ELUCIDATION OF MECHANISMS IN MAMMARY TUMORIGENESIS AND IDENTIFICATION OF DRIVING EVENTS AND SUSCEPTIBILITY LOCI

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Marsha Diane Wallace

August 2012

ELUCIDATION OF MECHANISMS IN MAMMARY TUMORIGENESIS AND IDENTIFICATION OF DRIVING EVENTS AND SUSCEPTIBILITY LOCI

Marsha Diane Wallace, Ph. D.

Cornell University 2012

Abstract

Human breast cancer research faces limitations necessitating a comparative oncogenomic approach and use of models, in which the environment and genetic background can be controlled. The *Mcm4*^{Chaos3} (*Chaos3*) mouse model contains the only endogenous mutation in mice known to lead exclusively to spontaneous development of mammary adenocarcinomas. My comparative analysis of *Chaos3* and human oncogenomic data implicate *NF1* deficiency as a major driver of breast cancer. Traditionally known for its tumor suppressive role in preventing neurofibromas, I find *NF1* is deficient in *Chaos3* mammary tumors and 27.7% of human breast tumors, including >40% of Her2-enriched and basal-like subtypes. NF1 loss triggered hyperactivation of the RAS oncogene, and these tumor cells were sensitive to RAS pathway drugs. As NF1 deficiency confers increased resistance to standard tamoxifen treatment, my findings have considerable implications for NF1 testing and personalized treatment that we project impacts ~383,230 women who develop breast cancer with *NF1* deficiency annually. ~25% of breast cancer cases have a heritable/familial basis, but underlying susceptibility genes remain largely unknown. While, *Chaos3*-C3H mice develop mammary tumors, *Chaos3*-C57BL/6 mice develop lymphomas and histiocytic sarcomas, indicating that cancer type is highly influenced

by background strain. We utilized *Chaos3* mice of mixed backgrounds to identify mammary tumor susceptibility and resistance loci. Quantitative trait loci (QTL) analysis of ~200 C3H x C57BL/6 F2 *Chaos3* females revealed candidate genes involved in cell proliferation, DNA repair, cell signaling, and cancer-associated genes. One locus contained *Tln1*, a gene required for integrin activation, in which a germline mutation was discovered. *Chaos3 Tln1* mutants were aged, and a significantly higher proportion developed mammary tumors, validating *Tln1* impact on mammary tumor susceptibility. In another set of studies, I found that DNA damage response (DDR) deficiency and reproductive hormones have a significant impact on carcinogenesis when MCM DNA replication machinery is defective. ATM pathway deficiency in *Chaos3* mice resulted in decreased tumor latency and/or increased tumor susceptibility. Oophorectomized *Chaos3* mice had decreased mammary tumor incidence, but increased susceptibility to other cancer types. Together, my results in four areas of breast cancer research demonstrate significant advancement in the understanding of mechanisms involved in mammary tumorigenesis.

BIOGRAPHICAL SKETCH

Marsha was born in Nashville, TN and raised in Smyrna, TN. In high school, she volunteered at Middle Tennessee Medical Center, was a member of the chess club, worked as an editor on the school newspaper in journalism, and served on the student council as student body historian during her senior year. She graduated near the top of her class in 2002, and was accepted into Middle Tennessee State University (MTSU) on a presidential scholarship. She double majored in Biology and Psychology, while earning a minor in Chemistry. In her third year, she won the Ann Harrison McClary and Richard E. McClary Scholarship for Academic Excellence. While at MTSU, she worked at a veterinary hospital as an outpatient nurse and then at The Nashville Zoo as an overnight instructor. She graduated MTSU *summa cum laude* in 2006, and she accepted Cornell's offer into the field of Genetics and Development for graduate school. Her final rotation in Dr. John Schimenti's lab focused on the *Chaos3* cancer project that became the heart of her dissertation. In 2010, she won the 24th International Mammalian Genome Conference Outstanding Poster Award in Crete, Greece for her cancer research. She was awarded positions on NIH training grants in 2008 and 2011, including a Cornell Genomics Scholars Award for the training grant in Developmental Genomics. Now, in 2012 she is graduating with her Ph.D. in genetics. For the future, Marsha is looking forward to a career in genetic testing with possible connections to cancer and personalized treatments in order to help alleviate human disease.

This dissertation is dedicated to my family for their lifelong support, to cancer patients across the globe, for which the research herein is focused, and to research scientists in every field of every generation, who have dedicated their lives to the betterment of humanity.

ACKNOWLEDGMENTS

I would like to extend my thanks to Dr. John Schimenti for accepting me into his lab and allowing me to tackle the *Chaos3* project from a genomics mindset. He challenged me in every aspect of the scientific process, from experimental design to conclusions. He taught me that 1, 5, and 20 revision versions were never enough for manuscript editing.

I wish to acknowledge my committee members Dr. Andrew Clark and Dr. Robert Weiss for their ability to expand my perspective of the research. Dr. Clark brought his insight of genomics to the research. Dr. Weiss taught an engaging course in Genomic Integrity that greatly paralleled my own research and reinforced my knowledge in the field. I would also like to thank Dr. Charles Aquadro for pioneering the Personalized Genomics class and allowing me to be involved in the Cornell Genographic Project.

I would like to thank all of the Schimenti Lab members, past and present, for their attentive ears and input over the years. I would especially like to thank Chen-Hua Chuang for his work on the *Chaos3* project and our scientific collaboration.

I am grateful for the funding I have been awarded in my time as a graduate student. Funding has been provided by the Developmental Genomics Training Program and the Genetics and Development Training Program through the NIH (IT32HDO57854 and 5T32GM007617).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Breast cancer is the most prevalent cancer in women (29% of all cases), with an expected 226,870 new cases and 40,000 deaths in 2012 in the United States alone [1]. Efforts are being placed on identifying events and factors driving carcinogenesis as well as determining loci contributing to breast cancer susceptibility. This dissertation will focus on four main areas of research in breast cancer:

1. Genomic alterations in mammary tumors to identify drivers
2. Identification of mammary tumor susceptibility loci
3. Impact of DNA damage response perturbation on mammary carcinogenesis
4. Role of reproductive hormones in mammary carcinogenesis

1.1 Areas of Breast Cancer Research

One challenge facing cancer researchers is that cancer is not a singular disease, rather a complex set of diseases. Human breast tumors can be divided into multiple subtypes based on reproductive hormone receptor status and expression signature [2-4]. Estrogen and Progesterone receptor positive (ER+, PR+) tumors can respond to the presence of estrogen and progesterone. Luminal A breast tumors tend to be ER+ and low grade (well differentiated) [5,6]. Luminal B mammary tumors also tend to be ER+ but are often high grade (poorly differentiated) [5,6]. In the HER2/neu+ subtype, epidermal growth factor receptor Her2/neu/ErbB2 is amplified [5,6]. Basal-like mammary tumors are triple negative (ER-, PR- and HER2-) and associated with

poorer prognosis [7]. Claudin-low mammary tumors are often triple-negative, and additionally have low expression of cell-cell junction proteins (E-cadherin), which is associated with increased invasion and metastasis [8].

The majority of breast cancer cases (~75%) are believed to have a sporadic, rather than heritable basis [9]. Therefore, research focus is directed at genomic analysis of somatic tissue, contrasting normal tissue to the tumors to identify genes and pathways that are frequently spontaneously mutated, misregulated, or have copy number alterations (CNAs), which may thus represent cancer "drivers" with causative roles. However, the prevalence of passenger mutations, heterogeneity, and the diversity of tumor etiologies and subtypes complicates conclusions about genes identified in these studies [4,10,11]. These genes are only putative drivers, being established solely by statistical association, and mechanistic validation remains to be tested directly.

Twin and family studies indicate that ~25% of breast cancer cases have a heritable basis [9]. However, mutations in the most penetrant susceptibility genes known, *BRCA1* and *BRCA2*, account for only ~5% of breast cancer cases in the general population [12-14]. The majority of susceptibility genes underlying heritable breast cancer remain unknown. To identify additional susceptibility genes and drivers, large-scale genome wide association studies and genome cancer resequencing projects have been conducted [10,15-18]. Overall, it appears that the majority of genetically-based breast cancers are caused by low-penetrance modifier alleles [19]. This complicates attempts at genetic mapping by GWAS in humans.

DNA damage response (DDR) pathways are responsible for helping maintain genomic stability and suppressing tumorigenesis. Ataxia Telangiectasia Mutated (ATM) as well as ATM and Ataxia-Telangiectasia-and-Rad3-related (ATR) are DNA damage sensors that head DNA checkpoint and repair pathways, including signaling to Tumor Protein 53 (TP53), the tumor suppressor most frequently compromised in human cancer [20,21]. Genomic studies have shown that many genes are misregulated or altered at low frequencies in human cancers, but together comprise significant alterations in key pathways, specifically DDR pathways [15,22,23].

Reproductive hormones and their receptors also have a profound impact on mammary tumorigenesis. Approximately 75% of all human breast tumors are positive for estrogen receptor (ER+), and growth of these tumors can be stimulated by estrogen [24]. Nulliparous women have twice the risk of developing breast cancer as women who have undergone a full term pregnancy before 20 years of age [25]. In women, multiple early pregnancies confer a lifelong reduced risk of breast cancer [26,27]. This protection can be mimicked in rodents through administration of estrogen and progesterone treatments [28,29], and hormonal treatment causes increased long-term expression of *Trp53* and other pro-apoptotic genes [26,27].

1.2 Mouse Models

With the challenges and limitations of human studies, mouse cancer models are powerful for untangling the genomic basis of cancers. Genetic backgrounds can be precisely defined, and phenotypic variation can be minimized, allowing mouse models to be used to identify genetic loci that modify cancer risk. Whereas the

heterogeneity of human populations and the diversity in breast cancer etiology complicates genetic analysis, within a given cancer mouse model, the tumors that arise have a consistent underlying basis. Additionally, mice can be experimentally manipulated to solidify evidence of candidate genes and validate genomic findings.

However, the majority of mouse models have drawbacks for breast cancer research. With the singular eitiology of mammary tumors within a mouse model in a single strain, the universality of effects of a mutation across different backgrounds is unknown. In the worst case, results could be restricted to that background. Currently, the most widely used mouse models of mammary cancer are transgenics, in which the Mouse Mammary Tumor Virus (MMTV) is driving overexpression of an oncogene [30,31]. This is artificial and may not be relevant to the human situation. Also, despite the powerful genetics in mice, there has been little success in cloning modifier loci or identifying new mammary cancer drivers on a large scale.

1.3 Utilizing the *Chaos3* Mouse Model to Study Mammary Carcinogenesis

The *Chaos3* (Chromosome aberrations occurring spontaneously 3) mouse model has several advantages to elucidate mechanisms of carcinogenesis, including that tumors arise spontaneously rather than the mice being genetically engineered or treated with carcinogens. The $mcm4^{Chaos3}$ (*Chaos3*) nonsynonomous point mutation in the C3H genetic background is the only endogenous gene mutation in mice that leads exclusively to mammary carcinogenesis [32]. Nearly all *Chaos3* homozygous nulliparous females in the C3H strain inbred background succumb to mammary adenocarcinomas with a mean latency of 12 months [32]. The *Chaos3* cancer model was

isolated in an *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis screen for mutations causing genomic instability (GIN) [32]. The nonsynonomous mutation identified in *mcm4* (minichromosome maintenance 4) causes a dramatic increase in micronuclei (a hallmark of GIN), a destabilized MCM helicase, pan-reduction of all MCMs, and a decreased number of dormant origins [33-37]. MCM4 is a highly conserved subunit of the MCM2-7 DNA replicative helicase, an essential component of pre-replication (pre-RC) complexes [38]. These complexes are "licensed" at replication origins for activation in S phase, and regulatory mechanisms inhibit reloading of the MCMs during S phase to prevent re-replication of the genome [38].

Here, I took a comparative oncogenomic approach to identify breast cancer drivers, utilizing the *Chaos3*-C3H mouse model. The controlled genetic background and singular tumor etiology allows identification of recurrent mutational events likely to be involved in driving tumorigenesis.

While the *Chaos3* mice were being bred to be congenic in C3H, some *Chaos3* mice of mixed background between C3H and C57BL/6J ("B6") developed lymphomas [39], as did *Chaos3* MCM2-deficient mice in a mixed C3H x B6 background (Figure A1-1). This suggested that *Chaos3* predisposes to cancer, but other loci in the genome determine tumor type susceptibility. These other loci could either be mammary tumor predisposition genes (in C3H) and/or mammary tumor preventative genes (in B6). Thus, we can exploit the *Chaos3* model to identify breast cancer modifiers.

In *Chaos3* mouse embryonic fibroblasts (MEFs), significant chromosome breakage (compared to wild-type controls) occurred only under conditions of replication stress, indicating that the damage was a consequence of a defect(s) in some

aspect of DNA replication [32,33]. Work in yeast carrying the *Chaos3* mutation suggests that the stress can cause the replication fork to collapse, leading to double strand breaks which then activates the HR (homologous recombination) pathway [40]. Evidence from other model systems support the conclusion that MCM dysfunction can cause DNA damage and rearrangements [41]. DDR genes target components of the core DNA replication complex, including the MCM helicase. MCM2 is a direct target of ATR, and MCM3 is a target of ATM [42,43]. Additionally, *Chaos3* cells demonstrate elevated activation of DDR pathways in the form of increased levels of RAD51 and BLM foci [36], and upregulation of p53 and p21 are observed in *Chaos3* MEFs [37]. *Chaos3* animals also deficient for p53 have decreased time to cancer onset [37]. This evidence suggests the *Chaos3* model may be sensitive to DDR gene perturbation. Here we generate double mutant lines between *Chaos3* and an additional DDR gene to examine the impact of components in DDR pathways on carcinogenesis when the core DNA replication machinery is defective.

Additional mechanistic variables, such as reproductive hormones and their receptors, may tie MCMs to *Chaos3* carcinogenesis and mammary tumor specificity. The reproductive hormones estrogen and progesterone control DNA replication in uterine epithelial cells by regulating MCM proteins [44]. Progesterone inhibits DNA synthesis by decreasing *Mcm* transcription (particularly *Mcm4*), MCM protein levels, and CDT1, the protein that facilitates loading MCMs onto replication origins. Progesterone also leads to the sequestration of MCMs into the cytoplasm even though these proteins are primarily nuclear [44]. Progesterone may regulate the MCMs through miRNAs in uterine epithelial cells [45], and normal MCM levels are restored in *Chaos3*

cells with the dicer or drosha pathways are knocked down [35]. In normal human breast cells, progesterone increases transcription of the Mcms and other DNA licensing factors [46].

Together, the characteristics of the *Chaos3* model make it an excellent tool to study mammary carcinogenesis and the consequences of defects in the core DNA replication machinery on cancer.

1.4 References

1. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J Clin* **62**, 10-29 (2012).
2. Carey, L.A. et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama* **295**, 2492-502 (2006).
3. Sorlie, T. et al. Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. *BMC Genomics* **7**, 127 (2006).
4. Perou, C.M. et al. Molecular portraits of human breast tumours. *Nature* **406**, 747-52 (2000).
5. Sorlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **98**, 10869-74 (2001).
6. Sorlie, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* **100**, 8418-23 (2003).
7. O'Brien, K.M. et al. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. *Clin Cancer Res* **16**, 6100-10 (2010).
8. Prat, A. et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).
9. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* **343**, 78-85 (2000).
10. Sjoblom, T. et al. The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268-74 (2006).
11. Fox, E.J., Salk, J.J. & Loeb, L.A. Cancer genome sequencing--an interim analysis. *Cancer Res* **69**, 4948-50 (2009).
12. Ford, D. et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **62**, 676-89 (1998).
13. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. *Br J Cancer* **83**, 1301-8 (2000).
14. Antoniou, A.C. et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* (2008).
15. Wood, L.D. et al. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108-13 (2007).
16. de Tayrac, M. et al. Integrative genome-wide analysis reveals a robust genomic glioblastoma signature associated with copy number driving changes in gene expression. *Genes Chromosomes Cancer* **48**, 55-68 (2009).
17. Jones, S. et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* **324**, 217 (2009).
18. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-12 (2008).

19. Pharoah, P.D. et al. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* **31**, 33-6 (2002).
20. Moynahan, M.E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* **11**, 196-207 (2010).
21. Lane, D. & Levine, A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol* **2**, a000893 (2010).
22. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).
23. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
24. Li, C.I., Daling, J.R. & Malone, K.E. Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. *J Clin Oncol* **21**, 28-34 (2003).
25. MacMahon, B. et al. Age at first birth and breast cancer risk. *Bull World Health Organ* **43**, 209-21 (1970).
26. Russo, I.H. & Russo, J. Hormonal approach to breast cancer prevention. *J Cell Biochem Suppl* **34**, 1-6 (2000).
27. Medina, D. Breast cancer: the protective effect of pregnancy. *Clin Cancer Res* **10**, 380S-4S (2004).
28. Sinha, D.K., Pazik, J.E. & Dao, T.L. Prevention of mammary carcinogenesis in rats by pregnancy: effect of full-term and interrupted pregnancy. *Br J Cancer* **57**, 390-4 (1988).
29. Guzman, R.C. et al. Hormonal prevention of breast cancer: mimicking the protective effect of pregnancy. *Proc Natl Acad Sci U S A* **96**, 2520-5 (1999).
30. Green, J.E. & Hudson, T. The promise of genetically engineered mice for cancer prevention studies. *Nat Rev Cancer* **5**, 184-98 (2005).
31. Ottewell, P.D., Coleman, R.E. & Holen, I. From genetic abnormality to metastases: murine models of breast cancer and their use in the development of anticancer therapies. *Breast Cancer Res Treat* **96**, 101-13 (2006).
32. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-8 (2007).
33. Shima, N., Buske, T.R. & Schimenti, J.C. Genetic screen for chromosome instability in mice: Mcm4 and breast cancer. *Cell Cycle* **6**, 1135-40 (2007).
34. Chuang, C.H., Wallace, M.D., Abratte, C., Southard, T. & Schimenti, J.C. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* **6**(2010).
35. Chuang, C.H. et al. Post-transcriptional homeostasis and regulation of MCM2-7 in mammalian cells. *Nucleic Acids Res* (2012).
36. Kawabata, T. et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41**, 543-53 (2011).
37. Kawabata, T. et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. *Mamm Genome* **22**, 506-17 (2011).

38. Sclafani, R.A. & Holzen, T.M. Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**, 237-80 (2007).
39. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-98 (2007).
40. Li, X.C. & Tye, B.K. Ploidy dictates repair pathway choice under DNA replication stress. *Genetics* **187**, 1031-40 (2011).
41. Bailis, J.M. & Forsburg, S.L. MCM proteins: DNA damage, mutagenesis and repair. *Curr Opin Genet Dev* **14**, 17-21 (2004).
42. Cortez, D., Glick, G. & Elledge, S.J. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A* **101**, 10078-83 (2004).
43. Shechter, D. & Gautier, J. MCM proteins and checkpoint kinases get together at the fork. *Proc Natl Acad Sci U S A* **101**, 10845-6 (2004).
44. Pan, H., Deng, Y. & Pollard, J.W. Progesterone blocks estrogen-induced DNA synthesis through the inhibition of replication licensing. *Proc Natl Acad Sci U S A* **103**, 14021-6 (2006).
45. Kuokkanen, S. et al. Genomic profiling of microRNAs and messenger RNAs reveals hormonal regulation in microRNA expression in human endometrium. *Biol Reprod* **82**, 791-801 (2010).
46. Graham, J.D. et al. DNA replication licensing and progenitor numbers are increased by progesterone in normal human breast. *Endocrinology* **150**, 3318-26 (2009).

CHAPTER 2

Evidence of NF1 Deficiency as a Leading Breast Cancer Driver

Marsha D. Wallace[1], Adam D. Pfefferle[2], Lishuang Shen[1], Adrian J. McNairn[1], Ethan G. Cerami[3], Barbara L. Fallon[1], Vera D. Rinaldi[1], Teresa L. Southard[4], Charles M. Perou[2,5], John C. Schimenti[1,6]

Affiliations:

[1]Department of Biomedical Sciences, Cornell University, Ithaca, NY.

[2]Department of Genetics, University of North Carolina, Chapel Hill, NC.

[3]Memorial Sloan-Kettering Cancer Center, New York, NY.

[4]Section of Anatomic Pathology, Cornell University, Ithaca, NY.

[5]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC.

[6]Center for Vertebrate Genomics, Cornell University, Ithaca, NY.

One Sentence Summary: Comparative analysis of a mouse mammary tumor model and human oncogenomic data implicate *NF1* deficiency as a major driver of breast cancer.

2.1 Abstract

Most large-scale genomic studies of human breast cancer lack experimental evidence to support computationally implicated driver genes, a task complicated by the genetic diversity amongst tumors and people. To overcome these issues, we incorporated human genomic tumor data with experimental data from the C3H-*Chaos3* mouse model to provide evidence of *NF1* deficiency as a leading driver in breast cancer. Recurrent *Chaos3* mammary tumor copy number alterations (CNAs) overlap with those found in human breast cancer, most strikingly loss of the *Nf1* tumor suppressor in nearly all cases. Analysis of The Cancer Genome Atlas (TCGA) data revealed *NF1* deficiency in 27.7% of all human breast tumors, including >40% of Her2-enriched and basal-like subtypes. We show that NF1 loss triggers hyperactivation of the RAS oncogene in *Chaos3* tumors and cell lines, rendering them sensitive to drugs targeting the RAS pathway. These data implicate *NF1* deficiency as a major breast cancer driver that we project to impact ~383,330 women annually, a finding that will be informative for personalized treatments.

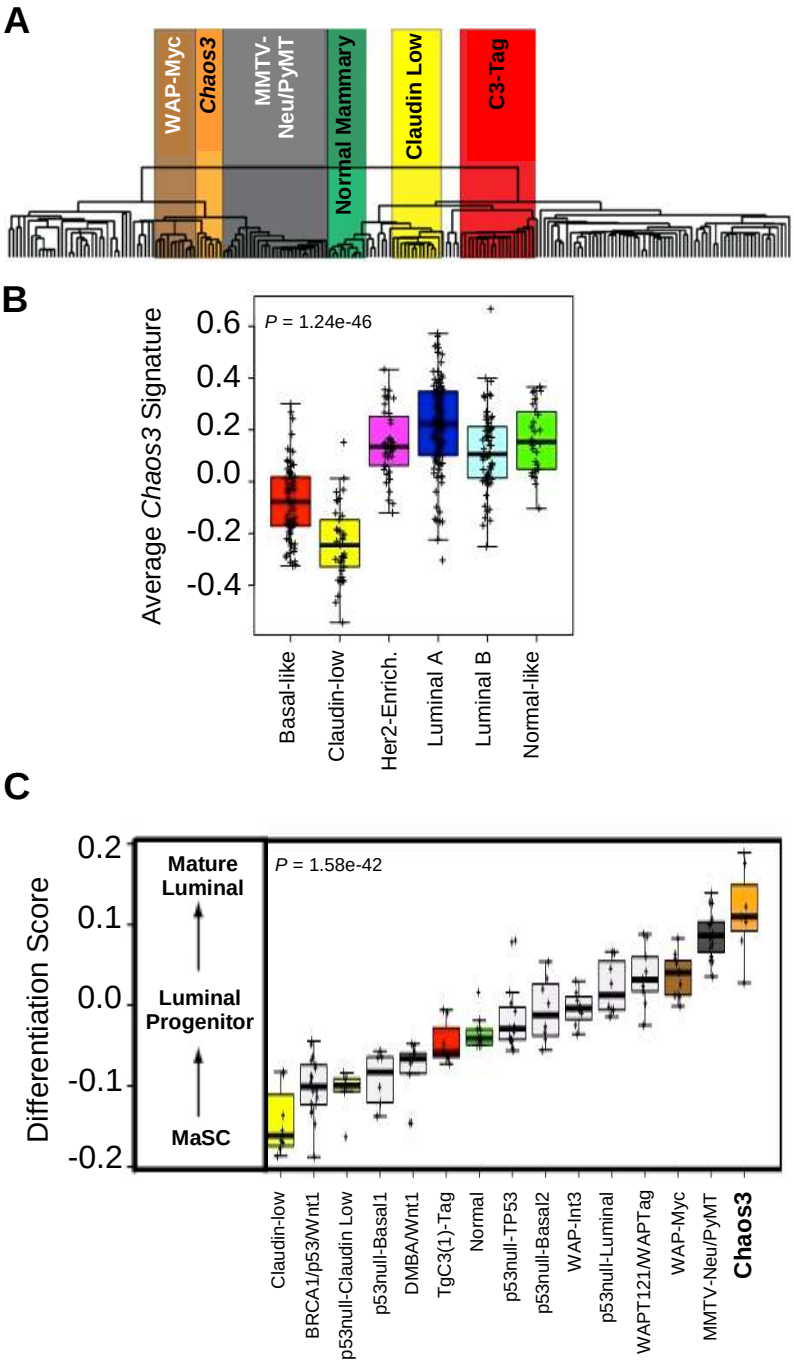
2.2 Main Text

Twin and family studies indicate that only ~25% of breast cancer cases have a heritable/familial basis, and thus the majority (~75%) appear to be "sporadic" [1]. Hence, much effort is being placed on genomic analysis of breast tumors and other cancers. The goal is to identify genes and pathways that are commonly altered and which may thus represent cancer "drivers" with causative roles. However, the prevalence of passenger mutations, genetic heterogeneity, and the diversity of tumor etiologies and subtypes complicates unequivocal identification of drivers,

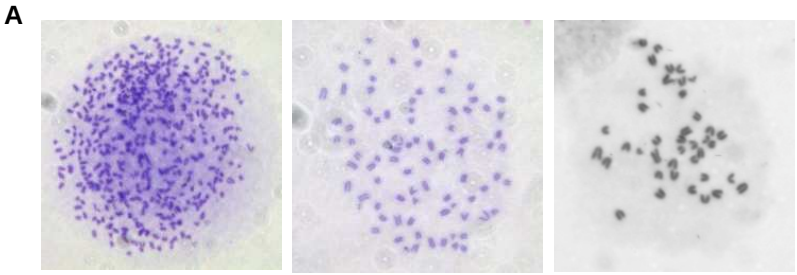
necessitating experimental validation. Here, we took a comparative oncogenomic approach for breast cancer driver identification, exploiting a highly relevant mouse model, C3H-*Chaos3*. These mice bear a point mutation in the minichromosome maintenance 4 (*Mcm4*) gene that destabilizes the MCM2-7 helicase that is essential for faithful DNA replication. The resulting genomic instability (GIN) causes >80% of nulliparous females to develop mammary adenocarcinomas exclusively [2]. The controlled genetic background and singular tumor etiology allows identification of recurrent mutational events likely to be involved in driving tumorigenesis.

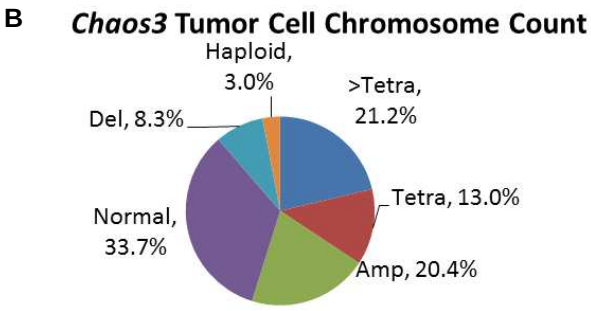
Human breast tumors can be classified into subtypes using gene expression signatures that are also present within mouse models of mammary cancers [3,4]. Expression profiling of *Chaos3* mammary tumors revealed that they cluster near three luminal adenocarcinoma mouse models (Figure 2-1A). Consistent with this, the *Chaos3* gene signature was most highly expressed in the Human Luminal A subtype, and was also high in HER2-enriched and Luminal B tumors (Figure 2-1B). Luminal breast tumors are the most prevalent type in humans [5]. Significance Analysis of Microarray (SAM) revealed that *Chaos3* tumors have a distinct gene expression pattern from all other mouse models, including dramatic upregulation of *Mucl1*, a diagnostic marker in human breast cancer (Table A2-1) [6]. Tumor differentiation score (D-Score) analysis showed that *Chaos3* tumors more closely resemble mature human luminal cells than any mouse model analyzed to date (Figure 2-1C). Together, these results show that *Chaos3* mice, which are not genetically engineered, are an excellent human breast cancer model.

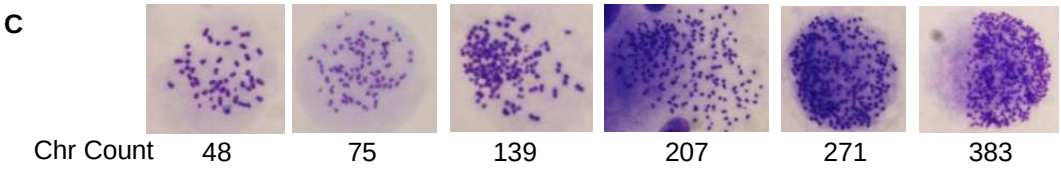
Figure 2-1: *Chaos3* tumors model key human features. (**A**) Expression microarray dendrogram of *Chaos3* mammary tumors and 185 other mouse mammary carcinomas and normal mammary tissue samples. The *Chaos3* tumors cluster together as a distinct group near luminal mouse models: WAP-MYC, PyMT, and Her2/Neu. (**B**) Boxplot of the *Chaos3* gene signature in the UNC337 human breast tumor dataset. *Chaos3* tumors have higher signature expression in human luminal, HER2-enriched, and normal-like intrinsic subtypes. (**C**) *Chaos3* Differentiation Score (D score) in relationship with other GEMMs. The high D-Score shows that *Chaos3* tumors more closely resembles the expression signature of mature human luminal cells relative to all other mouse models analyzed. (**B**, **C**) P-values reflect statistical significance of ANOVAs. Key: MaSC - Mammary Stem Cell.


A
WAP-Myc
Chaos3
MMTV-Neu/PyMT
Normal Mammary
Claudin Low
C3-Tag
B
P = 1.24e-46
Average Chaos3 Signature
0.6
0.4
0.2
0.0
-0.2
-0.4
Basal-like
Claudin-low
Her2-Enrich.
Luminal A
Luminal B
Normal-like
C
P = 1.58e-42
Differentiation Score
0.2
0.1
0.0
-0.1
-0.2
Mature Luminal
Luminal Progenitor
MaSC
Claudin-low
BRCA1/p53/Wnt1
p53null-Claudin Low
p53null-Basal1
DMBA/Wnt1
TgC3(1)-Tag
Normal
p53null-TP53
p53null-Basal2
WAP-Int3
p53null-Luminal
WAPT121/WAPTag
WAP-Myc
MMTV-Neu/PyMT
Chaos3

Primary *Chaos3* cells have increased stalled replication forks that persist through metaphase, leading to chromosome breaks and improper chromosomal segregation [2,7]. Similar to human breast tumors [8], *Chaos3* tumors had high levels of aneuploidy and drastic variation in chromosome number, even within the cells of a single tumor (Figure 2-2). With such intratumor variation, we expect that only early and/or highly selected mutations would be readily detectable and highly recurrent across multiple cases. To uncover mutations potentially driving carcinogenesis in *Chaos3* mice, we first performed partial exome resequencing of mammary tumors (Figure 2-3; Table A2-2 through Table A2-7). Surprisingly, we discovered few somatic point mutations in the targeted exonic regions and calculated the mutation rate at 1.1×10^{-7}, or 0.25 mutations/Mb, which is not above the background rate in other genomic studies of breast cancer (Table A2-8) [9,10]. The mutated genes are involved in diverse functions, and together they do not implicate a commonly affected pathway underlying carcinogenesis (Figure 2-3). These results indicated that elevated intragenic mutagenesis is not the primary mechanism driving *Chaos3* carcinogenesis, suggesting that other initiators such as CNAs may be responsible.







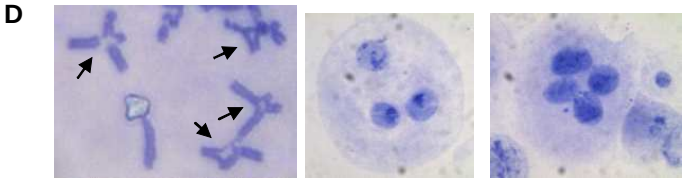
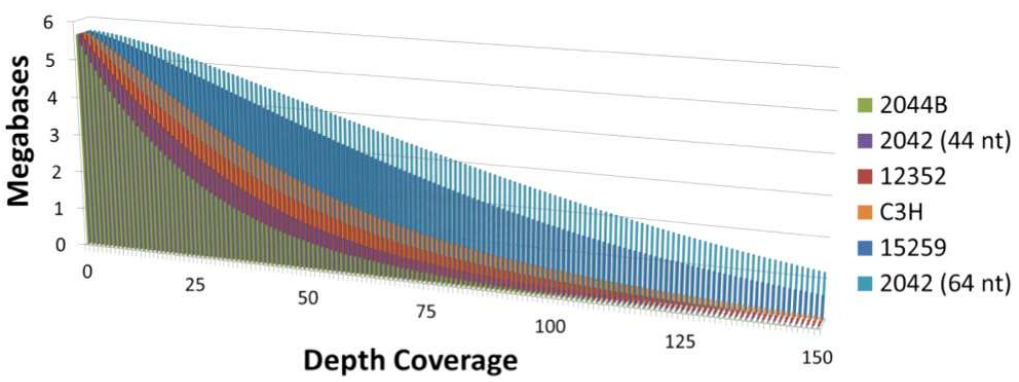


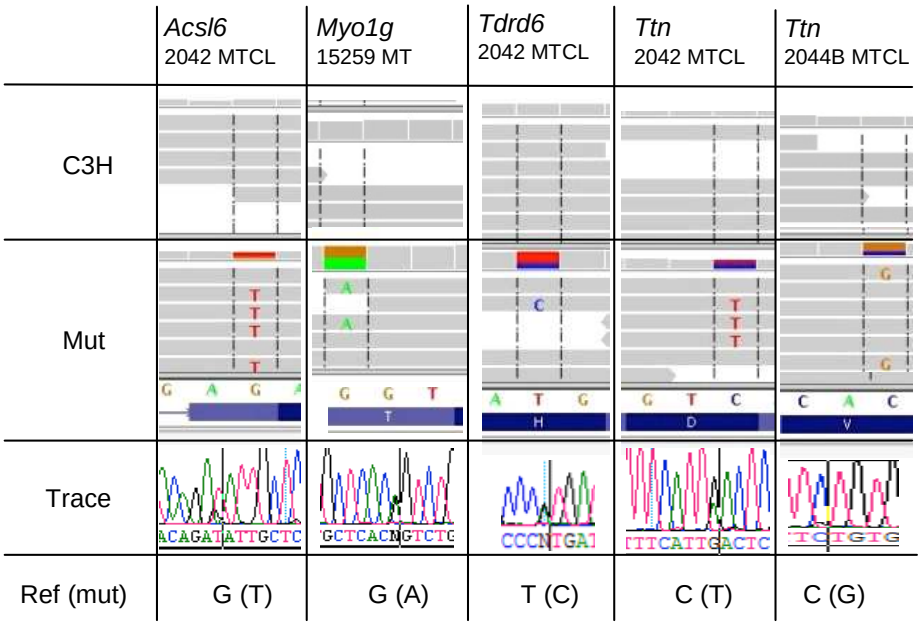
Figure 2-2: *Chaos3* tumors demonstrate high levels of GIN and aneuploidy. (**A**) Metaphase spreads from cells of 3 *Chaos3* mammary tumors. Note aneuploidy in left and middle spreads compared to the normal 40 chromosomes (left to right: 414, 83, 40). (**B**) Examination of 16 *Chaos3* tumors reveal a normal chromosome count in an average of only 1/3 of the cells (>Tetra= Beyond Tetraploidy; Tetra=Tetraploid; Amp=Amplification; Del=Deletion). (**C**) Metaphase spreads from one *Chaos3* mammary tumor (16864a). Chromosome count is indicated beneath the images. Note the extreme variation of aneuploidy found within a single tumor. (**D**) Additional abnormal features displayed by tumor cells, including: cruciform structures (left) and abnormal multi-nucleated cells (middle and right).

Figure 2-3: Low point mutation rate in coding regions of *Chaos3* mammary tumors. (**A**) Partial exome resequencing depth coverage for Sequence Capture on-target reads. An average read depth coverage of 52.6 was achieved for the target regions, with >85% of the bases in the capture target region represented by at least 5x coverage. NT=nucleotide. A single 88 nt read length run was conducted on sample 2042, and reads were subsequently shortened during analysis to both 64 and 44 nt to reduce error rate. The 44 nt length was used for SNP and mutation calling. (**B**) Somatic mutations in *Chaos3* tumors. Shown are aligned Sequence Capture reads and validated sequence trace from Sanger sequencing, reflecting the five somatic mutations discovered in target regions from four *Chaos3* tumor samples: *Acsl6*, *Myo1g*,*Tdrd6*, *Ttn*. Note: one *Chaos3* primary mammary tumor (12352) had no validated somatic point mutations. C3H is the wild type control; Mut is the tumor. Nucleotide positions in gray indicate no deviation from wild type C57BL/6 reference. MT= Mammary tumor; CL= Cell Line. The *Chaos3* calculated tumor mutation rate is 1.1 x 10-7, or 0.25 Mutations/Mb. *Chaos3* tumors do not exhibit an increased mutation rate above background breast tumor mutation rates.

A


Megabases
6
5
4
3
2
1
0
0
25
50
75
100
125
150
Depth Coverage
2044B
2042 (44 nt)
12352
C3H
15259
2042 (64 nt)

B


Acsl6
2042 MTCL
Myo1g
15259 MT
Tdrd6
2042 MTCL
Ttn
2042 MTCL
Ttn
2044B MTCL
C3H
Mut
Trace
Ref (mut)
G (T)
G (A)
T (C)
C (T)
C (G)

To examine genomic copy number changes, we performed array comparative genomic hybridization (aCGH) on twelve *Chaos3* tumors, including 9 *Chaos3* mammary tumors and 3 non-mammary tumors, and two MMTV-*Neu* mammary tumors. *Chaos3* non-mammary tumors can be obtained by genetic perturbations or altering the strain background [11,12]. Strikingly, the *Chaos3* tumors exhibited recurrent chromosomal aberrations. Nearly all had specific amplifications on Chromosomes (Chr) 12 and 16 (Figure 2-4A; Table A2-9 through Table A2-11). CNAs on Chrs 4, 5, and 11 were found in mammary tumors specifically (Table A2-9 through Table A2-12). We screened breast cancer data from The Cancer Genome Atlas (TCGA) and the Catalogue of Somatic Mutations in Cancer (COSMIC) databases and found overlapping syntenic CNAs in human mammary tumors (Figure 2-4B; Table A2-13 through Table A2-15). The Chr 12 amplification has remarkably precise breakpoints that flank an Immunoglobulin (Ig) gene locus, and curiously, the Chr 16 amplified regions are also replete with immunity-related genes (Table A2-9 and Table A2-10). Additional genes in these regions have roles in metastasis, pluripotency, signal transduction, or are known to be upregulated in cancer (Table 2-1; Table A2-10). Genes in the deleted regions function in apoptosis/necrosis, DNA checkpoint/repair, signal transduction, and tumor suppression (Table 2-1; Table A2-10).

Figure 2-4. Recurrent CNAs in *Chaos3* mammary tumors are frequently altered in human breast cancer, including one quarter of human cases with *Nf1* deficiency. (**A**) KCSmart analysis of combined aCGH data from 12 *Chaos3* tumors, (9 mammary and 3 non-mammary). The most significant amplification peaks (red) lie on Chrs 12 and 16, and deletions (green) on Chrs 4, 5, and 11. (**B**) Overlap of mouse (Mmu) *Chaos3* recurrent mammary tumor deletions (thick red bars) with recurrent human (Hs) breast tumor segmental CNAs (thick black bars). Human gene orders are shown. Refer to Table A2-13 and Table A2-14 for complete comparison. Asterisks indicate juxtaposed and contiguous sequences in the mouse genome. See Figure 2-8 and Methods for details. (**C**) Percentage of *NF1* CNA and mutation in 511 human breast tumors, including 57 Her2-Enriched and 93 Basal breast tumors. Note that 27.7% of human breast tumors have *NF1* deletion or mutation, and HER2-Enriched and Basal breast tumor subtypes have >40% of cases with *NF1* deletion or mutation. (**D**) Boxplot of *NF1* mRNA expression (microarray) vs. copy number (GISTIC analysis) in human breast cancer. Horizontal gray bars are the means of each group. Blue X's represent individual tumor or normal samples. Homdel = homozygous deletion; Hetloss = heterozygous deletion; Amp = high level amplification. Expression levels significantly correlate with genomic copy number status (ANOVA between Hetloss and Diploid groups, $p=3.32 * 10^{-13}$). Human data were from unpublished TCGA (see Materials and Methods).

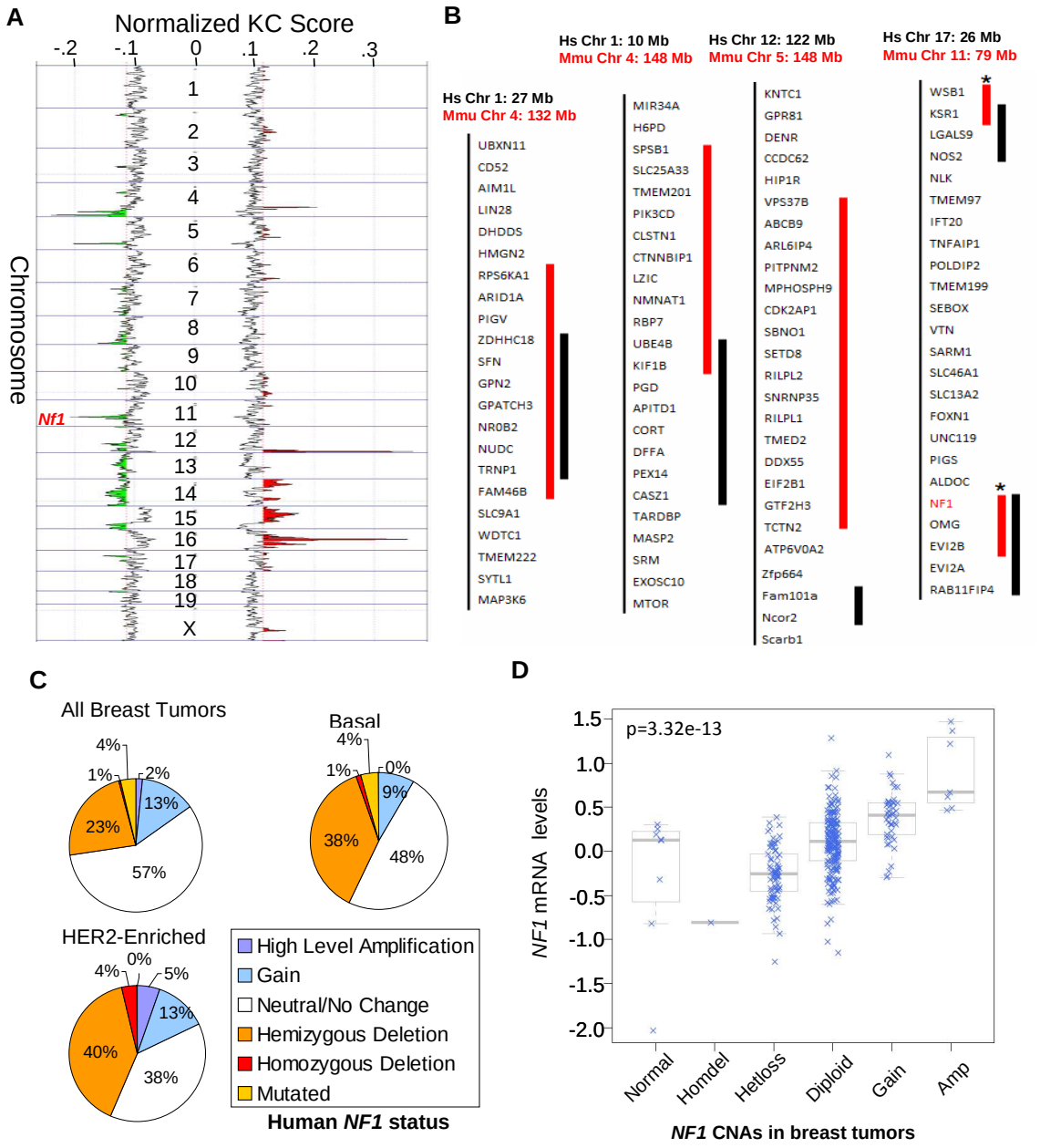

A
Normalized KC Score
-.2
-.1
0
.1
.2
.3
Chromosome
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
X
Nf1
B
Hs Chr 1: 27 Mb
Mmu Chr 4: 132 Mb
UBXN11
CD52
AIM1L
LIN28
DHDDS
HMGN2
RPS6KA1
ARID1A
PIGV
ZDHHC18
SFN
GPN2
GPATCH3
NR0B2
NUDC
TRNP1
FAM46B
SLC9A1
WDTC1
TMEM222
SYTL1
MAP3K6
Hs Chr 1: 10 Mb
Mmu Chr 4: 148 Mb
MIR34A
H6PD
SPSB1
SLC25A33
TMEM201
PIK3CD
CLSTN1
CTNNBIP1
LZIC
NMNAT1
RBP7
UBE4B
KIF1B
PGD
APITD1
CORT
DFFA
PEX14
CASZ1
TARDBP
MASP2
SRM
EXOSC10
MTOR
Hs Chr 12: 122 Mb
Mmu Chr 5: 148 Mb
KNTC1
GPR81
DENR
CCDC62
HIP1R
VPS37B
ABCB9
ARL6IP4
PITPNM2
MPHOSPH9
CDK2AP1
SBNO1
SETD8
RILPL2
SNRNP35
RILPL1
TMED2
DDX55
EIF2B1
GTF2H3
TCTN2
ATP6V0A2
Zfp664
Fam101a
Ncor2
Scarb1
Hs Chr 17: 26 Mb
Mmu Chr 11: 79 Mb
WSB1
KSR1
LGALS9
NOS2
NLK
TMEM97
IFT20
TNFAIP1
POLDIP2
TMEM199
SEBOX
VTN
SARM1
SLC46A1
SLC13A2
FOXN1
UNC119
PIGS
ALDOC
NF1
OMG
EVI2B
EVI2A
RAB11FIP4
C
All Breast Tumors
4%
1%
2%
13%
23%
57%
Basal
4%
1%
0%
9%
38%
48%
HER2-Enriched
4%
0%
5%
13%
40%
38%
High Level Amplification
Gain
Neutral/No Change
Hemizygous Deletion
Homozygous Deletion
Mutated
Human NF1 status
D
p=3.32e-13
NF1 mRNA levels
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
Normal
Homdel
Hetloss
Diploid
Gain
Amp
NF1 CNAs in breast tumors

Table 2-1: Cancer and Immunity Related Genes in *Chaos3* CNAs

	Amplified	
Function	**Chr 16**	**Chr 12**
Pluripotency	**Dppa4**, **Dppa2**	
Signal Transduction		***Adam6***
Immunity/ Inflammation	**Pvrl3**, **Retnlb**, **Retnla**	***Ig/abParts***
Upregulated in Cancer	**Igsf11**	

	Deleted				
Function	**Chr 4**	**Chr 5**	**Chr 11**	**Chr 10**	**Chr 19**
Tumor Suppressor		**Cdk2ap1**	***Nf1***		
DNA Checkpoint/ Repair		Kntc1, Gtf2h3, **Setd8**			**Rad9**
Apoptosis/ Necrosis	Dffa, ***Ube4b***, ***Kif1b***,			***Oaz1***	
Signal Transduction	**Pik3cd**,		***Ksr1***	***Csnk1g2***, ***Mknk2***	
Immunity/ Inflammation				***Lingo3***	
Other Cancer Related	**Arid1a**, ***Sfn***	**Sbno1**,			***Minpp1***

Legend: Genes altered in a high percentage of *Chaos3* mammary tumors specifically are underlined. Critical regions of central overlap across multiple *Chaos3* tumors were defined within CNAs (Refer to Figure 2-4B, Table A2-13, and Supplementary Methods). Genes within critical regions of CNAs are bolded, and these genes that additionally have CNAs in human breast cancers are italicized. Ig/abParts= Ig locus and Antibody Parts gene feature conserved between mice and humans. See Table A2-10, Table A2-13, and Table A2-14 for extended lists and complete analysis.

Of particular interest is a set of *Chaos3* deletions on Chr 11 that overlaps with a recurring cluster of CNAs on human Chr 17. All *Chaos3* mammary tumors examined by aCGH but none of the MMTV-*Neu* driven mammary tumors or *Chaos3* non-mammary tumors contained Chr 11 deletions (Figure 2-5; Table A2-9, Table A2-11, and Table A2-12). The small deletions have nested breakpoints that define a commonly-deleted region containing the tumor suppressor Neurofibromin 1 (*Nf1)* (Figure 2-5B). We then analyzed the DNA of these and additional *Chaos3* mammary tumors by qPCR. Overall, 59/60 contained *Nf1* deletions, with 51.6% appearing homozygous and 46.6% heterozygous (Table A2-12). *Nf1*-deleted tumors showed absence or severe reduction of mRNA and protein (Figure 2-6A, Figure 2-5C). NF1 negatively regulates RAS, which controls proliferation, differentiation, cell adhesion, apoptosis, and cell migration through the MAPK and PI3K signal transduction pathways (Figure 2-6B). The RAS pathway is misregulated in many cancer types, including recent studies implicating it in breast cancer [13-15]. RAS deregulation leads to increased invasion and metastasis and decreased apoptosis [16].

Figure 2-5: *Nf1* is deleted in *Chaos3* mammary tumors. (**A**) Recurrent significant deletion detected by aCGH on Chr 11 at ~79 Mb, specific to *Chaos3* mammary tumors. The broken red line indicates significant log2 ratios. 17883 is a mediastinal lymphoma/leukemic tumor, 16862 is a histiocytic sarcoma in the uterus, 10658 is a bone tumor, and the other tumors are mammary. Note that 16168 and 12352 mammary tumors did not have significant detectable deletion by Nimblegen aCGH software, but deletion was determined by qPCR (Table A2-11 and Table A2-12). (**B**) Top: Shown are aCGH results of 2 primary *Chaos3* mammary tumors and 1 *Chaos3* mammary tumor cell line. Dots substantially above the log2 ratio line correspond to loci amplified in the tumor, and dots below are underrepresented. Arrows mark loci commonly amplified in *Chaos3* tumors regardless of tumor type, and asterisks mark commonly deleted loci segregating specifically with mammary tumors. Bottom: Expanded view of Chr 11 deletion. Red bars indicate aCGH or qPCR confirmed deletion in all 9 *Chaos3* mammary tumors overlapping the *Nf1* tumor suppressor gene. Note MMTV-neu mammary tumors and *Chaos3* non-mammary tumors do not demonstrate *Nf1* deletion. (**C**) qRT-PCR analysis of *Nf1* mRNA levels across the transcript in *Chaos3* tumors. Percent expression is relative to an MMTV-PyVT tumor as control, which does not have loss of *Nf1*. Error bars show Standard Error of the Mean (SEM). Mammary tumor 15259 is classified as being heterozygously deleted for *Nf1*, and the other mammary tumors are homozygously deleted. Residual signal may reflect biopsy contamination or tumor heterogeneity.

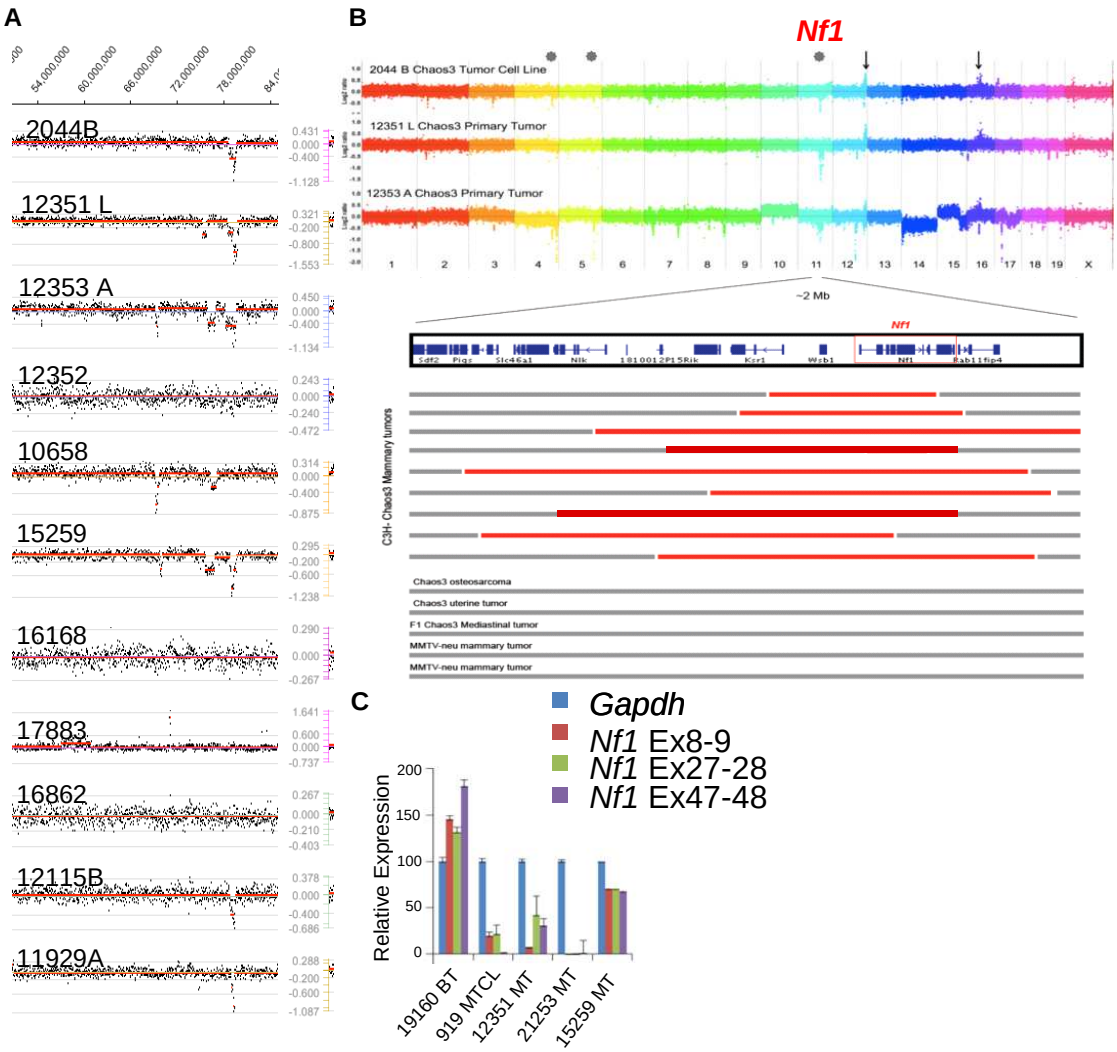
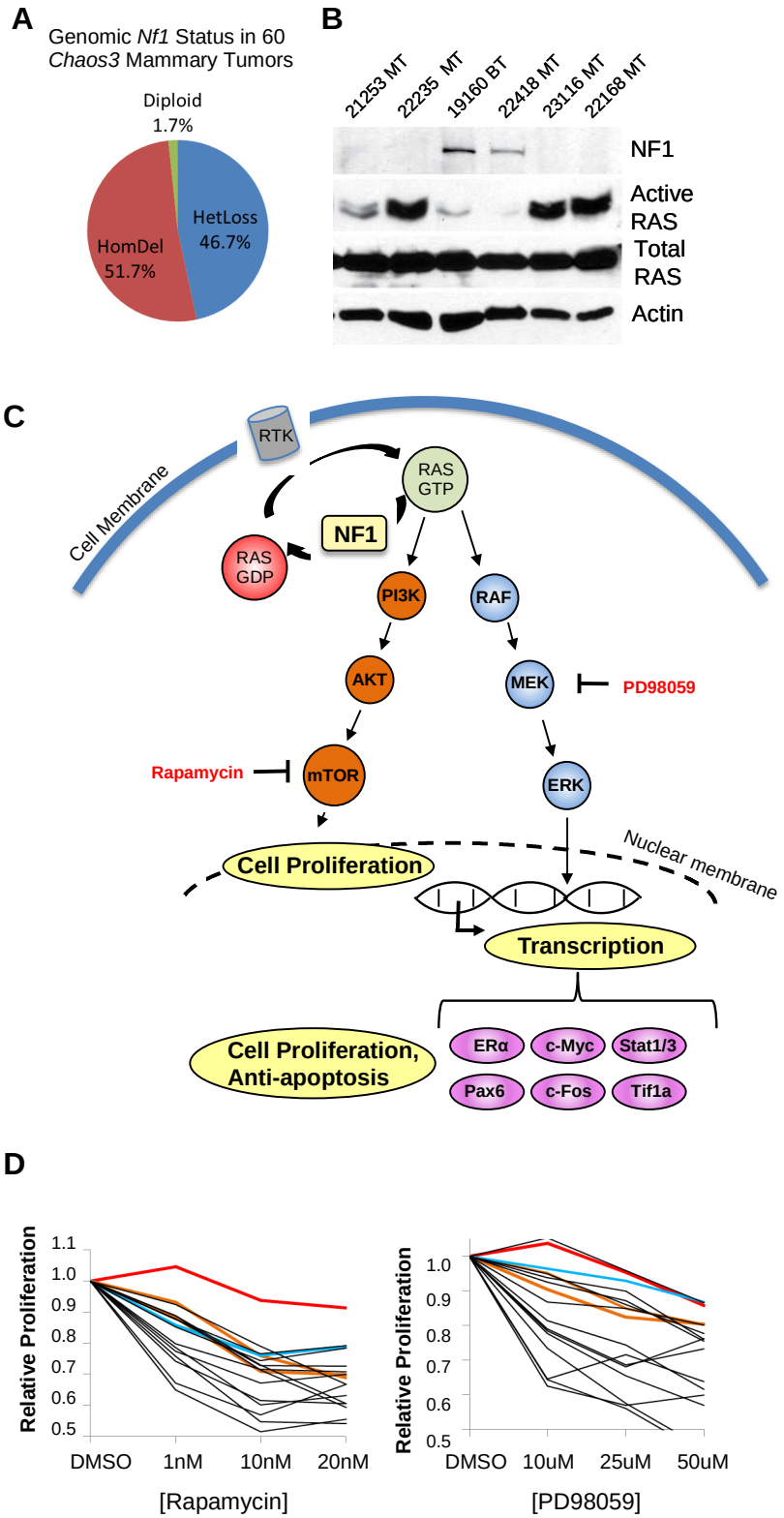

A
54,000,000
60,000,000
66,000,000
72,000,000
78,000,000
2044B
12351 L
12353 A
12352
10658
15259
16168
17883
16862
12115B
11929A
B
Nf1
2044 B Chaos3 Tumor Cell Line
12351 L Chaos3 Primary Tumor
12353 A Chaos3 Primary Tumor
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
~2 Mb
Nf1
Sdf2 Pigs Slc46a1 Nlrk 1810012P15Rik Kcnr1 Wnb1 Nf1 Rab11fip4
C3H- Chaos3 Mammary tumors
Chaos3 osteosarcoma
Chaos3 uterine tumor
F1 Chaos3 Mediastinal tumor
MMTV-neu mammary tumor
MMTV-neu mammary tumor
C
Gapdh
Nf1 Ex8-9
Nf1 Ex27-28
Nf1 Ex47-48
Relative Expression
200
150
100
50
0
19160 BT
919 MTCL
12351 MT
21253 MT
15259 MT

Figure 2-6: *Nf1* deletion leads to increased activated RAS and sensitivity to PI3K and MAPK inhibitors. (**A**) Western Blot analysis of *Chaos3* tumors for NF1 and active RAS levels. Mammary tumors without detectable NF1 have homozygous deletions of *Nf1*, whereas the bone tumor and mammary tumor 22418 have both genomic copies of *Nf1* (Table A2-12). The presence of NF1 protein is inversely proportional with the level of activated (GTP-bound) RAS. (**B**) NF1 loss leads to increased cell proliferation and transcription of anti-apoptosis genes. Inhibitors used in this study to slow proliferation of NF1-deficient tumor cells are shown in red type. Not all downstream targets are shown. RTK = receptor tyrosine kinase. (**C**) Cell proliferation assays showing sensitivity of *Chaos3* tumors to Rapamycin and MEK1 Inhibitor PD98059. Line colors: Red=HeLa, Brown=MCF-7 and MDA-MB231, Blue=PyVT, and Black=*Chaos3*. BT= bone tumor; MT= mammary tumor; MTCL= mammary tumor cell line. Zero concentration is DMSO solvent only.


A
Genomic Nf1 Status in 60
Chaos3 Mammary Tumors
Diploid
1.7%
HetLoss
46.7%
HomDel
51.7%
B
21253 MT
22235 MT
19160 BT
22418 MT
23116 MT
22168 MT
NF1
Active
RAS
Total
RAS
Actin
C
Cell Membrane
RTK
RAS
GTP
NF1
RAS
GDP
PI3K
RAF
AKT
MEK
PD98059
Rapamycin
mTOR
ERK
Nuclear membrane
Cell Proliferation
Transcription
Cell Proliferation,
Anti-apoptosis
ERα
c-Myc
Stat1/3
Pax6
c-Fos
Tif1a
D
Relative Proliferation
1.1
1.0
0.9
0.8
0.7
0.6
0.5
DMSO
1nM
10nM
20nM
[Rapamycin]
Relative Proliferation
1.0
0.9
0.8
0.7
0.6
0.5
DMSO
10uM
25uM
50uM
[PD98059]

Best known for causing neurofibromas in the autosomal dominant genetic disorder Neurofibromatosis type 1, women with inherited NF1 deficiency also have an increased risk of, or association with, breast cancer [17,18]. Though there are few reports implicating spontaneous *Nf1* loss in breast tumorigenesis [19,20], upon screening TCGA breast cancer datasets we found that 27.7% of human breast tumors have *NF1* deletions or mutations, most being hemizygous (Figure 2-4C; Table A2-16). Furthermore, >40% of Basal and HER2-enriched tumor subtypes have *NF1* loss or mutations (Figure 2-4C, Table A2-16). Genomic *NF1* deficiency in human breast tumors significantly correlated with decreased expression levels ($p=3.32 * 10^{-13}$) (Figure 2-4D). Canonically, tumor suppressors are thought to require loss of both copies to have functional impact. However, there is accumulating evidence that haploinsufficiency or reduced expression of tumor suppressor genes can have a carcinogenic impact [21]. Together these data indicate that *NF1* loss in conjunction with other CNAs is important for initiation and maintenance of mammary tumorigenesis in *Chaos3* mice and a substantial subset of human patients.

Cancer genome resequencing studies are finding evidence that *NF1* is mutated at significant rates in multiple cancers. *NF1* is the third most prevalently mutated or deleted gene in Glioblastoma multiforme (GBM) [22], one of the most significantly mutated genes in lung adenocarcinoma [23], and the 4th most (intragenically) mutated gene in ovarian carcinoma [24]. We examined *NF1* status in TCGA datasets available from 20 types of cancer. While most cancer types rarely contained *NF1*losses (<5% of cases), five cancer types showed >10% of cases with deletions (Table 2-2),

including an astounding 85.8% of serous ovarian cancers which also correlated with decreased *NF1* mRNA expression ($p=4.22 * 10^{-8}$) and patient survival (p=0.05) (Table 2-2, Figure 2-7). Only the 8% of cases with *NF1* homozygous deletion was emphasized in the TCGA serous ovarian publication [24].

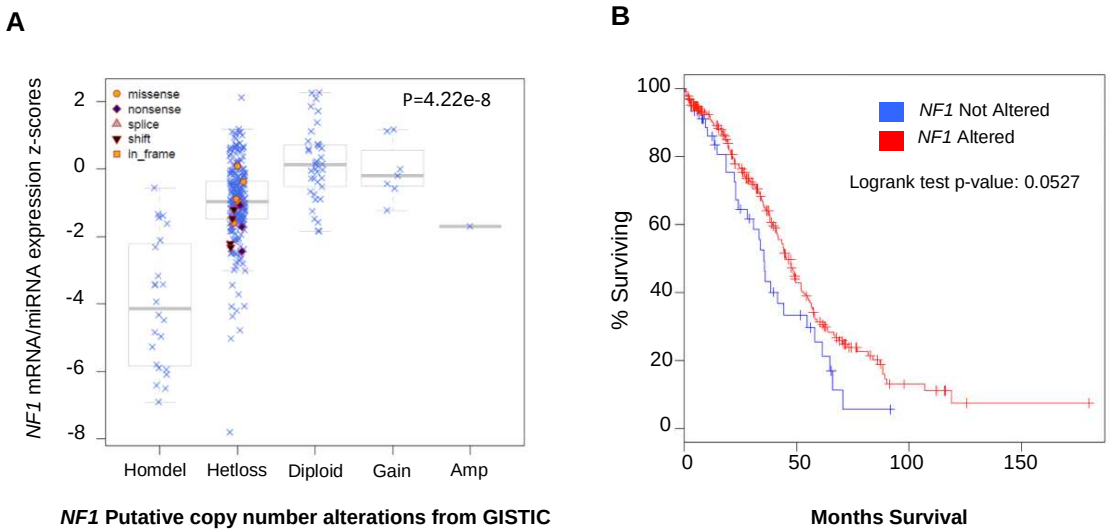


Figure 2-7: *NF1* alteration in human ovarian cancer. (**A**) Boxplot of *NF1* mRNA expression vs. copy number in human ovarian cancer. Data are from TCGA. Mutations are denoted with special symbols. 86% of 316 human ovarian cancer cases demonstrate *NF1* deletion (Table 2-2). P value is for ANOVA between Hetloss and Diploid groups, indicating expression levels significantly correlate with genomic deletion status. (**B**) Overall survival plot of *NF1* altered vs. unaltered human ovarian cancer cases. Patients with *NF1* alteration have longer overall survival. Key: Horizontal gray bars are the means of each group. Blue X's represent individual tumor or normal samples. Homdel = homozygous deletion, Hetloss = heterozygous deletion, Amp = high level amplification. Graphs were generated using the cBio Cancer Genomics Portal (www.cbioportal.org/public-portal/).

Table 2-2: *NF1* Copy Number Alteration in Human Cancers

Studies with Mutation Data	Data Status	Homozygous and Heterozygous Deletion
Colon and Rectum Adenocarcinoma	(TCGA, Provisional)	8.80%
Glioblastoma	(TCGA)	9.80%
Prostate Cancer	(MSKCC)	2.40%
Sarcoma	(MSKCC/Broad)	21.50%
Serous Ovarian Cancer	(TCGA)	85.80%
Studies without Mutation Data		
Bladder Urothelial Carcinoma	(TCGA, Provisional)	2.60%
Brain Lower Grade Glioma	(TCGA, Provisional)	1.20%
Breast Invasive Carcinoma	(TCGA, Provisional)	23.10%
Cervical Squamous Cell Carcinoma	(TCGA, Provisional)	2.80%
Head and Neck Squamous Cell Carcinoma	(TCGA, Provisional)	4.20%
Kidney Renal Clear Cell Carcinoma	(TCGA, Provisional)	4.10%
Kidney Renal Papillary Cell Carcinoma	(TCGA, Provisional)	0%
Liver Hepatocellular Carcinoma	(TCGA, Provisional)	5.70%
Lung Adenocarcinoma	(TCGA, Provisional)	3.20%
Lung Squamous Cell Carcinoma	(TCGA, Provisional)	14.60%
Pancreatic Adenocarcinoma	(TCGA, Provisional)	0%
Prostate Adenocarcinoma	(TCGA, Provisional)	4.90%
Stomach Adenocarcinoma	(TCGA, Provisional)	2.90%
Thyroid Carcinoma	(TCGA, Provisional)	0%
Uterine Corpus Endometrioid Carcinoma	(TCGA, Provisional)	11.80%

Legend: Copy numbers calls are made by GISTIC. Cancers having >10% of cases with *NF1* CNA are underlined.

NF1 is a negative regulator of the RAS signaling pathway that stimulates the GTPase activity of RAS, pushing it to the inactive state. NF1 is important for negatively regulating the pro-growth factor mTOR, which is stimulated by RAS (Figure 2-6B). Tumor cells of patients with Acute Myelogenous Leukemia (AML) having *NF1* deficiency demonstrate an elevated level of activated RAS and sensitivity to the mTOR inhibitor Rapamycin [25]. To assess the functional impact of *Nf1* deletion, we examined the level of activated RAS and found it to be dramatically higher in *Chaos3* mammary tumor cells deleted for *Nf1* (Figure 2-6A). We hypothesized that if the elevation of RAS signaling in *Nf1*-deleted mammary tumor cells is important for their maintenance, then inhibition of downstream pathways would compromise the growth of these cells. *Chaos3* mammary tumor cell lines were markedly sensitive to MAPK/MEK1 and/or mTOR inhibitors, PD98059 and Rapamycin respectively (Figure 2-6C). Identification of *NF1* as a tumor driver in a subset of breast cancers, and possibly other cancer types such as ovarian cancer, can provide guidance for patient treatment. Firstly, suppression of the RAS pathway would be an appropriate target. Secondly, there is reason to believe that tamoxifen, the estrogen receptor (ER) inhibitor that is standard treatment for ER+ breast cancers, may not be appropriate for women whose cancers involve *NF1* mutations. NF1 depletion was reported to confer resistance of human breast cancer (MCF7) cells to tamoxifen, and tamoxifen-treated patients whose tumors had lower *NF1* expression levels had poorer clinical outcomes [26]. Based on global cancer statistics [27,28] and the frequency of *NF1* mutation and deletion (Figure 2-4C), we project that ~383,330 women (~63,450 in the United

States) will develop breast cancer with *NF1* deficiency annually, underscoring the need for NF1 testing in the clinic.

The mechanism responsible for generating recurrent CNAs in *Chaos3* mice is likely related to the destabilized MCM2-7 replicative helicase [29], which may be predisposed to stalling at particular genomic regions that are difficult to replicate. Frequent deletion of *NF1* may be due to a combination of factors including fragile sites (Figure 2-8), a complex chromatin structure, and/or its large genomic size. Indeed, replication fork stalling near *Nf1* has been noted at a 5 kb isochore transition zone conserved between human and mouse, separating early and late replicating chromatin [30]. Furthermore, collisions between replication and transcription complexes cause instability at fragile sites in the longest human genes [31]. Loss or decrease of NF1 may trigger more than RAS pathway activation, as NF1 has been shown to bind to Focal Adhesion Kinase (FAK) and has multiple isoforms of unknown functions [32]. Additionally, siRNA-mediated *NF1* knockdown in epithelial-like breast cancer cells induced the expression of epithelial-to-mesenchymal transition-related transcription factors [33]. In addition to *Nf1* deletion, *Ube4b* and *Kif1b* were also frequently deleted in *Chaos3* and MMTV-neu mammary tumors, as in human breast tumors (26%) (Table 2-1, Figure 2-9). Genes in these regions (Table 2-1; Table A2-10) are excellent candidates to validate susceptibility genes underlying spontaneous or heritable forms of breast cancer.

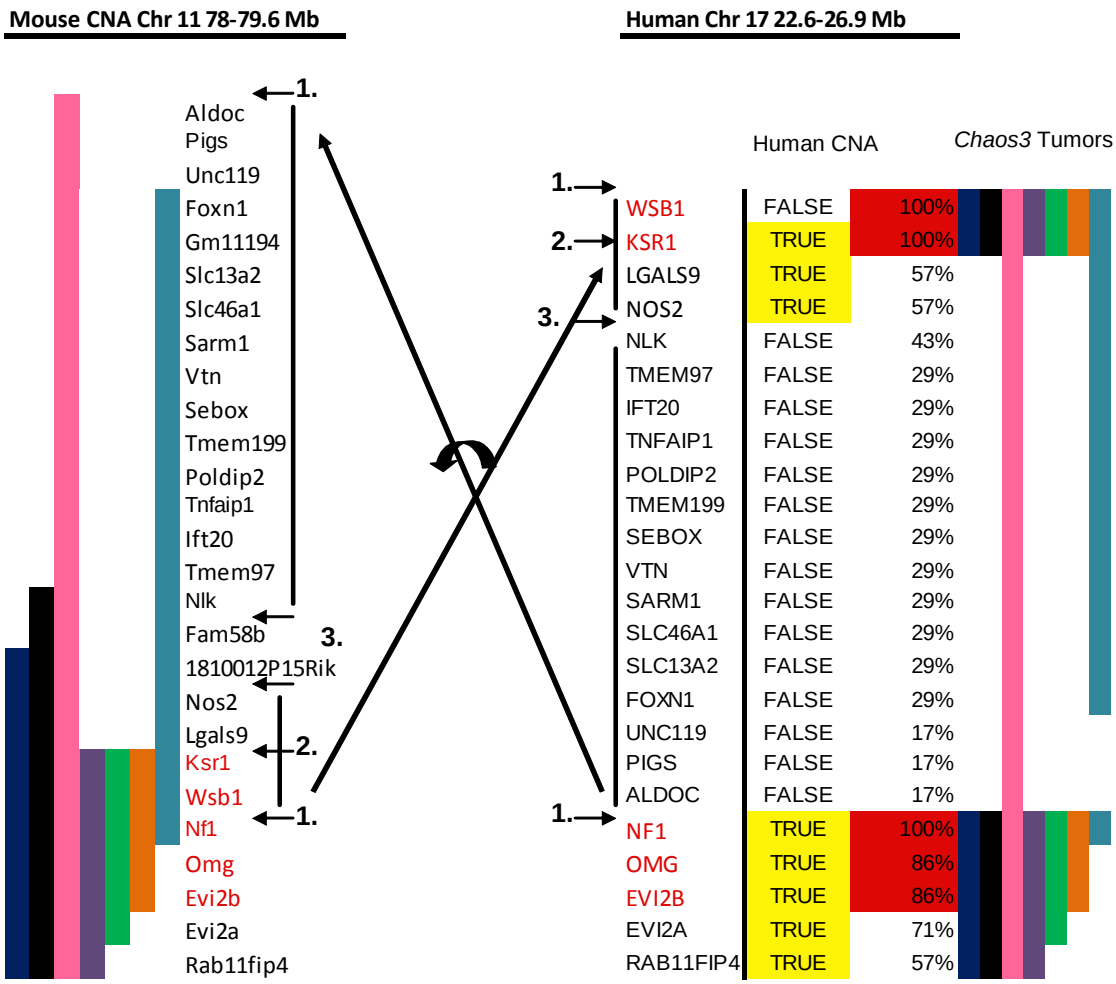


Figure 2-8: Genomic sequence around *Nf1* is prone to CNA and contains a genomic rearrangement. Colored vertical bars represent the deleted region in 7 *Chaos3* mammary tumors as detected by aCGH, and the percentages reflect how many of these tumors contain CNA for a given mouse gene. Gene names in red denote the *Chaos3* critical region. Mouse and human genomic orientations of the *Nf1* region are depicted. TRUE/FALSE indicates TCGA Level 4 (unpublished limited dataset) analysis of a subset of invasive breast carcinomas for segmental CNAs; it is possible that the intervals between *NOS2* and *NF1* are actually part of more inclusive deletion events. Numbers in bold with small arrows indicate positions of interest: **1.** Proximal to *Nf1*, a breakpoint of chromosomal inversion between human and mouse occurred between and including *Wsb1* to *Aldoc*. This is a site of both human and mouse tumor CNA, and the human CNA begins with *NF1*. **2.** The mouse critical CNA begins at *Ksr1*, which has flipped orientation in humans and starts/forms a second smaller CNA, with the caveats mentioned above. **3.** The mouse genome has gene insertion between *Nlk* and *Nos2*, where human CNA ends.

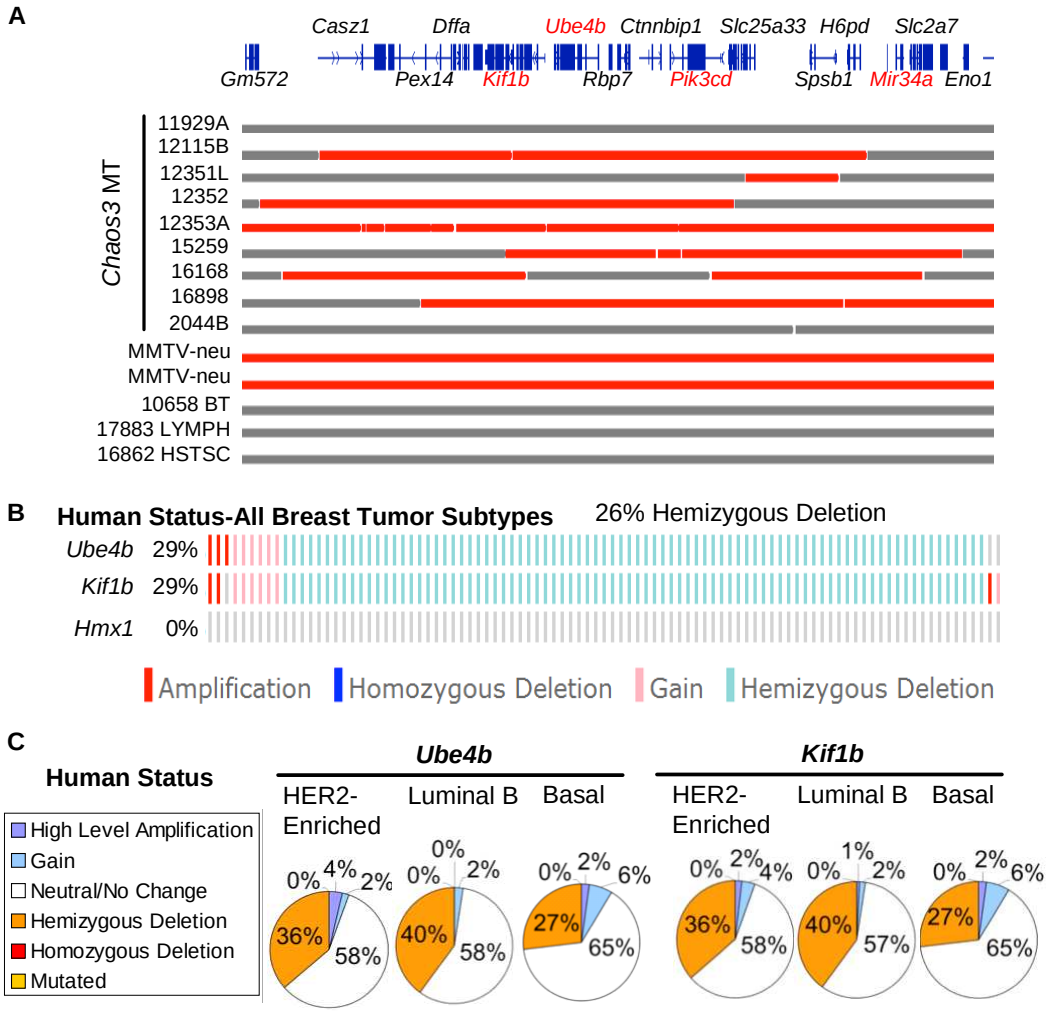


Figure 2-9: *Ube4b* and *Kif1b*, deleted in over half of *Chaos3* mammary tumors, show frequent deletion in human breast tumors. (**A**) Recurrent Chr 4 deletions specific to mammary tumors (MT). Horizontal bars represent tumors examined by aCGH. Red portions of bars indicate deleted regions in *Chaos3* and MMTV-neu mammary tumors. Cancer-related genes are in red. Note that *Chaos3* non-mammary tumors do not demonstrate this deletion. (**B**) Oncoprints of *Ube4b* and *Kif1b* alterations in 320 human breast tumors (unpublished TCGA). Rows contain bars representing individual tumors, and samples are aligned for visualization of alterations within the same tumor across multiple genes. *Hmx1* does not have a known role in cancer and was used as a control gene for qPCR. (**C**) Percentage of *Ube4b* and *Kif1b* CNA in 55 Her2-Enriched, 125 Luminal B, and 93 Basal human breast tumors (unpublished TCGA). Note that ~40% of HER2-Enriched and Luminal B tumors have hemizygous deletion of *Ube4b* and *Kif1b*, and 27% of Basal breast tumors.

2.3 Methods and Notes

Animals—*Chaos3* mammary tumors originated in mice congenic in C3HeB/FeJ except 16898. 16898 arose in a mixed C57BL/6J and C3HeB/FeJ background [7,11]. MMTV-neu and PyVT mammary tumors arose in FVB.

Microarray Expression Profiling—RNA was hybridized to custom murine Agilent microarrays and normalized as described [4,34]. Data were deposited into GEO (Accession # GSE36240). *Chaos3* tumors were clustered in relation to other GEMMs using an unsupervised analysis, and differentiation score was calculated as described [4,35]. SAM results were used to define a *Chaos3* gene signature (upregulated, FDR 0%) and compared to the UNC337 human tumor dataset [35].

Partial Exome Resequencing—A custom mouse 5Mb Sequence Capture array (NimbleGen) was used to enrich DNA corresponding to ~1200 breast cancer candidate gene exons (Table A2-4), followed by Illumina GAIIx sequencing. Candidate genes were selected and ranked based on breast cancer specificity and frequency in primary literature, existing cancer arrays, and cancer databases; see Table A2-2 through Table A2-4.

Capture Array Handling—Genomic DNA libraries of ~200 bp fragment size were constructed for four *Chaos3* mammary tumors and one inbred C3H WT spleen following the standard protocol of Illumina (San Diego, CA). One microgram of tumor and control library DNA was hybridized to the 385K or 720K capture array using an X1 mixer on the NimbleGen Hybridization System (Roche-NimbleGen) at 42°C for 3 days. Arrays were washed; then the captured molecules were eluted from

the slides using a NimbleGen Elution Station. Eluted molecules were vacuum-dried and amplified by LM-PCR. Real-time PCR of eight control amplicons was performed in the pre-capture library and post-capture library to estimate the target fold enrichment, which varied from 30-744x (Table A2-5).

Computational Analysis—The read data from each sample were aligned to the mouse C57BL/6, NCBI Build 37 (mm9) reference sequence using Novoalign (http://novocraft.com, v 2.05, academic version). Default alignment settings were used, but non-uniquely mapped reads or reads failing on alignment quality were discarded (-r NONE -Q 9). The percentage of on-target reads for mutant samples ranged from 34.5% to 62.9%, reflecting a 230 fold average enrichment for the target breast cancer candidate genes (Table A2-5 and Table A2-6). Genome Analysis Toolkit (GATK) version 1.04413 was used sequentially for base quality recalibration, depth of coverage estimation (Table A2-6 and Table A2-7), variant calling, and variant evaluation [36]. Substitution variants discovery and genotyping were performed with the GATK Unified Genotyper across all samples simultaneously. Single sample SNP calling was used to complement joint-sample SNP calling. The raw SNP calls were filtered per GATK recommendations with standard hard filtering parameters or variant quality score recalibration [36]. Criterion required SNP loci to have ≥5x coverage, variant frequency in ≥25% of reads, missing bases < 30%, no significant strand bias, and not overlapping indels. Indels were called with GATK IndelGenotyperV2 under both single sample and paired sample modes using C3H as the "normal" tissue to identify novel indels against C3H. No novel indels were identified in targeted coding regions. Known SNPs between C3H and C57/B6 were

mined from the Mouse Genome Database (http://www.informatics.jax.org/mgihome/projects/overview.shtml#snp), dbSNP [37], and Sanger Mouse Genome Project [38] (http://www.sanger.ac.uk/resources/mouse/genomes/). There were 3097 known C3H SNPs in seqcap target regions from traditional Sanger sequencing. GATK joint estimation from in-house data identified 2990 filtered SNPs, representing a 96.6% sensitivity. Known C3H SNPs were filtered out, and novel SNPs were identified for annotation and validation. Variation consequence was annotated with Ensembl Variation API (http://www.ensembl.org/info/docs/api/variation/index.html) and custom perl scripts. BAM, BED and VCF files were generated to visualize alignments and variations using the Integrative Genomics Viewer (IGV) software [39]. Variants were manually examined in IGV before proceeding to Sanger sequence validation.

Validation—Sequence reads of putative mutations were manually viewed using Integrative Genomics Viewer (IGV, Broad Institute). Variant positions were amplified in corresponding tumor samples and inbred C3H control genomic DNA. Following Fast AP and Exo1 (Fermentas) treatment, PCR products were Sanger sequenced and analyzed using SeqMan. GeneCard, Ingenuity Pathway Tool, Biocarta, and KEGG databases were used to annotate genes.

aCGH data analysis, and data sources—Genomic DNA from tumor and reference samples were hybridized to Nimblegen 3x720K mouse CGH arrays. Two reference samples were used independently. CNAs were visualized using Nimblegen, IGV, and KCsmart software [40]. Select genes were validated via qPCR (Table A2-11

and Table A2-12). Critical regions within each *Chaos3* CNA were identified as the region with the greatest overlap across multiple *Chaos3* tumors and compared to human datasets (Table A2-13 and Table A2-14).

Five micrograms of genomic DNA from tumor and reference samples were labeled and hybridized to 3x720K mouse Nimblegen CGH whole genome tiling arrays. The arrays consist of 50-75mer probes and a median spacing of 3.5kb, with a subset of probes concentrated on exons. Two reference samples were used independently to ensure recurring CNAs were not artifacts caused by the reference sample. The first reference sample was collected from a C3H WT inbred mouse, and run with samples 2044b, 12351, and 12353. The second reference sample selected originated from a C3H congenic *Chaos3*+/+ mouse and run as the reference for the remaining samples. DNA labeling, hybridization, and post-hybridization processing was performed according to the manufacturer's protocol. Nimblegen software was used to normalize test/reference ratios and perform background correction. Copy number changes were identified and segmented with Nimblegen CGH-segMNT algorithm using unaveraged and 10x averaging windows. The significance threshold was set at +/- 0.15 Log_2 ratio and required a minimum of two consecutive probes to exhibit change in order to call a segment. Amplifications and deletions were visualized using Nimblegen software and confirmed by manually examining Log_2 ratios for regions of interest. In addition to using Nimblegen software, the normalized log_2 ratio data were also analyzed using KC-smart software [40], to identify significantly recurrent CNAs. The kernel width was 1 Mb, and the resolution of the sample point matrix was 5 Kb. Simple Bonferroni multiple testing correction p<=0.025 was used as

threshold for declaring significant regions. Select genes within CNAs were validated via qPCR. See Table A2-17 for the primer list. Critical regions within each *Chaos3* CNA were identified as the region with the greatest overlap across multiple *Chaos3* tumors (Table A2-13 and Table A2-14). Of *Chaos3* tumors with CNAs in recurring regions, the percentage of those containing the critical region is as follows: Chr4 132M= 86%; Chr4 148M= 71%; Chr5 = 86%; Chr11 = 86% (100% for *Nf1*, *Ksr1*, *Wsb1*).

Human breast cancer data and CNA calls for comparison with *Chaos3* CNAs (Figure 2-4B; Table A2-13 and Table A2-14) were taken from the publicly available unpublished TCGA portal (https://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp) 2010 update. The regions considered to have undergone segmental deletions by the unpublished TCGA analysis ("Level 4" dataset) are those indicated in Figure 2-4B, Table A2-13, and Table A2-14. The MSKCC cBio portal provides a breakdown by mammary tumor subtype for individual genes (http://www.cbioportal.org/public-portal/index.do). Recurrent segmental CNA data in human breast cancers were pulled from limited level 4 unpublished dataset from TCGA). The Chr 11 deletions are single events in mice, and it is possible that the interval between *NOS2* and *NF1* may also be deleted as single events in human breast cancers, since the intervening genes are present in the hemizygous state in a high percentage of tumors according to extended TCGA datasets. According to the extended data available as of March 2012 through the MSKCC cBioPortal, the genes between *NOS2* and *NF1* interval (which were not classified as significantly segmentally deleted in the limited Level 4 dataset mentioned above), are hemizygously deleted at rates similar to *Nf1* itself.

Active RAS Pull-down and Western Blotting—Levels of activated RAS were obtained using an active RAS pull-down kit (Thermo Scientific). Rabbit anti-NF1 (Novus Biologicals) was used at 2ug/ml for Western analyses.

Cell Culture, Karyotyping, and Drug Treatment—Primary *Chaos3* tumor biopsies were homogenized, cultured, treated with colcemid, and metaphase spreads were made [2]. Imaged chromosomes were counted using ImageJ. Tumor cell lines were treated with the MEK1 Inhibitor PD98059 or MTOR inhibitor Rapamycin [41,42]. Cell Proliferation was assessed via MTT assay (Sigma) and values read on a 96-well ELISA plate reader.

Acknowledgments

Chaos3 Expression microarray data were deposited into GEO (Accession # GSE36240). Human breast cancer segmental CNA data (Figure 2-4B) are available through TCGA at https://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp. The MSKCC cBio portal provided extended datasets, including breakdown by mammary tumor subtype for individual genes (http://www.cbioportal.org/public-portal/index.do). This study was supported by NIH training grants IT32HDO57854 and 5T32GM007617 that supported M.D.W.; Empire State Stem Cell Fund contract numbers C026442 and C024174 to J.C.S.; and C.M.P. and A.D.P. were supported by NCI Breast SPORE program (P50-CA58223-09A1), by U24-CA143848, and by the Breast Cancer Research Foundation.

We thank R. Weiss for providing the MMTV-*Neu* tumor samples and the MDA-MB231 cell line; R. Davisson for providing the MCF-7 and HeLa cell lines; the

Cornell Genomics Core Facility for processing our genomic samples; and the Cornell Computational Biology Service Unit for consultation on data analysis.

Author Contributions

M.D.W. and J.C.S designed experiments; experiments were conducted and data analyzed primarily by M.D.W.; L.S. analyzed SeqCap data and conducted the KCSmart analysis; A.J.M. and V.D.R. preformed Western Blot analyses; E.G.C. led SeqCap chip design; A.D.P. conducted expression microarray analyses overseen by C.M.P.; B.L.F. assisted in CNA validation; T.L.S. conducted histopathology; J.C.S. oversaw the *Chaos3* project.

2.4 References

1. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* **343**, 78-85 (2000).
2. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-8 (2007).
3. Perou, C.M. et al. Molecular portraits of human breast tumours. *Nature* **406**, 747-52 (2000).
4. Herschkowitz, J.I. et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8**, R76 (2007).
5. Carey, L.A. et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama* **295**, 2492-502 (2006).
6. Hube, F., Mutawe, M., Leygue, E. & Myal, Y. Human small breast epithelial mucin: the promise of a new breast tumor biomarker. *DNA Cell Biol* **23**, 842-9 (2004).
7. Kawabata, T. et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41**, 543-53 (2011).
8. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).
9. Greenman, C. et al. Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153-8 (2007).
10. Kan, Z. et al. Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466**, 869-73 (2010).
11. Chuang, C.H., Wallace, M.D., Abratte, C., Southard, T. & Schimenti, J.C. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* **6**(2010).
12. Kawabata, T. et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. *Mamm Genome* **22**, 506-17 (2011).
13. Baljuls, A. et al. The tumor suppressor DiRas3 forms a complex with H-Ras and C-RAF and regulates localization, dimerization and kinase activity of C-RAF. *J Biol Chem* (2012).
14. Eastman, B., Jo, M., Webb, D., Takimoto, S. & Gonias, S. A transformation in the mechanism by which the urokinase receptor signals provide a selection advantage for estrogen receptor-expressing breast cancer cells in the absence of estrogen. *Cellular Signal* (2012).
15. Xue, M. et al. Insulin-like Growth Factor-1 Receptor (IGF-1R) Kinase Inhibitors in Cancer Therapy: Advances and Perspectives. *Curr Pharm Des* **18**, 2901-13 (2012).
16. Pylayeva-Gupta, Y., Grabocka, E. & Bar-Sagi, D. RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* **11**, 761-74 (2011).

17. Sharif, S. et al. Women with neurofibromatosis 1 are at a moderately increased risk of developing breast cancer and should be considered for early screening. *J Med Genet* **44**, 481-4 (2007).
18. Salemis, N.S., Nakos, G., Sambaziotis, D. & Gourgiotis, S. Breast cancer associated with type 1 neurofibromatosis. *Breast Cancer* **17**, 306-9 (2010).
19. Guran, S. & Safali, M. A case of neurofibromatosis and breast cancer: loss of heterozygosity of NF1 in breast cancer. *Cancer Genet Cytogenet* **156**, 86-8 (2005).
20. Lee, J. et al. Loss of SDHB and NF1 genes in a malignant phyllodes tumor of the breast as detected by oligo-array comparative genomic hybridization. *Cancer Genet Cytogenet* **196**, 179-83 (2010).
21. Berger, A.H., Knudson, A.G. & Pandolfi, P.P. A continuum model for tumour suppression. *Nature* **476**, 163-9 (2011).
22. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).
23. Ding, L. et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* **455**, 1069-75 (2008).
24. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
25. Parkin, B. et al. NF1 inactivation in adult acute myelogenous leukemia. *Clin Cancer Res* **16**, 4135-47 (2010).
26. Mendes-Pereira, A.M. et al. Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci U S A* (2012).
27. Jemal, A. et al. Global cancer statistics. *CA Cancer J Clin* **61**, 69-90 (2011).
28. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J Clin* **62**, 10-29 (2012).
29. Chuang, C.H. et al. Post-transcriptional homeostasis and regulation of MCM2-7 in mammalian cells. *Nucleic Acids Res* (2012).
30. Schmegner, C., Berger, A., Vogel, W., Hameister, H. & Assum, G. An isochore transition zone in the NF1 gene region is a conserved landmark of chromosome structure and function. *Genomics* **86**, 439-45 (2005).
31. Helmrich, A., Ballarino, M. & Tora, L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Mol Cell* **44**, 966-77 (2011).
32. Kweh, F. et al. Neurofibromin physically interacts with the N-terminal domain of focal adhesion kinase. *Mol Carcinog* **48**, 1005-17 (2009).
33. Arima, Y. et al. Decreased expression of neurofibromin contributes to epithelial-mesenchymal transition in neurofibromatosis type 1. *Exp Dermatol* **19**, e136-41 (2009).
34. Herschkowitz, J.I. et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. *Proc Natl Acad Sci U S A* **109**, 2778-83 (2012).
35. Prat, A. et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).

36. DePristo, M.A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-8 (2011).
37. Sherry, S.T. et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-11 (2001).
38. Keane, T.M. et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289-94 (2011).
39. Robinson, J.T. et al. Integrative genomics viewer. *Nat Biotechnol* **29**, 24-6 (2011).
40. Klijn, C. et al. Identification of cancer genes using a statistical framework for multiexperiment analysis of nondiscretized array CGH data. *Nucleic Acids Res* **36**(2008).
41. Chang, S.B., Miron, P., Miron, A. & Iglehart, J.D. Rapamycin inhibits proliferation of estrogen-receptor-positive breast cancer cells. *J Surg Res* **138**, 37-44 (2007).
42. Leong, D.T. et al. Cancer-related ectopic expression of the bone-related transcription factor RUNX2 in non-osseous metastatic tumor cells is linked to cell proliferation and motility. *Breast Cancer Res* **12**, R89 (2010).

CHAPTER 3

Tln1 and other genetic susceptibility and resistance loci in mammary adenocarcinomas

Marsha D. Wallace[1], Pavel Korniliev[2], Lishuang Shen[1], Teresa L. Southard[3], Jason G. Mezey[2], John C. Schimenti[1,4]

Affiliations:

[1]Department of Biomedical Sciences, Cornell University, Ithaca, NY.

[2]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY.

[3]Section of Anatomic Pathology, Cornell University, Ithaca, NY.

[4]Center for Vertebrate Genomics, Cornell University, Ithaca, NY.

One Sentence Summary: *Tln1* and other cancer susceptibility and resistance loci are identified using cancer-prone *Chaos3* mice in controlled environmental and genetic backgrounds.

3.1 Abstract

An estimated ~25% of breast cancer cases have a familial inherited basis, but the majority of susceptibility genes underlying these heritable cases remain unknown. Identification of genomic variants contributing to cancer susceptibility is complicated by both the breadth of genetic diversity between individuals and populations as well as differing environmental factors. These issues can be overcome using cancer mouse models with defined genetic backgrounds in a controlled environment. Here we utilize C3H x C57BL/6 F2 *Chaos3* mice, which bear a point mutation in the *Mcm4* DNA replication gene that leads the animals to spontaneously develop tumors, in order to identify susceptibility and resistance loci of mammary tumorigenesis and other cancer types. Conducting a quantitative trait loci (QTL) analysis, we find mammary tumor susceptibility and resistance loci which contain genes involved in cell proliferation (*Fgfr3*), DNA repair (*Msh4*, *Fancg*, *Fancc*, *Rad51ap1*), cell signaling (*Fbxw7*, *Sfrp1*, *Ptc1*, *Tln1*, *Pax5*), and cancer associated genes (*Rab2a*, *Rab28*, *Styk1*, *Mycbp2*). A Chr 4 mammary tumor susceptibility locus contains *Tln1*, a gene involved in integrin activation, in which a germline point mutation was discovered in the *Chaos3*-C3H stock line. C3H congenic *Chaos3* mice of both mutant and wild-type *Tln1* status were aged, and a significantly higher proportion of *Chaos3 Tln1* mutants developed mammary tumors compared to mice with the *Chaos3* mutation alone, validating *Tln1* impact on mammary tumor susceptibility.

.

Introduction

Twin and family studies indicate that ~25% of breast cancer cases have a familial genetic basis [1]. However, mutations in the most penetrant known susceptibility genes, *BRCA1* and *BRCA2*, account for only ~5% of breast cancer cases in the general population [2-4]. The majority of susceptibility genes underlying heritable breast cancer remain unknown.

To identify additional susceptibility genes, large-scale cancer genome resequencing projects have been conducted [5-9]. The emergent picture is that many mutations, most of which occur *de novo* and not inherited, collectively contribute to a given neoplasm [5]. Overall, it appears that the majority of genetically-based breast cancers are caused by low-penetrance modifier alleles [10], and also a large number of relatively rare breast cancer predisposition (driver) alleles. This complicates attempts at genetic mapping by GWAS in humans, and also confounds the goals of cancer genome resequencing efforts.

Whereas the heterogeneity of human populations and the diversity in breast cancer etiology complicates genetic analysis, inbred mouse strains are a powerful tool for studying cancer genetics because the genetic backgrounds are precisely defined. Within a given cancer mouse model, the tumors that arise have a consistent underlying basis. These models can be used to identify genetic loci that modify cancer risk. Mice bearing a mutation in the *Apc* gene (*Min*) have been used to discover a modifier of Min (*Mom1* or *Pla2g2a*) that affects tumor multiplicity and size of intestinal tumors [11]. Other studies have mapped genetic modifiers affecting mammary tumor frequency, multiplicity and latency in *Trp53* mutant animals [12-14].

However, the majority of mouse models have drawbacks for breast cancer research. Currently, the most widely used mouse models of mammary cancer are transgenics in which the Mouse Mammary Tumor Virus (MMTV) is driving overexpression of an oncogene [15,16]. This is artificial and may not be relevant to the human situation. Additionally, despite the powerful genetics in mice, there has been little success in cloning modifier loci or identifying new mammary cancer drivers on a large scale. To identify cancer modifier loci, we have utilized the *Chaos3* mouse model, which is not genetically engineered or treated with carcinogens, in order to overcome the disadvantages of traditional mouse models.

We isolated the *Chaos3* breast cancer mouse model in a genetic mutagenesis screen for mutations causing genomic instability (GIN) [17]. *Chaos3* is an allele of *Mcm4*, and the protein is a subunit of the highly conserved MCM2-7 DNA replicative helicase, essential for DNA replication and involved in preventing cells from over-replicating DNA. Most remarkably, nearly all *Chaos3* homozygous nulliparous females in the C3H strain inbred background succumbed to mammary adenocarcinomas with a mean latency of 12 months. However, while the mice were being bred to be congenic in C3H, some *Chaos3* mice of mixed background between C3H and C57BL/6J ("B6") developed lymphomas [18]. This suggested that *Chaos3* predisposes to cancer, but other loci in the genome determine tumor type susceptibility. These other loci could either be mammary tumor predisposition genes (in C3H) and/or mammary tumor preventative genes (in B6). As described below, tumor type is strongly influenced by genetic background, and we exploited this to identify breast cancer modifiers.

3.3 Results

Genetic background strongly influences tumor type and latency

To determine the impact of genetic background on *Chaos3* tumorigenesis, an F2 cross between *Chaos3* C3H vs. C57BL/6J strains, was conducted (Figure 3-1A). Animals from the two congenic populations were intercrossed to make F1s, then F1s were intercrossed to make F2s. From these crosses, 19 F1 females and 219 F2 females were aged to a terminal endpoint of 16 months or until tumors developed.

F1s had low mammary tumor incidence (23.8%), suggesting that either B6 has dominant suppressors, or that C3H has recessive susceptibility loci. F2s developed a range of disease phenotypes, predominantly histiocytic sarcomas (50.5%), while mammary tumor incidence was 15.1% (Figure 3-1B-C, Table A3-1). F2s had statistically significantly different time to onset of tumorigenesis associated with specific tumor types as well as compared to *Chaos3*-C3H mammary tumor latency. *Chaos3* C3HxB6 have significantly increased time to mammary tumor onset (14.5 months) over *Chaos3*-C3H mammary tumors (12.3 months) (Log-rank/Mantel-Cox Test, p=.002; Gehan-Breslow-Wilcoxon Test, p=.006) (Figure 3-1D). Median tumor free survival for *Chaos3* C3HxB6 mice developing other tumor types is also increased: Osteosarcomas (13.3 mo), Lymphoma (15.5 mo), Histiocytic Sarcoma (16.1 mo).

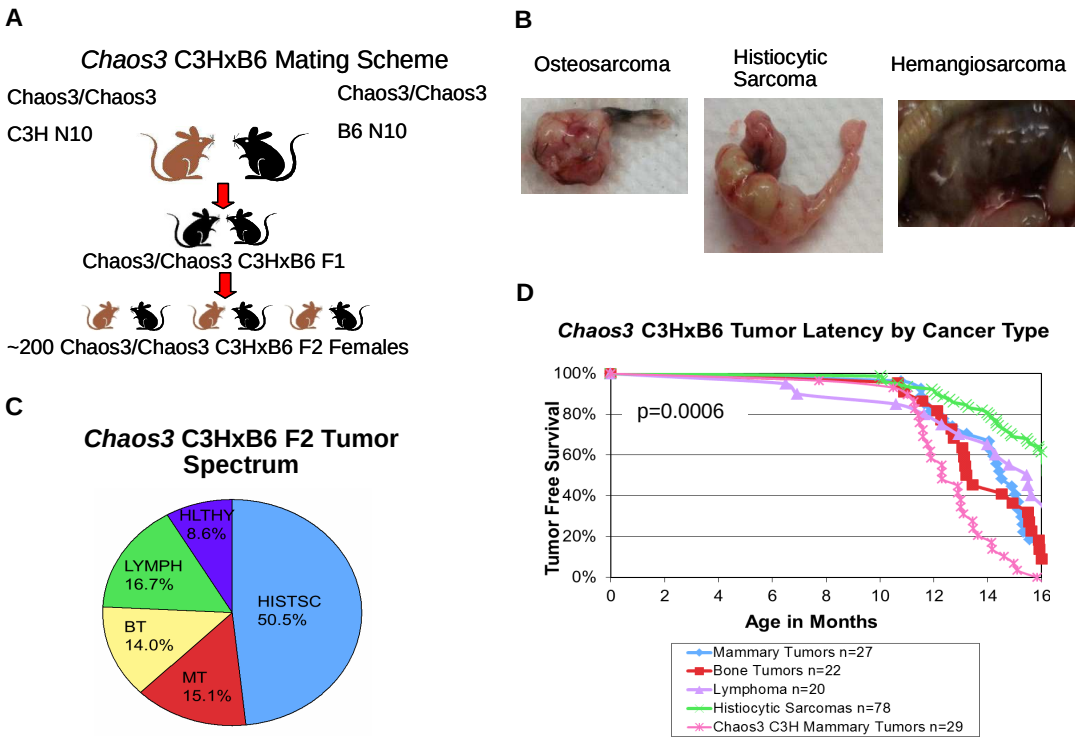


Figure 3-1: Altering background strain impacts tumor latency, tumor susceptibility, and eliminates mammary tumor specificity. (**A**) *Chaos3* C3HxB6 Mating Scheme. (**B**) *Chaos3* C3HxB6 F2 non-mammary tumors. (**C**) Tumor spectrum of *Chaos3* C3HxB6 F2 animals. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, HLTHY=healthy (no detectable cancer). (**D**) *Chaos3* tumor latency by cancer type. *Chaos3* C3HxB6 F2s have significantly increased time to mammary tumor onset (14.5 mo median tumor free survival) compared to 12.3 month median Chaos3-C3H mammary tumorigenesis (Log-rank/Mantel-Cox Test, p=.002; Gehan-Breslow-Wilcoxon Test, p=.006). Median tumor free survival for *Chaos3* C3HxB6 mice developing other tumor types is also increased: Osteosarcomas (13.3 mo), Lymphoma (15.5 mo), Histiocytic Sarcoma (16.1 mo).

Identification of loci containing mammary tumor susceptibility and resistance genes

The decreased proportion of mammary tumors in the F2s, compared to C3H, suggested a segregation of modifier alleles. A QTL analysis was conducted to identify modifier loci that cause resistance or susceptibility to mammary tumors. SNP Chips were used to genotype the mice at 377 SNPs (217 informative between C3H vs. C57BL/6J) genome-wide (Figure A3-1). This represented an average spacing of about 16 Mb, ideal for F2 QTL or modifier mapping crosses, which have relatively few recombination events. The recombination fraction (RF) plot shows successful high quality genotyping and mapping (Figure A3-2A).

Composite Interval Mapping (CIM), Multiple QTL Mapping (MQM), and SMA (single marker analysis) were used to analyze the data. The significance of mammary adenocarcinoma association for each marker was calculated as a LOD score. Figures 3-2A and A3-2B show the effects identified for makers at/near QTLs identified by cross-controlled single QTL analysis. Animals homozygous for C3H alleles are consistently associated with increased mammary tumor susceptibility (Figure 3-2B). Notably, studies of other mammary tumors models in which mice with *Trp53* or Apc^{Min} mutations are mammary tumor-susceptible on some backgrounds (FvB, BALB/c) but not on B6, have all identified recessive susceptibility loci [12-14]. A genome-wide two-dimensional QTL scan was performed on chromosomes with strong or suggestive QTLs found in at least 1 method used in the one-dimensional scan. As shown in Figures 2C and A3-2C, there is evidence of QTL epistasis. Remarkably, in

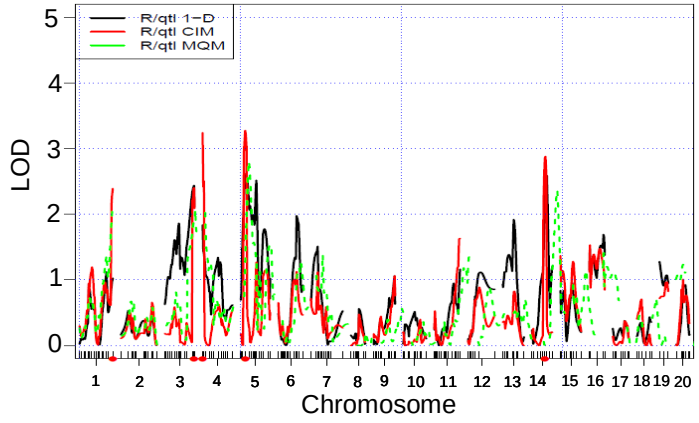
animals inheriting some single modifying alleles, mammary tumor incidence is doubled from 15.1% to 30% or reduced to <5% (Figure 3-3).

QTL analysis revealed specific regions on several chromosomes associated with modifying mammary tumor susceptibility (Figure 3-2A). Candidate genes in these regions are involved in cell proliferation (*Fgfr3*), DNA repair (*Msh4*, *Fancg*, *Fancc*, *Rad51ap1*), cell signaling (*Fbxw7*, *Sfrp1*, *Ptc1*, *Tln1*, *Pax5*), and are associated with cancer (R*ab2a*, *Rab28*, *Styk1*, *Mycbp2*) (Table 3-1 and Table A3-2). Though SNP data for the substrain C3HeB/FeJ is not available, *Fancc*, *Tln1*, *Styk1*, and *Cdkl2* have known non-synonomous SNPs between strains C3H/HEJ and C57BL/6J (Table 3-1). Examination of these genes in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database provided additional insight into candidates where amino acid change is associated with cancer.

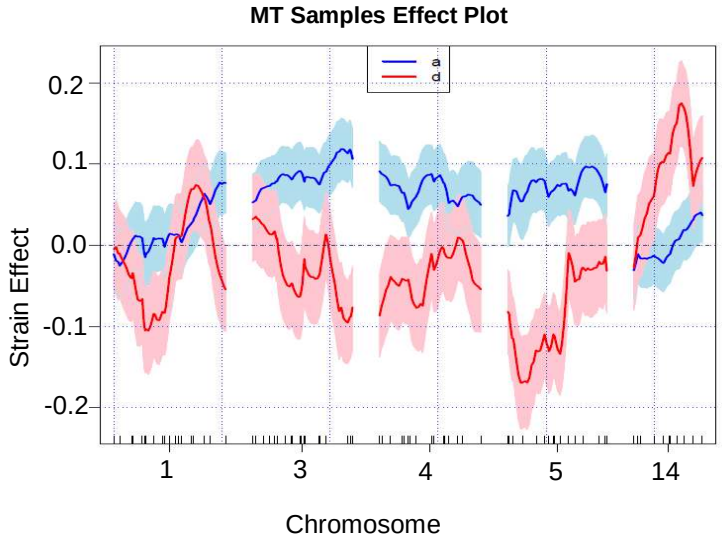
Osteosarcomas arose in 14% of the F2 animals, and QTL analysis for modifying alleles was conducted as it was for mammary tumors. Table A3-3 shows associated loci. Interestingly, two loci where B6 homozygosity was associated with increased bone tumor incidence, C3H homozygosity was associated with mammary tumor susceptibility (Figure 3-3B).

Figure 3-2: Mammary Tumor QTL analysis. (**A**) Comparison of 1-dimensional QTL scan, composite interval mapping, and multiple QTL mapping methods of QTL analysis. Red dots indicate the consensus of the most significant genomic regions associating with mammary tumors. (**B**) Effect plot of top mammary tumor QTL chromosomes. Additive effects are in blue; positive additive values indicate the increasing allele arises from the C3H background strain, and negative additive values indicate the increasing allele originates from C57/B6. Note that the top mammary tumor QTLs arise as expected from the C3H background, indicating these regions are either C3H mammary tumor susceptibility loci or mark an absence of C57/B6 mammary tumor resistance loci. Dominance effects are in red; positive values indicate a dominant effect of the increasing allele, while negative values indicate a recessive effect of the increasing allele. Note that the QTL on chromosome 14 appears to have a dominant effect from C3H. (**C**) Calculated Main and epistasis effects for QTLs using Bayes Factor analysis.

A


R/qtl 1-D
R/qtl CIM
R/qtl MQM
LOD
Chromosome

B


MT Samples Effect Plot
a
d
Strain Effect
Chromosome

C

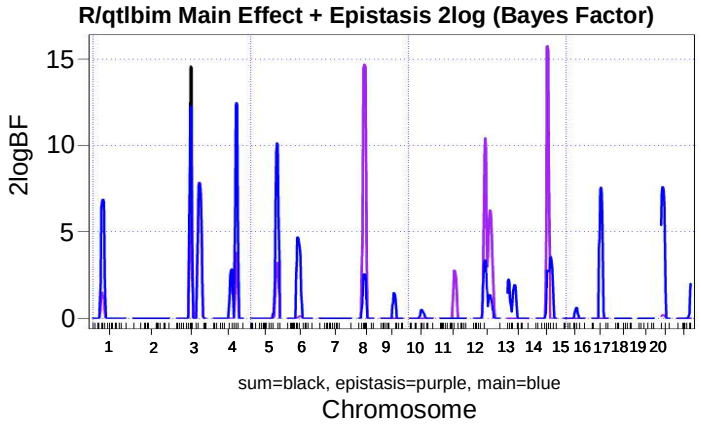

R/qtlbim Main Effect + Epistasis 2log (Bayes Factor)
2logBF
sum=black, epistasis=purple, main=blue
Chromosome

Table 3-1: Mammary Tumor QTL Candidate Genes (Condensed)

Location (Mb)	Size (Mb)	Effect Type	Gene	Can Gene Description	COSMIC Mut
chr1:194.3-197.2	2.9	C3H RS	Lamb3	Laminin beta 3 (cell attachment, migration, and organization)	18
			Hhat	Hedgehog acyltransferase (SHH signaling)	13
			Kcnh1	Potassium voltage-gated channel, subfamily H (Overexpression growth advantage to cancer cells)	16
			Irf6	Interferon regulatory factor 6 (Regulates mammary epithelial cell proliferation by similarity)	7
			Traf3ip3	TRAF3 interacting Jun N terminal kinase (JNK) activating modulator	12
chr3:79-87	8.4	C3H RS	Fbxw7	F-box/WD repeat-containing protein 7 (Notch Signaling)	289
chr3:148-154	6.0	C3H RS	Msh4	MutS protein homolog 4 (DNA damage repair)	2
			Fubp1	Far upstream element-binding protein 1 (c-myc activation)	3
chr4:41.6-46.6	5.1	C3H RS	Tln1	Talin 1 (Integrin Activation) (rs27831179)	29
			Fancg	Fanconi anemia group G protein homolog	3
			Pax5	Paired box protein Pax-5 (TF for B cell lineage)	135
chr5:33.6-45.4	11.7	C3H RS	Fgfr3	Fibroblast growth factor receptor 3	2398
			Bodl	biorientation of chromosomes in cell division	6
			Rab28	Ras-related protein Rab-28 (TGF-B signaling, Ras family)	2
chr5:79.1-93.7	14.6	C3H RS	Rassf6	Ras association domain-containing protein	8
			Cdkl2	Cyclin-dependent kinase-like 2 (rs31762832)	12
			Epha5	Ephrin type-A receptor 5 (tyrosine kinase)	50
			G3bp2	Ras GTPase-activating protein-binding protein 2	8
chr5:127-140	12.9	C3H RS	Mcm7	DNA replication licensing factor	5
			Rasa4	Ras GTPase-activating protein 4	0
			Cldn15	Claudin 15	3
chr6:120.5-131.9	11.4	B6 RR	Styk1	Tyrosine protein-kinase STYK1 (apoptosis, angiogenesis) (rs30473070)	6
			Rad51ap1	RAD51-associated protein 1	0
chr7:37.6-47.7	10.1	B6 RR	Ccne1	G1/S-specific cyclin-E1	1
chr11:110-112	10.0*	C3H RS	Map2k6	Dual specificity mitogen-activated protein kinase kinase 6	9
			Socs3	Suppressor of cytokine signaling 3 (JAK-STAT, anti-apoptosis)	1
			Aatk	Serine/threonine-protein kinase LMTK1	10
			Sumo2	Small ubiquitin-related modifier 2	1
			Cbx4	E3 SUMO-protein ligase CBX4	3
			Sox9	Transcription factor SOX-9	8

			Srsf2	Serine/arginine-rich splicing factor 2 (pre-mRNA splicing, apoptosis)	95
chr13:60.3-70.3	10.0*	C3H RS, B6 RR	Fancc	Fanconi anemia, complementation group C (rs33850907)	1
			Ptc1	Patched homolog 1 (SHH Receptor)	333
			Hsd17b3	Testosterone 17-beta-dehydrogenase 3 (Androgen, estrogene, progesterone biosynthesis->Estradiol 17beta-dehydrogenase)	0
chr14:87.2-106.9	19.6	C3H DS	Pibf1	Progesterone immunomodulatory binding factor 1	7
			Dach1	Dachshund 1 (Transcription factor, lost in some forms of metastatic cancer, correlated with poor prognosis)	8
			Mycbp2	MYC binding protein 2 (MYC regulation/activation)	37

COSMIC = Catalogue Of Somatic Mutations In Cancer; Mut=Mutation. Underlined genes have known non-synonomous SNPs between C3H/HEJ and C57BL/6J with dbSNP in ( ). *Single informative SNP in large region. RS=Recessive Susceptibility, RR=Recessive Resistance, DS=Dominant Susceptibility.

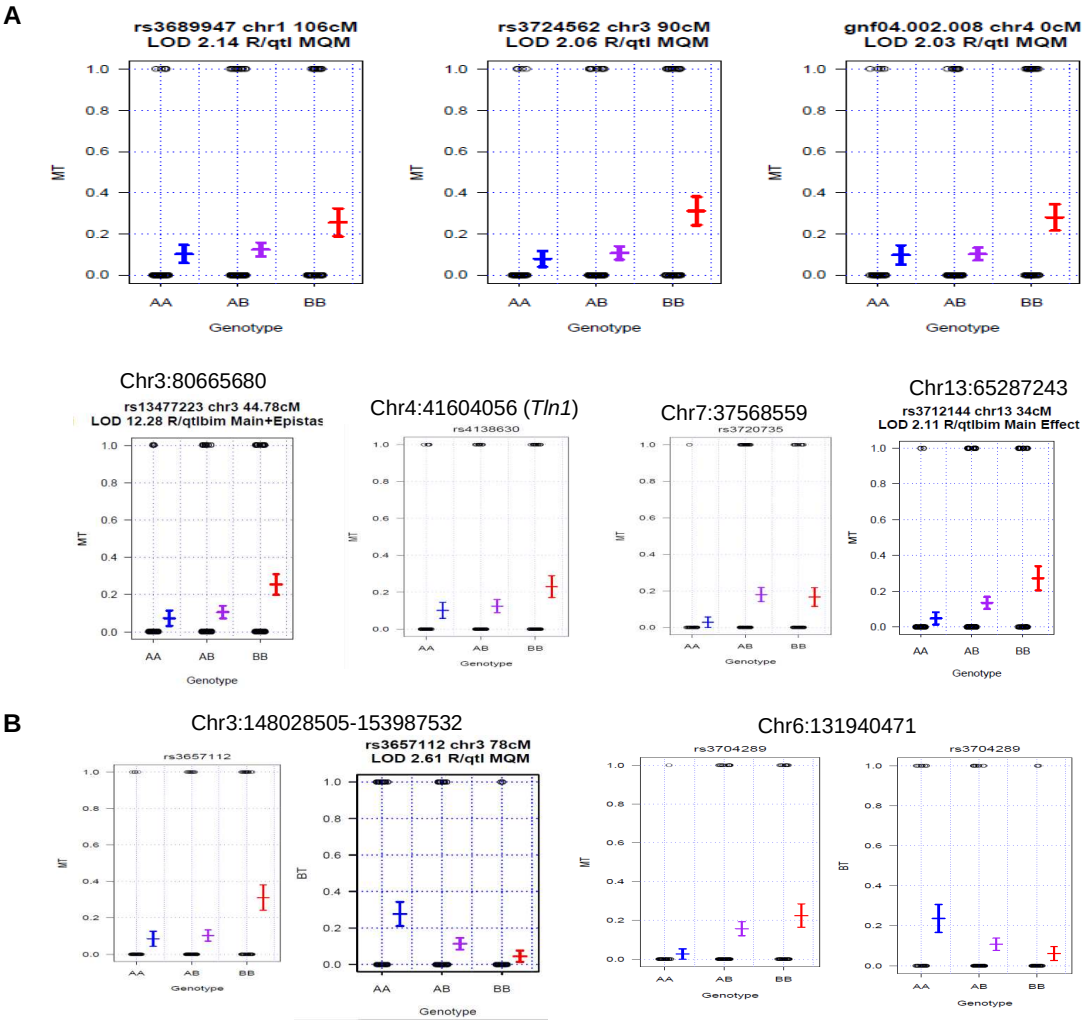


Figure 3-3: Effect of individual alleles of significance on tumor formation. (A) Effect plots of mammary tumor susceptibility and resistance loci. The *y*-axis is the percentage of the total number of animals that develop mammary tumors for an allelic genotype given on the *x*-axis. AA represents allelic homozygosity from the C57/B6 strain, AB represents heterozygosity between C3H and C57/B6, and BB represents homozygosity from C3H. Given F2 mammary tumor incidence in 15.1% of cases, genotypes deviating from the rate of incidence indicate loci of mammary tumor susceptibility or resistance. Note the markers for loci on chromosomes 1, 3, and 4 where less than 10% of the of the animals of the AA and AB genotypes develop mammary tumors, but ~30% of the animals with C3H homozygosity develop mammary tumors. The homozygous B6 genotype at SNP rs3720735 on Chr 7 and rs3712144 on Chr 13 are associated with mammary tumor resistance. C3H homozygosity on the Chr 13 SNP is also associated with mammary tumor susceptibility. **(B)** Effect plots comparing mammary tumor and bone tumor incidence at two loci. F2 bone tumor incidence was 14%. At two loci where C3H homozygosity was associated with mammary tumor susceptibility, B6 homozygosity confers bone tumor susceptibility. MT=Mammary Tumor, BT=Bone Tumor.

Identification of *Talin1* as a mammary tumor susceptibility gene

Within a Chr 4 mammary susceptibility locus, we identified a nonsynonomous *Talin 1* (*Tln1)* mutation at Chr4: 43550665 in the *Chaos3*–C3H stock mice. Analysis of 18 mouse strains, using the MPD/Jax Lab SNP database, detected no known SNPs in the *Tln1* gene at the Chr4: 43550665 locus. Additional Sanger sequencing validated that inbred C3H, B6, and 129 mice were homozygous WT (Table A3-4). The *Tln1* mutation was detected in both tumor and normal tissues, indicating the *Tln1* mutation was present in the germline. Testing both the *Chaos3*-C3H and *Chaos3*-B6 colonies revealed the *Tln1* germline mutation arose in the *Chaos3*-C3H population (Figure 3-4A, Table A3-4), likely late and spontaneously in crossing *Chaos3* into C3H to make the *Chaos3* mutation congenic in the strain. As few animals are used to generate the next generation, the *Tln1* mutation swept through the colony in the resulting progeny.

This *Tln1* mutation, occurring at a residue conserved in vertebrates, causes an E1910K glutamine to lysine, polar acidic to polar basic amino acid change. *Tln1* is responsible for linking vinculin to integrin. This bound complex creates mechanical force, triggering conformational changes, coupling integrin to cytoskeletal actin. It is the last common element of cellular signaling cascades that control integrin activation [19].

Tln1 mutation bore a significant impact on tumor type. Given C3HxB6 *Chaos3/Chaos3 Tln1/+* F1 parents, the expected *Tln1* genotype ratios in the F2 generation are 1:2:1 under a null hypothesis (that the *Tln1* mutation has no effect on tumor type). However, the F2 cohort of animals that developed mammary tumors

demonstrated twice the genotype frequency for *Tln1* mutant homozygosity (p=0.03) and increased mutant allele frequency (p=0.006) (Table A3-5).

To determine the role of *Tln1* as a mammary tumor modifier, *Chaos3*-C3H mice with homozygous mutant, heterozygous, and wild-type *Tln1* status were aged. While tumor latency was not significantly impacted, *Tln1* status significantly altered *Chaos3* mammary tumor incidence, where 92% of homozygous mutants developed mammary tumors compared to 72% of heterozygous animals (p=0.002) and 55% *Tln1* wild-type animals ($p=5.5 \times 10^{-5}$) that developed mammary tumors. (Figure 3-4B-C). However, wild-type *Tln1* alleles did not convey a resistance to all cancer types, as the animals developed a different type of cancer instead of mammary (Figure 3-4C). These results validate QTL findings and show that *Tln1* is a mammary tumor modifier.

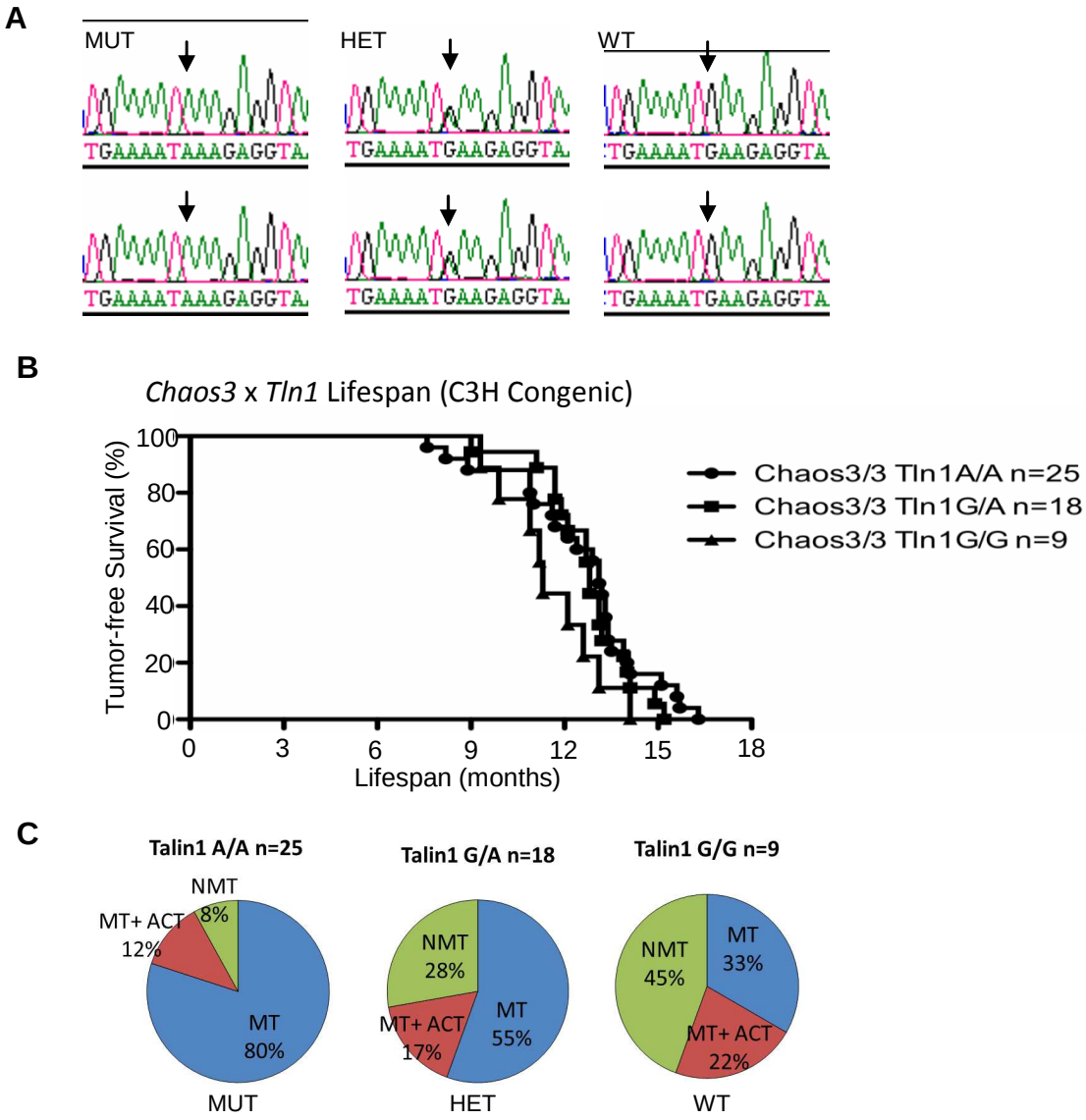


Figure 3-4: *Tln1* mutation increases mammary tumor susceptibility. (**A**) *Tln1* trace files showing germline *Tln1* mutation in the *Chaos3*-C3H colony. Refer to Tables A3-4 and A3-5. (**B**) Tumor-free survival of *Chaos3* x *Tln1* animals. The survival curves do not significantly differ (Log-rank/Mantel-Cox Test, p=0.14). (**C**) *Chaos3* x *Tln1* C3H-N10 tumor spectrum. All animals are congenic in C3H. The 'A' allele is mutant; the 'G' allele is Wild-type; Mut=Homozygous Mutant, Het=Heterozygous, WT=Wild-type. MT=Mammary Tumor; ACT =Additional Cancer Type; NMT=Non-Mammary Tumor. Note that *Tln1* status has significant impact on mammary tumor susceptibility (Chi Square values compared to Tln1 homozygous mutant mammary tumor incidence: Tln1 HET p=0.002; Tln1 WT p=5.5 x 10^{-5}).

3.4 Discussion

Tln1, validated here as a modifier of mammary tumorigenesis, has previously been shown to promote cancer progression and metastasis [20,21]. *Tln1* overexpression results in pro-survival, resistance to anoikis, and progression to metastasis, and *TLN1* has been suggested as a biomarker for tumor progression [20]. *TLN1* expression significantly increased in prostate cancer and has also been identified as a candidate driver in human breast cancer with low passenger mutation probability, as calculated by Wood et al. [6,20]. The COSMIC database shows *TLN1* mutation in 29 of 616 tumors, with 5 being in breast tumors and 3 in ovarian tumors (Figure A3-3). Our newly discovered *Tln1* mutation at the 1910 amino acid position falls within a cryptic vinculin binding domain [22]. This suggests a mutation at this residue may alter how *Tln1* interacts with vinculin and thus affect integrin activation and downstream signaling processes such as cell growth, adhesion, migration, division, survival, differentiation, and apoptosis. The rs27831179 SNP also causes a change in the TLN1 coding sequence between the C3H/HEJ and C57BL/6J strains. In the dbSNP database, there are 11 nonsynonomous SNPs and 2 SNPs in the UTRs of TLN1 present in >1% of the human population, which would be interesting to examine for association with reproductive cancers (Table A3-6).

Candidate genes in other loci identified in this study are frequently mutated in cancer, such as *FGFR3* particularly in urothelial carcinoma [23]. *Fancg* and *fancc* are Fanconi anemia (FA) genes, which are involved in the homologous recombination (HR) pathway for DNA repair. The HR pathway is altered in 51% of all ovarian cancer cases [24], and this pathway also includes BRCA1 and BRCA2, the most

penetrant genes known in heritable forms of breast cancer. Lastly, 5 RAS-related genes appeared in 4 of the 13 QTL regions. The RAS pathway is misregulated in many cancer types, including recent studies implicating it in breast cancer [25-27]. Together our findings identify loci and implicate candidate genes that modify susceptibility and resistance to mammary adenocarcinomas.

3.5 Methods and Notes

Chaos3 C3HeB/FeJ N10 animals were crossed to *Chaos3* C57BL/6J N10 animals to generate C3HxB6 F1s, which were intercrossed to generate the F2 generation. ~200 F2s were aged to a terminal endpoint of 16 months or until animals showed signs of disease. Tumor samples were collected, fixed in 10% formalin, embedded in paraffin, and examined histologically to classify tumor type.

Genomic DNA was collected from 189 *Chaos3* C3HxC57/B6 F2 females and hybridized to Goldengate Mouse LD Linkage BeadChips along with C3HxC57/B6 F1, C3H inbred, and C57/B6 inbred control DNA. Informative SNPs between C3H and C57/B6 were filtered for analysis. Allelic ratios were examined for accordance with expected 1:2:1 distribution and Hardy-Weinberg Equilibrium.

Animal SNP data were categorically analyzed between mammary tumor and non-mammary tumor groups. R/QTL version 1.19 [28] and R/qtlbim [29] were used to statistically analyze the SNP data. Composite Interval Mapping (CIM), Multiple QTL Mapping (MQM), and SMA (single marker analysis) methods were used to analyze the data.

Genome-wide one-dimensional scan

Pseudo-markers were generated at 2-cM spacing for each chromosome, and a whole genome scan. QTL with thresholds above LOD score 2 were treated as strong QTLs. Both full and additive models were analyzed. Bayesian mapping in R/qtlbim was used [29], treated as binary traits, cross as the GxE used, and epistasis was included. Interval mapping in R/QTL was used. Full model includes QTL, covariates (Talin) and QTL*covariates interaction effects. The additive model includes only additive

QTL and covariates effects. Binary trait coding was used. MQM interval mapping in R/QTL was used. Co-factors selected from top QTLs identified with R/qtl additive models were used as starting set, and the default setting to trim down the set of cofactors. Alternatively, Co-factors were selected automatically by picking 50 markers and then backward eliminated. Binary trait coding used. Composite interval mapping in R/QTL was used. Traits were coded as numerical traits as CIM not support binary traits.

Genome-wide two--dimensional scan

Pair-wise scans were performed on all-markers. The likelihood from the full model (pseudo-marker pair and the interaction between them) and the null model (no genetic effect) was compared and LOD scores were calculated. In addition, LOD scores from comparing the likelihood from the full model and the additive model (with only the main effects of pseudo-markers and but no interaction) were also calculated. qb.scantwo, implemented in R/qtlbim, scanned the Baysian models for epistasis.

Multiple regression

QTL and possible QTL*QTL interactions identified from a one and two-dimensional scans were fit into multiple regression models. From the model, variations of the phenotype in the models were estimated. Probabilities (p-values) for the significance of terms in the multiple regression models were calculated.

SMA

A simple linear regression model for genotypic tests was used to provide a general test of association in disease-by-genotype tables.

Acknowledgements

This study was supported by NIH training grants IT32HDO57854 and 5T32GM007617 that supported M.D.W.; Empire State Stem Cell Fund contract numbers C026442 and C024174 to J.C.S.

Author Contributions

M.D.W. and J.C.S designed experiments; experiments were conducted by M.D.W.; P.K. and L.S. performed QTL analyses; T.L.S. conducted histopathology; J.G.M. oversaw QTL analyses; J.C.S. oversaw the *Chaos3* project.

3.6 References

1. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* **343**, 78-85 (2000).
2. Ford, D. et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **62**, 676-89 (1998).
3. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. *Br J Cancer* **83**, 1301-8 (2000).
4. Antoniou, A.C. et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* (2008).
5. Sjoblom, T. et al. The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268-74 (2006).
6. Wood, L.D. et al. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108-13 (2007).
7. de Tayrac, M. et al. Integrative genome-wide analysis reveals a robust genomic glioblastoma signature associated with copy number driving changes in gene expression. *Genes Chromosomes Cancer* **48**, 55-68 (2009).
8. Jones, S. et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* **324**, 217 (2009).
9. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-12 (2008).
10. Pharoah, P.D. et al. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* **31**, 33-6 (2002).
11. Cormier, R.T. et al. Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. *Nat Genet* **17**, 88-91 (1997).
12. Blackburn, A.C. et al. Genetic mapping in mice identifies DMBT1 as a candidate modifier of mammary tumors and breast cancer risk. *Am J Pathol* **170**, 2030-41 (2007).
13. Wang, H. et al. Identification of novel modifier loci of Apc Min affecting mammary tumor development. *Cancer Res* **67**, 11226-33 (2007).
14. Koch, J.G. et al. Mammary tumor modifiers in BALB/cJ mice heterozygous for p53. *Mamm Genome* **18**, 300-9 (2007).
15. Green, J.E. & Hudson, T. The promise of genetically engineered mice for cancer prevention studies. *Nat Rev Cancer* **5**, 184-98 (2005).
16. Ottewell, P.D., Coleman, R.E. & Holen, I. From genetic abnormality to metastases: murine models of breast cancer and their use in the development of anticancer therapies. *Breast Cancer Res Treat* **96**, 101-13 (2006).
17. Shima, N. et al. Phenotype based identification of mouse chromosome instability mutants. *Genetics* **163**, 1031-1040 (2003).
18. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-98 (2007).

19. Tadokoro, S. et al. Talin binding to integrin beta tails: a final common step in integrin activation. *Science* **302**, 103-6 (2003).
20. Desiniotis, A. & Kyprianou, N. Significance of talin in cancer progression and metastasis. *Int Rev Cell Mol Biol* **289**, 117-47 (2011).
21. Sakamoto, S., McCann, R.O., Dhir, R. & Kyprianou, N. Talin1 promotes tumor invasion and metastasis via focal adhesion signaling and anoikis resistance. *Cancer Res* **70**, 1885-95 (2010).
22. Goult, B.T. et al. The domain structure of talin: residues 1815-1973 form a five-helix bundle containing a cryptic vinculin-binding site. *FEBS Lett* **584**, 2237-41 (2010).
23. Tomlinson, D.C., Baldo, O., Harnden, P. & Knowles, M.A. FGFR3 protein expression and its relationship to mutation status and prognostic variables in bladder cancer. *J Pathol* **213**, 91-8 (2007).
24. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
25. Baljuls, A. et al. The tumor suppressor DiRas3 forms a complex with H-Ras and C-RAF and regulates localization, dimerization and kinase activity of C-RAF. *J Biol Chem* (2012).
26. Eastman, B., Jo, M., Webb, D., Takimoto, S. & Gonias, S. A transformation in the mechanism by which the urokinase receptor signals provide a selection advantage for estrogen receptor-expressing breast cancer cells in the absence of estrogen. *Cellular Signal* (2012).
27. Xue, M. et al. Insulin-like Growth Factor-1 Receptor (IGF-1R) Kinase Inhibitors in Cancer Therapy: Advances and Perspectives. *Curr Pharm Des* **18**, 2901-13 (2012).
28. Broman, K.W., Wu, H., Sen, S. & Churchill, G.A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889-90 (2003).
29. Banerjee, S., Yandell, B.S. & Yi, N. Bayesian quantitative trait loci mapping for multiple traits. *Genetics* **179**, 2275-89 (2008).

CHAPTER 4

Impact of DNA checkpoint and repair deficiency on carcinogenesis in MCM-deficient mice

Marsha D. Wallace[1,2], Xin Li[2], Teresa L. Southard[1,3], John C. Schimenti[1,2,4]

Affiliations:

[1]Department of Biomedical Sciences

[2]Department of Molecular Biology and Genetics

[3]Section of Anatomic Pathology

[4]Center for Vertebrate Genomics

Cornell University, Ithaca, NY 14853, USA

One Sentence Summary: Perturbation of DNA damage response/repair pathways impacts carcinogenesis in a gender-specific manner in mice with defective DNA replication machinery.

4.1 Abstract

Many genes are misregulated or altered at low frequencies in human cancers, but together comprise significant alterations in key pathways, particularly DNA checkpoint and repair pathways. Recent studies show that defective DNA replication machinery can result in genomic instability (GIN) and carcinogenesis. To understand the roles of DNA damage repair (DDR) genes on carcinogenesis in mutants defective for core DNA replication machinery, we utilized the novel *Mcm4*$^{Chaos3/Chaos3}$ ("*Chaos3*" or "C3") cancer mouse model. MCM4 is a highly conserved subunit of the MCM2-7 DNA replicative helicase, required for DNA replication and DNA licensing mechanisms to prevent re-replication of the genome. We generated double mutant lines between *Chaos3* and *Atm*, *p21*, *Chk2*, *Hus1*, and *Blm*. We find that *Chaos3* animals deficient in the *Atm* pathway have decreased tumor latency and/or increased tumor susceptibility. Tumor latency and susceptibility differed between genders, with females demonstrating an overall greater cancer susceptibility to *Atm* and *p21* deficiency than males. These findings indicate that deficiency in the ATM DDR pathway impacts tumor susceptibility and latency in MCM-deficient cancer-prone mice, and this impact is modified by gender.

4.2 Introduction

Genomic studies have shown that many genes are misregulated or altered at low frequencies in human cancers, but together comprise significant alterations in key pathways [1-3]. Of particular interest is the role that DNA checkpoint and repair pathways play. Breast and ovarian cancer type 1 susceptibility and 2 (*BRCA1* and *BRCA2*) are known susceptibility genes in inherited forms of human breast and ovarian cancer [4-6]. These two genes are altered in 33% of serous ovarian cancer cases and contribute to an overall 51% of cases deficient in the Homologous Recombination (HR) pathway [2]. In HR, DNA damage sensors such as Ataxia Telangiectasia Mutated (ATM) detect DNA damage including double-strand breaks and trigger BRCA1, the Fanconi Anemia (FA) core complex, and subsequent downstream signaling for HR-mediated repair [7]. ATM and Ataxia-Telangiectasia-and-Rad3-related (ATR) head additional DNA checkpoint and repair pathways, including signaling to Tumor Protein 53 (TP53), the tumor suppressor most frequently compromised in human cancer (Figure 4-1) [8]. DNA damage response (DDR) pathways are responsible for helping maintain genomic stability and suppressing tumorigenesis. DDR genes also target components of the core DNA replication complex, including the MCM helicase. MCM2 is a direct target of ATR, and MCM3 is a target of ATM [9,10].

Accumulating evidence shows associations between defects in, or missregulation of, core replication machinery and cancer. The highly conserved MCM2-7 DNA replicative helicase is an essential component of pre-replication (pre-RC) complexes [11]. These complexes are “licensed” at replication origins for activation

in S phase, and regulatory mechanisms inhibit reloading of the MCMs during S phase to prevent re-replication of the genome [11]. MCM mutation and deficiency results in phenotypes of synthetic lethality, growth retardation, decreased cellular proliferation, GIN, and early onset of cancer in mice [12]. MCM2 deficiency leads to GIN and aggressive tumor susceptibility in mice, specifically lymphomas [12-15]. Elevated expression of CDT1, a protein that helps load MCMs onto the origins of replication, is associated with increased chromosomal instability and tumor growth in lung cancers when p53 is mutated [16].

To understand the roles of DDR genes on cancer incidence and tumor-free survival in MCM deficient mice, we utilized the *Mcm4*$^{Chaos3/Chaos3}$ ("*Chaos3*") cancer mouse model. The *Mcm4*^{Chaos3} mutation demonstrated that mutant alleles of essential replication proteins can cause GIN and cancer [17,18]. The *Chaos3* cancer model was isolated in an *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis screen for mutations causing genomic instability (GIN) [17]. The nonsynonomous mutation identified in *Mcm4* caused a dramatic (20 fold) increase in micronuclei, a hallmark of GIN [18]. In *Chaos3* mouse embryonic fibroblasts (MEFs), significant chromosome breakage (compared to wild-type controls) occurred only under conditions of replication stress, indicating that the damage was a consequence of a defect(s) in some aspect of DNA replication [17,18]. Work in yeast carrying the *Chaos3* mutation suggests that the stress can cause the replication fork to collapse, leading to double-strand breaks which then activates the HR (homologous recombination) pathway [19]. Evidence from other model systems support the conclusion that MCM dysfunction can cause DNA damage and rearrangements [20]. *Chaos3* cells demonstrate increased levels of RAD51 and BLM

foci [21]. Additionally, upregulation of p53 and p21 are observed in *Chaos3* MEFs [22]. *Chaos3* mice succumb to cancer at a mean latency of 12 months [17]. *Chaos3* animals also deficient for p53 have decreased time to cancer onset [22]. This evidence indicates that defective DNA replication machinery sensitizes cells to replication stress, and DDR pathways are activated. Therefore, models in which the DNA replication machinery is defective may be sensitive to DDR perturbation. Here we generate double mutant lines between *Chaos3* and an additional DDR gene to examine the impact of components in DDR pathways on carcinogenesis.

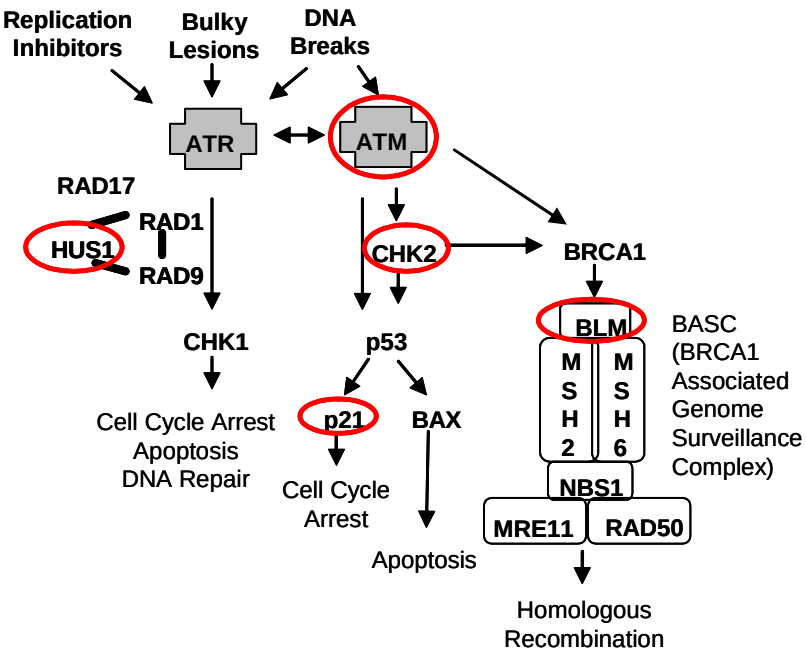


Figure 4-1: DNA Damage Response Pathways. Key genes in DDR pathways are shown with ATR and ATM DNA damage sensors emphasized in gray boxes. Genes perturbed in this study to determine impact on carcinogenesis are indicated by red ovals.

4.3 Results

To determine the impact of DDR deficiency on carcinogenesis in *Chaos3* mice, we generated *Chaos3* mice deficient for *Atm*, *p21* (*Cdkn1a*), *Chk2* (*checkpoint kinase 2, Chek2*), *Hus1*, or *Blm*. Mice were aged for eighteen months or until animals showed clinical signs of disease.

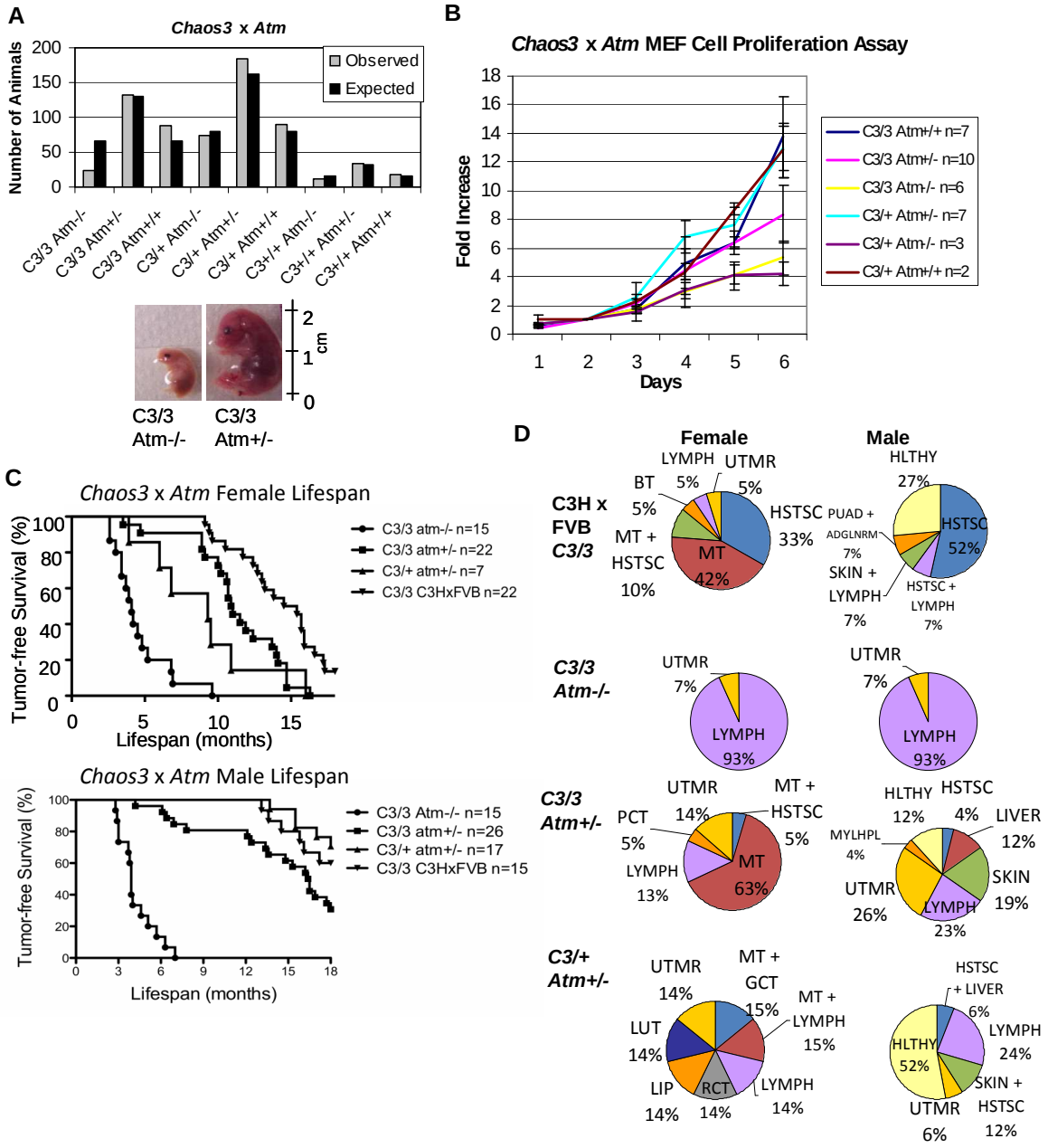
Atm deficiency impacts *Chaos3* viability, cell proliferation, tumor latency, and tumor susceptibility

ATM is a serine/threonine kinase that senses DNA double-strand breaks (DSBs) and triggers several key events, such as H2AX phosphorylation at the site of breaks and phosphorylation of downstream targets such as CHK2, to activate the DNA damage checkpoint leading to cell cycle arrest or apoptosis (Figure 4-1) [23]. ATM deficiency is associated with the development of lymphomas and leukemias in humans and mice, with $Atm^{-/-}$ mice developing malignant lymphomas at 2-4 months of age [24,25]. Generating the *Chaos3* x *Atm* mice, the *Chaos3/Chaos3* (*C3/3*) *Atm-/-* genotype resulted in semi-lethality with only 25 observed animals out of 65 expected in 648 weanlings (Figure 4-2A). Timed-matings indicated that embryonic lethality occurred between E13.5 and E18.5, with small and underdeveloped *C3/3 Atm-/-* embryos observed at E18.5 (Figure 4-2A). This semi-lethality suggested that cells were accumulating unrepaired DNA damage beyond the point of viability. Alternatively, a separate checkpoint pathway may have been activated and triggered cell death.

To determine the impact of Atm deficiency on *Chaos3* cell proliferation, cell proliferation assays were conducted on *Chaos3* x *Atm* MEFs. *C3/3 Atm+/-* cells grew significantly slower than *C3/3*, *C3/+ Atm+/-*, or *C3/+ Atm+/+* cells (Figure 4-2 B).

When *Chaos3* x *Atm* mice were aged, the strong lymphoma phenotype of Atm homozygous mutants obscured the effect on mammary tumorigenesis. *C3/3 Atm-/-* mice have significantly decreased time to tumor onset compared to *C3/3* alone, but they succumb to lymphoma at ~2-4 months of age, similar to *Atm-/-* mice (Figure 4-2C, Table A4-1) [24]. However, the impact of *Atm* deficiency on mammary tumorigenesis was evident in *C3/3 Atm+/-* and *C3/+ Atm+/-* animals. *C3/3 Atm+/-* females have a median tumor latency of 10.95 months and *C3/+ Atm+/-* females have a median tumor latency of 9.3 months, a significantly decreased latency compared to *C3/3* alone (14.95 months) (Respectively: LRMCT p=0.001, p=0.0027; GBWT p=0.0031, p=0.0005). *C3/3 Atm+/-* males neared statistical significance for decreased tumor latency (LRMCT p=0.0751; GBWT p=0.0729), and *C3/+ Atm+/-* male tumor latency was similar to *C3/3* alone (LRMCT p=0.472; GBWTp=0.4339) (Figure 4-2C, Table A4-1). The tumor spectrum of *Chaos3* x *Atm* mice shifts away from histiocytic sarcomas (from 33% to ≤5% in females and from 52% to ≤6% in males) to an increased incidence of lymphoma and other cancer types. The spectra also differ between genotype and gender, with *C3/+ Atm+/-* females being more susceptible to cancer than males.

Figure 4-2: *Atm* deficiency impacts *Chaos3* viability, cell proliferation, tumor latency, and tumor susceptibility. (**A**) **Top-**Genotype distribution of 648 *Chaos3* x *Atm* weanlings. The *C3/3 Atm-/-* genotype causes semi-lethality (Chi square p=. $6.03*10^{-6}$; Fisher's Exact p=$4.5*10^{-5}$). **Bottom-** Littermate pups from embryonic day 18.5. Note the poor development of the *C3/3 Atm-/-* embryo. (**B**) *Chaos3* x *Atm* MEF cell proliferation assay. *C3/3 Atm+/-* cells grow significantly slower than *C3/3*, *C3/+ Atm+/-*, or *C3/+ Atm+/+* cells. (**C**) *Chaos3* x *Atm* tumor latency. *C3/3 Atm-/-* mice have significantly decreased time to tumor onset, similar to *Atm-/-*. *C3/3 Atm+/-* and *C3/+ Atm+/-* females have significantly decreased tumor latency compared to *C3/3* alone (Respectively: LRMCT p=0.001, p=0.0027; GBWT p=0.0031, p=0.0005). *C3/3 Atm+/-* males neared statistical significance for decreased tumor latency (LRMCT p=0.0751; GBWT p=0.0729), and *C3/+ Atm+/-* male tumor latency was similar to *C3/3* alone (LRMCT p=0.472; GBWTp=0.4339). LRMCT= Log-rank/Mantel-Cox Test; GBWT= Gehan-Breslow-Wilcoxon Test. (**D**) Tumor spectrum of *Chaos3* x *Atm* mice. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, HLTHY=healthy (no detectable cancer), PCT=plasma cell tumor, RCT=round cell tumor, GCT=granulosa cell tumor, LIP=lipoma, LUT=luteoma, SKIN=skin tumor, LIVER=liver tumor, PUAD=pulmonary adenoma, MYLHPL=myeloid hyperplasia, ADGLNRM=adrenal ganglioneuroma, UTMR=unknown tumor type. Note that tumor spectrum differs between genotype and gender, and that *C3/+ Atm+/-* females are more susceptible to cancer than males.


A
Chaos3 x Atm
Number of Animals
Observed
Expected
C3/3 Atm-/-
C3/3 Atm+/-
C3/3 Atm+/+
C3/+ Atm-/-
C3/+ Atm+/-
C3/+ Atm+/+
C3+/+ Atm-/-
C3+/+ Atm+/-
C3+/+ Atm+/+
cm
C3/3 Atm-/-
C3/3 Atm+/-
B
Chaos3 x Atm MEF Cell Proliferation Assay
Fold Increase
Days
C3/3 Atm+/+ n=7
C3/3 Atm+/- n=10
C3/3 Atm-/- n=6
C3/+ Atm+/- n=7
C3/+ Atm-/- n=3
C3/+ Atm+/+ n=2
C
Chaos3 x Atm Female Lifespan
Tumor-free Survival (%)
Lifespan (months)
C3/3 atm-/- n=15
C3/3 atm+/- n=22
C3/+ atm+/- n=7
C3/3 C3HxFVB n=22
Chaos3 x Atm Male Lifespan
Tumor-free Survival (%)
Lifespan (months)
C3/3 Atm-/- n=15
C3/3 atm+/- n=26
C3/+ atm+/- n=17
C3/3 C3HxFVB n=15
D
Female
Male
C3H x FVB C3/3
LYMPH 5%
UTMR 5%
BT 5%
HSTSC 33%
MT + HSTSC 10%
MT 42%
HLTHY 27%
PUAD + ADGLNRM 7%
HSTSC 52%
SKIN + LYMPH 7%
HSTSC + LYMPH 7%
C3/3 Atm-/-
UTMR 7%
LYMPH 93%
UTMR 7%
LYMPH 93%
C3/3 Atm+/-
UTMR 14%
MT + HSTSC 5%
PCT 5%
LYMPH 13%
MT 63%
HLTHY 12%
HSTSC 4%
LIVER 12%
MYLHPL 4%
SKIN 19%
UTMR 26%
LYMPH 23%
C3/+ Atm+/-
UTMR 14%
MT + GCT 15%
MT + LYMPH 15%
LUT 14%
LIP 14%
RCT 14%
LYMPH 14%
HSTSC + LIVER 6%
LYMPH 24%
HLTHY 52%
SKIN + HSTSC 12%
UTMR 6%

Chk2 deficiency impacts tumor latency in *Chaos3* females and cancer susceptibility in males

CHK2 is a phosphorylation target of ATM and serves as a downstream effector of the DSB checkpoint response [26]. When activated, the CHK2 can phosphorylate p53, protecting it from ubiquitination by MDM2 and subsequent degradation [26]. In addition to p53, CHK2 can also phosphorylate BRCA1, meaning that CHK2 activation can lead to cell cycle arrest through p21 inhibition of cyclin-dependent kinases (CDKs), apoptosis through BAX, or homologous recombination through BRCA (Figure 4-1). Unlike Atm null mice, Chk2 null mice do not spontaneously develop tumors [27]. *Chaos3* x *Chk2* mice were aged, and *C3/3 Chk2-/-* females were found to have decreased time to tumor onset compared to *C3/3* alone (LRMCT p=0.0189, GBWT p=0.027) (Figure 4-3A, Figure A4-1, Table A4-1). Interestingly, *C3/+ Chk2+/-* male mice are more susceptible to cancer than *C3/+* alone (Figure 4-3B).

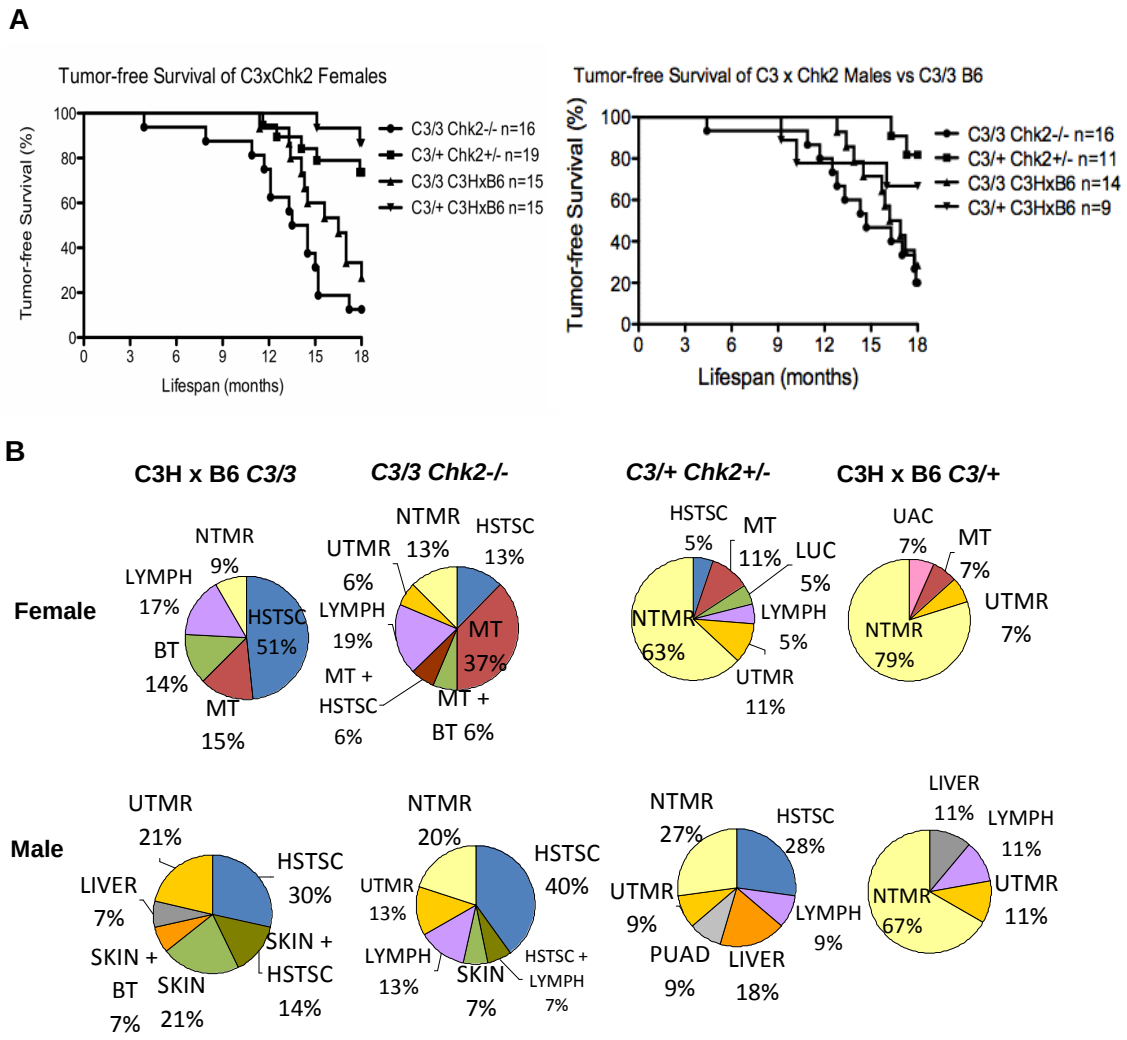


Figure 4-3: *Chk2* deficiency impacts tumor latency in *Chaos3* females and cancer susceptibility in males. (**A**) *Chaos3* x *Chk2* tumor latency. *C3/3 Chk2-/-* female mice have significantly decreased time to tumor onset than *C3/3* alone ("DDR C3/3 cohort" GBWT p=0.0581; "C3/3 C3HxB6 F2 cohort" LRMCT p=0.0189, GBWT p=0.027; see Methods and Figure A4-1. GBWT=Gehan-Breslow-Wilcoxon Test ; LRMCT= Log-rank Mantel-Cox Test. (**B**) Tumor spectrum of *Chaos3* x *Chk2* mice. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, NTMR=no tumor, SKIN=skin tumor, LIVER=liver tumor, PUAD=pulmonary adenocarcinoma, UAC=uterine adenocarcinoma, LUC=lung carcinoma, UTMR=unknown tumor type. Note that *C3/+ Chk2+/-* male mice are more susceptible to cancer than *C3/+* alone.

p21 deficiency impacts tumor latency in *Chaos3* mice and tumor susceptibility in *Chaos3* females

p21 is a cyclin-dependent kinase inhibitor and downstream target of p53 that halts cell cycle progression when activated (Figure 4-1). It functions by blocking the activity of cyclin-CDK complexes (CDK2 and CDC2), and can inhibit proliferating cell nuclear antigen (PCNA) and therefore DNA replication [28]. Unlike *Atm* deficient mice, *p21* deficient mice do not have an increased susceptibility to cancer [29]. A previous study suggested that p21 upregulation was unlikely to contribute to tumor suppression in *Chaos3*-C3HxB6 mice because the mean tumor latency was similar to *Chaos3*-C3H mice [22]. However, this did not take the delayed cancer onset of *C3/3* in B6 or C3H x B6 backgrounds into consideration, and *C3/3* C3H x B6 or B6 controls were not included in that report. Here we expand that study to include both male and female mice and additional genotypes, including *C3/3* controls, as well as examination of cancer incidence. When *Chaos3* x *p21* mice were aged, both *C3/3 p21-/-* males and females had significantly decreased time to tumor onset than *C3/3* alone (GBWT male p=0.046, female p=0.0055) (Figure 4-4A, Table A4-1). *C3/3 p21+/-* females, but not males, also had significantly decreased tumor latency compared to *C3/3* alone (GBWT female p=0.0223, male p= 0.3813) (Figure 4-4A, Table A4-1). Female *C3/+ p21-/-* and *C3/+ p21+/-* mice had increased cancer susceptibility. Cancer developed in 55% of *C3/+ p21-/-* females and 42% of *C3/+ p21+/-* females, compared to only 21% of female *C3/+* alone and 18% of male *C3/+ p21+/-* mice (Figure 4-4B).

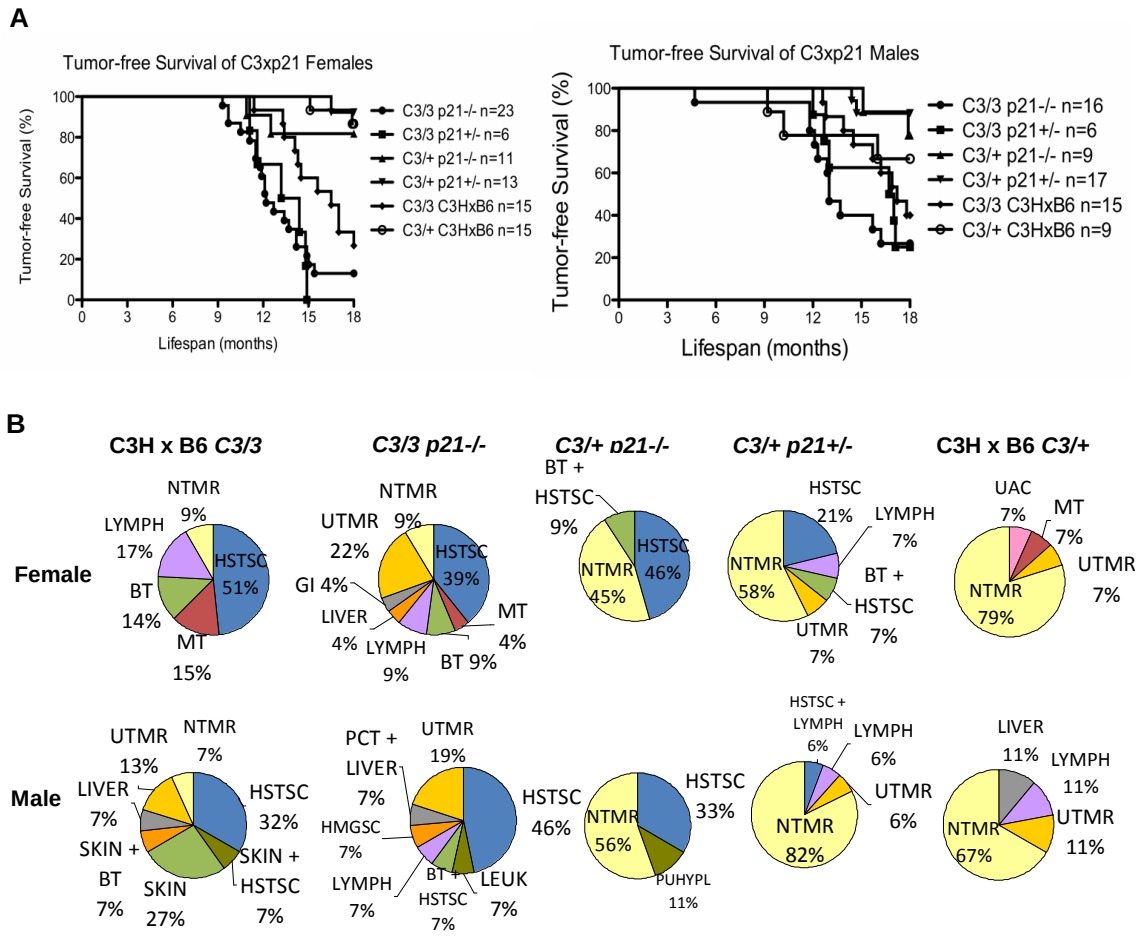
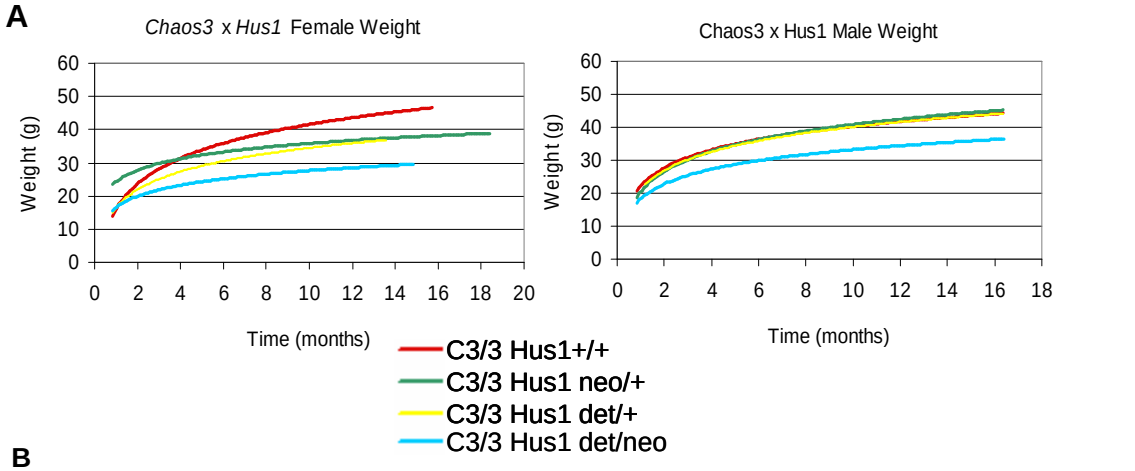


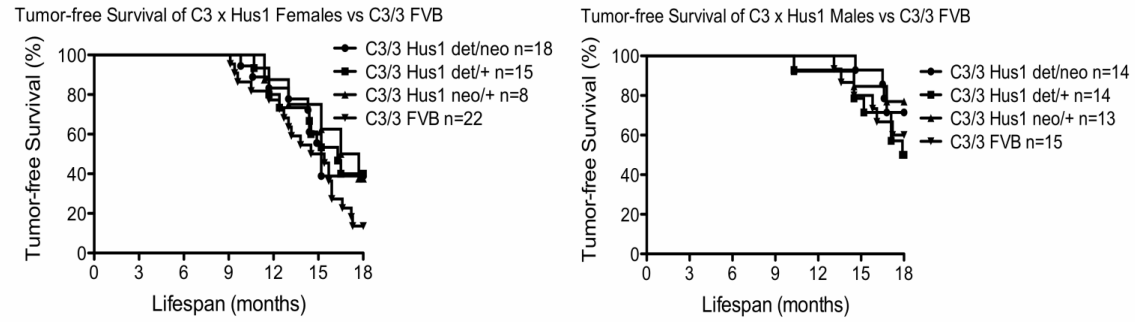
Figure 4-4: *p21* deficiency impacts *Chaos3* tumor latency in males and females and tumor susceptibility in females. (**A**) *Chaos3* x *p21* tumor latency. *C3/3 p21-/-* male and female mice have significantly decreased time to tumor onset than *C3/3* alone (GBWT male p=0.046, female p=0.0055). *C3/3 p21+/-* females, but not males, also have significantly decreased tumor latency compared to *C3/3* alone (GBWT p=0.0223, p= 0.3813). GBWT= Gehan-Breslow-Wilcoxon Test . (**B**) Tumor spectrum of *Chaos3* x *p21* mice. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, NTMR=no tumor, PCT=plasma cell tumor, SKIN=skin tumor, LIVER=liver tumor, PUHYPL=pulmonary hyperplasia, HMGSC=hemangiosarcoma, GI=Gastro-intestinal, UAC=uterine adenocarcinoma, LEUK=leukemia, UTMR=unknown tumor type. Note that *C3/+ p21-/-* females are more susceptible to cancer than *C3/+* alone, and *C3/+ p21+/-* females are more susceptible to cancer than males.

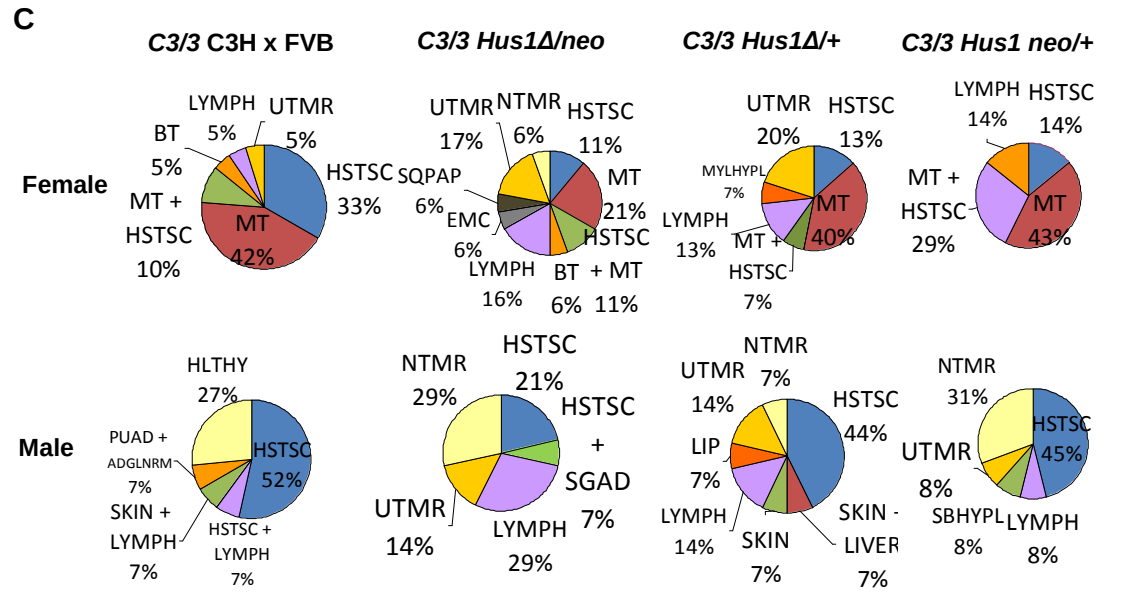
Hus1 deficiency impacts body size but not tumor latency or cancer susceptibility in *Chaos3/3* mice

The study of genes in the ATR pathway is complicated by lethality that arises when any of the components are null. However, mice with hypomorphic alleles of *Hus1* are viable [30]. Compared to wild-type levels, *Hus1* mutant expression is as follows: *Hus1neo/+* 71.4%, *Hus1Δ/+* 43.5%, and *Hus1Δ/neo* 20.8% [30]. Using hypomorphic allele combinations of *Hus1*, we were able to examine the effects of ATR pathway deficiency on *Chaos3* development and carcinogenesis. *Chaos3 Hus1Δ/neo* mice have dwarfed body size with abnormal craniofacial features, similar to *Hus1* x *Atm* mice [31], and significantly reduced body weights (Figure 4-5A). *Chaos3 Hus1Δ/+* females also have significantly lower body weights (Figure 4-5A). However, *C3/3 x Hus1* mice do not have significantly different time to tumor onset than *C3/3* alone (Figure 4-5B, Table A4-1). Interestingly, *Chaos3* males in the C3H x FVB background are more resistant to tumorigenesis than females (Chi Square p=0.004) (Figure 4-5C).

Figure 4-5: *Hus1* deficiency does not impact tumor latency or cancer susceptibility in *Chaos3/3* mice. (**A**) *Chaos3* x *Hus1* body weight. *C3/3* x *Hus1 det/neo* animals have significantly lower body weights than C3/3 alone (ANCOVA: male $p=5.96*10^{-08}$; female $p=4.04*10^{-12}$). *C3/3 Hus1 det/+* females also have significantly lower weights than *C3/3*, but greater than *C3/3 Hus1 det/neo* (ANCOVA: respectively $p=4.31*10^{-06}$; $p=7.32*10^{-4}$). (**B**) *Chaos3* x *Hus1* tumor latency. *C3/3 x Hus1* mice do not have significantly different time to tumor onset than *C3/3* alone (Female: Log-rank Mantel-Cox Test p=0.2872, Gehan-Breslow-Wilcoxon Test p=0.1402; Male: Log-rank Mantel-Cox Test p=0.5117, Gehan-Breslow-Wilcoxon Test p=0.7731). (**C**) Tumor spectrum of *Chaos3* x *Hus1* mice. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, PCT=plasma cell tumor, RCT=round cell tumor, GCT=granulosa cell tumor, LIP=lipoma, LUT=luteoma, SKIN=skin tumor, LIVER=liver tumor, PUAD=pulmonary adenoma, MYLHPL=myeloid hyperplasia, ADGLNRM=adrenal ganglioneuroma, PUAD=pulmonary adenocarcinoma, UAC=uterine adenocarcinoma, LUC=lung carcinoma, UTMR=unknown tumor type, NTMR=no tumor. Note that males in the C3H x FVB background are more resistant to *Chaos3* tumorigenesis than females (Chi Square p=0.004).


A
Chaos3 x Hus1 Female Weight
Weight (g)
Time (months)
Chaos3 x Hus1 Male Weight
Weight (g)
Time (months)
C3/3 Hus1+/+
C3/3 Hus1 neo/+
C3/3 Hus1 det/+
C3/3 Hus1 det/neo


B
Tumor-free Survival of C3 x Hus1 Females vs C3/3 FVB
Tumor-free Survival (%)
Lifespan (months)
C3/3 Hus1 det/neo n=18
C3/3 Hus1 det/+ n=15
C3/3 Hus1 neo/+ n=8
C3/3 FVB n=22
Tumor-free Survival of C3 x Hus1 Males vs C3/3 FVB
Tumor-free Survival (%)
Lifespan (months)
C3/3 Hus1 det/neo n=14
C3/3 Hus1 det/+ n=14
C3/3 Hus1 neo/+ n=13
C3/3 FVB n=15


C
C3/3 C3H x FVB
C3/3 Hus1Δ/neo
C3/3 Hus1Δ/+
C3/3 Hus1 neo/+
Female
LYMPH 5%
UTMR 5%
BT 5%
HSTSC 33%
MT + HSTSC 10%
MT 42%
UTMR 17%
NTMR 6%
HSTSC 11%
SQPAP 6%
MT 21%
EMC 6%
HSTSC 6%
LYMPH 16%
BT + MT 11%
UTMR 20%
HSTSC 13%
MYLHYPL 7%
LYMPH 13%
MT 40%
MT + HSTSC 7%
LYMPH 14%
HSTSC 14%
MT + HSTSC 29%
MT 43%
Male
HLTHY 27%
PUAD + ADGLNRM 7%
HSTSC 52%
SKIN + LYMPH 7%
HSTSC + LYMPH 7%
NTMR 29%
HSTSC 21%
HSTSC + SGAD 7%
UTMR 14%
LYMPH 29%
NTMR 7%
UTMR 14%
HSTSC 44%
LIP 7%
LYMPH 14%
SKIN 7%
SKIN - LIVER 7%
NTMR 31%
HSTSC 45%
UTMR 8%
SBHYPL 8%
LYMPH 8%

Blm deficiency has cryptic effects on tumor latency and cancer susceptibility in *Chaos3* mice

BLM is a member of the Rec-Q helicase family and component of the BRCA1 Associated Genome Surveillance Complex (BASC) involved in control of homologous recombination [32]. Inherited autosomal recessive *Blm* mutation in humans causes Bloom's Syndrome, and affected individuals have a significantly increased risk of developing cancer with early onset [33,34]. Complete absence of *Blm* in mice results in lethality (site ref), so *C3/3 Blm+/-* and *C3/+ Blm+/-* were aged. *C3/3 Blm+/-* females do not have significantly different time to tumor onset than *C3/3* alone (LRMCT p=0.8469, GBWT p=0.9013) (Figure 4-6A, Table A4-1). Interestingly however, *C3/3 Blm+/-* males near statistical significance for delayed tumor onset (LRMCT p=0.059, GBWT p=0.081) (Figure 4-6A, Table A4-1). Additionally, male *C3/3 Blm+/-* mice may have increased resistance to tumorigenesis compared to their female counterparts or to *C3/3* alone. However, *C3/+ Blm+/-* males do not have a delayed tumor onset or and have increased cancer susceptibility compared to *C3/+ Blm+/-* females or *C3/+* alone. Therefore, the role of *Blm* deficiency on tumor onset and progression in *Chaos3* cancer-prone males is unclear.

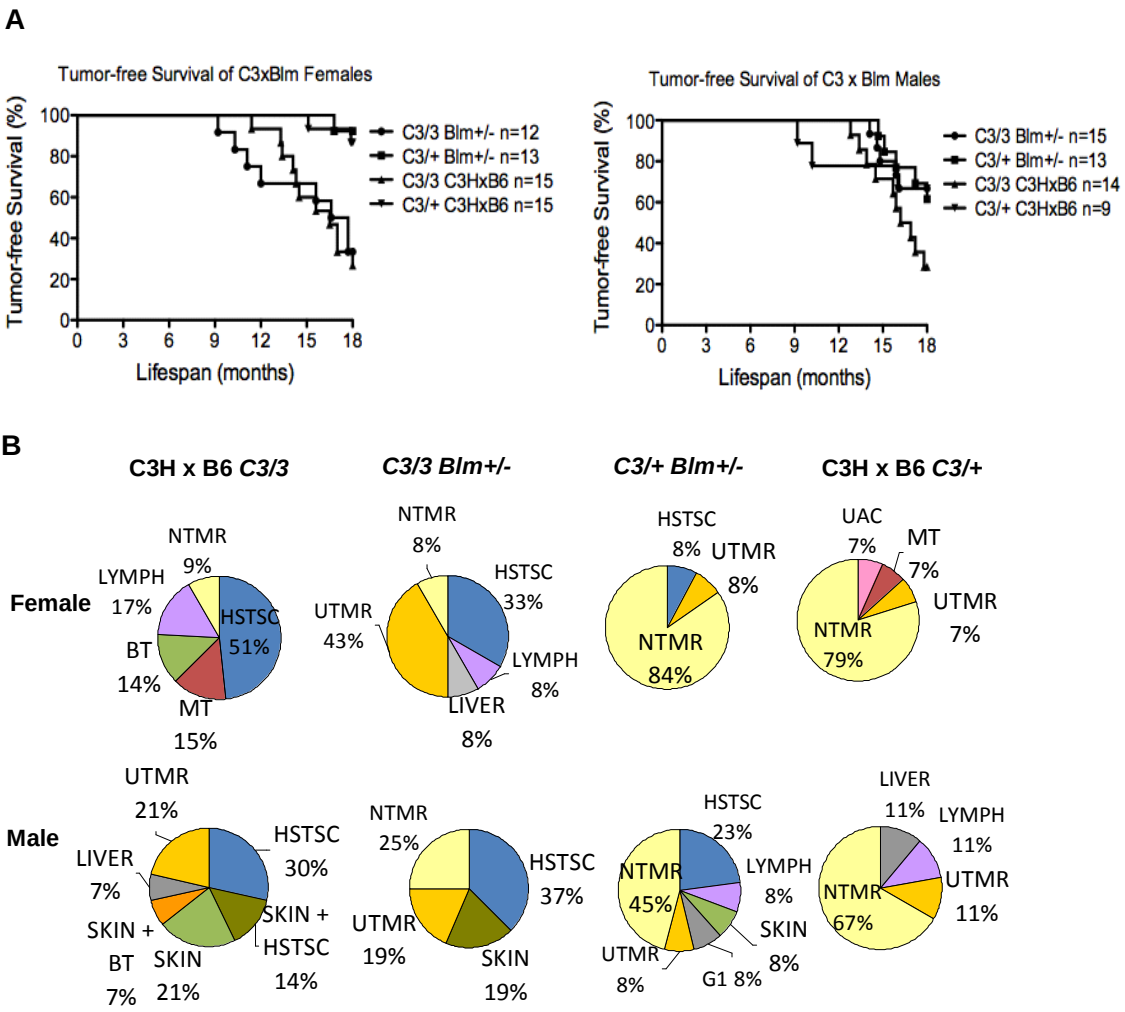


Figure 4-6: *Blm* deficiency has cryptic effects on tumor latency and cancer susceptibility in *Chaos3* mice. (**A**) *Chaos3* x *Blm* tumor latency. *C3/3 Blm+/-* females do not have significantly different time to tumor onset than *C3/3* alone, but males near statistical significance for delayed tumor onset (Female: LRMCT p=0.8469, GBWT p=0.9013; Male: LRMCT p=0.059, GBWT p=0.081). LRMCT=Log-rank/Mantel-Cox Test; GBWT=Gehan-Breslow-Wilcoxon Test. (**B**) Tumor spectrum of *Chaos3* x *Blm* mice. HSTSC=histiocytic sarcoma, MT=mammary tumor, BT=bone tumor, LYMPH=lymphoma, SKIN=skin tumor, LIVER=liver tumor, UTMR=unknown tumor type , NTMR=no tumor. Male *C3/3 Blm+/-* mice may have increased resistance to tumorigenesis compared to their female counterparts or to *C3/3* alone. However, *C3/+ Blm+/-* males may equally have increased cancer susceptibility compared to females or *C3/+* alone.

4.4 Discussion

Together, the *Chaos3* x DDR mutants consistently show that deficiency in the *Atm* DDR pathway results in decreased tumor latency and/or increased tumor susceptibility. This evidence supports the dependence upon an intact p53 pathway and replication arrest in response to damage occurring in the *Chaos3* mice.

Interestingly, tumor latency and susceptibility differ between genders, with females demonstrating an overall greater cancer susceptibility to *Atm* and *p21* deficiency than males. Cancer incidence was 24% and 48% higher in *C3/+ p21+/-* and *C3/+ Atm+/-* females respectively compared to males. *C3/+ p21-/-* females had a 34% increase in cancer incidence compared to *C3/+*, whereas *C3/+ p21-/-* males only showed an 11% increase. *Atm* and *p21* inherited polymorphisms in humans lead to decreased DDR response and efficiency, which is associated with significantly increased risk of developing lung cancer specifically in African American women [35]. Interestingly, *C3/+ Chk2+/-* males showed a marked 40% increase in cancer incidence over *C3/+* alone compared to females' 16%. These results not only mark the importance of intact DDR pathways in protection from carcinogenesis, but underscore that gender and genetic background significantly impact cancer susceptibility and latency when DDR pathways are compromised.

C3H x B6 F2 *C3/3* females are largely resistant to mammary tumors (15% incidence), compared to *C3/3* C3H congenic animals (>80% incidence). Due to the strong impact of background strain on the type of cancer that develops, shifts in tumor spectrum must be analyzed with caution. It is interesting to note however, a shift in mammary tumor susceptibility in the *Chaos3* (C3H) x *Chk2* (B6) mice. The

mammary tumor incidence in *C3/3 Chk2-/-* females increased from 15% to 37%. Littermate *C3/+ Chk2+/-* females do not demonstrate this shift in mammary tumorigenesis, suggesting that rather than a factor of genetic background effect, the increased mammary tumorigenesis observed in the *C3/3 Chk2-/-* females may be due to *Chk2* homozygosity itself. In women, inherited *Chk2* mutation is known to convey a 2-3 fold increased breast cancer risk [36].

Overall tumor latency and susceptibility were not altered in *Chaos3* mice deficient for *Blm* or *Hus1* as they were when the *Chaos3* mice lacked *Atm*, *Chk2*, or *p21*. This suggests that DNA damage occurring in the *Chaos3* mice may not rely as heavily on the ATR or homologous recombination pathways for repair, or that these two pathways have sufficient alternative sub-routes to bypass the need for *Blm* and *Hus1* for DNA damage repair. There were however, specific genotypes that appeared impacted but conflicted with results from other genotypes. *C3/3 Hus1Δ/+* males had a ~25% increase in cancer incidence compared to all other *Chaos3* x *Hus1* genotypes. If this cancer susceptibility is due to genotype, it is unclear why the *C3/3 Hus1 neo/+* or *C3/3 Hus1 Δ/neo* males would not also share susceptibility, or at least differing susceptibilities from each other and/or to the *Hus1+/+* males. Additionally, *C3/3 Blm+/-* males had a 25% decrease in cancer incidence compared to *C3/3* alone, while *C3/+ Blm+/-* mice had a 22% increased incidence over *C3/+* alone. It is possible that there are thresholds for damage that some allelic combinations surpass (such as *C3/3 Blm+/-* or semi-lethality in *C3/3 atm-/-* animals), taking cells beyond the threshold for viability rather than leading to tumorigenesis. This kind of damage threshold has been found to occur in BRCA1/2 deficient tumor cells that are treated with poly(ADP-

ribose) polymerase (PARP) inhibitors [37-39]. When PARP is inhibited, DNA single-strand breaks (SSBs) degenerate to DSBs, requiring homologous recombination for repair. PARP inhibition becomes synthetic lethal to cells that are impaired for DDR genes essential for homologous recombination, such as BRCA or CDK1 [40,41]. This has proven to be a successful therapeutic strategy in the clinic [42-44]. In the context of decreased tumor incidence in *C3/3 Blm+/-* male mice, it is possible that Blm deficiency sufficiently compromises the HR pathway to resemble the synthetic lethal phenotype observed in PARP inhibited/HR deficient cells. Our results suggest that gender and genetic background may impact the efficacy of these therapeutic treatments.

4.5 Methods and Notes

Animals & Samples

p21 mice (B6;129S2-Cdkn1atm1Tyj/J) were purchased from the Jackson Laboratory. *Hus1* mice were obtained from R. Weiss [30]. *Chk2*, *p21*, and *Blm* mutants were congenic in C57BL/6J. *Atm* and *Hus1* animals were congenic in FVB. *Chaos3* C3HeB/FeJ congenic animals were crossed to DDR mutants to generate double mutant animals that were of mixed genetic background. Progeny were genotyped as described by Leveitt et al.[30] and the Jackson laboratory (http://jaxmice.jax.org). Double mutants and littermates of the same gender were aged to a terminal endpoint of eighteen months or until animals showed clinical signs of disease. Prism (GraphPad 5) statistical software was used to analyze survival curves and generate Kaplan-Meier plots.

MEFs

Timed matings were conducted for *Chaos3* x *Atm* animals. Embryos were collected at embryonic day 12.5, 13.5, and 18.5. MEFs were generated, cultured, and cell proliferation assays were conducted as previously described [17].

Histology

Tumor samples were formalin-fixed and embedded in paraffin for sectioning and histological analysis. Slides were stained with hematoxylin and eosin (H&E) and examined for pathology.

Acknowledgements

We would like to thank R. Weiss for providing the *Hus1* mice.

This study was supported by NIH training grants IT32HDO57854 and 5T32GM007617 that supported M.D.W.; Empire State Stem Cell Fund contract numbers C026442 and C024174 to J.C.S.;

Author Contributions

M.D.W. and J.C.S designed experiments; experiments were conducted and data analyzed primarily by M.D.W.; X.L. generated *Chaos3* x *Blm* mice and conducted initial necropsies; T.L.S. conducted histopathology; J.C.S. oversaw the *Chaos3* project.

4.6 References

1. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).
2. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).
3. Wood, L.D. et al. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108-13 (2007).
4. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. *Br J Cancer* **83**, 1301-8 (2000).
5. Antoniou, A.C. et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* **98**, 1457-66 (2008).
6. Ford, D. et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **62**, 676-89 (1998).
7. Moynahan, M.E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* **11**, 196-207 (2010).
8. Lane, D. & Levine, A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol* **2**, a000893 (2010).
9. Cortez, D., Glick, G. & Elledge, S.J. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A* **101**, 10078-83 (2004).
10. Shechter, D. & Gautier, J. MCM proteins and checkpoint kinases get together at the fork. *Proc Natl Acad Sci U S A* **101**, 10845-6 (2004).
11. Sclafani, R.A. & Holzen, T.M. Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**, 237-80 (2007).
12. Chuang, C.H., Wallace, M.D., Abratte, C., Southard, T. & Schimenti, J.C. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* **6**(2010).
13. Kunnev, D. et al. DNA damage response and tumorigenesis in Mcm2-deficient mice. *Oncogene* **29**, 3630-8 (2010).
14. Pruitt, S.C., Bailey, K.J. & Freeland, A. Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. *Stem Cells* **25**, 3121-32 (2007).
15. Rusiniak, M.E., Kunnev, D., Freeland, A., Cady, G.K. & Pruitt, S.C. Mcm2 deficiency results in short deletions allowing high resolution identification of genes contributing to lymphoblastic lymphoma. *Oncogene* (2011).
16. Karakaidos, P. et al. Overexpression of the replication licensing regulators hCdt1 and hCdc6 characterizes a subset of non-small-cell lung carcinomas:

synergistic effect with mutant p53 on tumor growth and chromosomal instability--evidence of E2F-1 transcriptional control over hCdt1. *Am J Pathol* **165**, 1351-65 (2004).
17. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-8 (2007).
18. Shima, N., Buske, T.R. & Schimenti, J.C. Genetic screen for chromosome instability in mice: Mcm4 and breast cancer. *Cell Cycle* **6**, 1135-40 (2007).
19. Li, X.C. & Tye, B.K. Ploidy dictates repair pathway choice under DNA replication stress. *Genetics* **187**, 1031-40 (2011).
20. Bailis, J.M. & Forsburg, S.L. MCM proteins: DNA damage, mutagenesis and repair. *Curr Opin Genet Dev* **14**, 17-21 (2004).
21. Kawabata, T. et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41**, 543-53 (2011).
22. Kawabata, T. et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. *Mamm Genome* **22**, 506-17 (2011).
23. Bhatti, S. et al. ATM protein kinase: the linchpin of cellular defenses to stress. *Cell Mol Life Sci* **68**, 2977-3006.
24. Barlow, C. et al. Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell* **86**, 159-71 (1996).
25. Reiman, A. et al. Lymphoid tumours and breast cancer in ataxia telangiectasia; substantial protective effect of residual ATM kinase activity against childhood tumours. *Br J Cancer* **105**, 586-91.
26. Bartek, J. & Lukas, J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell* **3**, 421-9 (2003).
27. Hirao, A. et al. Chk2 is a tumor suppressor that regulates apoptosis in both an ataxia telangiectasia mutated (ATM)-dependent and an ATM-independent manner. *Mol Cell Biol* **22**, 6521-32 (2002).
28. Cazzalini, O., Scovassi, A.I., Savio, M., Stivala, L.A. & Prosperi, E. Multiple roles of the cell cycle inhibitor p21(CDKN1A) in the DNA damage response. *Mutat Res* **704**, 12-20.
29. Deng, C., Zhang, P., Harper, J.W., Elledge, S.J. & Leder, P. Mice lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. *Cell* **82**, 675-84 (1995).
30. Levitt, P.S. et al. Genome maintenance defects in cultured cells and mice following partial inactivation of the essential cell cycle checkpoint gene Hus1. *Mol Cell Biol* **27**, 2189-201 (2007).
31. Balmus, G. et al. Disease severity in a mouse model of ataxia telangiectasia is modulated by the DNA damage checkpoint gene Hus1. *Hum Mol Genet*.
32. Wang, Y. et al. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* **14**, 927-39 (2000).
33. Amor-Gueret, M. Bloom syndrome, genomic instability and cancer: the SOS-like hypothesis. *Cancer Lett* **236**, 1-12 (2006).

34. German, J. Bloom's syndrome. XX. The first 100 cancers. *Cancer Genet Cytogenet* **93**, 100-6 (1997).
35. Zheng, Y.L. et al. Elevated lung cancer risk is associated with deficiencies in cell cycle checkpoints: genotype and phenotype analyses from a case-control study. *Int J Cancer* **126**, 2199-210 (2010).
36. Meijers-Heijboer, H. et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* **31**, 55-9 (2002).
37. Bryant, H.E. et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913-7 (2005).
38. Farmer, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-21 (2005).
39. McCabe, N. et al. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res* **66**, 8109-15 (2006).
40. Ashworth, A. A synthetic lethal therapeutic approach: poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. *J Clin Oncol* **26**, 3785-90 (2008).
41. Johnson, N. et al. Compromised CDK1 activity sensitizes BRCA-proficient cancers to PARP inhibition. *Nat Med* **17**, 875-82 (2011).
42. Fong, P.C. et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med* **361**, 123-34 (2009).
43. Tutt, A. et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. *Lancet* **376**, 235-44 (2010).
44. Audeh, M.W. et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. *Lancet* **376**, 245-51 (2010).

CHAPTER 5

Hormonal involvement and the protective effect of oophorectomy against mammary adenocarcinomas in MCM4-deficient mice

Marsha D. Wallace[1,2], Teresa L. Southard[1,3], John C. Schimenti[1,2,4]

Affiliations:

[1]Department of Biomedical Sciences

[2]Department of Molecular Biology and Genetics

[3]Section of Anatomic Pathology

[4]Center for Vertebrate Genomics

Cornell University, Ithaca, NY 14853, USA

One Sentence Summary: *Mcm4*^{Chaos3}-C3H is an ER+ mammary tumor mouse model that mimics the human protective effect of oophorectomy against breast cancer.

5.1 Abstract

Almost all female *Mcm4*^{Chaos3}-C3H (*Chaos3*) mice succumb to spontaneous mammary tumors at a median of twelve months of age. Estrogen and progesterone control DNA replication in uterine epithelial cells by regulating *Mcm4* transcription, MCM proteins, and CDT1, the protein that facilitates loading the MCMs onto replication origins where the MCM complex acts as the DNA helicase during S phase. Estrogen and progesterone's role in regulating MCMs in other tissues is poorly understood. We examined the impact of reproductive hormones in mammary tumorigenesis in *Chaos3* mice. *Chaos3* females have higher levels of estrogen at twelve months of age compared to wild-type. The majority of *Chaos3* mammary tumors expressed ERα, though the mammary tumors of most mouse models do not. We find that oophorectomy, but not pregnancy, has a profound protective effect against mammary tumorigenesis in *Chaos3* mice. However, *Chaos3* oophorectomized mice were significantly more susceptible to developing other types of cancer, including histiocytic sarcomas, lymphomas, and osteosarcomas. *Chaos3* mammary tumor cells injected into oophorectomized WT mice did not reestablish as well as WT mice with intact ovaries. Together, our results signify the involvement of estrogen and progesterone in mammary carcinogenesis in these MCM-deficient mice.

5.2 Introduction

The *mcm4*^{Chaos3} (*Chaos3*) nonsynonomous point mutation in the C3H genetic background is the only endogenous gene mutation in mice that leads exclusively to mammary carcinogenesis [1]. MCM4 is a highly conserved subunit of the MCM2-7 DNA replicative helicase, an essential component of pre-replication (pre-RC) complexes [2]. These complexes are "licensed" for activation in S phase, and regulatory mechanisms inhibit reloading of the MCMs during S phase to prevent re-replication of the genome [2]. *Chaos3* mice have high levels of Genomic instability (GIN), a destabilized MCM helicase, a pan-reduction of all MCMs, a decreased number of dormant origins, and ultimately succumb to cancer at a mean latency of 12 months [3-7]. GIN, recurring copy number alterations (CNAs), background strain, additional mutations in mammary tumor modifier genes, and loss of DNA damage response (DDR) pathways, are all variables that contribute to *Chaos3* carcinogenesis (Wallace, Manuscript Chapters 2-4) [1,6,7]. However, additional mechanistic variables, such as reproductive hormones and their receptors, may tie MCMs to *Chaos3* carcinogenesis and mammary tumor specificity.

The reproductive hormones estrogen and progesterone control DNA replication in uterine epithelial cells by regulating MCM proteins [8]. Progesterone inhibits DNA synthesis by decreasing *Mcm* transcription (particularly *Mcm4*), MCM protein levels, and CDT1, the protein that facilitates loading MCMs onto replication origins. Progesterone also leads to the sequestration of MCMs into the cytoplasm even though these proteins are primarily nuclear [8]. Progesterone may regulate the

MCMs through miRNAs in uterine epithelial cells [9], and normal MCM levels are restored in *Chaos3* cells with the dicer or drosha pathways are knocked down [5]. In contrast to uterine epithelial cells, mammary tissue increases proliferation in response to progesterone [10]. In mammary tissue, progesterone causes ductal cell proliferation, leading to ductal enlargement or widening and side branching, while estrogen concentrates cell proliferation at the terminal end buds (TEBs) and ductal elongation and branching [10]. In normal human breast cells, progesterone increases transcription of the Mcms and other DNA licensing factors [11]. Mammary epithelial cells also demonstrate increased proliferation with an increase in Progesterone Receptor (PR). Breast cancer type 1 susceptibility (BRCA1) is involved in posttranscriptional downregulation of PR, and mice lacking BRCA1 exhibit increased PR, and subsequent increased mammary ductal growth and tumors [12].

In a cancer susceptibility study including 186 *Chaos3* C3HxB6 F2 females, 28% demonstrated cystic endometrial hyperplasia (CEH) or other abnormal uterine cell growth before 16 months of age (Wallace, QTL Manuscript—Chapter 3). This is particularly interesting in the context of reproductive hormonal regulation of the MCMs in uterine cells and suggests the *Chaos3* mutation may perturb these interactions, be it through the destabilized helicase, the pan-reduction of MCMs, or an additional mechanism triggered directly or indirectly by the mutation. Most cases of CEH result from high levels of estrogens, or insufficient levels of progesterone, which ordinarily counteracts estrogen's proliferative effects in uterine tissue [13]. Related conditions such as pyometra and mucometra, which were observed in the *Chaos3* F2s, also arise due to endometrial hyperplasia (Wallace, QTL Manuscript—Chapter 3) [13].

Endometrial hyperplasia is a significant risk factor for the development of endometrial cancer, and one *Chaos3* F2 female progressed to scirrhous endometrial carcinoma (Wallace, QTL Manuscript—Chapter 3) [13,14]. Scirrhous endometrial carcinoma is a type of uterine cancer associated with excessive estrogen exposure and often develops with CEH [15]. The high incidence of endometrial hyperplasia in *Chaos3* mice suggests the presence of high levels of estrogen, insufficient levels of progesterone, or that progesterone is incapable of downregulating *Chaos3* MCMs and thereby replication in uterine tissue, resulting in hyperplasia. Here we explore the impact of reproductive hormones on carcinogenesis in the *Chaos3* mouse model.

5.3 Results

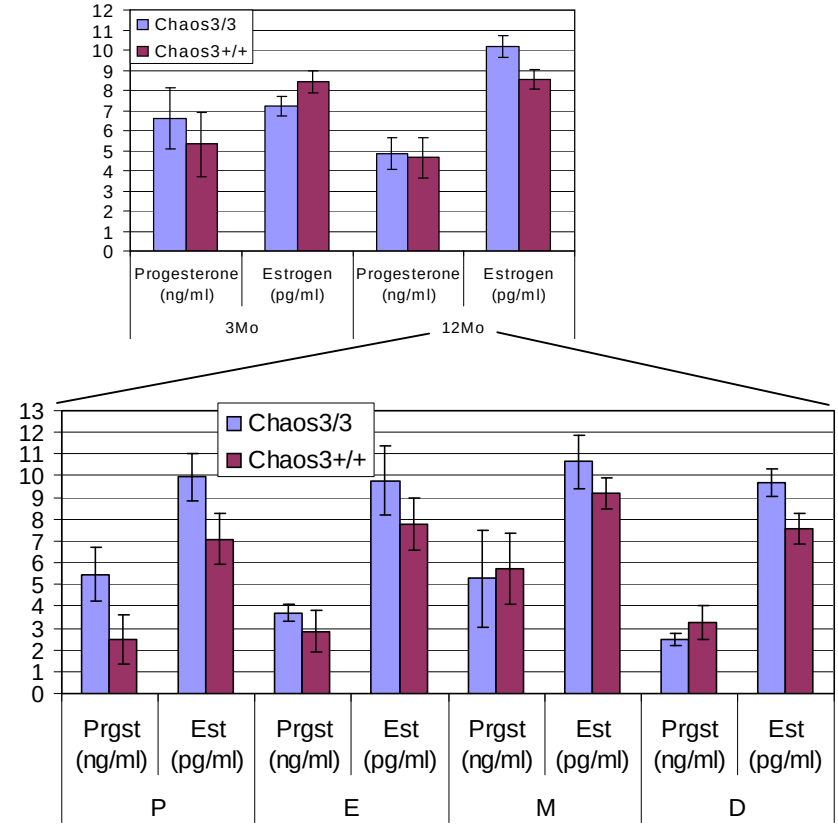
Chaos3 mice have higher levels of estrogen, and *Chaos3* mammary tumors express ER α

To examine the importance of reproductive hormones in the *Chaos3* model, blood samples were collected from *Chaos3*-C3H virgins at 3 and 12 months of age, and estrogen and progesterone levels were quantified by ELISA and RIA, respectively. *Chaos3* animals have an overall higher level of estrogen than WT at 12 months (Figure 5-1A). Mice have four day cyclic estrous cycles with bursts of ductal development. Ducts proliferate into the fat pad during proestrus and estrus, then regress or involute during metestrus and diestrus [16-19]. Due to variability in hormone levels within each stage of estrous [17,20], distinguishing differences based on stage is

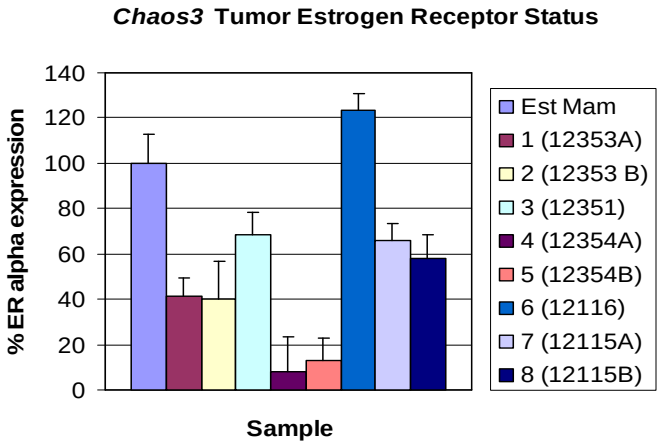
difficult. However, significant differences in increased hormone levels are observed in *Chaos3* females during proestrous and diestrous (Figure 5-1A). Approximately 75% of all human breast tumors are ER+, and growth of these tumors can be stimulated by estrogen [21]. qRT-PCR analysis of ER α mRNA levels in *Chaos3* tumors was conducted. Six out of eight (75%) *Chaos3* mammary tumors examined showed expression of ER α .

Figure 5-1: *Chaos3* animals have higher levels of estrogen at 12 months, and mammary tumors express ER α . (**A**) **Top**: *Chaos3* estrogen and progesterone levels in blood at 3 and 12 months of age. *C3/3*: 3 months (n=20), 12 months (n=29). *C3+/+*: 3months (n=20), 12 months (n=26). Error bars show Standard Error of the Mean (SEM). *Chaos3* animals have higher levels of estrogen than WT at 12 months. **Bottom**: Estrogen and progesterone levels in blood at 12 months by estrous stage. Prgst=Progesterone, Est=Estrogen. P=Proestrous, E=Estrous, M=Metestrous, D=Diestrous. (**B**) qRT-PCR analysis of ER α mRNA levels in *Chaos3* tumors. Levels were normalized to Actin and compared to expression of a normal estrous mammary gland. Error bars show Standard Error of the Mean (SEM). Note the similarity in ER α expression in tumors from the same animal (denoted with 'A' and 'B').

A


Chaos3/3
Chaos3+/+
Progesterone (ng/ml)
Estrogen (pg/ml)
Progesterone (ng/ml)
Estrogen (pg/ml)
3Mo
12Mo
Prgst (ng/ml)
Est (pg/ml)
P
E
M
D

B


Chaos3 Tumor Estrogen Receptor Status
% ER alpha expression
Sample
Est Mam
1 (12353A)
2 (12353 B)
3 (12351)
4 (12354A)
5 (12354B)
6 (12116)
7 (12115A)
8 (12115B)

Pregnancy does not strongly impact mammary tumorigenesis in *Chaos3* mice

Reproductive hormones and their receptors have a profound impact on mammary tumorigenesis. Nulliparous women have twice the risk of developing breast cancer as women who have undergone a full term pregnancy before 20 years of age [22]. In women, multiple early pregnancies confer a lifelong reduced risk of breast cancer [23,24]. This protection can be mimicked in rodents through administration of estrogen and progesterone treatments [25,26], and hormonal treatment causes increased long-term expression of *Trp53* and other pro-apoptotic genes [23,24].

To determine if *Chaos3* mammary tumor growth is impacted by reproductive hormones, we aged *Chaos3* breeder females. Breeders were classified into two groups: Group A and B. Group A breeders were required to have at least two litters, with their first litter being birthed before the mother reached three months of age. Group B breeders had their first litter after three months of age and were only required to have one litter. Thus, Group A represents a cohort of breeders with multiple, early pregnancies, while Group B females represent a cohort of later-life pregnancies. When *Chaos3* breeders were aged, tumor latency did not differ between Group A breeders (12.5 months) and virgin controls (13.1 months) (LRMCT p=0.8453; GBWT p=0.5173) **(Figure 5-2A)**. Tumor latency was decreased in Group B (11.5 months) in GBWT analysis (p=0.0344) but not LRMCT (p=0.1046) **(Figure 5-2A)**. Mammary tumor latency specifically was similarly decreased in Group B (GBWT p=0.0471; LRMCT p=0.0548). Incidence of mammary tumors did not significantly differ between *Chaos3* breeders (A=77%, B=81%) and virgins (92%) (FET group A

p=0.1385, group B p=0.4189) **(Figure 5-2B)**. These data show that pregnancy does not confer a strong protective effect against mammary tumorigenesis in *Chaos3* mice.

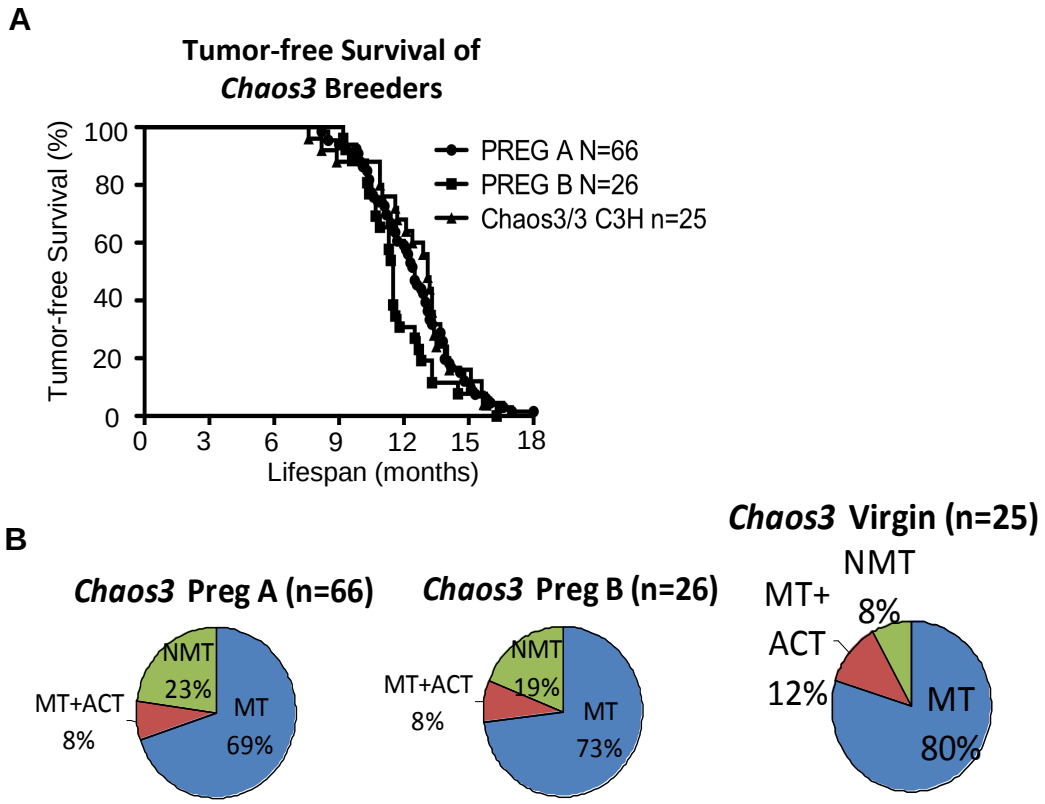


Figure 5-2: Multiple, early pregnancies do not reduce tumor latency or significantly reduce mammary tumor susceptibility. (**A**) Tumor-free survival of *Chaos3* breeders. PREG A=females with 2 or more litters, with their first litter being born before the females reached 3 months of age. PREG B=female breeders not fulfilling the previous criteria. Tumor latency did not differ between Group A breeders (12.5 months) and virgin controls (13.1 months) (LRMCT p=0.8453; GBWT p=0.5173). Tumor latency was decreased in Group B (11.5 months) in GBWT analysis (p=0.0344) but not LRMCT (p=0.1046). LRMCT=Log-Rank Mantel-Cox; GBWT=Gehan-Breslow-Wilcoxon Test. (**B**) Tumor Spectrum of *Chaos3* breeders. MT=mammary tumor; ACT=additional cancer type; NMT=non-mammary tumor. Incidence of mammary tumors did not significantly differ between *Chaos3* breeders and virgins (FET group A p=0.1385, group B p=0.4189).

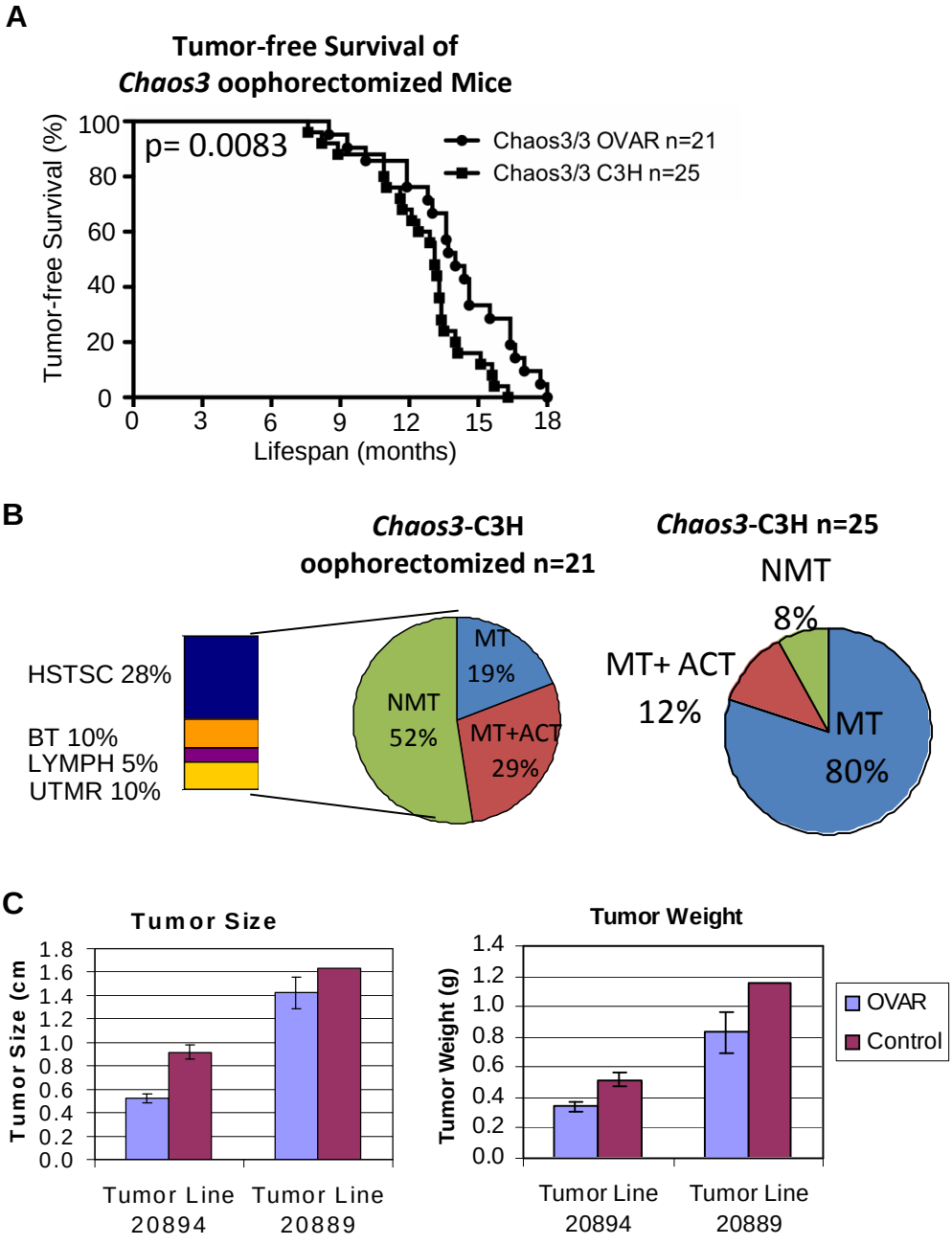
Oophorectomy strongly protects against mammary tumorigenesis in *Chaos3* mice

Premenopausal women who have undergone oophorectomy have a significant reduction in breast cancer risk [27,28]. The lifelong protective effect of oophorectomy extends to BRCA mutant carriers, with BRCA1 carriers having a 56% reduction in breast cancer risk and 15% increase in survival to age 70, and 46% breast cancer reduction and 7% increase in survival for BRCA2 carriers [29-31]. Mice with a conditional *Brca1* mutation also have reduced breast cancer incidence following oophorectomy [32].

To determine if *Chaos3* mammary tumor growth is impacted by or dependent upon reproductive hormones, we bilaterally oophorectomized 21 *Chaos3* female weanlings. *Chaos3* oophorectomized mice had significantly delayed tumor onset (14 months) compared to unoophorectomized mice (13.1 months) (Log-Rank Mantel-Cox p=0.0083; Gehan-Breslow-Wilcoxon Test p=0.0379) **(Figure 5-3A)**. Oophorectomy significantly impacted mammary tumorigenesis (FET p=0.0012), with mammary tumor incidence decreased by 44% **(Figure 5-3B)**. Partial ovarian tissue regrowth was observed histologically in 6 of the 21 oophorectomized mice, 5 of the 10 total animals in which mammary tumors were observed. This suggests that oophorectomy may have an even stronger protective effect against mammary tumorigenesis in the mice. However, cancer susceptibility remains constant, with oophorectomized animals developing non-mammary tumors such as bone tumors, lymphomas, and histiocytic sarcomas **(Figure 5-3B)**. To further explore the protective effect of oophorectomy, we examined the impact of oophorectomy on recapitulation of mammary tumors. *Chaos3* tumor cells were surgically implanted into the right abdominal mammary

gland of WT C3H mice. Oophorectomized mice formed smaller tumors, by size and weight, than unoophorectomized controls **(Figure 5-3C)**. These results show that oophorectomy confers a lifelong protective effect against mammary tumorigenesis in *Chaos3* mice.

Figure 5-3: Oophorectomy delays tumor onset in *Chaos3* mice and decreases susceptibility to mammary tumorigenesis, but mice are susceptible to other cancer types. (**A**) Tumor-free survival of *Chaos3* oophorectomized mice. oophorectomized mice had significantly delayed tumor onset (14 months) compared to unoophorectomized mice (13.1 months) (Log-Rank Mantel-Cox p=0.0083; Gehan-Breslow-Wilcoxon Test p=0.0379). (**B**) Tumor Spectrum of *Chaos3* ovariecomized mice. MT=mammary tumor; ACT=additional cancer type; NMT=non-mammary tumor; HSTSC=histiocytic sarcoma, BT=bone tumor; LYMPH=lymphoma; UTMR=undetermined tumor of non-mammary origin. Note that incidence of mammary tumors in *Chaos3* oophorectomized mice is significantly impacted (FET p=0.0012), decreasing by 44%. However, cancer susceptibility remains constant, with oophorectomized animals developing non-mammary tumors such as bone tumors, lymphomas, and histiocytic sarcomas. (**C**) Impact of oophorectomy on recapitulation of mammary tumors by size and weight. Error bars show Standard Error of the Mean (SEM). *Chaos3* mammary tumor cells injected into C3H WT mice formed smaller tumors in oophorectomized (n=5) vs. unoophorectomized controls (n=3).


A
Tumor-free Survival of
Chaos3 oophorectomized Mice
Tumor-free Survival (%)
100
80
60
40
20
0
p= 0.0083
Chaos3/3 OVAR n=21
Chaos3/3 C3H n=25
0
3
6
9
12
15
18
Lifespan (months)
B
Chaos3-C3H
oophorectomized n=21
HSTSC 28%
BT 10%
LYMPH 5%
UTMR 10%
NMT
52%
MT
19%
MT+ACT
29%
Chaos3-C3H n=25
NMT
8%
MT+ ACT
12%
MT
80%
C
Tumor Size
Tumor Size (cm
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Tumor Line
20894
Tumor Line
20889
Tumor Weight
Tumor Weight (g)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Tumor Line
20894
Tumor Line
20889
OVAR
Control

5.4 Discussion

The dual nature of the protective effect of reproductive hormones in pregnancy or their absence in oophorectomy highlights the complexities surrounding breast cancer mechanisms. Understanding the mechanisms of these effects is an area of active research. During pregnancy, the mammary gland is exposed to high levels of ovarian (estrogens and progestins), pituitary (prolactin and growth hormone), and placental (placental lactogens) hormones [33]. One hypothesis is that pregnancy removes a highly cancer-susceptible mammary epithelial cell population by inducing differentiation, removal, or modification of the cells [33,34]. Alternatively, after pregnancy, mammary cells have enhanced DNA repair capabilities [33]. Post-partum there is also a reduction in hormone levels and receptors in the mammary gland [35]. Preventative treatment therapy for young women against breast cancer by administering hormonal injection during the teenage years is controversial, but research and drug treatments are moving forward in clinical trials [36].

Interestingly, pregnancy Group B *Chaos3* females who bore their first litter later in life showed decreased tumor latency by GBWT analysis compared to virgins **(Figure 5-2A)**. Women over the age of 35 having their first full-term pregnancy have an increased risk of developing breast cancer [37,38], and women over age 30 have an even higher risk of breast cancer soon after delivery [39-42]. This increased susceptibility to breast cancer extends up to 20 years after first delivery in women greater than 30 years of age [43,44]. These risk effects are limited to hormone-responseive breast cancers (ER+/PR+) [45]. This is consistent with *Chaos3* Group B mammary tumorigenesis occurring earlier than in virgins (median tumor-free survival 11.5 vs 13.1 months).

Notably, while *Chaos3* oophorectomized mice had significantly decreased mammary tumor incidence, the animals developed other forms of cancer instead **(Figure 5-3B)**. One human study with 24 years of follow-up finds that females with bilateral oophorectomy have decreased risk of breast and ovarian cancer, but these women have an increased risk of all-cause mortality, including fatal coronary heart disease and lung cancer [46]. This raises the prospect that humans prone to reproductive cancers who undergo prophylactic-oophorectomy may subsequently increase their risk for developing other disease phenotypes.

Mammary tumors in most mouse models are ER-, indicating that the tumor growth is not stimulated by estrogen as in the majority of human breast cancer cases [47,48]. We have shown that *Chaos3* mammary tumors express ER α and that mammary tumor incidence is greatly attenuated in *Chaos3* oophorectomized mice. Additionally, the *Chaos3* mammary tumor expression signature more closely resembles the differentiation signature of mature human luminal cells than all other mouse models analyzed, and the tumors develop recurring copy number alterations (CNAs) that overlap with human breast tumors (Wallace, Nf1 Manuscript—Chapter 2). Together, these characteristics make *Chaos3* a highly relevant model to study human breast cancer.

5.5 Methods and Notes

Animals

For oophorectomy studies, *Chaos3* C3HeB/FeJ congenic and inbred C3HeB/FeJ animals were aged to a terminal endpoint of eighteen months or until animals showed

clinical signs of disease. Prism (GraphPad 5) statistical software was used to analyze survival curves and generate Kaplan-Meier plots.

Estrogen and Progesterone Levels

Blood was collected from *Chaos3* C3HeB/FeJ congenic and inbred C3HeB/FeJ mice at 3 and 12 months of age. Blood was allowed to clot at room temperature for 90 minutes and then centrifuged at 2000 g for 15 minutes. Serum was removed and stored at -20. Mouse estradiol was quantified by ELISA (Calbiotech), and progesterone was quantified by RIA, both at the UVA ligand core. Estrous stage was determined from histological examination of H&E sections from the reproductive tracts.

Surgeries

Mice were anesthetized with an intraperitoneal (IP) injection of 0.15 ml Avertin/10 g. Bilateral oophorectomy survival surgeries were performed following standard SOPs when females were three weeks of age, before the onset of sexual maturity. An incision through the peritoneal wall above the fat pad was made, the bursa opened, and then the ovaries were removed. The incision was sutured and a wound clip was used to close the skin. For WT mice in the tumor recapitulation study, *Chaos3* tumor cells were injected using Hamilton syringe directly into the right abdominal mammary fat pad. Mammary glands were not cleared. Wound clips were used to close the skin. Mice were placed on a slide warmer tray while recovering from anesthesia and then placed in a clean cage. After surgery, 0.1 ml Ketoprofen was injected subcutaneously (SubQ). 24 hours after surgery, 0.1 ml Ketoprofen injected SubQ. The mice and surgery site were observed daily for nine days until the wound clips were removed.

Histology

Tumor samples were formalin-fixed and embedded in paraffin (FFPE) for sectioning and histological analysis. Slides were stained with hematoxylin and eosin (H&E) and examined for pathology.

Acknowledgements

This study was supported by NIH training grants IT32HDO57854 and 5T32GM007617 that supported M.D.W.; Empire State Stem Cell Fund contract numbers C026442 and C024174 to J.C.S.

We would like to thank the University of Virginia for Research in Reproduction Ligand Assay and Analysis Core for hormone level quantification (NICHD SCCPRR Grant U54-HD28934); R. Munroe for mouse oophorectomy surgery services.

Author Contributions

M.D.W. and J.C.S designed experiments; experiments were conducted and data analyzed by M.D.W.; T.L.S. conducted histopathology; J.C.S. oversaw the *Chaos3* project.

5.6 References

1. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-8 (2007).
2. Sclafani, R.A. & Holzen, T.M. Cell cycle regulation of DNA replication. *Annu Rev Genet* **41**, 237-80 (2007).
3. Shima, N., Buske, T.R. & Schimenti, J.C. Genetic screen for chromosome instability in mice: Mcm4 and breast cancer. *Cell Cycle* **6**, 1135-40 (2007).
4. Chuang, C.H., Wallace, M.D., Abratte, C., Southard, T. & Schimenti, J.C. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* **6**(2010).
5. Chuang, C.H. et al. Post-transcriptional homeostasis and regulation of MCM2-7 in mammalian cells. *Nucleic Acids Res* (2012).
6. Kawabata, T. et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41**, 543-53 (2011).
7. Kawabata, T. et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. *Mamm Genome* **22**, 506-17 (2011).
8. Pan, H., Deng, Y. & Pollard, J.W. Progesterone blocks estrogen-induced DNA synthesis through the inhibition of replication licensing. *Proc Natl Acad Sci U S A* **103**, 14021-6 (2006).
9. Kuokkanen, S. et al. Genomic profiling of microRNAs and messenger RNAs reveals hormonal regulation in microRNA expression in human endometrium. *Biol Reprod* **82**, 791-801 (2010).
10. Brisken, C. & O'Malley, B. Hormone action in the mammary gland. *Cold Spring Harb Perspect Biol* **2**, a003178 (2010).
11. Graham, J.D. et al. DNA replication licensing and progenitor numbers are increased by progesterone in normal human breast. *Endocrinology* **150**, 3318-26 (2009).
12. Poole, A.J. et al. Prevention of Brca1-mediated mammary tumorigenesis in mice by a progesterone antagonist. *Science* **314**, 1467-70 (2006).
13. Kim, J.J. & Chapman-Davis, E. Role of progesterone in endometrial cancer. *Semin Reprod Med* **28**, 81-90 (2010).
14. Kurman, R.J., Kaminski, P.F. & Norris, H.J. The behavior of endometrial hyperplasia. A long-term study of "untreated" hyperplasia in 170 patients. *Cancer* **56**, 403-12 (1985).
15. Bokhman, J.V. Two pathogenetic types of endometrial carcinoma. *Gynecol Oncol* **15**, 10-7 (1983).
16. Daniel, C.W. & Smith, G.H. The mammary gland: a model for development. *J Mammary Gland Biol Neoplasia* **4**, 3-8 (1999).

17. Fata, J.E., Chaudhary, V. & Khokha, R. Cellular turnover in the mammary gland is correlated with systemic levels of progesterone and not 17beta-estradiol during the estrous cycle. *Biol Reprod* **65**, 680-8 (2001).
18. Nelson, J.F., Felicio, L.S., Randall, P.K., Sims, C. & Finch, C.E. A longitudinal study of estrous cyclicity in aging C57BL/6J mice: I. Cycle frequency, length and vaginal cytology. *Biol Reprod* **27**, 327-39 (1982).
19. Richert, M.M., Schwertfeger, K.L., Ryder, J.W. & Anderson, S.M. An atlas of mouse mammary gland development. *J Mammary Gland Biol Neoplasia* **5**, 227-41 (2000).
20. Hurn, P.D. & Macrae, I.M. Estrogen as a Neuroprotectant in Stroke. *Journal of Cerebral Blood Flow & Metabolism* **20**, 631-652 (2000).
21. Li, C.I., Daling, J.R. & Malone, K.E. Incidence of invasive breast cancer by hormone receptor status from 1992 to 1998. *J Clin Oncol* **21**, 28-34 (2003).
22. MacMahon, B. et al. Age at first birth and breast cancer risk. *Bull World Health Organ* **43**, 209-21 (1970).
23. Russo, I.H. & Russo, J. Hormonal approach to breast cancer prevention. *J Cell Biochem Suppl* **34**, 1-6 (2000).
24. Medina, D. Breast cancer: the protective effect of pregnancy. *Clin Cancer Res* **10**, 380S-4S (2004).
25. Sinha, D.K., Pazik, J.E. & Dao, T.L. Prevention of mammary carcinogenesis in rats by pregnancy: effect of full-term and interrupted pregnancy. *Br J Cancer* **57**, 390-4 (1988).
26. Guzman, R.C. et al. Hormonal prevention of breast cancer: mimicking the protective effect of pregnancy. *Proc Natl Acad Sci U S A* **96**, 2520-5 (1999).
27. Trichopoulos, D., MacMahon, B. & Cole, P. Menopause and breast cancer risk. *J Natl Cancer Inst* **48**, 605-13 (1972).
28. Kelsey, J.L., Gammon, M.D. & John, E.M. Reproductive factors and breast cancer. *Epidemiol Rev* **15**, 36-47 (1993).
29. Eisen, A. et al. Breast cancer risk following bilateral oophorectomy in BRCA1 and BRCA2 mutation carriers: an international case-control study. *J Clin Oncol* **23**, 7491-6 (2005).
30. Kurian, A.W., Sigal, B.M. & Plevritis, S.K. Survival analysis of cancer risk reduction strategies for BRCA1/2 mutation carriers. *J Clin Oncol* **28**, 222-31 (2010).
31. Kauff, N.D. et al. Multi-center prospective analysis of risk-reducing salpingo-oophorectomy to prevent BRCA-associated breast and ovarian cancer. *Journal of Clinical Oncology* **24**(2006).
32. Bachelier, R. et al. Effect of bilateral oophorectomy on mammary tumor formation in BRCA1 mutant mice. *Oncol Rep* **14**, 1117-20 (2005).
33. Russo, I.H. & Russo, J. Mammary gland neoplasia in long-term rodent studies. *Environ Health Perspect* **104**, 938-67 (1996).
34. Grubbs, C.J., Peckham, J.C. & McDonough, K.D. Effect of ovarian hormones on the induction of 1-methyl-1-nitrosourea-induced mammary cancer. *Carcinogenesis* **4**, 495-7 (1983).

35. Thordarson, G. et al. Refractoriness to mammary tumorigenesis in parous rats: is it caused by persistent changes in the hormonal environment or permanent biochemical alterations in the mammary epithelia? *Carcinogenesis* **16**, 2847-53 (1995).
36. Cuzick, J. et al. Preventive therapy for breast cancer: a consensus statement. *Lancet Oncol* **12**, 496-503 (2011).
37. Britt, K., Ashworth, A. & Smalley, M. Pregnancy and the risk of breast cancer. *Endocr Relat Cancer* **14**, 907-33 (2007).
38. Trichopoulos, D. et al. Age at any birth and breast cancer risk. *Int J Cancer* **31**, 701-4 (1983).
39. Janerich, D.T. & Hoff, M.B. Evidence for a crossover in breast cancer risk factors. *Am J Epidemiol* **116**, 737-42 (1982).
40. Lambe, M. et al. Transient increase in the risk of breast cancer after giving birth. *N Engl J Med* **331**, 5-9 (1994).
41. Lambe, M. et al. Parity, age at first birth and the risk of carcinoma in situ of the breast. *Int J Cancer* **77**, 330-2 (1998).
42. Schedin, P. Pregnancy-associated breast cancer and metastasis. *Nat Rev Cancer* **6**, 281-91 (2006).
43. Albrektsen, G., Heuch, I., Hansen, S. & Kvale, G. Breast cancer risk by age at birth, time since birth and time intervals between births: exploring interaction effects. *Br J Cancer* **92**, 167-75 (2005).
44. Rosner, B., Colditz, G.A. & Willett, W.C. Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study. *Am J Epidemiol* **139**, 819-35 (1994).
45. Ma, H., Bernstein, L., Pike, M.C. & Ursin, G. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res* **8**, R43 (2006).
46. Parker, W.H. et al. Ovarian conservation at the time of hysterectomy and long-term health outcomes in the nurses' health study. *Obstet Gynecol* **113**, 1027-37 (2009).
47. Mohibi, S., Mirza, S., Band, H. & Band, V. Mouse models of estrogen receptor-positive breast cancer. *J Carcinog* **10**, 35 (2011).
48. Wagner, K.U. Models of breast cancer: quo vadis, animal modeling? *Breast Cancer Res* **6**, 31-8 (2004).

CHAPTER 6

Summary & Discussion

6.1 Mammary Tumor Carcinogenesis

Breast cancer research faces many challenges and limitations in the human system. In human tumors, the prevalence of passenger mutations, heterogeneity, and the diversity of tumor etiologies and subtypes complicates conclusions about putative drivers [1-3]. The majority of genetically-based breast cancers are caused by low-penetrance modifier alleles, complicating attempts at genetic mapping by GWAS in humans [4]. These challenges necessitate a comparative oncogenomic approach, for which mouse models are a powerful tool. The environment and genetic backgrounds can be precisely defined, and within a given cancer mouse model, the tumors that arise have a consistent underlying basis. Additionally, mice can be experimentally manipulated to solidify evidence of candidate genes and validate genomic findings.

Here I have utilized the *Chaos3* mouse model, which is not genetically engineered or treated with carcinogens, in order to study spontaneous mammary tumorigenesis. Using this model I have been able to identify mammary tumor drivers and susceptibility loci as well as examine the role of DDR and reproductive hormones on carcinogenesis when the core DNA replication machinery is defective.

6.2 Genomic Analysis of Mammary Tumors and Identification of Driving Mechanisms

Genomic analysis of *Chaos3* tumors culminated in several important findings. *Chaos3* tumors model key human features. Luminal breast tumors are the most prevalent type in humans [5]. Expression profiling of *Chaos3* mammary tumors revealed that they cluster near luminal mouse models, and the *Chaos3* gene signature was the most highly expressed in the human luminal subtypes. *Chaos3* tumors have a distinct gene expression pattern from all other mouse models, including dramatic upregulation of *Mucl1*, a diagnostic marker in human breast cancer [6]. Also, *Chaos3* tumors more closely resemble mature human luminal cells than any mouse model analyzed to date. Partial exome resequencing of *Chaos3* mammary tumors revealed few somatic point mutations, indicating that elevated intragenic mutagenesis is not the primary mechanism driving *Chaos3* carcinogenesis. aCGH showed that *Chaos3* mammary tumors have recurring CNAs that overlap with those found in human breast cancer, including loss of the *Nf1* tumor suppressor. aCGH using higher density chips (1.4 million probes *versus* 385 or 720 thousand) has now been conducted on an additional 15 *Chaos3* tumors, and the data are currently being analyzed. The higher density chips will allow for higher resolution detection of small amplifications and deletions to identify additional candidate drivers. The recurrence of *Chaos3* CNAs, despite high levels of aneuploidy in tumor cells, and similarity to human CNAs indicated that recurring amplifications and deletions were a driving mechanism of mammary tumorigenesis in *Chaos3* mice.

6.3 *Nf1*, a prevalent breast cancer driver

Nf1 has not previously been implicated in spontaneous breast cancer. I found that *Nf1* is lost in almost all *Chaos3* mammary tumors, and this loss leads to hyperactivation of the RAS oncogene, which controls cell proliferation and anti-apoptosis pathways. These tumors are sensitive to drugs targeting the RAS pathway. Analysis of unpublished TCGA data revealed that *NF1* deficiency is present in 27.7% of all human breast tumors, including >40% of Her2-enriched and basal-like subtypes. These data implicate *NF1* deficiency as a major breast cancer driver, and should inform treatment of the ~63,450 Americans who develop breast cancer with *NF1* deficiency annually.

An *Nf1* knockdown experiment is currently underway in C3H WT inbred females that have been surgically injected with *Chaos3* mammary epithelial cells containing a construct of *Nf1*-shRNA. These cells were injected into the cleared right abdominal mammary gland fat pad, while untreated *Chaos3* cells were injected into the cleared left abdominal mammary gland fat pad as a control. If animals develop tumors more frequently in the right abdominal mammary gland, it will provide further evidence that *Nf1* deficiency is driving tumorigenesis. Long-term, the role of *Nf1* deficiency as a driver of mammary tumorigenesis will be addressed by generating and aging *Chaos3/3* x *Nf1*+/- and *Chaos3*/+ x *Nf1*+/- C3H females. 75% of animals heterozygous for *Nf1*, with neo cassette replacement of a portion of exon 30 and 31, develop tumors over 27 months, compared to 15% of wild-type animals [7]. Furthermore, the tumor spectrum observed in the *Nf1*+/- animals was similar to those

in wild-type mice, suggesting that the mutation may accelerate development of tumor types to which the animals are already susceptible [7].

In the near future, an *Nf1* add-back/rescue experiment will be conducted similar to the *Nf1* knockdown experiment. For the rescue experiment, an *Nf1* construct will be transfected into *Chaos3* tumor cells. Tumor cells with and without the *Nf1* construct will be surgically injected into female C3H mammary fat pads to determine the impact of *Nf1* on mammary tumor recapitulation and growth. The *Nf1* construct has almost all intronic regions removed, and given the gene's large size and multiple isoforms, proper expression of the gene is a potential technical problem for the experiment. However, if *Nf1*-bearing cells show delayed tumor reformation and growth, it will contribute evidence of the role of *Nf1* in mammary tumor suppression.

Chaos3 tumor cells are sensitive to rapamycin (MTOR inhibitor) and PD98059 (MEK1 inhibitor). There are additional drugs that also target the RAS pathway, including Tipifarnib, Sorafenib, LY294002, Salirasib, Temsirolimus, and Everolimus **(Fig 1)**. Among these, it would be beneficial to determine (in *Chaos3* and human *Nf1* deficient tumor lines) which combination has the greatest impact on cell proliferation. Additional markers predicting sensitivity could be extrapolated from tumor profiles. Based on drug treatment outcome in cell lines, *Chaos3* mice could then be treated to determine efficacy *in vivo*.

Tipifarnib was being tested in phase II clinical trials for use in stage II and stage III breast cancer as well as hormone-receptor positive breast cancer [8,9]. Despite promising results, clinical trials were suspended after the FDA disapproved the drug for treatment of acute myeloid leukemia (AML), citing the lack of a randomized study

and that many of the patients were suitable candidates for standard induction therapy [10]. A phase II trial of salirasib for treatment of lung adenocarcinomas patients with KRAS mutations found that the given dosage and schedule had insufficient activity to warrant further evaluation [11]. Salirasib would have been the most ideal inhibitor as it specifically targets RAS and could block progression of both the PI3K and MAPK pathways. Sorafenib is currently approved to treat hepatocellular carcinoma and advanced renal cancer. Temsirolimus and Everolimus are two other mTOR inhibitors that are approved for the treatment of renal cell carcinoma.

In the absence of an approved RAS inhibitor in humans, a powerful strategy for treatment of Nf1-defcicient mammary tumors may be a combination of drugs to target both the PI3K and MAPK pathways. Additionally, clinical trials may advance more quickly by using drugs that have already been approved for treatment of other cancer types. Thus, the most promising combinations to test are Sorafenib/Temsirolimus and Sorafenib/Everolimus.

As evidence accumulates of *Nf1* as a major breast cancer driver, impacting ~1/4 of all human cases, clinical testing of *Nf1* will be critical for personalized treatment. Tamoxifen, the estrogen receptor (ER) inhibitor that is standard treatment for ER+ breast cancers, may not be appropriate for women whose cancers involve *NF1* deficiency. NF1 depletion confers resistance of human breast cancer (MCF7) cells to tamoxifen, and tamoxifen-treated patients whose tumors had lower *NF1* expression levels had poorer clinical outcomes [12].

Frequent *NF1* deletion may be due to a combination of factors including fragile sites (Fig. S6), a complex chromatin structure, and/or its large genomic size.

Indeed, replication fork stalling near *Nf1* has been noted at a 5 kb isochore transition zone conserved between human and mouse, separating early and late replicating chromatin [13]. Furthermore, collisions between replication and transcription complexes cause instability at fragile sites in the longest human genes [14]. Due to these factors, in context of the *Chaos3* mutation, the unstable DNA replication helicase may have difficulty progressing through this region, leading to frequent deletion, which subsequently contributes to driving carcinogenesis.

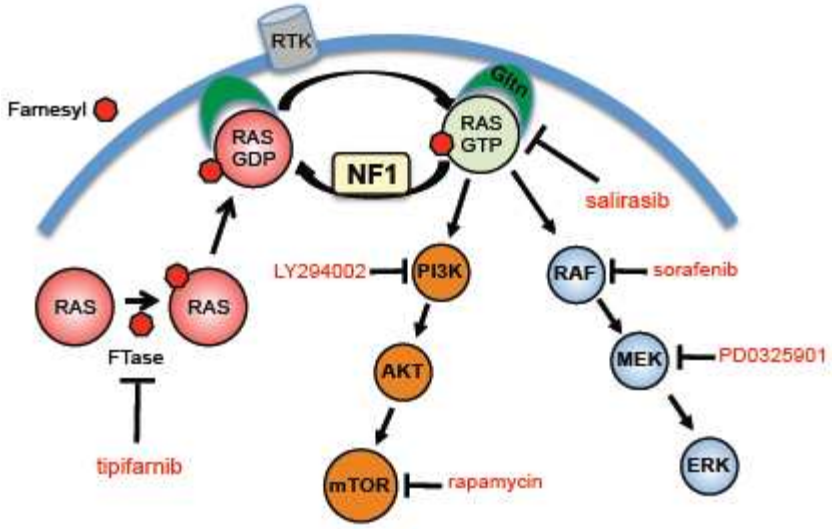


Figure 6-1: RAS signaling pathway and inhibitors. Drug inhibitors are shown in red type. Note, Temsirolimus and Everolimus are mTOR inhibitors, like rapamycin, but not shown. Ftase = farnesyltransferase. RTK = receptor tyrosine kinase. The curved blue line near the top depicts the cell membrane. Gltn = Galectin, a protein that binds farnesylated Ras to the cell membrane. Not all downstream targets are shown.

6.4 Mammary Tumor Susceptibility and Resistance Loci

While an estimated ~25% of breast cancer cases have an inherited familial basis, the majority of susceptibility genes underlying these heritable cases remain unknown [15]. QTL analysis of *Chaos3* C3H x B6 F2s allowed for identification of mammary tumor susceptibility and resistance loci. Candidate genes are involved in

cell proliferation (*Fgfr3*), DNA repair (*Msh4*, *Fancg*, *Fancc*, *Rad51ap1*), cell signaling (*Fbxw7*, *Frp1*, *Ptc1*, *Tln1*, *Pax5*), and cancer associated genes (*Rab2a*, *Rab28*, *Styk1*, *Mycbp2*). To bolster the number of mammary tumors analyzed, a new cohort of mixed C3H, B6, and FVB backgrounds from the DDR experiments was genotyped on a 3rd SNP chip of higher density. Analysis is currently underway to identify additional cancer susceptibility and resistance loci.

Tln1, a gene required for integrin activation, was identified in the *Chaos3* QTL experiment as a candidate gene for mammary tumor susceptibility. *Tln1* was previously identified as a candidate driver in human breast cancer with a low passenger mutation probability [16]. Within the *Chaos3*-C3H colony, *Tln1* was discovered to have a nonsynonomous germline point mutation. When aged, *Chaos3 Tln1* mutants had a significantly higher mammary tumor incidence than *Chaos3* alone, validating *Tln1* impact on mammary tumor susceptibility. The *Tln1* mutation falls within a cryptic vinculin binding domain [17], which may interfere with how the proteins interact with integrin for activation. As a first step to understand the mechanism by which *Tln1* is contributing to mammary susceptibility, components of the Integrin pathway could be examined for altered levels of activation, predicting hyperactivation. However, understanding this mechanism may prove to be a long-term endeavor, as the mechanism of BRCA1 and BRCA2 mutations on specifically breast and ovarian cancer susceptibility remains unclear despite intense research. Other genes identified in the *Chaos3* aCGH and F2-QTL experiments are also excellent candidates for future studies.

6.5 Interaction between TLN1/Integrin and NF1/RAS pathways

There is growing evidence for the use of *TLN1* expression as a prognostic marker in tumor progression and as a therapeutic target [18]. *TLN1* overexpression leads to pro-survival and progression to metastasis [19]. There is significant cross-talk between the talin/integrin and RAS pathways (Figure 6-2). TLN1-induced integrin activation triggers the formation and activation of the focal-adhesion complex, including FAK, which then activates the PI3K/AKT survival pathway (Figure 6-2) [18]. This results in anoikis resistance, angiogenesis, and survival after cellular detachment from the extra cellular matrix (ECM) [18]. The focal-adhesion complex also activates the MAPK pathway through phosphorylation of ERK1/2. NF1 also binds FAK, but the functional consequence of this is unknown [20]. Ras GTPases mediate integrin–talin interactions [21]. Furthermore, binding of GRB2/SOS to FAK plays a significant role in activating the RAS pathway [22]. Additional RAS family proteins impact integrin activation as well, including R-RAS, RAP1, RIAM, and Rho-GTPases [23-25]. Interaction between caveolin-1 and Rho-GTPases promotes metastasis by controlling the expression of alpha5-integrin and the activation of Src, Ras and Erk [23]. Together, this suggests that concurrent loss of NF1 and TLN1 mutation has a synergistic effect to promote carcinogenesis via the PI3K and MAPK pathways.

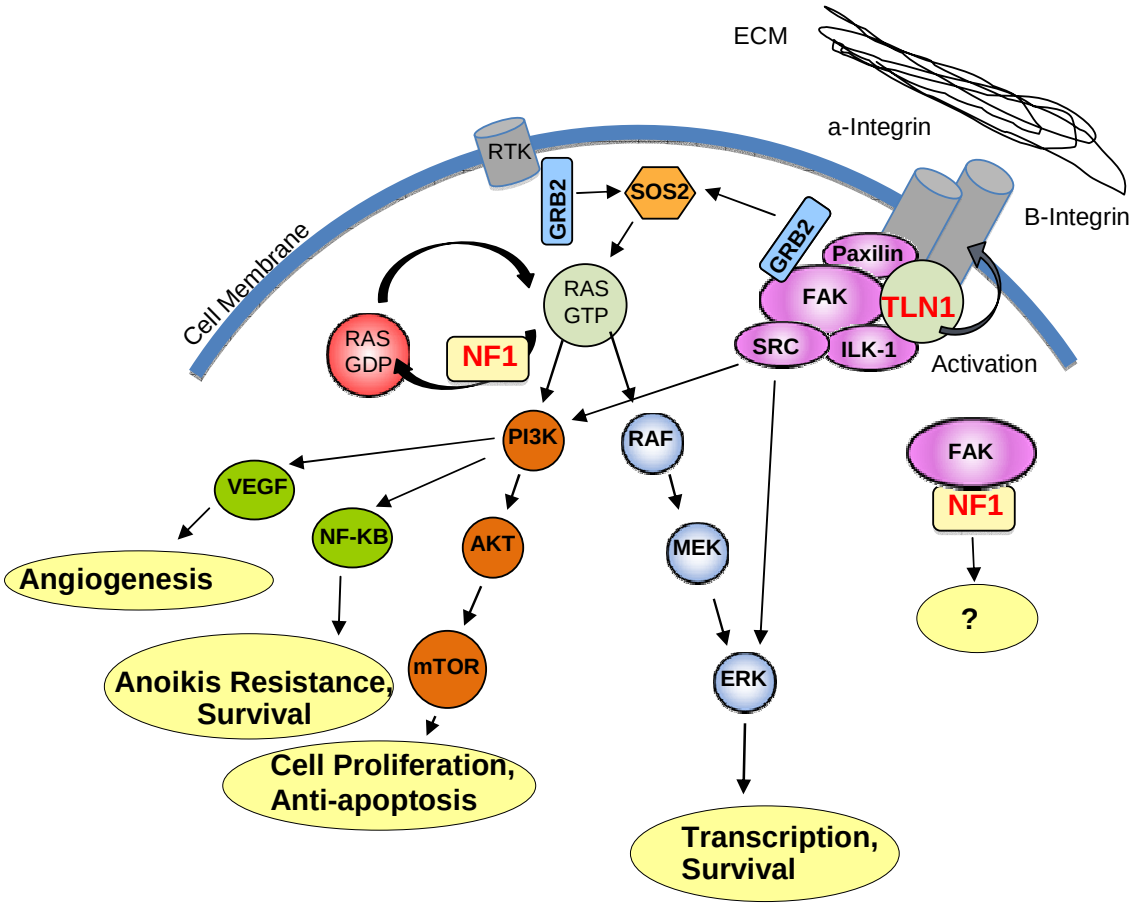


Figure 6-2: Interaction between TLN1/Integrin and NF1/RAS pathways. RTK = receptor tyrosine kinase. ECM=extra cellular matrix. NF1 inhibits activated RAS-GTP by increasing its conversion rate into the inactive RAS-GDP state. Loss of NF1 leads to increased activated RAS and subsequent downstream activation of PI3K and MAPK signal transduction pathways. TLN1 activates integrin, triggering activation of the focal-adhesion complex (including FAK) and subsequent activation of the PI3K and MAPK pathways. These pathways control cell proliferation and transcription of anti-apoptosis/pro-survival genes. NF1 also binds FAK, but the functional consequence of this is not known. Not all downstream targets are shown.

6.6 The role of DDR on tumor suppression when core DNA replication machinery is defective

Defects in the DNA replication machinery are only recently being appreciated for their role in driving carcinogenesis, specifically MCM deficiency [26-29]. MCMs are direct targets of ATM and ATR [30,31], the damage sensors in the DDR pathways [32,33], which are frequently altered in human cancers [16,34,35]. To understand the role of the DDR pathways on carcinogenesis when there is a deficiency of MCMs, I generated double mutants of *Chaos3* with an additional DDR mutation. Together, the *Chaos3* x DDR mutants consistently show that deficiency in the *Atm* DDR pathway results in decreased tumor latency and/or increased tumor susceptibility. This evidence supports the importance of an intact p53 pathway for tumor suppression. Tumor latency and susceptibility differed between genders, with females demonstrating an overall greater cancer susceptibility to *Atm* and *p21* deficiency than males. *Chk2* deficiency increased mammary tumor incidence and decreased tumor latency in *Chaos3* females and increased cancer susceptibility in males. These results mark the importance of intact DDR pathways in protection from carcinogenesis when the DNA replication machinery is defective, and underscore that gender and genetic background significantly impact cancer susceptibility and latency when DDR pathways are compromised.

6.7 The role of reproductive hormones and their receptors on carcinogenesis in MCM-deficient mice

Estrogen and progesterone control DNA replication in uterine epithelial cells by regulating the MCMs [36], but the role of reproductive hormones in regulating MCMs in other tissues is poorly understood. I examined the impact of estrogen and progesterone on mammary tumorigenesis in *Chaos3* mice. *Chaos3* females have higher levels of estrogen at twelve months of age compared to wild-type. This is consistent with the high incidence of CEH observed in the mice. The majority of *Chaos3* mammary tumors expressed ER α , similar to humans, though the mammary tumors of most mouse models do not [37,38]. A tissue microarray of 100 *Chaos3* tumor and mammary gland samples is currently being constructed in order to examine ER and PR by standard IHC methods. Oophorectomy, but not pregnancy, had a profound protective effect against mammary tumorigenesis in *Chaos3* mice. However, *Chaos3* oophorectomized mice were significantly more susceptible to developing other types of cancer. One human study with 24 years of follow-up finds that females with bilateral oophorectomy have decreased risk of breast and ovarian cancer, but these women have an increased risk of all-cause mortality, including fatal coronary heart disease and lung cancer [39]. This raises the prospect that humans prone to reproductive cancers who undergo prophylactic-oophorectomy may subsequently increase their risk for developing other disease phenotypes.

Estrogen itself may be playing a larger role than inducing proliferation in *Chaos3* tumorigenesis. In tissues prone to estrogen-induced cancer, estrogens can be metabolically activated to 4-hydroxylated metabolites [40]. Studies indicate the

predominant 4-hydroxylation of estrogen occurs in organs prone to estrogen-associated cancer [40]. Evidence suggests that these metabolites and their semiquinone/quinone oxidation products can induce DNA damage by forming DNA adducts, causing single-strand breaks, and 8-hydroxylation of guanine bases [40]. Estrogen has been shown to induce genetic mutations including aneuploidy, structural chromosomal aberrations, gene mutations, gene amplification and deletion, and microsatellite instability [40]. These studies show that estrogen-induced carcinogenesis does not occur solely by receptor-mediated cell proliferation. *Chaos3* cells are significantly more sensitive to added replication stress such as aphidicolin treatment [29,41]. Therefore in the context of the *Chaos3* mutation, estrogen may act as a potent genotoxic stress in *Chaos3* mammary epithelial tissue where there are fewer backup replication origins to fire due to the decreased number of dormant origins.

6.8 *Chaos3* as a highly relevant model for ovarian cancer

Nf1 is frequently lost in both human mammary and ovarian cancers [34,35], suggesting that *Nf1* deficiency may be a prominent driver in ovarian cancer as well. However, *Chaos3*-C3H mice do not frequently develop ovarian tumors. *Nf1* may work differently in ovarian tissue in mice *versus* humans or C3H and B6 background may not be susceptible to or may be protecting mice from ovarian tumors.

Ovarian cancer is the most lethal of the female reproductive cancers, with 22,280 cases in the United States annually and 15,500 deaths [42]. The epithelial subtype constitutes 90% of the cases, and ~5% are the granulosa cell subtype [43,44]. Currently, there are no ideal mouse models for either subtype due to the artificial

context of carcinogens, non-ovary specific driving promoters (leaky systems), infertile females, or failure of offspring to develop tumors [43].

However, the NZC strain has a 17% spontaneous incidence of the ovarian granulosa cell tumor subtype at ~2 years of age in virgins (15% breeders) [45]. For comparison, C3H inbred female virgins have a 2% spontaneous mammary tumor incidence (12% breeders) at typically >18 months of age, which increases to >80% incidence at ~12 months when in the presence of *Chaos3* homozygosity [29,46]. Interestingly, all NZC mice developing spontaneous ovarian granulosa cell tumors also developed uterine cystic hyperplasia [45]. A high incidence of uterine cystic endometrial hyperplasia was observed in *Chaos3*-C3H mice and quantified at 28% in C3HxB6 F2s. Given that *Chaos3* tumor type is dependent upon background strain, the driving mechanism of recurrent *Nf1* loss in breast and ovarian cancers, and the similarity of CEH phenotype, it is plausible that *Chaos3*-NZC congenic mice would demonstrate an extremely high incidence of ovarian granulosa cell tumors. *Chaos3* tumorigenesis occurs earlier than spontaneous tumors typically develop naturally in the background strains (12-18 months *vs.* >18 months). Thus, the ovarian tumor latency in *Chaos3*-NZC would also likely be decreased. As an alternative, NZO mice have a 5.5% spontaneous incidence [47]; 4% in virgins and 14.3% breeder incidence has also been reported [45]. Therefore, generation of a *Chaos3* granulosa cell tumor model of ovarian cancer may be fairly straight forward.

Unfortunately, spontaneous tumors of the epithelial subtype have not been observed/reported in any specific mouse strain. Interestingly however, in the 6 of 21 *Chaos3* oophorectomized females in which ovarian tissue re-grew, all were

accompanied by ovarian disease. One animal demonstrated only an ovarian cyst, 3 animals developed ovarian cysts and tubulostromal hyperplasia, 1 animal developed a luteoma (a rare type of ovarian tumor), and 1 animal developed an ovarian adenocarcinoma (the epithelial type that comprises 90% of all human ovarian cancers) [43]. This indicates it is at least possible to generate ovarian tumors of interest in *Chaos3* mice.

p53 is mutated in 96% of serous ovarian cancers, indicating it is a critical step in the development of ovarian tumors [35]. It has been shown that *Chaos3* x *p53* mice develop bone tumors, lymphomas, and histiocytic sarcomas. However, the study was conducted in the B6 background. Given the dependency of tumor type on background strain (particularly C3H *vs*. B6), deficiency of *p53* in strains amenable to reproductive cancers would be expected to significantly shift the tumor spectrum. If *C3/+ p53+/-* and *C3/3 p53+/-* animals were aged in a C3H, FVB, or NZC background, ovarian tumors and the incidence of other reproductive cancers would be expected to increase, and tumor latency would decrease. Incidence of ovarian tumors may be enhanced further if *Chaos3* could be introduced in an ovarian-specific Cre mouse line as well to a LoxP (floxed) *p53* mouse. *Chaos3/3* Ovar-Cre could then be crossed to *Chaos3/3* LoxP *p53*, resulting in mice mutant for *Chaos3* that had ovarian-specific *p53* deficiency. It may be equally beneficial to generate *Nf1+/-* Ovar-Cre LoxP *p53* mice to achieve the same ends. While spontaneous models of epithelial ovarian cancer may be more difficult to achieve, eventual *Chaos3* and *Nf1* models of ovarian cancer are promising.

6.9 *Chaos3* Carcinogenesis

How does a point mutation in a DNA replication gene lead exclusively to mammary carcinogenesis? The defective *Chaos3* helicase is unstable [48]. The consistency of recurring CNAs, as opposed to somatic point mutations, in *Chaos3* tumors indicates that the mutant helicase may be predisposed to stalling at particular genomic regions that are difficult to replicate. This is consistent with recurring CNAs in yeast flanking long terminating repeat (LTR) sites [49]. *Chaos3* cells have fewer dormant replication origins to fire in response to fork stalls or collapse [50,51]. Despite activation of multiple fork recovery pathways, these replication intermediates persist into M phase, increasing the number of abnormal anaphase cells with lagging chromosomes and/or acentric fragments [50,51]. This results in GIN and the micronuclei phenotype originally observed in the *Chaos3* mice. This also helps account for the high levels of aneuploidy observed in *Chaos3* tumor cells. *Chaos3* animals develop tumors in a specific window of time, rarely before ten months of age without additional perturbation. This indicates that other age-related factors are involved that may or may not relate to changing hormonal conditions, and this aspect of carcinogenesis remains to be explored. There is a pan-reduction of MCMs in *Chaos3* cells, which can be restored by knocking down Dicer and Drosha [48]. With MCMs downregulated, cell cycle is delayed. Hormones regulate the MCMs through miRNAs [52]. This suggests that either hormones cause the pan-reduction of MCMs, or the imbalance of the MCMs leaves the hormonal system attempting to resolve the imbalance, perhaps through the increased levels of estrogen observed in *Chaos3* mice at 12 months. When hormonal conditions and potentially other age-related factors are

in place, and tumor suppressors such as *Nf1* are lost due to problematic replication, the cells transform. There are several areas that require further research, but these studies have helped to make the overall picture of *Chaos3* carcinogenesis is much clearer. Overall, the *Chaos3* model is highly relevant for the study of human cancer and the consequences of DNA replication defects on carcinogenesis.

6.10 References

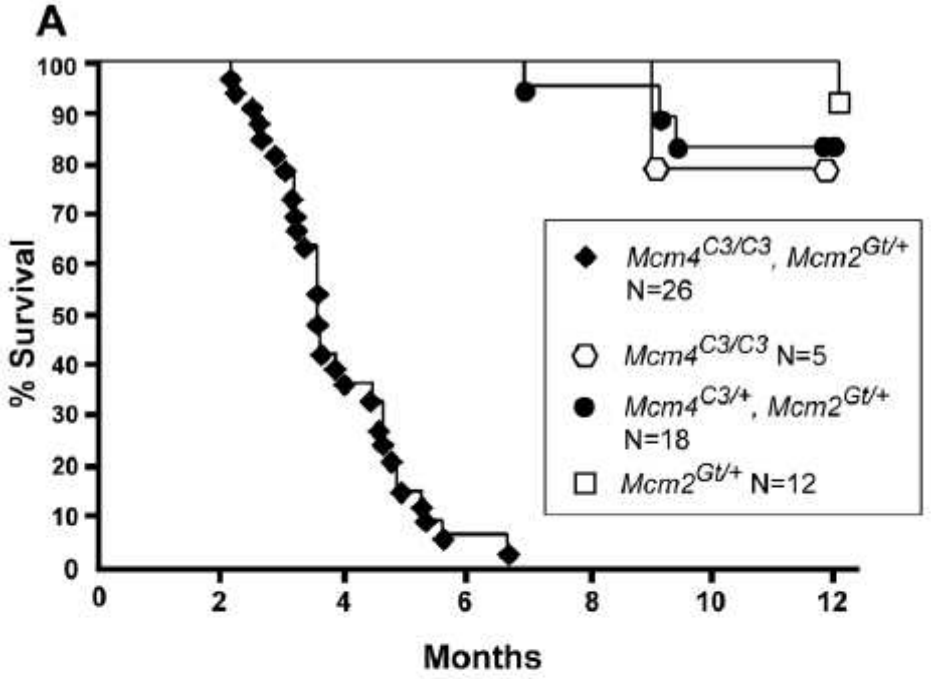
1. Fox, E.J., Salk, J.J. & Loeb, L.A. Cancer genome sequencing--an interim analysis. *Cancer Res* **69**, 4948-50 (2009).
2. Perou, C.M. et al. Molecular portraits of human breast tumours. *Nature* **406**, 747-52 (2000).
3. Sjoblom, T. et al. The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268-74 (2006).
4. Pharoah, P.D. et al. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* **31**, 33-6 (2002).
5. Carey, L.A. et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *Jama* **295**, 2492-502 (2006).
6. Hube, F., Mutawe, M., Leygue, E. & Myal, Y. Human small breast epithelial mucin: the promise of a new breast tumor biomarker. *DNA Cell Biol* **23**, 842-9 (2004).
7. Jacks, T. et al. Tumour predisposition in mice heterozygous for a targeted mutation in Nf1. *Nat Genet* **7**, 353-61 (1994).
8. Li, T. et al. Phase II trial of the farnesyltransferase inhibitor tipifarnib plus fulvestrant in hormone receptor-positive metastatic breast cancer: New York Cancer Consortium Trial P6205. *Ann Oncol* **20**, 642-7 (2009).
9. Sparano, J.A. et al. Phase II trial of tipifarnib plus neoadjuvant doxorubicin-cyclophosphamide in patients with clinical stage IIB-IIIC breast cancer. *Clin Cancer Res* **15**, 2942-8 (2009).
10. FDA. FOOD AND DRUG ADMINISTRATION CENTER FOR DRUG EVALUATION AND RESEARCH ONCOLOGIC DRUGS ADVISORY MEETING. (2009).
11. Riely, G.J. et al. A phase II trial of Salirasib in patients with lung adenocarcinomas with KRAS mutations. *J Thorac Oncol* **6**, 1435-7 (2011).
12. Mendes-Pereira, A.M. et al. Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci U S A* (2012).
13. Schmegner, C., Berger, A., Vogel, W., Hameister, H. & Assum, G. An isochore transition zone in the NF1 gene region is a conserved landmark of chromosome structure and function. *Genomics* **86**, 439-45 (2005).
14. Helmrich, A., Ballarino, M. & Tora, L. Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. *Mol Cell* **44**, 966-77 (2011).
15. Lichtenstein, P. et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* **343**, 78-85 (2000).
16. Wood, L.D. et al. The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108-13 (2007).
17. Goult, B.T. et al. The domain structure of talin: residues 1815-1973 form a five-helix bundle containing a cryptic vinculin-binding site. *FEBS Lett* **584**, 2237-41 (2010).

18. Desiniotis, A. & Kyprianou, N. Significance of talin in cancer progression and metastasis. *Int Rev Cell Mol Biol* **289**, 117-47 (2011).
19. Sakamoto, S., McCann, R.O., Dhir, R. & Kyprianou, N. Talin1 promotes tumor invasion and metastasis via focal adhesion signaling and anoikis resistance. *Cancer Res* **70**, 1885-95 (2010).
20. Kweh, F. et al. Neurofibromin physically interacts with the N-terminal domain of focal adhesion kinase. *Mol Carcinog* **48**, 1005-17 (2009).
21. Kinbara, K., Goldfinger, L.E., Hansen, M., Chou, F.L. & Ginsberg, M.H. Ras GTPases: integrins' friends or foes? *Nat Rev Mol Cell Biol* **4**, 767-76 (2003).
22. Schlaepfer, D.D., Mitra, S.K. & Ilic, D. Control of motile and invasive cell phenotypes by focal adhesion kinase. *Biochim Biophys Acta* **1692**, 77-102 (2004).
23. Arpaia, E. et al. The interaction between caveolin-1 and Rho-GTPases promotes metastasis by controlling the expression of alpha5-integrin and the activation of Src, Ras and Erk. *Oncogene* **31**, 884-96 (2012).
24. Banno, A. & Ginsberg, M.H. Integrin activation. *Biochem Soc Trans* **36**, 229-34 (2008).
25. Conklin, M.W., Ada-Nguema, A., Parsons, M., Riching, K.M. & Keely, P.J. R-Ras regulates beta1-integrin trafficking via effects on membrane ruffling and endocytosis. *BMC Cell Biol* **11**, 14 (2012).
26. Chuang, C.H., Wallace, M.D., Abratte, C., Southard, T. & Schimenti, J.C. Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. *PLoS Genet* **6**(2010).
27. Kunnev, D. et al. DNA damage response and tumorigenesis in Mcm2-deficient mice. *Oncogene* **29**, 3630-8 (2010).
28. Pruitt, S.C., Bailey, K.J. & Freeland, A. Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. *Stem Cells* **25**, 3121-32 (2007).
29. Shima, N. et al. A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nat Genet* **39**, 93-8 (2007).
30. Cortez, D., Glick, G. & Elledge, S.J. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A* **101**, 10078-83 (2004).
31. Shechter, D. & Gautier, J. MCM proteins and checkpoint kinases get together at the fork. *Proc Natl Acad Sci U S A* **101**, 10845-6 (2004).
32. Lane, D. & Levine, A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol* **2**, a000893 (2010).
33. Moynahan, M.E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* **11**, 196-207 (2010).
34. TCGA. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061-8 (2008).
35. TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609-15 (2011).

36. Pan, H., Deng, Y. & Pollard, J.W. Progesterone blocks estrogen-induced DNA synthesis through the inhibition of replication licensing. *Proc Natl Acad Sci U S A* **103**, 14021-6 (2006).
37. Mohibi, S., Mirza, S., Band, H. & Band, V. Mouse models of estrogen receptor-positive breast cancer. *J Carcinog* **10**, 35 (2011).
38. Wagner, K.U. Models of breast cancer: quo vadis, animal modeling? *Breast Cancer Res* **6**, 31-8 (2004).
39. Parker, W.H. et al. Ovarian conservation at the time of hysterectomy and long-term health outcomes in the nurses' health study. *Obstet Gynecol* **113**, 1027-37 (2009).
40. Liehr, J.G. Role of DNA adducts in hormonal carcinogenesis. *Regul Toxicol Pharmacol* **32**, 276-82 (2000).
41. Shima, N., Buske, T.R. & Schimenti, J.C. Genetic screen for chromosome instability in mice: Mcm4 and breast cancer. *Cell Cycle* **6**, 1135-40 (2007).
42. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J Clin* **62**, 10-29 (2012).
43. Fong, M.Y. & Kakar, S.S. Ovarian cancer mouse models: a summary of current models and their limitations. *J Ovarian Res* **2**, 12 (2009).
44. Jamieson, S. & Fuller, P.J. Molecular pathogenesis of granulosa cell tumors of the ovary. *Endocr Rev* **33**, 109-44 (2012).
45. Bielschowsky, M. & D'Ath, E.F. Spontaneous granulosa cell tumours in mice of strains NZC-B1, NZO-B1, NZY-B1 and NZB-B1. *Pathology* **5**, 303-10 (1973).
46. Hoag, W.G. Spontaneous Cancer in Mice. *Ann N Y Acad Sci* **108**, 805-31 (1963).
47. Goodall, C.M., Bielschowsky, M., Forster, D.R. & D'Ath, E.F. Oncological and survival reference data for NZO-B1 inbred mice. *Lab Anim* **7**, 65-71 (1973).
48. Chuang, C.H. et al. Post-transcriptional homeostasis and regulation of MCM2-7 in mammalian cells. *Nucleic Acids Res* (2012).
49. Li, X.C., Schimenti, J.C. & Tye, B.K. Aneuploidy and improved growth are coincident but not causal in a yeast cancer model. *PLoS Biol* **7**, e1000161 (2009).
50. Kawabata, T. et al. Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Mol Cell* **41**, 543-53 (2011).
51. Kawabata, T. et al. A reduction of licensed origins reveals strain-specific replication dynamics in mice. *Mamm Genome* **22**, 506-17 (2011).
52. Kuokkanen, S. et al. Genomic profiling of microRNAs and messenger RNAs reveals hormonal regulation in microRNA expression in human endometrium. *Biol Reprod* **82**, 791-801 (2010).

APPENDIX

Supplementary Figure A1-1: Premature morbidity and cancer susceptibility in *Mcm4Chaos3/Chaos3 Mcm2Gt/+* mice. (A) Kaplan-Meier survival plot of the indicated genotypes. Animals of both sexes are combined. ‘‘C3’’ = Chaos3. **(B)** Spleen and liver histopathology of a Mcm4C3/C3 Mcm2Gt/+ male diagnosed with T cell leukemic lymphoma. i. H&E stained spleen. Neoplastic cells have abundant cytoplasm, 1–2 nucleoli and a high mitotic rate, consistent with lymphoblastic lymphoma. Bar = 20 mm. ii. Neoplastic cells in spleen demonstrate immunoreactivity with anti-CD3 (brown; immunoperoxidase staining with DAB chromogen & hematoxalin counterstain), indicating T lymphocytes. Bar = 200mm. iii. In spleen, immunoreactivity (brown) with anti-PAX-5 (a B cell marker) is limited to follicular remnants and scattered individual cells. Bar = 200 mm. iv. In liver, neoplastic cells surround central veins and expand sinusoids (see also Figure S4) and demonstrate immunoreactivity (brown) with the anti-CD3 T lymphocyte marker. Bar = 50 mm. doi:10.1371/journal.pgen.1001110.g004. **(C)** X-Inactivation in *Chaos3 Mcm2* 10.5 day female embryos. Extreme semi-lethality of *Chaos3 Mcm2* females is not due to improper X-inactivation. X-GFP in *Chaos3* X-GFP mice was measured by flow cytometry.


A
100
90
80
70
60
50
40
30
20
10
0
% Survival
0
2
4
6
8
10
12
Months
Mcm4C3/C3, Mcm2Gt/+ N=26
Mcm4C3/C3 N=5
Mcm4C3/+, Mcm2Gt/+ N=18
Mcm2Gt/+ N=12

B

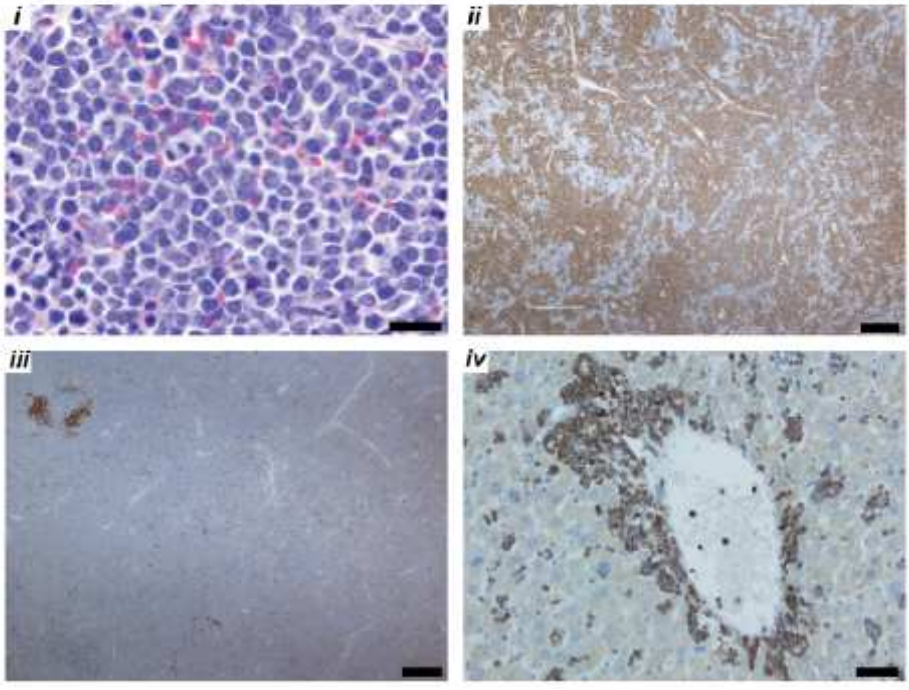
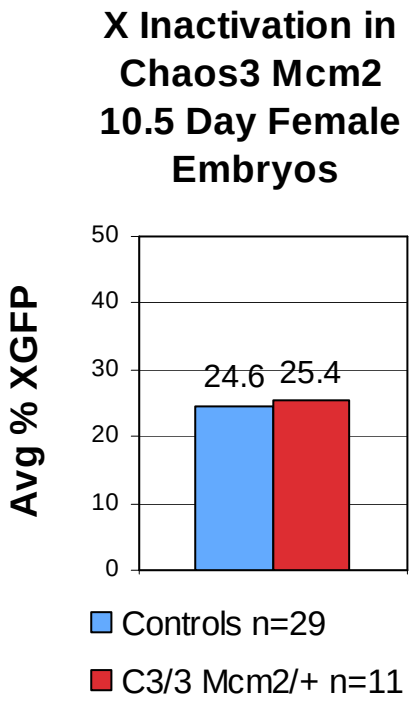

i
ii
iii
iv

Figure A1-1 Continued

C



Supplementary Table A2-1: Genes Significantly Differentially Expressed Between *Chaos3* and Other GEMMs

(FDR 0%) Identified in 2-Class SAM Analysis

Current settings

Input parameters

Data type?	Two class unpaired
Arrays centered?	FALSE
Delta	1.691
Minimum fold change	0.000
Test statistic	standard
Are data are log scale?	TRUE
Number of permutations	500.000
Input percentile for exchangeability factor s0	Automatic choice
Number of neighbors for KNN	10.000
Seed for Random number generator	1234567

Computed values

Estimate of pi0 (proportion of null genes)	0.636
Exchangibility factor s0	0.051
s0 percentile	0.000
False Discovery Rate (%)	0.000

List of Significant Genes for Delta = 1.691

Upregulated Genes

Row	Gene Name	Score(d)	Numer.	Denom.	Fold Change
6516	Mucl1	9.910	5.158	0.520	35.708
4734	H28	6.810	3.373	0.495	10.358
779	Abpb	4.267	3.053	0.715	8.299

54	1110030E23Rik	8.393	2.984	0.356	7.911
5922	Ltf	4.305	2.868	0.666	7.298
7770	Pip	4.378	2.836	0.648	7.143
6512	Muc4	4.760	2.759	0.580	6.769
3512	Egfbp2	6.809	2.681	0.394	6.413
5976	Mageb16	8.122	2.650	0.326	6.276
106	1700008P20Rik	7.928	2.563	0.323	5.910
7527	Pclo	10.081	2.353	0.233	5.107
836	Acsm3	6.716	2.246	0.334	4.742
5560	Klk1b3	7.104	2.077	0.292	4.221
9437	Slc35f3	5.065	1.989	0.393	3.968
5561	Klk1b3	6.134	1.973	0.322	3.927
9222	Sh2d4a	5.558	1.879	0.338	3.677
3238	Dmrta2	6.273	1.835	0.292	3.566
7669	Pglyrp1	4.743	1.829	0.386	3.552
366	2610305D13Rik	4.023	1.762	0.438	3.392
1183	Ankrd43	5.677	1.734	0.305	3.327
4100	Fmn2	6.127	1.710	0.279	3.271
1699	BC051142	8.433	1.686	0.200	3.218
978	Agr3	4.307	1.674	0.389	3.191
5645	Lalba	3.809	1.668	0.438	3.178
11256	Xkr6	8.043	1.632	0.203	3.100
5559	Klk1b27	4.452	1.630	0.366	3.094
802	Acbd7	4.332	1.628	0.376	3.092
3877	Fam20c	4.977	1.521	0.306	2.870
10243	Thrb	8.051	1.479	0.184	2.787
4021	Fer1l4	7.674	1.464	0.191	2.759
4591	Gpr98	5.610	1.443	0.257	2.719
1121	Amhr2	3.747	1.434	0.383	2.701
7051	Nxph3	5.409	1.429	0.264	2.693
3112	Ddx60	3.722	1.425	0.383	2.684
2481	Cldn1	3.894	1.366	0.351	2.577
166	1700042B14Rik	5.658	1.365	0.241	2.577
453	4833424O15Rik	5.429	1.364	0.251	2.575
5562	Klk1b4	4.234	1.348	0.318	2.545
888	Adamts18	4.107	1.345	0.328	2.541
1587	AY026312	4.905	1.344	0.274	2.538
4415	Gm14446	5.898	1.339	0.227	2.530
7676	Pgp	3.841	1.338	0.348	2.528
10317	Tmc5	3.537	1.315	0.372	2.489
6158	Mei1	5.012	1.309	0.261	2.478
7604	Pdxk	3.846	1.308	0.340	2.476

5139	Igkv14-111	4.636	1.278	0.276	2.425
1899	C920025E04Rik	3.525	1.274	0.361	2.418
233	2010106G01Rik	4.395	1.261	0.287	2.397
5099	Ifng	6.865	1.258	0.183	2.392
5863	Lrrc26	3.606	1.253	0.348	2.384
1568	Atxn7l3	5.382	1.252	0.233	2.382
1032	Ak7	4.249	1.251	0.294	2.381
6665	Nbea	3.664	1.244	0.340	2.369
6324	Mospd4	3.711	1.229	0.331	2.344
2949	Cyp4f41-ps	6.664	1.227	0.184	2.341
5851	Lrp6	4.437	1.197	0.270	2.293
2576	Cnnm1	5.311	1.165	0.219	2.243
5373	Jmjd5	3.689	1.164	0.316	2.241
9249	Shisa7	4.889	1.158	0.237	2.231
7055	Oas1d	6.581	1.155	0.176	2.227
3332	Dpp10	4.101	1.152	0.281	2.222
4295	Gcc2	4.134	1.132	0.274	2.192
11321	Zbtb38	5.078	1.127	0.222	2.185
4354	Gjb1	6.640	1.126	0.170	2.183
2555	Cmpk2	4.544	1.118	0.246	2.170
10415	Tmem44	4.198	1.112	0.265	2.162
6827	Nhsl1	4.174	1.107	0.265	2.154
4066	Fign	3.874	1.106	0.286	2.153
6959	Nrcam	3.953	1.103	0.279	2.148
6657	Nat1	5.526	1.095	0.198	2.136
3111	Ddx60	3.750	1.088	0.290	2.126
1685	BC026762	3.543	1.078	0.304	2.112
9322	Slc16a12	4.005	1.075	0.268	2.106
8377	R3hdm1	4.607	1.074	0.233	2.106
1893	C77370	5.644	1.071	0.190	2.100
10954	Unc80	3.920	1.062	0.271	2.087
6452	Mtap1b	4.977	1.043	0.210	2.061
6188	Mfap3l	4.263	1.043	0.245	2.060
3774	F5	3.575	1.042	0.291	2.059
6487	Mtmr9	5.510	1.036	0.188	2.050
3099	Ddx4	5.152	1.035	0.201	2.049
11602	Zfyve20	4.174	1.034	0.248	2.048
5161	Il12b	4.429	1.019	0.230	2.026
697	A430061O12Rik	4.502	1.017	0.226	2.024
4417	Gm15800	4.456	1.001	0.225	2.001
5897	Lrrtm3	4.127	0.994	0.241	1.992
4209	Fzd10	4.794	0.991	0.207	1.988

8636	Rgn	3.632	0.988	0.272	1.983
5415	Kcnh8	3.574	0.985	0.275	1.979
310	2310057N15Rik	4.812	0.983	0.204	1.977
5251	Ints4	4.387	0.980	0.223	1.973
1204	Anp32a	3.758	0.977	0.260	1.968
8709	Rnase10	5.454	0.977	0.179	1.968
5875	Lrrc46	4.278	0.974	0.228	1.964
5769	Lingo2	3.857	0.964	0.250	1.951
1979	Card14	4.302	0.960	0.223	1.945
11243	Wtap	4.247	0.935	0.220	1.912
1337	Arhgap29	3.841	0.929	0.242	1.904
9335	Slc18a1	4.212	0.917	0.218	1.889
630	8430419L09Rik	4.385	0.913	0.208	1.883
7311	Olfr920	5.146	0.908	0.176	1.876
351	2610021K21Rik	3.679	0.900	0.245	1.866
9751	Spire2	5.121	0.866	0.169	1.822
1826	Brwd3	5.013	0.865	0.173	1.822
9045	Scyl3	3.948	0.852	0.216	1.805
10471	Tmtc4	3.764	0.852	0.226	1.805
2746	Crip3	3.964	0.847	0.214	1.798
2966	D10Bwg1379e	3.906	0.845	0.216	1.796
1144	Angel1	3.955	0.844	0.214	1.796
11560	Zfp786	4.417	0.844	0.191	1.794
3363	Dsg2	3.724	0.836	0.225	1.786
2298	Ceacam1	4.248	0.816	0.192	1.761
994	Ahnak	3.601	0.801	0.222	1.743
6395	Mrps12	3.657	0.800	0.219	1.741
4180	Fstl4	4.475	0.798	0.178	1.738
6939	Nr0b1	5.011	0.794	0.158	1.734
11069	Vgll1	5.150	0.784	0.152	1.721
5799	Lnx2	3.842	0.781	0.203	1.718
6946	Nr2e3	4.434	0.780	0.176	1.717
1551	Atp8b1	4.161	0.780	0.187	1.717
397	2810474C18Rik	4.197	0.769	0.183	1.704
5442	Kctd12b	3.544	0.765	0.216	1.699
10344	Tmem116	4.038	0.753	0.186	1.685
8857	Rps14	3.614	0.750	0.208	1.682
3730	Etl4	3.550	0.740	0.209	1.670
10829	Tyrp1	3.996	0.740	0.185	1.670
2221	Cdc25a	3.711	0.733	0.198	1.663
8322	Ptprt	3.630	0.728	0.200	1.656
1336	Arhgap28	4.236	0.723	0.171	1.650

6609	N4bp2	4.312	0.723	0.168	1.650
10872	Ube2o	3.764	0.719	0.191	1.646
921	Adh4	4.185	0.713	0.170	1.640
11269	Xrcc1	3.781	0.713	0.189	1.639
6144	Med29	3.544	0.712	0.201	1.638
7060	Oas3	3.624	0.708	0.195	1.634
2273	Cdk6	3.721	0.706	0.190	1.632
9095	Selp	4.641	0.690	0.149	1.613
572	5430411K18Rik	3.904	0.654	0.167	1.573
9841	Sstr2	3.581	0.646	0.180	1.565
11275	Xrn1	3.652	0.624	0.171	1.541
10026	Synrg	4.093	0.618	0.151	1.534
5688	Lce1g	3.658	0.607	0.166	1.523
3278	Dnajc5b	4.332	0.604	0.139	1.520
712	A830093I24Rik	3.691	0.581	0.157	1.496
11014	Usp42	3.587	0.567	0.158	1.481
9191	Sfmbt2	3.547	0.562	0.158	1.476
739	Abca8b	3.941	0.551	0.140	1.466
4520	Gp5	3.532	0.512	0.145	1.426

Downregulated genes (137)

Row	Gene Name	Score(d)	Numer.	Denom.	Fold Change
8317	Ptprg	-13.185	-4.928	0.374	0.033
7755	Pik3c2g	-6.858	-2.721	0.397	0.152
4759	H60a	-6.257	-2.472	0.395	0.180
885	Adamdec1	-4.578	-2.288	0.500	0.205
2648	Comp	-3.843	-2.128	0.554	0.229
9159	Serpinh1	-5.786	-2.099	0.363	0.233
3583	Elf2	-6.086	-1.974	0.324	0.255
10060	Taf15	-5.207	-1.860	0.357	0.275
654	9430002A10Rik	-9.927	-1.738	0.175	0.300
2509	Clec4n	-4.766	-1.732	0.363	0.301
6378	Mrpl35	-8.219	-1.725	0.210	0.302
3842	Fam160a1	-5.243	-1.707	0.326	0.306
8511	Rasl12	-4.652	-1.701	0.366	0.308
5789	Lmnb1	-3.472	-1.698	0.489	0.308
6923	Nphp4	-4.563	-1.673	0.367	0.314
8202	Psap	-5.455	-1.618	0.297	0.326
9767	Spr	-6.492	-1.617	0.249	0.326
9742	Spic	-4.740	-1.615	0.341	0.326

6927	Npm3	-3.767	-1.590	0.422	0.332
6925	Npl	-4.721	-1.590	0.337	0.332
1390	Arl6ip5	-3.847	-1.558	0.405	0.340
4072	Fkbp10	-3.453	-1.453	0.421	0.365
3315	Dok7	-6.717	-1.386	0.206	0.383
4409	Gm11818	-4.220	-1.380	0.327	0.384
10879	Ube2s	-5.367	-1.371	0.255	0.387
5508	Kifc5b	-3.478	-1.369	0.393	0.387
8465	Raly	-4.057	-1.366	0.337	0.388
6782	Nes	-3.886	-1.364	0.351	0.388
9859	St8sia5	-4.712	-1.342	0.285	0.394
10500	Tnik	-6.163	-1.337	0.217	0.396
11242	Wsb1	-4.822	-1.333	0.276	0.397
585	5730437N04Rik	-4.442	-1.305	0.294	0.405
103	1700001P01Rik	-5.067	-1.296	0.256	0.407
7583	Pdgfrb	-3.575	-1.291	0.361	0.409
747	Abcc3	-5.324	-1.288	0.242	0.409
685	A130022J15Rik	-3.553	-1.268	0.357	0.415
10292	Tlr1	-3.868	-1.263	0.326	0.417
872	Adam12	-5.826	-1.243	0.213	0.423
4270	Gas5	-5.356	-1.238	0.231	0.424
4912	Hmox1	-3.637	-1.227	0.337	0.427
7071	Ociad1	-5.039	-1.225	0.243	0.428
8643	Rgs16	-3.616	-1.223	0.338	0.428
3089	Ddx25	-3.671	-1.216	0.331	0.431
10958	Unkl	-3.478	-1.199	0.345	0.436
1690	BC030307	-4.613	-1.191	0.258	0.438
9861	Stab1	-3.679	-1.188	0.323	0.439
2179	Cd300lf	-4.422	-1.184	0.268	0.440
914	Adcy7	-4.062	-1.183	0.291	0.440
2698	Cplx2	-3.866	-1.173	0.304	0.443
9675	Sord	-3.549	-1.172	0.330	0.444
8402	Rab31	-3.513	-1.162	0.331	0.447
5737	Lhfp	-3.706	-1.158	0.312	0.448
1280	Apoe	-3.882	-1.154	0.297	0.449
8567	Rbpms2	-3.590	-1.150	0.320	0.451
5325	Itgb3	-4.426	-1.128	0.255	0.458
8881	Rraga	-3.971	-1.124	0.283	0.459
3150	Dfna5	-3.736	-1.123	0.301	0.459
8965	Samd14	-4.163	-1.115	0.268	0.462
5843	Lrmp	-3.505	-1.112	0.317	0.463
9397	Slc27a6	-4.500	-1.108	0.246	0.464

10069	Taf8	-5.439	-1.106	0.203	0.465
10878	Ube2s	-4.357	-1.105	0.254	0.465
3839	Fam154b	-5.723	-1.103	0.193	0.466
10160	Tctex1d2	-4.348	-1.093	0.251	0.469
4400	Gm10083	-8.231	-1.092	0.133	0.469
10943	Ulbp1	-3.481	-1.085	0.312	0.472
2614	Col15a1	-3.622	-1.084	0.299	0.472
1924	Cadm4	-4.735	-1.084	0.229	0.472
4226	Gabbr1	-5.145	-1.071	0.208	0.476
619	6430590A07Rik	-6.043	-1.063	0.176	0.479
10493	Tnfrsf4	-4.348	-1.051	0.242	0.483
2168	Cd207	-3.591	-1.050	0.292	0.483
8623	Rftn2	-3.531	-1.047	0.297	0.484
919	Add2	-3.548	-1.025	0.289	0.491
136	1700020L24Rik	-4.172	-1.020	0.244	0.493
2781	Csdc2	-3.615	-1.017	0.281	0.494
7564	Pde1a	-3.517	-1.012	0.288	0.496
541	4933408B17Rik	-4.936	-1.010	0.205	0.497
364	2610301B20Rik	-4.919	-1.007	0.205	0.497
9206	Sfxn4	-4.818	-1.005	0.209	0.498
5354	Izumo4	-5.728	-0.993	0.173	0.502
391	2810055G20Rik	-4.385	-0.993	0.226	0.503
1730	Bcl6b	-3.705	-0.992	0.268	0.503
3756	Exosc8	-4.339	-0.988	0.228	0.504
8421	Rab5c	-3.595	-0.983	0.273	0.506
5760	Limk1	-3.939	-0.971	0.247	0.510
6300	Mobkl2c	-3.509	-0.966	0.275	0.512
1672	BC017612	-3.931	-0.955	0.243	0.516
4046	Fgfr1	-4.476	-0.955	0.213	0.516
4011	Fcrls	-3.522	-0.953	0.271	0.517
1789	Bmp8b	-4.028	-0.950	0.236	0.518
2837	Ctnnbip1	-3.760	-0.939	0.250	0.522
6999	Nub1	-3.570	-0.929	0.260	0.525
4137	Foxj3	-4.929	-0.925	0.188	0.527
1955	Cand2	-4.082	-0.921	0.226	0.528
6456	Mtap6	-3.757	-0.915	0.244	0.530
941	Adrb1	-3.530	-0.914	0.259	0.531
10306	Tm7sf4	-3.712	-0.914	0.246	0.531
11236	Wnt5a	-3.638	-0.913	0.251	0.531
3995	Fbxw2	-4.693	-0.909	0.194	0.533
9667	Sod3	-3.913	-0.897	0.229	0.537
10180	Tekt2	-3.533	-0.897	0.254	0.537

2292	Cdrt4	-5.797	-0.896	0.155	0.537
447	4831426I19Rik	-3.724	-0.891	0.239	0.539
8254	Ptbp1	-4.349	-0.888	0.204	0.540
11484	Zfp459	-3.798	-0.883	0.232	0.542
8516	Rassf4	-3.489	-0.868	0.249	0.548
716	A930005I04Rik	-3.757	-0.867	0.231	0.548
428	3110070M22Rik	-5.114	-0.860	0.168	0.551
3825	Fam125b	-3.478	-0.849	0.244	0.555
8417	Rab44	-3.869	-0.836	0.216	0.560
1843	Btk	-3.539	-0.835	0.236	0.561
8570	Rcbtb2	-3.948	-0.833	0.211	0.561
5958	Lzic	-3.662	-0.826	0.226	0.564
556	4933439C10Rik	-3.485	-0.825	0.237	0.565
2334	Cep135	-3.937	-0.816	0.207	0.568
10737	Ttc12	-3.762	-0.806	0.214	0.572
2854	Cttn	-3.565	-0.782	0.219	0.581
5861	Lrrc23	-4.208	-0.781	0.186	0.582
321	2410002O22Rik	-3.948	-0.772	0.195	0.586
2178	Cd300ld	-4.191	-0.771	0.184	0.586
5783	Lmbr1l	-4.270	-0.768	0.180	0.587
304	2310047B19Rik	-4.005	-0.759	0.190	0.591
4618	Grb10	-3.448	-0.757	0.220	0.592
6003	Manbal	-3.599	-0.737	0.205	0.600
10669	Trpc2	-3.928	-0.729	0.186	0.603
10492	Tnfrsf25	-4.752	-0.722	0.152	0.606
1407	Arpc3	-3.718	-0.715	0.192	0.609
450	4833420G17Rik	-3.810	-0.695	0.182	0.618
4868	Hip1	-3.553	-0.667	0.188	0.630
959	Aftph	-3.806	-0.652	0.171	0.636
10862	Ube2i	-3.493	-0.650	0.186	0.637
846	Actg-ps1	-3.553	-0.640	0.180	0.642
3254	Dnaja4	-3.447	-0.626	0.182	0.648
421	3110021A11Rik	-3.574	-0.624	0.175	0.649
5443	Kctd13	-3.792	-0.565	0.149	0.676
9048	Sdc3	-3.921	-0.534	0.136	0.691

Estimated Miss rates for Delta=1.69072742642906

Quantiles	Cutpoints	Miss Rate(%)
0 -> 0.05	-3.413 -> -2.091	94.320
0.05 -> 0.1	-2.091 -> -	81.430

	1.626	
0.1 -> 0.15	-1.626 -> -1.317	66.320
0.15 -> 0.2	-1.317 -> -1.052	45.540
0.2 -> 0.25	-1.052 -> -0.822	25.010
0.25 -> 0.75	-0.822 -> 0.859	6.100
0.75 -> 0.8	0.859 -> 1.077	36.040
0.8 -> 0.85	1.077 -> 1.327	50.870
0.85 -> 0.9	1.327 -> 1.653	64.380
0.9 -> 0.95	1.653 -> 2.154	78.840
0.95 -> 1	2.154 -> 3.507	92.930

Supplementary Table A2-2: Lines of Evidence for Nimblegen Sequence Capture Design		
Name	**Description**	**Number of genes**
Breast Cancer: Biomarker BioScience SuperArray	BioScience Super Array Oligo GEArray Human Breast Cancer Biomarker Microarray (OHS-402). 264 genes useful as molecular markers in the prognosis and diagnosis of breast cancer. This array is designed with clinical researchers and research pathologists in mind. The genes in the diagnosis markers group are highly associated with breast cancer based on the published literature. The genes in the prognosis markers group have been successfully used to predict the clinical outcome of breast cancer. [Array Details]	262
Breast Cancer: Mouse Mammary Cancer MMV	Theodorou, et. al., MMTV insertional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. Nature Genetics 39, 759 - 769 (2007) [PubMed.]	33
Breast Cancer: Mutation Profiling	This list includes only those genes which were found to be mutated in breast cancer, as derived from: Roman Thomas, et. al, High-throughput oncogene mutation profiling in human cancer. Nat Genet. 2007 Mar;39(3):347-51. [PubMed].	6
Breast Cancer: Novel Breast Cancer Suceptibility Loci	List derived from DF Easton et. al, Genome-wide association study identifies novel breast cancer susceptibility loci. Nature. 2007 Jun 28;447(7148):1087-93. [PubMed]	11
Breast Cancer: CAN-genes	Breast CAN-genes, derived from Wood LD, et. al. The genomic landscapes of human breast and colorectal cancers., 2007 Nov 16;318(5853):1108-13. Epub 2007 Oct 11. [PubMed]. Score is weighted by CaMP ranking: 6.0 + (140-rank)/140. The CaMP score (cancer mutation prevelance) score reflects the probability that the number of mutations observed in a gene reflects a mutation frequency that is higher than that expected to be observed by chance.	136

Breast Cancer: Protein Kinase Screen	This list includes all mutations found in primary breast cancer samples and breast cancer cell lines, in a survey of 518 protein kinases. The list does not include mutations from the hyper-mutator sample PD0119. Derived from Philip Stephens, et. al, **A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer.** Nature Genetics 37, 590 - 592 (2005). [PubMed]. Primary data obtained from: Sanger Cancer Genome Project.	36
Breast Cancer: Susceptibility	This list includes all genes which have been linked to breast cancer susceptibility, as of Dec. 2007. Derived from Michael R Stratton, Nazneen Rahman, The emerging landscape of breast cancer susceptibility. Nature Genetics, Vol. 40, No. 1. (27 December 2007) [PubMed]	15
General Cancer: Atlas Human Cancer 1.2 Array	**Atlas Human Cancer 1.2 Array** from Clontech. [Array Details]	1114
General Cancer: BioScience SuperArray	BioScience Super Array Oligo GEArray Human Cancer Microarray (OHS-802). The Oligo GEArray Human Cancer Microarray profiles profiles the expression of 440 genes that include members of several different pathways frequently altered during the progression of cancer. This array allows you to examine specific aspects of tumor progression. The genes represented by this array include but are not limited to tumor suppressors, oncogenes, signal transduction molecules, growth factors, growth factor receptors, and angiogenesis factors [Array Details]	440
General Cancer: DNA Damage Model	Halazonetis TD, et. al, An oncogene-induced DNA damage model for cancer development. Science. 2008 Mar 7;319(5868):1352-5 [PubMed.]	15
General Cancer: Land Colon Cancer Synergy	McMurray HR, et. al, Synergistic response to oncogenic mutations defines gene class critical to cancer phenotype. Nature. 2008 Jun 19;453(7198):1112-6. Epub 2008 May 25. [PubMed.]	95
General Cancer: Mutagenesis in p19ARF and p53 deficient mice	This list includes the top 346 Common Insertion Sites (CIS) identified in: AG Uren et. al., Large-scale mutagenesis in p19(ARF)- and p53-deficient mice identifies cancer genes and their collaborative networks. Cell. 2008 May 16;133(4):727-41. [PubMed]	309

General Cancer: Review Paper by Hahn and Weinberg	Hahn WC, Weinberg RA., **Modelling the molecular circuitry of cancer.** Nat Rev Cancer. 2002 May;2(5):331-41. [PubMed]. Gene list derived from MSKCC CancerGenes Resource.	41
General Cancer: Sanger Cancer Gene Census	**Sanger Cancer Gene Census**, [Census Details] The Cancer Gene Census is an ongoing effort to catalogue those genes for which mutations have been causally implicated in cancer. The original census and analysis was published in Nature Reviews Cancer and supplemental analysis information related to the paper is also available.	340
General Cancer: Review Paper by Volgestein and Kinzler	Vogelstein B, Kinzler KW., **Cancer genes and the pathways they control.** Nat Med. 2004 Aug;10(8):789-99. [PubMed]. Gene list derived from MSKCC CancerGenes Resource. If implicated with breast cancer, score is set to 3.5. Otherwise, defaults to 3.	112
Lung Cancer: Comb Kinase Survey	Genes in this paper were ranked according to activity level and frequency of activation in cell lines and tumors. Special Scoring to incorporate this ranking may be needed. Rikova K, et. al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell. 2007 Dec 14;131(6):1190-203. [PubMed]	25
Lung Cancer: Myerson	Weir BA, et. al., Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007 Dec 6;450(7171):893-8. Epub 2007 Nov 4. [PubMed]	24
MCM Associated Genes	Gins complex, mcm, damage checkpoints with dual roles in fork stability and fork stalling response, etc. Diffley Genes and Dev. 2008, Gambus Trends in Cell biology 2007. Forsburg Biochemical Society Transactions 2008. Forsburg Molecular and Cellular Biology 2008. List manually curated.	22
TCGGA: GBM Phase 1	TCGA Target List for Phase 1 of Glioblastoma multiforme.	583

Supplementary Table A2-3: MCAD Basic Default Scoring	
Category	**Score**
Mutation or genomic association study: breast cancer specific.	6
Mutation study: general cancer.	5
Biological pathways or interactions: breast cancer or MCM4 specific.	4
Biological pathways or background: general cancer.	3
Custom microarray design: breast cancer specific.	2
Custom microarray design: general cancer.	1

Supplementary Table A2-4: Sequence Capture Chip Gene List					
1200003C05Rik	Adsl	Arnt	Bcl9	Cbfa2t3	Cdc6
1300001I01Rik	Aff1	Asl	Bcl9l	Cbfb	Cdc73
1600014C10Rik	Aff3	Asns	Bcr	Cbl	Cdh1
2010106G01Rik	Aff4	Aspm	Bex1	Ccdc6	Cdh11
2310044G17Rik	Aif1	Aspscr1	Bgn	Ccdc69	Cdh20
2610207I05Rik	Aim1	Astn2	Bid	Ccl9	Cdk10
2810002O09Rik	Ak1	Atad2	Bik	Ccnb1	Cdk2
3110082I17Rik	Akap1	Atf1	Birc3	Ccnb1ip1	Cdk4
4930588N13Rik	Akap9	Atic	Birc5	Ccnb2	Cdk5
4932417H02Rik	Akna	Atm	Blm	Ccnd1	Cdk6
5330417C22Rik	Akt1	Atn1	Blmh	Ccnd2	Cdk8
5430405G05Rik	Akt2	Atp2b1	Bmp6	Ccnd3	Cdk9
5430411K18Rik	Alb	Atp8a1	Bmpr1a	Ccne1	Cdkl1
5730559C18Rik	Aldh4a1	Atp8b1	Bnip3	Ccne2	Cdkl2
6330527O06Rik	Aldh6a1	Atr	Braf	Ccr7	Cdkn1a
6330548G22Rik	Alk	Aurka	Brca1	Ccr9	Cdkn1b
6720460F02Rik	Amfr	Aurkb	Brca2	Ccrk	Cdkn1c
A2m	Angpt1	Auts2	Brd4	Cct4	Cdkn2a
Abat	Ank	Axin2	Brip1	Cd163l1	Cdkn2b
Abca1	Ankk1	Axl	Btg1	Cd2	Cdkn2d
Abca3	Ankrd1	B2m	Btg2	Cd27	Cdt1
Abcb1a	Anpep	Bad	Btk	Cd34	Cdx2
Abcb8	Anxa7	Bag1	Bub1	Cd3e	Ceacam5
Abcc4	Ap2b1	Bag3	Bub1b	Cd48	Cebpa
Abcg2	Apc	Bard1	C030046E11Rik	Cd55	Cenpa
Abi1	Apc2	Bat2	C130038G02Rik	Cd59b	Cenpf
Abi2	Ar	Bax	C330024D12Rik	Cd68	Cenpn
Abl1	Araf	Bbc3	C630035N08Rik	Cd70	Centb1
Abl2	Arfgef2	Bbs7	Cacna1f	Cd72	Centd3
Abp1	Arhgap1	BC033915	Cadm1	Cd97	Centg1
Acadm	Arhgap19	BC059842	Calcr	Cdc123	Centg2
Acads	Arhgap24	Bcl10	Camk1g	Cdc25a	Cep97
Acsl6	Arhgap25	Bcl11a	Car9	Cdc25b	Cfp
Actb	Arhgap26	Bcl11b	Cars	Cdc25c	Cgrrf1
Acy1	Arhgef12	Bcl2	Casc5	Cdc2l5	Chaf1a
Ada	Arhgef2	Bcl2l1	Casp8	Cdc42	Chd1
Adam12	Arid4a	Bcl3	Casp9	Cdc42bpa	Chd5
Adm	Armc1	Bcl6	Catsperb	Cdc42bpb	Chek1

Adrbk1	Armc7	Bcl7a	Cav1	Cdc45l	Chek2
Chic2	Cp	Ddit3	Eif2c2	Etv4	Fgfr1op2
Chl1	Cpz	Ddr1	Eif4a2	Etv6	Fgfr2
Chn1	Creb1	Ddr2	Eif4e	Evi1	Fgfr3
Chpt1	Crebbp	Ddx10	Elavl2	Evl	Fgr
Chst1	Crtc3	Ddx6	Elf1	Ewsr1	Fh1
Cic	Csf1	Degs1	Elf4	Ext1	Fhit
Ciita	Csf1r	Dek	Elk3	Ext2	Fhod3
Cirbp	Csf3	Dennd3	Ell	Ezh2	Fip1l1
Cit	Csk	Depdc1a	Eml4	Ezr	Fkbp10
Ckmt1	Csnk1d	Dffb	Eng	F2rl1	Flcn
Cks2	Csnk1g2	Dgka	Eno3	Fanca	Fli1
Clcn3	Cspp1	Dhx40	Ep300	Fancc	Flna
Cldn10a	Cstad	Dhx58	Epha1	Fancd2	Flnb
Clec12a	Ctnna1	Dip2c	Epha10	Fance	Flnc
Clic1	Ctnnb1	Dixdc1	Epha2	Fancf	Flt1
Clk1	Ctps	Dlg3	Epha4	Fancg	Flt3
Clk2	Ctsb	Dpagt1	Epha5	Fas	Flt4
Clk3	Ctsc	Dpp10	Ephb1	Fastk	Fn1
Clp1	Ctsd	Dpp4	Ephb2	Fbp1	Fnbp1
Cltc	Ctse	Dpyd	Ephb3	Fbxo15	Fos
Cnbp	Cugbp2	Dst	Ephb4	Fbxo31	Fosl1
Cnbp2	Cxcl1	Dtl	Eps15	Fbxo5	Fosl2
Cnnm4	Cxcl15	Dusp15	Eras	Fbxw7	Foxc2
Cnot8	Cxcr7	Dusp2	Erbb2	Fcgr2b	Foxo1
Cnr2	Cyb5	Dvl1	Erbb3	Fchsd2	Foxo3
Cntn3	Cyb5r4	Dvl3	Erbb4	Fert2	Foxo4
Cntn6	Cyc1	Dyrk2	Erc1	Fes	Foxp1
Col11a1	Cyld	E2f1	Ercc2	Fev	Frap1
Col19a1	Cyp19a1	E2f2	Ercc3	Fgd6	Frem1
Col1a1	Cyp1a1	E2f3	Ercc4	Fgf1	Frzb
Col1a2	Cyr61	E2f5	Ercc5	Fgf10	Fstl3
Col3a1	D430039N05Rik	E330009J07Rik	Erg	Fgf18	Fus
Col4a2	Dapk1	Ebf4	Ergic1	Fgf3	Fut8
Col6a2	Dbf4	Ebi2	Esm1	Fgf4	Fzd2
Col6a3	Dbn1	Ect2	Espn	Fgf6	Fzd5
Col7a1	Dcc	Egf	Esr1	Fgf7	Fzd9
Col9a3	Dck	Egfr	Esr2	Fgf8	Gab1
Coro1a	Dclk3	Egln1	Etl4	Fgf9	Gabbr1
Coro2a	Dcn	Egr1	Ets1	Fgfr1	Gadd45b

Cox6c	Ddb2	Egr3	Etv1	Fgfr1op	Galnt5
Gapdh	Grb2	Hoxa3	Il21r	Kif21a	Ly6e
Garnl3	Grb7	Hoxa9	Il22ra2	Kif21b	Lyl1
Gas6	Grin2d	Hoxc11	Il2ra	Kif3b	Lyn
Gas7	Grk6	Hoxc13	Il6st	Kit	Mad2l1
Gata1	Gsdmd	Hoxd11	Ilk	Kitl	Maf
Gbe1	Gse1	Hoxd13	Ing1	Klf15	Mafb
Gca	Gsn	Hps5	Inhbe	Klf3	Mafk
Gcn1l1	Gspt1	Hras1	Ins2	Klf6	Magee1
Gdf15	Gstm1	Hrasls	Ints7	Klk13	Malt1
Gen1	Gstm3	Hrb	Iqgap1	Kpna2	Mamdc4
Gfi1	Gucy2f	Hrbl	Irf3	Kras	Maml2
Gfi1b	H2afy	Hsp90aa1	Irf4	Krt18	Man2b1
Gga1	Haao	Hsp90ab1	Irs4	Krt19	Map2k1
Ggh	Hbegf	Hspa5	Itga3	Krt2	Map2k2
Gins1	Hck	Hspa8	Itga6	Krt9	Map2k4
Gins2	Hdac1	Hspa9	Itga9	Ktn1	Map2k7
Gins3	Hdac4	Htatip	Itgal	Lama1	Map3k1
Gins4	Hdgf	Icam5	Itgb1	Lamb1-1	Map3k14
Git1	Hdlbp	Id1	Itgb3	Lasp1	Map3k8
Glb1	Helz	Id2	Itgb4	Lass4	Mapk10
Glcci1	Hes1	Id4	Itk	Lck	Mapk11
Gli1	Hey2	Ifitm1	Itpkb	Lcn2	Mapk12
Gm672	Hgf	Ifitm2	Ivns1abp	Lcp1	Mapk13
Gm694	Hgfac	Ifnar1	Jag1	Ldha	Mapk14
Gmps	Hif1a	Igf1	Jag2	Ldhb	Mapk3
Gnai1	Hip1	Igf1r	Jak2	Lef1	Mapk7
Gnas	Hist1h4i	Igf2	Jarid1a	Lfng	Mapk8
Gnaz	Hjurp	Igf2r	Jarid1b	Lhfp	Mapk9
Golga5	Hlf	Igfbp2	Jazf1	Lifr	Mapkbp1
Gopc	Hmga1	Igfbp3	Jdp2	Litaf	Mapre1
Gpc3	Hmga2	Igfbp4	Jmjd1c	Lmo1	Mapre2
Gphn	Hmgb2	Igfbp5	Jun	Lmo2	Mark4
Gpr126	Hmgb3	Ihpk2	Junb	Lpp	Mas1
Gpr149	Hmgn2	Ikbkb	Jund	Lrba	Mast2
Gpr180	Hmmr	Ikzf1	Kctd15	Lrpap1	Matn3
Gpr19	Hnf1a	Ikzf3	Kdr	Lrrc7	Mbd2
Gpr56	Hnrnpa2b1	Il13	Kdsr	Lrrc8d	Mbip
Gpsm2	Hnrnpu	Il15	Keap1	Lrrfip1	Mcam
Grap2	Hoxa11	Il1b	Khdrbs1	Lsp1	Mcc

Grb10	Hoxa13	Il2	Kif14	Ltf	Mccc1
Mcl1	Mmp13	Myd88	Nmu	Paqr3=	Plac8
Mcm10	Mmp14	Myh1	Nono	Pard6g	Plag1
Mcm2	Mmp15	Myh11	Nos2	Parvb	Pld2
Mcm3	Mmp17	Myh9	Notch1	Patz1	Plekha6
Mcm4	Mmp1a	Mylc2pl	Notch2	Pax3	Plekha8
Mcm5	Mmp2	Mylk	Notch3	Pax5	Plg
Mcm6	Mmp9	Myo18b	Notch4	Pax7	Pltp
Mcm7	Mn1	Myo1g	Npm1	Pax8	Plxdc2
Mcm8	Mnx1	Myo3a	Nr4a1	Pbx1	Pmaip1
Mcm9	Morc3	Myrip	Nr4a3	Pcdhb22	Pml
Mdm2	Mos	Myst4	Nras	Pcm1	Pms1
Mdm4	Mpl	Naca	Nrg1	Pcna	Pms2
Mds1	Mpp7	Nbea	Nsd1	Pcsk7	Pot1a
Med13	Mprip	Nbn	Nsmce1	Pctk1	Pou2af1
Med24	Mpzl2	Nckipsd	Ntrk1	Pctk2	Pou5f1
Mef2d	Mrc1	Ncoa2	Ntrk3	Pdcd11	Ppapdc3
Melk	Mre11a	Ncoa4	Numa1	Pde4d	Pparg
Men1	Mrpl13	Ncoa6	Nup133	Pde4dip	Ppm1e
Mertk	Mrpl15	Ndc80	Nup210	Pdgfb	Ppp1r10
Met	Ms4a10	Ndrg1	Nup214	Pdgfra	Ppp2cb
Mgat1	Ms4a7	Ndufa4l2	Nup98	Pdgfrb	Prcc
Mgat4a	Msh2	Nedd4l	Nusap1	Pdpk1	Prdm16
Mgmt	Msh6	Nf1	Nut	Pecam1	Prkaa2
Mib1	Msi2	Nf2	Oaf	Per1	Prkar1a
Mif	Msn	Nfatc3	Odz1	Perp	Prkca
Mki67	Mt3	Nfe2	Odz3	Phf10	Prkcb1
Mkl1	Mta3	Nfkb1	Olig2	Phox2b	Prkcbp1
Mknk2	Mtcp1	Nfkb2	Omd	Picalm	Prkcc
Mlf1	Mtdh	Nfkbib	Orai2	Pigv	Prkcd
Mlf1ip	Mtmr2	Nfkbil1	Orc6l	Pik3ca	Prkcz
Mlf2	Mtus1	Nid1	Otof	Pik3cb	Prkd1
Mlh1	Muc1	Nin	Otud7b	Pik3cd	Prkd2
Mll1	Muc19	Ninj1	Oxct1	Pik3cg	Prkg1
Mllt1	Mutyh	Nkd2	P2ry5	Pik3r5	Pros1
Mllt10	Mx2	Nkx2-1	Pa2g4	Pim1	Prpf4b
Mllt11	Myb	Nmb	Padi1	Pim2	Prps1
Mllt3	Mybl2	Nmbr	Pafah1b2	Pir	Prr3
Mllt4	Myc	Nme1	Pag1	Pitx2	Prrx1
Mllt6	Mycl1	Nme2	Palb2	Pkn1	Psap

Mmp11	Mycn	Nme3	Paqr3	Pla2g7	Psip1
Ptbp1	Rcsd1	Runx2	Slc6a3	Syk	Thoc5
Ptch1	Recql4	S1pr1	Slc8a3	Tacc2	Thrap3
Ptcra	Recql5	Satb1	Slc9a10	Taf1	Timeless
Pten	Rel	Sbds	Slc9a8	Taf15	Timp1
Ptgis	Ret	Sbk1	Smad2	Tal1	Tk1
Ptma	Rfc4	Sbsn	Smad4	Tal2	Tln1
Ptn	Rfx2	Scml4	Smad7	Taok3	Tlx1
Ptp4a3	Rgl1	Scn3b	Smarcb1	Tap2	Tlx3
Ptpn1	Rgs2	Sdhb	Smg6	Tax1bp1	Tmem123
Ptpn11	Rhoa	Sdhc	Smo	Tbxa2r	Tmem16f
Ptpn22	Rhod	Sdhd	Sms	Tcea1	Tmem50a
Ptprc	Rhof	Sell	Socs1	Tcf12	Tmhs
Ptprcap	Rhoh	Selplg	Sod3	Tcf7	Tmie
Ptprd	Rims2	Sema3d	Sorl1	Tcfap4	Tmprss2
Ptpre	Rnf19a	Sema4b	Sp1	Tcfe2a	Tmprss6
Pvrl4	Rock1	Sema4d	Sp110	Tcfe3	Tnf
Rab11fip4	Rorc	Sema5b	Specc1	Tcfeb	Tnfaip8
Rab17	Ros1	Sema7a	Spn	Tcl1	Tnfrsf10b
Rab37	Rp1l1	Sephs2	Spna2	Tcof1	Tnfrsf17
Rab40b	Rpa1	Serinc3	Src	Tdrd6	Tnfrsf18
Rabep1	Rpgrip1	Serpinb2	Srgn	Tec	Tnfrsf1b
Rac2	Rpl11	Serpine1	Ss18	Tecta	Tnnt2
Rad21	Rpl22	Set	Ss18l1	Tek	Top1
Rad51	Rpn1	Sfmbt2	Ssbp3	Terf1	Tox3
Rad51l1	Rprm	Sfpq	St6galnac5	Terf2	Tpm4
Raf1	Rps27a	Sfrp2	St7	Tert	Tpr
Rai2	Rps6ka1	Sfrs3	Stam	Tesk1	Trim24
Ranbp17	Rps6ka2	Sgms1	Stard10	Tesk2	Trim33
Rap1gds1	Rps6ka3	Sh3d19	Stard8	Tex15	Triobp
Raph1	Rps6ka4	Sh3gl1	Stat1	Tfg	Trip11
Rara	Rps6kb1	Shh	Stat3	Tfpt	Trp53
Rasal2	Rps6kc1	Shisa5	Stat5a	Tfrc	Trp53bp1
Rasgrp1	Rras2	Six4	Stil	Tg	Trpm7
Rasgrp2	Rreb1	Skil	Stk11	Tgfa	Tsc1
Rasl11a	Rrm2	Sla2	Stk3	Tgfb1	Tsc2
Rb1	Rspo2	Slamf6	Stk32b	Tgfb2	Tsg101
Rbl2	Rspo3	Slc14a1	Stmn4	Tgfb3	Tshr
Rbm10	Rtn4ip1	Slc27a3	Sufu	Tgfbr1	Tspan14
Rbm15	Runx1	Slc2a3	Sulf2	Tgfbr2	Ttl

Rcbtb2	Runx1t1	Slc45a3	Suz12	Thbs3	Ttll10
Ttll12	Zfp668	Vim	Wnk1	Wrn	Zeb1
Ttn	Zfp710	Vpreb2	Wnt1	Wt1	Zeb2
Tyk2	Zfp74	Was	Wnt2	Xdh	Zfp384
Tyro3	Zfyve26	Wasf2	Wnt3	Xirp1	Zfp385a
Ubxd5	Zmym2	Wdr91	Wnt3a	Xpa	Zfp64
Uhrf1	Znrf1	Wee1	Wnt5a	Xpc	Zfp646
Unc45b	Veph1	Whsc1	Wnt8b	Xrcc6	
Vegfa	Vhl	Whsc1l1	Wnt9a	Zbtb16	

Supplementary Table A2-5: Sequence Capture Fold Enrichment Validation by qPCR

	Nimblegen Internal Control Genes				Breast Cancer Candidate Genes		Negative Control Genes	
Sample	1 (Runx2)	2 (Prkg1)	3 (Smg1)	4 (Nlk)	1 (Brca1)	2 (Egfr)	1 (Hmx1)	2 (Pgr)
2044B	77.28	324.03	369.73	80.32	136.21	277.88	0.01	0.00
12352					409.58	567.71	0.70	0.00
15259					487.30	744.43	1.16	0.01
2042	30.49	79.27	95.96	38.69	29.63	98.31	0.29	0.10
C3H WT	59.96	198.59	207.55	111.56	65.36	265.64	0.24	0.13

Supplementary Table A2-6: Sequence Capture Summary Statistics

Sample	Total Reads	Number & % Aligned Reads	% Aligned Reads On-Target	Fold Enrichment of Target Regions	Average Fold Coverage
2044b	8072818	7560670 (93.7%)	62.90%	301.1	36.92
2042	14254508	44 nt- 12188052 (85.5%)	34.50%	165	32.61
		64 nt- 10592292 (74.3%)	35.10%	168.1	41.83
12352	24635029	18975145 (77.0%)	56.50%	270.3	81.29
15259	23479616	21039263 (89.6%)	58.90%	282	94.02
C3H WT	10352987	9137591 (88.3%)	41.00%	196.4	29.1
Average	16158992	13248836 (84.7%)	48.10%	230	52.6

A single 88 nt read length run was conducted on sample 2042, and reads were subsequently shortened during analysis to both 64 and 44 nt to reduce error rate. The 44 nt length was used for SNP and mutation calling. Fold enrichment was calculated as (% Reads on-target)/(Target region in bases/mm9 genome size in bases), where target region = 5.69 Mb, and mm9 genome size = 2,726 Mb.

Supplementary Table A2-7: Proportions of Fold Coverage Depth for On-Target Sequence Capture Regions

	Fold Coverage				
Sample	≥1	≥3	≥5	≥10	≥25
15259	100%	99%	98%	95%	86%
12352	100%	99%	98%	96%	85%
2044B	99%	95%	92%	82%	55%
2042 (44 nt)	97%	92%	87%	75%	49%
2042 (64 nt)	98%	93%	89%	80%	57%
C3H	97%	91%	85%	72%	44%

A single 88 nt read length run was conducted on sample 2042, and reads were subsequently shortened during analysis to both 64 and 44 nt to reduce error rate. The 44 nt length was used for SNP and mutation calling.

Supplementary Table A2-8: Validated Gene Mutations in *Chaos3* Mammary Tumors

Sample	Name	Mutation	Effect	Description	Function
15259	Myo1g	G/A	Splice Site	Myosin-Ig	Precursor of minor histocompatibility antigen HA-2
2042	Acsl6	G/T	E>D	Long-chain-fatty-acid--CoA ligase 6	Catalyze formation of acyl-CoA from fatty acids, ATP, and CoA.
2042	Tdrd6	T/C	H>R	Tudor domain-containing protein 6	Required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression.
2042	Ttn	C/T	D>N	Titin (Connectin)	Cardiac and skeletal muscle protein. Disease Associations: Familial Cardiomyopathy, Tibial muscular dystrophy
2044b	Ttn	C/G	V>L	Titin (Connectin)	

Supplementary Table A2-9: Recurrent Copy Number Alterations in *Chaos3* Tumors

Amplifications

Tumor/Sample Type	Sample Name	Boundaries (Mb)		Boundaries (Mb)		Boundaries (Mb)		Boundaries (Mb)		Boundaries (Mb)		Boundaries (Mb)	
		Chr16 45-54		Chr16 38-39		Chr12 114-117		Chr17 13-14		Chr17 77-78		Chr16 27M	
Chaos3 MT	15259	45.8	50.4	38.8	39.6	114.5	117.1	13.4	13.7	-	-	-	-
Chaos3 MT	12351L	45.8	50.5	38.8	39.0	114.7	117.1	-	-	-	-	-	-
Chaos3 MT	12353A	45.0	50.5	38.8	39.0	114.7	117.1	-	-	-	-	-	-
Chaos3 MT	12115B	44.9	53.9	38.8	39.5	*qPCR	*qPCR	-	-	-	-	-	-
Chaos3 MT	12352	45.1	49.6	38.8	39.0	114.7	117.1	-	-	77.5	78.0	27.6	27.7
Chaos3 MT CL	2044B	45.8	50.5	38.8	39.5	114.6	117.1	13.3	13.7	-	-	-	-
Chaos3 MT	11929A	45.8	49.7	38.8	39.6	114.7	117.1	13.2	13.7	-	-	-	-
Chaos3 MT	16168	47.3	47.3	-	-	114.7	117.1	-	-	78.1	78.1	27.6	27.7
Chaos3 MT	16898	46.3	49.3	-	-	114.7	117.1	13.2	13.6	76.3	78.5	-	-
Chaos3 Mdl	17883	45.1	49.6	38.8	39.6	114.7	117.1	13.3	13.6	77.2	78.4	27.6	27.7
Chaos3 HS Uterus	16862	45.8	50.4	38.8	39.6	114.7	115.1	-	-	-	-	27.6	28.0
Chaos3 BT	10658	45.8	48.6	38.8	39.6	114.7	117.1	3.5	27.5	67.2	95.3	27.6	27.7
MMTV-Neu MT	3750	45.8	49.3	-	-	115.2	117.1	-	-	76.4	78.9	-	-
MMTV-Neu MT	3744	46.3	49.3	38.8	39.6	114.8	117.2	13.2	13.7	77.6	78.5	-	-

Deletions

		Chr11 78-79		Chr5 124-125		Chr4 148-149		Chr4 132-133		Chr10 79-80		Chr19 3-7		Chr19 32-33	
Chaos3 MT	15259	78.9	79.6	124.3	125.1	148.6	149.6	132.7	133.3	79.9	80.3	4.0	4.5	-	-
Chaos3 MT	12351L	78.5	79.7	-	-	149.1	149.3	-	-	-	-	3.9	4.6	-	-
Chaos3 MT	12353A	78.1	79.6	123.5	125.1	147.6	150.1	133.0	134.5	78.1	78.9	-	-	-	-
Chaos3 MT	12115B	78.9	79.4	122.7	125.3	148.6	149.4	132.9	133.8	-	-	-	-	-	-
Chaos3 MT	12352	*qPCR	*qPCR	124.5	125.1	148.1	149.1	131.9	132.2	80.6	80.7	-	-	-	-
Chaos3 MT CL	2044B	78.6	79.6	123.8	125.1	*qPCR	*qPCR	133.0	134.5	79.8	80.3	4.2	4.8	33.1	33.5
Chaos3 MT	11929A	78.9	79.3	124.3	125.3	-	-	133.2	133.6	79.0	81.0	-	-	-	-
Chaos3 MT	16168	*qPCR	*qPCR	-	-	148.1	149.5	-	-	-	-	-	-	32.2	32.6
Chaos3 F1 MT	16898	78.2	79.2	124.9	125.8	148.4	149.3	132.6	133.5	79.0	80.4	3.0	7.5	-	-
Chaos3 Mdl	17883	-	-	-	-	-	-	-	-	-	-	-	-	32.6	33.1
Chaos3 HS Uterus	16862	-	-	-	-	-	-	131.2	133.6	79.4	79.6	-	-	-	-
Chaos3 BT	10658	-	-	-	-	-	-	131.3	132.6	-	-	-	-	-	-
MMTV-Neu MT	3750	-	-	-	-	147.2	155.6	122.2	140.2	79.0	81.2	3.0	9.2	26.0	61.3
MMTV-Neu MT	3744	-	-	-	-	122.2	155.6	122.2	155.6	79.0	81.1	-	-	-	-

Legend. Recurring CNAs amongst 12 Chaos3 tumors and 2 MMTV-Neu mammary tumors identified by aCGH. The values are physical locations of the deleted or amplified regions according in the mouse genome sequence. In some cases, map locations were refined by qPCR as indicated. Samples analyzed are primary tumors except where indicated. MT = mammary tumor. CL = Cell line; HS = Histiocytic sarcoma; Mdl = Mediastinal tumor; BT = bone tumor.

Supplementary Table A2-10: Cancer and Immunity Related Genes in *Chaos3* CNAs (Extended)

	Amplified		
Function	**Chr 16**	**Chr 12**	**Chr 17**
Metastasis	Tmprss7		
Pluripotency	**Dppa4**, **Dppa2**		
Apoptosis/ Necrosis	Ift57		
Signal Transduction (Integrin, PiK3)	Trat1	**Adam6**	
Immunity/ Inflammation	Btla, CD200, CD96, **Pvrl3**, Cd47, **Retnlb**, **Retnla**	**Ig/abParts**	
Upregulated in Cancer	**Igsf11**, Upk1b, Gcet2	Tle6-like	
New Candidate Region			**Tcp10**
Cancer & Immunity Related Genes	15	6	

Supplementary Table A2-10: (Continued)

	Deleted				
Function	**Chr 4**	**Chr 5**	**Chr 11**	**Chr 10**	**Chr 19**
Tumor Suppressor		**Cdk2ap1**	**Nf1**	Apc2	Cdk2ap2, Pten
DNA Checkpoint/ Repair	Mtor/Frap1	Rad9b, Kntc1, **Gtf2h3**, **Setd8**		Stk11	**Rad9**
Apoptosis/ Necrosis	Dffa, Aptid1, **Ube4b**, Rere, **Kif1b**, Wdtc1	Diablo	Tnfaip1, Lgals9	Atcay, Cdc34, Dapk3, Gadd45b, **Oaz1**, Tmprss9	
Signal Transduction (MapK, Integrin, Wnt, PiK3)	**Pik3cd**, Map3k6	Il31	Nlk, **Ksr1**	**Csnk1g2**, Gna15, Gng7, **Mknk2**	Rps6kb2, Coro1b, Map2k2, Ndufs8
Immunity/ Inflammation	Ptafr, Pafah2			Tbxa2r, Gpx4, **Lingo3**	Tcirg1, Clcf1
Other Cancer Related	Eno1, **Arid1a**, Fgr, Pdik1l, **Sfn**	Atp2a2, Anapc7, Bcl7a, Anapc5, **Sbno1**, P2rx7		Matk, Mum1, Shc2	**Minpp1**, Rhod
New Candidate Region					**Syt12**
Cancer & Immunity Related Genes	16	13	5	18	11

Supplementary Table A2-11 Legend: *Chaos3* CGH Deleted Regions (Top) and Amplified Regions (Bottom). qPCR values are presented in percent of genomic DNA compared to C3H wild-type. Orange coloration indicates deletion (<80%), and blue indicates amplification (>115%). Cancer-related genes, deleted at high frequency in mammary tumors specifically, are underlined. *Nf1* deletion was validated at the 5' and 3' ends of the gene (*Omg* lies within an Nf1 intron near the 3' end). Note that copy number differences between *Nf1* 5' and 3' are observed in some tumors, indicating a breakpoint within the *Nf1* gene. Deletion calls were made as follows: Heterozygous = 15-80%; Homozygous = <15%; If either *Nf1* or *Omg* were <15%, the tumor sample was called homozygous deleted because full *Nf1* transcripts cannot be made. Nucleotide positions are from the mm9 mouse assembly.

Supplementary Table A2-11: qPCR Analysis of Amplifications and Deletions in *Chaos3* Mammary Tumors

		Chaos3 Mammary Tumor Samples							
		1	2	3	4	5	6	7	8
Deletion Regions		15259	12351 L	12353 A	12352	2044B CL	11929 A	16168	16898
Chr 4	Kif1b	53.0	141.1	77.8	25.9	61.3		96.5	
	Pik3cd	55.8	134.6	89.9	29.2	55.4		103.3	
	eno1	58.7	87.3	105.3	39.3	51.5		108.4	
	Rere	51.7	108.9	104.4	32.4	48.9		100.6	
Chr 5	Rad9b	67.9		62.9	23.7	90.9	70.5		93.0
	Anapc7	94.5		96.8	34.7	130.5	97.7		
	Atp2a2	82.5		93.0	46.5	124.4	97.5		
	Bcl7a	92.9		65.8	30.3	56.9	100.0		
	Il31	93.6		55.3	15.4	52.1	78.6		
	Diablo	87.1		94.1	41.6	55.5	103.6		
	Kntc1	44.7		52.5	39.5	55.2	97.8		
	Setd8	59.0		47.8	8.0	80.7	44.6		
	Gtf2h3	51.7		50.6	7.4	60.6	42.6		
Chr11	Slc46a1	82.2							36.7
	Tnfaip1	92.9	71.4	79.3		65.0	83.4	59.8	
	Nlk	103.6	63.5	105.4		67.5	116.0	50.5	
	Nos2	105.1	142.37	48.2	14.8	108.6	87.8		60.36
	Lgals9	34.3	107.0	67.6		109.3	56.0	53.8	
	Nf1 (5')	17.4	35.5	25.5	12.7	63.7	16.7	9.2	9.6
	Omg (Nf1 3')	52.9	12.0	16.7	0.7	48.0	16.3		75.1
	Rab11fip4	49.8							65.9
Amplification Regions									
Chr 12	Tle6-like	89.2	92.1	113.1	133.0	107.5	77.4	93.4	746.2
	Adam6b	148.1	176.7	173.7	10.8	150.9	119.4	170.0	
Chr 16	Cd47	98.9	137.0	84.9	121.5	115.8	109.1	88.7	
	Cd200	96.0	123.3	59.7	19.1	135.4	92.5	106.0	
	Igsf11	85.8	153.3	86.2	119.2	106.2	106.6	94.6	
	Btla	96.5	105.4	193.9	295.1	89.8	232.7	106.2	

Supplementary Table A2-12 Legend: qPCR values are presented as the percentage vs C3H DNA. *Nf1* probes were at the 5' and 3' ends of the gene. The 3' probe corresponds to the *Omg* gene that lies within an *Nf1* intron near the 3' end. Note that copy number differences between the *Nf1* 5' and 3' are observed in some tumors, indicating a breakpoint within *Nf1*. Deletion calls were made as follows: Heterozygous = 15-80%; Homozygous = <15%. If either *Nf1* or *Omg* were <15%, the tumor sample was called as homozygous deleted because full *Nf1* transcripts cannot be made. MT = mammary tumor.

Supplementary Table A2-12: qPCR analysis of *Nf1* locus in tumors

Geno & Type	Tumor #	*Nf1* 5'	*Nf1* 3' (*Omg*)
Chaos3 MT	15259	17.4	52.9
Chaos3 MT	12351 L	35.5	12.0
Chaos3 MT	12353A	25.5	16.7
Chaos3 MT	12352	12.7	0.7
Chaos3 MT	2044B CL	63.7	48.0
Chaos3 MT	11929A	16.7	16.3
Chaos3 MT	16168	9.2	ND
Chaos3 MT	16898	9.6	75.1
Chaos3 MT	12115B	31.2	38.7
Chaos3 MT	2042 CL	78.9	92.3
Chaos3 MT	919 CL	51.8	0.1
Chaos3 MT	21040	0.1	0.1
Chaos3 MT	21253	0.3	0.2
Chaos3 MT	20317	29.5	48.7
Chaos3 MT	19957	32.7	5.4
Chaos3 MT	19958	11.2	12.4
Chaos3 MT	19959	14.4	72.4
Chaos3 MT	20783	7.0	10.4
Chaos3 MT	20164	20.6	22.8
Chaos3 MT	20888	27.3	26.0
Chaos3 MT	20892	6.4	34.9
Chaos3 MT	20893	7.4	24.5
Chaos3 MT	20138	41.1	10.8
Chaos3 MT	21039	68.4	66.7
Chaos3 MT	21809	62.0	60.0
Chaos3 MT	20894	85.1	67.1
Chaos3 MT	20889	12.0	18.4
Chaos3 MT	21333	40.8	71.4
Chaos3 MT	20626	36.7	22.6
Chaos3 MT	20318	14.5	24.9
Chaos3 MT	20890	13.6	38.9
Chaos3 MT	20891	2.5	1.4
Chaos3 MT	21123	34.7	39.0
Chaos3 MT	19660	44.8	48.3
Chaos3 MT	20459	53.7	69.8
Chaos3 MT	21597	28.6	5.3
Chaos3 MT	22182	62.6	78.2
Chaos3 MT	21416	63.8	56.8
Chaos3 MT	22236	31.6	8.0
Chaos3 MT	22418	24.9	24.2
Chaos3 MT	22180	23.9	49.0
Chaos3 MT	22235	8.1	8.2
Chaos3 MT	22166	21.7	21.0
Chaos3 MT	22168	8.0	7.6

Chaos3 MT	22238	3.4	3.4
Chaos3 MT	21417	3.3	14.8
Chaos3 MT	21419	19.2	50.6
Chaos3 MT	21255	7.0	6.1
Chaos3 MT	21124	8.5	29.1
Chaos3 MT	21810	26.3	27.0
Chaos3 MT	21811	47.3	51.7
Chaos3 MT	21254	9.5	7.9
Chaos3 MT	21041	24.4	21.2
Chaos3 MT	22420	40.5	17.1
Chaos3 MT	22476	75.5	70.7
Chaos3 MT	22414	3.3	17.1
Chaos3 MT	22416	3.5	14.6
Chaos3 MT	22417	7.1	19.4
Chaos3 MT	23116	11.2	10.8
Chaos3 MT	22418	115.0	110.3
Chaos3 non-MT	19160	198.2	207.5
Chaos3 non-MT	10658	87.1	ND
Chaos3 non-MT	16862	98.5	101.2
Chaos3 non-MT	17883	97.0	88.0
PyVT		96.3	105.1
MMTV-neu1		108.6	93.2
MMTV-neu2		103.8	104.4
Chaos3 +/+ MT		107.3	109.4

Supplementary Table A2-13 Legend: x= deleted. Tumor Codes: A: 2044B; B: 12353A; C: 12351L; D: 12352; E: 15259; F: 16168; G: 12115B; H: 16898; I: 11929A. Mmu = *Mus musculus*. Some of the deletions extend further than indicated. The True (deleted) and False (not deleted) calls for human gene deletions are from TCGA level 4 data (see Methods) and refer to whether that locus is deleted at levels statistically above background. Human genes in red are potentially cancer-relevant. Red shaded regions are the "critical regions" of a deletion set. Note that the Mmu Chr 11 deletion cluster is organized in the human genome order, which is inverted and has an insertion. Thus, the critical region is actually contiguous. The "Chaos3 CNA" column refers to the % of Chaos3 mammary tumors analyzed by aCGH that contained deletions of that particular locus.

Supplementary Table A2-13: *Chaos3* -specific and Mammary Tumor-specific Recurrent Deletions Overlapping Human Breast Cancer CNAs

					Tumors								
Mouse CNA	**Gene**	**Hum Chr**	**Human CNA**	***Chaos 3* CNA**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Mmu Chr 4 148.4-149.5 Mb	SLC2A7	1	FALSE	33%		x			x				
	SLC2A5	1	FALSE	50%		x			x	x			
	GPR157	1	FALSE	50%		x			x	x			
	MIR34A	1	FALSE	50%		x			x	x			
	H6PD	1	FALSE	57%		x			x	x	x		
	SPSB1	1	FALSE	86%		x	x		x	x	x	x	
	SLC25A33	1	FALSE	86%		x	x		x	x	x	x	
	TMEM201	1	FALSE	86%		x	x		x	x	x	x	
	PIK3CD	1	FALSE	86%		x		x	x	x	x	x	
	CLSTN1	1	FALSE	71%		x		x	x		x	x	
	CTNNBIP1	1	FALSE	71%		x		x	x		x	x	
	LZIC	1	FALSE	71%		x		x	x		x	x	
	NMNAT1	1	FALSE	71%		x		x	x		x	x	
	RBP7	1	FALSE	71%		x		x	x		x	x	
	UBE4B	1	TRUE	71%		x		x	x		x	x	
	KIF1B	1	TRUE	71%		x		x	x		x	x	
	PGD	1	TRUE	33%		x		x				x	
	APITD1	1	TRUE	33%		x		x				x	
	CORT	1	TRUE	33%		x		x				x	
	DFFA	1	TRUE	33%		x		x				x	
	PEX14	1	TRUE	33%		x		x				x	
	CASZ1	1	TRUE	33%		x		x					
	TARDBP	1	FALSE	17%		x							
	MASP2	1	FALSE	17%		x							
	SRM	1	FALSE	17%		x							
Mmu Chr 5 122-125 Mb	CLIP1	12	FALSE	43%	x	x					x		
	ZCCHC8	12	FALSE	43%	x	x					x		
	RSRC2	12	FALSE	43%	x	x					x		
	KNTC1	12	FALSE	71%	x	x		x			x		x
	GPR81	12	FALSE	71%	x	x		x			x		x
	DENR	12	FALSE	71%	x	x		x			x		x
	CCDC62	12	FALSE	71%	x	x		x			x		x
	HIP1R	12	FALSE	71%	x	x		x			x		x
	VPS37B	12	FALSE	86%	x	x	x	x			x		x
	ABCB9	12	FALSE	86%	x	x	x	x			x		x
	OGFOD2	12	FALSE	86%	x	x	x	x			x		x
	ARL6IP4	12	FALSE	86%	x	x	x	x			x		x
	PITPNM2	12	FALSE	86%	x	x	x	x			x		x

	MPHOSPH9	12	FALSE	86%	x	x	x	x			x		x
	CDK2AP1	12	FALSE	86%	x	x	x	x			x		x
	SBNO1	12	FALSE	86%	x	x	x	x			x		x
	SETD8	12	FALSE	86%	x	x	x	x			x		x
	RILPL2	12	FALSE	86%	x	x	x	x			x		x
	SNRNP35	12	FALSE	100%	x	x	x	x			x	x	x
	RILPL1	12	FALSE	100%	x	x	x	x			x	x	x
	TMED2	12	FALSE	100%	x	x	x	x			x	x	x
	DDX55	12	FALSE	100%	x	x	x	x			x	x	x
	EIF2B1	12	FALSE	100%	x	x	x	x			x	x	x
	GTF2H3	12	FALSE	100%	x	x	x	x			x	x	x
	TCTN2	12	FALSE	100%	x	x	x	x			x	x	x
	ATP6V0A2	12	FALSE	43%						x		x	x
	CCDC92	12	FALSE	29%								x	x
	Zfp664	12	FALSE	14%								x	
	Fam101a	12	TRUE	14%								x	
	Ncor2	12	TRUE	14%								x	
	Scarb1	12	FALSE	14%								x	
Mmu Chr 11													
	WSB1	17	FALSE	100%	x	x	x		x		x	x	
78-79.6 Mb	KSR1	17	TRUE	100%	x	x	x		x		x	x	
	LGALS9	17	TRUE	57%		x						x	
	NOS2	17	TRUE	57%		x						x	
	NLK	17	FALSE	43%		x						x	
	TMEM97	17	FALSE	29%		x						x	
	IFT20	17	FALSE	29%		x						x	
	TNFAIP1	17	FALSE	29%		x						x	
	POLDIP2	17	FALSE	29%		x						x	
	TMEM199	17	FALSE	29%		x						x	
	SEBOX	17	FALSE	29%		x						x	
	VTN	17	FALSE	29%		x						x	
	SARM1	17	FALSE	29%		x						x	
	SLC46A1	17	FALSE	29%		x						x	
	SLC13A2	17	FALSE	29%		x						x	
	FOXN1	17	FALSE	29%		x						x	
	UNC119	17	FALSE	17%		x							
	PIGS	17	FALSE	17%		x							
	ALDOC	17	FALSE	17%		x							
	NF1	17	TRUE	100%	x	x	x		x		x	x	x
	OMG	17	TRUE	86%	x	x	x		x		x		x
	EVI2B	17	TRUE	86%	x	x	x		x		x		x
	EVI2A	17	TRUE	71%	x	x	x		x		x		
	RAB11FIP4	17	TRUE	57%	x	x	x		x				

Supplementary Table A2-14 Legend: x= deleted. Tumor Codes: A: 15259; B: 12353A; C: 12115B; D: 11929A; E: 16898; F: 2044B; G: 16892. Mmu = *Mus musculus*. Hs = *Homo sapiens*. Some of the deletions extend further than indicated. The True (deleted) and False (not deleted) calls for human gene deletions are from TCGA level 4 data (see Methods) and refer to whether that locus is deleted at levels statistically above background. Human genes in red are potentially cancer-relevant. Red shaded regions are the "critical regions" of a deletion set. The "Chaos3 CNA" column refers to the % of Chaos3 mammary tumors analyzed by aCGH that contained deletions of that particular locus.

Supplementary Table A2-14: *Chaos3* Mammary Tumor Non-specific Recurrent Deletions Overlapping Human Breast Cancer CNAs

Mouse Region	Human Gene	Hs Chr	Human CNA	*Chaos3* CNA	A	B	C	D	E	F	G
Mmu Chr 4 132.4-133.5 Mb	AIM1L	1	FALSE	43%		x	x			x	
	LIN28	1	FALSE	71%		x	x	x		x	x
	DHDDS	1	FALSE	71%		x	x	x		x	x
	HMGN2	1	FALSE	71%		x	x	x		x	x
	RPS6KA1	1	FALSE	86%		x	x	x	x	x	x
	ARID1A	1	FALSE	100%	x	x	x	x	x	x	x
	PIGV	1	FALSE	100%	x	x	x	x	x	x	x
	ZDHHC18	1	TRUE	86%	x	x	x		x	x	x
	SFN	1	TRUE	86%	x	x	x		x	x	x
	GPN2	1	TRUE	86%	x	x	x		x	x	x
	GPATCH3	1	TRUE	86%	x	x	x		x	x	x
	NR0B2	1	TRUE	86%	x	x	x		x	x	x
	NUDC	1	TRUE	86%	x	x	x		x	x	x
	TRNP1	1	TRUE	86%	x	x	x		x	x	x
	FAM46B	1	FALSE	86%	x	x	x		x	x	x
	SLC9A1	1	FALSE	71%	x		x		x	x	x
	WDTC1	1	FALSE	43%	x				x		x
	TMEM222	1	FALSE	43%	x				x		x
	SYTL1	1	FALSE	43%	x				x		x
	MAP3K6	1	FALSE	43%	x				x		x
	CD164L2	1	FALSE	43%	x				x		x
	GPR3	1	FALSE	43%	x				x		x
Mmu Chr 10 79.4-80.2 Mb	PPAP2C	19	TRUE	33%				x	x		
	MIER2	19	TRUE	33%				x	x		
	THEG	19	TRUE	33%				x	x		
	C2CD4C	19	TRUE	33%				x	x		
	SHC2	19	TRUE	33%				x	x		
	ODF3L2	19	TRUE	33%				x	x		
	MADCAM1	19	TRUE	33%				x	x		
	CDC34	19	TRUE	33%				x	x		
	GZMM	19	TRUE	33%				x	x		
	BSG	19	TRUE	33%				x	x		
	HCN2	19	TRUE	33%				x	x		
	POLRMT	19	TRUE	33%				x	x		
	FGF22	19	TRUE	33%				x	x		
	RNF126	19	TRUE	33%				x	x		
	FSTL3	19	TRUE	33%				x	x		
	PRSSL1	19	TRUE	33%				x	x		
	PALM	19	TRUE	33%				x	x		

PTBP1	19	TRUE	33%		x	x		
PRTN3	19	TRUE	33%		x	x		
ELANE	19	TRUE	33%		x	x		
CFD	19	TRUE	33%		x	x		
MED16	19	TRUE	33%		x	x		
KISS1R	19	TRUE	50%		x	x		x
ARID3A	19	TRUE	50%		x	x		x
WDR18	19	TRUE	50%		x	x		x
GRIN3B	19	TRUE	50%		x	x		x
CNN2	19	TRUE	50%		x	x		x
ABCA7	19	TRUE	50%		x	x		x
HMHA1	19	TRUE	50%		x	x		x
POLR2E	19	TRUE	50%		x	x		x
GPX4	19	TRUE	50%		x	x		x
SBNO2	19	TRUE	50%		x	x		x
STK11	19	TRUE	50%		x	x		x
ATP5D	19	TRUE	50%		x	x		x
MIDN	19	TRUE	50%		x	x		x
CIRBP	19	TRUE	33%		x	x		
EFNA2	19	TRUE	33%		x	x		
MUM1	19	TRUE	33%		x	x		
NDUFS7	19	TRUE	33%		x	x		
GAMT	19	TRUE	33%		x	x		
DAZAP1	19	TRUE	33%		x	x		
RPS15	19	TRUE	33%		x	x		
APC2	19	TRUE	33%		x	x		
PCSK4	19	TRUE	33%		x	x		
REEP6	19	TRUE	33%		x	x		
ADAMTSL5	19	TRUE	33%		x	x		
MEX3D	19	TRUE	50%		x	x	x	
MBD3	19	TRUE	50%		x	x	x	
TCF3	19	TRUE	50%		x	x	x	
ONECUT3	19	TRUE	67%	x	x	x	x	
ATP8B3	19	TRUE	67%	x	x	x	x	
REXO1	19	TRUE	67%	x	x	x	x	
KLF16	19	TRUE	67%	x	x	x	x	
SCAMP4	19	TRUE	67%	x	x	x	x	
ADAT3	19	TRUE	67%	x	x	x	x	
CSNK1G2	19	TRUE	67%	x	x	x	x	
BTBD2	19	TRUE	67%	x	x	x	x	
MKNK2	19	TRUE	67%	x	x	x	x	
MOBKL2A	19	TRUE	67%	x	x	x	x	
AP3D1	19	TRUE	67%	x	x	x	x	
DOT1L	19	TRUE	67%	x	x	x	x	
PLEKHJ1	19	TRUE	67%	x	x	x	x	

	SF3A2	19	TRUE	67%	x	x	x	x
	AMH	19	TRUE	67%	x	x	x	x
	JSRP1	19	TRUE	67%	x	x	x	x
	OAZ1	19	TRUE	67%	x	x	x	x
	LINGO3	19	TRUE	67%	x	x	x	x
	LSM7	19	TRUE	33%		x	x	
	TMPRSS9	19	FALSE	33%		x	x	
	TIMM13	19	FALSE	33%		x	x	
	LMNB2	19	FALSE	33%		x	x	
	GADD45B	19	FALSE	17%		x		

Supplementary Table A2-15: Comparison of *Chaos3* Commonly Deleted Genes to COSMIC Database

***Chaos3* CNAs**		**COSMIC Mutations**		
Chr	**# Genes in Region(s)**	**Total Mutations**	**Prominantly Mutated**	**Mutated in Breast Cancer**
4	105	62	MTOR (22), ARID1A (8), MAP3K6 (6)	MAP3K6 (4), PTCHD2 (1), RHD (1), STX12 (1)
5	58	30	ANAPC5 (6), SBNO1 (6), KNTC1 (5), P2RX7 (4)	ANAPC5 (2), CLIP1 (1), GPR81 (2), P2RX7 (1), SBNO1 (5)
10	122	243	STK11 (214)	ADAT3 (1), APC2 (3), DAZAP1 (1), UQCR11 (1)
11	30	277	NF1 (266)	NF1 (2), NOS2 (1), PIGS (2)
19	50	2158	PTEN (2140)	PTEN (66), RCE1 (1)

Supplementary Table A2-16 Legend: MSKCC -Nf1 CGH. TCGA Breast Cancer Samples. * "Complete Tumors" are tumor samples that have mRNA, CNA, and sequencing data.

Supplementary Table A2-16: *NF1* Loss in Human Mammary Tumors

Nf1		Complete Tumors*	
2	High Level Amplification	7	2.19%
1	Gain	45	14.06%
0	Neutral/No Change	194	60.63%
-1	Hemizygous Deletion	73	22.81%
-2	Homozygous Deletion	1	0.31%
	Mutated		
	Total # of Samples	**320**	

Nf1		All Breast Tumors	
2	High Level Amplification	13	1.66%
1	Gain	116	14.81%
0	Neutral/No Change	479	61.17%
-1	Hemizygous Deletion	170	21.71%
-2	Homozygous Deletion	5	0.64%
	Mutated		
	Total # of Samples	**783**	

Nf1		TCGA Manuscript Luminal A/B	
2	High Level Amplification	6	1.90%
1	Gain	47	14.87%
0	Neutral/No Change	210	66.46%
-1	Hemizygous Deletion	53	16.77%
-2	Homozygous Deletion	0	0.00%
	Mutated		
	Total # of Samples	**316**	

Nf1		PAM50 Luminal A	
2	High Level Amplification	2	0.91%
1	Gain	23	10.45%
0	Neutral/No Change	157	71.36%
-1	Hemizygous Deletion	31	14.09%
-2	Homozygous Deletion	0	0.00%
	Mutated	7	3.18%
	Total # of Samples	**220**	

Nf1		PAM50 Luminal B	
2	High Level Amplification	4	3.15%
1	Gain	31	24.41%
0	Neutral/No Change	59	51.18%
-1	Hemizygous Deletion	27	21.26%
-2	Homozygous Deletion	0	0.00%
	Mutated	6	4.72%
	Total # of Samples	**127**	

Nf1		PAM50 Basal	
2	High Level Amplification	0	0.00%
1	Gain	8	8.60%
0	Neutral/No Change	45	48.39%
-1	Hemizygous Deletion	35	37.63%
-2	Homozygous Deletion	1	1.08%
	Mutated	4	4.30%
	Total # of Samples	**93**	

Nf1		PAM50 Claudin Low	
2	High Level Amplification	0	0.00%
1	Gain	0	0.00%
0	Neutral/No Change	6	75.00%
-1	Hemizygous Deletion	2	25.00%
-2	Homozygous Deletion	0	0.00%
	Mutated		
	Total # of Samples	**8**	

Nf1		PAM50 Her2 Enriched	
2	High Level Amplification	3	5.45%
1	Gain	7	12.73%
0	Neutral/No Change	21	38.18%
-1	Hemizygous Deletion	22	40.00%
-2	Homozygous Deletion	2	3.64%
	Mutated	0	0.00%
	Total # of Samples	**55**	

Nf1		PAM50 Normal-like	
2	High Level Amplification	0	0.00%
1	Gain	0	0.00%
0	Neutral/No Change	5	62.50%
-1	Hemizygous Deletion	2	25.00%
-2	Homozygous Deletion	0	0.00%
	Mutated	1	12.50%
	Total # of Samples	8	

Nf1		Total Breast Tumors (TCGA Manuscript 511)	
2	High Level Amplification	9	1.78%
1	Gain	70	13.61%
0	Neutral/No Change	296	57.89%
-1	Hemizygous Deletion	120	23.56%
-2	Homozygous Deletion	3	0.60%
	Mutated	18	3.55%
	Total # of Samples	**511**	**aCGH**

Supplementary Table A2-17: List of PCR & qPCR Primers

Gene Name	Primer Name	Primer Sequence
1300001I01Rik	1300001I01RikExon1_L	GTGTTTTGTCCCTCGTCGTC
1300001I01Rik	1300001I01RikExon1_R	GAAGACGGTTGTTCCCACC
Acsl6	Acsl6Exon2_L	CACTAGCGATCCCTCCACAG
Acsl6	Acsl6Exon2_R	TGGCTTCTACCTTCTGGAGC
Adam6b	Adam6b_F	AGGCCATCCTACTTGGCGTTACAT
Adam6b	Adam6b_R	GCAACAAGTGTCTGGGCCAAATGA
AK156805, AL837506.3-2	AK156805_L	ATGCCGGTTTCCAACGTAT
AK156805, AL837506.3-2	AK156805_R	TGGAAACAGATGATTTCGTCA
Akap9	Akap9Exon27_L	TTTTAAATTGGAGTGAGAATTACTGC
Akap9	Akap9Exon27_R	CTTCGCTTTCAGACAGCATC
AL837506.3-2	AL837506.3-2_L	ACCTACTTTGGAAAGAAAACTG
AL837506.3-2	AL837506.3-2_R	TGGAAACAGATGATTTCGTCA
Anapc10	Anapc10_F	GCTGAAGCAAAGCATGTGCTGAGA
Anapc10	Anapc10_R	AGTTCAGCCTCCGTCATTGTGAGT
Anapc7	Anapc7_F	GCTGTGCCAGGTTTAGCGACATTT
Anapc7	Anapc7_R	ACGAGGCAGACTGCTGAGCTTATT
Ap2b1	Ap2b1Exon11_L	AACCTTTGCAGCTTGGAATC
Ap2b1	Ap2b1Exon11_R	AAACGATCATCCCTTTTGAGAC
Apc	ApcExon16_10_L	GCAGAGCCCTAGAAAAGTGG
Apc	ApcExon16_10_R	AGAGGCTGGAGTCTTGAGGG
Arid4a	Arid4a_L	GACGACCAGCTGGAGAAGAG
Arid4a	Arid4a_R	GCCTGAGTCCTGTCGTTTTC
Atm	Atm_L	TCATGACCTCTGGAAGAGCA
Atm	Atm_R	GCAGAAGTCCAACATTTGAAGA
Atp2a2	Atp2a2_F	TGTGTTCTGTCTGCTGGGATGGAT
Atp2a2	Atp2a2_R	TGCACTAGGTGCCACAGAGCTAAA
Bcl7a	Bcl7a_F	TCCAGAAGGTTCCTTGGCAGTGAT
Bcl7a	Bcl7a_R	TGACTTCCCACAGTTCCCTGTGTT
Brca1	BRCA1_Exon9_F	AAT CTC AAG GGC CTG TCA ATC CCA
Brca1	BRCA1_Exon9_R	GCA CTG TCA GAC ATT TGG CGT GTT
Btla	Btla_F	TGACTGATGCCCAGTGGCTTCTAA
Btla	Btla_R	AGGACTGCATTGTCCCTTTGGAGA
Cd200	Cd200_F	AAGGGTAAGCTGTCTGGCATCTGT
Cd200	Cd200_R	ACTGTGTGGCTCTTTCATAGCCCT

Cd47	Cd47_F	TAATTGGACTTGAAGCCCGAGGGT
Cd47	Cd47_R	AATGCACCATCTGAAGCAGCCTTG
Diablo	Diablo_F	ACACAGACTCAACAGGGAAAGGCT
Diablo	Diablo_R	TCACCAGAGCAAGCTAGCAAGTGA
Egfr	EGFR_Exon28_F	TAG TGG GTT TAA CAG CCC TGC ACT
Egfr	EGFR_Exon28_R	TCT AGG CTC ATT TGG TGA CTG CCT
Egln	Egln1Exon5_1_L	GGGCTCTCTGGTGCTTTATG
Egln	Egln1Exon5_1_R	GCCTTGAGGGCAAAATTTAATAG
Eno1	Eno1_F	TCTGTATCCATGCCACCAGAGCAA
Eno1	Eno1_R	CAGGTGCTTTCAGAAGCAAAGGCA
Eps15	Eps15_L	GGAATTCTTTGTTGCTTTACGG
Eps15	Eps15_R	GCCAGCCCTCATGTTTACAA
Fchsd2	Fchsd2Exon21_3_L	AGGACAAAGGCCTACATTGC
Fchsd2	Fchsd2Exon21_3_R	TCTATTATAAGCACTTTGAAGATCAGC
Fhod3	Fhod3_L	ACAGACCTAGGGGAGGAGGA
Fhod3	Fhod3_R	AAGCCCTGAGGAGTGTTGAA
Gtf2h3	Gtf2h3_F	AAGGACCATGAAGCTCTGGTGTCA
Gtf2h3	Gtf2h3_R	TTGATCACCTGCAAACACAGCCAC
Hjurp	HjurpExon1-2_L	ctccccttctcgtttccac
Hjurp	HjurpExon1-2_R	attgcaaattctaggcccac
Hjurp	HjurpExon8_9_L	CACGAGGCCTCTGATGCTAC
Hjurp	HjurpExon8_9_R	GTTGTCTGGGCTCAACACTG
Hmx1	Hmx1_Intron1_F	TGG GAG CCC ACG GGA ACT TAT TTA
Hmx1	Hmx1_Intron1_R	ATT CAG GGC GTA CAA GGG ATG TCA
Igsf11	Igsf11_F	AGCAGAGCAACAAACGAACTGCTG
Igsf11	Igsf11_R	ACCCTGGGATTAGCATTTCCACCT
Il31	Il31_F	TTACCGTCGCCATGATCTTCCACA
Il31	Il31_R	AGTTACAACAGCCTCTGTCCAGCA
Jarid1b	Jarid1b_L	TCTCTAGTGCCCCCTTCTGA
Jarid1b	Jarid1b_R	TCCATCACTGGCATGTTGTT
Kif1b	Kif1b_F	ACTAACGCCCAGAAGTGAAGTGGT
Kif1b	Kif1b_R	ACATGCCTGGCAAGCTCTACCTTA
Klk13	Klk13_L	GAGCTATCATGTGGCCCCTA
Klk13	Klk13_R	CTGGGAAATCAGGTCACACA
Kntc1	Kntc1_F	GTGGGAGATTGGTTTGGCTTTGCT
Kntc1	Kntc1_R	AAACACAACACTAGGAGGTCGGCT
Lgals9	Lgals9_F	TTTGAGCTTTGCTTCCTGGTGCAG
Lgals9	Lgals9_R	AGCAAGAGGTGCCTCTGCAGATTA
Mapre2	Mapre2Exon8_1_L	ATTAGGCCTGACACCCCAC
Mapre2	Mapre2Exon8_1_R	TCTGACCAAATTTATGAGAGGAG

Mcm4	Mcm4_Intron7_F	ATG CTG TTT GCA TCT GTG ACT GGC
Mcm4	Mcm4_Intron7_R	ACT GGC TCA GAA ATC CCT GCT TCA
Mrpl48-ps, AL772394.7	Mrpl48_L	CCCTCACGGACAATCTGACT
Mrpl48-ps, AL772394.7	Mrpl48_R	CCAGACCAAGGCAACAAAAT
Myo1g	Myo1g_L	CATGCACAGAGCTGGCTTAG
Myo1g	Myo1g_R	ATCCTGAAGCAAGAGCAGGA
Nf1	Nf1_F	GTGCTGTTTGTGCTGAGCTGTGAA
Nf1	Nf1_R	TCTATTGAACTGCCCATACCCGCA
Nlk	Nlk_F	TGACTGCTGCTGCGTATAGAGCTT
Nlk	Nlk_R	ATGACGCTGACTGGAGATTCCGTT
Nlk	NSC-0272_Nlk_F	CAG CCC CAG CTC AGG TAC AG
Nlk	NSC-0272_Nlk_R	ATG ATG CGA GTG CTG ATG ATG
Nos2	Nos2_F	TCTTTGACGCTCGGAATGTAGCA
Nos2	Nos2_R	ACCTGATGTTGCCATTGTTGGTGG
Omg	Omg_F	CAAGCCTCACCAGTGCAACAAAGT
Omg	Omg_R	TGCAGTGGTTTCTTCCCTCCGTAA
Pgr	PGR_Exon1_F	TCC TTG GAG CAA GAC TCT CCC ATT
Pgr	PGR_Exon1_R	GCA GGA TGG GCA CAT GGA TGA AAT
Pik3cd	Pik3cd_F	GTTGTTGACTTCTTGCTGCCCACA
Pik3cd	Pik3cd_R	TCTGGTCAGTTCCCACACACAGTT
pisd-ps1,3	pisd-ps1,3_L	aactcagcctcctgcactgt
pisd-ps1,3	pisd-ps1,3_R	GAAAAACTCGCTGAGGTTGC
Prkg1	NSC-0247_Prkg1_F	CCC ACC GCC TTC GAC AT
Prkg1	NSC-0247_Prkg1_R	CCT GCT TAC TGT GGG CTC TTG
Ptn	PtnExon3_L	TCCACAATCAAGGTGCTGAG
Ptn	PtnExon3_R	tttaatataaaCATTAGCTGTCCTTGG
Rab11fip4	Rab11fip4_F	AACCTTTCAGGCCACCTTTCAAGC
Rab11fip4	Rab11fip4_R	TACGATGGACTTAGCCCAAAGGCA
Rad9b	Rad9b_F	AGCATGACTCTGGAGCAAAGCCTA
Rad9b	Rad9b_R	AGGTTCAAGTGCGAGTGGATGAGT
Rere	Rere_F	TGCTTGGCTCTCTGGTGCATAGAT
Rere	Rere_R	AAGGGATCTTTGTCAGGTCAGGCA
Runx2	NSC-0237_Runx2_F	CGC ATT CCT CAT CCC AGT ATG
Runx2	NSC-0237_Runx2_R	AAA GGA CTT GGT GCA GAG TTC AG
Setd8	Setd8_F	TGGTCCTCGCTGCATTGTTCTACT
Setd8	Setd8_R	ACACATGGCTTTAATGCCCAGCAG
Sfi1	Sfi1_L	GCAATGCATGTACCAACAGG
Sfi1	Sfi1_R	CGTAGGTGGCTGGAGCTAAG
sik3, BC033915	sik3_L	GGAGGTTTCACTGCACAACA

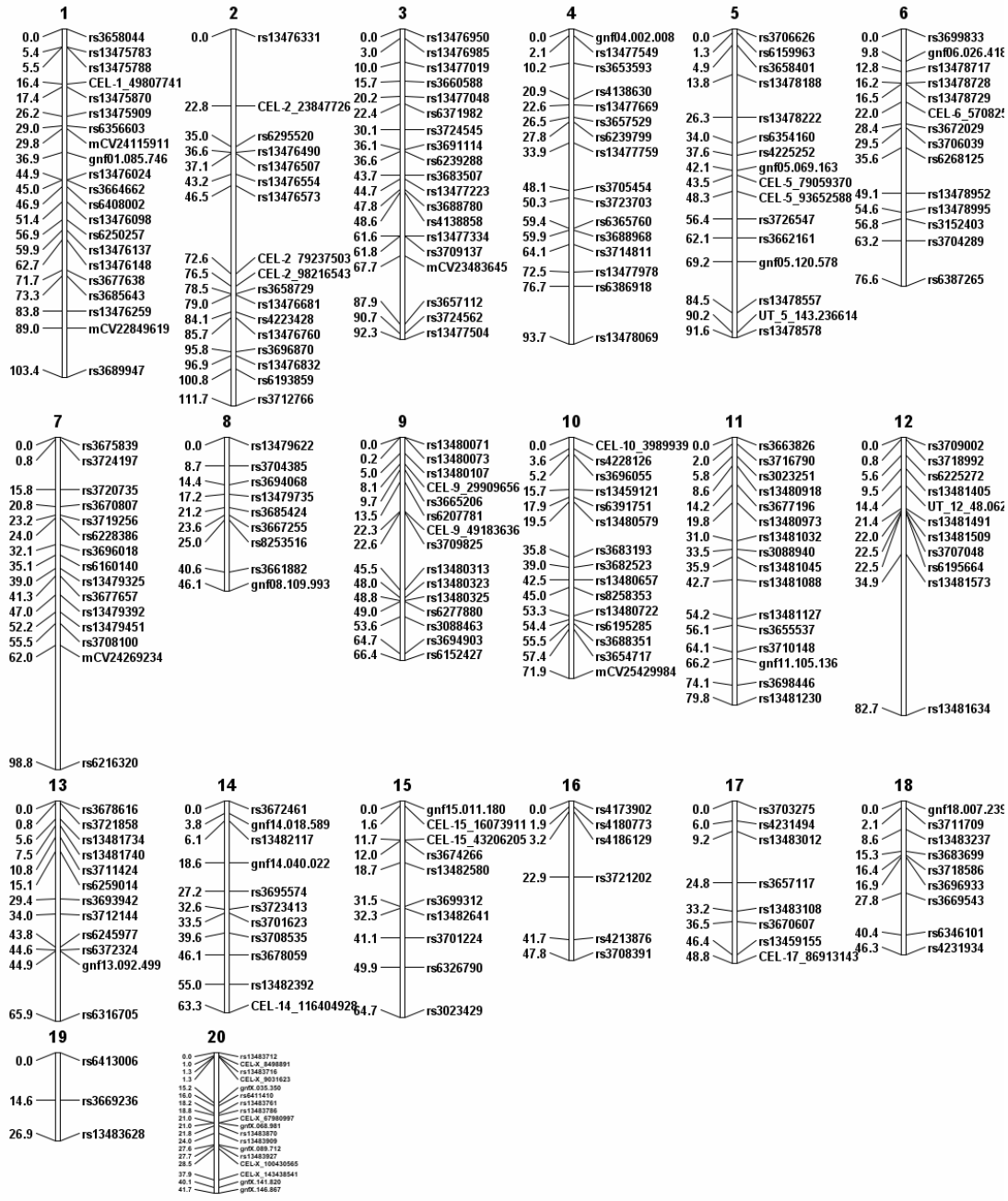
sik3, BC033915	sik3_R	GGTGCGGTCGATCTCGTA
Slc46a1	Slc46a1_F	TGGGACTCCAAGCTGATTGGCTAT
Slc46a1	Slc46a1_R	CAAACGCAAAGACCACCATTCCCA
Smg1	NSC-0268_Smg1_F	CTC GCT TAA CCA GAC TCA TCT ACT GT
Smg1	NSC-0268_Smg1_R	ACT TGG CTC AGC TGT ATG AAG GT
Supt6h	Supt6h_F	TCTGCTAAGGGAAACGTGCT
Supt6h	Supt6h_R	TTGCAAAGAGAAAGCCAGGT
Tdrd6	Tdrd6Exon1_8_L	GCCTTGCCAGATATACCGAC
Tdrd6	Tdrd6Exon1_8_R	CTAGGTTGGCCTTTGCTGAG
Tle6-like	Tle6-like_F	ACACTATCTTAGGCCTCAAGTTCTCTC
Tle6-like	Tle6-like_R	AGTCATGCCATAGCATCTGACAGT
Tle6-like	Tle6-like_F	ACACTATCTTAGGCTCAAGTTCTCTC
Tle6-like	Tle6-like_R	AGTCATGCCATAGCATCTGACAGT
Tln1	Tln1Exon43_L	GTGCCCTTAGGGAGCTGAG
Tln1	Tln1Exon43_R	AACAAGggccatgaactgg
Tnfaip1	Tnfaip1_F	AGCTCAGACTTCATCTGGCACCAA
Tnfaip1	Tnfaip1_R	ACCCTACAGCATCCCACAACTGAA
Ttn	Ttn15259_L	CTTCCTACCTTCCACGGACA
Ttn	Ttn15259_R	CCTGGTACAAGGAGGAGCAG
Ttn	TtnExon66_L	TTGGACTTTTGTGATTCATTGC
Ttn	TtnExon66_R	AAGACACCATGAGACCCAGG
Tyro	Tyro3Exon19_1_L	CTCAGCGTCATATGGTCGG
Tyro	Tyro3Exon19_1_R	CAATTTGGTCCGTCCAGG
Utp6	Utp6_F	CAGATGATGGCTGATGATGG
Utp6	Utp6_R	CCTGCTCACAAAAAGGAAGC
Zeb2	Zeb2_L	GGGCCTCTTCTTACCGTTTT
Zeb2	Zeb2_R	CGCTGTGTTTGGTTGCTAGA

Supplementary Table A3-1: *Chaos3* F1 and F2 Disease Incidence

Abbreviation	Term	Incidence (186 F2)	Percentage (186 F2)
CE-HPL	Cystic Endometrial Hyperplasia	46	24.7%
PYM	Pyometra	5	2.7%
MUM	Mucometra	8	4.3%
SCEC	Scirrhous Endometrial Carcinoma	1	0.5%
EMP	Endometrial polyp	1	0.5%
EMP-HPL	Hyperplastic Endometrial Polyp	1	0.5%
FCC	Ovarian, Bursal, or Follicular Cyst	12	6.5%
EMH	Extramedullary hematopoiesis	39	21.0%
MYL-HPL	Myeloid Hyperplasia	7	3.8%
LPD	Lipidosis	2	1.1%
HEP	Hepatitis	1	0.5%
HMGSC	hemangiosarcoma	2	1.1%
SCT	Spindle Cell Tumor	1	0.5%
LPMA	Lipoma	6	3.2%
HCC	Hepatocellular Carcinoma	1	0.5%
FBRSC	Fibrosarcoma	2	1.1%
GMS	Granulomatous steatitis	1	0.5%
SM-HPL	Smooth Muscle Hyperplasia	1	0.5%
BA-HPL	Bronchioloalveolar hyperplasia	1	0.5%
HSTCY	Histiocytosis	1	0.5%
BM-HPL	Bone Marrow Hyperplasia	2	1.1%
LEUK	Leukemia	1	0.5%
HISTSC	Histiocytic Sarcoma	94	50.5%
MT	Mammary Carcinoma	28	15.1%
BT	Osteosarcoma	26	14.0%
LYMPH	Lymphoma	31	16.7%
HLTHY	Healthy	16	8.6%
CE-HPL*	Cystic Endometrial Hyperplasia & Related (CE-HPL+PYM+MUM+EMP+EMP-HPL, SCEC non overlapping)	52	28.0%

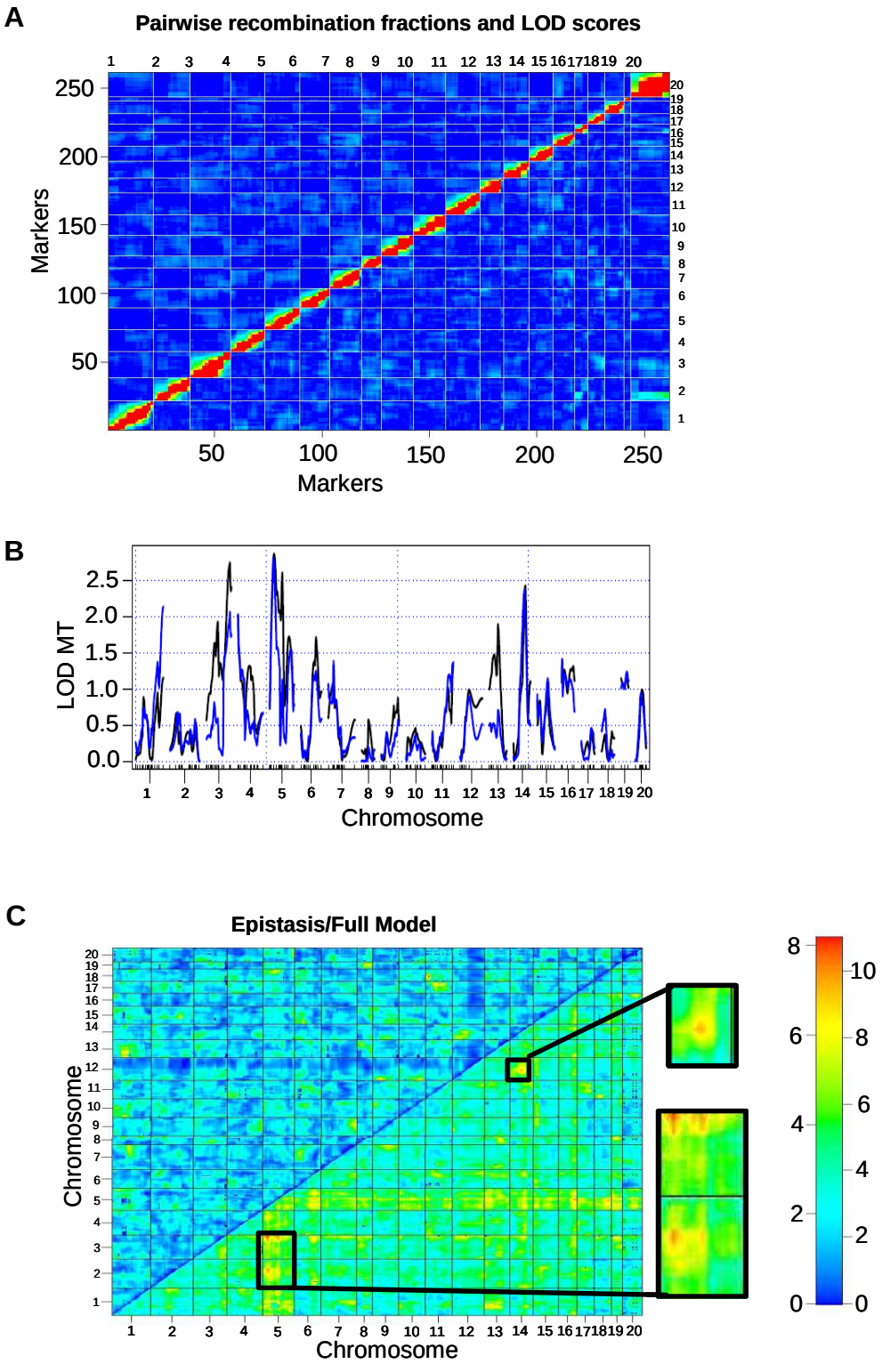
Abbreviation	Term	Incidence (21 F1)	Percentage (21 F1)
RBMSC	Rhabdomyosarcoma	1	4.8%

MD	Myodegeneration	1	4.8%
EMH	Extramedullary hematopoiesis	2	9.5%
HISTSC	Histiocytic Sarcoma	10	47.6%
MT	Mammary Carcinoma	5	23.8%
BT	Osteosarcoma	2	9.5%
LYMPH	Lymphoma	3	14.3%
CE-HPL	Cystic Endometrial Hyperplasia	9	42.9%



Supplementary Figure A3-1: SNPs assayed for QTL analysis. Shown are the 263 SNPs used on the Goldengate Beadchip. The informative C3H vs. C57/B6 SNP subset was used for subsequent QTL analysis.

Supplementary Figure A3-2: Mammary Tumor QTL analysis. (**A**) Recombination Fraction (RF) plot quality control showing successful genotyping and high-quality mapping. Markers that are close in physical proximity are strongly linked (red); Markers that are distant in physical proximity are independently assorting (blue). (**B**) Multiple QTL Mapping (MQM) analysis of mammary tumor cofactors. LOD score significance threshold was set at 2.0. Note significant peaks at Chr 1, 3, 4, 5, and 14 representing genomic loci associated specifically with mammary tumor formation. (**C**) Two-dimensional QTL scan plot. The upper and lower triagonal matrix represents the LOD scores contributed by the epistasis/interactive QTL model. Note the significant epistatic effect between chromosome 12 and 14 as well as between chromosomes 2 and 3 with chromosome 5.


A
Pairwise recombination fractions and LOD scores
Markers
Markers
B
LOD MT
Chromosome
C
Epistasis/Full Model
Chromosome
Chromosome

Supplementary Table A3-2: Mammary Tumor QTL Candidate Genes (Extended)

Location	Gene	Can Gene Description	Total # COSMIC Mut	Repro* COSMIC Mut
chr1:194325877-197195432 (2.9 Mb)	Lamb3	Laminin beta 3 (cell attachment, migration and organization)	18	4
	Hhat	Hedgehog acyltransferase (SHH signaling)	13	3
	G0s2	G0/G1switch (Potential oncogene and regulator of latent HIV)	0	0
	Kcnh1	Potassium voltage-gated channel, subfamily H (Overexpression may confer growth advantage to cancer cells and favor tumor cell proliferation)	16	4
	Irf6	Interferon regulatory factor 6 (Plays a role in regulating mammary epithelial cell proliferation by similarity)	7	2
	Traf3ip3	TRAF3 interacting Jun N terminal kinase (JNK) activating modulator	12	3
chr3:79004814-87446447 (8.4 Mb); C3H Recess Suscept	Fbxw7	F-box/WD repeat-containing protein 7 (Notch Signaling)	289	25
	Sfrp1	Secreted frizzled-related protein 2 (Wnt Signaling)	2	1
chr3:148028504-153987532 (6 Mb); C3H Recess Suscept	Msh4	MutS protein homolog 4 (DNA damage repair)	2	1
	Fubp1	Far upstream element-binding protein 1 (c-myc activation)	3	0
Chr4:5013576-8730829 (3.7 Mb); C3H Recess Suscept	Chd7	Chromodomain-helicase-DNA-binding protein 7	8	6
	Rab2a	Ras-related protein Rab-2A (Ras oncogene family)	0	0
chr4:41604056-46655694 (5.1 Mb); C3H Recess Suscept	Tln1	Talin 1 (Integrin Activation)	4	3
	Fancg	Fanconi anemia group G protein homolog	3	2
	Pax5	Paired box protein Pax-5 (TF for B cell lineage)	135	0
chr5:33660148-45362659 (11.7 Mb); C3H Recess Suscept	Fgfr3	Fibroblast growth factor receptor 3	2398	6
	Bodl	biorientation of chromosomes in cell division	6	6
	Rab28	Ras-related protein Rab-28 (TGF-B signaling, Ras family)	2	1
chr5:79059370-93652588 (14.6 Mb); C3H Recess Suscept	Rassf6	Ras association domain-containing protein	8	1
	Cdkl2	Cyclin-dependent kinase-like 2	12	5
	Epha5	Ephrin type-A receptor 5 (tyrosine kinase)	50	5
	G3bp2	Ras GTPase-activating protein-binding protein 2	8	2
chr5:127039809-139970688 (12.9 Mb); C3H Recess Suscept	Fis1	Mitochondrial fission 1 protein (apoptosis)	1	0
	Mcm7	DNA replication licensing factor	5	0
	Rasa4	Ras GTPase-activating protein 4	0	0
	Cldn15	Claudin 15	3	1

chr6:120528826-131940471 (11.4 Mb); C57/B6 Recess Resist	Styk1	Tyrosine protein-kinase STYK1 (apoptosis, angiogenesis)	6	2
	Rad51ap1	RAD51-associated protein 1	0	0
	*	Bcl2l13, Ccnd2, Fgf23, Fgf6, Ing4, Ltbr, Mlf2, Tnfrsf1a	low freq	low freq
chr7:37568559-47672477 (10.1 Mb); C57/B6 Recess Resist	Ccne1	G1/S-specific cyclin-E1	1	0
chr11:110022477-112022477 (10 Mb **); C3H Recess Suscept	Map2k6	Dual specificity mitogen-activated protein kinase kinase 6	9	1
	Socs3	Suppressor of cytokine signaling 3 (JAK-STAT, anti-apoptosis)	1	1
	Aatk	Serine/threonine-protein kinase LMTK1	10	3
	Sumo2	Small ubiquitin-related modifier 2	1	0
	Cbx4	E3 SUMO-protein ligase CBX4	3	1
	Sox9	Transcription factor SOX-9	8	1
	Srsf2	Serine/arginine-rich splicing factor 2 (pre-mRNA splicing, apoptosis)	95	0
chr13:60287243-70287243 (10 Mb**); C3H Recess Suscept; C57/B6 Recess Resist	Fancc	Fanconi anemia, complementation group C	1	0
	Ptc1	Patched homolog 1 (SHH Receptor)	333	6
	Hsd17b3	Testosterone 17-beta-dehydrogenase 3 (Androgen, estrogene, progesterone biosynthesis->Estradiol 17beta-dehydrogenase)	0	0
chr14:87238704-106808620 (19.6 Mb); C3H Dominant Suscept	Pibf1	Progesterone immunomodulatory binding factor 1	7	3
	Dach1	Dachshund 1 (Transcription factor, lost in some forms of metastatic cancer, and is correlated with poor prognosis)	8	3
	Mycbp2	MYC binding protein 2 (MYC regulation/activation)	37	9
	Commd6	(Down-regulates activation of NF-kappa-B, inhibits TNF-induced NFKB1 activation)	0	0
	Klf5	Kruppel-like factor 5 (transcription factor, maintenance of embryonic stem cell pluripotency, regulation of cell proliferation in intestinal epithelium)	7	2
	Spry2	Sprouty homolog 2 (antagonist of fibroblast growth factor FGF)	3	1

COSMIC = Catalogue Of Somatic Mutations In Cancer; *Repro-Tumors in female reproductive tissues; breast, ovary, uterus; **Single informative SNP in large region.

Supplementary Table A3-3: Bone Tumor QTLs

Chr	Left Position	Right Position	Size (Mb)	Effect Type	SNP(s)
3	3785419	3785419	*	C3H RS, B6 RR	rs13476950
3	14887895	14887895	*	B6 RS, C3H RR	rs13476985
3	23723842	23723842	*	B6 RS, C3H RR	rs13477019
3	30227687	34034360	3.8	B6 RS, C3H RR	rs3660588, rs13477048, rs6371982
3	123767337	153987532	30.2	B6 RS, C3H RR	mCV23483645, rs3657112, rs3724562, rs13477504
6	131940471	145693849	13.8	B6 RS, C3H RR	rs3704289, rs6387265
8	16233576	35610881	19.4	B6 RS	rs13479622, rs3704385, rs3694068
15	43206205	44053945	0.8	C3H RR	CEL-15_43206205, rs3674266
15	71166963	78182168	7.0	C3H RR	rs3699312, rs13482641, rs3701224

*Single informative SNP in large region. RS=Recessive Susceptibility, RR=Recessive Resistance.

Supplementary Table A3-4: Tln1 Mutation Status

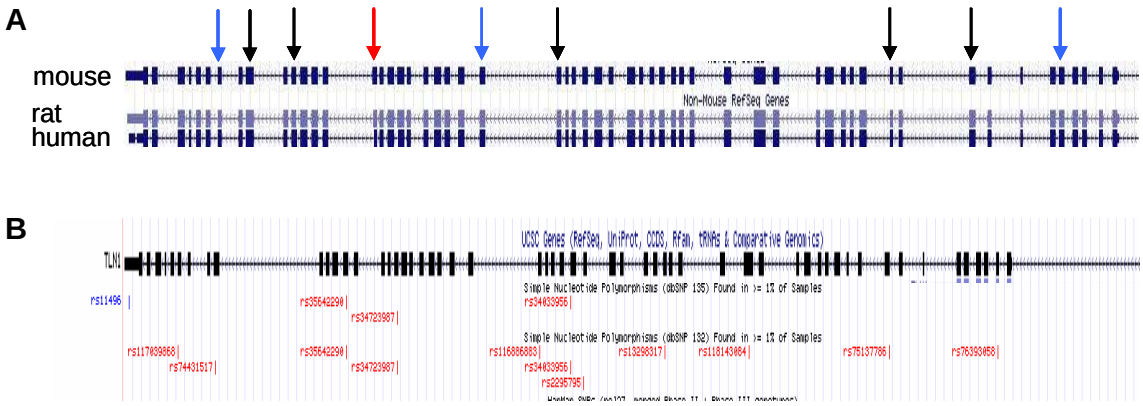
Sample	Chaos3 Genotype	Talin1 Genotype
	Wild Type Controls	
UCSC	**WT**	**G/G**
C3H Inbred	**WT**	**G/G**
B6 Inbred Male	**WT**	**G/G**
B6 Inbred Female	**WT**	**G/G**
129 Inbred	**WT**	**G/G**
	***Chaos3* Controls**	
19223 (3 mo, spleen)	**C3+/+ C3H N10F1**	**G/A**
20685 (1 mo, spleen)	**C3/3 C3H N10F2**	**A/A**
Male (tail)	**C3/3 B6 N10**	**G/G**
Female (tail)	**C3/3 B6 N10**	**G/G**
	***Chaos3* Tumors**	
12352 Mammary	**C3/3 C3H N9F2**	**A/A**
15259 Mammary	**C3/3 C3H N10**	**G/A**
2042 Mammary Cell Line	**C3/3 C3H N8F2**	**A/A**
2044B Mammary Cell Line	**C3/3 C3H N8F2**	**G/A**
17883 Mediastinal* Tumor	**C3/3 C3HxB6 F2**	**G/G**

Legend: mo=month. *Mediastinal Tumor represents either lymphoma or histiocytic sarcoma

Supplementary Table A3-5: *Chaos3* F2 *Talin 1* Mutation

Total *Talin1* Genotyping (n=129)				
0.28		39	A/A	Mut
0.42		58	G/A	Het
0.30		42	G/G	WT
	A= .49		G=.51	
Animals not developing mammary tumors (n=105)				
0.25		26	A/A	Mut
0.43		45	G/A	Het
0.32		34	G/G	WT
	A= .46		G=.54	
ChiSquare, Genotype Freq.			p=	0.74025
Fisher Exact, Allele Frequency			p=	0.7771
Animals developing mammary tumors (n=24)				
0.50		12	A/A	Mut
0.38		9	G/A	Het
0.13		3	G/G	WT
	A= .69		G=.31	
ChiSquare, Genotype Freq.			p=	0.03484
Fisher Exact, Allele Frequency			p=	0.0061

Legend: All genotyped animals have histopathology calls.



Supplementary Figure A3-3: Nonsynonomous Mutations and SNPs in *TLN1*. (**A**) *TLN1* missense mutations in tumors listed in the COSMIC database. Mutations are aligned to mouse orientation with the *Chaos3 Tln1* mutation indicated by a red arrow. Black arrows indicate locations of mutations in human breast tumors. Blue arrows represent locations of mutations in human ovarian tumors. (**B**) *TLN1* Nonsysnonomous and UTR SNPs present in >1% of the human population. Red text = nonsysnonomous SNPs; blue text = UTR SNPs.

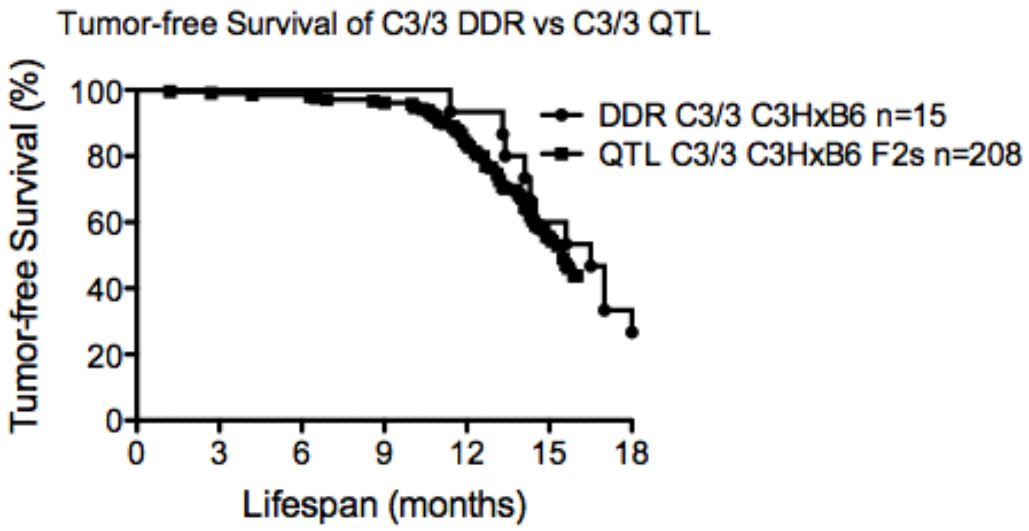
Supplementary Table A3-6: *TLN1* Non-Synonomous & UTR SNPs in > 1% of the Human Population						
dbSNP	**Chr. Position**	**Major Allele**	**Major %**	**Minor Allele**	**Minor %**	**Function**
rs11496	Chr9:35697470	T	91.89%	C	8.11%	3' UTR
rs117039868	Chr9:35699043	G	98.33%	A	1.67%	missense R (CGG) → W (TGG)
rs74431517	Chr9:35700238	T	96.59%	G	3.41%	missense T (ACA) → P (CCA)
rs35642290	Chr9:35704426	G	94.74%	A	5.26%	missense A (GCT) →T (ACT)
rs34723987	Chr9:35706047	A	98.72%	G	1.28%	missense E (GAG) →G (GGG)
rs34033956	Chr9:35711593	C	98.49%	T	1.52%	missense P (CCG) →L (CTG)
rs2295795	Chr9:35712003	G	80.46%	A	19.43%	missense S (TCG) →L (TTG)
rs13298317	Chr9:35714611	G	75.00%	A	25.00%	missense A (GCT) →V (GTT)
rs75137786	Chr9:35721781	T	66.67%	G	33.33%	missense N (AAC) →T (ACC)
rs116509251	Chr9:35732321	G	98.27%	A	1.74%	5' UTR
rs116886883	Chr9:35710605	A	80.83%	G	19.17%	missense S (TCC) →P (CCC)
rs118143084	Chr9:35717293	C	87.50%	G	12.50%	missense A (GCA) →P (CCA)
rs76393058	Chr9:35725267	C	98.61%	T	1.39%	nonsense W (TGG) → * (TAG)

Table A4-1: Median Tumor Latency of *Chaos3* x DDR Mice

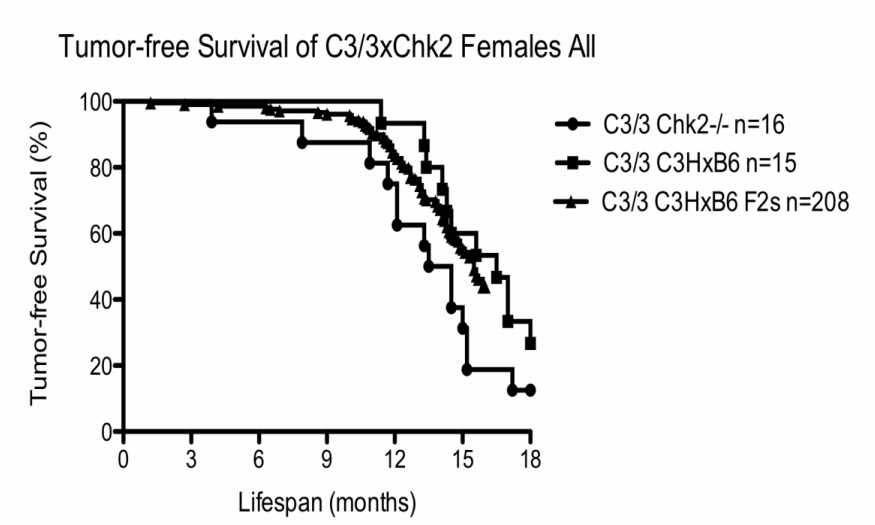
Genotype	Gender	Strain	Median Tumor Latency (mo.)	LRMCT p-value	GBWT p-value
C3/3	**M**	**C3H x FVB**	**18.00**	**-**	**-**
C3/3 Atm-/-	M	C3H x FVB	3.90	<0.0001	<0.0001
C3/3 Atm+/-	M	C3H x FVB	16.45	0.0751	0.0729
C3/+ Atm+/-	M	C3H x FVB	18.00	0.472	0.4339
C3 x Hus1	M	C3H x FVB	-	0.5117	-
C3/3 Hus1 det/neo	M	C3H x FVB	18.00	0.4334	0.355
C3/3 Hus1 det/+	M	C3H x FVB	17.95	0.6739	0.7204
C3/3 Hus1 neo/+	M	C3H x FVB	18.00	0.3714	0.4063
C3/3	**M**	**C3H x B6**	**17.20**	**-**	**-**
C3/3 p21-/-	M	B6 N8 (C3H mix)	13.00	0.1325	0.046
C3/3 p21+/-	M	B6 N8 (C3H mix)	16.85	0.405	0.3813
C3/3	**M**	**C3H x B6**	**16.55**	**-**	**-**
C3/3 Chk2-/-	M	C3H x B6	14.70	0.4272	0.2542
C3/3 Blm+/-	M	C3H x B6	18.00	0.059	0.081
C3/+	**M**	**C3H x B6**	**18.00**	**-**	**-**
C3/+ Blm+/-	M	C3H x B6	18.00	0.9512	0.9068
C3/3 (DDR)	**F**	**C3H x B6**	**16.50**	**-**	**-**
C3/3 p21-/-	F	B6 N8 (C3H mix)	12.20	0.0207	0.0055
C3/3 p21+/-	F	B6 N8 (C3H mix)	13.80	0.0115	0.0223
C3/3 Blm+/-	F	C3H x B6	17.15	0.8469	0.9013
C3/3 Chk2-/-	F	C3H x B6	14.00	0.1008	0.0581
C3/3 (QTL)	**F**	**C3H x B6 F2**	**15.50**	**-**	**-**
C3/3 DDR	F	C3H x B6	16.50	0.4236	0.3606
C3/3 Chk2-/-	F	C3H x B6	14.00	0.0189	0.027
C3/3	**F**	**C3H x FVB**	**14.95**	**-**	**-**
C3/3 Atm-/-	F	C3H x FVB	4.10	<0.0001	<0.0001
C3/3 Atm+/-	F	C3H x FVB	10.95	0.001	0.0031
C3/+ Atm+/-	F	C3H x FVB	9.30	0.0027	0.0005
C3 x Hus1	F	C3H x FVB	-	0.4756	-
C3/3 Hus1 +/+	**F**	**C3H x FVB**	**15.80**	**-**	**-**
C3/3 Hus1 det/neo	F	C3H x FVB	15.20	0.8054	0.7447
C3/3 Hus1 det/+	F	C3H x FVB	16.30	0.5765	0.9122
C3/3 Hus1 neo/+	F	C3H x FVB	17.10	0.4596	0.5618

Legend: LRMCT=Log-rank Mantel-Cox Test; GBWT=Gehan-Breslow-Wilcoxon Test.

A



B



Supplementary Figure A4-1: Tumor latency compared to 208 *Chaos3* C3H x B6 F2 females. (**A**) *C3/3* C3H x B6 female tumor latency of DDR crosses littermate controls ("DDR *C3/3*) compared to 208 *Chaos3* C3H x B6 F2 females ("*C3/3* QTL") from Wallace et al. 2012 (QTL study). The *C3/3* survival curves are not significantly different from each other (LRMCT p= 0.4236; GBWT p= 0.3606). (**B**) Female *C3/3 Chk2-/-* tumor latency. The limited number of the DDR *C3/3* cohort had less statistical power to detect smaller effects of DDR deficiency on tumor latency. Given the similarity of the DDR *C3/3* and QTL *C3/3* survival plots, the QTL *C3/3* cohort was utilized to gain statistical power to distinguish smaller effects of DDR impact. *C3/3 Chk2-/-* mice have significantly decreased time to tumor onset than *C3/3* C3H x B6 F2s (LRMCT p=0.0189; GBWT p=0.027). (**A**) and (**B**) LRMCT=Log-rank/Mantel-Cox Test; GBWT=Gehan-Breslow-Wilcoxon Test.