Gene expression profiling of Met receptor tyrosine kinase-induced mouse mammary tumors

by

Marisa Grace Ponzo

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Department of Medicine

Department of Experimental Medicine

McGill University

Montréal, Québec, Canada

Your file Votre référence
ISBN: 978-0-494-66661-6
Our file Notre référence
ISBN: 978-0-494-66661-6

ABSTRACT

Breast cancer is a heterogeneous disease comprised of distinct biological entities that correlate with diverse clinical outcomes. Gene expression profiling has divided this heterogeneity into luminal, ERBB2+ and basal molecular subtypes. Basal breast cancers are difficult to treat as they lack expression of candidates suitable for targeted therapies and are associated with poor outcome.

Elevated protein level of the hepatocyte growth factor receptor, MET, is observed in 20% of human breast cancers and correlates with poor prognosis. However, the role of MET in mammary tumorigenesis is poorly understood. To address this, we generated a murine model that expresses weakly oncogenic mutants of Met (Met^{mt}) in the mammary epithelium under the transcriptional control of the mouse mammary tumor virus promoter. We demonstrate that Met^{mt} induces mammary carcinomas with diverse phenotypes and used gene expression microarrays to elucidate gene expression changes induced by Met. Since mammary tumors contained variable contents of epithelium and stroma, we used laser capture microdissection to procure epithelial cells for microarray analysis. Based on immunohistochemistry and expression profiling, we show that Met^{mt} produces tumors with luminal or basal characteristics. From hierarchical clustering, Met^{mt}-induced basal tumors clustered with murine models that share features of epithelial to mesenchymal transition and human basal breast cancers. Moreover, Met^{mt} basal tumors clustered with human basal breast cancer. The status of MET among the human breast cancer subtypes has not previously been addressed. We demonstrate that MET levels are variable across molecular subtypes but show elevation in the basal subtype and correlates with poor outcome. We used a candidate gene approach derived from microarray data to gain an understanding of signals required for Met-dependent tumorigenesis. We investigated Nck adaptor proteins and demonstrate a role for Nck in cell motility and actin dynamics of Met-dependent breast carcinoma cells and show elevated expression in human basal breast cancers. By generating a unique mouse model in which Met is expressed in mammary epithelia, with the examination of MET levels in human breast cancer, we have established a novel link between MET and basal breast cancer. This work identifies poor outcome basal breast cancers that may benefit from anti-MET therapies.

ABRÉGÉ

Le cancer du sein est une maladie hétérogène englobant plusieurs phenomènes biologiques qui mènent à diverses manifestations cliniques. Le profilage de l'expression génique a permit de diviser cette hétérogéneité en trois sous-types : luminal, ERBB2+ et basal. Les cancers du sein de type basal sont particulièrement difficiles à traiter étant donné leur manque de cibles adéquates pour des thérapies ciblées et sont par conséquent associés à un faible pronostic de survie.

La surexpression du récepteur au facteur de croissance hépatocyte (HGF, Hepatocyte Growth Factor), MET, est observée dans 20% des cancers du sein chez l'humain et est associée à un faible pronostic. Toutefois, le rôle de Met dans la tumorigenèse mammaire chez la souris n'est pas encore bien compris. Nous avons créé un modèle murin qui exprime un mutant faiblement oncogènique de Met (Met^{mt}) dans l'épithélium mammaire sous le contrôle transcriptionnel du promoteur du virus de tumeur mammaire de souris (MMTV, Mouse Mammary Tumor Virus). Nous démontrons que Met^{mt} induit la formation de carcinomes mammaires avec divers phénotypes. Nous avons en outre utilisé le profilage génétique afin d'élucider les variations de l'expression génique induits par Met. Étant donné que les tumeurs mammaires contiennent des quantités variables d'épithélium et de stroma, nous avons utilisé la technique de microdissection par laser afin d'obtenir une population pure de cellules épithéliales pour le profilage génique. En s'appuyant sur les études d'immunohistochimie et sur les profils d'expression, nous montrons que Met^{mt} entraîne la formation de tumeurs de type luminales ou basales. Les tumeurs de type basal induites par Met^{mt} montrent un profil d'expression similaire à celui de modèles murins présentant des caractéristiques de transition epithélio-mesenchymale et de cancers du sein humain de type basal. De plus, les tumeurs basales Met^{mt} se rapprochent des cancers du sein chez l'humain. Le statut de MET dans les différents sous-types de cancers du sein chez l'humain n'avait pas été étudié précédemment. Nous démontrons que les niveaux de MET sont variables en fonction du sous-type moléculaire mais une élévation peut être observée dans le sous-type basal en association avec un faible pronostic de survie. Afin de comprendre les signaux nécessaires à la tumorigenèse induite par MET, nous avons entrepris d'étudier les

protéines adaptatrices Nck après l'analyse des données du profilage génique. Nous démontrons un rôle de Nck dans la motilité et dans la dynamique de l'actine dans les cellules de carcinome mammaire Met dépendant. De plus, nous montrons une élévation de l'expression de Nck dans les cancers du sein humains de type basal. En générant un modèle murin unique dans lequel Met est exprimé dans l'épithélium mammaire et en examinant les niveaux de MET dans les cancers du sein humains, nous avons établi un nouveau lien entre MET et les cancers du sein humain de type basal. Cette étude identifie donc des cancers du sein de type basal à faible pronostic qui pourraient bénéficier de thérapies ciblées contre le récepteur MET.

ACKNOWLEDGEMENTS

This work would not have come to completion without the support of many people along the way.

I would like to first thank my supervisor and mentor Dr. Morag Park. I am extremely appreciative for being assigned to such an inspiring project that has allowed me to integrate basic science research with a clinical aspect over the past 4 years. I am honored to have had the privilege to work with a supervisor who exhibits a strong passion for science and learning, and who is a constant inspiration.

I would like to thank the members of the MD/PhD program. Dr. Jacquetta Trasler who was encouraging since the beginning. Dr. Nada Jabado for her guidance and inspiring passion for research and patient care. I would also like to acknowledge the friendship and support of fellow MD/PhD colleagues, Brett Burstein and Geneviève Cadieux.

I would like to thank collaborators, Robert Lesurf and Dr. Mike Hallett at the McGill Centre for Bioinformatics. Thank you for a great collaboration.

Dr. Louise Larose for her collaboration, advice for the Nck portion of my project and for critical reading of the Nck portion of my thesis.

I wish to acknowledge Drs. Alain Nepveu, Bill Muller, Peter Siegel for helpful discussion throughout my studies and for feedback on my manuscript in Chapter 4.

I wish to recognize the collaboration with MUHC pathologists, Drs. Atilla Omeroglu and Karim Khetani. Thank you for answering my questions and for providing tissue with which I used to optimize my staining conditions.

Thank you to Dr. Babette Schade for advice early on regarding microarray experiments and helpful discussion from April Rose on the same topic-we were the few doing these experiments at the beginning in our department and your support is appreciated.

Dr. Robert Cardiff – you were pleasure to collaborate with and always encouraged me in my endeavors.

I would like to acknowledge the members of the Park Lab – it has been a pleasure to work with you and will truly miss the fun and positive environment you created. I need to thank several individuals specifically:

Stephanie Petkiewicz – Thank you for starting this great project and for your help since the very beginning of my PhD and throughout.

Jennifer Knight –Also, thank you for your help with last experiments and for helpful comments on my thesis.

Naila Chughtai – you put in so much work and effort in helping with cell line maintenance and microarrays. Thank you for putting your all in the work that you do.

Monica Naujokas – thank you for the fun and cheerful environment that you helped create in the lab. I will miss our time in tissue culture and long discussions about cells.

Anie Monast – thank you for taking care of the precious mice with me and for teaching me along the way.

Dongmei Zuo – you were a pleasure to work with. Thank you for the beautiful confocal images and for the time you spent optimizing assays.

Julie Laferrière – thank you very much for helpful discussion throughout my studies and regarding my thesis.

David Germain - for translating my abstract to French.

Jasmine Abella – You were the first student I worked with in the Park Lab, and the great experience impacted my decision to pursue a PhD in the lab. It was wonderful to be able to work with you again at the end. Thank you for your friendship.

I am also grateful to the present and former members of the BCFGG – At the very beginning you took care to teach me everything you could about LCM to microarrays – thank you for this, especially Stéphane Dumont, Margarita Souleimanova, Hong Zhao, and Haiying Chen.

I need to especially thank my supportive family and friends: Mom-Tina, Dad-Ralph, brother-Angelo, grandparents, aunts, uncles and cousins. Thank you for listening to me and always knowing what to say. Even though you may not have understood the science, you understood me and knew the words of encouragement to say to keep me

going. Thank you Alisha Wali and Desiree Corso for constant encouragement, laughter and friendship. Alisha, I will cherish talking over tea at midnight.

Thank you to Marylin Hares, Dominique Besso and Dr. Hugh Bennett in the Department of Experimental Medicine, who created a welcoming environment and were always so helpful and encouraging.

I'd like to thank Greg Paliouras, who could arguably be one of the greatest results of my Ph.D. Thank you for your unconditional love and support.

Funding: This work was supported by the CIHR MD/PhD studentship and Richard H. Tomlinson Doctoral Fellowship.

CONTRIBUTIONS OF AUTHORS

In order to achieve the quality of data presented in this thesis, several collaborations were established with experts outside of my immediate field of study (pathologists, statisticians and bioinformaticians). Author contributions are listed below by chapter. I performed all other experiments and coordinated those that I did not directly perform myself. I made all figures shown in this thesis and wrote the manuscript together with my supervisor Dr. Morag Park from the data shown in Chapter 4.

Chapter 3: Optimization of methods for gene expression profiling of MMTV/Met^{mt} murine mammary tumor and normal tissue

In collaboration with Greg Finak (former PhD student, Mike Hallett's lab), data analysis for Figures 3.13 and 3.14 was performed to test for amplification and labeling batch effects in microarray data. Greg Finak also contributed to the experimental design set up.

Chapter 4: Mammary tumors induced by the Met receptor exhibit diverse histology and are associated with poor outcome and basal breast cancers

Stephanie Petkiewicz and Pascal Peschard (former PhD students, Morag Park's lab) generated the MMTV/Met^{mt} mouse model (Peschard 2005; Petkiewicz 2007). Anie Monast (animal technician) and I performed the animal work (sacrifices, and transplantation experiments). I organized, maintained and collected data on the mouse transgenic lines.

Using the methods that I optimized in Chapter 3, Naila Chughtai (research assistant, Morag Park's lab) performed laser capture microdissection, RNA isolation, amplification, hybridization on the second set of tumor samples: 562T, 567T, 4425T, 4695T, 5033-4T, 5482T, 5612T.

Comparison of gene expression data from MMTV/Met^{mt} mouse mammary tumors to matched-normal, to other mouse models of breast cancer, to human breast cancer and pathway analyses were performed in collaboration with Robert Lesurf (PhD student, Mike Hallett's lab).

MET immunostaining on the ANN human breast cancer cohort, and analysis was performed in collaboration with Irene Andrulis, Frances O'Malley, Shelley Bull, and Dushanthi Pinnaduwage (University of Toronto).

Robert D. Cardiff, who is a pathologist specializing in the study of transgenic animal models and their comparison to human breast cancer, provided a phenotypic description of each MMTV/Met^{mt} mouse tumor (University of California-Davis).

Atilla Omeroglu and Karim Khetani (current and former pathologists with the McGill University Health Centre, respectively) reviewed H&E stained sections with me prior to laser capture microdissection.

Dongmei Zuo (research assistant, Morag Park's lab) performed immunofluorescence staining and captured confocal images shown in Figures 4.10, 4.11, and I assembled the figures.

David Germain (former post-doctoral student, Morag Park's lab) performed quantitative RT-PCR for Snail on HGF-stimulated MDCK cells (Figure 4.16), and I assembled the figures.

Chapter 5: The adaptor protein Nck couples the Met receptor to cell migration and actin dynamics.

Although I optimized the conditions for these experiments, the representative blots shown in Figure 5.10, and the data used for the invasion and migration assay in Figure 5.6 were performed by Jennifer Knight (post-doctoral student, Morag Park's lab). I captured the images and assembled the figures.

Gene expression analysis of *Nck1* and *Nck2* in MMTV/Met^{mt} mouse tumor and matched-normal tissue, and of the status of *NCK1* and *NCK2* in human breast cancer was performed by Robert Lesurf (PhD student, Mike Hallett's lab). I assembled the figures.

Dongmei Zuo performed immunofluorescence staining and captured confocal images shown in Figures 5.7 and 5.8. I assembled the figures.

PUBLICATIONS ARISING FROM THIS WORK

1) Ponzo MG,* Lesurf R,* Petkiewicz S, O'Malley FP, Pinnaduwage D, Andrulis IL, Bull SB, Chughtai N, Zuo D, Germain D, Omeroglu A, Cardiff RD, Hallett M, and Park M. **Mammary tumors induced by the Met receptor exhibit diverse histology and are associated with poor outcome and basal breast cancers.** Manuscript in press with PNAS; *co-first authors.

2) Ponzo MG, Knight JF, Lesurf R, Larose L, Hallett M and Park M. **The adaptor protein Nck couples the Met receptor to cell migration and actin dynamics.** Manuscript in preparation.

3) Ponzo MG, Park M. **Murine models of basal breast cancer: involvement of the Met receptor tyrosine kinase.** Manuscript (review) in preparation.

4) Petkiewicz S, Ponzo MG, Peschard P, Musallam L, Monast A, Park M. **Ubiquitylation suppresses the transforming activity of an oncogenic Met receptor tyrosine kinase *in vivo*.** Manuscript in preparation.

LIST OF ABBREVIATIONS

Abl Abelson murine leukemia viral (v-abl) oncogene homolog 1

Akt1 v-akt murine thymoma viral oncogene homolog 1

Bcr breakpoint cluster region

BLG beta-lactoglobulin

BRCA1 breast cancer 1, early onset

Cbl Casitas B-lineage lymphoma

CD cluster of differentiation

Cdc42 cell division cycle 42 (GTP binding protein, 25kDa)

cDNA complementary DNA

Crk v-crk sarcoma virus CT10 oncogene homolog (avian)

Cy cyanine

Dab1 disabled homolog 1

DCIS ductal carcinoma *in situ*

DNA deoxyribonucleic acid

dNTP Deoxyribonucleotide triphosphate

Dock180 dedicator of cytokinesis 1

Dok downstream of tyrosine kinase

EGF epidermal growth factor

EGFR epidermal growth factor receptor

EMT epithelial to mesenchymal transtition

EPEC Enteropathogenic *Escherichia coli*

ERBB2 v-erb-b2 erythroblastic leukemia viral oncogene homolog 2

ERBB3 v-erb-b2 erythroblastic leukemia viral oncogene homolog 3

ESR1 estrogen receptor

FACS fluorescence activated cell sorting

FAK focal adhesion kinase

FISH fluorescence *in situ* hybridization

FLT3 fms-related tyrosine kinase 3

FVB/N Friend leukemia virus "b" strain/National Institutes of Health mouse

Gab1 growth factor receptor bound protein 2-associated protein 1

GAPDH glyceraldehydes-3-phosphate dehydrogenase

GFP green fluorescent protein

GITC guanidinium isothiocyanate

Grb2 growth factor receptor-bound protein 2

Grb4 growth factor receptor-bound protein 4

H_2O_2 hydrogen peroxide

H_2O water

HA hemagglutinin tag

H&E hematoxylin and eosin

HGF hepatocyte growth factor

HOS human osteosarcoma cell line

ILK integrin-linked kinase

IP immunoprecipitation

Kit v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog

Krt keratin

LCM laser capture microdissection

LCuM laser cutting microdissection

LTR long terminal repeat

MAPK mitogen activated protein kinase

MET hepatocyte growth factor receptor

Met^{mt} MMTV/Met containing mutations M1248T and Y1003F/M1248T

MMTV mouse mammary tumor virus

MNNG *N*-methyl-*N'*-nitro-*N*-nitroso-guanidine

MT metallothionein

NCK non-catalytic region of tyrosine kinase

NOS not otherwise specified

NPHS1 congenital nephrotic syndrome of the Finnish type

NPI Nottingham Prognostic Index

OCT optimal cutting temperature embedding medium

Pak1 p21 protein (Cdc42/Rac)-activated kinase 1

PDGF-B platelet-derived growth factor-B

PDGFR platelet-derived growth factor receptor

PI3K Phosphoinositide 3-kinase

PINCH LIM and senescent cell antigen-like domains 1 (Lims1)

PCR polymerase chain reaction

PGR progesterone receptor

PTB phosphotyrosine-binding

PyMT polyoma virus middle T

qRT-PCR quantitative reverse transcriptase polymerase chain reaction

H-ras v-Ha-ras Harvey rat sarcoma viral oncogene homolog

K-ras v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog

N-ras neuroblastoma RAS viral oncogene homolog

Rac1 ras-related C3 botulinum toxin substrate 1

Rho ras homolog gene family

RNA Ribonucleic acid

Ron Recepteur d'Origine Nantais

SF scatter factor

SH2 Src homology 2

SH3 Src homology 3

Shc1 (Src homology 2 domain containing) transforming protein 1

SNP single nucleotide polymorphism

Src v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)

TGF-β Transforming growth factor beta

Tir translocated intimin receptor

TNM tumor, lymph node, metastases

TPR translocated promoter region

Vav1 vav 1 guanine nucleotide exchange factor

VEGFR1 vascular endothelial growth factor 1

VEGFR2 vascular endothelial growth factor receptor 2

WAP whey acidic protein

WASP Wiskott-Aldrich syndrome (eczema-thrombocytopenia)

N-WASP Wiskott-Aldrich syndrome-like

WAVE1 WAS protein family, member 1

WCL whole cell lysate

WIP1 WAS/WASL interacting protein family, member 1

Note on nomenclature for genes and proteins of mouse or human origin.

Throughout this thesis, effort was made to use proper nomenclature when referring to DNA, RNA and protein of mouse or human origin. Genes and proteins of human origin are written in upper case letters. DNA and RNA of human origin are italicized, while protein products are not. Genes and proteins of mouse origin are written with the first letter in upper case, and following letters in lower case. DNA and RNA of mouse origin are italicized, while protein products are not.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

1 Literature Review

1.1 Breast cancer introduction

1.1.1 Global burden of breast cancer

Epithelial cancers arising in tissues such as breast, lung, colon, prostate and ovary constitute approximately 80% of all cancers. Of these, breast cancer is the fifth leading cause of cancer mortality in the world.* Breast cancer represents the most commonly diagnosed female cancer (Parkin 1998), with the highest incidence and prevalence being in North America and Western Europe (Lacey, Devesa et al. 2002). While age-adjusted mortality rates are increasing in developing countries, these rates are decreasing in the United Kingdom and United States due to improved detection and treatment.*

1.1.2 Canadian statistics on breast cancer

In 2008, the Canadian Cancer Society reported a breast cancer incidence in Canada of 22 400 and mortality rate of 5 300 people per year (2008). Breast cancer accounts for the greatest number of new cancer cases among Canadian women and is twice that reported for lung cancer, but second to lung cancer in mortality rate. The lifetime risk of developing breast cancer for Canadian women is 1 in 9 (11%). A total of 1 in 28 (4%) breast cancer sufferers will succumb to the disease.

1.1.3 Incidence rate

Based on the 2008 Canadian statistics, breast cancer incidences appear to be consistent across Canadian provinces (2008). Between 1979 and 1999, breast cancer incidence has gradually increased, an observation likely attributed to earlier detection

---

* World Health Organization, http://www.who.int/mediacentre/factsheets/fs297/en/index.html

following the introduction of population-based mammography screening. Since 1999 however, the incidence of breast cancer has significantly declined at a rate of 1.7% per year. The recent decline suggests that the prevalent cases in the screened population have been exhausted, with the result that rates have returned closer to pre-screening levels, as was the case for prostate cancer in Canadian men. However, changes in risk and preventive factors such as childbearing patterns, and hormonal levels are also predicted to have had an impact on incidence rates (2008).

1.1.4 Mortality rate

Breast cancer mortality has decreased steadily since the mid-1980s. A similar decline has also been observed in other developed countries including the United Kingdom, United States and Australia. The age-adjusted mortality rate has decreased since 1986 from 32 to 23.1 per 100 000 (more than a 25% reduction). Death rates have declined at 1.6% per year since 1999. This is due to the implementation of screening programs (i.e.: mammography and self-examination) and the use of adjuvant therapies following surgery.

1.1.5 Age-distribution

Breast cancer predominantly affects women between the ages of 50 and 69 (52% of breast cancers are diagnosed in this age group), considerably younger than the age distribution for lung and colorectal cancers (ages 70 or more) (2008). However, the majority of deaths from breast cancer occur in the 80 or older age group, demonstrating the positive effects of prevention and treatment strategies.

1.1.6 Population-based survival rate

The contribution of breast cancer to overall Canadian cancer burden can be

estimated using a population-based indicator of severity of disease, known as the five-year survival rate. The five-year survival rate for breast cancer is relatively equal across Canadian provinces. The best prognosis for breast cancer was observed between the ages 40 and 79 with a five-year survival rate of 88-89%, while the worst prognosis was associated with ages between 20 and 39 (81%) and between 80 and 99 (80%) (2008). Although population-based survival statistics are useful as average indicators, they are not necessarily representative of an individual's survival rate since a host of factors such as specific disease characteristics must be taken into account.

1.1.7 Risk factors associated with breast cancer

Although there are many risk factors associated with developing breast cancer, the most clinically relevant are increasing age, inheritance of predisposing genetic factors (i.e.: "breast cancer 1, early onset," *BRCA1*), family history (regardless of known mutation), atypical hyperplasia and prior exposure to radiation (O'Malley and Pinder 2006).

1.1.8 Breast cancer diagnosis

Lesions of the female breast usually become apparent as palpable nodules or masses that may or may not be painful. Although the majority of breast lumps are identified as benign, all lumps are investigated by triple assessment (ultrasound, mammography and magnetic resonance imaging) (O'Malley and Pinder 2006). For suspected malignant breast lesions, a pre-operative diagnosis is performed to decide on appropriate treatment planning. Samples are taken by fine needle aspirate cytology, core needle biopsy.

1.1.9 Challenges in the clinical management of breast cancer

Breast cancers are phenotypically and clinically heterogeneous and it is currently not possible to accurately identify which patients will develop aggressive, metastatic disease. More than 80% of women with breast cancer currently receive adjuvant chemotherapy, although only 40% of patients relapse and succumb to metastatic disease (Weigelt, Peterse et al. 2005). Thus, a large number of breast cancer sufferers undergo aggressive therapy and would have otherwise survived without it.

As will be discussed in this section, traditional prognostic (a situation or condition, or a characteristic of a patient, that can be used to estimate the chance of recovery from a disease or the chance of the disease recurring) and predictive indicators (a situation or condition that may increase a person's risk of developing a certain disease or disorder) include tumor morphology, expression of molecular markers, and information about the extent of spread of the cancer throughout the body. However, novel prognostic, and predictive markers are needed to identify patients at high risk of developing metastases. Individualized treatment will improve quality of life by minimizing patient 'over-treatment' and provide targeted therapies for the subgroup of patients who will benefit from it. As discussed later, gene expression analyses of primary breast cancers is a current method used to help identify patients with poor outcome that would benefit from adjuvant treatment. Additionally, gene expression analyses have unveiled novel biomarkers for patient prognostication and as candidates for targeted therapy.

1.1.10 Prognostic and predictive indicators of breast cancer

Prognostic indices such as the TNM classification and the Nottingham Prognostic Index (NPI) help to stratify patients into tiered risk groups, which aid in planning treatment strategies.

1.1.10.1 TNM classification system

TNM staging (T=tumor, N=lymph node, M=metastases) was developed in 1942 by Pierre Denoix based on major morphologic characteristics of tumors that determine their behavior; tumor size (T), presence and number of regional lymph node involvement (N), and presence of distant metastases (M). The 10-year survival data for TNM staging is 95% for Stage 0, 88% for Stage I, 65% for Stage II, 35% for Stage III, and 5% for Stage IV (Elston, Ellis et al. 1999).

1.1.10.2 Nottingham Prognostic Index

In 1992, Ellis and Elston devised the Nottingham Prognostic Index (NPI) based on a tumor registry composed of 387 patients as a means to predict prognosis (Galea, Blamey et al. 1992). It has since been validated prospectively (Robertson, Dixon et al. 1992), independently (Balslev, Axelsson et al. 1994; Sundquist, Thorstenson et al. 1999; D'Eredita, Giardina et al. 2001), as a prognostic tool to stratify patients into risk groups: excellent, good, moderate and poor prognosis. This index combines the prognostic power of the three best prognostic factors that will be discussed in the following sections: lymph node stage, histological grade, and tumor size. But there are concerns for patients falling within all tiers. Three-tiered risk predictors are most helpful when a patient's projected risk is at either end of the spectrum. However, for patients falling into the moderate risk group, a decision must be made regarding the benefit of undergoing several months of chemotherapy or hormonal therapy with little chance of improvement. Notably, within the high-risk groups, many women will have undergone aggressive adjuvant treatment yet a portion will never develop recurrence. In the opposite scenario, a portion women in the low-risk group will go on to develop occult metastases without having received aggressive therapy.

1.1.10.3 Traditional factors (morphological)

The purpose of prognostic factors in breast cancer is to identify patients whose tumors are likely to have good or poor outcome. Several putative prognostic factors have been identified over the years, however none predict with absolute certainty the clinical course of breast tumors. Traditional morphological factors that are most clinically useful include four main groups; tumor size, differentiation (histological type, histological grade), vascular invasion, lymph node stage (Elston, Ellis et al. 1999).

1.1.10.3.1 Tumor size

Tumor size is initially measured macroscopically in the fresh state in three planes, then again after fixation when tumor margins are more clearly defined. The greatest diameter is taken as the final tumor size. The size of tumors is a time-dependent dependent prognostic factor in that over time, the tumor is expected to grow larger. It is thus not surprising that many studies have correlated large tumor size to worse long-term survival (Elston, Ellis et al. 1999). Projected relapse-free survival rates over a 20 year period after initial treatment are found to be: <10 mm – 88%, 11-13 mm – 73%, 14-16 mm – 65%, 17-22mm – 59%. It was first reported in the early nineteenth century that the microscopic appearance of a tumor correlated with malignancy, followed by the decision to divide epithelial tumors of the breast into two main divisions, *in situ* (non-invasive) versus invasive carcinoma (Elston, Ellis et al. 1999).

1.1.10.3.2 Differentiation-histological type

The degree of differentiation of a breast tumor is valued as a reliable prognostic factor. It is determined by the histological type of the tumor and by the grade of differentiation. Histological types are assigned based on the architectural pattern of the tumor. Several histological tumor types are associated with favorable outcome, such as

tumors of 'specified type', including tubular, mucinous, invasive cribriform, medullary, infiltrating lobular and tubulo lobular in contrast to invasive carcinomas of 'not otherwise specified' type (NOS) which are associated with poor outcome. The histological types with favorable outcome only account for about 5-10% of breast tumors (Weigelt, Peterse et al. 2005). Unfortunately, the great majority (50-80%) of breast cancer cases fall into the invasive carcinoma (NOS) group and theses cancers are associated with poor prognosis and a 10-year survival rate of 35-50% (Weigelt, Peterse et al. 2005). Lobular invasive carcinoma accounts for 5-15% of breast cancers and is also associated with poor prognosis. Importantly, when compared in multivariate analysis with histological grade, histological type does not qualify as an independent prognostic factor (Pereira, Pinder et al. 1995). Nevertheless, histological type can provide clues on the biology of the cancer.

1.1.10.3.3 Differentiation-histological grade

Histological grade is a powerful prognostic factor and researchers have independently validated the correlation between increasing grade and patient survival (Elston 1987; Henson, Ries et al. 1991). The Scarff-Bloom-Richardson grading system assesses grade using a numerical scale (1-3) that assigns points to three specific histological features and the overall grade is based on the sum of these points (Bloom and Richardson 1957). The histological features are glandular (tubule) formation, nuclear pleomorphism, and mitotic count. Grade I tumors are well differentiated, grade II tumors are moderately differentiated and grade III are poorly differentiated. The 10-year survival rate for grade I tumors is very good at 85% (3-5 points), for grade II tumors is 55% (6-7 points) and is significantly reduced for grade III tumors at 45% (8-9 points) (Elston, Ellis et al. 1999).

1.1.10.3.4 Vascular invasion

Vascular invasion is another prognostic factor with importance in breast cancer (Bettelheim, Penman et al. 1984). Vascular invasion is the presence of tumor emboli in vascular spaces. As a prognostic factor, vascular invasion is not as informative as lymph node stage and however, has been found to correlate with early local disease recurrence even in cases that are lymph node status negative (Pinder, Ellis et al. 1994).

1.1.10.3.5 Lymph node stage

Lymph node staging assesses the extent of cancer spread or metastasis to the regional lymph nodes, and is currently one of the most important prognostic factors in breast cancer. The closest set of lymph nodes to the breast is found in the axilla, and these drain the breast of lymph fluid. Sentinel lymph nodes are a portion of axillary lymph nodes in which breast cancer will likely metastasize first. Patients with the presence of lymph node metastases have significantly poorer prognosis (Carter, Allen et al. 1989; Page 1991). Patients with primary operable invasive breast cancer either undergo axillary lymph node dissection (ALND), where several lymph nodes are removed and histologically examined, or sentinel lymph node biopsy (SLNB), where only the sentinel lymph nodes are removed. Overall, patients with confirmed lymph node metastases have a 10-year survival rate of 25-30% compared to those without, 75% (Elston, Ellis et al. 1999). Although surgical dissection of the axillary lymph nodes is valuable prognostically, it is associated with morbidity (i.e.: lymphedema). SLNB involves the removal of relatively fewer lymph nodes and is associated with less morbidity and comparable prognostic accuracy to ALND (Krag, Anderson et al. 2007).

1.1.10.3.6 Single molecular markers: Hormone receptors

The majority of breast tumors are immunopositive for estrogen (ESR1) and/or for

progesterone receptors (PGR) and this property is exploited pharmacologically with the use of endocrine therapy. About 50-60% of ESR1+ breast tumors are responsive to anti-estrogen therapy while less than 10% of ESR1- breast tumors respond (1980). Furthermore, low levels of hormone receptor (ESR1/PGR-) are associated with metastasis and poor outcome (Page 1991).

1.1.10.3.7 Single molecular markers: ERBB2 and TP53

In recent years there have been numerable molecular markers under investigation for prognostic or predictive power in breast cancer. However, in order for a new prognostic factor to be considered, it must provide information independently from existing clinicopathologic prognostic factors (see TNM and NPI). Perhaps the best example of these to date is the HER2/Neu/ERBB2 receptor tyrosine kinase. About 20-30% of breast cancers overexpress and amplify the *ERBB2* locus, and these tend to be very aggressive tumors that lack expression of hormone receptors. ERBB2 protein levels or *ERBB2* amplification are routinely assessed in treatment centers by either immunohistochemistry or fluorescent *in situ* hybridization (FISH), respectively. Although ERBB2 was found to be prognostic for lymph node positive women, this was not the case for lymph node negative, and thus was termed as a weak to moderate prognostic factor by the World Health Organization (WHO) (Weigelt, Peterse et al. 2005).

TP53 is a tumor suppressor gene and DNA-binding protein that has a multifunctional role in DNA repair, differentiation, and apoptosis. *TP53* is frequently mutated in human cancers (Rakha, Putti et al. 2006), and correlates with poor prognosis in breast cancers. Under physiological conditions, TP53 protein is rapidly degraded, however, *TP53* mutations results in inefficiently degradation of TP53 protein whereby

elevated levels are found within the cell and can be detected by immunohistochemistry. Over the past 15 years, TP53 had been studied in more than 10 000 patients, in over 29 different studies to assess the correlation between *TP53* mutation and breast cancer prognosis. Unfortunately, several different antibodies were used for these immunohistochemical analyses producing different results. A meta-analysis of 37 publications involving over 9 800 patients, high TP53 immunohistochemical staining was correlated with weak prognostic and predictive value (Barbareschi 1996). However, sequencing data from more than 25 studies involving over 6 000 patients have shown that *TP53* mutations are of strong prognostic significance (Borresen-Dale 2003). The predictive value of *TP53* mutations as a clinical biomarker in breast cancers has been investigated extensively, however, a consensus has not been found.

1.2 "-OMICS" and breast cancer

As discussed, clinical management of breast tumors currently relies on traditional clinicopathological parameters, such as lymph node status, histological grade, as well as the expression of specific markers, estrogen receptor (ESR1) and human epidermal growth factor receptor 2 (ERBB2) (Elston, Ellis et al. 1999). However, traditional markers fail to consistently identify those with unfavorable outcome who should receive aggressive treatment, from those with favorable outcome. Within the past decade, whole genome approaches, such as gene expression profiling and array-based comparative genomic hybridization, have provided novel prognostic and predictive information with greater accuracy than traditional clinicopathological parameters. In addition, these large-scale screens have identified novel biomarkers for which targeted therapy may be designed.

1.2.1 Gene expression profiling

Gene expression profiling is a method in which the relative abundance of mRNA species within a sample, the transcriptome, can be measured in a single experiment. A large number of samples are processed in parallel on microarray platforms and the data is compared. The major advantages of using this technique is that it is high-throughput, where mRNA species can be compared simultaneously between samples, and expression data can be linked to clinical data to address specific biological questions.

1.2.1.1 Molecular subtypes of breast cancer

As mentioned, breast cancers are phenotypically and clinically heterogeneous. Gene expression profiling has divided this heterogeneity into molecular subtypes that have been reproducibly associated with different clinical outcomes (Perou, Sorlie et al. 2000; Sorlie, Perou et al. 2001). Using an unsupervised hierarchical clustering approach, in which biologically diverse breast cancers are stratified into subtypes with gene expression profile similarities, natural 'groupings' or molecular subtypes of breast cancers have emerged that correlate with clinical outcome (Figure 1.1). The majority of human breast cancers (60%) are ESR1 positive and fall within luminal-like A and B molecular subtypes. The luminal subtype is named as such to reflect a similarity in gene expression profiling to morphologically normal luminal epithelial cells which line mammary ducts and lobules, and typical luminal markers include keratins 8/18 and transcription factor GATA binding protein 3 (GATA3). Luminal A tumors tend to have relatively good prognosis and respond to hormonal therapies, whereas luminal B tumors have lower ESR1 protein levels, a higher proliferative index and are resistant to therapy (Brenton, Carey et al. 2005). The remaining breast cancers are ESR1-negative and fall in two distinct subgroups that are associated with poor prognosis; ERBB2-positive and the

basal-like subtypes. The ERBB2-positive class is immunopositive for ERBB2 protein, and/or amplification of the *ERBB2* locus as well as other genes in the *ERBB2* amplicon located on chromosome 17q21 (i.e.: *GRB7*). These tumors are treated with a regimen that involves humanized monoclonal anti-ERBB2 antibodies or small molecule kinase inhibitors, and significantly, about 50% of patients respond to treatment. The term "basal" reflects the epithelial layer located basal to the luminal layer, and closest to the basement membrane of the morphologically normal mammary gland, which includes mammary stem cells as well as myoepithelial cells. Basal breast cancers are named as such as they share gene expression similarities to cells within the basal layer of the mammary gland and markers include keratins 5/6, 14, 17 and others. The basal subtype represents between 15-25% of all breast cancers, as defined by immunohistochemistry or gene expression profiling, and this subtype lacks expression of ESR1 and ERBB2 (Perou, Sorlie et al. 2000; Sorlie, Perou et al. 2001; Sorlie, Tibshirani et al. 2003; Sotiriou, Neo et al. 2003; Abd El-Rehim, Pinder et al. 2004; Nielsen, Hsu et al. 2004; Potemski, Kusinska et al. 2005). Thus, targeted therapies are not available for tumors falling within this molecular subtype and chemotherapy is the only therapeutic option for basal-like breast tumors. However, their efficacy is currently under investigation in clinical trials (Rakha, Reis-Filho et al. 2008).

1.2.1.2 70-gene prognosis signature

In an attempt to identify gene expression differences between breast cancer sufferers who are likely to experience disease recurrence compared to those who are not, gene expression profiling was performed on a cohort of 295 axillary lymph node negative/positive, ESR1-positive/negative early stage (stage I or II) breast cancer samples

and data was linked to disease recurrence (van 't Veer, Dai et al. 2002; van de Vijver, He et al. 2002). Using a supervised clustering approach, two pre-determined breast cancer groupings were identified with distinct clinical outcome (i.e.: breast cancer with and without recurrence) and gene expression profiles between these groups were compared. From this, 70-gene prognosis signature was identified that could predict disease recurrence. Significantly, the 70-gene predictor was found to be a strong predictor of metastatic risk independent of traditional clinicopathological parameters, with an overall hazard ratio of 4.6 (95% CI:2.3-9.2, $p<0.001$), and a sensitivity of 91% and specificity of 73%. The 70-gene prognosis signature represents the first example of the development of an array-based diagnostic test (MammaPrint, by Agendia) from gene expression studies that has been translated for clinical use. Large-scale validation studies are currently underway to assess its utility in the clinic (Buyse, Loi et al. 2006).

1.2.1.3 Stromal-derived prognostic predictor

One caveat to the above-mentioned studies is the use of whole tissue samples to generate gene expression profiles. Thus, the resulting profiles represent a mixture of tissue compartments including tumor epithelium and tumor stroma, which in itself contains several different cellular populations (fibroblasts, endothelial cells, macrophages, and lymphocytes). Using an approach that will be described in further detail in Chapter 3, tumor stroma from 53 patients was dissected, subjected to gene expression data and linked to clinical outcome. Intriguingly, a stromal-derived prognostic predictor (SDPP) was derived that stratifies disease outcome independently of traditional clinicopathological parameters, and independently of the aforementioned gene expression predictors as wells as others (Finak, Bertos et al. 2008).

1.2.2 Array-based comparative genomic hybridization (aCGH)

Some breast tumors have been shown to exhibit high-level amplifications and deletions at the level of genomic DNA that correlate with adverse outcome (Isola, Kallioniemi et al. 1995; Jain, Chin et al. 2001; Al-Kuraya, Schraml et al. 2004; Chin, DeVries et al. 2006). Traditionally, these studies have been done using fluorescence *in situ* hybridization (FISH) (Kallioniemi, Kallioniemi et al. 1992; Tanner, Tirkkonen et al. 1994; Al-Kuraya, Schraml et al. 2004), and comparative genomic hybridization on metaphase chromosomes (Kallioniemi, Kallioniemi et al. 1994). Within the past decade, array-based approaches to identify regions of DNA amplification and deletions in the whole genome of breast tumors have been developed (Pinkel, Segraves et al. 1998; Pollack, Perou et al. 1999; Loo, Grove et al. 2004; Naylor, Greshock et al. 2005). The association of gene expression molecular subtypes, gene copy aberrations and clinical endpoints has led to improved patient stratification (Pollack, Sorlie et al. 2002; Chin, DeVries et al. 2006; Neve, Chin et al. 2006). Obviously, the ERBB2 molecular subtype is associated with high-level amplification of *ERBB2*, in addition to increased copy numbers at 1q, 7p, 8q, 16p, and 20q and reduced copy number at 1p, 8p, 13q, and 18q (Chin, DeVries et al. 2006). The luminal A molecular subtype associates with increased copy number at 1q and 16p and reduced copy number at 16q, as well as high-level amplifications at 8p11-12, 11q13-14, 12q13-14, 17q11-12, 17q21-24, and 20q13 which are relatively common for luminal A tumors (Chin, DeVries et al. 2006). Luminal B tumors show chromosomal gains at 1q, 8q, 17q, and 20q and losses involving portions of 1p, 8p, 13q, 16q, 17p, and 22q (Chin, DeVries et al. 2006). Notably, high-level copy number aberrations are infrequent in the basal molecular subtype, and it is associated with frequent low-level DNA amplification (involving 3q, 8q, and 10p) and losses (3p, 4p, 4q,

5q, 12q, 13q, 14q, and 15q), (Chin, DeVries et al. 2006). Interestingly, the great majority gene expression changes were associated with low-level changes in genome copy number, while only 10% of genes contained expression levels that were significantly associated with genome copy number changes. Comparison of gene expression studies with gene copy number has provided insights on the mechanisms of breast tumorigenesis.

1.2.3 Other "-OMIC" approaches

MicroRNAs (miRNAs), which are a class of short non-coding RNAs, frequently act to inhibit gene expression. miRNA expression profiling of breast tumors has begun and differences between molecular subtypes have been observed although these need to be further validated in independent studies (Blenkiron, Goldstein et al. 2007; Adams, Guttilla et al. 2008). Additionally, studies have begun to analyze the protein profiles of human breast cancerss and cell lines (Goncalves, Charafe-Jauffret et al. 2008). Within the next decade, we will undoubtedly see an integration of human breast cancer profiling strategies. This will hopefully provide a better characterization of tumor subtypes with poor prognosis, as well as highlight candidates suitable for targeted treatment and personalized care.

1.3 Mammary development and stem cells

1.3.1 The definition of a normal stem cell

The properties and behavior of stem cells have been best characterized in the hematopoietic system, partly due to the relatively easy accessibility of blood and bone marrow, and the need to understand the hierarchy of the bone marrow and blood system for successful bone marrow transplants. The understanding of the hematopoeitic hierarchy has proved to be of tremendous value as it has aided in revealing the etiology of

diseases within the blood system (i.e.: leukemia) and facilitated the development of treatment. However, the understanding of stem cells in solid organs is far less developed. Like all stem cells, adult stem cells from solid organs must fulfill certain requirements. They need to be *self-renewing*, which means that each cell division must produce at least one daughter cell-a new multipotent stem cell that replenishes the stem cell pool of the organ (example shown in Figure 1.2B). However, stem cells must also be able to undergo *asymmetrical division*, where one of the daughter cells in a cell division loses part of its stem cell abilities and begins differentiation to produce a cell lineage found in the mature organ (example shown in Figure 1.2B). This daughter cell is termed a progenitor cell, and unlike the parent cell, this cell has no self-renewal capacity but has the ability to undergo population expansion to increase the number of fully differentiated cells of the tissue. Studies have shown that the stromal microenvironment of a stem cell compartment (i.e.: extracellular matrix consisting of collagen, diffusible growth factors and cytokines, and cell types such as fibroblasts, endothelial and inflammatory cells), or stem cell niche, is a major regulator of stem cell behavior (Miller, Lavker et al. 2005). The stem cell niche maintains the stem cells in a quiescent state, until they are activated to fulfill their roles in tissue repair, replenishment and regeneration (Bissell and Labarge 2005). Adult stem cells have been identified in epithelium such as the epidermis of the skin (Tumbar, Guasch et al. 2004), endoderm of the gut (Radtke and Clevers 2005) and the breast (see below).

1.3.2 Normal mammary gland development

Studies on normal mammary gland development in the human female breast have been limited for ethical reasons, thus the majority of investigations have been performed

on rodent animal models. The adult rodent mammary gland is made up of a tree-like structure composed of hollow branches termed ducts, which are embedded within adipose tissue. At the ends of the ducts are grape-like out pouches that are called alveolar sacs, and are responsible for milk production and secretion. The milk is subsequently channeled through the ducts and out the nipple during lactation. The mammary ducts and alveoli are formed by a bilayer of mammary epithelium, composed of two cell lineages. The inner *luminal* layer is formed by simple columnar epithelium that produces milk secretion during lactation, and the outer *myoepithelial* layer contracts to push the milk along the ducts in response to the hormone oxytocin.

The mammary gland is a unique organ in that it undergoes several cycles of growth, morphogenesis, and cell differentiation during post-natal and adult life. At birth, the rodent mammary gland consists of 15-20 branching ducts connected to the nipple by a single primary duct. At the onset of puberty at 3 weeks of age, the rudimentary ducts elongate extensively and invade the mammary fat pad through dichotomous branching. Elongation of the mammary ducts begins from the nipple and extends towards the lymph node. By 12-weeks of age, the mammary gland is filled with ducts. This occurs as a result of proliferative activity in the mammary stroma and the terminal end bud, which is a highly specialized region at the invading tip region of the ducts (Figure 1.2). The ductal pattern generated is a result of regulated proliferation of terminal end buds. Just as in mature ducts, a bilayer of epithelium with inner luminal and outer myoepithelial cells is found in the terminal end bud structures (Figure 1.2). Within the tip of the terminal end bud are cap cells, which are putative stem cells specifically for these structures (Smalley and Ashworth 2003). They divide by symmetrical division to replenish the tips with more cap cells, but also undergo asymmetrical division to produce progenitor cells called body

cells that differentiate to form the luminal lineage (Figure 1.2). Cap cells are also thought to produce progenitor cells that form myoepithelium. Simplistically, the role of the anterior region of the terminal end buds is to penetrate the fatty stroma, while the posterior duct is meant to supply differentiating luminal and myoepithelial cells to elongate the ducts (Silberstein and Daniel 1982). How these structures branch through and invade the mammary fat pad during puberty is not well understood.

In a similar manner to the human breast, the mature virgin mouse mammary gland consists of ducts with branches embedded within adipose tissue with scarce alveolar buds. During pregnancy, secondary ductal branches arise which subsequently increases the number of alveoli, such that during lactation, the mammary gland is full of secretory alveoli. After pregnancy, a process termed involution occurs, whereby epithelial cells undergo apoptosis and the gland undergoes remodeling to resemble a virgin-like state. Thus, it can be concluded that normal mammary gland development involves a tightly regulated morphogenic program. This morphogenic program is tightly regulated by stromal interactions and by systemic steroids, and growth factors (Topper and Freeman 1980).

1.3.3 Mammary epithelial cell hierarchy

The signals that govern epithelial cell maintenance within the mammary gland are under intense research however, the current understanding of the mammary epithelial hierarchy is shown in Figure 1.2. Luminal and myoepithelial cell lineages are thought to derive from a common progenitor cell (Figure 1.2). Although grossly oversimplified, the progenitor cells give rise to lineage committed myoepithelial and luminal progenitor cells (Stingl, Eaves et al. 1998; Stingl, Eaves et al. 2001; Gudjonsson, Ronnov-Jessen et al.

2002). These produce the differentiated myoepithelial and luminal cells (Figure 1.2). The cellular cues and signals that dictate differentiation along distinct epithelial cell lineages, and what maintains progenitor cells in an undifferentiated state are currently unknown. Scientific evidence supporting this model is presented in the following section.

1.3.4 Evidence for the existence of mammary stem cells

Several lines of evidence suggest that mammary stem cells exist and that luminal and myoepithelial layers arise from a common progenitor cell. Light and electron microscope studies on human and murine mammary tissue have identified "small light cells" that reside in a suprabasal position between the luminal and myoepithelial layers (Smith and Chepko 2001). These cells contain few organelles, where the number of organelles is associated with the level of differentiation. In the 1950s, DeOme *et al.* first demonstrated the existence of stem cells in mammary tissue by limiting dilution transplantation experiments (Deome, Faulkin et al. 1959). Here, clonal progenitors were isolated from a normal murine mammary gland and transplanted into a cleared (devoid of endogenous mammary epithelium, but containing stroma) mammary fat pad at limiting dilution to generate functional mammary outgrowths containing both luminal and myoepithelial cell lineages. The longevity of mammary stem cells was demonstrated when this experiment was reproduced at several stages throughout the life of the mouse (Smith and Medina 1988). The isolation as well as exhibiting the ability to regenerate the mammary fat pad upon transplantation and thus self-renew are sufficient criteria to identify such cells as mammary stem cells. In the mouse, cleared-fat pad transplantation has become a standard technique in mammary gland biology, however the method of progenitor cell isolation has been refined over the years. Much controversy lies in

methods used to isolate stem cells as these methods are based on several assumptions (Woodward, Chen et al. 2005). One method of stem cell isolation relies on the assumption that since multipotent stem cells are undifferentiated, they should be negative for lineage-specific markers typically found in differentiated luminal and myoepithelial cells, while being positive for "stem cell markers". A matter of large debate is which markers should be used to identify and isolate stem cells and their derivatives. Two reports recently described the isolation of putative breast stem cells, using $CD24^{med}CD49f^{high}$ as markers for fluorescence-activated cell sorting (FACS), which formed a functional mammary gland containing both luminal and myoepithelial cells upon transplantation in cleared fat pads (Shackleton, Vaillant et al. 2006; Stingl, Eirew et al. 2006). These and other studies suggest that mammary stem cells do exist, and that the luminal and myoepithelial layers arise from a common progenitor cell (Figure 1.2). However, the mammary stem cell field is far from having a detailed understanding of the events regulating mammopoiesis of the mammary gland (Visvader and Lindeman 2006).

1.3.5 Stem cell markers for isolation

The field of mammary stem cells is rapidly evolving, and novel stem cell markers for the mammary gland are constantly identified. In mouse mammary epithelium, Stingl *et al.* used the expression of CD24, a protein anchored to the plasma membrane by glycosyl phosphatidylinositol, high expression of CD49f, α_6 integrin, and of CD29, integrin β1, (cumulatively $CD24^{+}$ Stem cell antigen-1 $[Sca\text{-}1]^{low}$ $CD49f^{high}$ $CD29^{high}$) as markers for progenitor cells (Stingl, Eirew et al. 2006). These are cell surface markers used for isolating stem cells with the FACS technology. For human mammary epithelium, markers which have been used to identify multi-lineage progenitors are intermediate

filaments, keratins 19^{+}, 14^{+}, a cell adhesion molecule, Epithelial cell adhesion molecule $(EpCAM)^{high}$, which is found basolaterally in most epithelial cells, $CD49f^{high}$, a luminal cell marker $MUC1^{-}$, which is a glycosylated cell surface marker, and stage-specific embryonic antigen-4 $(SSEA\text{-}4)^{high}$, which is an embryonic stem cell marker found in teratomas. However, there have been contradicting reports about stem cell makers, and further work will be needed to validate these. Recent studies have begun to investigate the functional role of stem cell markers in stem cell biology (Taddei, Deugnier et al. 2008).

Other markers that have been used to enrich for stem cells include murine hematopoeitic stem cell antigen-1, Sca-1, and Hoescht dye. Hoescht dye is a DNA-intercalating fluorescent dye that is more efficiently effluxed from stem cells using the p-glycoprotein ABC transporter protein ABCG2. These populations of cells were termed the 'side population' (Welm, Tepera et al. 2002). However, both have proven to not be as sensitive and specific as the aforementioned markers. It was found that some populations within $Sca\text{-}1^{+}$, or side population cells could not reconstitute the mammary fat pad, and thus these isolates included other cells within the mammary gland that did not exhibit stem cell properties.

1.3.6 Immunohistochemical markers of distinct cell lineages

Traditionally, several immunohistochemical markers of myoepithelial, luminal and progenitor cells have been identified. Progenitor cells stain positively for high molecular weight keratins (intermediate filaments) 5/6, while luminal cells stain for low molecular weight keratins 8/18, and myoepithelial cells stain positive for keratin 14 and

smooth muscle actin (Bocker, Moll et al. 2002; Birnbaum, Bertucci et al. 2004) (Figure1.2).

1.3.7 Cancer stem cells

Sporadic breast cancers are thought to arise from a series of genetic alterations that accumulate over a woman's lifetime. For example, in the 1920s to the 1950s, adolescent girls exposed to toxic doses of irradiation by an X-ray fluoroscope as a screening method for tuberculosis were reported to have higher incidence of breast cancer 25-30 years later (Bissell and Labarge 2005). Although there is much cellular turnover within the breast during a woman's lifetime, the cells that remain quiescent, exhibit longevity, likely are mammary stem cells. Stem cells, then, became attractive candidate cells for the originators of breast cancer. The presence of stem or progenitor markers in mammary tumors suggests that tumors arose from theses immature cells. Cancer stem cells by definition refer to a subset of tumor cells that has the ability to self-renew and generate the diverse cells that comprise the tumor (Bonnet and Dick 1997; Reya, Morrison et al. 2001). Stem cells utilize signaling pathways that are thought to mediate self-renewal, such as the Wnt/β-catenin, Notch, Hedgehog, and Transforming Growth Factor (TGF)-β pathways, and these pathways are also deregulated in a subset of human breast cancers (Woodward, Chen et al. 2005).

Although the idea of tumors originating from transformed stem cells is an old concept (Reya, Morrison et al. 2001), this was first shown by the isolation of cancer stem cells from acute myeloid leukemia (Lapidot, Sirard et al. 1994; Bonnet and Dick 1997). The existence of cancer stem cells in solid tumors came later, when they were isolated from breast, brain, lung and recently, intestinal tumors, and shown that they can re-

establish the tumor upon transplantation (Al-Hajj, Wicha et al. 2003; Fomchenko and Holland 2005; Kim, Jackson et al. 2005; O'Brien, Pollett et al. 2007). The first isolation of cancer stem cells from solid tumors came from Al-Hajj *et al.* who demonstrated that only a minority of human breast cancer cells (as few as 100 cells) had the propensity to regenerate tumors in a human xenograft model (Al-Hajj, Wicha et al. 2003). These breast cancer stem cells were isolated by flow cytometry and were positive for markers (CD44) and negative for lineage-specific markers (CD24).

1.3.8 The stem cell niche: an epithelial-stromal interaction involved in breast cancer development

The stroma has been traditionally thought to play a bystander role in the transformation of breast epithelium. However, several studies have shown the stroma surrounding normal epithelium undergoes changes during tumorigenesis and is histologically and molecularly different from the stroma surrounding tumor epithelium (Allinen, Beroukhim et al. 2004; Bhowmick, Neilson et al. 2004; Orimo, Gupta et al. 2005). Fibroblasts, which are the main stromal cells responsible for the synthesis, deposition and remodeling of much of the extracellular matrix in tumor stroma, produce a variety of paracrine growth factors that stimulate growth, proliferation and survival of epithelial cells, such as hepatocyte growth factor (HGF) and TGF-β (Kuperwasser, Chavarria et al. 2004). Kuperwasser *et al.* have shown that human fibroblasts engineered to overexpress HGF and TGF-β induce the transformation of human epithelial cells in a xenograft model. Therefore, it can be postulated that the stem cell niche, which is simply a specialized stroma surrounding stem cells, has the propensity to incur a growth and proliferative influence on stem cells (Bissell and Labarge 2005). For example, Hedgehog ligand is believed to mediate interactions between the stroma and epithelium in the

developing mammary gland (Woodward, Chen et al. 2005). However, little is known about the contribution of the stem cell niche to cancer stem cell development.

1.3.9 The cancer stem cell theory versus clonal evolution as models of mammary tumorigenesis

The role of cancer stem cells in tumorigenesis, tumor heterogeneity and drug-resistance is under intense scientific observation. Two models, the cancer stem cell theory and clonal evolution have emerged (Figure 1.3). Both models have in common the notion that tumors arise from single cells that have acquired multiple genetic mutations and unlimited proliferative capacity. The differences between these models lie within the mechanisms used to explain intra-tumoral heterogeneity. In the cancer stem cell theory, only cancer stem cells are susceptible to genetic mutation and have unlimited self-renewing division, whereas in the clonal evolution model, any cell in the mammary hierarchy is susceptible to genetic and epigenetic insults, and any cell can undergo self-renewing division. Thus, in the cancer stem cell theory, cancer stem cells are the drivers of tumor initiation and progression, and the resulting bulk of the tumor mass is composed of cells that have undergone aberrant differentiation along luminal and myoepithelial lineages (Polyak 2007).

In the clonal evolution model, any cell-type in the mammary epithelial hierarchy can acquire selective advantage and drive tumorigenesis. In this case, the tumor is composed of a mixture of different mammary epithelial cell types that have undergone transformation and produces intra-tumoral heterogeneity. The two models explain therapy resistance differently. Cancer stem cells are found to be inherently drug resistant, while in the clonal evolution model, resistant clones are selected for during treatment. However, both models are not necessarily mutually exclusive. For instance, a tumor may be

initiated by a cancer stem cell (CSC_1) and following tumor progression it may undergo clonal evolution to give rise to a second cancer stem cell (CSC_2). This may arise due to the acquisition of additional genetic or epigenetic events within the CSC_1. In this case, the CSC_2 is more aggressive and becomes the driver of tumor formation. Indeed, a recent study has supported the combined existence of both the cancer stem cell and clonal evolution theories (Shipitsin, Campbell et al. 2007). Gene expression profiles of putative stem and progenitor cells from normal and cancerous breast tissue were isolated using CD24 and CD44 cell surface markers that were associated with stem cell-like properties (Al-Hajj, Wicha et al. 2003). Gene expression profiling showed that indeed, $CD44^+$ cells were less differentiated and more stem cell-like in their expression pattern than $CD24^+$ cells which were more differentiated. Additionally, cells of the same type (i.e.: normal CD44+ and cancerous CD44+) were more similar to each other than cells of the same tissue (i.e.: normal CD24+ and normal CD44+). Through single nucleotide polymorphism (SNP) and FISH analyses, it was determined that CD24+ and CD44+ cells from the same tumors were clonally related but not identical. This contradicts the cancer stem cell theory and supports the clonal evolution model, or a combination of the two, in that a cancer stem cell may have undergone clone evolution during tumor progression, or as a result of clonal selection from therapy. In addition, the stem cell-like phenotype (i.e.: activation of self-renewal pathways, unlimited growth potential) may be a characteristic that drives selection of cancer stem cells.

1.3.10 A model for the generation of basal, luminal and ERBB2+ molecular subtypes of breast cancer

As mentioned, breast cancer is a heterogeneous disease both at the molecular and clinical levels. In comparison, breast cancer is arguably more heterogeneous than other

epithelial cancers such as colon and pancreatic in which a dominant pathway mutation during tumorigenesis has been identified (Kinzler and Vogelstein 1996; Deramaudt and Rustgi 2005; Segditsas and Tomlinson 2006). In breast cancer, there is no single dominant pathway, and there are multiple histological presentations. Gene expression profiling has revealed the existence of three molecular subtypes of breast cancer that are correlated with clinical outcome that had been identified previously by pathologists: luminal-good prognosis, ERBB2-poor prognosis, basal-poor prognosis. Interestingly, these molecular subtypes are consistent across ethnic groups, and are also evident at a pre-invasive stage of breast cancer, called ductal carcinoma *in situ* (DCIS) (Yu, Lee et al. 2004), which suggests that the molecular subtypes may arise by a common mechanism.

The cancer stem cell and clonal evolution theories can be adapted to hypothesize the mechanism of tumorigenesis of luminal, basal, and ErbB2 subtypes of breast cancer (Figure 1.3). Based on the cancer stem cell theory, each tumor subtype is initiated by a different progenitor cell (Figure 1.3A). As mentioned, different mammary epithelial cell lineages within the normal breast tissue express distinct cell markers (Figure 1.4). The tumors, which are derived from these distinct epithelial cell lineages, are thought to retain the expression of those markers. For example, luminal and ErbB2+ breast cancers are thought to arise from luminal cell progenitors, and consistent with this, are known to express luminal markers, keratins 8/18 (Figure 1.4). Basal breast cancers, derived from a more primitive cell type, express basal markers, keratins 5/6. In the clonal evolution model, all three molecular subtypes are derived from a common progenitor cell (Figure 1.3B). Acquisition of the acquired genetic or epigenetic changes within the common progenitor will dictate the final tumor subtype that is generated. Once again, it is likely

that a combination of both cancer stem cell and clonal evolution will explain the mechanism of tumorigenesis in breast cancer however this remains to be validated.

1.4 Receptor tyrosine kinases

1.4.1 Identification of the MET receptor tyrosine kinase

The identification of the MET receptor tyrosine kinase in the 1980s was driven by the search for transforming genes in human cancer, termed oncogenes. Most oncogenes are the transforming counterparts of "normal" genes within the cell, known as protooncogenes (*c-onc*), some of which were acquired by RNA tumor viruses (*v-onc*). Transforming genes were identified from DNA transfection assays where genomic DNA prepared from human tumors or tumor-derived cell lines was transferred into recipient mouse fibroblasts, NIH3T3 cells (Hill and Hillova 1972; Graham and van der Eb 1973). Cells were monitored for transformation by changes in cell culture (loss of contact inhibition and production of foci). Human-derived transforming genes within a rodent background were identified and subsequently cloned by virtue of their association with the *alu* family of repetitive sequences, which are specific to the human genome, and thus act as human-specific 'tags' associated with the transforming gene in a background of mouse genomic DNA (Pulciani, Santos et al. 1982; Schmid and Jelinek 1982).

During this time, it was also shown that transforming genes could be identified in chemically-induced rodent tumors and transformed cells lines *in vitro* through the use of chemical carcinogens. Many of the oncogenes identified in this manner were related to the *ras* family of oncogenes (Balmain and Pragnell 1983; Eva and Aaronson 1983; Guerrero, Calzada et al. 1984). In some cases, it was possible to transform human cell lines using chemical carcinogens (Kakunaga 1978). As evidence of this, DNA isolated

from chemically-treated human tumor cell lines also transformed NIH3T3 cells, demonstrating that transforming genes were present in the chemically-treated human cells (Cooper, Blair et al. 1984).

The chemical carcinogen, *N*-methyl-*N'*-nitro-*N*-nitroso-guanidine (MNNG), was used to transform the human osteosarcoma (HOS) cell line, and a dominant transforming gene was determined to exist through DNA transfection assays (Cooper, Blair et al. 1984). Additionally, the transforming gene shared no homology to H-*ras*, K-*ras* or N-*ras*, which were frequently activated in chemically-treated cells. Using Southern analysis and karyotyping strategies with rodent-human somatic cell hybrids, the transforming gene was mapped to chromosome 7 (7q11.4-7qter), providing additional support that the transforming gene represented a different oncogene than those known to map on other chromosomes (Cooper, Park et al. 1984). Since *erbB* localized to 7pter-7q22, it could not be excluded at the time that the MNNG-HOS transforming gene was not *erbB*, although *erbB* probes failed to bind to the MNNG-HOS gene (Cooper, Park et al. 1984). The MNNG-HOS transforming gene was given the three-letter abbreviation 'MET', named after the "methyl" group from the name of the chemical carcinogen, MNNG. MET was one of the first oncogenes found that had not been previously identified as a retroviral equivalent. It is erroneous to term the cellular MET protooncogene as "c-*met*", since there is no need to distinguish it from a viral counterpart that was never identified (v-*met*).

A cDNA fragment of the MNNG-HOS transforming gene was subsequently sequenced, and was determined to be highly homologous to tyrosine kinase domains of growth factor receptors and tyrosine kinase oncogenes, such as insulin receptor, *erbB*, v-*abl*, v-*src*, and v-*mos* (Dean, Park et al. 1985). By *in situ* hybridization, the localization of *MET* on chromosome 7 was further refined to 7q21-q31 (Dean, Park et al. 1985). In 1986,

the MNNG-HOS transforming gene was found to be a chromosomal rearrangement in which a 5' portion of DNA from chromosome 1q25 was fused to a 3' portion of DNA from the *MET* gene on chromosome 7q31 (Park, Dean et al. 1986). The DNA locus from chromosome 1q25 thus served as the promoter for the *MET* gene, and this was termed *TPR*, or translocated promoter region. The resulting 5 kb RNA species coded for the active MET protein product that was responsible for transformation. Additionally, probes produced from the MET oncogene recognized multiple RNA species in MNNG-HOS transformed cells. A 9 kb species, which was determined to be the MET protooncogene was found to have restricted expression in a cell-type specific manner, while a 10 kb species, found in all cells tested, was the *TPR* locus. The discovery of *TPR-MET* was the first description of an activated gene fusion product from cells treated with a chemical carcinogen (Park, Dean et al. 1986) and would be the prototype of receptor tyrosine kinase-derived oncogenes generated from chromosomal translocation (Figure 1.5) (Rodrigues and Park 1993). Additionally, this report came soon after the publication of *BCR-ABL*, the chromosomal rearrangement responsible for chronic myelogenous leukemia (de Klein, van Kessel et al. 1982), and thus further validated chromosomal rearrangements as a mechanism for oncogenesis. In the years following, the breakpoint region within *TPR-MET* was characterized (Dean, Park et al. 1987), and the sequence of full-length MET was published (Park, Dean et al. 1987). The mechanism of *TPR-MET* activation was found to be a result of constitutive dimerization caused by the TPR leucine zipper, which results in activation in the absence of ligand (Rodrigues and Park 1993).

1.4.2 The structure of the MET receptor

MET is a single-pass transmembrane receptor composed of disulphide-linked 50 kDa α- and 145 kDa β-subunits (Figure 1.5) (Giordano, Ponzetto et al. 1989). The mature α/β heterodimer is formed from proteolytic cleavage of a precursor in the post-Golgi compartment (Giordano, Di Renzo et al. 1989). The extracellular portion of MET is composed of the Sema domain, which encompasses the whole α domain and part of the β domain, and is structurally similar to domains found in semaphorin and plexin family members, the PSI domain, which are also common to domains found in plexin, semaphorin and integrins, and four immunoglobulin-like domains (Figure 1.5) (Gherardi, Youles et al. 2003). The intracellular portion of MET begins at the juxtamembrane domain, followed by the kinase domain, and C-terminal tail that contain multiple tyrosine residues that serve as multi-substrate docking sites (Figure 1.5).

1.4.3 Hepatocyte growth factor/scatter factor

Hepatocyte growth factor (HGF) was originally identified as a blood-borne factor that was a mitogen for liver parenchyma in culture, and induced liver regeneration following injury caused by partial hepatectomy or toxins *in vivo* (Zarnegar and Michalopoulos 1989; Ishiki, Ohnishi et al. 1992). HGF was produced by the liver itself and was found at increased levels in the plasma 24 hours following liver injury (Nakamura, Nawa et al. 1984). The growth factor was purified, the primary amino acid sequence deduced, and cloned from cDNA (Miyazawa, Tsubouchi et al. 1989; Nakamura, Nishizawa et al. 1989; Zarnegar and Michalopoulos 1989; Rubin, Chan et al. 1991; Weidner, Arakaki et al. 1991). An independent group described "scatter factor" which was produced from fibroblasts and induced scattering of tight epithelial sheets from a variety of normal epithelial tissues or cell lines, of which the Madin Darby Canine

Kidney (MDCK) epithelial cell line was the most sensitive (Stoker, Gherardi et al. 1987). Scatter factor and HGF were later determined to be encoded by same gene, and thus represented same molecule (Gherardi and Stoker 1990; Weidner, Arakaki et al. 1991). HGF/SF was then established as the ligand for MET, or hepatocyte growth factor receptor (Bottaro, Rubin et al. 1991; Naldini, Weidner et al. 1991). The MET receptor tyrosine kinase is found in many tissues in the adult, including breast epithelium (Beviglia, Matsumoto et al. 1997). MET is expressed in tissues of epithelial and endothelial origin during development and at maturity, while HGF is found in mesenchymal cells (Rosario and Birchmeier 2003). HGF is produced by fibroblasts within the stroma, and activates the MET receptor in epithelial cells in a paracrine manner.

HGF is secreted as a biologically inert, single chain precursor that is cleaved to produce a biologically active form by extracellular proteases. Biologically active HGF is composed of α and β subunits connected through a disulphide bond. The α-chain contains a hair-loop and four kringle domains (K1-K4) (Lokker, Mark et al. 1992; Okigaki, Komada et al. 1992), while the β-chain is structurally similar to serine proteases of the fibrinolysis system, plasminogens, but is itself proteolytically inactive (Lokker, Mark et al. 1992; Donate, Gherardi et al. 1994). The K1 portion mediates receptor binding, where K1 is thought to engage two MET molecules leading to receptor clustering and activation (Lokker, Presta et al. 1994; Chirgadze, Hepple et al. 1999; Gherardi, Sandin et al. 2006).

1.4.4 Biological responses of the MET/HGF signaling axis

The MET/HGF pathway is a major regulator of motogenesis, morphogenesis, cell survival, and invasion. The MET/HGF signaling pathway coordinates a complex

morphogenic program known as branching tubulogenesis (Rosario and Birchmeier 2003). This involves cell proliferation, motility, cytoskeletal and cell-cell junction rearrangements and remodeling of the cell matrix. Altogether, this process is thought to occur *in vivo* during kidney, lung and breast organogenesis, and MDCK cells grown in collagen is a well-established method of studying this process (Montesano, Matsumoto et al. 1991). During tubulogenesis, cells form polarized, hollow cyst-like structures and upon the addition of HGF, cells extend and arrange in chains away from the cyst, transiently losing their polarity (Fournier, Kamikura et al. 1996; Maroun, Holgado-Madruga et al. 1999; Rosario and Birchmeier 2003). The tubules subsequently regain polarity and form hollow branches. The ability of MET signaling to generate tubes in three-dimensional matrix requires a major phosphoprotein and binding partner, the adaptor protein GRB2-associated binding protein 1, or Gab1 (Weidner, Di Cesare et al. 1996; Maroun, Holgado-Madruga et al. 1999).

MET also regulates pathways involved in actin cytoskeletal dynamics, cell adhesion and migration through kinases Src, Rho family of small GTPases (Rho and Rac), and serine/threonine kinase Pak1 (Rahimi, Hung et al. 1998; Royal, Lamarche-Vane et al. 2000). These biological activities are mediated through Gab1 by association to SH2/SH3 adaptor protein Crk, which results in breakdown of cell-cell junctions and cell dispersal (Lamorte, Royal et al. 2002; Rodrigues, Fathers et al. 2005).

The above-mentioned processes, tubulogenesis, cell migration and scattering, involve a process known as epithelial to mesenchymal transition (EMT). EMT was originally described by developmental biologists as a morphological conversion occurring at specific sites in embryonic epithelia to give rise to individual migratory cells (Hay 1995). As indicated, HGF or 'scatter factor' was in part identified through its ability

to induce a tight colony of epithelial cells to disperse or scatter, which is a two-dimensional representation of an EMT. HGF/MET pathway activation induces EMT in a variety of epithelial cells and tissues (Stoker, Gherardi et al. 1987). In the adult organism, MET and HGF expression is upregulated following tissue injury in restricted sites such as the liver, kidney, and heart, suggesting that the MET/HGF pathway is utilized for tissue repair and wound healing and these processes also involve EMT (Michalopoulos and DeFrances 1997; Nakamura, Mizuno et al. 2000; Matsumoto and Nakamura 2001). Under physiological conditions, HGF plays a regenerative role in tissues in addition to the liver (Ishiki, Ohnishi et al. 1992), including lung (Yanagita, Matsumoto et al. 1993), and kidney epithelium (Kawaida, Matsumoto et al. 1994). A role for the Met receptor in liver regeneration, repair and tissue remodeling in adult animals was demonstrated with *Met* conditional knockout models which showed a defect in liver repair following liver injury induced either chemically (Huh, Factor et al. 2004) or by partial hepatectomy (Borowiak, Garratt et al. 2004). EMT is thought to occur *in vivo* during embryogenesis, wound healing and tissue remodeling, or under aberrant conditions, such as tumor invasion and metastasis.

1.4.5 MET/HGF signal transduction

The role of receptor tyrosine kinases is to transmit signals from extracellular cues within the cell to mediate biological activities. The MET receptor becomes activated following HGF binding, resulting in receptor clustering and autophosphorylation on tyrosine residues, Y1234/1235, within the kinase domain which results in kinase activation, as well as tyrosines in the C-terminal tail, Y1349, Y1356 and Y1365, which serve as docking sites for protein-protein interactions (Figure 1.6). Early studies using a

chimeric MET receptor show that the cytoplasmic domain of the MET receptor is sufficient to orchestrate the biological responses induced by HGF in epithelial cells (Weidner, Sachs et al. 1993; Zhu, Naujokas et al. 1994). Additionally, tyrosine phosphorylation of Y1349 and Y1356 residues in the C-terminus are necessary and sufficient to mediate all biological activities downstream from the MET receptor (Ponzetto, Bardelli et al. 1994; Zhu, Naujokas et al. 1994; Fixman, Naujokas et al. 1995). Phosphotyrosine residues of the activated MET receptor serve as sites for binding to Src homology 2- (SH2) domain- or phosphotyrosine binding domain (PTB)-containing proteins. These include the adaptor proteins GRB2 and SHC, the scaffold protein GAB1, which is the major phosphoprotein downstream from the activated MET receptor (Figure 1.6) (Fixman, Fournier et al. 1996; Nguyen, Holgado-Madruga et al. 1997). GAB1 is part of a family of docking proteins that lack enzymatic activity and includes GAB2, daughter-of-sevenless (DOS), and more distantly related insulin-receptor substrate 1 (IRS1), downstream-of-kinases (DOK) and fibroblast growth factor receptor substrate 2 (FRS2) (Liu and Rohrschneider 2002). Genetic (Sachs, Brohmann et al. 2000) and cell-based assays (Maroun, Holgado-Madruga et al. 1999) have shown that GAB1 is required for biological outcomes downstream from MET (Figure 1.6). Indeed, *Gab1*$^{-/-}$ mouse models exhibit a similar phenotype to *Met*$^{-/-}$ and *Hgf*$^{-/-}$ embryos (Sachs, Brohmann et al. 2000). GAB1 is recruited directly to MET on phosphotyrosine Y1349 by a binding site within GAB1 known as the MET-binding domain (MBD), but also indirectly, through GRB2 (Schaeper, Gehring et al. 2000; Lock, Frigault et al. 2003). The GAB1 MBD is dissimilar to SH2 or PTB domains and represents a unique binding site between a receptor tyrosine kinase and substrate (Lock, Frigault et al. 2003). Both indirect and direct methods of GAB1 recruitment to MET result in a robust association and prolonged

phosphorylation of GAB1 (in comparison to EGFR-GAB1), which is required for an invasive morphogenic response (Maroun, Naujokas et al. 2000). Once GAB1 is tyrosine phosphorylated, it recruits the P85 subunit of PI3K, PLCγ, the tyrosine phosphatase SHP2 and the CRK adaptor protein to the MET complex (Figure 1.6) (Nguyen, Holgado-Madruga et al. 1997; Maroun, Holgado-Madruga et al. 1999; Lamorte, Kamikura et al. 2000; Maroun, Naujokas et al. 2000; Schaeper, Gehring et al. 2000). In addition, activation of MET recruits STAT3 (Boccaccio, Ando et al. 1998), the tyrosine kinase SRC (Ponzetto, Bardelli et al. 1994), focal adhesion kinase FAK (Chen and Chen 2006), and lipid kinase SHIP1 (Stefan, Koch et al. 2001) among others. Hence, the intracellular domain of the MET receptor allows for the assembly of multiprotein complexes, which act to diversify the signals downstream from the receptor. The MET receptor can cooperate with signaling pathways generated from other cell surface receptors, such as for β4 integrin (Trusolino, Bertotti et al. 2001), Plexin B (Conrotto, Corso et al. 2004), as well as hyaluronan receptor CD44 (Orian-Rousseau, Chen et al. 2002), possibly leading to further diversification of signaling.

1.4.6 MET and HGF in mammalian development

Signaling from the MET/HGF pathway is essential during embryogenesis as *Met*$^{-/-}$ and *Hgf*$^{-/-}$ transgenic animals are embryonic lethal at approximately day 14.5 (Bladt, Riethmacher et al. 1995; Uehara, Minowa et al. 1995). The *Met*$^{-/-}$ murine studies phenocopy those of the *Hgf*$^{-/-}$ studies, and demonstrate that during embryogenesis, the Met receptor is the only receptor for HGF, and HGF is its only ligand. Defects were found in tissues that normally express high levels of Met and HGF, such as the placenta,

liver, kidney and skeletal muscle cell precursors (Bladt, Riethmacher et al. 1995; Schmidt, Bladt et al. 1995; Uehara, Minowa et al. 1995).

Often during development, cells migrate over long distances to their final destination where they form the intended tissue or organ. Interestingly, both *Met*$^{-/-}$ and *Hgf*$^{-/-}$ mice showed a defect in all muscle tissue derived from long-range muscle progenitor cells (Bladt, Riethmacher et al. 1995). Normally, through the activation of the Met/HGF pathway in a paracrine manner, the muscle progenitor cells undergo an EMT from the epithelial dermomyotome structure, migrate to the limbs, tongue and diaphragm to produce skeletal muscle. During delamination of muscle progenitor cells, HGF is initially produced by the mesenchyme located proximally to the migrating cells to allow for release of these cells, then along the routes of migration and at the final destination to drive the migration. Similar processes are involved in neural crest cell and mesoderm formation (Takayama, La Rochelle et al. 1996).

1.4.7 MET/HGF in human cancer

When the MET/HGF signaling axis is deregulated, a program of uncontrolled proliferation and invasive growth ensues. MET and HGF are deregulated in multiple types of human cancers (http://www.vai.org/met/). Various methods of MET/HGF-mediated tumorigenesis have been investigated (Figure 1.7). Co-expression of MET and HGF in an autocrine manner is a method used frequently in tumors of mesenchymal origin to drive sarcomagenesis as seen in human sarcomas, osteosarcoma, and rhabdomyosarcoma (Figure 1.7) (Cortner, Vande Woude et al. 1995). Similarly, elevated expression of MET and HGF, and activation through a paracrine manner is evidenced in tumors of epithelial origin such as breast, pancreatic, and colorectal cancers (Ghoussoub,

Dillon et al. 1998; La Rosa, Uccella et al. 2000; Baldus, Kort et al. 2007). Additionally, amplification of the MET receptor has been identified in gastric and colorectal cancers (Figure 1.7) (Di Renzo, Olivero et al. 1995; Nakajima, Sawada et al. 1999). Both overexpression and amplification of the MET receptor can result in catalytic activation of the kinase domain in the absence of HGF (Figure 1.7). However, since HGF is ubiquitously expressed in the stroma, cancer cells with elevated MET levels become responsive to sub-threshold levels of ligand, producing a paracrine loop of activation. Although occurring less frequently, activating mutations in the kinase domain of the receptor represents another mode of MET deregulation (Figure 1.7). Missense mutations are seen in all patients with hereditary papillary renal cell carcinoma (PRCC) and in 13% of sporadic PRCC cases (Schmidt, Duh et al. 1997). In lung and gastric cancer, point mutations and exon deletions (exon 14) within the juxtamembrane domain of the MET receptor have also been found, resulting in delayed degradation and sustained signaling of the receptor (Figure 1.7) (Lee, Han et al. 2000; Kong-Beltran, Seshagiri et al. 2006). These data support a causative role of the MET/HGF signaling axis in tumorigenesis.

Another mechanism that MET has been shown to result in oncogenesis is through crosstalk with other receptor tyrosine kinases. The MET signaling pathway has been shown to crosstalk with other receptor tyrosine kinases, such as RON, ERBB2, ERBB3, and EGFR (Khoury, Naujokas et al. 2005; Bonine-Summers, Aakre et al. 2007; Engelman, Zejnullahu et al. 2007; Cheng, Chytil et al. 2008; Lee, Chow et al. 2008). The MET receptor has been shown to transphosphorylate its family member, RON, and vice versa (Follenzi, Bakovic et al. 2000). MET can synergize with ERBB2 to promote a malignant and invasive phenotype by disrupting epithelial morphogenesis, resulting in the breakdown of cell-cell junctions and causing dispersal and invasion of single cells

(Khoury, Naujokas et al. 2005). This is intriguing since a portion of breast tumors co-express MET and ERBB2, and this may represent a potent mechanism of malignancy in these poorly studied tumors (Lengyel, Prechtel et al. 2005).

Recently, MET was identified as a target of an oncogenic variant of EGFR with a truncated extracellular domain, EGFRvIII, is commonly found in a form of brain cancer with poor prognosis, glioblastoma multiforme. The level of tyrosine phosphorylation of MET authophosphorylation sites, Y1234/1235, was proportional to increasing expression of EGFRvIII, suggesting that MET can be transactivated by EGFRvIII (Huang, Mukasa et al. 2007). EGFRvIII-expressing cells responded (decreased cell survival) when a combination of MET receptor inhibitors (SU11274) together with EGFR inhibitors (erlotinib, AG1478) or chemotherapy were used (Huang, Mukasa et al. 2007; Stommel, Kimmelman et al. 2007). A novel mechanism of tumor resistance to therapy through cross-talk with the MET receptor has recently been unveiled for non-small cell lung cancer (NSCLC), which accounts for 80% of all lung cancers. About 25-30% of NSCLC tumors containing EGFR-activating mutations (ie: T790M) show intrinsic resistance to anti-EGFR therapies (erlotinib, gefitinib), and those that are responsive eventually acquire resistance after variable periods of time (Inoue, Suzuki et al. 2006; Mitsudomi and Yatabe 2007). Tumors exhibiting acquired resistance to EGFR-inhibitors were found to have selected for amplification of the *MET* locus (Figure 1.7) (Bean, Brennan et al. 2007; Engelman, Zejnullahu et al. 2007; Rikova, Guo et al. 2007; Guo, Villen et al. 2008). The mechanism of acquired resistance involved restoring survival pathways from the PI3K/AKT pathway downstream from ERBB3, and amplification of the *MET* receptor engaged ERBB3 to do so. Only when NSCLC-resistant cells received a combination of MET and EGFR inhibitors did they undergo a therapeutic response.

Notably, in patients that had acquired resistance to EGFR therapies, 50% exhibited T790M mutation and 20% had amplification of *MET* (Kobayashi, Boggon et al. 2005; Pao, Miller et al. 2005; Bean, Brennan et al. 2007; Engelman, Zejnullahu et al. 2007; Mitsudomi and Yatabe 2007). More recently, HGF-induced activation of MET was found to be another mechanism resulting in acquired and intrinsic resistance of NSCLC by restoring PI3K/AKT activity in the absence of both ERBB3 and MET amplification (Yano, Wang et al. 2008). By immunohistochemical analysis, cancer cells were found to show immunoreactivity to HGF and MET, thus suggesting an autocrine mechanism of activation (Yano, Wang et al. 2008). Altogether, these data provide strong rationale for the use of MET therapies in the treatment of a number of solid tumors.

1.4.8 MET and HGF therapeutics

Given that MET and HGF are involved in multiple human cancers, strategies to block their action through therapeutics have been devised. The main therapeutic strategies include, 1) HGF and MET biological antagonists, 2) antibodies against HGF and MET, 3) small-molecule inhibitors targeted to the MET catalytic domain.

Decoy MET is a recombinant protein that mimics the structure of the full extracellular domain of MET (Michieli, Mazzone et al. 2004). It functions to bind and sequester active HGF, which titres out ligand availability for the MET receptor. Additionally, decoy MET binds to the MET receptor and prevents dimerization. The isolated Sema domain of MET acts in a similar manner to decoy MET, by competitively inhibiting ligand binding to MET (Kong-Beltran, Stamos et al. 2004). Examples of HGF competitive analogues are NK2 (Chan, Rubin et al. 1991; Montesano, Soriano et al. 1998), and NK4 (Matsumoto and Nakamura 2003; Matsumoto and Nakamura 2008),

which are naturally occurring or truncated forms of HGF. These act by binding to the MET extracellular domain but fail to elicit a biological response. HGF and MET biological antagonists are in the preclinical stage of development.

Anti-HGF antibodies (various from Amgen, L2G7 from Galaxy Biotech, AMG102 from Amgen) and anti-MET antibodies (OA-5D5 by Genentech, DN30 by Metheresis) are currently in the early stages of clinical development (Burgess, Coxon et al. 2006; Kim, Wang et al. 2006; Martens, Schmidt et al. 2006; Petrelli, Circosta et al. 2006; Jun, Sun et al. 2007). Anti-HGF antibodies act in a similar manner to HGF antagonists described above, by sequestering HGF and preventing it from binding and activating the MET receptor. Amgen has produced various fully humanized anti-HGF antibodies, and one of these, AMG102, is under evaluation in progressed glioblastoma and kidney cancer cases in a Phase II clinical trial. Several anti-MET antibodies have been produced however their mechanism of action remains poorly understood. Anti-MET antibodies have been largely unsuccessful as they act as agonists, inducing activation of MET by virtue of the bivalent ("two-armed") structure of the immunoglobulin itself (Prat, Crepaldi et al. 1998). Bivalent anti-MET antibodies induced the biological effects elicited by HGF using *in vitro* assays (Prat, Crepaldi et al. 1998). This problem has been avoided with the development of monovalent ("one-armed") antibodies (Martens, Schmidt et al. 2006). The primary mode of neutralization by current clinical candidates is by blockade of ligand binding (Kong-Beltran, Seshagiri et al. 2006; Jin, Yang et al. 2008). A secondary effect is receptor down-regulation after prolonged exposure to antibody, which impairs signal transduction, abolishes invasive growth response *in vitro*, and interferes with tumorigenic and metastatic potential of tumor cells *in vivo* (Petrelli, Circosta et al. 2006).

Small-molecule inhibitors compete with ATP for binding to the active site of the kinase. A broad-spectrum kinase inhibitor, the staurosporine analogue, K252a, was shown to inhibit the MET receptor (Morotti, Mila et al. 2002). Small-molecule inhibitors with more specificity for the MET receptor have been developed and are the compounds, SU11274 (Sugen, now Pfizer) (Sattler, Pride et al. 2003; Berthou, Aebersold et al. 2004), and PHA665752 (Pfizer) (Christensen, Schreck et al. 2003; Smolen, Sordella et al. 2006). However, these small-molecule inhibitors did not perform well in clinical trials due to poor pharmacokinetic properties (i.e.: low oral bioavailability), although they continue to be used for cell-based and some animal-based experiments. Several novel small-molecule inhibitors have been generated with greater oral bioavailability by different pharmaceutical companies and are currently in preclinical, Phase I or II clinical trials. Small-molecule inhibitors differ in their selectivity for the MET receptor. As mentioned earlier, receptor tyrosine kinases can cross-talk therefore inhibiting multiple receptors with a single agent could be an advantage, as is the case with compound XL880 (Exelixis), which mainly targets MET and VEGFR2, and MP470 (SuperGen), which targets MET, Ret, and mutant forms of Kit, PDGFR and FLT3 (Comoglio, Giordano et al. 2008).

1.5 MET and breast cancer

1.5.1 MET/HGF in normal mammary gland development

To date, there are no published reports of a conditional knockout animal model of Met or HGF in the mammary gland epithelium. Furthermore, embryonic lethality of knockout Met and HGF mice has precluded the study of mammary development post-natally (Bladt, Riethmacher et al. 1995; Schmidt, Bladt et al. 1995; Uehara, Minowa et al.

1995). Despite this, an experimental model was used to investigate the role of HGF in mammary epithelial development (Yang, Spitzer et al. 1995). Mammary gland organ cultures derived from excised mammary fat pads of 4-week old animals contained a scarcely developed epithelium and were cultured. Using an anti-sense oligonucleotide directed against HGF, ductal branching was inhibited (Yang, Spitzer et al. 1995). When primary mammary epithelial cells were infected with HGF-expressing DNA, cells had an increased growth rate in culture, and three-dimensional collagen matrices formed branching tubes within 6 days. When these were transplanted into a cleared mammary fat pad, an increase in ductal end bud size and number, abnormal end bud structures ("clover-shaped") as well as hyperplastic ductal branching were observed (Yant, Buluwela et al. 1998). Furthermore, using an experimental animal model that recapitulates human mammary epithelial morphogenesis, human fibroblasts were engineered to stably express HGF and these were implanted into cleared mammary fat pads. When morphologically normal human mammary epithelial cells from reduction mammoplasties were subsequently engrafted in the HGF-expressing stroma, human invasive ductal carcinomas were generated (Kuperwasser, Chavarria et al. 2004). These data suggest a role for Met and HGF in mammary gland development although genetically engineered animal models will be needed to directly test their roles.

Met and HGF expression levels are coordinately regulated in the rodent mammary gland during development (Niranjan, Buluwela et al. 1995; Pepper, Soriano et al. 1995; Soriano, Pepper et al. 1995; Yang, Spitzer et al. 1995). Met and HGF RNA expression is weakly detectable in the 6-week old virgin glands, and increases at 12-weeks of age. High levels of expression were consistently maintained from 2.5-12.5 days into pregnancy but begin to decrease near the end of pregnancy (Niranjan, Buluwela et al. 1995). At the end

of pregnancy and beginning of lactation, the expression is undetectable and this coincides with an increase in milk protein β-casein, and whey acidic protein (Niranjan, Buluwela et al. 1995; Yang, Spitzer et al. 1995; Niemann, Brinkmann et al. 1998). During lactation, Met and HGF expression are undetectable and this has is likely under neuroendocrine control as it coincides with an increase in prolactin levels which were shown to decrease Met expression in NMuMG cells (Pepper, Soriano et al. 1995). These data are in agreement with analysis of Met and HGF RNA expression in developmental stages of the rat mammary gland (Pepper, Soriano et al. 1995). The levels of Met and HGF subsequently increase during involution, a period during which the gland undergoes extensive remodeling. Thus, HGF and Met are specifically elevated during mammary gland development and periods of extensive differentiation and remodeling that require elongation and tubulogenesis, but are diminished during lactation suggesting that the Met/HGF signaling axis does not play a role in alveologenesis.

In vitro experiments using mammary epithelial cells EpH4, NMuMG, and MCF-10A have been shown to undergo mammary tubulogenesis in collagen or matrigel matrices following HGF stimulation (Soriano, Pepper et al. 1995; Montesano, Soriano et al. 1998; Montesano, Soriano et al. 1998). Using Eph4 cells in matrigel as a model, branching morphogenesis was determined to require PI3K signals, since inhibitors (wortmannin and LY294002) reduced their ability to form tubules (Niemann, Brinkmann et al. 1998). These signals are likely to arise from Gab1, since Eph4 cells stably expressing Gab1 were able to produce tubes in the absence of HGF stimulation (Niemann, Brinkmann et al. 1998). Additionally, Eph4 cells stably expressing HGF formed tubes in the absence of HGF, and this was blocked using an antibody directed

against HGF, showing that HGF induces the formation of ductular structures from Eph4 mouse mammary epithelial cells (Brinkmann, Foroutan et al. 1995). These data illustrate the requirement of HGF to form tubes in mammary epithelial cells, and delineates signaling pathways that mediate tubulogenesis.

Interestingly, the specific behavior of mammary luminal and myoepithelial cells during branching morphogenesis has been investigated (Niranjan, Buluwela et al. 1995; Haslam, Drolet et al. 2008). A time course analysis of HGF-induced tubule formation was performed using murine luminal and myoepithelial cells in three-dimensional collagen (Haslam, Drolet et al. 2008). Over a period of 72 hours, luminal cells were found to initiate the event by forming cellular extensions followed by the formation of chains, and finally the development of cords with two layers of luminal cells. The final stage was the hollowing of the lumen within the cords. During this process, myoepithelial cells lagged behind luminal cells during each stage of tubule development, and were found to migrate behind the leading edge of the luminal cells. Proliferation was examined by quantification of BrdU staining within these structures, and luminal cells were more proliferative then myoepithelial cells upon stimulation with HGF. Interestingly, myoepithelial cells were found by immunofluorescence to have higher levels of Met protein than luminal cells. These data are the direct opposite of an earlier report by Niranjan *et al.* (Niranjan, Buluwela et al. 1995), which isolated and examined human luminal and myoepithelial cells in separate collagen assays. Here, myoepithelial cells were found to form tubules upon HGF-stimulation, whereas luminal cells were unresponsive. Luminal cells were more proliferative than myoepithelial cells, and had higher levels of Met RNA expression. Contrasting data between the two studies could be due to differences between mouse and human tissue, however, this more likely reflects that luminal and

myoepithelial have a role together in maintaining proper epithelial cell polarity during tubulogenesis. The more recent study (Haslam, Drolet et al. 2008) might better reflect mammary biology *in vivo* as progenitor cells were likely included in their isolation of mammary epithelial cells.

1.5.2 MET in human breast cancer

MET protein is elevated in 15-20% of human breast cancers (Camp, Rimm et al. 1999) and is reported to be an independent prognostic factor of poor outcome (Ghoussoub, Dillon et al. 1998; Lengyel, Prechtel et al. 2005). Elevated MET protein levels and downstream signal transduction proteins FAK and PI3K were found to correlate with decreased survival and increased metastatic spread using tissue microarray analysis of 930 breast tumors (Garcia, Dales et al. 2007). Similarly, elevated HGF protein levels are found in the serum of breast cancer patients and is associated with a shorter disease-free interval following surgery (Yamashita, Ogawa et al. 1994) and a higher tumor/lymph node/metastasis score (TNM) (Taniguchi, Toi et al. 1995).

Although the MET receptor has been associated with poor outcome tumors, the link between MET and a specific breast cancer subtype had not previously been fully examined. A report by Charafe-Jauffret *et al.* in 2006 examined the gene expression profiles of 31 human breast cancer cell lines (Charafe-Jauffret, Ginestier et al. 2006). Among the eight cell lines classified by whole genome DNA microarrays as basal-like, MET was among a group of genes upregulated and differentially expressed between basal and luminal breast cancer cell lines. Similarly, in a proteomic screen of breast cancers, MET was found to be present in a group of basal-like tumors (Goncalves, Charafe-Jauffret et al. 2008). Recently, Wu *et al.* performed immunohistochemistry for basal

markers on ten primary breast cancers and metastases from different organs of the same patients (Wu, Fackler et al. 2008). Four primary breast cancers and corresponding metastases were found to have a basal phenotype, being ESR1/PGR/ERBB2-negative and staining positively for basal markers such as epidermal growth factor receptor, EGFR, keratins 5/6 and notably exhibited high MET protein levels (Wu, Fackler et al. 2008). Furthermore, in a study involving inflammatory breast cancers, which are highly aggressive poor outcome tumors with a 5-year survival rate of 3-50%, all 41 cases were MET positive, compared to 480 control tumor tissues which were negative (Garcia, Dales et al. 2007). When molecular subtypes were analyzed, it was found that inflammatory breast cancers fall into basal and ERBB2-positive groups as determined by immunohistochemistry staining. These data support a role for the MET receptor in aggressive basal, and poor outcome breast tumors.

1.5.3 Met/HGF mouse models of breast cancer

The Met receptor was cloned in the mid-1980s, and the first mouse models of Met and HGF pathway-induced mammary tumorigenesis were developed a decade later. Several different mouse models with deregulation of the Met receptor and HGF have produced mammary tumors (Table 1.1). These models have been examined to a limited extent, however, from the data available, general consistencies can be found. Mouse models whose transgenes are under the control of hormonally regulated promoters (i.e.: WAP, MMTV-LTR) generally require multiple rounds of pregnancy and lactation to induce tumor formation (Table 1.1) (Cardiff and Muller 1993). Notably, several models have been generated using the ubiquitously expressed metallothionein (MT) promoter (i.e.: MT/Tpr-Met, MT/HGF, MT/NK1, MT/M1248T) as well as a model that expresses

an oncogenic variant of Met from the endogenous *Met* locus (Met M1248T/L1193V, Carrie Graveel, personal communication) in an FVB/N genetic background (Liang, Reid et al. 1996; Takayama, LaRochelle et al. 1997; Jakubczak, LaRochelle et al. 1998). Data from these models suggests that the mammary epithelium is naturally susceptible to tumorigenesis by disruption of the Met/HGF pathway. Furthermore, mammary tumors tend to form at a long latency and low to moderate penetrance (Table 1.1) when compared to those of the well-established MMTV/ErbB2 or MMTV/PyMT models (Guy, Cardiff et al. 1992; Guy, Webster et al. 1992). These data suggest that other genetic events are required for the generation of Met/HGF pathway-induced mammary tumors.

Where described in detail, another common feature among Met/HGF mouse models of breast cancer is the production of tumors with multiple pathologies including a squamous component (Table 1.1). The MT/Tpr-Met model, induced tumors with papillary, scirrhous and nodular solid pathologies. This suggests that deregulation of the Met signaling pathway targets an early progenitor cell that results in transformation to these different phenotypes. In support of this, WAP/HGF tumors showed features of squamous metaplasia and stained positively for basal marker keratin 6 (Gallego, Bierie et al. 2003). Mammary adenocarcinomas generated by Welm *et al.* by infecting primary mammary epithelial cells from inducible Met transgenics with a modified stem cell retrovirus that expressed Myc resulted in mammary tumors that displayed not only keratin 14 positive cells but expansion of the progenitor population, as determined by keratin 6 staining (Welm, Kim et al. 2005).

1.5.4 The MMTV/Met^{mt} mouse model of breast cancer

In an effort to understand the role of the Met receptor specifically in the mammary epithelium, we have generated a mouse model that expresses wild type and weakly oncogenic variants of Met receptor under the expression of the MMTV-LTR promoter (Table 1.1, Figure 1.8) (Petkiewicz 2007). In addition to the wild type Met receptor, Met receptor variants included an activating mutation in the kinase domain originally identified in hereditary papillary renal cell carcinoma (PRCC), M1248T (Schmidt, Duh et al. 1997), a juxtamembrane domain mutation, Y1003F, which results in a loss of the binding site for the E3 ubiquitin ligase c-Cbl (Peschard, Fournier et al. 2001), as well as a combination of these two mutations, Y1003F/M1248T (see the following section for biochemical and biological aspects of the Met receptor variants) (Figure 1.8). Our laboratory has previously shown that the Y1003F mutation results in a loss of receptor ubiquitination, degradation and results in prolonged signaling (Peschard, Fournier et al. 2001; Abella, Peschard et al. 2005). MMTV/Met^{mt} mammary tumors were generated at long latency and low penetrance, however the latency was significantly reduced and penetrance was significantly increased in the Met Y1003F/M1248T model (Petkiewicz 2007). We termed the mouse models generated by the M1248T and Y1003F/M1248T mutations as Met^{mt} and examined these further as they represented the highest incidence of tumor formation among the Met models overall (Figure 1.8). Altogether, this data demonstrates a direct role in Met/HGF pathway-induced mammary tumorigenesis.

1.5.5 *In vitro* and *in vivo* data on the Met mutations used to generate the MMTV/Met^{mt} model of breast cancer

The Y1003 site is the tyrosine residue within the juxtamembrane domain of the Met receptor responsible for binding to the tyrosine kinase-binding (TKB) domain of the

E3 ligase Cbl (Peschard, Fournier et al. 2001). Cbl is a negative regulator of several receptor tyrosine kinases, including Met, and acts to downregulate signaling from the receptor by targeting Met for degradation through clathrin-mediated endocytosis (Abella, Peschard et al. 2005). Cbl mediates ubiquitination of the Met receptor, acting as a post-translational modification that marks the receptor as cargo for lysosomal degradation in the endocytic pathway (Abella, Peschard et al. 2005). Substitution of a tyrosine to a phenylalanine residue (Y1003F) converts the receptor to an oncogenic Met receptor variant (Peschard, Fournier et al. 2001) (Figure 1.8). Cells expressing the Y1003F Met receptor variant were transforming *in vitro* and *in vivo* as evidenced by focus-forming assays and tumorigenesis assays in immunocompromised nude mice (Abella, Peschard et al. 2005). In comparison to the wild type Met receptor, the Y1003F Met mutant was internalized and underwent endosomal trafficking with comparable kinetics, however, was inefficiently degraded. As a result, the Met Y1003F mutant was retained in an endocytic compartment within the cell with sustained signaling through the ras/MAPK signaling pathway. Although a Met Y1003F mutation as not yet been found in human tumors, a comparable mutation in non-small cell lung cancer (NSCLC) has recently been found where the exon coding for the Y1003 is deleted (delta exon 14) (Kong-Beltran, Seshagiri et al. 2006).

The Met M1248T mutation was originally identified in hereditary and sporadic human renal cell carcinoma (Schmidt, Duh et al. 1997). The methionine to threonine substitution occurs within the catalytic region of the kinase domain, specifically in the p + 1 activation loop (Maritano, Accornero et al. 2000) (Figure 1.8). This mutation was found to be tumorigenic as determined by focus forming assays *in vitro* and tumor growth assays in nude mice (Jeffers, Schmidt et al. 1997; Jeffers, Fiscella et al. 1998; Nakaigawa,

Weirich et al. 2000). Notably, KIT and RET receptor tyrosine kinases were found to exhibit similar mutations of this conserved residue, illustrating its importance as a mechanism of oncogenesis. Overall, the Met M1248T mutation has been shown to result in increased tyrosine phosphorylation through increased kinase activity (Jeffers, Schmidt et al. 1997; Maritano, Accornero et al. 2000). One study suggested that the M1248T substitution leads to altered substrate specificity compared to wild type Met (Yuan, Guan et al. 2004). Furthermore, our laboratory has generated a Met Y1003F/M1248T double mutant receptor and investigated its effects (Peschard 2005; Petkiewicz 2007) (Figure 1.8).

From biochemical analyses, cells stably expressing the Met M1248T showed similar receptor degradation kinetics and tyrosine phosphorylation status in response to HGF when compared to the wild type Met receptor. In contrast, both the Y1003F and Y1003F/M1248T Met receptor variants exhibited comparable delays in receptor degradation and showed sustained tyrosine phosphorylation (Petkiewicz 2007). Downstream signaling pathways where sustained for the Met Y1003F and Y1003F/M1248T receptor mutants compared to wild type and Met M1248T Met receptors as evidenced by prolonged ras/MAPK (phospho-ERK1/2, phospho-MEK1/2), and PI3K/AKT (phospho-AKT) signaling (Petkiewicz 2007). In biological assays, mouse mammary epithelial cells (Eph4) expressing the Met receptor variants where injected into the mammary fat pads of nude mice and tumor growth was measured over time. Although the Met Y1003F and Met M1248T receptor variants were weakly oncogenic, the combination of both activating and loss of negative regulation mutations of the Met Y1003F/M1248T double mutant receptor showed increased tumor growth kinetics (100% penetrance, 70 day latency)(Petkiewicz 2007).

1.6 Mouse models

1.6.1 The utility of mouse models in studying breast cancer

Multiple mouse models have been designed to study the contribution of specific signaling pathways to mammary tumorigenesis. As discussed, breast cancer is a heterogeneous disease that arises as a result of multiple genetic and epigenetic insults. Since studies are performed in a uniform genetic background, animal models allow the dissection of individual signaling pathways that contribute to tumor initiation and development that would not be possible from clinical studies alone (Bild, Yao et al. 2006; Ursini-Siegel, Schade et al. 2007). Furthermore, mouse models of mammary intraepithelial neoplasia have been developed that mimic human premalignant breast lesions both morphologically and biologically (Medina 2000). This greatly facilitates the elucidation of genetic and epigenetic contributors in the step-wise progression from a benign lesion to a malignant tumor, an area poorly understood in humans. Thus, although animal models do not perfectly recapitulate human breast cancer in all aspects, they nevertheless provide an understanding of the disease that would not be possible in any other system.

1.6.2 FVB/N mouse strain

The vast majority of mouse models of breast cancer have been developed using the FVB/N mouse strain. These mice were originally derived from an outbred colony of Swiss mice N:GP (NIH General Purpose mouse) in 1935 at the National Institutes of Health. In 1966, two strains were selected and further developed based on their sensitivity and resistance to histamine following sensitization to the *Bordetella pertussis* (the bacteria that causes Whooping cough) vaccine. At the time, children received prophylactic vaccination, however a small portion experienced severe adverse

anaphylactic reactions when challenged with histamine (Kind 1953). These mice were used as a model to further investigate this adverse effect as they mimicked the phenotype. Thus, the two populations of these mice were developed based on their sensitivity or resistance to histamine and termed histamine sensitivity factor sensitive (HSFS/N) and histamine sensitivity factor resistance (HSFR/N). As part of a strain survey, which were routinely carried out, the eighth inbred generation from the HSFS/N line was found to carry the *Fv-1*^{b} allele, which confers sensitivity to the B strain of Friend leukemia virus; a virus of interest to cancer research. Mice were inbred to obtain homozygosity for the *Fv-1*^{b} allele and maintained without selection for histamine sensitivity and this line was termed "FVB/N" (Friend virus B-type susceptibility/NIH).

The FVB/N mouse strain has become of widespread use in transgenic experiments because of its well-defined inbred background, vigorous reproductive performance and consistently large litter size. Moreover, fertilized FVB/N eggs contain large and prominent pronuclei, which facilitate microinjection of DNA (Taketo, Schroeder et al. 1991). FVB/N mice have very low susceptibility to spontaneous mammary carcinoma in nulliparous animals (Nieto, Shyamala et al. 2003). One study identified that 6 mammary carcinomas were found in 5 animals from a cohort of 500 wild type FVB/N mice (Nieto, Shyamala et al. 2003). Two were identified as adenosquamous carcinomas, three were squamous cell carcinomas, and one adenocarcinoma. Tumor development occurred at various ages in these mice ranging from 7 to 13 months of age. Mammary hyperplasias occurred more frequently within these animals (Nieto, Shyamala et al. 2003). FVB/N mice were reported to have increased incidence of mammary hyperplasia with increasing age and parity. By 13 months of age, 40% of animals exhibited lobuloalveolar hyperplasia, and this incidence increased with age (Wakefield, Thordarson et al. 2003).

Hyperplasias were correlated with pituitary anomalies, specifically proliferating, prolactin-secreting lesions of the pituitary gland. In mice aged 18 to 23 months, hyperplasia of the pituitary and adenomas were observed in 52% and 19% of mice, respectively. Notably, when pituitary anomalies were present, this correlated with high incidence of spontaneous mammary tumors in aged multiparous mice (4/6 mice affected) (Wakefield, Thordarson et al. 2003). Thus, background phenotypes must be taken into account when studying transgenic manipulation of FVB/N mice through the use of age-matched, and parity-matched controls. Furthermore, pituitaries from aged multiparous animals should be checked for pituitary hyperplasia or adenoma.

When studying phenotypes generated by transgenic mouse models, care must be taken to compare phenotypes within animals of the same genetic background. The first mouse model of ErbB2 overexpression in the mammary epithelium used an activated variant of the ErbB2 rat homologue, Neu-NT, under the transcriptional control of the MMTV promoter in FVB/N mice (Muller, Sinn et al. 1988). This produced mammary tumors with a latency of only 3 months in virgin animals. When the same transgene with the same promoter was overexpressed in BALB/c mice, there was a significant delay in tumor latency, with tumors appearing between 5 to 10 months of age and after multiple rounds of pregnancy (Bouchard, Lamarre et al. 1989). This illustrates the importance of the genetic background in mouse model studies.

1.6.3 Mouse model designs: transgenic and knockout strategies

The majority of genetic alterations in human breast cancers fall in two categories, oncogenes and tumor suppressor genes, whose gain and loss of function, respectively, contribute to tumorigenesis. Oncogenes are involved in supplying signals implicated in

cell growth, proliferation and survival while tumor suppressor genes provide a mechanism to prevent unrestricted cell proliferation. Thus, there are two main methods to study the role of a gene in mammary gland tumorigenesis using animal models (Hutchinson and Muller 2000). Generally, transgenic mice are used to study gain of function of a specific gene in target tissues, while knockout mice are used to study genes ablated by homologous recombination. A combination of these techniques may also be used to create bigenic animals using the knockin and conditional tissue specific knockout animals.

Several promoters have been used in the creation of transgenic mice that target expression of the transgene to the mammary epithelium. The two most commonly used promoters are the mouse mammary tumor virus long terminal repeat (MMTV-LTR) or the whey acidic protein promoter (WAP). The mouse mammary tumor virus is a retrovirus and when infected in mice results in mammary carcinoma. Due to its specificity for the mammary epithelium, its promoter has been used to assist in targeting the expression of genes of interest to the mammary gland (Cardiff and Muller 1993). The MMTV-LTR promoter transcribed high levels of transgene expression throughout mammary development with an increase in pregnancy (Wagner, McAllister et al. 2001). Transgene expression is initially detected 6 to 22 days post partum and is expressed by both ductal and alveolar luminal cells as well as myoepithelial cells (Wagner, McAllister et al. 2001). Although expression is mainly restricted to the mammary epithelium, other tissues have been reported to exhibit variable levels of expression, such as salivary glands, oocytes, epidermis, B and T cells (Wagner, McAllister et al. 2001). In comparison to the MMTV-LTR, the WAP promoter is transcriptionally active only during mid-pregnancy. The MMTV-LTR promoter is responsive to progesterone, prolactin,

androgen, glucocorticoid, and dihydrotestosterone, but not to estrogen (Otten, Sanders et al. 1988; Glover and Darbre 1989; Haraguchi, Good et al. 1992; Inazawa, Tanabe et al. 2001). The MMTV-WAP promoter is responsive to estradiol-17 beta and insulin, and to a lesser extent progesterone and corticosterone, but not to testosterone, hydrocortisone and prolactin (Inuzuka, Yamanouchi et al. 2001; Booth, Boulanger et al. 2007). Although both promoters have been used to generate the majority of animal models of breast cancer, like any tool they have some limitations. Since both are hormonally responsive, it can be a challenge to uncouple oncogene-induced biological effects from steroid hormone-induced biological effects within the mammary gland. However, gene expression and histopathological analyses of tumors generated from the same transgene but under the control of different promoters are in fact similar, suggesting that overall, the transgene has a greater effect on the resulting tumor than the promoter (Rosner, Miyoshi et al. 2002; Herschkowitz, Simin et al. 2007). Another limitation is that multiple copies of the transgene are randomly integrated into the mouse genome and thus, multiple lines of transgenic mice should be used to rule out that the phenotypes are not a result of a specific integration site. Transgenics are made by pronuclear injection of DNA into a fertilized zygote, and this method gives rise to random integration (Nagy, Gertsenstein et al. 2003). In an effort to overcome these drawbacks, the 'knockin' technology was developed, where transgenics express the transgene from their endogenous promoter. This is achieved through the use of homologous recombination of the transgene at their endogenous loci (Nagy, Gertsenstein et al. 2003).

To study the loss of function of tumor suppressor genes in mammary tumorigenesis, knockout technology was developed. Here, endogenous genes are altered by homologous recombination in embryonic stem cells (Nagy, Gertsenstein et al. 2003).

One problem with this approach is that depending on the target gene, the viability and survival of the mice may be affected, precluding analysis of mammary tumor biology. For example, knockout mice of the *BRCA1* gene, which is a DNA repair gene strongly correlated with familial and heritable breast cancer, produced knockout mice that were embryonic lethal (Liu, Flesken-Nikitin et al. 1996). As a solution to this problem, conditional knockout animals were developed using the Cre-Lox system (Sauer 1998). Briefly, the target gene is flanked by *loxP* sites, which are DNA sequences recognized by the Cre recombinase enzyme from temperate phage P1. By expressing Cre downstream from a hormonally regulated promoter such as MMTV or WAP, Cre recombinase-mediated excision of the target gene occurs in the mammary epithelium. It should be noted that there are other methods of producing genetically engineered mice, however the basic strategies have been discussed here. Furthermore, mammary tumorigenesis is a multi-step process involving numerous genetic insults within the cell. Transgenic, knockin, and knockout approaches can be combined to study multiple cellular aberrations in tumor progression. For example, TP53 is the most commonly altered gene in human breast cancer either by deletion or inactivation (Elledge and Allred 1994), and interbreeding mice expressing oncogenes with Trp53 knockout animals allows for an understanding of the collaboration between the genes in mammary tumorigenesis (Xu, Qiao et al. 2001).

1.6.4 Pathway pathology

Drs. Robert D. Cardiff and Phil Leder first reported the concept of "pathway pathology" from extensive analysis of genetically-engineered mouse mammary tumors that are currently archived in the University of California Davis Mutant Mouse Pathology

Archives (Cardiff, Sinn et al. 1991; Rosner, Miyoshi et al. 2002). This archive holds data from over 3000 genetically-engineered mouse mammary tumors and allows for comparison of histological phenotypes induced by various oncogenes and tumor suppressor genes.

Hematoxylin and eosin-stained sections from ErbB/ras pathway-induced tumors were compared to Wnt pathway-induced tumors. It was found that mammary tumors induced by a receptor tyrosine kinase (i.e.: *ErbB2*), its surrogate (i.e.: polyoma virus middle T, *PyMT*), and *ras*, induced tumors with solid nodular pathology (Rosner, Miyoshi et al. 2002). These tended to form solid nodules consisting of poorly differentiated cells with abundant cytoplasm, scarce stroma and a lack of myoepithelial or squamous differentiation. On the other hand, tumors induced by the Wnt, myc, glycogen synthase kinase 3-β (GSK3β), β-catenin pathways produced a very different histological phenotype. This was characterized by a combination of myoepithelial, acinar, or glandular differentiation, squamous metaplasia, presence of branched tubular structures, well-developed stroma, inflammatory infiltration, and pushing margins. Altogether, these findings demonstrated that different transgenes that induce similar signaling pathways give rise to tumors with similar pathology. Thus, the expression of transgenes that induce the ras/MAPK pathway produce solid nodular tumors, while Wnt pathway tumors produce a heterogeneous, mixed pathology phenotype.

Interestingly, gene expression profiling analyses of various mouse models of breast cancer have been consistent with the concept of pathway pathology. Unsupervised hierarchical clustering showed that models with similar genotypes express similar patterns of gene expression and thus cluster together, and these tended to have similar

phenotypes (Desai, Xiao et al. 2002; Herschkowitz, Simin et al. 2007). Using this unbiased analysis further illustrates that genotype predicts phenotype. In these analyses, mouse models of breast cancer were compared with human breast cancers (Herschkowitz, Simin et al. 2007). It was found that mouse models with solid nodular phenotypes (i.e.: ErbB2, PyMT) were most similar to luminal human breast cancer, while mouse models with perturbations of the Wnt, Brca1, and Trp53, pRB (i.e.: mouse model $C3(1)T_{ag}$) pathways were most similar to human basal breast cancer. Thus, the various mouse models of breast cancer are representative of the spectrum of tumors seen in human breast cancer.

1.6.5 Mouse models of basal breast cancer

As mentioned, human breast cancer is subdivided into different molecular subtypes that correlate with clinical outcome. These are the luminal, ERBB2+ and basal subtypes, where the basal subtype exhibits the worst outcome and lacks targeted therapies. Mouse models representative of these subtypes of breast cancer have been developed (Andrechek, Laing et al. 2003; Herschkowitz, Simin et al. 2007).

The molecular mechanisms of human basal breast cancer are to date poorly understood. Several mouse models have provided insight into this clinical challenge. Transgenic mouse models with basal-like features are those which contain Brca1, Trp53 and pRB deficiencies or perturbation of the Wnt/ß-catenin pathway (Herschkowitz, Simin et al. 2007). Furthermore, several transgenic mouse models of breast cancer have been reported to have basal-like features either by gene expression analyses (WAP/T_{121}, MMTV/Wnt, $Brca1^{co/co}$; MMTV/Cre; $Trp53^{+/-}$, K14/Cre; $Trp53^{fl/fl}$, $Brca1^{fl/fl}$), and/or by comparative pathology (BLG/Cre; $Brca1^{F22\text{-}24/F22\text{-}24}$; $Trp53^{+/-}$) (Xu, Wagner et al. 1999;

Brodie, Xu et al. 2001; Herschkowitz, Simin et al. 2007; Liu, Holstege et al. 2007; McCarthy, Savage et al. 2007). Although mouse models deficient in Brca1 and Trp53 are useful tools to study human basal-like breast cancer with *BRCA1* germline mutations, *BRCA1* somatic mutations have not been detected in sporadic basal tumors (Futreal, Liu et al. 1994; Richardson, Wang et al. 2006). It is intriguing that mice containing diverse genetic alterations produce mammary tumors with a similar phenotype. However, there are likely other genetic events implicated in the generation of a basal-like tumor phenotype, and it is necessary to identify novel animal models of basal breast cancer to further study the molecular mechanisms of tumorigenesis.

1.7 NCK proteins

1.7.1 Introduction

Extracellular cues are in part transmitted within the cell through receptor tyrosine protein kinases and produce distinct biological outcomes. Signal transduction from receptor tyrosine kinases are mediated by adaptor proteins, which act as scaffolds to recruit multiple protein complexes to specific locations within the cell (Schlessinger 1994). Ligand-induced activation of receptor tyrosine kinases results in oligomerization and phosphorylation on tyrosine residues, creating docking sites for Src homology 2 (SH2) or phosphotyrosine-binding (PTB) domain-containing proteins. SH2 domains recognize specific phosphotyrosine residues-containing amino acid sequences (Pawson and Schlessingert 1993). This results in subsequent recruitment of signaling molecules that give rise to signal amplification, and disseminates through effector proteins to produce specific biological outcomes. There was an interest to further understand the protein molecules that coupled signals from activated receptor tyrosine kinases to

downstream effectors. Many novel molecules were identified through screens during this time. The number of signaling molecules currently exceeds 200, and can be subdivided into three categories 1) enzymes, such as Src family tyrosine kinases, 2) regulator molecules, such as the Vav family of Rho guanine nucleotide exchange factors, and 3) adaptor proteins with no catalytic activity. The NCK proteins belong to the adaptor proteins category (Buday, Wunderlich et al. 2002).

1.7.2 Identification of NCK genes

In 1990, NCK1 was cloned from a human melanoma cell line expression library using monoclonal antibodies against the melanoma-associated antigen MUC18 (Lehmann, Riethmuller et al. 1990). NCK2 was subsequently identified using a bacterial expression/cloning system called CORT (cloning of receptor targets), which was used to identify novel SH2 domain-containing proteins interacting with the phosphorylated EGFR carboxyl terminus (Margolis, Silvennoinen et al. 1992). A partial murine cDNA that revealed 74% amino acid identity with human NCK1 was found and at the time was termed growth factor receptor-bound 4, or *Grb4* (Margolis, Silvennoinen et al. 1992). However, *Grb4* and *NCK2* represent the same gene (Chen, She et al. 1998; Braverman and Quilliam 1999).

1.7.3 Structure of NCK genes

Human *NCK1* maps on chromosome 3q21, whereas human *NCK2* maps on chromosome 2q12, indicating that NCK1 and NCK2 are two different gene products (Huebner, Kastury et al. 1994; Chen, She et al. 1998). NCK1 and NCK2 contain three N-terminal SH3 domains followed by one C-terminal SH2 domain, and lack any other known enzymatic or functional motifs. SH3 domains bind to proline-rich sequences

containing hydrophobic residues and act as constitutively associated protein-recruiting sites (Cicchetti, Mayer et al. 1992). Both proteins exhibit 68% amino acid identity in their SH2 and SH3 domains, but differ in the linker regions outside of the SH2 and SH3 domains. NCK1 and NCK2 are 47-kDa and 48-kDa, respectively, and mainly cytosolic proteins although NCK1 has also been detected in the nucleus (Lawe, Hahn et al. 1997; Kremer, Adang et al. 2007).

1.7.4 Tissue expression of NCK RNA and proteins

The expression pattern of *NCK* genes have been examined in human and murine tissues however, inconsistent results regarding the expression of *Nck2* have been reported (Chen, She et al. 1998; Braverman and Quilliam 1999; Bladt, Aippersbach et al. 2003). *Nck1* was ubiquitously expressed at the mRNA and protein levels within murine spleen, thymus, brain, liver, kidney, lung, heart, placenta and muscle (Bladt, Aippersbach et al. 2003). In agreement, Braverman *et al.* found that *NCK1* mRNA was expressed in 16 different human tissues tested (Braverman and Quilliam 1999). *Nck2* mRNA was highly expressed in murine lung, skeletal muscle, kidney, and in human skeletal muscle, placenta, pancreas, and heart (Chen, She et al. 1998). Another study contradicted these findings and found elevated Nck2 protein in the murine thymus, spleen and lungs, reduced in the heart, kidney and brain, and undetectable levels in liver and skeletal muscle (Bladt, Aippersbach et al. 2003). Braverman *et al.* reported that the *NCK2* mRNA levels were fairly abundant in the 16 human tissues studied but noticeably lower in lung, liver, and kidney although their Northern blot analyses lacked loading controls (Braverman and Quilliam 1999). Overall, it can be concluded that Nck1 is ubiquitously expressed but Nck2 expression is relatively more restricted. Nck proteins have distinct

but overlapping tissue distribution, suggesting that they may have both compensatory and unique biological roles.

1.7.5 Nck proteins in development

Vertebrate and mammalian models have been generated to study the role of Nck in development. *Caenorhabditis elegans, Drosophila melanogaster* and *Xenopus laevis* have a single *Nck* gene, whereas mammals express both *Nck* family members (Chen, She et al. 1998). The majority of the phenotypes described affect dorsal-ventral development, formation of growth cone receptors, and nervous system development. These phenotypes share a common feature of aberrantly affecting the actin cytoskeleton and cell differentiation.

The *Drosophila melanogaster* orthologue of Nck is Dreadlocks (Dock). In *Drosophila*, signaling through serine/threonine kinase Pak1 mediates proper targeting of the photoreceptor growth cones and axonal guidance (Garrity, Rao et al. 1996; Hing, Xiao et al. 1999). Further studies showed that the association of Dock with Misshapen, an Nck-interacting serine/threonine kinase, contributes to photoreceptor axon targeting (Ruan, Pang et al. 1999; Su, Maurel-Zaffran et al. 2000). The function of Nck in vertebrates was addressed using dominant inhibitory SH2 and SH3 domain mutants, and Nck was implicated in dorsoventral axis development during embryogenesis in *Xenopus laevis* (Tanaka, Lu et al. 1997).

In order to address the physiological functions of *Nck1* and *Nck2* in mammalian development, *Nck1*$^{-/-}$ and *Nck2*$^{-/-}$ knockout mouse models were generated (Bladt, Aippersbach et al. 2003). Loss of either *Nck1* or *Nck2* alone did not result in any gross morphological or functional abnormalities, suggesting that Nck proteins have redundant

and compensatory functions in the mammalian system (Bladt, Aippersbach et al. 2003). However, loss of both *Nck1* and *Nck2* in doubly mutant animals resulted in embryonic lethality at day 9.5 (Bladt, Aippersbach et al. 2003). Defects were found in mesoderm structures, and specifically in the notochord. Compared to littermate controls, $Nck1^{-/-}$ $Nck2^{-/-}$ embryos suffered from a developmental delay, were noticeably smaller, exhibited a lack of chorion allantoic fusion (resulting in a lack of chorioallantoic placenta and thus a lack of embryonic circulation), impaired closure of the cephalic neural folds, and deficient axial rotation. Axial rotation is a process that begins at embryonic day 8.5 and is necessary to revert the orientation of the germ layers. Based on these data, Nck proteins are required for this event to occur. Thus, Nck proteins play a necessary role in mammalian embryogenesis, and specifically, in the formation of mesoderm-derived structures.

The same group subsequently generated *Nck1* and *Nck2* conditional knockout animal models to examine the role of Nck proteins in the nervous system development since embryonic lethality precluded analysis in the $Nck1^{-/-}Nck2^{-/-}$ knockout mice (Fawcett, Georgiou et al. 2007). Animals were generated using a conditional allele for *Nck2* (floxed), and a null allele for *Nck1*. These animals were crossed with Nestin-Cre transgenic animals to target excision of *Nck2* to the mammalian nervous system, specifically in neuronal and glial cell precursors. Animals with inactivation of both *Nck1* and *Nck2* showed a hopping gait phenotype as early as postnatal day 2, a similar phenotype as that observed in EphA4- and ephrinB3-null animals (Fawcett, Georgiou et al. 2007). The hopping gait phenotype was attributed to a defect in the spinal central pattern generator, which controls rhythmic flexor and extensor muscles needed for walking. This defect was explained by synchronous excitation from bilateral ventral

motor neurons. Mice also showed abnormal projections of corticospinal tract axons where axons aberrantly decussated the spinal midline at the thoracic level, and impaired development of the posterior tract of the anterior commissure.

1.7.6 NCK-mediated signal transduction

A role for NCK as signaling adaptor proteins was initially identified downstream from multiple receptor tyrosine kinases including EGF, and PDGF receptors (Chou, Fajardo et al. 1992; Li, Hu et al. 1992; Meisenhelder and Hunter 1992; Park and Rhee 1992). Notably, NCK was phosphorylated on tyrosine following stimulation with EGF and PDGF (Li, Hu et al. 1992; Meisenhelder and Hunter 1992; Park and Rhee 1992). However, in the absence of growth factor-induced phosphorylation, NCK was basally phosphorylated on serine (Li, Hu et al. 1992).

NCK proteins act by coupling tyrosine-phosphorylated proteins through SH2 domains to downstream effectors through SH3 domains that bind proline-rich motifs, effectively recruiting multi-protein complexes to localized areas within the cell (Figure 1.9) (Li and She 2000). NCK proteins has been shown to associate with various tyrosine phosphorylated growth factor receptors, including EGFR (Li, Hu et al. 1992; Park and Rhee 1992), PDGFR (Nishimura, Li et al. 1993), vascular endothelial growth factor receptors, VEGFR1 and VEGFR2 (Ito, Wernstedt et al. 1998), EPHB1 (Stein, Huynh-Do et al. 1998), RET (Bocciardi, Mograbi et al. 1997), as well as tyrosine phosphorylated proteins DOK1 (Holland, Gale et al. 1997), DOK2 (Jones and Dumont 1998), and P130CAS (Rivera, Antoku et al. 2006) and others (Buday, Wunderlich et al. 2002). The SH2 domain of NCK proteins preferentially binds phospho-YDXV motif (Jones, Blasutig et al. 2006; Fawcett, Georgiou et al. 2007). Numerous SH3 domain effector proteins have

been identified to interact with NCK proteins (Buday, Wunderlich et al. 2002). Several have a unified role in regulating actin polymerization including PAK1 (Galisteo, Chernoff et al. 1996), N-WASP (Rivero-Lezcano, Marcilla et al. 1995), and WASP-interacting protein (WIP) (Anton, Lu et al. 1998). PAK1 is activated by GTP-bound CDC42 and Rac, and transmits signals to mitogen-activated protein kinase family members c-JUN amino terminal kinase (JNK) and p38 (Minden, Lin et al. 1995). Furthermore, NCK proteins have been found to bind N-WASP, which is a proline-rich CDC42/RAC-binding cytoskeletal protein (Rivero-Lezcano, Marcilla et al. 1995). GTPases Rac and CDC42 activate N-WASP, WIP, and WAVE, and they interact with the actin related proteins ARP2 and ARP 3 (ARP2/3) to induce actin nucleation (Jaffe and Hall 2005). Altogether, these data implicate NCK downstream from CDC42/RAC signaling pathways (Figure 1.9).

1.7.7 Differences in NCK1- and NCK2-interacting proteins and signaling pathways

Although NCK proteins perform similar functions, they exhibit some differences in signaling and biology. The phosphotyrosine binding specificities of NCK1 and NCK2 SH2 domains have been recently studied and found to be very similar (Frese, Schubert et al. 2006). Thus, the SH2 domain of either protein cannot account for these differences. Nevertheless, there are examples where NCK1 and NCK2 are found to bind to different phosphotyrosine sites on the same phospho-protein. In the case of the PDGFR, NCK2 binds to a different phosphotyrosine residue (Y1009) than NCK1 (Y751) (Nishimura, Li et al. 1993; Chen, She et al. 2000). In addition, some NCK2-specific binding partners have been identified, namely, EPHRINB1 (Cowan and Henkemeyer 2001; Bong, Park et

al. 2004), EPHRINB2 (Su, Xu et al. 2004), Disabled-1 (DAB1) (Pramatarova, Ochalski et al. 2003), which are involved in neuronal signaling.

Furthermore, the SH3 domains of NCK proteins bind overlapping yet distinct proline-rich effector molecules (Buday, Wunderlich et al. 2002). In addition to those previously discussed (PAK1, N-WASP) other SH3-binding partners in common to NCK1 and NCK2 include CBL (Rivero-Lezcano, Sameshima et al. 1994; Wunderlich, Goher et al. 1999), FAK (Choudhury, Marra et al. 1996), VAV (Ramos-Morales, Druker et al. 1994), ABL (Ren, Ye et al. 1994; Smith, Katz et al. 1999), and BCR-ABL (Wunderlich, Goher et al. 1999; Coutinho, Jahn et al. 2000), among others (Buday, Wunderlich et al. 2002). Notably, putative NCK2-specific interacting proteins have been found and include the adaptor proteins PINCH (Tu, Li et al. 1998), IRS1 (Tu, Liang et al. 2001), paxillin (Guan, Chen et al. 2007), and guanine nucleotide exchange factor DOCK180 (Tu, Kucik et al. 2001) (Figure 1.9).

An increasing number of studies have described non-overlapping roles for NCK1 and NCK2 and these appear to be cell-type dependent. Guan *et al.* showed that NCK2 but not NCK1 was required for NGF-induced neuritogenesis in an *in vitro* mammalian model (Guan, Chen et al. 2007). In this system, NCK2 was required for proper neuronal cell differentiation. This involved paxillin levels, such that with NCK2 knock down, paxillin levels concomitantly decreased and resulted in a block in differentiation (Guan, Chen et al. 2007). Interestingly, the relationship between NCK2 and paxillin was not found in the majority of cell lines from various origins, including mammary breast cancer epithelial cells (MCF7), fibroblasts, and endothelial cells (Guan, Chen et al. 2007).

A specific role for NCK2 in bridging growth factor and integrin signaling has been reported (Figure 1.9). The SH2 domain of NCK2 can bind directly to

phosphotyrosine on a receptor tyrosine kinase or indirectly through adaptor protein IRS1, while the third SH3 domain interacts with PINCH (Tu, Li et al. 1998). PINCH (particularly interesting new cysteine-histidine proteins) contains five cysteine-rich consensus sequences called LIM domains, which mediate protein-protein interactions (Rearden 1994). Among the binding partners of PINCH is integrin-linked kinase (ILK), a serine/threonine kinase involved in cell adhesion and mammary tumorigenesis downstream from integrin engagement (Tu, Li et al. 1999; White, Cardiff et al. 2001). Further investigation is necessary to understand the physiological significance of this pathway, and the synergistic value of integrating growth factor and integrin signaling branches, although data from knockout studies imply a role in focal adhesion formation (Li, Zhang et al. 1999; Legate, Montanez et al. 2006).

1.7.8 NCK-mediated cellular biology

Many of the signaling pathways involving NCK result in biological outcomes that relate to actin cytoskeleton reorganization. Actin cytoskeletal rearrangements are necessary for several cellular processes including cell shape change, motility, and intracellular vesicle trafficking. These are mediated in part by coupling tyrosine phosphorylation induced by extracellular signals with downstream regulators of actin dynamics. Mouse embryonic fibroblasts (MEFs) derived from *Nck1*$^{-/-}$*Nck2*$^{-/-}$ embryos have been used extensively to study the biological roles of Nck. The initial studies using double knockout MEFs showed defects in migration as assessed by *in vitro* wound healing assays, but no change in proliferation. This defect was partially rescued by re-expression of Nck1 (Bladt, Aippersbach et al. 2003). Actin cytoskeletal rearrangement defects were investigated as an explanation for the defective migratory phenotype, and

Nck1$^{-/-}$*Nck2*$^{-/-}$ MEFs formed truncated and irregularly-shaped lamellipodia, with actin filaments appearing shorter, less branched and sparser in the area immediately after the leading edge (Bladt, Aippersbach et al. 2003).

Subsequent studies used *Nck1*$^{-/-}$*Nck2*$^{-/-}$ MEFs and examined other biological defects resulting from loss of Nck proteins. It was found that Nck1 and Nck2 were required for signaling from the PDGF receptor for cytoskeletal organization and chemotaxis upon stimulation with ligand PDGF-B (Rivera, Antoku et al. 2006). Under these conditions, Crk-associated protein, p130Cas, was the major SH2-domain phosphoprotein associating with Nck to mediate these biological effects and both were seen to colocalize in membrane ruffles induced by PDGF-B. In response to PDGF-B, Nck and p130Cas deficient cells failed to show cytoskeletal rearrangements that are normally seen in wild-type cells, such as membrane ruffling, and disassembly of actin bundles. In a recent study using the same cells the individual contribution of Nck1 and Nck2 toward membrane ruffling, specifically dorsal ruffling, migration, chemotaxis, cell adhesion and focal adhesion formation were examined in response to PDGF-B (Ruusala, Pawson et al. 2008). The presence of both Nck1 and Nck2 were required for dorsal ruffle formation, however Nck2 was found to rescue ruffling to a greater degree than Nck1. Nck knockout cells exhibited reduced chemotactic and migratory activity following PDGF-B stimulation, and reduced cell attachment properties following serum stimulation. These biological effects were attributed to defective Rho GTPase signaling, and indeed Rac1, and RhoA were implicated. In addition to lamellipodia formation, Nck proteins have also been shown to mediate invadopodia formation through Cdc42 downstream from EGF stimulation (Yamaguchi, Lorenz et al. 2005), and filopodia formation during cell attachment downstream from integrin signaling (Antoku, Saksela et al. 2008).

1.7.9 Transforming properties of NCK proteins

Nck was shown to have oncogenic potential when Nck-expressing NIH3T3 fibroblasts were subjected to focus forming assays. Since Nck2 had not been identified at the time, these assays were performed with Nck1, and it is currently unknown whether Nck2 would behave similarly. Using N-*ras* as a positive control, the efficiency of focus formation induced by Nck was approximately 20-fold lower than focus formation induced by N-*ras*. Interestingly, elevation of Nck protein levels was sufficient to induce cellular transformation of fibroblasts grown in serum (Li, Hu et al. 1992), and tumor formation in mice (Chou, Fajardo et al. 1992). These criteria allowed the Nck gene product to be called an oncoprotein, albeit weaker than the oncogenic potential of v-*src* and N-*ras* (Chou, Fajardo et al. 1992).

At the time, it was hypothesized that Nck-overexpression would promote increased levels of phosphotyrosine proteins when compared to parental cells based on the proposed mechanism of cell transformation by the *v-crk* oncoprotein which shares a similar SH2/SH3 domain structure with Nck. Although phosphotyrosine levels were not elevated in Nck-transformed cell lines it was proposed that Nck altered tyrosine kinase pathways within the cells (Chou, Fajardo et al. 1992), although this remains controversial (Braverman and Quilliam 1999). From these data, and consistent with the finding that either Nck1 or Nck2 alone are weakly oncogenic, Nck proteins can cooperate with other signaling molecules to promote cellular transformation.

1.7.10 NCK proteins and human disease

NCK proteins have been shown to have a role in infection resulting from Enteropathogenic *Escherichia coli* (EPEC), which is a major cause of severe infant diarrhea, and vaccinia virus. Both pathogens use NCK proteins to assist in transducing

pathological signals within the cell. EPEC attaches to the membrane surface of intestinal enterocytes and inserts the translocated intimin receptor, Tir, into the host membrane, which acts as a means to induce the remodeling of cellular microfilaments and eventual suppression of microvilli (Gruenheid, DeVinney et al. 2001). Clustering of Tir proteins induces phosphorylation on its cytoplasmic tail and recruitment of NCK proteins (Phillips, Hayward et al. 2004). Through NCK-mediated pathways, actin polymerization is induced resulting in protrusions at the cell membrane known as pedestals (Rosenshine, Ruschkowski et al. 1996).

A crucial role for NCK proteins has also been shown for a form of renal disease known as congenital nephrotic syndrome of the Finnish type (NPHS1) (Jones, Blasutig et al. 2006; Verma, Kovari et al. 2006). Within kidney glomeruli, the exterior of capillary walls are enveloped in actin-rich foot processes called podocytes, which are extensions of adjacent kidney epithelium (the cells of which are called podocytes). Between the podocyte foot processes are spaces that allow for the filtration of waste and toxins from the blood to be passed into the urinary system. The waste passes through a "sieve", which is a specialized intracellular junction known as the slit diaphragm (Pavenstadt, Kriz et al. 2003). When the integrity of the slit diaphragm is compromised, unwanted blood proteins, and nutrients pass through the kidneys resulting in disease. A mutation in nephrin, which is a transmembrane adhesion molecule, is one example of a genetic defect that gives rise to this syndrome (Kestila, Lenkkeri et al. 1998). It was found that nephrin maintains proper actin cytoskeletal organization in podocytes specifically through Nck proteins, and that the nephrin-Nck interaction is required for this process (Jones, Blasutig et al. 2006).

1.7.11 NCK proteins and human cancer

Despite its role in transformation of fibroblast cell lines, reports of the role of NCK proteins in cancer have been lacking. This is unexpected given earlier reports that NCK1 is located at 3q21, a site of frequent insertions, inversions, deletions and translocations in hematopoeitic malignancies and inherited disorders (Mitelman, Kaneko et al. 1990). The chromosomal location 2q12 of NCK2 however, has not been associated with human malignancies thus far. The role of Nck in mammary tumorigenesis has not been investigated neither by *in vitro* nor by *in vivo*-based assays although other SH2/SH3 adaptor proteins have been implicated in mammary tumorigenesis (Dankort, Maslikowski et al. 2001; Rodrigues, Fathers et al. 2005). Additionally, since many receptor tyrosine kinases are frequently deregulated in human cancers, it would be reasonable to consider that their effector proteins, such as SH2/SH3 domain containing proteins, as being implicated in an oncogenic signaling pathway, thus, NCK might be a valuable signaling node in mammary tumorigenesis and warrants further study.

Transgene	Promoter	Mouse strain	Pregnancy required	Latency (days)	Penetrance (%)	Mammary tumor histology	Reference
Tpr-Met	MT	FVB/N	yes	381	42 (11/26)	Scirrhous, papillary and solid nodular	Liang TJ *et al.* J Clin Inv 1996
HGF	MT	FVB/N	yes and no	~450	41 (11/27)	Adenosquamous and adenocarcinoma	Takayama H *et al.* PNAS. 1997
NK1	MT	FVB/N	NA	~630	5 (1/19)	NA	Jakubczak JL *et al.* MCB 1998
M1248T	MT	C57BL/6NCR x C3H/HeNCR F2	yes	~300	5 (1/21)	Adenocarcinoma	Jeffers M *et al.* PNAS 1998
HGF	WAP	FVB/N	yes	~300	89 (49/55)	Glandular adenocarcinoma and squamous metaplasia	Gallego MI *et al.* Oncogene 2003
Met	MMTV	FVB/N	yes	475	9 (2/23)	Squamous, adenomyoepithelial, secretory, papillary, spindle cell, glandular tumors	Ponzo MG *et al. under review* 2008
Y1003F	MMTV	FVB/N	yes	442	8 (2/26)	Squamous, glandular, solid nodular, spindle cell, large cell solid adenocarcinoma	Ponzo MG *et al. under review* 2008
M1248T	MMTV	FVB/N	yes	474	23 (5/22)	Adenosquamous	Ponzo MG *et al. under review* 2008
Y1003F/M 1248T	MMTV	FVB/N	yes	397	41 (16/39)	Solid nodular, scirrhous, papillary, secretory, microacinar, glandular adenosquamous, spindle cell tumors	Ponzo MG *et al. under review* 2008

NA=not available

Table 1.1 Met/HGF signaling-induced mouse models of mammary tumorigenesis.

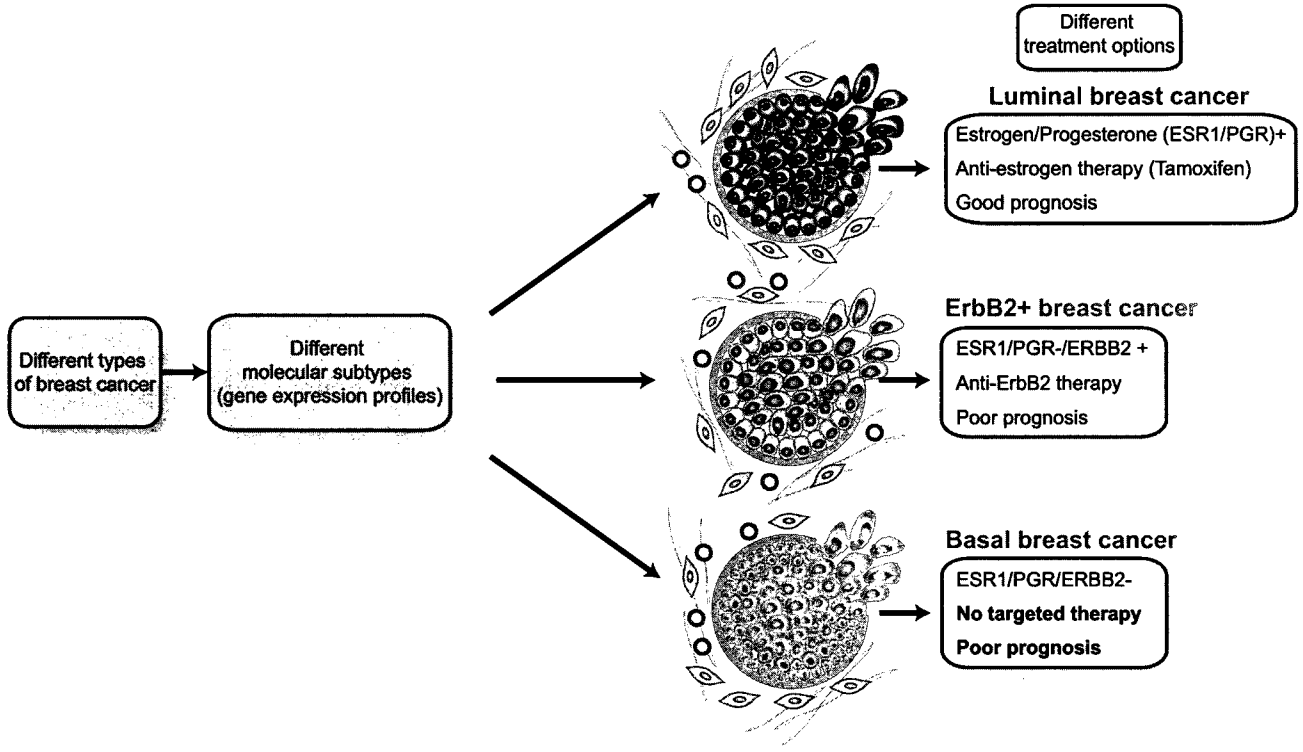


Figure 1.1: Gene expression profiling has defined luminal, ERBB2+ and basal molecular subtypes. Breast cancers are heterogenous both biologically and clinically. Gene expression profiling of human breast cancers have resolved this heterogeneity into three predominant groups, luminal (good prognosis), ERBB2+ (poor prognosis) and basal (poor prognosis). Targeted therapies exist for the luminal and ERBB2+ subtypes however, none are available for the basal subtype.

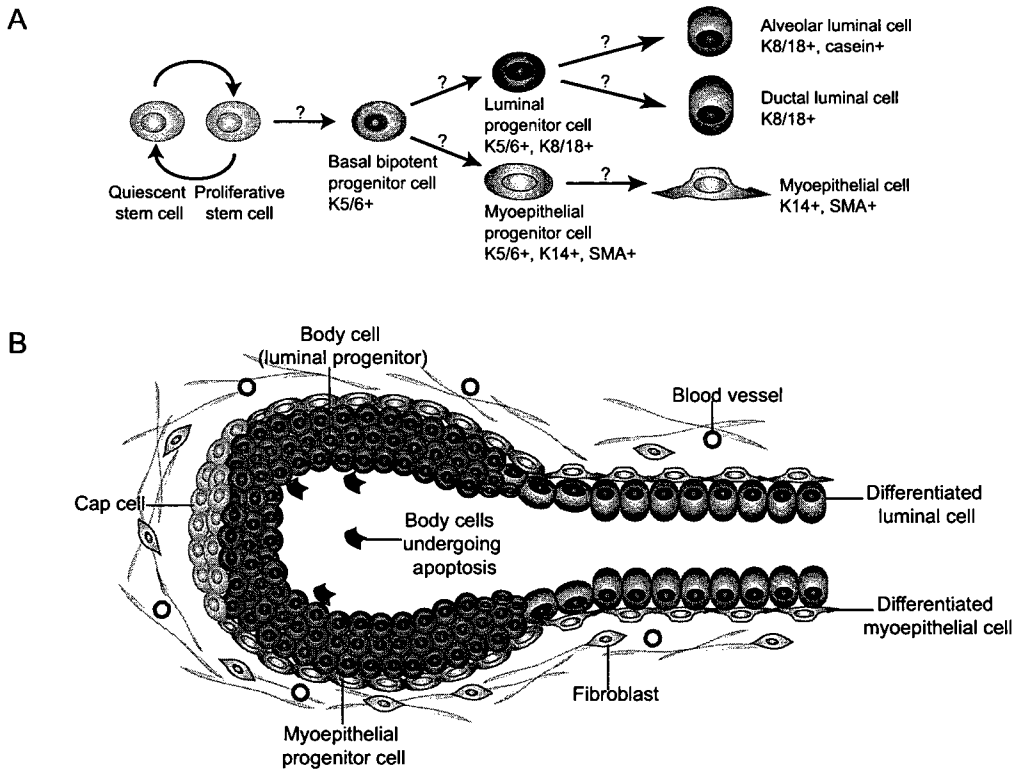


Figure 1.2: A model of the mammary epithelial hierarchy. (A) The adult mammary gland is composed of a bilayer of epithelial cells (luminal and myoepithelial), which are thought to arise from a common progenitor population. The mechanism regulating these processes are unknown (question marks). Markers of distinct cell lineage are indicated (SMA-smooth muscle actin, K-keratin). (B) The terminal end bud structure as it moves through the mammary fat pad during mammary gland development. The cap cells (orange) are thought to give rise to both luminal progenitor cells (body cells, blue) and myoepithelial progenitor cells (purple). Cap cells undergo symmetrical division to regenerate more cap cells, and asymmetrical division to produce luminal and myoepithelial progenitor cells.

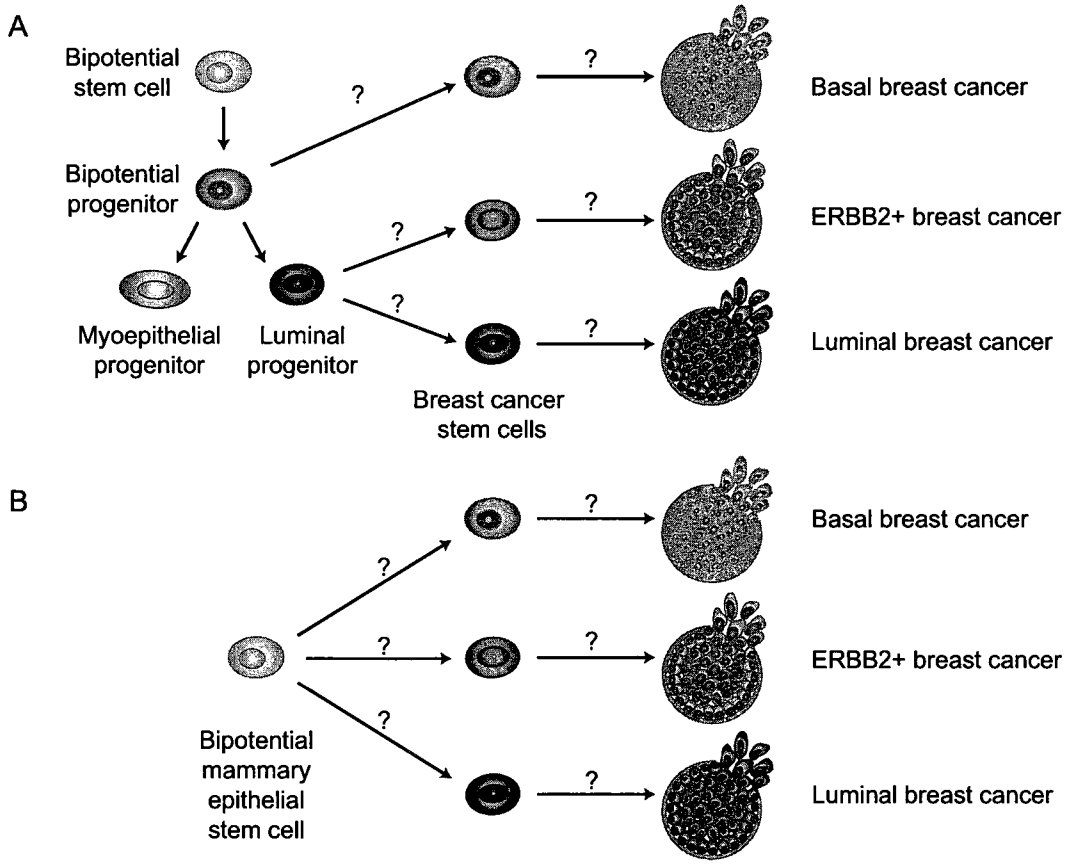


Figure 1.3: Two models of mammary tumorigenesis: cancer stem cell theory and clonal evolution. (A) In the cancer stem cell model, each molecular subtype originates from a progenitor cell that has undergone transformation. (B) In the clonal evolution model, all cells in the mammary epithelial hierarchy (not only progenitor cells) have equal ability to undergo transformation. The resulting molecular subtypes are thus determined by acquired genetic, and epigenetic events. The processes governing each step are as yet unknown (question marks).

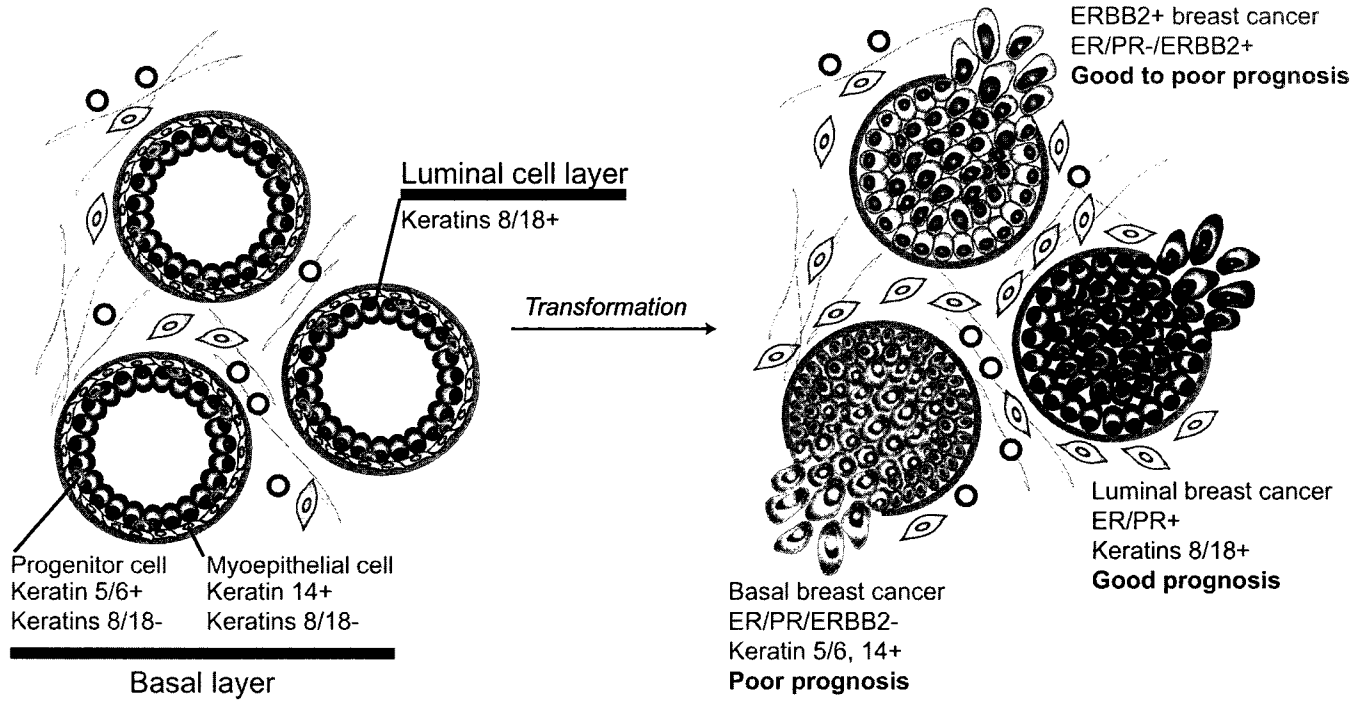


Figure 1.4: Molecular subtypes of human breast cancer arise from the transformation and expansion of distinct cell lineages within normal epithelium. While luminal (blue) and ERBB2+ (purple) breast cancers are thought to arise from the luminal cell lineage, basal breast cancers (red) are thought to arise from the 'basal' layer of the normal breast epithelium which is composed of progenitor and myoepithelial cells.

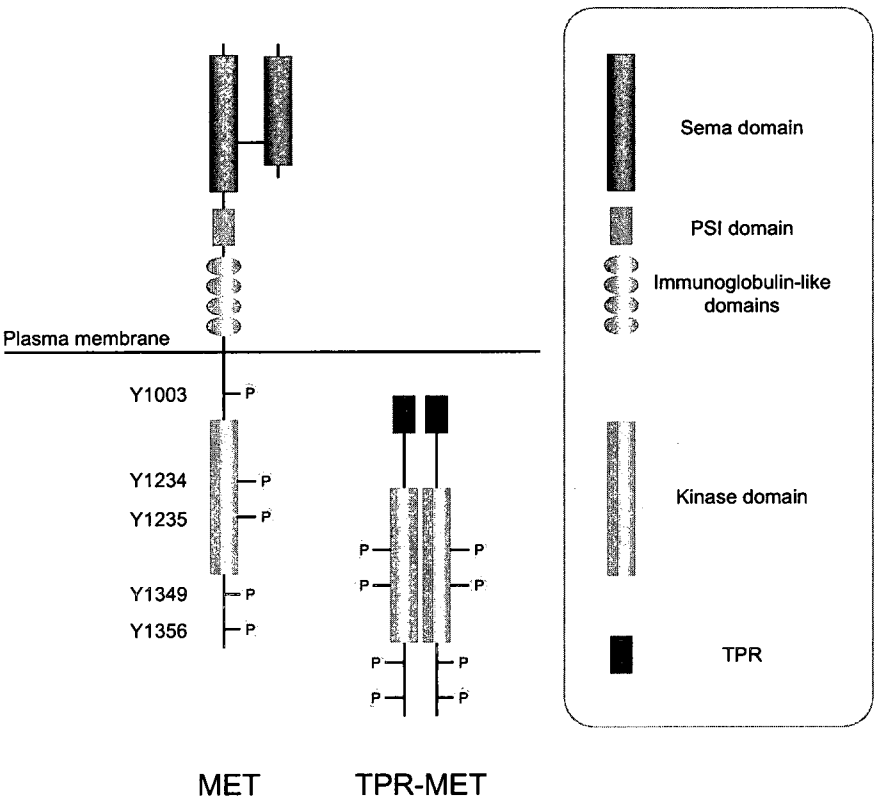


Figure 1.5: The MET receptor tyrosine kinase and TPR-MET. The extracellular domain of the MET receptor contains a Sema, plexin/semaphorin/integrin (PSI), and immunoglobulin-like domains, while the intracellular domain contains the juxtamembrane, kinase and C-terminal domains. TPR-MET is an oncogenic fusion protein that joins the *TPR* locus with the intracellular portion of *MET*.

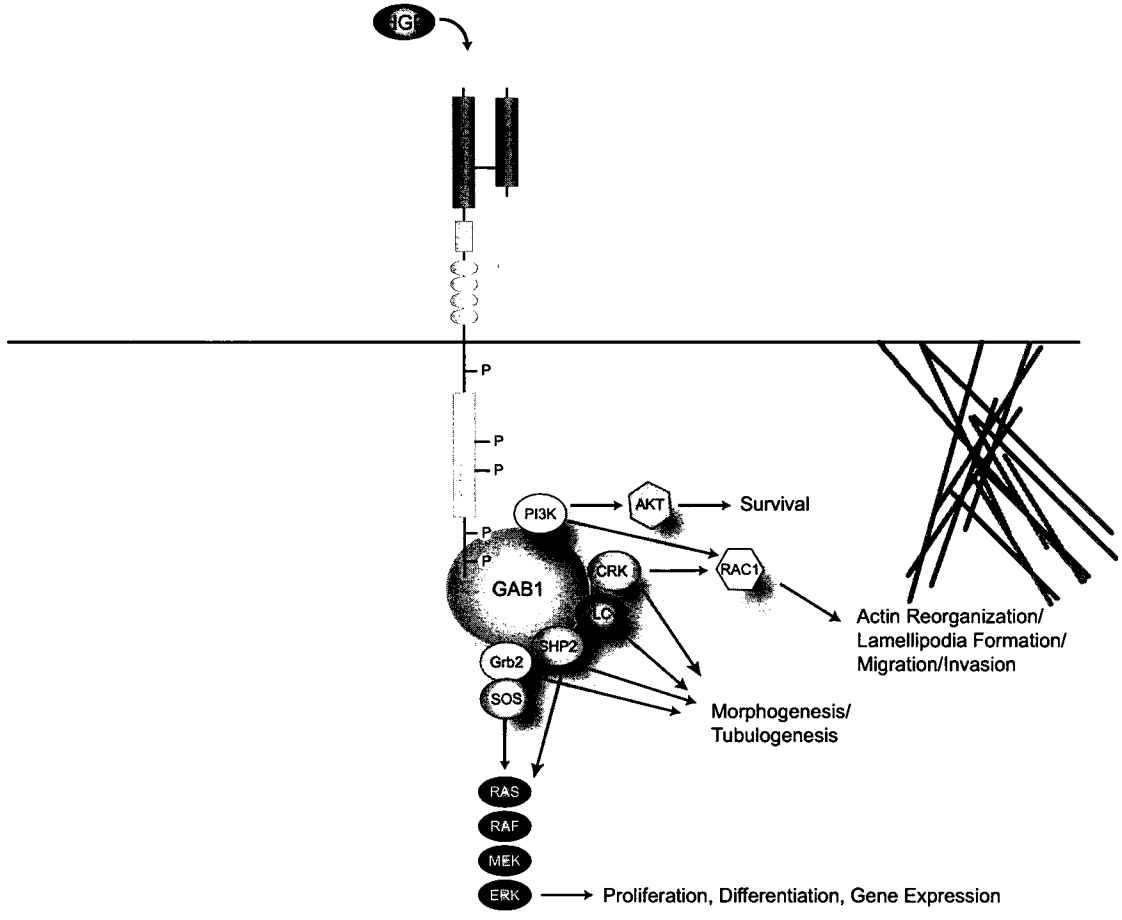


Figure 1.6: MET/HGF signaling axis. Upon activation of the MET receptor, phosphorylation occurs on tyrosines which results in the recruitment of SH2-domain-containing signaling proteins. The major phosphoprotein to interact with MET is GAB1, which acts as a scaffold to recruit other signaling proteins and activate a number of signaling pathways such as ras/MAPK, PI3K/AKT, and Rac1. Signaling pathways result in the biological responses indicated in the figure.

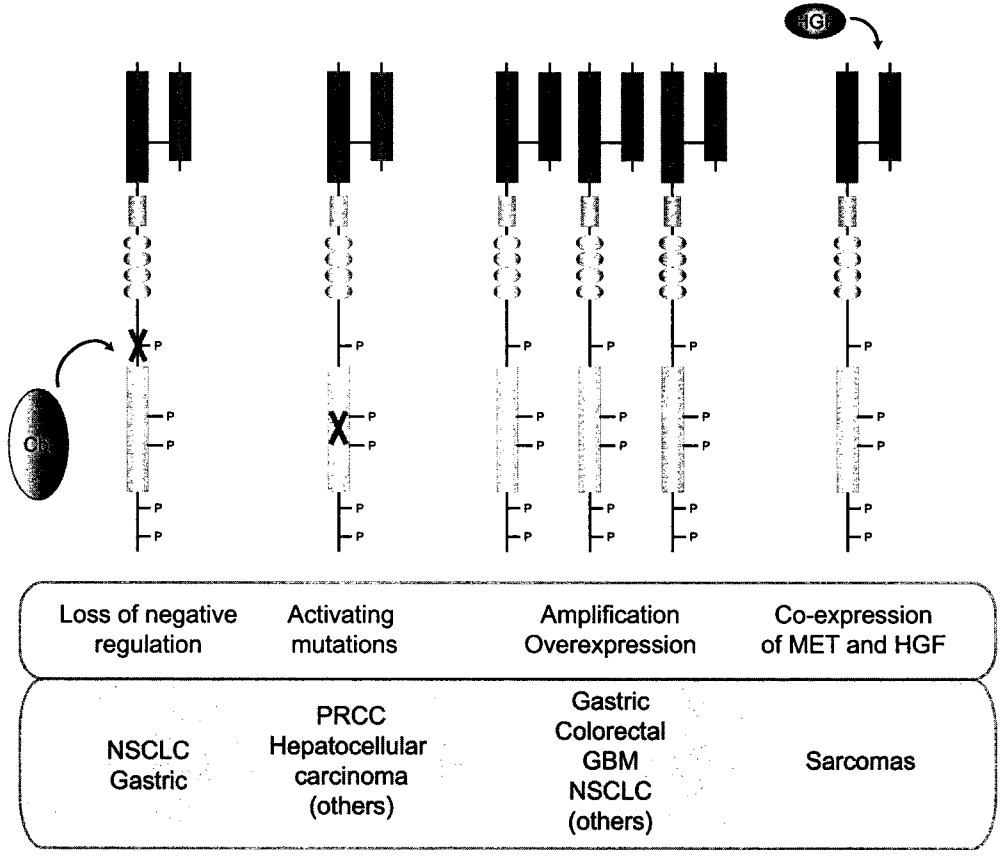


Figure 1.7: Mechanism of MET oncogenesis in human cancers. Overexpression of the *MET* receptor as observed in breast cancer, is the predominant mechanism of MET deregulation in human cancers. Deletion of exon 14 results in a loss of the Cbl-binding site Y1003, and a loss in downregulation of the receptor, which is observed in non-small cell lung cancers (NSCLC). Activating mutations within the kinase domain were originally found in hereditary and sporadic renal cell carcinoma (PRCC). *MET* is frequenty amplified in gastric, and colorectal cancers. MET amplification has recently been identified as a mechanism of acquired resistance in NSCLC and glioblastoma multiforme (GBM) brain tumors.

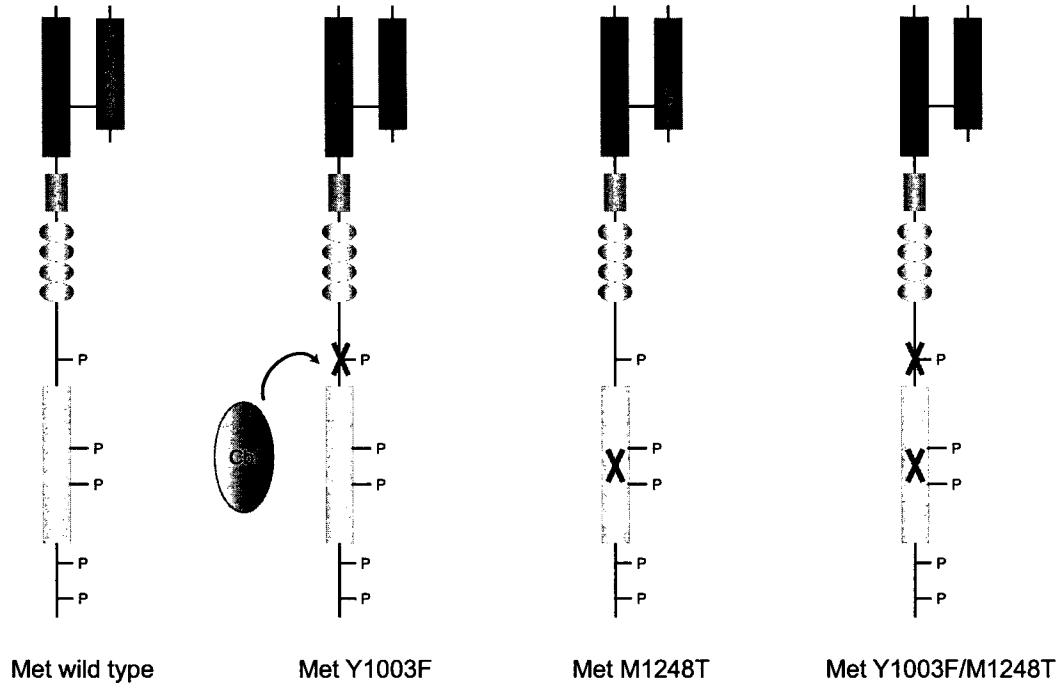


Figure 1.8: Met receptor variants used to generate the MMTV/Met mouse model of breast cancer. Oncogenic variants of the Met receptor (wild type, Y1003F, M1248T, and Y1003F/M1248T) were overexpressed under the transcriptional control of the MMTV promoter in the FVB/N mouse strain (see text for details). Note that "MMTV/Met^{mt}" refers to Met M1248T and Y1003F/M1248T models combined (Chapter 4).

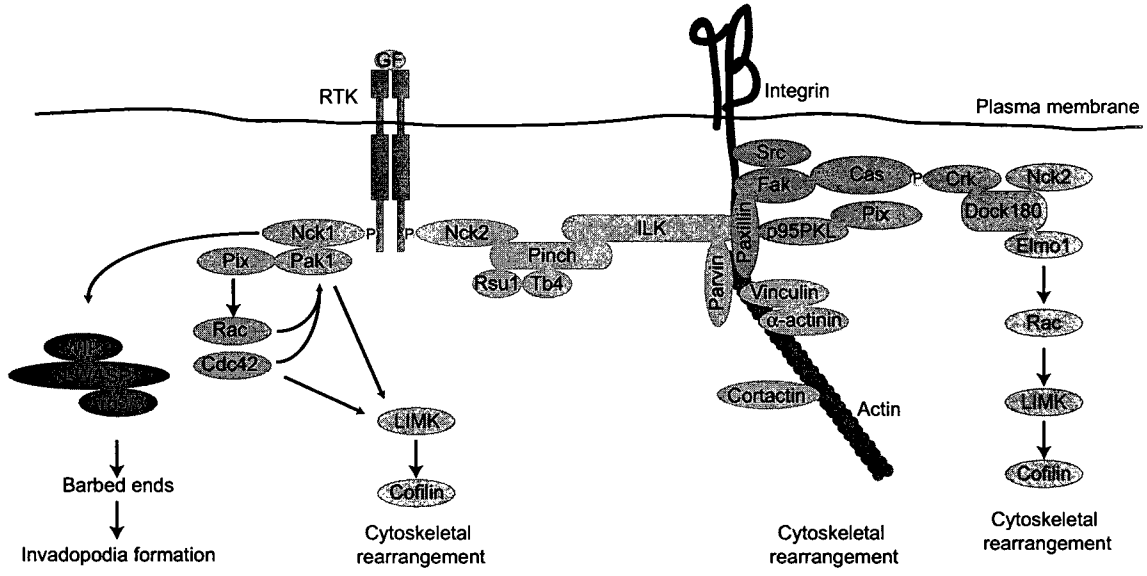


Figure 1.9: Nck-mediated signaling pathways. Nck1 and Nck2 are SH2/SH3 adaptor proteins which couple growth factor receptor and integrin signals to actin cytoskeletal rearrangement. Biological outcomes include filopodia, lamellipodia, and invadopodia formation which impact cell-matrix contacts, migration and invasion. Nck has been implicated in receptor-mediated endocytosis. Signaling pathways are highlighted by different colors, however, these are not necessarily mutually exclusive. RTK, receptor tyrosine kinase; GF, growth factor.

Chapter 2

2 Materials and Methods

2.1 Materials

Rabbit polyclonal antibodies anti-Nck1 #2383 and anti-Nck2 #3313, pan-Nck #1698 and cDNA constructs GFP-NCK1, GFP-NCK2, HA-NCK1, and HA-NCK2 were kind gifts from Dr. Louise Larose. Met antibody 148 was raised against a carboxy-terminal peptide of the human Met protein as previously described (Rodrigues, Naujokas et al. 1991). Monoclonal Met19S antibody was a kind gift from Dr. George Vande Woude.

2.2 Transgenic mice

MMTV/Met mice were generated as described (Petkiewicz 2007). The 8.7 kb fragment containing the Met cDNA, MMTV promoter and SV40 polyA was injected into the pronuclei of FVB/N zygotes and implanted into FVB/N hosts by the McIntyre Transgenic Core Facility (Montréal, Canada). Mice were housed in accordance with McGill University Animal Ethics Committee guidelines.

2.3 Precautions used while working with RNA

All glassware used for RNA work had been baked in an oven at 210°C for a minimum of 48 hours. Pipette tips were either purchased pre-sterilized, or were autoclaved. Eppendorf tubes were autoclaved. Prior to working with RNA, all work surfaces (centrifuges, bench top, tube racks, etc.) were treated with RNase*Zap* (Ambion) and rinsed twice with RNase-free water. All solutions that required preparation were done so with RNase-free water using oven-baked glass graduated cylinders and bottles. All solutions were freshly-dispensed and never reused. Protocols were carried out as quickly

as possible while respecting the incubation times for each step. Disposable gloves were worn and changed frequently throughout procedures. Prior to staining, all glass baths were treated with 3 % H_2O_2 for at least 10 minutes and rinsed twice with RNase-free water to deactivate RNases.

2.4 Cryopreservation and cryosectioning of mouse mammary tissue

Upon sacrifice of the animal, mouse mammary tumor and matched-normal tissues were embedded in 1 ml TissueTek OCT (Somagen, Edmonton, Alberta, Canada) in cryovials (Nalgene) and flash frozen in liquid nitrogen. Samples were stored at -80°C or in liquid nitrogen until use. Frozen specimens were cryosectioned using the Microm HM505E or Microm HM550 (Thermo Fisher Scientific) cryostats. Mouse mammary tumors were cut at a temperature of -25°C and mouse mammary normal tissues were cut at -33°C. Tissues were removed from cryovials by slight warming of the vial to facilitate removal of tissue-embedded OCT with forceps. The tissue was placed in the cryostat and mounted on a pre-cooled specimen stage located in the fast-freezing station with a thin layer of OCT. OCT was added to the specimen to create an even cutting surface and placed in the specimen holder. The OCT block was first trimmed at a thickness of 20 μm until the tissue was reached. Sections were cut at a thickness of 9 to 10 μm, and mounted on electrostatically adherent, non-coated glass slides (Thermo Fisher Scientific). Slides were then placed in pre-cooled (on dry ice) slide boxes and subsequently placed within the cryostat. Approximately 50 slides (1 section per slide) were prepared for mouse mammary normal tissue, and 25 slides (1 section per slide) for mouse mammary tumor tissue. Slides were labeled with the mouse identification number, slide number, and were dated using a pencil within the cryostat. Slides were stored at -80°C until use for laser

capture microdissection and were always transported with an adequate amount of dry ice to prevent thawing and freezing of the sections. To minimize contamination between samples, cutting blades were replaced, waste was vacuumed and all surfaces were cleaned with acetone.

2.5 H&E staining (non-RNase-free conditions)

Slides numbered 1, 25 and 50 for normal mouse mammary tissue, and slides numbered 1 and 25 for mouse mammary tumors were subjected to H&E staining under non-RNase-free conditions to analyse tissue histology. Frozen sections were thawed evenly at room temperature on a glass surface for 30 seconds, fixed in 70 % ethanol for 30 seconds, rinsed in single-distilled water for 30 seconds, and stained in Harris hematoxylin (Surgipath) for 2 minutes. Hematoxylin was vacuum-filtered prior to each use using a 0.22 μm filter (Steritop system, Millipore) to remove precipitate. Slides were rinsed in water for 30 seconds and blued in 0.3 % ammonium hydroxide (Thermo Fisher Scientific) solution for 30 seconds, followed by subsequent rinsing in single-distilled water for 30 seconds. Sections were immersed in eosin (Surgipath) for 2 minutes, followed by 60-second dehydration steps in 95 % ethanol (two changes), and 100 % ethanol. Sections were cleared in two changes of xylene for 5 minutes each, air dried in the fumehood for 20 minutes and mounted on glass slides with Permount (Thermo Fisher Scientific). Slides were examined by clinical pathologists (Karim Khetani, Atilla Omeroglu) to determine regions of morphologically tumor and normal epithelium suitable for LCM. Using slides numbered 1, 25 and 50 as a guide, slides numbered 2 to 24 (for mouse mammary tumors) and slides numbered 26 to 49 (for mouse mammary normal tissue) were used for LCM.

2.6 H&E staining (short protocol) under RNase-free conditions for LCM

Care was taken to discourage drying of the sections during the staining process. Harris hematoxylin (Surgipath) was filtered prior to use using a 0.22 μm filter (Steritop system, Millipore) to remove precipitate. Slides were thawed at room temperature on an RNase-free glass surface to ensure even thawing for a maximum of 30 seconds. Slides were fixed with 70 % ethanol for 30 seconds. Following two quick dips in RNase-free water, slides were submerged in hematoxylin for 30 seconds. Slides were rinsed with a quick dip in RNase-free water, followed by bluing in 0.3% ammonium hydroxide solution for 30 seconds. Slides were submerged for two 30-second periods in 70 % ethanol followed by 95 % ethanol prior to staining with eosin (Surgipath) for 20 seconds. Slides were dehydrated to xylene through 30-second steps in 95 (twice), and 100 % ethanol and two final 1-minute, and 1.5-minute dehydration steps in xylene. Slides were air-dried for 10 minutes in a fumehood and placed in a slide box within a dessicator, next to a 4°C-cold block. Under these conditions, RNA quality was not significantly degraded during the 3-hour period in which LCM was performed (Chapter 3).

2.7 H&E staining (long protocol) under RNase-free conditions for LCM

This protocol was used by our laboratory for the staining of human breast tissue for LCM (Finak, Sadekova et al. 2006). Harris hematoxylin (Surgipath) was filtered prior to use using a 0.22 μm filter (Steritop system, Millipore) to remove precipitate. Slides were thawed at room temperature on an RNase-free glass surface to ensure even thawing for 30 seconds (or longer for larger sections). Slides were fixed in 70 % ethanol for 30 seconds. Following a 30 second rinse in RNase-free water, slides were submerged in hematoxylin for 45 seconds. Slides were rinsed for 30 seconds in RNase-free water, followed by

bluing in 0.3% ammonium hydroxide for 30 seconds. Slides were submerged for two 60-second periods in 70 % ethanol followed by 95 % ethanol prior to staining with eosin (Surgipath) for 45 seconds. Slides were dehydrated to xylene through three 60-second steps in 95, 95, and 100 % ethanol and two final 5-minute dehydration steps in fresh xylene. Slides were air-dried for 20 minutes in a fumehood and placed in a slide box within a dessicator, next to a 4°C-cold block until use.

2.8 HistoGene staining protocol for LCM

This protocol was performed as directed by manufacturers (User Guide: "HistoGene LCM frozen section staining kit", Molecular Devices). Slides were thawed evenly on a glass surface for 30 seconds, followed by fixation in 75 % ethanol for 30 seconds and rinsing in RNase-free water for 30 seconds. The HistoGene staining solution was then applied to each section (100 µl per section) for 20 seconds on a glass surface. Slides were rinsed in RNase-free water for 30 seconds and dehydrated in ethanol solutions, 75 % ethanol for 30 seconds, 95 % ethanol for 30 seconds, and 100% ethanol for 30 seconds. Sections were cleared in xylene for 5 minutes and air-dried in a fumehood for 5 minutes. Slides were placed in a dessicator next to 4°C-cold block until use.

2.9 Methyl Green staining protocol for LCM

Methyl Green (Dako) was used as a 1% w/v solution in RNase-free water. Slides were thawed evenly on a glass surface for 30 seconds, followed by fixation in 75 % ethanol for 30 seconds and rinsing in RNase-free water for 30 seconds. Sections were stained with methyl green (100 µl per slide) for 30 seconds and rinsed in RNase-free water for 30 seconds. Sections were dehydrated through incubations with ethanol solutions, 70 % ethanol for 30 seconds, 95 % ethanol for 30 seconds, and 100 % ethanol

for 30 seconds. Sections were immersed in xylene for 5 minutes. Slides were air-dried for 20 minutes in the fumehood and placed in a dessicator next to 4°C-cold block until use.

2.10 Cresyl violet staining protocol for LCM

1 % Cresyl Violet (w/v) (Sigma) was prepared using 100% ethanol and allowed to stir overnight until completely dissolved. All solutions used in this protocol were kept on ice. Slides were thawed evenly on a glass surface for 30 seconds, fixed in 70 % ethanol for 2 minutes, placed on a pre-cooled glass plate while 1 % cresyl violet (100 µl per slide) was applied to each section. Slides were dipped twice in 70 % ethanol, followed by dipping in 100 % ethanol twice and air-dried for 2 minutes.

2.11 Total RNA extraction from frozen tissue sections

This method of total RNA extraction was used to prepare "non-stained," "H&E-stained/before LCM," and "H&E-stained/after LCM" correlative controls for LCM samples (Chapter 3). Total RNA was extracted from frozen tissue sections using TRIzol reagent (Invitrogen) according to manufacturer's protocol. All steps were performed at room temperature unless otherwise stated. A volume of 1 ml of TRIzol was used for each mouse mammary normal or tumor tissue sample. About 200 µl of TRIzol was pipetted onto the section, and by pipetting up and down over the section, the material was transferred from the slide to the TRIzol-containing tube. Total RNA from four sections were used for tumor tissue and eight to ten sections were extract from normal mouse mammary tissue. Samples were vortexed briefly to homogenize tissue, and subsequently stored at -80°C until ready to proceed with total RNA extraction. Samples were thawed and incubated at room temperature for 5 minutes to permit complete dissociation of nucleoprotein complexes. Phase separation was carried out by adding 200 µl of

chloroform (Thermo Scientific Fisher) and shaking vigorously for 15 seconds by hand, followed by incubation for 3 minutes. Samples were centrifuged at 12 000g at 4°C for 15 minutes. The upper aqueous phase containing total RNA was transferred (about 400 to 450 µl) to a new 1.5 ml-eppendorf tube. RNA was precipitated by adding 500 µl of isopropanol per sample, 2 µl of glycogen (GenHunter) to facilitate precipitation of small quantities of RNA, mixed by inversion 10 times and incubated for 10 minutes. RNA was pelleted by centrifugation at 12 000g for 10 minutes at 4°C. The supernatant was carefully removed so as to not disturb the RNA pellet. The RNA pellet was washed with 1 ml of ice cold 75 % ethanol was added, tubes were shaken, and subsequently precipitated at 7 500g for 5 minute at 4°C. Ethanol was discarded and RNA pellets were air-dried for 10 to 15 minutes until the ethanol had evaporated. Total RNA was then resuspended in 15 µl RNase-free water, and dissolution was facilitated by incubation at 55°C for 10 minutes.

2.12 Bioanalysis of RNA quality

The Agilent 2100 BioAnalyzer instrument (Agilent Technologies) and Agilent RNA 6000 Pico kit (Agilent Technologies) were used to assess the quality of total RNA or amplified RNA. The protocol was performed according to manufacturer's instructions (Agilent RNA 6000 Pico Assay Reagent Kit Guide, edition 01/2006). All procedures were performed at room temperature unless otherwise indicated. Agilent 2100 BioAnalyzer electrodes were cleaned before and after each assay to minimize contamination and ensure proper functioning of the instrument. The gel-dye mix was prepared by vortexing RNA 6000 Pico dye concentrate for 10 seconds and adding 1 µl of this to 65 µl of filtered gel, and mixing by vortexing. The gel-dye mix was centrifuged at

13 000g for 10 minutes, kept away from light until use and equilibrated at room temperature for 30 minutes. Samples were diluted with nuclease-free water to a concentration within the range of 200-5000 pg/μl for total RNA, and 500-5000 pg/μl for mRNA (or amplified RNA). RNA samples and RNA 6000 Pico Ladder were heat denatured before use for 2 minutes at 70°C and then kept on ice. Nine microliters of the gel-dye mix was pipetted into the first well labeled "G" in the RNA Pico chip. The chip was placed in the chip priming station and was pressurized for 30 seconds using the attached syringe and plunger. Gel-dye mix was pipetted in the last two wells (9 μl each) labeled "G," and 9 μl RNA 6000 Pico Conditioning solution was added to the well labeled "CS." RNA 6000 Pico marker (5 μl) was added to all 11 sample wells and in the ladder well. RNA 6000 Pico Ladder (1 μl) was added to the ladder well, and samples (1 μl each) were pipetted into the sample wells. The RNA 6000 Pico Ladder produced distinct peaks of 0.2, 0.5, 1.0, 2.0, and 6.0 kb which facilitated sizing of total RNA and mRNA. The well contents were mixed by vortexing in the IKA vortex mixer for 1 minute at the 2400 setting. Bioanalysis was performed using the Eukaryote Total RNA Pico Series II (for total RNA) or mRNA Pico Series II (for mRNA) assays within the Agilent 2100 Expert Software (Agilent Technologies).

2.13 Laser capture microdissection (LCM)

Slides intended for LCM were stored at -80°C and were less than 2 months old from the date of cryosectiong. Sections were reviewed by pathologists (K.K., A.O.) prior to microdissection using the PixCell IIe LCM system (Arcturus). Laser settings used were a spot size of 15 μm, target voltage of 0.2V, current of 4.4mA, temp rd 21.3°C, and power of 25mW. The relative humidity in the room was maintained at ≤60% with a

dehumidifier and air conditioner. Care was taken not to over-, or under-clean LCM caps (CapSure Macro LCM caps, Molecular Devices) with the CapSure Pad (Molecular Devices) following microdissection to minimize sample loss and contamination, respectively. One cap was used per slide to minimize the time that captured cells were exposed to air and moisture. 12,000-24,000 morphologically normal mammary epithelial cells and 8,000-16,000 tumorigenic epithelial cells were laser captured. Images were acquired before and after LCM, and of the captured cells on the cap to estimate cellular homogeneity. Each population was >98% "homogeneous", as determined by microscopic visualization of the captured cells and verification by pathologists (K.K., A.O.).

2.14 Total RNA isolation from LCM samples

Total RNA from each population of laser-captured cells was extracted using a phenol-chloroform protocol described by Sgroi *et al.* (Sgroi, Teng et al. 1999). LCM caps, containing transfer film and adherent epithelial cells were incubated in 200 µl GITC (guanidinium isothiocyanate) RNA extraction buffer (4 M GITC, 25 mM sodium citrate pH 7.0, 0.1 M β-mercaptoethanol, 0.5% N-lauroylsarcosine) with 1.6 µl β-mercaptoethanol for 5 minutes at room temperature. Samples were subsequently stored at -80°C until ready to proceed with RNA isolation. RNA isolation from LCM-dissected material was a two-day procedure involving two organic/aqueous phase extractions and treatment with DNaseI to digest contaminating genomic DNA. All steps were performed at room temperature unless otherwise stated. Water-saturated phenol was prepared by combining 1 ml phenol (3.75:1, v/v, Invitrogen) with 400 µl nuclease-free water (Ambion), vortexing, spinning at 12 000 rpm for 2 minutes, and letting stand for 10 minutes. For the first organic/aqueous phase extraction, 20 µl 2 M sodium acetate pH 4.0,

220 µl water-saturated phenol and 60 µl chloroform-isoamyl alcohol (23:1) were added to the extraction buffer. Following a 15-minute incubation on ice and centrifugation (12 000 rpm, 30 minutes), the aqueous phase was transferred to a new 1.5-ml tube and RNA was precipitated with 2 µl glycogen (10mg/ml, GenHunter), 200 µl isopropanol and mixed by inversion. Samples were placed on dry ice for 30 minutes, followed by centrifugation at 4°C (12 000 rpm) for 30 minutes to pellet RNA. Although the RNA pellet was not always visible at this step, the supernatant was carefully discarded so as not to disturb it, and RNA was washed twice with 400 µl ice cold 70% ethanol and centrifuged at 4°C (12 000 rpm) for 5 minutes. Following air-drying for 10 to 15 minutes to evaporate residual ethanol, samples were resuspended in 15 µl nuclease-free water and heated at 55°C for 10 minutes to facilitate dissolution of the RNA pellet. RNA was subjected to DNaseI treatment. Per reaction, samples were mixed with 1 µl RNase inhibitor (Invitrogen), 2 µl DNaseI (Roche) and 2 µl 10x reaction buffer (GenHunter) which was incubated at 37°C for 2 hours. Following DNaseI treatment, samples were either placed at -80°C or on dry ice overnight to continue with the protocol the next day, or used immediately for re-extraction. Samples were re-extracted as described above to further purify the RNA and remove residual protein contamination from the DNaseI treatment, and were resuspended in 10 µl nuclease-free water and stored at -80°C.

2.15 Total RNA quantification of LCM samples

Total RNA from LCM samples was quantified using RiboGreen RNA Quantitation Reagent (Molecular Probes) as directed by manufacturer except for minor changes. A standard curve was generated using concentrations of 0, 20 pg/µl, 0.1 ng/µl, 0.25 ng/µl, 0.5 ng/µl, and 1.0 ng/µl in 1x TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5), which

spanned a lower range than the "low range" assay described by the manufacturer. Instead of using ribosomal RNA (16S and 23S rRNA from *E. Coli*) for the standard curve, we used total RNA derived from a tumor specimen that underwent RNA isolation using the same protocol as that used for the LCM samples. RiboGreen reagent was diluted 200-fold into 1x TE and 10 µl was added to LightCycler capillaries (Roche) in duplicate. Samples were diluted 20-fold into 1x TE with Ribogreen reagent (1 µl sample, 9 µl 1x TE, 10 µl of 200-fold diluted Ribogreen reagent) and added to LightCycler capillaries in duplicate. As an external control, an RNA sample of known concentration (total RNA isolated from HeLa cell lines, 5 ng/µl) was used with each assay. To control for background RFU readings that could be introduced from organic solvents during RNA isolation, three negative control samples (0, 00, 000) containing no RNA were subjected to RNA isolation and an average of these readings was subtracted from each RFU value obtained from samples. RFU values were obtained from the LightCycler instrument (Roche) using the LightCycler3 software (Roche), and an average of the duplicate measurements was used.

2.16 RNA amplification of LCM samples

Two rounds of linear T7-based RNA amplification was performed using the Amino Allyl MessageAmp kit (Ambion) as directed by the manufacturer with minor changes. Although an input quantity of 100 ng to 2000 ng total RNA is recommended by manufacturers, we demonstrated that good RNA yield and quality could be obtained from 4 ng of total RNA. All steps were performed at room temperature unless otherwise stated. RNA amplification reactions were carried out in a Biometra $T_{Gradient}$ thermocycler (Biometra, Montréal Biotech Inc, Montréal, Québec, Canada). Reverse transcription of

mRNA was carried out by first annealing T7 oligo(dT) primers (1 µl) to mRNA (4 ng in a volume of 11 µl) at 70°C for 10 minutes, followed by submerging samples on ice. First strand cDNA synthesis was performed by adding 2 µl 10x first strand buffer, 1 µl ribonuclease inhibitor, 4 µl dNTP mix and 1 µl reverse transcriptase to each sample and incubating at 42°C for 2 hours. The T7 promoter primer, introduced in the annealing step, was used to convert single-stranded cDNA to a double-stranded DNA template for transcription. Second strand cDNA synthesis was performed by adding 63 µl nuclease-free water, 10 µl 10x second strand buffer, 4 µl 5mM dNTP mix, 2 µl DNA polymerase and 1 µl RNase H to each sample and incubating with the thermocycler lid open at 16°C for 2 hours. The resulting cDNA was purified by the addition of 250 µl cDNA binding buffer to each sample, and subsequent filtration through an equilibrated (50 µl cDNA binding buffer for 5 minutes) filter cartridge. Samples were passed through the filter cartridge by centrifugation at 10 300g for 1 minute and flow-through was discarded. cDNA-bound to the filter was washed with 500 µl cDNA wash buffer, centrifuged at 10 300g for 1 minute, and flow-through was discarded. The samples were centrifuged again at 10 300g for 1 minute to remove traces of ethanol from the cDNA wash buffer before transferring filter cartridges to cDNA elution tubes. cDNA was eluted by the addition of 9 µl preheated (50°C) water to the center of the column, incubation for 2 minutes and centrifugation at 10 300g for 2 minutes. To ensure complete elution of cDNA, this step was repeated. The volume of eluted cDNA was taken up to 14 µl with nuclease-free water, and transferred to PCR tubes. Samples underwent *in vitro* transcription with T7 RNA polymerase, which produced antisense RNA (aRNA) from mRNA. In an overnight incubation (12 hours) at 37°C, 14 µl of cDNA sample was combined with 12 µl ATP, CTP, GTP mix (25 mM each), 6 µl UTP, 4 µl T7 10x reaction buffer and 4 µl T7 enzyme

mix. The following day, cDNA was digested by the addition to each sample of 2 µl DNaseI at 37°C for 30 minutes. Sample volume was brought to 100 µl with nuclease-free water, and aRNA purification ensued. aRNA binding buffer (350 µl) and 250 µl 100% ethanol were added to each sample and passed through a filter cartridge by centrifugation at 10 300g for 1 minute. Flow-through was discarded and 650 µl aRNA wash buffer was pipetted onto each filter, followed by centrifugation at 10 300g for 1 minute. Flow-through was discarded and filter cartridges were centrifuged again at 10 300g for 1 minute to remove traces of ethanol, and filter cartridges were subsequently transferred to aRNA collection tubes. aRNA was eluted by the addition of 50 µl preheated (50°C) water to the center of the filter, incubation for 2 minutes and centrifugation at 10 300g for 2 minutes. This step was repeated with 100 µl of preheated water. aRNA was concentrated by vacuum drying for approximately 80 minutes until the volume was reduced to 10 µl and samples were transferred to PCR tubes. To obtain enough aRNA for microarray analysis, a second round of RNA amplification was performed using methodology consistent with the first round of amplification with the exception of minor changes. The major difference was that second round primers consisting of random primer hexamers were used. Synthesis of the first strand of cDNA was initiated by the addition of 2 µl of second round primers and were allowed to anneal at 70°C for 10 minutes followed by submersion on ice. Reverse transcription was performed in the same manner, however following the 2-hour incubation at 42°C, aRNA in RNA-DNA molecules was digested by the addition of 1 µl RNase H to each sample by incubating at 37°C for 30 minutes, leaving the cDNA as a template for T7 oligo (dT) primers. To these tubes, 5 µl T7 oligo (dT) primers were annealed by incubating at 70°C for 10 minutes, and samples were placed on ice before the next step. Second strand cDNA synthesis and cDNA purification

was carried out using the same protocol as described in the first round of amplification. During the *in vitro* transcription step of the second round of amplification, a modified nucleotide, 5-(3-aminoallyl)-UTP (aaUTP) was incorporated into the aRNA. aaUTP contains a reactive primary amino group on the C5 position of uracil, used for coupling to Cy3/Cy5 dyes in the labeling step of the protocol. Specifically, *in vitro* transcription was carried out by the addition of 12 µl ATP, CTP, GTP (25 mM each), 3 µl 50 mM UTP, 3 µl 50 mM aaUTP, 4 µl T7 10x reaction buffer, 4 µl T7 enzyme mix to the 14 µl of double-stranded cDNA. Following overnight (12 hours) incubation at 37°C, cDNA was digested using DNaseI, and aRNA purification ensued in the same manner as previously described. aRNA is vacuum-dried to a volume of approximately 33 µl. Samples (559N, 559T, 562N, 562T, 567T, 3229N, 3229T, 4164N, 4164T, 4425T, 4691N, 4691T, 4695T, 5033N, 5033T, 5033-4T, 5154N, this method.

2.17 Bioanalysis of RNA following amplification

Quality inspection of aRNA using the Agilent 2100 Bioanalyzer (Agilent Technologies) was carried out as described in Section 2.12, however the mRNA Pico Series II assay was used. An average peak size of 0.2-0.5 kb was acceptable to proceed with microarray analysis.

2.18 RNA labeling with Cy3/Cy5 dyes

Amino Allyl MessageAmp kit (Ambion) was used for dye coupling to aRNA as directed by manufacturer. The reactive primary amino group on the aaUTP in aRNA was chemically coupled to N-hydroxysuccinimidyl ester-derivitized reactive dyes (NHS ester dyes), specifically, CyDye Post labeling Reactive Dyes (Amersham Biosciences). Half of the aRNA yield per sample was coupled to mono-reactive Cy3 (5 µg aRNA), and the

other half (5 μg aRNA) to Cy5 dye. All protocol steps were performed at room temperature. One package of Cy3 and Cy5 dye was used per sample. To each Cy dye tube was added 4 μl DMSO and was centrifuged briefly to collect the dye at the bottom of the tubes. Cy dyes were kept away from light until needed. Coupling buffer (9 μl) was added to each aRNA sample and mixed thoroughly by pipetting. Cy dyes (4 μl) were transferred to the samples, vortexed and spun down. The coupling reaction took place by incubating for 30 minutes away from light. Hydroxylamine reagent (4M, 4.5 μl) was added to each sample and mixed well by vortexing, followed by brief centrifugation. Samples were incubated for 15 minutes away from light. The total volume was increased to 100 μl by the addition of 75.5μl nuclease-free water. The Cy dye-coupled aRNA was purified using the same protocol described in Section 2.16 for aRNA purification, was eluted twice with a volume of 55 μl nuclease-free water, and concentrated by vacuum drying to a final volume of 25-30 μl. Efficiency of dye incorporation was assayed using the Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific) under the "microarray" setting within the Nanodrop 1000 version 3.1 software. Dye incorporation were well within the acceptable range recommended by manufacturers (20-50 dye molecules per 1000nt). Percentage of labeled RNA recovery was between 80 and 90%.

2.19 Preparation of the universal mouse reference for microarray experiments

Universal mouse reference total RNA (Stratagene, ID #740100) was used as a common reference for microarray gene expression profiling experiments. Universal mouse reference RNA was provided as a solution of 70 % ethanol and 0.1 M sodium acetate. It was prepared for use by centrifuging at 12 000g for 15 minutes at 4°C, removing the supernatant, washing the pellet in 70 % ethanol, followed by centrifugation

at 12 000g for 15 minutes at 4°C, drying the RNA pellet at room temperature for 30 minutes and resuspending in RNase-free water to a final concentration of 1 μg/μl. Universal mouse reference underwent one round of linear T7 amplification (Amino Allyl MessageAmp kit) in a single batch to minimize batch effect bias, using an input of 10 μg total RNA per reaction and the amplified product was subsequently pooled and labeled with Cy Dyes. Half of the yield was labeled with Cy3 dye and the other half with Cy5 dye in one batch. Aliquots from this batch were used for all microarray experiments.

2.20 Hybridization on whole genome mouse expression microarrays

Initial samples (559N, 559T, 562N, 562T, 3229N, 3229T, 4164N, 4164T, 4691N, 4691T, 5033N, 5033T, 5156N, 5156T, 6696N, 6696T, 5154N, 5154T) were analyzed using 1 x 44k Whole Mouse Genome Expression arrays (Agilent Technologies, product G4122A), however, during the course of this study, Agilent discontinued production of these expression arrays and subsequent samples (567T, 4425T, 4695T, 5033-4T, 5482T, 5612T) were hybridized onto the 4 x 44k Whole Mouse Genome Expression arrays (Agilent Technologies, product G4122F). Agilent's Two-Color Microarray-Based Gene Expression Analysis Protocol (Version 5.5, February 2007) was used for hybridization onto 4 x44k expression arrays. A dye-swap replication design was used and each biological sample was hybridized against universal mouse reference RNA (Stratagene) that was amplified and labeled identically. The 10x blocking agent was prepared by adding 1.25 ml nuclease-free water to the tube and mixed by vortexing. To prepare the samples for hybridization on 4 x 44K arrays, 825 ng of Cy3 and 825 ng of Cy5 amplified RNA were combined in a tube with 11 μl 10x blocking agent, 2.2 μl 25x fragmentation buffer to a final volume of 52.8 μl. RNA was fragmented by incubation for 30 minutes at

60°C, the reaction was subsequently stopped by the addition of 55 µl GEx hybridization buffer Hi-RPM and gently mixing by pipetting (taking care not to introduce bubbles into the sample). Samples were centrifuged at 13 000 rpm for 1 minute and were placed on ice until ready for use. Agilent SureHyb chambers were used according to the "Agilent microarray hybridization chamber user guide" (version 1.0, August 2003), and gasket slides were placed in each. A volume of 100 µl was applied to the gasket slide (Agilent Technologies) in a "drag and dispense" manner, without making contact between the pipette tip and the glass surface. The expression array was placed with the 'active side' facing down on top of the gasket slide. The slide chambers were assembled by placing the SureHyb chamber cover over the expression array sandwich and the clamp was hand-tightened. The assembled slide chambers were placed in a hybridization oven set to rotate at 10 rpm at 65°C, for 17 hours. The following day, microarrays were washed under in ozone-free tent containing a HEPA filter that was turned on an hour before beginning. Microarray slides were disassembled in gene expression wash buffer 1 (Agilent Technologies). Once disassembled, microarray slides were placed in a slide rack submerged in gene expression wash buffer 1 over a magnetic stir plate, and were washed for 1 minute. Microarray slides were next washed in pre-warmed (37°C) gene expression wash buffer 2 (Agilent Technologies) for 1 minute. Microarrays were immediately processed for scanning by first removing the slides from the wash buffer and placing them in slide holders within the scanner carousel. Microarray slides were scanned using Agilent's dual-laser microarray scanner (Model G2505B), and Cy5/Cy3-signals were quantified using Agilent's Feature Extraction software (version 9.5.1, Agilent Technologies).

2.21 Gene Expression Microarray Data Preprocessing

All data preprocessing of Met^{mt} tumors was carried out as in Finak *et al.* (Finak, Sadekova et al. 2006). Platform-specific differences between the Agilent 1 x 44k and the 4 x 44k arrays were corrected using the Limma package (Smyth 2004) by fitting a model for replicate samples on each array type. Preprocessing of the NKI and Herschkowitz *et al.* datasets was performed by background subtracting raw feature intensities, and further normalized using spatial and intensity-dependent loess on the log2-ratios. As these dataset do not contain dye swaps, Rquantile scaling was used to normalize between arrays. Missing values were estimated using the impute package (Gentleman, Carey et al. 2004).

2.22 Gene Expression Microarray Data analysis

The significance of clusters formed during class discovery was obtained using the pvclust package (Suzuki and Shimodaira 2006) with genes that had a variance of at least 2 across all samples and using 1000 bootstrap iterations. Class distinction was performed using the non-parametric RankProd package (Hong, Breitling et al. 2006) with 100 bootstrap iterations. Overrepresented GO and KEGG categories were computed by means of a standard hypergeometric test on the set of genes overexpressed in a group of tumors, and was carried out using the GOstats package (Gentleman, Carey et al. 2004). For the Gene Set Enrichment Analysis (GSEA), genes were ranked using the program's built in algorithm, with the sample classes undergoing 1000 permutations. Symbols in the gene sets were translated from human to mouse using the biomaRt package (Mootha, Lindgren et al. 2003; Durinck, Moreau et al. 2005; Subramanian, Tamayo et al. 2005). All unsupervised hierarchical clustering was performed by randomly removing duplicate probes, using the Pearson correlation as a distance measure, and with Ward's minimum

variance agglomeration method. Integration of separate datasets was made by scaling each gene to ~N(0,1) across all samples, under the assumption that each gene in the intrinsic gene list should represent a similar distribution of values. To test the null hypothesis of gene expression populations being equal within tumor subtypes, the Kruskal-Wallis rank sum test was used, which does not make any assumptions about the distributions of the populations. Human subtype centroids were calculated by taking the mean value of each gene across the set of samples belonging to the class. Of the 15 Met^{mt} mammary tumors, we LCM-dissected 9 matched-normal tissues and determined that continued dissection of matched-normal tissue was not necessary and that a 'pooled normal' sample could be used instead. We performed class distinction analysis on tumor versus matched-normal and on tumor versus pooled normal using Limma package for R with false discovery rate <0.05. We compared differentially expressed genes from the two analyses and found that on average 68% of the genes belonged to both sets. While assessing the genes that were present in only one of the two analyses, we found that 97% were differentially expressed in at most 2 samples out of 8. The data reported in this thesis have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE10450).

2.23 Mouse and human tissue sample processing

Banking and processing of human and mouse tissue specimens have been described (Finak, Sadekova et al. 2006) respectively (Ursini-Siegel, Hardy et al. 2008). Paraffin sections were stained with H&E. For the human ANN breast cancer cases, a prospectively-ascertained consecutive series was enrolled as previously described (Andrulis, Bull et al. 1998). Written consent was obtained from all of the patients

included in the IRB approved study (Protocol title: Molecular genetic alterations in breast cancer. Mount Sinai Hospital REB number: 01-0313-U).

2.24 Tissue microarray processing and immunohistochemical analysis

Areas of invasive carcinoma were selected from an H&E-stained section of each tumor and two 0.6 mm cores of tissue were embedded in a paraffin block. For MET and SNAIL staining on human breast tissue, 4 μm sections were rehydrated to water, treated with 3% hydrogen peroxide, and microwave antigen retrieved using the Micromed T/T Mega Microwave Processing Lab Station (ESBE Scientific) for 12 min with Tris-HCl pH 9 buffer for MET, and 10 mM citrate buffer pH 6 for SNAIL. Sections were blocked with 1% horse serum (Gibco) containing Avidin (Vector labs) for 10 min. MET (1/30, Novocastra) and SNAIL antibodies (1/100, Santa Cruz) were applied for 1 hr at room temperature. For MET staining, horse anti-mouse antibodies (1/200, Vector labs) were applied for 30 min, and for SNAIL staining, the goat detection kit (R&D systems) was used as described by the manufacturer. Sections were developed with DAB and counterstained in Mayer's hematoxylin. As shown previously (Lengyel, Prechtel et al. 2005), tumor epithelium was positive for MET immunostaining (cytoplasmic and perinuclear staining), with an absence of staining in the stroma (data not shown). Each of the immunohistochemical stained sections were scored using the Allred scoring method (Mulligan, Pinnaduwage et al. 2008). The raw score data were processed using a TMA deconvoluter software program (Liu, Prapong et al. 2002).

2.25 Tissue microarray data analysis

Association of Disease-Free Survival (DFS) with MET protein status was evaluated in a large cohort of axillary node negative (ANN) patients using Kaplan-Meier plots and

Cox proportional hazards models with and without including traditional clinicopathological factors as covariates, as described previously (Andrulis, Bull et al. 1998; Mulligan, Pinnaduwage et al. 2008). Patients with tumors expressing high levels of MET (MET +; n=292) have a shorter DFS when compared to those with low MET (MET-; n=376) (Log-rank test p-value=3.94e-02). For the table in Figure 4.14 the basal group was characterized as ERBB2-/ESR1-/KERATIN 5+; Total sample size used was n=509; restricted to the dataset with ERBB2, ESR1, and KERATIN 5. Patients with tumors expressing high levels of both Met and Snail (MET+/SNAIL+; n=76) show reduced DFS in comparison to the other groups (MET-/SNAIL-, MET-/SNAIL+ and MET+/SNAIL-; n=456) (Log-rank p-value=7.40e-03). To assess the association of DFS with the MET and SNAIL protein status jointly, we compared survival of patients with tumors positive for both MET and SNAIL to a combined group with tumors positive for neither or only one of MET and SNAIL. Analysis was performed using SAS v9.1.3 (Cary, NC) with PROC LIFETEST and PROC PHREG. Investigation of differences between basal and non-basal subtypes was performed by a Cox proportional hazards model including an interaction term to statistically assess homogeneity of the association of DFS with MET and MET/SNAIL. Due to low event counts, inference was based on a method designed for small samples implemented in the SAS macro FC (Heinze and Schemper 2001). See supplementary text for details.

2.26 Immunohistochemical analysis of murine tissue

Following antigen retrieval using the specified solution, primary antibodies were incubated at the following dilutions in 2% BSA/PBS: Met (1/50; Assay Designs; 10 mM sodium citrate, pH 6), keratin 5 (1/1000; Covance; 10 mM sodium citrate, pH 6), keratin

6 (1/600; Covance; 10 mM sodium citrate, pH 6), keratin 14 (1/1000; Covance; 10 mM sodium citrate, pH 6), keratins 8/18 (1/1000; Fitzgerald Industries International; 10 mM sodium citrate, pH 6), ß-catenin (1/200; BD Transduction; 10 mM sodium citrate, pH 6), Trp53 (1/300; Santa Cruz; 10 mM sodium citrate, pH 6), Trp63 (1/50; Abcam; 10 mM sodium citrate, pH 6), ErbB2 (1/100; Cell Signaling; 1mM EDTA, pH 8), Smooth muscle actin (1/500; Sigma-Aldrich; 10 mM sodium citrate, pH 6), Vimentin (1/800; Novus Biologicals; 10 mM sodium citrate, pH 6) and Egfr (1/250; Epitomics; 10 mM sodium citrate, pH 6). Secondary antibodies, goat anti-mouse HRP (Jackson Immunoresearch Laboratories), goat anti-rabbit (Vector labs) were diluted in 2% BSA/PBS and the signal was amplified using the Vector Elite ABC kit (Vector Labs). For Trp63 staining, the M.O.M. kit (Vector labs) was used according to manufacturer's protocol. All immunohistochemistry stains were visualized using DAB (Dakocytomation-Glostrup). Sections were counterstained with hematoxylin, dehydrated and mounted using Acrytol (Surgipath). Antibody staining was optimized using appropriate positive and negative control tissues. Slides were scanned using an Aperio ScanScope XT (Aperio Technologies) with a 20x objective and images were extracted using the ImageScope image viewer (Aperio Technologies).

2.27 Immunofluorescence analysis of mouse tissue and cells

Paraffin-embedded mouse tissues were antigen retrieved with 10mM Tris base/1mM EDTA pH 8 for 15 minutes in a high pressure cooker and incubated with keratins 5 (1/500, Covance), 8/18 (1/200, Fitzgerald Industries International), and Met (1/50, R&D systems) antibodies. The Alexa Fluor donkey anti-goat, and donkey anti-rabbit 647 antibodies were purchased from Molecular Probes. Donkey anti-guinea pig IgG (H+L)

Fab'2:FITC-labeled secondary antibody was purchased from Jackson Immunoresearch Laboratories. Actin staining was visualized using Alexa 488 conjugated phalloidin (Molecular Probes) in Met^{mt} breast epithelial cell cultures as described (Frigault, Naujokas et al. 2008). Nuclei were counterstained using a 0.5 ng/ml DAPI dilution in water and cover slips were mounted onto microscope slides with immu-mount (Thermo-Shandon). Z-sections and regular images were taken using a Zeiss Meta confocal microscope (model LSM 510, Carl Zeiss).

2.28 Reverse transcription (for quantitative RT-PCR)

First strand cDNA synthesis was performed using Superscript II RT (Invitrogen). To a volume of 10 µl total RNA was added 1 µl oligo (dT) primers (Invitrogen), and 1 µl 10 mM dNTP mix (Invitrogen) to produce a final volume of 12 µl. For Met^{mt} tumor and normal tissue, 100 ng of total RNA was used for reverse transcription. Oligo (dT) primers were annealed at 65°C for 5 min in Biometra's $T_{Gradient}$ thermocycler and samples were chilled on ice. A master mix containing 4 µl 5x First Strand Buffer (Invitrogen), 2 µl 0.1M DTT (Invitrogen), 1 µl RNase Inhibitor (10 units/µl, Invitrogen)) was added to each sample, mixed by pipetting, and incubated at 42°C for 2 minutes. Superscript II RT (Invitrogen) was subsequently added to each tube (1 µl, 200 units), mixed by pipetting and incubated at 42°C for 50 minutes. The reaction was inactivated by heating at 70°C for 15 minutes. Negative controls included a sample with no RNA, or no Superscript II RT enzyme.

2.29 Quantitative RT-PCR (qRT-PCR) analysis

To validate levels of *Nck1* in Met^{mt} mouse mammary tumors compared to normal tissues, cDNA was first prepared as described in Section 2.28 from 100 ng of total RNA

(isolated using TRIzol) and was diluted 2-fold for qRT-PCR analysis. Quantitative PCR was performed using Quantitect SYBR Green kit (Qiagen) in a LightCycler (Roche). Primers were designed using Primer3 Output software. qRT-PCR was performed in duplicate, and transcript levels were normalized to glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) as levels of *Gapdh* were similar in normal and tumor tissues. Primer sequences were as follows: *Nck1* sense, 5′- ATT CAG CAC CAT GGA GGA AC-3′; *Nck1* antisense, 5′- AAG TAC AAT TGG CCC AGC AC-3′; *Gapdh* sense, 5′-CAT CAA GAA GGT GAA GC-3′; *Gapdh* antisense, 5′-CCC TCA ACG ACA ACT TCA GC-3′. MDCK cells were stimulated with 135 ng/ml HGF for the indicated times (Chapter 4) and RNA was isolated using TRIzol (Invitrogen). cDNA from MDCK epithelial cells was synthesized using the SuperScript II reverse transcriptase (Invitrogen) and oligo(dT) (Invitrogen). Quantitative PCR was performed using Quantitect SYBR Green kit (Qiagen) in a RG3000 Rotor Gene (Corbett Research). *SNAIL* expression levels were normalized using *GAPDH* as levels were relatively equal across different samples analysed. The following primer sequences were used: *SNAIL* sense, 5'-CGT CCT TCT CTT CCA CCT CA-3'; *SNAIL* antisense, 5'-AAG GTT CGG GAA CAG GTC TT-3'; *GAPDH* sense, 5'-ACC ACC GTC CAT GCC ATC AC-3'; *GAPDH* antisense, 5'-TCC ACC ACC CGG TTG CTG TA-3'. Values are the result of three independent experiments.

2.30 PCR analysis from genomic DNA

Genomic DNA from Met^{mt} tumor and matched-normal mammary gland tissue was isolated and analyzed by PCR. The Met transgene was identified by PCR using SV40 primers. The following primers were used: SV40 sense: 5'-GGA ACC TTA CTT CTG

TGG TGT-3' and SV40 antisense: 5'-GGA AAG TCC TTG GGG TCT TCT-3'. PCR conditions were as follows: 5 minutes at 94°C, followed by 35 cycles (45 sec, 94°C; 45 sec, 60°C; 45 sec, 72°C), and 7 minutes extension at 72°C. Per 50 µl reaction, the PCR reaction conditions were 5 µl PCR buffer (Invitrogen), 0.125 µl forward primer (1 µg/µl), 0.125 µl reverse primer (1 µg/µl), 1 µl 50 mM $MgCl_2$ (Invitrogen), 1 µl 10mM dNTP (Invitrogen), 0.5 µl Taq Polymerase (Invitrogen), 50 ng of genomic DNA and water. *Gapdh* was used as the control as levels were comparable between samples. The following primers were used: *Gapdh* forward 5'-AAC TTT GGC ATT GTG GAA GG-3' and *Gapdh* reverse 5'-ACA CAT TGG GGG TAG GAA CA-3'. PCR conditions were as follows: 5 minutes at 95°C, followed by 30 cycles (30 seconds, 95°C; 30 seconds, 55°C; 45 seconds, 72°C). For a 25 µl reaction, PCR reaction conditions were 1 µl 10 µM forward primer, 1 µl 10 µM reverse primer, 2.5 µl 10x PCR reaction buffer (Invitrogen), 1 µl 50mM $MgCl_2$ (Invitrogen), 1.5 µl 5mM dNTP (Invitrogen), 0.2 µl Taq polymerase (Invitrogen), 50 ng of genomic DNA and water. PCR reaction products were 300 bp (SV40), and 223 bp (*Gapdh*).

2.31 Cell culture of human cell lines

Human embryonic kidney 293 cell line variant containing the SV40 Large T antigen (HEK293T) and Madin Darby Canine Kidney (MDCK) cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) containing 10% fetal bovine serum (FBS) and 50 µg/ml Gentamicin (Invitrogen).

2.32 Transient transfections of human cell lines

1 x 10^6 HEK293T cells were seeded 24 hours prior to performing transient transfections using Lipofectamine Plus (Invitrogen) according to manufacturers

instructions. Media was replaced 3 hours post-transfection, and cells were lysed 48 hours post-transfection in radioimmunoprecipitation assay (RIPA) buffer (0.05% SDS, 50 mM Tris, pH 8.0, 150 nM NaCl, 1% Nonidet P-40, 0.05% sodium deoxycholate) supplemented with 1 mM PMSF, 1 mM sodium vanadate, 10 μg/ml aprotinin, and 10 μg/ml leupeptin. Homogenates were centrifuged at 13 000 rpm for 15 minutes to remove debris.

2.33 Immunoprecipitation assays

500 μg of HEK 293T cell protein lysates were used for each immunoprecipitation. Antibodies were allowed to bind for 1 h at 4°C (2 μl/sample for Met148, and 0.5 μl/sample for GFP antibodies), after which 15 μl of Protein A sepharose beads were added to collect immune complexes. Beads were washed three times in 1 % Triton lysis buffer (50 mM HEPES, 150 mM NaCl, 1.5 mM MgCl2, 1mM EGTA, 1% Triton X-100, 10% glycerol), then resuspended in 15 μl of 6x Laemmli sample buffer and boiled for 2 minutes. Samples were separated on 10 % SDS-PAGE gels. Proteins were then transferred onto nitrocellulose membranes (Hybond-C, Amersham). Mouse mammary tumor and normal tissue protein lysates required 1 mg per immunoprecipitation reaction (3 μl for Nck1 #2383, and Nck2 #3313 antibodies). Protein lysates of tissue origin were pre-cleared with Pansorbin (Calbiochem) for 1 hour at 4°C and subsequently processed for immunoprecipitation as described above.

2.34 SDS-PAGE and immunoblot analysis

Met^{mt} tumor-derived epithelial cells were harvested in RIPA buffer supplemented with 1 mM PMSF, 1 mM sodium vanadate, 10 μg/ml aprotinin, and 10 μg/ml leupeptin. Protein lysates were resolved by SDS-PAGE (10% acrylamide gels), and transferred to

nitrocellulose membrane. Membranes were blocked in 3% BSA in TBST (10 mM Tris, pH 8.0, 150 mM NaCl, 2.5 mM EDTA, 0.1% Tween 20) for 1 hour and incubated with primary and horseradish peroxidase-linked secondary (Amersham) antibodies in TBST overnight at 4°C or for 1 hour at room temperature, respectively. Antibodies used for western blotting included phospho-Met (pTyr1234/1235, Cell Signaling #3126), phosphotyrosine (PY100, Cell Signaling #9411), total Erk (Cell Signaling #9102), phospho-Akt (pSer473, Cell Signaling #9271), total Akt (Cell Signaling #9272), GFP (Molecular Probes A-6455), actin (Santa Cruz, I-19), Nck (BD biosciences), phospho-specific Erk1/2 (pThr202/pTyr204, Santa Cruz, E-4), Met148 and Met19S (see Section 2.1). Immunoblotting with Nck1 and Nck2-specific antibodies required a different set of conditions. Membranes were blocked with 10% skim milk in TBST for 15 minutes at room temperature and Nck1 (1:750 in 10% milk/TBST) and Nck2 (1:250 in 10% milk/TBST) antibodies were incubated overnight at 4°C. Horseradish-peroxidase donkey anti-rabbit (1:10 000 in 10% milk/TBST) and goat-anti rabbit (1: 10 000 in 10% milk/TBST) were incubated at room temperature for 1 hour for Nck1 and Nck2 immunoblotting, respectively. Protein bands were visualized using an ECL detection kit (Amersham).

2.35 Establishment of primary culture from mouse mammary tumors

Two methods were used to establish primary culture from mouse mammary tumors. The first method affords large quantities of epithelial cells compared to the second method. Tumors were minced well with a razor blade until a slurry was produced. Slurries were incubated with 20 ml of DMEM (in a 50 ml conical tube, Starstedt) containing 3 mg/ml of collagenase A (Roche, product no. 11088793001) and 3 mg/ml

Dispase II (Roche, product no. 04942078001) and nutated at 37°C for 2-3 hours. Samples were spun at low speed (700 rpm for 5 minutes), and supernatants containing fibroblasts, were discarded. Epithelial cells were washed with 0.02% EDTA/PBS pH 8, and centrifuged (700 rpm for 5 minutes) until this was repeated until supernatants were clear (as opposed to reddish from blood). After the last spin, cells were counted and seeded in T25 (2 x 10^6 cells), T75 (4 x 10^6 cells) flasks, 6-cm (2 x 10^6 cells) or 10-cm (4 x 10^6 cells) plates culture dishes (NUNC Nalgene) in 5 % complete medium (see Section 2.36). The second method selects for epithelial cell growth and minimizes fibroblast contamination and was adapted from a previously described procedure (Pei, Noble et al. 2004). This method is suitable for culturing cell from mammary tumors with basal-like phenotypes, which are notoriously difficult to culture (Pei, Noble et al. 2004). Here, 6-cm plates were FBS-coated by covering the bottom of the dishes with 100 % FBS and aspirating the excess serum. Plates were subsequently placed in an incubator at 37°C for 30 minutes. Tumors were cut into 1-mm^3 size pieces. Care was taken to avoid pieces with visible adjacent normal mammary gland. About 1-2 pieces were transferred to each 6-cm plate, followed by the addition of 5 % FBS in complete medium (2ml – just enough to cover the bottom of the plate and so that tumor pieces are able to adhere). Once the tumors had adhered (48-72 hours later), the cells were fed with 4 ml of fresh medium. Epithelial cells beginning to grow out of the tumor pieces can be seen at this stage (Figure 2.1A-B). If pure epithelial cells grew out of tumor pieces, then medium was kept in 5 % FBS complete medium. If fibroblasts grew out of tumor pieces, the FBS content was reduced to 0.5 % FBS for several weeks to months until fibroblasts died (Figure 2.1C-D). Once the 6-cm plates became confluent with epithelial cells, cells were passaged into one 10-cm plate. This was done by rinsing plates in 2 ml 0.05 % trypsin, aspirating, and

incubating in 2 ml 0.05 % trypsin at 37°C until all cells were detached. Cells were spun down at 700g for 5 minutes to remove trypsin, resuspended and plated in 5 % FBS complete medium. Cell culture purity was qualitatively assessed by co-immunofluorescence staining with epithelial markers keratins 8/18 and fibroblast marker smooth muscle actin prior to their functional characterization.

2.36 Mouse mammary tumor culture medium

Met^{mt}-murine tumor-derived epithelial cells were maintained in DMEM containing 0.5-5% FBS (Invitrogen), 5 μg/ml of insulin (Sigma-Aldrich), 1 μg/ml of hydrocortisone (Sigma-Aldrich), 35 μg/ml of bovine pituitary extract (J.R. Scientific), and 50 μg/ml Gentamicin (Invitrogen).

2.37 siRNA knockdown assays

siRNA-mediated knockdown of the Met receptor in Met^{mt}-derived tumor cells (6030T) was accomplished using the Fast-Forward Transfection of Adherent Cells with siRNA in 24-Well Plates (Qiagen) as per the manufacturer's protocol. Briefly, 1 x 10^5 cells were seeded per well in 500 μl of 5% FBS in complete medium with 7.5 μl HiPerfect transfection reagent and a final concentration of 50nM Met siRNA duplexes (Met_3 sense siRNA: 5'-CAA GGG AAG AAG UGU UUA A-3'; Met_3 anti-sense siRNA: 3'-UUA AAC ACU UCU UCC CUU G-5'). Cells were harvested 24, 48 and 72 hours post-transfection in 200 μl of RIPA lysis buffer. A similar protocol was used for Nck1 and Nck2 knockdown in Met^{mt}-derived tumor cells (6030T). Cells were seeded at a density of 5 x 10^5 cells per well of a 6-well dishes in 1852 μl of 5% FBS in complete medium with 30 μl Hiperfect transfection reagent (Qiagen) and a final concentration of 50nM of Nck1-1 and Nck2-4 siRNA duplexes (Qiagen) (Nck1_1 sense siRNA: 5'-CCU

AAA UAC GGG UCA AGU A-3'; Nck1_1 antisense siRNA: 3'-UAC UUG ACC CGU AUU UAG G-5'; Nck2_4 sense siRNA: 5'-CAG AAU GUU UAU GAA GAA A-3'; Nck2_4 antisense siRNA: 3'-UUU CUU CAU AAA CAU UCU G-5'). After 48 hours, a second round of siRNA transfection was performed with Nck1_1 and Nck2_4 siRNA. In some cases, coverslips were added to the bottom of 6-well dishes. After 48 hours, the cells were used for assays (invasion/migration, confocal) or lysed in RIPA buffer for western blotting. Appropriate concentrations of AllStars Negative control siRNA (Qiagen) were used in all experiments.

2.38 Invasion and migration assays

Met^{mt} tumor-derived cells (6030T) were counted (5×10^4 or 1×10^5) and resuspended in 5% FBS in complete medium containing either 0.1 μM of Met inhibitor PHA-665752 (Pfizer Pharmaceuticals) or 0.1 μM of DMSO (Sigma). Cells were seeded directly onto 6.5-mm diameter Corning Costar transwells for migration or transwells coated with 100 $\mu g/cm^2$ of matrigel (BD Biosciences) for invasion assays. Complete media was added to the bottom wells and plates were incubated at 37°C for 24 hrs. For HGF stimulations, 34 ng/ml of HGF was added to the bottom wells in 5% FBS in complete medium. Following incubation, cells were fixed using 10% neutral-buffered formalin (Surgipath) for 20 minute at room temperature. After washing with distilled water, cells were stained with 0.1% crystal violet in 20% methanol for 20 min at room temperature. Cells on the top layer were scraped and membranes were left to dry overnight. Images were captured using a Zeiss Axioskop microscope (Carl Zeiss). Quantification was done using a constant threshold for all images and measuring pixels using a Scion Image-NIH equivalent program for Microsoft Windows (Scion Company).

Data for each insert are represented as the average pixel count from the five images. All experiments were done a minimum of thrice.

2.39 Tumor transplantation

Transgenic mammary tumors (~2mm^3) were routinely transplanted into the 4th mammary fat pad of FVB/N recipient mice that were less than 4 weeks of age. Animals were sacrificed and transplanted tumors were harvested when total tumor volume was ~1cm^3. Transplanted tumors were representative of the histology from the originating transgenic mammary tumor, and this was used as a method to generate more tumor material for experimentation.

2.40 Cryopreservation of mammary tumors and tumor-derived cells

Mammary tumors and tumor-derived cell lines were frozen in 95 % FBS, 5 % DMSO and stored in liquid nitrogen. Mammary tumors could be thawed at a later time and cultured as described in the second method of Section 2.35. These were also found to grow upon transplantation in FVB/N or nude mice.

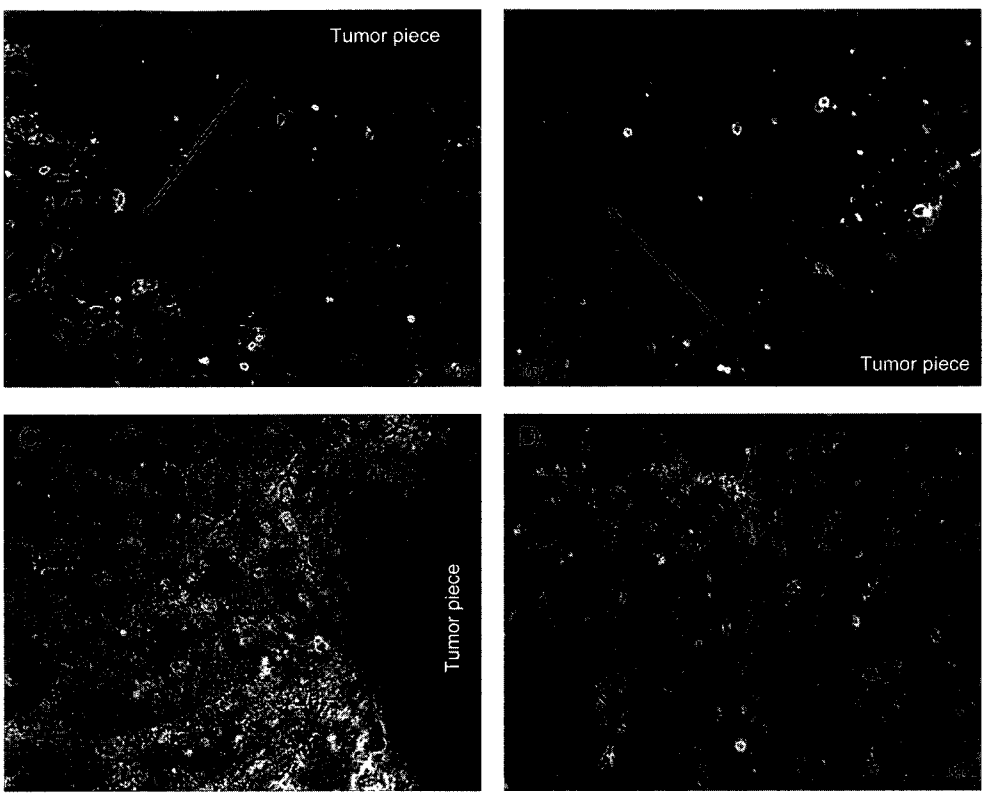


Figure 2.1: MMTV/Met^{mt} mammary tumors are cultured to select for growth of epithelial cells. (A-B) Pure populations of epithelials are seen growing from tumor pieces. The direction of cell growth is shown with an arrow, and the tumor piece is indicated. (C) An example of a fibroblast-rich culture growing from a tumor piece. This tumor piece was kept in low percentage of serum (0.5%) in complete medium, which allows epithelial cells to survive, but results in fibroblast death. (D) Image shows fibroblasts undergoing cell death after prolonged periods (3 months) in low serum conditions. Images were captured using the Axiovert (Zeiss) microscope with 10x objective.

Chapter 3

3 Optimization of methods for gene expression profiling of MMTV/Met^{mt} murine mammary tumor and normal tissue

3.1 Introduction

The role of the Met receptor in breast cancer is poorly understood. We have previously generated a mouse model in which oncogenic variants of Met (MMTV/Met^{mt}) are expressed in the mammary epithelium, which produced mammary tumors with diverse phenotypes. We used a microarray approach to gain an understanding of the global gene expression changes involved in MMTV/Met^{mt}-induced mammary tumorigenesis. In order to generate gene expression profiles of tumor epithelium that were not affected by contaminating stromal cells, we used laser capture microdissection (LCM) as a method of cell isolation. At the time during which this work was conducted, no reports had been published using mouse mammary tissue for LCM and subsequent molecular analysis. Therefore, the goal of this chapter was to optimize a protocol in which epithelial cells from MMTV/Met^{mt} murine mammary tumors could be dissected and analyzed with microarrays. Importantly, this protocol needed to be conducted under RNA-preserving conditions. Following LCM, RNA was isolated, amplified, labeled and hybridized onto whole genome mouse arrays (Figure 3.1). Here, a description of these procedures as well as important factors to consider for each step is outlined.

3.1.1 "Laser capture" and "laser cutting" microdissection

The development of laser microdissection emerged from a need to isolate pure populations of cells from tissues with cellular heterogeneity. During the mid-1990s there was a rise in microarray experiments and the projected goal was to be able to use

dissected cells for analysis in a future diagnostic setting. Manual methods of tissue dissection included using a scalpel blade to cut away regions of undesired frozen tissue, irradiation of manually ink-stained areas to destroy unwanted tissue, and more precise methods such as using a very fine stainless steel needle attached to a micromanipulator (Fearon, Hamilton et al. 1987; Shibata, Hawes et al. 1992; Radford, Fair et al. 1993; Going and Lamb 1996; Hernandez and Lloreta 2006). However, these methods are prone to contamination and are time-consuming, making desired macromolecules (i.e.: RNA) sensitive to degradation. The development of laser-based microdissection has completely changed molecular analysis of normal versus tumor tissues, allowing for a better understanding of the biological processes that differ between the two.

LCM encompasses two distinct laser-based methodologies for isolating specific cell types from whole tissue. LCM was the first of the two methodologies to be developed in the mid-1990s by Dr. Emmert-Buck at the National Institutes of Health (Emmert-Buck, Bonner et al. 1996) and was subsequently made commercially available by Arcturus (now Molecular Devices). This method involves an inverted microscope, an infrared laser beam and a manually controlled stage (Figure 3.2A). A small plastic cap, which resembles the cap of an eppendorf tube, is lowered over the area of interest to lie flatly upon the tissue section (Figure 3.2A steps 1 and 2). On the underside of the cap is a transparent, thin (100 μm) thermoplastic film (ethylene vinyl acetate polymer). Once the infrared laser is focused and fired on the cells of interest, the laser activates the thermoplastic film causing it to melt, infiltrate the cells and within milliseconds solidify and thus encapsulate the cells (Figure 3.2A step 3). The user can select the diameter of the laser beam to use between 7.5, 15, and 30 μm, depending on the level of precision that is required. Once the LCM cap is lifted, the captured cells that are adhered to the film

detach, while the surrounding undesired tissue remains on the microscope slide (Figure 3.2 step 4). One important advantage of this methodology is that the LCM cap can be viewed microscopically to determine if contamination by non-targeted cells has occurred. Contaminating cells can be removed by lightly touching the LCM cap against an adhesive strip (CapSure Pad). Following this, the LCM cap is placed in a collection tube containing total RNA extraction buffer and processed for downstream applications (Figure 3.2 step 5). The infrared laser generates minimal heat and can be activated many times using the same LCM cap on the same tissue section without causing unwanted modifications to nuclear material and cellular proteins (Emmert-Buck, Bonner et al. 1996).

The second methodology is called "laser cutting microdissection" (LCuM) and uses a motorized-stage, a 'cold' nitrogen ultraviolet (UV) laser, and an inverted microscope. The basic principle of this technique differs from LCM in that the user first defines the cells of interest on a computer screen by drawing around them with a cursor. A very narrow (1 μm) UV laser is then used to 'cut out' targeted cells from undesired tissue. This method was developed in Europe and has been made commercially available by European-based microscope manufacturers (Bohm, Wieland et al. 1997; Schutze and Lahr 1998). The desired cells are then placed into a collection tube by different methods depending on the manufacturer. In the case of the PALM instrument by Zeiss, dissected cells are catapulted into the collection tube using a pressure catapulting system. One advantage of this methodology over LCM is that this is a completely non-contact system. Therefore, the theoretical disadvantage of melting and solidifying of a thermoplastic film over the tissue section, and the potential of picking up contaminating cells with the cap are avoided with LCuM. LCuM can also be more time-efficient if large regions of tissue

are to be dissected, as the user can manually delineate multiple large regions of cells within a section, and then instruct the microscope to collect these areas. In contrast, a user would need to repeatedly fire the laser to capture the same area with LCM. On the other hand, a major disadvantage of the LCuM method is that the desired cells are immediately "catapulted" into the collection tube once the UV laser has finished cutting out the section of interest. At no point may the user verify that the desired cells have in fact been properly dissected out of the tissue section and that non-contaminating cells have not been cut out with them.

3.1.2 Staining of mouse mammary tissue for LCM

The theoretical aims of tissue preparation for LCM are to preserve tissue morphology, to visualize the cells of interest using a nuclear dye or immunohistochemistry techinque while maintaining RNA integrity. However, these are conflicting aims since the best method for preserving tissue morphology, namely formalin fixation and paraffin embedding, is a poor method for maintaining RNA integrity. Formalin fixation results in RNA fragmentation and induces several chemical modifications such as cross-linking between amino groups, which results in poor PCR amplification (Burgemeister, Gangnus et al. 2003). Thus, frozen tissue embedded in optimal cutting temperature (OCT) medium is often used for LCM as it preserves RNA integrity to a better degree than samples fixed in formalin and embedded in paraffin. Tissue morphology is also maintained, although to a lesser extent than in formalin-fixed paraffin embedded samples.

In order to perform LCM, a staining method is needed to distinguish between epithelium and stroma. Immunohistochemistry and staining with nuclear dyes (i.e.:

hematoxylin, cresyl violet, methyl green) are two methods that can be used. However, immunohistochemical staining for the purpose of LCM is generally a more lengthy procedure than staining with nuclear dyes. If specific antigen recognition with antibodies is not required, it is best to use the least time-consuming method possible to minimize RNA degradation. A goal for staining is to minimize the time in which the tissue is exposed to aqueous solutions at room temperature, since this causes reactivation of endogenous RNases (Fleige and Pfaffl 2006). In comparison to DNases, RNases are very stable, and active enzymes that do not generally require metal cations to function (Fleige and Pfaffl 2006). They are ubiquitous, difficult to inactivate, and minute amounts are sufficient to destroy RNA. Also, RNA is soluble in aqueous solutions and can be lost, thus lowering final yields (Fleige and Pfaffl 2006). However, nuclear dyes bind RNA and can become problematic for sensitive applications (Murray 2007).

3.1.3 Assessment of RNA quality and quantity

3.1.3.1 Determination of RNA quality

The most common way to assess total RNA quality is to run an aliquot of the sample on a denaturing agarose gel. However, at least 200 ng of total RNA is needed to load on a gel for visualization. Since such low quantities of total RNA are extracted from LCM, it is often not possible to spare 200 ng, and to have enough material left over to continue with microarray analysis. As an alternative, Agilent developed the 2100 BioAnalyzer that uses microfluidics-based technology to assess RNA quality of up to 12 samples in one run from as little as 200 pg of sample. Using this method, RNA samples are electrophoretically separated on a micro-fabricated chip using a gel-dye mix. The fluorescent dye intercalates with the RNA and is subsequently read by laser-induced fluorescence detection. This is translated into an electropherogram or gel-like image

(Figure 3.3). An RNA ladder is included for estimation of sample band size during electrophoresis. The time in seconds on the x-axis is the equivalent to the migration time of RNA bands on a regular denaturing gel, where the size of RNA species increases from left to right; the y-axis displays fluorescence units (Figure 3.3). An example of an electropherogram and gel-like image produced from a pure source of total RNA (i.e.: HeLa) that was exposed to minimal RNA degradation is shown in Figure 3.3A. There are two ribosomal RNA (rRNA) bands, 28S and 18S, and a small peak at about 27 seconds representing 5S rRNA. Messenger RNA (mRNA) represents only about 1-5 % (0.1-1 pg) of all RNA within the cell, and thus it is too low in abundance to visualize by any means. As an alternative, the relative ratio of 28S to 18S rRNA is used as a surrogate to approximate the quality of mRNA. mRNA is assumed to be of excellent quality if the ratio of 28S to 18S rRNA approaches 2 (Skrypina, Timofeeva et al. 2003) (Figure 3.3A). From the gel-like image, two bands are apparent where the 28S rRNA band is about twice the intensity of the 18S band, with an absence of smaller degradation bands closer to the bottom of the gel (Figure 3.3A). A 28S/18S rRNA ratio of 2 is considered perfect, yet this is rarely attainable, especially from clinical or other tissue samples (Fleige and Pfaffl 2006). A degraded RNA sample would be one that shows a decrease in 28S/18S ratio, with an increase in abundance of more peaks representing degraded RNA fragments toward the left of the electropherogram, and more banding or smearing at the bottom of the gel-like image (Figure 3.3B).

3.1.3.2 Determination of RNA quantity

As mentioned earlier, small yields of total RNA are obtained following LCM, and these are far below the limit of detection of regular laboratory spectrophotometers. The

Nanodrop1000 UV spectrophotometer is one instrument that may be used to quantify small RNA yields. This instrument uses one microliter of sample, however, its lower limit of detection is 1.5 ng/μl of total RNA, and may not be suitable. A more sensitive assay that quantifies picograms of sample from a single microliter is RiboGreen (Molecular Probes). RiboGreen assay is accurate and reproducible when RNA yields are in the picogram range.

3.1.4 Factors influencing RNA amplification

Depending on the platform, microarray experiments generally require a minimum of 2-4 μg of poly(A) RNA or 25-50 μg of total RNA for hybridization. Thus, a method to amplify the amount of RNA isolated from LCM samples is necessary in order to perform a single microarray experiment. Although there has been some success with polymerase chain reaction (PCR)-based methods of RNA amplification (Iscove, Barbara et al. 2002; Smith, Underhill et al. 2003), the most commonly used method is a T7-based linear amplification (Van Gelder, von Zastrow et al. 1990; Eberwine, Yeh et al. 1992; Phillips and Eberwine 1996). This method uses oligo dT primers containing a phage T7 RNA polymerase promoter to prime the synthesis of first strand cDNA by reverse transcription of the poly(A) tail within mRNA (Figure 3.4). In this manner, the T7 promoter of RNA polymerase becomes incorporated within double stranded DNA. T7 RNA polymerases then catalyze the transcription of multiple copies of antisense RNA from the cDNA template in a linear fashion. However, there is a drawback to this method of RNA amplification. The reverse transcription enzyme is not completely efficient at reverse transcribing full length of poly(A) RNA and ultimately generates transcribed RNA products with a 3' end bias (Phillips and Eberwine 1996; Baugh, Hill et al. 2001). The

average length of amplified RNA transcripts are thus further shortened when a second round of amplification is introduced (Figure 3.4). Since the amplification bias is the same for all amplified RNA, the relative abundance of RNA between samples that are processed using the same method may still be compared (Schneider, Buness et al. 2004). Several companies, including Agilent Technologies, designed the majority of probes on their microarray platforms to map to the 3' end in order to compensate for the bias. Multiple studies have validated T7-based linear RNA amplification by comparing unamplified and amplified material from the same samples and the concordances reported were high, demonstrating that RNA amplification is reproducible and large systematic biases are not introduced (Luzzi, Holtschlag et al. 2001; Zhao, Hastie et al. 2002; Luzzi, Mahadevappa et al. 2003; Schneider, Buness et al. 2004).

3.1.5 Factors influencing RNA labeling with fluorophores

3.1.5.1 Principles of RNA labeling

Microarray experiments require that samples be labeled with fluorophores. In the case of one-color experiments, each sample is labeled with one fluorophore and hybridized to a single microarray containing probes that represent different genes. Other samples are treated the same, and the relative levels of fluorescence, which represent relative levels of transcript abundance, are compared between microarray. With two-color experiments, two samples are labeled with two different fluorophores, and co-hybridized to the same array. The relative levels of expression of each are compared. Fluorescent labeling with cyanine 3 (Cy3) and 5 (Cy5) are a commonly used fluorophores for two-color microarray experiments as they emit strong fluorescence, share similar chemical properties, have well-separated fluorescence spectra, and display minimal background adherence to the glass slides (Ehrenreich 2006).

3.1.5.2 Advantages and disadvantages of direct and indirect labeling methods

Labeling can be done in two ways, through a direct or indirect labeling approach. For direct labeling, Cy3 or Cy5 fluorescently conjugated nucleotides (i.e.: Cy3-dUTP and Cy5-dUTP) are incorporated into amplified RNA during the T7 *in vitro* transcription step. For indirect labeling, a modified nucleotide, 5-(3-aminoallyl)-UTP (aaUTP), containing a reactive primary amino group on the C5 position of uracil is incorporated during this step instead (Figure 3.4). The aminoallyl-labeled RNA is subsequently coupled to N-hydroxysuccinimidyl (NHS) ester-derivitized reactive Cy3 or Cy5 dyes in a separate chemical reaction. Indirect labeling is often favored, as RNA polymerases do not efficiently incorporate the large and bulky Cy3 and Cy5 nucleotide conjugates in the transcribed strand (Ehrenreich 2006). Moreover, the Cy5 fluorophore is even larger than the Cy3 fluorophore and steric hindrance results in less efficient dye incorporation of Cy5 compared to Cy3 in the amplified RNA product. This could result in a global dye bias, where Cy3-labelled transcripts are mistakenly interpreted as being overexpressed due to stronger fluorescence emitted from these samples. This is much less of a problem using indirect labeling as the aminoallyl moiety of dUTP is small and RNA polymerases are able to incorporate aaUTP efficiently into the growing strands of RNA. In a separate step, the Cy3 and Cy5 dye-tags are added by chemical coupling to the aminoallyl moiety increasing the likelihood that both dyes are incorporated equally into the final product and thus reducing the probability of global dye biases. The drawbacks to indirect labeling include that the NHS ester-derived dyes are sensitive to moisture, and require more bench work to perform the additional chemical coupling step (Ehrenreich 2006). Overall, direct labeling results in the incorporation of 1 fluorophore for every 6 nucleotides whereas indirect labeling results in 1 fluorophore for every 20-50 nucleotides (Lee and Saeed

2007). Dye incorporations are monitored after labeling step to determine whether the samples are suitable for microarray hybridization.

3.1.6 Microarray analysis

3.1.6.1 Microarray platforms

There are two main types of microarrays, (1) oligonucleotide microarrays (i.e.: Affymetrix, Agilent), and (2) double stranded DNA microarrays. DNA microarrays were the first to be developed. DNA probes of 200- to 800-bp and as large as 1.3 kb are designed, amplified, purified by electrophoresis and spotted onto glass slides. An advantage to double stranded DNA microarray platforms is that since the probe lengths are larger, they allow for greater specificity and affinity to their targets under stringent conditions. They are also cost-efficient. However, the drawbacks are that the fabrication of these arrays is laborious and error-prone. Oligonucleotide arrays use 25-mer probes and are called "high density" arrays as their probes cover an organism's entire genome. Genes are represented by multiple different probes on these arrays. Due to the short size of the probes, reduced specificity and sensitivity becomes an issue. To compensate, Affymetrix has designed a set of probes for each gene that are "perfect matches" and a set of probes for the same gene that are "mismatched' by one nucleotide. The mismatched probes, which should not bind any transcripts, act as a built-in control for non-specific binding for each gene (Lipshutz, Fodor et al. 1999). Agilent uses longer 60-mer probes, providing greater specificity and sensitivity than Affymetrix arrays, and obviates the need for "mismatch" probes. Instead, more probes could be added to the array in the extra space (Ehrenreich 2006).

3.1.6.2 Study designs for two-color experiments

There are four design strategies commonly used for two-color experiments: direct comparison, balanced block-design, reference sample design, and loop design (Lee and Saeed 2007). By contrast, one-color experiments obviate the need for design strategies as each sample is directly hybridized to their respective array and analyzed.

The direct comparison strategy is used when RNA is not limiting, and when a maximum of two different experimental samples require analysis. The greatest advantage of using the direct design strategy is explained by its name. Both samples are co-hybridized on the same array and thus, analysis of differentially expressed genes is more accurately measured (Lee and Saeed 2007). Gene-specific dye biases are a problem but can be avoided by performing dye swaps, where each sample is divided into two aliquots and reversely labeled. When RNA is not unlimited, the balanced block design is a viable option and dye swaps are not needed. This employs a similar strategy as direct comparison, but the direction of labeling is switched among the hybridizations, and each hybridization represents one biological replicate. The loop design is used when two or more samples are compared to each other in a "head to tail" fashion. In this way, dye biases are not a problem because each sample in the loop is used once as the head (Cy3) and once as the tail (Cy5). This design however is less robust when even a single array is of poor quality as the entire loop breaks down. Although the loop design can be used with more than two samples, a far more efficient design is the reference design. Here, each sample is hybridized against a common reference on a single array. The goal of the common reference is to provide fluorescently labeled target RNA that covers most (>90%) probes on the array (Lee and Saeed 2007). This is typically accomplished by pooling multiple cell lines of the same species (i.e.: mouse). The scanner reads

fluorescence as a ratio between the two dyes, and in order to get a proper scan, each probe needs to generate fluorescence signal. The reference design is most commonly used when multiple experimental samples require comparison (i.e.: multiple tumor and normal samples) and it is not economical, or feasible (due to i.e., low RNA quantity) to use the direct comparison design. With the reference design, data generated from any sample co-hybridized to a common reference can be compared. Ratio values for gene expression changes are obtained indirectly, where the mathematical effect of the common reference is cancelled out of the equation (ie: $[\mathbf{T_1}/R]/[\mathbf{N_1}/R]=[\mathbf{T_1}/\mathbf{N_1}]$, where T_1=tumor sample 1, N_1=normal sample 1, and R=common reference). The drawback to the reference design is that since the ratio values are inferred through a common reference, the data exhibits greater variance in expression measurements than if the two samples were directly compared on a single microarray (Lee and Saeed 2007). Although dye swaps are not necessary they increase the accuracy of the data.

3.2 Results

3.2.1 Determination of conditions for cryopreservation and cryosectioning

Cryopreservation of tissue samples is important for the maintenance of tissue architecture and for the preservation of genomic material in unfixed samples. Samples were embedded in OCT medium and flash frozen in liquid nitrogen, similar to methods previously described (Finak, Sadekova et al. 2006). This resulted in preservation of tissue histology and total RNA integrity of MMTV/Met^{mt} mouse mammary normal and tumor samples (Figure 3.3C).

An important step for LCM is obtaining good quality tissue sections with minimal tissue ripping or folding. During LCM, the cap must lie flatly on the tissue section or this

will prohibit the infrared laser from melting the transfer film and adhering to the captured cells (Figure 3.2A). In order to determine the optimal temperature at which to cut frozen sections of mouse mammary tumor and normal tissue that would minimize ripping or folding, we first tested temperature settings suggested in the Microm cryostat user guide. Since the majority of normal mouse mammary gland is composed of fat tissue, we approximated that the temperature (-33°C) suggested for human fat could be used for cutting normal mouse mammary gland tissue. Such a low temperature was needed to help solidify the soft and oily fat tissue, and prevent the sections from folding, wrinkling, and sticking to the cutting stage and anti-roll plate which can be problematic. We found that cooling the cutting stage with a piece of dry ice facilitated cutting of normal mammary gland tissue. We used a slight warmer temperature of -25°C for cryosectioning mouse mammary tumor tissue which was more denser and solid than fat, making it easier to section. Overall, these conditions were suitable for crysectioning of MMTV/Met^{mt} tumor and matched-normal tissue. In practice, RNA quality similar to that shown in Figure 3.3C was routinely obtained from Met^{mt} mouse mammary tumor and normal tissues.

3.2.2 Determination of an RNA-preserving tissue staining protocol for LCM

The MMTV/Met^{mt} mouse model generates mammary tumors with variable contents of epithelium and stroma (Figure 3.5) (Petkiewicz 2007). In order to enrich for epithelial cells, tissue required staining under RNA-preserving conditions. The criteria for good staining was determined to be one that would easily allow distinction between epithelium and stroma without a coverslip, and that would preserve RNA integrity for downstream applications. We tested five different staining methods using four different pathological stains for frozen tissue. These included a variety of routinely used nuclear

dyes (hematoxylin and eosin [H&E], methyl green, cresyl violet) and a commercially available staining kit made specifically for LCM (Histogene) (Sgroi, Teng et al. 1999; Mikulowska-Mennis, Taylor et al. 2002; Okuducu, Janzen et al. 2003; Ginsberg and Che 2004; Finak, Sadekova et al. 2006; Finak, Bertos et al. 2008). Specifically, we tested the previously described H&E (H&E-long protocol) (Finak, Sadekova et al. 2006), and a modified H&E staining protocol (H&E-short protocol), in addition to methyl green, cresyl violet, and the HistoGene staining kit. The same mouse mammary tumor sample (4691T) was used for assessing each stain, and consecutive sections were used in order to compare RNA yields generated following staining. RNA extracted from tumor 4691T was first tested on the Agilent 2100 BioAnalyzer to determine its quality (Figure 3.6A "non-stained"). The impact of the H&E-long protocol on RNA integrity was subsequently assessed using mouse tissue. In comparison to non-stained tissue, tissue stained with the H&E-long protocol yielded RNA of sub-optimal quality and quantity (Figure 3.6B). The 28S/18S ratio significantly decreased immediately following staining. This protocol was originally optimized for the staining of human samples (Finak, Sadekova et al. 2006), which tend to involve larger tissue sections, on average, than mouse tissue section. We found that slides with large sections required more time to thaw prior to staining (over 30 seconds). Additionally, lengthy incubation periods in staining solutions (>30 seconds) were necessary to stain larger sections evenly. Since mouse tissues were a fraction of the size, we speculated that incubation periods of the H&E-long protocol were too long for small tissue sections, resulting in increased RNA degradation due to prolonged periods in aqueous solutions at room temperature. We next used the HistoGene protocol recommended for LCM by Molecular Devices. However, this resulted in the poorest RNA quality and a low quantity (Figure 3.6D). Additionally, staining using the

HistoGene protocol produced a gray-like stain over the entire tissue section and made identification of epithelial cells for LCM-dissection a challenge. Methyl green, a basic dye that binds preferentially to DNA and then RNA, stains nuclei light green (Figure 3.7A). Nuclear staining with methyl green was slightly faint without a glass coverslip, precluding easy identification of epithelial cells and RNA quality and quantity were poor (Figures 3.6F and 3.7A). Cresyl violet, a basic dye that binds preferably to acidic components such as RNA-rich ribosomes, nuclei, and nucleoli, stains nuclei deep purple (Figure 3.7A). Cresyl violet was a satisfactory in that it stained nuclei to the same intensity as hematoxylin and made identification of epithelial cells relatively easy. Interestingly, cresyl violet resulted in the best RNA yield (Figure 3.6E). Conventional staining protocols use aqueous solutions and this can reactivate endogenous RNases within the tissues that degrade RNA. In contrast, the cresyl violet staining protocol was carried out in ethanol or ethanol-based solutions, which may account for the greater yields (see Chapter 2 for protocol details). However, bioanalysis profiles showed additional peaks of low size consistent with RNA degradation (Figure 3.6E). This was puzzling because when this sample was subjected to electrophoresis on a 1% agarose gel, no significant degradation could be detected (data not shown). It is possible that organic solvents or other impurities may affect the readings from the BioAnalyzer instrument. In order to rule this out, 100 % ethanol or cresyl violet controls were subjected to bioanalysis, however no background peaks were observed (Figure 3.8). This suggests that either RNA degradation had occurred and the BioAnalyzer could detect it due to the increased sensitivity of this assay over agarose gel electrophoresis, or that an interaction between unknown impurities and the RNA itself produced background peaks.

In comparing the H&E-long staining protocol to other staining procedures for LCM, we reasoned that long incubations could be shortened to increase RNA quality and quantity at minimal cost to staining quality. We called this modified protocol H&E-short (Figure 3.6C). Staining quality was not compromised, and this protocol produced the greatest quality and sufficient quantity of RNA (Figure 3.6C and 3.7B). Although, cresyl violet gave the greatest RNA yields, the presence of additional peaks of unknown orign in the BioAnalyzer profile were concerning in regard to downstream applications that require pure RNA (i.e.: amplification). Since RNA isolated from H&E-stained tissue has previously been used successfully for microarray analysis (Finak, Sadekova et al. 2006), and given that cresyl violet staining produced incomprehensible in electropherogram traces, we chose to use the H&E-short protocol for our application.

3.2.3 Determination of RNA quality during LCM using correlative controls

We found that LCM samples contained trace amounts of salts and other impurities that interacted with total RNA and created background fluorescence readings on electropherograms (data not shown). Furthermore, LCM samples inherently contain low quantities of RNA, and it was not always possible to spare one microliter for bioanalysis of the LCM sample itself. Therefore, we needed to establish a method to routinely assess RNA quality of LCM samples by an "indirect" means before proceeding with costly downstream applications. We used correlative controls for each LCM sample to assess RNA quality from the starting material before it had been H&E-stained (non-stained), immediately after H&E-staining and before LCM (H&E-stained/before LCM), and following H&E staining and a 3-hour incubation during which time LCM is meant to be performed (H&E-stained/after LCM) (Figure 3.9). Correlative controls represented serial

sections of those used for LCM. Four slides of tumor tissue and eight to ten slides of matched-normal tissue were used for each grouping. Twice the number of normal sections was needed compared to tumor sections to account for low epithelial content in mammary fat pads. Non-stained control slides were thawed for the same amount of time as those that were processed for staining, since this reflects an important step where endogenous RNases can be reactivated (Fleige and Pfaffl 2006). H&E-stained/before LCM slides underwent RNA extracted immediately following H&E staining. Lastly, H&E-stained/after LCM slides were placed in extraction buffer after a 3-hour incubation representing the time allocated for performing LCM. Manufacturers recommended that LCM should be performed within a 3-hour period to minimize RNA degradation, and we adopted this time frame. Thus, total RNA quality after each step in the staining and LCM process was assessed with the BioAnalyzer. Bioanalysis showed that although there was a decrease in RNA quality post-staining (H&E-stained/before LCM) and post-LCM (H&E-stained/after LCM) compared to non-stained tissue, RNA quality was nevertheless acceptable for further processing of the sample (Figure 3.9). Furthermore, these data confirmed that 3 hours is a suitable time frame in which to perform LCM on murine tissue. Correlative controls were processed on all samples and subjected to bioanalysis for assessment of RNA integrity.

3.2.4 Determination of cell numbers and RNA yield from microdissected murine mammary normal and tumor tissue

After determining the best method to stain the tissue and to assess RNA quality, we needed to determine the number of infrared laser pulses that were required to obtain sufficient quantities of total RNA for downstream applications. Using two different mouse mammary tumor samples (4691T, 4164T), we determined that one laser pulse

captures 1.4-2.3 tumor epithelial cells (Table 3.1A). We next established that one laser pulse captures 0.6-1.1 normal mammary epithelial cells (data not shown).

We next asked how many laser pulses were necessary to obtain sufficient quantities of total RNA from mouse mammary tumor and matched-normal epithelium. Using four different tumor samples (559T, 4164T, 4691T, 4474T) in duplicate, we assessed how much total RNA could be obtained from 2 000, 5 000, 10 000, and 20 000 pulses (Table 3.1B). On average, four different tumor samples gave variable RNA yields for each category (Table 3.1B). We found that 2 000 pulses yielded 33.56ng, 5 000 pulses yielded 51.86ng, 10 000 pulses yielded 67.2ng, and 20 000 pulses yielded 137ng. Thus, a minimum of 2 000 pulses was sufficient to isolate total RNA from mouse mammary tumor epithelium. However, 2 000 pulses could easily be obtained from a single mammary tumor section on a single slide. We reasoned that generating gene expression profiles from a single tumor section may reflect biological biases found within the tumor microenvironment of that particular section which are not representative of the tumor as a whole. In order to generate gene expression profiles that were more representative of whole tumor epithelium, we pooled cells from multiple tissue sections of the same sample until a quota of 5 000 pulses was reached. This corresponded to 8 000 to 16 000 tumorigenic epithelial tumor cells. Since we had determined that half of the number of morphologically normal epithelial cells could be captured per infrared laser pulse, we dissected double the number of normal epithelial cells (12,000-24,000 morphologically normal mammary epithelial cells). We captured images of each mouse mammary tumor and normal sample before LCM, after LCM, and of dissected cells on the LCM cap to monitor for potential contamination from non-targeted cells (Figure 3.2B).

3.2.5 Determination of the amount of total RNA required for linear RNA amplification

We used Agilent whole genome expression microarrays, as these platforms required the lowest amount of amplified RNA for hybridization that were available at the start of the project. Using these microarrays, at least 2-3 μg of mRNA were required for labeling with Cy3 and Cy5 dyes and subsequent hybridization. RNA amplification was thus a prerequisite for gene expression microarray analyses since insufficient quantities of total RNA are isolated from LCM.

Although the lowest starting amount of total RNA for amplification that is recommended by several manufacturers is 100 ng, our laboratory and others have shown that 4 ng of total RNA or less is sufficient to generate robust and reproducible amplification from human breast tissue (Baugh, Hill et al. 2001; Finak, Sadekova et al. 2006). As shown in Figure 3.10A, an ideal amplified RNA profile (i.e.: HeLa) is one that shows a broad peak, which signifies transcripts of different lengths. These transcript range in size from 0.2-5.0 kb, with the majority being ~1 kb (Figure 3.10A). In order to generate sufficient quantities of RNA for hybridization, it was necessary to process mouse tumor and normal LCM samples through two rounds of RNA amplification. To test the effect of introducing a second round of amplification on the RNA bioanalysis profile, we used total RNA from a pure source (i.e.: HeLa). As expected due to a 3' bias, the broad peak from Figure 3.10A became narrower (Figure 3.10B) following two rounds of amplification. Transcript lengths spanned 0.2 kb-1.0 kb following two rounds of amplification (Figure 3.10B).

We wished to validate the use of 4 ng of starting material from LCM samples of mouse mammary tissue. We tested RNA quantities at higher (10 ng) and lower limits (2

ng), and found that this yielded similar bioanalysis profiles with similar transcript lengths (Figure 3.11). This indicated that as little as 2 ng of total RNA from LCM samples could be amplified without a significant loss in amplified RNA quality (Figure 3.11). However, amplified RNA yields were variable at 2 ng, therefore, in order to ensure that sufficient yields were reproducibly generated, the input amount of total RNA was increased to 4 ng. Furthermore, data from Table 3.1B indicated that sufficient total RNA yields are isolated from 5000 pulses in order to spare 4 ng per amplification reaction. Indeed, from 4 ng of total RNA, we routinely obtained 15-20 μg of amplified RNA (Figure 3.4).

3.2.6 LCM does not introduce a bias

We described the use of correlative controls as a means to indirectly assess the quality of total RNA from a given LCM sample. To further verify that total RNA isolated from LCM-dissected cells is not further degraded during the LCM process compared to non-dissected cells of the same sample, we used BioAnalyzer profiles of amplified RNA profiles as a read-out. Amplified RNA was isolated from a single tumor sample (4164T) that was either H&E-stained and LCM-dissected (LCM-dissected), or H&E-stained but not LCM-dissected (non-LCM-dissected) (Figure 3.12). Following two rounds of amplification, we observed that the bioanalysis profiles were remarkably similar between LCM-dissected and non-LCM-dissected samples with transcript lengths appearing between 0.2 and 0.5 kb (Figure 3.12). Thus, we determined that at this level of detection, LCM did not introduce bias in biological samples.

3.2.7 Preparation and testing of the common mouse reference

Since we were interested in comparing gene expression profiles from multiple mouse tumor and matched-normal samples, we used the common reference design

strategy. A commercially available mouse reference (Stratagene) was used for all hybridizations and consisted of 11 pooled murine cell lines. To determine whether the labeling biases were detectable within the mouse reference, we amplified 10 μg of mouse reference total RNA and labeled half of the sample with Cy3 and the other half with Cy5 and performed a self-hybridization (Figure 3.13A). The scatter plot shows that a labeling bias was not detected in the mouse reference samples labeled with different dyes (r^2=0.99, Figure 3.13B). Additionally, all hybridization experiments described in this thesis were co-hybridized with a mouse reference that was amplified and labeled in a single batch, aliquoted and stored until use. This provided confidence that the mouse reference would not introduce an additional source of error in microarray analyses.

3.2.8 Gene expression profiles can be generated from samples amplified, labeled and hybridized in different batches

Since biological samples would need to be processed in different amplification, labeling and hybridization experiments, we needed to determine whether this introduced a batch effect in the resulting gene expression data. To do this, total RNA from the same LCM-dissected mouse mammary tumor sample (4164T) was divided in three, and prepared for hybridization in parallel (Figure 3.14A). Each sample was hybridized against a common mouse reference in a two-color hybridization experiment with dye swaps (Figure 3.14A). Two samples were amplified in different batches, but labeled in the same batch (Figure 3.14A). Two samples were amplified in the same batch, but labeled in different batches (Figure 3.14A). This data showed that although amplification is a greater source of variability than labeling, both sources of bias were negligible (r^2=0.95 and 0.98, respectively, Figure 3.14B and C).

3.3 Discussion

The technologies and tools involved in gene expression profiling have progressed rapidly since its emergence in the 1990s (Fodor, Read et al. 1991; Schena, Shalon et al. 1995). Microarray technology has allowed for genome-wide quantification of mRNA abundance for a large number of biological specimens in parallel. This has enabled the evaluation of thousands of genes in a single experiment, and for the comparison of data across laboratories. Gene expression profiling has helped to understand the diversity and heterogeneity seen in human cancers and other disease states at a molecular level (Ramaswamy, Ross et al. 2003). In comparison to human colon and pancreatic tumors in which dominant oncogenic pathways have been defined, breast cancer is among the most heterogeneous in that no dominant pathway has emerged (Tavassoli 2003). Gene expression analyses have allowed for five molecular subtypes of breast cancer to be defined that correlate with distinct clinical outcomes (Sorlie, Perou et al. 2001; Sorlie, Tibshirani et al. 2003; Sotiriou, Neo et al. 2003). Furthermore, this technology has identified gene sets (i.e.: 70-gene prognosis profile) that predict the likelihood of disease recurrence (van 't Veer, Dai et al. 2002; van de Vijver, He et al. 2002; Ma, Wang et al. 2004).

One of the limitations of gene expression analyses is the heterogeneous nature of the source from which the RNA is isolated. Human breast cancers are composed of tumor stroma, epithelium and variable amounts of inflammatory cells. Thus, the resultant gene expression profiles represent an average of RNA abundance between different tumor compartments and this factor must be considered when interpreting gene expression results. LCM has addressed this issue by allowing for the dissection of specific populations of cells (Emmert-Buck, Bonner et al. 1996). Using this strategy, specific

biological questions have been investigated, including analysis of gene expression changes in breast cancer progression from the pre-invasive stages of atypical ductal hyperplasia (ADH) and ductal carcinoma *in situ* (DCIS) to invasive carcinoma (Ma, Salunga et al. 2003). Furthermore, gene expression analyses have been conducted to address factors implicated in metastasis by comparing gene profiles from morphologically normal epithelium, invasive carcinoma, and metastases (Sgroi, Teng et al. 1999). Studies have largely focused on the role of epithelial tissue in mammary tumorigenesis however, LCM was recently used for the procurement of tumor stroma from breast cancers (Finak, Bertos et al. 2008). Indeed, the role of the tumor stroma in human breast cancer was highlighted with the identification of the stroma-derived prognostic predictor (SDPP), a 26-gene prognostic factor that accurately identifies poor outcome individuals independently of published predictors and standard clinical prognostic factors.

Cell- and animal-based models of breast cancer have added another layer of understanding to mammary tumorigenesis that could not have been achieved by studying expression profiles in primary breast tumors alone. These *in vitro* and *in vivo* studies allow the contribution of specific oncogenic pathways to mammary tumorigenesis to be addressed in a genetically consistent background (Bild, Yao et al. 2006; Neve, Chin et al. 2006). Gene expression profiles of several well-established transgenic mouse models of breast cancer have been published and were found to cluster in an unbiased manner (unsupervised hierarchical clustering) with the molecular subtypes of human breast cancer, demonstrating that mouse models are representative of the spectrum of tumors seen in the human disease (Herschkowitz, Simin et al. 2007).

Although multiple mouse models of breast cancer with deregulated Met/HGF signaling have been reported (Table 1.1), the contribution of the Met/HGF oncogenic pathway to mammary tumorigenesis is poorly understood. We have used gene expression profiling to gain an understanding of the global gene expression changes involved in MMTV/Met^{mt}-induced mammary tumorigenesis compared to matched-normal tissue. Previous gene expression profiling studies using mouse models of breast cancer and normal tissue have used total RNA isolated from the whole tissue (Landis, Seachrist et al. 2005; Herschkowitz, Simin et al. 2007). Since tumors are composed of epithelium and stroma, and normal mammary tissue is mainly composed of fat, the resulting gene expression profiles represent a mixture of cellular components. In order to generate gene expression profiles driven by MMTV/Met^{mt} in mammary epithelium, we developed a microdissection approach to procure pure epithelial populations from tumor and matched-normal tissue. To our knowledge, this is the first description of gene expression analyses generated from a mouse model of breast cancer using LCM. We have shown that both mouse mammary normal and tumor tissue can be microdissected from which high quality total RNA can be isolated (Figures 3.3C), amplified through two rounds of linear amplification (Figure 3.10), labeled, and co-hybridized with a common mouse reference on whole genome mouse microarrays (Figure 3.14). We have described an H&E-staining protocol for frozen sections of mouse tissue that can be used for LCM that generates isolated RNA with high quality that is usable for downstream applications (Figures 3.2, 3.7B and 3.9). Furthermore, since multiple samples required analysis, we addressed whether LCM, amplification or labeling are subject to batch effect biases and determined that they were negligible (Figures 3.12 and 3.14). A common reference design (Figure 3.13) allows for the comparison between the MMTV/Met^{mt} mouse model and any other

mouse model co-hybridized to the same reference sample (Schade, Lam et al. 2007; Dourdin, Schade et al. 2008). Furthermore, on-going projects include the analysis of gene signatures from the tumor stroma of mouse models of breast cancer and their comparison to human datasets.

Overall, the data presented in Chapter 3 describes the methodology that was used to generate gene expression profiles from pure populations of cells from mouse mammary tumor and normal tissue. In the following chapter (Chapter 4), we analyze gene expression data generated from the MMTV/Met^{mt} mouse model of breast cancer. We compare the MMTV/Met^{mt} model to other mouse models of breast cancer, and additionally, determine its utility as a preclinical model of human breast cancer. In Chapter 5, we use MMTV/Met^{mt} gene expression data to validate a functional and biochemical role of candidate genes that may be involved in Met-mediated tumorigenesis.

A

Tumor	# cells counted	# pulses	# cells/pulse
4691T	500	0-609	0.82
	540	609-994	1.4
	600	994-1218	2.68
	544	0-441	1.23
	756	441-896	1.66
	644	896-1321	1.51
	585	49-546	1.18
	600	546-961	1.44
	620	961-1365	1.53
Total	**5392**	**3855**	**1.4**

4164T	724	0-303	2.39
	558	303-558	2.19
	637	558-866	2.07
	579	0-285	2.03
	758	285-551	2.85
	698	551-850	2.33
	622	0-305	2.04
	722	305-600	2.45
	600	600-803	2.12
Total	**5898**	**2599**	**2.3**

B

# of pulses	Total RNA yield (ng)
2000	33.56
5000	51.86
10000	67.2
20000	137

Table 3.1 Determination of cell number per laser pulse and RNA yield.
(A) The number of cells within nine 20x fields from two H&E-stained mammary tumors (4691T and 4164T) were counted (# cells counted). Three slides were used per mammary tumor (total=6 slides). Following this, all cells within same 20x fields were laser-captured and the number of laser pulses required to do so was recorded (# of pulses). The number of cells per pulse (# cells/pulse) was subsequently determined by dividing the total number of cells over the total number of pulses. Thus about 1.4-2.3 cells are dissected per laser pulse. (B) Using 4 different mammary tumors (559T, 4164T, 4474T, 4691T), 2000, 5000, 10 000, and 20 000 laser pulses were fired to determine total RNA yield.

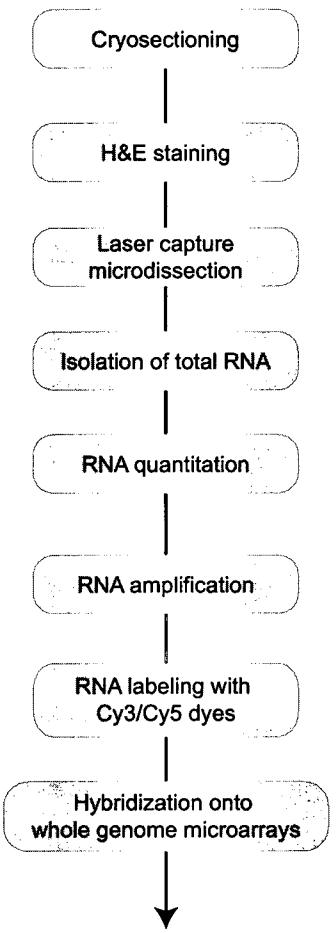


Figure 3.1: Schematic of methods from tissue cryosectioning to hybridization onto microarrays. Mouse mammary tumors and matched-normal tissues are cryosectioned, H&E-stained, laser capture microdissected, from which RNA is isolated, quantified, amplified, labeled with Cy3/Cy5 dyes and co-hybridized with mouse reference on Agilent's whole genome mouse expression arrays.

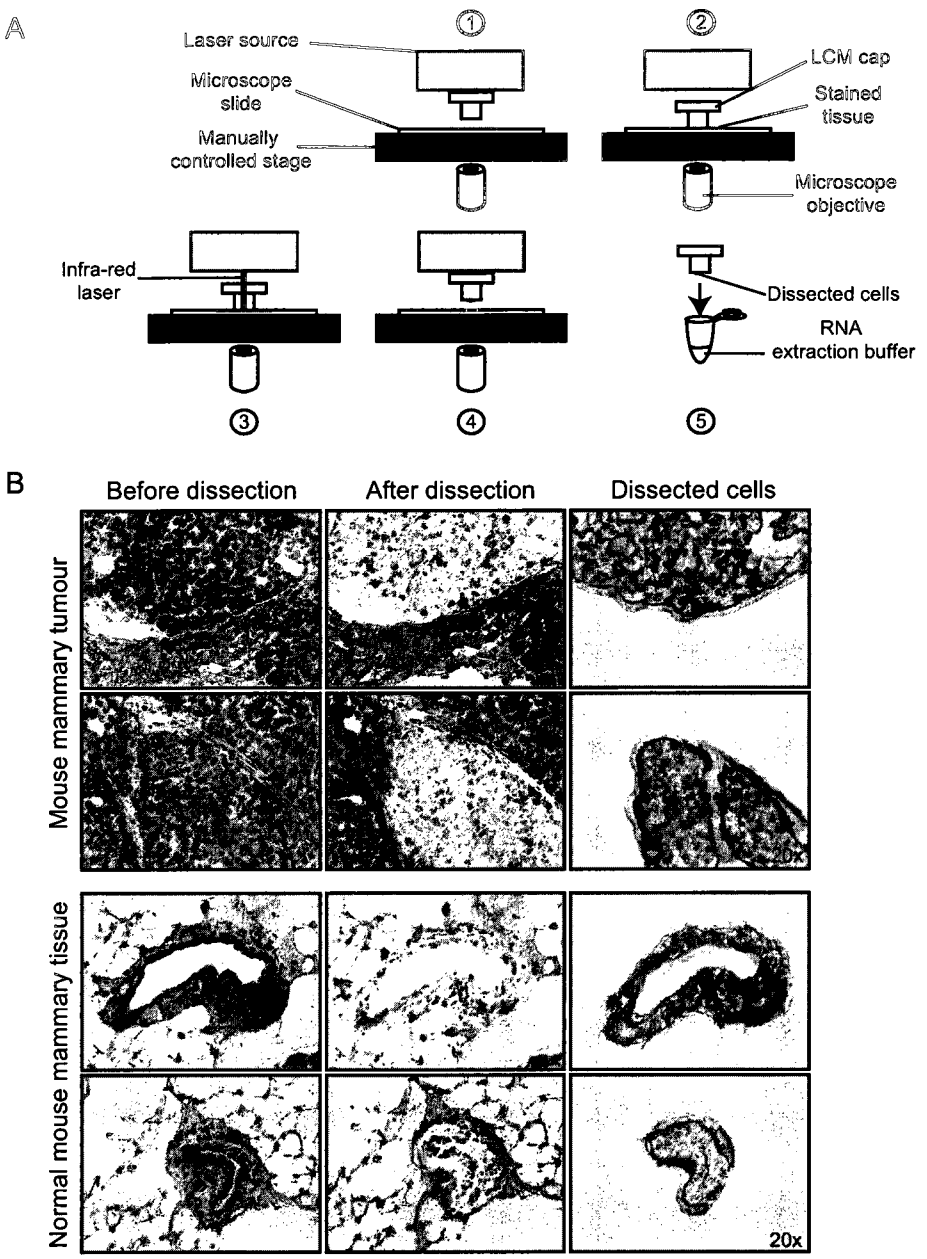


Figure 3.2: Laser capture microdissection of mouse mammary tissues. (A) The principles of laser capture microdissection (LCM) are represented in steps 1 to 5: *1*, Cells of interest are located on the microscope slide. *2*, LCM cap is lowered over cells of interest. *3*, An infra-red laser is fired and cells are captured. *4*, LCM cap and captured cells are lifted off the microscope slide. *5*, Captured cells are subjected to total RNA isolation by submerging in RNA extraction buffer. (B) Representative H&E-stained images of cryosectioned (9 μm) mouse mammary tumor and matched-normal tissues are shown before and after dissection, and of captured cells (magnification 20x).

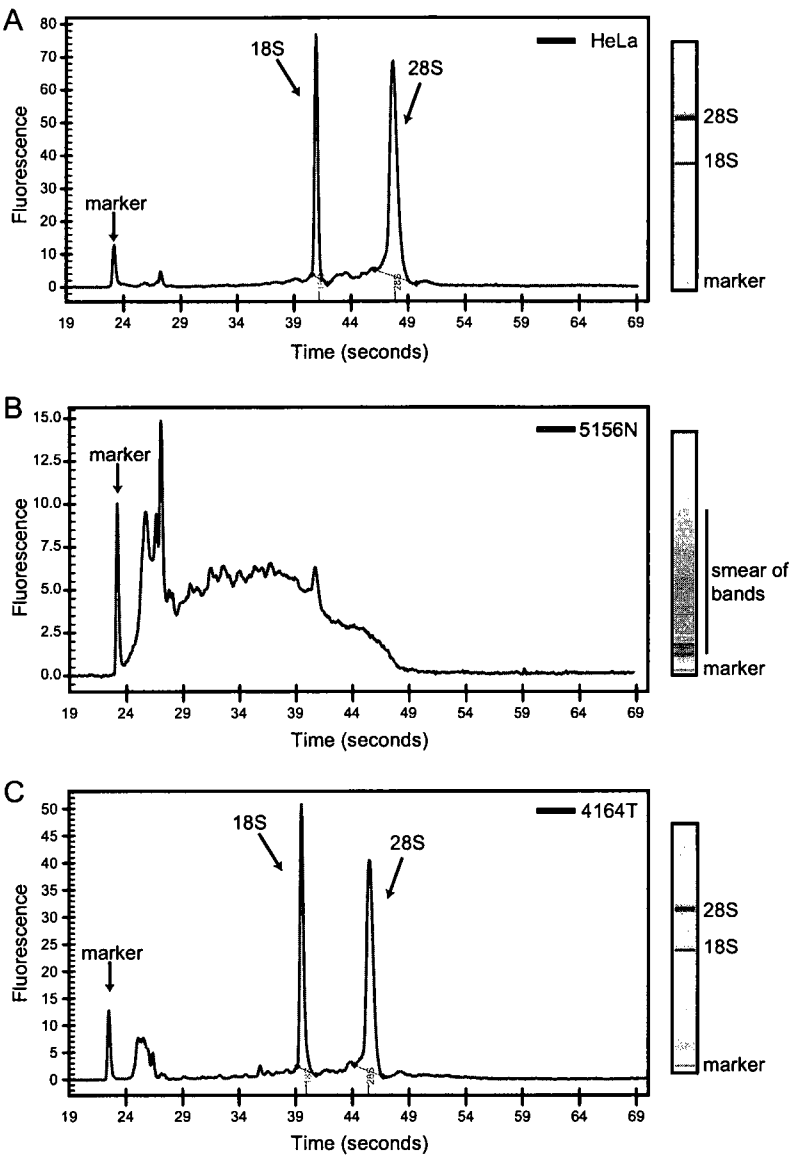


Figure 3.3: Bioanalysis profiles are used to assess total RNA quality. Electropherograms and gel-like images are shown for (A) a high quality sample of total RNA (HeLa cell line), (B) a low quality, degraded sample of total RNA from mouse tissue (5156N) and, (C) a high quality sample of total RNA from mouse tissue (4164T). See text for details.

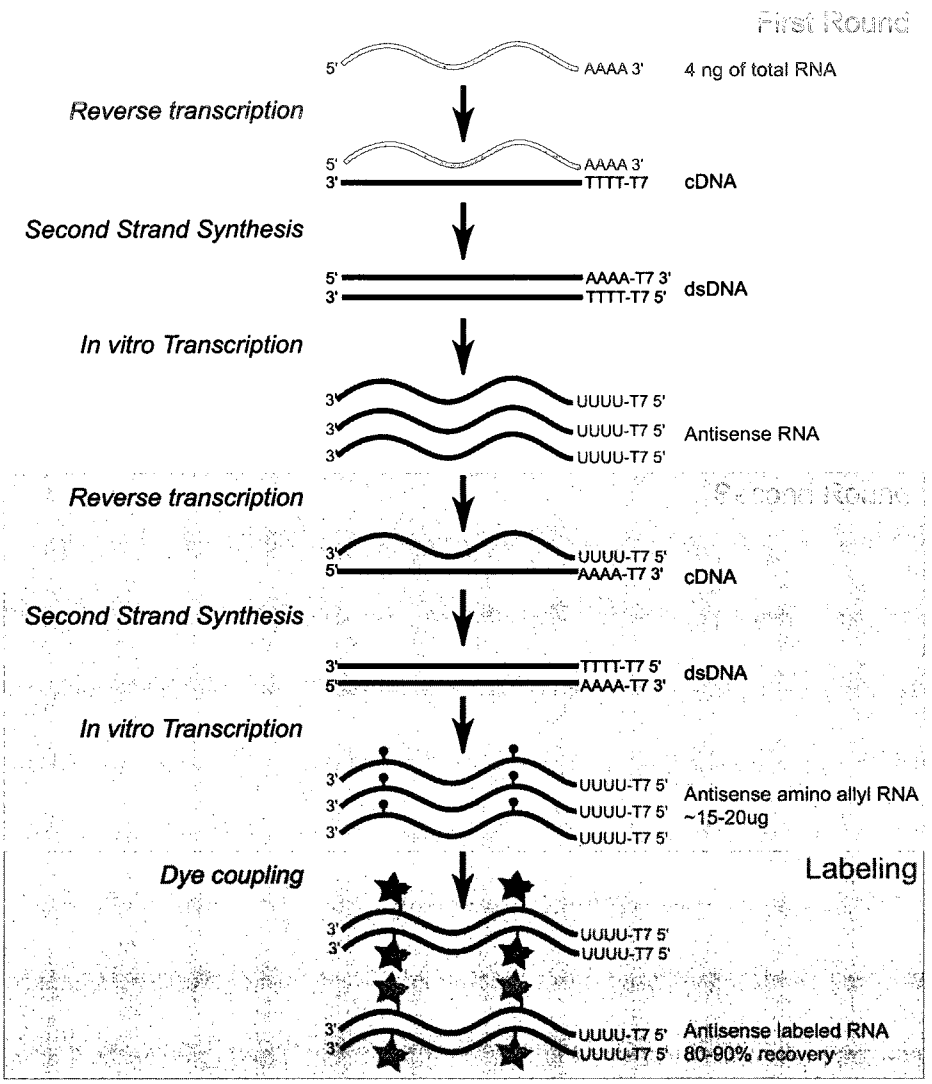


Figure 3.4: Two rounds of linear T7-based RNA amplification are required to process LCM-dissected tissue for microarray analysis. An input of 4 ng total RNA from LCM-dissected mouse mammary tumor or normal tissue is subjected to two rounds of linear amplification, producing between 15 and 20 μg of amplified RNA. Half of the amplified RNA yield is coupled to Cy3 and the other half to Cy5 dyes with 80-90% sample recovery.

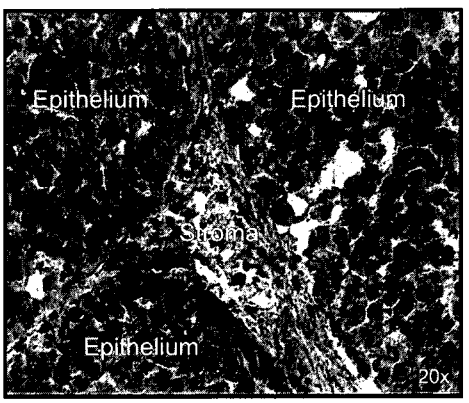


Figure 3.5: Mouse mammary tumors contain both epithelial and stromal tissue components. Frozen sections (9 μm) were stained with hematoxylin and eosin (H&E) (magnification 20x).

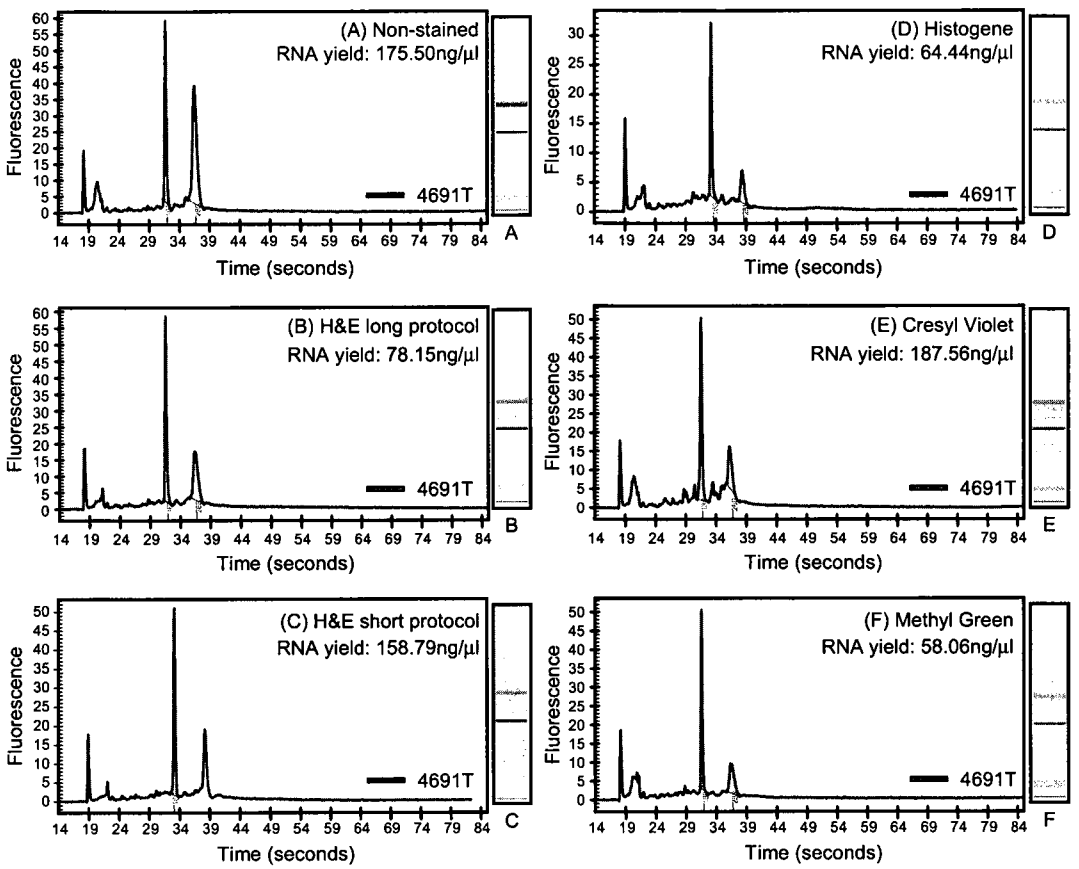


Figure 3.6: Tissue staining protocols were tested for their ability to preserve RNA integrity. Bioanalysis profiles are shown for total RNA isolated from mouse mammary tumor (4691T) that had been stained with (B) hematoxylin and eosin (H&E)-long, (C) H&E-short, (D) HistoGene (Arcturus), (E) cresyl violet, or (F) methyl green protocols. Bioanalysis profile from non-stained tissue is shown as a control (A). Total RNA yields obtained from each staining method are shown as determined by Nanodrop.

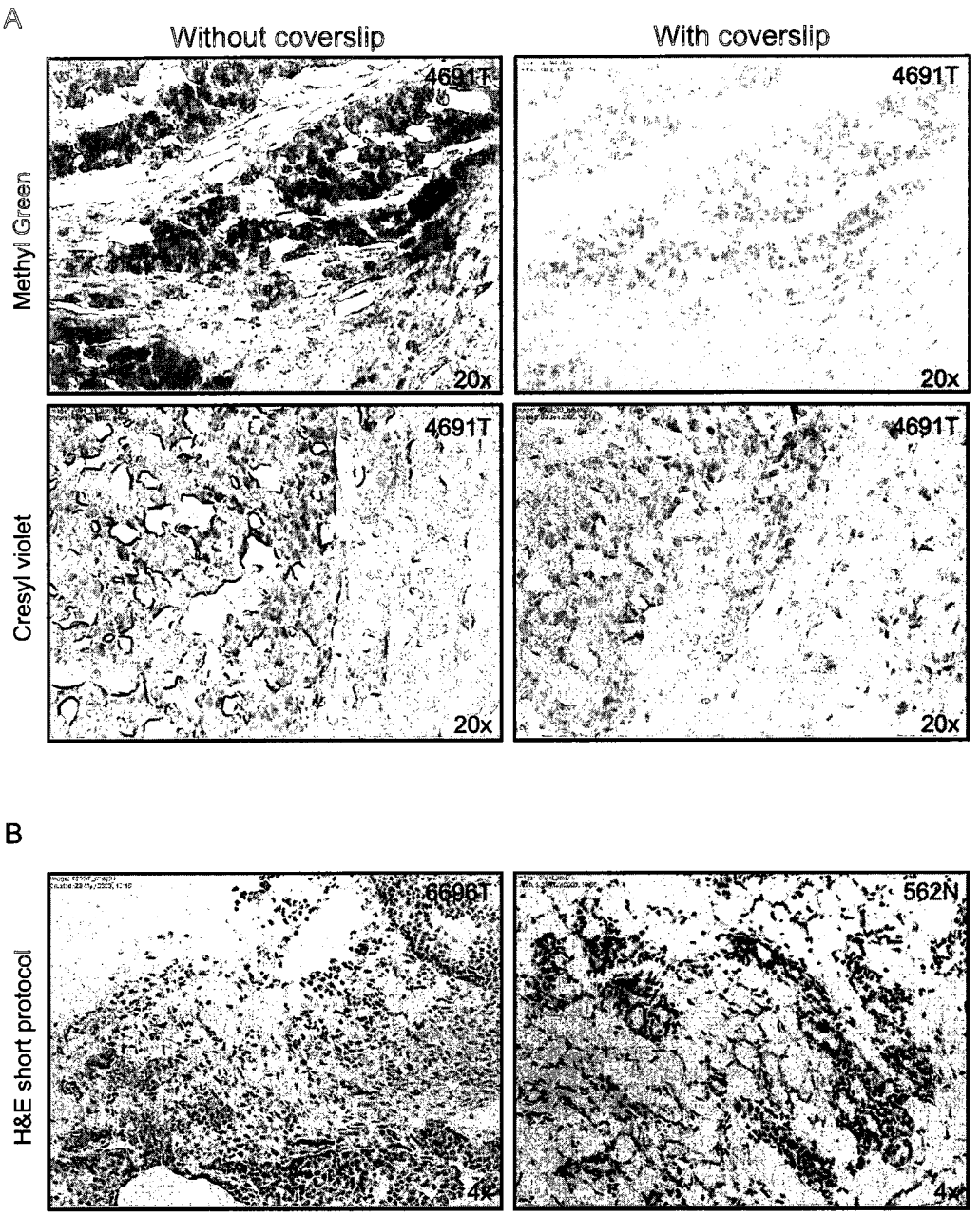


Figure 3.7: Staining protocols were tested for their suitability for LCM. (A) Mouse mammary frozen sections (9 μm thickness, 4691T) were stained with methyl green and cresyl violet, and images were captured with and without coverslips the Arcturus Pixcell II microscope station (magnification 20x). (B) Mouse mammary tumor (6696T) and normal (562N) tissues were stained with the 'short' hematoxylin and eosin (H&E) protocol and imaged with coverslips (magnification 4x).

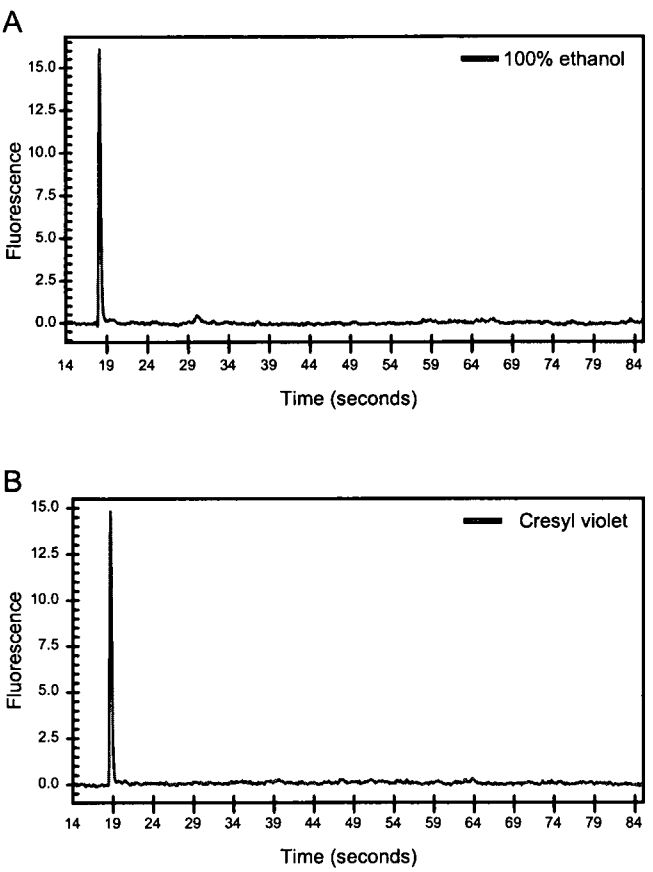


Figure 3.8: Bioanalysis profiles of 100% ethanol and cresyl violet. In an attempt to understand the additional peaks seen in the electrophoretic trace and gel-like image of Figure 3.6E, the bioanalysis profiles of potential contaminants, (A) 100% ethanol and (B) cresyl violet dye were determined.

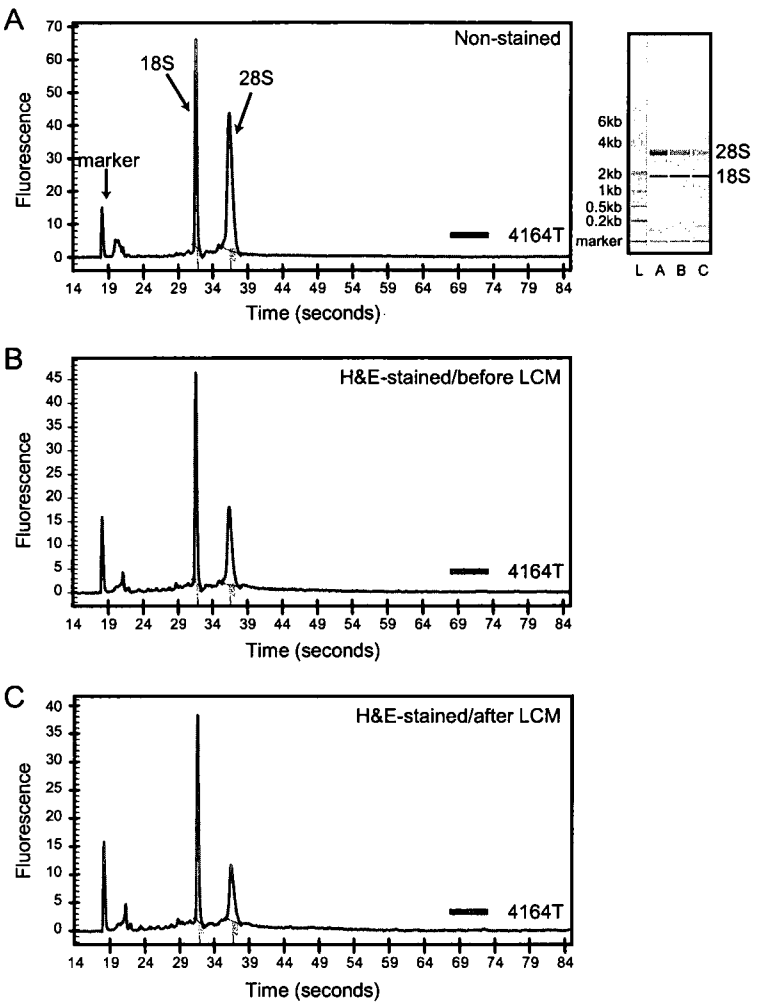


Figure 3.9: Correlative controls are used to assess RNA quality of the LCM-dissected sample. Electropherograms and gel-like images of total RNA are shown for a representative mouse tumor sample (4164T) (A) before staining (non-stained), (B) after H&E-staining and before LCM and, (C) after LCM. RNA quality, as judged by the ratio of 28S to 18S rRNA peaks, is acceptable for downstream microarray analysis.

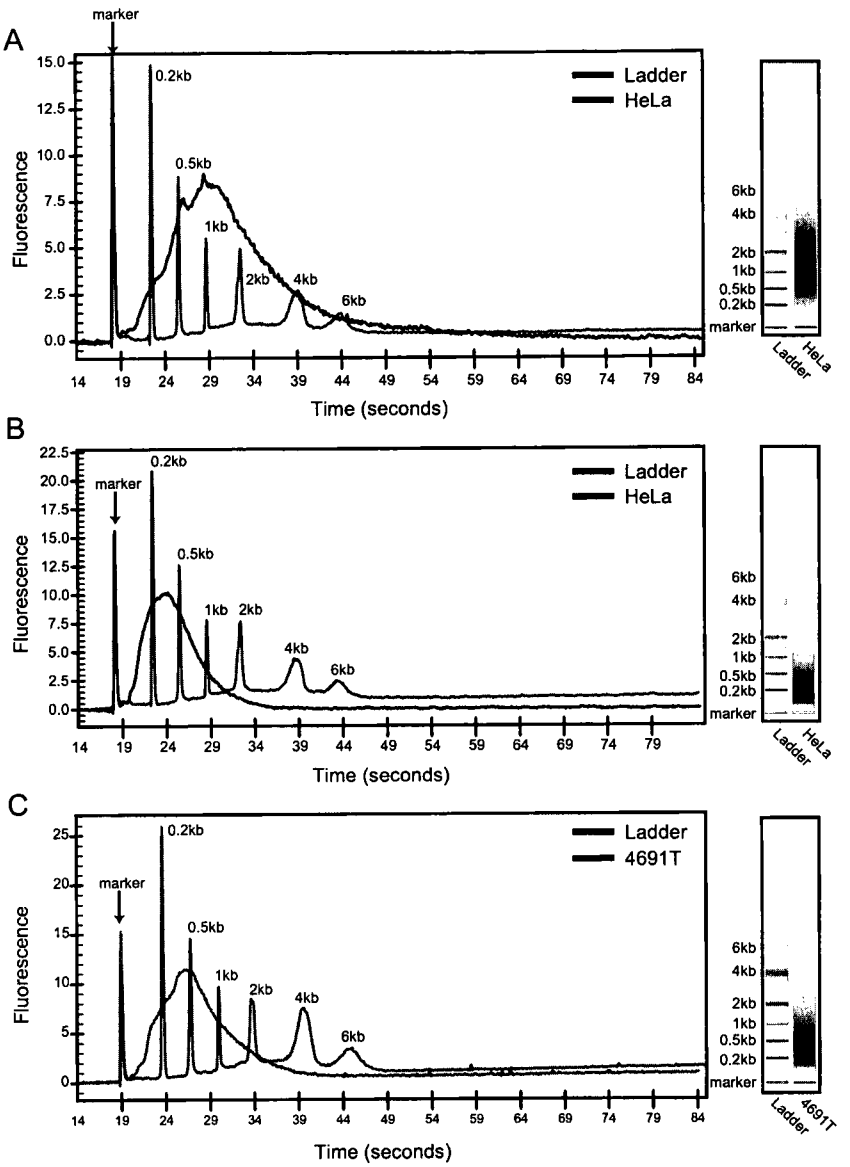


Figure 3.10: Bioanalysis profiles are used to assess quality of amplified RNA. Electropherograms and gel-like images are shown for (A) a high quality sample of amplified RNA that has undergone one round of amplification (HeLa cell line), (B) two rounds of amplification and, (C) a high quality sample of amplified RNA from mouse tissue (4691T) that has undergone two rounds of amplification.

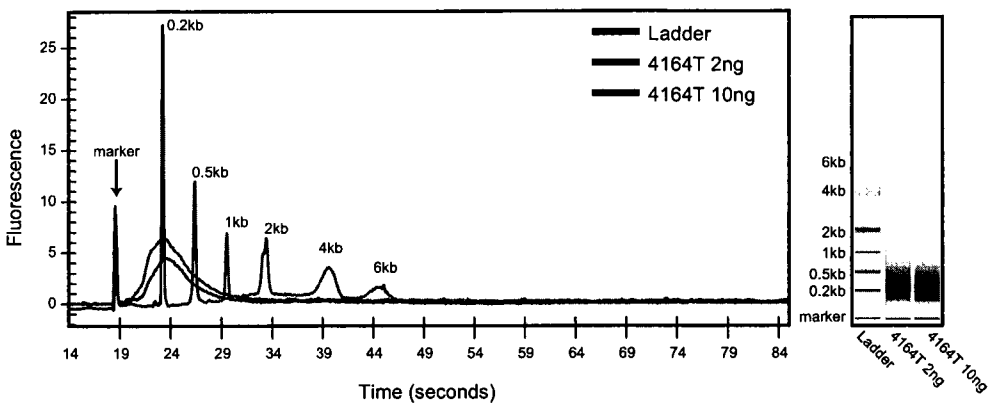


Figure 3.11: Determination of the amount of total RNA to use for amplification of LCM-dissected mouse mammary tissue. To determine the amount of input total RNA required to produce good quality amplified RNA, input quantities of 2 ng (purple trace) and 10 ng (green trace) were tested from the same mouse tumor (4164T). Electropherograms and gel-like images are shown.

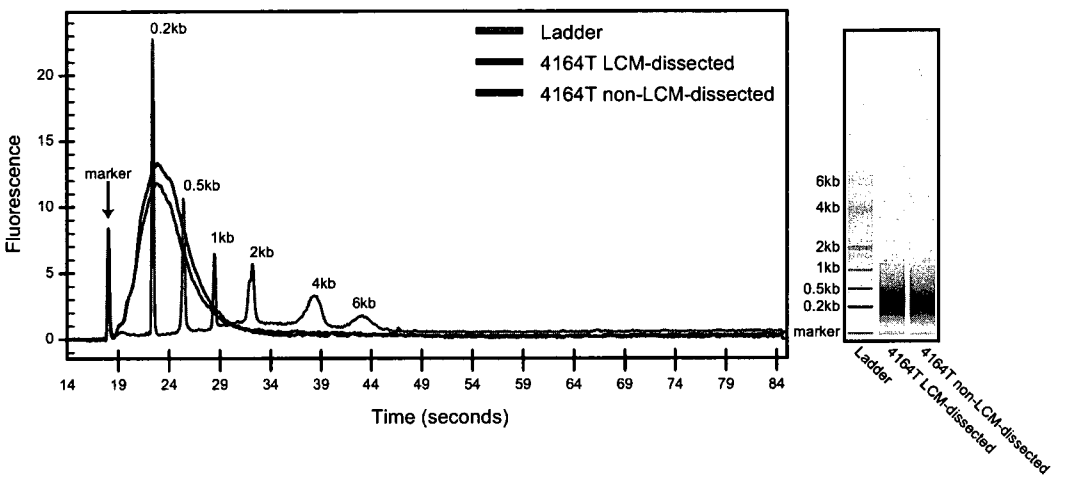
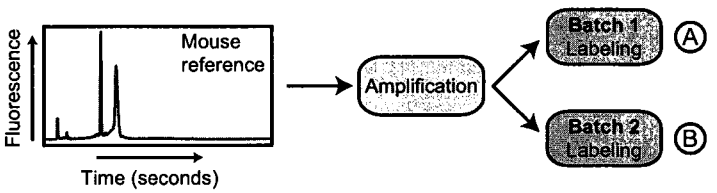


Figure 3.12: LCM does not introduce a bias as determined by bioanalysis of amplified RNA. RNA from mouse mammary tumor (4164T) was isolated from a laser capture microdissected (blue trace) and a non-LCM-dissected (brown trace) sample. Electropherograms and gel-like images are shown.

A



B

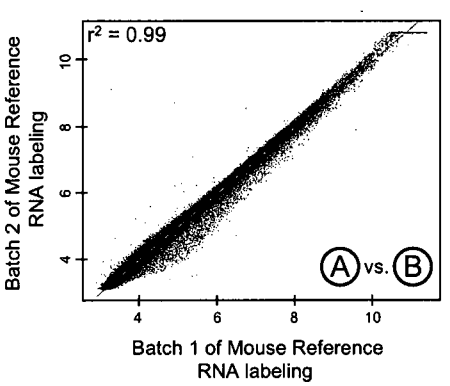


Figure 3.13: Labeling biases are not detected in mouse reference RNA. (A) Schematic showing the experimental design to test for labeling bias in mouse reference RNA. (B) RNA from mouse reference was amplified in one batch, and labeled in two different batches and co-hybridized to whole genome mouse microarrays.

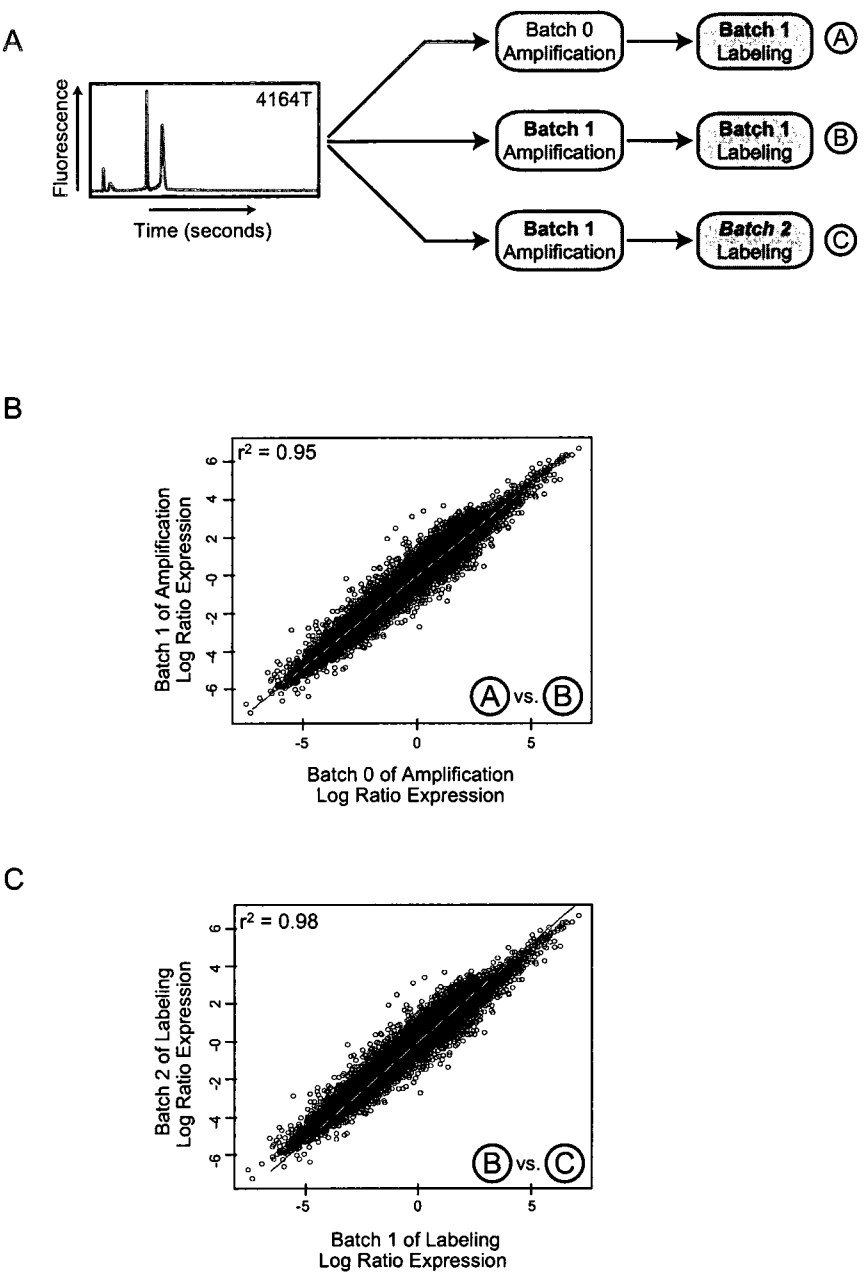


Figure 3.14: RNA amplification and labeling do not introduce batch effect biases. (A) Schematic showing experimental design to test for amplification and labeling batch effects. (B) RNA from a mouse mammary tumor (4164T) was amplified in two different batches, labeled in the same batch, and hybridized on whole genome mouse microarrays. (C) RNA from a mouse mammary tumor (4164T) was amplified in the same batch, labeled in different batches and hybridized onto microarrays. Mouse mammary tumor samples were co-hybridized with a common mouse reference.

Chapter 4

4 Mammary tumors induced by the Met receptor exhibit diverse histology and are associated with poor outcome and basal breast cancers

4.1 Introduction

Despite advances in early detection and the implementation of screening guidelines, breast cancer remains the second most frequent cause of cancer mortality in women. Breast cancer is a heterogeneous disease that comprises distinct biological entities that are correlated with diverse clinical outcomes and responses to treatment. Gene expression profiling as well as molecular pathology have revealed that the heterogeneity of breast cancers naturally divides into the luminal, ERBB2 positive, and basal-like subtypes (Perou, Sorlie et al. 2000; Sorlie, Perou et al. 2001; Da Silva, Clarke et al. 2007). These subtypes were named to reflect gene expression patterns of the two principal cell types of the differentiated breast, luminal epithelial cells lining the duct and lobule, and myoepithelial cells, which form a single layer surrounding the luminal cells. The luminal subtype comprises approximately 60% of breast cancers, is estrogen receptor (ESR1) positive, expresses ESR1 responsive genes and luminal markers such as keratin 8/18. Up to 30% of breast cancers are identified with overexpression/amplification of the ERBB2 receptor tyrosine kinase and these tumors are generally ESR1/progesterone receptor (PGR) negative. The basal group is characterized as ESR1/PGR/ERBB2 negative and is frequently positive for basal keratins 5/6 (Nielsen, Hsu et al. 2004; Livasy, Karaca et al. 2006). Breast cancers within the luminal subtype receive anti-estrogen therapies and tend to have good prognosis. Although ERBB2 amplification is associated with aggressive disease and poor outcome, anti-ERBB2 therapies are increasing the survival of this subgroup. Due to the lack of treatment options, patients

within the basal subtype historically have poor prognosis (Perou, Sorlie et al. 2000; Sorlie, Perou et al. 2001). Hence, an understanding of the signalling pathways active in these tumors is crucial for the generation of targeted therapies.

The Met receptor tyrosine kinase, which is the receptor for Hepatocyte Growth Factor/scatter factor (HGF/SF), is expressed at elevated levels in 15-20% of human breast cancers (Camp, Rimm et al. 1999), and is a prognostic factor for poor outcome (Ghoussoub, Dillon et al. 1998; Lengyel, Prechtel et al. 2005). High levels of the Met receptor ligand, HGF/SF, in the serum of breast cancer patients is also correlated with a shorter disease-free interval following surgery (Yamashita, Ogawa et al. 1994) and a higher tumor/lymph node/metastasis score (Taniguchi, Toi et al. 1995).

Met controls a program of invasive epithelial growth through the co-ordination of cell proliferation, survival, cell migration and epithelial morphogenesis (Birchmeier, Birchmeier et al. 2003). These processes are important during embryogenesis and for organ regeneration in adults. However, when deregulated, the HGF/Met signalling axis contributes to tumorigenesis and metastasis (Peschard and Park 2007). Although several transgenic mouse models have examined the tumorigenic capacity of Met receptor signalling, none have targeted the expression of Met specifically to the mammary epithelium. However, when either HGF or an activated Met receptor is expressed under the constitutively active metallothionein promoter, transgenic mice develop malignancies, including mammary tumors (Liang, Reid et al. 1996; Takayama, LaRochelle et al. 1997; Jeffers, Fiscella et al. 1998). Additionally, multiparous transgenic mice expressing HGF under the mammary-specific whey acidic protein (WAP) promoter developed mammary tumors (Gallego, Bierie et al. 2003). These transgenic mice highlight the susceptibility of the mammary epithelium to transformation by an enhanced Met/HGF signal.

Here we describe a murine model of breast cancer generated by the expression of weakly transforming mutants of the Met receptor tyrosine kinase in the mammary epithelium. We demonstrate that Met^{mt} induces mammary carcinomas with diverse histopathologies, which based on immunohistochemistry and expression profiling, includes tumors with basal and luminal characteristics. By performing microarray analyses, we reveal that Met^{mt}-induced basal tumors show expression of basal keratins, enrichment for markers of the Wnt pathway, display features indicative of epithelial to mesenchymal transition (EMT), cluster with human basal breast cancer and murine models of basal-like breast cancer. In human breast cancer, Met shows consistent elevated expression in the basal subtype and identifies patients with poor outcome. Our transgenic mouse model, coupled with human breast cancer data, identifies MET as a rational therapeutic target for both basal and aggressive breast cancer. Furthermore, this model provides a tool to understand a molecular mechanism that can induce basal breast cancer and such models are currently lacking.

4.2 Results

4.2.1 MMTV/Met^{mt} induces tumors with multiple histopathologies

To study the role of the Met receptor in mammary tumorigenesis, we have generated a transgenic mouse model that expresses the Met receptor in mammary epithelium driven by the murine mammary tumor virus (MMTV) promoter. Mice transgenic for wild type Met and oncogenic variants of the Met receptor (M1248T, Y1003F/M1248T), hereafter called Met^{mt} (Jeffers, Fiscella et al. 1998; Abella, Peschard et al. 2005), develop tumors with moderate penetrance (9-40%) and long latency (381-475 days) (Table 4.1) (Petkiewicz 2007). MMTV expression is greatly enhanced

postpartum, during puberty, pregnancy and involution (Cardiff and Muller 1993), and consistent with this, tumor induction was observed predominantly in multiparous mice (Table 4.1). Notably, in addition to carcinomas with solid nodular histopathology (~47%) that are most common to MMTV/Neu and MMTV/PyMT models (Cardiff and Muller 1993; Rosner, Miyoshi et al. 2002), Met^{mt} also induced tumors with papillary, scirrhous, adenosquamous, and spindle cell histopathologies (~53%) (Figure 4.1A, Table 4.2).

To confirm that mammary tumors in Met^{mt} mice resulted from integration of the transgene, we first performed PCR analysis on DNA prepared from tumors and matched-normal tissue and demonstrated the presence of the Met transgene in both (Figure 4.2A). Using an antibody specific for the transgene, Met^{mt} protein was detected only in tumor tissue at variable levels, but was undetectable in tumors possessing spindle cell pathology (Figure 4.2B). This is consistent with previous observations demonstrating suppression of the MMTV promoter in spindle cell tumors arising in MMTV/ILK transgenic mice and in others with an EMT spindle tumor phenotype (White, Cardiff et al. 2001; Damonte, Gregg et al. 2007).

To address the utility of the Met^{mt} transgenic mice as a preclinical model, we examined the ability of epithelial cell lines derived from Met^{mt} tumors to invade and migrate (Figure 4.3). Using a specific small molecule inhibitor for the Met receptor tyrosine kinase (PHA-665752) as well as siRNA to knockdown Met expression, we demonstrate that elevated Akt and Erk1 and 2 activity is dependent on the Met receptor (Figure 4.3E and F). In addition, the invasive (Figure 4.3A and B) and migratory (Figure 4.3C and D) capacity of these cell lines is also dependent on Met. These results provide strong support that tumors derived from the Met^{mt} transgenic mice are indeed Met dependent, a criterion required for effective preclinical models.

To better understand the molecular characteristics associated with the distinct histopathologies found in Met^{mt} tumors, we performed gene expression analyses using Agilent whole genome mouse arrays. Tumors of the various histopathologies and matched-normal epithelium were analyzed. Since Met^{mt} tumors contained variable stromal content, LCM was utilized to enrich for tumor cells (Figure 3.2B). Unsupervised hierarchical clustering revealed that the arrayed samples naturally fell into three distinct classes. The first major division within the data separated all normal samples ("normal epithelium", green) from tumor samples (Figure 4.1B). Within the group of tumor samples, two subclusters formed; the first subcluster consisted of tumors with a solid phenotype ("solid", purple), and the second consisted of tumors with papillary, scirrhous, adenosquamous, and spindle cell pathologies ("mixed pathology", red) (Figure 4.1B).

To understand the differences between these tumor types, genes differentially expressed between the clusters were analyzed for overrepresentation of specific biological pathways (Figure 4.4). Genes overrepresented in Met^{mt} mixed pathology tumors (red, Figure 4.1B) compared to the Met^{mt} solid tumors (purple, Figure 4.1B) were related to gene ontology (GO) categories involving epithelial and mesenchymal cell differentiation, including genes linked to EMT. Met^{mt} mixed pathology tumors also showed overrepresentation of genes involved in tissue remodelling, collagen production, cytokine signalling, cell migration, angiogenesis and inflammatory response, and showed enrichment for Wnt, integrin, and TGFß signalling pathways (Figure 4.4A, Table 4.3). In contrast to this, Met^{mt} tumors with solid histopathologies (purple, Figure 4.1B) showed enrichment for GO terms related to regulation of apoptosis, cell adhesion, and small GTPase signal transduction (Figure 4.4B, Table 4.4), supporting the interpretation that distinct histopathologies are associated with gene expression profiles that represent

different biological processes. Similar results were obtained when molecular pathway analyses was carried out using GSEA and KEGG (Tables 4.5 to 4.8).

4.2.2 Met^{mt} mixed pathology tumors cluster with basal-like murine mammary tumors

Numerous transgenic mouse mammary tumor models display diverse histopathologies that reflect changes observed in human breast cancer (Cardiff 2001). Many of these have recently been analysed at the gene expression level and were found to correlate with the subtypes defined in human breast tumors including those with a luminal as well as those with a basal gene expression profile (Herschkowitz, Simin et al. 2007). To determine whether Met^{mt}-induced tumors resemble a specific subtype of mammary tumors, we utilized the "866 intrinsic gene" set found to be consistently representative of classes of murine mammary tumors (Herschkowitz, Simin et al. 2007). Our samples were combined with this dataset, which contains 13 other mouse mammary tumor models, and the samples were hierarchically clustered over the intrinsic gene set. This analysis further confirmed that the two subtypes of Met^{mt} tumors represent distinct molecularly defined groups (Figure 4.5). Met^{mt} tumors with solid histopathologies clustered closely together with well-established mouse models also having solid histopathologies, including the MMTV/Neu and MMTV/PyMT transgenic models (Figure 4.5). In contrast, the majority of mixed pathology Met^{mt} tumors clustered within a group that included $Trp53^{-/-}$ transplants, $Trp53^{+/-}$ irradiated (IR), $C3(1)/T_{ag}$, and WAP/T_{121} tumors, which have been shown to express basal and mesenchymal related genes (Herschkowitz, Simin et al. 2007). These models are consistent with EMT and mesenchymal pathways enriched in the mixed pathology Met^{mt} tumors (Figures 4.4A, 4.9A).

To assess if the Met^{mt} mixed pathology tumors have large-scale transcriptional similarities to human breast cancers, we developed a hierarchical cluster of these Met^{mt} tumors with 172 human breast tumors using the cross-species intrinsic gene set (Herschkowitz, Simin et al. 2007). Of the 106 genes in the original intrinsic gene set, 104 were found to be present in our dataset. The results of the analysis support the interpretation that Met^{mt} mixed pathology tumors reflect a basal subtype, as these tumors clustered with human basal breast cancers, whereas Met^{mt} tumors with solid histopathologies clustered among human tumors of the luminal subtype (Figure 4.6). To further explore these relationships, the correlation between each Met^{mt} murine tumor sample and each centroid for the three human subtypes was quantified (Figure 4.7). Consistent with the hierarchical clustering results, the Met^{mt} mixed pathology mouse tumors all correlated highest with the human basal subtype, while the Met^{mt} solid mouse tumors all correlated highest with the human luminal subtype (Figure 4.7). Hence, these results provide support that the Met^{mt} mixed pathology tumors reflect the human basal subtype, whereas Met^{mt} tumors with solid histopathologies reflect human luminal tumors.

4.2.3 Met^{mt} mixed pathology tumors express markers of EMT and basal-like breast cancer

Tumors with solid histopathologies are most commonly observed when receptor tyrosine kinases, such as ErbB2, or oncoproteins that activate the Ras pathway are driven by the MMTV promoter (Rosner, Miyoshi et al. 2002). Thus, tumors of mixed pathology induced by MMTV/Met^{mt} were of interest since they may reflect a response to specific Met signals. To better understand the cellular composition of mixed pathology Met^{mt} tumors, we investigated markers of distinct cell lineages. Examination of genes in the array data showed enrichment for expression of several basal markers, including keratins

(*Krt)* 5, 6a, 14, 15, and 17 and, from validation, Met^{mt} mixed pathology tumors, but not solid tumors, stained immunopositive for basal keratins *(5, 6, 14)* (Figures 4.8 and 4.9). In contrast, expression data from Met^{mt} tumors with solid histopathologies were enriched in luminal markers, such as *Gata3*, keratins 8 and 18 (Figure 4.9A), and were negative by immunostaining for markers of basal breast cancer (Figure 4.8). Interestingly, cells within regions of the mixed pathology Met^{mt} tumors showed co-expression of Met^{mt} transgene with basal marker keratin 5 (Figure 4.10B), indicating a correlation between Met transgene expression and basal markers, in contrast to Met^{mt} solid tumor controls (Figure 4.11).

Additional genes associated with a basal-like phenotype, such as *Egfr, Trp63 (*Trp63), members of the Wnt pathway, ß-catenin (*Ctnnb1*), and transcription factors Tcf/Lef (*Tcf4, Tcf7, Lef1*) (Nielsen, Hsu et al. 2004; Livasy, Karaca et al. 2006), were elevated in Met^{mt} mixed pathology tumors, but not Met^{mt} tumors with solid pathologies (Figure 4.8 and 4.9). In serial sections, regions of Trp63 immunopositivity corresponded to regions of Trp53 nuclear staining (Figure 4.9B), a protein frequently mutated in basal breast cancer and often associated with high nuclear grade (Rakha, Putti et al. 2006). Both Trp53 and Trp63 were immunonegative in Met^{mt} solid tumors, which reflect a luminal lineage (Figure 4.8). In support of activation of the Wnt pathway, nuclear localization of ß-catenin was observed in Met^{mt} tumors of mixed pathology (Figure 4.9B), whereas membranous localization of ß-catenin was observed in Met^{mt} solid tumors (Figure 4.8). Moreover, mixed pathology Met^{mt} tumors showed elevated expression (*Vim, Snail, Snai2, Acta2*) and immunopositivity of EMT and mesenchymal-associated genes (*Vim, Acta2*) (Figure 4.9), and these were absent in Met^{mt} solid tumors (Figure 4.8). Consistent with features found in human basal-like breast cancers (Livasy, Karaca et al.

2006), the Met^{mt} mixed pathology tumors were poorly differentiated, some showed elements of squamous metaplasia (Figure 4.10A*i*), geographical necrosis (Figure 4.10A*ii*), lymphocytic infiltration (Figure 4.10A*iii*), high mitotic index (Figure 4.10A*iv*), atypical mitotic figures (Figure 4.10A*v*), and nuclear pleomorphism (Figure 4.10A*vi*). Together these data demonstrate that mixed pathology Met^{mt} tumors share histopathological features and express basal protein markers consistent with human basal breast cancers.

4.2.4 Met correlates with human basal breast cancer and poor prognosis

To establish if *MET* expression correlates with specific subtypes in human breast cancer we examined gene expression data obtained from tumor epithelium of human breast cancer patients. Patients included in this dataset were lymph node status positive and negative and patient characteristics are published elsewhere (Finak, Sadekova et al. 2006; Finak, Bertos et al. 2008). The 54 human breast cancers within this dataset segregated into three main molecular subtypes as defined by immunohistochemistry: luminal (ESR1+ and/or PGR+), ERBB2+ (ERBB2+ and ESR1-/PGR-) and basal (ESR1-/PGR-/ERBB2-). Consistent with these subdivisions, the mean expression of *ESR1* and *PGR*, as well as luminal-specific genes *KRT8* and *KRT18*, were highest among luminal samples, and were significantly different among the three subtypes. As expected, expression of *ERBB2* was highest among the *ERBB2* positive subtype (Figure 4.12). In contrast to this, the mean expression of basal-specific markers, *KRT6B* and *EGFR* were highest among the basal population (Figure 4.12). Notably, the mean expression of *MET* was found to be highest in basal samples, with significant differences among the groups (p-value=3.64e-06, Figure 4.12). However, when luminal samples were excluded from

the analysis, the expression of *MET* was not significantly different between the basal and *ERBB2* positive samples (p-value=2.89e-01), indicating that *MET* expression may impact both breast cancer subtypes.

While MET itself was discovered to have an association with the human basal subtype, we also wished to determine whether the same is true of the downstream transcriptional targets of the MET/HGF signaling axis. To establish if the MET molecular pathway is activated in a particular subtype of human breast cancer, we applied a MET transcriptional signature (Kaposi-Novak, Lee et al. 2006) to an independent human breast cancer dataset (NKI dataset) (van de Vijver, He et al. 2002). The MET signature clearly separated human breast tumors into two primary clusters (Figure 4.13), with the smaller cluster (turquoise) determined to represent an induction of the MET transcriptional response when compared to the larger cluster (orange). Genes expected to be overexpressed during MET activation were significantly higher in the turquoise cluster (Figure 4.13B, p-value<2.20e-16), while genes expected to be repressed by MET activation were significantly lower in this cluster (Figure 4.13B, p-value=1.90e-09). Notably, tumors belonging to the basal subtype of breast cancer were overrepresented in the turquoise cluster identified by the MET transcriptional signature (p-value=2.99e-24), supporting the association of MET with the basal group. The cluster of all MET activated tumors had significantly worse overall prognosis (Figure 4.13C, p-value=1.00e-04), and furthermore the MET signature was able to identify poor outcome within the basal subtype (Figure 4.13D, p-value=3.19e-02). Unfortunately, there were insufficient luminal and ERBB2 positive samples in the MET activated group (n=3, n=2, respectively) to assess whether the MET signature was associated with poor prognosis in these subtypes.

In support of our results, elevated MET protein levels have previously been correlated with poor outcome in human breast cancer (Ghoussoub, Dillon et al. 1998; Tolgay Ocal, Dolled-Filhart et al. 2003; Lengyel, Prechtel et al. 2005). Furthermore, in cDNA microarray analyses conducted in a prospectively accrued cohort of women with axillary lymph node negative (ANN) breast cancer (Andrulis, Bull et al. 1998), our collaborators (Irene Andrulis, Shelley Bull) identified *MET* as a gene differentially expressed between tumors from 48 women who experienced disease recurrence within 4 years of diagnosis compared to 41 who remained disease-free for over ten years (Table 4.9, p-value=1.33e-02 by standard t-test, p-value=8.70e-03 by non-parametric Wilcoxon rank sum test in univariate analysis), an association that was independent of traditional clinicopathological parameters, (Supplemental text, p-value=3.10e-02).

To further investigate the correlation between transcriptional activation of *MET* and the basal subtype, we stained for MET protein in a cohort of 668 ANN human breast cancer cases (in collaboration with Frances O'Malley, Irene Andrulis, Shelley Bull) (Mulligan, Pinnaduwage et al. 2008). As shown previously (Lengyel, Prechtel et al. 2005), tumor epithelium was positive for MET immunostaining (cytoplasmic and perinuclear staining), with an absence of staining in the stroma (Figure 4.14A). High MET immunostaining was observed in all histological subtypes but a significantly greater proportion of basal subtype tumors were MET positive (65%) as opposed to MET negative (35%), when compared to the ERBB2 positive and luminal subtypes (in collaboration with Frances O'Malley, Irene Andrulis, Shelley Bull) (Figure 4.14A, inserted Table, p-value=6.51e-03). Basal tumors were defined as ESR1, ERBB2 negative, and keratin 5 positive. Consistent with previous results, the Kaplan-Meier plot

demonstrates that MET positive tumor status correlates with poor disease-free survival outcome among ANN patients (Figure 4.14A, Log-Rank p-value=3.94e-02). Additionally, MET positive protein status appears to be associated with poor outcome irrespective of breast cancer subtype (Supplementary text and Figure 4.15, p-value=6.31e-01 by Cox model test for differences between basal and non-basal, Table 4.10 and Figure 4.15A, B), which is consistent with the idea that MET expression is causally related to an aggressive phenotype.

Since we observed a strong induction of EMT genes including *Snail* in the murine Met^{mt} mixed pathology tumors (Figure 4.9A), as well as the induction of *SNAIL (SNAI1)* mRNA following stimulation of normal epithelial cells with HGF (Figure 4.16), we investigated (in collaboration with Frances O'Malley, Irene Andrulis, Shelley Bull) co-expression of MET with the downstream EMT target, SNAIL. Notably, tumors exhibiting co-expression of MET with SNAIL showed a significant correlation with poor outcome when compared to those negative for either MET and/or SNAIL (Figure 4.14B, Log-Rank p-value=7.40e-03), demonstrating that the combination of MET with an EMT signal, SNAIL, is a strong predictor of poor outcome (Figure 4.14B). This association persisted after adjustment for traditional histopathologic prognostic factors (Tables 4.11 and 4.12) and for basal/non-basal status (Supplementary text, Table 4.13). Together these results strongly support a role for MET signaling both in human basal breast cancers and breast cancers with poor outcome.

4.3 Discussion

Understanding the oncogenic pathways that distinguish subsets of human cancer is critical for the development of new therapies. Gene expression analyses of human

breast cancers have provided new understanding of the different molecular subtypes of the disease. At least three major tumor subtypes have been identified, defined as ESR1 positive (luminal), ERBB2 positive, or ESR1/PGR/ERBB2 negative (basal) breast cancers. Tumors of the basal subtype lack targeted treatment and are a clinical challenge. Here we show that mammary-specific expression of weakly oncogenic forms of the Met receptor promotes the formation of mammary carcinomas with the induction of mammary tumors with histologies, gene expression and immunohistochemical profiles similar to human basal breast cancers.

Although elevated MET protein levels have been associated with poor outcome in human breast cancer, the role for the MET receptor tyrosine kinase in the induction and development of breast cancer was poorly understood. Mammary-specific expression of MMTV/Met^{mt} surprisingly induced tumors with distinct histological phenotypes (solid and mixed pathology). The expression profile derived from the Met^{mt} mixed pathology tumor subtype clusters these with human basal breast cancers and murine models of basal-like tumors (Figures 4.5, 4.6). This observation is consistent with findings that these Met^{mt} murine tumors display histopathologies similar to the human basal subtype (Figure 4.10A) and express a series of known basal markers (Figure 4.9A, B).

When compared with other murine models, gene expression analyses revealed that the Met^{mt} mixed pathology tumors clustered with murine tumors having basal characteristics including MMTV/Cre;$Trp53^{+/-}$, TgWAP/T_{121}, TgWAP/T_{ag} and TgC3(1)/T_{ag} tumors (Figure 4.5). Within this class, Met^{mt} tumors shared expression and immunopositivity of genes implicated in EMT, including *Tgfbi*, *Snai1, Snai2 and Twist1, Sdc1*, (Damonte, Gregg et al. 2007) in addition to basal-like markers including *Keratins 5, 6b, 14, and 17*, as well as *Kras* (Figures 4.5, 4.9A). Although a detailed analysis of

previous Met/HGF transgenic mouse models is lacking, models with constitutively activated Met, such as Tpr-Met or WAP/HGF, (Liang, Reid et al. 1996; Gallego, Bierie et al. 2003; Welm, Kim et al. 2005) also produce mammary tumors with heterogeneous histological phenotypes. In some cases, these tumors were shown to be keratin 6 positive (Welm, Kim et al. 2005) with accumulation of nuclear ß-catenin (Gallego, Bierie et al. 2003), consistent with the nuclear localisation of ß-catenin as observed in our Met^{mt} basal-like tumors (Figure 4.9A, B).

The observation that Met^{mt}-induced tumors with solid pathologies resemble MMTV/Neu, MMTV/PyMT and MMTV/ras tumors within a luminal subtype (Figure 4.5. 4.9A) may reflect the ability of the Met^{mt} carrying the Y1003F mutation to promote enhanced activity of the Ras pathway (Abella, Peschard et al. 2005). The presence of both luminal and basal-like tumor subtypes raises the possibility that Met may target a multipotent progenitor cell with the capacity to differentiate towards each lineage. In support of this, co-expression of Met with basal marker keratin 5 and luminal marker keratin 8/18 was observed in cells within Met^{mt} mixed pathology basal-like tumors but not in solid luminal-like tumors (Figures 4.10B and 4.11). Moreover, cells immunopositive for nuclear Trp63 in Met^{mt} basal-like tumors (Figure 4.9B) are indicative of a stem/progenitor cell population (DiRenzo, Signoretti et al. 2002). These cells also show elevated immunostaining for Trp53, which is frequently elevated and mutated in human basal breast cancers (Rakha, Putti et al. 2006).

Elevated MET protein has been associated with poor outcome in human breast cancer (Ghoussoub, Dillon et al. 1998; Tolgay Ocal, Dolled-Filhart et al. 2003; Lengyel, Prechtel et al. 2005). Consistent with this, we show that elevated MET protein is associated with poor outcome in patients who are lymph node negative at presentation

(Figure 4.14A). Specifically, within these patients, elevated MET protein is significantly correlated with the basal subtype of breast cancers (p-value=6.51e-03, Figure 4.14A). Among the biological processes highlighted in gene expression studies with the Met^{mt} basal tumors are genes involved in EMT (Figures 4.4A, 4.9). Within EMT-related genes, one of the most differentially regulated genes in Met^{mt} basal tumors is the *Snail* transcription factor, which regulates EMT during development. Strikingly, elevated levels of both MET and SNAIL in tumors is significantly correlated with poor outcome in lymph node negative patients when compared to either protein alone (Log-Rank p-value=7.40e-03 and Figure 4.14B). Induction of an EMT program is associated with reduced cellular differentiation as observed in tumors within the mixed pathology subtype and is found to confer on carcinoma cells many of the phenotypes required for invasion and metastasis (Thiery 2002; Scheel, Onder et al. 2007).

This is a first report of an MMTV-driven receptor tyrosine kinase mouse model that produces tumors resembling human basal breast cancer. This provides strong support that activation of Met signaling pathways play an important role in the induction of tumors of the basal subtype. Molecular mechanisms driving basal tumors are poorly understood (Yehiely, Moyano et al. 2006), and the Met^{mt} mouse model provides a useful tool to dissect signaling pathways of the basal phenotype. Since human basal breast cancer is especially difficult to treat due to the lack of understanding behind the genes and processes involved in its induction, developing a model for the dissection of this subtype is of utmost importance. Here, we establish that the MET receptor is a marker of basal and poor outcome breast cancer, and may prove useful as a therapeutic target in the management of these patients.

4.4 Supplemental Text – Statistical Analysis of Clinical Outcomes in the Axillary Node Negative (ANN) Cohort

4.4.1 Differential expression analysis of cDNA Array Gene Expression

Gene expression/microarray studies in the Andrulis *et al.* (Andrulis, Bull et al. 1998) ANN cohort data were conducted in a subgroup of 89 tumors selected from a prospectively accrued cohort of newly diagnosed women with axillary node negative disease – 48 were from women who had experienced disease recurrence within 4 years of diagnosis compared to 41 without recurrence within 10 years of diagnosis. A probe for MET values but not for SNAIL was available on the arrays.

Relative expression derived from log ratio means and fold change ratio, as well as the t-statistic and p-value for the mean differences in log ratio MET values (Table 4.9) demonstrate significant differences between the two groups of tumours (early recurrence vs. long follow up without recurrence) - this is a univariate analysis without accounting for any prognostic factors. In additional multivariate analysis, we fit a linear regression of the MET log ratio gene expression values, including as covariates an indicator for early recurrence versus long disease-free follow up together with the traditional prognostic factors of menopausal status, tumor size, histological grade, ER status, lymphatic invasion, age at diagnosis, and adjuvant treatment received. We found a statistically significant difference between the survival groups, independent of these traditional clinico-pathological parameters ($p=0.031$).

4.4.2 Disease-free survival (DFS) analysis of Tissue Microarrays (TMA) for MET and SNAIL Protein Levels

Analyses of the association of DFS with MET protein status were conducted in a large group of the ANN cohort patients for whom tissue microarrays had been

constructed (Andrulis, Bull et al. 1998; Mulligan, Pinnaduwage et al. 2008) using a standard Cox proportional hazards model with and without including traditional clinicopathological factors as covariates (Tables 4.11-4.12). To assess the association of DFS with the MET and SNAIL protein status jointly, we compared survival of patients with tumors positive for both MET and SNAIL to a combined group with tumors positive for neither or only one of MET and SNAIL (Table 4.12).

We also investigated whether MET protein is similarly associated with poor outcome for basal and non-basal breast cancer subtypes (n=60 and n=446, respectively), using a Cox proportional hazards model with an interaction term to statistically assess homogeneity of the MET association with poor prognosis between basal and non-basal groups (Table 4.10). Although the available sample size limits somewhat the precision of the subgroup estimates (HR = 1.57 in the non-basal group [95% CI=0.92-2.71] versus HR=2.44 in the basal group [95% CI=0.51-23.4]), significant differences between hazard ratios were not detected (p-value=6.3e-01 in univariate analysis, Table 4.10). Inference was based on a method designed for small samples (Heinze and Schemper 2001). Kaplan-Meier plots demonstrate the similarity between basal and non-basal subtypes in the pattern of the association of MET positive protein status with reduced disease-free survival (Figure 4.15A, log-rank p-value=2.15e-01 in basal and Figure 4.15B, log-rank p-value=9.49e-02 in non-basal).

To assess whether the expression of MET/SNAIL correlates with poor outcomes in both basal and non-basal subtype groups (n=55 and 401, respectively), we compared survival of patients with tumors positive for both MET and SNAIL to a combined group with tumors positive for neither or only one of MET and SNAIL (Table 4.13). We found

some evidence for association within the non-basal subtype (p=0.10) as well as in the basal subtype (p=0.02). A suggestive homogeneity test p-value in a Cox model (p=0.15) for differences between basal and non-basal in the association of MET with time to recurrence: (HR = 1.83 in the non-basal group [95% CI=0.88-3.50] versus HR=6.08 in the basal group [95% CI=1.35-34.9]), indicated that the basal/MET+/SNAIL+ subgroup may be at especially high risk of recurrence. However, HR values as large as 3.5 cannot be excluded in the non-basal group, and the MET/SNAIL association with DFS was similar in the basal and non-basal subgroups.

Strain	No. of mice with tumors/mice (%)	Average Latency (days)
FVB	0/34 (0)	0
Met wt	2/22 (9)	475
Met M1248T	2/17 (12)	478
Met Y1003F/M1248T	14/35 (40)	381

Table 4.1. Mutationally activated Met^{mt} transgene induces mammary tumors.
Data for MMTV/Met wild type (wt), Met^{mt} (M1248T, Y1003F/M1248T) and FVB multiparous controls are shown. Mammary tumor development was observed in multiparous Met^{mt} females.

Met mutation status	Mouse ID	Mammary fat pad no.	Diagnosis
Y1003F/M1248T	559*	MFP 2/3L	papillary carcinoma
Y1003F/M1248T	559*	MFP 2/3R	solid nodular carcinoma
Y1003F/M1248T	562	MFP 2/3L	MIN with solid, nodular intra lobular polyps
Y1003F/M1248T	567	MFP 2L	adenocarcinoma with sclerosis
M1248T	3229	MFP 2L	cystic adenosquamous carcinoma
M1248T	3229	MFP 2R	adenosquamous carcinoma
Y1003F/M1248T	4164	MFP 4L	solid nodular carcinoma
M1248T	4425	MFP 2/3R	adenocarcinoma
Y1003F/M1248T	4474	MFP 1L	solid nodular carcinoma
Y1003F/M1248T	4474	MFP 2/3L	solid nodular carcinoma
Y1003F/M1248T	4474	MFP 5L	solid nodular carcinoma
M1248T	4600	MFP 2/3R	adenocarcinoma with cribriform pattern, central necrosis and squamous metaplasia
Y1003F/M1248T	4691	MFP 2/3R	solid nodular carcinoma and papillary tumor
Y1003F/M1248T	4691	MFP 1R	solid nodular carcinoma
Y1003F/M1248T	4695	MFP 2L	scirrhous adenocarcinoma
Y1003F/M1248T	4695	MFP 4L	scirrhous adenocarcinoma
Y1003F/M1248T	4815	MFP 3R	cystic adenosquamous carcinoma and inflammatory nodules
Y1003F/M1248T	4996	MFP 2/3R	adenosquamous carcinoma
Y1003F/M1248T	5033	MFP 2/3R	solid nodular carcinoma
Y1003F/M1248T	5033	MFP 4R	papillary carcinoma
M1248T	5154	MFP 2/3R	adenosquamous carcinoma with cellular stroma
M1248T	5154	MFP 3L	adenosquamous carcinoma
Y1003F/M1248T	5156	MFP 1R	solid nodular carcinoma
Y1003F/M1248T	5156	MFP 3R	solid nodular carcinoma
Wild type	5382	MFP 1R	solid nodular carcinoma with spindle cell component
Wild type	5385	MFP 1L	papillary and micropapillary adenocarcinomas
Y1003F/M1248T	5482	MFP 2/3L	spindle cell tumor and lobular-alveolar hyperplasia
Y1003F/M1248T	5482	MFP 4L	type A microacinar carcinoma with squamous metaplasia
Y1003F/M1248T	5484	MFP 1L	scirrhous adenosquamous carcinoma
Y1003F/M1248T	5484	MFP 3L	adenosquamous carcinoma
Wild type	5488	MFP 3R	adenomyoepithelial carcinoma
Y1003F/M1248T	5546	MFP 1R	partially necrotic papillary adenocarcinoma
Wild type	5612	MFP 1R	adenocarcinoma
Y1003F/M1248T	6030	MFP 2R	solid nodular carcinoma
Y1003F/M1248T	6030	MFP 3R	solid nodular carcinoma
Y1003F/M1248T	6032	MFP 1L	adenosquamous carcinoma and solid nodular carcinoma
Y1003F/M1248T	6615	MFP 2R	solid nodular carcinoma
Y1003F/M1248T	6615	MFP 3R	necrotic solid nodular carcinoma
Y1003F/M1248T	6615	MFP 4R	solid nodular carcinoma
Y1003F/M1248T	6616	MFP 2R	papillary adenocarcinoma with sclerosis
Wild type	6646	MFP 4L	squamous cell carcinoma and lobular-alveolar hyperplasia
Y1003F/M1248T	6696*	MFP 1La/b	solid nodular carcinoma
Wild type	7122	MFP 2/3R	glandular and papillary adenocarcinoma
Y1003F/M1248T	7325	MFP 2R	solid nodular carcinoma
Y1003F/M1248T	7325	MFP 4L	solid nodular carcinoma
Y1003F/M1248T	7327	MFP 1R	papillary adenocarcinoma

Table 4.2. Summary of Met^{mt} transgenic mammary tumors.
Several mice contained tumors in multiple mammary fat pads per animal. Mammary fat pad number, and side (R=right, L=left) in which the tumor was found is noted. MIN=mammary intraepithelial neoplasia. The asterisk (*) identifies animals that developed lung metastases. The notation "a, b" for mouse ID 6696 refers to two separate tumor nodules within the same mammary gland.

Table 4.3. Overrepresented Biological Process GO categories in genes elevated in Met^{mt} mixed pathology vs. solid tumors.

GO ID	Pvalue	ExpCount	Count	Size	Description
GO:0050875	2.55E-15	812.9833	930	8157	cellular physiological process
GO:0043170	8.29E-13	329.3993	436	3305	macromolecule metabolism
GO:0044238	7.17E-12	533.9158	645	5357	primary metabolism
GO:0044237	1.19E-10	555.6432	660	5575	cellular metabolism
GO:0008152	1.85E-10	591.8223	695	5938	metabolism
GO:0019538	3.92E-10	241.792	326	2426	protein metabolism
GO:0044267	4.84E-10	227.9383	310	2287	cellular protein metabolism
GO:0044260	1.09E-09	231.327	312	2321	cellular macromolecule metabolism
GO:0006412	4.20E-09	48.63747	90	488	protein biosynthesis
GO:0007275	1.00E-08	177.4072	246	1780	development
GO:0009058	7.10E-08	96.67694	147	970	biosynthesis
GO:0009059	8.06E-08	53.91982	93	541	macromolecule biosynthesis
GO:0006996	2.14E-07	78.73689	123	790	organelle organization and biogenesis
GO:0044249	2.27E-07	85.3149	131	856	cellular biosynthesis
GO:0016043	3.42E-07	146.8094	203	1473	cell organization and biogenesis
GO:0050789	3.43E-07	283.6521	355	2846	regulation of biological process
GO:0009653	1.02E-06	81.02923	123	813	morphogenesis
GO:0016049	3.26E-06	7.076353	21	71	cell growth
GO:0048513	3.30E-06	86.21191	127	865	organ development
GO:0040007	3.99E-06	13.75404	32	138	growth
GO:0008361	6.74E-06	7.375354	21	74	regulation of cell size
GO:0042254	7.90E-06	12.25903	29	123	ribosome biogenesis and assembly
GO:0007155	1.21E-05	47.54113	77	477	cell adhesion
GO:0001558	1.31E-05	5.980017	18	60	regulation of cell growth
GO:0050794	1.50E-05	261.7254	320	2626	regulation of cellular process
GO:0007028	1.65E-05	13.35537	30	134	cytoplasm organization and biogenesis
GO:0050791	1.76E-05	254.7487	312	2556	regulation of physiological process
GO:0009887	1.87E-05	45.74713	74	459	organ morphogenesis
GO:0048514	2.26E-05	12.25903	28	123	blood vessel morphogenesis
GO:0051726	2.75E-05	27.80708	50	279	regulation of cell cycle
GO:0000074	2.75E-05	27.80708	50	279	regulation of progression through cell cycle
GO:0040008	3.58E-05	9.368693	23	94	regulation of growth
GO:0006817	3.60E-05	5.282348	16	53	phosphate transport
GO:0051258	4.41E-05	3.289009	12	33	protein polymerization
GO:0000902	5.21E-05	31.49475	54	316	cellular morphogenesis
GO:0007046	5.81E-05	11.56137	26	116	ribosome biogenesis
GO:0007049	7.71E-05	55.81349	84	560	cell cycle
GO:0045807	0.000102	1.694338	8	17	positive regulation of endocytosis
GO:0051244	0.000118	244.8817	295	2457	regulation of cellular physiological process
GO:0007346	0.000175	1.395337	7	14	regulation of progression through mitotic cell cycle
GO:0050764	0.000175	1.395337	7	14	regulation of phagocytosis
GO:0050766	0.000175	1.395337	7	14	positive regulation of phagocytosis
GO:0042475	0.000211	2.29234	9	23	odontogenesis (sensu Vertebrata)
GO:0042476	0.000211	2.29234	9	23	odontogenesis
GO:0006897	0.000217	13.15604	27	132	endocytosis

GO:0006909	0.000228	2.790674	10	28	phagocytosis
GO:0009611	0.000245	22.6244	40	227	response to wounding
GO:0001525	0.000279	9.966694	22	100	angiogenesis
GO:0001944	0.00028	14.75071	29	148	vasculature development
GO:0030574	0.000301	1.495004	7	15	collagen catabolism
GO:0006937	0.000308	2.392007	9	24	regulation of muscle contraction
GO:0007010	0.00039	35.5811	56	357	cytoskeleton organization and biogenesis
GO:0048518	0.000414	59.50117	85	597	positive regulation of biological process
GO:0051216	0.000433	2.990008	10	30	cartilage development
GO:0045637	0.000439	2.491674	9	25	regulation of myeloid cell differentiation
GO:0007167	0.000453	19.53472	35	196	enzyme linked receptor protein signaling pathway
GO:0001568	0.000454	14.45171	28	145	blood vessel development
GO:0043283	0.000662	207.8056	249	2085	biopolymer metabolism
GO:0043037	0.000662	13.35537	26	134	translation
GO:0048519	0.000786	64.98285	90	652	negative regulation of biological process
GO:0030100	0.000838	2.691007	9	27	regulation of endocytosis
GO:0006954	0.000849	12.85704	25	129	inflammatory response
GO:0048041	0.000988	0.299001	3	3	focal adhesion formation
GO:0007582	0.001116	975.241	1014	9785	physiological process
GO:0016331	0.001158	4.485012	12	45	morphogenesis of embryonic epithelium
GO:0009605	0.001212	32.39176	50	325	response to external stimulus
GO:0050793	0.00126	14.65104	27	147	regulation of development
GO:0045595	0.00141	11.1627	22	112	regulation of cell differentiation
GO:0046916	0.001491	2.890341	9	29	transition metal ion homeostasis
GO:0048522	0.001571	51.02948	72	512	positive regulation of cellular process
GO:0016477	0.001648	21.72739	36	218	cell migration
GO:0048523	0.001917	59.1025	81	593	negative regulation of cellular process
GO:0042127	0.001965	22.72406	37	228	regulation of cell proliferation
GO:0007169	0.002041	10.76403	21	108	transmembrane receptor protein tyrosine kinase signaling pathway
GO:0002009	0.002321	8.770691	18	88	morphogenesis of an epithelium
GO:0051246	0.002505	16.84371	29	169	regulation of protein metabolism
GO:0000278	0.002505	16.84371	29	169	mitotic cell cycle
GO:0035239	0.002541	7.475021	16	75	tube morphogenesis
GO:0008283	0.002643	33.58776	50	337	cell proliferation
GO:0006457	0.002769	20.03306	33	201	protein folding
GO:0016070	0.002785	36.97644	54	371	RNA metabolism
GO:0051050	0.002886	2.591341	8	26	positive regulation of transport
GO:0050790	0.003223	14.85037	26	149	regulation of enzyme activity
GO:0051674	0.003584	25.11607	39	252	localization of cell
GO:0006928	0.003584	25.11607	39	252	cell motility
GO:0030099	0.003619	6.378684	14	64	myeloid cell differentiation
GO:0006936	0.003619	6.378684	14	64	muscle contraction
GO:0045022	0.003657	0.398668	3	4	early endosome to late endosome transport
GO:0042487	0.003657	0.398668	3	4	regulation of odontogenesis (sensu Vertebrata)
GO:0042759	0.003657	0.398668	3	4	long-chain fatty acid biosynthesis
GO:0043087	0.003749	2.691007	8	27	regulation of GTPase activity
GO:0006139	0.003922	254.35	291	2552	nucleobase, nucleoside, nucleotide and nucleic acid metabolism
GO:0008154	0.004002	3.887011	10	39	actin polymerization and/or

					depolymerization
GO:0051242	0.004272	47.0428	65	472	positive regulation of cellular physiological process
GO:0043062	0.004329	5.182681	12	52	extracellular structure organization and biogenesis
GO:0030198	0.004329	5.182681	12	52	extracellular matrix organization and biogenesis
GO:0051336	0.004329	5.182681	12	52	regulation of hydrolase activity
GO:0045765	0.004558	1.694338	6	17	regulation of angiogenesis
GO:0040011	0.004713	25.51474	39	256	locomotion
GO:0051094	0.0051	5.282348	12	53	positive regulation of development
GO:0051049	0.005407	5.980017	13	60	regulation of transport
GO:0014032	0.005618	2.29234	7	23	neural crest cell development
GO:0014033	0.005618	2.29234	7	23	neural crest cell differentiation
GO:0006413	0.005878	4.086345	10	41	translational initiation
GO:0008015	0.005977	5.382015	12	54	circulation
GO:0051239	0.006254	17.14271	28	172	regulation of organismal physiological process
GO:0009260	0.00626	6.079684	13	61	ribonucleotide biosynthesis
GO:0006164	0.00626	6.079684	13	61	purine nucleotide biosynthesis
GO:0050801	0.006288	8.870358	17	89	ion homeostasis
GO:0043119	0.006664	48.8368	66	490	positive regulation of physiological process
GO:0006955	0.006872	48.03947	65	482	immune response
GO:0048534	0.007002	14.95004	25	150	hemopoietic or lymphoid organ development
GO:0045597	0.007041	4.186012	10	42	positive regulation of cell differentiation
GO:0042060	0.007217	6.179351	13	62	wound healing
GO:0016055	0.007332	10.46503	19	105	Wnt receptor signaling pathway
GO:0043118	0.007583	51.72714	69	519	negative regulation of physiological process
GO:0009152	0.008085	5.581349	12	56	purine ribonucleotide biosynthesis
GO:0006910	0.008205	0.897002	4	9	phagocytosis, recognition
GO:0051093	0.008285	6.279017	13	63	negative regulation of development
GO:0048754	0.008375	4.285679	10	43	branching morphogenesis of a tube
GO:0009954	0.008463	0.498335	3	5	proximal/distal pattern formation
GO:0006882	0.008463	0.498335	3	5	zinc ion homeostasis
GO:0008354	0.008463	0.498335	3	5	germ cell migration
GO:0015074	0.008463	0.498335	3	5	DNA integration
GO:0045185	0.008463	0.498335	3	5	maintenance of protein localization
GO:0050920	0.008463	0.498335	3	5	regulation of chemotaxis
GO:0050921	0.008463	0.498335	3	5	positive regulation of chemotaxis
GO:0050926	0.008463	0.498335	3	5	regulation of positive chemotaxis
GO:0050927	0.008463	0.498335	3	5	positive regulation of positive chemotaxis
GO:0050918	0.008463	0.498335	3	5	positive chemotaxis
GO:0050930	0.008463	0.498335	3	5	induction of positive chemotaxis
GO:0009142	0.008991	4.983347	11	50	nucleoside triphosphate biosynthesis
GO:0006958	0.009234	2.491674	7	25	complement activation, classical pathway
GO:0030509	0.009234	2.491674	7	25	BMP signaling pathway
GO:0030003	0.009454	7.076353	14	71	cation homeostasis
GO:0030005	0.009475	6.378684	13	64	di-, tri-valent inorganic cation homeostasis

GO:0009888	0.009807	25.0164	37	251	tissue development
GO:0001501	0.009908	19.33539	30	194	skeletal development
GO:0035295	0.009912	10.76403	19	108	tube development
GO:0030069	0.009926	0.199334	2	2	lysogeny
GO:0001935	0.009926	0.199334	2	2	endothelial cell proliferation
GO:0001936	0.009926	0.199334	2	2	regulation of endothelial cell proliferation
GO:0001937	0.009926	0.199334	2	2	negative regulation of endothelial cell proliferation
GO:0008228	0.009926	0.199334	2	2	opsonization
GO:0007090	0.009926	0.199334	2	2	regulation of S phase of mitotic cell cycle
GO:0048384	0.009926	0.199334	2	2	retinoic acid receptor signaling pathway
GO:0019047	0.009926	0.199334	2	2	provirus integration
GO:0050919	0.009926	0.199334	2	2	negative chemotaxis
GO:0042481	0.009926	0.199334	2	2	regulation of odontogenesis
GO:0016071	0.010471	21.02973	32	211	mRNA metabolism
GO:0009408	0.011009	1.993339	6	20	response to heat
GO:0051707	0.011311	32.79042	46	329	response to other organism
GO:0048468	0.01147	28.60441	41	287	cell development
GO:0001763	0.011616	4.485012	10	45	morphogenesis of a branching structure
GO:0030029	0.011667	12.45837	21	125	actin filament-based process
GO:0006396	0.011848	29.50142	42	296	RNA processing
GO:0045321	0.011927	13.2557	22	133	immune cell activation
GO:0006885	0.012491	1.495004	5	15	regulation of pH
GO:0048593	0.012608	0.996669	4	10	eye morphogenesis (sensu Vertebrata)
GO:0048595	0.012608	0.996669	4	10	eye morphogenesis (sensu Mammalia)
GO:0045576	0.012608	0.996669	4	10	mast cell activation
GO:0051243	0.012688	49.4348	65	496	negative regulation of cellular physiological process
GO:0001775	0.012971	13.35537	22	134	cell activation
GO:0009790	0.013126	28.00641	40	281	embryonic development
GO:0030036	0.013563	11.86037	20	119	actin cytoskeleton organization and biogenesis
GO:0030097	0.014086	13.45504	22	135	hemopoiesis
GO:0050727	0.01414	2.093006	6	21	regulation of inflammatory response
GO:0001747	0.015027	3.986678	9	40	eye development (sensu Mammalia)
GO:0006873	0.01517	7.475021	14	75	cell ion homeostasis
GO:0006397	0.015227	18.33872	28	184	mRNA processing
GO:0046849	0.01546	15.14938	24	152	bone remodeling
GO:0009968	0.015615	6.777352	13	68	negative regulation of signal transduction
GO:0006950	0.015674	71.36153	89	716	response to stress
GO:0000080	0.015678	0.598002	3	6	G1 phase of mitotic cell cycle
GO:0030856	0.015678	0.598002	3	6	regulation of epithelial cell differentiation
GO:0050819	0.015678	0.598002	3	6	negative regulation of coagulation
GO:0045669	0.015678	0.598002	3	6	positive regulation of osteoblast differentiation
GO:0051318	0.015678	0.598002	3	6	G1 phase
GO:0009145	0.015724	4.684346	10	47	purine nucleoside triphosphate biosynthesis
GO:0007596	0.015724	4.684346	10	47	blood coagulation
GO:0009201	0.015724	4.684346	10	47	ribonucleoside triphosphate biosynthesis

GO:0009206	0.015724	4.684346	10	47	purine ribonucleoside triphosphate biosynthesis
GO:0006956	0.016409	3.388676	8	34	complement activation
GO:0001503	0.016635	14.45171	23	145	ossification
GO:0031214	0.016635	14.45171	23	145	biomineral formation
GO:0030041	0.016706	1.594671	5	16	actin filament polymerization
GO:0009613	0.017097	31.89342	44	320	response to pest, pathogen or parasite
GO:0048762	0.017483	2.790674	7	28	mesenchymal cell differentiation
GO:0014031	0.017483	2.790674	7	28	mesenchymal cell development
GO:0019221	0.017483	2.790674	7	28	cytokine and chemokine mediated signaling pathway
GO:0006875	0.017541	6.877019	13	69	metal ion homeostasis
GO:0009259	0.017541	6.877019	13	69	ribonucleotide metabolism
GO:0006163	0.017541	6.877019	13	69	purine nucleotide metabolism
GO:0009150	0.017993	6.179351	12	62	purine ribonucleotide metabolism
GO:0050817	0.018142	4.784013	10	48	coagulation
GO:0006911	0.018273	1.096336	4	11	phagocytosis, engulfment
GO:0045667	0.018273	1.096336	4	11	regulation of osteoblast differentiation
GO:0048731	0.018681	47.74047	62	479	system development
GO:0009165	0.018772	11.4617	19	115	nucleotide biosynthesis
GO:0006935	0.018887	9.169359	16	92	chemotaxis
GO:0042330	0.018887	9.169359	16	92	taxis
GO:0051325	0.019466	3.488343	8	35	interphase
GO:0051329	0.019466	3.488343	8	35	interphase of mitotic cell cycle
GO:0019222	0.019887	187.6729	213	1883	regulation of metabolism
GO:0043010	0.02048	4.186012	9	42	eye development (sensu Vertebrata)
GO:0048771	0.02072	15.54804	24	156	tissue remodeling
GO:0007160	0.020822	4.88368	10	49	cell-matrix adhesion
GO:0007599	0.020822	4.88368	10	49	hemostasis
GO:0030855	0.021118	2.890341	7	29	epithelial cell differentiation
GO:0006959	0.021925	7.076353	13	71	humoral immune response
GO:0050767	0.022156	2.29234	6	23	regulation of neurogenesis
GO:0009141	0.023423	5.681016	11	57	nucleoside triphosphate metabolism
GO:0043038	0.023684	4.285679	9	43	amino acid activation
GO:0043039	0.023684	4.285679	9	43	tRNA aminoacylation
GO:0006418	0.023684	4.285679	9	43	tRNA aminoacylation for protein translation
GO:0031589	0.02378	4.983347	10	50	cell-substrate adhesion
GO:0009205	0.02378	4.983347	10	50	purine ribonucleoside triphosphate metabolism
GO:0008284	0.024065	11.7607	19	118	positive regulation of cell proliferation
GO:0048598	0.02423	13.35537	21	134	embryonic morphogenesis
GO:0008064	0.025247	2.990008	7	30	regulation of actin polymerization and/or depolymerization
GO:0048592	0.025247	2.990008	7	30	eye morphogenesis
GO:0001676	0.025288	1.196003	4	12	long-chain fatty acid metabolism
GO:0006692	0.025288	1.196003	4	12	prostanoid metabolism
GO:0006693	0.025288	1.196003	4	12	prostaglandin metabolism
GO:0048048	0.025288	1.196003	4	12	embryonic eye morphogenesis
GO:0050729	0.025288	1.196003	4	12	positive regulation of inflammatory response
GO:0051272	0.025423	0.697669	3	7	positive regulation of cell motility
GO:0051235	0.025423	0.697669	3	7	maintenance of localization
GO:0042573	0.025423	0.697669	3	7	retinoic acid metabolism

GO:0030330	0.025423	0.697669	3	7	DNA damage response, signal transduction by p53 class mediator
GO:0040017	0.025423	0.697669	3	7	positive regulation of locomotion
GO:0050818	0.025423	0.697669	3	7	regulation of coagulation
GO:0016032	0.025423	0.697669	3	7	viral life cycle
GO:0030705	0.025662	7.973356	14	80	cytoskeleton-dependent intracellular transport
GO:0006259	0.026313	46.84346	60	470	DNA metabolism
GO:0009199	0.027029	5.083014	10	51	ribonucleoside triphosphate metabolism
GO:0016064	0.027029	5.083014	10	51	humoral defense mechanism (sensu Vertebrata)
GO:0007229	0.027076	7.275687	13	73	integrin-mediated signaling pathway
GO:0007517	0.027298	9.568027	16	96	muscle development
GO:0030278	0.027729	1.794005	5	18	regulation of ossification
GO:0001755	0.027729	1.794005	5	18	neural crest cell migration
GO:0009892	0.027789	24.31873	34	244	negative regulation of metabolism
GO:0001878	0.027802	0.299001	2	3	response to yeast
GO:0001578	0.027802	0.299001	2	3	microtubule bundle formation
GO:0018401	0.027802	0.299001	2	3	peptidyl-proline hydroxylation to 4-hydroxy-L-proline
GO:0002026	0.027802	0.299001	2	3	cardiac inotropy
GO:0007089	0.027802	0.299001	2	3	traversing start control point of mitotic cell cycle
GO:0008593	0.027802	0.299001	2	3	regulation of Notch signaling pathway
GO:0046627	0.027802	0.299001	2	3	negative regulation of insulin receptor signaling pathway
GO:0030512	0.027802	0.299001	2	3	negative regulation of transforming growth factor beta receptor signaling pathway
GO:0045428	0.027802	0.299001	2	3	regulation of nitric oxide biosynthesis
GO:0019471	0.027802	0.299001	2	3	4-hydroxyproline metabolism
GO:0042448	0.027802	0.299001	2	3	progesterone metabolism
GO:0019511	0.027802	0.299001	2	3	peptidyl-proline hydroxylation
GO:0007399	0.027981	43.45479	56	436	nervous system development
GO:0030154	0.02812	81.42789	98	817	cell differentiation
GO:0050878	0.029625	5.88035	11	59	regulation of body fluids
GO:0030832	0.029895	3.089675	7	31	regulation of actin filament length
GO:0007018	0.02996	7.375354	13	74	microtubule-based movement
GO:0009144	0.030583	5.182681	10	52	purine nucleoside triphosphate metabolism
GO:0045596	0.030583	5.182681	10	52	negative regulation of cell differentiation
GO:0051128	0.031133	4.485012	9	45	regulation of cell organization and biogenesis
GO:0001654	0.031599	6.677685	12	67	eye development
GO:0009889	0.032725	10.5647	17	106	regulation of biosynthesis
GO:0006776	0.03371	1.29567	4	13	vitamin A metabolism
GO:0009147	0.03371	1.29567	4	13	pyrimidine nucleoside triphosphate metabolism
GO:0045638	0.03371	1.29567	4	13	negative regulation of myeloid cell differentiation
GO:0051169	0.034107	8.272356	14	83	nuclear transport
GO:0008016	0.034632	1.893672	5	19	regulation of heart contraction
GO:0045664	0.034632	1.893672	5	19	regulation of neuron differentiation

GO:0000904	0.035049	13.05637	20	131	cellular morphogenesis during differentiation
GO:0001516	0.037709	0.797336	3	8	prostaglandin biosynthesis
GO:0046457	0.037709	0.797336	3	8	prostanoid biosynthesis
GO:0030199	0.037709	0.797336	3	8	collagen fibril organization
GO:0007389	0.039072	17.34205	25	174	pattern specification
GO:0009266	0.04084	3.289009	7	33	response to temperature stimulus
GO:0016310	0.04141	55.31515	68	555	phosphorylation
GO:0016192	0.041475	31.19575	41	313	vesicle-mediated transport
GO:0006221	0.042506	1.993339	5	20	pyrimidine nucleotide biosynthesis
GO:0008286	0.042506	1.993339	5	20	insulin receptor signaling pathway
GO:0046850	0.042506	1.993339	5	20	regulation of bone remodeling
GO:0006913	0.044705	9.368693	15	94	nucleocytoplasmic transport
GO:0006807	0.045465	27.90674	37	280	nitrogen compound metabolism
GO:0031324	0.046277	21.02973	29	211	negative regulation of cellular metabolism
GO:0006468	0.046488	50.23214	62	504	protein amino acid phosphorylation
GO:0006521	0.047174	3.388676	7	34	regulation of amino acid metabolism
GO:0001932	0.047174	3.388676	7	34	regulation of protein amino acid phosphorylation
GO:0031323	0.049399	181.8922	202	1825	regulation of cellular metabolism

Table 4.4. Overrepresented Biological Process GO categories in genes elevated in Met^{mt} solid vs. mixed pathology tumors.

GO ID	Pvalue	ExpCount	Count	Size	Description
GO:0050875	2.09E-05	713.1432	772	8157	cellular physiological process
GO:0019883	6.79E-05	0.874271	6	10	antigen presentation, endogenous antigen
GO:0045184	7.01E-05	51.49459	79	589	establishment of protein localization
GO:0006810	9.43E-05	200.995	248	2299	transport
GO:0015031	9.83E-05	49.57119	76	567	protein transport
GO:0008104	0.000125	54.02998	81	618	protein localization
GO:0019882	0.000126	3.05995	11	35	antigen presentation
GO:0009132	0.000256	1.049126	6	12	nucleoside diphosphate metabolism
GO:0051234	0.000515	223.2015	266	2553	establishment of localization
GO:0016559	0.000667	0.262281	3	3	peroxisome fission
GO:0016584	0.000667	0.262281	3	3	nucleosome spacing
GO:0051179	0.000698	225.2123	267	2576	localization
GO:0009185	0.000755	0.524563	4	6	ribonucleoside diphosphate metabolism
GO:0009134	0.000755	0.524563	4	6	nucleoside diphosphate catabolism
GO:0006470	0.000771	9.267277	20	106	protein amino acid dephosphorylation
GO:0019885	0.000877	0.874271	5	10	antigen processing, endogenous antigen via MHC class I
GO:0016311	0.000984	9.442132	20	108	dephosphorylation
GO:0045045	0.001451	13.98834	26	160	secretory pathway
GO:0006334	0.002015	5.333056	13	61	nucleosome assembly
GO:0030333	0.002113	2.44796	8	28	antigen processing
GO:0048285	0.002492	0.349709	3	4	organelle fission
GO:0009179	0.002492	0.349709	3	4	purine ribonucleoside diphosphate metabolism
GO:0009135	0.002492	0.349709	3	4	purine nucleoside diphosphate metabolism
GO:0006955	0.003152	42.13988	60	482	immune response
GO:0006333	0.00333	8.305579	17	95	chromatin assembly or disassembly
GO:0031497	0.003406	6.294754	14	72	chromatin assembly
GO:0006325	0.003433	17.13572	29	196	establishment and/or maintenance of chromatin architecture
GO:0016337	0.003627	10.49126	20	120	cell-cell adhesion
GO:0006887	0.003638	5.682764	13	65	exocytosis
GO:0007155	0.003913	41.70275	59	477	cell adhesion
GO:0006915	0.004845	38.73022	55	443	apoptosis
GO:0006469	0.00572	2.273106	7	26	negative regulation of protein kinase activity
GO:0009191	0.005824	0.437136	3	5	ribonucleoside diphosphate catabolism
GO:0009143	0.005824	0.437136	3	5	nucleoside triphosphate catabolism
GO:0006323	0.0061	17.83514	29	204	DNA packaging
GO:0012501	0.006669	39.34221	55	450	programmed cell death
GO:0000004	0.007034	96.9567	120	1109	biological process unknown
GO:0051348	0.00715	2.360533	7	27	negative regulation of transferase activity
GO:0009203	0.007637	0.174854	2	2	ribonucleoside triphosphate catabolism
GO:0007616	0.007637	0.174854	2	2	long-term memory
GO:0045087	0.007807	3.584513	9	41	innate immune response
GO:0046903	0.007996	17.398	28	199	secretion
GO:0016265	0.0081	41.44047	57	474	death
GO:0008219	0.0081	41.44047	57	474	cell death
GO:0007001	0.008387	19.84596	31	227	chromosome organization and biogenesis (sensu Eukaryota)
GO:0007160	0.008812	4.28393	10	49	cell-matrix adhesion
GO:0043086	0.008829	2.44796	7	28	negative regulation of enzyme activity

GO:0007156	0.008956	6.994172	14	80	homophilic cell adhesion
GO:0042981	0.009249	24.04246	36	275	regulation of apoptosis
GO:0016192	0.009392	27.3647	40	313	vesicle-mediated transport
GO:0051276	0.009933	20.89509	32	239	chromosome organization and biogenesis
GO:0031589	0.010175	4.371357	10	50	cell-substrate adhesion
GO:0016043	0.01089	128.7802	153	1473	cell organization and biogenesis
GO:0009261	0.010893	0.524563	3	6	ribonucleotide catabolism
GO:0043067	0.011491	24.39217	36	279	regulation of programmed cell death
GO:0007031	0.012843	1.486261	5	17	peroxisome organization and biogenesis
GO:0006917	0.014553	8.130724	15	93	induction of apoptosis
GO:0012502	0.014553	8.130724	15	93	induction of programmed cell death
GO:0040008	0.015964	8.218152	15	94	regulation of growth
GO:0043407	0.016247	1.049126	4	12	negative regulation of MAPK activity
GO:0006071	0.016247	1.049126	4	12	glycerol metabolism
GO:0007613	0.016247	1.049126	4	12	memory
GO:0019751	0.016247	1.049126	4	12	polyol metabolism
GO:0007185	0.016529	1.573689	5	18	transmembrane receptor protein tyrosine phosphatase signaling pathway
GO:0001910	0.017833	0.61199	3	7	regulation of immune cell mediated cytotoxicity
GO:0030168	0.017833	0.61199	3	7	platelet activation
GO:0031341	0.017833	0.61199	3	7	regulation of cell killing
GO:0046907	0.020473	47.0358	61	538	intracellular transport
GO:0045931	0.021578	0.262281	2	3	positive regulation of progression through mitotic cell cycle
GO:0015780	0.021578	0.262281	2	3	nucleotide-sugar transport
GO:0007184	0.021578	0.262281	2	3	SMAD protein nuclear translocation
GO:0009181	0.021578	0.262281	2	3	purine ribonucleoside diphosphate catabolism
GO:0009137	0.021578	0.262281	2	3	purine nucleoside diphosphate catabolism
GO:0042415	0.021578	0.262281	2	3	norepinephrine metabolism
GO:0050803	0.021875	1.136553	4	13	regulation of synapse structure and function
GO:0009056	0.02239	38.46794	51	440	catabolism
GO:0051649	0.025105	47.56037	61	544	establishment of cellular localization
GO:0009607	0.02617	54.81682	69	627	response to biotic stimulus
GO:0035264	0.026424	2.360533	6	27	body growth
GO:0040014	0.026424	2.360533	6	27	regulation of body size
GO:0006940	0.026702	0.699417	3	8	regulation of smooth muscle contraction
GO:0006066	0.026853	16.69858	25	191	alcohol metabolism
GO:0006939	0.028553	1.22398	4	14	smooth muscle contraction
GO:0051641	0.029576	47.9975	61	549	cellular localization
GO:0006952	0.03127	52.63114	66	602	defense response
GO:0009166	0.031567	1.83597	5	21	nucleotide catabolism
GO:0015837	0.034722	4.546211	9	52	amine transport
GO:0044248	0.035182	31.64863	42	362	cellular catabolism
GO:0007167	0.035338	17.13572	25	196	enzyme linked receptor protein signaling pathway
GO:0006730	0.036555	2.535387	6	29	one-carbon compound metabolism
GO:0040018	0.037496	0.786844	3	9	positive regulation of body size
GO:0045927	0.037496	0.786844	3	9	positive regulation of growth
GO:0000082	0.038003	1.923397	5	22	G1/S transition of mitotic cell cycle
GO:0051056	0.038709	4.633639	9	53	regulation of small GTPase mediated signal transduction
GO:0009615	0.03873	3.234804	7	37	response to virus
GO:0016322	0.040663	0.349709	2	4	neuron remodeling
GO:0001912	0.040663	0.349709	2	4	positive regulation of immune cell mediated

					cytotoxicity
GO:0015858	0.040663	0.349709	2	4	nucleoside transport
GO:0040016	0.040663	0.349709	2	4	embryonic cleavage
GO:0009154	0.040663	0.349709	2	4	purine ribonucleotide catabolism
GO:0009133	0.040663	0.349709	2	4	nucleoside diphosphate biosynthesis
GO:0031343	0.040663	0.349709	2	4	positive regulation of cell killing
GO:0009595	0.040663	0.349709	2	4	detection of biotic stimulus
GO:0007264	0.041087	22.46878	31	257	small GTPase mediated signal transduction
GO:0030384	0.042437	2.622814	6	30	phosphoinositide metabolism
GO:0016125	0.042995	4.721066	9	54	sterol metabolism
GO:0030029	0.044439	10.92839	17	125	actin filament-based process
GO:0006826	0.045185	2.010824	5	23	iron ion transport
GO:0006464	0.04573	108.4097	125	1240	protein modification
GO:0006807	0.047391	24.4796	33	280	nitrogen compound metabolism
GO:0035023	0.04888	2.710241	6	31	regulation of Rho protein signal transduction
GO:0006650	0.049781	3.409659	7	39	glycerophospholipid metabolism
GO:0008154	0.049781	3.409659	7	39	actin polymerization and/or depolymerization
GO:0007611	0.049781	3.409659	7	39	learning and/or memory

Table 4.5. GSEA gene sets elevated in Met^{mt} mixed pathology vs. solid tumors.

Gene Set	Size	ES	NES	NOM p-val	FDR q-val	FWER p-val	Rank at Max	Leading Edge
IDX_TSA_DN_CLUSTER2	58	0.6943	2.8812	0	0	0	5024	tags=66%, list=18%, signal=80%
RIBOSOMAL_PROTEINS	53	0.6947	2.8132	0	0	0	5756	tags=72%, list=21%, signal=91%
PENG_GLUTAMINE_DN	227	0.5308	2.7364	0	0	0	5940	tags=55%, list=22%, signal=70%
JECHLINGER_EMT_UP	51	0.6821	2.7255	0	0	0	4509	tags=67%, list=16%, signal=80%
PENG_LEUCINE_DN	135	0.5585	2.6513	0	0	0	5940	tags=54%, list=22%, signal=69%
LI_FETAL_VS_WT_KIDNEY_DN	147	0.5388	2.6414	0	0	0	5180	tags=50%, list=19%, signal=62%
EMT_UP	56	0.6465	2.6256	0	0	0	4509	tags=61%, list=16%, signal=73%
AGUIRRE_PANCREAS_CHR22	44	0.6727	2.6113	0	0	0	8727	tags=93%, list=32%, signal=137%
ZUCCHI_EPITHELIAL_UP	35	0.6973	2.5751	0	0	0	5552	tags=74%, list=20%, signal=93%
SANA_IFNG_ENDOTHELIAL_DN	76	0.5936	2.5655	0	0	0	4932	tags=57%, list=18%, signal=69%
BLEO_MOUSE_LYMPH_HIGH_24HRS_DN	32	0.6951	2.5314	0	0	0	4086	tags=63%, list=15%, signal=73%
IDX_TSA_DN_CLUSTER1	40	0.672	2.5267	0	0	0	2675	tags=50%, list=10%, signal=55%
BRCA1_OVEREXP_DN	93	0.5579	2.5096	0	0	0	5720	tags=54%, list=21%, signal=68%
FALT_BCLL_UP	43	0.6541	2.5027	0	0	0	5225	tags=72%, list=19%, signal=89%
ADIP_VS_PREADIP_DN	34	0.6683	2.4775	0	0	0	3289	tags=56%, list=12%, signal=63%
POD1_KO_UP	353	0.4569	2.4519	0	0	0	5413	tags=44%, list=20%, signal=55%
AGED_MOUSE_HYPOTH_DN	41	0.6435	2.441	0	0	0	4705	tags=63%, list=17%, signal=76%
TARTE_PLASMA_BLASTIC	278	0.4631	2.4407	0	0	0	6307	tags=52%, list=23%, signal=67%
MRNA_PROCESSING_REACTOME	92	0.5386	2.4394	0	0	0	7003	tags=61%, list=26%, signal=82%

NING_COPD_UP	128	0.5176	2.4361	0	0	0	5813	tags=53%, list=21%, signal=67%
PRMT5_KD_UP	161	0.4973	2.4326	0	0	0	6254	tags=54%, list=23%, signal=70%
PENG_RAPAMYCIN_DN	176	0.4904	2.4225	0	0	0	5734	tags=49%, list=21%, signal=62%
MENSSEN_MYC_UP	25	0.7197	2.4117	0	0	0	3017	tags=60%, list=11%, signal=67%
BASSO_REGULATORY_HUBS	119	0.5055	2.3808	0	0	0	7285	tags=60%, list=27%, signal=81%
NADLER_OBESITY_UP	44	0.6099	2.38	0	0	0	4065	tags=61%, list=15%, signal=72%
STEMCELL_EMBRYONIC_UP	1212	0.3977	2.3749	0	0	0	6752	tags=44%, list=25%, signal=56%
CANCER_UNDIFFERENTIATED_META_UP	59	0.5706	2.3713	0	0	0	6254	tags=56%, list=23%, signal=72%
HCC_SURVIVAL_GOOD_VS_POOR_DN	101	0.5121	2.3554	0	0	0	6307	tags=50%, list=23%, signal=64%
COLLER_MYC_UP	15	0.794	2.3428	0	0	0	3291	tags=73%, list=12%, signal=83%
HESS_HOXAANMEIS1_DN	64	0.5514	2.3212	0	0	0	6223	tags=55%, list=23%, signal=71%
AGEING_KIDNEY_UP	333	0.4326	2.316	0	0	0	6842	tags=49%, list=25%, signal=64%
ADIP_HUMAN_DN	23	0.6894	2.3108	0	0	0	3146	tags=48%, list=11%, signal=54%
ADIP_VS_FIBRO_DN	25	0.6781	2.3078	0	0	0	1954	tags=52%, list=7%, signal=56%
DNA_REPLICATION_REACTOME	41	0.6041	2.3025	0	0	0	6636	tags=66%, list=24%, signal=87%
ADIP_DIFF_CLUSTER4	44	0.588	2.2997	0	0	0	4254	tags=48%, list=16%, signal=56%
CHANG_SERUM_RESPONSE_UP	123	0.4843	2.2995	0	0	0	5710	tags=51%, list=21%, signal=64%
HESS_HOXAANMEIS1_UP	64	0.5514	2.2987	0	0	0	6223	tags=55%, list=23%, signal=71%
LE_MYELIN_UP	87	0.505	2.2913	0	0	0	7938	tags=55%, list=29%, signal=77%
AGED_MOUSE_MUSCLE_DN	29	0.6508	2.2869	0	0	0	7070	tags=79%, list=26%, signal=107%

ST_B_CELL_ANTIGEN_RECEPTOR	35	0.6234	2.2856	0	0	0	6229	tags=71%, list=23%, signal=92%
LEI_MYB_REGULATED_GENES	293	0.4313	2.2781	0	0	0	6668	tags=46%, list=24%, signal=60%
RNA_TRANSCRIPTION_REACTOME	35	0.6235	2.2779	0	0	0	7202	tags=69%, list=26%, signal=93%
FLECHNER_KIDNEY_TRANSPLANT_REJECTION_UP	81	0.5223	2.2763	0	0	0	5020	tags=47%, list=18%, signal=57%
IRS_KO_ADIP_DN	37	0.6064	2.2668	0	4E-05	0.001	6359	tags=59%, list=23%, signal=77%
MYC_TARGETS	37	0.6067	2.2382	0	6E-05	0.002	4124	tags=49%, list=15%, signal=57%
BHATTACHARYA_ESC_UP	56	0.5532	2.2409	0	6E-05	0.002	6631	tags=61%, list=24%, signal=80%
FLOTHO_CASP8AP2_MRD_DIFF	66	0.5243	2.2526	0	7E-05	0.002	4433	tags=45%, list=16%, signal=54%
GAY_YY1_DN	250	0.4306	2.2584	0	7E-05	0.002	4887	tags=38%, list=18%, signal=46%
MA_ATRA_EMP_UP	35	0.6132	2.2591	0	7E-05	0.002	5422	tags=60%, list=20%, signal=75%
STEMCELL_NEURAL_UP	1648	0.3739	2.2593	0	7E-05	0.002	6558	tags=42%, list=24%, signal=52%
SERUM_FIBROBLAST_CORE_UP	168	0.4529	2.2328	0	9E-05	0.003	5710	tags=46%, list=21%, signal=58%
IGLESIAS_E2FMINUS_UP	133	0.4691	2.2246	0	0.0001	0.004	4637	tags=42%, list=17%, signal=50%
IDX_TSA_UP_CLUSTER4	39	0.5912	2.2263	0	0.0001	0.004	3010	tags=49%, list=11%, signal=55%
CELL_CYCLE	74	0.5205	2.2302	0	0.0001	0.004	5904	tags=50%, list=22%, signal=64%
IDX_TSA_UP_CLUSTER2	59	0.5373	2.2306	0	0.0001	0.004	3340	tags=42%, list=12%, signal=48%
ZELLER_MYC_UP	19	0.7063	2.2015	0	0.0002	0.008	2732	tags=53%, list=10%, signal=58%
SHIPP_FL_VS_DLBCL_DN	34	0.5991	2.2021	0	0.0002	0.008	3022	tags=44%, list=11%, signal=50%
SERUM_FIBROBLAST_CELLCYCLE	115	0.4748	2.2065	0	0.0002	0.008	6254	tags=47%, list=23%, signal=61%
TGFBETA_C1_UP	15	0.727	2.2112	0	0.0002	0.008	4138	tags=60%, list=15%, signal=71%

SCHUMACHER_MYC_UP	44	0.5699	2.1969	0	0.0002	0.009	6447	tags=64%, list=24%, signal=83%
KAMMINGA_EZH2_TARGETS	34	0.59	2.1895	0	0.0002	0.01	6636	tags=53%, list=24%, signal=70%
BOQUEST_CD31PLUS_VS_CD31MINUS_DN	246	0.4202	2.1897	0	0.0003	0.01	6763	tags=45%, list=25%, signal=59%
MRNA_SPLICING	40	0.5842	2.1911	0	0.0003	0.01	3448	tags=50%, list=13%, signal=57%
INTEGRINPATHWAY	34	0.5972	2.1959	0	0.0003	0.01	5843	tags=68%, list=21%, signal=86%
FERRANDO_MLL_T_ALL_DN	72	0.5144	2.185	0	0.0003	0.011	5714	tags=47%, list=21%, signal=60%
ET743_SARCOMA_DN	238	0.4207	2.1745	0	0.0003	0.012	7439	tags=50%, list=27%, signal=68%
HDACI_COLON_CUR48HRS_UP	57	0.5336	2.1745	0	0.0003	0.012	4099	tags=46%, list=15%, signal=54%
NING_COPD_DN	105	0.4777	2.1756	0	0.0003	0.012	5449	tags=46%, list=20%, signal=57%
CROONQUIST_IL6_STROMA_UP	36	0.5883	2.1773	0	0.0003	0.012	6853	tags=75%, list=25%, signal=100%
NAKAJIMA_MCSMBP_MAST	41	0.5685	2.1803	0	0.0003	0.012	5819	tags=56%, list=21%, signal=71%
AGEING_KIDNEY_SPECIFIC_UP	156	0.443	2.1707	0	0.0003	0.014	6847	tags=51%, list=25%, signal=67%
TGFBETA_ALL_UP	72	0.5086	2.169	0	0.0003	0.015	5332	tags=50%, list=19%, signal=62%
AGUIRRE_PANCREAS_CHR8	54	0.535	2.1697	0	0.0003	0.015	8172	tags=70%, list=30%, signal=100%
ETSPATHWAY	20	0.6765	2.1524	0	0.0005	0.024	4446	tags=55%, list=16%, signal=66%
REN_E2F1_TARGETS	33	0.589	2.1585	0	0.0005	0.023	5713	tags=55%, list=21%, signal=69%
CANCER_NEOPLASTIC_META_UP	54	0.5309	2.1526	0	0.0005	0.024	3354	tags=44%, list=12%, signal=51%
TNFA_NFKB_DEP_UP	17	0.696	2.1544	0.002	0.0005	0.024	6075	tags=65%, list=22%, signal=83%
GOLDRATH_CELLCYCLE	32	0.5813	2.1573	0	0.0005	0.024	6254	tags=56%, list=23%, signal=73%
LEE_MYC_TGFA_UP	58	0.5195	2.1495	0	0.0005	0.026	6517	tags=52%, list=24%, signal=68%

ADIP_DIFF_CLUSTER1	56	0.5233	2.15	0	0.0005	0.026	6625	tags=61%, list=24%, signal=80%
PASSERINI_EM	31	0.6059	2.1385	0	0.0006	0.03	4797	tags=52%, list=18%, signal=63%
IDX_TSA_UP_CLUSTER3	87	0.4808	2.1408	0	0.0006	0.03	7181	tags=52%, list=26%, signal=70%
CROONQUIST_IL6_STARVE_UP	29	0.6048	2.1414	0	0.0006	0.03	7832	tags=69%, list=29%, signal=97%
PARP_KO_DN	14	0.7314	2.134	0	0.0006	0.032	6636	tags=79%, list=24%, signal=104%
TGFBETA_C2_UP	16	0.7255	2.1314	0	0.0007	0.037	2561	tags=63%, list=9%, signal=69%
FSH_GRANULOSA_DN	67	0.4939	2.1209	0	0.0008	0.044	5533	tags=46%, list=20%, signal=58%
MANALO_HYPOXIA_DN	74	0.4928	2.1181	0	0.0009	0.047	7950	tags=54%, list=29%, signal=76%
IGFR_IR_UP	16	0.7165	2.1173	0	0.0009	0.048	5940	tags=75%, list=22%, signal=96%
ZHAN_MULTIPLE_MYELOMA_SUBCLASSES_DIFF	25	0.6257	2.1154	0	0.0009	0.051	7285	tags=72%, list=27%, signal=98%
IRS1_KO_ADIP_DN	101	0.4594	2.1119	0	0.001	0.054	7509	tags=53%, list=27%, signal=73%
PENG_LEUCINE_UP	98	0.4666	2.1128	0	0.001	0.054	5965	tags=49%, list=22%, signal=62%
OLDAGE_DN	42	0.5491	2.107	0	0.001	0.056	6254	tags=57%, list=23%, signal=74%
BRCA_ER_NEG	863	0.3589	2.1017	0	0.001	0.06	6763	tags=42%, list=25%, signal=54%
AGUIRRE_PANCREAS_CHR19	62	0.5104	2.1001	0	0.001	0.061	7618	tags=61%, list=28%, signal=85%
HIPPOCAMPUS_DEVELOPMENT_PRENATAL	25	0.6226	2.1072	0	0.001	0.056	6307	tags=68%, list=23%, signal=88%
FERNANDEZ_MYC_TARGETS	173	0.4291	2.1018	0	0.001	0.06	5882	tags=43%, list=21%, signal=55%
UVB_NHEK3_ALL	356	0.3881	2.0984	0	0.001	0.062	6612	tags=46%, list=24%, signal=60%
LH_GRANULOSA_DN	67	0.4939	2.1027	0	0.001	0.06	5533	tags=46%, list=20%, signal=58%
ET743_SARCOMA_72HRS_DN	197	0.4161	2.106	0	0.001	0.059	5967	tags=42%, list=22%, signal=53%

PARK_MSCS_DIFF	29	0.5905	2.1036	0	0.001	0.06	5732	tags=55%, list=21%, signal=70%
PROSTAGLANDIN_SYNTHESIS_REGULATION	28	0.5958	2.1063	0	0.001	0.059	4182	tags=43%, list=15%, signal=51%
ZHAN_MMPC_SIM_BC_AND_MM	39	0.5535	2.0926	0	0.0011	0.066	7126	tags=64%, list=26%, signal=87%
HG_PROGERIA_DN	22	0.6369	2.0894	0	0.0011	0.067	6254	tags=73%, list=23%, signal=94%
GUO_HEX_UP	71	0.4797	2.0858	0	0.0011	0.07	5984	tags=54%, list=22%, signal=68%
STEMCELL_HEMATOPOIETIC_UP	1305	0.3478	2.0861	0	0.0011	0.07	6854	tags=42%, list=25%, signal=53%
BRCA2_BRCA1_UP	43	0.5329	2.0866	0	0.0011	0.07	7236	tags=67%, list=26%, signal=92%
TGFBETA_EARLY_UP	42	0.5526	2.087	0	0.0011	0.07	2561	tags=45%, list=9%, signal=50%
ATRIA_UP	173	0.42	2.0821	0	0.0011	0.073	6374	tags=47%, list=23%, signal=61%
MMS_HUMAN_LYMPH_HIGH_24HRS_UP	15	0.7087	2.0806	0	0.0011	0.075	1989	tags=33%, list=7%, signal=36%
STRIATED_MUSCLE_CONTRACTION	34	0.5605	2.0677	0	0.0014	0.092	8364	tags=68%, list=31%, signal=97%
AMINOACYL_TRNA_BIOSYNTHESIS	20	0.6551	2.0551	0	0.0016	0.109	4285	tags=65%, list=16%, signal=77%
CMV_IE86_UP	47	0.5228	2.0602	0	0.0016	0.104	5881	tags=55%, list=21%, signal=70%
BRENTANI_CELL_CYCLE	79	0.4731	2.0539	0	0.0016	0.112	7202	tags=53%, list=26%, signal=72%
VHL_RCC_UP	105	0.4414	2.0555	0	0.0016	0.109	5922	tags=46%, list=22%, signal=58%
HDACI_COLON_CUR24HRS_UP	34	0.5558	2.057	0	0.0016	0.107	5834	tags=56%, list=21%, signal=71%
VERNELL_PRB_CLSTR1	62	0.4952	2.054	0	0.0016	0.112	6080	tags=48%, list=22%, signal=62%
HOFFMANN_BIVSBII_BI_TABLE2	192	0.4071	2.056	0	0.0016	0.109	6535	tags=43%, list=24%, signal=56%
CMV_ALL_DN	90	0.4598	2.0577	0	0.0016	0.107	6019	tags=48%, list=22%, signal=61%
GALINDO_ACT_UP	77	0.4741	2.0518	0	0.0016	0.115	6335	tags=55%, list=23%, signal=71%

FLECHNER_KIDNEY_TRANSPLANT_WELL_PBL_UP	129	0.4365	2.0532	0	0.0016	0.114	7844	tags=53%, list=29%, signal=74%
CELL_CYCLE_KEGG	81	0.4686	2.0481	0	0.0016	0.118	5904	tags=47%, list=22%, signal=60%
TSA_CD4_UP	23	0.6164	2.049	0	0.0017	0.118	4915	tags=52%, list=18%, signal=64%
MRNA_PROCESSING	36	0.5552	2.0452	0	0.0017	0.123	7315	tags=64%, list=27%, signal=87%
IFN_BETA_GLIOMA_DN	36	0.5514	2.0458	0	0.0017	0.122	5714	tags=53%, list=21%, signal=67%
RECKPATHWAY	11	0.7603	2.046	0	0.0017	0.122	4127	tags=82%, list=15%, signal=96%
SANSOM_APC_LOSS5_UP	77	0.4681	2.0425	0	0.0018	0.131	6140	tags=48%, list=22%, signal=62%
CROONQUIST_RAS_STROMA_DN	17	0.668	2.0401	0.0019	0.0019	0.138	6853	tags=76%, list=25%, signal=102%
LEE_MYC_UP	52	0.509	2.0378	0	0.0019	0.141	6224	tags=52%, list=23%, signal=67%
TRNA_SYNTHETASES	18	0.6656	2.0359	0.0019	0.0019	0.143	4285	tags=67%, list=16%, signal=79%
HDACI_COLON_CUR_UP	98	0.4454	2.0285	0	0.002	0.154	4234	tags=41%, list=15%, signal=48%
ST_PHOSPHOINOSITIDE_3_KINASE_PATHWAY	33	0.5688	2.0287	0	0.002	0.154	6229	tags=55%, list=23%, signal=71%
ESR_FIBROBLAST_UP	46	0.5093	2.0292	0	0.0021	0.154	5229	tags=50%, list=19%, signal=62%
BRENTANI_REPAIR	35	0.54	2.0271	0	0.0021	0.158	6988	tags=60%, list=26%, signal=80%
ET743_SARCOMA_48HRS_DN	170	0.4089	2.0207	0	0.0021	0.169	7439	tags=48%, list=27%, signal=65%
BREASTCA_TWO_CLASSES	129	0.4229	2.026	0	0.0021	0.165	7184	tags=51%, list=26%, signal=69%
ERK5PATHWAY	19	0.647	2.0217	0.0019	0.0021	0.169	2673	tags=53%, list=10%, signal=58%
PARP_KO_UP	30	0.5841	2.0223	0	0.0021	0.168	6280	tags=60%, list=23%, signal=78%
MUNSHI_MM_UP	61	0.4827	2.0233	0	0.0021	0.167	5813	tags=48%, list=21%, signal=60%
YAGI_AML_PROGNOSIS	30	0.5717	2.016	0	0.0022	0.173	3531	tags=53%, list=13%, signal=61%

FALT_BCLL_DN	39	0.5288	2.0139	0	0.00 22	0.178	7138	tags=67%, list=26%, signal=90%
CELL_MOTILITY	102	0.4416	2.0099	0	0.00 23	0.189	6728	tags=46%, list=25%, signal=61%
ZUCCHI_EPITHELIAL_DN	35	0.5447	2.0099	0.0018	0.00 24	0.189	6543	tags=66%, list=24%, signal=86%
WNT_SIGNALING	57	0.4888	2.0029	0	0.00 26	0.202	8371	tags=54%, list=31%, signal=78%
LEI_HOXC8_UP	10	0.7562	2.0026	0	0.00 26	0.205	2965	tags=60%, list=11%, signal=67%
NAB_LUNG_DN	49	0.4943	2.0027	0	0.00 26	0.205	6254	tags=49%, list=23%, signal=63%
OXSTRESS_RPETWO_DN	100	0.4361	1.9987	0	0.00 27	0.214	5542	tags=46%, list=20%, signal=57%
SA_TRKA_RECEPTOR	18	0.649	1.9991	0	0.00 27	0.214	6229	tags=78%, list=23%, signal=101%
TGF_BETA_SIGNALING_PATHWAY	49	0.5042	1.9962	0	0.00 27	0.218	4445	tags=43%, list=16%, signal=51%
BRCA_BRCA1_POS	94	0.4408	1.9956	0	0.00 27	0.221	7249	tags=55%, list=26%, signal=75%
DNA_DAMAGE_SIGNALING	83	0.452	1.995	0	0.00 27	0.223	6988	tags=53%, list=26%, signal=71%
IGF1_NIH3T3_UP	31	0.5659	1.9918	0	0.00 28	0.231	7556	tags=61%, list=28%, signal=85%
MTORPATHWAY	20	0.6158	1.9889	0	0.00 28	0.235	7614	tags=75%, list=28%, signal=104%
VHL_NORMAL_UP	365	0.365	1.9892	0	0.00 28	0.234	6607	tags=41%, list=24%, signal=53%
IRITANI_ADPROX_VASC	135	0.4216	1.9899	0	0.00 28	0.234	5713	tags=44%, list=21%, signal=56%
ERKPATHWAY	45	0.5125	1.9845	0	0.00 3	0.247	5843	tags=49%, list=21%, signal=62%
UVB_NHEK3_C6	26	0.5764	1.9835	0	0.00 31	0.253	4350	tags=50%, list=16%, signal=59%
ROME_INSULIN_2F_UP	169	0.4039	1.9809	0	0.00 32	0.264	6454	tags=46%, list=24%, signal=60%
CHEN_HOXA5_TARGETS_DN	40	0.5188	1.9809	0	0.00 32	0.264	8388	tags=68%, list=31%, signal=97%
IGF1PATHWAY	22	0.6048	1.9757	0.0019	0.00 33	0.274	5843	tags=68%, list=21%, signal=87%

SARCOMAS_HISTIOCYTOMA_UP	9	0.7926	1.9767	0	0.0033	0.272	5332	tags=89%, list=19%, signal=110%
TNFALPHA_30MIN_UP	38	0.5206	1.9744	0	0.0034	0.281	3299	tags=39%, list=12%, signal=45%
RCC_NL_UP	442	0.3574	1.9748	0	0.0034	0.28	7004	tags=44%, list=26%, signal=58%
GERY_CEBP_TARGETS	108	0.4273	1.9722	0	0.0034	0.289	6447	tags=44%, list=24%, signal=58%
ALZHEIMERS_DISEASE_UP	1295	0.3296	1.9709	0	0.0035	0.293	6288	tags=36%, list=23%, signal=45%
HOFMANN_MDS_CD34_LOW_RISK	56	0.4813	1.9699	0	0.0035	0.296	4565	tags=46%, list=17%, signal=56%
HBX_NL_UP	21	0.6031	1.9684	0	0.0036	0.304	3070	tags=38%, list=11%, signal=43%
KANG_TERT_UP	76	0.4539	1.9657	0	0.0036	0.312	5792	tags=41%, list=21%, signal=52%
CARM_ERPATHWAY	23	0.5989	1.9664	0	0.0036	0.312	4773	tags=48%, list=17%, signal=58%
CHEN_LUNG_SURVIVAL	18	0.6293	1.9665	0	0.0036	0.312	4584	tags=56%, list=17%, signal=67%
RNA_POLYMERASE	13	0.6875	1.9666	0	0.0037	0.312	5402	tags=69%, list=20%, signal=86%
CMV_24HRS_DN	63	0.4751	1.9632	0	0.0037	0.322	4218	tags=43%, list=15%, signal=51%
PEART_HISTONE_DN	69	0.4636	1.9616	0	0.0038	0.33	4864	tags=43%, list=18%, signal=53%
ECMPATHWAY	24	0.5789	1.962	0.0019	0.0038	0.329	6427	tags=75%, list=23%, signal=98%
XU_ATRA_PLUSNSC_DN	13	0.6983	1.9578	0	0.0039	0.344	3735	tags=54%, list=14%, signal=62%
PYRIMIDINE_METABOLISM	54	0.4798	1.9519	0	0.004	0.366	5402	tags=44%, list=20%, signal=55%
TPA_SENS_MIDDLE_UP	61	0.4705	1.9521	0	0.004	0.365	4192	tags=49%, list=15%, signal=58%
SHEPARD_CELL_PROLIFERATION	198	0.3869	1.9522	0	0.004	0.364	5141	tags=37%, list=19%, signal=46%
ET743_RESIST_DN	39	0.5114	1.9526	0	0.004	0.363	2488	tags=38%, list=9%, signal=42%
IRS1_KO_ADIP_UP	90	0.4355	1.9537	0	0.0041	0.36	6473	tags=52%, list=24%, signal=68%

METHOTREXATE_PROBCELL_DN	10	0.7396	1.9527	0.0018	0.0041	0.363	5834	tags=80%, list=21%, signal=102%
GILDEA_BLADDER_UP	33	0.5455	1.9544	0	0.0041	0.359	4238	tags=48%, list=15%, signal=57%
DAVIES_MGUS_MM	31	0.5475	1.9545	0.0018	0.0041	0.359	6853	tags=65%, list=25%, signal=86%
BYSTRYKH_HSC_BRAIN_TRANS_GLOCUS	156	0.3984	1.9484	0	0.0042	0.378	5173	tags=37%, list=19%, signal=46%
TCA	15	0.671	1.9467	0	0.0042	0.385	3929	tags=53%, list=14%, signal=62%
DOX_RESIST_GASTRIC_UP	36	0.5207	1.9454	0	0.0042	0.391	5007	tags=44%, list=18%, signal=54%
GN_CAMP_GRANULOSA_DN	50	0.4872	1.9455	0	0.0043	0.39	5435	tags=46%, list=20%, signal=57%
SIG_INSULIN_RECEPTOR_PATHWAY_IN_CARDIAC_MYOCYTES	50	0.4965	1.9435	0	0.0043	0.396	6939	tags=54%, list=25%, signal=72%
INNEREAR_UP	39	0.5213	1.9446	0	0.0043	0.394	2596	tags=33%, list=9%, signal=37%
CELL_PROLIFERATION	198	0.3869	1.9437	0	0.0043	0.395	5141	tags=37%, list=19%, signal=46%
PENG_GLUTAMINE_UP	205	0.3808	1.9404	0	0.0043	0.405	7047	tags=46%, list=26%, signal=61%
BCNU_GLIOMA_NOMGMT_48HRS_DN	26	0.5622	1.935	0	0.0046	0.423	6549	tags=58%, list=24%, signal=76%
TPA_SENS_LATE_UP	51	0.4803	1.9356	0	0.0046	0.422	4136	tags=49%, list=15%, signal=58%
HIPPOCAMPUS_DEVELOPMENT_NEONATAL	26	0.5619	1.9332	0.0019	0.0046	0.43	3403	tags=50%, list=12%, signal=57%
STEMCELL_COMMON_UP	162	0.3907	1.9298	0	0.0048	0.446	5958	tags=39%, list=22%, signal=49%
REFRACTORY_GASTRIC_UP	67	0.4531	1.9283	0	0.0049	0.453	5230	tags=40%, list=19%, signal=50%
ROTH_HTERT_DIFF	29	0.5408	1.9283	0.0041	0.0049	0.453	6377	tags=66%, list=23%, signal=85%
NEMETH_TNF_UP	74	0.4508	1.9254	0	0.0049	0.458	6311	tags=49%, list=23%, signal=63%
ZHAN_MM_CD138_CD1_VS_REST	48	0.4848	1.9216	0	0.0051	0.464	5427	tags=42%, list=20%, signal=52%
SPPAPATHWAY	24	0.5852	1.9208	0.0019	0.0051	0.465	5843	tags=67%, list=21%, signal=85%

MIDDLEAGE_DN	14	0.6594	1.9193	0	0.00 52	0.471	8958	tags=93%, list=33%, signal=138%

Table 4.6. GSEA gene sets elevated in Met^{mt} solid vs. mixed pathology tumors.

Gene Set	Size	ES	NES	NOM p-val	FDR q-val	FWER p-val	Rank at Max	Leading Edge
TGZ_ADIP_UP	14	-0.557	-1.609	0.0263	0.7245	1	5291	tags=50%, list=19%, signal=62%
AGEING_KIDNEY_DN	102	-0.38	-1.814	0.0023	0.7367	0.911	4826	tags=37%, list=18%, signal=45%
FEEDERPATHWAY	8	-0.553	-1.345	0.1447	0.7687	1	1023	tags=38%, list=4%, signal=39%
SULFUR_METABOLISM	8	-0.535	-1.316	0.1563	0.7716	1	8210	tags=50%, list=30%, signal=71%
TNFALPHA_ADIP_DN	59	-0.313	-1.321	0.0773	0.7718	1	5582	tags=29%, list=20%, signal=36%
FSH_HUMAN_GRANULOSA_UP	11	-0.493	-1.349	0.1142	0.7779	1	8342	tags=45%, list=30%, signal=65%
HUMAN_TISSUE_SALIVARY	8	-0.539	-1.334	0.1481	0.7882	1	4189	tags=50%, list=15%, signal=59%
GH_EXOGENOUS_LATE_DN	62	-0.315	-1.353	0.063	0.7888	1	5245	tags=32%, list=19%, signal=40%
AS3_FIBRO_C4	18	-0.424	-1.321	0.1175	0.7949	1	2328	tags=22%, list=9%, signal=24%
GPCRDB_CLASS_B_SECRETIN_LIKE	19	-0.428	-1.357	0.1013	0.8021	1	4175	tags=32%, list=15%, signal=37%
4NQO_UNIQUE_FIBRO_DN	9	-0.522	-1.303	0.173	0.8033	1	4723	tags=33%, list=17%, signal=40%
NITROGEN_METABOLISM	20	-0.404	-1.324	0.1384	0.8076	1	4856	tags=30%, list=18%, signal=36%
HDACI_COLON_SUL12HRS_DN	25	-0.481	-1.682	0.0165	0.8097	0.999	5167	tags=40%, list=19%, signal=49%
KIM_TH_CELLS_UP	42	-0.396	-1.571	0.011	0.819	1	3442	tags=29%, list=13%, signal=33%
VANTVEER_BREAST_OUTCOME_GOOD_VS_POOR_UP	25	-0.384	-1.36	0.1087	0.8198	1	5430	tags=40%, list=20%, signal=50%
AGEING_KIDNEY_SPECIFIC_DN	107	-0.291	-1.362	0.0362	0.8428	1	4761	tags=31%, list=17%, signal=37%
LEE_DENA_DN	87	-0.349	-1.612	0.0046	0.8493	1	5852	tags=34%, list=21%, signal=44%
TCAPOPTOSISPATHWAY	6	-0.698	-1.545	0.0424	0.8554	1	2204	tags=50%, list=8%, signal=54%

TOB1PATHWAY	17	-0.545	-1.721	0.0099	0.8714	0.995	2204	tags=35%, list=8%, signal=38%
LEE_MYC_TGFA_DN	64	-0.313	-1.363	0.0574	0.8728	1	5938	tags=31%, list=22%, signal=40%
GPCRS_CLASS_B_SECRETIN_LIKE	19	-0.428	-1.368	0.0826	0.8861	1	4175	tags=32%, list=15%, signal=37%
HINATA_NFKB_DN	18	-0.399	-1.256	0.1582	0.8943	1	5205	tags=33%, list=19%, signal=41%
BRCA_ER_POS	463	-0.218	-1.249	0.013	0.9052	1	6548	tags=31%, list=24%, signal=40%
ZMPSTE24_KO_DN	45	-0.318	-1.257	0.1456	0.9107	1	3110	tags=22%, list=11%, signal=25%
SARCOMAS_GISTROMAL_UP	13	-0.448	-1.273	0.1701	0.911	1	3009	tags=38%, list=11%, signal=43%
POD1_KO_DN	657	-0.212	-1.262	0.0054	0.9136	1	4323	tags=22%, list=16%, signal=26%
CMV_HCMV_TIMECOURSE_16HRS_DN	18	-0.391	-1.219	0.2096	0.9156	1	4939	tags=33%, list=18%, signal=41%
GH_IGF_CHONDROCYTES_UP	10	-0.483	-1.242	0.2112	0.916	1	1654	tags=30%, list=6%, signal=32%
NELSON_ANDROGEN_UP	77	-0.311	-1.369	0.0457	0.9178	1	3139	tags=23%, list=11%, signal=26%
CPR_LOW_LIVER_UP	34	-0.325	-1.223	0.1871	0.9203	1	10507	tags=50%, list=38%, signal=81%
MARSHALL_SPLEEN_BAL	21	-0.374	-1.227	0.183	0.9206	1	5000	tags=38%, list=18%, signal=47%
ADIP_DIFF_UP	76	-0.278	-1.264	0.0826	0.9265	1	5582	tags=29%, list=20%, signal=36%
AGUIRRE_PANCREAS_CHR18	13	-0.415	-1.212	0.2091	0.9277	1	2996	tags=31%, list=11%, signal=35%
GH_EXOGENOUS_ANY_DN	73	-0.274	-1.234	0.1351	0.9295	1	5245	tags=30%, list=19%, signal=37%
BRENTANI_CANCER_TESTIS	14	-0.416	-1.228	0.2049	0.9361	1	2632	tags=21%, list=10%, signal=24%
CHESLER_D6MIT150_CIS_GLOCUS	7	-0.574	-1.373	0.1419	0.9424	1	5497	tags=71%, list=20%, signal=89%
JISON_SICKLE_CELL	31	-0.267	-0.974	0.4832	0.9436	1	2775	tags=23%, list=10%, signal=25%
GLYCOSAMINOGLYCAN_DEGRADATION	10	-0.455	-1.2	0.2342	0.9439	1	1298	tags=30%, list=5%, signal=31%

BRCA1KO_MEF_DN	74	-0.219	-0.98	0.4797	0.9443	1	2495	tags=16%, list=9%, signal=18%
INSULIN_ADIP_INSENS_UP	21	-0.296	-0.976	0.486	0.9446	1	5127	tags=33%, list=19%, signal=41%
NOVA2_KO_SPLICING	36	-0.257	-0.984	0.4875	0.9461	1	6810	tags=44%, list=25%, signal=59%
CMV_HCMV_TIMECOURSE_10HRS_DN	12	-0.356	-0.981	0.4716	0.947	1	626	tags=17%, list=2%, signal=17%
ZHAN_MM_CD1_VS_CD2_UP	74	-0.222	-0.985	0.4786	0.9493	1	3510	tags=20%, list=13%, signal=23%
HOGERKORP_ANTI_CD44_DN	11	-0.368	-0.976	0.4787	0.9494	1	832	tags=18%, list=3%, signal=19%
SANSOM_APC_4_DN	65	-0.221	-0.969	0.4621	0.9514	1	5541	tags=25%, list=20%, signal=31%
CITED1_KO_WT_UP	12	-0.351	-0.986	0.4669	0.9516	1	1151	tags=17%, list=4%, signal=17%
AGED_MOUSE_RETINA_ANY_DN	12	-0.502	-1.38	0.1107	0.9517	1	6077	tags=67%, list=22%, signal=86%
TGFBETA_C3_UP	11	-0.449	-1.202	0.2516	0.9575	1	3276	tags=45%, list=12%, signal=52%
BYSTRYKH_SCP2_CIS_TRANSCRIPTS_GLOCUS	5	-0.458	-0.965	0.4876	0.9577	1	4625	tags=40%, list=17%, signal=48%
TNFALPHA_TGZ_ADIP_UP	21	-0.306	-0.987	0.4754	0.9578	1	8220	tags=48%, list=30%, signal=68%
SHIPP_DLBCL_CURED_UP	41	-0.246	-0.962	0.5078	0.9618	1	6091	tags=32%, list=22%, signal=41%
WILLERT_WNT_NCCIT_ALL_UP	17	-0.316	-0.988	0.4505	0.9631	1	2889	tags=35%, list=11%, signal=39%
METASTASIS_ADENOCARC_DN	30	-0.257	-0.916	0.6141	0.9657	1	5602	tags=30%, list=20%, signal=38%
HADDAD_HSC_CD7_DN	79	-0.202	-0.912	0.6623	0.9658	1	3457	tags=20%, list=13%, signal=23%
KIM_TH_CELLS_DN	15	-0.46	-1.387	0.1071	0.9667	1	4254	tags=47%, list=16%, signal=55%
GPCRDB_CLASS_A_RHODOPSIN_LIKE	164	-0.181	-0.913	0.6981	0.9692	1	2682	tags=12%, list=10%, signal=13%
RETT_DN	13	-0.341	-0.988	0.4842	0.9698	1	6353	tags=46%, list=23%, signal=60%
GLYCOGEN	17	-0.301	-0.927	0.5569	0.9705	1	344	tags=12%, list=1%, signal=12%

FALT_BCLL_IG_MUTATED_VS_WT_DN	40	-0.236	-0.917	0.6112	0.9708	1	1654	tags=15%, list=6%, signal=16%
IFNGPATHWAY	6	-0.42	-0.938	0.5236	0.9714	1	314	tags=17%, list=1%, signal=17%
MENSE_HYPOXIA_TRANSPORTER_GENES	45	-0.253	-0.989	0.4859	0.9722	1	2148	tags=16%, list=8%, signal=17%
HISTIDINE_METABOLISM	27	-0.263	-0.918	0.5603	0.9725	1	3427	tags=22%, list=13%, signal=25%
BRCA_PROGNOSIS_POS	35	-0.373	-1.395	0.0609	0.9735	1	7993	tags=54%, list=29%, signal=77%
ABBUD_LIF_DN	23	-0.279	-0.939	0.5388	0.9746	1	9791	tags=70%, list=36%, signal=108%
CHESLER_D6MIT150_NEURAL_TARGETS_GLOCUS	7	-0.421	-0.993	0.4555	0.9747	1	4500	tags=43%, list=16%, signal=51%
TAKEDA_NUP8_HOXA9_10D_UP	164	-0.179	-0.922	0.6791	0.975	1	4413	tags=20%, list=16%, signal=23%
HADDAD_CD45CD7_PLUS_VS_MINUS_DN	79	-0.202	-0.919	0.6032	0.9752	1	3457	tags=20%, list=13%, signal=23%
HOFMANN_MANTEL_LYMPHOMA_VS_LYMPH_NODES_DN	40	-0.244	-0.932	0.5817	0.9752	1	3975	tags=25%, list=15%, signal=29%
GH_AUTOCRINE_DN	109	-0.195	-0.928	0.6312	0.9763	1	6653	tags=34%, list=24%, signal=45%
TAKEDA_NUP8_HOXA9_3D_UP	148	-0.189	-0.934	0.6159	0.9765	1	4974	tags=24%, list=18%, signal=29%
CMV_HCMV_TIMECOURSE_48HRS_UP	64	-0.213	-0.93	0.6146	0.9766	1	4096	tags=23%, list=15%, signal=27%
INOSITOL_PHOSPHATE_METABOLISM	22	-0.279	-0.923	0.5539	0.977	1	3784	tags=23%, list=14%, signal=26%
TAKEDA_NUP8_HOXA9_6H_UP	74	-0.221	-0.99	0.4551	0.979	1	4413	tags=22%, list=16%, signal=26%
CHESLER_HIGHEST_FOLD_RANGE_GENES	37	-0.252	-0.944	0.5517	0.9799	1	4775	tags=24%, list=17%, signal=29%
POD1_KO_MOST_DN	24	-0.355	-1.188	0.2149	0.9801	1	6517	tags=42%, list=24%, signal=55%
AMINOSUGARS_METABOLISM	15	-0.314	-0.939	0.5184	0.9801	1	7225	tags=47%, list=26%, signal=63%
ST_INTERFERON_GAMMA_PATHWAY	10	-0.361	-0.952	0.5118	0.9809	1	7995	tags=50%, list=29%, signal=71%
N_GLYCAN_DEGRADATION	13	-0.323	-0.941	0.5213	0.9809	1	7276	tags=46%, list=27%, signal=63%

GABAPATHWAY	12	-0.344	-0.95	0.4969	0.9817	1	8161	tags=50%, list=30%, signal=71%
IL17PATHWAY	12	-0.337	-0.945	0.5031	0.982	1	3266	tags=33%, list=12%, signal=38%
GH_GHRHR_KO_6HRS_DN	31	-0.281	-1.019	0.4283	0.9824	1	8537	tags=45%, list=31%, signal=66%
ELECTRON_TRANSPORT	80	-0.198	-0.895	0.6797	0.9826	1	7934	tags=38%, list=29%, signal=53%
ST_JAK_STAT_PATHWAY	10	-0.383	-0.993	0.4716	0.9827	1	6498	tags=50%, list=24%, signal=66%
GANGLIOSIDE_BIOSYNTHESIS	8	-0.367	-0.9	0.5846	0.9829	1	1881	tags=25%, list=7%, signal=27%
GPCRS_CLASS_A_RHODOPSIN_LIKE_2	10	-0.357	-0.954	0.5162	0.983	1	4017	tags=20%, list=15%, signal=23%
LEE_TCELLS9_UP	24	-0.265	-0.903	0.622	0.9844	1	2759	tags=21%, list=10%, signal=23%
CISPLATIN_PROBCELL_UP	15	-0.342	-1.004	0.4295	0.9848	1	3225	tags=27%, list=12%, signal=30%
CMV_HCMV_TIMECOURSE_18HRS_DN	19	-0.284	-0.896	0.6227	0.9867	1	2950	tags=26%, list=11%, signal=29%
LIZUKA_L1_GR_G1	20	-0.309	-1.02	0.4093	0.987	1	4008	tags=35%, list=15%, signal=41%
HEMATOP_STEM_ALL_UP	35	-0.188	-0.697	0.9204	0.9882	1	2908	tags=14%, list=11%, signal=16%
CTLPATHWAY	11	-0.348	-0.946	0.5235	0.9882	1	4091	tags=36%, list=15%, signal=43%
GPCRDB_OTHER	54	-0.217	-0.9	0.6739	0.9883	1	6061	tags=24%, list=22%, signal=31%
MONOAMINE_GPCRS	28	-0.272	-0.994	0.4507	0.9887	1	10315	tags=50%, list=38%, signal=80%
LEE_TCELLS4_UP	37	-0.229	-0.884	0.6539	0.9901	1	3457	tags=22%, list=13%, signal=25%
HASLINGER_B_CLL_MUTATED	27	-0.245	-0.873	0.6438	0.9903	1	4186	tags=19%, list=15%, signal=22%
C21_STEROID_HORMONE_METABOLISM	22	-0.209	-0.699	0.884	0.9905	1	7469	tags=32%, list=27%, signal=44%
ZHAN_MM_MOLECULAR_CLASSI_DN	44	-0.253	-1.005	0.4391	0.9912	1	3399	tags=23%, list=12%, signal=26%
ST_STAT3_PATHWAY	12	-0.311	-0.87	0.6373	0.992	1	6498	tags=50%, list=24%, signal=66%

AGED_MOUSE_RETINA_ANY_UP	18	-0.274	-0.885	0.6364	0.9922	1	5672	tags=39%, list=21%, signal=49%
NELSON_ANDROGEN_DN	13	-0.344	-0.995	0.4482	0.9922	1	2590	tags=31%, list=9%, signal=34%
STOSSI_ER_UP	46	-0.218	-0.874	0.6825	0.9927	1	6353	tags=35%, list=23%, signal=45%
HDACI_COLON_SUL2HRS_UP	10	-0.265	-0.702	0.8299	0.9927	1	6057	tags=40%, list=22%, signal=51%
HBX_NL_DN	11	-0.328	-0.889	0.6124	0.9947	1	968	tags=18%, list=4%, signal=19%
CPR_NULL-LOW_LIVER_UP	37	-0.268	-1.009	0.4444	0.9949	1	6747	tags=38%, list=25%, signal=50%
TSADAC_HYPERMETH_HYPERAC_OVCA_UP	7	-0.378	-0.876	0.625	0.9949	1	80	tags=14%, list=0%, signal=14%
TH1TH2PATHWAY	16	-0.292	-0.886	0.6167	0.9953	1	592	tags=13%, list=2%, signal=13%
ELECTRON_TRANSPORT_CHAIN	84	-0.219	-0.999	0.4596	0.9953	1	3355	tags=20%, list=12%, signal=23%
LEE_MYC_DN	70	-0.234	-1.02	0.4095	0.9955	1	5938	tags=30%, list=22%, signal=38%
VOBESITYPATHWAY	6	-0.246	-0.545	0.9706	0.9956	1	1685	tags=17%, list=6%, signal=18%
HEME_BIOSYNTHESIS	9	-0.273	-0.684	0.8806	0.9956	1	377	tags=11%, list=1%, signal=11%
CREMPATHWAY	7	-0.284	-0.673	0.8524	0.9959	1	619	tags=14%, list=2%, signal=15%
HUMAN_TISSUE_LIVER	42	-0.217	-0.861	0.6992	0.9962	1	6170	tags=24%, list=23%, signal=31%
REELINPATHWAY	6	-0.446	-0.996	0.4573	0.9963	1	1322	tags=33%, list=5%, signal=35%
PEPIPATHWAY	5	-0.277	-0.572	0.9589	0.9966	1	1971	tags=20%, list=7%, signal=22%
HUMAN_TISSUE_KIDNEY	11	-0.203	-0.556	0.9738	0.9969	1	7207	tags=27%, list=26%, signal=37%
PELP1PATHWAY	6	-0.39	-0.852	0.6486	0.997	1	3411	tags=33%, list=12%, signal=38%
LEE_ACOX1_DN	73	-0.156	-0.702	0.9609	0.997	1	5202	tags=19%, list=19%, signal=24%
STEROID_BIOSYNTHESIS	13	-0.308	-0.879	0.6099	0.9973	1	7885	tags=46%, list=29%, signal=65%

BRENTANI_HORMONAL_FUNCTION	12	-0.296	-0.859	0.6304	0.9974	1	3985	tags=33%, list=15%, signal=39%
TFF2_KO_UP	21	-0.306	-1.01	0.4451	0.9976	1	1594	tags=14%, list=6%, signal=15%
HDACI_COLON_SUL2HRS_DN	11	-0.26	-0.705	0.848	0.9977	1	4619	tags=36%, list=17%, signal=44%
AHSPPATHWAY	9	-0.329	-0.856	0.6196	0.9978	1	6523	tags=44%, list=24%, signal=58%
SARCOMAS_LIPOSARCOMA_UP	10	-0.33	-0.865	0.636	0.9981	1	7160	tags=50%, list=26%, signal=68%
AS3_FIBRO_DN	30	-0.238	-0.863	0.6915	0.9981	1	2328	tags=13%, list=9%, signal=15%
WELCH_GATA1	19	-0.143	-0.457	0.998	0.9982	1	9292	tags=37%, list=34%, signal=56%
NEUROTRANSMITTERSPATHWAY	6	-0.286	-0.644	0.8676	0.9984	1	19539	tags=100%, list=71%, signal=350%
TERCPATHWAY	6	-0.277	-0.611	0.9188	0.9989	1	6979	tags=50%, list=26%, signal=67%
TNFALPHA_ADIP_UP	17	-0.19	-0.577	0.9739	0.9993	1	8220	tags=47%, list=30%, signal=67%
VEGF_HUVEC_UP	12	-0.304	-0.853	0.6693	0.9994	1	6653	tags=33%, list=24%, signal=44%
GOLDRATH_CYTOLYTIC	19	-0.274	-0.876	0.6507	0.9999	1	6171	tags=32%, list=23%, signal=41%
BENNETT_SLE_UP	31	-0.127	-0.459	0.9979	1	1	8807	tags=39%, list=32%, signal=57%
NUMATA_G_CSF_DIFF	16	-0.191	-0.579	0.9669	1	1	4999	tags=25%, list=18%, signal=31%
NO2IL12PATHWAY	14	-0.201	-0.589	0.9607	1	1	1119	tags=7%, list=4%, signal=7%
CANTHARIDIN_UP	11	-0.228	-0.616	0.9298	1	1	5814	tags=27%, list=21%, signal=35%
BIOGENIC_AMINE_SYNTHESIS	14	-0.211	-0.621	0.9281	1	1	467	tags=7%, list=2%, signal=7%
LIZUKA_G1_GR_G2	11	-0.232	-0.625	0.9096	1	1	7995	tags=45%, list=29%, signal=64%
RIBOFLAVIN_METABOLISM	9	-0.252	-0.647	0.918	1	1	3661	tags=22%, list=13%, signal=26%
EPONFKBPATHWAY	11	-0.242	-0.65	0.9071	1	1	5076	tags=27%, list=19%, signal=33%

IFNA_UV-CMV_COMMON_HCMV_6HRS_UP	31	-0.18	-0.654	0.962	1	1	7131	tags=35%, list=26%, signal=48%
IFNALPHA_NL_HCC_UP	29	-0.184	-0.673	0.9406	1	1	3321	tags=10%, list=12%, signal=12%
CITED1_KO_WT_DN	13	-0.246	-0.706	0.8514	1	1	7659	tags=46%, list=28%, signal=64%
SHHPATHWAY	13	-0.248	-0.707	0.8704	1	1	2086	tags=15%, list=8%, signal=17%
IFNALPHA_NL_UP	37	-0.189	-0.712	0.9104	1	1	4309	tags=14%, list=16%, signal=16%
CAMPTOTHECIN_PROBCELL_UP	19	-0.221	-0.715	0.8843	1	1	7424	tags=42%, list=27%, signal=58%
HDACI_COLON_BUT30MIN_DN	29	-0.199	-0.716	0.8747	1	1	7116	tags=34%, list=26%, signal=47%
GPCRDB_CLASS_C_METABOTROPIC_GLUTAMATE_PHEROMONE	12	-0.255	-0.717	0.8641	1	1	7160	tags=33%, list=26%, signal=45%
SA_DIACYLGLYCEROL_SIGNALING	10	-0.271	-0.718	0.8143	1	1	1074	tags=10%, list=4%, signal=10%
CHESLER_D6MIT150_TRANSCRIPTION_FACTORS_GLOCUS	8	-0.3	-0.732	0.8	1	1	478	tags=13%, list=2%, signal=13%
DISTECHE_XINACTIVATED_GENES	18	-0.233	-0.733	0.8017	1	1	20994	tags=100%, list=77%, signal=430%
ADIP_HUMAN_UP	54	-0.18	-0.736	0.9109	1	1	5394	tags=24%, list=20%, signal=30%
CANCERDRUGS_PROBCELL_UP	17	-0.237	-0.737	0.8337	1	1	7424	tags=41%, list=27%, signal=56%
CMV_HCMV_6HRS_UP	23	-0.218	-0.738	0.8559	1	1	7660	tags=39%, list=28%, signal=54%
MAPKKK_CASCADE	9	-0.296	-0.739	0.7805	1	1	2710	tags=22%, list=10%, signal=25%
MMS_HUMAN_LYMPH_LOW_4HRS_DN	15	-0.243	-0.743	0.7854	1	1	8284	tags=40%, list=30%, signal=57%
WANG_HOXA9_VS_MEIS1_UP	27	-0.213	-0.747	0.8353	1	1	2635	tags=15%, list=10%, signal=16%
RABPATHWAY	9	-0.292	-0.75	0.7958	1	1	2328	tags=22%, list=9%, signal=24%
ASBCELLPATHWAY	8	-0.301	-0.751	0.7653	1	1	592	tags=13%, list=2%, signal=13%
SELENOAMINO_ACID_METABOLISM	11	-0.285	-0.756	0.7879	1	1	109	tags=9%, list=0%, signal=9%

IGFR_IR_DN	7	-0.326	-0.759	0.7643	1	1	9101	tags=71%, list=33%, signal=107%
CPR_NULL_LIVER_UP	51	-0.186	-0.768	0.8693	1	1	6747	tags=29%, list=25%, signal=39%
GRANDVAUX_IRF3_UP	10	-0.29	-0.772	0.7559	1	1	6645	tags=40%, list=24%, signal=53%
SCHURINGA_STAT5A_DN	17	-0.256	-0.775	0.7973	1	1	1292	tags=12%, list=5%, signal=12%
TAKEDA_NUP8_HOXA9_8D_UP	136	-0.158	-0.777	0.9289	1	1	8177	tags=35%, list=30%, signal=50%
ST_TYPE_I_INTERFERON_PATHWAY	9	-0.302	-0.778	0.7356	1	1	7995	tags=56%, list=29%, signal=78%
NF90_UP	21	-0.233	-0.779	0.7578	1	1	8807	tags=52%, list=32%, signal=77%
HDACI_COLON_TSA48HRS_DN	7	-0.336	-0.779	0.7286	1	1	34	tags=14%, list=0%, signal=14%
GH_EXOGENOUS_LATE_UP	66	-0.183	-0.782	0.8561	1	1	4367	tags=17%, list=16%, signal=20%
EEA1PATHWAY	7	-0.337	-0.787	0.7149	1	1	2298	tags=29%, list=8%, signal=31%
HEPARAN_SULFATE_BIOSYNTHESIS	8	-0.326	-0.79	0.7308	1	1	1486	tags=25%, list=5%, signal=26%
MOUSE_TISSUE_KIDNEY	11	-0.287	-0.795	0.7206	1	1	10019	tags=64%, list=37%, signal=100%
ET743_SARCOMA_24HRS_UP	8	-0.318	-0.8	0.7082	1	1	156	tags=13%, list=1%, signal=13%
MARTINELLI_IFNS_DIFF	18	-0.253	-0.8	0.7835	1	1	20449	tags=100%, list=75%, signal=396%
CMV_HCMV_TIMECOURSE_12HRS_UP	28	-0.221	-0.8	0.7906	1	1	7131	tags=39%, list=26%, signal=53%
UV_UNIQUE_FIBRO_DN	27	-0.232	-0.803	0.7665	1	1	1555	tags=11%, list=6%, signal=12%
INSULIN_ADIP_INSENS_DN	14	-0.274	-0.805	0.7364	1	1	5675	tags=36%, list=21%, signal=45%
ABBUD_LIF_GH3_DN	5	-0.381	-0.807	0.6867	1	1	4429	tags=40%, list=16%, signal=48%
CTLA4PATHWAY	18	-0.258	-0.812	0.7239	1	1	2204	tags=17%, list=8%, signal=18%
HOFMANN_MDS_CD34_HIGH_RISK	21	-0.252	-0.814	0.7381	1	1	7690	tags=57%, list=28%, signal=79%

CHONDROITIN	8	-0.326	-0.816	0.7079	1	1	1486	tags=25%, list=5%, signal=26%
AD12_24HRS_DN	19	-0.253	-0.817	0.706	1	1	3652	tags=21%, list=13%, signal=24%
NOTCHPATHWAY	5	-0.393	-0.818	0.709	1	1	1660 0	tags=100%, list=61%, signal=254%
PITUITARY_FETAL_UP	17	-0.264	-0.82	0.7332	1	1	3791	tags=24%, list=14%, signal=27%
PENTOSE_PHOSPHATE_PATHWAY	22	-0.243	-0.823	0.7433	1	1	5193	tags=32%, list=19%, signal=39%
CYTOKINEPATHWAY	26	-0.236	-0.831	0.7395	1	1	8438	tags=38%, list=31%, signal=56%
HDACI_COLON_SUL30MIN_DN	33	-0.224	-0.834	0.7039	1	1	7116	tags=33%, list=26%, signal=45%
CPR_NULL_LIVER_DN	22	-0.253	-0.835	0.7044	1	1	7842	tags=50%, list=29%, signal=70%
LEE_CIP_DN	80	-0.186	-0.845	0.7816	1	1	7753	tags=35%, list=28%, signal=49%
HCC_SURVIVAL_GOOD_VS_POOR_UP	152	-0.202	-1.005	0.4113	1	1	4682	tags=19%, list=17%, signal=23%
IGLESIAS_E2FMINUS_DN	20	-0.331	-1.012	0.465	1	1	8342	tags=45%, list=30%, signal=65%
HUMAN_TISSUE_PANCREAS	31	-0.283	-1.021	0.4289	1	1	5700	tags=35%, list=21%, signal=45%
BECKER_TAMOXIFEN_RESISTANT_DN	45	-0.256	-1.022	0.3987	1	1	1828	tags=18%, list=7%, signal=19%
ZHAN_PCS_MULTIPLE_MYELOMA_SPKD	21	-0.312	-1.023	0.4222	1	1	3321	tags=29%, list=12%, signal=32%
IFNA_HCMV_6HRS_UP	52	-0.249	-1.025	0.397	1	1	4438	tags=27%, list=16%, signal=32%
DAVIES_N	10	-0.383	-1.026	0.4238	1	1	5194	tags=40%, list=19%, signal=49%
HSIAO_LIVER_SPECIFIC_GENES	281	-0.188	-1.026	0.4029	1	1	5887	tags=25%, list=22%, signal=31%
HPV31_DN	43	-0.256	-1.028	0.4066	1	1	2619	tags=19%, list=10%, signal=21%
VEGF_HUVEC_2HRS_UP	25	-0.3	-1.028	0.4196	1	1	5747	tags=28%, list=21%, signal=35%
MMS_MOUSE_LYMPH_HIGH_24HRS_UP	21	-0.311	-1.028	0.3949	1	1	2680	tags=29%, list=10%, signal=32%

RORIE_ES_PNET_DN	27	-0.296	-1.03	0.436	1	1	1892	tags=19%, list=7%, signal=20%
TRYPTOPHAN_METABOLISM	67	-0.238	-1.039	0.3681	1	1	7934	tags=42%, list=29%, signal=59%

KEGG ID	P-value	ExpCount	Count	Size	Description
03010	1.75E-07	10.72388	29	84	Ribosome
04540	6.37E-06	11.23454	27	88	Gap junction
04510	1.31E-05	23.36274	44	183	Focal adhesion
01430	0.000301	14.42617	28	113	Cell Communication
04512	0.000703	10.85155	22	85	ECM-receptor interaction
00970	0.004672	3.446962	9	27	Aminoacyl-tRNA biosynthesis
05210	0.006595	9.064232	17	71	Colorectal cancer
04610	0.011369	8.808902	16	69	Complement and coagulation cascades
04110	0.012874	13.53252	22	106	Cell cycle
04730	0.013059	8.936567	16	70	Long-term depression
04310	0.01971	18.12846	27	142	Wnt signaling pathway
04670	0.023732	14.29851	22	112	Leukocyte transendothelial migration
04662	0.028105	8.170576	14	64	B cell receptor signaling pathway
00400	0.029281	1.276652	4	10	Phenylalanine, tyrosine and tryptophan biosynthesis
00790	0.043299	4.085288	8	32	Folate biosynthesis
04350	0.047123	10.34088	16	81	TGF-beta signaling pathway

Table 4.7. Overrepresented KEGG categories in genes elevated in Met^{mt} mixed pathology vs. solid tumors.

KEGG ID	P-value	ExpCount	Count	Size	Description
04514	8.42E-06	13.06237	30	145	Cell adhesion molecules (CAMs)
04940	6.16E-05	5.495203	16	61	Type I diabetes mellitus
04612	0.000308	7.477079	18	83	Antigen processing and presentation
04530	0.010267	10.08955	18	112	Tight junction
04640	0.010931	7.206823	14	80	Hematopoietic cell lineage
04910	0.023965	11.80117	19	131	Insulin signaling pathway
00604	0.049336	1.441365	4	16	Glycosphingolipid biosynthesis - ganglioseries

Table 4.8. Overrepresented KEGG categories in genes elevated in Met^{mt} solid vs. mixed pathology tumors.

Table 4.9. MET levels from microarray data.

	GENE ANNOTATIONS				*SAMPLE SIZES*		*CLASS MEANS*			*STANDARD T-TEST*		
UHN ID	*GB Acc*	*Symbol*	*Name*	*Cytoband*	*Early recurrence (n_1=48)*	*Long follow-up (n_2=41)*	*Early recurrence*	*Long follow-up*	*Fold difference*	*Rank*	*t-statistic*	*p-value*
UHNhcdna0015084	AA005247	MET	MET	7q31	48	41	0.947	0.387	2.446	1054	2.53	1.33E-02

Table 4.10. Disease-free Survival Analysis by Cox Proportional Hazards Model (with Firth Correction) for MET protein status in basal, non-basal subgroups of ANN breast cancer (n=506)

Prognostic Factor		Univariate				Multivariate*			
	Subgroup	RR	95% CI		P-value	RR	95% CI		P-value
MET protein positive vs. negative	Basal	2.44	0.51	23.41	0.2840	3.02	0.62	29.296	0.1818
MET protein positive vs. negative	Non-Basal	1.57	0.92	2.71	0.0961	1.49	0.86	2.62	0.1586
Relative HR Difference between Basal and Non-basal		1.55	0.30	15.62	0.6309	2.02	0.38	20.62	0.4336

*Adjusting for menopausal status, tumor size, grade, lymphatic invasion, age at diagnosis, and adjuvant treatment

Table 4.11. Association between clinical/pathological characteristics and MET protein in axillary node negative (ANN) breast cancer.

Characteristic		MET Low (0-3†) (n=376)		MET High (4-8†) (n=292)		P-value
		Number	%	Number	%	
Menopausal status	Pre	122	32.4	102	34.9	0.6589
	Peri	15	4.0	14	4.8	
	Post	239	63.6	176	60.3	
Tumor Size	< 0.5 cm	10	2.6	3	1.0	0.0103
	0.5 to < 1.0 cm	60	16.0	30	10.3	
	1.0 to < 2.0 cm	169	45.0	117	40.1	
	2 to 5 cm	123	32.7	130	44.5	
	> 5 cm	14	3.7	12	4.1	
Estrogen receptor	Positive	240	63.8	176	60.3	0.0006
	Negative/Equivocal	61	16.2	87	29.8	
	ND‡	75	20.0	29	9.9	
Progesterone receptor	Positive	224	59.6	148	50.7	<0.0001
	Negative/Equivocal	77	20.5	115	39.4	
	ND‡	75	19.9	29	9.9	
Histological grade	1*	142	37.8	67	23.0	<0.0001
	2	134	35.6	97	33.2	
	3	73	19.4	97	33.2	
	ND‡	27	7.2	31	10.6	
Lymphatic invasion	Absent	331	88.0	255	87.3	0.8744
	Present	45	12.0	36	12.3	
	ND‡	0	0.0	1	0.4	
Adjuvant treatment	Hormonal	172	45.7	102	34.9	0.0042
	Chemotherapy	48	12.8	57	19.5	
	Both	7	1.9	13	4.5	
	None	149	39.6	120	41.1	
ERBB2	Negative	337	92.8	257	91.8	0.6182
	Positive	26	7.2	23	8.2	
	NA‡	13	---	12	---	
TP53	Negative	289	83.5	189	68.7	<0.0001
	Positive	57	16.5	86	31.3	
	NA‡	30	---	17	---	
EGFR	Negative	309	97.2	223	89.9	0.0003
	Positive	9	2.8	25	10.1	
	NA‡	58	---	44	---	
SNAIL	Negative	203	70.2	171	69.2	0.7994
	Positive	86	29.8	76	30.8	
	NA‡	87	---	45	---	
Age (years)	Mean	56.2		55.4		
	SD	11.5		11.6		
	Minimum	25.5		28.6		
	Maximum	75.8		75.8		

†Allred score
‡ND: Unknown, not done or missing (related to small tumor size, excluded from test of association)
NA: Not available due to insufficient material or IMH failure
*Includes mucinous, lobular and tubular subtypes

Table 4.12. Results of Disease-free Survival Analysis by Cox Proportional Hazards Model for MET and SNAIL protein levels in ANN breast cancer.

Prognostic Factor		Univariate				Multivariate*			
	Sample Size	RR	95% CI		P-value	RR	95% CI		P-value
MET protein positive vs. negative**	668	1.57	1.02	2.42	0.0411	1.35	0.87	2.11	0.1851
Positive for both MET and SNAIL vs. negative for either of MET or SNAIL	532	2.12	1.21	3.74	0.0089	1.96	1.08	3.56	0.0270

* Adjusting for menopausal status (pre/peri vs. post), tumor size (>5cm vs. <2cm, 2-5cm vs. <2cm), ESR1 receptor status (neg/equiv vs. pos/ND), Grade (grade 2/3 vs. grade1/subtype, ND vs. grade1/subtype), Lymphatic invasion, Age at diagnosis, and Adjuvant treatment (chemo, hormonal, none).
**Reduction in RR in the multivariate model can be explained by the correlation of MET protein status with hormone receptor status, grade, and tumor size.

Table 4.13. Disease-free Survival Analysis by Cox Proportional Hazards Model (with Firth Correction) for MET+/SNAIL+ protein status in basal, non-basal subgroups of ANN breast cancer (n=456)

Prognostic Factor	Subgroup	Univariate				Multivariate*			
		RR	95% CI		P-value	RR	95% CI		P-value
MET+/SNAIL+ vs. else	Basal	6.08	1.35	34.90	0.0195	6.42	1.40	37.411	0.0175
MET+/SNAIL+ vs. else	Non-Basal	1.83	0.88	3.48	0.1034	1.67	0.78	3.28	0.1759
Relative HR Difference between Basal and Non-basal		3.33	0.64	21.49	0.1521	3.85	0.74	24.98	0.1103

*Adjusting for menopausal status, tumor size, grade, lymphatic invasion, age at diagnosis, and adjuvant treatment

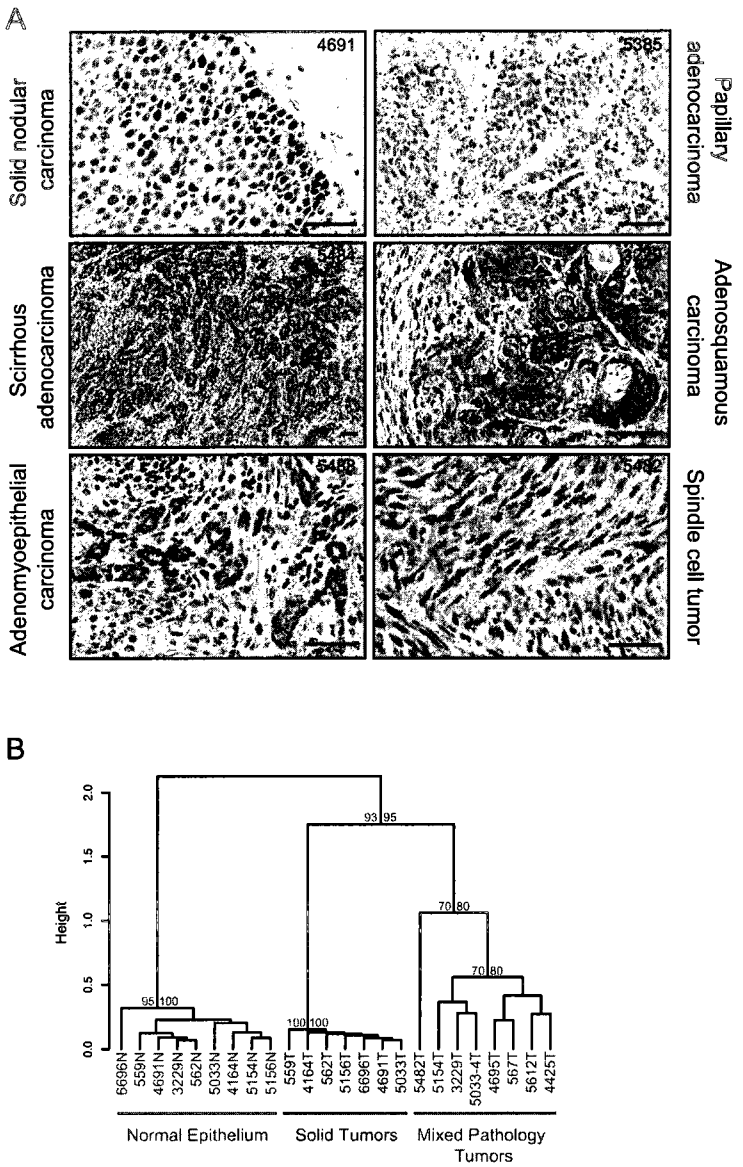


Figure 4.1: Met^{mt} mice develop mammary tumors with a variety of phenotypes that cluster into two distinct groups by microarray analysis. (A) Histopathology of tumors derived from Met^{mt} transgenic mice show a spectrum of histological phenotypes. Scale bars represent 50 μm. (B) Class discovery analysis of microdissected Met^{mt} tumor and matched-normal epithelium segregate samples into three main clusters: normal (green), solid tumors (purple) and mixed pathology tumors (red). The significance of clusters is calculated using multiscale bootstrap resampling, and is provided for both approximately unbiased (brown) and bootstrap probability (gray) p-values. The height of the dendrogram represents the distance between cluster centroids.

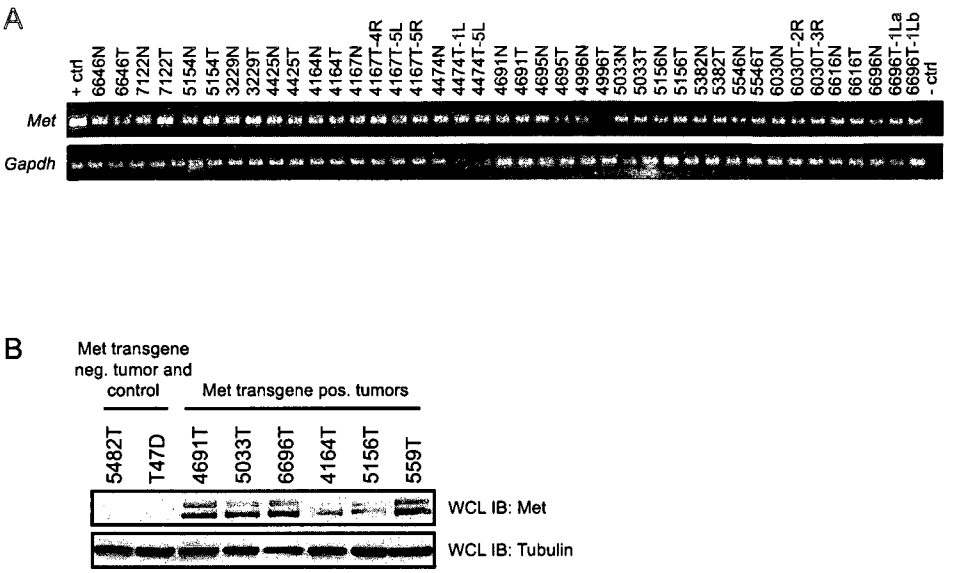


Figure 4.2: Analysis of Met^{mt} mammary tumors and matched-normal for transgene integration and protein. (A) PCR analysis of DNA isolated from tumor and matched-normal tissue for the presence of Met transgene using Gapdh as a control. Negative and positive controls for transgene are shown. (B) Met transgene protein was undetectable in spindle cell tumor (5482T) but was detectable at variable levels in other tumors (4691T, 5033T, 6696T, 4164T, 5156T, 559T). The T47D human breast cancer cell line was used as a negative control for Met transgene protein.

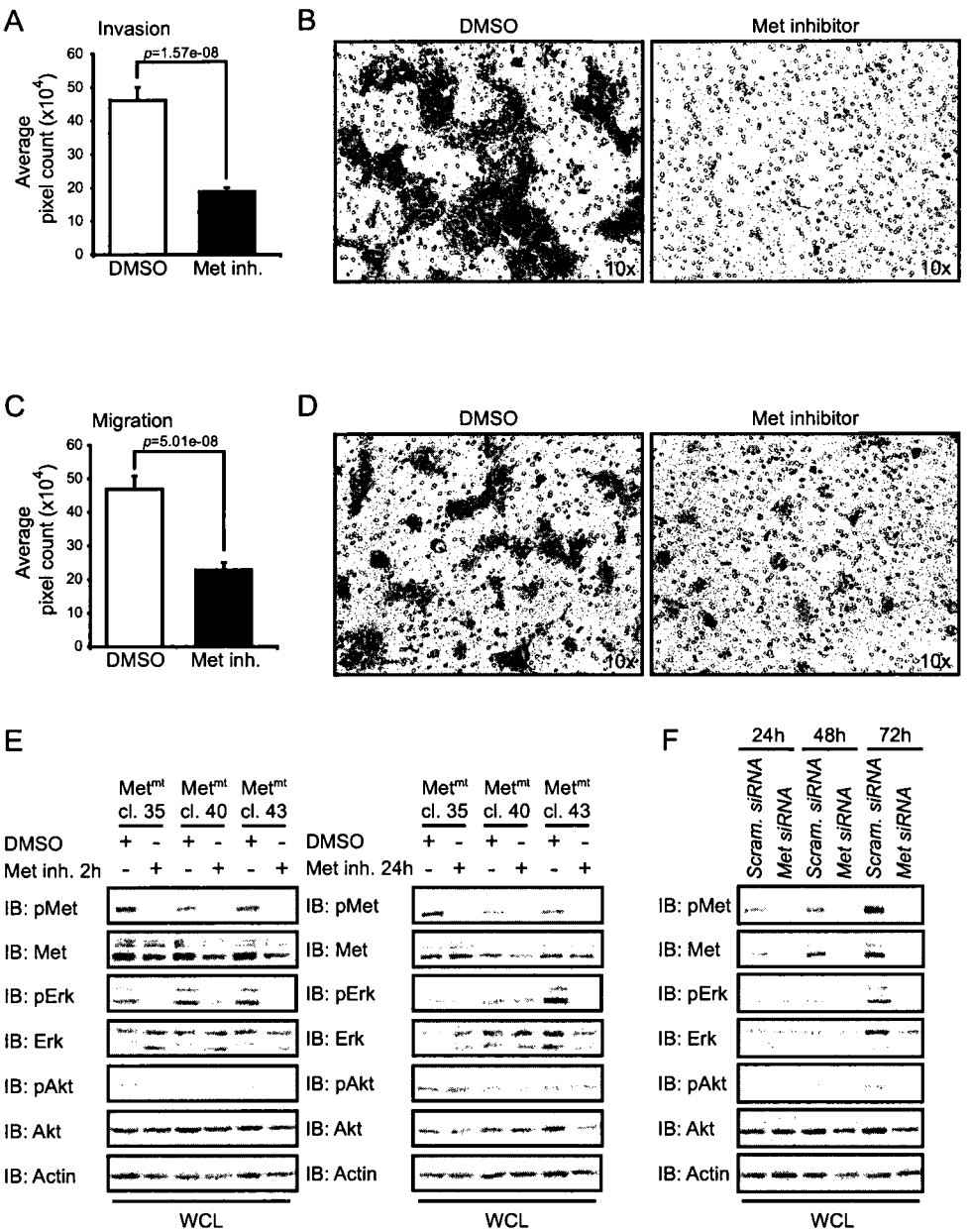


Figure 4.3: Met-dependent invasion and migration of Met^{mt} tumor-derived cell lines. (A and C) Invasion and migration data from three independent clones derived from Met^{mt} tumor cell lines in the presence (black bars) or absence (white bars) of Met inhibitor (PHA-665752, 0.1 μM) for 24 hrs. Each bar represents the mean ± SEM. (B and D) Representative images of invading and migrating cells taken from the underside of transwells in the absence and presence of Met inhibitor (Images were captured at 10x). (E) Immunoblot analysis of the same Met^{mt} tumor cell clones in *A* after 2 hr and 24 hr incubation with Met inhibitor. (F) Immunoblot analysis (25 μg) of Met^{mt} tumor cell lines treated with siRNA to Met.

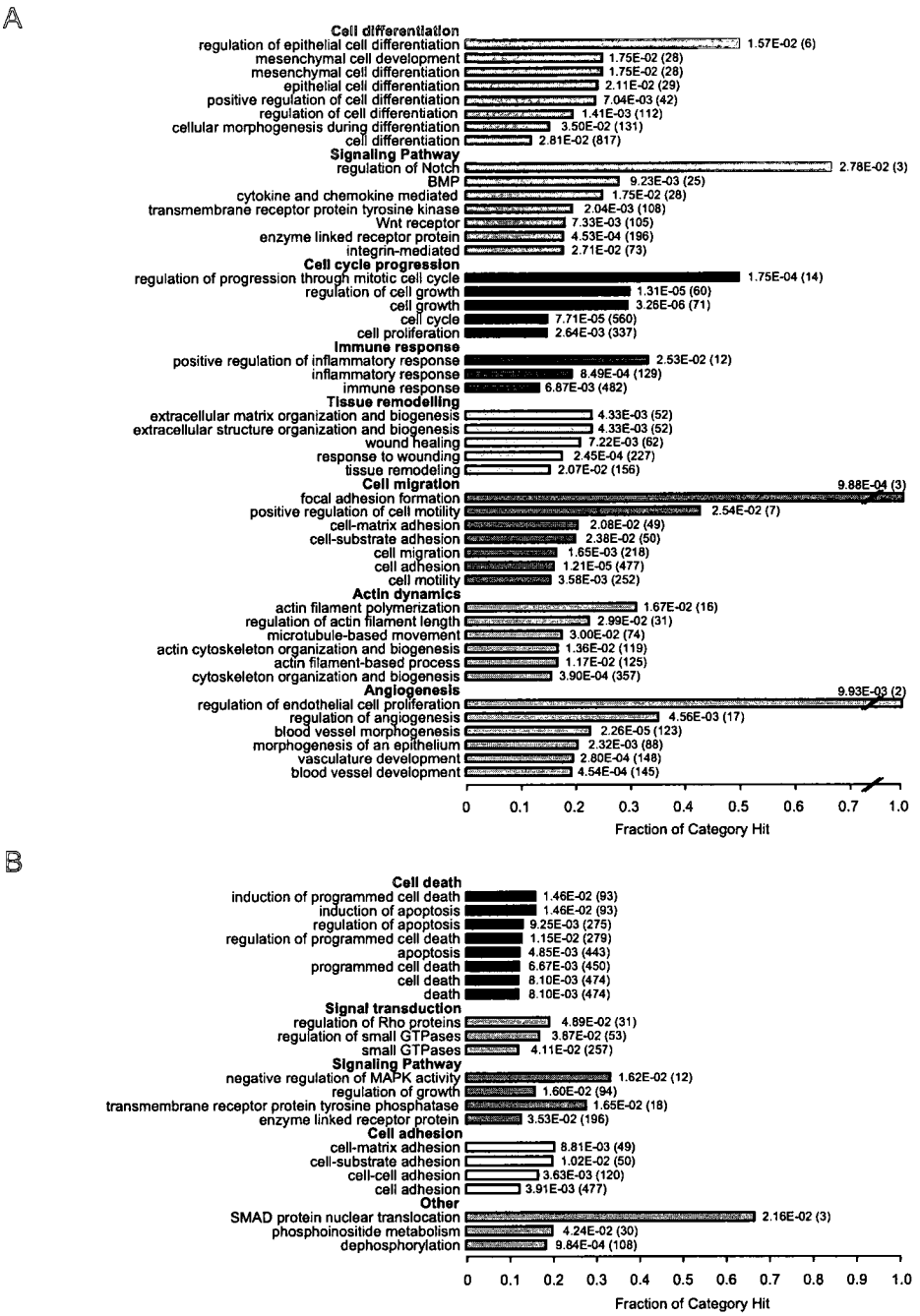


Figure 4.4: Gene ontology (GO) terms overrepresented in upregulated genes of Met^{mt} mixed pathology and solid tumors. (A) GO terms overrepresented in the upregulated genes of Met^{mt} mixed pathology vs. solid tumors. (B) GO terms overrepresented in the upregulated genes of Met^{mt} solid vs. mixed pathology tumors. GO terms from a common ancestor are grouped by color. The number of terms per category is shown in brackets.

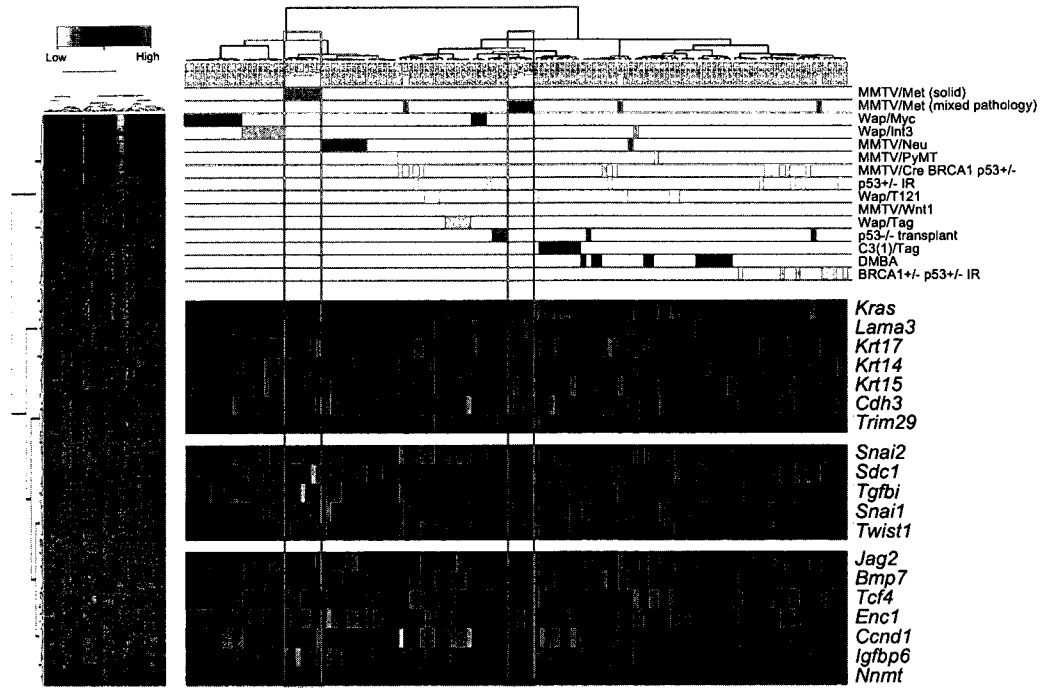


Figure 4.5: Comparison of microarray data from Met^{mt} tumors with mouse models of breast cancer. Met^{mt}-induced tumors cluster with other mouse models with similar histopathology. Mouse models are color-coded as indicated in the figure. Expression levels are color-coded from green (low expression) to red (high expression) according to the row Z-score across all tumors.

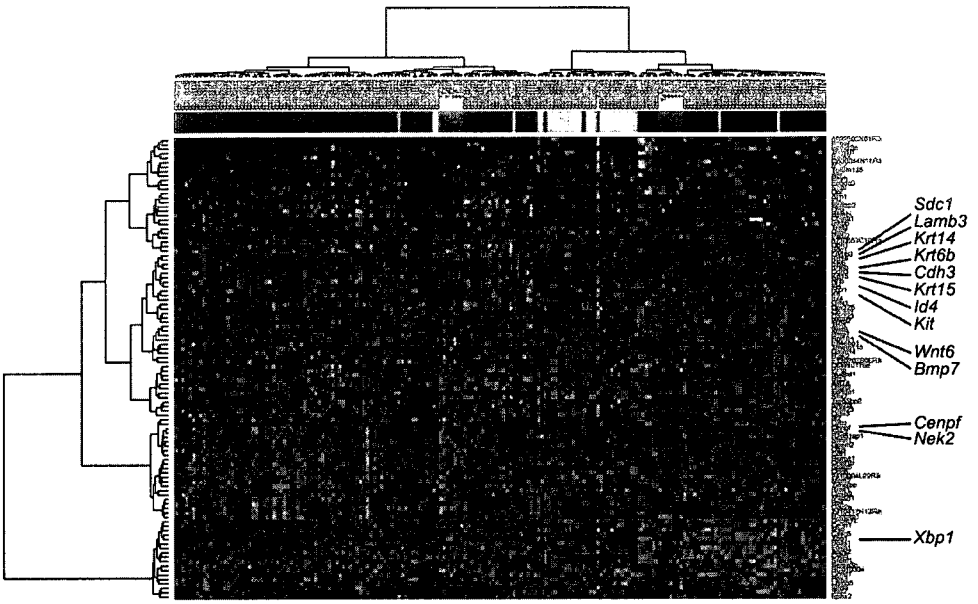


Figure 4.6: Comparison of microarray data from Met^{mt} tumors with human breast cancers. Met^{mt} solid tumors cluster with luminal human breast cancers and Met^{mt} mixed pathology tumors cluster with human basal-like tumors. Samples are color-coded as follows: Met^{mt} solid tumors (purple), Met^{mt} mixed pathology tumors (red), human luminal (blue), human ERBB2 (pink), and human basal (deep red) breast tumors.

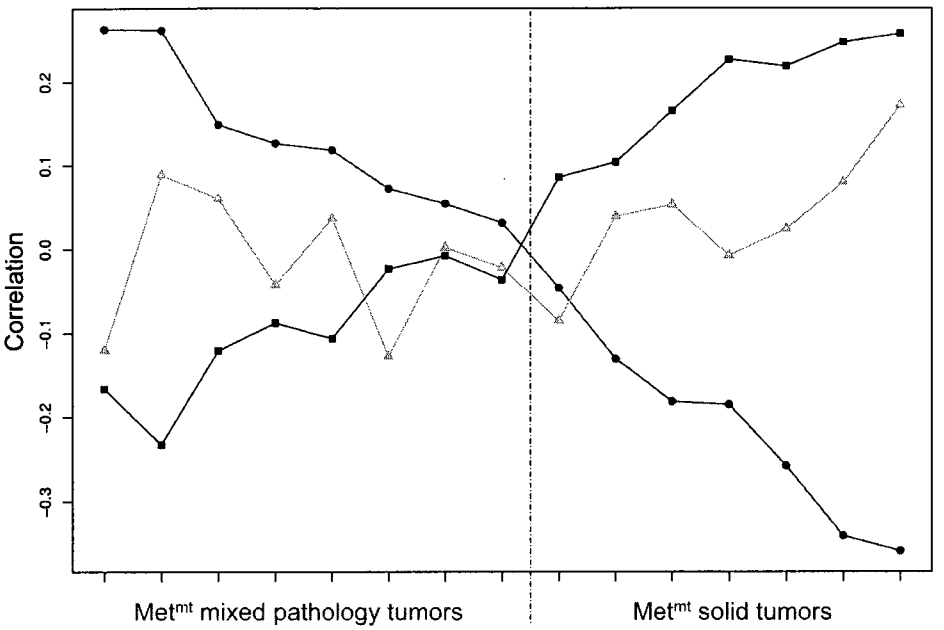


Figure 4.7: Correlation of Met^{mt} murine tumors with human subtype centroids. Correlations are colored as follows: basal (maroon circles), ERBB2 (pink triangles), and luminal (blue squares). Grey line represents division between Met^{mt} mixed pathology and solid tumors. All of the Met^{mt} mixed pathology murine tumors correlated highest with the human basal subtype, while all of the Met^{mt} solid murine tumors correlated highest with the human luminal subtype.

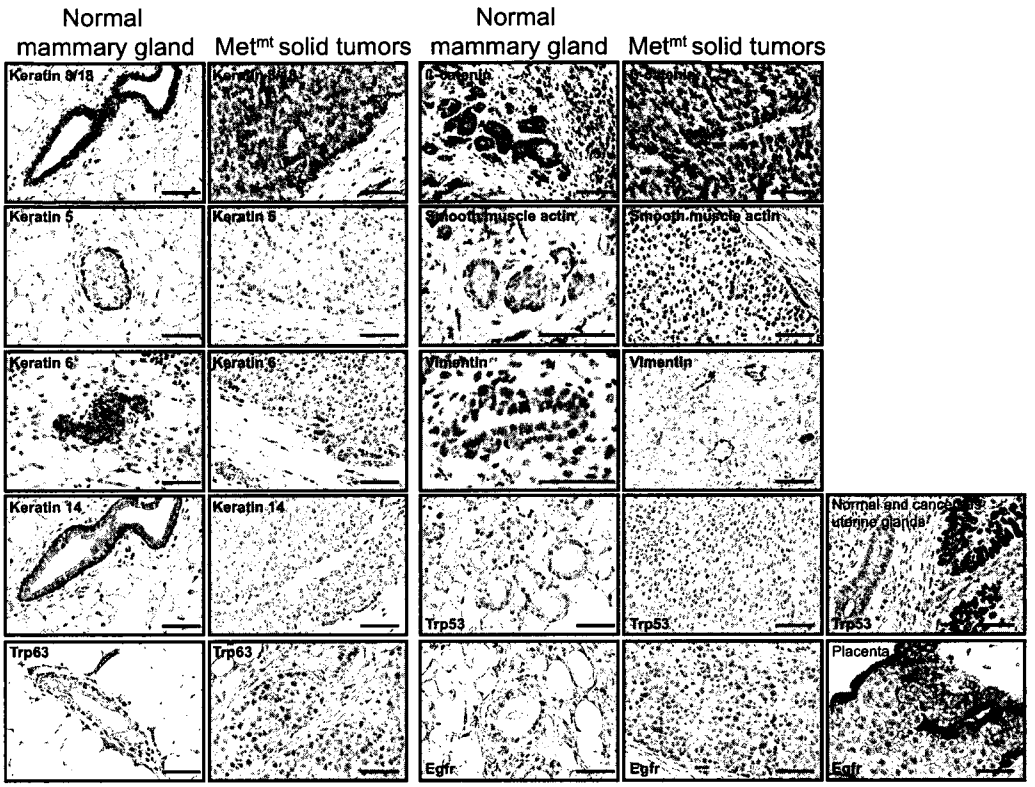


Figure 4.8: Immunohistochemical staining Met^{mt} solid tumor and normal mammary gland tissue controls. For Egfr, and Trp53 antibodies that show negative immunostaining, positive control tissues are shown, human placenta and human uterine cancer, respectively. Scale bars represent 50 μm.

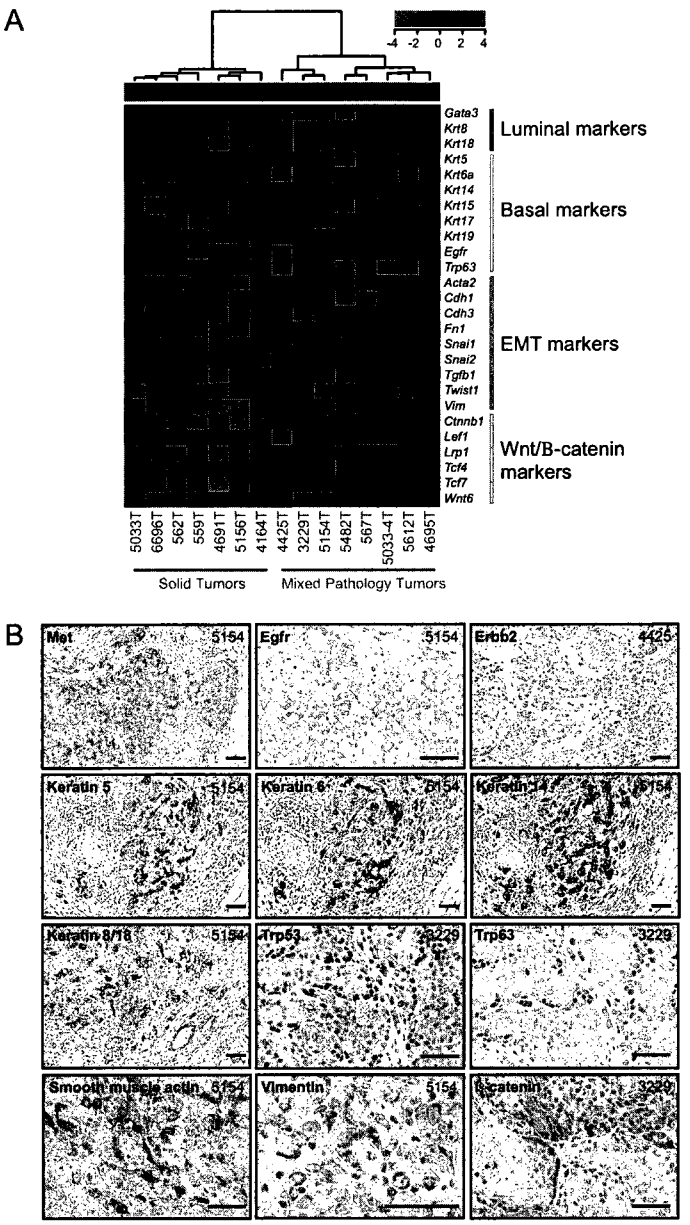
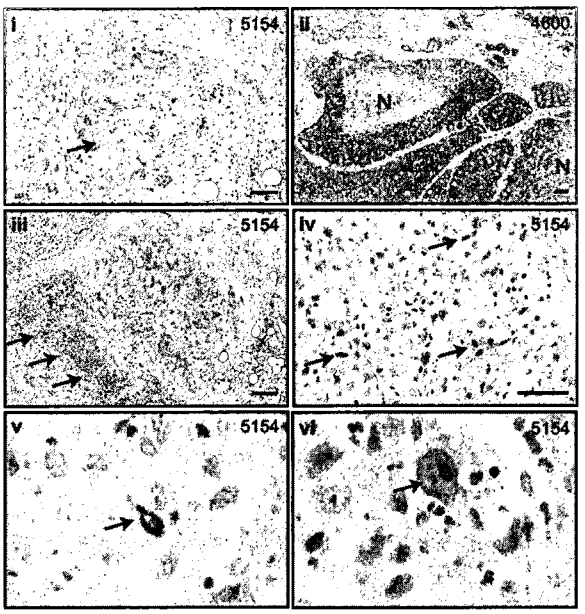


Figure 4.9: Microarray and immunostaining of basal markers in Met^{mt}-induced mammary tumors. (A) Hierarchical clustering of Met^{mt} tumors using a subset of differentially expressed genes. (B) Met^{mt} mixed pathology tumors show immunopositivity for markers of EMT and the basal phenotype. Basal markers (Krt5, 6, 14, and Trp63, Trp53, Egfr), EMT markers (smooth muscle actin, vimentin), luminal markers (Krt8/Krt18), Wnt/ß-catenin pathway marker, ß-catenin (ß-catenin) and Erbb2 are shown. Scale bars represent 50 µm.

A



B

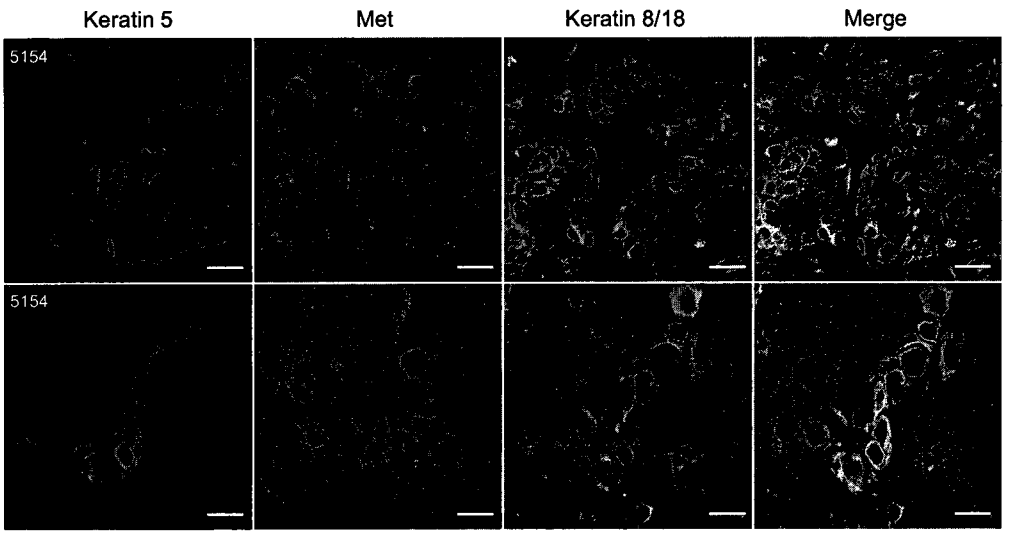


Figure 4.10: Met^{mt} mixed pathology tumors exhibit hallmarks of basal breast cancer. (A) Met^{mt} mixed pathology tumors show hallmarks representative of human basal breast cancer, including the presence of squamous metaplasia (arrow, i), geographic necrosis (N, ii), lymphocytic infiltration (arrows, iii), a high mitotic index (arrows, iv), nuclear atypia (arrow, v), and nuclear pleomorphism (arrow, vi). Scale bars represent 50 μm; images v, vi, were captured at 400x. (B) Co-expression of Met (red) with basal keratin 5 (purple) and luminal keratin 8/18 (green) in Met^{mt} mixed pathology tumors. Tumor regions of co-expression are represented by pink staining in the 'merge' images. Scale bars represent 20 μm.

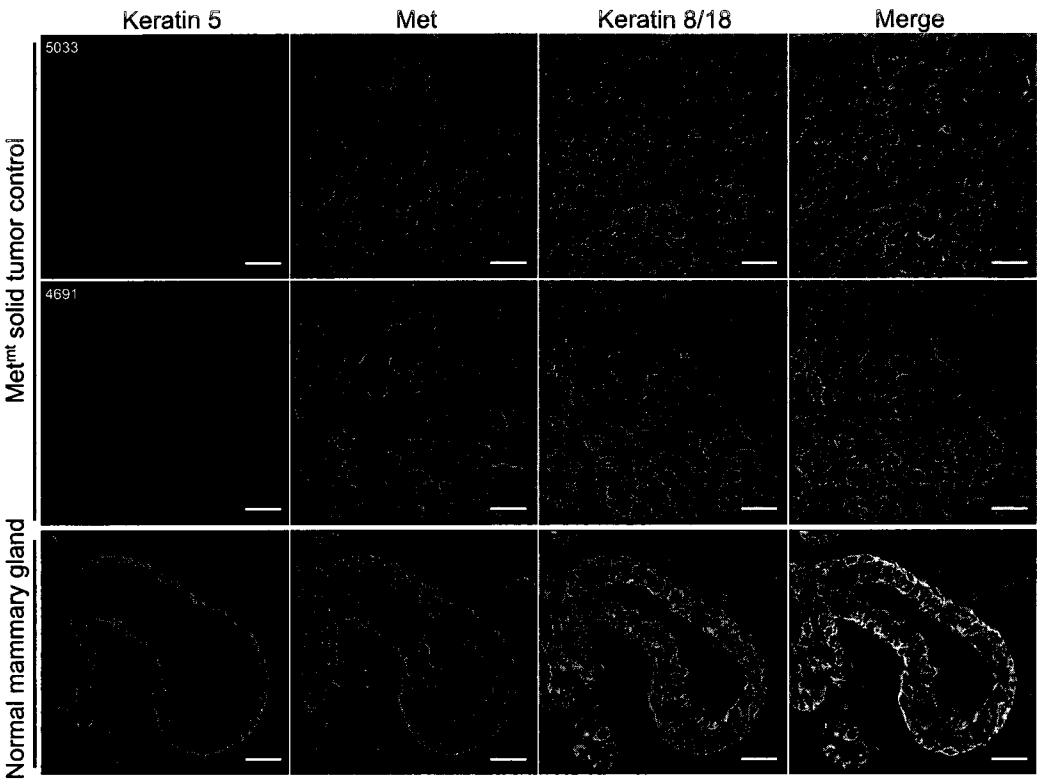


Figure 4.11: Immunofluorescence staining of Met^{mt} solid tumor and normal mammary gland tissue controls for Met transgene, basal marker keratin 5 and luminal marker keratin 8/18. Representative images of Met^{mt} solid tumors and normal murine mammary gland are shown. Scale bars represent 20 μm.

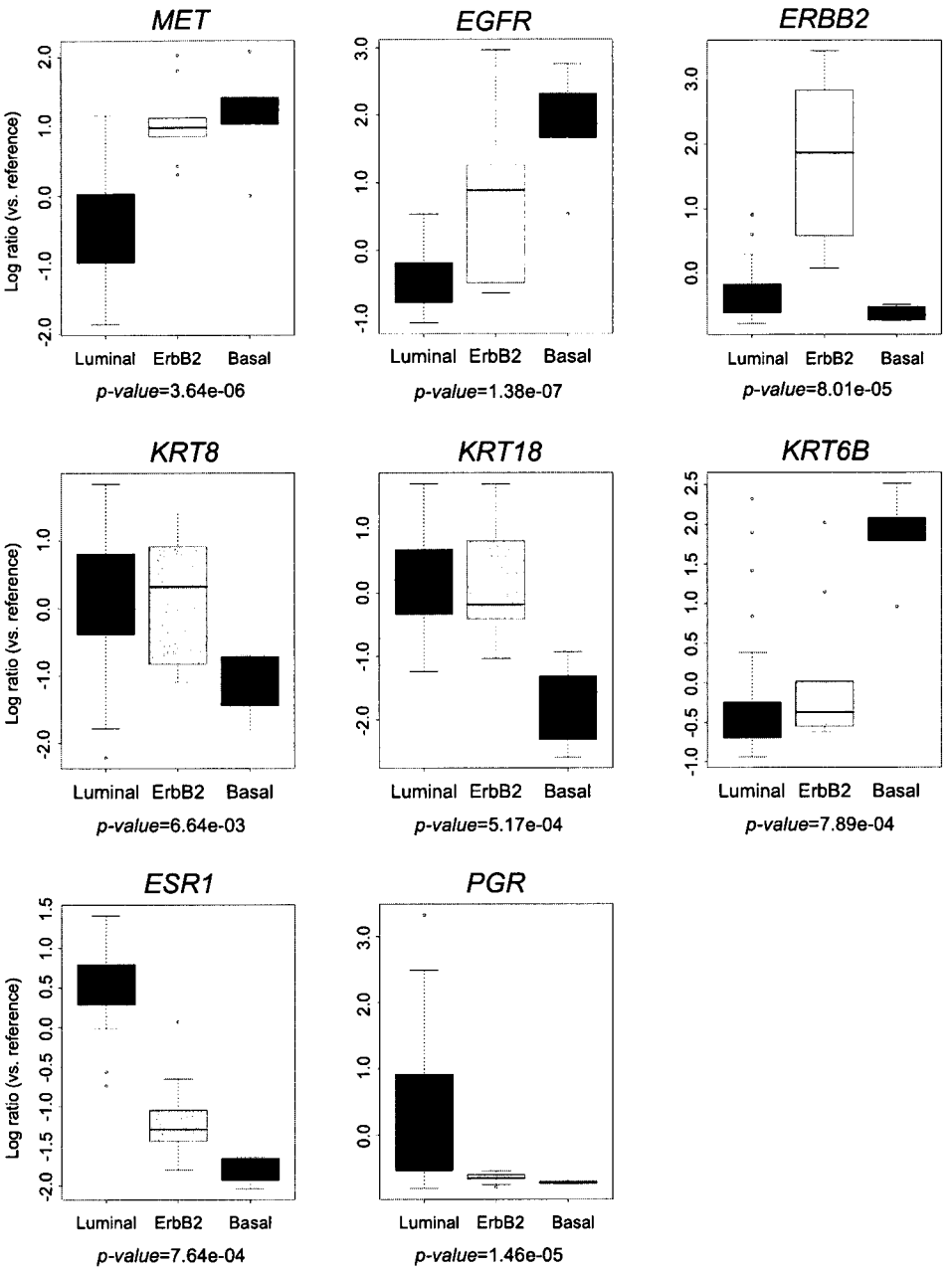


Figure 4.12: High *MET* mRNA levels in human breast cancers correlate with the basal subtype. (A) *MET* mRNA levels in human breast cancer subtypes. Box and whisker plots represent levels of mRNA for *MET*, *EGFR*, *ERBB2*, *KRT8*, *KRT18*, *KRT6B*, *ESR1*, and *PGR*. The Kruskal-Wallis rank sum test was used as the statistic.

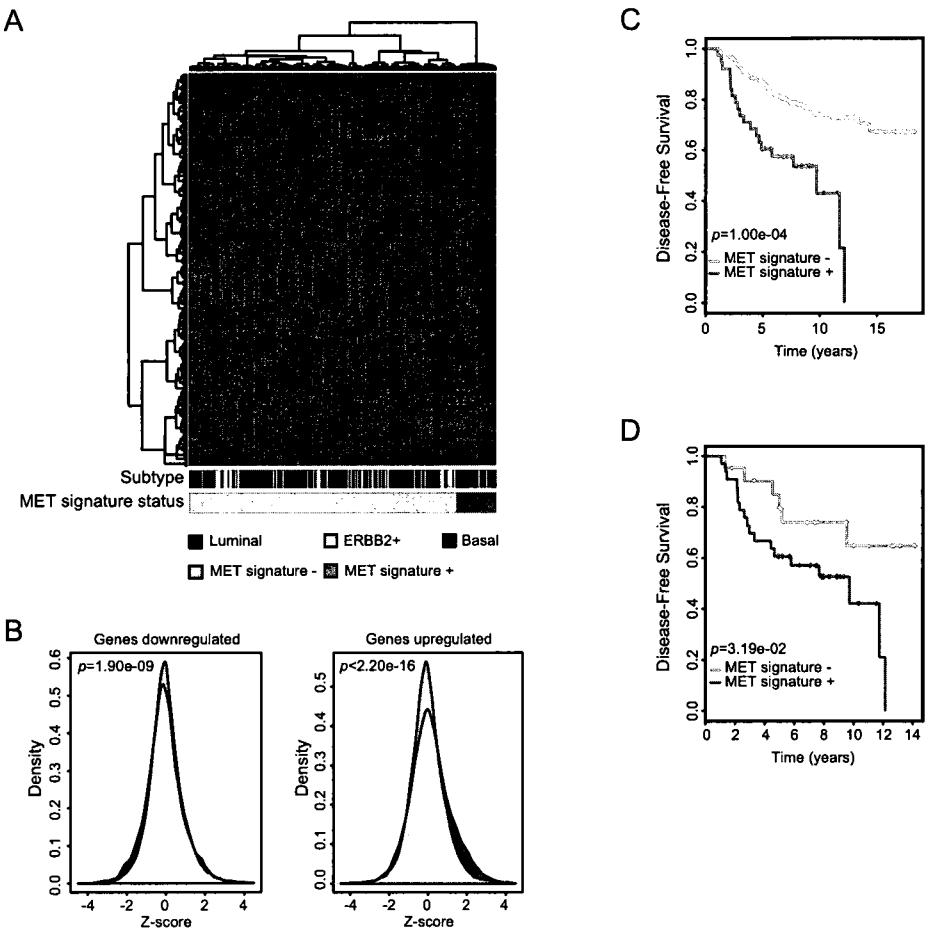
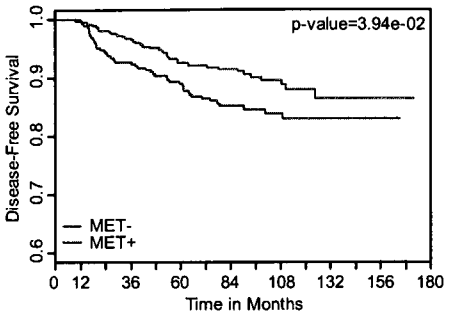


Figure 4.13. A MET gene expression signature clusters tumors primarily of the basal phenotype and correlates with poor outcome. (A) A Met signature from HGF/Met stimulated hepatocytes clusters the Netherlands Cancer Institute (NKI) microarray dataset comprising 295 breast cancer patients, into two main groups: a MET signature + cluster (turquoise), and a MET signature - cluster (orange). Breast cancers within the MET signature + cluster are primarily of the basal subtype (p-value=3.94e-26). Color legend at the bottom refers to the breast cancer subtypes, luminal (blue), ERBB2+ (pink), basal (red) and to the MET signature status, either positive (turquoise), or negative (orange). (B) Representative histograms of data observed in (A). Genes expected to be overexpressed following Met activation were significantly higher in the turquoise cluster (p-value<2.20e-16) and genes expected to be underexpressed where significantly lower in this cluster (p-value=1.90e-09). (C) Kaplan-Meier analysis based on Disease-Free Survival (DFS) of the patient clusters generated in (A) (p-value=1.00e-04). (D) Kaplan-Meier analysis based on DFS of patients from only the basal subtype, grouped by the clusters generated in (A) (p-value=3.19e-02).

A



MET status	Subtype					
	Basal		ErbB2		Luminal	
	Number	%	Number	%	Number	%
Low MET (0-3)	21	35.0	22	53.7	232	56.9
High MET (4-8)	39	65.0	19	46.3	176	43.1

p-value=6.51e-03

B

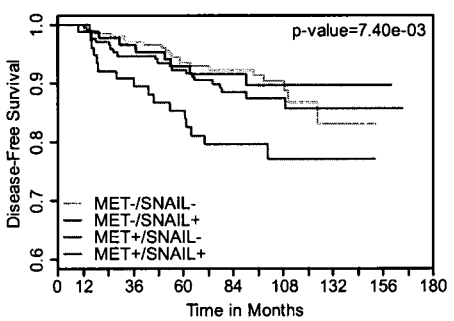


Figure 4.14: Elevated MET protein levels correlate with human basal breast cancer and poor outcome. (A) Kaplan-Meier curves show the relationship between MET protein level and Disease-Free Survival (DFS). Patients with tumors expressing high levels of MET (MET +; n=292) have a shorter DFS when compared to those with low MET (MET-; n=376) (Log-rank test p-value=3.94e-02). Representative image of high MET levels (magnification 40x). Table shows the association between MET protein status and breast cancer subtype. Basal group was characterized as ERBB2-/ESR1-/KERATIN 5+; Total sample size used was n=509; restricted to the dataset with ERBB2, ESR1, and KERATIN 5. (B) Kaplan-Meier curves showing relationship between MET and SNAIL protein level with DFS. Patients with tumors expressing high levels of both Met and Snail (MET+/SNAIL+; n=76) show reduced DFS in comparison to the other groups (MET-/SNAIL-, MET-/SNAIL+ and MET+/SNAIL-; n=456) (Log-

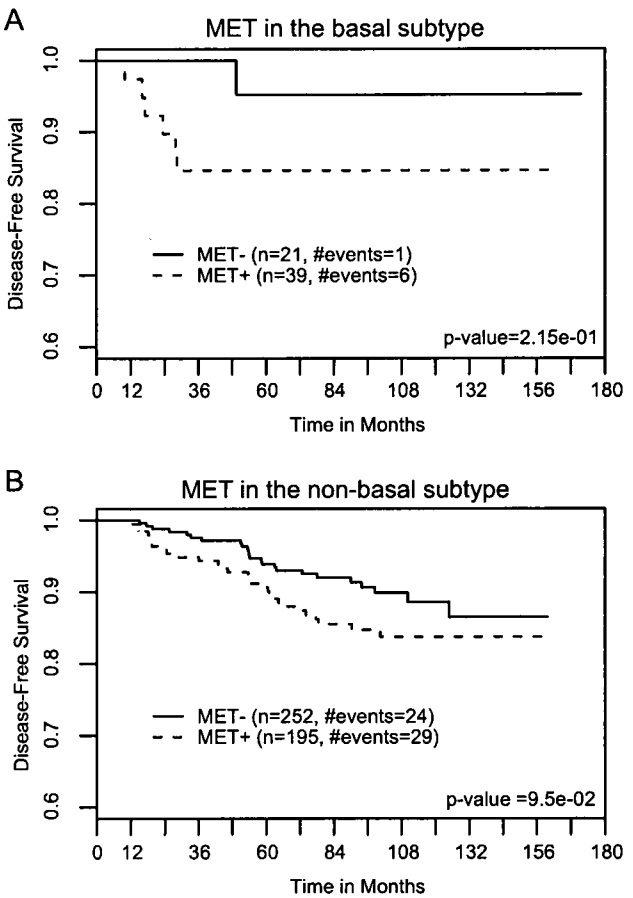


Figure 4.15: High MET protein correlates with poor outcome in ANN patients similarly for basal and non-basal tumors. Kaplan-Meier curves show the relationship between MET protein level and Disease-Free Survival (DFS) stratified by basal/non-basal tumor status. Basal group was characterized as ERBB2-/ESR1-/KERATIN 5+; Total sample size used was n=507, restricted to the dataset with ERBB2, ESR1, and KERATIN 5. There is no evidence that association with MET protein differs between basal and non-basal groups (Cox model test for interaction p-value=6.3e-01). (A) Patients with basal tumors and high levels of MET (MET +; n=39) tend to have shorter DFS when compared to those with low MET (MET-; n=21) (Log-rank test p-value=2.15e-01). (B) Patients with non-basal tumors and high levels of MET (MET +; n=195) also tend to have shorter DFS when compared to those with low MET (MET-; n=252) (Log-rank test p-value=9.5e-02).

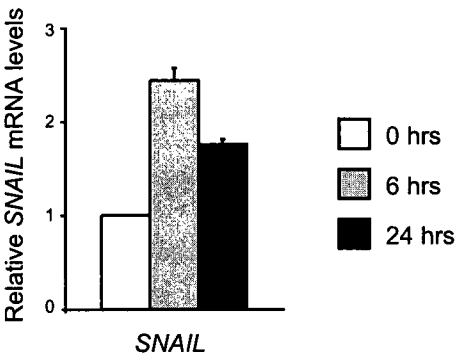


Figure 4.16: Induction of EMT transcription factor SNAIL upon HGF/MET stimulation. Expression of SNAIL mRNA in MDCK cells following 0 (white), 6 (gray) and 24 hrs (black) of HGF stimulation. Expression of SNAIL mRNA was measured by quantitative real-time PCR. The expression level of SNAIL following 0 hrs of stimulation was used as a baseline. Values are the result of three independent experiments. Each bar represents the mean ± SEM.

Chapter 5

5 The adaptor protein Nck couples the Met receptor to cell migration and actin dynamics

5.1 Introduction

We have described a mouse model in which weakly oncogenic forms of the Met receptor tyrosine kinase were expressed in the mammary epithelium under the regulation of the mouse mammary tumor virus (MMTV) promoter (Chapter 4). This murine model, MMTV/Met^{mt}, produced mammary tumors with diverse phenotypes with long latency and moderate penetrance (Chapter 4). In order to elucidate the gene expression changes involved in Met^{mt}-induced mammary tumorigenesis, we undertook a microarray approach. Gene expression profiles were generated from mammary tumors in comparison to matched-normal epithelium. In this way, significant changes in gene expression between normal and tumor tissue can be analyzed and the pathways implicated in tumorigenesis deduced.

Extracellular signals induce biological responses in part through receptor tyrosine kinases. Such signals are transduced within the cell by adaptor proteins and kinases through protein-protein interactions. NCK1 and NCK2 are SH2/SH3 adaptor proteins that have been implicated downstream from several receptor tyrosine kinases including EGFR (Li, Hu et al. 1992; Park and Rhee 1992), PDGFR (Nishimura, Li et al. 1993; Chen, She et al. 2000), and MET (Kochhar and Iyer 1996), and function to couple receptor tyrosine kinase signals to changes in the actin cytoskeleton (Buday, Wunderlich et al. 2002). NCKs are weakly oncogenic when overexpressed in NIH3T3 and 3Y1 fibroblasts in culture, and NCK-overexpressing cells produce tumors upon subcutaneous injection in

mice (Chou, Fajardo et al. 1992; Li, Hu et al. 1992). To date, a role for NCK in mammary development or tumorigenesis has not been addressed.

Although a role for the MET receptor in poor outcome basal breast cancers has been identified in Chapter 4, the mechanisms responsible for this are poorly understood. In an attempt to elucidate these mechanisms, we have examined the role of Nck proteins, Nck1 and Nck2, which were found to be elevated in Met^{mt}-induced mammary tumors when compared to matched-normal tissue.

5.2 Results

5.2.1 Nck1 and Nck2 mRNA and proteins are elevated in Met^{mt}-induced mouse mammary tumors when compared to matched-normal tissue

We used whole genome mouse expression microarrays to identify genes whose expression is altered in Met-induced mammary tumors. Microarray analysis identified multiple differentially expressed genes, which were ranked in order of decreasing fold change. Among the top genes differentially expressed from analysis of multiple microarray probe features were the transcripts coding for SH2/SH3 domain containing adaptor protein *Nck1* and *Nck2* (5.23 and 6.61-fold elevation respectively in Met^{mt} tumor when compared to matched normal, Figure 5.1).

Nck mRNA levels were first validated using qRT-PCR (Figure 5.2). Total RNA was isolated from a subset of tumor and matched-normal tissues used in microarray analysis. Using primers specific for *Nck1*, an 8-fold elevation in Met^{mt} tumor compared to matched normal was observed (Figure 5.2). Although multiple primer sets and conditions were tested for *Nck2*, we were not able to produce amplification of an *Nck2* product from cDNA by qRT-PCR. Nck1 and Nck2 proteins levels were subsequently validated with western blot analysis using Nck1 and Nck2-specific antibodies (Figure 5.3). The

specificity of antibodies against Nck1 and Nck2 was demonstrated by western blot analysis of transiently expressed tagged versions of NCK1 and NCK2 in human embryonic kidney (HEK293T) cells (Figure 5.3A). Since whole tissue extracts were used for western blot analysis and normal mammary gland tissue contains a high fat to epithelium ratio, epithelial cell marker keratin 8 was used as an additional loading control for epithelial cell content (Figure 5.3B). Using densitometric analysis, optical densities for tumor and matched-normal tissue from Nck1 and Nck2 blots were compared to both keratin 8 and α-tubulin loading controls. These data demonstrated a significant increase in Nck1 and Nck2 protein levels when compared with keratin 8 and α-tubulin as controls (Figure 5.3C). Thus, Nck1 and Nck2 protein levels and *Nck1* mRNA levels were elevated in Met^{mt}-induced mouse mammary tumors compared to matched-normal tissue.

5.2.2 Nck1 and Nck2 associate with a phosphorylated Met receptor in mammary tumors

To elucidate the signaling mechanisms that may implicate Nck1 and Nck2 proteins in mammary tumorigenesis downstream from the Met receptor, phosphotyrosine profiles from mammary tumor lysates containing Met receptor expression were compared to those that did not. Antibody specificity was established by transiently transfecting GFP-tagged versions of NCK1 and NCK2 in HEK293T cells and immunoblotting with a GFP antibody (Figure 5.4A), demonstrating that Nck1 and Nck2 antibodies were indeed specific. Immunoprecipitation of Nck1 and Nck2 from Met^{mt}-expressing tumor lysates and non-Met^{mt} expressing controls, including a tumor that does not express Met^{mt} transgene (5482T) and the human breast epithelial cell line T47D which contains undetectable levels of Met, was performed (Figure 5.4B). To establish if Nck proteins or associated proteins are tyrosine phosphorylated we utilized an antibody that recognizes

tyrosine phosphorylated proteins (Figure 5.4B, immunoblot PY100). Notably, in both Nck1 and Nck2 immunoprecipitates, tyrosine phosphorylated proteins of ~145 kDa and 170 kDa, and ~100 kDa were observed (Figure 5.4B). No other differences in tyrosine phosphoproteins were found between Met^{mt}-expressing and non-Met^{mt}-expressing lysates at other molecular weights. The doublet at ~145 kDa and 170 kDa is reminiscent of the processed and unprocessed forms of the Met receptor (Chapter 4). Subsequent immunoblotting of membranes using an antibody that specifically recognized the Met transgene confirmed the presence of an activated Met receptor in tumor lysates (Figure 5.4B, immunoblot Met). Thus, these data support that Nck1 and 2 are engaged in a complex with the phosphorylated Met receptor tyrosine kinase in Met^{mt}-induced mouse mammary tumors.

5.2.3 Nck knockdown results in decreased migration and invasion in Met^{mt} mammary tumor epithelial cells

Nck1 and Nck2 proteins have been shown to play a role in migration and invasion of fibroblasts (Rivera, Antoku et al. 2006; Ruusala, Pawson et al. 2008). However, the role of Nck1 and Nck2 proteins in epithelial cell migration, especially mammary epithelia, has not been extensively studied and fibroblasts have been the main system used to study these biological phenotypes. To test the functional significance of an association between the Met receptor and Nck proteins in tumor biology, tumor-derived epithelial cells generated from Met^{mt} mammary tumors were used as a model system (Chapter 2). To determine whether Met^{mt} tumor-derived epithelial cell lines were representative of their tumor counterparts, clones from cell lines were generated and analyzed for protein levels of Met, phosphorylation status of Met, Nck1 and Nck2 (Figure 5.5). Indeed, Met, Nck1 and Nck2 proteins were detected in all clones tested, and in

addition, the Met receptor was activated as determined by phosphorylation of tyrosine residues 1234/1235 in the catalytic domain (Figure 5.5). We have previously shown the Met^{mt}-tumor derived cell lines are invasive and migratory in boyden chamber assays and this was determined to be a Met-dependent event, since blocking Met receptor signaling with a Met inhibitor (PHA-665752, 0.1 μM) or siRNA targeted to Met significantly reduced the invasive and migratory capability of these cells (Chapter 4, Figure 4.3). To address whether Nck1 and Nck2 knockdown would result in decreased migration and invasion, we used an siRNA approach to knockdown protein levels of Nck1 and Nck2 using two different siRNA constructs targeting Nck1, or Nck2 (Figure 5.6A). Knockdown of Nck1 and Nck2 in Met^{mt} tumor-derived epithelial cells (6030T) was consistently observed to be >80% from 48 to 96 hours post-transfection within the time frame during which all experiments were performed (Figure 5.6A). Interestingly, a significant decrease in migration and invasion were observed between scrambled and Nck1/2 siRNA-treated Met^{mt} transformed cells in the absence of HGF (Figure 5.6B-E). Significantly, a more dramatic decrease in cellular migration (1.7-fold, Fig. 5.6B, black bars) and invasion (1.7-fold, Fig. 5.6D, black bars) was observed in cells treated with Nck1/2 siRNA compared to scrambled siRNA controls in response to HGF (Fig. 5.6B-E). These data show that Nck1 and 2 proteins play a role in Met-dependent cellular migration and invasion in mammary tumor epithelial cells.

5.2.4 Nck1 and Nck2 protein knockdown disturbs actin rearrangement and dynamics in Met^{mt} mammary tumor cell lines

Nck proteins have been shown to act as adaptors that connect extracellular cues from activated receptor tyrosine kinases to actin cytoskeletal changes (Rivera, Antoku et al. 2006; Ruusala, Pawson et al. 2008). Actin dynamics and rearrangement is a necessary

prerequisite for cellular migration and invasion (Bladt, Aippersbach et al. 2003; Jaffe and Hall 2005). Since a decrease in Met^{mt}-dependent cellular migration and invasion was observed following Nck1/2 knockdown, we wished to determine whether this might be explained by a defect in actin organization and dynamics within Met^{mt}-tumor derived cells. To assess the role of Nck proteins in mediating actin cytoskeletal rearrangements following siRNA knockdown of both Nck1 and Nck2 proteins, Met^{mt} tumor epithelial cells (6030T) were stained with phalloidin and imaged by confocal microscopy (Figure 5.7A). In scrambled siRNA-treated cells, prominent actin-rich protrusions were observed at the cellular edges (Figure 5.7A, arrowheads). In Nck1/2 siRNA-treated cells, we observed markedly fewer actin-rich protrusions (Figure 5.7A). Additionally, gross morphological differences were found, where Met^{mt} tumor epithelial cells treated with Nck1/2 siRNA were flatter and more spread than scrambled siRNA-treated controls (Figure 5.7B). Scrambled siRNA-treated control cells were determined to have a height of 7.6 μm in z-stack confocal images, whereas the height of Nck1/2 siRNA-treated cells was 2.63 μm (Figure 5.8). These data demonstrate that Nck proteins play a key role in mediating actin cytoskeletal dynamics, which affect membrane protrusions, and cell shape in the context of an activated Met receptor.

5.2.5 NCK1 and NCK2 associate with the activated MET receptor and become tyrosine phosphorylated in a MET-dependent manner

Since we observed biological differences in Met^{mt} tumor epithelial cells with loss of Nck1/2 proteins, to further elucidate the mechanism responsible, we tested the association between Nck and the Met receptor. Following transient transfection of HEK293T cells with increasing levels of wild type MET receptor along with consistent amounts of GFP-NCK1, the ability of these proteins to co-immunoprecipitate was

investigated (Figure 5.9). We have previously shown that overexpression of the MET receptor leads to its activation, thus providing an alternative to HGF stimulation (Rodrigues, Naujokas et al. 1991) (see phosphotyrosine blot, PY100, in Figure 5.9). We indeed observed tyrosine phosphorylation, and thus activation of the MET receptor when transiently expressed in HEK293T in the absence of HGF stimulation (Figure 5.9 IP MET immunoblot PY100). With increasing amounts of wild type MET receptor, we detected increasing tyrosine phosphorylation of MET and increasing co-immunoprecipitation with GFP-NCK1, demonstrating that NCK1 can form a complex with the activated wild type MET receptor (Figure 5.9). Co-immunoprecipitation of MET and NCK1 was demonstrated using both MET and GFP antibodies to immunoprecipitate (Figure 5.9, see IP MET immunoblot GFP, and IP GFP immunoblot MET).

NCK proteins have been shown to become tyrosine phosphorylated downstream from receptor tyrosine kinases including EGFR and PDGFR (Li, Hu et al. 1992). To determine whether NCK was tyrosine phosphorylated downstream from the MET receptor we examined NCK1 following transient co-transfection with MET. Tyrosine phosphorylation of NCK1 was enhanced with increasing amounts of wild type Met receptor, demonstrating a role for the MET receptor in NCK1 tyrosine phosphorylation (Figure 5.9, IP GFP immunoblot PY100). These data demonstrate that NCK is a target of phosphorylation in a pathway downstream from the MET receptor tyrosine kinase. Consistent with the idea that Nck proteins may be implicated in Met-induced tumorigenesis, lysates of MMTV/Met^{mt} tumors were analyzed using mass spectrometry for presence of phosphotyrosine proteins, and both Nck1 and Nck2 were present in this screen (Forest White, personal communication).

Since the SH2 domains of NCK1 and NCK2 adaptor proteins have been previously shown to interact with phosphotyrosine residues of receptor tyrosine kinases including the PDGF, and VEGF receptors (Li, Hu et al. 1992; Ito, Wernstedt et al. 1998), we hypothesized that NCK1 and NCK2 proteins would interact with the MET receptor in a similar manner. To test this, GFP-NCK1 or GFP-NCK2 were transiently co-transfected with wild type MET or a kinase inactive MET (K1110A) receptor mutant (Figure 5.10). Both GFP-NCK1 (Figure 5.10A) and GFP-NCK2 (Figure 5.10B) co-immunoprecipitated with the activated wild type MET receptor, but this association was abolished with the kinase inactive MET K1110A receptor mutant. Notably, both NCK1 and NCK2 were tyrosine phosphorylated only in the presence of the activated wild type MET receptor (Figure 5.10, IP GFP, IB PY100), and not in the presence of the kinase inactive, K1110A, Met receptor mutant, further validating that tyrosine phosphorylation of NCK1 and NCK2 is Met-dependent.

5.2.6 NCK1 and NCK2 interact with phosphoprotein GAB1 in a MET-dependent manner

To further elucidate the mechanism by which Met receptor activation leads to the induction of tyrosine phosphorylation of Nck1 and Nck2, we examined Nck1 and Nck2 associated phosphoproteins in Met^{mt} mammary tumors. Nck1 and Nck2 co-immunoprecipitated with a phosphotyrosine-containing protein with an approximate molecular weight of 100 kDa (Figure 5.4B). The major phosphoprotein downstream from an active Met receptor is known to be the scaffold/adaptor protein growth factor receptor bound protein 2-associated protein 1, GAB1 (Fixman, Naujokas et al. 1995; Nguyen, Holgado-Madruga et al. 1997). To determine the identity of the ~100 kDa protein, the membrane in Figure 5.4B was stripped and immunoblotted with an anti-GAB1 antibody.

Although a faint band overlapping with the ~100 kDa species was detected with an anti-GAB1 blotting solution, the data was inconclusive (data not shown). However, an online resource, Scansite, which searches for motifs within proteins predicted to be binding sites for known SH2 and SH3 domains, was used to determine whether an association of GAB1 with NCK was predicted (Obenauer, Cantley et al. 2003). Indeed, Scansite identified putative NCK SH2 domain binding sites on GAB1. These were localized to tyrosines 242 and 307, which we have previously determined are CRK, and both CRK and ABL binding sites, respectively. To establish whether GAB1 could associate with NCK proteins, we first tested the association by transient transfection in HEK293T cells using wild type MET, GFP-NCK1, and HA-GAB1 constructs (Figure 5.11). Following co-transfection, NCK1 and GAB1 co-immunoprecipitated but only in the presence of activated wild type Met receptor (Figure 5.11) indicating that tyrosine phosphorylation by MET may be required. Co-immunoprecipitation of GAB1 with NCK1 was evidenced whether HA antibodies or GFP antibodies were used to immunoprecipitate the complexes (Figure 5.11, IP HA immunoblot GFP or IP GFP immunoblot HA). Tyrosine phosphorylation of GAB1 was induced in the presence of the activated MET (Figure 5.11, IP HA immunoblot PY100). Similarly, NCK1 tyrosine phosphorylation was observed in the presence of activated MET but not downstream from HA-GAB1 alone (Figure 5.11, IP GFP immunoblot PY100). The association of NCK1 with GAB1 identifies a novel Gab1 signaling complex downstream from the MET receptor. Altogether, these data show that an activated MET receptor is required for the recruitment of NCK1 and NCK2 to the GAB1 scaffold protein and for tyrosine phosphorylation of NCK1 and NCK2. In addition, we have demonstrated that NCK signals from protein complexes that include the active MET receptor and tyrosine phosphorylated GAB1.

5.2.7 Elevated *NCK1* and *NCK2* mRNA levels correlate with human basal breast cancers

We have previously shown that elevated *MET* receptor mRNA and MET protein levels are found in human basal breast cancers and correlates with poor prognosis (Chapter 4). Within the last decade, gene expression profiling has classified breast cancers into categories that are correlated with distinct clinical outcome: luminal, ERBB2+, and basal (see Chapters 1 and 4 for background). The luminal subgroup correlates with good outcome and patients generally respond to anti-estrogen therapy, whereas basal and ERBB2+ patient subgroups correlate with poor outcome. However, since targeted therapy is not available for the basal subgroup, this remains the most difficult patient subgroup to manage clinically.

From our observed increase in Nck1 and Nck2 in Met^{mt} tumors, it was conceivable that Nck adaptor proteins might also be elevated in human breast cancer. Levels of *NCK1* and *NCK2* mRNA were examined in three publicly available datasets of human breast cancer that contained suitable probes for *NCK1* and *NCK2* (Figure 5.12) (van de Vijver, He et al. 2002; Chin, DeVries et al. 2006; Finak, Sadekova et al. 2006). In our McGill dataset of breast cancer gene expression data (Chapter 4) (Finak, Sadekova et al. 2006), luminal, ERBB2+ and basal subtypes were defined by immunohistochemistry. Subtypes were defined by expression profiling for the NKI and Chin datasets. The NKI dataset included 295 cases with axillary lymph node positive and negative early stage (I and II) breast cancers occurring in a young cohort of women (less than an average of 53 years old) (van de Vijver, He et al. 2002). The Chin dataset included 130 predominantly early stage (I and II) breast cancers that were axillary lymph node positive and negative, a portion of which had received tamoxifen and adjuvant therapy prior to surgery (Chin,

DeVries et al. 2006). We were intrigued to find that mRNA levels of *NCK1* and *NCK2* were significantly elevated in the basal subgroup across all three datasets, in comparison to the luminal and ERBB2+ groups (Fig. 5.12). However, immunohistochemical staining will be necessary to further validate these data and this is presently underway.

5.3 Discussion

For multiple human cancers, receptor tyrosine kinases have been shown to initiate signaling pathways that drive tumorigenesis. It is therefore crucial to understand the signaling pathways induced in tumors relative to normal tissue. We used a microarray approach to gain an understanding of the gene expression changes occurring downstream from the Met receptor tyrosine kinase, specifically in mammary epithelium (see Chapters 3 and 4). We identified SH2/SH3 adaptor proteins Nck1 and Nck2 as elevated in Met^{mt}-induced mammary tumors compared to matched-normal epithelium (Figure 5.1) and further investigated their functional and biological roles in mammary tumorigenesis.

Surprisingly, the role of Nck proteins in signaling, invasion, migration, cell adhesion and actin cytoskeletal dynamics have largely been studied in rodent fibroblast cell lines such as mouse embryonic fibroblasts (MEFs), NIH3T3 and rat 3Y1 (Chou, Fajardo et al. 1992; Li, Hu et al. 1992; Bladt, Aippersbach et al. 2003; Rivera, Antoku et al. 2006; Ruusala, Pawson et al. 2008). Whether the same pathways exist for epithelial cells is unknown. Here, we used a tumor epithelial cell model system, derived from Met^{mt}-induced mammary tumors to study the role of Nck proteins downstream from the activated Met receptor (Figure 5.5). These cell lines were established using protocols described in Chapter 2 and were representative of the tumors from which they were

generated in that they expressed activated Met receptor and several other proteins identified in their originating tumors (Figure 5.5 and data not shown).

Earlier reports stated that Nck is weakly oncogenic in NIH3T3 and 3Y1 fibroblasts, and that Nck proteins can cooperate with oncoproteins such as *ras*, *src* and *abl* to enhance tumorigenic capability (Chou, Fajardo et al. 1992; Li, Hu et al. 1992; Braverman and Quilliam 1999). In the context of human basal breast cancer, the MET receptor may cooperate with downstream proteins such as NCK to promote tumorigenesis.

We used this model system to investigate the biological and functional role of Nck proteins in epithelial cell lines with activated Met receptor. In order for cell motility to occur, cells must organize a branching actin meshwork at the leading edge forming a protrusion or lamellipodium, establish new adhesions at the front of the cell, shift forward through cell body contraction, and detach cell adhesions at the rear of the cell (Raftopoulou and Hall 2004). These steps involve reorganization of the actin cytoskeleton through a dynamic process of actin assembly and disassembly in a spatially- and temporally-coordinated manner in order to generate directional movement. Consistent with data from Nck-null MEFs (Rivera, Antoku et al. 2006; Ruusala, Pawson et al. 2008), we observed a marked decrease in actin-rich protrusions and peripheral ruffles in Met^{mt} tumor-derived epithelial cell lines treated with Nck1/2 siRNA compared to scramble siRNA-treated controls (Figure 5.7). Thus, within the context of an activated Met receptor in tumor epithelial cells, both Nck1 and Nck2 are required for the production of actin-rich cellular protrusions (Figure 5.7). In agreement with a defect in actin cytoskeletal rearrangement, we observed a decreased capability of Met^{mt} tumor-derived epithelial cell lines to migrate and invade in boyden chamber assays (Figure 5.6). We

have previously established that cell migration and invasion is a Met-dependent event in these cells through the use of Met inhibitor and siRNA targeted to the Met receptor (Chapter 4). Nck1 and Nck2 knockdown also resulted in decreased cell migration and invasion in the absence of HGF, albeit to a lesser degree than that observed with Met knockdown (Figure 5.6, compared to Figure 4.3). Nck proteins can thus be implicated in a signaling pathway downstream from the active Met receptor in cell migration and invasion. Given that Nck1 and Nck2 knockdown alone is not able to decrease invasion and migration to the same extent as treatment with Met inhibitor or Met siRNA, this suggests that other Met-dependent pathways are involved in addition to Nck proteins in mediating cellular migration and invasion. Although multiple siRNA were tested for knockdown in Met^{mt} primary cell culture, only one siRNA construct for Nck1 and for Nck2 produced efficient knockdown. Thus, future goals include validating these data with additional Nck1 and Nck2 siRNA constructs to rule out off-target effects. However, none of the siRNA tested that failed to knockdown Nck1/2 had any effect on cell morphology or migration further supporting a causal role for Nck in these processes downstream of Met. The data presented are consistent with studies using rodent cell lines and Nck-null mouse models, in that the roles of Nck1 and Nck2 appear to be redundant and compensatory in Met^{mt}-tumor derived epithelial cells. We observed no differences in behavior of Nck1 and Nck2 in biochemical assays, and required knockdown of both Nck1 and Nck2 to observe differences in biological assays.

Nck1 and Nck2 have been shown to associate with receptor tyrosine kinases VEGFR, and PDGFR through their SH2 domains, and Nck proteins are phosphorylated downstream from receptor activation (Li, Hu et al. 1992; Park and Rhee 1992; Nishimura, Li et al. 1993; Ito, Wernstedt et al. 1998). Consistent with these data, we found that

NCK1 and NCK2 were tyrosine phosphorylated upon activation of the MET receptor (Figures 5.9 and 5.10). The association was determined to be phosphotyrosine-dependent, since a kinase inactive MET receptor mutant (K1110A) not only abolished phosphorylation of NCK1 and NCK2, but also abolished association between NCK proteins and MET (Figure 5.10). Significantly, we discovered a novel interaction between NCK1 and GAB1, the major scaffold protein downstream from the Met receptor (Figure 5.11).

To further define the role of Nck in Met^{mt}-induced mammary tumors, we analyzed the role of other phosphoproteins co-immunoprecipitating with Nck proteins. Activation of the MET receptor tyrosine kinase is known to promote recruitment and tyrosine phosphorylation of GAB1, which provides docking sites for multiple SH2 domain containing signaling molecules, including the p85 subunit of PI(3)kinase (PI3K), PLC-γ, the adaptor protein CRK and the tyrosine phosphatase SHP2 (Maroun, Holgado-Madruga et al. 1999; Gual, Giordano et al. 2000; Maroun, Naujokas et al. 2000; Sakkab, Lewitzky et al. 2000; Schaeper, Gehring et al. 2000; Lamorte, Royal et al. 2002; Lock, Maroun et al. 2002). This is the first demonstration that GAB1 interacts with the NCK adaptor protein and has significant implications for MET-dependent biologies in epithelial cells (Figure 5.11). Significantly, the association of NCK proteins to GAB1 may be unique to the MET/HGF signaling axis, as NCK was previously not found to associate with GAB1 downstream from the EGF receptor (Holgado-Madruga, Emlet et al. 1996). Although the functional significance of a GAB1-NCK interaction remains to be tested, GAB1 is required downstream from the MET receptor for cell migration, invasion (Maroun, Holgado-Madruga et al. 1999), lamellipodia (Frigault, Naujokas et al. 2008), and dorsal

ruffle formation (Abella, Frigault et al. 2008) and NCK may mediate a pathway linking MET to these biological outcomes (Figure 5.13).

Upon activation of the MET receptor, epithelial colonies break down tight and adherens junctions, acquire a fibroblast-like shape and "scatter", and thus become motile (Stoker, Gherardi et al. 1987; Montesano, Matsumoto et al. 1991). This morphogenic event downstream from the MET/HGF signaling axis is dependent on the GAB1 adaptor protein (Maroun, Holgado-Madruga et al. 1999). Some but not all of these functions have been attributed to recruitment of the CRK adaptor protein to GAB1. Overexpression of CRK enhances lamellipodia formation and cell spreading in response to HGF and consequently enhances cell motility and promotes epithelial dispersal (Lamorte, Royal et al. 2002; Lamorte, Rodrigues et al. 2003). NCK proteins may represent a novel pathway connecting MET/HGF signaling to complexes containing PAK1, N-WASP, WIP, which may subsequently interact with Rac and Cdc42 to promote actin polymerization by Apr2/3. However, further studies are needed to determine if this is the case.

Given that Nck1 was originally identified and cloned from a human melanoma cell library (Lehmann, Riethmuller et al. 1990), it is surprising that reports implicating NCK1 or NCK2 in any form of human disease, such as cancer, have not been published. Here, we report for the first time that *NCK1* and *NCK2* mRNA levels were elevated in poor outcome human basal breast cancers in three independent datasets of human breast cancer (Figure 5.12). The basal subtype of human breast cancer is a molecular subtype that has recently received much attention as tumors of this type lack targeted therapy and exhibit poor outcome. We have previously demonstrated that *MET* receptor mRNA and MET protein levels are also elevated in the basal subtype of breast cancer and that high MET protein correlates with poor prognosis (Chapter 4). Thus, NCK1 and NCK2 as well

as Met levels are co-elevated in the same form of poor outcome human breast cancers and it is interesting to speculate from the data presented here that MET may cooperate with NCK1/2 to promote tumorigenesis (Figure 5.13).

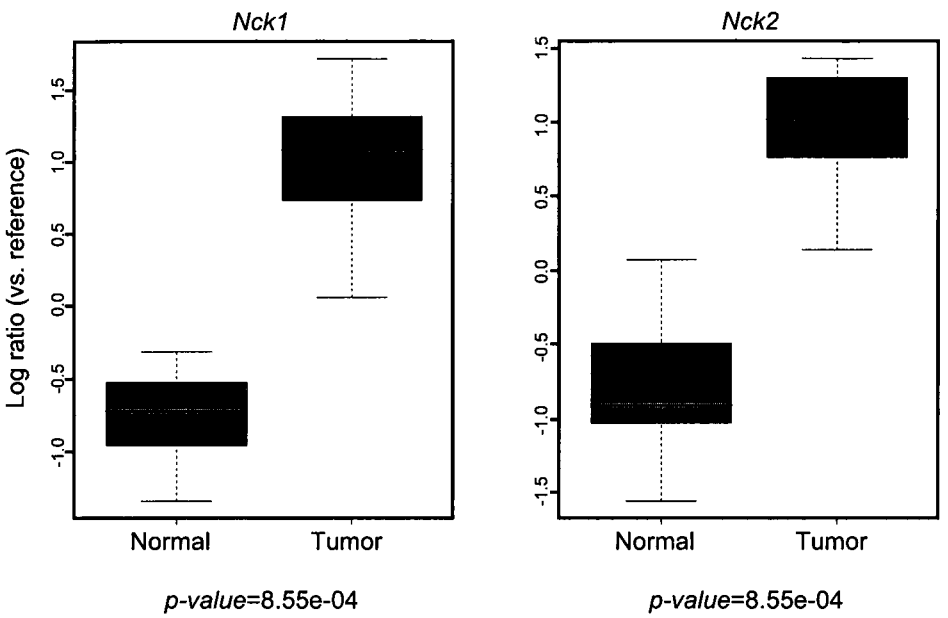


Figure 5.1: Microarray analysis identified *Nck1* and *Nck2* to be upregulated in Met^{mt}-induced mouse mammary tumor compared to normal tissue. *Nck1* and *Nck2* mRNA levels from microarray data in Met^{mt} mammary tumors and matched-normal tissue. Values for 7 tumors (purple), and 9 normal tissues (green) were averaged to produce the box and whisker plots. The Kruskal-Wallis rank sum test was used as the statistic.

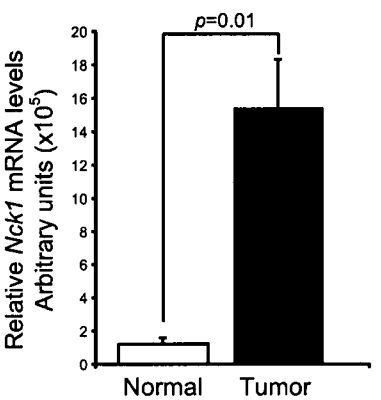


Figure 5.2: Validation of *Nck1* mRNA levels using quantitative RT-PCR. Relative levels of *Nck1* mRNA were measured by RT-PCR in Met^{mt} mammary tumors compared to matched-normal tissue (559, 562, 5033, 5156). *Nck1* levels were compared to the housekeeping gene, Gapdh. Each bar represents the mean ± SEM.

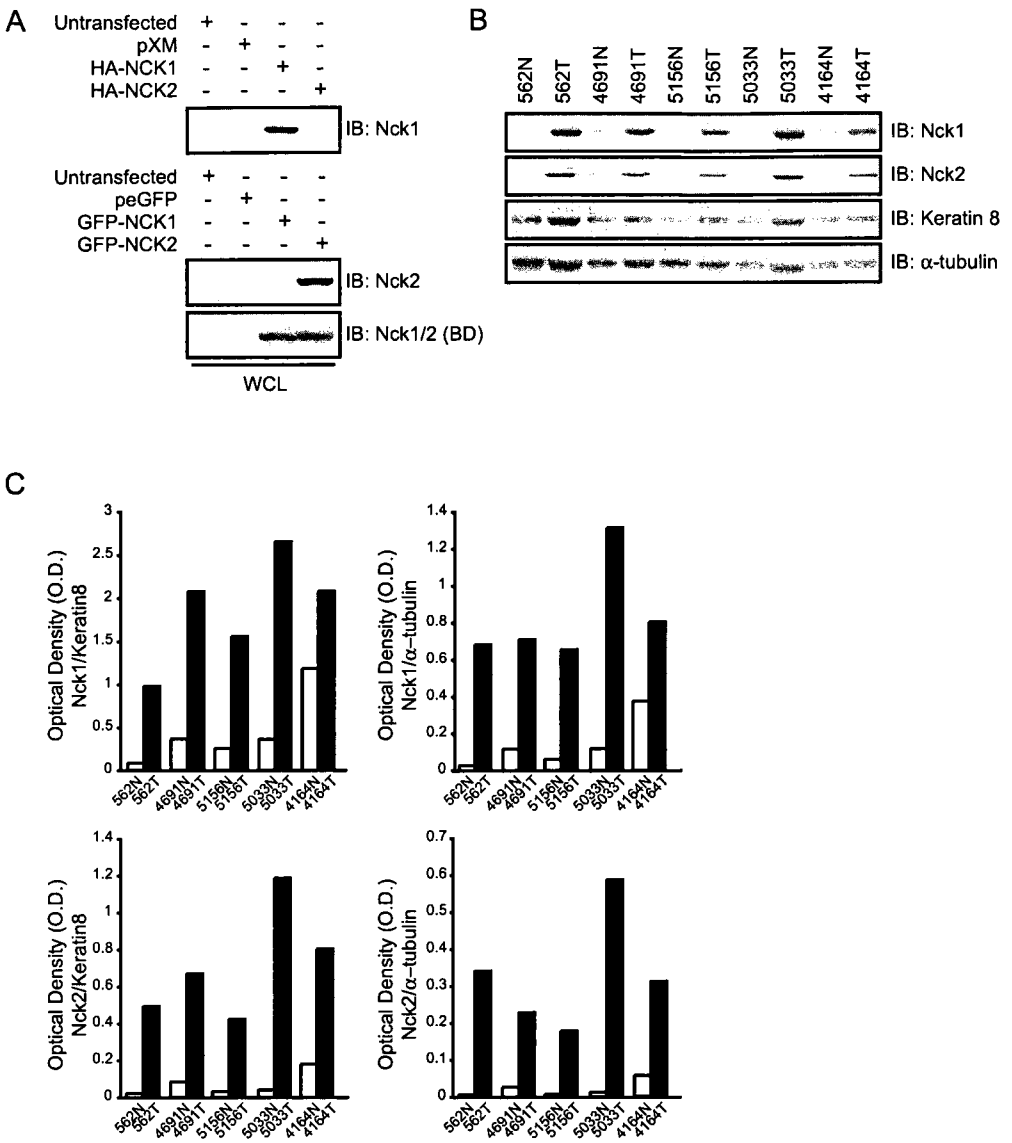


Figure 5.3: Validation of Nck1 and Nck2 protein levels using western blot analysis. (A) Nck1 (#2383) and Nck2 (#3313) antibodies were tested for their specificity in western blot analysis using HEK293T lysates transiently expressing GFP-tagged or HA-tagged versions of NCK1 and NCK2, vector alone or untransfected. (B) Nck1 and Nck2 endogenous protein levels in Met^{mt} tumor and matched-normal lysates (40 μg) were examined by immunoblotting with antibodies tested in *A* for their specificity. Keratin 8 was used as a marker for epithelial content and α-tubulin was used as a loading control. (C) Densitometric analysis of Nck1 and Nck2 levels compared to keratin 8 or tubulin controls from *B* as measured by ImageJ.

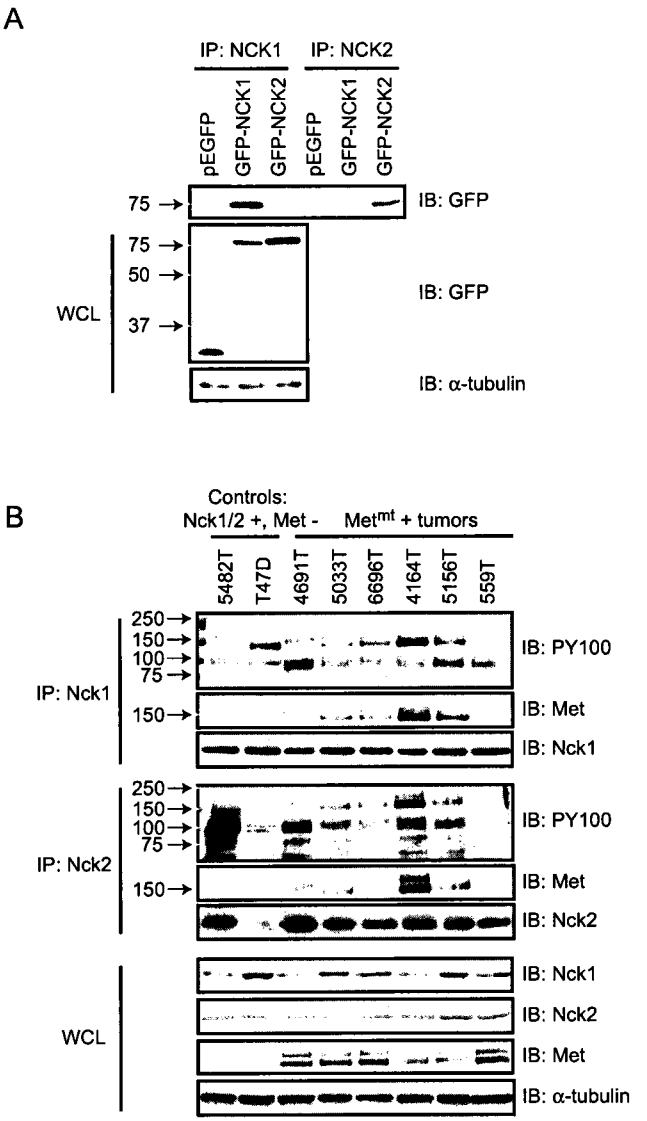


Figure 5.4: Nck1 and Nck2 associate with the Met receptor in Met^{mt} mammary tumors. (A) The specificities of Nck1 (#2383) and Nck2 (#3313) antibodies were tested for their ability to immunoprecipitate NCK1 and NCK2, respectively. HEK293T cells were transfected with GFP-tagged versions of NCK1 and NCK2 constructs and either Nck1 (#2383) or Nck2 (#3313) antibodies were used for immunoprecipitation. Proteins collected in immune complexes were visualized by western blot analysis using GFP antibodies. (B) Met^{mt}-mammary tumors (1 mg) were immunoprecipitated with Nck1 and Nck2-specific antibodies tested in *A* and immunoblotted with indicated antibodies (phosphotyrosine PY100, Met [Met19S], Nck1 [#2383], Nck2 [#3313], and α-tubulin).

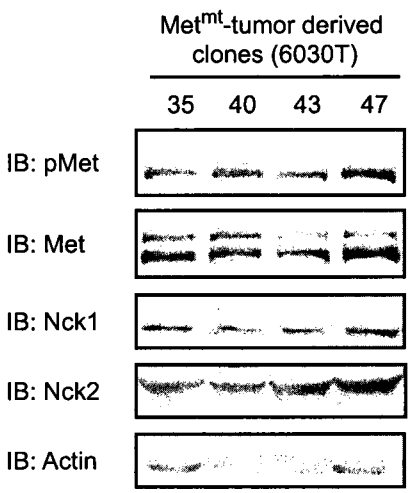


Figure 5.5: Tumor epithelial clones derived from Met^{mt} mammary tumor 6030T. Clones were generated as described in Chapter 2 and subsequently characterized by western blot analysis using antibodies recognizing phospho-tyrosines 1234/1235 within the kinase domain of the Met receptor (pMet), Met (Met148), Nck1 (#2383), Nck2 (#3313), and actin.

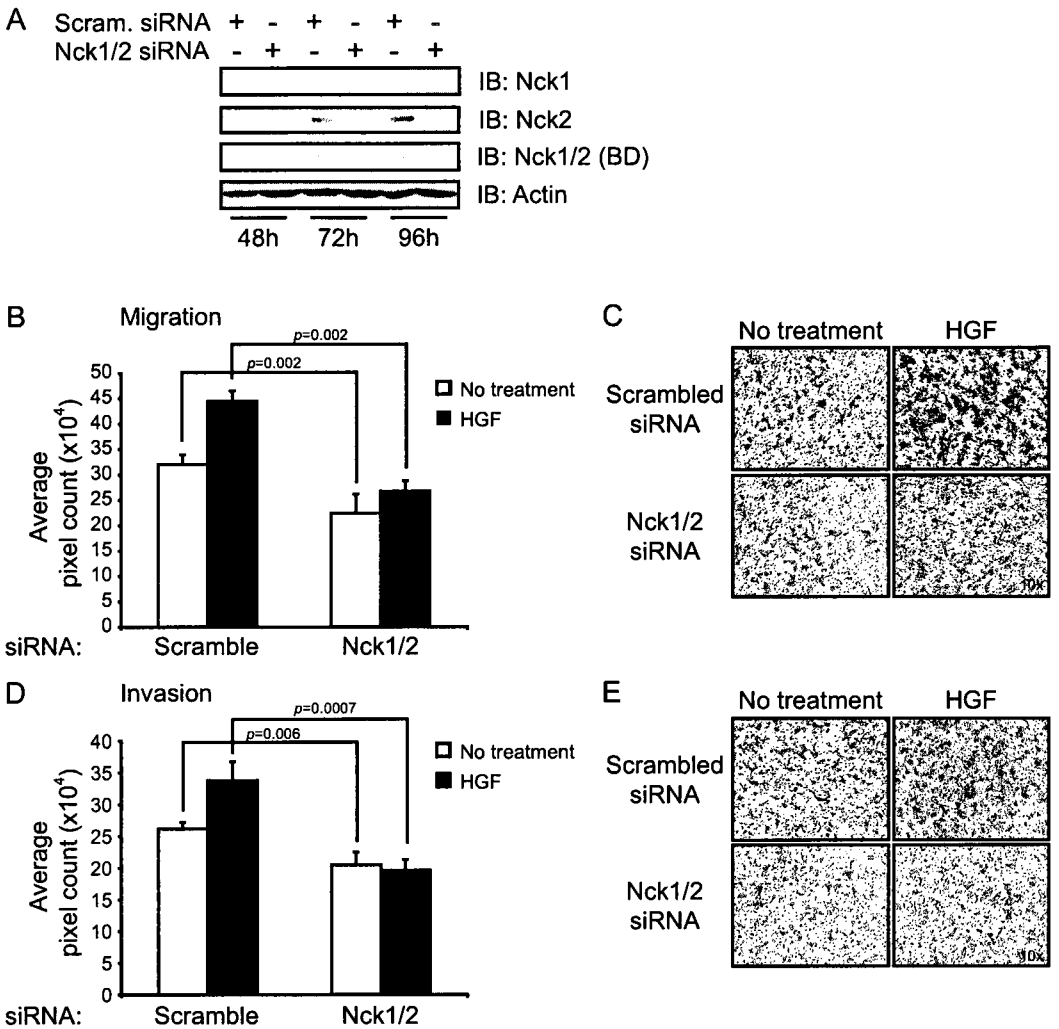


Figure 5.6: Nck1 and Nck2 knockdown decreases invasion and migration of Met^{mt} tumor-derived cells. Met^{mt} tumor-derived cells (6030T) with knockdown of Nck1 and Nck2 were analysed for their invasive and migratory capacities with or without HGF (34 ng/ml). (A) Whole cell lysates (25 μg) were collected at 48, 72, and 96 hours after the second round of siRNA transfection, seperated by SDS-PAGE and immunblotted with antibodies recognizing Nck1 or Nck2, antibodies recognizing both Nck1 and Nck2 (Nck BD) and actin. Met^{mt} tumor-derived cells (5 x 10^5) were plated in 6-well dishes and transfected with either 50 nM of Nck1 and 50 nM of Nck2 siRNA or 100 nM scrambled siRNA. After 48 hours, a second round of siRNA transfection was performed. Cells were trypsinized after 48 hours, seeded (5 x 10^4) onto boyden chambers (B and C) or matrigel-coated boyden chambers (D and E) and assays were stopped 24 hours later. Each bar represents the mean ± SEM. Using a Zeiss Axioskop microscope, bottom layers of transwells were imaged in five seperate fields for each condition using a 10x objective in phase contrast (C and E). Image analysis was carried out using Scion Image.

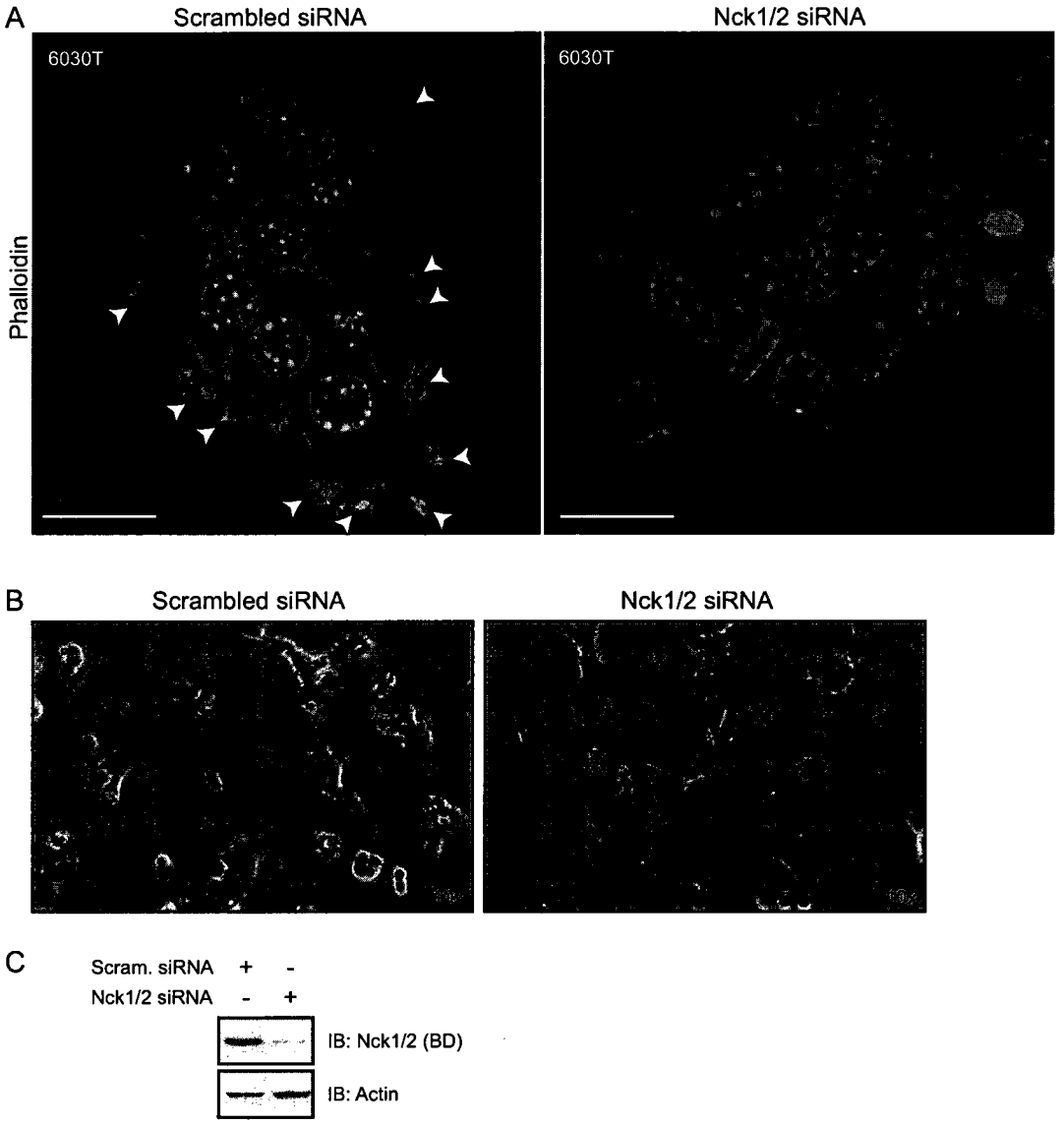


Figure 5.7: Nck1 and Nck2 knockdown decreases actin dynamics in Met^{mt} tumor-derived cells. (A) Phalloidin staining of Met^{mt} tumor-derived cells (6030T). Met^{mt} tumor-derived cells (5 x 10^5) were plated in 6-well dishes and were either transfected with 50 nM of Nck1 and 50 nM of Nck2 siRNA or 100 nM scrambled siRNA. After 48 hours, a second round of siRNA transfection was performed and cells were plated on coverslips. Cells were fixed, stained and imaged using a Zeiss 510 Meta laser scanning confocal microscope 48 hours after the second transfection. Scale bars represent 20 μm. (B) Met^{mt} tumor-derived cells (6030T) from the same experiment as in *A* were imaged after fixation using a Zeiss Axiovision 135 microscope with a 10x objective. (C) Whole cell lysates (25 μg) from the experiment shown in *A* and *B* were immunblotted with an antibody (Nck BD) recognizing Nck1 and Nck2, and actin. Experiments were performed a minimum of three times.

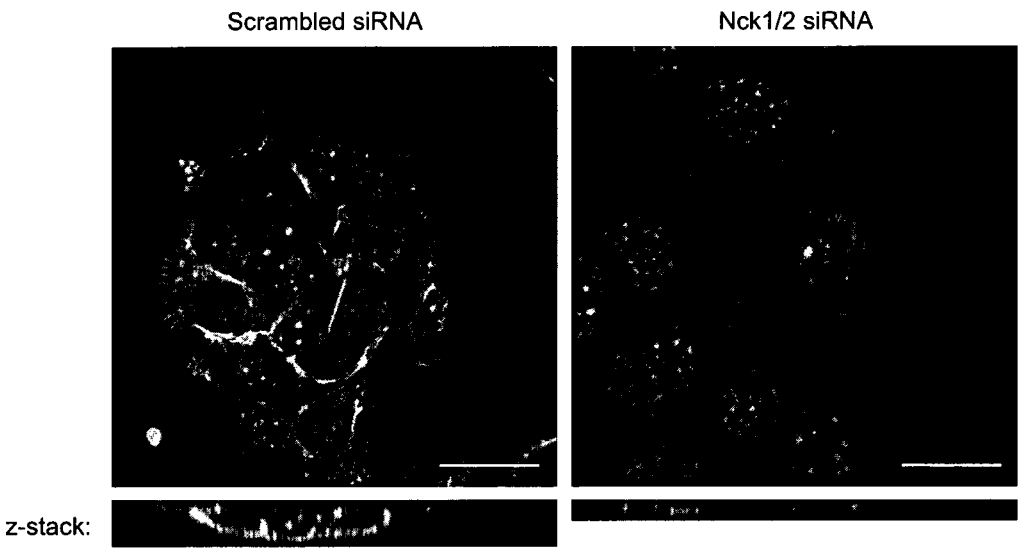


Figure 5.8: Nck1 and Nck2 knockdown affects cell shape in Met^{mt} tumor-derived cells. Phalloidin (green) and Nck (red) staining of Met^{mt} tumor-derived cells (6030T) from the same experiment as in Figure 5.7. Z-stack images were captured using a Zeiss 510 Meta laser scanning confocal microscope with a 100x oil immersion objective. Scramble siRNA control treated cells have a height of 7.6 μm while Nck1/2 siRNA treated cells have a height of 2.63 μm. Nck1/2 siRNA treated cells are also larger than scrambled siRNA treated cells. Scale bars represent 20 μm.

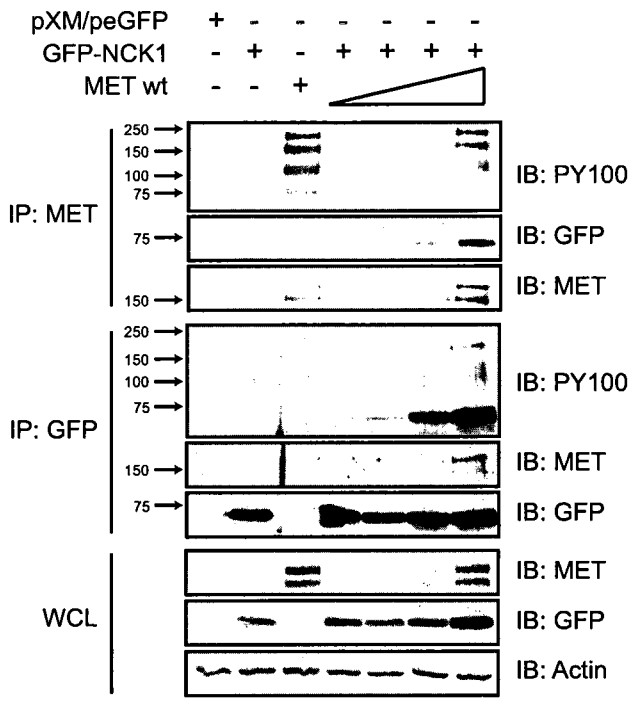


Figure 5.9: NCK1 is tyrosine phosphorylated in a MET-dependent manner. HEK293T lysates (500 μg) transiently expressing increasing wild type (wt) MET protein levels and consistent levels of GFP-NCK1 were tested for their ability to associate through co-immunoprecipitation and immunoblotted with antibodies recognizing phospho-tyrosine (PY100), MET and GFP. Whole cell lysates (25 μg) were separated by SDS-PAGE and immunoblotted with MET (Met148), GFP and actin antibodies.

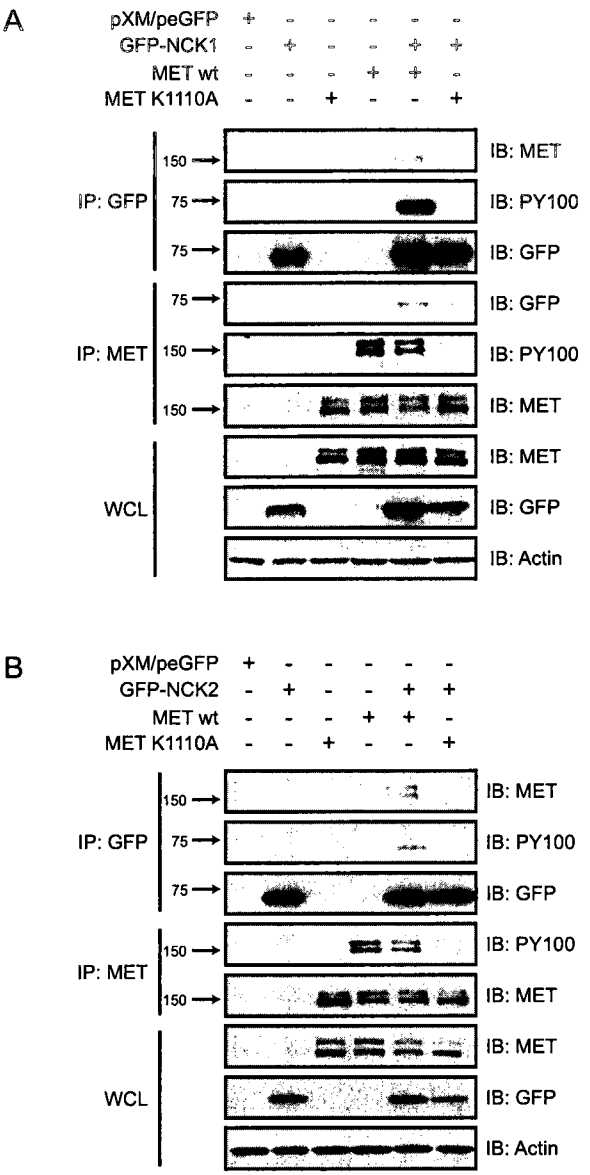


Figure 5.10: NCK1 and NCK2 proteins associate with the activated Met receptor. HEK293T lysates (500 µg) transiently expressing wild type (wt) MET or a kinase inactive MET mutant (K1110A) were tested for their ability to co-immunoprecipitate with GFP-NCK1 (A) or GFP-NCK2 (B) and immunoblotted with antibodies recognizing phospho-tyrosine (PY100), MET (Met148) and GFP. Whole cell lysates (WCL, 25 µg) were separated by SDS-PAGE and immunoblotted with Met (Met148), GFP and actin antibodies.

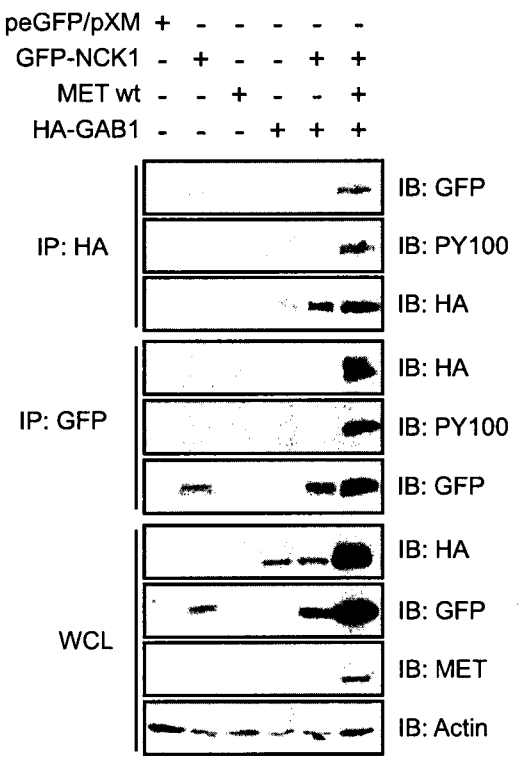


Figure 5.11: NCK1 engages with a GAB1 complex downstream from the activated MET receptor. HEK293T lysates (500 µg) transiently expressing HA-tagged GAB1 were tested for co-immunoprecipitation with GFP-NCK1 in the presence and absence of the activated wild type (wt) MET receptor, and immunoblotted with antibodies recognizing phospho-tyrosine (PY100), HA and GFP. Whole cell lysates (25 µg) were separated by SDS-PAGE and immunoblotted with MET (Met148), GFP and actin antibodies.

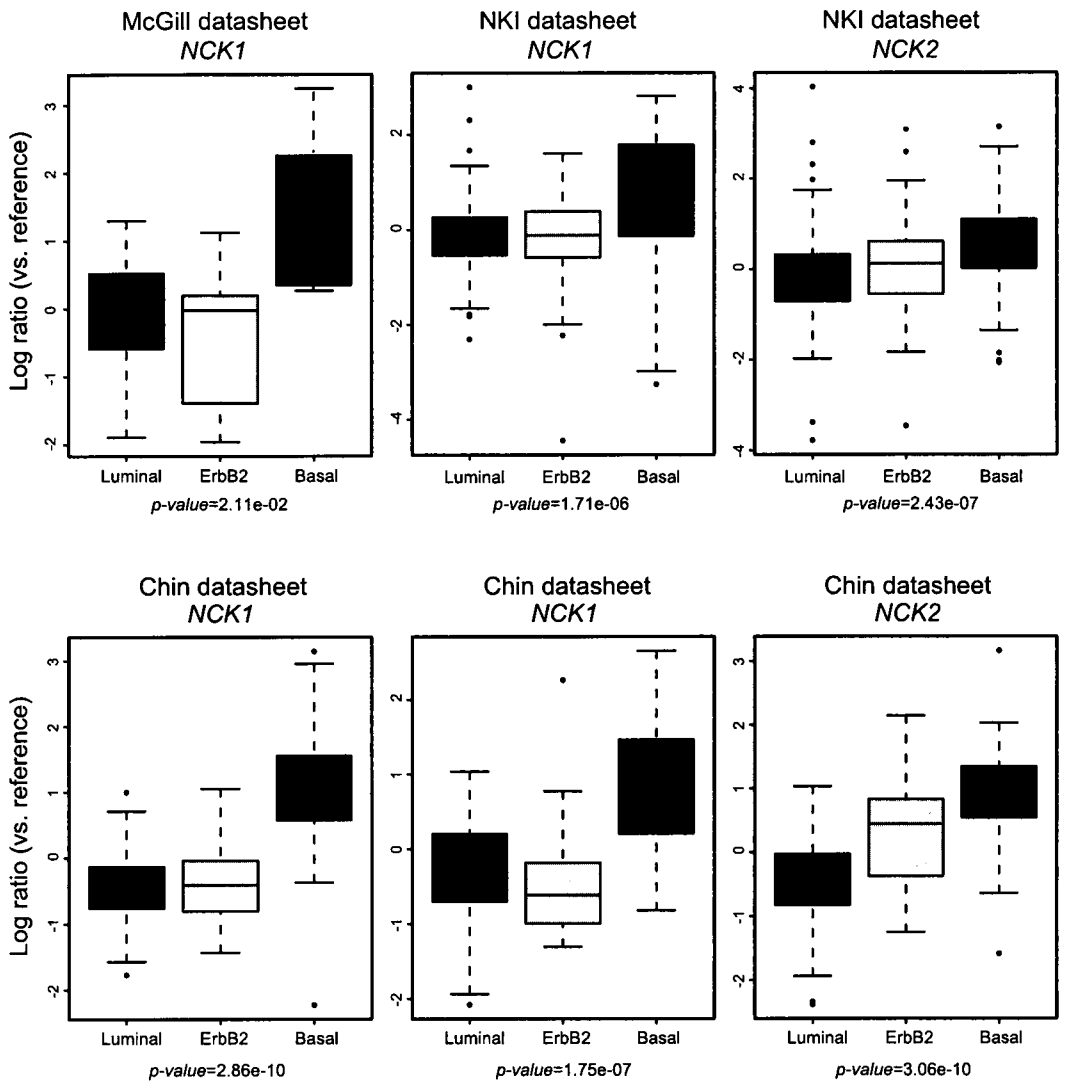


Figure 5.12: *NCK1* and *NCK2* mRNA levels are elevated in human basal breast cancers. *NCK1* and *NCK2* RNA levels in human breast cancer subtypes: luminal (blue), ErbB2 (pink), basal (red). Box and whisker plots represent levels of mRNA for *NCK1*, and *NCK2* in the McGill, NKI, and Chin datasets (see text for details). Two suitable probes for *NCK1* were available from the Chin dataset. The Kruskal-Wallis rank sum test was used as the statistic.

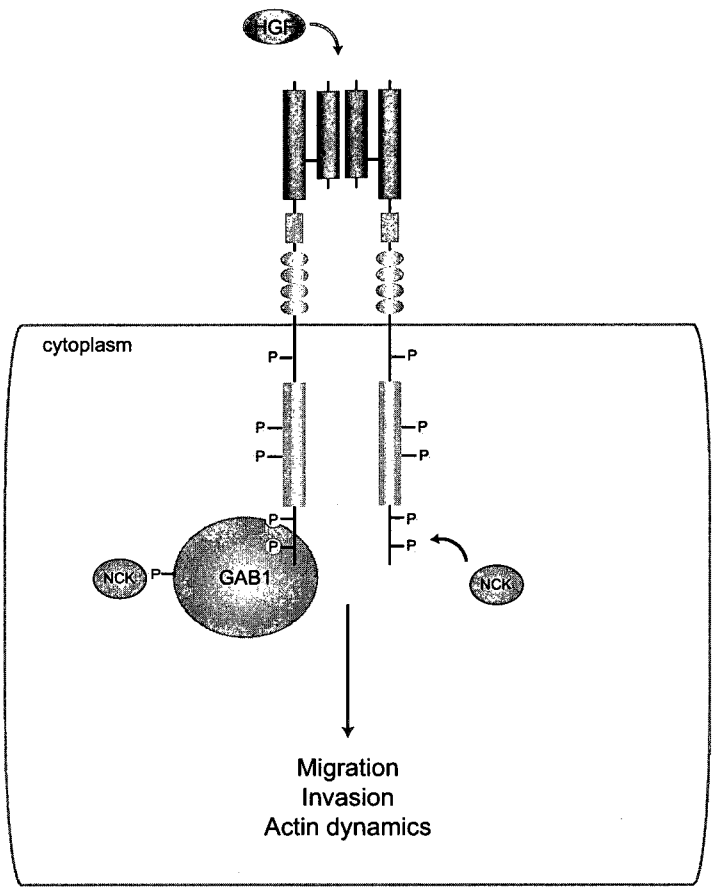


Figure 5.13: A model for an NCK-GAB1-MET containing complex. Upon activation of the MET receptor either by overexpression or binding to its ligand, HGF, the adaptor/scaffold protein GAB1 is recruited and tyrosine phosphorylated. Since the GAB1-NCK and MET-NCK associations are dependent on an activated MET receptor, NCK proteins be recruited to MET directly, or through tyrosine phosphorylated GAB1.

Chapter 6

6 General Discussion

6.1 Introductory statement

In this thesis, I have presented the characterization of a novel mouse model of breast cancer induced by an oncogenic variant of the Met receptor, expressed in the mammary epithelium under the control of the mouse mammary tumor virus promoter (MMTV/Met^{mt}). I have established a new protocol to procure epithelial cells from mouse mammary tumor and normal tissue, which affords high quality RNA for microarray analysis (Chapter 3). I have demonstrated through gene expression and immunohistochemistry approaches that the MMTV/Met^{mt} model produces mammary tumors with two predominant phenotypes, luminal ("solid") and basal-like ("mixed pathology") (Chapter 4). Furthermore, I have shown that MMTV/Met^{mt} basal tumors cluster with mouse models of breast cancer exhibiting features of epithelial to mesenchymal (EMT) transition and human basal breast cancer, while MMTV/Met^{mt} luminal tumors cluster with mouse models exhibiting solid nodular pathology and human luminal breast cancer. Significantly, I have demonstrated that in human breast cancer MET expression and protein levels are variable across all molecular subtypes, but consistently elevated in the human basal subtype. In a large dataset consisting of breast cancers from patients with axillary lymph node negative disease, elevated MET protein correlated with poor prognosis in basal, and other poor outcome tumors. When MET was co-elevated with EMT transcription factor, SNAIL, this correlated with poorer outcome than elevated levels of either MET or SNAIL alone. To further investigate the molecular mechanisms involved in Met-mediated mammary tumorigenesis, I have established a role

for two SH2/SH3 adaptor proteins, Nck1 and Nck2, which were co-elevated with Met in MMTV/Met^{mt} murine mammary tumors and in human basal breast cancers (Chapter 5). I have demonstrated a role for Nck proteins in cellular motility and actin dynamics using MMTV/Met^{mt} tumor-derived epithelial cell cultures established from MMTV/Met^{mt} tumors (Chapter 2). In this discussion, I will address outstanding scientific questions regarding the data presented and propose how these can be investigated in future experiments.

6.2 Pathway pathology: MMTV/Met^{mt} mouse model produces tumors with multiple phenotypes

As described in Chapter 4, the MMTV/Met^{mt} model generated mammary tumors with multiple phenotypes including solid nodular, adenosquamous, adenomyoepithelial, papillary, scirrhous, and spindle cell carcinomas. This is a relatively unique characteristic of a breast cancer mouse model driven by a receptor tyrosine kinase under the transcriptional control of the MMTV promoter. Several MMTV-driven receptor tyrosine kinase mouse models of breast cancers, which strongly induce the ras/MAPK/ERK signaling pathway, have been found to characteristically produce tumors with a solid nodular phenotype (Rosner, Miyoshi et al. 2002). The observation that genotype predicts phenotype is one of the premises of "pathway pathology" (Rosner, Miyoshi et al. 2002). It was thus intriguing that the MMTV/Met^{mt} mouse model produced 53% of tumors with a mixed phenotype and 47% of tumors with solid nodular pathology. We wondered why this was the case and explore reasons for the phenotypic diversity in the sections that follow.

6.3 Why does the MMTV/Met^{mt} model produce mammary tumors with multiple phenotypes?

The majority of MMTV/Met^{mt} mammary tumors were induced from the Met receptor tyrosine kinase expressing either the M1248T or Y1003F/M1248T mutations. Although transgenic animals were also generated by overexpression of the wild type and Y1003F variants of the Met receptor (Chapter 4, Table 4.1), we observed greater tumor penetrance with the M1248T and Y1003F/M1248T mutations and examined these further with gene expression analysis. As described earlier, Y1003 is a tyrosine residue within the juxtamembrane domain of the Met receptor that is responsible for binding to the tyrosine kinase-binding (TKB) domain of the E3 ligase Cbl (Peschard, Fournier et al. 2001). Cbl is a negative regulator of several receptor tyrosine kinases, including Met, and acts to downregulate signal transduction from the receptor by targeting Met for degradation through clathrin-mediated endocytosis (Abella, Peschard et al. 2005). Interestingly, upon time course stimulation of cells expressing either the Y1003F or Y1003F/M1248T Met receptor variants with HGF, a delay in receptor degradation and sustained ras/MAPK signaling was observed compared to cells expressing wild type or M1248T Met receptors (Petkiewicz 2007). Thus, is it conceivable that the Met Y1003F/M1248T variant may drive the induction of solid nodular tumors through sustained activation of the ras/MAPK pathway, while the Met M1248T preferentially produces non-solid, mixed pathology phenotypes (Chapter 4).

However, there is one caveat to this explanation of the phenotypic variance observed among Met Y1003F/M1248T-expressing tumors. Mixed pathology phenotypes were also observed among Met Y1003F/M1248T-expressing tumors. Thus, mutational status of the Met receptor alone cannot explain the phenotypic differences generated by

the MMTV/Met^{mt} mouse model of breast cancer. Since the MMTV promoter can be expressed in both luminal and myoepithelial cellular lineages, we considered whether a combination of Met receptor mutational status and the cell of origin expressing the transgene might account for the phenotypic differences. There are currently two predominant models on the generation of mammary tumors, the cancer stem cell and clonal evolution theories (described in Chapter 1), and I will explain our rationale in the context of these models.

6.4 Model 1: Cancer stem cell theory

In Chapter 4, we demonstrated that the MMTV/Met^{mt} model generates mammary tumors expressing markers of distinct cell lineages (Figures 4.8 to 4.11). Additionally, the Met^{mt} transgene was co-expressed with basal marker keratin 5 in mixed pathology tumors (Figure 4.10), whereas Met^{mt} was co-expressed with luminal marker keratins 8/18 in solid pathology tumors (Figure 4.11). These data argue that the Met^{mt} transgene may target progenitor cells for transformation. Transformed progenitor cells, or cancer stem cells, would have the capacity to differentiate, albeit aberrantly, toward distinct cell lineages. Using the strictest definition of the cancer stem cell theory, only progenitor cells have the ability to undergo transformation (Polyak 2007). Progenitor cells expressing the Met Y1003F/M1248T mutation may be driven to differentiate towards the luminal cell lineage owing to potent ras/MAPK signals. In contrast, a progenitor cell expressing the Met M1248T mutant may be selected for differentiation toward a mixed basal, myoepithelial lineage. Earlier results from our laboratory showed that mixed pathology tumors (i.e.: papillary, scirrhous phenotypes) induced by the Met Y1003F/M1248T mutation expressed lower levels of Met^{mt} transgene protein and thus lower levels of Met^{mt} tyrosine

phosphorylation compared to solid nodular tumors induced by the same mutant receptor (Petkiewicz 2007). Therefore, differences in signals generated from a single mutant receptor could result in distinct cellular differentiation, thus producing diverse phenotypes. Whereas a potent ras/MAPK signal driven by a highly expressed receptor tyrosine kinase can push a progenitor cell toward differentiation and expansion of a luminal cell lineage, a weaker signal from a less activated (tyrosine phosphorylated) receptor may allow for differentiation along other cellular lineages.

Several lines of evidence support this explanation for the spectrum of mammary tumors generated by the MMTV/Met^{mt} model of breast cancer. An earlier report demonstrated that transgenic MMTV/Wnt-1, MMTV/myc and MMTV/ß-catenin mice produce tumors as the result of expansion of a progenitor cell compartment, whereas strong inducers of the ras/MAPK pathway, models MMTV/PyMT, MMTV/ras and MMTV/ErbB2 did not, and predominantly expanded the luminal cell compartment (Li, Welm et al. 2003). The ability of the Met receptor to induce transformation in a progenitor cell population was demonstrated using an experimental animal model system (Welm, Kim et al. 2005). By virtue of a retroviral vector system (murine stem cell virus, pMSCV/pMIG) that specifically transduces progenitor cells, Met and transcription factor/oncogene myc were co-expressed in murine mammary progenitor cells and transplanted into cleared mammary fat pads of mice. This generated mammary tumors that expressed markers of distinct cell lineage, with cellular regions positive for basal marker keratin 6, myoepithelial maker keratin 14, and luminal markers keratins 8/18. Expression of wild type Met alone, or Myc alone did not produce mammary tumors, but co-operation from both induced tumors with heterogeneous phenotypes and cellular

diversity. These data illustrate that different oncogenic signals expressed in progenitor cells can drive differentiation along distinct cellular lineages.

Experiments to test this theory for the Met^{mt} model include the isolation of progenitor cells from the mammary gland, viral infection of wild type, Y1003F, M1248T, and Y1003F/M1248T Met receptor variants within these cells, and subsequent transplantation into cleared mammary fat pads. Different clones of Met Y1003F/M1248T-expressing progenitor cells can be selected with varying levels of receptor expression. Tumor explants can be compared for differing tumor phenotypes and correlated back to Met mutational status. Alternatively, an *in vivo* approach to test this scientific question is to generate transgenic animals that target transgene expression under the transcriptional control of the keratin 5 or 14 promoters. These have been shown to target expression of transgenes in the basal layer of the mammary gland and produce tumors exhibiting markers of distinct cell lineages (Jonkers, Meuwissen et al. 2001; Teuliere, Faraldo et al. 2005). Another layer of complexity is that there may exist differences in the signaling pathways downstream of the Met M1248T and Y1003F/M1248T receptors in terms of diversity in the engagement of protein complexes that are as yet unknown that may also account for phenotypic differences in Met^{mt} mammary tumors.

6.5 Model 2: Clonal evolution

An alternative explanation is possible within the context of the clonal evolution model of mammary tumorigenesis. Within the clonal evolution theory, all cell types, no matter the state of differentiation, are equally susceptible to genetic or epigenetic 'hits' and transformation. It is thus feasible, that a genetic, epigenetic or extracellular factor

from microenvironment can induce transformation in a differentiated cell within the mammary epithelium and cause a conversion of that differentiated cell into one with progenitor-like features (Polyak 2007). Thus, in contrast to the cancer stem cell theory, here the cell of origin is a differentiated cell which, following a molecular insult, induces behavioral and phenotypic changes to those resembling stem cells. This scenario reflects 'dedifferentiation,' where a differentiated cell reverts to an undifferentiated state. Indeed, recent evidence supports this hypothesis. Mani *et al.* demonstrated a specific link between EMT, the stem cell phenotype and cellular plasticity (Mani, Guo et al. 2008). Upon induction of an EMT in differentiated human immortalized mammary epithelial cells, these cells acquired a mesenchymal morphology consistent with an EMT, expressed EMT markers, but also exhibited stem cell-like properties (Mani, Guo et al. 2008). In this system, an EMT was induced by different means, including prolonged exposure (12 days) to TGF-β, and stable expression of the EMT transcription factors Twist or Snail. Cells that underwent an EMT had increased ability to form mammospheres in culture, and increased expression of stem cell markers $CD44^{high}CD24^{low}$ (Mani, Guo et al. 2008). These data suggest that differentiated cells have a degree of plasticity in that they can be converted to a more undifferentiated state through EMT (Radisky and LaBarge 2008).

As discussed in Chapter 1, HGF is an inducer of EMT, resulting in cell scattering and motility. During embryogenesis, the MET/HGF signaling axis is implicated in processes governed by EMT such as gastrulation, and in the migration of neural crest cells from the dermomyotome to the limb buds. Thus, it is possible that other inducers of EMT, such as HGF or activated Met receptor variants, can cause a switch to a "stem cell-like," less differentiated phenotype. To test this, different Met receptor variants can be

expressed in human immortalized mammary epithelial cells and tested for their ability to convert these cells to stem cell-like cells by monitoring an increase proportion of $CD44^{high}CD24^{low}$ cells, and an increase in mammosphere-forming abilities (which are non-adherent, sphere-forming cell cultures derived from progenitor cells). Low Met Y1003F/M1248T or M1248T-expressing cells may induce a switch to the stem cell phenotype, whereas highly expressing Met Y1003F/M1248T mutants may do so to a lesser extent if not at all. These cells can subsequently be injected into mammary fat pads and tumor phenotypes may be analyzed for differences.

6.6 Concluding comments on Models 1 and 2

I have outlined two different mechanisms of Met^{mt}-induced mammary tumorigenesis that could account for the spectrum of phenotypes observed from mutation of a single receptor tyrosine kinase. However, these models are not necessarily mutually exclusive (Polyak 2007; Shipitsin, Campbell et al. 2007). For instance, mammary tumorigenesis may be initiated through targeted expression of Met receptor variants to a cancer stem cell (CSC_1), which subsequently undergoes clonal evolution by acquiring additional genetic events to form CSC_2. Although this is a simplistic model, diverse phenotypes may arise as a result of differing secondary mutations acquired by different cancer stem cell populations. These would then give rise to tumors with differing phenotypes. Indeed, we found that the mixed pathology tumors, which were generated in part by MMTV/Met Y1003F/M1248T and and MMTV/M1248T animals, showed immunopositivity for Trp53 (Chapter 4). Since the $MMTV/Met^{mt}$ model generates tumors at long latency it is likely that multiple additional genetic events are acquired in this time (Chapter 4). We have also shown that the adaptor proteins Nck1 and Nck2 are co-

elevated with Met in the MMTV/Met^{mt} tumors and these may represent other contributing factors. To explore other genetic events that may be contributing factors, our laboratory aims to analyze aberrations at the genomic DNA level, through array-based comparative genomics hybridization (aCGH) and through mirRNA analysis of MMTV/Met^{mt} tumors.

6.7 Which patients would benefit from anti-MET therapies?

In Chapter 4, we shown that *MET* is expressed at variable levels across the molecular subtypes of breast cancer, but that it is consistently and significantly elevated in the basal subtype. When we examined protein levels of MET in a cohort of 668 breast cancer patients with axillary lymph node negative disease, we found that firstly, high MET protein levels correlated with the basal subtype where 65% of tumors were MET positive. Secondly, although MET protein was primarily elevated in the basal subtype of breast cancers, it was present at variable levels across all subtypes. Thirdly, high MET protein levels correlated with poor prognosis and with TP53 immunopositivity. Our interpretation is that elevated MET status is a marker of poor outcome breast cancers regardless of basal and non-basal subtypes, but since a large portion of poor outcome tumors are of the basal subtype, that MET correlates with the basal group as well. One question we might ask is, who would benefit from anti-MET therapies? I will speculate and propose answers to this question, based on similarities and correlations observed during the course of these studies.

6.7.1 Human breast cancer with EMT features

A portion of Met^{mt} mixed pathology tumors exhibited adenosquamous, myoepithelial and spindle cell tumor pathologies, and had morphological features resembling human basal breast cancer such as squamous metaplasia, central necrosis,

pushing margins, high grade, lymphocytic infiltration, nuclear atypia and pleomorphism (Livasy, Karaca et al. 2006). Some of these histological features are similar to those of medullary (lymphocytic infiltrate and pushing margins) or metaplastic (spindle cells or squamous metaplasia) breast carcinomas, two rare types of breast cancer with a basal-like immunohistochemistry profile (Jacquemier, Padovani et al. 2005; Reis-Filho, Milanezi et al. 2006). These tumors are also known as carcinosarcomas, which are thought to contain features of EMT (Damonte, Gregg et al. 2007). Gene expression studies of medullary and metaplastic breast tumors have identified an EMT-signature and they cluster among sporadic basal-like breast cancers (Lien, Hsiao et al. 2007; Sarrio, Rodriguez-Pinilla et al. 2008). Thus, human breast cancers displaying features of EMT may also represent tumors with characteristically elevated levels of MET protein.

6.7.2 Human breast cancers with elevated levels of Met and EGFR

About 65% of basal breast cancers exhibit elevated levels of the EGF receptor tyrosine kinase (Nielsen, Hsu et al. 2004). As introduced in Chapter 1, MET receptor has been shown to co-operate with EGFR in multiple aggressive human malignancies, such as glioblastoma multiforme, non-small cell lung cancer, and an *in vitro* model of mammary tumorigenesis (Bonine-Summers, Aakre et al. 2007; Engelman, Zejnullahu et al. 2007; Huang, Mukasa et al. 2007; Stommel, Kimmelman et al. 2007). Notably, EGFR inhibitors have been used for the treatment of human breast cancers but with little success (Blagosklonny and Darzynkiewicz 2003; Twombly 2005). One study showed that MET might be involved in a mechanism of acquired resistance (Mueller, Hunter et al. 2008). Data consistent with a co-operation between MET and the kinase, SRC, to compensate for the loss of functional EGFR in breast cancer cells treated with EGFR inhibitors was

demonstrated (Mueller, Hunter et al. 2008). Although MET has been shown to be a prognostic factor independent of both EGFR and ERBB2 (Tolgay Ocal, Dolled-Filhart et al. 2003; Lengyel, Prechtel et al. 2005), a subset of human breast cancers co-express MET and ERBB2 (Lengyel, Prechtel et al. 2005). Thus, human breast cancers in which both EGFR and MET are co-elevated may benefit from a combination of EGFR and MET inhibitors.

6.7.3 Human breast cancers with mutated TP53

The tumor suppressor TP53 is frequently mutated in human cancer (Rakha, Putti et al. 2006). TP53 protein levels are normally very low within the cell due to rapid turnover within the cell, but in response to various stressful stimuli, such as DNA damage or hypoxia, TP53 protein is stabilized and activated to perform its functions (Blackburn and Jerry 2002). MET and TP53 have been found to cooperate in Li-Fraumeni syndrome (Rong, Donehower et al. 1995). In support of this, Smolen *et al.* observed amplification of the *Met* locus in 73% of mammary tumors induced by a conditional Brca1/Trp53-deficient mouse model (Smolen, Muir et al. 2006). Additionally, *in vitro* studies have shown that the expression of a family of miRNAs, miR-34a-c, is regulated by TP53 and that one of the transcriptional targets of the miR-34 family is the MET receptor (He, He et al. 2007). In the context of a non-functional TP53, the miR-34 family is downregulated, and endogenous *MET* RNA expression results. Overall, these results support genetic cooperation between Met and TP53, and thus, in tumors in which TP53 is mutated may also exhibit elevated MET levels.

6.7.4 Hereditary BRCA1-germline mutation breast cancers

Notably, a phenotypic similarity between sporadic basal-like tumors and *BRCA1* germline mutation breast cancers has been identified. About 80 – 90% of *BRCA1* germline mutation breast cancers are phenotypically basal-like raising the possibility of a similar mechanism of tumorigenesis between these tumors and sporadic basal breast cancers (Turner and Reis-Filho 2006). Although somatic mutations in the *BRCA1* locus have not been found in sporadic basal tumors, defects in the BRCA1-DNA repair pathway have been observed. *BRCA1*-promoter methylation, which is frequently accompanied by loss of heterozygosity of the other *BRCA1* allele results in reduced or undetectable expression in 11 – 30% of sporadic breast cancers (Catteau, Harris et al. 1999; Wei, Grushko et al. 2005; Turner, Reis-Filho et al. 2007). A reduction in BRCA1 nuclear expression has been found in sporadic basal subtypes compared to other subtypes (Abd El-Rehim, Pinder et al. 2004). *BRCA1* mRNA expression is negatively regulated by the helix-loop-helix transcription factor, ID4, and ID4 is elevated in sporadic basal tumors as well as in MMTV/Met^{mt} mixed pathology tumors (Chapter 4) (Perou, Sorlie et al. 2000; Beger, Pierce et al. 2001; Sorlie, Perou et al. 2001). Furthermore, in tumors where BRCA1 function is not directly abrogated, components upstream or downstream in the BRCA1 pathway may be affected in sporadic basal-like tumors such as Fanconi anemia complex, which has a net result of abrogating the BRCA1-DNA repair pathway (Garcia-Higuera, Taniguchi et al. 2001; Guenard, Labrie et al. 2008). Therefore, since *BRCA1* mutated tumors produce basal-like phenotypes, these may also have elevated Met levels, however, this hypothesis remains to be shown.

Overall, the histological phenotypes of tumors that may benefit from anti- MET therapeutics include basal invasive ductal carcinomas, inflammatory breast cancers, and

carcinosarcomas. Furthermore, elevated MET protein levels may be more likely to be found in tumors exhibiting BRCA1 or TP53 mutations, and high levels of EGFR. Identification of novel biomarkers, such as the Met receptor will aid in individualizing patient treatment. Thus, depending on molecular markers expressed by a patient's breast tumor, a combination of kinase inhibitors, anti-MET, anti-EGFR and anti-ERBB2 inhibitors may be of benefit. Although MET inhibitors are currently under investigation in clinical trials, (Liu, Yao et al. 2008) there is undoubtedly a number of tumor types, including breast that may benefit from their use.

6.8 Conclusion

Studies will be needed to further examine the role of MET and MET inhibitors in the context of human breast cancers. By generating a unique mouse model in which the Met receptor tyrosine kinase is expressed in the mammary epithelia, in conjunction with the examination of MET expression in human breast cancer, we have established a specific link between MET and basal-like breast cancer. This work identifies basal breast cancers, as well as additional poor outcome breast cancers, which may benefit from anti-MET receptor therapies and provides a useful model with which to study the molecular basis through which MET contributes to this disease.

7 References

(1980). "Steroid receptors in breast cancer: an NIH Consensus Development Conference, Bethesda, Maryland, June 27-29, 1979." Cancer **46**(12 Suppl): 2759-963.

(2008). Canadian Cancer Statistics 2008. C. C. S. N. C. I. o. Canada, Toronto, Canada.

Abd El-Rehim, D. M., S. E. Pinder, et al. (2004). "Expression of luminal and basal cytokeratins in human breast carcinoma." J Pathol **203**(2): 661-71.

Abella, J. V., M. M. Frigault, et al. (2008). "The Gab1 scaffold is required for RTK signal polarisation to dorsal ruffles."

Abella, J. V., P. Peschard, et al. (2005). "Met/Hepatocyte growth factor receptor ubiquitination suppresses transformation and is required for Hrs phosphorylation." Mol Cell Biol **25**(21): 9632-45.

Adams, B. D., I. K. Guttilla, et al. (2008). "Involvement of microRNAs in breast cancer." Semin Reprod Med **26**(6): 522-36.

Al-Hajj, M., M. S. Wicha, et al. (2003). "Prospective identification of tumorigenic breast cancer cells." Proc Natl Acad Sci U S A **100**(7): 3983-8.

Al-Kuraya, K., P. Schraml, et al. (2004). "Prognostic relevance of gene amplifications and coamplifications in breast cancer." Cancer Res **64**(23): 8534-40.

Allinen, M., R. Beroukhim, et al. (2004). "Molecular characterization of the tumor microenvironment in breast cancer." Cancer Cell **6**(1): 17-32.

Andrechek, E. R., M. A. Laing, et al. (2003). "Gene expression profiling of neu-induced mammary tumors from transgenic mice reveals genetic and morphological similarities to ErbB2-expressing human breast cancers." Cancer Res **63**(16): 4920-6.

Andrulis, I. L., S. B. Bull, et al. (1998). "neu/erbB-2 amplification identifies a poor-prognosis group of women with node-negative breast cancer. Toronto Breast Cancer Study Group." J Clin Oncol **16**(4): 1340-9.

Antoku, S., K. Saksela, et al. (2008). "A crucial role in cell spreading for the interaction of Abl PxxP motifs with Crk and Nck adaptors." J Cell Sci **121**(Pt 18): 3071-82.

Anton, I. M., W. Lu, et al. (1998). "The Wiskott-Aldrich syndrome protein-interacting protein (WIP) binds to the adaptor protein Nck." J Biol Chem **273**(33): 20992-5.

Baldus, S. E., E. J. Kort, et al. (2007). "Quantification of MET and hepatocyte growth factor/scatter factor expression in colorectal adenomas, carcinomas and non-neoplastic epithelia by quantitative laser scanning microscopy." Int J Oncol **31**(1): 199-204.

Balmain, A. and I. B. Pragnell (1983). "Mouse skin carcinomas induced in vivo by chemical carcinogens have a transforming Harvey-ras oncogene." Nature **303**(5912): 72-4.

Balslev, I., C. K. Axelsson, et al. (1994). "The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG)." Breast Cancer Res Treat **32**(3): 281-90.

Barbareschi, M. (1996). Prognostic value of the immunohistochemical expression of p53 in breast carcinomas - a review of the literature involving over 9,000 patients. Appl Immunohistochem. **4:** 106-116.

Baugh, L. R., A. A. Hill, et al. (2001). "Quantitative analysis of mRNA amplification by in vitro transcription." Nucleic Acids Res **29**(5): E29.

Bean, J., C. Brennan, et al. (2007). "MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib." Proc Natl Acad Sci U S A **104**(52): 20932-7.

Beger, C., L. N. Pierce, et al. (2001). "Identification of Id4 as a regulator of BRCA1 expression by using a ribozyme-library-based inverse genomics approach." Proc Natl Acad Sci U S A **98**(1): 130-5.

Berthou, S., D. M. Aebersold, et al. (2004). "The Met kinase inhibitor SU11274 exhibits a selective inhibition pattern toward different receptor mutated variants." Oncogene **23**(31): 5387-93.

Bettelheim, R., H. G. Penman, et al. (1984). "Prognostic significance of peritumoral vascular invasion in breast cancer." Br J Cancer **50**(6): 771-7.

Beviglia, L., K. Matsumoto, et al. (1997). "Expression of the c-Met/HGF receptor in human breast carcinoma: correlation with tumor progression." Int J Cancer **74**(3): 301-9.

Bhowmick, N. A., E. G. Neilson, et al. (2004). "Stromal fibroblasts in cancer initiation and progression." Nature **432**(7015): 332-7.

Bild, A. H., G. Yao, et al. (2006). "Oncogenic pathway signatures in human cancers as a guide to targeted therapies." Nature **439**(7074): 353-7.

Birchmeier, C., W. Birchmeier, et al. (2003). "Met, metastasis, motility and more." Nat Rev Mol Cell Biol **4**(12): 915-25.

Birnbaum, D., F. Bertucci, et al. (2004). "Basal and luminal breast cancers: basic or luminous? (review)." Int J Oncol **25**(2): 249-58.

Bissell, M. J. and M. A. Labarge (2005). "Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment?" Cancer Cell **7**(1): 17-23.

Blackburn, A. C. and D. J. Jerry (2002). "Knockout and transgenic mice of Trp53: what have we learned about p53 in breast cancer?" Breast Cancer Res **4**(3): 101-11.

Bladt, F., E. Aippersbach, et al. (2003). "The murine Nck SH2/SH3 adaptors are important for the development of mesoderm-derived embryonic structures and for regulating the cellular actin network." Mol Cell Biol **23**(13): 4586-97.

Bladt, F., D. Riethmacher, et al. (1995). "Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud." Nature **376**(6543): 768-71.

Blagosklonny, M. V. and Z. Darzynkiewicz (2003). "Why Iressa failed: toward novel use of kinase inhibitors (outlook)." Cancer Biol Ther **2**(2): 137-40.

Blenkiron, C., L. D. Goldstein, et al. (2007). "MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype." Genome Biol **8**(10): R214.

Bloom, H. J. and W. W. Richardson (1957). "Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years." Br J Cancer **11**(3): 359-77.

Boccaccio, C., M. Ando, et al. (1998). "Induction of epithelial tubules by growth factor HGF depends on the STAT pathway." Nature **391**(6664): 285-8.

Bocciardi, R., B. Mograbi, et al. (1997). "The multiple endocrine neoplasia type 2B point mutation switches the specificity of the Ret tyrosine kinase towards cellular substrates that are susceptible to interact with Crk and Nck." Oncogene **15**(19): 2257-65.

Bocker, W., R. Moll, et al. (2002). "Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept." Lab Invest **82**(6): 737-46.

Bohm, M., I. Wieland, et al. (1997). "Microbeam MOMeNT: non-contact laser microdissection of membrane-mounted native tissue." Am J Pathol **151**(1): 63-7.

Bong, Y. S., Y. H. Park, et al. (2004). "Tyr-298 in ephrinB1 is critical for an interaction with the Grb4 adaptor protein." Biochem J **377**(Pt 2): 499-507.

Bonine-Summers, A. R., M. E. Aakre, et al. (2007). "Epidermal growth factor receptor plays a significant role in hepatocyte growth factor mediated biological responses in mammary epithelial cells." Cancer Biol Ther **6**(4): 561-70.

Bonnet, D. and J. E. Dick (1997). "Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell." Nat Med **3**(7): 730-7.

Booth, B. W., C. A. Boulanger, et al. (2007). "Alveolar progenitor cells develop in mouse mammary glands independent of pregnancy and lactation." J Cell Physiol **212**(3): 729-36.

Borowiak, M., A. N. Garratt, et al. (2004). "Met provides essential signals for liver regeneration." Proc Natl Acad Sci U S A **101**(29): 10608-13.

Borresen-Dale, A. L. (2003). "TP53 and breast cancer." Hum Mutat **21**(3): 292-300.

Bottaro, D. P., J. S. Rubin, et al. (1991). "Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product." Science **251**(4995): 802-4.

Bouchard, L., L. Lamarre, et al. (1989). "Stochastic appearance of mammary tumors in transgenic mice carrying the MMTV/c-neu oncogene." Cell **57**(6): 931-6.

Braverman, L. E. and L. A. Quilliam (1999). "Identification of Grb4/Nckbeta, a src homology 2 and 3 domain-containing adapter protein having similar binding and biological properties to Nck." J Biol Chem **274**(9): 5542-9.

Brenton, J. D., L. A. Carey, et al. (2005). "Molecular classification and molecular forecasting of breast cancer: ready for clinical application?" J Clin Oncol **23**(29): 7350-60.

Brinkmann, V., H. Foroutan, et al. (1995). "Hepatocyte growth factor/scatter factor induces a variety of tissue-specific morphogenic programs in epithelial cells." J Cell Biol **131**(6 Pt 1): 1573-86.

Brodie, S. G., X. Xu, et al. (2001). "Multiple genetic changes are associated with mammary tumorigenesis in Brca1 conditional knockout mice." Oncogene **20**(51): 7514-23.

Buday, L., L. Wunderlich, et al. (2002). "The Nck family of adapter proteins: regulators of actin cytoskeleton." Cell Signal **14**(9): 723-31.

Burgemeister, R., R. Gangnus, et al. (2003). "High quality RNA retrieved from samples obtained by using LMPC (laser microdissection and pressure catapulting) technology." Pathol Res Pract **199**(6): 431-6.

Burgess, T., A. Coxon, et al. (2006). "Fully human monoclonal antibodies to hepatocyte growth factor with therapeutic potential against hepatocyte growth factor/c-Met-dependent human tumors." Cancer Res **66**(3): 1721-9.

Buyse, M., S. Loi, et al. (2006). "Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer." J Natl Cancer Inst **98**(17): 1183-92.

Camp, R. L., E. B. Rimm, et al. (1999). "Met expression is associated with poor outcome in patients with axillary lymph node negative breast carcinoma." Cancer **86**(11): 2259-65.

Cardiff, R. D. (2001). "Validity of mouse mammary tumour models for human breast cancer: comparative pathology." Microsc Res Tech **52**(2): 224-30.

Cardiff, R. D. and W. J. Muller (1993). "Transgenic mouse models of mammary tumorigenesis." Cancer Surv **16**: 97-113.

Cardiff, R. D., E. Sinn, et al. (1991). "Transgenic oncogene mice. Tumor phenotype predicts genotype." Am J Pathol **139**(3): 495-501.

Carter, C. L., C. Allen, et al. (1989). "Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases." Cancer **63**(1): 181-7.

Catteau, A., W. H. Harris, et al. (1999). "Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics." Oncogene **18**(11): 1957-65.

Chan, A. M., J. S. Rubin, et al. (1991). "Identification of a competitive HGF antagonist encoded by an alternative transcript." Science **254**(5036): 1382-5.

Charafe-Jauffret, E., C. Ginestier, et al. (2006). "Gene expression profiling of breast cell lines identifies potential new basal markers." Oncogene **25**(15): 2273-84.

Chen, M., H. She, et al. (1998). "Identification of Nck family genes, chromosomal localization, expression, and signaling specificity." J Biol Chem **273**(39): 25171-8.

Chen, M., H. She, et al. (2000). "Nckbeta adapter regulates actin polymerization in NIH 3T3 fibroblasts in response to platelet-derived growth factor bb." Mol Cell Biol **20**(21): 7867-80.

Chen, S. Y. and H. C. Chen (2006). "Direct interaction of focal adhesion kinase (FAK) with Met is required for FAK to promote hepatocyte growth factor-induced cell invasion." Mol Cell Biol **26**(13): 5155-67.

Cheng, N., A. Chytil, et al. (2008). "Transforming growth factor-beta signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion." Mol Cancer Res **6**(10): 1521-33.

Chin, K., S. DeVries, et al. (2006). "Genomic and transcriptional aberrations linked to breast cancer pathophysiologies." Cancer Cell **10**(6): 529-41.

Chirgadze, D. Y., J. P. Hepple, et al. (1999). "Crystal structure of the NK1 fragment of HGF/SF suggests a novel mode for growth factor dimerization and receptor binding." Nat Struct Biol **6**(1): 72-9.

Chou, M. M., J. E. Fajardo, et al. (1992). "The SH2- and SH3-containing Nck protein transforms mammalian fibroblasts in the absence of elevated phosphotyrosine levels." Mol Cell Biol **12**(12): 5834-42.

Choudhury, G. G., F. Marra, et al. (1996). "Thrombin stimulates association of src homology domain containing adaptor protein Nck with pp125FAK." Am J Physiol **270**(2 Pt 2): F295-300.

Christensen, J. G., R. Schreck, et al. (2003). "A selective small molecule inhibitor of c-Met kinase inhibits c-Met-dependent phenotypes in vitro and exhibits cytoreductive antitumor activity in vivo." Cancer Res **63**(21): 7345-55.

Cicchetti, P., B. J. Mayer, et al. (1992). "Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho." Science **257**(5071): 803-6.
Comoglio, P. M., S. Giordano, et al. (2008). "Drug development of MET inhibitors: targeting oncogene addiction and expedience." Nat Rev Drug Discov **7**(6): 504-16.
Conrotto, P., S. Corso, et al. (2004). "Interplay between scatter factor receptors and B plexins controls invasive growth." Oncogene **23**(30): 5131-7.
Cooper, C. S., D. G. Blair, et al. (1984). "Characterization of human transforming genes from chemically transformed, teratocarcinoma, and pancreatic carcinoma cell lines." Cancer Res **44**(1): 1-10.
Cooper, C. S., M. Park, et al. (1984). "Molecular cloning of a new transforming gene from a chemically transformed human cell line." Nature **311**(5981): 29-33.
Cortner, J., G. F. Vande Woude, et al. (1995). "The Met-HGF/SF autocrine signaling mechanism is involved in sarcomagenesis." Exs **74**: 89-121.
Coutinho, S., T. Jahn, et al. (2000). "Characterization of Ggrb4, an adapter protein interacting with Bcr-Abl." Blood **96**(2): 618-24.
Cowan, C. A. and M. Henkemeyer (2001). "The SH2/SH3 adaptor Grb4 transduces B-ephrin reverse signals." Nature **413**(6852): 174-9.
D'Eredita, G., C. Giardina, et al. (2001). "Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution." Eur J Cancer **37**(5): 591-6.
Da Silva, L., C. Clarke, et al. (2007). "Demystifying basal-like breast carcinomas." J Clin Pathol **60**(12): 1328-32.
Damonte, P., J. P. Gregg, et al. (2007). "EMT tumorigenesis in the mouse mammary gland." Lab Invest **87**(12): 1218-26.
Dankort, D., B. Maslikowski, et al. (2001). "Grb2 and Shc adapter proteins play distinct roles in Neu (ErbB-2)-induced mammary tumorigenesis: implications for human breast cancer." Mol Cell Biol **21**(5): 1540-51.
de Klein, A., A. G. van Kessel, et al. (1982). "A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia." Nature **300**(5894): 765-7.
Dean, M., M. Park, et al. (1985). "The human met oncogene is related to the tyrosine kinase oncogenes." Nature **318**(6044): 385-8.
Dean, M., M. Park, et al. (1987). "Characterization of the rearranged tpr-met oncogene breakpoint." Mol Cell Biol **7**(2): 921-4.
Deome, K. B., L. J. Faulkin, Jr., et al. (1959). "Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice." Cancer Res **19**(5): 515-20.
Deramaudt, T. and A. K. Rustgi (2005). "Mutant KRAS in the initiation of pancreatic cancer." Biochim Biophys Acta **1756**(2): 97-101.
Desai, K. V., N. Xiao, et al. (2002). "Initiating oncogenic event determines gene-expression patterns of human breast cancer models." Proc Natl Acad Sci U S A **99**(10): 6967-72.
Di Renzo, M. F., M. Olivero, et al. (1995). "Overexpression and amplification of the met/HGF receptor gene during the progression of colorectal cancer." Clin Cancer Res **1**(2): 147-54.

DiRenzo, J., S. Signoretti, et al. (2002). "Growth factor requirements and basal phenotype of an immortalized mammary epithelial cell line." Cancer Res **62**(1): 89-98.
Donate, L. E., E. Gherardi, et al. (1994). "Molecular evolution and domain structure of plasminogen-related growth factors (HGF/SF and HGF1/MSP)." Protein Sci **3**(12): 2378-94.
Dourdin, N., B. Schade, et al. (2008). "Phosphatase and tensin homologue deleted on chromosome 10 deficiency accelerates tumor induction in a mouse model of ErbB-2 mammary tumorigenesis." Cancer Res **68**(7): 2122-31.
Durinck, S., Y. Moreau, et al. (2005). "BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis." Bioinformatics **21**(16): 3439-40.
Eberwine, J., H. Yeh, et al. (1992). "Analysis of gene expression in single live neurons." Proc Natl Acad Sci U S A **89**(7): 3010-4.
Ehrenreich, A. (2006). "DNA microarray technology for the microbiologist: an overview." Appl Microbiol Biotechnol **73**(2): 255-73.
Elledge, R. M. and D. C. Allred (1994). "The p53 tumor suppressor gene in breast cancer." Breast Cancer Res Treat **32**(1): 39-47.
Elston, C. W. (1987). Grading of invasive carcinoma of the breast. Diagnostic histopathology of the breast. D. L. A. Page, T. J. Edinburgh, Churchill Livingstone: 300-11.
Elston, C. W., I. O. Ellis, et al. (1999). "Pathological prognostic factors in breast cancer." Crit Rev Oncol Hematol **31**(3): 209-23.
Emmert-Buck, M. R., R. F. Bonner, et al. (1996). "Laser capture microdissection." Science **274**(5289): 998-1001.
Engelman, J. A., K. Zejnullahu, et al. (2007). "MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling." Science **316**(5827): 1039-43.
Eva, A. and S. A. Aaronson (1983). "Frequent activation of c-kis as a transforming gene in fibrosarcomas induced by methylcholanthrene." Science **220**(4600): 955-6.
Fawcett, J. P., J. Georgiou, et al. (2007). "Nck adaptor proteins control the organization of neuronal circuits important for walking." Proc Natl Acad Sci U S A **104**(52): 20973-8.
Fearon, E. R., S. R. Hamilton, et al. (1987). "Clonal analysis of human colorectal tumors." Science **238**(4824): 193-7.
Finak, G., N. Bertos, et al. (2008). "Stromal gene expression predicts clinical outcome in breast cancer." Nat Med **14**(5): 518-27.
Finak, G., S. Sadekova, et al. (2006). "Gene expression signatures of morphologically normal breast tissue identify basal-like tumors." Breast Cancer Res **8**(5): R58.
Fixman, E. D., T. M. Fournier, et al. (1996). "Pathways downstream of Shc and Grb2 are required for cell transformation by the tpr-Met oncoprotein." J Biol Chem **271**(22): 13116-22.
Fixman, E. D., M. A. Naujokas, et al. (1995). "Efficient cell transformation by the Tpr-Met oncoprotein is dependent upon tyrosine 489 in the carboxy-terminus." Oncogene **10**(2): 237-49.
Fleige, S. and M. W. Pfaffl (2006). "RNA integrity and the effect on the real-time qRT-PCR performance." Mol Aspects Med **27**(2-3): 126-39.

Fodor, S. P., J. L. Read, et al. (1991). "Light-directed, spatially addressable parallel chemical synthesis." Science **251**(4995): 767-73.
Follenzi, A., S. Bakovic, et al. (2000). "Cross-talk between the proto-oncogenes Met and Ron." Oncogene **19**(27): 3041-9.
Fomchenko, E. I. and E. C. Holland (2005). "Stem cells and brain cancer." Exp Cell Res **306**(2): 323-9.
Fournier, T. M., D. Kamikura, et al. (1996). "Branching tubulogenesis but not scatter of madin-darby canine kidney cells requires a functional Grb2 binding site in the Met receptor tyrosine kinase." J Biol Chem **271**(36): 22211-7.
Frese, S., W. D. Schubert, et al. (2006). "The phosphotyrosine peptide binding specificity of Nck1 and Nck2 Src homology 2 domains." J Biol Chem **281**(26): 18236-45.
Frigault, M. M., M. A. Naujokas, et al. (2008). "Gab2 requires membrane targeting and the Met binding motif to promote lamellipodia, cell scatter, and epithelial morphogenesis downstream from the Met receptor." J Cell Physiol **214**(3): 694-705.
Futreal, P. A., Q. Liu, et al. (1994). "BRCA1 mutations in primary breast and ovarian carcinomas." Science **266**(5182): 120-2.
Galea, M. H., R. W. Blamey, et al. (1992). "The Nottingham Prognostic Index in primary breast cancer." Breast Cancer Res Treat **22**(3): 207-19.
Galisteo, M. L., J. Chernoff, et al. (1996). "The adaptor protein Nck links receptor tyrosine kinases with the serine-threonine kinase Pak1." J Biol Chem **271**(35): 20997-1000.
Gallego, M. I., B. Bierie, et al. (2003). "Targeted expression of HGF/SF in mouse mammary epithelium leads to metastatic adenosquamous carcinomas through the activation of multiple signal transduction pathways." Oncogene **22**(52): 8498-508.
Garcia, S., J. P. Dales, et al. (2007). "Poor prognosis in breast carcinomas correlates with increased expression of targetable CD146 and c-Met and with proteomic basal-like phenotype." Hum Pathol **38**(6): 830-41.
Garcia-Higuera, I., T. Taniguchi, et al. (2001). "Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway." Mol Cell **7**(2): 249-62.
Garrity, P. A., Y. Rao, et al. (1996). "Drosophila photoreceptor axon guidance and targeting requires the dreadlocks SH2/SH3 adapter protein." Cell **85**(5): 639-50.
Gentleman, R. C., V. J. Carey, et al. (2004). "Bioconductor: open software development for computational biology and bioinformatics." Genome Biol **5**(10): R80.
Gherardi, E., S. Sandin, et al. (2006). "Structural basis of hepatocyte growth factor/scatter factor and MET signalling." Proc Natl Acad Sci U S A **103**(11): 4046-51.
Gherardi, E. and M. Stoker (1990). "Hepatocytes and scatter factor." Nature **346**(6281): 228.
Gherardi, E., M. E. Youles, et al. (2003). "Functional map and domain structure of MET, the product of the c-met protooncogene and receptor for hepatocyte growth factor/scatter factor." Proc Natl Acad Sci U S A **100**(21): 12039-44.
Ghoussoub, R. A., D. A. Dillon, et al. (1998). "Expression of c-met is a strong independent prognostic factor in breast carcinoma." Cancer **82**(8): 1513-20.
Ginsberg, S. D. and S. Che (2004). "Combined histochemical staining, RNA amplification, regional, and single cell cDNA analysis within the hippocampus." Lab Invest **84**(8): 952-62.

Giordano, S., M. F. Di Renzo, et al. (1989). "Biosynthesis of the protein encoded by the c-met proto-oncogene." Oncogene **4**(11): 1383-8.
Giordano, S., C. Ponzetto, et al. (1989). "Tyrosine kinase receptor indistinguishable from the c-met protein." Nature **339**(6220): 155-6.
Glover, J. F. and P. D. Darbre (1989). "Multihormone regulation of MMTV-LTR in transfected T-47-D human breast cancer cells." J Steroid Biochem **32**(3): 357-63.
Going, J. J. and R. F. Lamb (1996). "Practical histological microdissection for PCR analysis." J Pathol **179**(1): 121-4.
Goncalves, A., E. Charafe-Jauffret, et al. (2008). "Protein profiling of human breast tumor cells identifies novel biomarkers associated with molecular subtypes." Mol Cell Proteomics **7**(8): 1420-33.
Graham, F. L. and A. J. van der Eb (1973). "A new technique for the assay of infectivity of human adenovirus 5 DNA." Virology **52**(2): 456-67.
Gruenheid, S., R. DeVinney, et al. (2001). "Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells." Nat Cell Biol **3**(9): 856-9.
Gual, P., S. Giordano, et al. (2000). "Sustained recruitment of phospholipase C-gamma to Gab1 is required for HGF-induced branching tubulogenesis." Oncogene **19**(12): 1509-18.
Guan, S., M. Chen, et al. (2007). "Nckbeta adapter controls neuritogenesis by maintaining the cellular paxillin level." Mol Cell Biol **27**(17): 6001-11.
Gudjonsson, T., L. Ronnov-Jessen, et al. (2002). "Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition." J Cell Sci **115**(Pt 1): 39-50.
Guenard, F., Y. Labrie, et al. (2008). "Mutational analysis of the breast cancer susceptibility gene BRIP1 /BACH1/FANCJ in high-risk non-BRCA1/BRCA2 breast cancer families." J Hum Genet **53**(7): 579-91.
Guerrero, I., P. Calzada, et al. (1984). "A molecular approach to leukemogenesis: mouse lymphomas contain an activated c-ras oncogene." Proc Natl Acad Sci U S A **81**(1): 202-5.
Guo, A., J. Villen, et al. (2008). "Signaling networks assembled by oncogenic EGFR and c-Met." Proc Natl Acad Sci U S A **105**(2): 692-7.
Guy, C. T., R. D. Cardiff, et al. (1992). "Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease." Mol Cell Biol **12**(3): 954-61.
Guy, C. T., M. A. Webster, et al. (1992). "Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease." Proc Natl Acad Sci U S A **89**(22): 10578-82.
Haraguchi, S., R. A. Good, et al. (1992). "Human prolactin regulates transfected MMTV LTR-directed gene expression in a human breast-carcinoma cell line through synergistic interaction with steroid hormones." Int J Cancer **52**(6): 928-33.
Haslam, S. Z., A. Drolet, et al. (2008). "Progestin-regulated luminal cell and myoepithelial cell-specific responses in mammary organoid culture." Endocrinology **149**(5): 2098-107.
Hay, E. D. (1995). "An overview of epithelio-mesenchymal transformation." Acta Anat (Basel) **154**(1): 8-20.

He, L., X. He, et al. (2007). "A microRNA component of the p53 tumour suppressor network." Nature **447**(7148): 1130-4.
Heinze, G. and M. Schemper (2001). "A solution to the problem of monotone likelihood in Cox regression." Biometrics **57**(1): 114-9.
Henson, D. E., L. Ries, et al. (1991). "Relationship among outcome, stage of disease, and histologic grade for 22,616 cases of breast cancer. The basis for a prognostic index." Cancer **68**(10): 2142-9.
Hernandez, S. and J. Lloreta (2006). "Manual versus laser micro-dissection in molecular biology." Ultrastruct Pathol **30**(3): 221-8.
Herschkowitz, J. I., K. Simin, et al. (2007). "Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors." Genome Biol **8**(5): R76.
Hill, M. and J. Hillova (1972). "Virus recovery in chicken cells tested with Rous sarcoma cell DNA." Nat New Biol **237**(71): 35-9.
Hing, H., J. Xiao, et al. (1999). "Pak functions downstream of Dock to regulate photoreceptor axon guidance in Drosophila." Cell **97**(7): 853-63.
Holgado-Madruga, M., D. R. Emlet, et al. (1996). "A Grb2-associated docking protein in EGF- and insulin-receptor signalling." Nature **379**(6565): 560-4.
Holland, S. J., N. W. Gale, et al. (1997). "Juxtamembrane tyrosine residues couple the Eph family receptor EphB2/Nuk to specific SH2 domain proteins in neuronal cells." Embo J **16**(13): 3877-88.
Hong, F., R. Breitling, et al. (2006). "RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis." Bioinformatics **22**(22): 2825-7.
Huang, P. H., A. Mukasa, et al. (2007). "Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma." Proc Natl Acad Sci U S A **104**(31): 12867-72.
Huebner, K., K. Kastury, et al. (1994). "Chromosome locations of genes encoding human signal transduction adapter proteins, Nck (NCK), Shc (SHC1), and Grb2 (GRB2)." Genomics **22**(2): 281-7.
Huh, C. G., V. M. Factor, et al. (2004). "Hepatocyte growth factor/c-met signaling pathway is required for efficient liver regeneration and repair." Proc Natl Acad Sci U S A **101**(13): 4477-82.
Hutchinson, J. N. and W. J. Muller (2000). "Transgenic mouse models of human breast cancer." Oncogene **19**(53): 6130-7.
Inazawa, T., T. Tanabe, et al. (2001). "Glucocorticoid-regulated expression of exogenous human growth hormone gene in rats." Mol Ther **4**(3): 267-72.
Inoue, A., T. Suzuki, et al. (2006). "Prospective phase II study of gefitinib for chemotherapy-naive patients with advanced non-small-cell lung cancer with epidermal growth factor receptor gene mutations." J Clin Oncol **24**(21): 3340-6.
Inuzuka, H., K. Yamanouchi, et al. (2001). "A transgenic mouse model for investigating the response of the upstream region of whey acidic protein (WAP) gene to various steroid hormones." Exp Anim **50**(1): 1-7.
Iscove, N. N., M. Barbara, et al. (2002). "Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA." Nat Biotechnol **20**(9): 940-3.

Ishiki, Y., H. Ohnishi, et al. (1992). "Direct evidence that hepatocyte growth factor is a hepatotrophic factor for liver regeneration and has a potent antihepatitis effect in vivo." Hepatology **16**(5): 1227-35.
Isola, J. J., O. P. Kallioniemi, et al. (1995). "Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer." Am J Pathol **147**(4): 905-11.
Ito, N., C. Wernstedt, et al. (1998). "Identification of vascular endothelial growth factor receptor-1 tyrosine phosphorylation sites and binding of SH2 domain-containing molecules." J Biol Chem **273**(36): 23410-8.
Jacquemier, J., L. Padovani, et al. (2005). "Typical medullary breast carcinomas have a basal/myoepithelial phenotype." J Pathol **207**(3): 260-8.
Jaffe, A. B. and A. Hall (2005). "Rho GTPases: biochemistry and biology." Annu Rev Cell Dev Biol **21**: 247-69.
Jain, A. N., K. Chin, et al. (2001). "Quantitative analysis of chromosomal CGH in human breast tumors associates copy number abnormalities with p53 status and patient survival." Proc Natl Acad Sci U S A **98**(14): 7952-7.
Jakubczak, J. L., W. J. LaRochelle, et al. (1998). "NK1, a natural splice variant of hepatocyte growth factor/scatter factor, is a partial agonist in vivo." Mol Cell Biol **18**(3): 1275-83.
Jeffers, M., M. Fiscella, et al. (1998). "The mutationally activated Met receptor mediates motility and metastasis." Proc Natl Acad Sci U S A **95**(24): 14417-22.
Jeffers, M., L. Schmidt, et al. (1997). "Activating mutations for the met tyrosine kinase receptor in human cancer." Proc Natl Acad Sci U S A **94**(21): 11445-50.
Jin, H., R. Yang, et al. (2008). "MetMAb, the one-armed 5D5 anti-c-Met antibody, inhibits orthotopic pancreatic tumor growth and improves survival." Cancer Res **68**(11): 4360-8.
Jones, N., I. M. Blasutig, et al. (2006). "Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes." Nature **440**(7085): 818-23.
Jones, N. and D. J. Dumont (1998). "The Tek/Tie2 receptor signals through a novel Dok-related docking protein, Dok-R." Oncogene **17**(9): 1097-108.
Jonkers, J., R. Meuwissen, et al. (2001). "Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer." Nat Genet **29**(4): 418-25.
Jun, H. T., J. Sun, et al. (2007). "AMG 102, a fully human anti-hepatocyte growth factor/scatter factor neutralizing antibody, enhances the efficacy of temozolomide or docetaxel in U-87 MG cells and xenografts." Clin Cancer Res **13**(22 Pt 1): 6735-42.
Kakunaga, T. (1978). "Neoplastic transformation of human diploid fibroblast cells by chemical carcinogens." Proc Natl Acad Sci U S A **75**(3): 1334-8.
Kallioniemi, A., O. P. Kallioniemi, et al. (1994). "Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization." Proc Natl Acad Sci U S A **91**(6): 2156-60.
Kallioniemi, O. P., A. Kallioniemi, et al. (1992). "ERBB2 amplification in breast cancer analyzed by fluorescence in situ hybridization." Proc Natl Acad Sci U S A **89**(12): 5321-5.

Kaposi-Novak, P., J. S. Lee, et al. (2006). "Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype." J Clin Invest **116**(6): 1582-95.

Kawaida, K., K. Matsumoto, et al. (1994). "Hepatocyte growth factor prevents acute renal failure and accelerates renal regeneration in mice." Proc Natl Acad Sci U S A **91**(10): 4357-61.

Kestila, M., U. Lenkkeri, et al. (1998). "Positionally cloned gene for a novel glomerular protein--nephrin--is mutated in congenital nephrotic syndrome." Mol Cell **1**(4): 575-82.

Khoury, H., M. A. Naujokas, et al. (2005). "HGF converts ErbB2/Neu epithelial morphogenesis to cell invasion." Mol Biol Cell **16**(2): 550-61.

Kim, C. F., E. L. Jackson, et al. (2005). "Identification of bronchioalveolar stem cells in normal lung and lung cancer." Cell **121**(6): 823-35.

Kim, K. J., L. Wang, et al. (2006). "Systemic anti-hepatocyte growth factor monoclonal antibody therapy induces the regression of intracranial glioma xenografts." Clin Cancer Res **12**(4): 1292-8.

Kind, L. S. (1953). "The altered reactivity of mice after immunization with Hemophilus pertussis vaccine." J Immunol **70**(4): 411-20.

Kinzler, K. W. and B. Vogelstein (1996). "Lessons from hereditary colorectal cancer." Cell **87**(2): 159-70.

Kobayashi, S., T. J. Boggon, et al. (2005). "EGFR mutation and resistance of non-small-cell lung cancer to gefitinib." N Engl J Med **352**(8): 786-92.

Kochhar, K. S. and A. P. Iyer (1996). "Hepatocyte growth factor induces activation of Nck and phospholipase C-gamma in lung carcinoma cells." Cancer Lett **104**(2): 163-9.

Kong-Beltran, M., S. Seshagiri, et al. (2006). "Somatic mutations lead to an oncogenic deletion of met in lung cancer." Cancer Res **66**(1): 283-9.

Kong-Beltran, M., J. Stamos, et al. (2004). "The Sema domain of Met is necessary for receptor dimerization and activation." Cancer Cell **6**(1): 75-84.

Krag, D. N., S. J. Anderson, et al. (2007). "Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial." Lancet Oncol **8**(10): 881-8.

Kremer, B. E., L. A. Adang, et al. (2007). "Septins regulate actin organization and cell-cycle arrest through nuclear accumulation of NCK mediated by SOCS7." Cell **130**(5): 837-50.

Kuperwasser, C., T. Chavarria, et al. (2004). "Reconstruction of functionally normal and malignant human breast tissues in mice." Proc Natl Acad Sci U S A **101**(14): 4966-71.

La Rosa, S., S. Uccella, et al. (2000). "Localization of Hepatocyte Growth Factor and Its Receptor met in Endocrine Cells and Related Tumors of the Gut and Pancreas: An Immunohistochemical Study." Endocr Pathol **11**(4): 315-329.

Lacey, J. V., Jr., S. S. Devesa, et al. (2002). "Recent trends in breast cancer incidence and mortality." Environ Mol Mutagen **39**(2-3): 82-8.

Lamorte, L., D. M. Kamikura, et al. (2000). "A switch from p130Cas/Crk to Gab1/Crk signaling correlates with anchorage independent growth and JNK activation in cells transformed by the Met receptor oncoprotein." Oncogene **19**(52): 5973-81.

Lamorte, L., S. Rodrigues, et al. (2003). "Crk associates with a multimolecular Paxillin/GIT2/beta-PIX complex and promotes Rac-dependent relocalization of Paxillin to focal contacts." Mol Biol Cell **14**(7): 2818-31.

Lamorte, L., I. Royal, et al. (2002). "Crk adapter proteins promote an epithelial-mesenchymal-like transition and are required for HGF-mediated cell spreading and breakdown of epithelial adherens junctions." Mol Biol Cell **13**(5): 1449-61.

Landis, M. D., D. D. Seachrist, et al. (2005). "Gene expression profiling of cancer progression reveals intrinsic regulation of transforming growth factor-beta signaling in ErbB2/Neu-induced tumors from transgenic mice." Oncogene **24**(33): 5173-90.

Lapidot, T., C. Sirard, et al. (1994). "A cell initiating human acute myeloid leukaemia after transplantation into SCID mice." Nature **367**(6464): 645-8.

Lawe, D. C., C. Hahn, et al. (1997). "The Nck SH2/SH3 adaptor protein is present in the nucleus and associates with the nuclear protein SAM68." Oncogene **14**(2): 223-31.

Lee, C. T., N. H. Chow, et al. (2008). "The prognostic significance of RON and MET receptor coexpression in patients with colorectal cancer." Dis Colon Rectum **51**(8): 1268-74.

Lee, J. H., S. U. Han, et al. (2000). "A novel germ line juxtamembrane Met mutation in human gastric cancer." Oncogene **19**(43): 4947-53.

Lee, N. H. and A. I. Saeed (2007). "Microarrays: an overview." Methods Mol Biol **353**: 265-300.

Legate, K. R., E. Montanez, et al. (2006). "ILK, PINCH and parvin: the tIPP of integrin signalling." Nat Rev Mol Cell Biol **7**(1): 20-31.

Lehmann, J. M., G. Riethmuller, et al. (1990). "Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homology units SH2 and SH3." Nucleic Acids Res **18**(4): 1048.

Lengyel, E., D. Prechtel, et al. (2005). "C-Met overexpression in node-positive breast cancer identifies patients with poor clinical outcome independent of Her2/neu." Int J Cancer **113**(4): 678-82.

Li, F., Y. Zhang, et al. (1999). "Integrin-linked kinase is localized to cell-matrix focal adhesions but not cell-cell adhesion sites and the focal adhesion localization of integrin-linked kinase is regulated by the PINCH-binding ANK repeats." J Cell Sci **112 (Pt 24)**: 4589-99.

Li, W., P. Hu, et al. (1992). "The SH2 and SH3 domain-containing Nck protein is oncogenic and a common target for phosphorylation by different surface receptors." Mol Cell Biol **12**(12): 5824-33.

Li, W. and H. She (2000). "The SH2 and SH3 adapter Nck: a two-gene family and a linker between tyrosine kinases and multiple signaling networks." Histol Histopathol **15**(3): 947-55.

Li, Y., B. Welm, et al. (2003). "Evidence that transgenes encoding components of the Wnt signaling pathway preferentially induce mammary cancers from progenitor cells." Proc Natl Acad Sci U S A **100**(26): 15853-8.

Liang, T. J., A. E. Reid, et al. (1996). "Transgenic expression of tpr-met oncogene leads to development of mammary hyperplasia and tumors." J Clin Invest **97**(12): 2872-7.

Lien, H. C., Y. H. Hsiao, et al. (2007). "Molecular signatures of metaplastic carcinoma of the breast by large-scale transcriptional profiling: identification of genes potentially related to epithelial-mesenchymal transition." Oncogene.

Lipshutz, R. J., S. P. Fodor, et al. (1999). "High density synthetic oligonucleotide arrays." Nat Genet **21**(1 Suppl): 20-4.

Liu, C. L., W. Prapong, et al. (2002). "Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays." Am J Pathol **161**(5): 1557-65.

Liu, C. Y., A. Flesken-Nikitin, et al. (1996). "Inactivation of the mouse Brca1 gene leads to failure in the morphogenesis of the egg cylinder in early postimplantation development." Genes Dev **10**(14): 1835-43.

Liu, X., H. Holstege, et al. (2007). "Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer." Proc Natl Acad Sci U S A **104**(29): 12111-6.

Liu, X., W. Yao, et al. (2008). "Targeting the c-MET signaling pathway for cancer therapy." Expert Opin Investig Drugs **17**(7): 997-1011.

Liu, Y. and L. R. Rohrschneider (2002). "The gift of Gab." FEBS Lett **515**(1-3): 1-7.

Livasy, C. A., G. Karaca, et al. (2006). "Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma." Mod Pathol **19**(2): 264-71.

Lock, L. S., M. M. Frigault, et al. (2003). "Grb2-independent recruitment of Gab1 requires the C-terminal lobe and structural integrity of the Met receptor kinase domain." J Biol Chem **278**(32): 30083-90.

Lock, L. S., C. R. Maroun, et al. (2002). "Distinct recruitment and function of Gab1 and Gab2 in Met receptor-mediated epithelial morphogenesis." Mol Biol Cell **13**(6): 2132-46.

Lokker, N. A., M. R. Mark, et al. (1992). "Structure-function analysis of hepatocyte growth factor: identification of variants that lack mitogenic activity yet retain high affinity receptor binding." Embo J **11**(7): 2503-10.

Lokker, N. A., L. G. Presta, et al. (1994). "Mutational analysis and molecular modeling of the N-terminal kringle-containing domain of hepatocyte growth factor identifies amino acid side chains important for interaction with the c-Met receptor." Protein Eng **7**(7): 895-903.

Loo, L. W., D. I. Grove, et al. (2004). "Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes." Cancer Res **64**(23): 8541-9.

Luzzi, V., V. Holtschlag, et al. (2001). "Expression profiling of ductal carcinoma in situ by laser capture microdissection and high-density oligonucleotide arrays." Am J Pathol **158**(6): 2005-10.

Luzzi, V., M. Mahadevappa, et al. (2003). "Accurate and reproducible gene expression profiles from laser capture microdissection, transcript amplification, and high density oligonucleotide microarray analysis." J Mol Diagn **5**(1): 9-14.

Ma, X. J., R. Salunga, et al. (2003). "Gene expression profiles of human breast cancer progression." Proc Natl Acad Sci U S A **100**(10): 5974-9.

Ma, X. J., Z. Wang, et al. (2004). "A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen." Cancer Cell **5**(6): 607-16.

Mani, S. A., W. Guo, et al. (2008). "The epithelial-mesenchymal transition generates cells with properties of stem cells." Cell **133**(4): 704-15.

Margolis, B., O. Silvennoinen, et al. (1992). "High-efficiency expression/cloning of epidermal growth factor-receptor-binding proteins with Src homology 2 domains." Proc Natl Acad Sci U S A **89**(19): 8894-8.

Maritano, D., P. Accornero, et al. (2000). "Two mutations affecting conserved residues in the Met receptor operate via different mechanisms." Oncogene **19**(10): 1354-61.

Maroun, C. R., M. Holgado-Madruga, et al. (1999). "The Gab1 PH domain is required for localization of Gab1 at sites of cell-cell contact and epithelial morphogenesis downstream from the met receptor tyrosine kinase." Mol Cell Biol **19**(3): 1784-99.

Maroun, C. R., M. A. Naujokas, et al. (2000). "The tyrosine phosphatase SHP-2 is required for sustained activation of extracellular signal-regulated kinase and epithelial morphogenesis downstream from the met receptor tyrosine kinase." Mol Cell Biol **20**(22): 8513-25.

Martens, T., N. O. Schmidt, et al. (2006). "A novel one-armed anti-c-Met antibody inhibits glioblastoma growth in vivo." Clin Cancer Res **12**(20 Pt 1): 6144-52.

Matsumoto, K. and T. Nakamura (2001). "Hepatocyte growth factor: renotropic role and potential therapeutics for renal diseases." Kidney Int **59**(6): 2023-38.

Matsumoto, K. and T. Nakamura (2003). "NK4 (HGF-antagonist/angiogenesis inhibitor) in cancer biology and therapeutics." Cancer Sci **94**(4): 321-7.

Matsumoto, K. and T. Nakamura (2008). "NK4 gene therapy targeting HGF-Met and angiogenesis." Front Biosci **13**: 1943-51.

McCarthy, A., K. Savage, et al. (2007). "A mouse model of basal-like breast carcinoma with metaplastic elements." J Pathol **211**(4): 389-98.

Medina, D. (2000). "The preneoplastic phenotype in murine mammary tumorigenesis." J Mammary Gland Biol Neoplasia **5**(4): 393-407.

Meisenhelder, J. and T. Hunter (1992). "The SH2/SH3 domain-containing protein Nck is recognized by certain anti-phospholipase C-gamma 1 monoclonal antibodies, and its phosphorylation on tyrosine is stimulated by platelet-derived growth factor and epidermal growth factor treatment." Mol Cell Biol **12**(12): 5843-56.

Michalopoulos, G. K. and M. C. DeFrances (1997). "Liver regeneration." Science **276**(5309): 60-6.

Michieli, P., M. Mazzone, et al. (2004). "Targeting the tumor and its microenvironment by a dual-function decoy Met receptor." Cancer Cell **6**(1): 61-73.

Mikulowska-Mennis, A., T. B. Taylor, et al. (2002). "High-quality RNA from cells isolated by laser capture microdissection." Biotechniques **33**(1): 176-9.

Miller, S. J., R. M. Lavker, et al. (2005). "Interpreting epithelial cancer biology in the context of stem cells: tumor properties and therapeutic implications." Biochim Biophys Acta **1756**(1): 25-52.

Minden, A., A. Lin, et al. (1995). "Selective activation of the JNK signaling cascade and c-Jun transcriptional activity by the small GTPases Rac and Cdc42Hs." Cell **81**(7): 1147-57.

Mitelman, F., Y. Kaneko, et al. (1990). "Report of the committee on chromosome changes in neoplasia." Cytogenet Cell Genet **55**(1-4): 358-86.

Mitsudomi, T. and Y. Yatabe (2007). "Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer." Cancer Sci **98**(12): 1817-24.

Miyazawa, K., H. Tsubouchi, et al. (1989). "Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor." Biochem Biophys Res Commun **163**(2): 967-73.
Montesano, R., K. Matsumoto, et al. (1991). "Identification of a fibroblast-derived epithelial morphogen as hepatocyte growth factor." Cell **67**(5): 901-8.
Montesano, R., J. V. Soriano, et al. (1998). "Isolation of EpH4 mammary epithelial cell subpopulations which differ in their morphogenetic properties." In Vitro Cell Dev Biol Anim **34**(6): 468-77.
Montesano, R., J. V. Soriano, et al. (1998). "Differential effects of hepatocyte growth factor isoforms on epithelial and endothelial tubulogenesis." Cell Growth Differ **9**(5): 355-65.
Mootha, V. K., C. M. Lindgren, et al. (2003). "PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes." Nat Genet **34**(3): 267-73.
Morotti, A., S. Mila, et al. (2002). "K252a inhibits the oncogenic properties of Met, the HGF receptor." Oncogene **21**(32): 4885-93.
Mueller, K. L., L. A. Hunter, et al. (2008). "Met and c-Src cooperate to compensate for loss of epidermal growth factor receptor kinase activity in breast cancer cells." Cancer Res **68**(9): 3314-22.
Muller, W. J., E. Sinn, et al. (1988). "Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene." Cell **54**(1): 105-15.
Mulligan, A. M., D. Pinnaduwage, et al. (2008). "Prognostic effect of basal-like breast cancers is time dependent: evidence from tissue microarray studies on a lymph node-negative cohort." Clin Cancer Res **14**(13): 4168-74.
Murray, G. I. (2007). "An overview of laser microdissection technologies." Acta Histochem **109**(3): 171-6.
Nagy, A., M. Gertsenstein, et al. (2003). Manipulating the Mouse Embryo: A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.
Nakaigawa, N., G. Weirich, et al. (2000). "Tumorigenesis mediated by MET mutant M1268T is inhibited by dominant-negative Src." Oncogene **19**(26): 2996-3002.
Nakajima, M., H. Sawada, et al. (1999). "The prognostic significance of amplification and overexpression of c-met and c-erb B-2 in human gastric carcinomas." Cancer **85**(9): 1894-902.
Nakamura, T., S. Mizuno, et al. (2000). "Myocardial protection from ischemia/reperfusion injury by endogenous and exogenous HGF." J Clin Invest **106**(12): 1511-9.
Nakamura, T., K. Nawa, et al. (1984). "Partial purification and characterization of hepatocyte growth factor from serum of hepatectomized rats." Biochem Biophys Res Commun **122**(3): 1450-9.
Nakamura, T., T. Nishizawa, et al. (1989). "Molecular cloning and expression of human hepatocyte growth factor." Nature **342**(6248): 440-3.
Naldini, L., K. M. Weidner, et al. (1991). "Scatter factor and hepatocyte growth factor are indistinguishable ligands for the MET receptor." Embo J **10**(10): 2867-78.
Naylor, T. L., J. Greshock, et al. (2005). "High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization." Breast Cancer Res **7**(6): R1186-98.

Neve, R. M., K. Chin, et al. (2006). "A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes." Cancer Cell **10**(6): 515-27.
Nguyen, L., M. Holgado-Madruga, et al. (1997). "Association of the multisubstrate docking protein Gab1 with the hepatocyte growth factor receptor requires a functional Grb2 binding site involving tyrosine 1356." J Biol Chem **272**(33): 20811-9.
Nielsen, T. O., F. D. Hsu, et al. (2004). "Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma." Clin Cancer Res **10**(16): 5367-74.
Niemann, C., V. Brinkmann, et al. (1998). "Reconstitution of mammary gland development in vitro: requirement of c-met and c-erbB2 signaling for branching and alveolar morphogenesis." J Cell Biol **143**(2): 533-45.
Nieto, A. I., G. Shyamala, et al. (2003). "Persistent mammary hyperplasia in FVB/N mice." Comp Med **53**(4): 433-8.
Niranjan, B., L. Buluwela, et al. (1995). "HGF/SF: a potent cytokine for mammary growth, morphogenesis and development." Development **121**(9): 2897-908.
Nishimura, R., W. Li, et al. (1993). "Two signaling molecules share a phosphotyrosine-containing binding site in the platelet-derived growth factor receptor." Mol Cell Biol **13**(11): 6889-96.
O'Brien, C. A., A. Pollett, et al. (2007). "A human colon cancer cell capable of initiating tumour growth in immunodeficient mice." Nature **445**(7123): 106-10.
O'Malley, F. P. and S. E. Pinder (2006). Breast Pathology. Philadelphia, Elsevier Inc.
Obenauer, J. C., L. C. Cantley, et al. (2003). "Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs." Nucleic Acids Res **31**(13): 3635-41.
Okigaki, M., M. Komada, et al. (1992). "Functional characterization of human hepatocyte growth factor mutants obtained by deletion of structural domains." Biochemistry **31**(40): 9555-61.
Okuducu, A. F., V. Janzen, et al. (2003). "Influence of histochemical stains on quantitative gene expression analysis after laser-assisted microdissection." Int J Mol Med **11**(4): 449-53.
Orian-Rousseau, V., L. Chen, et al. (2002). "CD44 is required for two consecutive steps in HGF/c-Met signaling." Genes Dev **16**(23): 3074-86.
Orimo, A., P. B. Gupta, et al. (2005). "Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion." Cell **121**(3): 335-48.
Otten, A. D., M. M. Sanders, et al. (1988). "The MMTV LTR promoter is induced by progesterone and dihydrotestosterone but not by estrogen." Mol Endocrinol **2**(2): 143-7.
Page, D. L. (1991). "Prognosis and breast cancer. Recognition of lethal and favorable prognostic types." Am J Surg Pathol **15**(4): 334-49.
Pao, W., V. A. Miller, et al. (2005). "Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain." PLoS Med **2**(3): e73.
Park, D. and S. G. Rhee (1992). "Phosphorylation of Nck in response to a variety of receptors, phorbol myristate acetate, and cyclic AMP." Mol Cell Biol **12**(12): 5816-23.

Park, M., M. Dean, et al. (1986). "Mechanism of met oncogene activation." Cell **45**(6): 895-904.
Park, M., M. Dean, et al. (1987). "Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors." Proc Natl Acad Sci U S A **84**(18): 6379-83.
Parkin, D. M. (1998). "The global burden of cancer." Semin Cancer Biol **8**(4): 219-35.
Pavenstadt, H., W. Kriz, et al. (2003). "Cell biology of the glomerular podocyte." Physiol Rev **83**(1): 253-307.
Pawson, T. and J. Schlessingert (1993). "SH2 and SH3 domains." Curr Biol **3**(7): 434-42.
Pei, X. F., M. S. Noble, et al. (2004). "Explant-cell culture of primary mammary tumors from MMTV-c-Myc transgenic mice." In Vitro Cell Dev Biol Anim **40**(1-2): 14-21.
Pepper, M. S., J. V. Soriano, et al. (1995). "Modulation of hepatocyte growth factor and c-met in the rat mammary gland during pregnancy, lactation, and involution." Exp Cell Res **219**(1): 204-10.
Pereira, H., S. E. Pinder, et al. (1995). "Pathological prognostic factors in breast cancer. IV: Should you be a typer or a grader? A comparative study of two histological prognostic features in operable breast carcinoma." Histopathology **27**(3): 219-26.
Perou, C. M., T. Sorlie, et al. (2000). "Molecular portraits of human breast tumours." Nature **406**(6797): 747-52.
Peschard, P. (2005). The role of Cbl-mediated ubiquitination in the regulation of the Met receptor tyrosine kinase. Department of Biochemistry. Montreal, McGill University. **Doctoral degree**.
Peschard, P., T. M. Fournier, et al. (2001). "Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein." Mol Cell **8**(5): 995-1004.
Peschard, P. and M. Park (2007). "From Tpr-Met to Met, tumorigenesis and tubes." Oncogene **26**(9): 1276-85.
Petkiewicz, S. L. (2007). The Met receptor tyrosine kinase in mammary gland tumorigenesis and development. Division of Experimental Medicine. Montreal, McGill University. **Doctor of Philosophy:** 226.
Petrelli, A., P. Circosta, et al. (2006). "Ab-induced ectodomain shedding mediates hepatocyte growth factor receptor down-regulation and hampers biological activity." Proc Natl Acad Sci U S A **103**(13): 5090-5.
Phillips, J. and J. H. Eberwine (1996). "Antisense RNA Amplification: A Linear Amplification Method for Analyzing the mRNA Population from Single Living Cells." Methods **10**(3): 283-8.
Phillips, N., R. D. Hayward, et al. (2004). "Phosphorylation of the enteropathogenic E. coli receptor by the Src-family kinase c-Fyn triggers actin pedestal formation." Nat Cell Biol **6**(7): 618-25.
Pinder, S. E., I. O. Ellis, et al. (1994). "Pathological prognostic factors in breast cancer. III. Vascular invasion: relationship with recurrence and survival in a large study with long-term follow-up." Histopathology **24**(1): 41-7.
Pinkel, D., R. Segraves, et al. (1998). "High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays." Nat Genet **20**(2): 207-11.

Pollack, J. R., C. M. Perou, et al. (1999). "Genome-wide analysis of DNA copy-number changes using cDNA microarrays." Nat Genet **23**(1): 41-6.
Pollack, J. R., T. Sorlie, et al. (2002). "Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors." Proc Natl Acad Sci U S A **99**(20): 12963-8.
Polyak, K. (2007). "Breast cancer: origins and evolution." J Clin Invest **117**(11): 3155-63.
Ponzetto, C., A. Bardelli, et al. (1994). "A multifunctional docking site mediates signaling and transformation by the hepatocyte growth factor/scatter factor receptor family." Cell **77**(2): 261-71.
Potemski, P., R. Kusinska, et al. (2005). "Prognostic relevance of basal cytokeratin expression in operable breast cancer." Oncology **69**(6): 478-85.
Pramatarova, A., P. G. Ochalski, et al. (2003). "Nck beta interacts with tyrosine-phosphorylated disabled 1 and redistributes in Reelin-stimulated neurons." Mol Cell Biol **23**(20): 7210-21.
Prat, M., T. Crepaldi, et al. (1998). "Agonistic monoclonal antibodies against the Met receptor dissect the biological responses to HGF." J Cell Sci **111 (Pt 2)**: 237-47.
Pulciani, S., E. Santos, et al. (1982). "Oncogenes in solid human tumours." Nature **300**(5892): 539-42.
Radford, D. M., K. Fair, et al. (1993). "Allelic loss on a chromosome 17 in ductal carcinoma in situ of the breast." Cancer Res **53**(13): 2947-9.
Radisky, D. C. and M. A. LaBarge (2008). "Epithelial-mesenchymal transition and the stem cell phenotype." Cell Stem Cell **2**(6): 511-2.
Radtke, F. and H. Clevers (2005). "Self-renewal and cancer of the gut: two sides of a coin." Science **307**(5717): 1904-9.
Raftopoulou, M. and A. Hall (2004). "Cell migration: Rho GTPases lead the way." Dev Biol **265**(1): 23-32.
Rahimi, N., W. Hung, et al. (1998). "c-Src kinase activity is required for hepatocyte growth factor-induced motility and anchorage-independent growth of mammary carcinoma cells." J Biol Chem **273**(50): 33714-21.
Rakha, E. A., T. C. Putti, et al. (2006). "Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation." J Pathol **208**(4): 495-506.
Rakha, E. A., J. S. Reis-Filho, et al. (2008). "Basal-like breast cancer: a critical review." J Clin Oncol **26**(15): 2568-81.
Ramaswamy, S., K. N. Ross, et al. (2003). "A molecular signature of metastasis in primary solid tumors." Nat Genet **33**(1): 49-54.
Ramos-Morales, F., B. J. Druker, et al. (1994). "Vav binds to several SH2/SH3 containing proteins in activated lymphocytes." Oncogene **9**(7): 1917-23.
Rearden, A. (1994). "A new LIM protein containing an autoepitope homologous to "senescent cell antigen"." Biochem Biophys Res Commun **201**(3): 1124-31.
Reis-Filho, J. S., F. Milanezi, et al. (2006). "Metaplastic breast carcinomas are basal-like tumours." Histopathology **49**(1): 10-21.
Ren, R., Z. S. Ye, et al. (1994). "Abl protein-tyrosine kinase selects the Crk adapter as a substrate using SH3-binding sites." Genes Dev **8**(7): 783-95.
Reya, T., S. J. Morrison, et al. (2001). "Stem cells, cancer, and cancer stem cells." Nature **414**(6859): 105-11.

Richardson, A. L., Z. C. Wang, et al. (2006). "X chromosomal abnormalities in basal-like human breast cancer." Cancer Cell **9**(2): 121-32.
Rikova, K., A. Guo, et al. (2007). "Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer." Cell **131**(6): 1190-203.
Rivera, G. M., S. Antoku, et al. (2006). "Requirement of Nck adaptors for actin dynamics and cell migration stimulated by platelet-derived growth factor B." Proc Natl Acad Sci U S A **103**(25): 9536-41.
Rivero-Lezcano, O. M., A. Marcilla, et al. (1995). "Wiskott-Aldrich syndrome protein physically associates with Nck through Src homology 3 domains." Mol Cell Biol **15**(10): 5725-31.
Rivero-Lezcano, O. M., J. H. Sameshima, et al. (1994). "Physical association between Src homology 3 elements and the protein product of the c-cbl proto-oncogene." J Biol Chem **269**(26): 17363-6.
Robertson, J. F., A. R. Dixon, et al. (1992). "Confirmation of a prognostic index for patients with metastatic breast cancer treated by endocrine therapy." Breast Cancer Res Treat **22**(3): 221-7.
Rodrigues, G. A., M. A. Naujokas, et al. (1991). "Alternative splicing generates isoforms of the met receptor tyrosine kinase which undergo differential processing." Mol Cell Biol **11**(6): 2962-70.
Rodrigues, G. A. and M. Park (1993). "Dimerization mediated through a leucine zipper activates the oncogenic potential of the met receptor tyrosine kinase." Mol Cell Biol **13**(11): 6711-22.
Rodrigues, S. P., K. E. Fathers, et al. (2005). "CrkI and CrkII function as key signaling integrators for migration and invasion of cancer cells." Mol Cancer Res **3**(4): 183-94.
Rong, S., L. A. Donehower, et al. (1995). "Met proto-oncogene product is overexpressed in tumors of p53-deficient mice and tumors of Li-Fraumeni patients." Cancer Res **55**(9): 1963-70.
Rosario, M. and W. Birchmeier (2003). "How to make tubes: signaling by the Met receptor tyrosine kinase." Trends Cell Biol **13**(6): 328-35.
Rosenshine, I., S. Ruschkowski, et al. (1996). "A pathogenic bacterium triggers epithelial signals to form a functional bacterial receptor that mediates actin pseudopod formation." Embo J **15**(11): 2613-24.
Rosner, A., K. Miyoshi, et al. (2002). "Pathway pathology: histological differences between ErbB/Ras and Wnt pathway transgenic mammary tumors." Am J Pathol **161**(3): 1087-97.
Royal, I., N. Lamarche-Vane, et al. (2000). "Activation of cdc42, rac, PAK, and rho-kinase in response to hepatocyte growth factor differentially regulates epithelial cell colony spreading and dissociation." Mol Biol Cell **11**(5): 1709-25.
Ruan, W., P. Pang, et al. (1999). "The SH2/SH3 adaptor protein dock interacts with the Ste20-like kinase misshapen in controlling growth cone motility." Neuron **24**(3): 595-605.
Rubin, J. S., A. M. Chan, et al. (1991). "A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte growth factor." Proc Natl Acad Sci U S A **88**(2): 415-9.

Ruusala, A., T. Pawson, et al. (2008). "Nck adapters are involved in the formation of dorsal ruffles, cell migration, and Rho signaling downstream of the platelet-derived growth factor beta receptor." J Biol Chem **283**(44): 30034-44.
Sachs, M., H. Brohmann, et al. (2000). "Essential role of Gab1 for signaling by the c-Met receptor in vivo." J Cell Biol **150**(6): 1375-84.
Sakkab, D., M. Lewitzky, et al. (2000). "Signaling of hepatocyte growth factor/scatter factor (HGF) to the small GTPase Rap1 via the large docking protein Gab1 and the adapter protein CRKL." J Biol Chem **275**(15): 10772-8.
Sarrio, D., S. M. Rodriguez-Pinilla, et al. (2008). "Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype." Cancer Res **68**(4): 989-97.
Sattler, M., Y. B. Pride, et al. (2003). "A novel small molecule met inhibitor induces apoptosis in cells transformed by the oncogenic TPR-MET tyrosine kinase." Cancer Res **63**(17): 5462-9.
Sauer, B. (1998). "Inducible gene targeting in mice using the Cre/lox system." Methods **14**(4): 381-92.
Schade, B., S. H. Lam, et al. (2007). "Distinct ErbB-2 coupled signaling pathways promote mammary tumors with unique pathologic and transcriptional profiles." Cancer Res **67**(16): 7579-88.
Schaeper, U., N. H. Gehring, et al. (2000). "Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses." J Cell Biol **149**(7): 1419-32.
Scheel, C., T. Onder, et al. (2007). "Adaptation versus selection: the origins of metastatic behavior." Cancer Res **67**(24): 11476-9; discussion 11479-80.
Schena, M., D. Shalon, et al. (1995). "Quantitative monitoring of gene expression patterns with a complementary DNA microarray." Science **270**(5235): 467-70.
Schlessinger, J. (1994). "SH2/SH3 signaling proteins." Curr Opin Genet Dev **4**(1): 25-30.
Schmid, C. W. and W. R. Jelinek (1982). "The Alu family of dispersed repetitive sequences." Science **216**(4550): 1065-70.
Schmidt, C., F. Bladt, et al. (1995). "Scatter factor/hepatocyte growth factor is essential for liver development." Nature **373**(6516): 699-702.
Schmidt, L., F. M. Duh, et al. (1997). "Germline and somatic mutations in the tyrosine kinase domain of the MET proto-oncogene in papillary renal carcinomas." Nat Genet **16**(1): 68-73.
Schneider, J., A. Buness, et al. (2004). "Systematic analysis of T7 RNA polymerase based in vitro linear RNA amplification for use in microarray experiments." BMC Genomics **5**(1): 29.
Schutze, K. and G. Lahr (1998). "Identification of expressed genes by laser-mediated manipulation of single cells." Nat Biotechnol **16**(8): 737-42.
Segditsas, S. and I. Tomlinson (2006). "Colorectal cancer and genetic alterations in the Wnt pathway." Oncogene **25**(57): 7531-7.
Sgroi, D. C., S. Teng, et al. (1999). "In vivo gene expression profile analysis of human breast cancer progression." Cancer Res **59**(22): 5656-61.
Shackleton, M., F. Vaillant, et al. (2006). "Generation of a functional mammary gland from a single stem cell." Nature **439**(7072): 84-8.
Shibata, D., D. Hawes, et al. (1992). "Specific genetic analysis of microscopic tissue after selective ultraviolet radiation fractionation and the polymerase chain reaction." Am J Pathol **141**(3): 539-43.

Shipitsin, M., L. L. Campbell, et al. (2007). "Molecular definition of breast tumor heterogeneity." Cancer Cell **11**(3): 259-73.
Silberstein, G. B. and C. W. Daniel (1982). "Glycosaminoglycans in the basal lamina and extracellular matrix of the developing mouse mammary duct." Dev Biol **90**(1): 215-22.
Skrypina, N. A., A. V. Timofeeva, et al. (2003). "Total RNA suitable for molecular biology analysis." J Biotechnol **105**(1-2): 1-9.
Smalley, M. and A. Ashworth (2003). "Stem cells and breast cancer: A field in transit." Nat Rev Cancer **3**(11): 832-44.
Smith, G. H. and G. Chepko (2001). "Mammary epithelial stem cells." Microsc Res Tech **52**(2): 190-203.
Smith, G. H. and D. Medina (1988). "A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland." J Cell Sci **90 (Pt 1)**: 173-83.
Smith, J. M., S. Katz, et al. (1999). "Activation of the Abl tyrosine kinase in vivo by Src homology 3 domains from the Src homology 2/Src homology 3 adaptor Nck." J Biol Chem **274**(39): 27956-62.
Smith, L., P. Underhill, et al. (2003). "Single primer amplification (SPA) of cDNA for microarray expression analysis." Nucleic Acids Res **31**(3): e9.
Smolen, G. A., B. Muir, et al. (2006). "Frequent met oncogene amplification in a Brca1/Trp53 mouse model of mammary tumorigenesis." Cancer Res **66**(7): 3452-5.
Smolen, G. A., R. Sordella, et al. (2006). "Amplification of MET may identify a subset of cancers with extreme sensitivity to the selective tyrosine kinase inhibitor PHA-665752." Proc Natl Acad Sci U S A **103**(7): 2316-21.
Smyth, G. K. (2004). "Linear models and empirical bayes methods for assessing differential expression in microarray experiments." Stat Appl Genet Mol Biol **3**: Article3.
Soriano, J. V., M. S. Pepper, et al. (1995). "Hepatocyte growth factor stimulates extensive development of branching duct-like structures by cloned mammary gland epithelial cells." J Cell Sci **108 (Pt 2)**: 413-30.
Sorlie, T., C. M. Perou, et al. (2001). "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications." Proc Natl Acad Sci U S A **98**(19): 10869-74.
Sorlie, T., R. Tibshirani, et al. (2003). "Repeated observation of breast tumor subtypes in independent gene expression data sets." Proc Natl Acad Sci U S A **100**(14): 8418-23.
Sotiriou, C., S. Y. Neo, et al. (2003). "Breast cancer classification and prognosis based on gene expression profiles from a population-based study." Proc Natl Acad Sci U S A **100**(18): 10393-8.
Stefan, M., A. Koch, et al. (2001). "Src homology 2-containing inositol 5-phosphatase 1 binds to the multifunctional docking site of c-Met and potentiates hepatocyte growth factor-induced branching tubulogenesis." J Biol Chem **276**(5): 3017-23.
Stein, E., U. Huynh-Do, et al. (1998). "Nck recruitment to Eph receptor, EphB1/ELK, couples ligand activation to c-Jun kinase." J Biol Chem **273**(3): 1303-8.
Stingl, J., C. J. Eaves, et al. (1998). "Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast." Differentiation **63**(4): 201-13.

Stingl, J., C. J. Eaves, et al. (2001). "Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue." Breast Cancer Res Treat **67**(2): 93-109.

Stingl, J., P. Eirew, et al. (2006). "Purification and unique properties of mammary epithelial stem cells." Nature **439**(7079): 993-7.

Stoker, M., E. Gherardi, et al. (1987). "Scatter factor is a fibroblast-derived modulator of epithelial cell mobility." Nature **327**(6119): 239-42.

Stommel, J. M., A. C. Kimmelman, et al. (2007). "Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies." Science **318**(5848): 287-90.

Su, Y. C., C. Maurel-Zaffran, et al. (2000). "The Ste20 kinase misshapen regulates both photoreceptor axon targeting and dorsal closure, acting downstream of distinct signals." Mol Cell Biol **20**(13): 4736-44.

Su, Z., P. Xu, et al. (2004). "Single phosphorylation of Tyr304 in the cytoplasmic tail of ephrin B2 confers high-affinity and bifunctional binding to both the SH2 domain of Grb4 and the PDZ domain of the PDZ-RGS3 protein." Eur J Biochem **271**(9): 1725-36.

Subramanian, A., P. Tamayo, et al. (2005). "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles." Proc Natl Acad Sci U S A **102**(43): 15545-50.

Sundquist, M., S. Thorstenson, et al. (1999). "Applying the Nottingham Prognostic Index to a Swedish breast cancer population. South East Swedish Breast Cancer Study Group." Breast Cancer Res Treat **53**(1): 1-8.

Suzuki, R. and H. Shimodaira (2006). "Pvclust: an R package for assessing the uncertainty in hierarchical clustering." Bioinformatics **22**(12): 1540-2.

Taddei, I., M. A. Deugnier, et al. (2008). "Beta1 integrin deletion from the basal compartment of the mammary epithelium affects stem cells." Nat Cell Biol **10**(6): 716-22.

Takayama, H., W. J. La Rochelle, et al. (1996). "Scatter factor/hepatocyte growth factor as a regulator of skeletal muscle and neural crest development." Proc Natl Acad Sci U S A **93**(12): 5866-71.

Takayama, H., W. J. LaRochelle, et al. (1997). "Diverse tumorigenesis associated with aberrant development in mice overexpressing hepatocyte growth factor/scatter factor." Proc Natl Acad Sci U S A **94**(2): 701-6.

Taketo, M., A. C. Schroeder, et al. (1991). "FVB/N: an inbred mouse strain preferable for transgenic analyses." Proc Natl Acad Sci U S A **88**(6): 2065-9.

Tanaka, M., W. Lu, et al. (1997). "Expression of mutated Nck SH2/SH3 adaptor respecifies mesodermal cell fate in Xenopus laevis development." Proc Natl Acad Sci U S A **94**(9): 4493-8.

Taniguchi, T., M. Toi, et al. (1995). "Serum concentrations of hepatocyte growth factor in breast cancer patients." Clin Cancer Res **1**(9): 1031-4.

Tanner, M. M., M. Tirkkonen, et al. (1994). "Increased copy number at 20q13 in breast cancer: defining the critical region and exclusion of candidate genes." Cancer Res **54**(16): 4257-60.

Tavassoli, F. A. a. D., P. (2003). Pathology and Genetics of Tumours of the Breast and Femail Genital Organs. Lyon, Oxford Univ. Press.

Teuliere, J., M. M. Faraldo, et al. (2005). "Targeted activation of beta-catenin signaling in basal mammary epithelial cells affects mammary development and leads to hyperplasia." Development **132**(2): 267-77.
Thiery, J. P. (2002). "Epithelial-mesenchymal transitions in tumour progression." Nat Rev Cancer **2**(6): 442-54.
Tolgay Ocal, I., M. Dolled-Filhart, et al. (2003). "Tissue microarray-based studies of patients with lymph node negative breast carcinoma show that met expression is associated with worse outcome but is not correlated with epidermal growth factor family receptors." Cancer **97**(8): 1841-8.
Topper, Y. J. and C. S. Freeman (1980). "Multiple hormone interactions in the developmental biology of the mammary gland." Physiol Rev **60**(4): 1049-106.
Trusolino, L., A. Bertotti, et al. (2001). "A signaling adapter function for alpha6beta4 integrin in the control of HGF-dependent invasive growth." Cell **107**(5): 643-54.
Tu, Y., D. F. Kucik, et al. (2001). "Identification and kinetic analysis of the interaction between Nck-2 and DOCK180." FEBS Lett **491**(3): 193-9.
Tu, Y., F. Li, et al. (1999). "The LIM-only protein PINCH directly interacts with integrin-linked kinase and is recruited to integrin-rich sites in spreading cells." Mol Cell Biol **19**(3): 2425-34.
Tu, Y., F. Li, et al. (1998). "Nck-2, a novel Src homology2/3-containing adaptor protein that interacts with the LIM-only protein PINCH and components of growth factor receptor kinase-signaling pathways." Mol Biol Cell **9**(12): 3367-82.
Tu, Y., L. Liang, et al. (2001). "Src homology 3 domain-dependent interaction of Nck-2 with insulin receptor substrate-1." Biochem J **354**(Pt 2): 315-22.
Tumbar, T., G. Guasch, et al. (2004). "Defining the epithelial stem cell niche in skin." Science **303**(5656): 359-63.
Turner, N. C. and J. S. Reis-Filho (2006). "Basal-like breast cancer and the BRCA1 phenotype." Oncogene **25**(43): 5846-53.
Turner, N. C., J. S. Reis-Filho, et al. (2007). "BRCA1 dysfunction in sporadic basal-like breast cancer." Oncogene **26**(14): 2126-32.
Twombly, R. (2005). "Failing survival advantage in crucial trial, future of Iressa is in jeopardy." J Natl Cancer Inst **97**(4): 249-50.
Uehara, Y., O. Minowa, et al. (1995). "Placental defect and embryonic lethality in mice lacking hepatocyte growth factor/scatter factor." Nature **373**(6516): 702-5.
Ursini-Siegel, J., W. R. Hardy, et al. (2008). "ShcA signalling is essential for tumour progression in mouse models of human breast cancer." Embo J **27**(6): 910-20.
Ursini-Siegel, J., B. Schade, et al. (2007). "Insights from transgenic mouse models of ERBB2-induced breast cancer." Nat Rev Cancer **7**(5): 389-97.
van 't Veer, L. J., H. Dai, et al. (2002). "Gene expression profiling predicts clinical outcome of breast cancer." Nature **415**(6871): 530-6.
van de Vijver, M. J., Y. D. He, et al. (2002). "A gene-expression signature as a predictor of survival in breast cancer." N Engl J Med **347**(25): 1999-2009.
Van Gelder, R. N., M. E. von Zastrow, et al. (1990). "Amplified RNA synthesized from limited quantities of heterogeneous cDNA." Proc Natl Acad Sci U S A **87**(5): 1663-7.
Verma, R., I. Kovari, et al. (2006). "Nephrin ectodomain engagement results in Src kinase activation, nephrin phosphorylation, Nck recruitment, and actin polymerization." J Clin Invest **116**(5): 1346-59.

Visvader, J. E. and G. J. Lindeman (2006). "Mammary stem cells and mammopoiesis." Cancer Res **66**(20): 9798-801.
Wagner, K. U., K. McAllister, et al. (2001). "Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice." Transgenic Res **10**(6): 545-53.
Wakefield, L. M., G. Thordarson, et al. (2003). "Spontaneous pituitary abnormalities and mammary hyperplasia in FVB/NCr mice: implications for mouse modeling." Comp Med **53**(4): 424-32.
Wei, M., T. A. Grushko, et al. (2005). "BRCA1 promoter methylation in sporadic breast cancer is associated with reduced BRCA1 copy number and chromosome 17 aneusomy." Cancer Res **65**(23): 10692-9.
Weidner, K. M., N. Arakaki, et al. (1991). "Evidence for the identity of human scatter factor and human hepatocyte growth factor." Proc Natl Acad Sci U S A **88**(16): 7001-5.
Weidner, K. M., S. Di Cesare, et al. (1996). "Interaction between Gab1 and the c-Met receptor tyrosine kinase is responsible for epithelial morphogenesis." Nature **384**(6605): 173-6.
Weidner, K. M., M. Sachs, et al. (1993). "The Met receptor tyrosine kinase transduces motility, proliferation, and morphogenic signals of scatter factor/hepatocyte growth factor in epithelial cells." J Cell Biol **121**(1): 145-54.
Weigelt, B., J. L. Peterse, et al. (2005). "Breast cancer metastasis: markers and models." Nat Rev Cancer **5**(8): 591-602.
Welm, A. L., S. Kim, et al. (2005). "MET and MYC cooperate in mammary tumorigenesis." Proc Natl Acad Sci U S A **102**(12): 4324-9.
Welm, B. E., S. B. Tepera, et al. (2002). "Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population." Dev Biol **245**(1): 42-56.
White, D. E., R. D. Cardiff, et al. (2001). "Mammary epithelial-specific expression of the integrin-linked kinase (ILK) results in the induction of mammary gland hyperplasias and tumors in transgenic mice." Oncogene **20**(48): 7064-72.
Woodward, W. A., M. S. Chen, et al. (2005). "On mammary stem cells." J Cell Sci **118**(Pt 16): 3585-94.
Wu, J. M., M. J. Fackler, et al. (2008). "Heterogeneity of breast cancer metastases: comparison of therapeutic target expression and promoter methylation between primary tumors and their multifocal metastases." Clin Cancer Res **14**(7): 1938-46.
Wunderlich, L., A. Goher, et al. (1999). "Requirement of multiple SH3 domains of Nck for ligand binding." Cell Signal **11**(4): 253-62.
Xu, X., W. Qiao, et al. (2001). "Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis." Nat Genet **28**(3): 266-71.
Xu, X., K. U. Wagner, et al. (1999). "Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation." Nat Genet **22**(1): 37-43.
Yamaguchi, H., M. Lorenz, et al. (2005). "Molecular mechanisms of invadopodium formation: the role of the N-WASP-Arp2/3 complex pathway and cofilin." J Cell Biol **168**(3): 441-52.
Yamashita, J., M. Ogawa, et al. (1994). "Immunoreactive hepatocyte growth factor is a strong and independent predictor of recurrence and survival in human breast cancer." Cancer Res **54**(7): 1630-3.

Yanagita, K., K. Matsumoto, et al. (1993). "Hepatocyte growth factor may act as a pulmotrophic factor on lung regeneration after acute lung injury." J Biol Chem **268**(28): 21212-7.

Yang, Y., E. Spitzer, et al. (1995). "Sequential requirement of hepatocyte growth factor and neuregulin in the morphogenesis and differentiation of the mammary gland." J Cell Biol **131**(1): 215-26.

Yano, S., W. Wang, et al. (2008). "Hepatocyte growth factor induces gefitinib resistance of lung adenocarcinoma with epidermal growth factor receptor-activating mutations." Cancer Res **68**(22): 9479-87.

Yant, J., L. Buluwela, et al. (1998). "In vivo effects of hepatocyte growth factor/scatter factor on mouse mammary gland development." Exp Cell Res **241**(2): 476-81.

Yehiely, F., J. V. Moyano, et al. (2006). "Deconstructing the molecular portrait of basal-like breast cancer." Trends Mol Med **12**(11): 537-44.

Yu, K., C. H. Lee, et al. (2004). "Conservation of breast cancer molecular subtypes and transcriptional patterns of tumor progression across distinct ethnic populations." Clin Cancer Res **10**(16): 5508-17.

Yuan, Z. L., Y. J. Guan, et al. (2004). "Central role of the threonine residue within the p+1 loop of receptor tyrosine kinase in STAT3 constitutive phosphorylation in metastatic cancer cells." Mol Cell Biol **24**(21): 9390-400.

Zarnegar, R. and G. Michalopoulos (1989). "Purification and biological characterization of human hepatopoietin A, a polypeptide growth factor for hepatocytes." Cancer Res **49**(12): 3314-20.

Zhao, H., T. Hastie, et al. (2002). "Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis." BMC Genomics **3**(1): 31.

Zhu, H., M. A. Naujokas, et al. (1994). "Tyrosine 1356 in the carboxyl-terminal tail of the HGF/SF receptor is essential for the transduction of signals for cell motility and morphogenesis." J Biol Chem **269**(47): 29943-8.

Zhu, H., M. A. Naujokas, et al. (1994). "Receptor chimeras indicate that the met tyrosine kinase mediates the motility and morphogenic responses of hepatocyte growth/scatter factor." Cell Growth Differ **5**(4): 359-66.

breeding purposes.
MT-met mice : - there's 2 endpoints for the mice on special water : 35 days old and 8 months old.
- Hepatocarcinogenesis : The mice will be sacrificed at 52 weeks of age.
Nude mice : We did not work on this part yet, but once we have established a time frame that is optimal for our experiment (tumor formation or metastasis) we will establish a protocol with specific endpoints (see clinical endpoints section about tumor -bearing mice). Tumor tissue will be collected for histopathological evaluation after sacrificing the mice by cervical dislocation with anaesthesia.

Clinical endpoints :
No adverse effects are anticipated from the mammary manipulations involving the floxed-Met gene. The mice carrying the floxed-Met alleles should be phenotypically normal as are the floxed-lacZ and MMTV-Cre mice. Following MFP injection of the transplanted epithelial cells the recipients will be monitored daily for signs of rejection or adverse effects of the surgery. These include such manifestations as inflammation, swelling, malaise, weight loss >20%, labored breathing, lack of grooming and abnormal defecation (e.g. diarrhea). Regarding the experiments affecting liver, there can be signs of liver dysfunction such as jaundice, ascites and behavioral abnormalities. Should any of these signs be observed, the mice will be sacrificed. Following the 10 days post surgery the clips will be removed and the monitoring of the mice will decrease to twice weekly.

For any cells that have not been investigated for their ability to induce experimental metastases (the growth of cancer cells in the lungs), we will monitor the animals daily, looking for signs of distress or discomfort, weight loss greater than 20%, abnormal defecation, or respiratory distress. In ALL cases for all the procedures, animals exhibiting any of the signs listed above will be sacrificed immediately. Animals will be monitored at least twice per week and tumor-bearing animals will be sacrificed before tumor ulceration or before the tumor reaches a volume of 1.0 cm3.

8. Hazards (click here if none are used: ☐)

a) Are the hazards different from original protocol? (infectious, radioactive, toxic, carcinogen, tumours)

YES ☐ **NO** ☒ **if yes, supply details (material, risks, precautions):**

b) Have the cell lines been tested for human and animal pathogens? YES: ☒ **NO:** ☐ **None used:** ☐

9. Description of Animals to be used in the coming year (only):

Quality Control Assurance: To prevent introduction of infectious diseases into animal facilities, a health status report or veterinary inspection certificate may be required prior to receiving animals from all non-commercial sources or from commercial sources whose animal health status is unknown or questionable. Quarantine and further testing may be required for these animals. *If more than 6 columns are needed, please attach another page*

	Sp/strain 1	Sp/strain 2	Sp/strain 3	Sp/strain 4	Sp/strain 5	Sp/strain 6
Species	mouse	mouse	mouse	mouse	mouse	mouse
Supplier/Source	CR	In House (IH)	IH	IH	Dr W. Muller	CR
Strain	FVB	c-met (floxed Met)	mmtv Cre	GT Rosa	Neu / floxed-lacZ (activated ErbB2)	CD1 nu/nu
Sex	M/F	M/F	M/F	M/F	M	F
Age/Wt	3-12 weeks	varies	varies	varies	6-12 weeks	3-12 weeks
# To be purchased	max of 20	0	0	0	2	72

# Produced by in-house breeding	max of 100	max of 300	max of 200	max of 200	max of 100	0
# Other (e.g.field studies)	none	none	none	none	none	none
TOTAL# /YEAR	**120**	**300**	**200**	**200**	**102**	**72**

10. Justification of Animal Numbers:

BASED ON THE EXPERIMENTAL OBJECTIVES OF THE PROJECT, describe the number of animals required for one year. Include information on experimental and control groups, # per group, and failure rates. For breeding, specify how many adults are used, number of offspring produced, and how many offspring are used in experimental procedures. The arithmetic explaining how the total of animals for each column in the table above is calculated should be made clear.

Strain 1 - FVB
120 FVB mice is the number needed. The FVB mice are used for the mammary fat pad surgeries and the maintaining for all our lines. The maximum to be purchased is 20 because we will buy new mice only to refresh the genetic of our FVB colony (to avoid cosanguinity). All produced FVBs (100 mice) will be use such as control, as breeder or as a MFP surgery patient, none will be discarded at weaning unless it is sick.
20 + 100 = 120

Strain 2 - c-met
Last year, we accomplished the 7 back-crosses to FVB Background. Now, we have operational heterozygotes (+/-) c-met mice. Our experiments requires homozygotes (-/-) so from now on we will keep only the -/- mice. In the best situatuion, about 30% of the pups of a litter are -/- (often, it is only 10-20%). We need to produces several litters at begining of the year (here goes the first 150 mice but about 30 will be kept). From those 30 -/- mice, we will get progeny (30 mice) as backup for the colony. 10 mice will be used to cross with the Cre mice (see strain 3 for details). 10 mice will be used to cross with GT Rosa mice (See strain 4 for details). The last 100 mice goes for the experiments spreaded over the year. 10 or 20 mice are required at the time. They are euthanised at 12 or 16 weeks old and their mammary fat pads are harvested.
150 + 30 + 10 + 10 + 100 = 300

Strain 3 - mmtv Cre
100 mice will be used for the pure Cre colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with Cre mice it will be in the second half of the year. From the crossing between mmtv Cre mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 4 - GT Rosa
100 mice will be used for the pure GT Rosa colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with GT Rosa mice it will be in the second half of the year. From the crossing between GT Rosa mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 5 - Neu
The process of getting those mice is pending. The 2 males we got last year were negative. We will get replacements shortly. The number of mice requested reflects the requirement for identification and maintenance of the transgenic stocks. The number written is the number of mice to be generated, however, the majority of the mice will not be kept. Following genotyping between day 14 and 20, the unsuitable mice will be sacrificed at weaning.

7. Animal Data

7 a) Please justify the need for live animals versus alternate methods (e.g. tissue culture, computer simulation)

We will perform several of the assays in tissue culture, however, mammary gland development and the development of mammary tumors as well as the ability of of cells to colonize and metastasize cannot be reproduced by alternative methods. It is critical to study the mammary epithelium within the physiological environment that subjects it to the hormonal fluctuations that are responsible for the development and differentiation of the mammary epithelium. Correct development and differentiation of the mammary ductal system as well as accurate behavior of tumorigenic cell lines cannot occur outside of the animal.

7 b) Describe the characteristics of the animal species selected that justifies its use in the proposed study (consider characteristics such as body size, species, strain, data from previous studies or unique anatomic/physiological features)

These experiments are based on validated mouse models for mammary development, tumor progression and metastasis.

7 c) Description of animals

Quality Control Assurance: To prevent introduction of infectious diseases into animal facilities, a health status report or veterinary inspection certificate may be required prior to receiving animals from all non-commercial sources or from commercial sources whose animal health status is unknown or questionable. Quarantine and further testing may be required for these animals.

If more than 6 columns are needed, please attach another page

	Sp/strain 1	Sp/strain 2	Sp/strain 3	Sp/strain 4	Sp/strain 5	Sp/strain 6
Species	mouse	mouse	mouse	mouse	mouse	mouse
Supplier/Source	CR	In House (IH)	IH	IH	Dr W. Muller	CR
Strain	FVB	c-met (floxed Met)	mmtv Cre	GT Rosa	Neu / floxed-lacZ (activated ErbB2) Protocol #4613	CD1 nu/nu
Sex	M/F	M/F	M/F	M/F	M	F
Age/Wt	3-12 weeks	varies	varies	varies	6-12 weeks	3-12 weeks
# To be purchased	max of 20	0	0	0	2	72
# Produced by in-house breeding	max of 100	max of 300	max of 200	max of 200	max of 100	0
# Other (e.g.field studies)	none	none	none	none	none	none
#needed at one time	various	various	various	various	various	various
# per cage	5	5	5	5	5	5
TOTAL# /YEAR	120	300 (max)	200 (max)	200	102	72

7 d) Justification of Animal Usage: BASED ON THE EXPERIMENTAL OBJECTIVES OF THE PROJECT, describe the number of animals required for one year. Include information on experimental and control groups, # per group, and failure rates. For breeding, specify how many adults are used, number of offspring produced, and how many offspring are used in experimental procedures. Use the table below when applicable. The arithmetic explaining how the total of animals for each column in the table above is calculated should be made clear. *(Space will expand as needed)*

Strain 1 - FVB
120 FVB mice is the number we need. It is double than last year because we expanded. The FVB mice are used for the mammary fat pad surgeries and the maintaining for all our lines. The maximum to be purchased is 20 because we will buy new mice only to refresh the genetic of our FVB colony (to avoid cosanguinity). All of the FVB produced (100 mice) has an utility (as control, as breeder or as a MFP surgery patient), none will be discarded at weaning unless it is sick.
20 + 100 = 120

Strain 2 - c-met
We have operational heterozygotes (+/-) c-met mice. Our experiments requires homozygotes (-/-) so from now on we will keep only the -/- mice. In the best situatuion, about 30% of the pups of a litter are -/- (often, it is only 10-20%). We need to produces several litters at begining of the year (here goes the first 150 mice but about 30 will be kept). From those 30 -/- mice, we will get progeny (30 mice) as backup for the colony. 10 mice will be used to cross with the Cre mice (see strain 3 for details). 10 mice will be used to cross with GT Rosa mice (See strain 4 for details). The last 100 mice goes for the experiments spreaded over the year. 10 or 20 mice are required at the time. They are euthanised at 12 or 16 weeks old and their mammary fat pads are harvested.
150 + 30 + 10 + 10 + 100 = 300

Strain 3 - mmtv Cre
100 mice will be used for the pure Cre colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with Cre mice it will be in the second half of the year. From the crossing between mmtv Cre mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 4 - GT Rosa
100 mice will be used for the pure GT Rosa colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with GT Rosa mice it will be in the second half of the year. From the crossing between GT Rosa mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 5 - Neu
The process of getting those mice is pending. The 2 males we got last year were negative. We will get replacements shortly. The number of mice requested reflects the requirement for identification and maintenance of the transgenic stocks. The number written is the number of mice to be generated, however, the majority of the mice will not be kept. Following genotyping between day 14 and 20, the unsuitable mice will be sacrificed at weaning.
2 + 100 = 102

Theory related
The homozygous floxed-Met mice will also be bred with the floxed-lacZ mice in order to create a line of mice that carries two alleles of both the floxed-Met and floxed-lacZ so that when these mice are crossed with the MMTV-Cre mice the Met RTK will be knocked out in the mammary epithelium and the lacZ reporter gene will become functional in the epithelium. This will allow the knock out cells to be identified upon tissue harvest.

Strain 6 - CD1 nu/nu
In order to establish the role of Crk, a downstream signaling molecule from the Met RTK, in the metastatic and tumorigenic phenotype of human breast cancer cell lines, human breast cancer cell lines (MCF7 and T47D) will be transfected with plasmids to establish RNAi ablation of Crk expression. Cell lines where Crk is successfully ablated will be injected once subcutaneously or via tail vein into nude mice. For each injection 6 mice will be used. Total 12 mice per cell line repeated 3 times 36 mice per cell line.
(3 x 12) + (3 x 12) = 72
SC IV

7d table) ***The following table may help you explain the animal numbers listed in the 7c table:***

(Table will expand as needed)	**Sp/strain 1**	**Sp/strain 2**	**Sp/strain 3**	**Sp/strain 4**	**Sp/strain 5**	**Sp/strain 6**
Test agents or procedures	MCF7 cells	T47D cells				
# of animals per group	12	12				
Dosage / route of administration	SC	SC				

# of endpoints	3	3				
Other variables (sex,genotypes...)	N/A	N/A				
Total number of animals per year	**36**	**36**				

8. Animal Husbandry and Care

8 a) If projects involves non-standard cages, diet and/or handling, please specify

Filter top cages, sterilized food and water, SPF facility
*special water : we will provide the animal facility with special water (zinc sulfate added), they will have to sterilize their bottles and provide it to identified cages.

8 b) Is there any component to the proposed procedures which will result in immunosuppression or decreased immune function (e.g. stress, radiation, steroids, chemotherapeutics, genetic modification of the immune system)?

NO ☒ YES ☐ **if yes, specify:**

8 c) Indicate area(s) where animal use procedures will be conducted:

Building: RVH **Room:** H3

Indicate area(s) all facilities where animals will be housed:

Building: RVH **Room:** H3

If animal housing and animal use are in different locations, briefly describe procedures for transporting animals:
Applicable only to mice Crk I which will be generated at the ARC by Dr Tremblay's lab.
Animals will be transported by Animal Resources Centre service

9. Standard Operating Procedures (SOPs)

Complete this section if you plan to use any of the UACC SOPs listed below. IT IS UACC POLICY THAT THESE SOPS BE USED WHEN APPLICABLE. Any proposed variation of the SOPs must be described and justified. The Standard Operating Procedures can be found at the UACC website at www.mcgill.ca/research/compliance/animal/procedures . The completed and signed SOP form must be attached to the protocol.

Check all SOPs that will be used:

Blood Collection UACC#1	☐	**Collection of Amphibian Oocytes UACC#9**	☐
Anaesthesia in rodents UACC#2	☒	**Rodent Survival Surgery UACC#10**	☐
Analgesia in rodents UACC#3	☐	**Anaesthesia & Analgesia Neonatal Rodents UACC#11**	☐
Breeding transgenics/knockouts UACC#4	☒	**Stereotaxic Survival Surgery in Rodents UACC#12**	☐
Transgenic Generation UACC#5	☒	**Field Studies Form**	☐
Knockout/in Generation UACC#6	☐	**Phenotype Disclosure Form**	☐
Production of Monoclonal Antibodies UACC#7	☐	**Other, specify:**	☐
Production of Polyclonal Antibodies UACC#8	☐		☐

10. Description of Procedures

10 a) . IF A PROCEDURE IS COVERED BY AN SOP, WRITE "AS PER SOP", NO FURTHER DETAIL IS REQUIRED.

FOR EACH EXPERIMENTAL GROUP, DESCRIBE ALL PROCEDURES AND TECHNIQUES, WHICH ARE NOT PART OF THE SOPs, IN THE ORDER IN WHICH THEY WILL BE PERFORMED – surgical procedures, immunizations, behavioural tests, immobilization and restraint, food/water deprivation, requirements for post-operative care, sample collection, substance administration, special monitoring, etc Appendix 2 of the Guidelines (www.mcgill.ca/research/compliance/animal/forms) provides a sample list of points that should be addressed in this section.

Transplantation of mammary epithelial cells to a cleared mammary fat pad (MFP): Mammary fat pads will be harvested from floxed Met mice and the epithelial cells isolated in vitro. The epithelial cells will be infected in vitro with a replication incompetent adenovirus expressing cre recombinase. The treated epithelial cells will then be suspended in phosphate buffered saline and injected into a cleared MFP. The injection of the epithelial cells requires anaesthetizing a 3 week old female animal

with the rodent cocktail given intramuscularly at a dose of 0.1mL/100g body weight. Following confirmation that the animal is insensible an inverted "Y" incision is made in the skin of the abdomen allowing for exposure of the the number four MFP. The cells to be tranplanted to the MFP are injected into a distal portion of the MFP and the proximal portion of the gland is removed using a cauterizer, thus removing any endogenous mammary epithelium. The incisions is closed using clips which are removed 10 days later. The procedure should take no longer than ten minutes starting from the first incision. Once the procedure is complete the mouse will be placed near a heat lamp (3-4 feet depending on bulb strength) and will be monitored until partial consciousness. For the 10 days following surgery the animals will be observed daily for signs of infection and distress and if either of these occur the animal will be sacrificed.

Tumorigenesis and metastasis assay: To assay the ability of Met-expressing cells to form tumors or experimental metastases mice will be injected with tumor cells (10^4-10^7 in PBS or DMEM) either intravenously via the tail vein or subcutaneously.

Generation of transgenic animals: Performed as a fee for service at the McIntyre Transgenic Facility (Dr. Michel Tremblay's protocol #4437).

10 b) Experimental endpoint – for each experimental group indicate survival time

Mice receiving the epithlial transplant will be sacrificed at 13 weeks of age; this is 10 weeks following the MFP injection. Floxed-Met mice crossed with the MMTV-Cre expressing mice will either be sacrificed at 10 weeks of age in order to observe any mammary gland developmental defects or they will be bred and sacrificed mid-pregnancy or during lactation to observe any differentiation or lactational defects incurred by the loss of Met RTK expression. There will be no groups of aging mice, however, a few homozygous mice of floxed-Met, mmtv-cre, floxed-lacZ and the combined floxed Met/floxed lacZ will be kept for breeding purposes.

Nude mice : Once we have established a time frame that is optimal for our experiment (tumor formation or metastasis) we will establish a protocol with specific endpoints (see clinical endpoints section about tumor -bearing mice). Tumor tissue will be collected for histopathological evaluation after sacrificing the mice by cervical dislocation with anaesthesia.

10 c) Clinical endpoint – describe the conditions, complications, and criteria (e.g. >20% weight loss, maximum tumour size, vocalizing, lack of grooming) that would lead to euthanasia of an animal before the expected completion of the experiment (specify per species and project if multiple projects involved)

No adverse effects are anticipated from the mammary manipulations involving the floxed-Met gene. The mice carrying the floxed-Met alleles should be phenotypically normal as are the floxed-lacZ and MMTV-Cre mice. Following MFP injection of the transplanted epithelial cells the recipients will be monitored daily for signs of rejection or adverse effects of the surgery. These include such manifestations as inflammation, swelling, malaise, weight loss >20%, labored breathing, lack of grooming and abnormal defecation (e.g. diarrhea). Should any of these signs be observed, the mice will be sacrificed. Following the 10 days post surgery the clips will be removed and the monitoring of the mice will decrease to twice weekly.

For any cells that have not been investigated for their ability to induce experimental metastases (the growth of cancer cells in the lungs), we will monitor the animals daily, looking for signs of distress or discomfort, weight loss greater than 20%, abnormal defecation, or respiratory distress. In ALL cases for all the procedures, animals exhibiting any of the signs listed above will be sacrificed immediately. Animals will be monitored at least twice per week and tumor-bearing animals will be sacrificed before tumor ulceration or before the tumor reaches a volume of 1.0 cm3.

Frequency of monitoring: Twice weekly or daily following surgery. Daily following injection of a cell line that has not been investigated for its ability to induce experimental metastases.

10 d) Specify person(s) who will be responsible for animal monitoring and post-procedural care ***(must also be listed in section 4)***

Name: Anie Monast	**Phone #:** 483-1481

10 e) Pre-Anesthetic/Anaesthetic/Analgesic Agents: List all drugs that will be used to minimize pain, distress or discomfort. If covered in an SOP, write "*As per SOP*", no further details is required. *(Table will expand as needed)*

Species	Agent	Dosage (mg/kg)	Total volume(ml) per administration	Route	Frequency/Duration
mouse	rodent cocktail Ket / xyl / ace	50 / 5 / 1 mg/kg (0.1mL/100g body weight)	0.02 mL	IM	1 dose

*See SOP #2 for more details

10 f) Administration of ALL other substances: List all non-anaesthetic agents under study in the experimental component of the protocol, including but not limited to drugs, infectious agents, viruses. If covered in an SOP, write "*As per SOP*", no further details is required. *(Table will expand as needed)*

Species	Agent	Dosage (mg/kg)	Total volume(ml) per administration	Route	Frequency/Duration
mouse	transplanted mammary epithelial cells infected with adenovirus expressing Cre recombinase	N/A	0.05 mL	MFP injection	1 dose

10 g) Method of Euthanasia

Specify Species	
	☐ **Anaesthetic overdose, list agent/dose/route:**
	☐ **Exsanguination with anaesthesia, list agent/dose/route:**
	☐ **Decapitation without anaesthesia *** ☐ **Decapitation with anesthesia, list agent/dose/route (including CO_2):**
mouse	☐ **Cervical dislocation without anaesthesia *** ☒ **Cervical dislocation with anaesthesia, list agent/dose/route (including CO_2):** Isoflurane or CO_2 /to effect/ inhalation
	☐ **CO_2 chamber only**
	☐ **Other, specify:**
	☐ **Not applicable, explain:**

*** For physical method of euthanasia without anaesthesia, please justify:**

11. Category of Invasiveness: B ☐ C ☐ D ☒ E ☐

Categories of Invasiveness (from the CCAC *Categories of Invasiveness in Animal Experiments*). Please refer to this document for a more detailed description of categories.

Category A: Studies or experiments on most invertebrates or no entire living material.

Category B: Studies or experiments causing little or no discomfort or stress. ***These might include holding animals captive, injection, percutaneous blood sampling, accepted euthanasia for tissue harvest, acute non-survival experiments in which the animals are completely anaesthetized.***

Category C: Studies or experiments involving minor stress or pain of short duration. ***These might include cannulation or catheterizations of blood vessels or body cavities under anaesthesia, minor surgery under anaesthesia, such as biopsy; short periods of restraint, overnight food and/or water deprivation which exceed periods of abstinence in nature; behavioural experiments on conscious animals that involve short-term stressful restraint.***

Category D: Studies or experiments that involve moderate to severe distress or discomfort. ***These might include major surgery under anaesthesia with subsequent recovery, prolonged (several hours or more) periods of physical restraint; induction of behavioural stresses, immunization with complete Freund's adjuvant, application of noxious stimuli, procedures that produce pain, production of transgenics (in accordance with University policy).***

Category E: Procedures that involve inflicting severe pain, near, at or above the pain threshold of unanaesthetized, conscious animals. ***Not confined to but may include exposure to noxious stimuli or agents whose effects are unknown; exposure to drugs or chemicals at levels that (may) markedly impair physiological systems and which cause death, severe pain or extreme distress or physical trauma on unanaesthetized animals.*** **According to University policy, E level studies are not permitted.**

12. Potential Hazards to Personnel and Animals It is the responsibility of the investigator to obtain the necessary Biohazard and/or Radiation Safety permits before this protocol is submitted for review.

A copy of these certificates must be attached, if applicable.

No hazardous materials will be used in this study: ☐

12 a) Indicate which of the following will be used in animals:

☐ **Toxic chemicals** ☐ **Radioisotopes** ☐ **Carcinogens**
☒ **Infectious agents (includes vectors)** ☐ **Transplantable tumours**

12 b) Complete the following table for each agent to be used (use additional page as required):

Agent name	adenovirus Cre		
Dosage	50 uL		
Route of administration	in vitro cell infection injected into the MFP		
Frequency of administration	once		
Duration of administration	one day		
Number of animals involved	10		
Survival time after administration	varies		

12 c) After administration the animals will be housed in:

☒ **the animal care facility** ☐ **laboratory under supervision of laboratory personnel**

Please note that cages must be appropriately labeled at all times.

12 d) Describe potential health risk (s) to humans or animals:

All injections will be performed in a laminar flow hood. Since the Cre-expressing adenovirus is non-replicating, it will not be transmitted by infected cells. Should someone come in direct contact with the virus it should have no effect since the cre-recombinase induces recombination of segments of DNA surrounded by inserted floxP sequences which no human cells should carry in their genomes.

12 e) Describe measures that will be used to reduce risk to the environment and all project and animal facility personnel:

Barrier methods such as the use of gloves, masks and gowns (level 2 protection) will be used while handling the adenovirus and the cells that have been exposed to the adenovirus. The infections will take place inside a flow hood in H5 and all surfaces potentially exposed to the adenovirus will be cleaned with a disinfecting solution. As stated above, the adenovirus expressing Cre is non-replicative, thus any of the mammary cells infected in vitro with the adenovirus Cre will not produce virus thus there should be no superinfection in the mammary fat pad nor is there any risk of infection of the animal care personnel.

13. Reviewer's Modifications (to be completed by ACC only): The Animal Care Committee has made the following modification(s) to this animal use procedure protocol during the review process. Please make these changes to your copy and comply with the recommended changes as a condition of approval.

Guidelines for

7. Animal Data

7 c) Description of animals

Quality Control Assurance: To prevent introduction of infectious diseases into animal facilities, a health status report or veterinary inspection certificate may be required prior to receiving animals from all non-commercial sources or from commercial sources whose animal health status is unknown or questionable. Quarantine and further testing may be required for these animals.

If more than 6 columns are needed, please attach another page

	Sp/strain 7	Sp/strain 8	Sp/strain 9	Sp/strain 10	Sp/strain 11	Sp/strain 12
Species	mouse	mouse	mouse	mouse	mouse	mouse
Supplier/Source	In House (IH)	IH	IH	IH	IH	IH
Strain	Crk I	Crk II	Met wt	Met Y1003 F	Met M1268T	Met Y1003F/M1268T
Sex	M/F	M/F	M/F	M/F	M/F	M/F
Age/Wt	varies	varies	varies	varies	varies	varies
# To be purchased	0	0	0	0	0	0
# Produced by in-house breeding	300	300	60	60	60	100
# Other (e.g.field studies)	none	none	none	none	none	none
#needed at one time	various	various	various	various	various	various
# per cage	5	5	5	5	5	5
TOTAL# /YEAR	300	300	60	60	60	100

7 d) Justification of Animal Usage: BASED ON THE EXPERIMENTAL OBJECTIVES OF THE PROJECT, describe the number of animals required for one year. Include information on experimental and control groups, # per group, and failure rates. For breeding, specify how many adults are used, number of offspring produced, and how many offspring are used in experimental procedures. Use the table below when applicable. The arithmetic explaining how the total of animals for each column in the table above is calculated should be made clear. ***(Space will expand as needed)***

Strains 7 and 8 - Crk I and II

The involvment of the Crk adapter protein in mammary tumorigenesis and mammary development will be observed in transgenic mice expressing Crk I or Crk II adapter proteins under the control of the mammary specific MMTV promoter.MMTV/Crk transgenic MICE WILL BE GENERATED as a fee for service at the McIntyre Transgenic Core facility and mice will be transferred to the H3 facility by the staff of the McGill H3 facility. In the H3 facility the mice will be bred and progeny observed for tumor formation by palpation or sacrifice. At the moment, we are getting the Crk I mice. The Crk II mice are generated and started to breed in our animal facility. The number of mice requested reflects the necessity of identification and maintenance of transgenic stocks (approx. a max of 300 mice). We did obtain 13 transgenic animals from 30 injections for the Crk I construct. We estimate the production of crk I mice as a maximum of 300. We will initially breed 10 CrkII and 10 CrkI transgenics and F1s will be examined for transgene expression. Total of 10 male/female mice for breeding purposes. Two or three lines for each transgene will be selected and kept for maintenance of each line.

~~MMTV/CrkII transgenic (5 founders) MMTV/CrkI transgenic (5 founders)~~

Total 10 transgenic breeding lines.

Strain 9 - 10 - 11 - 12 : Met wt, Y1003F, M1268T and Y1003F/M1268T

Regarding the following strains : Met wt, Met Y1003F and Met M1268T, we are completing the required number of mice to get a portion of the colony as an aging colony and an other portion to breed for tissue colletion and to maintain it as well. (40 aging mice + 60 young mice = 100 mice per line to keep). The double mutant mice (Met Y1003F/M1268T) requires to produce more mice because we need to determine which of the ten founders transmit the transgene and get its progeny expressing the transgene in their mammary epithelium. For that purpose, some breeding females will be sacrifice at lactation stage (since the expression is present in the lactating mammary glands), and we need to breed FVB to get Foster parents. As other Met strain mentionned earlier, the double mutant strain requires aging and maintaining mice. So, the maximum to keep at all time for any of these lines is 200 mice (including aging mice from last year).

Strain 1 - FVB
120 FVB mice is the number we need. It is double than last year because we expanded. The FVB mice are used for the mammary fat pad surgeries and the maintaining for all our lines. The maximum to be purchased is 20 because we will buy new mice only to refresh the genetic of our FVB colony (to avoid cosanguinity). All of the FVB produced (100 mice) has an utility (as control, as breeder or as a MFP surgery patient), none will be discarded at weaning unless it is sick.
20 + 100 = 120

Strain 2 - c-met
We have operational heterozygotes (+/-) c-met mice. Our experiments requires homozygotes (-/-) so from now on we will keep only the -/- mice. In the best situatuion, about 30% of the pups of a litter are -/- (often, it is only 10-20%). We need to produces several litters at begining of the year (here goes the first 150 mice but about 30 will be kept). From those 30 -/- mice, we will get progeny (30 mice) as backup for the colony. 10 mice will be used to cross with the Cre mice (see strain 3 for details). 10 mice will be used to cross with GT Rosa mice (See strain 4 for details). The last 100 mice goes for the experiments spreaded over the year. 10 or 20 mice are required at the time. They are euthanised at 12 or 16 weeks old and their mammary fat pads are harvested.
150 + 30 + 10 + 10 + 100 = 300

Strain 3 - mmtv Cre
100 mice will be used for the pure Cre colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with Cre mice it will be in the second half of the year. From the crossing between mmtv Cre mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 4 - GT Rosa
100 mice will be used for the pure GT Rosa colony maintenance. All negative mice will be euthanised at weaning. By the time we get strain 2 ready for breeding with GT Rosa mice it will be in the second half of the year. From the crossing between GT Rosa mice and c-met mice, we estimate a production of 100 mice. No experiment is plan for these mice until next year. We are producing them this year.
100+100= 200

Strain 5 - Neu
The process of getting those mice is pending. The 2 males we got last year were negative. We will get replacements shortly. The number of mice requested reflects the requirement for identification and maintenance of the transgenic stocks. The number written is the number of mice to be generated, however, the majority of the mice will not be kept. Following genotyping between day 14 and 20, the unsuitable mice will be sacrificed at weaning.
2 + 100 = 102

Theory related
The homozygous floxed-Met mice will also be bred with the floxed-lacZ mice in order to create a line of mice that carries two alleles of both the floxed-Met and floxed-lacZ so that when these mice are crossed with the MMTV-Cre mice the Met RTK will be knocked out in the mammary epithelium and the lacZ reporter gene will become functional in the epithelium. This will allow the knock out cells to be identified upon tissue harvest.

Strain 6 - CD1 nu/nu
In order to establish the role of Crk, a downstream signaling molecule from the Met RTK, in the metastatic and tumorigenic phenotype of human breast cancer cell lines, human breast cancer cell lines (MCF7 and T47D) will be transfected with plasmids to establish RNAi ablation of Crk expression. Cell lines where Crk is successfully ablated will be injected once subcutaneously or via tail vein into nude mice. For each injection 6 mice will be used. Total 12 mice per cell line repeated 3 times 36 mice per cell line.
(3 x 12) + (3 x 12) = 72
SC IV
NOTE : Any detrimental phenotype will be reported to the FACC and ARC veterinarian at the time they are observed.

7d table) The following table may help you explain the animal numbers listed in the 7c table:

(Table will expand as needed)	Sp/strain 1	Sp/strain 2	Sp/strain 3	Sp/strain 4	Sp/strain 5	Sp/strain 6
Test agents or procedures	MCF7 cells	T47D cells				
# of animals per group	12	12				

Dosage / route of administration	SC	SC				
# of endpoints	3	3				
Other variables (sex,genotypes...)	N/A	N/A				
Total number of animals per year	36	36				

8. Animal Husbandry and Care

8 a) If projects involves non-standard cages, diet and/or handling, please specify

Filter top cages, sterilized food and water, SPF facility
*special water : we will provide the animal facility with special water (zinc sulfate added), they will have to sterilize their bottles and provide it to identified cages. The ZnSO4 is added in order to boost the transgene expression in the tissues of selected groups of MT-met mice.

8 b) Is there any component to the proposed procedures which will result in immunosuppression or decreased immune function (e.g. stress, radiation, steroids, chemotherapeutics, genetic modification of the immune system)?

NO ☒ YES ☐ **if yes, specify:**

8 c) Indicate area(s) where animal use procedures will be conducted:

Building: RVH **Room:** H3

Indicate area(s) all facilities where animals will be housed:

Building: RVH **Room:** H3

If animal housing and animal use are in different locations, briefly describe procedures for transporting animals:
Applicable only to mice Crk I which will be generated at the ARC by Dr Tremblay's lab.
Animals will be transported by Animal Resources Centre service

9. Standard Operating Procedures (SOPs)

Complete this section if you plan to use any of the UACC SOPs listed below. IT IS UACC POLICY THAT THESE SOPS BE USED WHEN APPLICABLE. Any proposed variation of the SOPs must be described and justified. The Standard Operating Procedures can be found at the UACC website at www.mcgill.ca/research/compliance/animal/procedures . The completed and signed SOP form must be attached to the protocol.

Check all SOPs that will be used:

Blood Collection UACC#1	☐	**Collection of Amphibian Oocytes UACC#9**	☐
Anaesthesia in rodents UACC#2	☒	**Rodent Survival Surgery UACC#10**	☐
Analgesia in rodents UACC#3	☐	**Anaesthesia & Analgesia Neonatal Rodents UACC#11**	☐
Breeding transgenics/knockouts UACC#4	☒	**Stereotaxic Survival Surgery in Rodents UACC#12**	☐
Transgenic Generation UACC#5	☒	**Field Studies Form**	☐
Knockout/in Generation UACC#6	☐	**Phenotype Disclosure Form**	☐
Production of Monoclonal Antibodies UACC#7	☐	**Other, specify:**	☐
Production of Polyclonal Antibodies UACC#8	☐		☐

10. Description of Procedures

10 a) . IF A PROCEDURE IS COVERED BY AN SOP, WRITE "AS PER SOP", NO FURTHER DETAIL IS REQUIRED.

FOR EACH EXPERIMENTAL GROUP, DESCRIBE ALL PROCEDURES AND TECHNIQUES, WHICH ARE NOT PART OF THE SOPs, IN THE ORDER IN WHICH THEY WILL BE PERFORMED – surgical procedures, immunizations, behavioural tests, immobilization and restraint, food/water deprivation, requirements for post-operative care, sample collection, substance administration, special monitoring, etc Appendix 2 of the Guidelines (www.mcgill.ca/research/compliance/animal/forms) provides a sample list of points that should be addressed in this section.

Transplantation of mammary epithelial cells to a cleared mammary fat pad (MFP): Mammary fat pads will be harvested from floxed Met mice and the epithelial cells isolated in vitro. The epithelial cells will be infected in vitro with a replication

Species	Agent	Dosage (mg/kg)	Total volume(ml) per administration	Route	Frequency/Duration
mouse	rodent cocktail Ket / xyl / ace	50 / 5 / 1 mg/kg (0.1mL/100g body weight)	0.02 mL	IM	1 dose

*See SOP #2 for more details

10 f) Administration of ALL other substances: List all non-anaesthetic agents under study in the experimental component of the protocol, including but not limited to drugs, infectious agents, viruses. If covered in an SOP, write "*As per SOP*", no further details is required. *(Table will expand as needed)*

Species	Agent	Dosage (mg/kg)	Total volume(ml) per administration	Route	Frequency/Duration
mouse	transplanted mammary epithelial cells infected with adenovirus expressing Cre recombinase	N/A	0.05 mL	MFP injection	1 dose
mouse	Zinc sulfate		drinking water	per os	all the time from 3 weeks of age until death

10 g) Method of Euthanasia

Specify Species	
	☐ **Anaesthetic overdose, list agent/dose/route:**
	☐ **Exsanguination with anaesthesia, list agent/dose/route:**
	☐ **Decapitation without anaesthesia *** ☐ **Decapitation with anesthesia, list agent/dose/route (including CO_2):**
mouse	☐ **Cervical dislocation without anaesthesia *** ☒ **Cervical dislocation with anaesthesia, list agent/dose/route (including CO_2):** Isoflurane or CO_2 /to effect/ inhalation
	☐ **CO_2 chamber only**
	☐ **Other, specify:**
	☐ **Not applicable, explain:**

*** For physical method of euthanasia without anaesthesia, please justify:**

11. Category of Invasiveness: B ☐ C ☐ D ☒ E ☐

Categories of Invasiveness (from the CCAC *Categories of Invasiveness in Animal Experiments*). Please refer to this document for a more detailed description of categories.

Category A: Studies or experiments on most invertebrates or no entire living material.

Category B: Studies or experiments causing little or no discomfort or stress. *These might include holding animals captive, injection, percutaneous blood sampling, accepted euthanasia for tissue harvest, acute non-survival experiments in which the animals are completely anaesthetized.*

Category C: Studies or experiments involving minor stress or pain of short duration. *These might include cannulation or catheterizations of blood vessels or body cavities under anaesthesia, minor surgery under anaesthesia, such as biopsy; short periods of restraint, overnight food and/or water deprivation which exceed periods of abstinence in nature; behavioural experiments on conscious animals that involve short-term stressful restraint.*

Category D: Studies or experiments that involve moderate to severe distress or discomfort. *These might include major surgery under anaesthesia with subsequent recovery, prolonged (several hours or more) periods of physical restraint; induction of behavioural stresses, immunization with complete Freund's adjuvant, application of noxious stimuli, procedures that produce pain, production of transgenics (in accordance with University policy).*

Category E: Procedures that involve inflicting severe pain, near, at or above the pain threshold of unanaesthetized, conscious animals. *Not confined to but may include exposure to noxious stimuli or agents whose effects are unknown; exposure to drugs or chemicals at levels that (may) markedly impair physiological systems and which cause death, severe pain or extreme distress or physical trauma on unanaesthetized animals.* According to University policy, E level studies are not permitted.

12. Potential Hazards to Personnel and Animals It is the responsibility of the investigator to obtain the necessary Biohazard and/or Radiation Safety permits before this protocol is submitted for review.
A copy of these certificates must be attached, if applicable.

No hazardous materials will be used in this study: ☐

12 a) Indicate which of the following will be used in animals:

☐ Toxic chemicals ☐ Radioisotopes ☐ Carcinogens
☒ Infectious agents (includes vectors) ☐ Transplantable tumours

12 b) Complete the following table for each agent to be used (use additional page as required):

Agent name	adenovirus Cre		
Dosage	50 uL		
Route of administration	in vitro cell infection injected into the MFP		
Frequency of administration	once		
Duration of administration	one day		
Number of animals involved	10		
Survival time after administration	varies		

12 c) After administration the animals will be housed in:

☒ **the animal care facility** ☐ **laboratory under supervision of laboratory personnel**

Please note that cages must be appropriately labeled at all times.

12 d) Describe potential health risk (s) to humans or animals:

All injections will be performed in a laminar flow hood. Since the Cre-expressing adenovirus is non-replicating, it will not be transmitted by infected cells. Should someone come in direct contact with the virus it should have no effect since the cre-recombinase induces recombination of segments of DNA surrounded by inserted floxP sequences which no human cells should carry in their genomes.

12 e) Describe measures that will be used to reduce risk to the environment and all project and animal facility personnel:

Barrier methods such as the use of gloves, masks and gowns (level 2 protection) will be used while handling the adenovirus and the cells that have been exposed to the adenovirus. The infections will take place inside a flow hood in H5 and all surfaces potentially exposed to the adenovirus will be cleaned with a disinfecting solution. As stated above, the adenovirus expressing Cre is non-replicative, thus any of the mammary cells infected in vitro with the adenovirus Cre will not produce virus thus there should be no superinfection in the mammary fat pad nor is there any risk of infection of the animal care personnel.

13. Reviewer's Modifications (to be completed by ACC only): The Animal Care Committee has made the following modification(s) to this animal use procedure protocol during the review process. Please make these changes to your copy and comply with the recommended changes as a condition of approval.

Guidelines for

7. Animal Data

7 c) Description of animals

Quality Control Assurance: To prevent introduction of infectious diseases into animal facilities, a health status report or veterinary inspection certificate may be required prior to receiving animals from all non-commercial sources or from commercial sources whose animal health status is unknown or questionable. Quarantine and further testing may be required for these animals.

If more than 6 columns are needed, please attach another page

	Sp/strain 7	Sp/strain 8	Sp/strain 9	Sp/strain 10	Sp/strain 11	Sp/strain 12
Species	mouse	mouse	mouse	mouse	mouse	mouse
Supplier/Source	In House (IH)	IH	IH	IH	IH	IH
Strain	Crk I	Crk II	Met wt	Met Y1003 F	Met M1268T	Met Y1003F/M1268T
Sex	M/F	M/F	M/F	M/F	M/F	M/F
Age/Wt	varies	varies	varies	varies	varies	varies
# To be purchased	0	0	0	0	0	0
# Produced by in-house breeding	300	300	60	60	60	100
# Other (e.g.field studies)	none	none	none	none	none	none
#needed at one time	various	various	various	various	various	various
# per cage	5	5	5	5	5	5
TOTAL# /YEAR	300	300	100	100	100	100

7 d) Justification of Animal Usage: BASED ON THE EXPERIMENTAL OBJECTIVES OF THE PROJECT, describe the number of animals required for one year. Include information on experimental and control groups, # per group, and failure rates. For breeding, specify how many adults are used, number of offspring produced, and how many offspring are used in experimental procedures. Use the table below when applicable. The arithmetic explaining how the total of animals for each column in the table above is calculated should be made clear. ***(Space will expand as needed)***

Strains 7 and 8 - Crk I and II

The involvment of the Crk adapter protein in mammary tumorigenesis and mammary development will be observed in transgenic mice expressing Crk I or Crk II adapter proteins under the control of the mammary specific MMTV promoter.MMTV/Crk transgenic MICE WILL BE GENERATED as a fee for service at the McIntyre Transgenic Core facility and mice will be transferred to the H3 facility by the staff of the McGill H3 facility. In the H3 facility the mice will be bred and progeny observed for tumor formation by palpation or sacrifice. At the moment, we are getting the Crk I mice. The Crk II mice are generated and started to breed in our animal facility. The number of mice requested reflects the necessity of identification and maintenance of transgenic stocks (approx. a max of 300 mice). We did obtain 13 transgenic animals from 30 injections for the Crk I construct. We estimate the production of crk I mice as a maximum of 300. We will initially breed 10 CrkII and 10 CrkI transgenics and F1s will be examined for transgene expression. Total of 10 male/female mice for breeding purposes. Two or three lines for each transgene will be selected and kept for maintenance of each line.

MMTV/CrkII transgenic (5 founders) MMTV/CrkI transgenic (5 founders)

Total 10 transgenic breeding lines.

Strain 9 - 10 - 11 - 12 : Met wt, Y1003F, M1268T and Y1003F/M1268T

Regarding the following strains : Met wt, Met Y1003F and Met M1268T, we are completing the required number of mice to get a portion of the colony as an aging colony and an other portion to breed for tissue colletion and to maintain it as well. (40 aging mice + 60 young mice = 100 mice per line to keep). The double mutant mice (Met Y1003F/M1268T) requires to produce more mice because we need to determine which of the ten founders transmit the transgene and get its progeny expressing the transgene in their mammary epithelium. For that purpose, some breeding females will be sacrifice at lactation stage (since the expression is present in the lactating mammary glands), and we need to breed FVB to get Foster parents. As other Met strain mentionned earlier, the double mutant strain requires aging and maintaining mice. So, the maximum to keep at all time for any of these lines is 200 mice (including aging mice from last year).

7. Animal Data

7 c) Description of animals

Quality Control Assurance: To prevent introduction of infectious diseases into animal facilities, a health status report or veterinary inspection certificate may be required prior to receiving animals from all non-commercial sources or from commercial sources whose animal health status is unknown or questionable. Quarantine and further testing may be required for these animals.

If more than 6 columns are needed, please attach another page

	Sp/strain 13	Sp/strain 14	Sp/strain 15	Sp/strain16	Sp/strain 17	Sp/strain 18
Species	mouse	mouse	mouse			
Supplier/Source	In House (IH)	IH	Dr Chuxia Deng			
Strain	MT-met wt	MT-met Y1003F	Brca1delta11/coTrp53 - Cre			
Sex	M/F	M/F	M/F			
Age/Wt	varies	varies	6 to 12 wks old			
# To be purchased	0	0	4			
# Produced by in-house breeding	200	200	200			
# Other (e.g.field studies)	none	none	none			
#needed at one time	various	various	various			
# per cage	5	5	5			
TOTAL# /YEAR	200	200	204			

7 d) Justification of Animal Usage: BASED ON THE EXPERIMENTAL OBJECTIVES OF THE PROJECT, describe the number of animals required for one year. Include information on experimental and control groups, # per group, and failure rates. For breeding, specify how many adults are used, number of offspring produced, and how many offspring are used in experimental procedures. Use the table below when applicable. The arithmetic explaining how the total of animals for each column in the table above is calculated should be made clear. ***(Space will expand as needed)***

Strain 13 -14 : MT-met wt and Y1003F
The metallothionein section (the MT-met mice). We have developed ten founders expressing an activated Met allele as well as 10 founder animals expressing a wild type Met allele under either the MMTV (mouse mammary tumor virus) or MT (metallothionein) promoter. Founders that transmit the transgene and whose progeny express the transgene in their mammary epithelium will be used to maintain lines. We have to select two to three lines per transgene to maintain and focus our work on. A part of the colony must age (positive and negative mice). The number of mice requested for strains 13 and 14 reflects the requirement for identification and maintenance of the transgenic stocks as well. We estimate our mouse production at 100 mice to cover all conditions for this year. We already have a total of 100 mice in the room (some are aging, some are stocks).
100 + 100 = 200 for each strain

Strain 15 : Brca1delta11/coTrp53 - Cre
We work on the Met receptor tyrosine kinase. The amplification of the murine Met locus in this mouse model has been demonstrated by Dr Chuxia Deng. As a collaboration, we are interested to examine the Met protein in primary tumor explants established from the tumors in these mice. We should get 2 males and 2 females from Dr Deng. We must generate some mice prior doing any experiement. Considering the complexity to get mice with all the transgenes and according to our experience, we estimate the breeding production to be at a maximum of 200 mice over the year.
200 + 2 + 2 = 204

Supplier's info :
Dr Chuxia Deng
Tel: (301) 402-7225
Fax: (301) 480-1135
Email: chuxiad@bdg10.niddk.nih.gov

National Institute of Diabetes and Digestive and Kidney Diseases
9000 Rockville Pike, Building 12A, Room 3011
Bethesda, MD 20892-5632