated "high" received a high fat diet (Research Diets D12451) that contained 45% of total calories from fat, 20% from protein and 35% from carbohydrate. Mice designated "control" received a matched control fat diet (Research Diets D12450B) that contained 10% of total calories from fat, 20% from protein and 70% from carbohydrates (Tables 1 and 2). Mice had *ad libitum* access to feed and water. Prior to starting specialized diets all animals received *ad libitum* access to a standard rodent diet (Teklad 8604 Rodent Chow). Total numbers of F_2 mice phenotyped within each of the 8 subclasses ([PyMT, non-PyMT] X [High, Control] X [Male, Female]) are presented in Figure 1.

Data collection: Body weight was measured for each mouse at 3, 6, and 9 wk of age and at sacrifice (~11 wk for females, ~14 wk for males). Body composition was analyzed at 7 wk of age using dual-energy X-ray absorptiometry (DEXA; GE Lunar PIXImus) while mice were anesthetized with Avertin (2,2,2 tribromoethanol, Aldrich), and again at sacrifice following CO_2 exposure. Age of onset of mammary tumor development in the F_2 population was evaluated in mice hemizygous for PyMT beginning at 4 wk of age. Mammary glands for each individual were palpated three times a week until age of onset was determined.

Tissue collection: Mice were sacrificed at approximately 1200 hours, 3 hr after removal of food to increase the accuracy of DEXA measurements by minimizing the amount of chow within the stomach. Blood was collected and the following tissues were dissected from all F_2 mice and snap frozen in liquid nitrogen: liver, epididymal fat for males, perimetrial fat for females (except from females with PyMT due to logistical constraints

associated with collecting the mammary tumors), skeletal muscle, hypothalamus, pituitary and tail. For F_2 mice with PyMT, additional tissues were collected, including one inguinal mammary tumor, one axillary mammary tumor, and one lobe of the lung (the other being fixed for evaluation of metastases). Liver weight was recorded for all mice, while weights of fat pads were recorded for all males and for females lacking PyMT.

Statistical Analysis: Proc Mixed from SAS (2002) was used to analyze data for effects of diet, PyMT and gender. Models for phenotypic traits measured prior to sacrifice (3, 6 and 9 wk weights, 7 wk lean mass and percent fat) included fixed effects of diet, gender and presence of PyMT, interactions of diet x gender, diet x PyMT and gender x PyMT, and random effects of replicate and dam. Traits measured at time of sacrifice (liver weight, fat pad weight, sacrifice weight, total fat, lean mass and percent fat) contained age at sacrifice (in days) as an additional covariate. Liver and fat pad weights were also expressed as a percentage of body weight at sacrifice. Correlations (with Bonferroni corrected P-values) among all traits were evaluated using the MANOVA procedure within Proc GLM (SAS) adjusted for fixed effects of diet, PyMT and gender.

Genotyping and Linkage Map: A total of 384 SNPs were genotyped across 552 tail DNA samples using a service provider (Illumina, San Diego, CA). These SNPs were selected to be relatively evenly spaced across the genome, and because they were predicted to be fully informative between M16i and FVB based on previous genotyping in M16 (the progenitor of M16i) and FVB (Cervino et al. 2005). No information existed on polymorphisms between M16i and FVB for the X chromosome, so only the autosomes were genotyped. The 552 DNA samples consisted of 2 representative M16i parents, 2 FVB, 2 F_1, most F_2 animals with PyMT (106 males, 158 females), the majority of non-PyMT F_2 females (n=178), a large

number of non-PyMT F_2 males (n=96), and 8 replicate samples (to verify genotyping accuracy). Remaining F_2 mice (n=77), containing the least relevant phenotypic information, were not genotyped due to sample size constraints in the genotyping process. Total numbers of F_2 mice genotyped within each of the 8 subclasses are presented in Figure 1.

Any SNPs found to have allele sharing between the genotyped M16i and FVB parental mice were discarded. All remaining SNPs were tested for deviation from the expected 1:2:1 F_2 genotypic distribution using chi-square analysis. This revealed 30 SNPs exhibiting segregation distortion, localized to three specific genomic regions (MMU1, MMU2, MMU3). In all cases the distortion was caused by under-representation of M16i alleles. Under the assumption that inbreeding was not complete in these regions for one or two of the parental M16i mice, we repeated the analysis after dropping data for these SNPs from F_2 mice originating from specific M16i parents. By this method segregation distortion was eliminated from the population with loss of only a small amount of data from the experiment.

A linkage map was created using Map Manager QTXb20 (Manly et al. 2001). Linkage groups were evaluated for consistency with known physical locations from Ensembl. Surplus SNPs (redundant markers) that were in complete linkage disequilibrium were removed, yielding a final linkage map consisting of 124 SNPs with average marker spacing of 10-15 cM (Table 4). The cM position for the first (proximal) SNP on each chromosome was estimated based on its physical location (Build 34) using the formula Mb x 1.6 (Shifman et al. 2006).

QTL Analyses: Phenotypic and genotypic data were merged to detect QTL with the web-based program QTL Express (Seaton et al. 2002) using the "F_2 inbred/Co-dominant Marker Analysis" option. One QTL per chromosome was fitted using the genetic model for additive plus dominance effects. Analyses included the additive and dominance effects, replicate, PyMT, gender and diet as fixed effects in the model as well as a QTL x diet interaction (interaction model). An additional model was tested for each trait with identical components as above but without the interaction term (non-interaction model). To test whether QTL x diet interaction effects were significant, the sum of squares error and the degrees of freedom for the peak position for each QTL in both the interaction and non-interaction models were calculated and used to estimate an F statistic. If the interaction was not significant at $p < 0.05$, the QTL x diet interaction effect was removed from the model. If an interaction was detected, then the mapping population was separated according to diet (animals fed either the high fat diet or matched control (lower) fat diet) and reanalyzed independently with using the non-interaction model to elucidate the cause of the interaction. Similar methods were used to test for gender x QTL interactions.

Genome-wide significance thresholds for QTL effects were determined using permutation testing. A bootstrap procedure was used to estimate confidence intervals for QTL positions. The percentage of phenotypic variation explained by each QTL was calculated as follows: [(residual variance of the reduced model) - (residual variance of the full model) / residual variance of the reduced model] x 100. The nonparametric bootstrap method described by Lebreton et al., (1998) was used to test for pleiotropy among multiple linked QTL.

Results:

F_2 Phenotypic Evaluation: No significant differences were found in body weight at 3 wk of age among groups destined to be fed different diets, or between PyMT genotype groups (Figure 2). Male pups were slightly heavier than female pups at weaning. By 6 wk of age males were 29.5% (difference between the two means as a percentage of the lower mean) heavier than females (P<.0001), individuals on the high fat diet had a modest 3.6% increase in body weight compared to those fed the matched control fat diet (P<.0001), while PyMT genotype still had no effect. At 9 wk of age males were 23.2% heavier than females (P<.0001), high fat diet led to a 5.3% (1.9 gram) increase in body weight (P<.0001), and the presence of PyMT was associated with a 3.5% increase in body weight (P<.001). At sacrifice (approximately 11 wk for females and 14 wk for males), there were 8.6% (3.3 grams) and 8.0% (3.1 grams) increases in body weight for F_2 mice fed high fat diet (P<.0001) and with PyMT (P<.0001), respectively. No interactions for body weight were detected at 3 or 6 wk. The only two way interaction detected for 9 wk weight was for sex x PyMT (P<.05), in which case females with PyMT weighed 6.1% (2.0 grams) more than females without PyMT, while PyMT only led to a 1.3% increase in male weight. The larger weight in animals with PyMT was most likely due to tumor mass. A similar gender x PyMT interaction was detected for weight at sacrifice (P<.0001), whereby the PyMT effect was greater in females than in males (14.5% and 2.8% increases, respectively), as a result of greater tumor mass in females relative to males.

Lean mass and adiposity were evaluated at 7 wk of age to examine short-term dietary effects and obtain baseline measures for the PyMT mice near or just after onset of mammary tumors, but prior to significant tumor development. Male mice had 27.8% (P<.0001) more

lean mass than females, whereas high-fat diet led to a modest 5.4% (P<.0001) increase (Figure 3). Presence of PyMT was associated with a slight (2.0%; P<.05) increase in lean mass (Figure 3). Measurements of adiposity revealed that female mice had an 8.4% increase in body fat relative to males (P<.0001), while high-fat diet led to an increase of 6.4% relative to the control fat diet (P<.01, Figure 4). There were no effects of PyMT on body fat in either female or male F_2 mice.

The main effects of PyMT, diet and gender were all significant for lean mass measured at sacrifice (P<.0001, Figure 3). A gender X PyMT interaction for lean mass at sacrifice was detected (P<.0001), in which PyMT males had a 2.6% increase and PyMT females had a much higher 28.5% increase in lean muscle mass relative to males and females without the transgene (Figure 3). For adiposity measurements at time of sacrifice, the main effects of diet (11.5% increase due to high fat; P<.0001) and PyMT (13.6% increase due to presence of PyMT; P<.0001) were significant (Figure 4). A significant interaction was detected between gender and PyMT (P<.0001), whereby females with PyMT had 35.6% relatively less body fat percentage compared to females without PyMT, while males with PyMT had slightly increased body fat percentage relative to males without PyMT (Figure 4). To ensure that the decrease in percent body fat at sacrifice was not a product of the large increase in body weight associated with tumor development, total fat in grams at sacrifice was calculated. The main effects of diet (21.8% increase due to high; P<.0001) and gender (26.4% increase in males vs. females; P<.05) were significant (Figure 4). A significant interaction was detected between gender and PyMT (P<.0001), whereby females with PyMT had 19.7% relatively less body fat compared to females without PyMT, while males with PyMT had slightly increased body fat relative to males without PyMT (Figure 4).

In addition to whole body composition, liver weight (all mice) and weights of the right epididymal (all males) or perimetrial (females without PyMT) fat pads were measured at sacrifice. Main effects for raw liver weight were all significant with the high fat diet, males and PyMT genotype all resulting in increased liver size (Figure 5). A significant gender x PyMT interaction was detected ($P<.0001$). Males and females with PyMT had livers that were 22.0% and 4.1% larger than mice of the same gender without PyMT. When livers were evaluated as a percentage of body weight, only the main effects of diet and PyMT were significant, with control fat diet and presence of PyMT both resulting in increased percent liver weight (Figure 5). Similarly to what was observed for raw liver weights, there was a significant gender x PyMT interaction effect ($P<.05$).

A strong dietary effect was detected for raw weight of perimetrial fat pad in females without PyMT (45.9% larger on high-fat diet vs. control fat diet; $P<.0001$; Figure 5). A comparable diet effect was seen for epididymal fat pad weight in males (Figure 5) but no influence of PyMT genotype was found (fat pad weight was not measured in females harboring PyMT). When fat pad weights were reanalyzed as percentage of body weight, strong dietary effects were still detected with higher values for mice fed the high fat diet (36.6% and 49.7% increase for high-fat vs. control fat diet in females and males respectively; $P<.0001$; Figure 5). No two way interactions were detected for fat pads measured as raw weights or as a percentage of body weight.

High positive phenotypic correlations were detected among all the body weight traits, although those between weight at 3 wk of age and the other time periods were of smaller magnitude (Table 3). Fatness measured at 7 wk of age and at sacrifice had a correlation of ~0.5. Likewise, correlations among body fat (measured by DEXA) and the fat pad

measurements were strongly positive. Correlations between body weight and fatness were positive and of moderate magnitude.

QTL evaluation: A total of 38 significant QTL (P<.05 genome-wise) were detected in the F_2 population (Table 5). These 38 QTL can be separated into four different categories, those associated with body weight, liver weight, lean mass and adiposity. For body weight, 18 QTL were identified across the four time points measured. Two were identified for weight at 3 wk of age, six for weight at 6 wk of age, five for weight at 9 wk of age and five for weight at sacrifice. QTL with the largest effects were mapped to MMU2, 9 and 11. All QTL effects were additive, with the M16i allele associated with increased body weight values.

Four QTL were detected for liver measurements, three for raw liver weights and one for liver as a percentage of body weight. The QTL detected for raw liver weight mapped to MMU2, 9 and 10 explaining 6.5%, 6.4% and 5.0% of the residual variance, respectively. Each demonstrated additive effects and in all cases the M16i allele was associated with increased liver weight. One QTL detected for liver as a percentage of body weight mapped to MMU9 and the effect was relatively small. This QTL exhibited an additive effect with the FVB allele associated with increased percent liver weight.

Analysis of lean mass revealed 7 QTL, 4 for lean mass at 7 wk of age and 3 for lean mass at sacrifice. The QTL at 7 wk mapped to MMU2, 6, 10 and 11, with the QTL on MMU2 explaining the largest percentage of residual variance (11.6%). The QTL detected for lean mass at sacrifice mapped to MMU2, 10 and 11 with LOD scores of 4.9, 4.34 and

3.77 respectively. Each QTL for lean mass exhibited an additive effect in which the FVB allele was associated with an increase in lean body mass relative to the M16i allele.

Adiposity was measured as total body fat (as a percentage of body weight) at 7 wk of age and at sacrifice, total fat at sacrifice (as grams of adipose), while a fat pad weight (raw weight and as a percentage of body weight) was also measured at sacrifice. Three QTL were found for adiposity at 7 wk on MMU2, 9 and 11. Four were detected for adiposity at sacrifice (2 for total fat and 2 for percent body fat) mapping to MMU2 and 9 in similar regions as those detected for 7 wk of age, with respective LOD scores of 4.97 and 4.5 for percent body fat and 10.65 and 5.43 for total fat. An additional two QTL were identified for fat pad measurements, one for raw fat pad weight and the other for fat pad as a percentage of body weight. Both fat pad weight QTL were located on the distal portion of MMU2 near the adiposity QTL, explaining 6.1% and 3.3% of the residual variance, respectively. Three of the QTL for adiposity (adiposity at 7 wk on MMU2 and 11 and raw fat pad weight on MMU2) were additive with the M16i allele leading to an increase in fat, two (fat pad as a percentage of body weight on MMU2 and adiposity at sacrifice on MMU2) were additive with the increasing allele inherited from FVB, while the remaining QTL (percent fat at 7 wk and total fat on MMU9 and percent fat at sacrifice on MMU2 and 9) exhibited dominance gene action.

Pleiotropy: The close proximity of QTL peak positions detected among highly correlated phenotypes suggested that a single locus may be affecting multiple traits. Results of a formal pleiotropy test revealed that of the original 38 QTL detected in the F_2 population, 9 distinct loci were present. Of these 9 loci, only two represented single traits (3 wk weight on MMU8 and 6 wk weight on MMU13, Table 5). The other 7 loci influences multiple phenotypes distributed across 5 different chromosomes (Figure 6). Typically, adiposity QTL

clustered together, while the body weight QTL clustered with liver and lean mass QTL, mirroring patterns of phenotypic correlations.

Diet and Gender X QTL Interactions: Given that PyMT leads to tumor development, we initially analyzed QTL for weight, lean mass and adiposity within PyMT subclasses to see if presence of tumors impacted QTL detection for body weight and composition. However, differential QTL did not appear to be present amongst PyMT genotypes, so all subsequent analyses of QTL x diet and QTL x gender interactions were performed across the full population. This revealed 17 QTL with significant interactions with diet and 7 QTL with significant gender interactions (Table 5). Interactions were identified for post 3 wk body weights and adiposity at seven weeks. Of the 17 QTL x diet interactions, 10 resulted from the detection of a significant QTL within one diet but not the other (8 of these had significant effects in control fat fed mice, while only 2 had a significant effect in high fat fed mice). The remaining QTL x diet interactions were caused by differential allelic effects within the two diets. Examples of two QTL with diet interactions were those found on chromosome 9 and 13 for body weight at 9 wk (Figure 7A and B), the first of which was caused by the presence of a significant QTL only in control mice, while the other was caused by the opposite scenario. Of the 7 QTL x gender interactions, only one resulted from the detection of a significant QTL within one gender but not the other (Figure 7C), while the remaining 6 interactions resulted from differential allelic effects within the two genders. While 24 interactions were detected between QTL and either sex or diet, several of these likely represented the same locus due to pleiotropy, as evidenced by patterns of interactions for specific traits. For example, the pleiotropic QTL affecting lean mass at 7 wk and body weights at 9 wk and at sacrifice on MMU6 exhibited a QTL x diet interaction for each trait.

Discussion:

Although the primary goal of these studies is to identify QTL and dietary fat x QTL interactions that influence mammary tumor development and subsequent metastatic activity, description of body weight and body composition phenotypes and QTL is important for several reasons. First, we show that the M16i x FVB/NJ-TgN (MMTV-PyMT)^{634Mul} F_2 population exhibited a broad range of phenotypic variation for body weight and fatness, caused by both segregating QTL and response to feeding of high versus control fat diets. This demonstrates that the cross we developed would be informative for testing the relationships between obesity and cancer, as described in the companion paper (Gordon et al., Companion Paper). Second, we have identified several novel findings, including specific QTL and QTL x diet and QTL x gender interactions that contribute to the growing knowledge base regarding polygenic control of body weight regulation. And third, with particular emphasis on traits measured at sacrifice, we have for the first time evaluated the effects of cancer on the genetic control of body weight and body composition, evaluated as QTL x PyMT interaction effects.

We have previously performed QTL detection in a variety of crosses using the M16 and M16i models of obesity (Allan et al. 2005; Rocha et al. 2004a, b; Yi et al. 2006). The present cross involves M16i and FVB, the latter being a seldom used line in QTL experimentation for traits related to energy balance. When comparing the present results with those published using M16 or M16i crosses, we found three QTL that had not been previously identified, explaining variation for adiposity at 7 wk of age and at sacrifice, and likely representing a single underlying gene on proximal MMU9 as indicated by pleiotropy testing. Independent studies using a variety of mouse crosses have found QTL for adiposity

in this region (see Table 1, (McDaniel et al. 2006)). There are several lines of evidence indicating that the MMU9 QTL identified in this M16i x FVB intercross could represent the *Obq5* gene causing the QTL originally described by Taylor et al., (1999). First, the QTL have been mapped in close proximity to each other. Second, *Obq5* and our MMU9 QTL for adiposity at sacrifice both have interactions with gender, with significantly greater phenotypic impact in females than in males. And third, *Obq5* impacts gonadal fat to a lesser extent than other fat depots (Taylor et al. 1999), and in the current study the MMU9 QTL was significant for total body fat but not for gonadal fat pad as a percentage of body weight.

Of the remaining QTL we detected, those with the biggest effects were located on the distal portion of Chromosome 2. This region has been routinely implicated as having large effects on body composition in mice (Allan et al. 2005; Horvat et al. 2000; Ishikawa et al. 2005; Jerez-Timaure et al. 2004; Rocha et al. 2004a, b).

Effect of high-fat diets on body composition in mice is variable, some strains having relatively little response while others are strongly impacted (Surwit et al. 1995; Svenson et al. 2007; West et al. 1992). In accordance with previous studies with the parental lines we used (FVB, (Yakar et al. 2006) and M16, (Allan et al. 2004)), we observed that M16 x FVB F_2 mice fed a high fat diet had modestly increased body weight and adiposity relative to those on the control fat diet. Despite the presence of widespread genetic variation in dietary-induced obesity, most QTL studies have been performed in the presence of either high or control fat diets, and thus only a few experiments have been suitably designed to analyze QTL x diet interaction effects. For example, Coulter et al., (2003) evaluated diet (high and low fat) x QTL interactions for Pgc-1α and Ucp1 levels in adipose, and detected scenarios in which QTL were found at a particular position in the presence of one diet but not the other.

This type of QTL x diet interaction was prevalent in the present study. Interestingly, we found that in the majority of cases a QTL for body weight or fat was detected in mice fed the control fat diet and not in those fed a high fat diet. This is essentially the opposite scenario to what was observed by Cheverud et al., (2004), who examined QTL in a set of LG x SM RIL's fed either high or control fat diets.

QTL x diet interactions are likely caused by differences in gene expression, either at the QTL itself or at a gene or genes regulating the QTL in *trans*, in response to the different diets. Diet-induced differential gene expression is a common phenomenon (e.g. (Jump and Clarke 1999; Kennedy et al. 2007)). QTL x gender interactions would likely be caused by similar mechanisms, where physiological changes between males and females lead to altered gene expression. Indeed, Yang et al., (2006) recently showed wide-spread sexually dimorphic gene expression across multiple tissues in mice. In the present study, several instances of QTL x gender interaction were detected. In the majority of cases, the interaction was caused by differential QTL effects within males and females, as opposed to finding a significant effect in one gender and not in the other. Such findings are not uncommon in QTL analysis for body weight and adiposity in mice (e.g. (Taylor et al. 1999)).

Understanding QTL x diet and QTL x gender interactions is important not only for understanding the genetic architecture of complex traits, but also for several practical reasons. First, the presence of such interactions should impact experimental design for QTL detection and discovery. Failure to appropriately account for interactions could lead to reduced power and increased type-II errors. Second, these interactions have relevance for development and application of genomics-based diagnostic and therapeutic tools.

Presence of the PyMT transgene in F_2 mice was associated with increased body weight when controlled for diet. Since this model of mammary tumor development is not previously thought to be associated with cachexia (Lifsted et al. 1998), the increase in body weight is likely due to the added weight of harbored tumors. Relative liver weights were larger in F_2 mice with PyMT, potentially due to the liver's function in detoxification of blood in the presence of tumor induced toxins and potentially circulating tumor cells (Canning et al. 2006).

In contrast, presence of PyMT led to a reduction in body fat in female F_2 mice. It is possible that females with tumors consume less food and utilize energy stores in an increased manner. Richardson and Davidson (Richardson and Davidson 2003) noted that humans with cancer can have increased energy demands and expenditures that can result in the loss of adipose tissue. den Broeder et al., (2001) found that the presence of a solid tumor in children was associated with an increased basal metabolic rate. Although most studies on cachexia have concentrated on muscle wasting, Ryden and Arner (2007) recently provided a synopsis of studies focusing on the loss of adipose tissue, concluding that adipocyte lipolysis is an important factor in the cachexic process.

The modest adipose-based cachexia we observed could also be a response to signaling factors secreted by the tumors (Bing et al. 2006), which in the case of mammary tumors are localized to areas with large quantities of adipose tissue. Speculatively, two possible candidates are lipid-mobilizing factor and zinc-α_2-glycoprotien, both of which are involved in initiation of lipolysis of adipose tissue (Russell and Tisdale 2002; Russell et al. 2004).

Despite the significant effects of PyMT increasing body weight, and decreasing body fat, in the F_2 population, there were no significant QTL x PyMT interaction effects on any of the traits studied, including weight and adiposity at sacrifice when the mammary tumor load was highest. This indicates that there were no QTL underlying variation in body weight and adiposity whose effects (and presumably, expression) were caused and/or altered either directly by PyMT expression or indirectly by the presence of mammary tumors.

While substantial progress has been made in understanding the genetic underpinnings of obesity in mice (Pomp 2007), and genetic variation in response to high-fat diet has been well characterized (e.g. (Taylor and Phillips 1997; West et al. 1994a; West et al. 1994b)), QTL x diet interactions have received very little research attention. Here we have shown that such interactions contribute significantly to the genetic architecture of body weight regulation, and that QTL results must be interpreted within the specific dietary environment(s) placed on the experimental samples. As pointed out by Schadt et al., (2003b), transcriptional responses to dietary fat intake represent significant heterogeneity that clearly demonstrates the complexity of underlying traits such as obesity. These findings highlight the relevance of accurately modeling the environmental exposures of human populations when conducting mouse genetical studies.

Table 2.1: Nutritional Composition of the high fat and matched control fat diets*

		D12451	(High Fat)			D12450B	(Control)
		gm%	**kcal%**			**gm%**	**kcal%**
Protein		24	20			19.2	20
Carbohydrate		41	35			67.3	70
Fat		24	45			4.3	10
	Total		100		**Total**		100
	kcal/gm	4.73			**kcal/gm**	3.85	
Ingredients		**gm**	**kcal**			**gm**	**kcal**
Casein, 80 Mesh		200	800			200	800
L-Cystine		3	12			3	12
Corn Starch		72.8	291			315	1260
Maltodextrin 10		100	400			35	140
Sucrose		172.8	691			350	1400
Cellulose, BW200		50	0			50	0
Soybean Oil		25	225			25	225
Lard		177.5	1598			20	180
Mineral Mix S10026		10	0			10	0
DiCalcium Phosphate		13	0			13	0
Calcium Carbonate		5.5	0			5.5	0
Potassium Citrate, 1 H20		16.5	0			16.5	0
Vitamin Mix V10001		10	40			10	40
Choline Bitartrate		2	0			2	0
FD&C Red Dye #40		0.05	0	Yellow Dye #5		0.05	0
	Total	858.15	4057		**Total**	1055.05	4057

*Diets were formulated and by E. A. Ulman, Ph.D., Research Diets, Inc. (New Brunswick, NJ).

Table 2.2: Composition of vitamin and mineral mixes in the high fat and matched control fat diets*

Vitamin Mix V10001				**Mineral Mix S10026**		
Ingredient	**gm**	**10 gm**		**Ingredient**	**gm**	**10 gm**
Vitamin A Palmitate 500,000 IU/gm	0.8	4,000 IU		Sodium Chloride 39.3% Na, 60.7% Cl	259	Na 1.0 gm Cl 1.6 gm
Vitamin D3 100,000 IU/gm	1.0	1,000 IU		Magnesium Oxide 60.3% Mg	41.9	Mg 0.5 gm
Vitamin E Acetate 500 IU/gm	10.0	50 IU		Magnesium Sulfate, 7 H2O 9.87% Mg, 13.0% S	257.6	S 0.33 gm
Menadione Sodium Bisulfite 62.5%Menadione	0.08	0.5 mg		Chromium K Sulfate, 12 H2O 10.4% Cr	1.925	Cr 2.0 mg
Biotin, 1.0%	2.0	0.2 mg		Cupric Carbonate, 57.5% Cu	1.05	Cu 6.0 mg
Cyancocobalamin, 0.1%	1.0	10 ug		Sodium Fluoride, 45.2% Fl	0.2	Fl 0.9 mg
Folic Acid	0.2	2 mg		Potassium Iodate, 59.3% I	0.035	I 0.2 mg
Nicotinic Acid	3.0	30 mg		Ferric Citrate, 21.2% Fe	21.0	45. mg
Calcium Pantothenate	1.6	16 mg		Manganous Carbonate, 47.8% Mn	12.25	59. mg
Pyridoxine-HCl	0.7	7 mg		Ammonium Molybdate, 4 H2O, 54.3% Mo	0.3	Mo 1.6 mg
Riboflavin	0.6	6 mg		Sodium Selenite, 45.7% Se	0.035	Se 0.16 mg
Thiamin HCl	0.6	6 mg		Zinc Carbonate, 52.1% Zn	5.6	Zn 29. mg
Sucrose	978.42			Sucrose	399.105	
Total	1000			Total	1000	

*The vitamin (V10001) and mineral mix (S10026) formulas were provided by Research Diets, Inc. (New Brunswick, NJ)

Table 2.3: Phenotypic correlations (and associated significance values adjusted for multiple comparisons) for traits measured in the M16i x FVB/NJ-TgN (MMTV-PyMT) ^{634Mul} F_2 population. Correlations are adjusted for the fixed effects of gender, diet and PyMT.

Trait	6W	9W	SW	PF7	PFS	FS	FP	PFP	LIV	PLIV	LM7	LMS
3W	0.45	0.39	0.35	0.19	0.11	0.22	0.24	0.20	0.25	0.01	0.33	0.26
	<.0001	<.0001	<.0001	<.0001	<.05	<.0001	<.0001	<.0001	<.0001	NS	<.0001	<.0001
6W		0.71	0.60	0.46	0.27	0.44	0.41	0.35	0.45	0.06	0.50	0.44
		<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	NS	<.0001	<.0001
9W			0.82	0.47	0.31	0.56	0.58	0.47	0.68	0.16	0.65	0.64
			<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.001	<.0001	<.0001
SW				0.44	0.28	0.62	0.72	0.56	0.78	0.13	0.54	0.78
				<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.01	<.0001	<.0001
PF7					0.50	0.57	0.58	0.55	0.22	-0.14	0.19	0.20
					<.0001	<.0001	<.0001	<.0001	<.0001	<.01	<.0001	<.0001
PFS						0.92	0.58	0.55	0.22	-0.14	0.17	-0.25
						<.0001	<.0001	<.0001	<.0001	<.01	<.0001	<.0001
FS							0.89	0.80	0.31	-0.17	0.10	0.35
							<.0001	<.0001	<.0001	<.0001	<.05	<.0001
FP								0.96	0.41	-0.22	0.32	0.34
								<.0001	<.0001	<.0001	<.0001	<.0001
PFP									0.28	-0.21	0.24	0.20
									<.0001	<.0001	<.0001	<.001
LIV										0.70	0.52	0.79
										<.0001	<.0001	<.0001
PLIV											0.25	0.37
											<.0001	<.0001
LM7												0.46
												<.0001

[a]3W: 3 week weight, 6W: 6 week weight, 9W: 9 week weight, SW: weight at sacrifice, PF7: percent fat measured at 7 weeks of age, PFS: percent fat measured at sacrifice, FS: Fat measured in grams at sacrifice, FP: raw weight of fat pad measured at sacrifice, PFP: fat pad as a percent of body weight, LIV: raw weight of liver at sacrifice, PLIV: liver as a percent of body weight, LM7: lean body mass at 7 weeks of age, LMS: lean body mass at sacrifice.

Table 2.4: List of 124 SNPs used in the final F_2 linkage map with known physical locations (Mb) from Ensembl and estimated linkage (cM) positions.

SNP Name	Chr	Mb Location	cM Location[a]	SNP Name	Chr	Mb Location	cM Location[a]
Rs6333200	1	6.2	3.9	rs3667341	8	131.5	74.0
Rs3659806	1	23.4	11.7	mCV24465575	9	6.7	4.2
Rs6401503	1	35.8	15.8	rs3665911	9	35.9	18.1
rs3696088	1	41.1	20.5	rs3703593	9	62.3	30.0
rs6181164	1	47.9	24.8	rs3714992	9	84.6	41.8
rs4222426	1	72.7	34.1	mCV24631499	9	106.5	55.6
mCV24377815	1	93.7	43.7	rs3692532	9	125.6	69.1
rs6163037	1	115.8	47.1	mCV25374719	10	8.3	5.2
rs6354736	1	128.2	52.0	rs3685588	10	10.7	14.1
rs3666261	1	128.2	60.5	rs8244562	10	61.3	28.7
rs3701630	1	159.8	67.1	rs3682523	10	69.4	39.5
mCV22660045	1	170.7	75.1	rs6284148	10	109.3	49.3
rs6208459	1	182.3	80.8	mCV24217147	10	117.5	57.5
mCV23204820	1	189.8	91.8	rs6171719	10	128.2	67.4
mCV25103560	2	2.9	1.8	rs3682937	11	4.6	2.9
rs4223189	2	61.7	40.2	rs3090752	11	26.7	12.3
rs3661811	2	93.0	51.7	rs8243015	11	51.9	26.5
rs8281186	2	120.0	61.9	rs3148189	11	70.8	38.2
rs3713952	2	146.4	73.2	rs3714299	11	92.9	51.0
rs8260429	2	163.2	83.2	rs4229101	11	102.0	58.0
mCV24846159	2	174.0	91.2	rs6192404	11	119.7	73.6
rs6310525	2	180.7	99.3	rs3689412	12	11.2	7.0
mCV25454657	3	7.1	4.4	rs3700857	12	50.7	18.7
rs3162526	3	36.9	14.2	rs6320805	12	75.6	32.8
rs6319642	3	56.1	23.7	rs3676085	12	88.6	44.5
rs6217010	3	95.5	37.8	rs3654706	12	112.8	61.7

mCV24793263	3	113.4	45.7	rs3686663	13	4.8	3.0
mCV23230498	3	133.9	59.6	mCV25144745	13	34.7	12.8
rs6157773	3	160.5	75.4	rs6411274	13	46.7	21.5
rs3693267	4	14.3	9.0	mCV22672997	13	70.2	32.9
rs3684104	4	38.5	22.0	rs6316213	13	88.3	42.5
mCV24089992	4	63.9	34.0	rs3711084	13	98.5	49.3
rs3712721	4	90.3	44.0	Rs3659752	13	114.0	60.3
rs3659850	4	110.3	51.7	rs6393665	14	11.0	6.9
rs3663950	4	135.0	71.5	rs3693175	14	27.0	16.6
rs3713685	4	147.3	79.5	rs4197422	14	49.5	26.7
rs3671575	5	21.8	13.6	rs3701623	14	67.5	39.9
rs6192958	5	48.1	29.3	rs3691209	14	98.8	49.8
rs3657238	5	74.3	40.6	rs6256423	14	106.9	57.0
mCV24416913	5	108.3	52.6	mCV25349597	15	10.3	6.4
mCV23328629	5	119.0	63.0	rs6165881	15	51.6	19.9
rs6377710	5	132.7	70.7	rs3667755	15	78.7	35.2
rs4225605	5	145.8	86.3	rs3722990	15	92.2	46.3
rs3710142	6	8.0	5.0	rs3722513	15	96.9	52.8
rs3662661	6	44.1	17.3	rs4170126	16	32.2	20.1
rs3699367	6	76.4	31.0	rs4206932	16	76.1	39.4
rs4226142	6	95.2	40.2	rs4215932	16	88.8	44.7
rs3722157	6	123.1	50.6	rs3693921	17	13.1	8.2
rs3673059	6	136.4	59.7	rs6278687	17	43.4	19.1
rs3677539	6	149.5	68.9	rs3712800	17	57.0	28.6
rs3688686	7	0.0	0.0	rs3710803	17	73.5	42.4
rs3023117	7	34.3	9.7	mCV23317493	17	84.4	54.5
rs8255275	7	46.0	18.9	rs6377403	18	11.4	7.1
rs3677657	7	68.5	31.0	rs3699816	18	39.5	17.9
rs3713052	7	101.6	42.0	rs3713935	18	58.7	29.1
rs4226910	7	122.3	54.8	rs3722524	18	68.9	38.6

rs3711570	8	3.6	2.2	rs8257628	19	4.0	2.5
rs3665640	8	33.5	20.5	rs3694495	19	20.0	12.8
rs3659789	8	55.7	30.4	rs3681148	19	30.7	23.3
rs6296403	8	92.0	40.5	rs3703918	19	42.8	32.8
rs3677807	8	106.0	52.6	rs6304326	19	53.4	42.6
rs3671292	8	118.4	60.7	rs3694663	19	58.1	51.6

[a] The cM position for the first (proximal) SNP on each chromosome was estimated based on its physical location (Build 34) using the formula Mb x 1.6 (Shifman et al. 2006)

Table 2.5: Descriptive statistics of QTL detected with genome-wide significance (P<.05).

Chr	Trait[a]	QTL Peak (cM)[b]	C.I.[c]	LOD[d]	Additive[e]	Dominance[f]	% Var[g]	Interaction[h]
2	LM7	47	42—54	14.02	1.43	-0.23	11.6	MC
2	LIV	56	43—74	7.55	0.16	0.01	6.5	
2	3W	58	13—85	4.23	0.28	0.63	3.6	
2	LMS	66	10.5—71	4.99	1.20	0.82	4.8	BD
2	6W	67	47—71	9.23	1.58	0.51	7.7	
2	9W	67	47—71	12.38	2.09	0.13	12.5	BD
2	SW	68	55—72	13.55	2.93	0.48	11.5	
2	PFP	72	6.5—90.5	2.61	-0.19	0.03	3.3	
2	FP	73	60—76	5.01	0.13	0.01	6.1	
2	FS	70	63—73	10.65	-1.43	-.20	10.0	BS
2	PFS	74	0—89	4.97	-0.18	-1.65	4.4	
2	PF7	75	65—82	6.88	1.35	-0.46	5.9	
6	6W	39	33—64	6.72	1.22	0.1	5.7	BS
6	LM7	41	27—65	5.0	0.84	0.18	4.3	BD
6	9W	44	35—58	5.88	1.34	-0.3	5.1	BD
6	SW	45	32—70	4.01	1.46	-0.62	3.5	BD
8	3W	50	21—58	4.24	0.4	0.31	3.6	
9	PF7	20	14—24	8.02	0.84	-1.74	6.8	MC
9	FS	21	11—52	5.44	-0.49	-0.50	5.0	MC
9	PFS	24	0—69	4.5	-0.29	0.55	4.0	F
9	9W	47	24—53	6.81	1.56	0.33	5.8	MC, BS
9	SW	47	21—54	6.08	1.96	0.52	5.3	MC, BS
9	PLIV	48	38.5—60	3.77	-0.14	-0.12	3.3	
9	LIV	49	42—55	7.34	0.15	-0.02	6.4	
9	6W	51	20—55	5.1	1.17	0.01	4.3	MC, BS
10	LMS	20	13—41	4.34	1.10	0.33	4.2	
10	9W	23	15—55	6.4	1.46	0.45	5.5	BD
10	LIV	23	16—45	5.76	0.13	0.02	5.0	
10	LM7	31	13—45	5.06	0.79	0.34	4.4	MC
10	6W	34	19—54	6.37	1.23	0.07	5.4	
10	SW	34	16—49	5.4	1.74	0.15	4.7	
11	LMS	27	16.5—52	3.77	1.03	-0.13	3.6	
11	LM7	35	23—42	5.87	0.64	-0.68	5.0	H
11	SW	35	25—46	6.1	1.72	-0.63	5.3	
11	PF7	37	28—63	4.37	1.04	-0.19	3.8	MC
11	9W	39	27—46	6.25	1.21	-0.81	5.4	MC
11	6W	40	30—45	8.41	1.13	-1.08	7.1	
13	6W	0	0—61	3.74	0.84	-0.49	3.2	H, BS

Table 2.5 Continued

[a] 3W: 3 week weight, 6W: 6 week weight, 9W: 9 week weight, SW: weight at sacrifice, PF7: percent fat measured at 7 weeks of age, PFS: percent fat measured at sacrifice, FS: Fat measured in grams at sacrifice, FP: raw weight of pad measured at sacrifice, PFP: fat pad as a percent of body weight, LIV: weight of liver at sacrifice, PLIV: Liver as a percent of body weight.

[b] Approximate peak QTL position. cM positions are adjusted to the linkage map presented in Table 4.

[c] 95% confidence interval for QTL peak (in cM).

[d] If an interaction was detected then the LOD score for the total genetic model was reported. If no interaction was detected then the LOD score for the genetic model without the interaction term was reported.

[e] Additive effect; a positive value indicates that the increasing allele originates from M16i.

[f] Dominance effect representing the heterozygous genotype in relation to the mean of the two homozygous genotypes: a positive value indicates that the heterozygote is larger than the mid-parent (mean of the parents).

[g] Percentage of phenotypic variance accounted for by the QTL effect

[h] Cause of the interaction. H: significant effect in high fat diet only, MC: significant effect in matched control fat diet only, BD: differential effects in high and matched control fat diets, F: significant effect in females only, BS: differential effects in females and males.

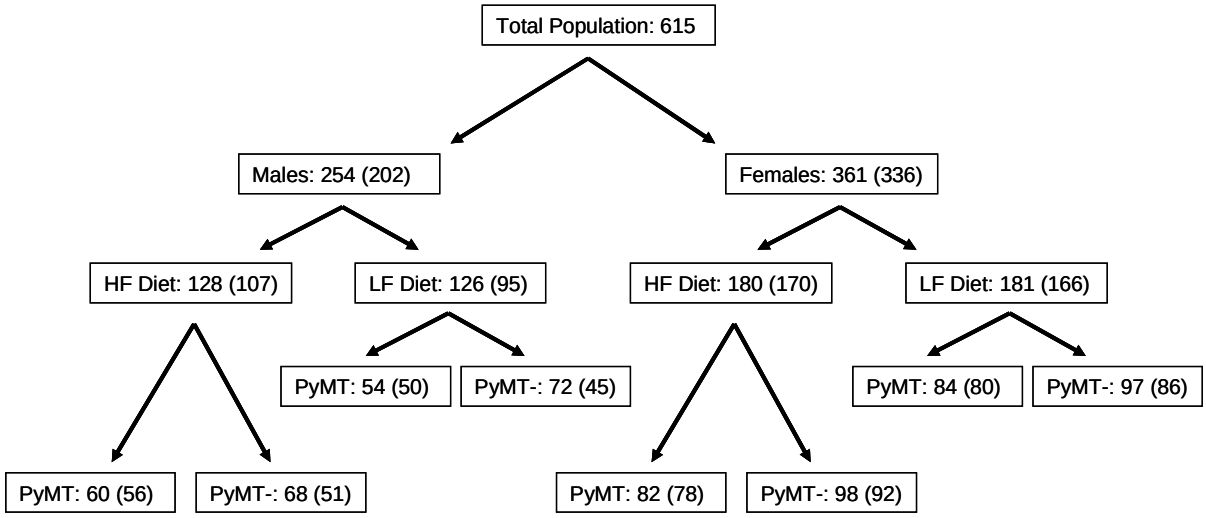
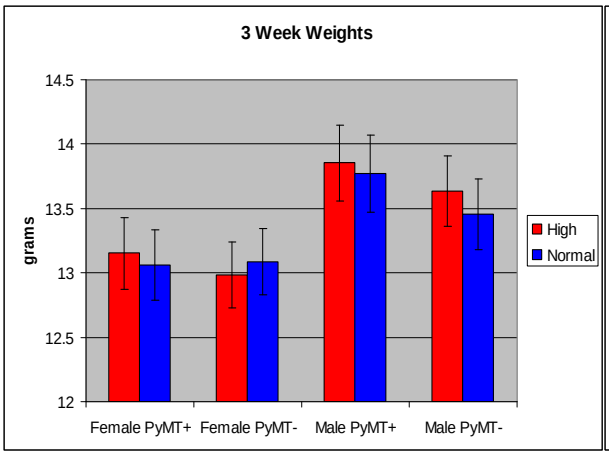
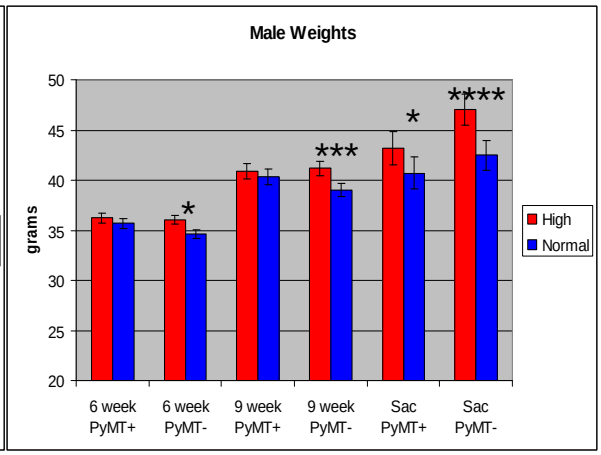


Figure 2.1: Numbers of animals phenotyped in each sub-group of the F_2 Population. HF designates animals fed a high fat diet, whereas LF designates animals fed the matched control fat diet. PyMT represents animals hemizygous for the PyMT transgene, while PyMT- represents individuals without the PyMT transgene. Numbers in parentheses indicate the number of animals genotyped per sub-group.

A.



B.



C.

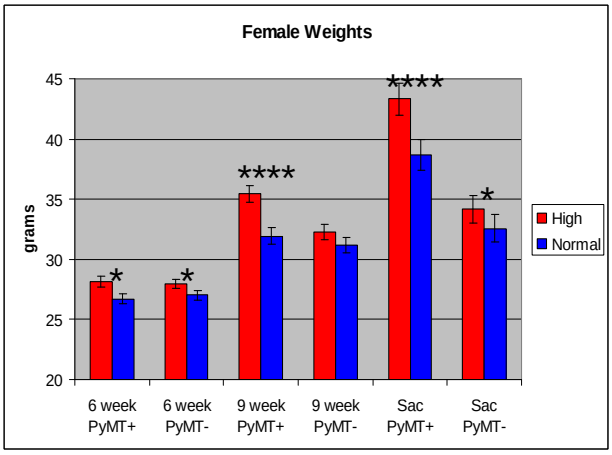
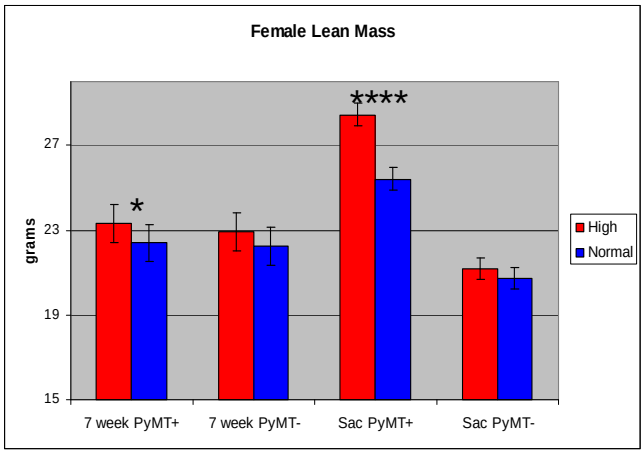


Figure 2.2: Least-squares means for body weight traits by gender and diet. PyMT+ signifies that animals were hemizygous for the PyMT transgene, whereas PyMT- indicates animals that did not harbor the PyMT transgene. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. A. Male and female body weight at 3 weeks of age. B. Female body weights at 6, 9 weeks of age and at sacrifice (~ 11 weeks). C. Male body weights at 6, 9 weeks of age and at sacrifice (~14 weeks). All * indicate the p value associated with the dietary effect: (*) Significant at P<.05, (**) Significant at P<.01, (***) Significant at P<.001, and (****) Significant at P<.0001.

A.



B.

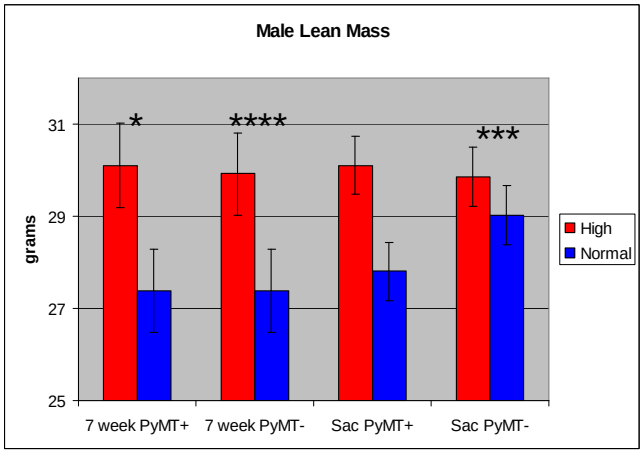
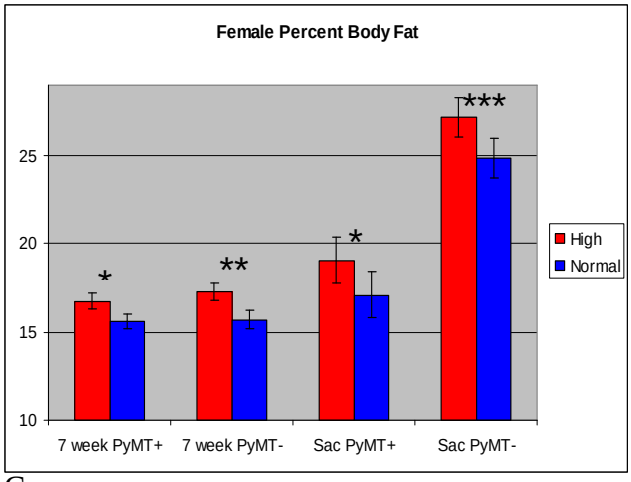
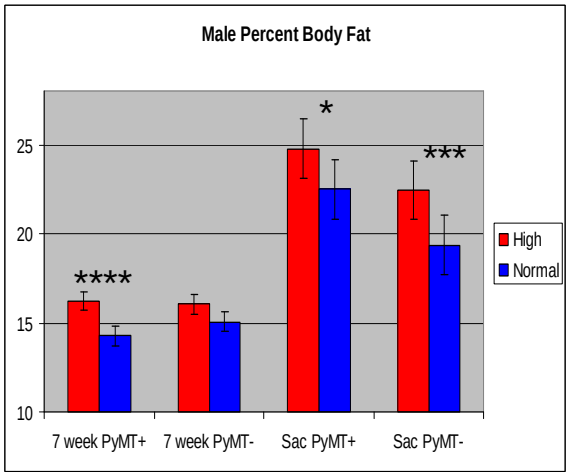


Figure 2.3: Least-squares means for lean mass measurements by gender and diet. PyMT+ signifies that animals were hemizygous for the PyMT transgene, whereas PyMT- indicates animals that did not harbor the PyMT transgene. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. A. Female lean body mass at seven weeks of age and at sacrifice (~ 11 weeks). B. Male lean body mass at seven weeks of age and at sacrifice (~ 14 weeks). All * indicate the p value associated with the dietary effect: (*) Significant at P<.05, (***) Significant at P<.001, and (****) Significant at P<.0001.

A.



B.



C.

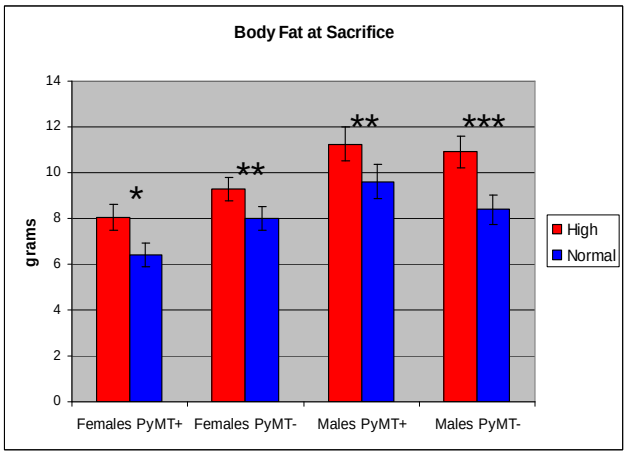
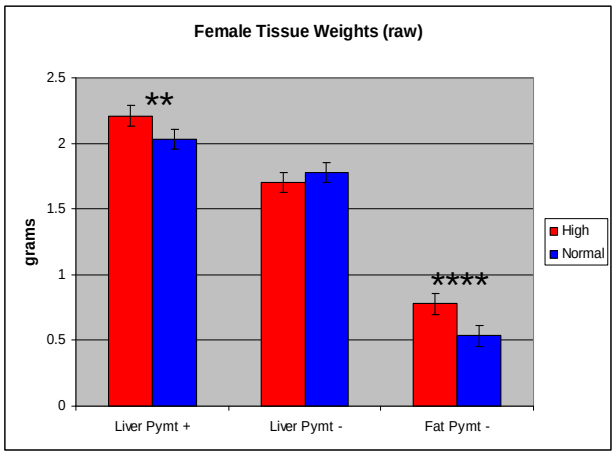
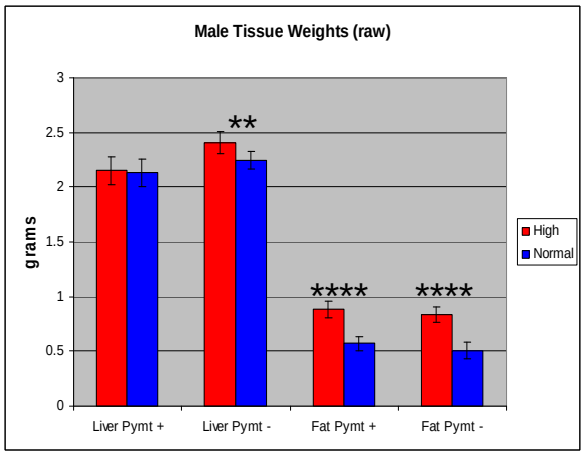


Figure 2.4: Least-squares means for adiposity traits by gender and diet. PyMT+ signifies that animals were hemizygous for the PyMT transgene, whereas PyMT- indicates animals that did not harbor the PyMT transgene. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. A. Female percent body fat at seven weeks of age and at sacrifice (~ 11 weeks). B. Male percent body fat at seven weeks of age and at sacrifice (~ 14 weeks). C. Body fat measured in grams at sacrifice. P values associated with the dietary effect: (*) Significant at P<.05, (**) Significant at P<.01, (***) Significant at P<.001, and (****) Significant at P<.0001.

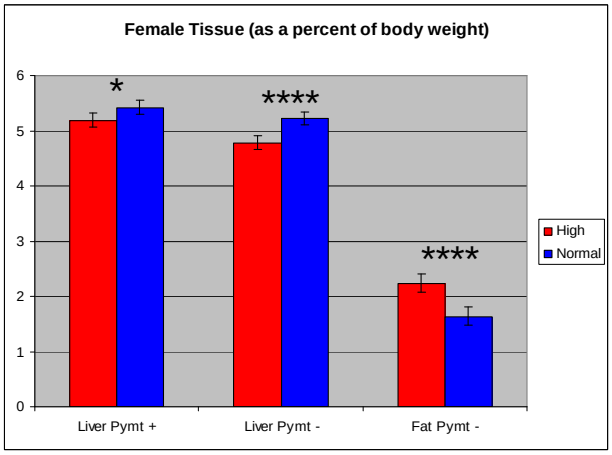
A.



B.



C.



D.

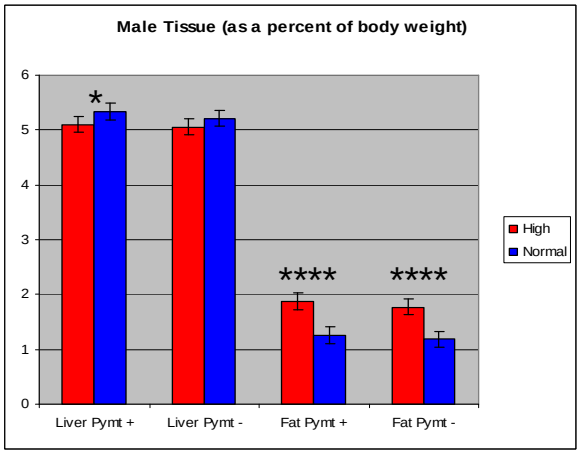


Figure 2.5: Least-squares means for liver and fat pad tissue weights by gender and diet. PyMT+ signifiesthat animals were hemizygous for the PyMT transgene, whereas PyMT- indicates animals that did not harbor the PyMT transgene. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. A. Female raw tissue weights. B. Male raw tissue weights. C. Female tissue as a percentage of body weight. D. Male tissue as a percentage of body weight. P values associated with the dietary effect: (*) Significant at P<.05, (**) Significant at P<.01, (***) Significant at P<.001, and (****) Significant at P<.0001.

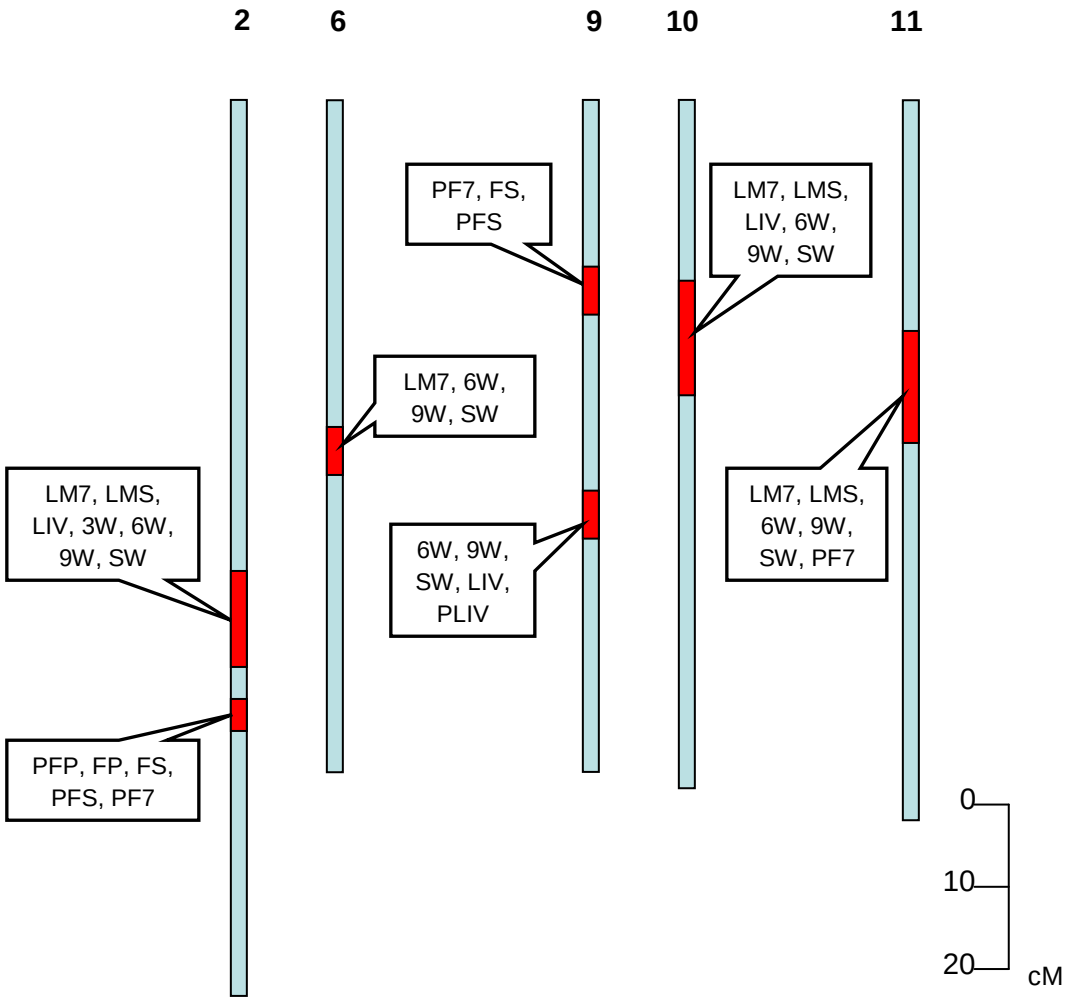


Figure 2.6: Chromosomal regions with QTL demonstrating pleiotropic effects. The seven regions designated in red (on MMU2, 6, 9, 10 and 11) represent segments of these chromosomes in which a single locus that accounts for variation in multiple traits is located. The respective textboxes indicate the traits that each of the loci impact. Trait symbols are as follows: 3W: 3 week weight, 6W: 6 week weight, 9W: 9 week weight, SW: weight at sacrifice, PF7: percent fat measured at 7 weeks of age, PFS: percent fat measured at sacrifice, FP: raw weight of fat pad measured at sacrifice, PFP: fat pad as a percent of body weight, FS: fat in grams at sacrifice, LIV: raw weight of liver at sacrifice, PLIV: liver as a percent of body weight, LM7: lean body mass at 7 weeks of age, LMS: lean body mass at sacrifice.

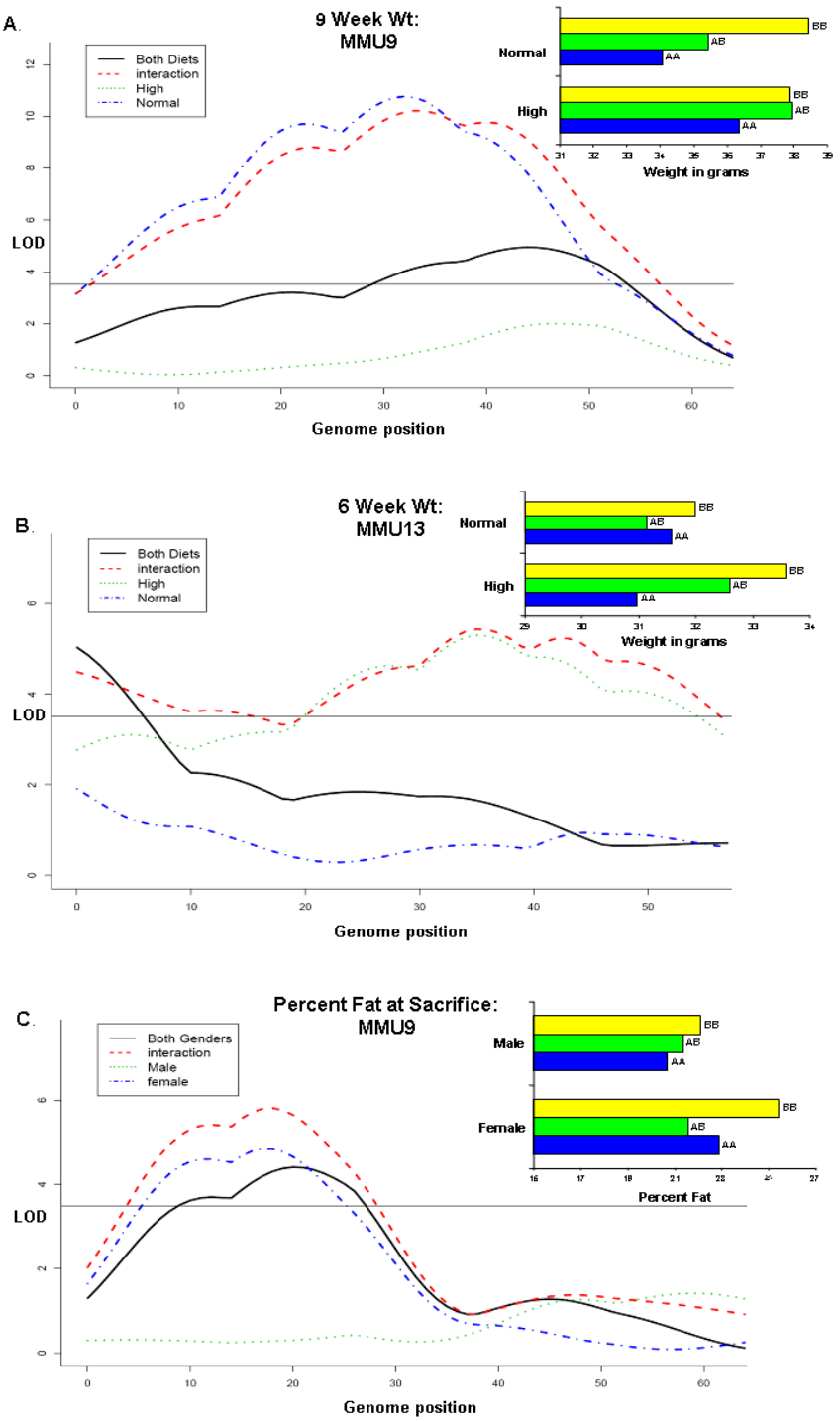


Figure 2.7: Examples of significant QTL interactions. Each figure includes graphical displays of the LOD curves that represent each component of the interaction. The solid black line is a representation of the resulting LOD curve when no interaction term is fitted in the model. The red dashed line represents the LOD curve when an interaction is included into the model. The remaining two lines represent the LOD curves of the two interaction components (either male/female or high fat diet/matched control fat diet). The threshold bar

represents an average threshold across the measured phenotypes. Additionally, the bar graphs represent the least-squares means of the trait of interest, for each allelic combination measured at the SNP marker closest to the QTL peak position (A allele: FVB; B allele: M16i). A. QTL x diet interaction for 9 week weight on chromosome 9. B. QTL x diet interaction for 6 week weight on chromosome 13. C. QTL x gender interaction for percent body fat at sacrifice on chromosome 9.

CHAPTER III

Genotype X Diet Interactions in Mice Predisposed to Mammary Cancer: II.

Tumors and Metastasis

Ryan Gordon, Kent W. Hunter, Michele la Merrill, Peter Sørensen, David W. Threadgill, Daniel Pomp

Mamm Genome. (2008) March;19(3):179-89

Abstract:

High dietary fat intake and obesity may increase the risk of susceptibility to certain forms of cancer. To study the interactions of dietary fat, obesity and metastatic mammary cancer, we created a population of F_2 mice cosegregating obesity QTL and the MMTV-PyMT transgene, we fed the F_2 mice either a very high fat or a matched control fat diet, and we measured growth, body composition, age at mammary tumor onset, tumor number and severity, and formation of pulmonary metastases. SNP genotyping across the genome facilitated analyses of QTL and QTL x diet interaction effects. Here we describe effects of diet on mammary tumor and metastases phenotypes, mapping of tumor/metastasis modifier genes, and the interaction between dietary fat levels and effects of cancer modifiers. Results demonstrate that animals fed a high fat diet are not only more likely to experience decreased mammary cancer latency, but increased tumor growth and pulmonary metastases occurrence over an equivalent time. We identified 25 modifier loci for mammary cancer and pulmonary

metastasis, likely representing 13 unique loci after accounting for pleiotropy, and novel QTL x diet interactions at a majority of these loci. These findings highlight the importance of accurately modeling not only the human cancer characteristics in mice, but also the environmental exposures of human populations.

Introduction:

According to National Institute of Health statistics, breast cancer is the most common cancer type and accounts for the second leading cause of cancer-related deaths in women, excluding skin cancers (National-Cancer-Institute 2005). Although breast cancer is thought to be a disease that primarily affects women, about 1% of all breast cancer cases diagnosed occur in men. It has been estimated that in 2007, 178,480 women and 2,030 men in the United States will be diagnosed with some form of breast cancer and that over 40,460 women and 450 men will die of this disease (American Cancer Society, 2007), typically from secondary metastatic disease (Sporn 1996). Furthermore, it has been estimated that each year breast cancer causes 502,000 deaths worldwide (World-Health-Organiziation 2006).

Differential susceptibility to breast cancer is thought to be mediated by two different types of cancer modifiers, either rare high-penetrance cancer-associated alleles or common low-penetrance polygenes (Balmain et al. 2003). While mutations in genes like *BRCA1*, *BRAC2*, *TRP53* and *PTEN* can result in increased breast cancer risk (King et al. 2003; Rohan et al. 2006; Song et al. 2006; Walsh et al. 2006), these inherited alleles contribute to only about 15-20% of all breast cancers (Balmain et al. 2003), suggesting that a polygenic etiology is responsible for the majority of breast cancer cases.

While genetics clearly influences breast cancer susceptibility, environmental factors play an equally important role. Recently, links between diet and incidence of breast cancer have been reported (Key et al. 2004). A broad spectrum of dietary components ranging from alcohol consumption (Terry et al. 2006) to fiber intake (Cade et al. 2007) has been investigated. Dietary fat has received increased attention of late, although reports to date have reached conflicting results. While some studies have shown a positive association between breast cancer and fat intake (Cho et al. 2003; Lee et al. 2005), others have failed to replicate these findings (Kim et al. 2006; Wakai et al. 2005). To clarify the relationship, we tested the hypothesis that mice predisposed to mammary cancer will show a positive association between mammary cancer development and dietary fat intake.

Numerous studies have investigated genetic or dietary links to breast cancer but few have focused on the gene x diet interactions that are likely to be major contributors to differential risk. We utilized an F_2 population of mice generated by intercrossing the FVB/NJ-TgN(MMTV-PyMT)^{634Mul} (PyMT) model of mammary cancer with the M16i polygenic obesity model and evaluated mammary cancer phenotypes that developed when the mice were maintained on one of two diets that differed in the level of dietary fat ((Gordon et al. 2008b), companion paper). Our objectives were to evaluate the phenotypic effects of dietary fat on tumor latency, severity and metastasis and to identify quantitative trait loci (QTL) and diet x QTL interactions associated with mammary cancer phenotypes.

Materials and Methods:

Population development: An F_2 population (n = 615) was generated by crossing M16i (Allan et al. 2005) and PyMT (Guy et al. 1992) mice in four replicate breeding cycles

(full details are provided in Gordon et al companion paper). F_2 mice were weaned at three weeks of age and randomly assigned, within litter, gender and genotype (PyMT or no PyMT) to receive one of two synthetic purified diets at four weeks of age. Mice had *ad libitum* access to water and their assigned feed, either a high fat diet (Research Diets D12451) containing 45% of total calories from fat, 20% from protein and 35% from carbohydrate or a matched control fat diet (Research Diets D12450B) containing 10% of total calories from fat, 20% from protein and 70% from carbohydrate. Prior to starting specialized diets all animals received *ad libitum* access to a standard rodent diet (Teklad 8604 Rodent Chow). Mice were evaluated for body weight and body composition as previously described ((Gordon et al. 2008b), companion paper).

Tumor analysis/Tissue collection: All F_2 mice that were confirmed carriers of the PyMT transgene were evaluated, starting at four weeks of age, to determine the age of onset for mammary tumors (TOID). All mammary glands from each individual were evaluated by palpation three times a week until tumor onset was determined, or until they were sacrificed at either 11 weeks (females) or 14 weeks (males were evaluated later because of slower development of cancer phenotypes). Upon dissection, total tumor number (TC) was recorded along with the combined weights of the inguinal tumors (TI) and the combined weights of the axillary tumors (TA). Total tumor weight was calculated by the addition of the total axillary and inguinal tumor weights (TTW). The following tissues were collected and snap frozen in liquid nitrogen: one inguinal and one axillary mammary tumor and one pulmonary lobe (the other being used for evaluation of metastasis).

Metastatic analysis: The metastatic burden was evaluated under a dissecting microscope by counting the total number of metastases present on the lung surface (MET).

The density of pulmonary metastases was determined in histological sections (AMD). One pulmonary lobe was fixed in 10% paraformaldehyde, embedded in paraffin, sectioned and stained with hemotoxylin-eosin. Three coronal nonadjacent sections of the lung, each separated by 100 µm, were prepared from each animal. The slides were examined using a Leica M420 Macroviewer with an Apozoom lens under 12x magnification. Three fields were scored for each slide, for a total of nine fields per animal. Pulmonary metastatic density was determined using a Leica Q500MC Image Analysis System, which aided in the elimination of alveolar space from the measurement of lung tissue area measurements to control for varying degrees of lung inflation at sacrifice. The metastasis index was measured as the number of multicellular metastatic lesions observed per square micron of lung tissue (AMN). Average metastasis size was calculated based on the total area of metastatic tissue on the slide divided by the number of metastases observed. All slides were read blind, and analyzed by a single operator to improve technical consistency.

Statistical Analysis: Proc Mixed REML in SAS (2002) was used to analyze all cancer phenotypes. The model for all male and female cancer traits except TOID contained a fixed effect of diet, random effects of replicate and dam, and used age at sacrifice as a covariate (TOID was analyzed without this covariate). Correlations (with Bonferroni corrected P-values) among traits were generated using the MANOVA procedure within Proc GLM in SAS (2002), with the fixed effect of diet.

QTL Analysis: As described previously ((Gordon et al. 2008b), companion paper), data were analyzed with the "F_2 inbred/Co-dominant Marker Analysis" option of the web-based program QTL Express (Seaton et al. 2002), fitting one QTL per chromosome with the genetic model for additive plus dominance effects. Given the dissimilarities in cancer

phenotypes between gender, males and females were evaluated as individual populations for all traits except TOID, which was also evaluated in a pooled gender population. In all analyses stratified by gender, the fixed effects of additive, dominance, replicate and diet as well as a QTL x diet interaction (interaction model) were included. In the pooled analyses the fixed effect of gender was included as well. An additional model was tested for each trait with identical fixed effects but without the interaction term (non-interaction model). To test whether QTL x diet interaction effects were significant, the sum of squares error and the degrees of freedom for the peak position for each QTL in both the interaction and non-interaction models were calculated and used to estimate an F statistic. If the interaction was not significant, the QTL x diet interaction effect was removed from the model. If an interaction was detected, then the mapping population was separated according to diet (animals fed either the high fat or matched control (lower) fat diet) and reanalyzed independently using the non-interaction model to elucidate the cause of the interaction. Genome-wide significance and suggestive thresholds for QTL effects were determined using permutation testing with 1000 iterations (permutated experiment-wide and chromosome-wide, respectively). A bootstrap procedure was used to estimate confidence intervals for QTL positions. The nonparametric bootstrap method described by Lebreton et al., (1998) was used to test for pleiotropy among multiple linked QTL.

Real Time PCR: PyMT expression was evaluated for all axillary tumors collected from the F_2 females using real time PCR to verify that differences observed in the phenotypes were not the result of varying levels of transgene expression. PyMT cDNA was amplified using primers AACGGCGGAGCGAGGAACTG and ATCGGGCTCAGCAACACAAG (Operon) and Taq polymerase (Qiagen). The real time

PCR results were evaluated using both the Proc Mixed and Proc Corr procedures of SAS (2002).

Results:

Phenotypic evaluation: Females fed the high fat diet were found to have a significantly earlier onset of mammary tumors, by approximately 3 days, than females fed the matched control fat diet (Figure 1A; $p < 0.001$). While results for the males also suggested an earlier onset of 2.4 days for high fat fed animals, this difference was not significant (Figure 1A). No significant dietary effect was detected in either the male or female populations for number of tumors (Figure 1B). The upper limit of tumor number that could be detected in mouse was ten, corresponding to the number of mammary pads. While many female mice had 10 tumors, the maximum we detected for any male was two.

After tumors were counted in females, they were removed and separated according to inguinal or axillary regions. Weights for these two regions were evaluated individually as well as together to form a total tumor weight. Females fed the control fat diet had 33.3% ($p < 0.0001$) (difference between the two means as a percentage of the lower mean) and 26.8% ($p < 0.01$) lower tumor weights in axillary and inguinal regions, respectively compared to females fed the high fat diet. When combined, this led to a reduction of 31.1% ($p < 0.001$) for total tumor weight (Figure 1C). No significant dietary effects were detected for male axillary tumor weight; no tumors developed in the inguinal region in males. In addition 20 females from replicate four were randomly selected to have their total tumor load measured for percent adipose in the tissue. These results (Figure 1D) suggested that animals fed the high

fat diet not only have larger tumor loads but their tumors contained less adipose tissue compared to tumors collected from animals fed the control fat diet (Figure 1D).

Less than half of females with PyMT (~42%) developed pulmonary metastases at the time of evaluation. Within these females, a 68.1% increase in initial surface count of metastases was observed in mice fed the high fat diet relative to mice fed the control fat diet (Figure 2; $p < 0.05$). For metastasis traits measured in lung sections, significant dietary effects were detected for average number of metastases (47.1% increase for animals fed the high fat diet; $p < 0.05$) and average density of metastases (46.1% increase for animals fed the high fat diet; $p < 0.05$) (Figure 2).

The correlation analysis of the female cancer traits (Table 1) was consistent with expectations. All measurements for tumor weights were highly correlated with each other and traits measuring metastasis had strong positive correlations. Additionally, the correlations among tumor weights and metastasis traits were all moderately positive, while tumor onset and tumor count were either weakly or not correlated with the other traits. However, the analysis of correlations among the cancer traits and the body composition traits for this population did reveal some interesting results (Table 2). Liver weight and lean mass at sacrifice were both highly correlated with tumor weights and a similar pattern was detected for body weight at sacrifice and tumor weights. Also, percent body fat measured at sacrifice and tumor weights had moderately negative correlations. In addition, the metastasis traits were positively correlated with both liver weight and sacrifice weight.

We further partitioned correlations by diet. Few changes due to diet were observed (data not shown), although reduction in power limited to some extent the ability to identify

significant correlations. A few examples of possible alterations in phenotypic relationships when mice are raised on high fat diet versus control fat diet include those between fat traits and metastasis traits. The correlation between amount of body fat at sacrifice and metastasis was not significant in the high fat diet group, but was significantly ($P<.05$) positive in the low fat group.

QTL evaluation: Two genome-wide significant QTL, mapping to Chr 1 and 9, were detected for tumor onset in days in the pooled population (Table 3). Each of these QTL had relatively large effects as shown by their contributions to the residual variance of 10.4% and 10.9%, respectively. When the data were evaluated at the suggestive $P < 0.05$ level, 23 additional QTL were detected (Table 3). Most of these QTL had high LOD scores (> 3.0), which likely did not reach genome-wide significance due to relatively small sample size. Given that we have shown that most, if not all, suggestive QTL for similar traits identified in our previous crosses were validated using chromosomal substitution lines (Lancaster et al. 2005), we include and discuss all detected QTL herein.

The QTL for tumor traits can be broken down into two categories, traits that are associated with the primary tumors and those that are associated with metastatic disease. A total of 15 QTL were detected across all traits associated with the primary tumors. Three QTL were detected for axillary tumor weight, located on Chr 9, 14 and 17. Alleles increasing tumor weight originated with both the M16 (Chr 9) and FVB (Chr 14, 17) lines. One QTL was detected for inguinal tumor weight on Chr 5 and three for tumor count at sacrifice on Chr 1, 9, and 14. The remaining eight QTL were all for TOID across seven different chromosomes (Table 3).

Analysis of our initial metastasis count at the time of sacrifice revealed four QTL in the female population that were located on Chr 1, 8, 11 and 19. The FVB allele was associated with increased values for the QTL detected on Chr 1 and 19, whereas the M16i allele was associated with increased values for Chr 8 and 11. The remaining metastatic QTL were detected for traits measured in a histological analysis of the formalin fixed lung of the PyMT positive females. Two were detected for average metastasis number, on Chr 8 and 13, and two for average metastasis density in similar locations. In all four cases the allele associated with increased metastasis was inherited from FVB.

Pleiotropy testing: The close proximity of QTL peak positions detected among highly correlated phenotypes suggested that a single locus may be affecting multiple traits. Results of a formal pleiotropy test revealed that of the original 25 QTL detected in the F_2 population, 13 distinct loci were present, with 6 pleiotropic loci representing many of the initial QTL detected (Table 3). For the most part, pleiotropic QTL influenced correlated traits within categories (e.g. tumor number/weight, or metastatic development). To determine if adiposity (companion paper) and female cancer phenotypes may be under joint control of a pleiotropy QTL, we first looked for overlap between QTL positions. Only 2 chromosomes (Chr 9 and 11) had overlapping QTL, and tests for these indicated a lack of pleiotropy.

Diet by QTL interactions: Testing for QTL x diet interactions was significant at 16 of the 25 cancer QTL identified (Table 3). Of these 16 interactions, four (for tumor onset, number of tumors at sacrifice and number of metastases at sacrifice) resulted from differential allelic effects within the two diets. The remaining QTL x diet interactions were a result of the detection of a significant QTL within one diet but not the other. Three of these had significant effects in mice fed the control fat diet, while nine had a significant effect in

mice fed the high fat diet. These interactions were detected amongst all traits evaluated except inguinal tumor weight and total tumor weight. Three examples of these interactions are depicted in Figure 4. The first two loci were located in close proximity to each other on Chr 8 for number of metastases at sacrifice and average metastasis density, resulting from the presence of a significant QTL only in mice fed high fat diet (Figure 4 A and B). The other locus displayed a similar interaction and was detected for the number of metastasis at sacrifice on Chr 19 (Figure 4C).

PyMT evaluation: In order to verify that the QTL and phenotypic differences detected resulted from both dietary and genetic variations were not influenced by varying levels of the PyMT transgene, we utilized real time PCR to measure PyMT expression levels. Results revealed that there were no significant differences in PyMT expression in the mammary tumors of females on the high fat diet versus those on the control fat diet (data not shown). Furthermore, correlation analyses confirmed that PyMT expression was not correlated with any measured phenotypic trait (data not shown).

Discussion:

There is a complex relationship between diet and genetics of mammary cancer susceptibility. While genetic predisposition to mammary cancer has been confirmed by identification of multiple, small-to-low effect familial risk factors, the evidence linking dietary fat to cancer susceptibility is less consistent. While results of some studies have suggested that there is no association between dietary fat intake and cancer development and progression (Salazar-Martinez et al. 2005; Smith-Warner et al. 2001), other studies have detected a positive association (Lee et al. 2005). To clarify the relationship, we tested the

hypothesis that mice predisposed to mammary cancer will show a positive association between mammary cancer development and dietary fat intake.

Our results show that animals fed a high fat diet are not only more likely to develop mammary cancer at an earlier time, but their tumors grow larger over an equivalent time. Increased levels of circulating estrogen as a result of increased adipose stores have been shown to be positively associated with age of cancer onset (Bray 2002; Key et al. 2003). In the current studies, mice fed the high fat diet not only had increased adipose percentages (percent body fat at seven weeks), but earlier onset of mammary cancer. Increased levels of insulin-like growth factor I (IGF-I) also have been implicated in the onset of tumors due to its ability to stimulate the proliferation in the mammary epithelium and inhibit apoptosis (Shi et al. 2004; Stoll 1998). Similarly, it has been suggested that obesity is associated with adverse features at diagnosis including larger tumor size, advanced grade, and increased nodal involvement as a result of increased levels of insulin (Boyd 2003). However, we did not detect a correlation between adiposity at seven weeks and cancer size or local progression. Nonetheless, we did observe a higher rate of metastasis in the high fat diet group than in the control fat diet cohort. These results were similar to those found in a study by Senzaki et al., (Senzaki et al. 1998), in which fat intake was positively associated with xenograft growth and metastasis. Consistent with these observations, Rose et al., (1991) found that animals fed a high fat diet were almost twice as likely to develop pulmonary metastasis.

Increased tumor load in mice fed the high fat diet was not the result of tumors with higher levels of adipose, as shown by the fact that tumors in mice fed high fat actually had lower percent fat content. This implies that the tumor epithelium may be more aggressive as a result of consumption of a high fat diet. Furthermore, the decreased fat content in tumors of

females on a high fat diet suggests that paracrine action of adipocyte signaling hypothesized elsewhere may not play a significant role in obesity-associated mammary cancer morbidity (Schaffler et al. 2007).

Genetic analysis of the F_2 population supported detection of several modifiers, some of which have not been previously reported including a modifier of tumor onset on Chr 1. We also confirmed previously detected modifiers such as the locus located on Chr 9 for tumor onset, encompassing the region harboring *Apmt2*, a locus associated with the acceleration of tumor latency (Le Voyer et al. 2000). In addition, the locus for tumor onset on Chr 7 appears to be in the same region as a previously detected suggestive locus for acceleration of tumor latency (Le Voyer et al. 2000), and a locus on Chr 17 to which we mapped a axillary tumor growth enhancer locus was recently implicated in harboring a tumor growth modifier (Lancaster et al. 2005). When our results were compared to the loci identified for metastasis by Hunter et al., (2001), it appears that the locus we identified on Chr 19 mapped to the same region as *Mtes1*. Additionally, we mapped a modifier to Chr 13 in a region previously implicated in metastasis (Hunter et al. 2001). Of the remaining loci, those modifiers residing in two regions appear particularly interesting. The first, located on Chr 1, is in close proximity to a locus detected for tumor onset suggesting that the modifier identified in this area represents the same underlying gene, which was supported by pleiotropy testing. The second, on Chr 8, influences three of the different metastatic traits and, because of their similarities and the pleiotropic effects detected in this region, it is appears that they too are all regulated by a common underlying gene.

Only a small number of the cancer modifiers detected were in similar locations to obesity QTL mapped in the same population ((Gordon et al. 2008b), companion paper), and

these do not appear to be due to pleiotropic loci. This may at first seem surprising, given the recent evidence that links obesity and cancer in humans (Hursting et al. 2007). Although this is just a single cross with limited power, it is possible that the obesity-cancer link has stronger environmental underpinnings relative to a genetic correlation. Our analyses were focused more on effects of dietary fat on cancer, and interaction between dietary fat and cancer modifiers.

Our experimental design supported testing the hypothesis that animals fed a high fat diet would have more advanced cancer phenotypes as a result of underlying gene x diet interactions. We found that 64% of all cancer modifiers detected had significant diet interactions, and of these interactions, nine were modifiers only detected in animals fed the high fat diet while three were detected in mice fed the control fat diet. These results implicate interactions of diet and modifier genes as a mechanism through which diet affects cancer.

As previously mentioned, the interaction detected with the Chr 19 locus controlling the number of metastasis counted at sacrifice mapped to almost an identical site as the *Mtes1* locus (Hunter et al. 2001). When we evaluated QTL for this trait without a diet interaction component it only yielded a suggestive modifier. However, when gene x diet interaction was added to the model, the modifier became highly significant with the LOD score rising from 2.56 to 4.21. These results suggest that the high fat diet may modulate the effect of the Chr 19 locus by increasing the development of pulmonary metastasis through the activity of this locus on a high fat diet. The work reported by Park et al., (2005) suggests that the *Sipa1* gene is a strong candidate for the *Mtes1* locus, which implies that a diet high in fat can alter the expression or activity of this gene. Such an interaction is plausible as *Sipa1* is correlated

to breast cancer metastasis (Crawford et al. 2006) and its binding partner Brd4 is implicated in the formation of subcutaneous fat (Houzelstein et al. 2002).

Another interesting interaction we detected was between diet and the Chr 9 modifier for tumor onset. The location of this modifier overlaps with the *Ampt2* locus, previously identified as a modifier of tumor onset (Le Voyer et al. 2000). However, unlike the locus on Chr 19, we were unable to detect any modifier at this position without the gene x diet interaction included in the model; when the interaction was fitted we detected a significant modifier at this location. When the two diets were evaluated separately, we observed that the LOD score in the high fat diet group was elevated above the pooled population, while the matched control fat diet suppressed detection of this modifier. The potential identification of the *Ampt2* locus as *Myc*, a gene shown to enhance tumor growth, suggests that a high fat diet could further enhance this gene's ability to cause increased tumor latency

For a small number of the interactions detected, we found that the control fat diet was responsible for increasing the significance of the modifier. However, at each of these modifier locations the differences in the LOD scores observed for the high and control fat diet populations, when evaluated individually, were relative small. Given that the majority of the interactions resulted from the detection of a significant modifier within the population that was fed a high fat diet and not within those fed control fat diet, we can speculate that mice predisposed to mammary cancer show a positive association between mammary cancer development and dietary fat intake due to diet-induced alterations in gene expression either underlying the actual modifier or regulating the modifier. How these alterations are manifested, e.g. through regulation of some global factors such as hormones (estrogen, IGF-1, IGFBP-3 (Kaklamani et al. 1999)), requires further investigation. However, these findings

enforce the importance of accurately modeling not only the human cancer characteristics in mice, but also the environmental exposures of human populations

Table 3.1: Phenotypic correlations (and associated significance values) for female cancer traits measured in the M16i x FVB/NJ-TgN (MMTV-PyMT) ^{634Mul} F_2 population.

Trait	TC	TA	TI	TTW	METS	AMN	AMD
TOID	-0.26 <.01	-0.23 <.01	-0.22 <.05	-0.25 <.01	-0.02 NS	-0.19 NS	-0.20 NS
TC		0.30 <.001	0.31 <.001	0.32 <.001	0.07 NS	0.04 NS	0.02 NS
TA			0.77 <.0001	0.97 <.0001	0.41 <.01	0.25 <.05	0.40 <.0001
TI				0.90 <.0001	0.34 <.05	0.21 NS	0.30 <.001
TTW					0.41 <.01	0.25 <.05	0.39 <.001
METS						0.59 <.0001	0.58 <.0001
AMN							0.67 <.0001

Correlations are adjusted for the fixed effect diet, and resulting P-values are adjusted for multiple comparisons. TOID: tumor onset measured in days, TC: number of tumors, TA: total weight of the axillary tumors, TI: total weight of the inguinal tumors, TTW: combined weight of axillary and inguinal tumors, METS: number of pulmonary metastasis counted at sacrifice, AMN: average metastasis number detected in internal sections, AMD: average metastatic density.

Table 3.2: Phenotypic correlations (and associated significance values) among female cancer and body composition traits measured in the M16i x FVB/NJ-TgN (MMTV-PyMT) ^{634Mul} F_2 population.

Trait	TOID	TC	TA	TI	TTW	METS	AMN	AMD
3W	-0.26 <.01	0.08 NS	0.20 <.05	0.12 NS	0.18 NS	-0.07 NS	0.13 NS	0.14 NS
6W	-0.09 NS	0.16 NS	0.32 <.001	0.33 <.05	0.35 <.001	-0.09 NS	0.09 NS	0.10 NS
9W	-0.35 <.0001	0.29 <.001	0.50 <.0001	0.44 <.0001	0.51 <.0001	0.11 NS	0.03 NS	0.29 <.01
SW	-0.26 <.001	0.33 <.0001	0.79 <.0001	0.73 <.0001	0.81 <.0001	0.23 NS	0.16 NS	0.37 <.001
PF7	-0.14 NS	-0.02 NS	0.11 NS	0.05 NS	0.10 NS	-0.22 NS	-0.07 NS	-0.06 NS
PFS	0.02 NS	-0.27 <.01	-0.51 <.0001	-0.58 <.0001	-0.57 <.0001	-0.28 <.05	-0.20 NS	-0.20 NS
FS	-0.23 <.01	-.07 NS	-0.14 NS	-0.23 <.01	-0.18 <.05	-0.13 NS	0.00 NS	0.01 NS
LIV	-0.29 <.0001	0.36 <.0001	0.77 <.0001	0.73 <.0001	0.80 <.0001	0.34 <.05	0.27 <.05	0.50 <.001
PLIV	-0.27 <.01	0.33 <.0001	0.51 <.0001	0.51 <.0001	0.54 <.0001	0.29 <.05	0.48 <.001	0.50 <.001
LM7	-0.37 <.0001	0.24 <.01	0.27 <.01	0.21 <.05	0.26 <.01	0.04 NS	0.26 <.05	0.26 <.01
LMS	-0.31 <.001	0.37 <.0001	0.89 <.0001	0.82 <.0001	0.91 <.0001	.034 <.01	.043 <.001	.042 <.001

Correlations are adjusted for the fixed effect diet, and resulting P-values are adjusted for multiple comparisons. TOID: tumor onset measured in days, TC: number of tumors, TA: total weight of the axillary tumors, TI: total weight of the inguinal tumors, TTW: combined weight of axillary and inguinal tumors, METS: number of pulmonary metastasis counted at sacrifice, AMN: average metastasis number detected in internal sections, AMD: average metastatic density, 3W: 3 week weight, 6W: 6 week weight, 9W: 9 week weight, SW: weight at sacrifice, PF7: percent fat measured at 7 weeks of age, PFS: percent fat measured at sacrifice, FP: raw weight of fat pad measured at sacrifice, PFP: fat pad as a percent of body weight, FS: fat in grams at sacrifice, LIV: raw weight of liver at sacrifice, PLIV: liver as a percent of body weight, LM7: lean body mass at 7 weeks of age, LMS: lean body mass at sacrifice.

Table 3.3: QTL detected at the experiment-wide and chromosome wide .05 levels and their respective statistics by chromosome.

Chr	Trait[a]	QTL Peak (cM)[b]	C.I.[c]	Add[d]	Dom[e]	Interaction[f]	LOD	Pleiotropy Group[g]
1	**TOID-A**	**4**	**4—89.5**	**-0.03**	**-3.51**	**BD**	**4.62**	**1**
1	MET-F	7	4—93	4.64	-3.11	BD	3.78	
1	TOID-M	17	4—37	-1.37	-9.9		3.74	1
1	TC-M	68	4—93	0.36	0.44	H	3.53	
1	TOID-F	69	4—90.5	0.31	-2.89		2.97	1
5	TI-F	39.5	24.5—85.5	0.51	-0.15		3.86	
7	TOID-F	42	17—46	-1.49	-2.67		3.82	
8	TOID-A	21	13.5—56	-1.59	2.35		3.17	
8	AMD-F	31	12—69	0.28	-3.32	H	2.89	2
8	AMN-F	31	8.5—70	0.03	-1.15	H	3.29	2
8	MET-F	33	2—74	-3.48	-5.3	H	3.74	2
9	TOID-F	56	4—69	1.2	-0.46	H	3	3
9	TA-M	56	8—63	-0.27	-0.14		2.36	3
9	TC-M	56	4—69	-0.42	0.07	H	3.53	3
9	**TOID-A**	**56**	**50—61.5**	**2.64**	**-0.91**		**4.86**	**3**
11	MET-F	14	5—73	-0.61	5.77	BD	3.17	
13	AMD-F	13	3—61	0.23	-2.49	H	3.09	4
13	AMN-F	13	3—61	0.1	-0.88	H	3.49	4
13	TOID-F	19	3—53.5	-2.1	-0.65		2.94	4
14	TA-M	27	10—53.5	0.28	-0.16	MC	3.2	5
14	TOID-F	40	7—57	0.65	-1.93	MC	3.07	5
14	TC-F	52	7—57	-0.09	0.5	BD	2.8	5
17	TA-F	42	8—47	0.17	1.18		2.63	6
17	TOID-M	51	8—54	-2.39	3.6	MC	2.87	6
19	MET-F	15.5	7.5—40.5	4.32	4.93	H	4.21	

Traits in bold were significant at the experiment wide level

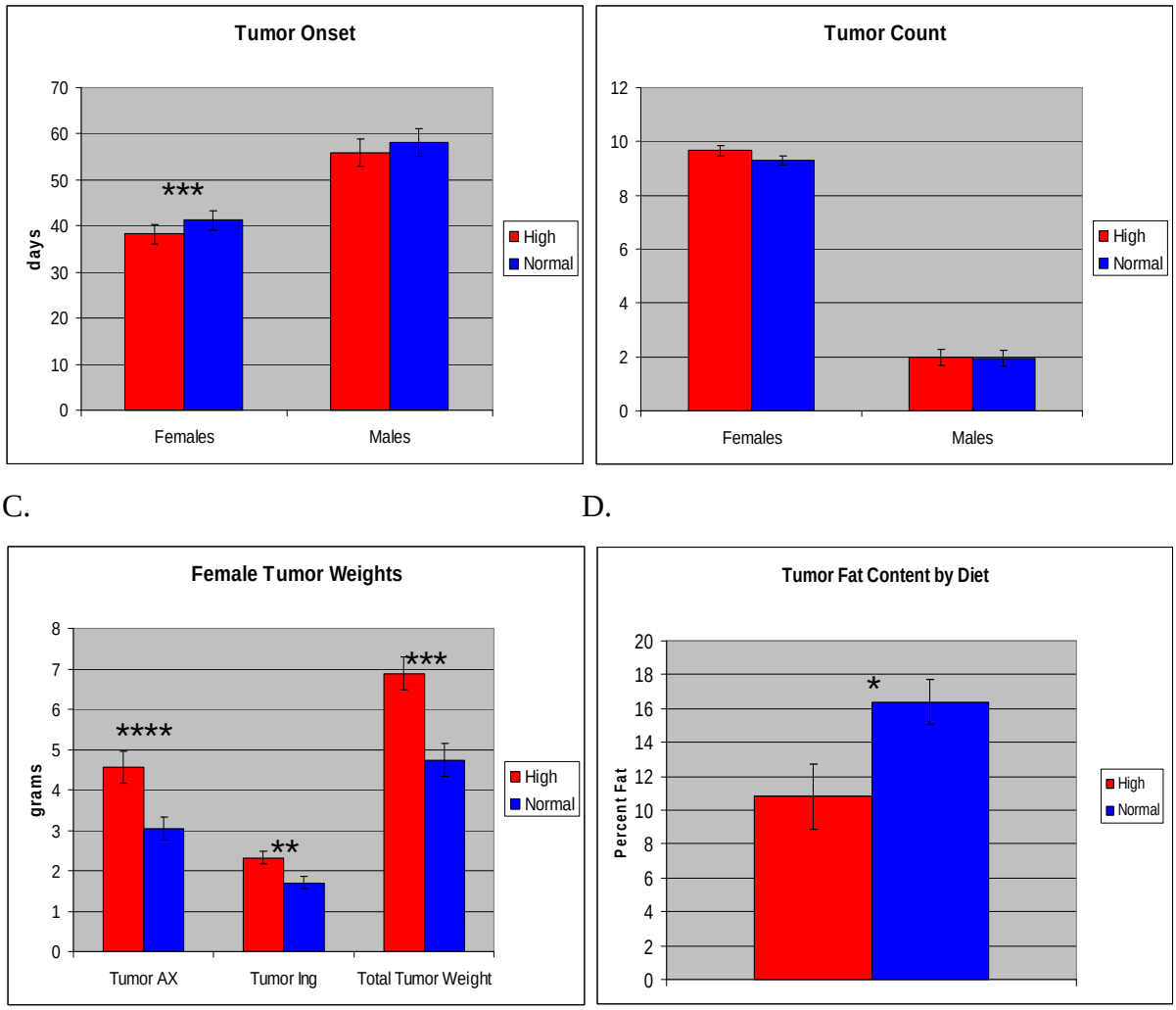
[a]TOID: tumor onset measured in days, TC: number of tumors, TA: total weight of the axillary tumors, TI: total weight of the inguinal tumors, MET: number of pulmonary metastasis counted at sacrifice, AMN: average metastasis number detected in internal sections, AMD: average metastatic density. An (A) signifies the QTL was detected in the pooled population, an (F) signifies the QTL was detected in the female population, an (M) signifies the QTL was detected in the male population.

[b]Approximate peak QTL position. cM positions are adjusted to the linkage map presented in Table 2 of Gordon et al. (companion paper). [c]95% confidence interval for QTL peak (in cM). [d]Additive effect determined by QTL Express, a positive value indicates that the increasing allele originates from the FVB. [e]Dominance effect representing the heterozygous genotype in

relation to the mean of the two homozygous genotypes: a positive value indicates that the heterozygote is larger than the mid-parent (mean of the parents). [f]Cause of the interaction. H: significant effect in high fat diet only, MC: significant effect in matched control fat diet only, BD: differential effects in high and matched control fat diets.

[g]Pleiotropic groups represent QTL that likely represent a single locus, as determined using the method described by Leberton et al., (1998). A cell without a number indicates that pleiotropy was not detected.

A. B.



C. D.

Figure 3.1: Least-squares means for tumor traits by diet and gender. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. (*) Significant at P<.05, (**) Significant at P<.01, (***) Significant at P<.001, and (****) Significant at P<.0001. **A.** Mammary cancer latency. **B.** Number of mammary tumors measured at the time of sacrifice. **C.** Tumor weights measured in the female population. Tumor Ax: weight of axillary tumors, Tumor Ing: weight of inguinal tumors. Total tumor weight: collective weight of axillary tumor and inguinal tumor. D. Percent fat of tumors.

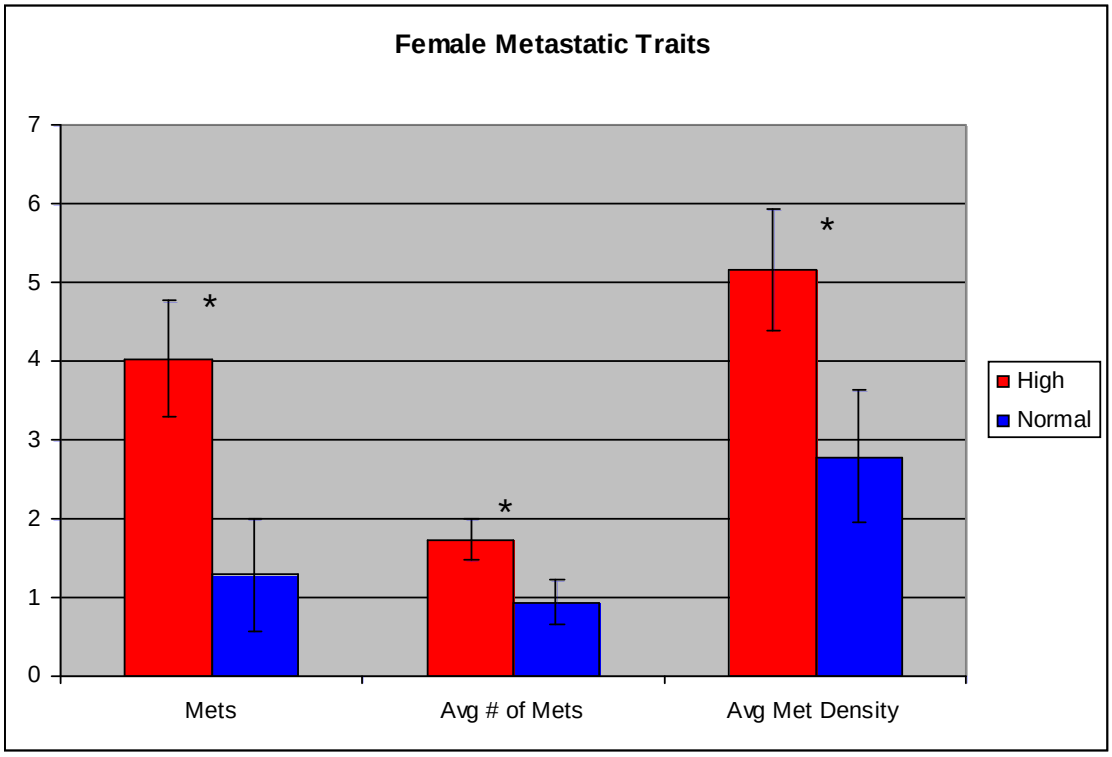


Figure 2. Least-squares means for metastatic traits when tested for dietary effects. High designates animals fed a high fat diet, whereas Normal designates animals fed the matched control fat diet. Mets: Surface metastatic burden determined at sacrifice, Avg # of Mets: average number of metastases per field of microscope view count, Avg Met Density: average metastases density calculated as the number of metastases per unit area of lung (number/square micron of lung x 1,000,000).

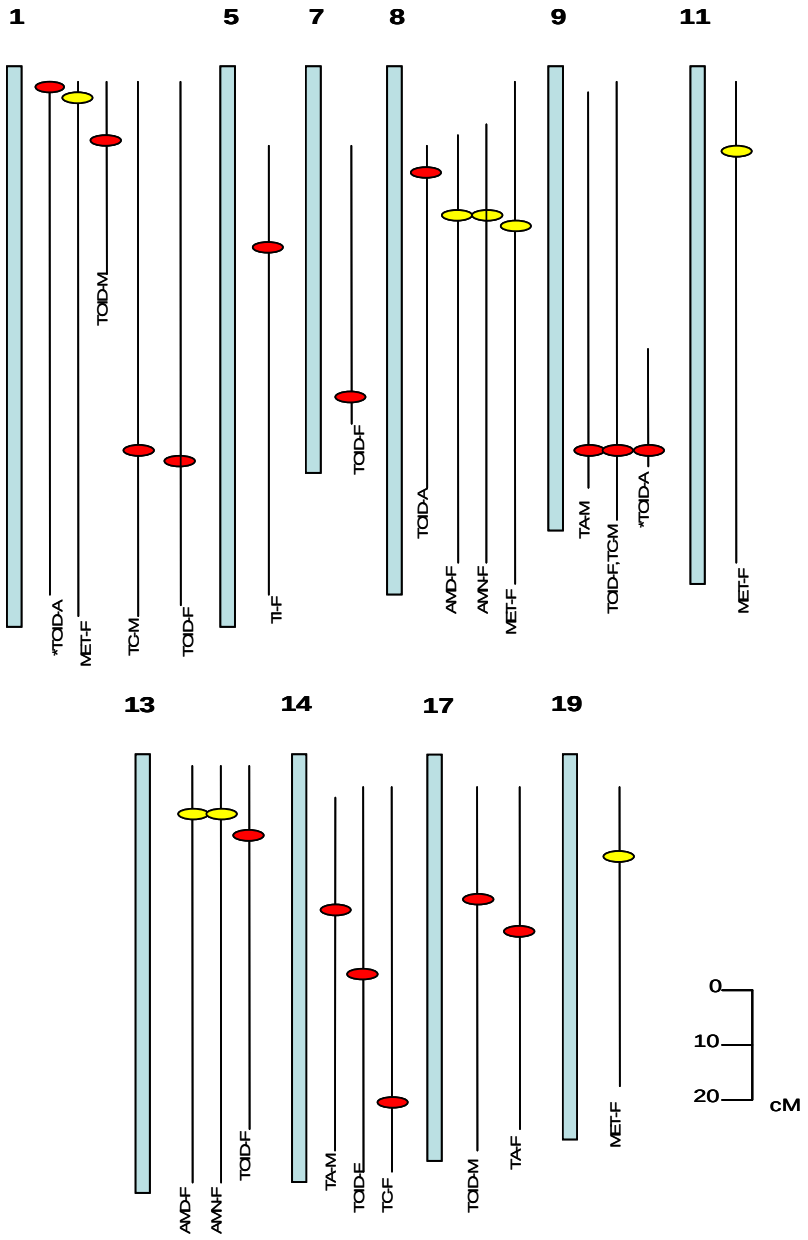
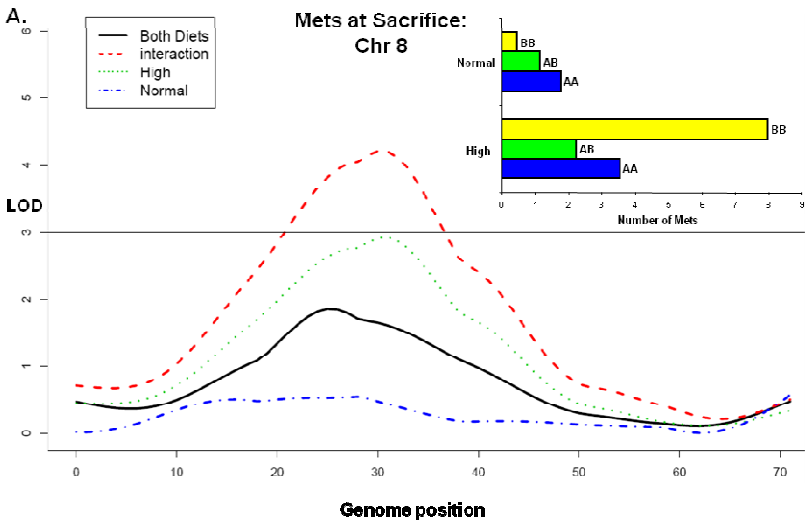
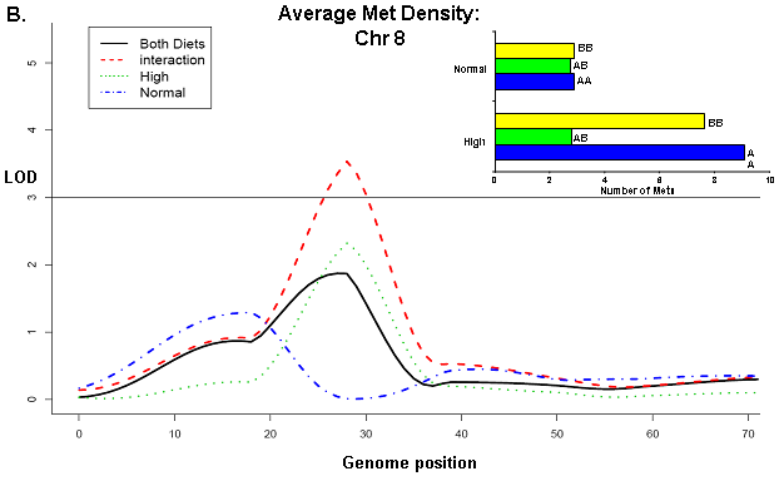


Figure 3.3: Positions and confidence intervals for cancer QTL. QTL detected for cancer phenotypes (AMD: average metastasis density, AMN: average metastasis number, MET: number of metastasis observed at sacrifice, TA: axillary tumor weight, TC: total tumors counted at sacrifice, TI: inguinal tumor weight, TOID: Tumor onset in days) at both the chromosomal P<.05 (suggestive) or at experimental P<.05 (significant). The presence of a A, F or M denotes that the QTL was detected in either the combined analysis, female analysis or the male analysis, respectively. The confidence intervals are indicated by the thin black lines and the estimated peak position of each QTL are denoted by the colored ovals. Yellow ovals represent metastasis traits while red oval represent traits related to primary tumors. (*) QTL detected at experimental P<.05


A.
Mets at Sacrifice:
Chr 8
Both Diets
interaction
High
Normal
Normal
High
BB
AB
AA
Number of Mets
LOD
Genome position


B.
Average Met Density:
Chr 8
Both Diets
interaction
High
Normal
Normal
High
BB
AB
Number of Mets
LOD
Genome position

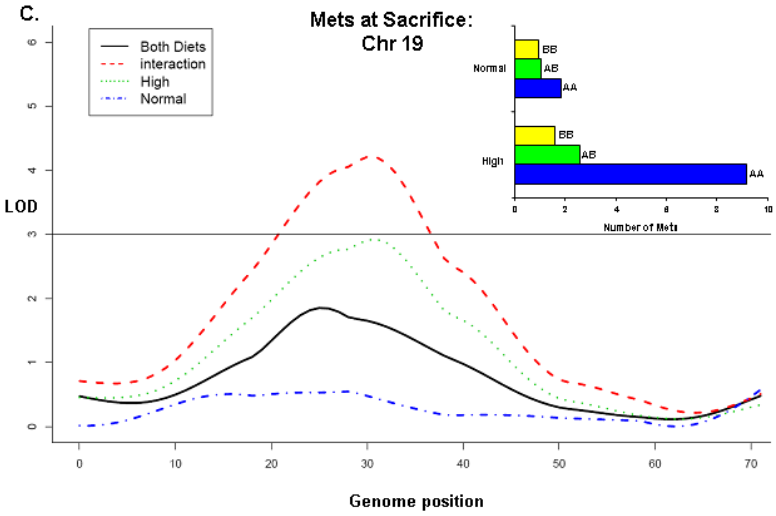

C.
Mets at Sacrifice:
Chr 19
Both Diets
interaction
High
Normal
Normal
High
BB
AB
AA
Number of Mets
LOD
Genome position

Figure 3.4: QTL x Diet interactions for select female cancer traits. Each figure includes graphical displays of the LOD curves that represent the individual components of the interaction. The solid black line is a representation of the resulting LOD curve when no interaction term is fitted in the model. The red dashed line represents the LOD curve when an interaction is included into the model. The remaining two lines represent the LOD curves of the two interaction components (high fat diet/matched control fat diet The threshold bar represents an average suggestive threshold across the cancer phenotypes. Additionally, the bar graphs represent the least-squares means of the trait of interest, for each allelic combination measured at the SNP marker closest to the QTL peak position (A allele: FVB; B allele: M16i). **A.** LOD intervals for the QTL detected for metastasis number at the time of sacrifice, on MMU8. **B.** LOD intervals for the QTL detected for average metastasis density, on MMU8.**C.** LOD intervals for the QTL detected for metastasis number at the time of sacrifice, on MMU19.

CHAPTER IV

Dietary fat alters pulmonary metastasis of mammary cancers through cancer autonomous and non-autonomous changes in gene expression

Michele La Merrill, Ryan R. Gordon, Kent W. Hunter, David W. Threadgill, Daniel Pomp

Preface:

The following chapter represents the work of a collaborative investigation between the laboratories of Daniel Pomp and David Threadgill. While, the primary author of this chapter is Michele la Merrill it was included in this dissertation because it describes experiments critical to the subsequent chapter. Additionally, many of the experiments/analyses described here were carried out by myself or performed in collaboration with Michele la Merrill. I specifically had a hand in following component described in this chapter: Isolation of RNA from the axillary tumors and livers, microarray analysis, data processing and normalization of expression profiles, differential expression analysis, eQTL evaluation of *Btn1a1*, writing of methodologies and additional manuscript assistance.

Abstract:

Metastasis virulence, a significant contributor to breast cancer prognosis, is influenced by environmental factors like diet. We previously demonstrated in an F2 mouse population generated from a cross between the M16i polygenic obese and MMTV-PyMT mammary cancer models that high fat diet (HFD) decreases mammary cancer latency and increases pulmonary metastases compared to a matched control diet (MCD). Genetic analysis detected eight modifier loci for pulmonary metastasis, and diet significantly interacted with all eight loci. Here, gene expression microarray analysis was performed on mammary cancers from these mice. Despite the substantial dietary impact on metastasis and its interaction with metastasis modifiers, HFD significantly altered the expression of only five genes in mammary tumors; four of which, including *serum amyloid A* (*Saa*), are downstream of the tumor suppressor PTEN. Conversely, HFD altered the expression of 211 hepatic genes in a set of tumor free F2 control mice. Independent of diet, pulmonary metastasis virulence correlates with mammary tumor expression of genes involved in endocrine cancers, inflammation, angiogenesis, and invasion. The most significant virulence-associated network harbored genes also found in human adipose or mammary tissue, and contained upregulated *Vegfa* as a central node. Additionally, expression of *Btn1a1*, a gene physically located near a putative cis-acting eQTL on Chromosome 13 and one of the metastasis modifiers, correlates with metastasis virulence. These data support the existence of diet- dependent and independent cancer modifier networks underlying differential susceptibility to mammary cancer metastasis and suggest that diet influences cancer metastasis virulence through tumor autonomous and non-autonomous mechanisms.

Introduction:

Breast cancer prognosis is largely determined by the level of metastasis and is influenced by non-genetic factors like diet, which is also a major contributor to obesity. A large prospective study of US women demonstrated that obese women in the highest quintile of body mass index (BMI) had twice the (Calle et al. 2003) death rate from breast cancer as did women in the lowest BMI quintile , possibly due to a higher risk for metastasis (Berclaz et al. 2004). The relationship between obesity and breast cancer risk has some genetic underpinnings; woman who have a family history of breast cancer are far more likely to develop breast cancer when obese rather than lean (Carpenter et al. 2003). Obesity can be linked to virulent breast cancer through both proliferative and inflammatory mechanisms (Lorincz and Sukumar 2006; Rose et al. 2004). For instance, adipocytes secrete the adipocytokine tumor necrosis factor (TNFa) and vascular endothelial growth factor (VEGF), which are both associated with breast cancer (Rose et al. 2004).

Much of the current research on breast cancer metastasis demonstrates an important role of angiogenesis and invasion. VEGF family members are associated with poor prognosis largely because of their angiogenic potency (Mohammed et al. 2007). Studies of breast cancer metastatic invasion frequently include members of the matrix metallopeptidase (MMP) family, which remodel the primary tumor through intravasation and seed lung metastasis by mediating extravasation (Gupta et al. 2007; Minn et al. 2007). Several obesity-modifying hormone pathways may interact in angiogenesis and invasion processes; for instance estradiol regulates MMP2 and tissue inhibitor of metalloproteinase 1 (TIMP1) (Nilsson et al. 2007). While there have been efforts to characterize the mechanistic events

driving metastasis and obesity-associated breast cancer, few have characterized the mechanistic relationships between obesity, breast cancer, and its metastasis.

In order to examine mechanisms underlying the interaction of obesity with breast cancer and its metastasis, we developed an obese mouse model of breast cancer metastasis (Gordon et al. 2008b). Mice from an F2 mouse population co-segregating obesity quantitative trait loci (QTL) and the MMTV-PyMT transgene had decreased mammary cancer latency and increased pulmonary metastases when fed a high fat diet (HFD) compared to those fed a matched control diet (MCD) (Gordon et al. 2008a). Genome-wide single nucleotide polymorphism (SNP) analyses reveled a strong genetic role in modifying susceptibility to pulmonary metastases, and diet significantly interacted with novel pulmonary metastasis QTL at all eight modifier loci detected (Gordon et al. 2008a). Here we examine relationships of dietary fat and the presence of metastasis on tumor gene expression signatures. Proliferation, inflammation, angiogenesis, and invasion processes were all significantly evident. For instance, pulmonary metastasis migration- associated serum amyloid A (*Saa2*) was upregulated in mammary tumors by HFD. Additionally, the milk component gene butyrophilin (*Btn1a1*) was identified on the metastasis virulence biomarker list, lies under a metastatic QTL that interacted with HFD, and mapped to an eQTL 7 cM away from its physical location on chromosome 13. These results suggest that *Btn1a1* should be further examined as a biomarker of breast cancer metastasis risk among women consuming a high fat diet.

Materials and Methods:

Husbandry and specimen collection: An F2 population was developed by mating M16i, a polygenic mouse model of obesity (Allan et al. 2004), with FVB/NJ-TgN(MMTV-PyMT)^{634Mul} (PyMT). PyMT is a mouse model of metastatic mammary cancer with primary tumor gene expression similar to the gene expression of luminal breast tumors (Guy et al. 1992; Herschkowitz et al. 2007). Full details of the generation and sampling of the F2 population are provided in Gordon et al. (Gordon et al. 2008b).

Briefly, F2 female mice hemizygous for PyMT received one of two synthetic purified diets, either HFD (n = 76 mice, D12451, Research Diets, New Brunswick, NJ) containing 45% of total calories from fat, 20% from protein and 35% from carbohydrate, or MCD (n = 79 mice, D12450B, Research Diets) containing 10% of total calories from fat, 20% from protein and 70% from carbohydrates, at four weeks of age and thereafter *ad libitum*. Mice were palpated three times weekly beginning at four weeks of age. Pulmonary metastasis (MET) was evaluated on the whole lung superficially by counting the number of foci visible under a dissecting scope. Subsequently, three coronal nonadjacent sections of one lung lobe per animal were examined under 12x magnification; the number of multicellular metastatic lesions observed per square micron of non-alveolar lung tissue was defined as the average pulmonary metastatic density (AMD). Axillary mammary tumors were flash frozen for microarray analyses (n = 64 and 67 mice fed HFD and MCD, respectively). To validate that HFD had expected strong physiological effects, livers were collected from randomly selected PyMT negative, non-tumor bearing- F2 female sib-pairs fed opposing diets (n = 12 sib-pairs fed HFD, n = 12 sib-pairs fed MCD).

RNA isolation and microarray analyses: RNA from both axillary tumor and liver was isolated by TRIzol reagent (Invitrogen, Carlsbad, CA), and amplified using the Illumina® TotalPrep RNA Amplification kit, both according to manufacturer's instructions (Ambion, Austin, TX). A solution containing 1.5 mg of highly purified biotinylated cRNA was applied to the Illumina Mouse 6 Sentrix array (version 1, Illumina, San Diego, CA) surface and hybridized at 55^{o}C for 17.5 hours. Following the hybridization period arrays were placed in High Temperature Wash Buffer (Illumina) for 10 min, E1BC Buffer (Illumina) for 5 min, 100% ethanol for 10 min, E1BC Buffer (Illumina) for 2 min, Block E1 Buffer for 10 min, and rocked with 2 mL of streptavidin-Cy3 (1 mg/mL in Block E1 Buffer, Illumina) for 10 min. Arrays were then washed in E1BC Buffer, dried, and evaluated on an Illumina Bead Scanner.

Microarray data processing: Raw data containing ~ 46,000 probe sets were log transformed and then normalized using a combination of the Loess and Quantile methods available in the R-based Lumi evaluation program for Illumina expression data (Du et al. 2008). Loess followed by Quantile normalization of identical pooled mammary tumor RNA ran across 11 chips produced a R^2 value of 0.95. In order to eliminate transcripts that were not significantly expressed above the background signal, data were filtered at an Illumina detection score of 0.95 and above.

Statistical analyses: Log transformed normalized data for all samples were run with 1000 permutations at an FDR q- value, or $Q < 0.05$ in SAM software using two class-unpaired analyses (Storey 2002; Tusher et al. 2001). We examined the effect of diet in tumors and livers by looking for significantly differential gene expression between HFD and MCD. To examine the relationship between metastasis virulence and gene expression in

primary cancers, we used a multi-tiered approach independent of diet. Samples were ordered such that 25% (n=33) of the total samples had no surface- or section- metastatic lesions and the longest tumor onset. These 33 samples were paired with 33 samples that had the most surface metastatic lesions detected (MET66), or with 33 samples that had the most sectional metastatic lesions detected (AMD66). Further analysis examined only those genes common to both MET66 and AMD66 with also known expression in human- mammary and/or adipose- tissues.

Functional analyses: Significant genes and their fold change values that were generated in SAM were imported into Ingenuity Pathways Analysis (IPA) 6.5-1602 (Redwood City, CA). RefSeq identifiers and their corresponding fold changes were mapped to corresponding gene objects in the IPA Knowledge Base (IPAKB). The curated IPAKB was used to generate functional analyses of significantly differential gene expression. IPAKB can also identify common pharmacological interactions through curation of Food & Drug Association data on approved pharmaceuticals, and of the National Institute of Health service ClinicalTrials.gov. The significance of functions and diseases to the gene set was determined by Fisher's exact test to calculate the probability (p-value, or P) that each biological function and/or disease assigned to the gene set was due to chance alone.

Candidate gene evaluation: Normalized expression profiles were analyzed with the F2 inbred/Co-dominant Marker Analysis option of the web-based program QTL Express (Seaton et al. 2002), fitting one expression QTL (eQTL) per chromosome. The genetic model included the additive plus dominance effects and fitted replicate and diet as fixed

effects. The resulting expression loci were classified in one of two categories, “cis” acting if they mapped within 10 cM of the physical location of the actual gene they represent; otherwise they were classified as “trans” acting. A genome-wide significance threshold for eQTL effects (LOD=3.5) was estimated using permutation testing with 1000 iterations (Churchill and Doerge 1994).

Results:

HFD alters expression of a limited repertoire of genes in mammary cancers: Genes differentially expressed between mammary primary cancers from mice fed HFD versus those from mice fed MCD were identified through comparison of microarray gene expression profiles. Despite the profound effect of diet on metastasis, only five genes were significantly upregulated in mammary cancers of mice fed HFD compared to those fed MCD, four of which were connected within a common network (Figure 1, Q = 0, *Mon1a* not shown). Reflective of their significant association with cancer ($P < 0.001$), the four genes of this network are downstream PTEN and/or TNF (Figure 1).

To determine if products of the genes associated with HFD correlate with existing therapies, we used the IPAKB to identify several pharmaceuticals that target products of HFD-induced genes. Hyperphosphorylation of eukaryotic translation initiation factor 4E binding protein 1 (EIF4EBP1) is induced by the breast cancer therapeutic paclitaxel (Greenberg and Zimmer 2005), and SCIO-469 blocks synthesis of TNFA, VEGF, and IL1B by inhibiting a MAPK14 complex that binds and phosphorylates EIF4EBP1 (clinicaltrials.gov ID NCT00744432). Given *Eif4ebp1* was upregulated by HFD in mammary

cancers, the efficacy of paclitaxel and SCIO-469 may be enhanced by combining therapy with dietary modifications reducing fat intake.

HFD causes extensive changes in liver gene expression: To confirm the validity of the modest HFD effect on mammary cancer gene expression (i.e. to demonstrate that this was not due to microarray technical issues), we examined global gene expression of livers from wild-type F2 littermates of the MMTV-PyMT transgenic mice, because obesity resulting from HFD is closely associated with substantial liver transcriptional changes.(Li et al. 2008; Morgan et al. 2008) Livers from mice fed HFD were compared to livers from mice fed MCD and significant gene expression differences were identified ($Q < 0.05$). A total of 211 genes met this criterion, of which 116 were significantly downregulated and 95 were upregulated (Supplemental Table 1). As expected, HFD deregulated numerous biological and toxicological functions in liver including carbohydrate-, nucleic acid-, and lipid- metabolism (Supplemental Table 2, $P < 0.05$). Despite the non-transgenic status of these mice, diet was associated with hepatic expression changes of 36 genes that are significantly associated with cancer phenotypes ($P < 0.01$).

Diet-independent transcriptional changes in primary cancers associated with metastasis: To investigate the relationship between gene expression in primary mammary cancers and metastasis, we identified genes with a $Q < 0.05$ by comparing global gene expression of primary cancers from mice with aggressive metastasis (MET66 and AMD66) to those from mice with no metastasis. A total of 478 genes met this criterion in MET66, of which 271 were significantly downregulated and 207 were upregulated (Supplemental Table 3). A total of 212 genes met this criterion in AMD66, of which 140 were significantly downregulated and 72 were upregulated (Supplemental Table 4). Together, all significant

genes found in each MET66 and AMD66 (n = 690 genes) were examined in IPA for functional analyses.

The top three most significant function and disease classes of metastasis-associated genes were endocrine system disorders, metabolic disease and cancer ($P < 0.05$). Tumorigenesis (101 genes), neoplasia (98 genes), and cell death (94 genes) were the groups most populated by genes differentially regulated based upon metastasis (Supplemental Table 5, $P < 0.05$). Many tissue remodeling activities, including the hematological system, were also significantly altered based upon metastasis ($P < 0.05$). However, only inflammation was evident among significant biological functions, diseases, and canonical pathways (e.g. glucocorticoid-, interferon-, and platelet derived growth factor- signaling canonical pathways, $P < 0.05$).

Development of candidate transcriptional biomarkers of metastasis: Given the large number of genes involved in various processes of metastasis, we restricted our analysis to transcripts that might later serve as biomarkers of human breast cancer metastasis risk. First, a list of candidate metastasis virulence biomarkers was developed from the significant gene lists generated by AMD66 (n = 147 biomarker filter eligible genes) and MET66 (n = 420 biomarker filter eligible genes) groups. The list of candidate biomarkers was developed by including only those genes whose expression occurs in mammary or adipose tissues of humans. Because AMD and MET are measurements of a similar phenotype, we further reduced the candidate biomarker list to those genes common to both gene sets (n = 128 genes). This list of candidate biomarkers was then subjected to IPA Core Analysis using the IPAKB as the Reference Set to assess significance of functions, pathways, and toxicological analyses.

Tumorigenesis (36 genes), cell death (36 genes), and neoplasia (35 genes) were the function- and disease- categories most populated in the candidate biomarker list (Supplemental Table 6). The top network includes upregulated *Vegfa*, whose product is implicated in many metastasis processes, including angiogenesis, and invasion (Figure 2, Supplemental Table 6).(Oshima et al. 2004) Other candidate metastasis virulence biomarkers are involved in angiogenesis of blood vessels (nine genes, $P < 10^{-4}$) and angiogenesis-related processes including neovascularization (six genes, $P < 10^{-4}$), as well as endothelial cell-migration (nine genes, $P < 10^{-4}$) and proliferation (six genes, $P < 0.01$, Supplemental Table 6). Further, molecules of the biomarker list were significantly associated with invasion (12 molecules, $P < 0.001$) and its processes, such as chemotaxis (10 molecules, $P < 0.01$), and breast cancer cell migration (six genes, $P < 10^{-4}$, Supplemental Table 6). The biomarker genes also associated with other diseases with an etiological basis in obesity, e.g. diabetes (15 genes, $P < 10^{-6}$), hypertension (six genes, $P < 0.01$) and atherosclerosis (ten genes, $P < 10^{-6}$, Supplemental Table 6).

To determine if products of the candidate transcriptional biomarkers correlate with existing therapies, we used the IPAKB to identify those biomarkers whose expression changed in direction as predicted by their clinical target. *Vegf* and *Vegfa* are upregulated in the biomarker list, and the latter is also a drug target for endocrine-, epithelial-, and metastatic- cancers, including breast cancer. Another upregulated gene in the biomarker list, endothelin receptor type B (*Ednrb*), is inhibited by atrasentan. Atrasentan is in phase II and III trials to treat various cancers as well as endothelial dysfunction (clinicaltrials.gov ID NCT00046943).

BTN1A1 as a candidate therapeutic target for metastasis: Milk fat component

Btn1a1 expression was significantly lower in primary cancers associated with metastasis compared to those without metastasis ($P < 0.01$). Similarly, a node just upstream of two genes upregulated by HFD in mammary tumors is butyric acid (Figure 1), another bioactive component of milk fat. Further, both *Btn1a1* and butyric acid are implicated in breast cancer (De los Santos et al. 2007; Woelfle et al. 2003). To evaluate the plausibility of these correlations, the IPAKB path explorer tool was used to determine whether butyric acid and BTN1A1 are functionally related. Through a number of pathways, butyric acid and BTN1A1 appear to be involved in a negative feedback loop (Figure 3A).

Given the non-invasive benefit of nipple aspirate as a potential biomarker of breast cancer metastasis risk, we confirmed the importance of BTN1A1 using eQTL analysis. We examined SNP markers across chromosome 13, and identified an eQTL on Chromosome 13 near the *Btn1a1* locus that regulates a significant amount of variation in *Btn1a1* mRNA abundance, suggesting the presence of a cis-acting eQTL (Figure 3B, LOD = 7.39). Further, *Btn1a1* colocalizes with a previously detected modifier for AMD that was only significant in mice fed HFD (Figure 3B) (Gordon et al. 2008a).

Discussion:

We previously demonstrated that diet plays a substantial role in tumor onset, tumor weight, and the extent of metastasis in an obese mouse model of breast cancer metastasis based on the MMTV-PyMT transgene (Gordon et al. 2008a). Yet the modest number of genes for which HFD significantly altered expression in primary tumors was unexpected given the strong HFD x QTL effects seen on mammary cancer metastasis phenotypes in this F2 mouse population (Gordon et al. 2008a). This result may be explained by: 1) the few but

highly significant transcript changes due to diet may have been sufficient to drive metastasis (Figure 1); 2) diet may have exerted its effects on metastasis through non-autonmous systemic changes in extra-mammary sites (Supplemental Table 2); and/or 3) HFD may have increased pulmonary metastasis through interaction between transcripts involved in the diet network (Figure 1) and transcripts involved in the metastasis networks (Supplemental Tables 3-4). Although HFD may have changed the expression of few genes because of experimental design limitations, the effect of HFD on the liver transcriptome strongly indicates that the experimental design was adequate to detect diet effects.

The expression of few genes in primary tumors change in response to diet: The upregulated diet-associated genes have an association with cancer (Figure 1, $P < 0.001$), perhaps because they are downstream from PTEN and/or TNF, both of which are implicated in advanced epithelial cancers and insulin resistance (Ikubo et al. 2008; Rosner et al. 2008). Further, upregulated *Eif4ebp1* is joined with PTEN in the phosphatidylinositol 3-kinase (PI3K)/ thymoma viral proto-oncogene (AKT) canonical pathway, a pathway relevant to both obesity and cancer metastasis, through the regulation of glucose uptake, as well as cellular proliferation and survival (Gingras et al. 1998). Among genes for which HFD significantly altered their primary tumor expression, only *Mon1a* does not have an existing link with cancer. Although MON1A is poorly characterized, it is involved in macrophage iron loading (Wang et al. 2007a), and thus may interact with the hemoglobin complex, one member of which was also upregulated in mammary cancers by HFD. Notwithstanding these associations with cancer processes, the modest number of HFD-induced changes in mammary cancer gene expression was surprising given the robust effects of HFD on tumor-

latency, size, metastasis, and modifier interaction in this F2 population (Gordon et al. 2008a; Gordon et al. 2008b).

Non-tumor-autonomous actions of diet may influences metastasis: Diet may have exerted its effects on metastasis through systemic changes in extra-mammary sites, such as the hepatic induction of the PI3K/AKT pathway genes associated with metastatic potential in the liver ($P < 0.01$, Supplemental Tables 1-2). PI3K/AKT hepatic signaling may have acted in concert with PI3K/AKT signaling in mammary tumors, given that in both tissues, the PI3K/AKT pathway was significantly altered by HFD ($P < 0.05$). A candidate PI3K signaling molecule driving such cancer non-autonomous action may be catenin beta 1 (*Ctnnb1*). *Ctnnb1* was upregulated in livers from mice fed HFD compared to those fed LFD ($P < 0.01$), and is implicated in breast cancer and its invasion (Adam et al. 2001; Michaelson and Leder 2001). Further, CTNNB1 binds N-myc downstream regulated gene 1 (NDRG1) (Tu et al. 2007), which was upregulated by HFD in mammary tumors (Supplemental Table 2, Figure 1). Cancer non-autonomous mechanisms may extend beyond the PI3K/AKT pathway. In liver from the non-transgenic F2 mice and elsewhere (Maxwell et al. 2003), HFD significantly upregulated hepatic ectonucleotide pyrophosphatase/ phosphodiesterase 2 (*Enpp2*, $P < 0.05$), a cell membrane enzyme associated with invasion and metastasis (Supplemental Tables 1-2) (Nam et al. 2000). These observations strongly suggest that part of the effect of HFD on cancer metastasis may be through cancer non-autonomous mechanisms.

Potential interactions between diet and metastasis networks: *Saa2* and *Saa3* were significantly downregulated in our previous functional genomic analysis of the PyMT model on FVB/J background compared to other strains of the PyMT model (Qiu et al. 2004). Yet

here *Saa2* was upregulated by HFD, suggesting HFD and/or polygenic obesity can overturn the negative regulation of SAA family expression associated with this genetic mammary cancer model. Indeed, SAA2 varies according to the metastatic potential of mouse models of breast cancer and is part of a gene expression signature that distinguishes breast cancer patient outcomes across independent breast cancer datasets (Lukes et al. 2009).

While the mechanism of SAA2 on metastasis has not yet been determined, the involvement of SAA2 in NFKB signaling in mammary epithelial cells has recently been demonstrated (Kho et al. 2008). Here, the second most significant network of the candidate panel of transcriptional biomarkers of metastasis has several up-regulated molecules in direct interaction with NFKB. Given the mechanism of SAA3 seeding metastasis in the pre-metastatic lung is attributed to regulation of chemoattractant secretion and resulting NFKB-mediated cell migration (Hiratsuka et al. 2008), diet-induced SAA2 may interact with the metastasis-induced NFKB pathway to increase metastatic virulence.

Milk fat components are another compelling link between the influences of diet and metastasis on mammary tumor gene expression. We found significantly lower expression of *Btn1a1*, a major component of milk fat droplets, in our candidate panel of transcriptional biomarkers of metastasis ($P < 0.01$, Supplemental Tables 3-4). Consistent with this finding, decreased *Btn1a1* expression was identified as part of the high virulent signature that previously characterized the effects of the MMTV-PyMT mammary cancer model (Qiu et al. 2004), and has also been associated with metastatic breast cancer in humans (Woelfle et al. 2003). Similarly, a node just upstream of two genes upregulated by HFD in mammary tumors is butyric acid (Figure 1), another bioactive component of milk fat.

Butyric acid is one molecule upstream of BTN1A1 binding partner, xanthine dehydrogenase (XDH, Figure 3A) (McManaman et al. 2002). The BTN1A1-XDH complex may be inhibited by butyric acid through the production of angiogenic- NADH and nitric oxide (Hewett et al. 1999; Park et al. 1998; Rinaldo et al. 1994; Thomsen and Miles 1998). Butyric acid decreases *interleukin 1 beta* (*IL1b*) expression (Joseph et al. 2004), which would also serve as negative feedback to the BTN1A1-XDH complex, normally stimulated by IL1b (Kocic et al. 1995). Because IL1b decreases inorganic pyrophosphate production, which synthesizes butyric acid, BTN1A1-XDH stimulus decreases butyric acid.

Together, the feedback loop of *Btn1a1* with butyric acid and the chromosomal associations of *Btn1a1* are suggestive of a mechanistic relationship between HFD-induced obesity and mammary metastasis virulence. Both BTN1A1 and butyric acid reside on the cell surface of mammary alveolar epithelial cells and are regulated by angiogenic Vegf (Figure 2) (McManaman et al. 2002; Rossiter et al. 2007). Butyric acid is currently thought to inhibit breast cancer through histone deacetylase (HDAC) inhibition (De los Santos et al. 2007). There is little known about the biological activity of BTN1A1, however its binding partner XDH increases secretion of MMP2 (Figure 2), which has been shown to increase metastasis through the degradation of extracellular matrix (Galli et al. 2005). Elsewhere XDH activation of NFKB mediates angiogenesis (Shenkar et al. 1996). Thus the Btn1a1-butyric acid feedback loop may be influencing metastasis through epigenetic-, motility- and angiogenic- activity.

Given the active chemotherapeutic research of butyric acid as a HDAC inhibitor, and the downregulation of *Btn1a1* reported here, our findings indicate that high levels of *Btn1a1* expression may protect rather than promote of breast cancer. BTN1A1 protein was

successfully measured in human nipple aspirate fluid, but was not deemed a suitable biomarker of cancer risk because no studies demonstrated a correlation between BTN1A1 and cancer at that time (Varnum et al. 2003). Consequently, *Btn1a1* levels in cancer patients treated by HDAC inhibitors may merit monitoring. Investigation of the biological function of differential *Btn1a1* expression in mouse models of mammary tumor metastasis is underway.

Study limitations:

A role of decreased carbohydrates cannot be eliminated as an influencing factor of gene expression- and metastatic- effects of HFD seen here because HFD was formulated to have the same caloric density as MCD through a relative decrease in carbohydrates. However, diets matched for every nutrient besides fat would require differences in caloric density, another imperfect experimental design.

Similarly, HFD may have increased mammary tumor gene expression through subtle, undetectable changes in gene expression in molecules upstream of those altered by HFD. TNFa, PTEN, and butyric acid are examples of genes directly upstream of at least two genes that HFD upregulated, and are also associated with breast cancer. However, the high R^2 value of identical, pooled, and normalized mammary tumor RNA ran across 11 microarray chips suggests minimal technical error. The strong influence of metastatic virulence on differential mammary tumor gene expression further indicates minimal technical error. As expected, these data indicate that HFD substantially altered liver gene transcription, and further suggest that cancer non-autonomous changes may contribute to the effects of HFD on mammary cancer in MMTV-PyMT transgenic mice. We also cannot exclude the possibility that the strong influence of HFD on metastasis seen in this mouse population was mediated

through gene expression changes in pulmonary metastasis that were not detectable in primary cancers.

Conclusion:

A diet high in fat is becoming a more prevalent occurrence. The resulting elevation in obesity prevalence is a substantial public health concern in part because of its association with breast cancer morbidity and mortality. Our data suggest that diet may increase breast cancer metastatic virulence through multifactoral cancer- autonomous and non- autonomous effects on cell migration, angiogenesis, and extracellular matrix breakdown. Transcript changes due to HFD, to metastasis, and to their interaction illuminate the influence of the primary mammary tumor on pulmonary metastatic virulence. Yet transcript changes in the liver indicate that HFD influenced metastasis through systemic, non-autonomous mechanisms as well. Our data suggest that the complex relationships between diet, mammary carcinogenesis, and its metastasis will become clearer if greater focus is placed on understanding the role of extra- mammary sites in carcinogenesis and metastasis.

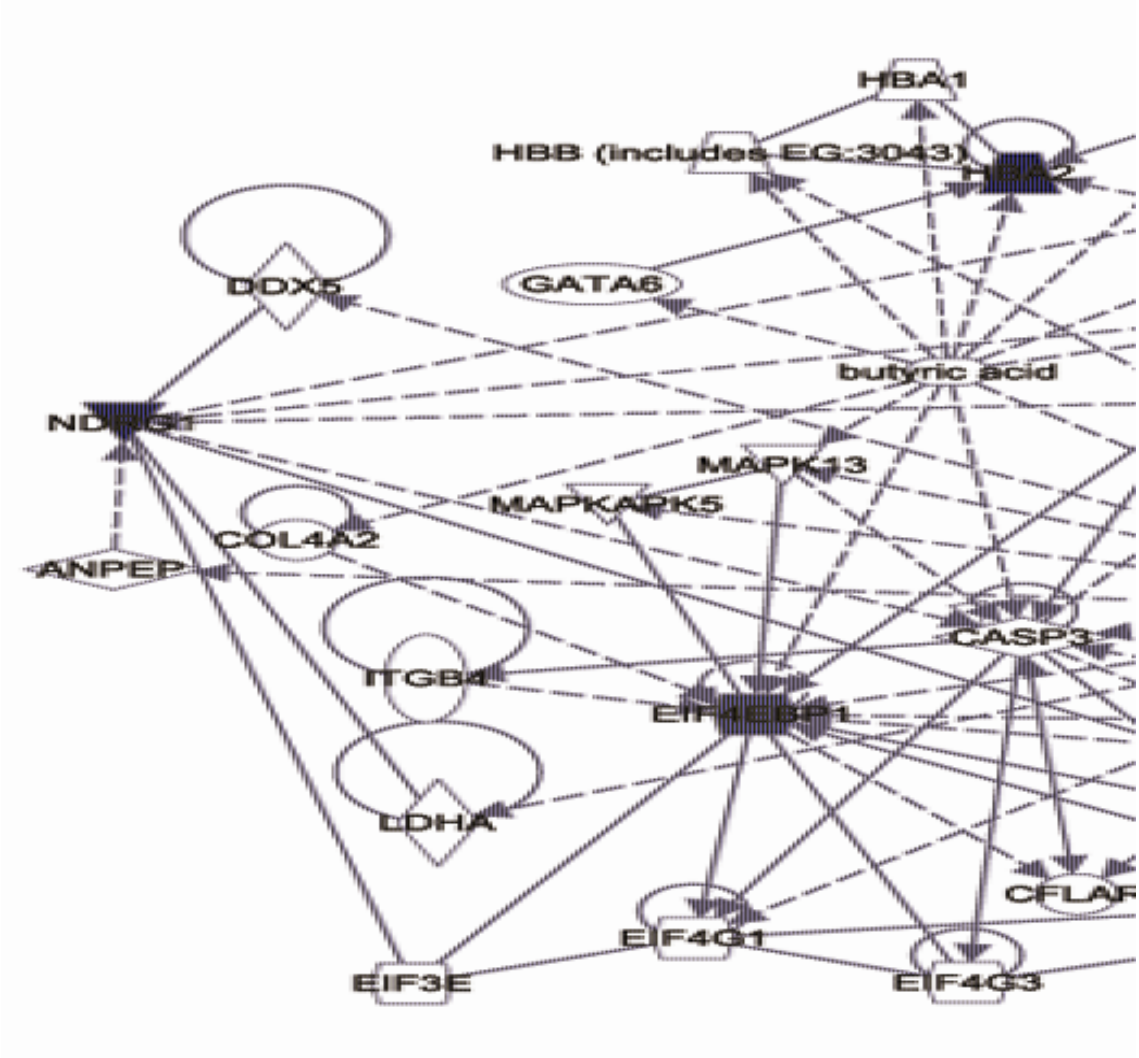


Figure 4.1. Top network depicting the influence of high fat diet on mammary tumor gene expression (Ingenuity Pathway Analysis Knowledge Base). Gene networks depict how the genes' products directly and indirectly interact with each other, including those genes not identified as significant on the microarrays, and thus networks are ranked such that the highest ranked network contains the highest number of significantly expressed genes. Red nodes denote significant upregulation comparing high fat diet- relative to matched control diet- fed mice (Q = 0).

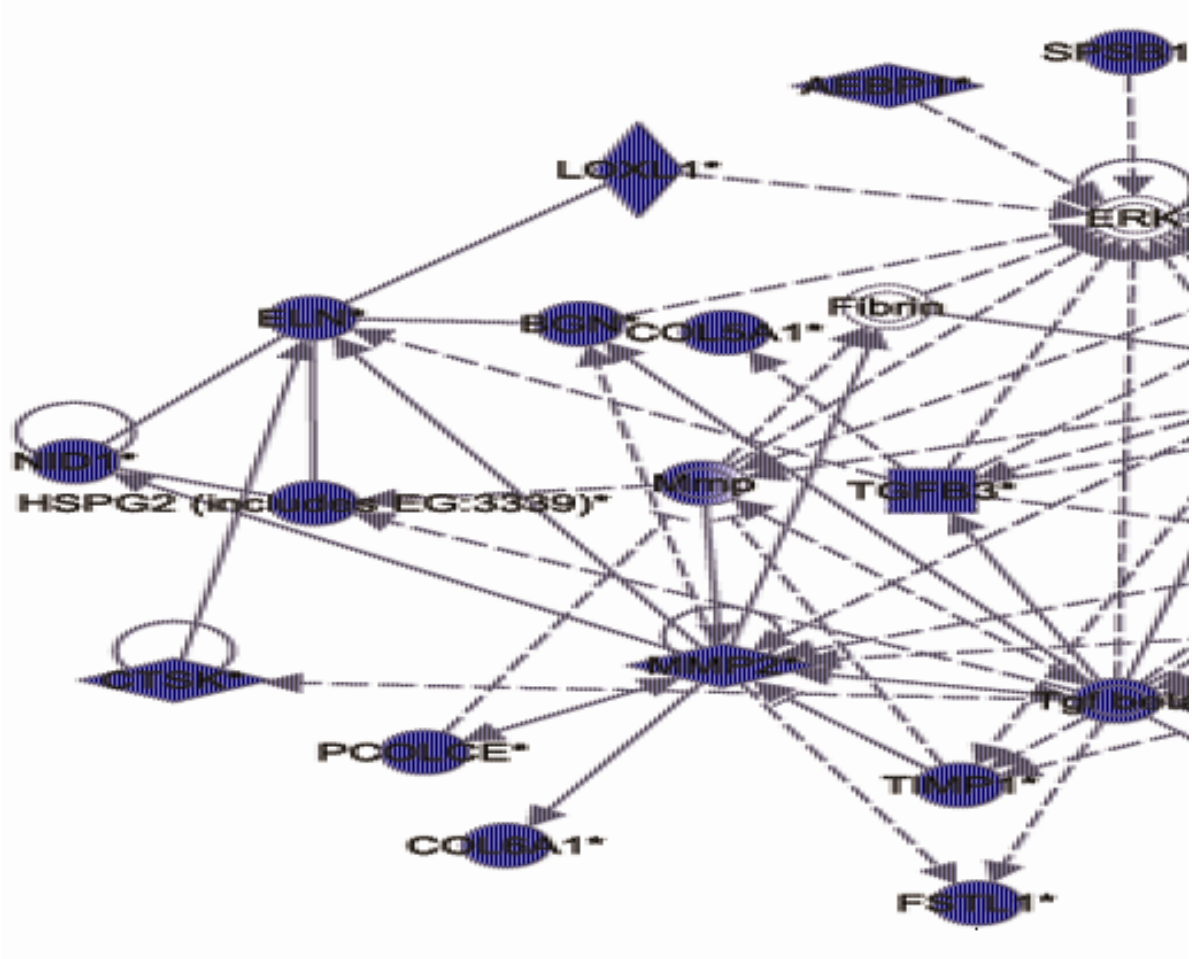


Figure 4.2. Top network of metastasis virulence biomarkers, depicting the genes of both AMD66 and MET66 that are expressed in human adipose or breast (Ingenuity Pathway Analysis Knowledge Base). Red nodes denote significant upregulation in high virulent mammary tumors relative to non- virulent mammary tumors ($Q < 0.05$). Green nodes denote significant downregulation in high virulent mammary tumors relative to non- virulent mammary tumors ($Q < 0.05$).

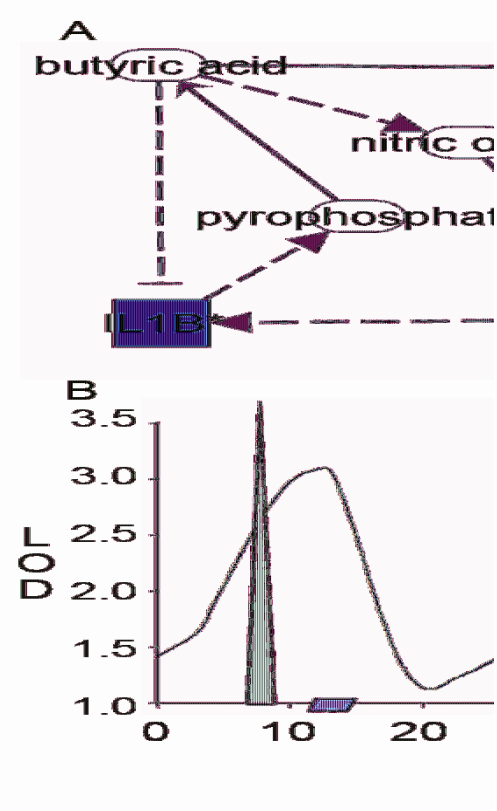


Figure 4.3. *Btn1a1*- gene expression network and quantitative trait locus in metastatic mammary cancer. A) *Btn1a1*, significantly downregulated by metastasis, and binding partner Xdh are regulated in a feedback loop with butyric acid, implicated in high fat diet effects on mammary tumors (Figure 1) through IL1B, also downregulated by metastasis ($P < 0.05$). B) Btn1a1 lies within the 95% confidence interval of a QTL associated with the trait average pulmonary metastatic density only among mice fed high fat diet. Variation in *Btn1a1* expression was partially explained by a significant expression quantitative trait locus also within the 95% confidence interval of the quantitative trait locus that describes the average pulmonary metastatic density x high fat diet interaction and near the physical location of the Btn1a1 gene.

CHAPTER V

Genetic architecture of tumor gene expression in dietary fat responsive mouse metastatic mammary cancer.

Ryan R Gordon, Michele La Merrill, Kent W Hunter, David W Threadgill, Daniel Pomp

Abstract

Breast cancer is the most common cancer type, and the second-leading cause of cancer-related deaths of women living in the United States. It has been estimated that in 2009, nearly 200,000 women in the United States will be diagnosed with some form of breast cancer and that over 40,000 will die of this disease, typically from secondary metastasis. Breast cancer is a complex disease resulting from a combination of, and interaction between, environmental and genetic factors. However, the underlying genetic architecture that results in differential susceptibility to this disease is poorly understood. Additionally, links between diet and incidence of breast cancer have been reported for a broad spectrum of dietary components, including fat.

We previously reported in an F_2 mouse intercross, segregating for both obesity and metastatic mammary cancer polygenes, that animals fed a high-fat diet not only have shorter mammary cancer latency but also increased tumor growth and more pulmonary metastases over an equivalent time. Subsequent genetic analysis identified several modifiers of

metastatic mammary cancer along with widespread QTL by dietary fat interactions. To further investigate the genetic underpinnings that modify mammary cancer and metastasis, global expression profiles of axillary tumors were characterized in F_2 mice, and expression QTL (eQTL), which are involved in the transcriptional networks of metastatic mammary cancer, were mapped. Several potential candidate genes colocalizing with previously detected metastatic cancer QTL were identified, while simultaneously accounting for copy number variation within the population. Additional analyses, such as eQTL by dietary fat interaction analysis, causality and database evaluations, helped to further refine the candidate loci to produce an enriched list of genes potentially involved in the pathogenesis of metastatic mammary cancer.

Introduction:

Breast cancer, a complex disease, results from a combination of environmental and genetic pressures. A substantial amount of effort has been expended to identify the many risk factors associated with breast cancer; however, the majority of underlying mechanisms that result in this altered disease state still remain unclear. Whereas some successful attempts to understand the genetic predisposition to mammary cancer have been achieved, such as identifying small-to-low effect familial risk factors (Rohan et al. 2006; Song et al. 2006; Walsh et al. 2006), the overall genetic architecture remains largely unknown. In addition, little success has been realized in understanding the interactions between genes and environmental components, such as diet, and the pathogenesis of metastatic mammary cancer.

Previously we generated an F_2 intercross (Gordon et al. 2008b) between the M16i polygenic mouse model of obesity (Allan et al. 2004) and FVB/NJ-TgN(MMTV-PyMT)^{634Mul} (PyMT), a mouse model of metastatic mammary cancer (Guy et al. 1992). Results demonstrated that animals fed a high-fat diet not only have decreased mammary cancer latency, but also increased tumor growth and pulmonary metastases over an equivalent time. Several modifier loci (i.e. quantitative trait loci; QTL) for metastatic mammary cancer were detected along with widespread QTL by dietary fat interactions (Gordon et al. 2008a). Subsequently, a subset of mammary tumors collected from F_2 female mice were evaluated for whole genome expression using the Illumina mouse Sentrix array. Analyses of the gene expression data surprisingly revealed that, globally, only five genes were differentially expressed between mice on the two dietary fat treatments (la Merrill et al., 2009). However, dietary fat was found to alter pulmonary metastasis of mammary cancers through cancer autonomous and non-autonomous changes in gene expression (la Merrill et al., 2009).

Whereas QTL analysis and differential expression may reveal pathways, candidate regions and genes potentially linked to disease phenotypes, these methods are both independent assessments of the paradigm. As such, inferring a relationship between the results of the two methods is difficult. An approach that can bridge this disconnect is to treat the expression of each transcript identified through microarray analysis as a quantitative trait. The traits can then be tested for associations with genotypic data to map what is known as an expression QTL (eQTL) (Jansen and Nap 2001; Schadt et al. 2003a). When running an eQTL analysis, two distinct classes of loci are detected. The first class is loci that map within close proximity to the actual physical location of the expressed gene (*cis*-acting), and the

second class is loci which map independently of the expressed genes physical location (*trans*-acting) (Pomp et al. 2008). Whereas *trans*-acting eQTL represent loci controlled by unknown regulators, *cis*-acting eQTL exhibit self-regulation (Alberts et al. 2005). Therefore, by overlaying these *cis*-acting eQTL with locations of traditional phenotypic QTL detected in the same mapping population, the potential exists to significantly narrow the pool of candidate genes that are both positional and functional in nature (Wang et al. 2007b). The field of research on metastatic breast cancer has produced only a few experiments utilizing this multifaceted approach (Crawford et al. 2008).

Recently a new source of genetic variation, known as copy number variants (CNV), has been identified that can potentially impact on disease processes. CNV are described as segments of DNA that are over-or under-represented because of insertions/deletions occurring naturally over time or acutely due to tissue-specific somatic mutations (Feuk et al. 2006). Approximately 12% of the human genome has been estimated to be affected by CNV (Beckmann et al. 2007) and this over/under-representation of chromosomal segments may have profound influences on the expression of the genes within these affected regions. Many diseases, such as Crohn's disease (Fellermann et al. 2006), lupus (Yang et al. 2007) and HIV (Gonzalez et al. 2005), have already been linked in part to CNV. CNV may also plausibly be linked to many common complex chronic diseases such as cancers (Shelling and Ferguson 2007), yet currently our knowledge of the relationship between CNV and of chronic diseases remains limited. Further, the influence of CNV on detection of eQTL has been largely ignored.

In this report, evaluation of our M16i x PyMT F_2 cancer mapping population for tumor eQTL potentially involved in the transcriptional networks of metastatic mammary

cancer is described. The genetic regulation of expression profiles of metastatic mammary tumors through an eQTL analysis was established initially. Potential candidate genes for the previously detected phenotypic cancer QTL were then identified, while simultaneously accounting for CNV within our population and their impact upon mammary cancer susceptibility and progression. With the aid of other analyses, such as causality and database evaluations, these candidate genes were further refined to produce a short enriched list of genes potentially involved in the pathogenesis of metastatic mammary cancer in this mouse population.

Materials and Methods:

Population development: An F_2 population (n = 615) was generated by crossing M16i, a polygenic obesity line (Allan et al. 2005), and FVB/NJ-TgN(MMTV-PyMT)634Mul (PyMT), a line transgenic for the Polyoma Middle T Oncoprotein leading to development of mammary tumors and subsequent pulmonary metastasis (Guy et al. 1992) (see (Gordon et al. 2008b) for full details of population development). F_2 mice were randomly assigned, within litter, gender, and genotype (PyMT or no PyMT), to receive one of two synthetic purified diets at 4 weeks of age, either a high-fat diet (HFD; Research Diets D12451) containing 45% of total calories from fat, 20% from protein, and 35% from carbohydrates, or a matched-control-fat diet (MCD; Research Diets D12450B) containing 10% of total calories from fat, 20% from protein, and 70% from carbohydrates. Mice were evaluated for various body weight, body composition and metastatic mammary cancer phenotypes as previously described (Gordon et al. 2008a; Gordon et al. 2008b). Axillary mammary tumors were

harvested from 131 F_2 female PyMT carriers (HFD = 64 and MCD = 67) for microarray analysis.

Microarray analysis of mouse mammary tumors: For complete details refer to la Merrill et al (2009). Briefly, total RNA was isolated from axillary tumors using TRIzol (Invitrogen, Carlsbad, CA) and preprocessed for array hybridization using the Illumina® TotalPrep RNA Amplification kit (Ambion, Austin, TX). Expression profiles were generated using the Illumina Muse 6 Sentrix arrays (Kuhn et al. 2004) (Illumina, San Diego, CA) and the resulting data were transformed and normalized using the R-based evaluation program for Illumina expression data Lumi (Du et al. 2008) prior to filtering at an Illumina detection score of 0.95 and above.

Correlation analyses: Pearson correlations were generated between all significantly expressed genes on the microarray and the metastatic cancer phenotypes previously measured in the F_2 population (Gordon et al. 2008a). The correlations were generated as three separate data sets: one for the whole population and the other two for animals on either the HFD or the MCD. All resulting p-values were adjusted with the FDR multiple comparisons test. Transcripts were then sorted by their strength of correlation within each phenotype.

Expression QTL (eQTL) analysis of mouse mammary tumors: eQTL were identified utilizing a customized bioconductor R-GUI based program. The eQTL models were fitted with the following effects: additive, dominance, breeding-replicate and diet. The resulting

eQTL were classified in one of two categories: "*cis*" if they mapped within 10 cM of the physical location of the gene they represent, and "trans" if they mapped elsewhere. Clusters containing 50 or more eQTL within a 5 cM interval (i.e. trans-bands) were designated as potential master regulatory regions. The significance threshold for all eQTL was set at a likelihood ratio statistic of 16.1 (LOD: 3.5). All eQTL classified as *cis*-acting were evaluated for the presence of eQTL x diet interactions by running an additional model which included all the aforementioned effects in addition to an eQTL x diet interaction term. The sum-of-squares error and the degrees of freedom for the peak position for each eQTL in both the interaction and non-interaction models were then calculated and used to estimate an F statistic. Confidence intervals were calculated using the bootstrap method with 1000 permutations in Grid QTL (Seaton et al. 2002) for *cis*-acting eQTL mapping in close proximity to master regulatory regions.

Pathway evaluation: All eQTLwere evaluated using Ingenuity Pathway Analysis (IPA; Ingenuity Systems Inc., Redwood City, CA). Using the IPA core analysis function, reports containing information regarding the function, regulation, known mutations, tissue expression patterns, cellular location, and disease implication were generated for all genes of interest. IPA also provides rankings for biological functions based upon the number of genes from a reference set (a list of genes provided by the user) that participate in the particular function, while considering the total size of the reference set, the total number of genes from the reference set eligible for function analysis and the total number of molecules know to participate in the function in question.

The IPA core comparison function was used to generate the predicted transcriptional networks between the *cis* and *trans*-acting eQTL in master regulatory regions. Initially, both gene lists were analyzed separately to generate within-set transcriptional networks, containing each gene within the list as well as all other molecules and genes with which they are known to interact (based on the IPA database). Then both transcriptional network files (*cis*-acting, *trans*-acting) were evaluated with the comparison function of IPA, to determine if any interactions existed among the gene sets.

Causality evaluation: Causality relationships between *cis*-acing eQTL and both cancer phenotypes and *trans*-acting eQTL were evaluated using the R-based package eQTL-TF (Sun et al. 2007). When estimating the relationships between *cis*-acting eQTL and cancer phenotypes three models were considered (Figure 1). The causal model is where genetic alterations (G) result in changes in the expression of a *cis*-acting eQTL (C), which in turn result in modification of the phenotype (or change in gene expression of a *trans*-acting eQTL) (P/T). In the reactive model, variation in G directly impacts P/T resulting in altered gene expression of C. The third model evaluates whether variation in G can result in changes in both P/T and C independently.

Oncomine evaluation: All *cis*-eQTL colocalizing with previously detected metastatic QTL (Gordon et al. 2008a) were entered into the Oncomine database (Rhodes et al. 2004) to determine if the genes they represent had been previously linked to metastatic cancer in humans. Specifically, all human breast cancer prognosis datasets in Oncomine were

evaluated at a p-value threshold of 0.01 for each gene. The prognosis category was evaluated because it best represents the possibility of having a metastatic cancer phenotype. If a gene was identified in any of the datasets in the prognosis category, it was further evaluated to determine if its expression is positively (no metastasis, alive, no disease) or negatively (metastasis, dead, relapse) associated with the clinical phenotype.

Evaluation of CNV in mouse mammary tumors: Copy number variation was evaluated using the NimbleGen mouse 385K whole genome tiling array (NimbleGen Systems, Inc. Madison, WI) with a median probe spacing of 5.7 kb, in a subset of the F_2 females selected to represent the largest spectrum of tumor mobilization capacity (propensity of the tumor to metastasize). This spectrum was achieved by selecting within each diet the 17 individuals with the highest number of observed pulmonary metastases and the 17 with the lowest number of observed pulmonary metastasis, for a total of 68 individuals. RNA-free genomic DNA was extracted using the Puregene Tissue Core Kit (Qiagen, Minneapolis, MN) from the axillary mammary tumors (test) and tails (reference) of all 68 individuals. DNA samples were fragmented and fluorescently labeled with either cy-3 (tumor DNA) or cy-5 (tail DNA). The labeled tumor samples were pooled with their reference and co-hybridized to a NimbleGen 385K CGH array. Arrays were evaluated for relative fluorescence to determine the copy number profile across the genome for each individual using the NimbleScan software (NimbleGen Systems, Inc. Madison, WI).

Resulting data were evaluated using SignalMap software (NimbleGen Systems, Inc. Madison, WI), to determine if patterns of CNV could be identified within the population. To provide common settings across the population for analysis in SignalMap, the $\log_2$ ratio scale

was set at -2 and 2, while the track height was set at 120. Segments that deviated from 0.0 by ± 0.2-0.4 indicated a single amplification or deletion. A segment that deviated from 0 by ± 0.4-0.6 indicated a double amplification or deletion. If CNV was identified on chromosomes where we previously detected metastatic QTL, Proc Mixed in SAS (SAS Institute, Cary, NC) was used to determine whether significant associations were present between the copy number change and the phenotype the locus represented. The model evaluated included the fixed effects of diet and CNV, the interaction of diet x CNV, and the random effect of breeding replicate. If an association between the phenotype and copy number change was detected then *cis*-eQTL on the Chr in question were also evaluated to determine if its gene expression was altered by the amplification or deletion.

Results:

Correlation evaluations: Correlations amongst gene expression and metastatic phenotypes were analyzed in a three-step process, the first of which utilized the entire population regardless of diet. While many genes were weakly correlated with the metastatic phenotypes, no genes surpassed the significance threshold of $p < 0.05$ after adjustment for multiple testing. Given the design of our experiment we initially attributed the lack of correlation to the fact that the animals were fed two different diets that have been shown to have divergent effects on the metastatic potential. As such, the animals were separated according to diet and reevaluated for correlations. While the correlations between gene expression and metastatic virulence were slightly stronger in animals only fed the HFD compared to those detected in the whole population, they still were not significant after

adjusting for multiple comparisons. However, when the correlations were evaluated in animals fed the MCD, 14 genes were found to have significant correlations with the metastatic phenotype (Table 1). When these genes were evaluated with IPA, the major network functions included cancer, genetic disorders, cellular assembly and organization, and connective tissue disorders. IPA identified direct and indirect connections between several of these genes within this set and *TP53*, a well known oncogene.

eQTL evaluation: Totals of 220 *cis*-acting eQTL and 890 *trans*-acting eQTL (Figure 2 A and B, and Supplemental Tables 1 and 2) were detected. The *cis*-acting eQTL were found across all chromosomes with Chr 11 and 17 harboring the most, while Chr 2 and 16 had the fewest. *Trans*-acting eQTL were likewise distributed across all the chromosomes with Chr 3, 8 and 19 containing the largest numbers. When averaged, the LRT score for all *cis*-acting eQTL was 45.9 with individual LRT scores ranging from 16.2 to 197.1. The average LRT score for all *trans*-acting eQTL was significantly less at 23.9. However, the range for individual LRT scores amongst the *trans*-acting eQTL was similar to that observed amongst *cis*-acting eQTL. If the *trans*-acting eQTL were evaluated as two separate groups, those mapping to the same chromosome (but not within 10 cM) as the actual gene they represent and those mapping to different chromosomes, the average LRT scores were 37.5 and 19.0 respectively.

While the *trans*-acting eQTL represent loci controlled by unknown regulators, *cis*-acting eQTL exhibit self-regulation (Alberts et al. 2005). As such, by comparing/overlaying these *cis*-acting eQTL with the phenotypic cancer QTL detected in this population, the

potential exists to reveal the unknown polygenes (or at least to provide gene candidates) that result in the metastatic mammary cancer phenotype. Utilization of this method enabled the detection of 76 potential candidates for the previously detected metastatic QTL and 95 potential candidates for tumor growth and latency QTL (Supplemental Table 3 and 4). These candidates were further refined based on their proximity to the phenotypic QTLs. If genes only within 15 cM of the phenotypic QTL peaks are considered, then the number of candidates is reduced from 76 to 44 and from 95 to 33 for metastatic and tumor growth/latency QTL respectively (Tables 2 and 3). We had previously identified metastatic modifiers on Chr 1, 8, 11, 13 and 19 (Gordon et al. 2008a) and the distribution of eQTL-based candidate genes relative to phenotypic QTL is portrayed in Figure 3. We has previously identified tumor growth and latency modifiers were identified on Chr 1, 5, 7, 9, 13, 14 and 17 (Gordon et al. 2008a) and the distribution of eQTL-based candidate genes relative to phenotypic QTL is portrayed in Figure 4.

To investigate the functions of all *cis* and *trans*-eQTL the two separate gene lists were entered into IPA and evaluated with the core analysis function. Of the 220 *cis*-eQTL, 206 were found in the IPA database and of those eQTL, 117 were eligible for network and bio-function analysis (Table 4). The top bio-functions as indicated by IPA in the *cis*-acting dataset were cancer and cellular movement. When the bio-functions were expanded to dissect more specifically the categories, it revealed that the genes in the cancer set fell into many subgroups, including apoptosis, tumor growth, migration, proliferation and invasion, along with categories representing a wide spectrum of specific cancer varieties. When the cellular movement group was expanded, the top functions were migration, invasion and localization. Additionally genes were partitioned into networks based on the known or predicted

interactions from the IPA database and ranked by IPA based on the number of candidate genes that appear within the dataset. Assessing the top functions of the highest ranked datasets revealed that many of these genes were involved in metabolism, cellular proliferation/death, and multiple processes of cancer.

These same analyses were performed on the *trans*-acting eQTL and out of the 890 eQTL, 808 were identified in the IPA database, and of those 515 were network and bio function eligible (Table 4). The top bio-functions represented in this dataset included carbohydrate metabolism, small molecule biochemistry and cancer. When the cancer bio-function category was expanded to provide further insight, the top functions of genes within this category were revealed to be involved in tumorigenesis, apoptosis and cell death. The expansion of the carbohydrate metabolism set identified genes involved in the metabolism of fructose-6-phosphate, glycosaminoglycan and UDP-N-acetylglucosamine The networks with the highest scores in the *trans*-eQTL data shared similar functions to the top networks in the *cis*-eQTL, such as the processes involved in metabolism, small molecule biochemistry and cancer. DNA replication and repair, gene expression as well as endocrine disorders were represented as well.

Diet interactions: Previously when evaluating this population for QTL we found that the majority of cancer related modifier loci exhibited interactions with dietary fat levels. For example, the metastatic modifier we previously identified on Chr 8 resulted from the presence of a significant QTL in mice fed the high–fat diet but not the control-diet (Gordon et al. 2008a). As such it was possible that the *cis*-acting eQTL would likewise show dietary

interactions. These dietary interactions could potentially be used as a filter to predict which candidates within the phenotypic QTL confidence intervals may be responsible for the phenotypic variation observed. However, only ~7% of the *cis*-eQTL had diet interactions (Table 5). Of the interactions detected, four resulted from differential allelic effects within the two diets, the first, *Slc43a3* mapped to Chr 2, the next, *Basp1* was identified on Chr 15, while *E4f1* was detected on Chr 17 and *Frmd8* mapped to Chr 19. All other eQTL by diet interactions resulted from the detection of a significant effect for animals fed one diet but not the other; the majority of significant effects were found in animals fed MCD.

Causality evaluation: One way to further refine the candidate gene list is to evaluate the relationship between the *cis*-eQTL and the observed metastatic modifier loci. If eQTL have a causal relationship with these metastatic loci then they may in turn represent the actual underlying genetic modifier of the phenotype of interest (Schadt et al. 2005). The results revealed that very few significant causal relationships existed between *cis*-eQTL and metastatic virulence QTL. In only a few situations were weak independent associations detected between eQTL and phenotypic QTL such as those observed on Chr 1, 8 and 19 (Table 6). In addition to the independent associations, one weak causal relationship, which bordered on the threshold of significance, was detected between *H2afv* and the metastatic modifier on chromosome 11 (Table 6).

Oncomine evaluation: The evaluation of the Oncomine database revealed that 22 of the 45 candidates for the metastatic QTL have been previously reported in human breast

cancer prognosis datasets (Table 7). Of the 22 candidates detected, five were found to have an association between both increased and decreased gene expression and the clinical phenotype among the different datasets. Six of the candidates were found to have an increase in their expression associated with a better clinical outcome. In this category the candidates with the greatest number of human studies supporting their association were *Dusp4* on Chr 8 and *Cxcl14* on Chr 13 with nine and five studies providing evidence, respectively. The final 11 candidates were found to have an increase in their expression be associated with a poorer clinical prognosis. The candidate with the most compelling evidence in this group was *H2afv* on Chr 11 with seven studies supporting this link.

Master Regulator Analysis: Another observed phenomenon in large transcriptome mapping studies are chromosomal segments enriched with *trans*-acting genes forming what are commonly referred to as trans-bands or master regulatory regions. When the *trans*-eQTL detected in this intercross were evaluated in 5 cM windows, two potential master regulatory regions were identified on Chr 3 and 19 (Figure 2 B). The region on Chr 3 contains 98 *trans*-acting eQTL from 14 cM through 18 cM. The region on Chr 19, from 13 cM through 17 cM, contains 60 *trans*-acting eQTL. To further evaluate these regions, the IPA comparison function was used to determine if any overlap exists between the co-localized *cis* and *trans*-acting eQTL. This analysis failed to reveal any known connections.

Given that IPA is limited to only identifying previously reported associations, the actual relationships between eQTL in the master regulatory regions may go undetected. As such, causality analysis was utilized to further investigate the relationships in master

regulatory regions between *cis*-eQTL and *trans*-eQTL. These analyses revealed potential candidates for the master regulators on Chrs 3 and 19. On Chr 3 the *cis*-eQTL *Pkia* formed significant causal relationships with 93 *trans*-acting eQTL. The eQTL peak for *Pkia* did not map within the master regulatory region on Chr 3. However, the confidence interval for this eQTL extended from 4.4 cM to 17.4 cM, encompassing almost the entire master regulator region. In the master regulator region on Chr 19 *Pkd2l1* was identified as being causally related to 54 *trans*-acting eQTL. Similar to *Pika*, the eQTL peak for *Pkd2l1* was not located within the master regulatory region on Chr 19, but again the confidence interval, which extended form 8 cM to 45 cM, contained the entire master regulatory region.

CNV evaluation: These results revealed that while situations of copy number variation did indeed exist (Figure 5), very few of these segments were shared among the animals tested. Four regions were identified that contained patterns of chromosomal rearrangements, the first of which was a short 0.5 Mb segment on Chr 8 at approximately 20 Mb. A 9 Mb segment at the distal end of Chr 11 was also detected at which 58% of the animals tested had an additional copy of this region. On Chr 13 a 1 Mb segment starting at 65 Mb was identified where 11% of the population had a single amplification and 17% had a single deletion. The final region detected was a 0.25 Mb segment staring at 9.6 Mb on Chr 19 where 8% had a single amplification and 20% of the animals had a single deletion.

When these copy number changes were evaluated to determine if they were associated with metastatic development, no relationship was identified within the regions detected on Chrs 13 and 19, but significant associations within the regions of chromosome 8

(Figure 6 A) and 11 (Figure 6 B) were found. On Chr 8 individuals with a deletion of the region had significantly less pulmonary metastases compared to individuals with no CNV, while individuals with an amplification of the region had almost a threefold increase in the number of metastasis detected when compared to individuals with no CNV. On the other hand, individuals with the amplification of the segment on Chr 11 had a significant reduction in metastatic development. This evaluation was expanded to investigate whether the expression levels of the *cis*-eQTL on these two Chrs were associated with the detected CNVs. Only one *cis*-eQTL, *Ascc2* which is located on Chr 11, was found to be associated with either of these copy number variations. A small but significant increase in expression of this gene was detected in animals that had amplification at the distal end of this chromosome (Figure 6 C).

Discussion:

Breast cancer is one of the most common cancer types diagnosed in the United States. In most cancer cases the primary tumor is considered nonfatal and if removed early enough, total remission should follow. Yet, in many instances as tumor growth progresses, abnormal cells invade the lymphatic system or other vasculature and metastasize to distant sites in the body, such as the brain, bones, and lungs. This advanced cancer is far more difficult to treat and typically results in mortality after a course of a few years or more (Murphy 2001). While the variability in the pathogenesis of breast cancer is most likely influenced by a combination of environmental elements (e.g. diet) and genetic predisposition, the underlying mechanisms resulting in the altered disease state still remain unclear. Since it has been estimated that 60-

70% of patients have progressed to metastatic disease by the time of their diagnosis (Eccles et al. 1994), the elucidation of the genetic underpinnings influencing metastatic cancer is essential for decreasing cancer mortality.

This pressing need to understand the genetic architecture of metastatic mammary cancer prompted the development of the F_2 intercross between mice genetically predisposed to obesity and metastatic mammary cancer (Gordon et al. 2008b). Utilizing this population, we characterized several primary tumor phenotypes, such as tumor latency and growth, as well as the development of pulmonary metastasis (Gordon et al. 2008b). In addition, we found that the consumption of a high fat diet not only contributes to increased tumor growth, but increased metastatic virulence as well. These cancer phenotypes along with the genotypic data obtained from this F_2 population facilitated the detection of multiple metastatic mammary cancer modifier loci, several of which exhibited QTL by diet interactions (Gordon et al. 2008b).

Interestingly, subsequent differential expression analyses within F_2 tumors revealed that very few genes were differentially expressed between the two dietary treatments (la Merrill et al, 2009). However, when the individuals were segregated based on metastatic tendencies, substantially more differentially expressed genes were identified (la Merrill et al, 2009). Given the wide range of metastatic phenotypes, the observed transcriptional differences in this genetically segregating population and the clinical importance of the metastatic process, further investigation into genetic architecture of metastatic mammary cancer was warranted.

Evaluation of candidates for metastatic QTL:

Detection of strong correlations between transcript expression and metastatic phenotypes was expected. This was a reasonable assumption, given the findings of the previous studies evaluating transcriptional data in large mapping populations that reported significant correlations between their observed clinical phenotypes and gene expression (Ghazalpour et al. 2005). Yet, after evaluating our data, very few genes expressed above background on our arrays were significantly correlated with any of the metastatic traits. In retrospect, this result was not entirely surprising given that previous studies have had conflicting success with experiments utilizing mice harboring the PyMT transgene (Crawford et al. 2008; Qiu et al. 2004).

One possibility for the limited correlations is that the metastatic process involves an intricate cascade of events (Carmeliet and Jain 2000; Howard et al. 2008; Kroemer and Pouyssegur 2008; Lunt et al. 2009), each of which may only require subtle changes in gene expression. If this were the case, then the combined network of genes could be significantly correlated with the clinical phenotype (Yvert et al. 2003), while the genes that comprise the network may individually be only weakly associated. The possibility also exists that genes expressed directly by the metastatic seed itself, or perhaps its host environment, could be more highly correlated with the metastatic phenotype than those expressed in the primary tumor tissue (Dong et al. 1999). However, Weigelt et al (2003), found that primary tumor cells and their matched distant metastatic cells had remarkably similar gene expression patterns. Weigelt et al findings suggest that tumors and metastasis are both relevant tissues for investigating the etiology of metastatic mammary cancer. Furthermore, it is more clinically relevant to be able to predict a tumor's ability to metastasize, prior to the event

actually occurring, and/or identify therapeutic targets which if perturbed may prevent the metastatic process altogether.

Relatively few studies have performed eQTL analyses utilizing cancer models in rodents. The populations that have been previously characterized evaluated models of skin tumors (Quigley et al. 2009), mammary cancer (Crawford et al. 2008) and prostate cancer (Yamashita et al. 2005). While the experiments by Crawford et al (2008) were successful in identifying a pathway altered in metastatic mammary cancer, they only focused on a select subset of genes which had been previously predicted to be important players in metastatic disease. Currently, no other full genome eQTL analysis for metastatic mammary cancer models has been reported. As such, we set out to expand this type of analysis on a larger scale by performing a whole genome eQTL analysis of metastatic mammary cancer, similar to analyses described previously in multiple other model species (Brem and Kruglyak 2005; Chesler et al. 2005; Schadt et al. 2003a; Yvert et al. 2003).

One caveat that needs to be considered when performing an eQTL study is the degree of CNV within the mapping population. Taking into account CNV is especially important when analyzing tumor tissue, given the substantial amount of evidence linking the accumulation of CNV to cancer pathogenesis (de Tayrac et al. 2009; Fridlyand et al. 2006; Namba et al. 2006; Reis-Filho et al. 2005; van Beers and Nederlof 2006; Yusenko et al. 2009). CNV can be classified into two categories, those that are inherited (germ-line) and those that are acquired during the replication of cells (somatic). In humans, germ-line CNVs are detected across all tissues in both healthy and diseased individuals (Shlien and Malkin 2009). The presence of CNV in genomic regions encoding cancer modifiers can lead to increased risk for the development of cancer (Albertson et al. 2003). Somatic CNV are

acquired during DNA replication and are not found uniformly throughout all tissue types, and their presence can impart a growth advantage to cells harboring them, resulting in disease (Greenman et al. 2007). Tsafrir et al. (2006) article suggests that as cancer progresses it is possible for tumors to continue acquiring somatic CNV, which could potentially alter their metastatic tendencies.

Our experimental design, in which we evaluated the tumor tissues against a reference of matched tail tissues, only provided us the ability to evaluate somatic CNV. The primary objective was to evaluate the impact of somatic CNV on both metastatic development and the detection of eQTL. The influence of CNV upon transcriptional studies has not been well characterized, yet clearly duplications or deletions of chromosomal segments, which can result in altered expression of genes residing within those boundaries, could influence the detection of eQTL. It is possible that without accounting for perturbations in gene expression mediated by CNV, eQTL results could be incorrectly interpreted. Additionally, given the findings of Williams et al (2009) where the *cis*-eQTL identified for *Glo1* was detected as a result of a duplicated region on CHR 17, CNV appears worth evaluating.

In our population all but one of the *cis*-eQTL acted independently of somatic CNV. *Ascc2*, a gene physically located outside the genomic boundaries of the CNV on Chr 11, was shown to have its gene expression significantly modified by a CNV (Figure 6C). It is possible that CNV interval we detected on Chr 11 contains an enhancer element that impacts the expression *Ascc2*, but is not the primary source of transcriptional variation for this candidate observed in the F_2 population. The presence of an enhancer element could potentially explain why the expression *Ascc2* could be modified by a distant CNV and still be classified as a *cis*-acting eQTL. Other investigators have likewise detected relationships

between copy number changes and the expression of genes located outside the CNV regions (Henrichsen et al. 2009; Stranger et al. 2007). As suggested by Reymond et al. (2007), this type of relationship indicates that transcriptional variation of genes encoded outside the boundaries of a CNV can potentially occur if a key regulatory element for that particular gene is encoded within the CNV interval. While we detected very little connection between *cis*-eQTL and CNV in our population, the link identified between the CNV on Chr 11 and *Ascc2*, highlights the importance of evaluating CNV in transcriptional studies in order properly interpret the results.

While several *cis* and *trans* regulated eQTL were identified across all chromosomes, the *cis*-acting eQTL mapping in close proximity to our previously detected metastatic cancer QTL were of primary interest. Prior experiments have suggested that *cis*-eQTL colocalizing with metastatic mammary cancer modifiers may represent logical positional and functional candidate genes for theses loci (Wang et al. 2007b; Yamashita et al. 2005). This is further supported by the discovery of *Rrp1b*, a *cis*-eQTL that colocalized with a previously detected metastatic QTL, and was later predicted to be clinically important in metastatic disease (Crawford et al. 2007). Filtering the *cis*-eQTL based on their proximity to our previously detected phenotypic QTL enabled the detection of 44 candidates colocalizing with metastatic modifier loci.

Causality testing has been touted as having the potential to predict whether a transcript is associated with a clinical phenotype locus. A *cis*-eQTL causally associated with and physically located close to a QTL could provide a logical candidate for that particular locus (Farber et al. 2009; Schadt et al. 2005). Using this method we detected one *cis*-eQTL that was causally associated with a metastatic locus, *H2afv* on Chr 11, coding for a protein

that is a member of the histone H2a family. Histone modifications have been detected in human cancers (Fraga et al. 2005) and can be used to provide insight into clinical outcomes (Seligson et al. 2005). Currently, very little information regarding the function of *H2afv* exists, but other members of the histone family H2A (*H2AX*) (Liu et al. 2007b) and proteins they encode (Buforin IIb) (Lee et al. 2008) have been shown to be potent inhibitors of cancer. While a casual relationship between *H2afv* and a metastatic locus on Chr 11 was detected, the overall limited number of observed associations (casual, reactive and independent) may be highlighting the complexity of the metastatic mammary cancer paradigm. For any particular complex trait, the intricate architecture that gives rise to the disease may be difficult to elucidate, because several genes and the networks they comprise likely contribute to the development of an endpoint phenotype. Additionally, if the metastatic QTL we detected were driven by coding changes not resulting in transcriptional variation then we would likely fail to detect causal relationships between that particular metastatic locus and the colocalizing *cis*-eQTL. It is also possible that the limited number of associations we detected may simply be a result of a lack of power. However, a lack of power seems a less likely scenario given that several other studies have detected causal relationships in utilizing smaller populations (Schadt et al. 2005; Sun et al. 2007). Whereas causality testing is useful tool for detecting whether a transcript is associated with a phenotypic locus, caution is required when interpreting a lack of association.

Given that previously QTL by diet interactions were identified in this F_2 population (Gordon et al. 2008a) we anticipated seeing interactions reflected in eQTL as well. A *cis*-acting eQTL interacting with diet and colocalizing with a metastatic QTL exhibiting a similar interaction may represent a likely candidate for that particular locus. However, only a small

number of interactions were detected in the transcriptome mapping population. Additionally, the eQTL for which we detected diet interactions colocalizing with a phenotypic QTL, such as *Gdi3* on Chr 13, *Frmd8* and *Ms4a6c* on Chr 19, resulted from significant eQTL effects in mice fed the MCD. The aforementioned results were not consistent with the interactions detected for the metastatic QTL that these eQTL colocalized with, which resulted from significant eQTL effects in mice fed the HFD. Whereas the limited number of interactions was initially surprising given our previous findings, this result seems consistent with the lack of differentially expressed genes between the two dietary treatments as we reported in la Merrill et al., (2009).

The limited number of interactions detected may indicate that indirect effects of diet upon the tumors may have been responsible for the dietary effect upon tumor growth and metastatic development that we previously observed. These effects may be mediated through transcriptional changes in another tissue such as the liver, known to respond metabolically to lipid consumption (Morgan et al. 2008). Links between hepatic expression of genes involved in the regulation of estrogenic compounds and breast cancer pathogenesis have been previously reported (Gong et al. 2007). This appears to be supported by our findings that 36 of the 211 genes differentially expressed in the liver between the two dietary treatments were involved in cancer processes (la Merrill et al 2009). It is also possible that the time point of tissue collection could be masking the diet effect. The axillary tumors were harvested at 11 weeks of age, therefore only providing a transcriptional picture of the tumors after they had been growing for several weeks. It is feasible that dietary induced changes are programmed into the tumor as it is beginning to develop, in which case we would likely fail to detect the presence of eQTL by diet interactions in axillary tumor collected at 11 weeks. While having

expression profiles from several time points could possibly provide insight into the etiology of dietary fat effects upon the tumors, these would come at the expense of the number of arrays evaluated per time point, reducing our power to detect eQTL.

The evaluation of the Oncomine database provides additional support for several of the candidates being involved in the development of metastatic mammary cancer. In addition the Oncomine results could potentially indicate which candidates might be clinically important in both humans and mice. Our analysis of this database yielded results that appear to support the link between *H2afv* and metastatic mammary cancer. Not only did we implicate *H2afv* as a candidate for the metastatic QTL on Chr 11 through causality analysis, the Oncomine evaluation supported a link between increased expression of this gene and a poorer clinical prognosis in humans. Another candidate, *Dusp4* was implicated by Oncomine as a modifier of metastatic mammary cancer in humans as well. *Dusp4*, a member of the dual-specificity phosphatase family, has been previously implicated as a potential tumor suppressor in a variety of cancer types (Chitale et al. 2009; Sieben et al. 2005) and may be potentially a strong candidate for the metastatic modifier identified on Chr 8 in our prior QTL analysis.

Whereas the evaluation of this database indentified several candidates that potentially influence the development of metastatic mammary cancer in both mice and humans, candidates not detected could also be modifiers of this disease in humans as well. One particular gene not identified in the Oncomine evaluation, which has been previously associated with metastatic mammary cancer in humans, is butyrophilin (*Btn1a1*) (Woelfle et al. 2003). *Btn1a1* was previously reported as being expressed significantly lower in mice with tumors metastasizing versus those without metastasis (la Merrill et al 2009). Subsequent

eQTL analysis provided evidence that *Btn1a1* is a potential modifier of the metastatic locus on Chr 13 (la Merrill et al. 2009).

Evaluation of master regulatory regions:

An additional use for causality testing is the ability to investigate genomic regions enriched with *trans*-eQTL. By applying the same methodologies used to detect *H2AFV* as a candidate for the metastatic locus on Chr 11 to these regions, we identified two *cis*-eQTL causally associated with a significant number of *trans*-eQTL. A *cis*-eQTL detected for *Pkia,* located at 5.4 cM on Chr 3, with a confidence interval extending from 4.4 cM to 17.4 cM, formed causal associations with 95% of the *trans*-eQTL located within this putative master regulator interval. The protein encoded by *Pkia* contributes to the inhibition of protein kinase A (PKA), a protein known to alter the functionality of numerous proteins via phosphorylation (Lum et al. 1999). The phosphorylation properties of PKA and the fact that it is a critical regulator of several metabolic pathways (Taylor et al. 2004) may provide the link between *Pkia* and the *trans*-eQTL mapping to Chr 3, which have top network functions that include several metabolic processes (Table 8) (as indicated by IPA) known to be regulated by PKA.

On Chr 19 the peak of the *cis*-eQTL for *Pkd2l1* mapped to 34.5 cM, well away from the master regulatory region at 13-17 cM. However, the master regulatory region was contained within the confidence interval detected for this eQTL. Results also indicated that it was casually associated with 90 % of the *trans*-eQTL in this master regulatory region. This gene encodes the protein *TRPP2*, a member of the transient receptor potential cation channel family. This protein is critical for the function of numerous biological processes across many

tissue types (Giamarchi et al. 2006). A link between this protein and the *trans*-eQTL appears to be supported by the fact that when the *trans*-eQTL are evaluated using IPA, several of the top functions are involved in the processes of the cell cycle, cellular growth and molecular transport (Table 8), all of which *TRPP2* could impact.

Confirmation of the relationship between both *Pkia* and *Pkd2l1* and the *trans*-eQTL they are causally associated with has yet to be confirmed. However, our results suggest that causality testing might potentially be a useful tool for investigating the structure of master regulatory regions. Additionally, these results suggest that it is important to not only evaluate the relationship between the *trans* and *cis*-eQTL mapping within the master regulatory region, but to evaluate *cis*-eQTL with confidence intervals that encompass the master regulatory region of interest as well.

Conclusions:

In conclusion, this work demonstrates the utility of using transcriptome mapping to identify candidates for previously detected QTL. This report represents is one of the most complete evaluations of metastatic mammary cancer to date, drawing upon on several techniques to provide insight into the genetic architecture of this disease. Utilizing CNV analysis we detected a somatic alteration on chromosome 11 that was associated with increased expression of *Ascc2*, a gene for which a *cis*-eQTL was detected as a potential candidate for a metastatic modifier on Chr 11. Additionally, to our knowledge this is the first eQTL analysis to report eQTL by diet interactions. Whereas only a few interactions were detected, their identification highlights the importance of accounting for eQTL by

environment interactions in transcriptional mapping studies in order to properly interpret the results. Furthermore, the use of causality testing provided insight into the relationship between the *cis*-eQTL detected for *H2afv* and the metastatic loci on Chr 11. The potential roles of *H2afv* and *Dusp4* in metastatic mammary cancer were further supported by our evaluation of the Oncomine database. The use of causality analyses also provide an interesting look at the structure of our master regulator regions on Chr 3 and 19 and helped identify two potential master regulator candidates, *Pika and Pkd2l1*. Ultimately, these findings may provide additional insight into the intricate cascade of metastatic mammary cancer, which can be potentially applied in a clinical setting. The increased understanding of this disease could provide both, a predictive analysis of cancer pathogenesis, and possible targets for therapeutic interventions.

Table 5.1: Significant correlations between gene expression and metastatic phenotype after adjustment for multiple test comparisons in animals fed the MCD

Gene	Chr	Correlation	Raw P-value	Adjusted	Function
Ubc	5	-0.53	0.000005	0.041	Unknown
S100a3	3	0.50	0.000014	0.041	calcium ion binding, protein binding
4933428G20Rik	11	-0.50	0.000017	0.041	Unknown
Arhgap17	7	-0.49	0.000023	0.041	protein binding
Nrm	17	-0.49	0.000029	0.041	Unknown
Bscl2	19	-0.49	0.000031	0.041	Unknown
Vps53	11	-0.49	0.000031	0.041	Unknown
Psmd3	11	-0.48	0.000033	0.041	protein binding , enzyme regulatory activity
Slc25a39	11	-0.48	0.000039	0.041	Binding
Sdhc	1	-0.48	0.000040	0.041	succinate dehydrogenase activity
Ctsk	3	-0.47	0.000051	0.042	protein binding , peptidase activity
Col1a1	11	-0.47	0.000056	0.042	ECM structure, protein binding
Col5a1	2	-0.47	0.000056	0.042	ECM structure, integrin binding
Sfpq	4	-0.47	0.000058	0.042	nucleotide binding, protein binding
Hist1h4m	13	0.46	0.000085	0.054	Unknown

Table 5.2: Candidate eQTL based on proximity to metastatic QTL

Accession	Symbol	Chr	LRT	LOD	eQTL peak (cM)[a]	eQTL to Gene (cM)[b]	eQTL to QTL (cM)[c]
NM_183019.1	*9330140K16Rik*	1	23.3	5.1	19.9	2.0	12.9
NM_133791.3	*Wwc2*	8	20.4	4.4	28.2	1.1	2.8
NM_023312.2	*Ndufa13*	8	91.6	19.9	36.2	7.0	5.2
NM_032544.2	*Gtpbp3*	8	24.9	5.4	36.2	8.0	5.2
NM_028993.2	*9130404D08Rik*	8	24.4	5.3	38.2	5.0	7.2
NM_176933.3	*Dusp4*	8	33.1	7.2	22.2	1.2	8.8
	9530006C21Rik	8	45.6	9.9	21.2	0.7	9.8
NM_019733	*Rbpms*	8	77.6	16.9	21.2	0.8	9.8
XM_109956	*Wwc1*	11	34.6	7.5	13.9	8.4	0.1
NM_177364	*Sh3pxd2b*	11	47.4	10.3	15.9	4.3	1.9
NM_029291.1	*Ascc2*	11	23.9	5.2	9.9	7.0	4.1
NM_009288.1	*Stk10*	11	18.4	4.0	19.9	0.4	5.9
NM_008698.1	*Nipsnap1*	11	22.4	4.9	7.9	4.9	6.1
XM_109868	*Tens1*	11	77.1	16.8	6.9	1.6	7.1
XM_488586	*2210015D19Rik*	11	93.6	20.3	6.9	3.4	7.1
XM_126043.3	*H2afv*	11	110.0	23.9	5.9	1.9	8.1
NM_134033.1	*Ccdc117*	11	67.6	14.7	3.9	0.5	10.1
NM_178187.2	*Hist1h2ae*	13	40.3	8.8	13	1.4	0
NM_024274.1	*Fars1*	13	61.4	13.3	15	7.5	2
NM_015786	*Hist1h1c*	13	134.9	29.3	10	4.5	3
NM_198093.2	*Elmo1*	13	16.7	3.6	9	3.6	4
NM_025387.1	*Tmem14c*	13	155.0	33.7	17	8.4	4
NM_026947.2	*1810022C23Rik*	13	18.6	4.0	18	3.5	5
NM_175655.1	***Hist1h4f*	13	28.2	6.1	18	3.6	5
NM_013483.1	***Btn1a1*	13	33.8	7.4	7	7.3	6

NM_178194.2	*Hist1h2be*	13	31.3	6.8	5	9.4	8
NM_008112.2	*Gdi3*	13	129.4	28.1	4	1.8	9
NM_009124.2	*Sca1*	13	49.0	10.6	25	3.2	12
NM_019568	*Cxcl14*	13	16.5	3.6	25	9.3	12
NM_173442.1	*Gcnt1*	19	23.5	5.1	14.5	4.2	1
NM_173442.1	*Gcnt1*	19	45.0	9.8	14.5	4.2	1
NM_026487.2	*Atad1*	19	16.2	3.5	13.5	6.5	2
NM_009199.1	*Slc1a1*	19	27.9	6.1	19.5	1.9	4
NM_028595	*Ms4a6c*	19	16.3	3.5	10.5	3.8	5
NM_021890	*Fads3*	19	39.0	8.5	10.5	4.7	5
NM_146097.1	*Cbwd1*	19	133.6	29.0	20.5	5.4	5
NM_013754.1	*Insl6*	19	55.7	12.1	21.5	3.6	6
NM_026169.3	*Frmd8*	19	20.3	4.4	8.5	5.0	7
NM_134154.1	*Slc25a45*	19	30.6	6.6	5.5	2.0	10
NM_021474.2	*Efemp2*	19	75.9	16.5	4.5	1.2	11
NM_019861.1	*Ctsf*	19	45.8	10.0	4.5	1.6	11
NM_016892.2	*Ccs*	19	16.8	3.7	4.5	1.6	11
AK032179	*Saps3*	19	23.5	5.1	3.5	1.5	12
NM_021460.1	*Lip1*	19	141.4	30.7	28.5	7.4	13

[a]eQTL peak: The estimated peak position of the eQTL in cM

[b]eQTL to gene: The distance in cM from the peak eQTL position to the position of the gene it represents.

[c]eQTL to QTL: The distance in cM from the peak eQTL position to the nearest metastatic QTL

[**] Genes previously reported as differentially expressed between tumors of varying metastatic tendencies.

Table 5.3: Candidate eQTL based on proximity to tumor growth and latency QTL

Accession	Symbol	Chr	LRT	LOD	eQTL peak (cM)[a]	eQTL to Gene (cM)[b]	eQTL to QTL (cM)[c]
NM_021099	*Kit*	5	48.7	10.6	43.6	3.0	4.1
NM_175270.2	*5730467H21Rik*	5	54.2	11.8	48.6	9.1	9.1
NM_007635.2	*Ccng2*	5	40.0	8.7	49.6	8.3	10.1
NM_016974.1	*Dbp*	7	16.7	3.6	32	7.1	10.0
AK030267	*4933439J20Rik*	7	18.4	4.0	54	6.1	12.0
XM_135023.2	*2610018I03Rik*	9	55.0	12.0	45.2	9.0	10.8
AK004616	*Slc21a2*	9	32.6	7.1	60.2	4.2	4.2
NM_009275.2	*Srprb*	9	33.5	7.3	62.2	2.3	6.2
NM_009153.1	*Sema3b*	9	69.3	15.1	60.2	7.1	4.2
NM_199195.1	*Bckdhb*	9	54.0	11.7	46.2	6.6	9.8
NM_008112.2	*Gdi3*	13	129.4	28.1	4	1.8	15.0
NM_178194.2	*Hist1h2be*	13	31.3	6.8	5	9.4	14.0
NM_013483.1	*Btn1a1*	13	33.8	7.4	7	7.3	12.0
NM_198093.2	*Elmo1*	13	16.7	3.6	9	3.6	10.0
NM_015786	*Hist1h1c*	13	134.9	29.3	10	4.5	9.0
NM_178187.2	*Hist1h2ae*	13	40.3	8.8	13	1.4	6.0
AK021333	*Btn1a1*	13	21.1	4.6	14	0.3	5.0
NM_024274.1	*Fars1*	13	61.4	13.3	15	7.5	4.0
NM_025387.1	*Tmem14c*	13	61.5	13.4	17	8.4	2.0
NM_025387.1	*Tmem14c*	13	155.0	33.7	17	8.4	2.0
NM_175655.1	*Hist1h4f*	13	28.2	6.1	18	3.6	1.0
NM_026947.2	*1810022C23Rik*	13	18.6	4.0	18	3.5	1.0
NM_009124.2	*Sca1*	13	49.0	10.6	25	3.2	6.0

NM_019568	*Cxcl14*	13	16.5	3.6	25	9.3	6.0
NM_023879.1	*Rpgrip1*	14	38.7	8.4	27.9	1.6	12.1
NM_027436.1	*Mipep*	14	91.4	19.9	27.9	6.7	12.1
NM_009029.1	*Rb1*	14	24.6	5.4	33.9	8.3	6.1
NM_016903.2	*Esd*	14	20.7	4.5	45.9	2.7	5.9
320628	*A130038J17Rik*	14	20.9	4.5	45.9	0.1	5.9
NM_008549.1	*Man2a1*	17	31.7	6.9	29.2	9.8	12.8
78592	*A330106M24Rik*	17	35.8	7.8	39.2	5.9	2.8
NM_152817.2	*2610511O17Rik*	17	18.0	3.9	44.2	1.2	2.2
NM_144802.2	*2810036L13Rik*	17	29.3	6.4	46.2	2.5	4.2

[a]eQTL peak: The estimated peak position of the eQTL in cM

[b]eQTL to gene: The distance in cM from the peak eQTL position to the position of the gene it represents.

[c]eQTL to QTL: The distance in cM from the peak eQTL position to the nearest tumor growth/latency QTL

Table 5.4: IPA Evaluation for *cis/trans*-eQTL

	Mapped Id's[a]	Eligible eQTL[b]	Bio functions[c]	Networks[d]
cis-acting eQTL	206	117	Cancer (25)	Cell death, Lipid metabolism, Small molecule biochemistry (22)
			Cellular movement (10)	Lipid metabolism, Small molecule biochemistry, cell morphology (20)
			Gastrointestinal Disease (6)	Genetic disorder, Neurological disease, Ophthalmic disease (14)
			Cell cycle (9)	Amino acid metabolism, Cancer, Carbohydrate metabolism (14)
trans-acting eQTL	808	515	Carbohydrate metabolism (14)	Gene expression, DNA replication and repair, Endocrine system disorders (32)
			Small molecule biochemistry (54)	Drug metabolism, Cancer, Lipid metabolism (31)
			Gene expression (21)	Carbohydrate metabolism, Small molecule biochemistry, immune response (31)
			Cancer (174)	Gene expression, Cellular development, Nervous system development/function (29)

[a] The number of candidates that were found in the IPA database

[b] The number of candidates that had corresponding bio-function and network information in the IPA database

[c] The top bio-functions as indicated by IPA . The number in () indicates how many of the candidates from the eligible list are involved in the function

[d] The functions of the top networks as indicated by IPA. The number in () indicates how many of the candidates are in the particular network

Table 5.5: Significant cis-eQTL by diet interactions

Symbol	Chr	Interaction[a]	LRT	eQTL peak
E030013I19Rik	2	MCD	18.48	13.8
Dnajc1	2	MC	23.40	7.8
Slc43a3	2	B	42.13	46.8
Accs	2	HFD	22.93	55.8
Wfdc3	2	MCD	47.08	94.8
Prdx2	8	MCD	16.70	52.2
Bckdhb	9	MCD	54.01	46.2
Gdi3	13	MCD	129.35	4
D14Ertd449e	14	HFD	85.90	15.9
Basp1	15	BD	23.16	16.4
E4f1	17	BD	18.25	13.2
Notch4	17	MCD	26.02	14.2
Frmd8	19	BD	20.31	8.5
Ms4a6c	19	MCD	16.26	10.5
Pkd2l1	19	MCD	24.48	34.5

[a]Interaction: (HFD) Significant effect in high-fat diet only; (MCD) Significant effect in the control diet only; (B) Differential effects in high-fat and control diets

Table 5.6: Causality results between cis-eQTL and metastatic phenotype

Marker (cM)[a]	Chr	eQTL	Phenotype	Relationship[b]	P-value
C1L10 (34.1)	1	*3222401M22Rik*	Met	Independent	0.04
C8L3 (20.5)	8	*Afg3|1*	Met	Independent	0.05
C8L6 (30.4)	8	*Gtbp3*	Met	Independent	0.05
C11L2 (2.9)	11	*H2afv*	Met	Causal	0.05
C19L6 (12.8)	19	*6430407L02Rik*	Met	Independent	0.04
C19L6 (12.8)	19	*AW491445*	Met	Independent	0.04
C19L6 (12.8)	19	*1200004M23Rik*	Met	Independent	0.04
C19L6 (12.8)	19	*Gcnt1*	Met	Independent	0.05
C19L9 (23.3)	19	*1200004M23Rik*	Met	Independent	0.05

[a] Marker: The genetic marker (Gordon et al. 2008b)used as the anchor for the analysis

[b] Relationship: Testing of the association between the marker, eQTL and phenotype

Table 5.7: Results from Oncomine database evaluation

Symbol	Chr	NF [a]	EE positive [b]	EE negative [c]
9330140K16Rik	1	x		
Wwc2	8			1
Ndufa13	8		2	2
Gtpbp3	8	x		
9130404D08Rik	8	x		
Dusp4	8		9	
9530006C21Rik	8	x		
Rbpms	8		4	
Wwc1	11			1
Sh3pxd2b	11	x		
Ascc2	11	x		
Stk10	11		1	4
Nipsnap1	11		1	1
Tens1	11			1
2210015D19Rik	11	x		
H2afv	11			7
Ccdc117	11	x		
Hist1h2ae	13			4
Btn1a1	13	x		
Fars1	13			1
Hist1h1c	13			4
Elmo1	13	x		
Tmem14c	13	x		
1810022C23Rik	13	x		
Hist1h4f	13	x		
Btn1a1	13	x		

Hist1h2be	13			3
Gdi3	13			1
Sca1	13		4	
Cxcl14	13		5	
Gcnt1	19		1	2
Gcnt1	19		1	2
Atad1	19	x		
Slc1a1	19		4	
Ms4a6c	19	x		
Fads3	19	x		
Cbwd1	19		1	
Insl6	19	x		
Frmd8	19	x		
Slc25a45	19			1
Efemp2	19			1
Ctsf	19	x		
Ccs	19	x		
Saps3	19	x		
Lip1	19	x		

The 45 candidates for the Metastatic QTL were evaluated in the Oncomine database to determine if a potential association between their expression and formation of metastasis has been previously reported in humans.

[a] NF: An x indicates that the particular candidate was not found in the Oncomine human breast cancer prognosis datasets

[b] EE positive: Indicates the number of studies reporting that increased expression is significantly associated with a positive (no metastasis, alive, disease free) clinical outcome.

[c] EE negative: Indicates the number of studies reporting that increased expression is significantly associated with a negative (metastasis, dead, relapse) clinical outcome.

Table 5.8: IPA Evaluation for *trans*-eQTL in Master regulator intervals

	Mapped Id's[a]	Eligible eQTL[b]	Networks[c]
MR Chr 3	85	65	Carbohydrate metabolism, Drug metabolism, Small molecule biochemistry (16)
			Cell cycle, cancer, Hematological disease (15)
			Carbohydrate metabolism, Small molecule biochemistry, Connective tissue development and function (14)
			Endocrine system development, Lipid metabolism, Small molecule biochemistry (14)
MR Chr 19	49	45	Cancer, Cell-to-cell signaling and interaction, Cellular assembly and organization (13)
			Cancer, Reproductive system disease, Tumor morphology (11)
			Cancer, Cellular function and maintenance, Respiratory disease (11)
			Amino acid metabolism, Molecular transport, Small molecule biochemistry (1)

[a] The number of candidates that were found in the IPA database

[b] The number of candidates that had corresponding network information in the IPA database

[c] The functions of the top networks as indicated by IPA. The number in () indicates how many of the candidates are in the particular network

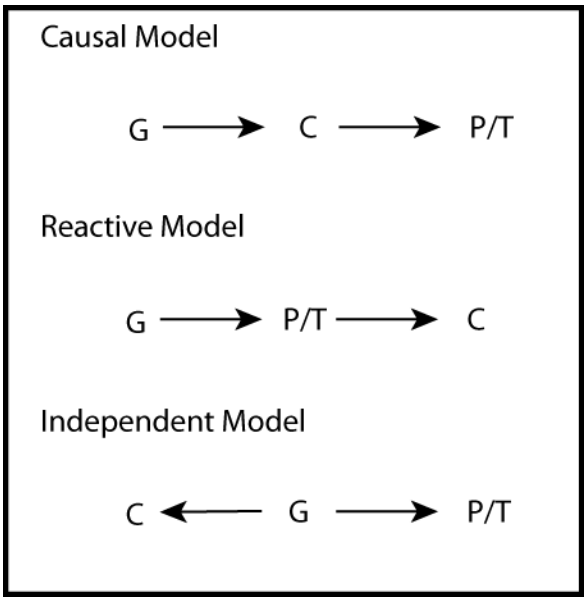
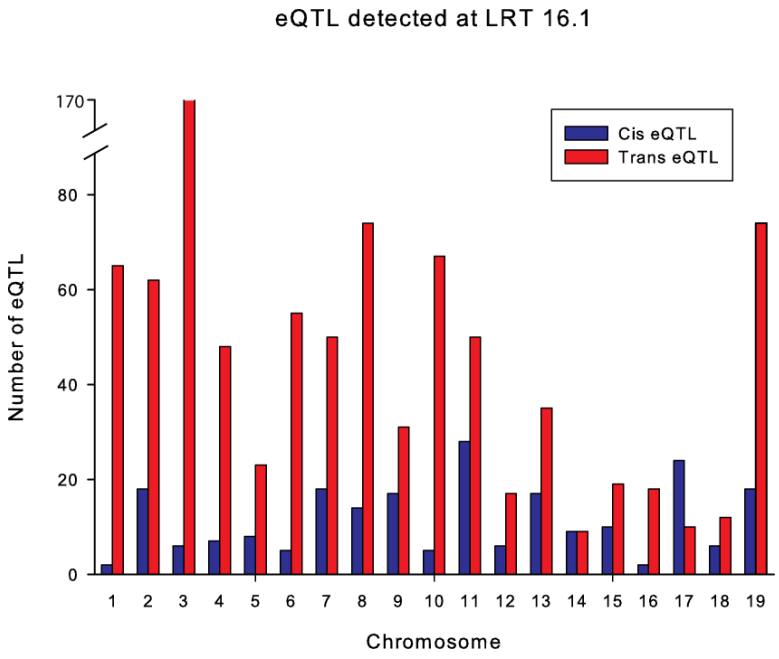


Figure 5.1: Causality Evaluation Models: The three causality models tested. The causal model is where genetic alterations (G) result in changes in the expression of an *cis*-acting eQTL (C), which in turn result in modification of the phenotype or change in gene expression of a *trans*-acting eQTL (P/T). In the reactive model, variation in G directly impacts P/T resulting in altered gene expression of C. The third model evaluates whether variation in G can result in changes in both P/T and C independently.

A. Number of *cis/trans* on each chromosome



B. Map of expression QTL (eQTL)

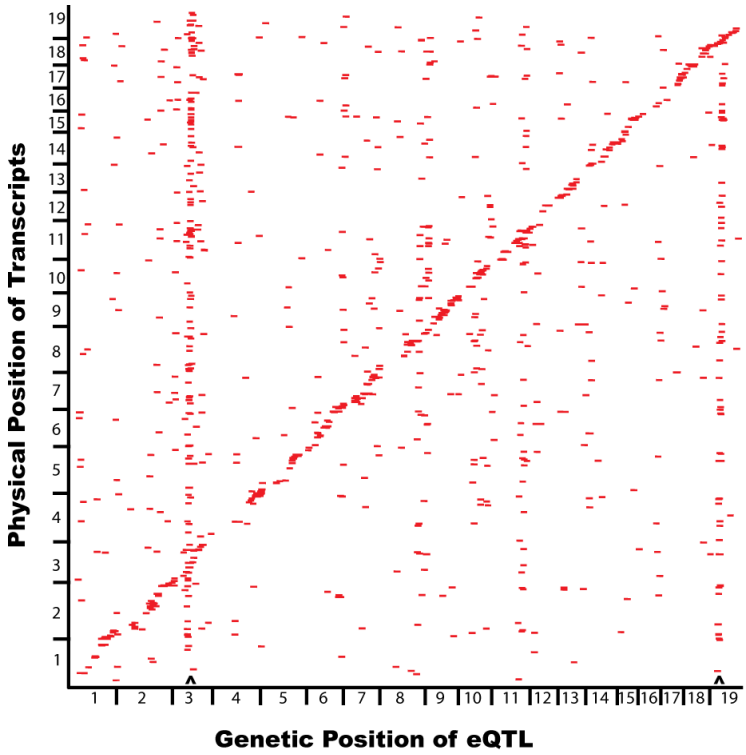
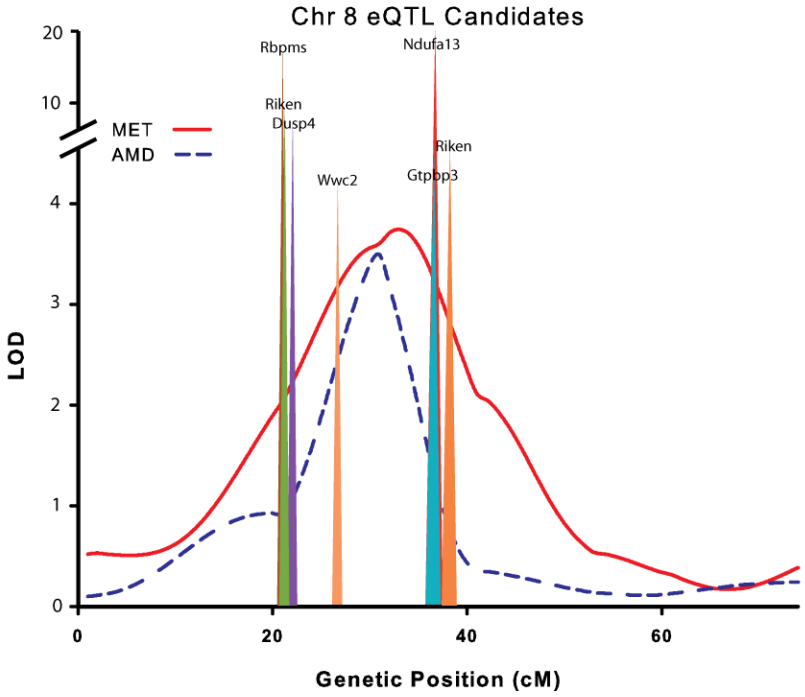
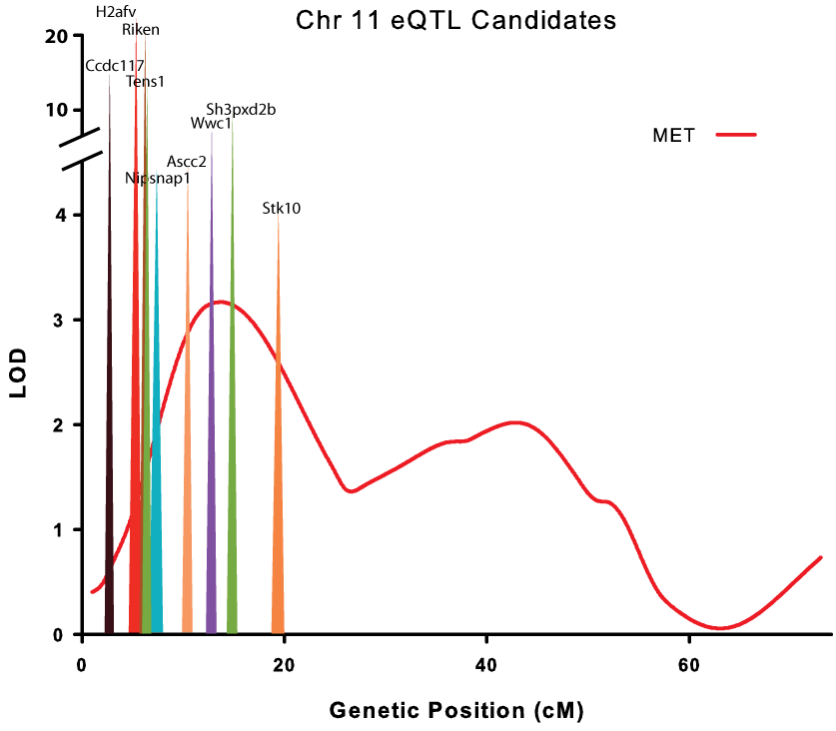


Figure 5.2: eQTL Mapping: The eQTL mapping results. A: The number of *cis and trans*-eQTL detected on each chromosome at the significance LRT threshold of 16.1. B: Mapping of expression QTL (eQTL). Physical gene positions on the microarray are plotted along the y-axis and the genetic locations of the QTL are plotted along the x-axis. eQTL along the cis diagonal map within 10 cM of the transcript that they represent. eQTL acting in trans map to a different chromosome than the transcript they represent. ^Master regulators are eQTL acting in cis that map to a region associated with many trans-acting eQTL.

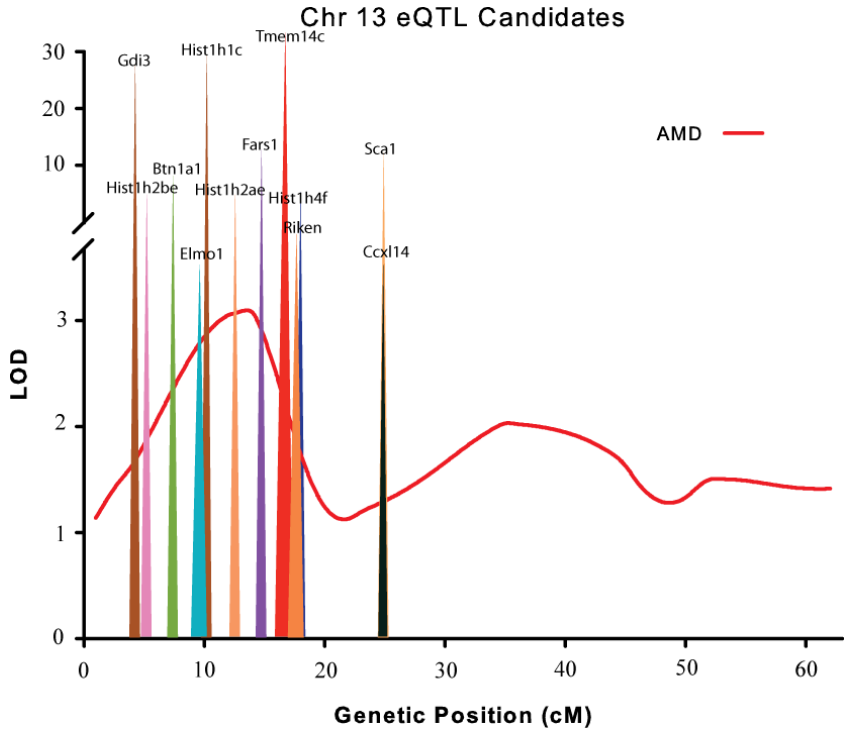
A. cis-eQTL colocalizing with the metastatic locus on Chr 8



B. *cis*-eQTL colocalizing with the metastatic locus on Chr 11



C. *cis*-eQTL colocalizing with the metastatic locus on Chr 13



D. *cis*-eQTL colocalizing with the metastatic locus on Chr 19

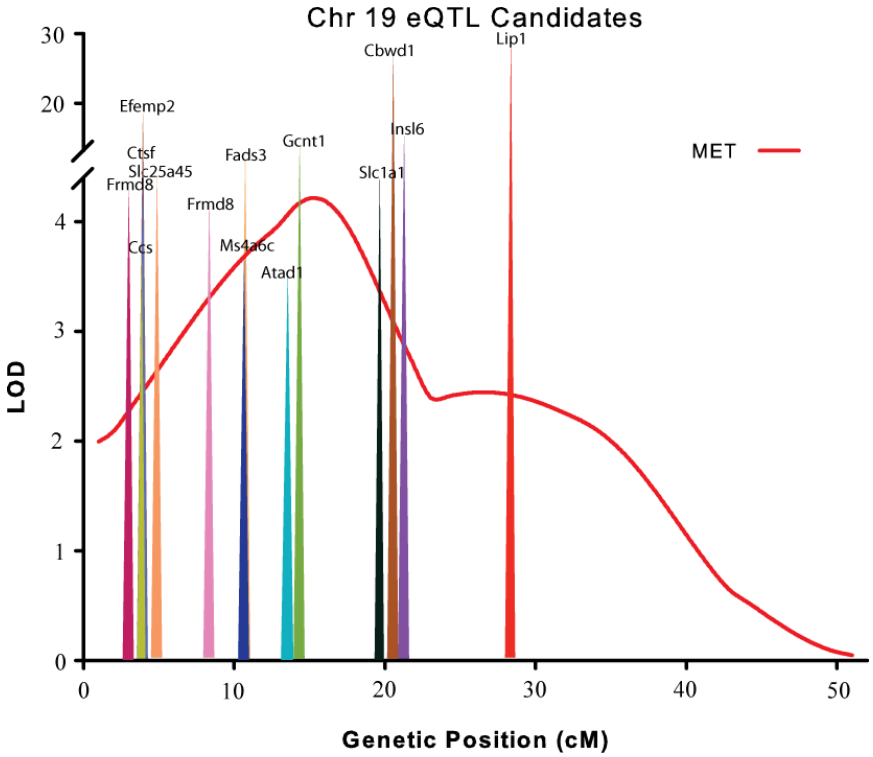
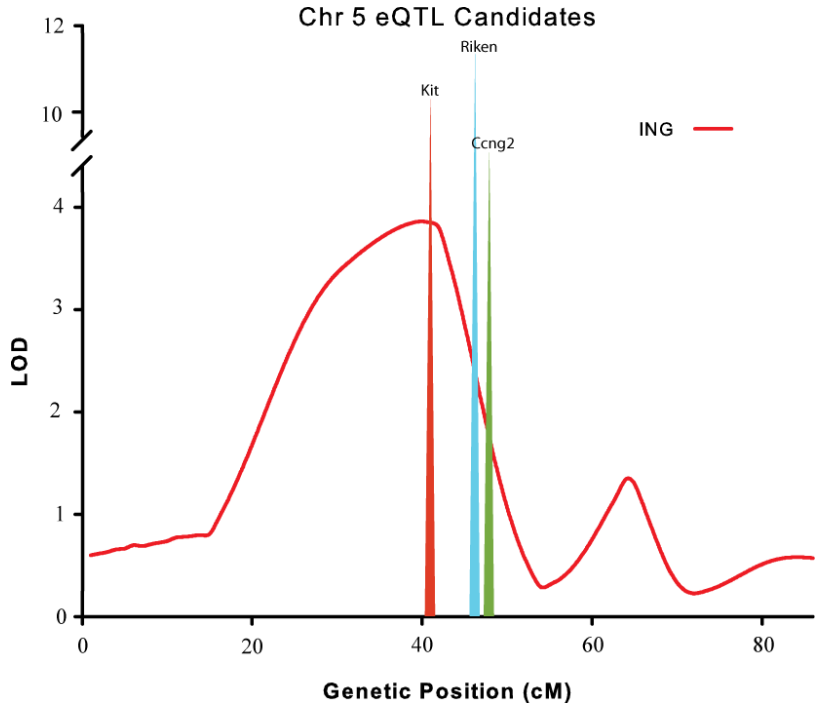
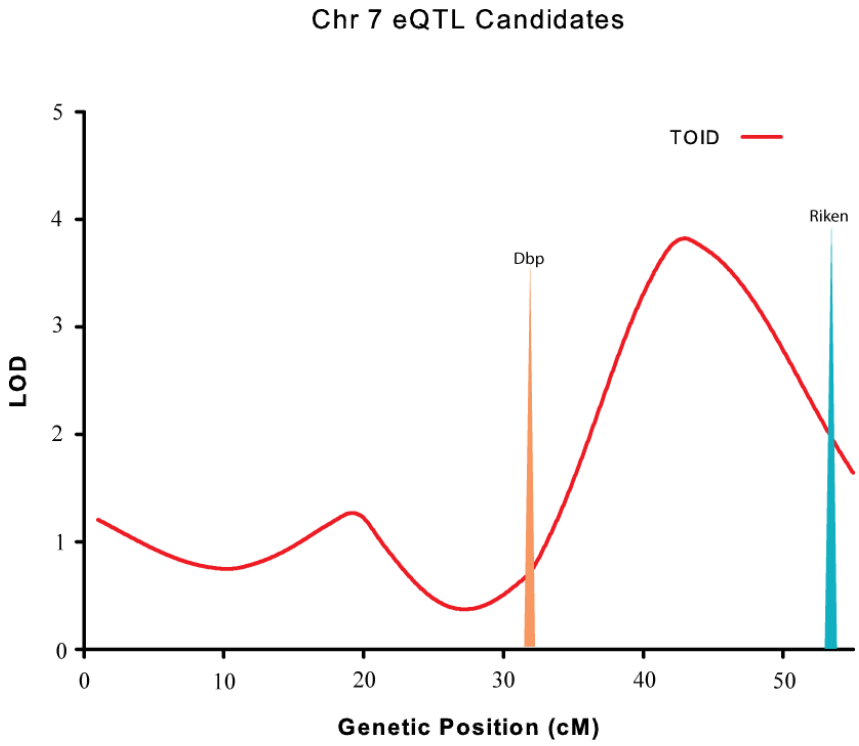


Figure 5.3: Metastatic QTL candidates: Metastatic QTL with colocalizing candidates. A: *Cis*-eQTL colocalizing with the metastatic locus on Chr 8. B: *Cis*-eQTL colocalizing with the metastatic locus on Chr 11. C: *Cis*-eQTL colocalizing with the metastatic locus on Chr 13. D: *Cis*-eQTL colocalizing with the metastatic locus on Chr 19.

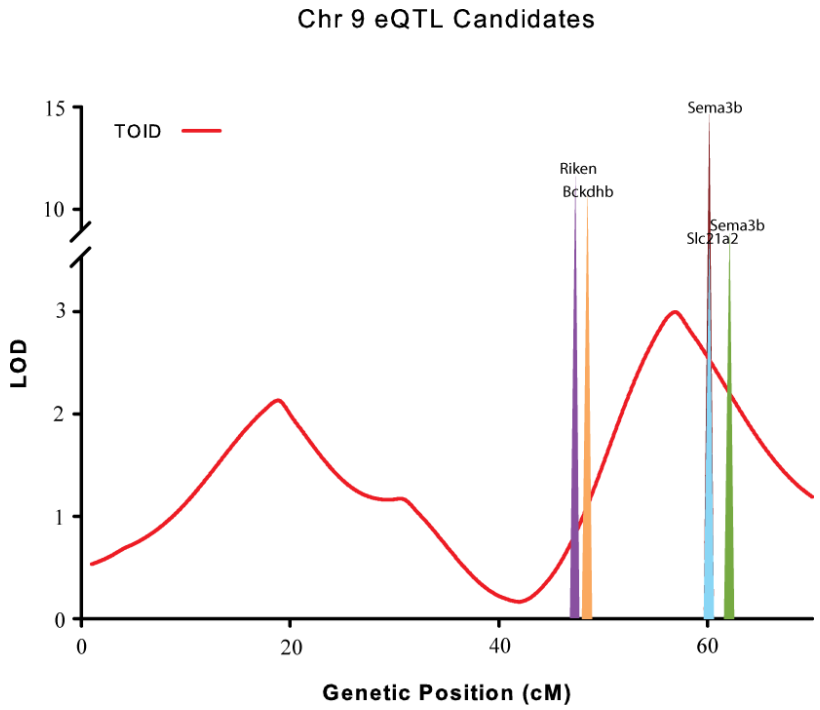
A. *cis*-eQTL colocalizing with the Inguinal tumor growth locus on Chr 5



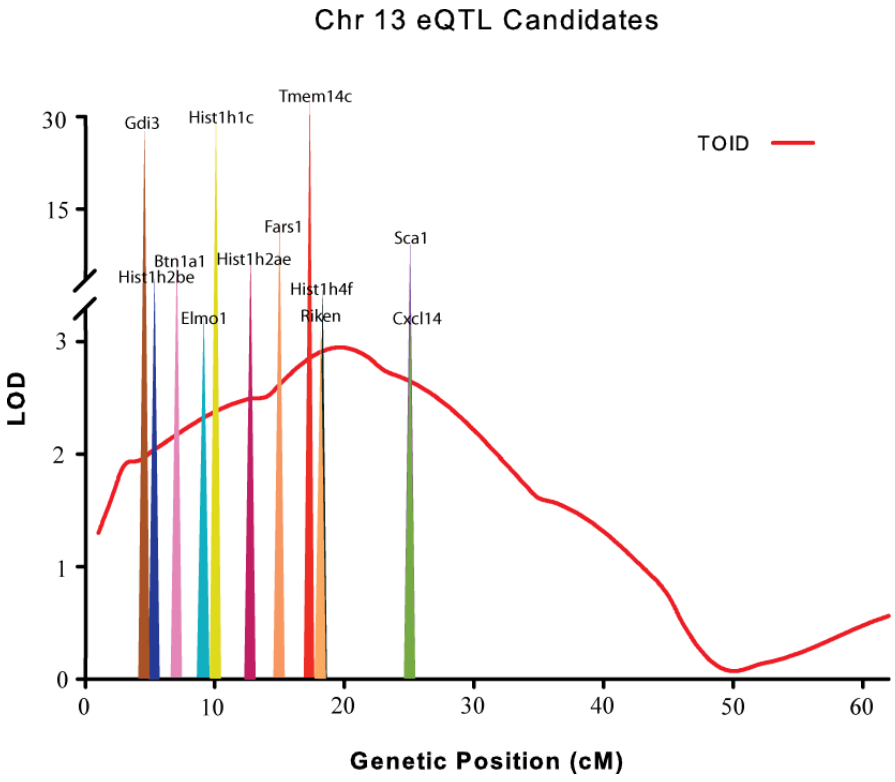
B. *cis*-eQTL colocalizing with the Tumor Latency locus on Chr 7



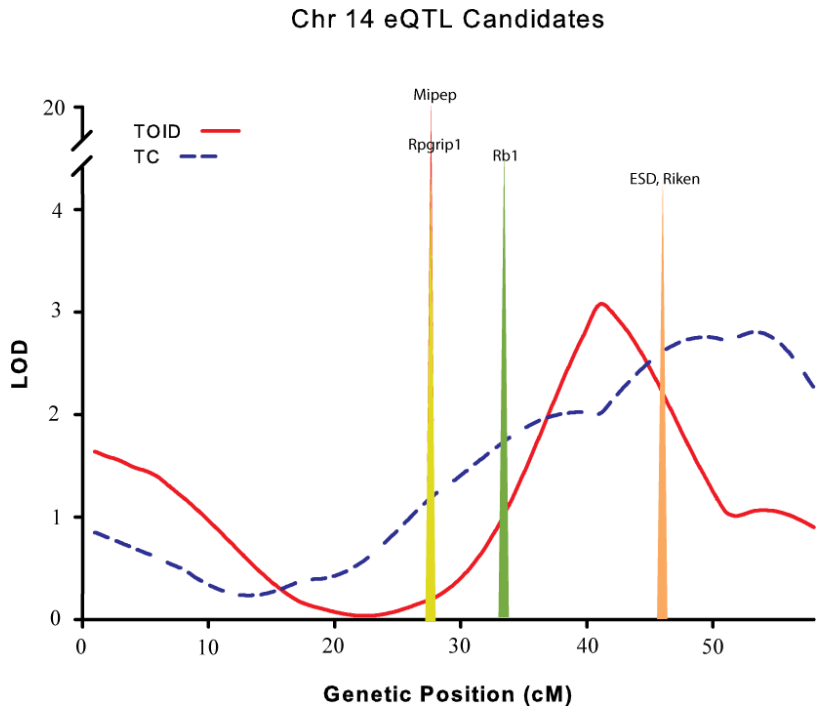
C. *cis*-eQTL colocalizing with the Tumor Latency locus on Chr 9



D. *cis*-eQTL colocalizing with the Tumor Latency locus on Chr 13



E. *cis*-eQTL colocalizing with the Latency and # of Tumors at Sacrifice loci on Chr 14



F. *cis*-eQTL colocalizing with the Axillary Tumor Growth locus on Chr 17

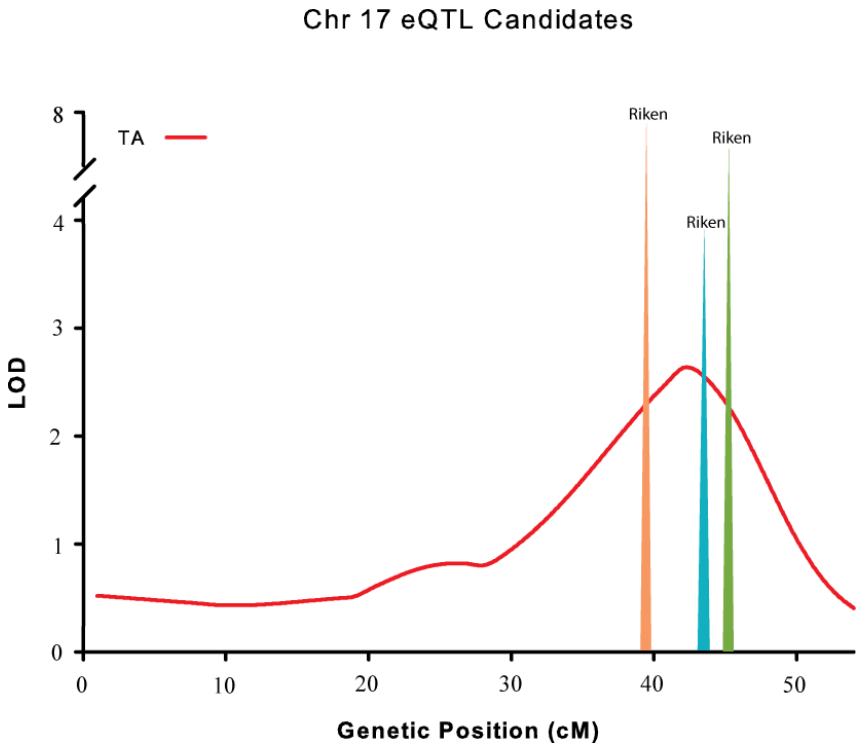


Figure 5.4: Primary Tumor QTL candidates: Primary Tumor QTL with colocalizing candidates. A: *Cis*-eQTL colocalizing with the inguinal tumor growth locus on Chr 5. B: *Cis*-eQTL colocalizing with the tumor latency locus on Chr 7. C: *Cis*-eQTL colocalizing with the tumor latency locus on Chr 9. D: *Cis*-eQTL colocalizing with the tumor latency locus on Chr 13. E: *Cis*-eQTL colocalizing with the tumor latency loci and total number of tumors at sacrifice loci on Chr 14. B: *Cis*-eQTL colocalizing with the axillary tumor growth locus on Chr 17.

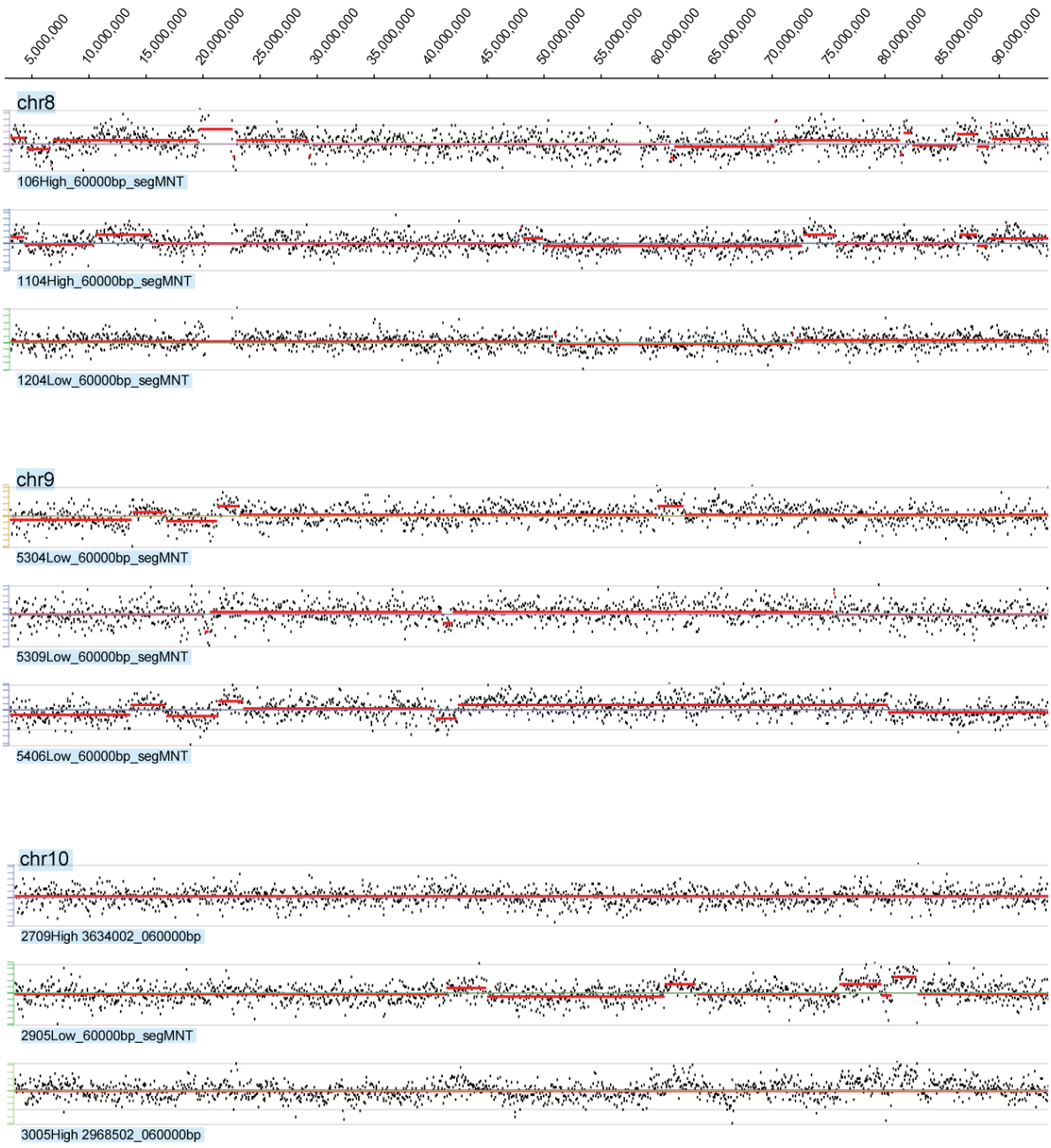
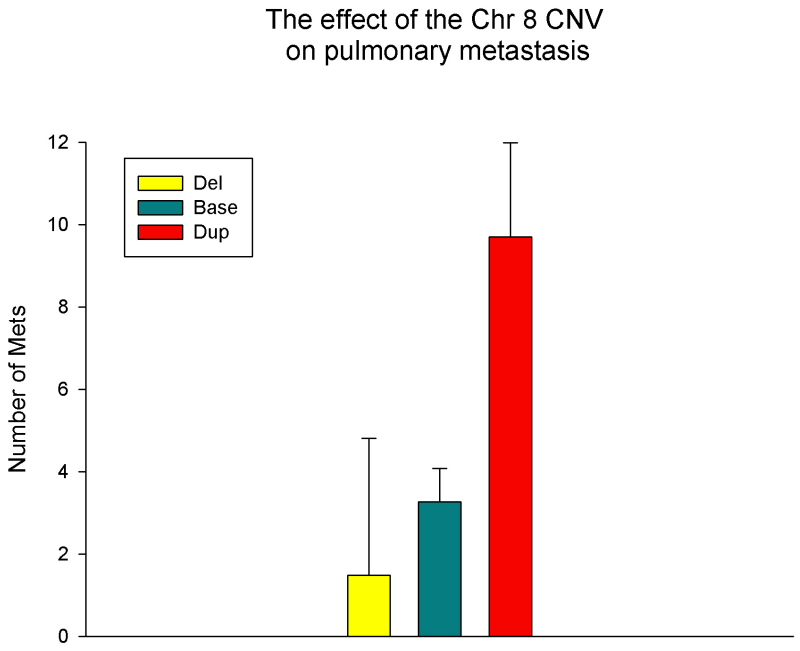
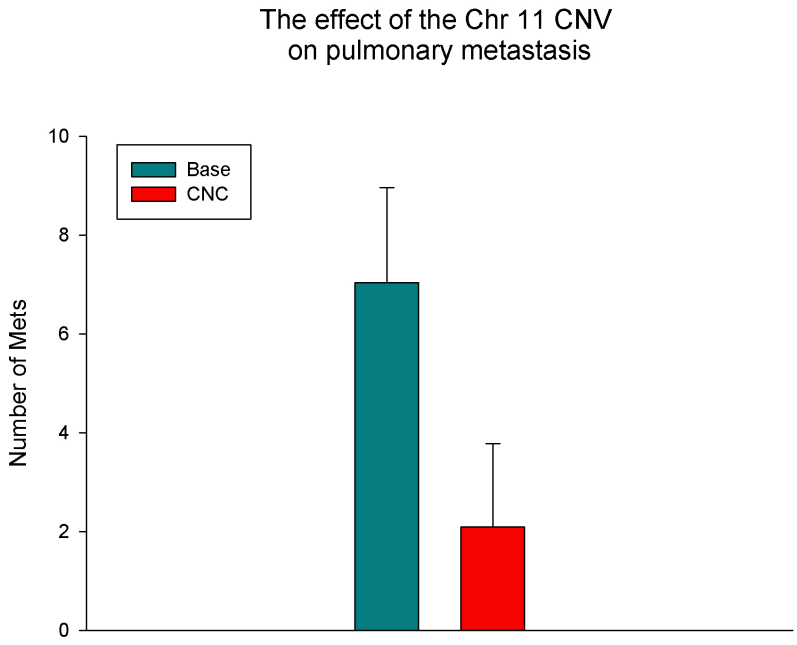

5,000,000
10,000,000
15,000,000
20,000,000
25,000,000
30,000,000
35,000,000
40,000,000
45,000,000
50,000,000
55,000,000
60,000,000
65,000,000
70,000,000
75,000,000
80,000,000
85,000,000
90,000,000
chr8
106High_60000bp_segMNT
1104High_60000bp_segMNT
1204Low_60000bp_segMNT
chr9
5304Low_60000bp_segMNT
5309Low_60000bp_segMNT
5406Low_60000bp_segMNT
chr10
2709High 3634002_060000bp
2905Low_60000bp_segMNT
3005High 2968502_060000bp

Figure 5.5: Examples of Copy Number Variation (CNV) in the F_2 Population: Examples of Copy Number Variation (CNV) found in the F_2 Population. A snapshot of the CNV found within the F_2 population. Examples from three separate chromosomes (8, 9 and 10) are shown.

A.


The effect of the Chr 8 CNV
on pulmonary metastasis
Number of Mets
0
2
4
6
8
10
12
Del
Base
Dup

B.


The effect of the Chr 11 CNV
on pulmonary metastasis
Number of Mets
0
2
4
6
8
10
Base
CNC

C.

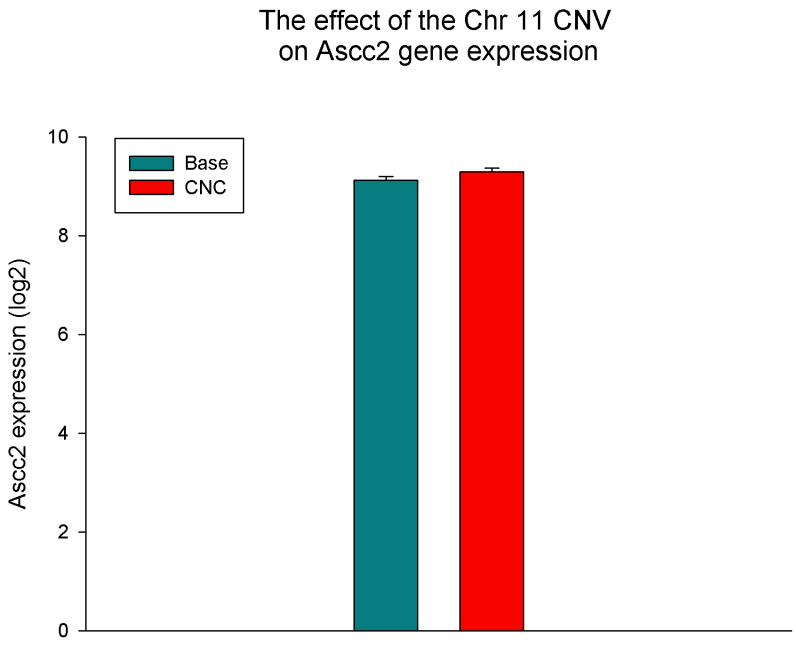


Figure 5.6: Evaluation of Copy Number Variation (CNV): Evaluation of Copy Number Variation (CNV). A: The impact of the Chr 8 CNV on the development of pulmonary metastasis. A duplication of this region was associated with a significant increase metastatic formation. B: The impact of the Chr 11 CNV on pulmonary metastasis development. The presence of an additional copy in this region was associated with a significant reduction in metastatic development. C: The impact of the Chr 11 CNV upon the expression of *Ascc2*. The presence of an additional copy in this region was associated with a significant increase in the expression of *Ascc2*.

CHAPTER VI

Synthesis

Introduction

Breast cancer is the most common cancer type, and the second-leading cause of cancer-related deaths in women (National-Cancer-Institute 2005). In 2008, over 180K women in the United States were diagnosed with some form of breast cancer and over 40K women died as a result of having this disease, typically from secondary metastatic malignancies (American-Cancer-Society 2009). Furthermore, it has been estimated that each year over $8B are spent on breast cancer treatment (National Cancer Institute 2007).

A complex relationship exists between diet and the genetic architecture of metastatic mammary cancer. Genetic predisposition to breast cancer has been confirmed by identification of multiple, small to low effect familial risk factors, but the evidence linking dietary fat to breast cancer susceptibility is less consistent. Whereas some human studies have shown a link between increased breast cancer and increased fat intake (Cho et al. 2003; Lee et al. 2005; Thiebaut et al. 2007), others have failed to replicate these findings (Kim et al. 2006; Wakai et al. 2005). Given this lack of congruity and the large number of individuals afflicted with this disease, a need exists to clarify further the linkage between diet and cancer. Therefore, we tested the hypothesis that mice predisposed to mammary tumor development and challenged with either high or normal levels of dietary fat will experience variability in

the pathogenesis of mammary cancer as a result of variations in genetic predisposition and gene expression. A summary of our experiments and a synthesis of our findings follows.

Population Development and Phenotype Characterization

An F_2 population (n = 615) was generated by crossing M16i, a polygenic obesity line (Allan et al. 2005), and FVB/NJ-TgN(MMTV-PyMT)634Mul (PyMT), a line transgenic for the Polyoma Middle T Oncoprotein, leading to the development of mammary tumors and subsequent pulmonary metastasis (Guy et al. 1992b). F_2 mice were weaned at 3 weeks of age and randomly assigned, within litter, sex, and genotype (PyMT or no PyMT), to receive one of two synthetic purified diets at 4 weeks of age. Mice had ad libitum access to water and to their assigned feed, either a high-fat diet (Research Diets D12451) containing 45% of total calories from fat, 20% from protein, and 35% from carbohydrates or a matched-control-fat diet (Research Diets D12450B) containing 10% of total calories from fat, 20% from protein, and 70% from carbohydrates.

The primary objective of our studies was to investigate the genetics of metastatic mammary cancer and the link between dietary fat and this disease. However, our experimental design also provided the opportunity to evaluate the effect of high-fat diet upon body weight and body composition in our genetically diverse F_2 population. Body weight was measured for each mouse at 3, 6, and 9 wk of age and at sacrifice (~11 wk for females, ~14 wk for males) and as expected the consumption of a high-fat diet resulted in increased weight gain at all time points regardless of gender and PyMT status, except for weight of both genders at 3 wk of age. Body composition measurements were collected at 7 wk of age

and at sacrifice using dual-energy X-ray absorptiometry. Similar to the results for body weight the animals fed the high-fat diet had significantly higher total body depositions of adipose tissue compared to animals fed the control-fat diet.

In order to ascertain the impact of dietary fat on the metastatic cancer paradigm, each mouse was evaluated for age at mammary tumor onset, tumor progression, and pulmonary metastases development. Results demonstrated that female mice fed a high-fat diet experienced a shortened period of mammary cancer latency of approximately 7.15% compared to female mice fed the control-fat diet. While the total number of tumors detected at sacrifice did not differ between the two dietary treatments, the weights of the axillary and inguinal tumors were significantly heavier in animals fed the high-fat diets. Interestingly, the increased tumor mass detected in mice fed the high-fat diet was not the result of increased adipose in tumors, as the tumors per se in mice fed a high-fat diet actually had a lower percent fat content. This finding suggests that the tumor epithelial growth may be more aggressive as a result of consumption of a high-fat diet.

In addition to the evaluation of the primary tumors, the pulmonary metastatic burden was evaluated at the time of sacrifice. Within the group of female mice that developed pulmonary metastases, a significant increase in metastatic load was observed in mice fed the high-fat diet relative to mice fed the control-fat diet.

QTL Evaluations

To elucidate the genetic underpinnings controlling both the body composition and the metastatic mammary cancer traits, and evaluate a possible genetic basis for the dietary effects

on cancer phenotypes, a large percentage of the F_2 population was genotyped for 384 SNP markers. Genomic regions of interest for metastatic mammary cancer and body composition traits were localized by integrating SNP genotypes with the aforementioned phenotypes to identify quantitative trait loci (QTL). This analysis supported identification of 38 modifier loci for body composition and 25 modifier loci for mammary cancer and pulmonary metastasis, likely representing 9 and 13 unique loci after accounting for pleiotropy, respectively. While the majority of QTL detected were located in regions previously implicated in body composition or metastatic mammary cancer, a few of these loci represented novel discoveries, including a modifier of tumor onset on Chr 1 and a metastatic modifier on Chr 8.

To evaluate the extent of overlap between obesity and metastatic cancer loci, the QTL positions from both phenotypes were compared and pleiotropy testing was performed. Only a small number of the cancer modifiers detected were in similar locations to obesity QTL that were mapped in the same population, and these cancer modifiers do not appear to result from pleiotropic effects of obesity loci. At first this seemed surprising, given the recent evidence that links obesity and cancer in humans (Hursting et al. 2007). Although this is only a single cross with limited power, it is possible that the obesity-cancer link has stronger environmental underpinnings relative to a genetic correlation. Our analyses were focused more on the effects of dietary fat on cancer, and the interaction between dietary fat and cancer modifiers, rather than the obesity-cancer linkage.

The experimental design of this F_2 population supported testing the hypothesis that animals fed a high-fat diet would have more advanced cancer phenotypes as a result of underlying QTL x diet interactions. This experiment revealed that 64% of all cancer

modifiers detected had significant diet interactions, and of these interactions, nine were caused by modifier effects being detected only in animals fed the high-fat diet, while three were detected only in mice fed the control-fat diet. These results implicate interactions of diet and modifier genes as a mechanism through which diet may affect breast cancer and metastasis.

Gene Expression Analysis

To investigate the complex genetic architecture that underlies metastatic mammary cancer in greater depth, i.e.; factors that modify the mammary diet-cancer-metastasis linkage, expression profiles of axillary tumors were characterized for a subset of the F_2 population with the Illumina Mouse-6 whole genome sentrix arrays. The differential expression analysis revealed an interesting primary result; at an False Discovery Rate (FDR) of $p < 0.05$, very few genes were differentially expressed in mice on the two diets. However, as the significance threshold was relaxed, a substantial number of differentially expressed genes, were identified. This finding suggests that dietary fat may modulate the expression of many genes in this paradigm, but only to a small extent.

The relationship between genes expressed in the axillary tumors and the metastatic phenotype was investigated by comparing the expression profiles from mice with low and high metastatic tendencies. A total 690 genes was differentially expressed at an FDR of $p < 0.05$. Entering the differentially expressed genes into the Ingenuity Pathway Analysis (IPA) revealed that the functions of many of these genes were involved in the processes of proliferation, inflammation, angiogenesis, and invasion. The multiple functions identified by

IPA provide a potential link between the genes identified by differential expression and metastatic mammary cancer.

eQTL Evaluation

Following differential expression analyses, the normalized data were entered into a systems-based analysis pipeline developed in R, facilitating a genome-wide expression QTL eQTL analysis as well as network and pathway QTL analyses. The resulting eQTL were classified in one of two categories: "*cis*" if they mapped within 10 cM of the physical location of the gene they represent, and "trans" if they mapped elsewhere. Clusters containing 50 or more eQTL within a 5 cM interval were designated as potential master regulatory regions. This analysis yielded the detection of 220 cis-acting eQTL and 890 trans-acting eQTL. Additionally, two potential master regulatory regions were identified on Chr 3 and 19. While the *trans*-acting eQTL represent loci controlled by unknown regulators, *cis*-acting eQTL exhibit self-regulation (Alberts et al. 2005). As such, by comparing/overlaying these *cis*-acting eQTL with the phenotypic cancer QTL detected in this population, the capability exists to reveal potential unknown polygenes (or at least to provide candidates) that influence the metastatic mammary cancer phenotype. Utilizing this method enabled the detection of 76 potential candidates for the previously detected metastatic QTL.

To further refine the candidate gene list of metastatic mammary cancer modifiers the relationship between the *cis*-eQTL and the observed metastatic loci was evaluated using causality analysis. An eQTL causally associated with a metastatic locus may represent the actual underlying genetic modifier of the phenotype of interest. The results revealed that very few significant relationships existed between *cis*-eQTL and metastatic virulence QTL.

However, one weak causal relationship bordering on the threshold of significance was detected between *H2afv* and the metastatic modifier on chromosome 11.

When previously evaluating this population for QTL, we found that the majority of cancer related loci exhibited interactions with diet. As such it was possible that the *cis*-acting eQTL would likewise interact with dietary fat. A *cis*-acting eQTL interacting with diet and colocalizing with a metastatic QTL exhibiting a similar interaction may represent a likely candidate for that particular locus. Our results however, demonstrated that very few of the *cis*-eQTL had diet interactions.

One caveat that needs to be considered when performing an eQTL study is the degree of copy number variation (CNV) within the mapping population. Taking CNV into account is especially important when analyzing tumor tissue, given the substantial amount of evidence linking the accumulation of CNV to cancer pathogenesis (Fridlyand et al. 2006; Reis-Filho et al. 2005; van Beers and Nederlof 2006). Additionally, it remains unclear how duplications or deletions of chromosomal segments that can result in altered expression of genes residing within those boundaries could influence the detection of eQTL. In our population all but one of the *cis*-eQTL acted independently of CNV. *Ascc2*, a gene physically located on Chr 11, was shown to have its gene expression significantly modified by a CNV. While we detected very little connection between *cis*-eQTL and CNV in our population, the link identified between the CNV on Chr 11 and *Ascc2*, highlights the importance of evaluating CNV in transcriptional studies in order properly interpret the results.

The evaluation of the publically available databases such as Oncomine provided support for several of the candidates being involved in the development of metastatic

mammary cancer. For example, our analyses yielded results that appear to support the link between *H2afv* and metastatic mammary cancer. Not only did we implicate *H2afv* as a candidate for the metastatic QTL on Chr 11 through causality analysis, the Oncomine evaluation supported a link between increased expression of this gene and a poorer clinical prognosis in humans. Another candidate, *Dusp4* was implicated by Oncomine as a modifier of metastatic mammary cancer in humans as well. *Dusp4*, a member of the dual-specificity phosphatase family, and a potential tumor suppressor, may be a strong candidate for the metastatic modifier we identified on Chr 8.

In addition to identifying a prioritized list of candidates for the metastatic modifiers detected through QTL analysis, the evaluations in Chapter V provided an interesting look at the master regulator regions on Chr 3 and 19. On Chr 3 98 *trans*-acting eQTL clustered within a 5 cM region at the proximal portion of the chromosome. Causality analysis revealed that a *cis*-eQTL detected for *Pkia* formed causal associations with 95% of the *trans*-eQTL located within this putative master regulator interval. This association between the *trans*-eQTL and *Pkia* was further supported by the IPA evaluation of gene functions. In the other master regulatory region on Chr 19, 60 *trans*-acting eQTL clustered in the 5 cM window which started at 13 cM and went through 17 cM. Using analyses similar to those that identified *Pika* as a potential candidate for the master regulator on Chr 3, we identified *Pkd2l1* as a potential candidate for the master regulator on Chr 19.

Future Directions

Several investigative approaches may be utilized in future experiments to gain a better understanding of the genetic architecture of metastatic breast cancer. These experiments should build upon the findings described in this dissertation.

Validation of Candidates: As described in the previous section, we were able to indentify several candidate genes that can potentially alter the pathogenesis of metastatic mammary cancer in mice. A logical direction for future experiments would be to explore further the role that these genes play in the processes of metastatic mammary cancer. In pursuit of this validation, we attempted to ligate several of the top priority genes into the pEF6/V5-His TOPO® TA mammalian expression vector (Invitrogen, Carlsbad, CA) for transfection into the Mvt1 cell line (Pei et al. 2004). Whereas we were able to develop vectors that over-expressed particular candidates of interest, the transfection into the Mvt1 cell line proved problematic. In culture, the Mvt1 cells were unable to survive with the increased expression of the candidate genes, and we were not able to characterize the effect of over-expressing these genes *in-vivo*.

While the over-expression of a candidate gene has the potential to be a viable technique for the validation of some genes, alternate approaches might be pursued as well. RNA interference (RNAi) could be used to induce the knockdown of candidate gene expression (Fire et al. 1998). A form of RNAi utilizing short hairpin RNA (shRNA) cassettes in a lentiviral-based vector has been shown to be effective in cell specific delivery, resulting

in permanent reduction of a targeted gene (Czauderna et al. 2003). This method could be potentially used to inhibit the expression of candidate genes in the Mvt1 cell line.

If the expression of a particular gene could be altered in the Mvt1 without compromising the cells, then the Mvt1 cells could be subcutaneously injected into the mammary gland of 6-week old virgin FVB/NJ female mice. Similar to a protocol previously described (Park et al. 2005), these mice would be allowed to age for four weeks, at which point they would be sacrificed by anesthetic overdose. Tumors would then be dissected and weighed. The lungs could then be removed and evaluated under a dissecting microscope to determine the number of pulmonary metastases on the surface. Both tumor weights and number of metastases would be compared to mice injected with Mvt1 cells transfected with a control vector. This comparison would provide the opportunity to physically evaluate the impact that altered-expression of a gene of interest has had upon both the tumor's growth and its metastatic potential.

Investigation of Master Regulators: Similar methods as described above could be used to further investigate the master regulator regions identified on Chr 3 and 19. *Pika* and *Pkd2l1* could be over-expressed or inhibited in a mammary cancer cell line via expression vectors or RNAi, respectively. If stable expression or knockdown of *Pika* and *Pkd2l1* were achieved, then it would be possible to evaluate with RT PCR the impact of the altered expression of these potential master regulators upon the expression of the genes residing in the master regulatory interval.

Alternate Tissue Evaluation: Another possible experiment that might provide additional insight into the processes of pulmonary metastatic mammary cancer would be to evaluate the expression profiles of the lungs collected from mice that developed pulmonary metastases. These lungs have previously been collected and snap-frozen in liquid nitrogen. The RNA from these lungs could be extracted and transcriptional profiles generated with the Illumina mouse-6-arrays. The generation of transcriptional profiles of the lungs would allow for several analyses. Evaluation of whether the lungs and the tumors had common expression patterns would be possible; genes differentially expressed between the two tissue types would also be analyzed. In addition to the comparisons between the expression patterns of the two tissue types, the correlation between lung gene expression and the development of pulmonary metastases could be measured as well. This analysis would help determine if any gene expressed in the lungs had a significant impact on the formation of metastatic breast cancer cells.

Furthermore, an eQTL analysis similar to one presented in chapter 5 could be carried out with the expression profiles of this tissue. An eQTL analysis of the lungs could help us identify additional candidates for the metastatic modifiers we identified in chapter 3. Several *cis*-eQTL colocalizing with the metastatic loci unique to the lungs could be identified, and this evaluation might detect *cis*-eQTL in common between the two datasets as well. If a *cis*-eQTL colocalizing with metastatic loci were detected in both the lung and tumor evaluations, then it would potentially be a high priority candidate for validation.

Alternate Time-point Collection: Evaluation of the tumor gene expression at different stages of tumor development could provide additional insight into the connection between cancer pathogenesis and dietary-fat. In the mouse population described in this dissertation, tumors were collected from the female mice at ~11 weeks of age, several weeks after tumors were first detected. When the tumors were evaluated at this time-point very few genes were differentially expressed between the two dietary treatments. Diet-induced gene changes may be programmed into the tumor at the beginning of tumor development, in which case the dietary effect on the axillary tumor could remain undetected at 11 weeks. To evaluate the expression of mammary tumors at several time-points, tumors could be surgically removed at different stages of their growth for genetic analysis. This time-dependent approach would provide the opportunity to assess the expression difference of tumors from the same individual as metastatic cancer progresses. Additionally, the impact of dietary fat upon the expression of genes at these different stages could then also be evaluated.

Overall Summary

In conclusion, this dissertation represents an innovative approach to understanding the paradigm of metastatic mammary cancer by evaluating the growth and metastatic potential of mammary tumors under the influence of genetic and environmental pressures, specifically modifier loci and dietary fat, respectively. In utilizing an intercross of mice predisposed to both mammary cancer and obesity, fed diets varying in fat percentage, we had the opportunity to critically evaluate the genetic and genomic underpinnings of metastatic mammary cancer in a controlled setting. This approach yielded several important outcomes.

First, we generated a genome-wide map of QTL and QTL by diet interactions for metastatic mammary cancer in a mammalian intercross. Second, we thoroughly evaluated the transcriptional networks for metastatic mammary cancer in this population and in the presence of two levels of dietary fat. The impact of somatic chromosomal aberrations upon gene expression was also explored. Finally, we prioritized genes harboring genetic variation that may explain deviation in mammary cancer development, severity and metastatic potential. These candidate genes can now be investigated using both *in vivo* and *in vitro* techniques to determine their impact upon mammalian metastatic mammary cancer. Understanding the genetic architecture and environmental interactions of this disease are critical given the widespread physical and financial impact that breast cancer has in our society. Ultimately, our data may provide additional insight into the intricate cascade of metastatic mammary cancer, and our results can be potentially applied in a clinical setting. This advance could provide both a predictive analysis of cancer pathogenesis and possible targets for therapeutic interventions.

References:

Adam L, Vadlamudi RK, McCrea P, Kumar R (2001) Tiam1 overexpression potentiates heregulin-induced lymphoid enhancer factor-1/beta -catenin nuclear signaling in breast cancer cells by modulating the intercellular stability. J Biol Chem 276, 28443-28450

Agnoli C, Berrino F, Abagnato CA, Muti P, Panico S, Crosignani P, Krogh V (2009) Metabolic syndrome and postmenopausal breast cancer in the ORDET cohort: A nested case-control study. Nutr Metab Cardiovasc Dis

Alberts R, Terpstra P, Bystrykh LV, de Haan G, Jansen RC (2005) A statistical multiprobe model for analyzing cis and trans genes in genetical genomics experiments with short-oligonucleotide arrays. Genetics 171, 1437-1439

Albertson DG, Collins C, McCormick F, Gray JW (2003) Chromosome aberrations in solid tumors. Nat Genet 34, 369-376

Allan MF, Eisen EJ, Pomp D (2004) The M16 mouse: an outbred animal model of early onset polygenic obesity and diabesity. Obesity research 12, 1397-1407

Allan MF, Eisen EJ, Pomp D (2005) Genomic mapping of direct and correlated responses to long-term selection for rapid growth rate in mice. Genetics 170, 1863-1877

American-Cancer-Society (2009) Detailed Guide: Breast Cancer.

American-Obesity-Association (2005) AOA fact sheets.

Anisimov VN, Ukraintseva SV, Yashin AI (2005) Cancer in rodents: does it tell us about cancer in humans? Nat.Rev.Cancer. 5, 807-819

Balmain A (2002) Cancer: new-age tumour suppressors. Nature 417, 235-237

Balmain A, Gray J, Ponder B (2003) The genetics and genomics of cancer. Nature genetics 33 Suppl, 238-244

Balmain A, Harris CC (2000) Carcinogenesis in mouse and human cells: parallels and paradoxes. Carcinogenesis 21, 371-377

Baracos VE (2006) Cancer-associated cachexia and underlying biological mechanisms. Annual Review of Nutrition 26, 435-461

Beckmann JS, Estivill X, Antonarakis SE (2007) Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. Nature reviews.Genetics 8, 639-646

Berclaz G, Li S, Price KN, Coates AS, Castiglione-Gertsch M, Rudenstam CM, Holmberg SB, Lindtner J, Erien D, Collins J, Snyder R, Thurlimann B, Fey MF, Mendiola C, Werner ID, Simoncini E, Crivellari D, Gelber RD, Goldhirsch A (2004) Body mass index as a prognostic feature in operable breast cancer: the International Breast Cancer Study Group experience. Ann Oncol 15, 875-884

Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, Olson JA, Jr., Marks JR, Dressman HK, West M, Nevins JR (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439, 353-357

Bing C, Russell S, Becket E, Pope M, Tisdale MJ, Trayhurn P, Jenkins JR (2006) Adipose atrophy in cancer cachexia: morphologic and molecular analysis of adipose tissue in tumour-bearing mice. British journal of cancer 95, 1028-1037

Bogenrieder T, Herlyn M (2003) Axis of evil: molecular mechanisms of cancer metastasis. Oncogene 22, 6524-6536

Borugian MJ, Sheps SB, Kim-Sing C, Olivotto IA, Van Patten C, Dunn BP, Coldman AJ, Potter JD, Gallagher RP, Hislop TG (2003) Waist-to-hip ratio and breast cancer mortality. American Journal of Epidemiology 158, 963-968

Bouloumie A, Sengenes C, Portolan G, Galitzky J, Lafontan M (2001) Adipocyte produces matrix metalloproteinases 2 and 9: involvement in adipose differentiation. Diabetes 50, 2080-2086

Bower AL, Lang DH, Vogler GP, Vandenbergh DJ, Blizard DA, Stout JT, McClearn GE, Sharkey NA (2006) QTL analysis of trabecular bone in BXD F2 and RI mice. J Bone Miner Res 21, 1267-1275

Boyd DB (2003) Insulin and cancer. Integr.Cancer.Ther. 2, 315-329

Bray GA (2002) The underlying basis for obesity: relationship to cancer. The Journal of nutrition 132, 3451S-3455S

Brem RB, Kruglyak L (2005) The landscape of genetic complexity across 5,700 gene expression traits in yeast. Proceedings of the National Academy of Sciences of the United States of America 102, 1572-1577

Cade JE, Burley VJ, Greenwood DC (2007) Dietary fibre and risk of breast cancer in the UK Women's Cohort Study. International journal of epidemiology

Callahan R, Smith GH (2000) MMTV-induced mammary tumorigenesis: gene discovery, progression to malignancy and cellular pathways. Oncogene 19, 992-1001

Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ (2003) Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. The New England journal of medicine 348, 1625-1638

Calle EE, Thun MJ (2004) Obesity and cancer. Oncogene 23, 6365-6378

Canning C, O'Brien M, Hegarty J, O'Farrelly C (2006) Liver immunity and tumour surveillance. Immunology letters 107, 83-88

Carmeliet P, Jain RK (2000) Angiogenesis in cancer and other diseases. Nature 407, 249-257

Carpenter CL, Ross RK, Paganini-Hill A, Bernstein L (2003) Effect of family history, obesity and exercise on breast cancer risk among postmenopausal women. Int J Cancer 106, 96-102

Cavanna T, Pokorna E, Vesely P, Gray C, Zicha D (2007) Evidence for protein 4.1B acting as a metastasis suppressor. J Cell Sci 120, 606-616

Cervino AC, Li G, Edwards S, Zhu J, Laurie C, Tokiwa G, Lum PY, Wang S, Castellini LW, Lusis AJ, Carlson S, Sachs AB, Schadt EE (2005) Integrating QTL and high-density SNP analyses in mice to identify Insig2 as a susceptibility gene for plasma cholesterol levels. Genomics 86, 505-517

Chambers AF, Groom AC, MacDonald IC (2002) Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2, 563-572

Chesler EJ, Lu L, Shou S, Qu Y, Gu J, Wang J, Hsu HC, Mountz JD, Baldwin NE, Langston MA, Threadgill DW, Manly KF, Williams RW (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. Nature genetics 37, 233-242

Cheverud JM, Ehrich TH, Hrbek T, Kenney JP, Pletscher LS, Semenkovich CF (2004) Quantitative trait loci for obesity- and diabetes-related traits and their dietary responses to high-fat feeding in LGXSM recombinant inbred mouse strains. Diabetes 53, 3328-3336

Chitale D, Gong Y, Taylor BS, Broderick S, Brennan C, Somwar R, Golas B, Wang L, Motoi N, Szoke J, Reinersman JM, Major J, Sander C, Seshan VE, Zakowski MF, Rusch V, Pao W, Gerald W, Ladanyi M (2009) An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. Oncogene 28, 2773-2783

Cho E, Spiegelman D, Hunter DJ, Chen WY, Stampfer MJ, Colditz GA, Willett WC (2003) Premenopausal fat intake and risk of breast cancer. Journal of the National Cancer Institute 95, 1079-1085

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138, 963-971

Cleary MP, Grande JP, Maihle NJ (2004) Effect of high fat diet on body weight and mammary tumor latency in MMTV-TGF-alpha mice. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity 28, 956-962

Comuzzie AG, Allison DB (1998) The search for human obesity genes. Science 280, 1374-1377

Connelly L, Robinson-Benion C, Chont M, Saint-Jean L, Li H, Polosukhin VV, Blackwell TS, Yull FE (2007) A transgenic model reveals important roles for the NF-kappaB alternative pathway (p100/p52) in mammary development and links to tumorigenesis. The Journal of biological chemistry 282, 10028-10035

Coulter AA, Bearden CM, Liu X, Koza RA, Kozak LP (2003) Dietary fat interacts with QTLs controlling induction of Pgc-1 alpha and Ucp1 during conversion of white to brown fat. Physiological genomics 14, 139-147

Cozma D, Lukes L, Rouse J, Qiu TH, Liu ET, Hunter KW (2002) A bioinformatics-based strategy identifies c-Myc and Cdc25A as candidates for the Apmt mammary tumor latency modifiers. Genome research 12, 969-975

Crawford NP, Qian X, Ziogas A, Papageorge AG, Boersma BJ, Walker RC, Lukes L, Rowe WL, Zhang J, Ambs S, Lowy DR, Anton-Culver H, Hunter KW (2007) Rrp1b, a new candidate susceptibility gene for breast cancer progression and metastasis. PLoS Genet 3, e214

Crawford NP, Walker RC, Lukes L, Officewala JS, Williams RW, Hunter KW (2008) The Diasporin Pathway: a tumor progression-related transcriptional network that predicts breast cancer survival. Clin Exp Metastasis 25, 357-369

Crawford NP, Ziogas A, Peel DJ, Hess J, Anton-Culver H, Hunter KW (2006) Germline polymorphisms in SIPA1 are associated with metastasis and other indicators of poor prognosis in breast cancer. Breast cancer research : BCR 8, R16

Cui Y, Whiteman MK, Flaws JA, Langenberg P, Tkaczuk KH, Bush TL (2002) Body mass and stage of breast cancer at diagnosis. International journal of cancer. Journal international du cancer 98, 279-283

Czauderna F, Santel A, Hinz M, Fechtner M, Durieux B, Fisch G, Leenders F, Arnold W, Giese K, Klippel A, Kaufmann J (2003) Inducible shRNA expression for application in a prostate cancer mouse model. Nucleic Acids Res 31, e127

Daniell HW (1988) Increased lymph node metastases at mastectomy for breast cancer associated with host obesity, cigarette smoking, age, and large tumor size. Cancer 62, 429-435

De los Santos M, Martinez-Iglesias O, Aranda A (2007) Anti-estrogenic actions of histone deacetylase inhibitors in MCF-7 breast cancer cells. Endocr Relat Cancer 14, 1021-1028

de Tayrac M, Etcheverry A, Aubry M, Saikali S, Hamlat A, Quillien V, Le Treut A, Galibert MD, Mosser J (2009) Integrative genome-wide analysis reveals a robust genomic glioblastoma signature associated with copy number driving changes in gene expression. Genes Chromosomes Cancer 48, 55-68

Demirkan B, Alacacioglu A, Yilmaz U (2007) Relation of body mass index (BMI) to disease free (DFS) and distant disease free survivals (DDFS) among Turkish women with operable breast carcinoma. Jpn J Clin Oncol 37, 256-265

den Broeder E, Oeseburg B, Lippens RJ, van Staveren WA, Sengers RC, van't Hof MA, Tolboom JJ (2001) Basal metabolic rate in children with a solid tumour. European journal of clinical nutrition 55, 673-681

Dewys WD, Begg C, Lavin PT, Band PR, Bennett JM, Bertino JR, Cohen MH, Douglass HO, Jr., Engstrom PF, Ezdinli EZ, Horton J, Johnson GJ, Moertel CG, Oken MM, Perlia C, Rosenbaum C, Silverstein MN, Skeel RT, Sponzo RW, Tormey DC (1980) Prognostic effect of weight loss prior to chemotherapy in cancer patients. Eastern Cooperative Oncology Group. The American Journal of Medicine 69, 491-497

Dong G, Chen Z, Kato T, Van Waes C (1999) The host environment promotes the constitutive activation of nuclear factor-kappaB and proinflammatory cytokine expression during metastatic tumor progression of murine squamous cell carcinoma. Cancer Res 59, 3495-3504

Douma S, Van Laar T, Zevenhoven J, Meuwissen R, Van Garderen E, Peeper DS (2004) Suppression of anoikis and induction of metastasis by the neurotrophic receptor TrkB. Nature 430, 1034-1039

Dragani TA (2003) 10 years of mouse cancer modifier loci: human relevance. Cancer Res 63, 3011-3018

Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. Bioinformatics 24, 1547-1548

Eccles SA, Box G, Court W, Sandle J, Dean CJ (1994) Preclinical models for the evaluation of targeted therapies of metastatic disease. Cell biophysics 24-25, 279-291

Eliassen AH, Missmer SA, Tworoger SS, Spiegelman D, Barbieri RL, Dowsett M, Hankinson SE (2006) Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. J Natl Cancer Inst 98, 1406-1415

Farber CR, Aten JE, Farber EA, de Vera V, Gularte R, Islas-Trejo A, Wen P, Horvath S, Lucero M, Lusis AJ, Medrano JF (2009) Genetic dissection of a major mouse obesity QTL (Carfhg2): integration of gene expression and causality modeling. Physiol Genomics 37, 294-302

Feigleson H, Jonas C, Teras L, Thun M, Calle E (2004) Weight gain, body mass index, hormone replacement therapy, and postmenopausal breast cancer in a large prospective study. Cancer Epidemiology, Biomarkers & Prevention 13, 220-224

Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL, Reinisch W, Teml A, Schwab M, Lichter P, Radlwimmer B, Stange EF (2006) A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. American Journal of Human Genetics 79, 439-448

Feuk L, Carson AR, Scherer SW (2006) Structural variation in the human genome. Nature reviews.Genetics 7, 85-97

Finkelstein EA, Fiebelkorn IC, Wang G (2004) State-level estimates of annual medical expenditures attributable to obesity. Obesity research 12, 18-24

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806-811

Flint J, Valdar W, Shifman S, Mott R (2005) Strategies for mapping and cloning quantitative trait genes in rodents. Nature reviews. Genetics 6, 271-286

Forman MR (2007) Changes in dietary fat and fiber and serum hormone concentrations: nutritional strategies for breast cancer prevention over the life course. The Journal of nutrition 137, 170S-174S

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 37, 391-400

Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN, McLennan J, Ziegler J, Chin K, Devries S, Feiler H, Gray JW, Waldman F, Pinkel D, Albertson DG (2006) Breast tumor copy number aberration phenotypes and genomic instability. BMC cancer 6, 96

Friedenreich CM, Courneya KS, Bryant HE (2002) Case-control study of anthropometric measures and breast cancer risk. International journal of cancer. Journal international du cancer 99, 445-452

Galli A, Svegliati-Baroni G, Ceni E, Milani S, Ridolfi F, Salzano R, Tarocchi M, Grappone C, Pellegrini G, Benedetti A, Surrenti C, Casini A (2005) Oxidative stress stimulates proliferation and invasiveness of hepatic stellate cells via a MMP2-mediated mechanism. Hepatology 41, 1074-1084

Geiger TR, Peeper DS (2005) The neurotrophic receptor TrkB in anoikis resistance and metastasis: a perspective. Cancer Res 65, 7033-7036

Ghazalpour A, Doss S, Sheth SS, Ingram-Drake LA, Schadt EE, Lusis AJ, Drake TA (2005) Genomic analysis of metabolic pathway gene expression in mice. Genome Biol 6, R59

Giamarchi A, Padilla F, Coste B, Raoux M, Crest M, Honore E, Delmas P (2006) The versatile nature of the calcium-permeable cation channel TRPP2. EMBO Rep 7, 787-793

Gingras AC, Kennedy SG, O'Leary MA, Sonenberg N, Hay N (1998) 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev 12, 502-513

Glinskii OV, Huxley VH, Glinsky GV, Pienta KJ, Raz A, Glinsky VV (2005) Mechanical entrapment is insufficient and intercellular adhesion is essential for metastatic cell arrest in distant organs. Neoplasia 7, 522-527

Gong H, Guo P, Zhai Y, Zhou J, Uppal H, Jarzynka MJ, Song WC, Cheng SY, Xie W (2007) Estrogen deprivation and inhibition of breast cancer growth in vivo through activation of the orphan nuclear receptor liver X receptor. Molecular endocrinology (Baltimore, Md 21, 1781-1790

Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'Connell R J, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK (2005) The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science 307, 1434-1440

Gordon JN, Green SR, Goggin PM (2005) Cancer cachexia. QJM : monthly journal of the Association of Physicians 98, 779-788

Gordon RR, Hunter KW, La Merrill M, Sorensen P, Threadgill DW, Pomp D (2008a) Genotype X diet interactions in mice predisposed to mammary cancer: II. Tumors and metastasis. Mamm Genome 19, 179-189

Gordon RR, Hunter KW, Sorensen P, Pomp D (2008b) Genotype X diet interactions in mice predisposed to mammary cancer. I. Body weight and fat. Mamm Genome 19, 163-178

Greenberg VL, Zimmer SG (2005) Paclitaxel induces the phosphorylation of the eukaryotic translation initiation factor 4E-binding protein 1 through a Cdk1-dependent mechanism. Oncogene 24, 4851-4860

Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR (2007) Patterns of somatic mutation in human cancer genomes. Nature 446, 153-158

Gupta GP, Nguyen DX, Chiang AC, Bos PD, Kim JY, Nadal C, Gomis RR, Manova-Todorova K, Massague J (2007) Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. Nature 446, 765-770

Gusella J, MacDonald M (2002) No post-genetics era in human disease research. Nature reviews. Genetics 3, 72-79

Guy CT, Cardiff RD, Muller WJ (1992) Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Molecular and cellular biology 12, 954-961

Hankinson SE, Willett WC, Manson JE, Colditz GA, Hunter DJ, Spiegelman D, Barbieri RL, Speizer FE (1998) Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J Natl Cancer Inst 90, 1292-1299

Hanrahan JP, Eisen EJ, Lagates JE (1973) Effects of population size and selection intensity of short-term response to selection for postweaning gain in mice. Genetics 73, 513-530

Hennighausen L (2000) Mouse models for breast cancer. Breast Cancer Res 2, 2-7

Henrichsen CN, Vinckenbosch N, Zollner S, Chaignat E, Pradervand S, Schutz F, Ruedi M, Kaessmann H, Reymond A (2009) Segmental copy number variation shapes tissue transcriptomes. Nat Genet 41, 424-429

Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S, Backlund MG, Yin Y, Khramtsov AI, Bastein R, Quackenbush J, Glazer RI, Brown PH, Green JE, Kopelovich L, Furth PA, Palazzo JP, Olopade OI, Bernard PS, Churchill GA, Van Dyke T, Perou CM (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 8, R76

Hewett P, Popplewell A, Finney H, Murray JC (1999) Changes in microvessel endothelial cell gene expression in an in vitro human breast tumour endothelial cell model. Angiogenesis 3, 221-229

Hiratsuka S, Watanabe A, Sakurai Y, Akashi-Takamura S, Ishibashi S, Miyake K, Shibuya M, Akira S, Aburatani H, Maru Y (2008) The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase. Nat Cell Biol 10, 1349-1355

Honda H, Ohi Y, Umekita Y, Takasaki T, Kuriwaki K, Ohyabu I, Yoshioka T, Yoshida A, Taguchi S, Ninomiya K, Akiba S, Nomura S, Sagara Y, Yoshida H (1998) Obesity affects expression of progesterone receptors and node metastasis of mammary carcinomas in postmenopausal women without a family history. Pathology International 49, 198-202

Horvat S, Bunger L, Falconer VM, Mackay P, Law A, Bulfield G, Keightley PD (2000) Mapping of obesity QTLs in a cross between mouse lines divergently selected on fat content. Mammalian genome : official journal of the International Mammalian Genome Society 11, 2-7

Houzelstein D, Bullock SL, Lynch DE, Grigorieva EF, Wilson VA, Beddington RS (2002) Growth and early postimplantation defects in mice deficient for the bromodomain-containing protein Brd4. Molecular and cellular biology 22, 3794-3802

Howard EW, Leung SC, Yuen HF, Chua CW, Lee DT, Chan KW, Wang X, Wong YC (2008) Decreased adhesiveness, resistance to anoikis and suppression of GRP94 are integral to the survival of circulating tumor cells in prostate cancer. Clin Exp Metastasis 25, 497-508

Hunter KW, Broman KW, Voyer TL, Lukes L, Cozma D, Debies MT, Rouse J, Welch DR (2001) Predisposition to efficient mammary tumor metastatic progression is linked to the breast cancer metastasis suppressor gene Brms1. Cancer research 61, 8866-8872

Huo D, Adebamowo CA, Ogundiran TO, Akang EE, Campbell O, Adenipekun A, Cummings S, Fackenthal J, Ademuyiwa F, Ahsan H, Olopade OI (2008) Parity and breastfeeding are protective against breast cancer in Nigerian women. Br J Cancer 98, 992-996

Hursting SD, Nunez NP, Varticovski L, Vinson C (2007) The obesity-cancer link: lessons learned from a fatless mouse. Cancer Res 67, 2391-2393

Ikubo M, Wada T, Fukui K, Ishiki M, Ishihara H, Asano T, Tsuneki H, Sasaka T (2008) Impact of lipid phosphateases SHIP2 and PTEN on the time- and Akt isoform-specific amelioration of TNFalpha-induced insulin resistance in 3T3-L1 adipocytes. Am J Physiol Endocrinol Metab in press

Ishikawa A, Hatada S, Nagamine Y, Namikawa T (2005) Further mapping of quantitative trait loci for postnatal growth in an inter-sub-specific backcross of wild Mus musculus castaneus and C57BL/6J mice. Genetical Research 85, 127-137

Jansen RC, Nap JP (2001) Genetical genomics: the added value from segregation. Trends in genetics : TIG 17, 388-391

Jerez-Timaure NC, Kearney F, Simpson EB, Eisen EJ, Pomp D (2004) Characterization of QTL with major effects on fatness and growth on mouse chromosome 2. Obesity research 12, 1408-1420

Joseph J, Mudduluru G, Antony S, Vashistha S, Ajitkumar P, Somasundaram K (2004) Expression profiling of sodium butyrate (NaB)-treated cells: identification of regulation of genes related to cytokine signaling and cancer metastasis by NaB. Oncogene 23, 6304-6315

Jump DB, Clarke SD (1999) Regulation of gene expression by dietary fat. Annual Review of Nutrition 19, 63-90

Kaaks R, Rinaldi S, Key TJ, Berrino F, Peeters PH, Biessy C, Dossus L, Lukanova A, Bingham S, Khaw KT, Allen NE, Bueno-de-Mesquita HB, van Gils CH, Grobbee D, Boeing H, Lahmann PH, Nagel G, Chang-Claude J, Clavel-Chapelon F, Fournier A, Thiebaut A, Gonzalez CA, Quiros JR, Tormo MJ, Ardanaz E, Amiano P, Krogh V, Palli D, Panico S, Tumino R, Vineis P, Trichopoulou A, Kalapothaki V, Trichopoulos D, Ferrari P, Norat T, Saracci R, Riboli E (2005) Postmenopausal serum androgens, oestrogens and breast cancer risk: the European prospective investigation into cancer and nutrition. Endocrine-related cancer 12, 1071-1082

Kaklamani VG, Linos A, Kaklamani E, Markaki I, Koumantaki Y, Mantzoros CS (1999) Dietary fat and carbohydrates are independently associated with circulating insulin-like growth factor 1 and insulin-like growth factor-binding protein 3 concentrations in healthy adults. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 17, 3291-3298

Kennedy AR, Pissios P, Otu HH, Xue B, Asakura K, Furukawa N, Marino FE, Liu FF, Kahn BB, Libermann TA, Maratos-Flier E (2007) A High Fat, Ketogenic Diet, Induces a Unique Metabolic State in Mice. American Journal of Physiology. Endocrinology and Metabolism

Kerver JM, Yang EJ, Obayashi S, Bianchi L, Song WO (2006) Meal and snack patterns are associated with dietary intake of energy and nutrients in US adults. J Am Diet Assoc 106, 46-53

Key TJ, Appleby PN, Reeves GK, Roddam A, Dorgan JF, Longcope C, Stanczyk FZ, Stephenson HE, Jr., Falk RT, Miller R, Schatzkin A, Allen DS, Fentiman IS, Wang DY, Dowsett M, Thomas HV, Hankinson SE, Toniolo P, Akhmedkhanov A, Koenig K, Shore RE, Zeleniuch-Jacquotte A, Berrino F, Muti P, Micheli A, Krogh V, Sieri S, Pala V, Venturelli E, Secreto G, Barrett-Connor E, Laughlin GA, Kabuto M, Akiba S, Stevens RG, Neriishi K, Land CE, Cauley JA, Kuller LH, Cummings SR, Helzlsouer KJ, Alberg AJ, Bush TL, Comstock GW, Gordon GB, Miller SR, Endogenous Hormones Breast Cancer Collaborative G (2003) Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. Journal of the National Cancer Institute 95, 1218-1226

Key TJ, Schatzkin A, Willett WC, Allen NE, Spencer EA, Travis RC (2004) Diet, nutrition and the prevention of cancer. Public health nutrition 7, 187-200

Kho Y, Kim S, Yoon BS, Moon JH, Kim B, Kwak S, Woo J, Oh S, Hong K, Kim H, You S, Choi Y (2008) Induction of serum amyloid A genes is associated with growth and apoptosis of HC11 mammary epithelial cells. Biosci Biotechnol Biochem 72, 70-81

Kim EH, Willett WC, Colditz GA, Hankinson SE, Stampfer MJ, Hunter DJ, Rosner B, Holmes MD (2006) Dietary fat and risk of postmenopausal breast cancer in a 20-year follow-up. American Journal of Epidemiology 164, 990-997

King MC, Marks JH, Mandell JB, New York Breast Cancer Study G (2003) Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science 302, 643-646

Kocic G, Vlahovic P, Dordevic V, Bjelakovic G, Koracevic D, Savic V (1995) Effects of growth factors on the enzymes of purine metabolism in culture of regenerating rat liver cells. Arch Physiol Biochem 103, 715-719

Korstanje R, DiPetrillo K (2004) Unraveling the genetics of chronic kidney disease using animal models. Am J Physiol Renal Physiol 287, F347-352

Kroemer G, Pouyssegur J (2008) Tumor cell metabolism: cancer's Achilles' heel. Cancer Cell 13, 472-482

Kuhn K, Baker SC, Chudin E, Lieu MH, Oeser S, Bennett H, Rigault P, Barker D, McDaniel TK, Chee MS (2004) A novel, high-performance random array platform for quantitative gene expression profiling. Genome research 14, 2347-2356

Lahmann PH, Hoffmann K, Allen N, van Gils CH, Khaw KT, Tehard B, Berrino F, Tjonneland A, Bigaard J, Olsen A, Overvad K, Clavel-Chapelon F, Nagel G, Boeing H, Trichopoulos D, Economou G, Bellos G, Palli D, Tumino R, Panico S, Sacerdote C, Krogh V, Peeters PH, Bueno-de-Mesquita HB, Lund E, Ardanaz E, Amiano P, Pera G, Quiros JR, Martinez C, Tormo MJ, Wirfalt E, Berglund G, Hallmans G, Key TJ, Reeves G, Bingham S, Norat T, Biessy C, Kaaks R, Riboli E (2004) Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). International journal of cancer. Journal international du cancer 111, 762-771

Lam JB, Chow KH, Xu A, Lam KS, Liu J, Wong NS, Moon RT, Shepherd PR, Cooper GJ, Wang Y (2009) Adiponectin haploinsufficiency promotes mammary tumor development in MMTV-PyVT mice by modulation of phosphatase and tensin homolog activities. PLoS ONE 4, e4968

Lancaster M, Rouse J, Hunter KW (2005) Modifiers of mammary tumor progression and metastasis on mouse chromosomes 7, 9, and 17. Mammalian genome : official journal of the International Mammalian Genome Society 16, 120-126

Le Voyer T, Lu Z, Babb J, Lifsted T, Williams M, Hunter K (2000) An epistatic interaction controls the latency of a transgene-induced mammary tumor. Mammalian genome : official journal of the International Mammalian Genome Society 11, 883-889

Le Voyer T, Rouse J, Lu Z, Lifsted T, Williams M, Hunter KW (2001) Three loci modify growth of a transgene-induced mammary tumor: suppression of proliferation associated with decreased microvessel density. Genomics 74, 253-261

Lebreton CM, Visscher PM, Haley CS, Semikhodskii A, Quarrie SA (1998) A nonparametric bootstrap method for testing close linkage vs. pleiotropy of coincident quantitative trait loci. Genetics 150, 931-943

Lee HS, Park CB, Kim JM, Jang SA, Park IY, Kim MS, Cho JH, Kim SC (2008) Mechanism of anticancer activity of buforin IIb, a histone H2A-derived peptide. Cancer Lett 271, 47-55

Lee MM, Chang IY, Horng CF, Chang JS, Cheng SH, Huang A (2005) Breast cancer and dietary factors in Taiwanese women. Cancer causes & control : CCC 16, 929-937

Lew JQ, Freedman ND, Leitzmann MF, Brinton LA, Hoover RN, Hollenbeck AR, Schatzkin A, Park Y (2009) Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study. Am J Epidemiol 170, 308-317

Li H, Xie Z, Lin J, Song H, Wang Q, Wang K, Su M, Qiu Y, Zhao T, Song K, Wang X, Zhou M, Liu P, Zhao G, Zhang Q, Jia W (2008) Transcriptomic and Metabonomic Profiling of Obesity-Prone and Obesity-Resistant Rats under High Fat Diet. J Proteome Res 7, 4775-4783

Liao D, Corle C, Seagroves TN, Johnson RS (2007) Hypoxia-inducible factor-1alpha is a key regulator of metastasis in a transgenic model of cancer initiation and progression. Cancer research 67, 563-572

Lifsted T, Le Voyer T, Williams M, Muller W, Klein-Szanto A, Buetow KH, Hunter KW (1998) Identification of inbred mouse strains harboring genetic modifiers of mammary tumor age of onset and metastatic progression. International journal of cancer. Journal international du cancer 77, 640-644

Lin EY, Jones JG, Li P, Zhu L, Whitney KD, Muller WJ, Pollard JW (2003) Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. American Journal of Pathology 163, 2113-2126

Liu P, Vikis H, Lu Y, Wang D, You M (2007a) Large-scale in silico mapping of complex quantitative traits in inbred mice. PLoS ONE 2, e651

Liu Y, Tseng M, Perdreau SA, Rossi F, Antonescu C, Besmer P, Fletcher JA, Duensing S, Duensing A (2007b) Histone H2AX is a mediator of gastrointestinal stromal tumor cell apoptosis following treatment with imatinib mesylate. Cancer Res 67, 2685-2692

Lorincz AM, Sukumar S (2006) Molecular links between obesity and breast cancer. Endocr Relat Cancer 13, 279-292

Lukanova A, Lundin E, Micheli A, Arslan A, Ferrari P, Rinaldi S, Krogh V, Lenner P, Shore RE, Biessy C, Muti P, Riboli E, Koenig KL, Levitz M, Stattin P, Berrino F, Hallmans G, Kaaks R, Toniolo P, Zeleniuch-Jacquotte A (2004) Circulating levels of sex steroid hormones and risk of endometrial cancer in postmenopausal women. Int J Cancer 108, 425-432

Lukes L, Crawford NP, Walker R, Hunter KW (2009) The origins of breast cancer prognostic gene expression profiles. Cancer Res 69, 310-318

Lum H, Jaffe HA, Schulz IT, Masood A, RayChaudhury A, Green RD (1999) Expression of PKA inhibitor (PKI) gene abolishes cAMP-mediated protection to endothelial barrier dysfunction. Am J Physiol 277, C580-588

Lunt SJ, Chaudary N, Hill RP (2009) The tumor microenvironment and metastatic disease. Clin Exp Metastasis 26, 19-34

Manly KF, Cudmore RH, Jr., Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. Mammalian genome : official journal of the International Mammalian Genome Society 12, 930-932

Maxwell KN, Soccio RE, Duncan EM, Sehayek E, Breslow JL (2003) Novel putative SREBP and LXR target genes identified by microarray analysis in liver of cholesterol-fed mice. J Lipid Res 44, 2109-2119

McDaniel AH, Li X, Tordoff MG, Bachmanov AA, Reed DR (2006) A locus on mouse Chromosome 9 (Adip5) affects the relative weight of the gonadal but not retroperitoneal adipose depot. Mammalian genome : official journal of the International Mammalian Genome Society 17, 1078-1092

McManaman JL, Palmer CA, Wright RM, Neville MC (2002) Functional regulation of xanthine oxidoreductase expression and localization in the mouse mammary gland: evidence of a role in lipid secretion. J Physiol 545, 567-579

McTiernan A, Rajan KB, Tworoger SS, Irwin M, Bernstein L, Baumgartner R, Gilliland F, Stanczyk FZ, Yasui Y, Ballard-Barbash R (2003) Adiposity and sex hormones in postmenopausal breast cancer survivors. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 21, 1961-1966

McTiernan A, Wu L, Chen C, Chlebowski R, Mossavar-Rahmani Y, Modugno F, Perri MG, Stanczyk FZ, Van Horn L, Wang CY (2006) Relation of BMI and physical activity to sex hormones in postmenopausal women. Obesity (Silver Spring) 14, 1662-1677

Michaelson JS, Leder P (2001) beta-catenin is a downstream effector of Wnt-mediated tumorigenesis in the mammary gland. Oncogene 20, 5093-5099

Miles CM, Wayne M (2008) Quantitative trait locus (QTL) analysis. Nature Education 1

Miles FL, Pruitt FL, van Golen KL, Cooper CR (2008) Stepping out of the flow: capillary extravasation in cancer metastasis. Clin Exp Metastasis 25, 305-324

Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, Nuyten D, Kreike B, Zhang Y, Wang Y, Ishwaran H, Foekens JA, van de Vijver M, Massague J (2007) Lung metastasis genes couple breast tumor size and metastatic spread. Proc Natl Acad Sci U S A 104, 6740-6745

Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, Massague J (2005) Genes that mediate breast cancer metastasis to lung. Nature 436, 518-524

Mohammed RA, Green A, El-Shikh S, Paish EC, Ellis IO, Martin SG (2007) Prognostic significance of vascular endothelial cell growth factors -A, -C and -D in breast cancer and their relationship with angio- and lymphangiogenesis. Br J Cancer 96, 1092-1100

Moore KJ, Nagle DL (2000) Complex trait analysis in the mouse: The strengths, the limitations and the promise yet to come. Annu Rev Genet 34, 653-686

Morgan K, Uyuni A, Nandgiri G, Mao L, Castaneda L, Kathirvel E, French SW, Morgan TR (2008) Altered expression of transcription factors and genes regulating lipogenesis in liver and adipose tissue of mice with high fat diet-induced obesity and nonalcoholic fatty liver disease. Eur J Gastroenterol Hepatol 20, 843-854

Murphy PM (2001) Chemokines and the molecular basis of cancer metastasis. The New England journal of medicine 345, 833-835

Murtaugh MA, Sweeney C, Giuliano AR, Herrick JS, Hines L, Byers T, Baumgartner KB, Slattery ML (2008) Diet patterns and breast cancer risk in Hispanic and non-Hispanic white women: the Four-Corners Breast Cancer Study. Am J Clin Nutr 87, 978-984

Nam SW, Clair T, Campo CK, Lee HY, Liotta LA, Stracke ML (2000) Autotaxin (ATX), a potent tumor motogen, augments invasive and metastatic potential of ras-transformed cells. Oncogene 19, 241-247

Namba R, Maglione JE, Davis RR, Baron CA, Liu S, Carmack CE, Young LJ, Borowsky AD, Cardiff RD, Gregg JP (2006) Heterogeneity of mammary lesions represent molecular differences. BMC Cancer 6, 275

National-Cancer-Institute (2005) What you need to know about breast cancer.

Nehrenberg DL, Hua K, Estrada-Smith D, Garland T, Jr., Pomp D (2009) Voluntary exercise and its effects on body composition depend on genetic selection history. Obesity (Silver Spring) 17, 1402-1409

Nguyen DX, Massague J (2007) Genetic determinants of cancer metastasis. Nature reviews.Genetics 8, 341-352

Nilsson UW, Garvin S, Dabrosin C (2007) MMP-2 and MMP-9 activity is regulated by estradiol and tamoxifen in cultured human breast cancer cells. Breast Cancer Res Treat 102, 253-261

Nowell PC (1976) The clonal evolution of tumor cell populations. Science (New York, N.Y.) 194, 23-28

Oestreicher N, White E, Lehman CD, Mandelson MT, Porter PL, Taplin SH (2002) Predictors of sensitivity of clinical breast examination (CBE). Breast Cancer Res Treat 76, 73-81

Oshima RG, Lesperance J, Munoz V, Hebbard L, Ranscht B, Sharan N, Muller WJ, Hauser CA, Cardiff RD (2004) Angiogenic acceleration of Neu induced mammary tumor progression and metastasis. Cancer Res 64, 169-179

Park JW, Chun YS, Kim MS, Park YC, Kwak SJ, Park SC (1998) Metabolic modulation of cellular redox potential can improve cardiac recovery from ischemia-reperfusion injury. Int J Cardiol 65, 139-147

Park YG, Clifford R, Buetow KH, Hunter KW (2003) Multiple cross and inbred strain haplotype mapping of complex-trait candidate genes. Genome research 13, 118-121

Park YG, Zhao X, Lesueur F, Lowy DR, Lancaster M, Pharoah P, Qian X, Hunter KW (2005) Sipa1 is a candidate for underlying the metastasis efficiency modifier locus Mtes1. Nature genetics 37, 1055-1062

Pei XF, Noble MS, Davoli MA, Rosfjord E, Tilli MT, Furth PA, Russell R, Johnson MD, Dickson RB (2004) Explant-cell culture of primary mammary tumors from MMTV-c-Myc transgenic mice. In Vitro Cell Dev Biol Anim 40, 14-21

Petrelli JM, Calle EE, Rodriguez C, Thun MJ (2002) Body mass index, height, and postmenopausal breast cancer mortality in a prospective cohort of US women. Cancer causes & control : CCC 13, 325-332

Pomp D (1997) Genetic dissection of obesity in polygenic animal models. Behavior genetics 27, 285-306

Pomp D (2007) Natural Polygenic Models. In Obesity: Genomics and Postgenomics., S.T. Clement K, ed. (London: Informa Healthcare)

Pomp D, Murray, J.D. and Medrano, J.F. (1991) Single day detection of transgenic mice by PCR of toe-clips. Mouse Genome 89, 1

Pomp D, Nehrenberg D, Estrada-Smith D (2008) Complex genetics of obesity in mouse models. Annu Rev Nutr 28, 331-345

Qiu TH, Chandramouli GV, Hunter KW, Alkharouf NW, Green JE, Liu ET (2004) Global expression profiling identifies signatures of tumor virulence in MMTV-PyMT-transgenic mice: correlation to human disease. Cancer Res 64, 5973-5981

Quigley DA, To MD, Perez-Losada J, Pelorosso FG, Mao JH, Nagase H, Ginzinger DG, Balmain A (2009) Genetic architecture of mouse skin inflammation and tumour susceptibility. Nature 458, 505-508

Rajkumar L, Kittrell FS, Guzman RC, Brown PH, Nandi S, Medina D (2007) Hormone-induced protection of mammary tumorigenesis in genetically engineered mouse models. Breast cancer research : BCR 9, R12

Ramaswamy S, Ross KN, Lander ES, Golub TR (2003) A molecular signature of metastasis in primary solid tumors. Nature genetics 33, 49-54

Rankinen T, Zuberi A, Chagnon YC, Weisnagel SJ, Argyropoulos G, Walts B, Perusse L, Bouchard C (2006) The human obesity gene map: the 2005 update. Obesity (Silver Spring, Md.) 14, 529-644

Reeves KW, Faulkner K, Modugno F, Hillier TA, Bauer DC, Ensrud KE, Cauley JA (2007) Body mass index and mortality among older breast cancer survivors in the Study of Osteoporotic Fractures. Cancer Epidemiol Biomarkers Prev 16, 1468-1473

Rega G, Kaun C, Demyanets S, Pfaffenberger S, Rychli K, Hohensinner PJ, Kastl SP, Speidl WS, Weiss TW, Breuss JM, Furnkranz A, Uhrin P, Zaujec J, Zilberfarb V, Frey M, Roehle R, Maurer G, Huber K, Wojta J (2007) Vascular endothelial growth factor is induced by the inflammatory cytokines interleukin-6 and oncostatin m in human adipose tissue in vitro and in murine adipose tissue in vivo. Arterioscler Thromb Vasc Biol 27, 1587-1595

Reis-Filho JS, Simpson PT, Gale T, Lakhani SR (2005) The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. Pathology, research and practice 201, 713-725

Reymond A, Henrichsen CN, Harewood L, Merla G (2007) Side effects of genome structural changes. Current opinion in genetics & development 17, 381-386

Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 6, 1-6

Richardson RA, Davidson HI (2003) Nutritional demands in acute and chronic illness. The Proceedings of the Nutrition Society 62, 777-781

Rinaldo JE, Clark M, Parinello J, Shepherd VL (1994) Nitric oxide inactivates xanthine dehydrogenase and xanthine oxidase in interferon-gamma-stimulated macrophages. Am J Respir Cell Mol Biol 11, 625-630

Rocha JL, Eisen EJ, Van Vleck LD, Pomp D (2004a) A large-sample QTL study in mice: I. Growth. Mammalian genome : official journal of the International Mammalian Genome Society 15, 83-99

Rocha JL, Eisen EJ, Van Vleck LD, Pomp D (2004b) A large-sample QTL study in mice: II. Body composition. Mammalian genome : official journal of the International Mammalian Genome Society 15, 100-113

Rohan TE, Li SQ, Hartwick R, Kandel RA (2006) p53 Alterations and protein accumulation in benign breast tissue and breast cancer risk: a cohort study. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 15, 1316-1323

Rose DP, Connolly JM, Meschter CL (1991) Effect of dietary fat on human breast cancer growth and lung metastasis in nude mice. Journal of the National Cancer Institute 83, 1491-1495

Rose DP, Komninou D, Stephenson GD (2004) Obesity, adipocytokines, and insulin resistance in breast cancer. Obes Rev 5, 153-165

Rosner M, Hanneder M, Siegel N, Valli A, Fuchs C, Hengstschlager M (2008) The mTOR pathway and its role in human genetic diseases. Mutat Res 659, 284-292

Rossiter H, Barresi C, Ghannadan M, Gruber F, Mildner M, Fodinger D, Tschachler E (2007) Inactivation of VEGF in mammary gland epithelium severely compromises mammary gland development and function. FASEB J 21, 3994-4004

Russell ST, Tisdale MJ (2002) Effect of a tumour-derived lipid-mobilising factor on glucose and lipid metabolism in vivo. British journal of cancer 87, 580-584

Russell ST, Zimmerman TP, Domin BA, Tisdale MJ (2004) Induction of lipolysis in vitro and loss of body fat in vivo by zinc-alpha2-glycoprotein. Biochimica et biophysica acta 1636, 59-68

Ryden M, Arner P (2007) Fat loss in cachexia--is there a role for adipocyte lipolysis? Clinical nutrition (Edinburgh, Scotland) 26, 1-6

Salazar-Martinez E, Lazcano-Ponce E, Sanchez-Zamorano LM, Gonzalez-Lira G, Escudero-De Los Rios P, Hernandez-Avila M (2005) Dietary factors and endometrial cancer risk. Results of a case-control study in Mexico. International journal of gynecological cancer : official journal of the International Gynecological Cancer Society 15, 938-945

Santen RJ, Boyd NF, Chlebowski RT, Cummings S, Cuzick J, Dowsett M, Easton D, Forbes JF, Key T, Hankinson SE, Howell A, Ingle J (2007) Critical assessment of new risk factors for breast cancer: considerations for development of an improved risk prediction model. Endocrine-related cancer 14, 169-187

Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D, Sieberts SK, Monks S, Reitman M, Zhang C, Lum PY, Leonardson A, Thieringer R, Metzger JM, Yang L, Castle J, Zhu H, Kash SF, Drake TA, Sachs A, Lusis AJ (2005) An integrative genomics approach to infer causal associations between gene expression and disease. Nature genetics 37, 710-717

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH (2003a) Genetics of gene expression surveyed in maize, mouse and man. Nature 422, 297-302

Schadt EE, Monks SA, Friend SH (2003b) A new paradigm for drug discovery: integrating clinical, genetic, genomic and molecular phenotype data to identify drug targets. Biochemical Society transactions 31, 437-443

Schaffler A, Scholmerich J, Buechler C (2007) Mechanisms of disease: adipokines and breast cancer - endocrine and paracrine mechanisms that connect adiposity and breast cancer. Nature clinical practice.Endocrinology & metabolism 3, 345-354

Seaton G, Haley CS, Knott SA, Kearsey M, Visscher PM (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics (Oxford, England) 18, 339-340

Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, Kurdistani SK (2005) Global histone modification patterns predict risk of prostate cancer recurrence. Nature 435, 1262-1266

Senzaki H, Iwamoto S, Ogura E, Kiyozuka Y, Arita S, Kurebayashi J, Takada H, Hioki K, Tsubura A (1998) Dietary effects of fatty acids on growth and metastasis of KPL-1 human breast cancer cells in vivo and in vitro. Anticancer Research 18, 1621-1627

Shelling AN, Ferguson LR (2007) Genetic variation in human disease and a new role for copy number variants. Mutat Res 622, 33-41

Shenkar R, Schwartz MD, Terada LS, Repine JE, McCord J, Abraham E (1996) Hemorrhage activates NF-kappa B in murine lung mononuclear cells in vivo. Am J Physiol 270, L729-735

Shi R, Yu H, McLarty J, Glass J (2004) IGF-1 and breast cancer: a meta-analysis. International Journal of Cancer 111, 418-423

Shifman S, Bell JT, Copley RR, Taylor MS, Williams RW, Mott R, Flint J (2006) A high-resolution single nucleotide polymorphism genetic map of the mouse genome. PLoS biology 4, e395

Shlien A, Malkin D (2009) Copy number variations and cancer. Genome medicine 1, 62

Sieben NL, Oosting J, Flanagan AM, Prat J, Roemen GM, Kolkman-Uljee SM, van Eijk R, Cornelisse CJ, Fleuren GJ, van Engeland M (2005) Differential gene expression in ovarian tumors reveals Dusp 4

and Serpina 5 as key regulators for benign behavior of serous borderline tumors. J Clin Oncol 23, 7257-7264

Singer JB, Hill AE, Burrage LC, Olszens KR, Song J, Justice M, O'Brien WE, Conti DV, Witte JS, Lander ES, Nadeau JH (2004) Genetic dissection of complex traits with chromosome substitution strains of mice. Science 304, 445-448

Slate J (2005) Quantitative trait locus mapping in natural populations: progress, caveats and future directions. Molecular ecology 14, 363-379

Smith-Warner SA, Spiegelman D, Adami HO, Beeson WL, van den Brandt PA, Folsom AR, Fraser GE, Freudenheim JL, Goldbohm RA, Graham S, Kushi LH, Miller AB, Rohan TE, Speizer FE, Toniolo P, Willett WC, Wolk A, Zeleniuch-Jacquotte A, Hunter DJ (2001) Types of dietary fat and breast cancer: a pooled analysis of cohort studies. International journal of cancer.Journal international du cancer 92, 767-774

Song CG, Hu Z, Wu J, Luo JM, Shen ZZ, Huang W, Shao ZM (2006) The prevalence of BRCA1 and BRCA2 mutations in eastern Chinese women with breast cancer. Journal of cancer research and clinical oncology 132, 617-626

Sporn MB (1996) The war on cancer. Lancet 347, 1377-1381

Steeg PS (2006) Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12, 895-904

Stoll BA (1998) Western diet, early puberty, and breast cancer risk. Breast cancer research and treatment 49, 187-193

Storey JD (2002) A direct approach to false discovery rates. J Roy Stat Soc Ser B 64, 479-498

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurles ME, Dermitzakis ET (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science (New York, N.Y.) 315, 848-853

Sun W, Yu T, Li KC (2007) Detection of eQTL modules mediated by activity levels of transcription factors. Bioinformatics 23, 2290-2297

Surwit RS, Feinglos MN, Rodin J, Sutherland A, Petro AE, Opara EC, Kuhn CM, Rebuffe-Scrive M (1995) Differential effects of fat and sucrose on the development of obesity and diabetes in C57BL/6J and A/J mice. Metabolism: clinical and experimental 44, 645-651

Svenson KL, Von Smith R, Magnani PA, Suetin HR, Paigen B, Naggert JK, Li R, Churchill GA, Peters LL (2007) Multiple Trait Measurements in 43 Inbred Mouse Strains Captures the Phenotypic Diversity Characteristic of Human Populations. Journal of applied physiology: respiratory, environmental and exercise physiology

Tannenbaum A (1942) The Genesis and Growth of Tumors. III. Effects of a High-Fat Diet. Cancer research 2, 468--475

Taylor BA, Phillips SJ (1997) Obesity QTLs on mouse chromosomes 2 and 17. Genomics 43, 249-257

Taylor BA, Tarantino LM, Phillips SJ (1999) Gender-influenced obesity QTLs identified in a cross involving the KK type II diabetes-prone mouse strain. Mammalian genome : official journal of the International Mammalian Genome Society 10, 963-968

Taylor SS, Yang J, Wu J, Haste NM, Radzio-Andzelm E, Anand G (2004) PKA: a portrait of protein kinase dynamics. Biochim Biophys Acta 1697, 259-269

Terry MB, Zhang FF, Kabat G, Britton JA, Teitelbaum SL, Neugut AI, Gammon MD (2006) Lifetime alcohol intake and breast cancer risk. Annals of Epidemiology 16, 230-240

Thiebaut AC, Kipnis V, Chang SC, Subar AF, Thompson FE, Rosenberg PS, Hollenbeck AR, Leitzmann M, Schatzkin A (2007) Dietary fat and postmenopausal invasive breast cancer in the National Institutes of Health-AARP Diet and Health Study cohort. Journal of the National Cancer Institute 99, 451-462

Thomas T, Faaland CA, Adhikarakunnathu S, Thomas TJ (1996) Structure-activity relations of S-adenosylmethionine decarboxylase inhibitors on the growth of MCF-7 breast cancer cells. Breast Cancer Res Treat 39, 293-306

Thomsen LL, Miles DW (1998) Role of nitric oxide in tumour progression: lessons from human tumours. Cancer Metastasis Rev 17, 107-118

Threadgill DW (2005) Metastatic potential as a heritable trait. Nature genetics 37, 1026-1027

Tisdale MJ (1997) Biology of cachexia. Journal of the National Cancer Institute 89, 1763-1773

Tisdale MJ (2002) Cachexia in cancer patients. Nature reviews.Cancer 2, 862-871

Tsafrir D, Bacolod M, Selvanayagam Z, Tsafrir I, Shia J, Zeng Z, Liu H, Krier C, Stengel RF, Barany F, Gerald WL, Paty PB, Domany E, Notterman DA (2006) Relationship of gene expression and chromosomal abnormalities in colorectal cancer. Cancer Res 66, 2129-2137

Tu LC, Yan X, Hood L, Lin B (2007) Proteomics analysis of the interactome of N-myc downstream regulated gene 1 and its interactions with the androgen response program in prostate cancer cells. Mol Cell Proteomics 6, 575-588

Turnbull C, Rahman N (2008) Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet 9, 321-345

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98, 5116-5121

van Beers EH, Nederlof PM (2006) Array-CGH and breast cancer. Breast Cancer Res 8, 210

Varnum SM, Covington CC, Woodbury RL, Petritis K, Kangas LJ, Abdullah MS, Pounds JG, Smith RD, Zangar RC (2003) Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. Breast Cancer Res Treat 80, 87-97

Vona-Davis L, Rose DP (2009) Angiogenesis, adipokines and breast cancer. Cytokine Growth Factor Rev 20, 193-201

Wakai K, Tamakoshi K, Date C, Fukui M, Suzuki S, Lin Y, Niwa Y, Nishio K, Yatsuya H, Kondo T, Tokudome S, Yamamoto A, Toyoshima H, Tamakoshi A, Group JS (2005) Dietary intakes of fat and fatty acids and risk of breast cancer: a prospective study in Japan. Cancer science 96, 590-599

Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, Roach KC, Mandell J, Lee MK, Ciernikova S, Foretova L, Soucek P, King MC (2006) Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. JAMA : the journal of the American Medical Association 295, 1379-1388

Wang F, Paradkar PN, Custodio AO, McVey Ward D, Fleming MD, Campagna D, Roberts KA, Boyartchuk V, Dietrich WF, Kaplan J, Andrews NC (2007a) Genetic variation in Mon1a affects protein trafficking and modifies macrophage iron loading in mice. Nat Genet 39, 1025-1032

Wang SS, Schadt EE, Wang H, Wang X, Ingram-Drake L, Shi W, Drake TA, Lusis AJ (2007b) Identification of pathways for atherosclerosis in mice: integration of quantitative trait locus analysis and global gene expression data. Circ Res 101, e11-30

Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ (2003) Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci U S A 100, 15901-15905

West DB, Boozer CN, Moody DL, Atkinson RL (1992) Dietary obesity in nine inbred mouse strains. The American Journal of Physiology 262, R1025-1032

West DB, Goudey-Lefevre J, York B, Truett GE (1994a) Dietary obesity linked to genetic loci on chromosomes 9 and 15 in a polygenic mouse model. The Journal of clinical investigation 94, 1410-1416

West DB, Waguespack J, York B, Goudey-Lefevre J, Price RA (1994b) Genetics of dietary obesity in AKR/J x SWR/J mice: segregation of the trait and identification of a linked locus on chromosome 4. Mammalian genome : official journal of the International Mammalian Genome Society 5, 546-552

Weyer C, Funahashi T, Tanaka S, Hotta K, Matsuzawa Y, Pratley RE, Tataranni PA (2001) Hypoadiponectinemia in obesity and type 2 diabetes: close association with insulin resistance and hyperinsulinemia. J Clin Endocrinol Metab 86, 1930-1935

Whiteman MK, Hillis SD, Curtis KM, McDonald JA, Wingo PA, Marchbanks PA (2005) Body mass and mortality after breast cancer diagnosis. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 14, 2009-2014

Williams Rt, Lim JE, Harr B, Wing C, Walters R, Distler MG, Teschke M, Wu C, Wiltshire T, Su AI, Sokoloff G, Tarantino LM, Borevitz JO, Palmer AA (2009) A common and unstable copy number variant is associated with differences in Glo1 expression and anxiety-like behavior. PLoS ONE 4, e4649

Woelfle U, Cloos J, Sauter G, Riethdorf L, Janicke F, van Diest P, Brakenhoff R, Pantel K (2003) Molecular signature associated with bone marrow micrometastasis in human breast cancer. Cancer Res 63, 5679-5684

World-Health-Organiziation (2006) Cancer.

Yakar S, Nunez NP, Pennisi P, Brodt P, Sun H, Fallavollita L, Zhao H, Scavo L, Novosyadlyy R, Kurshan N, Stannard B, East-Palmer J, Smith NC, Perkins SN, Fuchs-Young R, Barrett JC, Hursting SD, LeRoith D (2006) Increased tumor growth in mice with diet-induced obesity: impact of ovarian hormones. Endocrinology 147, 5826-5834

Yamashita S, Wakazono K, Nomoto T, Tsujino Y, Kuramoto T, Ushijima T (2005) Expression quantitative trait loci analysis of 13 genes in the rat prostate. Genetics 171, 1231-1238

Yang H, Crawford N, Lukes L, Finney R, Lancaster M, Hunter KW (2005) Metastasis predictive signature profiles pre-exist in normal tissues. Clin Exp Metastasis 22, 593-603

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ (2006) Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome research 16, 995-1004

Yang Y, Chung EK, Wu YL, Savelli SL, Nagaraja HN, Zhou B, Hebert M, Jones KN, Shu Y, Kitzmiller K, Blanchong CA, McBride KL, Higgins GC, Rennebohm RM, Rice RR, Hackshaw KV, Roubey RA, Grossman JM, Tsao BP, Birmingham DJ, Rovin BH, Hebert LA, Yu CY (2007) Gene copy-number variation and associated polymorphisms of complement component C4 in human systemic lupus erythematosus (SLE): low copy number is a risk factor for and high copy number is a protective factor against SLE susceptibility in European Americans. American Journal of Human Genetics 80, 1037-1054

Yi N, Zinniel DK, Kim K, Eisen EJ, Bartolucci A, Allison DB, Pomp D (2006) Bayesian analyses of multiple epistatic QTL models for body weight and body composition in mice. Genetical Research 87, 45-60

Yusenko MV, Kuiper RP, Boethe T, Ljungberg B, van Kessel AG, Kovacs G (2009) High-resolution DNA copy number and gene expression analyses distinguish chromophobe renal cell carcinomas and renal oncocytomas. BMC cancer 9, 152

Yvert G, Brem RB, Whittle J, Akey JM, Foss E, Smith EN, Mackelprang R, Kruglyak L (2003) Trans-acting regulatory variation in Saccharomyces cerevisiae and the role of transcription factors. Nature genetics 35, 57-64

Zhan M, Zhao H, Han ZC (2004) Signalling mechanisms of anoikis. Histol Histopathol 19, 973-983

Zografos GC, Panou M, Panou N (2004) Common risk factors of breast and ovarian cancer: recent view. International journal of gynecological cancer : official journal of the International Gynecological Cancer Society 14, 721-740