

AWARD NUMBER: W81XWH-11-1-0616

TITLE: From Breast to Bone: Tracking Gene Expression Changes Responsible for Breast Cancer Metastasis in a Humanized Mouse Model with Molecular Imaging

PRINCIPAL INVESTIGATOR: Emily Powell

CONTRACTING ORGANIZATION: The University of Texas MD Anderson Cancer Center
Houston, TX 77030

REPORT DATE: November 2015

TYPE OF REPORT: Final Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE November 2015		2. REPORT TYPE Final Annual Summary		3. DATES COVERED 1Sep2011 - 31Aug2015	
4. TITLE AND SUBTITLE From Breast to Bone: Tracking Gene Expression Changes Responsible for Breast Cancer Metastasis in a Humanized Mouse Model with Molecular Imaging				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-11-1-0616	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Emily C. Powell, Ph.D. E-Mail:epowell1@mdanderson.org				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The University of Texas MD Anderson Cancer Center 1515 Holcombe Boulevard, Unit 1906 Houston, TX 77030-4009				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Breast cancer is the leading cause of cancer-related death in women worldwide, and metastasis is responsible for the majority of these deaths. Triple-negative breast cancer (TNBC) is an aggressively metastatic subtype exhibiting a disproportionate degree of <i>TP53</i> mutation compared with other breast cancer subtypes. The major goals of the studies described here were to (i) determine how p53 loss contributes to the metastatic potential of breast cancer cells; (ii) identify transcriptional changes that enhance breast cancer metastasis to the lungs; (iii) functionally characterize how these changes contribute to breast cancer metastasis; and (iv) validate our findings in human breast cancer samples. We used tumor models of metastatic TNBC that differed only in p53 status to characterize the effect of p53 loss in various stages of metastasis. We isolated metastatic tumor subpopulations from lungs and passaged them in vivo to enrich for metastasis. The isolation and transcriptional profiling of primary and metastatic subpopulations allowed the identification of genes that are differentially expressed between primary tumors and metastatic lesions. A subset of these genes is currently being functionally validated in gain-of-function in vivo screens. These studies are expected to stimulate the development of therapeutic strategies to predict and prevent metastasis.					
15. SUBJECT TERMS triple-negative breast cancer, metastasis, p53, BTG2, PDX Models					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	Unclassified	37	

Table of Contents

	<u>Page</u>
1. Introduction	4
2. Keywords	4
3. Overall Project Summary	5
4. Key Research Accomplishments	21
5. Conclusion	22
6. Publications, Abstracts, and Presentations	23
7. Inventions, Patents, and Licenses	24
8. Reportable Outcomes	24
9. Other Achievements	24
10. References	24
11. Appendices	25

1. INTRODUCTION:

Metastasis is responsible for the majority of deaths due to breast cancer. Recent data suggest that p53 may play a role in metastasis. Triple-negative (i.e., negative for estrogen receptor, progesterone receptor and HER2 gene amplification) breast cancer (TNBC) is an aggressively metastatic subtype that preferentially metastasizes to visceral organs including the lungs. Metastasis occurs through a sequence of steps involving the dissociation of cells from the primary tumor, local invasion and migration through the basement membrane and into the circulation, survival in the circulation, extravasation into distant organs, tumor dormancy, and finally tumor growth in the new organ [1]. Despite the urgent need to understand the mechanisms governing metastasis, few model systems recapitulate all stages of metastasis. To enable us to study multiple stages of metastasis, we use patient-derived xenograft (PDX) models of triple-negative breast cancer (TNBC) in our research. In several of these models, cells metastasize from humanized mammary fat pads to distant organs, including the lungs. The objective of the funded work was to identify proteins that enhance TNBC metastasis to the lungs and to determine their mechanism of action. Furthermore, we sought to investigate the role of p53 in metastasis.

2. KEYWORDS: triple-negative breast cancer, metastasis, p53, BTG2, PDX Models

3. OVERALL PROJECT SUMMARY:

Task 1. Establish human-in-mouse breast cancer models that enable non-invasive and repetitive imaging of bone and lung metastases (months 1-12).

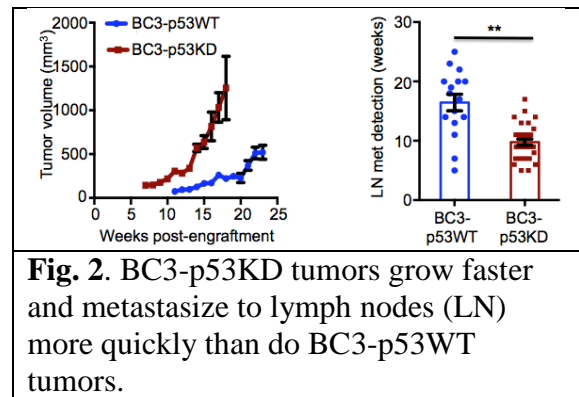
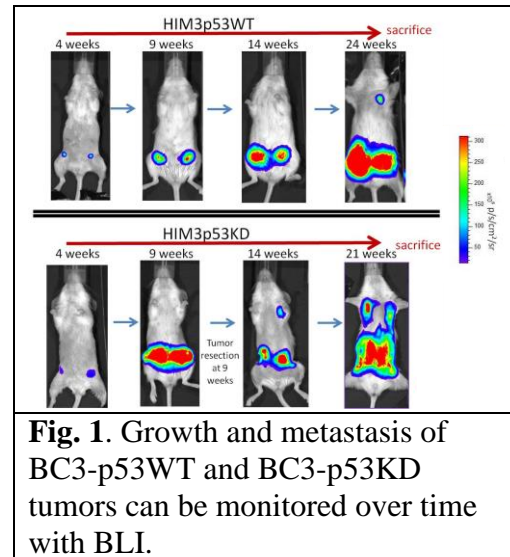
a. Transduce primary human tumor cell lines with lentivirus that encodes bioluminescent (click beetle red luciferase [CBR-luc]) and fluorescent (mCherry) markers for non-invasive bioluminescent imaging (BLI) (*month 1*).

b. Humanize the mammary glands of female NOD/SCID mice and xenograft human tumor cells (*month 1*).

c. Monitor primary human tumor cell metastasis using non-invasive BLI (*months 2-4*).

d. Purify RNA, lyse cells for whole protein analysis, mount tissues for immunohistochemical analysis, and serially transplant metastasized bone and lung tumors to enrich for cells that selectively home to the bones or lungs (*months 4-12*).

Metastatic human-in-mouse (HIM) models that stably express bioluminescent (CBR-luc) and fluorescent (mCherry) markers were generated. First, WU-BC3 tumor cells were infected with control retroviruses or retroviruses that encoded p53-specific small hairpin (sh)RNAs, giving rise to BC3-p53WT (wild-type) and BC3-p53KD (knockdown) cells, respectively [2]. BC3-p53WT and BC3-p53KD cells were then engineered to express both CBR-luc and mCherry by lentiviral transduction, which allows for non-invasive and repetitive monitoring of metastasis in the mice (Fig. 1).



One million transduced BC3-p53WT and BC3-p53KD cells were implanted into the fourth mammary fat pads of NOD/SCID mice, along with immortalized human stromal fibroblasts. BLI was used to monitor tumor growth and metastasis in anesthetized animals. As seen in Fig. 2, p53 loss significantly increased the growth rate of BC3 cells (left panel) and led to earlier detection of lymph node metastases (right panel). However, this does not necessarily mean that BC3-p53KD cells can escape the mammary tumor more quickly than BC3-p53WT cells can. We therefore sought to determine at what point circulating tumor cells (CTCs) from BC3-p53WT and BC3-p53KD tumors were first detected in mouse blood.

To assess the ability of tumor cells to enter into the bloodstream, CTCs present in the blood of tumor-bearing animals were quantified by flow cytometry at various times following tumor engraftment. The number of CTCs able to metastasize out of the mammary glands of tumor-bearing mice increased as a function of time, and greater numbers of CTCs were detected in mice harboring BC3-p53KD tumors

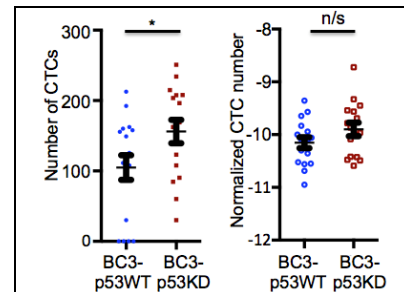


Fig. 3. BC3-p53KD tumor cells exhibit increased escape to blood circulation as a result of increased tumor growth.

than in those harboring BC3-p53WT tumors at each time point examined. CTCs were not detected in either

line earlier than 9 weeks after implantation. The total number of CTCs released over the 18-week period

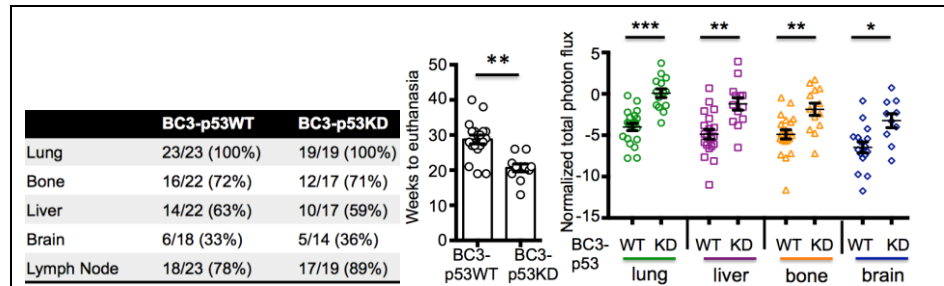


Fig. 4. BC3-p53KD tumors exhibit increased bioluminescence in multiple metastatic sites, but the frequency of metastasis is unchanged.

by mice harboring BC3-p53KD tumors was significantly higher than the total number released by mice harboring BC3-p53WT tumors (Fig. 3, left panel). However, when total CTC numbers were normalized to the photon flux of the corresponding mammary tumor (Fig. 3, right panel), the difference was not statistically significant. Thus, faster tumor growth induced by p53 loss was likely responsible for increases in CTC release from mammary tumors.

To quantify metastasis, animals were imaged with BLI just before euthanasia, and their lymph nodes, lungs, livers, bones, and brains were imaged ex vivo. Knockdown of p53 did not alter the frequency of metastasis to any organ (Fig. 4, left panel).

Because the mice were euthanized when they reached morbidity, the end points varied (Fig. 4, middle panel). To mathematically correct for this difference, the photon flux from each organ was normalized to

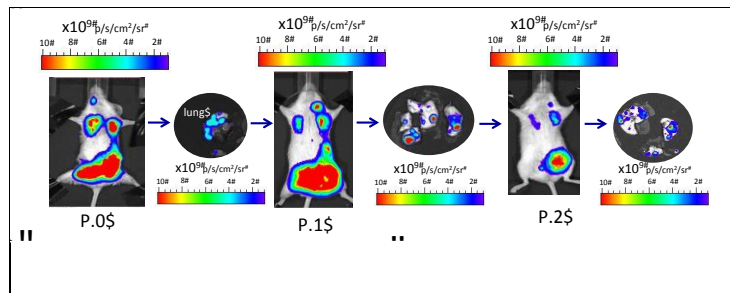


Fig. 5. In vivo passing of metastatic subpopulations to enrich the metastases for metastasis (passaging of lung metastases is shown).

the time of animal euthanasia. The bioluminescence of the lungs, livers, bones, and brains was significantly greater in mice bearing p53-deficient

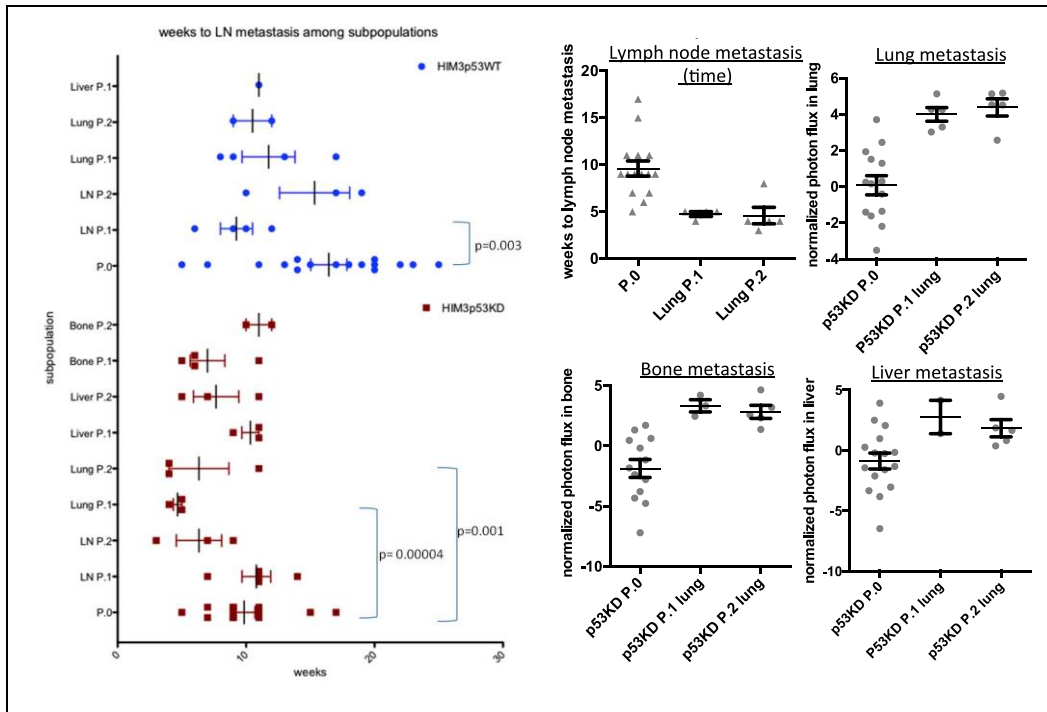


Fig. 6. Earliest detection of lymph node (LN) metastases in serially passaged metastatic tumor subpopulations (left panel). In vivo passaging of lung metastases enriches the tumor cells [←AU: OK?]for metastasis to multiple sites (right panels).[AU: I suggest moving the caption above the left panel to replace “weeks” as the x-axis label. Please capitalize the first letter of all x- and y-axis labels for consistency with other figures.]

tumors, even when the photon flux was normalized to the date of euthanasia (Fig. 4, right panel). Normalization to the photon flux of the mammary tumors was not possible because they were resected in a survival surgery prior to the study end point. This suggests that p53 silencing

enhanced the metastatic capacity of BC3 tumors. Statistical analysis was performed in collaboration with Dr. Kim-Anh Do at The University of Texas MD Anderson Cancer Center. Because the BC3-p53KD tumors metastasized more robustly, it was possible to harvest metastases from multiple sites before the primary tumors became prohibitively large and the animals required euthanasia. Thus, we focused on BC3-p53KD tumors to generate metastasis signatures. To enrich for the genes responsible for metastasis, the metastatic tumors were serially passaged in vivo (Fig. 5). After the mice were euthanized, metastatic lesions (denoted P.0 for passage 0) were visualized with BLI (Fig. 5), excised, and digested to organoids and single cells. Tumors growing in the mammary glands were also collected, snap-frozen in liquid nitrogen to preserve RNA integrity, and banked. Cells from metastatic lesions were counted using an automated cell counter with fluorescence microscopy (only the mCherry-tagged cells were counted). A maximum of 1×10^6 mCherry-positive tumor cells were then implanted into the fourth mammary fat pads of recipient mice along with human stromal fibroblasts. The mice were subjected to bi-weekly BLI to track metastasis, and at the time of euthanasia, all organs were subjected to BLI ex vivo. Metastatic lesions from these mice (denoted P.1) were digested to organoids, and a maximum of 1×10^6 mCherry-positive tumor cells were implanted into the fourth mammary fat pads of recipient NOD/SCID mice along with human stromal fibroblasts. The mice were subjected to bi-weekly BLI to track metastasis, and at the time of euthanasia, all organs were subjected to BLI ex vivo. Metastatic lesions and mammary tumors from these mice (denoted P.2) were then banked. *Passaging to P.2 was performed to enrich for lung metastases and to obtain sufficient tumor tissue for a comprehensive analysis by RNA sequencing.* In addition, tumor tissue was collected for Western blot and immunohistochemical analyses. We also collected tissue from additional metastatic sites; these samples will be used to validate our lung metastasis signature. An example of our enrichment procedure is shown in Figure 5 (passaging of lung metastases is shown for simplicity; metastases from several sites were serially passaged).

Fig. 6 (left panel) shows the time to appearance in the lymph nodes for each enriched metastatic subpopulation. Appearance in the lymph nodes indicated that the subpopulation could exit the mammary gland. The only BC3-p53KD metastatic subpopulations that appeared in the lymph nodes faster than the P.0 p53KD parental population were the P.1 and P.2 lung subpopulations ($p = 0.00004$ and $p = 0.001$, respectively, 2-tailed Student t -test). This finding

suggested that the metastatic subpopulations isolated from BC3-p53KD lungs were more aggressive than are their P.0 parental counterparts and that a more metastatic subpopulation had been enriched. We therefore focused on these BC3-p53KD lung metastatic subpopulations for transcriptional profiling. Enrichment for lung metastasis resulted in enrichment for metastasis to the lymph nodes, bones, and liver (Fig. 6, right panels). The bones and lungs from mice that harbored metastases were embedded in paraffin and mounted for immunohistochemical analysis to confirm the RNA sequencing data at the protein level. These experiments and the resulting data fulfilled Task 1.

Task 2: Determine the transcriptome of breast cancer cells that selectively home to the lungs and bones (months 12-18).

a. Perform expression profiling to determine changes in gene expression during metastasis (*months 12-15*).

b. Validate expression profiling results using quantitative reverse transcription polymerase chain reaction (qRT-PCR), Western blotting, and immunohistochemical analysis (*months 15-18*).

RNA sequencing was used instead of the originally described microarray platform for transcriptional profiling analysis. RNA sequencing is a superior platform for these studies because it provides an unbiased approach to identifying genes that are differentially expressed as tumors metastasize. Transcriptomes that are identified via microarray are biased toward the presence of probes on the chip, but RNA sequencing allows the identification of all human genes, including uncharacterized and novel genes that are known only by their locus in the genome. RNA sequencing is also more sensitive at detecting less abundant transcripts, has a better dynamic range for measuring transcript abundance, and allows the evaluation of allele-specific (single nucleotide polymorphism) expression and quantitation of splicing. Furthermore, the mouse and human genomes can be profiled simultaneously from the same sample, allowing full examination of the interaction between the tumor and its microenvironment at both the primary (mammary gland) and metastatic sites.

Determination of the transcriptome of p53-proficient and –deficient mammary tumors and functionalization of a lead candidate gene.

The sequencing of 4 BC3-p53KD P.0 mammary tumors and 3 BC3-p53WT P.0 mammary tumors has been completed. Differential expression analysis was performed on these mammary tumors, and a gene expression profile was generated using a cutoff value of $p \leq 0.05$ and \log_2 fold change ≥ 1.5 (Appendix 1). p53 silencing resulted in a 10.9-fold reduction in *TP53* expression in BC3-p53KD tumors relative to BC3-p53WT tumors. GeneGo pathway analysis revealed that the top enriched processes differentially affected by p53 silencing were those related to interactions with the surrounding microenvironment, including extracellular matrix (ECM) remodeling, connective tissue degradation, cell-matrix interactions, cell adhesion,

epithelial-mesenchymal transition (EMT), and apoptosis (Appendix 2). Alterations in components of these pathways are known to contribute to various stages of metastasis, including growth in primary and metastatic sites [3, 4]. Several genes that were previously demonstrated to be regulated by p53 appear on this signature; we speculated that these genes may be responsible for the increased metastasis

observed in BC3-p53KD tumors. The gene expression signature derived in this study was therefore reviewed to create a mini-signature of genes known to be in the p53 pathway (Appendix 3).

Because changes in gene expression that accompany disruptions in the p53 pathway are associated with poor prognosis in breast cancer patients [5-11], the dataset from The Cancer Genome Atlas (TCGA) [12] was used to determine if the expression (or lack thereof) of each gene in the mini-signature was predictive of overall survival for breast cancer patients ($n = 1022$) as well as for those patients with TNBC ($n = 119$). Datasets from San Diego ($n = 286$) [13], NKI ($n = 295$) [14], and Oxford cohorts ($n = 210$) [15] were used to correlate gene

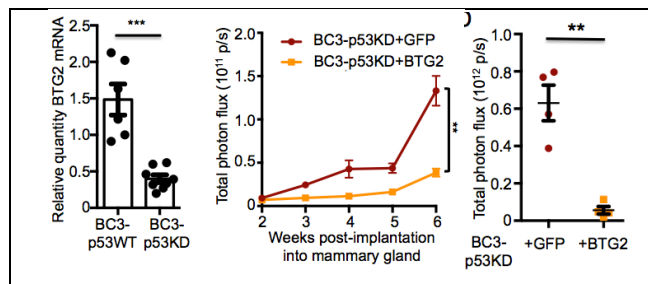


Fig. 7. BTG2 expression is decreased upon p53 silencing (left panel). Exogenous expression of BTG2 decreases tumor growth following implantation of tumor cells to mouse mammary glands (middle panel) or injection to the tail vein (right panel).

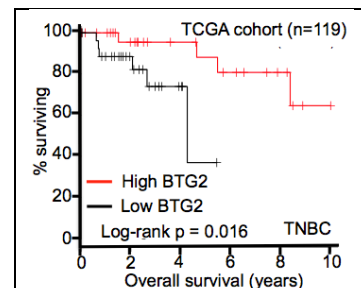


Fig. 8. Low BTG2 expression correlates with shorter overall survival rates among TNBC patients.

expression with metastasis-free survival (Appendix 3). BTG2 was chosen from the mini-gene signature for validation and further analysis.

BTG2 is a direct transcriptional target of p53, and quantitative RT-PCR validated that basal BTG2 expression was dependent on the p53 status of the originating cell lines (Fig. 7, left panel) [16-18]. We then ectopically expressed BTG2 in BC3-p53KD cells, using green fluorescent protein (GFP) expression as a control. Ectopic expression of BTG2 reduced the growth of BC3-p53KD mammary tumors in vivo relative to control cells (Fig. 7, middle panel). Reduced tumor growth was also observed in the lungs of mice after tail vein injection (Fig. 7, right panel). Taken together, these results suggest that BTG2 functions downstream of p53 to inhibit the growth of both primary breast tumors and metastatic lesions.

To explore the prognostic relevance of BTG2 expression, survival analyses were performed. Low BTG2 expression correlated with lower rates of overall survival in patients with TNBC (Fig. 8). Thus, BTG2 expression may be a prognostic marker for TNBC.

Determination of the transcriptome of breast cancer cells that selectively home to the lungs and bones.

P.2 lung metastases and corresponding mammary tumors from 3 BC3-p53KD and 2 HIM2p53WT tumors were also sequenced. Bone metastases from BC3-p53KD tumors were passaged through 2 mice, and the resulting P.2 tumors were sequenced. Because BC3-p53WT tumors do not metastasize well, it was not possible to enrich BC3-p53WT bone metastases by in vivo passaging; therefore, bone metastases were isolated from P.0 mice engrafted with BC3-p53WT tumors. In addition, bioluminescent cells from lung metastases of P.0 BC3-p53KD tumors were isolated. Non-adherent mammospheres were also cultured from lung metastases to enrich the metastases for putative stem-like or tumor-initiating cells. These populations were also submitted for RNA sequencing. The total number of replicates for each tumor or metastasis that has been sequenced is listed in Table 1. Please note that for each enriched (P.2) metastasis, the corresponding mammary glands from each mouse were also sequenced.

TABLE 1.

Subpopulation	Number of replicates*	
	BC3-p53WT	BC3-p53KD
P.0 mammary gland	3	4
P.2 lung metastases (corresponding mammary glands were also sequenced)	2	3
P.2 bone metastases (corresponding mammary glands were also sequenced)	0	2
P.0 lung metastases	0	2
P.0 bone metastases	3	1
Mammospheres from P.0 mammary gland tumors	4	2
Mammospheres from P.2 lung metastases	0	2
Mammospheres from P.0 bone metastases	1	1
TOTAL (including corresponding mammary glands)	15	22

*Each replicate is from a different mouse.

A bioinformatic pipeline was developed in collaboration with the laboratory of Dr. John Edwards at Washington University. This pipeline was validated and re-optimized in collaboration with the laboratory of Dr. Han Liang at MD Anderson. This pipeline maps transcript sequence reads to the mouse and human genomes to filter out mouse reads and determine the transcriptional profile of human epithelial tumors without contamination from mouse transcripts. The mouse reads are set aside for future analysis. The sequence reads are clustered according to replicates within a tissue type, indicating tight agreement among replicate samples and separation from other subpopulations (Fig. 9).

A statistical analysis of replicates with a false discovery rate < 0.25 was used to define the genes that were significantly differentially expressed between the P.2 lung metastases and P.0 parental mammary glands. This analysis yielded a list of genes that were up- or down-regulated in P.2 lung metastases relative to P.0 parental mammary glands. A “lung metastasis signature” was thus defined using a \log_2 fold change cutoff of 3x (Appendix 4, columns A-C). Note that a

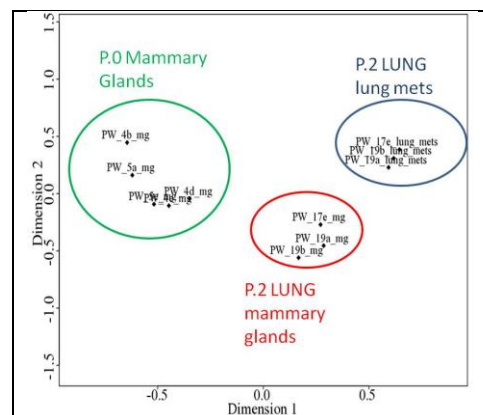


Fig. 9. Clustering of RNA sequencing reads indicates agreement of various populations.

\log_2 fold change cutoff of 3x, which corresponds to a fold change cutoff of 8x, is more stringent than the 1.4x-3.0x cutoff traditionally used in gene expression analysis. This more stringent cutoff was used to narrow the list of genes for these studies.

To control for the site of isolation, a gene expression signature was compiled for differentially expressed genes in the P.2 mammary gland. The genes that were dysregulated in both the P.2 lung and the P.2 mammary gland (Appendix 4, columns D-F) and in both P.2 lung metastases and mammospheres derived from lung metastases (Appendix 4, columns G-I) were identified. This was done because culturing breast tumor cells as mammospheres is believed to enrich for tumor-initiating cells. We speculated that these dysregulated genes might be required for seeding metastases. We hypothesized that because these mammospheres contain enriched populations of tumor-initiating cells, the enriched genes that were identified in the mammosphere RNA sequencing analysis are novel targets for metastasis prevention or are more effective at destroying metastatic cells while minimizing the risk of relapse. As shown in Appendix 4, several genes appeared on all three signatures. These genes may represent novel anti-lung metastasis targets that act at multiple steps of the metastatic cascade.

The p53KD lung metastasis signature was compared with the p53KD bone metastasis signature, which was derived in a similar manner. However, the RNA sequencing data derived from the bone metastases were not as robust because of a low representation of human cells relative to mouse cells in each bone metastasis sample. This low number was the result of low rates of bone metastasis and difficulties in achieving high RNA yields, possibly because releasing vascularized metastases from the bony matrix is technically difficult. As shown in Appendix 5, relatively few genes appeared on both the lung metastasis and bone metastasis signatures. This finding may reflect the low human:mouse representation ratio seen in bone metastasis; alternatively, it could reflect differences in the microenvironments of metastatic sites. The genes shown in Appendix 5 are of great interest because they may be non-site selective and therefore might serve as therapeutic targets of metastasis to multiple organs. It is worth noting that none of the genes that were dysregulated in lung metastases, lung metastasis mammospheres, or enriched P.2 mammary glands overlapped with the KD bone metastasis signature genes.

At the end of year 1, we attempted to enrich more P.2 bone metastasis populations (BC3-p53KD). Although we observed bone metastasis using our sensitive BLI techniques, bone metastasis was not robust enough to consistently enrich to P.2 using our in vivo passaging technique. Therefore, P.0 bone metastases were isolated for profiling by RNA sequencing (Table 2). None of the genes that overlapped between enriched bone metastases and enriched lung metastases appeared in the analysis for non-enriched (P.0) p53WT or p53KD bone metastases.

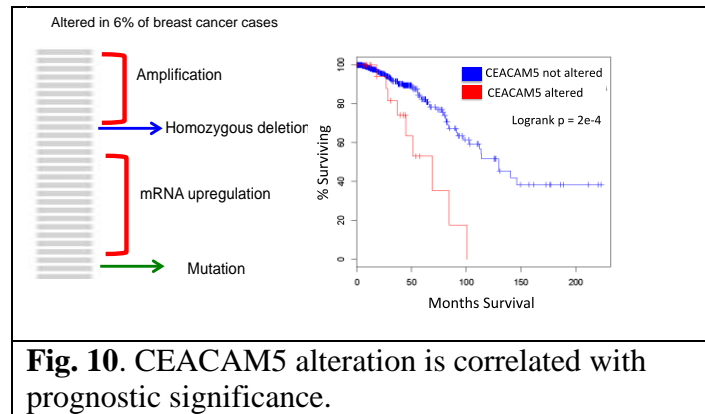


Fig. 10. CEACAM5 alteration is correlated with prognostic significance.

GeneGo pathway analysis of the BC3-p53KD lung metastasis signature revealed several dysregulated processes. Genes involved in epithelial-mesenchymal transition, extracellular matrix remodeling, cell adhesion, cytoskeletal remodeling, and blood coagulation were down-regulated in enriched lung metastases compared with in P.0 mammary glands. In addition to this automated analysis, we evaluated the signatures manually to identify the genes involved in common processes. Mucins and mucin-like proteins were dysregulated in several populations. For example, PARM1, a mucin-like, androgen-regulated gene that conveys cellular proliferation in prostate cancer, was up-regulated in both BC3-p53KD and BC3-p53WT lung metastases. PARM1, TFF1, MUC2, MUC6, MUC20, and MUC5AC were all up-regulated in BC3-p53KD lung metastases and corresponding P.2 mammary glands. MUC4 and MUC3B were also found to be dysregulated in P.2 BC3-p53KD bone metastases. MUC5AC was up-regulated in mammospheres derived from lung metastases. PARM1, MUC2, and MUC6 were dysregulated in BC3-p53WT lung metastases. These results suggest that mucins may play an important role in lung metastasis. Similarly, several common genes among these signatures clustered to other processes, including extracellular matrix remodeling and calcium homeostasis.

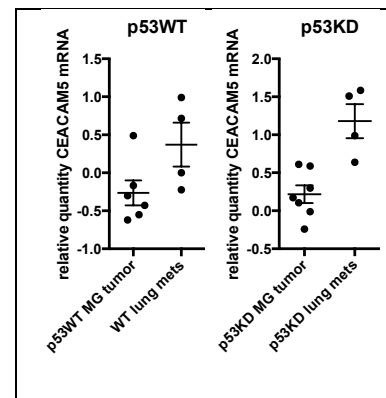


Fig. 11. CEACAM5 mRNA expression in BC3-p53WT (left) and BC3-p53KD (right) metastatic populations.

In silico analyses were performed to determine whether any of the genes on the lung metastasis signature had clinical relevance. Each gene on the signature was entered into the TCGA CBioPortal interface to determine whether an alteration (i.e., mutation, up-regulation, or down-regulation compared with normal tissue) of each gene was predictive of decreased survival. Alteration of CEACAM5 was predictive of decreased survival in breast cancer patients (Fig. 10). Six percent of the breast cancers evaluated exhibited an alteration in CEACAM5; the majority of these alterations were amplifications or mRNA up-regulation.

To confirm the expression of up-regulated genes at the mRNA level, qRT-PCR was performed on samples isolated from lung metastases and mammary tumors. CEACAM5 is shown as a

representative example (Fig. 11). qRT-PCR was also performed using primers that recognize KERA, PARM1, and ZP4. Data were normalized to GAPDH, and all primers used were specific for the human form of the gene. The direction and magnitude of regulation for each gene from RNA sequencing were recapitulated using qRT-PCR.

To confirm the expression of up-regulated genes in lung metastases at the protein level, immunohistochemical staining was performed. As shown in Figure 12, CEACAM5 and PARM1 were expressed in lung metastases of mice engrafted with BC3-p53KD tumors. The antibody that recognizes CK18 is a human-specific marker. Similar staining was performed to confirm the lung expression of ZP4, ARMC3, and Neu4. Task 2 was completed, and the lung metastasis gene signatures that were generated were evaluated for functional relevance in Task 4.

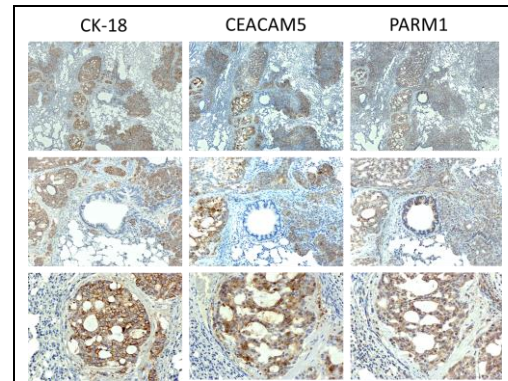


Fig. 12. Immunohistochemical staining shows that CEACAM5 and PARM1 are expressed in the lung metastases of mice engrafted with BC3-p53KD tumors.

TABLE 2

Mouse #	HIM model	P.0 mice	P.2 mice	P.0 metastases
4B	BC3-p53 WT	a,c		
4E	BC3-p53 WT	a,c,d	a,b,c,f	
4D	BC3-p53 WT	a,c,d	a,b,c	
21C	BC3-p53 WT	a,c		Lungs, bones
28A	BC3-p53 WT	e,f,h		
28B	BC3-p53 WT	e,f		
24A	BC3-p53 WT	e,f		
24B	BC3-p53 WT	e,f		
28E	BC3-p53 WT			Lungs, bones
21A	BC3-p53 WT			Lungs
21B	BC3-p53 WT			Bones
5A	BC3-p53KD	a,c		
5B	BC3-p53KD	a,c,d	a,b,c	Lungs
5E	BC3-p53KD	a,c		Lungs, liver
22A	BC3-p53KD	a,c,d	a,b,c,f	Lungs, bones
18D	BC3-p53KD	a,c		Lungs, liver
22B	BC3-p53KD	a,c		Liver, bones
5D	BC3-p53KD	c,d	a,b,c	
29B	BC3-p53KD	e,g,h		
29C	BC3-p53KD	f		
29E	BC3-p53KD	e,f		

a: Frozen mammary tumors for RNA sequencing and Western blot analyses; **b:** frozen lung metastases for RNA sequencing and Western blot analyses; **c:** tumor block for immunohistochemical analysis; **d:** successful passage of lung metastases to P.2; **e:** mammosphere cultures of mammary tumors; **f:** mammosphere cultures of lung metastases; **g:** mammosphere cultures of liver metastases; **h:** mammosphere cultures of bone metastases.

Task 4: Functionally characterize a subset of proteins that are encoded by up- or down-regulated genes in homing-enriched populations of tumor cells (months 18-36).

a. Assess the metastatic contribution of bone and lung metastasis drivers in HIM models by silencing gene targets (*months 18-36*).

b. Assess the metastatic contribution of bone or lung metastasis suppressors in HIM models by restoring the expression of gene targets (*months 18-36*).

As explained above, the most robust RNA sequencing data came from our lung metastasis models. Our imaging data also indicate that the lungs are primary metastasis sites in these models. Therefore, functional screens are being performed using our lung metastasis

signature and lung metastasis model. BC3-p53KD is used because it metastasizes more robustly than does BC3-p53WT. However, priority will be given to any genes from bone metastasis signatures that are able to drive metastasis to the lungs as well. In summer 2013, we relocated our research program from Washington University in St. Louis, Missouri, to MD Anderson Cancer Center in Houston, Texas. At this point we established a collaboration with Dr. Timothy Heffernan at MD Anderson for high-throughput functional screens. Dr. Heffernan's laboratory provided expression constructs for the top genes in our lung metastasis signature. This screening will allow the identification of *functional* metastasis genes identified in Task 2. In brief, BC3-p53KD is transduced with pooled lentivirus-encoding genes in the KD lung metastasis signature. Lentiviral libraries are generated that express the open reading frames (ORFs) of all genes that are upregulated in the lung metastasis signature. Cells are infected with a multiplicity of infection such that each cell receives at least one copy of the plasmid. These plasmids also encode GFP. The transduced population is then implanted into the fourth mammary glands of recipient-humanized NOD/SCID mice and allowed to metastasize. After the mice are euthanized, the metastases are isolated, and genomic sequencing is performed to evaluate the expression of the ORF. The presence of an ORF at a metastatic site suggests that the ORF can drive metastasis. This gain-of-function (GOF) screening provides an unbiased platform to functionalize genes on the lung metastasis signature.

Conditions were optimized for GOF screening of the BC3-p53KD lung metastasis signature, as defined in the statement of work (Task 4). The ongoing GOF screening uses ORFs that are up-regulated in lung metastases because these genes are potential therapeutic targets.

As a proof-of-concept, experiments were performed to determine whether known metastasis drivers increased the metastasis of BC3-p53KD cells

from the mammary tumor to the lymph nodes and lungs in vivo. To accomplish this, BC3-p53KD cells were transduced with lentiviruses expressing ORFs for HOXA1, ACP5, ID1, or GFP (control). The expression of HOXA1 and ACP5 decreased the time to metastasis to the

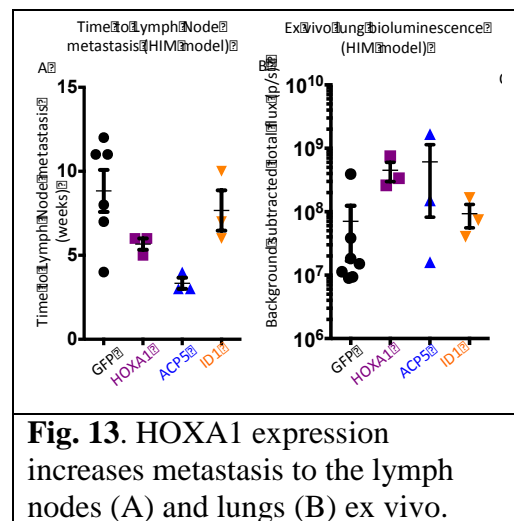


Fig. 13. HOXA1 expression increases metastasis to the lymph nodes (A) and lungs (B) ex vivo.

lymph nodes (Fig. 13, left panel), and HOXA1 increased the total bioluminescent signal in the lungs (Fig. 13, right panel) compared with cells expressing eGFP. This

demonstrates the feasibility of this assay and model system for identifying proteins that promote breast cancer metastasis.

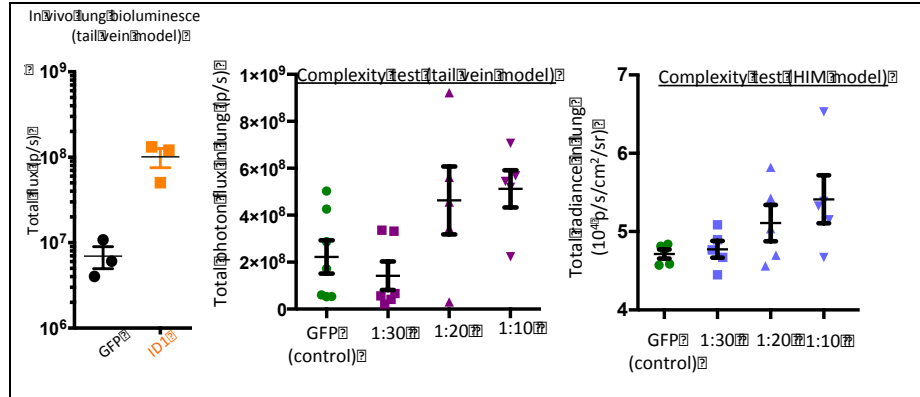


Fig. 14. Expression of ID1 in BC3-p53KD cells results in increased tumor outgrowth in the lungs after injection to the tail vein (in vivo lung bioluminescence, left panel). Dilution of ID1 with GFP control vector (ID1:GFP) shows that a true positive driver of metastasis can be observed with 20 or fewer ORFs per pool (tail vein model complexity test, middle panel; HIM model complexity test, right panel).

Prior to initiating high-throughput GOF screens in vivo, a complexity test was performed. The purpose of this test was to determine how many viruses to include in a single viral pool. ID1 was chosen as a positive control for this assay because ID1 expression in BC3-p53KD resulted in increased outgrowth in the lungs when cells were injected into the tail vein (Fig. 14, left panel).

To this end, lentiviruses encoding GFP (negative control), ID1 (positive control), or dilutions of ID1:GFP (1:10, 1:20, and 1:40) were generated. The bioluminescent signal in the lungs was increased when ID1 was expressed in ratios of 1:10 or 1:20 with GFP compared with the GFP control (Fig. 14, middle panel). This finding indicated that lentiviral pools

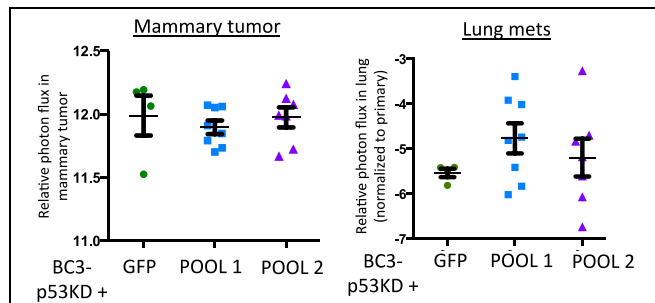


Fig. 15. Expression of candidate metastasis enhancers in a pooled format increases lung metastasis (left panel) without increasing mammary tumor growth (right panel).

should contain no more than 10-20 ORFs each for GOF screens. Because the screen is performed in orthotopic models, we elected to repeat this complexity test in HIM models. Results from this assay indicated that lentiviral pools should contain no more than 10-20 pools (Fig. 14, right panel), in keeping with findings from the tail vein complexity test. Thus, validation and optimization of our GOF screening platform has been completed.

High-throughput in vivo GOF screens are well under way. BC3-p53KD cells are transduced with lentivirus encoding a single ORF to create individual cell lines expressing a single ORF. Transduced cells are co-implanted to mammary fat pads of recipient mice along with mammary stromal fibroblasts. Pooling the lentiviral libraries (as opposed to analyzing one cDNA at a time) minimizes the number of animals needed for the screen and allows all up-regulated genes represented in the metastasis signature to be evaluated more rapidly and in an unbiased fashion. On the basis of

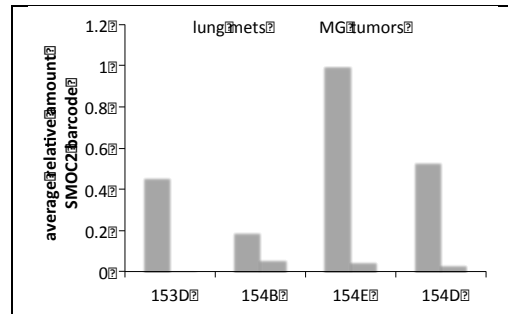


Fig. 16. The average relative amount of SMOC2 gDNA in lung metastases and mammary gland (MG) tumors is shown for 4 replicate mice (153D, 154B, 154E, and 154D).

complexity test data, pools consist of 13 ORFs (12 putative metastasis enhancers and 1 GFP as an internal control). Each ORF is associated with a unique nucleic acid “barcode” sequence and is identified by quantitative sequencing of lung metastases and mammary tumors. As a control, cells are transduced with lentivirus encoding eGFP and implanted to a separate cohort of mice. Metastasis is monitored using BLI (as described in Aim 1). Thirteen weeks after engraftment, mice are euthanized and tumor specimens are visualized with BLI (Fig. 16), harvested from all organs, and snap frozen in liquid nitrogen. We have chosen 13 weeks as an end point for this screen

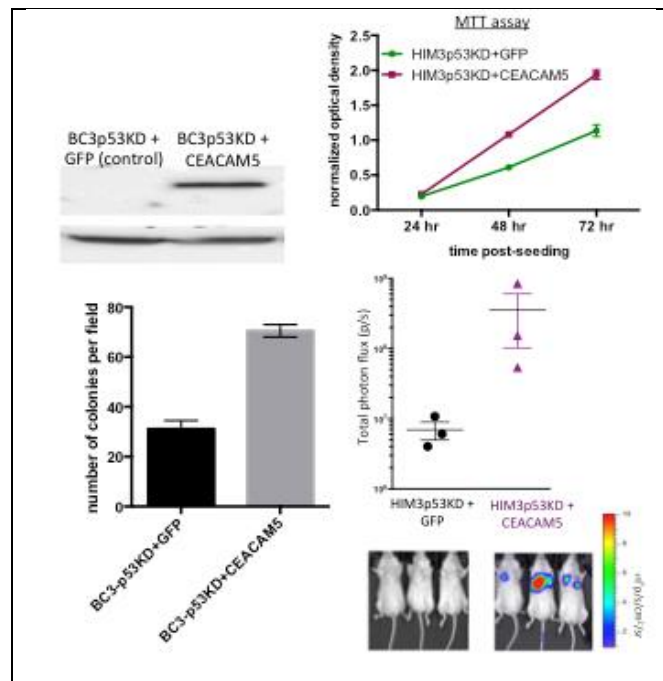


Fig. 17. Exogenous expression of CEACAM5 (upper left panel) can drive cell proliferation in vitro (upper right panel), colony formation in 3D (lower left panel), and the outgrowth of BC3-p53KD cells in the lungs (lower right panel).

because BC3-p53KD cells have not yet metastasized robustly to the lungs at this point. Genomic DNA (gDNA) is isolated and submitted for quantitative sequencing to assess enrichment for expressed ORFs relative to the total population retained in the mammary tumor. The preparation

of gDNA libraries is optimized, and sequencing analysis is performed in collaboration with the laboratory of Dr. Kenneth Scott at Baylor College of Medicine. From the first two pools of screening, we have already identified a potential functional enhancer of metastasis (Fig. 16). The barcode sequence associated with SMOC2 is enriched in lung metastases compared with mammary tumors and is thus a lead candidate to be a functional enhancer of metastasis. We will continue this high-throughput screening in vivo until the top 288 ORFs from the lung metastasis signature have been evaluated. ORFs that are significantly enriched in metastatic lesions compared with the primary mammary tumor will be confirmed as singletons in secondary in vivo screens. Genes that are validated as singletons in secondary screens will be prioritized for follow-up loss-of-function screening and for functional/clinical/pathological validation.

While optimizing the parameters for GOF screening, CEACAM5 was functionalized as a singleton in vitro and in vivo. ORFs that encode CEACAM5 or GFP (control) were expressed in BC3-p53KD cells (Fig. 17, upper left panel), and cell proliferation was evaluated in vitro in 2D (Fig. 17, upper right panel) and in a 3D matrix of collagen (Fig. 17 lower left panel). To characterize tumor growth in vivo, cells were injected into the tail veins of recipient mice. CEACAM5 expression resulted in an increase in bioluminescence in the lungs (Fig. 17, lower right panel), indicating that CEACAM5 expression in BC3-p53KD cells can promote the growth of breast cancer cells in the lungs. Experiments are currently under way using our HIM models to determine whether CEACAM5 can increase the rate of metastasis from the mammary glands to the lungs. We also plan to evaluate the mechanism by which CEACAM5 increases tumor growth and metastasis.

Attempts to establish HIM2 and HIM-Aft7 models that stably express CBR-luc and mCherry were not successful in previous years; however, this problem was revisited in the past year, and we were successful in labeling a new PDX line, PIM001-P, with CBR-luc and mCherry. PIM001-P was infected with highly concentrated lentiviruses encoding CBR-luc and mCherry. mCherry-positive cells were purified by fluorescence-activated cell sorting and implanted directly into mouse mammary glands. PIM001-P+CBRLuc will be used to validate the relevance of the genes validated as singletons in our secondary screen to BC3 metastasis.

4. KEY RESEARCH ACCOMPLISHMENTS:

- BC3-p53WT and BC3-p53KD tumors that expressed CBR-luc and mCherry were established in HIM models, and their metastasis was monitored using bi-weekly BLI.
- BC3-p53KD mammary tumors grow faster, escape to blood circulation earlier, and metastasize to the lungs, liver, bones, brain, and lymph nodes more robustly than do BC3-p53WT tumors.
- Silencing p53 in these models did not significantly alter the frequency of metastasis to any organ.
- A gene expression signature was derived that compared BC3-p53WT and BC3-p53KD mammary tumors.
- BTG2 was identified as a p53-regulated suppressor of TNBC tumor growth with potential utility as a prognostic indicator for TNBC patients.
- BC3-p53KD lung metastases were serially passaged through humanized mammary glands to enrich for metastasis. Lymph node metastasis was detected earlier in enriched (P.1 and P.2) populations than in parental (P.0) populations.
- Metastases from the lungs, bones, liver, and brain were passaged through the humanized mammary glands of NOD/SCID mice twice to enrich for genes that facilitated metastasis to each organ. Metastases and mammary tumors were snap-frozen in liquid nitrogen for RNA extraction.
- RNA sequencing analysis was performed on lung metastases and mammary tumors from mice bearing enriched metastatic tumors (P.2) and parental (non-metastatic) mammary tumors (P.0). Mammospheres were also derived from lung metastases to assess the contribution of tumor-initiating cells to metastasis; RNA sequencing analysis was performed on these populations.
- A bioinformatics pipeline was developed to distinguish mouse (stromal) transcripts from human (tumor) transcripts.
- Gene expression signatures generated from lung metastases, bone metastases, and mammospheres established from lung metastases were compiled. The genes in various signatures were compared to identify overlaps.

- Pilot experiments for functional screens showed that ORF expression in BC3-p53KD cells increased lung metastasis from mammary tumors, validating our pooled format functional screening platform in vivo.
- High-throughput screening is successfully under way and is leading to the identification of functional enhancers of metastasis.
- Preliminary data suggest that CEACAM5 is a driver of lung metastasis.

5. CONCLUSION:

The work completed during the funding period has led to the establishment of p53-proficient and p53-deficient PDX models of metastatic TNBC that express CBR-luc and mCherry markers; these models are valuable for elucidating the mechanisms by which TNBC escapes the breast and flourishes in distant organs. These models were used to evaluate the contribution of p53 loss to various stages of metastasis. We found that p53 loss led to increased tumor growth and the concurrent escape of tumor cells to blood circulation. We identified BTG2 as a p53-regulated suppressor of primary and metastatic tumors and a potential biomarker for TNBC.

These studies have also led to the identification of genes whose altered expression may enhance the metastasis of TNBC. We enriched for metastatic subpopulations that home to metastatic sites. Throughout this enrichment process, BLI in vivo showed that enriched metastatic subpopulations were more metastatic than the BC3-p53KD parental P.0 population, suggesting selection for a more aggressive subpopulation. Transcriptional profiling of these enriched subpopulations that home to the lungs provided a robust lung metastasis signature that was composed of genes that likely facilitate or abrogate lung metastasis. This signature was supplemented with transcriptional profiling on bone and liver metastases and mammospheres isolated from lung metastases. BC3-p53WT lung metastases also provided robust RNA sequencing data. Having so many varied populations has allowed us to search for commonly dysregulated genes in various aspects of metastasis. We performed a GeneGo MetaCore pathway analysis to identify the main processes dysregulated by these metastatic subpopulations and narrow our list of genes for functionalization studies. The top candidate genes were validated using qRT-PCR and immunohistochemical analysis to confirm the RNA sequencing findings.

The final stage of funding allowed us to optimize conditions for high-throughput functional screens in vivo and establish additional HIM models with bioluminescent and fluorescent markers. Two putative enhancers of metastasis, CEACAM5 and SMOC2, are currently being evaluated in mechanistic studies. Preliminary evidence suggests that CEACAM5 may drive TNBC metastasis; future work will focus on elucidating the mechanism by which CEACAM5 functions in vivo to control tumor growth and metastasis. We will also complete high-throughput screening to identify additional functional metastasis enhancers and evaluate their mechanisms of action.

6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS:

Powell E, Piwnica-Worms D, and Piwnica-Worms H. Contribution of p53 to Metastasis. *Cancer Discov.* 2014;4:405-14.

Powell E, Shao J, Yuan Y, Chen HS, Cai S, Echeverria GV, Mistry N, Decker KF, Schlosberg C, Do KA, Edwards JR, Liang H, Piwnica-Worms D, and Piwnica-Worms H. p53 deficiency linked to BTG2 loss enhances metastatic potential by promoting tumor growth in primary and metastatic sites in PDX models of triple-negative breast cancer. Submitted.

Invited lectures:

Baylor College of Medicine Breast Center Retreat. Title: p53 deficiency enhances metastatic potential of triple-negative breast cancer by promoting growth in primary and metastatic sites. 08/27/2015.

Poster presentations:

Hallmarks of Cancer Symposium. Title: p53 deficiency enhances metastatic potential of triple-negative breast cancer by promoting growth in primary and metastatic sites. 10/20/2015

Metastatic Breast Cancer Conference. Title: p53 deficiency enhances metastatic potential of triple-negative breast cancer by promoting growth in primary and metastatic sites. 09/18/2015

Gordon Research Conference: Mammary Gland Biology. Title: p53 deficiency enhances metastatic potential of triple-negative breast cancer by promoting growth in primary and metastatic sites. 06/09/2015

7. INVENTIONS, PATENTS AND LICENSES: Nothing to report

8. REPORTABLE OUTCOMES: Nothing to report

9. OTHER ACHIEVEMENTS: The funding period has allowed us to successfully establish models of human TNBC metastasis that can be tracked longitudinally in vivo with BLI. BC3-p53WT and BC3-p53KD are isogenic PDX lines that express click beetle red luciferase (CBR-luc) and mCherry. These lines metastasize to physiologically relevant sites in NOD/SCID mice including the lungs, liver, bones, brain, and lymph nodes.

10. REFERENCES:

1. Powell E, Piwnica-Worms D, Piwnica-Worms H. Contribution of p53 to Metastasis. *Cancer discovery*. 2014;4:405-14.
2. Ma CX, Cai S, Li S, Ryan CE, Guo Z, Schaiff WT, Lin L, Hoog J, Goiffon RJ, Prat A *et al*. Targeting Chk1 in p53-deficient triple-negative breast cancer is therapeutically beneficial in human-in-mouse tumor models. *J Clin Invest*. 2012;122:1541-52.
3. Schedin P, Keely PJ. Mammary gland ECM remodeling, stiffness, and mechanosignaling in normal development and tumor progression. *Cold Spring Harb Perspect Biol*. 2011;3:a003228.
4. Bourboulia D, Stetler-Stevenson WG. Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion. *Semin Cancer Biol*. 2010;20:161-8.
5. Borresen AL, Andersen TI, Eyfjord JE, Cornelis RS, Thorlacius S, Borg A, Johansson U, Theillet C, Scherneck S, Hartman S *et al*. TP53 mutations and breast cancer prognosis: particularly poor survival rates for cases with mutations in the zinc-binding domains. *Genes Chromosomes Cancer*. 1995;14:71-5.
6. Bergh J, Norberg T, Sjogren S, Lindgren A, Holmberg L. Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med*. 1995;1:1029-34.
7. Pharoah PD, Day NE, Caldas C. Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer*. 1999;80:1968-73.
8. Troester MA, Herschkowitz JI, Oh DS, He X, Hoadley KA, Barbier CS, Perou CM. Gene expression patterns associated with p53 status in breast cancer. *BMC Cancer*. 2006;6:276.
9. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET *et al*. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A*. 2005;102:13550-5.
10. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98:10869-74.

11. Langerod A, Zhao H, Borgan O, Nesland JM, Bukholm IR, Ikdahl T, Karesen R, Borresen-Dale AL, Jeffrey SS. TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Res.* 2007;9:R30.
12. Network. CGA. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012;490:61-70.
13. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet.* 2005;365:671-9.
14. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* 2002;347:1999-2009.
15. Buffa FM, Camps C, Winchester L, Snell CE, Gee HE, Sheldon H, Taylor M, Harris AL, Ragoussis J. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. *Cancer Res.* 2011;71:5635-45.
16. Boiko AD, Porteous S, Razorenova OV, Krivokrysenko VI, Williams BR, Gudkov AV. A systematic search for downstream mediators of tumor suppressor function of p53 reveals a major role of BTG2 in suppression of Ras-induced transformation. *Genes Dev.* 2006;20:236-52.
17. Matsuda S, Rouault J, Magaud J, Berthet C. In search of a function for the TIS21/PC3/BTG1/TOB family. *FEBS Lett.* 2001;497:67-72.
18. Rouault JP, Falette N, Guehenneux F, Guillot C, Rimokh R, Wang Q, Berthet C, Moyret-Lalle C, Savatier P, Pain B *et al.* Identification of BTG2, an antiproliferative p53-dependent component of the DNA damage cellular response pathway. *Nature genetics.* 1996;14:482-6.

11. APPENDICES:

Professional development activities: The Department of Cancer Biology and the community at MD Anderson Cancer Center have provided me with exceptional training opportunities. I have established strong collaborations with world-class experts in the fields of breast cancer and metastasis. I participate in a weekly departmental journal club and research seminar, the weekly lab meetings organized by Dr. Piwnica-Worms's group, and a multi-group lab meeting organized by Dr. Jeffrey Rosen at Baylor College of Medicine. I also meet one-on-one with Dr. Piwnica-Worms and our faculty collaborators on an as-needed basis as we collect data. In addition, I attend and present my ongoing work at professional scientific conferences, where I engage in discussions with the broader research community. These conferences also provide professional networking opportunities. These conferences include, but are not limited to, the San Antonio Breast Cancer Symposium, the recurring Gordon Conference on Mammary Gland Biology, and relevant AACR conferences such as the upcoming "Tumor Metastasis."

Development of mentoring and leadership skills: MD Anderson, its affiliated Postdoctoral Association, and the Association for Women in Science (AWIS) offer courses and programs including workshops on grant writing, communication, leadership, and mentoring skills. I take

advantage of these programs to improve the skills that I will need as I transition to independence. I am also the sole active research mentor to an undergraduate researcher, Hector Picon, who spends 9 hours per week in the Piwnica-Worms laboratory.

COLLABORATIVE AWARDS: Nothing to report

QUAD CHARTS: Nothing to report

MARKING OF PROPRIETARY INFORMATION: Nothing to report

Table S1. RNA-Seq signature of genes that are de-regulated upon p53 loss in BC3 mammary tumors

BC3-p53WT vs. BC3-p53KD mammary tumor gene expression signature				
<u>Gene Symbol</u>	<u>log2 Fold Change</u>	<u>actual fold change</u>	<u>pvalue (Wald test p-value)</u>	<u>expression increased or decreased when p53 is silenced</u>
TP53	3.4	10.9	7.84E-30	decreased
SGCZ	3.3	9.9	6.36E-07	decreased
NARR	3.2	9.5	6.33E-07	decreased
INPP5D	3.0	7.7	3.72E-13	decreased
EDA2R	2.8	7.2	1.23E-09	decreased
MIR3655	2.7	6.4	1.73E-05	decreased
RAD51L3-RFFL	2.5	5.5	2.29E-04	decreased
LINC01021	2.3	5.0	7.07E-05	decreased
SLC22A11	2.1	4.3	4.63E-05	decreased
CA9	2.1	4.2	8.08E-06	decreased
ABCC6P1	2.0	4.0	3.08E-03	decreased
TREML3P	2.0	4.0	1.05E-04	decreased
PPP1R14D	2.0	4.0	4.05E-04	decreased
GABRG1	2.0	3.9	3.80E-03	decreased
ADAM21	1.9	3.8	2.55E-03	decreased
C19orf83	1.9	3.7	1.00E-04	decreased
ANP32D	1.9	3.7	5.00E-03	decreased
SPATA18	1.8	3.6	4.52E-04	decreased
GDF15	1.8	3.5	2.30E-08	decreased
CCDC178	1.8	3.4	8.87E-03	decreased
C18orf56	1.8	3.4	1.57E-04	decreased
VWCE	1.8	3.4	4.55E-03	decreased
CECR7	1.8	3.4	1.83E-04	decreased
TMEM82	1.7	3.4	2.32E-03	decreased
SLC51B	1.7	3.3	3.44E-04	decreased
APOC1	1.7	3.2	7.64E-04	decreased
PSG3	1.7	3.2	9.60E-05	decreased
ARX	1.7	3.1	6.23E-03	decreased
ZNF141	1.6	3.1	1.15E-02	decreased
HABP2	1.6	3.1	1.08E-02	decreased
CD163	1.6	3.1	1.70E-02	decreased
CREB3L1	1.6	3.0	8.32E-03	decreased
ADORA1	1.6	3.0	3.38E-04	decreased

VSIG10L	1.6	3.0	5.63E-04	decreased
FAM27A	1.6	3.0	2.87E-03	decreased
NDUFA4L2	1.5	2.9	2.54E-03	decreased
RPS14P3	1.5	2.9	6.64E-03	decreased
WNT10B	1.5	2.9	2.39E-04	decreased
EIF4EBP3	1.5	2.9	6.72E-03	decreased
TNFRSF1B	1.5	2.9	2.33E-03	decreased
BTG2	1.5	2.9	1.92E-05	decreased
NME1-NME2	1.5	2.8	2.51E-02	decreased
ACTA2	1.5	2.8	3.00E-04	decreased
CACNA2D4	1.5	2.8	2.48E-03	decreased
PRAP1	1.5	2.8	5.81E-03	decreased
ASGR2	1.5	2.8	7.71E-03	decreased
PALD1	1.5	2.8	2.41E-03	decreased
BATF2	1.5	2.8	8.63E-03	decreased
LINC00261	1.5	2.7	1.59E-02	decreased
LOC102467081	1.5	2.7	2.73E-02	decreased
RPL22L1	1.5	2.7	1.00E-02	decreased
DSG3	3.2	9.0	3.63E-10	increased
ISPD	2.7	6.7	3.41E-05	increased
DDX11L9	2.7	6.5	7.16E-05	increased
ITGB8	2.6	5.9	2.53E-06	increased
EDIL3	2.6	5.9	3.05E-05	increased
PKP1	2.6	5.9	1.07E-04	increased
HOXC11	2.4	5.1	1.88E-04	increased
KLF8	2.3	5.1	2.26E-05	increased
BHLHE41	2.3	5.0	1.96E-04	increased
EPB41L3	2.3	4.9	2.15E-04	increased
CLIC6	2.2	4.5	1.24E-03	increased
MMP2	2.1	4.4	1.08E-04	increased
NTS	2.1	4.4	1.59E-03	increased
MB21D1	2.1	4.4	1.72E-03	increased
LRRTM3	2.1	4.2	2.03E-03	increased
PRR16	2.1	4.2	1.99E-03	increased
BLACAT1	2.0	4.1	1.11E-04	increased
NPM2	2.0	4.1	2.41E-03	increased
PCDH1	2.0	4.0	5.26E-05	increased
C10orf10	2.0	4.0	1.77E-04	increased
MFAP5	2.0	4.0	3.54E-04	increased
PTX3	2.0	3.9	3.92E-03	increased
PCBP3	1.9	3.8	2.04E-03	increased

C15orf56	1.9	3.8	4.11E-03	increased
TM4SF18	1.9	3.8	4.79E-03	increased
KRT6A	1.9	3.8	1.74E-03	increased
SERPINA3	1.9	3.7	4.88E-04	increased
ACTL8	1.9	3.7	1.86E-03	increased
CCL28	1.9	3.6	6.01E-04	increased
MAP3K7CL	1.8	3.6	1.74E-03	increased
FAP	1.8	3.6	3.35E-03	increased
ARID5B	1.8	3.6	1.83E-03	increased
SMIM10	1.8	3.5	7.20E-03	increased
SEMA3A	1.8	3.5	4.48E-04	increased
LOX	1.8	3.5	2.60E-06	increased
STL	1.8	3.5	5.99E-03	increased
LINGO2	1.8	3.4	5.87E-03	increased
RPS6KA2	1.8	3.4	1.38E-03	increased
GAS6-AS2	1.8	3.4	4.20E-03	increased
ACTR3C	1.8	3.4	8.44E-03	increased
TSPY26P	1.8	3.4	7.32E-03	increased
LUM	1.7	3.4	1.24E-03	increased
ZNF608	1.7	3.3	1.99E-03	increased
GJA1	1.7	3.3	3.88E-03	increased
MMP19	1.7	3.3	8.96E-03	increased
EPHA4	1.7	3.3	2.35E-03	increased
SPOCK3	1.7	3.3	9.60E-03	increased
EPHX4	1.7	3.3	3.01E-03	increased
CLIC5	1.7	3.3	2.78E-03	increased
GPR110	1.7	3.2	8.80E-03	increased
HOTAIR	1.7	3.2	1.27E-02	increased
MAOB	1.7	3.2	3.53E-03	increased
MMP7	1.7	3.2	1.33E-03	increased
PRKACB	1.7	3.2	6.96E-05	increased
RIMS1	1.7	3.2	1.34E-02	increased
TMPRSS4	1.7	3.2	1.20E-03	increased
KRT14	1.7	3.2	9.22E-03	increased
TNFSF14	1.7	3.2	1.11E-02	increased
NLRP3	1.7	3.2	9.49E-03	increased
TMEFF2	1.7	3.1	1.47E-02	increased
EDAR	1.6	3.1	1.86E-03	increased
CCDC85A	1.6	3.1	2.34E-04	increased
DUSP27	1.6	3.1	1.50E-02	increased
LOC340113	1.6	3.1	1.56E-02	increased

LOC100379224	1.6	3.1	4.11E-03	increased
FHOD3	1.6	3.1	1.45E-03	increased
IGFBP7	1.6	3.1	4.47E-03	increased
POU4F3	1.6	3.0	1.83E-02	increased
BMP6	1.6	3.0	7.39E-04	increased
SP7	1.6	3.0	1.59E-02	increased
RFTN2	1.6	3.0	1.96E-02	increased
KIAA1644	1.6	3.0	1.81E-02	increased
PCDH7	1.6	3.0	6.14E-03	increased
PELI1	1.6	3.0	2.62E-04	increased
GRIK1	1.6	3.0	1.88E-02	increased
MMP16	1.6	3.0	1.92E-02	increased
MOV10L1	1.6	3.0	1.74E-02	increased
HMCN1	1.6	3.0	1.92E-02	increased
KCNA7	1.6	2.9	1.11E-02	increased
MAML2	1.6	2.9	3.37E-04	increased
ZNF57	1.5	2.9	1.19E-02	increased
SOX8	1.5	2.9	1.79E-02	increased
LINC00163	1.5	2.9	1.25E-02	increased
THBS1	1.5	2.9	4.00E-03	increased
LOC100129027	1.5	2.9	1.71E-02	increased
MUC16	1.5	2.9	1.64E-02	increased
GRIN3B	1.5	2.8	2.77E-03	increased
GHRLOS	1.5	2.8	2.55E-02	increased
ROBO1	1.5	2.8	1.99E-02	increased
C4orf26	1.5	2.8	1.74E-02	increased
ZACN	1.5	2.8	2.65E-02	increased
FGB	1.5	2.8	1.66E-02	increased
RNF165	1.5	2.8	2.89E-02	increased
SLC16A7	1.5	2.8	5.92E-03	increased
CCL24	1.5	2.8	1.59E-02	increased
DNAH10	1.5	2.8	1.53E-02	increased
RUNX2	1.5	2.8	2.79E-02	increased
SLITRK6	1.5	2.8	2.80E-02	increased
IRAK3	1.5	2.8	3.02E-02	increased
PITPNM3	1.5	2.8	9.39E-03	increased
RYR2	1.5	2.8	2.97E-02	increased
B3GNT6	1.5	2.8	1.61E-02	increased
PAK6	1.5	2.8	2.08E-03	increased
PLCL1	1.5	2.8	2.46E-02	increased
JAM2	1.5	2.8	2.65E-02	increased

SH3TC2	1.5	2.8	9.77E-03	increased
UBOX5-AS1	1.5	2.7	2.94E-02	increased
SOX21-AS1	1.5	2.7	3.17E-02	increased
HR	1.5	2.7	9.98E-03	increased

Table S2. Pathway analysis of RNA-Seq expression signature consisting of genes de-regulated upon p53 knockdown

#	Networks	Total	p-value	FDR	# genes in data	Genes Deregulated
1	Proteolysis_ECM remodeling	85	9.426 E-06	5.314 E-04	7	MMP-7, MMP-19, MMP-16, SPOCK3, SERPINA3, MMP-2, Lumican
2	Proteolysis_Connective tissue degradation	119	9.489 E-06	5.314 E-04	8	MMP-7, MMP-19, MMP-16, SPOCK3, SERPINA3, MMP-2, Lumican, HABP2
3	Cell adhesion_Cell-matrix interactions	211	9.669 E-05	3.610 E-03	9	MMP-7, MMP-19, GJA1, MMP-16, MMP-2, THBS1, Lumican, HABP2, SGCZ
4	Development_Neurogenesis_Synaptogenesis	180	1.059 E-03	2.965 E-02	5	RIMS1, GRIK1, THBS1, WNT10B, ACTA2
5	Apoptosis_Anti-apoptosis mediated by external signals via NF-κB	111	3.040 E-03	6.413 E-02	5	PRKACB, RPS6KA2, p53, TNFRSF1B, ADORA1
6	Development_EMT_Regulation of epithelial-to-mesenchymal transition	225	3.774 E-03	6.413 E-02	6	KRT14, LOX, RPS6KA2, MMP-2, WNT10B, ACTA2
7	Development_Neurogenesis_Axonal guidance	230	4.259 E-03	6.413 E-02	7	EPHA4, CCL28, ROBO1, SEMA3A, PRKACB, ACTA2
8	Cell adhesion_Attractive and repulsive receptors	175	4.581 E-03	6.413 E-02	6	EPHA4, ROBO1, SEMA3A, PRKACB, ACTA2
9	Cell adhesion_Cadherins	180	5.252 E-03	6.512 E-02	6	PKP1, DSG3, PCDH7, ACTA2, WNT10B
10	Cytoskeleton_Intermediate filaments	81	5.814 E-03	6.512 E-02	4	PKP1, KRT14, KRT6A, ACTA2

green text indicates up-regulation in response to p53 knockdown

red text indicates down-regulation in response to p53 knockdown

Table S3. Mini-signature of genes that interact with the p53 pathway and associated p-values.

p53 associated genes	TCGA cohort (all breast cancer) Overall Survival (n=1022)	TCGA cohort (TNBC) overall survival (n=119)	San Diego cohort (metastasis free survival) (n=286)	NKI cohort (metastasis free survival) (n=295)	Oxford cohort (metastasis free survival) (n=210)
GDF15	0.618704	0.872751	0.512639	NA	NA
SPATA18	0.931971	0.909813	NA	NA	NA
FAM27A	0.795673	0.738823	NA	NA	NA
BTG2	0.000866	0.125703	0.181338	0.000003	0.000168
CA9	0.429666	0.538078	0.318494	0.001134	0.000787
TP53	0.406990	0.717982	0.892629	0.902928	0.748604
EDA2R	0.003220	0.493069	0.517135	NA	NA
NME1	0.394788	0.386684	0.043214	0.124441	0.020291
NME2	0.017816	0.401633	NA	0.407861	0.164076
TNFRSF1B	0.030622	0.737269	0.014245	0.967189	0.655479
ACTA2	0.128326	0.584963	0.010685	0.759370	0.633141
PRAP1	0.179881	0.469873	NA	NA	NA
DSG3	0.015148	0.281834	0.661268	0.404740	0.734274
EDIL3	0.019433	0.253810	0.138877	0.491648	0.323478
BHLHE41	0.192112	0.463182	0.868209	NA	NA
MMP2	0.744432	0.927997	0.779695	0.607070	0.558907
PCDH7	0.527964	0.473452	0.718918	0.506622	0.990493
THBS1	0.027219	0.234703	0.260597	0.413557	0.092741
PTX3	0.237603	0.747768	0.089691	0.659664	0.746536
SEMA3A	0.472948	0.127119	0.057343	0.955063	0.500024
LUM	0.577808	0.320572	0.436260	0.060627	0.726615
EPHA4	0.527416	0.561032	0.682572	0.121261	0.144562
KRT14	0.002357	0.723528	0.387805	0.700514	0.561103
TNFSF14	0.216639	0.738366	0.820042	0.708935	0.602292
NLRP3	0.276127	0.994360	0.876382	NA	NA
IGFBP7	0.098384	0.903900	0.071742	0.828768	0.605262
SP7	0.071958	0.034029	NA	NA	0.451514
MMP16	0.275221	0.149805	0.475863	0.439643	0.277535
ZACN	0.113341	0.285311	NA	NA	NA
RUNX2	0.190225	0.257235	0.362197	0.901616	0.616612
PAK6	0.046595	0.367752	0.684316	0.669276	0.328800

*Red shading indicates significant p-value

Appendix 4										
A	B	C		D	E	F		G	H	I
KD P2 lung vs KD P0 MG (original)	logFC in P2 lung	direction of regulation		KD P2 lung and KD P2 MG	logFC in P2 MG	direction of regulation		KD P2 lung, KD P2 MG, and KD P2 lung mammospheres	logFC in mammospheres	up/ down regulated
KERA	12.3	up		KERA	12.5	up			13.6	up
ARMC3	11.1	up		NEU4	5.8	up		SMOC2	7.8	up
ZP4	9.7	up		CYP2B6	4.1	up		RAB25	9.0	up
FAM101A	7.9	up		ARMC3	11.0	up		DPCR1	6.2	up
RYR2	7.1	up		COL9A3	3.9	up		TFF1	7.1	up
C21orf34	7.0	up		WDR72	4.5	up		SLC44A5	10.5	up
KRT20	6.8	up		SCG2	4.9	up		CEACAM5	5.7	up
NEU4	6.4	up		CCDC148	4.3	up		TSPAN7	5.1	up
FAM155A	6.3	up		ARL11	3.3	up		NAALADL2	4.8	up
ISX	6.3	up		GABRB3	3.8	up		IRX4	4.7	up
SCG3	6.0	up		SCG3	7.2	up		PRAP1	4.7	up
MUC2	5.7	up		FAM155A	6.0	up		CALCR	4.5	up
CYP2B6	5.6	up		CBFA2T3	4.2	up		KIAA0319	4.6	up
LGALS9C	5.5	up		RYR2	6.5	up		UPK1B	4.5	up
ST6GAL2	5.4	up		GPR133	4.3	up		GLYATL1	4.8	up
PARM1	5.1	up		MDK	3.4	up		NEU4	5.3	up
GPR133	5.1	up		FAM101A	6.9	up		ALPI	4.7	up
CDR1	5.0	up		SRGN	1.9	up		TFF3	4.0	up
WDR72	4.9	up		SLC44A5	2.6	up		MUC5AC	3.9	up
COL9A3	4.9	up		SLC1A7	3.2	up		ERN2	4.1	up
SCG2	4.9	up		SSTR5	4.0	up		TRPM2	3.8	up
SSTR5	4.7	up		ZP4	10.2	up		HUNK	3.8	up
WNT3A	4.7	up		C8G	2.3	up		CEACAM1	3.8	up
F7	4.6	up		IL2RG	3.1	up		SRMS	4.7	up
CCDC148	4.6	up		MUC20	2.5	up		MYO7B	3.8	up
MYBPC1	4.6	up		SLC17A4	2.4	up		CCDC148	4.5	up
SESN3	4.6	up		SYNGR4	2.2	up		ZP4	6.3	up
ARL11	4.5	up		SMOC2	2.2	up		SCG2	3.3	up
CBFA2T3	4.4	up		CPS1	1.5	up		STIM2	2.7	up
F2RL2	4.4	up		CREB3L1	2.1	up		FN1	4.4	down
MUC6	4.3	up		GAL3ST2	2.2	up		PAPPA	4.1	down
REG4	4.2	up		FABP1	3.5	up		TACSTD2	6.8	down
SCN8A	4.2	up		MUC5AC	2.1	up		KLF8	5.3	down
SLC1A7	4.2	up		TFF1	2.5	up		NRCAM	3.3	down
TTYH1	4.1	up		ST6GAL2	5.6	up		FGG	4.4	down
GPR114	4.1	up		IRX4	3.0	up		MMP14	7.0	down
MDK	4.0	up		SCN8A	3.4	up		FLJ22662	5.7	down
GABRB3	4.0	up		GUCY2C	2.7	up		CES8	4.9	down
IL2RG	4.0	up		F7	3.7	up		SYT12	3.6	down
PPP1R1B	3.9	up		TMEM82	1.4	up		DTX3	6.1	down
GPA33	3.9	up		TSPAN7	2.1	up		CD24	4.7	down
ZNF483	3.8	up		ALPI	3.3	up		WISP2	3.6	down
SFMBT2	3.8	up		SESN3	4.3	up				
MUC20	3.8	up		C5orf38	3.3	up				
GUCY2C	3.8	up		PARM1	4.4	up				
FXYD6	3.8	up		GPR114	3.0	up				
SLC22A11	3.8	up		PRAP1	2.3	up				
C8G	3.8	up		SYT5	3.2	up				
APCDD1	3.7	up		HOXC6	1.4	up				
SLC10A4	3.7	up		MUC6	3.2	up				
MUC17	3.7	up		ALX1	2.5	up				
SLC44A5	3.7	up		ALDOB	3.4	up				
RAB25	3.7	up		DPEP1	2.5	up				
ALPI	3.7	up		EBF3	2.4	up				
NEB	3.7	up		GLYATL1	2.1	up				
MPZ	3.7	up		CEACAM5	1.7	up				
LRRC19	3.7	up		C20orf151	2.3	up				
EGF	3.6	up		CALCR	2.5	up				
HTR2B	3.6	up		ZNF703	2.2	up				
FABP1	3.6	up		CACNA2D4	2.8	up				
C11orf53	3.5	up		EPHB3	2.0	up				
TSPAN7	3.5	up		RAB25	3.3	up				

SRGN	3.5	up	TAF3	1.9	up				
TFF1	3.4	up	DMRTA2	1.8	up				
SGK196	3.4	up	CLDN15	1.6	up				
SMOC2	3.4	up	HSD11B2	1.5	up				
FGF19	3.3	up	ERN2	1.2	up				
ZNF460	3.3	up	RAB40B	1.7	up				
SYNGR4	3.3	up	ESPN	1.7	up				
CREB3L1	3.3	up	MUC4	1.7	up				
SLC17A4	3.3	up	RNASE1	2.5	up				
WNT5A	3.3	up	CEACAM1	1.3	up				
IKZF1	3.3	up	MYO7B	2.2	up				
CXCL12	3.2	up	F10	2.2	up				
OCA2	3.2	up	GPER	4.2	up				
ELMO1	3.2	up	KIAA0319	1.7	up				
GAL3ST2	3.2	up	STIM2	2.0	up				
BTNL9	3.2	up	ARNT2	1.9	up				
FLI45803	3.2	up	HUNK	1.7	up				
ALDOB	3.2	up	TOX3	2.4	up				
RTN4RL2	3.2	up	FOXL1	2.3	up				
NNAT	3.2	up	SERPINA3	-3.5	down				
KIT	3.1	up	DSG3	-5.0	down				
EPHB3	3.1	up	TGFB2	-2.2	down				
RFTN1	3.1	up	ODZ2	-5.2	down				
DOC2A	3.1	up	NAV3	-5.3	down				
C5orf38	3.1	up	TACSTD2	-4.8	down				
CPS1	3.1	up	GPR87	-2.0	down				
MUC5AC	3.1	up	CDH3	-3.6	down				
C20orf151	3.0	up	RBP1	-1.5	down				
GLYATL1	3.0	up	LUM	-2.8	down				
FER1L6	3.0	up	FSTL1	-2.0	down				
DPEP1	3.0	up	LOXL2	-1.2	down				
FGG	-10.4	down	PXDNL	-2.7	down				
CTSF	-8.9	down	PAPPA	-3.2	down				
TGFB2	-8.7	down	CTSF	-9.0	down				
DSG3	-8.3	down	FGF2	-4.1	down				
MFAP5	-7.6	down	PTRF	-1.7	down				
SERPINA3	-7.4	down	FGB	-4.5	down				
SERPINE1	-6.9	down	BIRC7	-1.6	down				
PXDNL	-6.6	down	TRIML2	-2.9	down				
COL5A1	-6.5	down	FGG	-4.0	down				
IGFL2	-6.4	down	RRAD	-2.4	down				
GPR1	-6.3	down	CNN3	-2.2	down				
KANK4	-6.2	down	ANTXR1	-1.9	down				
EPHB2	-6.1	down	PCDH7	-2.9	down				
FSTL1	-6.1	down	CD24	-2.6	down				
KRT17	-6.0	down	GJC1	-4.0	down				
LUM	-6.0	down	DSE	-2.7	down				
NAV3	-5.9	down	MMP14	-3.4	down				
ODZ2	-5.8	down	RAB31	-3.0	down				
TACSTD2	-5.6	down	KLF8	-3.3	down				
MARCH4	-5.6	down	RAMP1	-3.4	down				
NNMT	-5.6	down	ID3	-1.8	down				
NPTX1	-5.5	down							
CCDC80	-5.5	down							
GPR87	-5.2	down							
CDH3	-5.2	down							
TIMP3	-5.2	down							
FGB	-5.2	down							
RBP1	-5.2	down							
SLC2A3	-5.1	down							
TRIML2	-5.1	down							
PTRF	-5.1	down							
CST6	-5.0	down							
CGB8	-4.9	down							
KRT6A	-4.8	down							
CD24	-4.8	down							
TGFB1	-4.8	down							
LOXL2	-4.8	down							
BIRC7	-4.7	down							
CAV1	-4.7	down							
TAGLN	-4.7	down							

Appendix 5

<u>GENE NAME</u>	<u>LOG2 fold change in P.2 LUNG</u>	<u>LOG2 fold change in P.2 bone (3x cutoff)</u>
KERA	up 12x	up 13x
ARMC3	up 11x	up 11x
ZP4	up 10x	up 10x
FAM101A	up 8x	up 7x
RYR2	up 7x	up 8x
NEU4	up 6x	up 5x
FAM155A	up 6x	up 7x
SCG3	up 6x	up 6x
CYP2B6	up 6x	up 6x
ST6GAL2	up 5x	up 7x
PARM1	up 5x	up 6x
COL9A3	up 5x	up 4x
SCG2	up 5x	up 5x
CCDC148	up 5x	up 6x
RAB25	up 4x	up 2x
CNN3	down 5x	down 4x
MMP14	down 5x	down 5x
FGF2	down 5x	down 11x
LRRN2	down 5x	down 4x
ANTXR1	down 6x	down 7x
RAB31	down 6x	down 3x
CD24	down 9x	down 10x
PTRF	down 10x	down 9x