

Award Number:  
**W81XWH-07-1-0402**

TITLE:  
***Role of PTEN in the Tumor Microenvironment***

PRINCIPAL INVESTIGATOR:  
**Gustavo Leone, PhD – Principal Investigator**  
**Michael Ostrowski, PhD – Co-Investigator**  
**Metin Gurcan, PhD – Co-Investigator**

CONTRACTING ORGANIZATION:  
**The Ohio State University Research Foundation**  
**Columbus, OH 43210**

REPORT DATE:  
**June 2010**

TYPE OF REPORT:  
**Final Report**

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.**

<b>1. REPORT DATE</b> 01-06-2010		<b>2. REPORT TYPE</b> Final Report		<b>3. DATES COVERED (From - To)</b> 15 MAY 2007 - 14 MAY 2010	
<b>4. TITLE AND SUBTITLE</b> Role of PTEN in the Tumor Microenvironment				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> W81XWH-07-1-0402	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Gustavo Leone, PhD				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  The Ohio State University Research Foundation  Columbus, OH 43210				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for public release					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> Recent scientific advances have revealed that a malignant tumor can be viewed as an organ consisting of different types of interacting cells. Different tumor cell types play different roles in the growth and development of the tumor and thus in the end, all cells within the tumor may help pave the way for further tumor growth in a patient with cancer. Current work in the field indicates that fibroblasts surrounding breast cancers are particularly important for cancer progression, but no one knows why. Our experimental approach to this problem is a direct one; we targeted the mutation of genes in stromal fibroblasts surrounding the tumor in order to learn whether these genes are important for cancer progression. Using this approach, we have shown that the PTEN gene in fibroblasts is a major gene involved in suppressing epithelial breast cancers. In order to understand how PTEN works in fibroblasts, we measured the immediate and long-term consequences of PTEN mutation on the biology of all the cells surrounding the tumor, including the tumor cells themselves, as well as the matrix that holds these cells together. Because the entire system is so complex, we studied how the PTEN gene behaves as a tumor suppressor by developing extremely detailed three-dimensional images of tumors where each image is annotated with detailed cancer related molecular information. From this work we were able to quantify major and minor changes in the extracellular matrix and cells in the microenvironment. We showed that Pten function in stromal fibroblasts is involved in reprogramming the entire tumor microenvironment to suppress mammary epithelial tumors. We anticipate that this work will lead to the design of novel therapeutic strategies that target fibroblasts and that could be used in combination with current therapies that target epithelial cells, to stop tumor growth and prevent reoccurrence of metastasis.					
<b>15. SUBJECT TERMS</b> Tumor microenvironment					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>  UU	<b>18. NUMBER OF PAGES</b>  99	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRMC
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			<b>19b. TELEPHONE NUMBER (include area code)</b>

## Table of Contents

	<u>Page</u>
<b>Introduction.....</b>	<b>4</b>
<b>Body.....</b>	<b>4</b>
<b>Key Research Accomplishments.....</b>	<b>5</b>
<b>Reportable Outcomes.....</b>	<b>5</b>
<b>Conclusion.....</b>	<b>6</b>
<b>References.....</b>	<b>6</b>
<b>Appendices.....</b>	<b>8</b>
<b>Principal Investigator’s Biographical Sketch.....</b>	<b>8</b>
<b>List of Personnel.....</b>	<b>16</b>
<b>Publications</b>	
– <b>Nature.....</b>	<b>17</b>
– <b>PNAS, USA.....</b>	<b>27</b>
– <b>Medical Image Analysis.....</b>	<b>33</b>
– <b>Cancer Research.....</b>	<b>44</b>
– <b>Journal of Biomedical Informatics.....</b>	<b>56</b>
<b>Animal Protocol.....</b>	<b>67</b>

## INTRODUCTION:

Emerging evidence points to a particularly important role of **fibroblasts** within in the tumor stroma in promoting phenotypic changes that allow the extracellular matrix to evolve and contribute to the *invasiveness of epithelial tumor cells*. However, how stromal fibroblasts that promote cancer progression are poorly understood and are only beginning to emerge. The Leone laboratory has developed a system to conditionally manipulate the mouse genome in stromal fibroblasts of the mammary gland (*FSP-cre*). Using this genetic system we present preliminary data suggesting that *Pten* function in mammary stromal fibroblasts plays a critical role in the progression of *ErbB2*-initiated mammary epithelial tumors. The task at hand is to now genetically and biochemically identify the key '*stroma-specific*' tumor suppressor activities of *Pten*, and to then relate these stroma-specific functions to the structural and molecular events taking place during tumor progression.

The immense genetic and cellular diversity of the tumor microenvironment, however, has made the problem of cancer progression difficult to address by traditional experimental approaches. The Leone and Saltz groups have therefore joined expertise to develop novel *genetic* and *high-resolution 3D bioinformatics models* that integrate *spatial* and *molecular* information relating to the genetics and biochemical properties of the tumor microenvironment. These models will provide an effective platform for identifying the relationships between the different cellular compartments of the tumor microenvironment that are critical for cancer progression. The proposed studies will fall into two major efforts, each divided into two phases. The first major effort involves the genetic analysis of *Pten* in mouse tumor models. The progress report outlined below describes the progress made in this phase of the study. The work is now being prepared for publication.

## BODY:

Coordinated signaling between different cell types of the '*normal stroma*' is required during embryonic and adult development ([Wiseman BS, Werb Z, 2002.](#)). While cellular and ECM activities of the stroma are maintained in balance throughout development, they can be appropriately activated in response to extreme but normal physiological cues, such as wounding, inflammation or pregnancy ([Schedin P, 2006](#)). The stroma can also be inappropriately activated, such as in cancer ([Nelson CM, Bissell MJ, 2006](#)). Fibroblasts are a principal constituent of the stroma responsible for the synthesis of growth and survival factors, chemokines, structural components of the ECM and enzymes that control its turnover. In breast tumors, stromal fibroblasts are believed to adapt and continuously co-evolve along with tumor epithelial cells (Littlepage et al., 2005). Fibroblasts are implicated in fostering transformation and tumor growth by providing factors that induce epithelial cell proliferation, ECM remodeling and blood vessel recruitment (Bhowmick et al., 2004b; Mueller and Fusenig, 2004; Kalluri and Zeisberg, 2006). Despite extensive evidence for a role of the tumor stroma in carcinogenesis, relatively little is known about the signaling pathways involved in the communication between the different cellular compartments of the tumor microenvironment that contribute to the cancer phenotype.

Phosphatase and tensin homolog (*PTEN*) is a tumor suppressor with lipid and protein phosphatase activity (Myers et al., 1998) that impacts several signaling pathways, including phosphoinositide 3-kinase (PI3K), and Ras-MAPK-Erk1/2 signaling pathways. *Pten* inactivation in mice and humans leads to a disruption in cell polarity, cell architecture, and chromosomal integrity as well as in the promotion of cell cycle progression, cell growth and stem cell self-

renewal (Di Cristofano and Pandolfi, 2000). Not surprisingly, its somatic or germline inactivation contributes to the genesis of many tumor types primarily epithelial in origin. While tremendous progress in understanding PTEN function in tumor cells has been made since its discovery over a decade ago, relatively little is known about its potential role in the tumor stroma. Patients with Cowden syndrome have germline mutations in *PTEN*, suggesting that the higher risk for developing breast cancer in these patients could be due to *PTEN* activity in either epithelial or stromal compartments. Within the context of this project, we have generated a mesenchymal-specific *Fsp-cre* transgene and conditional alleles of *Pten* (*Pten<sup>loxP</sup>*) in mice to ablate its function in mammary stromal fibroblasts *in vivo* and rigorously evaluate its role in the tumor microenvironment. We show that *Pten* ablation in mammary stromal fibroblasts results in the induction and activation of  $\beta$ -Ets2-P(T72), massive remodeling of the ECM, recruitment of innate immune cells, and enhanced progression and malignancy of mammary tumors of epithelial origin. These findings have been recently published in *Nature* and *PNAS* and we believe expand *Pten*'s repertoire as a tumor suppressor by identifying the fibroblast as a key site from which it exerts its powerful tumor suppressive influence on adjacent mammary epithelium. This novel function of *Pten* in maintaining homeostasis within the mammary microenvironment may also be relevant in the suppression of epithelial tumors of other organs. Such a role for *Pten* may extend beyond cancer, to conditions where the microenvironment is speculated to profoundly impact disease manifestation including in autoimmune syndromes, lung fibrosis and neurodegeneration. Interestingly, the stromal *Pten* expression signature identified here includes genes that have been causally linked with ECM deposition and inflammation in rheumatoid arthritis, lung fibrosis and neurodegeneration.

#### **KEY RESEARCH ACCOMPLISHMENTS:**

TASK 1 (month 3- 18): Manuscript published in *Nature* (Trimboli et al.) attached.

TASK 2 (months 3-18): Manuscript published in *PNAS* (Hui Wang et al.).

TASK 3 (month 12-18): See *Nature* (Trimboli et al.) article attached. This task is addressed in this article.

TASK 4 (month 18-24): See *Nature* (Trimboli et al.) article attached.

TASKS 5-7 (month 1- 24): The “segmentation” component has encountered significant technical hurdles dealing with image segmentation, so we have adapted our aim to focus on nuclear segmentation successfully.

#### **REPORTABLE OUTCOMES:**

##### 1. Manuscripts Published:

- A. Trimboli A.J., Cantemir-Stone C.Z., Li F., Wallace J.A., Merchant A., Creasap N., Thompson J.C., Caserta E., Wang H., Chong J-L., Naidu S., Wei G., Sharma S.M., Stephens J.A., Fernandez S.A., Gurcan M.N., Weinstein M.B., Barsky S.H., Yee L., Rosol T.J., Stromberg P.C., Robinson M.L., Pepin F., Hallett M., Park M., Ostrowski

- M.C., **Leone G.** (2009). Pten in stromal fibroblasts suppresses mammary epithelial tumors. *Nature* 461:1084-91.
- B. Wang H., Karikomi M., Naidu S., Rajmohan R., Caserta E., Chen H-Z., Rawahneh M., Moffitt J., Stephens J.A., Fernandez S.A., Weinstein M., Wang D., Sadee W., La Perle K., Stromberg P., Rosol T.J., Eng C., Ostrowski M.C., **Leone G.** (2010). Allele-specific tumor spectrum in *Pten* knockin mice. *Proc. Natl. Acad. Sci. USA.* 107:5142-5147.
- C. Mosaliganti K., Janoos F., Irganoglu O., Ridgway R., Machiraju R., Huang K., Saltz J., **Leone G.**, Ostrowski M.C. (2009). Tensor classification of N-point correlation function features for histology tissue segmentation. *Med. Image Anal.* 13:156-166.
- D. Zabuawala T., Taffany D.A., Sharma S.M., Merchant A., Adair B., Srinivasan R., Rosol T.J., Fernandez S.A., Huang K., **Leone G.**, Ostrowski M.C. (2010). An Ets2 specific transcriptional program in tumor-associated macrophages promotes tumor metastasis. *Cancer Res.* 70(4): 1323-1333.
- E. Mosaliganti K., Pan T., Ridgway R., Sharp R., Cooper L., Gulacy A., Sharma A., Irfanoglu O., Machiraju R., Kurc T., de Bruin A., Wenzel P., **Leone G.**, Saltz J, Huang K. (2008). An imaging workflow for characterizing phenotypical change in large histological mouse model datasets. *J. Biomed. Inform.* 41:863-873.
2. Submitted an NCI grant application (U01) that is related to this work. Not funded.
  3. Currently submitting an NCI grant application (U54 TMEN) that is related to this work.
  4. Currently submitting an NCI grant application (U54 TMEN) that is related to this work.

## CONCLUSIONS:

The mechanism by which Pten in the stroma regulates the expression of such an extensive set of genes likely involves complex transcriptional networks. The importance of the Pten-Ets2 axis in stromal fibroblasts is consistent with previous work from Oshima and colleagues that showed a critical cell non-autonomous role for Ets2 in mammary tumor growth (Tynan et al. 2005). Moreover, analysis of tumor-stroma gene expression signatures identified Ets2 activation as a key event associated with breast cancer patients having the worst prognosis (Park et al., 2007). In summary, this study offers a molecular basis for how altered signaling from the tumor stroma contributes to the most malignant characteristics of the breast cancer phenotype.

## REFERENCES:

1. [Wiseman BS, Werb Z.](#) (2002). Stromal effects on mammary gland development and breast cancer. *Science* 296, 1046-1049
2. [Schedin P.](#) (2006) Pregnancy-associated breast cancer and metastasis. *Nat Rev Cancer.* 6, 281-291.
3. [Nelson CM, Bissell MJ.](#) (2006). Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. *Annu Rev Cell Dev Biol.*, 287-309.

4. [Littlepage LE, Egeblad M, Werb Z. \(2005\). Coevolution of cancer and stromal cellular responses. Cancer Cell 7, 499-500.](#)
5. [Bhowmick NA, Neilson EG, Moses HL. \(2004b\). Stromal fibroblasts in cancer initiation and progression. Nature 432, 332-337.](#)
6. [Mueller MM, Fusenig NE. \(2004\). Friends or foes - bipolar effects of the tumor stroma in cancer. Nat Rev Cancer 4, 839-849.](#)
7. [Kalluri R, Zeisberg M. \(2006\). Fibroblasts in cancer. Nat Rev Cancer. 6, 392-401.](#)
8. [Myers MP, Pass I, Batty IH, Van der Kaay J, Stolarov JP, Hemmings BA, Wigler MH, Downes CP, Tonks NK. \(1998\). The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. Proc Natl Acad Sci U S A 95, 13513-13518.](#)
9. Di Cristofano A, Pandolfi PP. (2000). The multiple roles of PTEN in tumor suppression. Cell 100, 387-390.
10. [Tynan JA, Wen F, Muller W J, Oshima RG. \(2005\). Ets2-dependent microenvironmental support of mouse mammary tumors. Oncogene 24, 6870-6876.](#)
11. [Park ES, Lee JS, Woo HG, Zhan F, Shih JH, Shaughnessy JD Jr, Frederick Mushinski J. \(2007\). Heterologous tissue culture expression signature predicts human breast cancer prognosis. PLoS ONE 2, e145.](#)

BIOGRAPHICAL SKETCH

NAME P		POSITION TITLE	
Leone, Gustavo W.		Associate Professor	
eRA COMMONS USER NAME			
gleone			
EDUCATION/TRAINING			
INSTITUTION AND LOCATION	DEGREE (if applicable)	YEAR(s)	FIELD OF STUDY
University of Calgary, Calgary, Canada	B.Sc. 1	988	Biochemistry
University of Calgary, Calgary, Canada	Ph.D. 1	994	Molecular Virology
Duke University Medical Center, Durham, N.C	Postdoc 1	994-1998	Genetics
Ohio State University, Columbus, OH	Assoc Prof	1999-present	Cancer Genetics

**A. Personal Statement**

I am an active member in the Molecular Biology and Cancer Genetics program of the OSU Comprehensive Cancer Center. As such, I participate in program meetings, seminars and collaborative efforts such as programmatic grants and shared resource usage to further the goal of the OSUCCC of conducting basic and clinical research pertaining to the prevention, detection and therapy of cancer.

**B. Positions and Honors.**

*Academic Appointments*

1987 – 1988	Undergraduate Research Assistant, Department of Microbiology and Infectious Diseases, University of Calgary (Supervisor: Dr. Patrick W.K. Lee)
1988 – 1994	Graduate Student, Department of Microbiology and Infectious Diseases, University of Calgary (Supervisor: Dr. Patrick W.K. Lee)
1990 – 1994 in	Graduate student assistant, University of Calgary, Project Lab (intensive laboratory training for undergraduates in virology and molecular biology)
1994 – 1998	Postdoctoral Fellow, Department of Genetics, Duke University Medical Center (HHMI), Durham, North Carolina. Role of E2F in Cell Cycle Control (Mentor: Dr. Joseph R. Nevins)
1999–2006	Assistant Professor, Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio.
2006–present	Associate Professor, Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio.
2008–present	Director, Tumor Microenvironment Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio.

*Honors and Awards*

1989 – 1994	Alberta Heritage Foundation for Medical Research Studentship
1994 – 1997	Medical Research Council of Canada Fellowship and Alberta Heritage Foundation for Medical Research Fellowship
1997 – 1999	Medical Research Council of Canada Centennial Fellowship
1999	The Robert M. and Barbara R. Bell Basic Science of Cancer Award
1999 – 2001	Scholar Award, V - Foundation
2001 – 2005	Scholar Award, PEW Charitable Trust
2005 – 2009	Scholar Award, Leukemia and Lymphoma Society

**C. Selected peer-reviewed publications (from 91)**

1. Mah D.C.W., Leone G., Jankowski J.M., Lee P.W.K. (1990). The N-terminal quarter of Reovirus cell attachment protein sigma1 possesses intrinsic virion-anchoring function. *Virology* 79:95-103.



2. Duncan R., Horne D., Strong J.E., **Leone G.**, Pon R.T., Yeung M.C., Lee P.W.K. (1991). Conformational and functional analysis of the C-terminal globular head of the Reovirus cell attachment protein. *Virology* 182:810-819.
3. Strong J.E., **Leone G.**, Duncan R., Sharma R.K., Lee, P.W.K. (1991). Biochemical and biophysical characterization of the Reovirus cell attachment protein sigma1: Evidence that it is a homotrimer. *Virology* 184:23-32.
4. **Leone G.**, Duncan R., Lee P.W.K. (1991). Trimerization of the Reovirus cell attachment protein (sigma1) induces conformational changes in sigma1 necessary for its cell binding function *Virology* 184:758-761.
5. **Leone G.**, Mah D.C.W., Lee, P.W.K. (1991). The incorporation of Reovirus cell attachment protein sigma1 into virions requires the N-terminal hydrophobic tail and the adjacent heptad repeat region. *Virology* 182:346-350.
6. **Leone G.**, Duncan R., Mah D.C.W., Price A., Cashdollar L.W., Lee P.W.K. (1991). The N-terminal heptad repeat region of Reovirus cell attachment protein sigma1 is responsible for sigma1 oligomer stability and possesses intrinsic oligomerization function. *Virology* 182:336-345.
7. **Leone G.**, Maybaum L., Lee P.W.K. (1992). The Reovirus cell attachment protein possesses two independently active trimerization domains: Basis of dominant negative effects. *Cell* 71:479-488.
8. Fernandes J., Tang D., **Leone G.**, Lee P.W.K. (1994). Binding of Reovirus to receptor leads to conformational changes in viral capsid proteins that are reversible upon virus detachment. *J. Biol. Chem.* 269(25):17043-17047.
9. Lee P.W.K., **Leone G.** (1994). Reovirus protein sigma1: From cell attachment to protein oligomerization and folding mechanisms. *BioEssays* 16:199-206.
10. DeGregori J.\*, **Leone G.\***, Ohtani K., Miron A., Nevins J.R. (1995). E2F1 accumulation bypasses a G<sub>1</sub> arrest resulting from the inhibition of G<sub>1</sub> cyclin-dependent kinase activity. *Genes Dev.* 9:2873-2887. \*These authors contributed equally to this work.
11. Gilmore R., Coffey M., **Leone G.**, McLure K., Lee P.W.K. (1996). Co-translational trimerization of the Reovirus cell attachment protein. *The EMBO Journal* 15(11):2651-2658.
12. **Leone G.**, Coffey M., Gilmore R., Duncan R., Lee P.W.K. (1996) C-terminal trimerization, but not N-terminal trimerization, of the Reovirus cell attachment protein is a post-translational and hsp70/ATP dependent process. *J. Biol. Chem.* 271(14):8466-8471.
13. Smith E., **Leone G.**, DeGregori J., Jakoi L., Nevins J.R. (1996). The accumulation of an E2F-p130 transcriptional repressor distinguishes a G<sub>0</sub> cell state from a G<sub>1</sub> cell state. *Mol. Cell. Biol.* 16:6965-6976.
14. Ohtani K., DeGregori J., **Leone G.**, Herendeen D.R., Kelly T.J., Nevins J.R. (1996). Expression of the HsOrc1 gene, a human ORC1 homolog, is regulated by cell proliferation via the E2F transcription factor. *Mol. Cell. Biol.* 16:6977-6984.
15. DeGregori J., **Leone G.**, Miron A., Jakoi L., Nevins J.R. (1997). Distinct roles for E2F proteins in cell growth control and apoptosis. *Proc. Natl. Acad. Sci. USA.* 94:7245-7250.
16. Yan Z., DeGregori J., Shohet R., **Leone G.**, Stillman B., Nevins J.R., Williams R.S. (1997). Cdc6 is regulated by E2F and is essential for DNA replication in mammalian cells. *Proc. Natl. Acad. Sci. USA.* 95:3603-3608
17. **Leone G.**, DeGregori J., Sears R., Jakoi L., Nevins J.R. (1997). Myc and Ras collaborate in inducing accumulation of active cyclin E/Cdk2 and E2F. *Nature* 387:422-425.
18. Nevins J.R., **Leone G.**, DeGregori J., Jakoi L. (1997). Role of the Rb/E2F pathway in cell growth control. *J. Cell Physiol.* 173:233-236.
19. Nevins J.R., DeGregori J., Jakoi L., **Leone G.** (1997). Functional analysis of E2F transcription factor. *Meth. Enzymol.* 283:205-219.
20. Kowalik T.F., DeGregori J., **Leone G.**, Jakoi L., Nevins J.R. (1998). E2F1-Specific induction of apoptosis and p53 accumulation, which is blocked by Mdm2. *Cell Growth Differ.* 9:113-118.
21. Smith J., **Leone G.**, Nevins J.R. (1998). Distinct mechanisms control the accumulation of the Rb-related p107 and p130 proteins during cell growth. *Cell Growth Differ.* 9:297-303.
22. **Leone G.**, DeGregori J., Yan Z., Jakoi L., Ishida S., Williams R.S., Nevins J.R. (1998). E2F3 activity is regulated during the cell cycle and is required for the induction of S phase. *Genes Dev.* 12:2120-2130.
23. Sears R., **Leone G.**, DeGregori J., Nevins J.R. (1999). Ras enhances Myc protein stability. *Mol. Cell* 3:169-179.
24. **Leone G.**, DeGregori J., Jakoi L., Cook J.G., Nevins J.R. (1999). Collaborative role of E2F transcriptional activity and G<sub>1</sub> cyclin-dependent kinase activity in the induction of S phase. *Proc. Natl. Acad. Sci. USA.* 96:6626-6631.
25. Smith D.S., **Leone G.**, DeGregori J., Ahmed M.N., Qumsiyeh M.B., Nevins J.R. (2000). Induction of DNA replication in adult rat neurons by deregulation of the retinoblastoma/E2F G<sub>1</sub> cell cycle pathway. *Cell Growth Differ.* 11:625-633.
26. **Leone G.**, Nuckolls F., Ishida S., Adams M., Sears R., Jakoi L., Miron A., Nevins J.R. (2000). Identification of a novel E2F3 product suggests a mechanism for determining specificity of repression by Rb proteins. *Mol. Cell Biol.* 20:10, 3626-3632.

27. Adams M.R., Sears R., Nuckolls F., **Leone G.**, Nevins J.R. (2000). Complex transcriptional regulatory mechanisms control expression of the E2F3 locus. *Mol. Cell Biol.* 20:10, 3633-3639.
28. **Leone G.**, Sears R., Huang E., Rempel R., Nuckolls F., Park C., Giangrande P., Wu L., Saavedra H.I., Field S.J., Thompson M.A., Yang H., Fujiwara Y., Greenberg M.E., Orkin S., Smith C., Nevins J.R. (2001). Myc requires distinct E2F activities to induce S phase and apoptosis. *Mol. Cell* 8:105-113.
29. Weng L.P., Gimm O., Kum J.B., Smith W.M., Zhou X.P., Wynford-Thomas D, **Leone G.**, Eng C. (2001). Transient ectopic expression of *PTEN* in thyroid cancer cell lines induces cell cycle arrest and cell type-dependent cell death. *Hum. Mol. Genet.* 10:3, 251-258.
30. Tanner S.M., Austin J.L., **Leone G.**, Rush L.J., Plass C., Heinonen K., Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J., Caligiuri M.A., Bloomfield C.D., de la Chapelle A. (2001). BAALC, the human member of a novel mammalian neuroectoderm gene lineage, is implicated in hematopoiesis and acute leukemia. *Proc. Natl. Acad. Sci. USA.* 98:24,13901-13906.
31. Wu L., Timmers C., Maiti B., Saavedra H.I., Sang L., Chong G.T., Nuckolls F., Giangrande P., Wright F.A., Field S.J., Greenberg M.E., Orkin S., Nevins J.R., Robinson M.L., **Leone G.** (2001). The E2F1-3 transcription factors are essential for cellular proliferation. *Nature* 414:457-462.
32. Ishida S., Huang E., Zuzan H., Spang R., **Leone G.**, West M., Nevins, J.R. (2001). Role for E2F in control of both DNA replication and mitotic functions as revealed from DNA micro array analysis. *Mol. Cell. Biol.* 21: 14, 4684-4699.
33. Zhao Y., Gilmore R., **Leone G.**, Coffey M.C., Weber B., Lee P.W.K. (2001). Hsp90 phosphorylation is linked to its chaperoning function. *J. Biol. Chem.* 276:35 32822-32827.
34. Lan Z., Sever-Chroneos Z., Strobeck M.W., Park C., Baskaran R., Edelmann W., **Leone G.** Knudsen E. S. (2002). DNA damage invokes mismatch-repair-dependent cyclin D1 attenuation and retinoblastoma signaling pathways to inhibit cdk2. *J. Biol. Chem.* 277:10, 8372-8381.
35. Cook J.G., Park C., Burke T.W., **Leone G.**, DeGregori J., Engel A., Nevins J.R. (2002). Analysis of Cdc6 function in the assembly of mammalian prereplication complexes. *Proc. Natl. Acad. Sci. USA.* 99:3, 1347-1352.
36. Saavedra H.I., Wu L., de Bruin A., Timmers C., Rosol T.R., Weinstein M., Robinson M.L, **Leone G.** (2002). Specificity of E2F1, E2F2 and E2F3 in mediating phenotypes induced by loss of Rb. *Cell Growth Differ.* 13:5, 215-225. (cover of journal)
37. de la Puente A., Hall J., Wu Y.Z., **Leone G.**, Peters J., Yoon B.J., Soloway P., Plass C. (2002). Structural characterization of Rasgrf1 and a novel linked imprinted locus. *Gene* 291:287-297.
38. Toribio R.E., Kohn C.W., **Leone G.**, Capen C.C., Rosol T.J. (2003). Molecular cloning and expression of equine calcitonin, calcitonin gene-related peptide-I, and calcitonin gene-related peptide-II. *Mol. Cell Endocrinol.* 199(1-2):119-128.
39. Wu L., de Bruin A., Saavedra H.I., Trimboli A., Yang Y., Opavska J., Wilson P., Starovic M., Ostrowski M.C., Cross J.C., Weinstein M., Rosol T.J., Robinson M.L., **Leone G.** (2003). Extraembryonic function of Rb is essential for embryonic development and viability. *Nature* 421: 942-947. (N&V; 421: 903-004)
40. Saavedra H.I., Maiti B., Timmers C., Altura R., Fukasawa K., **Leone G.** (2003). Inactivation of *E2F3* results in premature centriole separation and centrosome amplification. *Cancer Cell* 3: 333-346. (Cover).
41. de Bruin A., Wu L., Saavedra H.I., Wilson P., Yang Y., Weinstein M., Rosol T.J., Robinson M.L., **Leone G.** (2003). Rb function in extraembryonic lineages is critical for the control of apoptosis in the central nervous system of Rb-deficient mice. *Proc. Natl. Acad. Sci. USA.* 100:11, 6546-6551.
42. de Bruin A., Maiti B., Jakoi L., Timmers C., **Leone G.** (2003). Identification and characterization of E2F7, a novel mammalian E2F family member capable of blocking cellular proliferation. *J. Biol. Chem.* 278:43, 42041-42049.
43. Cheng S., Hsia Y.C., **Leone G.**, Liou H-C. (2003). Cyclin E and Bcl-XL cooperatively induce cell cycle progression in c-Rel<sup>-/-</sup> B cells. *Oncogene* 22:8472-8486.
44. Zhang J., Gray J., Wu L., **Leone G.**, Rowan S., Cepko C.L., Zhu X., Craft C.M., Dyer M.A. (2004). Rb regulates proliferation and rod photoreceptor development in the mouse retina. *Nature Genetics* 36:4,1-10.
45. Dai Z., Popkie A.P., Zhu W., Timmers C., Raval A., Tannehill-Gregg S., Morrison C.D., Auer H., Kratzke R.A., Niehans G., Amatschek S., Sommergruber W., **Leone G.**, Rosol T., Otterson G.A., Plass C. (2004). Bone morphogenetic protein 3B silencing in non-small-cell lung cancer. *Oncogene* 23:3521-3529.
46. Jiang Y., Saavedra H.I., Holloway M.P., **Leone G.**, Altura R.A. (2004). Aberrant regulation of survivin by the Rb/E2F family of proteins. *J. Biol. Chem.* 279(39):40511-20.
47. Yu L., Liu C., Bennett K., Wu Y-Z., Dai Z., Vandeussen J., Opavsky R., Raval A., Trikha, P., Rodriguez B., Becknell B., Mao C., Lee S., Davuluri R.V., **Leone G.**, Van den Veyer I.B., Caligiuri M.A., Plass C. (2004). A NotI-EcoRV promoter library for studies of genetic and epigenetic alterations in mouse models of human malignancies. *Genomics.* 84(4):647-60.

48. Iavarone A., King E.R., Dai X., **Leone G.**, Stanley E.R., Lasorella A. (2004). Retinoblastoma promotes definitive erythropoiesis by repressing Id2 in fetal liver macrophages. *Nature* 432:1040-1045.
49. Chen Q., Liang D., Yang T., **Leone G.**, Overbeek P.A. (2004). Distinct capacities of individual E2Fs to induce cell cycle re-entry in postmitotic lens fiber cells of transgenic mice. *Dev. Neurosci.* 26(5-6):435-45.
50. Maiti B., Li J., de Bruin A., Gordon F., Timmers C., Opavsky R., Patil K., Tuttle J., Cleghorn W. **Leone G.** (2005). Cloning and characterization of mouse E2F8, a novel mammalian E2F family member capable of blocking cellular proliferation. *J. Biol. Chem.* 280:18211-20.
51. Logan N., Graham A., Zhao X., Fisher R., Maiti B., **Leone G.**, La Thangue N. (2005). E2F-8: an E2F family member with a similar organization of DNA binding domains to E2F-7. *Oncogene* 24: 5000-04.
52. Sharma N., Timmers C., Trikha P., Saavedra H.I., Obery A., Opavsky R., **Leone G.** (2006). Control of the p53-p21CIP1 axis by E2f1, E2f2 and E2f3 is essential for G1/S progression and cellular transformation. *J. Biol. Chem.* 281:36124-31.
53. Dorrance A., Liu S., Yuan W., Becknell B., Arnoczky K., Guimond M., Strout M., Feng L., Nakamura T., Yu L., Rush L., Weinstein M., **Leone G.**, Wu L., Ferketich A., Whitman S., Marcucci G., and Caligiuri M. (2006). The Mll partial tandem duplication induces aberrant Hox expression in vivo via specific epigenetic alterations. *J. Clin. Invest.* 116:2707-16.
54. Mosaliganti R., Pan T., Sharp R., Ridgway R., Iyengar S., Gulacy A., Wenzel P., de Bruin A., Machiraju R., Huang K., **Leone G.**, Saltz J. (2006). Registration and 3D visualization of large microscopy images, *Proc. SPIE Ann. Med. Im. Meet.* 923-34.
55. Sharp R., Ridgway R., Mosalignati K., Irfanoglu O., Wenzel P., Machiraju R., Pan T., de Bruin A., **Leone G.**, Huang K., Saltz J. (2006). Examining phenotype differences in mouse placenta with volume rendering and segmentation. *Proc. IEEE/NLM Life Sc. Syst. & Appl.* 70-71.
56. Timmers C., Opavsky R., Maiti B., Wu L., Wu J., Orringer D., Sharma N., Saavedra H.I., **Leone G.** (2007). E2f1-3 control E2F-target expression and cellular proliferation via a p53-dependent negative feedback loop. *Mol. Cell Biol.* 27:65-78.
57. Sharp R., Ridgway R., Mosaliganti K., Wenzel P., Pan T., de Bruin A., Machiraju R., Huang K., **Leone G.**, Saltz J. (2007). Volume rendering phenotype differences in mouse placenta microscopy data. *Comput. Sci. Eng.* 9:38-47.
58. Wenzel P., Wu L., de Bruin A., Chong J-L., Chen W-Y., Dureska G., Sites E., Pan T., Sharma A., Huang K., Ridgway R., Mosaliganti K., Sharp R., Machiraju R., Saltz J., Yamamoto H., Cross J., Robinson M.L., **Leone, G.** (2007). Rb is critical in a mammalian tissue stem cell population. *Genes Dev.* 21:85-97.
59. Wenzel P., **Leone G.** (2007). Expression of Cre recombinase in early diploid trophoblast cells of the mouse placenta. *Genesis* 45:129-134.
60. Mosaliganti K., Janoos F., Sharp R., Ridgway R., Machiraju R., Huang K., Wenzel P., deBruin A., **Leone G.**, Saltz J. (2007). Detection and visualization of surface-pockets to enable phenotyping studies. *IEEE Trans. Medical Imaging* 26:1283-1290.
61. Janoos F., Irfanoglu O., Mosaliganti K., Machiraju R., Huang K., Wenzel P., de Bruin A., **Leone G.** (2007). Multiple-resolution image segmentation using the 2-point correlation functions. *Proc. IEEE Internat. Symp. Medical Imaging* 300-303.
62. McClellan K., Ruzhynsky V., Douda D., Vanderluit J., Ferguson K., Chen D., Bremner R., Park D., **Leone G.**, Slack R. (2007). Unique requirement for Rb/E2F3 in neuronal migration: evidence for cell cycle-independent functions. *Mol. Cell Biol.* 27:4825-43.
63. Chen D., Opavsky R., Pacal M., Tanimoto N., Wenzel P., Seeliger M., **Leone G.**, Bremner R. (2007). Rb-mediated neuronal differentiation through cell cycle independent regulation of E2f3a. *PLoS Biology* 5:e179 1-16.
64. Opavsky R., Wang S-H., Trikha P., Raval A., Huang Y., Wu Y-A., Rodriguez B., Keller B., Liyanarachi S., Wei G., Davuluri R., Weinstein M., Felsher D., Ostrowski M.C., **Leone G.\***, Plass C.\*. (2007). CpG island methylation in cancer is driven by the genetic configuration of tumor cells. *PLoS Genetics* 3:1757-69. \*corresponding authors
65. Opavsky R., Tsai S-Y., Guimond M., Arora A., Opavska J., Becknell B., Kauffman M., Walton N., Stephens J., Fernandez S., Muthusamy N., Felsher D., Porcu P., Caligiuri M., **Leone G.** (2007). Specific tumor suppressor function for E2F2 in Myc-induced T cell lymphomagenesis. *Proc. Natl. Acad. Sci. USA.* 104:39, 15400-15405.
66. Saenz-Robles M., Markovics J.A., Chong J-L., Opavsky R., Whitehead R.H., **Leone G.**, Pipas J.M. (2007). Intestinal hyperplasia induced by SV40 large tumor antigen requires E2F2. *Virology* 81:13191-9.
67. Trimboli A.J., Fukino K., de Bruin A., Wei G., Shen L., Tanner S.M., Rosol T.J., Eng C., Ostrowski M.C., **Leone G.** (2008). In vivo visualization of Myc-induced epithelial-mesenchymal transitions in breast cancer. *Cancer Res.* 68(3) 937-945.
68. Li J., Ran C., Li E., Gordon F., Comstock G., Siddiqui H., Cleghorn W., Chen H-Z., Kornacker K., Liu C-G., Pandit S., Khanizadeh M., Weinstein M., **Leone G.\***, de Bruin A. (2008). Synergistic function of E2F7 and E2F8 is essential for cell survival and embryonic development. *Dev. Cell* 14:62-75. \*corresponding author

69. Tsai S-Y., Opavsky R., Sharma N., Wu L., Naidu S., Nolan E., Ferias-Arias E., Timmers C., Opavska J., de Bruin A., Chong J-L., Trikha P., Stromberg P., Rosol J.R., **Leone G.** (2008). Mouse development with a single E2F activator. *Nature* 454:1137-41.
70. Mosaliganti K., Pan T., Ridgway R., Sharp R., Cooper L., Gulacy A., Sharma A., Irfanoglu O., Machiraju R., Kurc T., de Bruin A., Wenzel P., **Leone G.**, Saltz J, Huang K. (2008). An imaging workflow for characterizing phenotypical change in large histological mouse model datasets. *J. Biomed. Inform.* 41:863-873.
71. Nadella K., Jones G., Trimboli A., Stratakis C., **Leone G.**, Kirschner L. (2008). Targeted deletion of Prkar1a reveals a role for protein kinase A in mesenchymal-to-epithelial transition. *Cancer Res.* 68:2671-2677.
72. Eiring A.M., Neviani P., Santhanam R., Oaks J.J., Chang J.S., Notari M., Willis W., Gambacorti-Passerini C., Volinia S., Marcucci G., Caligiuri M.A., **Leone G.**, Perrotti D. (2008). Identification of novel posttranscriptional targets of the BCR/ABL oncoprotein by ribonomics: requirement of E2F3 for BCR/ABL leukemogenesis. *Blood* 111:816-828.
73. Chong J-L., Tsai S-Y., Sharma N., Opavsky R., Price R., Wu L., Fernandez S., **Leone G.** (2009). E2f3a and E2f3b contribute to the control of cell proliferation and mouse development. *Mol. Cell Biol.* 29:414-424.
74. Lammens T., Li J., **Leone G.**, DeVeylder L. (2009). Atypical E2Fs: new players in the E2F transcription factor family. *Trends Cell Biol.* 19:111-118.
75. Trimboli A.J., Cantemir-Stone C.Z., Li F., Wallace J.A., Merchant A., Creasap N., Thompson J.C., Caserta E., Wang H., Chong J-L., Naidu S., Wei G., Sharma S.M., Stephens J.A., Fernandez S.A., Gurcan M.N., Weinstein M.B., Barsky S.H., Yee L., Rosol T.J., Stromberg P.C., Robinson M.L., Pepin F., Hallett M., Park M., Ostrowski M.C., **Leone G.** (2009). Pten in stromal fibroblasts suppresses mammary epithelial tumors. *Nature* 461:1084-91.
76. Chen D., Pacal M., Wenzel P., Knoepfler P.S., **Leone G.**, Bremner R. (2009). Division and apoptosis in the absence of activating E2fs. *Nature* 462:925-929.
77. Wei G., Srinivasan R., Cantemir-Stone C.Z., Sharma S.M., Santhanam R., Weinstein M., Muthusamy N., Man A.K. Oshima R.G., **Leone G.**, Ostrowski M.C. (2009). Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* 114:1123-1130.
78. Chong J-L., Wenzel P., Sáenz-Robles M., Nair V., Ferrey A., Hagan J., Gomez Y.M., Sharma N., Chen H-Z., Ouseph M., Wang S-H., Trikha P., Culp B., Mezache L., Winton D.J., Sansom O.J., Chen D., Bremner R., Cantalupo P.G., Robinson M.L., Pipas J.M., **Leone G.** (2009). A switch of E2F1-3 function from activators in progenitor cells to repressors in differentiating cells. *Nature* 462:930-934.
79. Chen H-Z., Tsai S-Y., **Leone G.** (2009). Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nature Reviews Cancer* 11:785-797.
80. Eng C., Ostrowski M.C., **Leone G.** (2009). Genomic alteration in tumor stroma. *Cancer Res.* 69:6759-6764.
81. Srinivasan R., Zabuawala T., Huang H., Zhang J., Gulati P., Fernandez S.A., Karlo J.C., Landreth G.E., **Leone G.**, Ostrowski M.C. (2009). Erk1 and Erk2 regulate endothelial cell proliferation and migration during mouse embryonic angiogenesis. *PLoS ONE* 4(12):e8283.
82. Mosaliganti K., Janoos F., Irganoglu O., Ridgway R., Machiraju R., Huang K., Saltz J., **Leone G.**, Ostrowski M.C. (2009). Tensor classification of N-point correlation function features for histology tissue segmentation. *Med. Image Anal.* 13:156-166.
83. Qamri Z., Preet A., Nasser M.W., Bass C.E., **Leone G.**, Barsky S.H., Ganju, R.K. (2009). Synthetic cannabinoid receptor agonists inhibit tumor growth and metastasis of breast cancer. *Mol. Cancer Ther.* 8:3117-29.
84. Martinez L.A., Goluszko E., Chen H-Z., **Leone G.**, Post S., Lozano G., Chen Z., Chauchereau A. (2010). E2F3a is a mediator of DNA damage induced apoptosis. *Mol. Cell Biol.* 30:524-536.
85. Zabuawala T., Taffany D.A., Sharma S.M., Merchant A., Adair B., Srinivasan R., Rosol T.J., Fernandez S.A., Huang K., **Leone G.**, Ostrowski M.C. (2010). An Ets2 specific transcriptional program in tumor-associated macrophages promotes tumor metastasis. *Cancer Res.* 70(4): 1323-1333.
86. Wang H., Karikomi M., Naidu S., Rajmohan R., Caserta E., Chen H-Z., Rawahneh M., Moffitt J., Stephens J.A., Fernandez S.A., Weinstein M., Wang D., Sadee W., La Perle K., Stromberg P., Rosol T.J., Eng C., Ostrowski M.C., **Leone G.** (2010). Allele-specific tumor spectrum in *Pten* knockin mice. *Proc. Natl. Acad. Sci. USA.* 107:5142-5147.
87. Wang H., Chen H-Z., Wu L., de Bruin A., House M., Ng S., Johnson J., Goldenberg L.E., Fang X., Fernandez S.A., Stephens J.A., Yu L., Naidu S., Stromberg P., Rosol T.J., Shapiro C.L., **Leone G.** Suppression of anti-tumor immunity by E2F3. (submitted; *Science*)
88. Li F., Cantemir C-Z., Wallace J., Sharma S.M., Naidu S., Wei G., Trimboli A.J., Weinstein M., Stromberg P., Rosol T.J., Fernandez S.A., Barsky S., Lesurf R., Pepin F., Hallett M., Park M., **Leone G.**, Ostrowski M.C. Ets2 in tumor fibroblasts regulates angiogenesis in breast cancer. (submitted; *Cancer Research*).

89. Trikha P., Sharma N., Opavsky R., Reyes R., Pena C., Ostrowski M.C., Roussel M.F., **Leone G.** E2F1-3 control cell survival and proliferation at distinct stages of myeloid development. (submitted; *J. EMBO*).
90. Wenzel P., Chong J-L., Saenz-Robles M.T., Ferrey A., Hagan J.P., Gomez Y.M., Rajmohan R., Sharma N., Chen H-Z., Pipas J.M., Robinson M.L., **Leone G.** Cell proliferation in the absence of E2F1-3. (submitted; *Dev. Biology*)
91. Trikha P., Sharma N., Pena C., Reyes R. and **Leone G.** Ablation of *E2f3* in macrophages reduces the incidence of lung metastasis. (in preparation; *Cancer Research*)

#### D. Research Support.

##### Current:

- |    |                                     |  |              |
|----|-------------------------------------|--|--------------|
| 1. | 06/01/09-04/30/14                   | R01CA098956-07 (U. of Pittsburgh subcontract)<br>Title: Expression of Simian Virus 40 T Antigens in Intestine<br>PI: James Pipas, PhD<br><b>Co-Investigator, Gustavo Leone PhD</b> | \$45,000/yr  |
| 2. | 12/10/07-11/30/12                   | NIH/NCI 1 R01 CA121275 \$2<br>Title: E2F7 & E2F8 in the control of transcription and cellular proliferation<br><b>PI: Gustavo Leone, PhD</b>                                       | 50,000/yr    |
| 3. | 09/15/04-07/31/11                   | NCI P01 CA97189 \$2<br>Title: Genetic Analysis of the breast tumor microenvironment.<br>PI: Michael Ostrowski, PhD<br><b>Project 2 Leader: Gustavo Leone, PhD</b>                  | 35,000/yr    |
| 4. | 08/01/04-07/31/14<br>Titl<br>D      | NCI T32 CA106196<br>e: Postdoctoral Training in Cancer Genetics<br>irector: Albert de la Chapelle, MD, PhD<br><b>Co-Directors: Gustavo Leone, PhD and Joanna Groden, PhD</b>       | \$448,323/yr |
| 5. | 07/01/10-06/31/15<br><br><b>PI:</b> | OSUCCC Pelotonia Fellowship Program \$2,<br>Title: Training Program in Cancer Biology<br><b>Gustavo Leone, PhD</b>   | 650,000/yr   |

##### Completed:

- |    |  |   |           |
|----|--|---|-----------|
| 1. | 07/01/99-06/30/01<br><br>an<br><b>PI:</b>      | The V Foundation Scholar Award<br>Title: Role of the E2F3 gene locus in the control of proliferation and tumorigenesis.<br><b>Gustavo Leone</b>                         | \$200,000 |
| 2. | 10/01/99-04/30/04<br><br><b>PI:</b>            | NIH/NCI R01 CA82259 \$6<br>Title: Role of E2F in the control of apoptosis<br><b>Gustavo Leone</b>   | 81,000    |
| 3. | 10/01/99-04/30/04<br><br><b>PI:</b>            | NIH/NCI R01 Supplement to CA82259 S1<br>Title: Role of E2F in the control of apoptosis<br><b>Gustavo Leone</b>  | \$123,000 |
| 4. | 03/01/01-02/28/04<br><br>breast<br><b>PIs:</b> | The V Foundation Group Grant<br>Title: Genetic analysis of the role of the microenvironment in tumor progression<br><b>Gustavo Leone, Michael Ostrowski, Charis Eng</b> | \$300,000 |
| 5. | 01/25/01-12/31/05                              | NIH/NCI 5 R01 CA85619 \$9<br>Title: E2F3 gene locus in the control of cellular proliferation  | 35,000    |

<b>PI:</b>	<b>Gustavo Leone</b>	
6. 04/01/01-03/30/04	6. NIH/NCI R01 Supplement to R01 CA85619 S1 Title: E2F3 gene locus in the control of cellular proliferation	\$136,000
<b>PI:</b>	<b>Gustavo Leone</b>	
7. 07/01/01 – 06/30/05	Pew Scholar Program \$2 Title: Role of E2F within the microenvironment of breast cancer	50,000
<b>PI:</b>	<b>Gustavo Leone</b>	
8. 06/01/02-05/31/07	NIH/NICHD 1 R01 HD42619 \$1, Title: Function of the tumor suppressor Retinoblastoma during placental and embryonic development PI: Michael Weinstein, Ph.D.	090,000
<b>C</b>	<b>o-PI: Gustavo Leone</b>	
9. 06/01/03-05/31/07	March of Dimes Birth Defects Foundation Title: Cell cycle control in the developing placenta <b>PI: Gustavo Leone</b>	\$400,000
10. 04/01/03-03/30/08	NIH/NCI K01 Grant \$6 Title: “Howard Temin Award” PI: Lizhao Wu	24,523
<b>Mentor:</b>	<b>Gustavo Leone</b>	
11. 04/01/03 – 03/30/08	NIH/NCI K01 \$ Title: “Minority Transition Grant” PI: Harold I. Saavedra	469,597
<b>Mentor:</b>	<b>Gustavo Leone</b>	
12. 04/01/04-03/30/09	NIH/NICHHD RO1 HD047470 \$1, Title: E2F3 and embryonic development	012,500
<b>PI:</b>	<b>Gustavo Leone</b>	
13. 07/22/04-04/30/08	NCI /5 R01 CA098933-03 \$1, Title: BAALC in neuropoiesis and hematopoiesis PI: Albert de la Chapelle <b>Co-I: Gustavo Leone</b>	001,000
14. 07/01/02-06/30/05	American Cancer Society \$1 Title: E2Fs in the control of cellular proliferation <b>Mentor: Gustavo Leone, Trainee: Cynthia Timmers, PhD.</b>	45,000
15. 10/01/06-10/30/09	Department of Defense – DOD-BC030892 Title: The role of Retinoblastoma within the microenvironment of Breast Cancer	\$442,500
<b>Mentor:</b>	<b>Gustavo Leone, Trainee: Alain de Bruin, DVM, PhD</b>	
16. 05/01/06-04/30/08	Komen Breast Cancer Foundation PDF0504264 Title: The role of E2f3 in breast cancer <b>Mentor: Gustavo Leone, Trainee: Shui-Huei Wang, PhD.</b>	\$130,000
17. 07/01/06-06/30/09	American Cancer Society PF-06-240-010 LIB \$1	45,000

lym	Title: E2F2b, a novel tumor suppressor in Myc-induced lymphomagenesis <b>Mentor: Gustavo Leone</b> , Trainee: Hui Wang, PhD.	
18. 10/01/06-10/30/09	Department of Defense – DOD-BC061730 Title: PTEN in stromal fibroblasts suppresses development of mammary epithelial tumors <b>Mentor: Gustavo Leone</b> , Trainee: Leon Chong, PhD	\$390,000
19. 01/01/00-12/31/09	Up on the Roof/Faculty/OSUCCC \$1, Title: Postdoctoral Training Program in Cancer Biology Research <b>PI: Gustavo Leone</b>	450,000
20. 07/01/04-06/30/09	NCI T32 CA106196 \$1, Title: Postdoctoral Training in Cancer Genetics Director: Albert de la Chapelle, MD, PhD <b>Co-Directors: Gustavo Leone, PhD</b> and Christoph Plass, PhD	346,015
21. 12/31/04-12/31/09	Lymphoma & Leukemia Society Scholar Program Title: Mechanisms of MYC-induced Lymphomagenesis <b>PI: Gustavo Leone, PhD</b>	\$100,000/yr

## Personnel Listing

<b>Employee Name</b>	<b>Title</b>
Leone,Gustavo W.	Associate Professor
Saltz,Joel H.	Chair
Ababneh,Sufyan	Biomedical Informatics Sp
Bae,Soo In	Post Doctoral Researcher
Bronisz-Kowalczyk,Agnieszka M	Research Scientist
Cantemir-Stone,Carmen Zenaida	Research Scientist
Chen,Hui-Zi	Graduate Research Associate
Gurcan,Metin Nafi	Associate Professor
Huang,Kun	Associate Professor
Machiraju,Raghu	Associate Professor
Sharma,Ashish	Research Assistant Professor
Sharma,Nidhi	Post Doctoral Fellow



## ARTICLES

# *Pten* in stromal fibroblasts suppresses mammary epithelial tumours

Anthony J. Trimboli<sup>1,2,\*</sup>, Carmen Z. Cantemir-Stone<sup>3,\*</sup>, Fu Li<sup>1,3,\*</sup>, Julie A. Wallace<sup>3</sup>, Anand Merchant<sup>3</sup>, Nicholas Creasap<sup>1,2</sup>, John C. Thompson<sup>1,2</sup>, Enrico Caserta<sup>1,2</sup>, Hui Wang<sup>1,2</sup>, Jean-Leon Chong<sup>1,2</sup>, Shan Naidu<sup>1,2,4</sup>, Guo Wei<sup>1,3</sup>, Sudarshana M. Sharma<sup>3</sup>, Julie A. Stephens<sup>5</sup>, Soledad A. Fernandez<sup>5</sup>, Metin N. Gurcan<sup>6</sup>, Michael B. Weinstein<sup>1,2</sup>, Sanford H. Barsky<sup>7,†</sup>, Lisa Yee<sup>8</sup>, Thomas J. Rosol<sup>4</sup>, Paul C. Stromberg<sup>4</sup>, Michael L. Robinson<sup>9,†</sup>, Francois Pepin<sup>10,11</sup>, Michael Hallett<sup>10,11</sup>, Morag Park<sup>10,12</sup>, Michael C. Ostrowski<sup>3,13</sup> & Gustavo Leone<sup>1,2,13</sup>

The tumour stroma is believed to contribute to some of the most malignant characteristics of epithelial tumours. However, signalling between stromal and tumour cells is complex and remains poorly understood. Here we show that the genetic inactivation of *Pten* in stromal fibroblasts of mouse mammary glands accelerated the initiation, progression and malignant transformation of mammary epithelial tumours. This was associated with the massive remodelling of the extracellular matrix (ECM), innate immune cell infiltration and increased angiogenesis. Loss of *Pten* in stromal fibroblasts led to increased expression, phosphorylation (T72) and recruitment of *Ets2* to target promoters known to be involved in these processes. Remarkably, *Ets2* inactivation in *Pten* stroma-deleted tumours ameliorated disruption of the tumour microenvironment and was sufficient to decrease tumour growth and progression. Global gene expression profiling of mammary stromal cells identified a *Pten*-specific signature that was highly represented in the tumour stroma of patients with breast cancer. These findings identify the *Pten*–*Ets2* axis as a critical stroma-specific signalling pathway that suppresses mammary epithelial tumours.

Coordinated signalling between different cell types of the ‘normal stroma’ is required during embryonic and adult development<sup>1</sup>. The stroma can be appropriately activated in response to extreme but normal physiological cues, such as wounding, inflammation or pregnancy<sup>2</sup>. The stroma can also be inappropriately activated in cancer<sup>3,4</sup>. In breast tumours, stromal fibroblasts are believed to adapt and continuously co-evolve with tumour epithelial cells to foster transformation and tumour growth<sup>5</sup>. Fibroblasts are a principal constituent of the stroma responsible for the synthesis of growth and survival factors, angiogenic and immunological chemokines, and structural components of the ECM as well as enzymes that control its turnover<sup>6,7</sup>. Despite extensive evidence for a role of the tumour stroma in carcinogenesis, relatively little is known about the signalling pathways involved in the communication between the different cellular compartments of the microenvironment that contribute to the cancer phenotype.

Alterations in the phosphoinositide 3-kinase (PI3K) pathway are associated with the activation of tumour-associated stroma<sup>8,9</sup>. One of the main regulators of PI3K signalling is the phosphatase and tensin homologue (PTEN), a tumour suppressor with lipid and protein phosphatase activity<sup>10,11</sup>. *PTEN* inactivation disrupts multiple cellular processes associated with cell polarity, cell architecture, chromosomal integrity, cell cycle progression, cell growth and stem cell self-renewal<sup>12,13</sup>. Germ-line inactivation of a single allele of *PTEN*

in both humans and mice contributes to the genesis of a variety of tumour types of epithelial origin<sup>14</sup>. Although tremendous progress in understanding PTEN function in tumour cells has been made since its discovery over a decade ago, relatively little is known about its potential role in the tumour stroma. Here, we show that *Pten* ablation in mammary stromal fibroblasts of mice results in massive remodelling of the ECM and tumour vasculature, recruitment of innate immune cells and increased malignancy of mammary epithelial tumours. Gene expression profiling of *Pten*-deleted stromal fibroblasts identified the activation of an *Ets2*-specific transcription program associated with many of these aggressive tumour phenotypes. Remarkably, the concomitant inactivation of *Ets2* in the mammary stroma reversed the increased malignancy caused by *Pten* deficiency. These findings expand *Pten*'s repertoire as a tumour suppressor by identifying the fibroblast as a key site from which it exerts its powerful tumour suppressive influence on the adjacent tumour epithelium.

## Stromal *Pten* suppresses mammary epithelial tumours

To evaluate rigorously the role of *Pten* in the tumour microenvironment of breast cancer, we generated mice containing a mesenchymal-specific *Fsp-cre* transgene<sup>15</sup> and conditional alleles of *Pten* (*Pten*<sup>loxP</sup>; Supplementary Fig. 1). Cell-type-marker analysis using a  $\beta$ -galactosidase *Rosa26*<sup>LoxP</sup> reporter allele showed specific *Fsp-cre* expression in stromal fibroblasts surrounding the mammary epithelial ducts, with

<sup>1</sup>Department of Molecular Genetics, College of Biological Sciences, <sup>2</sup>Department of Molecular Virology, Immunology and Medical Genetics, <sup>3</sup>Department of Molecular and Cellular Biochemistry, College of Medicine, <sup>4</sup>Department of Veterinary Biosciences, College of Veterinary Medicine, <sup>5</sup>Center for Biostatistics, Office of Health Sciences, <sup>6</sup>Department of Biomedical Informatics, <sup>7</sup>Department of Pathology and <sup>8</sup>Department of Surgery, School of Medicine, The Ohio State University, Columbus, Ohio 43210, USA. <sup>9</sup>Center for Molecular and Human Genetics, Columbus Children's Research Institute, Columbus, Ohio 43205, USA. <sup>10</sup>Department of Biochemistry, Rosalind and Morris Goodman Cancer Center, <sup>11</sup>McGill Center for Bioinformatics, <sup>12</sup>Department of Oncology, McGill University, Québec H3A 1A1, Canada. <sup>13</sup>Tumor Microenvironment Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210, USA. †Present addresses: Department of Pathology, University of Nevada School of Medicine, Reno, Nevada 89557, and Nevada Cancer Institute, Las Vegas, Nevada 89135, USA (S.H.B.); Department of Zoology, Miami University, Oxford, Ohio 45056, USA (M.L.R.).

\*These authors contributed equally to this work.

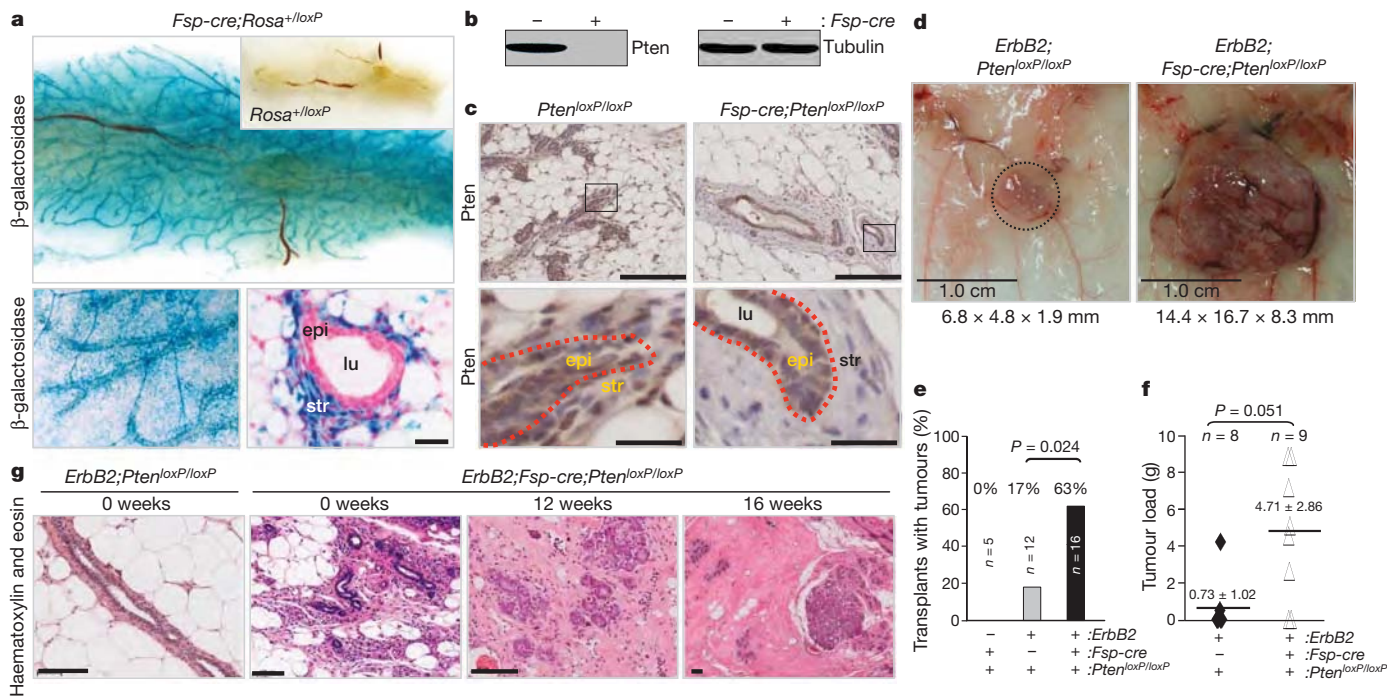
no expression in cytokeratin-positive epithelial cells, F4/80-positive macrophages and CD31-positive endothelial cells (Fig. 1a and Supplementary Fig. 2a, b). Western blot and PCR assays demonstrated efficient cre-mediated deletion of *Pten*<sup>loxP</sup> in stromal fibroblasts isolated from *Fsp-cre;Pten*<sup>loxP/loxP</sup> mammary glands (Fig. 1b and Supplementary Fig. 3a). Examination of mammary sections by immunohistochemistry (IHC) and immunofluorescence showed deletion of *Pten*<sup>loxP</sup> that was confined to stromal fibroblasts, with no collateral deletion in epithelial ducts or the adjacent myoepithelium (Fig. 1c and Supplementary Fig. 3b, c). Interestingly, this resulted in the expansion of the ECM, but did not lead to the transformation of the mammary epithelium (Fig. 1c, e).

We then examined the role of stromal *Pten* on mammary tumorigenesis using an established mouse model of breast cancer, *MMTV-ErbB2/neu* (*ErbB2*)<sup>16</sup>. To avoid possible confounding effects caused by *Pten* deletion in mesenchymal cells of other organs, mammary glands from *Fsp-cre;Pten*<sup>loxP/loxP</sup>, *ErbB2;Pten*<sup>loxP/loxP</sup> and *ErbB2;Fsp-cre;Pten*<sup>loxP/loxP</sup> donors were transplanted into syngeneic wild-type recipients<sup>17</sup> and tumour development was monitored over the course of several months. By genetically marking the stroma with the *Rosa26*<sup>LoxP</sup> reporter allele, we demonstrated that both the epithelium and its associated stroma were effectively transplanted into host female mice (Supplementary Fig. 4). Loss of *Pten* in stromal fibroblasts dramatically increased the incidence of *ErbB2*-driven mammary tumours (Fig. 1d–f). By 16 weeks post-transplantation, these lesions progressed to adenoma, carcinoma *in situ* and invasive carcinoma (Fig. 1g), and by 26 weeks most females met the criteria for early removal due to excessive tumour burden (Fig. 1f). Histological examination showed that *ErbB2*-tumour cells in *Pten* stromal-deleted tumours retained their typical oncogene-specific morphology, with

small nuclei, fine chromatin and abundant eosinophilic cytoplasm<sup>18</sup>. In contrast to non-deleted tumours<sup>18,19</sup>, *Pten* stromal-deleted tumours had a significant amount of stroma surrounding and infiltrating the epithelial masses (Fig. 1g). PCR-based and immunohistochemical assays confirmed that tumours had intact *Pten*<sup>loxP</sup> alleles in the epithelial compartment (Supplementary Fig. 5a, b and data not shown). Moreover, we used the *Rosa26*<sup>LoxP</sup> reporter allele to mark genetically early epithelial to mesenchymal transition events<sup>15</sup> and found no evidence of epithelial to mesenchymal transition in tumours that either contained or lacked *Pten* in stromal fibroblasts (data not shown). Thus, the analysis of the *ErbB2* breast-cancer tumour model identified a potent tumour suppressor role for *Pten* in stromal fibroblasts of the mammary gland.

**Stromal *Pten* controls ECM and innate immune functions**

To investigate the tumour suppressive mechanism of *Pten* action in stromal fibroblasts, we profiled the transcriptome of mammary stromal fibroblasts isolated from *Pten*<sup>loxP/loxP</sup> and *Fsp-cre;Pten*<sup>loxP/loxP</sup> females. Details of sample collection, processing of Affymetrix oligo-arrays and expression data are available in Methods. Briefly, we implemented class comparison analyses of all probe sets on the Affymetrix mouse genome 430 2.0 array to identify genes differentially expressed between the two genetic groups. We also used an unbiased approach similar to gene set enrichment analysis<sup>20</sup> to identify *a priori* defined groups of genes that were significantly differentially expressed. The analysis of over 14,000 mouse genes identified 129 upregulated and 21 downregulated unique genes in response to *Pten* deletion (Supplementary Fig. 6a, b; greater than fourfold at *P* < 0.001; Supplementary Tables 1 and 2). Reverse transcription followed by quantitative PCR (quantitative RT-PCR) assays on a subset of genes confirmed more than 85% of these expression



**Figure 1 | Stromal fibroblast-specific deletion of *Pten*.** **a**, Wholemount, X-gal-stained mammary glands from *Fsp-cre;Rosa*<sup>+/loxP</sup> and *Rosa*<sup>+/loxP</sup> (top, inset) mice. Higher magnification of wholemount gland (bottom left) and a histological cross section (bottom right); scale bar, 30  $\mu$ m. lu, Lumen; epi, epithelium; str, stroma. **b**, Representative western blot analysis of mammary fibroblast lysates derived from 8-week-old *Pten*<sup>loxP/loxP</sup> mice with (+) or without (-) *Fsp-cre*. **c**, Paraffin sections from 8-week-old female mammary glands stained with a *Pten*-specific antibody; lower panels represent higher magnifications of boxed areas; scale bars: top panels, 200  $\mu$ m; bottom panels, 30  $\mu$ m. lu, Lumen; epi, epithelial compartment; str, stromal compartment; red dotted line indicates the border between the two compartments.

**d**, Tumours collected at 26 weeks post-transplantation. **e**, Tumour development by 16 weeks in mammary glands with the indicated genotypes. Tumorigenicity was determined by palpation or histological presentation of adenoma/carcinoma at each implantation site and statistically analysed using Fisher's exact test. *n*, Total number of transplants. **f**, Total tumour burden at 26 weeks post-transplantation in mammary glands with the indicated genotypes. Values represent mean  $\pm$  s.d. Differences were tested using the non-parametric Wilcoxon rank sum test. **g**, H&E-stained sections of mammary glands harvested at time of transplantation (0 weeks) and indicated times post-transplantation; scale bars, 100  $\mu$ m.

changes using independent fibroblast samples (Supplementary Fig. 6c and Supplementary Table 3). Fibroblast samples used to probe the oligo-arrays lacked expression of macrophage-, endothelial- and epithelial-specific genes, confirming the purity of these fibroblast preparations (Supplementary Fig. 6d). Functional annotation<sup>21,22</sup> of Pten-responsive targets revealed a remarkable bias towards genes encoding proteins involved in ECM remodelling, wound healing and chronic inflammation<sup>21,22</sup> (Fig. 2a and Supplementary Tables 1 and 2). Given this unexpected convergence of function, we performed a more thorough cellular and molecular analysis of *Pten*-deleted stroma. Staining of consecutive mammary gland sections with haematoxylin and eosin (H&E) and Mason's trichrome stains indicated enhanced deposition of collagen in *Pten*-deleted stroma, which was independent of *ErbB2*-oncogene expression (Fig. 2b, c and Supplementary Fig. 7a). IHC and western blot assays using collagen type-specific antibodies showed that the non-cellular material consisted mostly of type-I collagen and not the basement membrane type-IV collagen (Fig. 2b, c and Supplementary Fig. 7b, c). There was significant infiltration of F4/80-positive macrophages into stromal *Pten*-deleted mammary glands (Fig. 2d, e), and this was independent of *ErbB2* expression (Supplementary Fig. 8a). The abundance of B and T cells did not change in response to stromal deletion of *Pten* (data not shown). From these experiments, we conclude that ablation of *Pten* in stromal fibroblasts recapitulates two key events associated with tumour malignancy: increased ECM deposition and innate immune cell infiltration.

### Stromal *Pten* loss activates *Ets2*

Along with the remarkable remodelling of the tumour microenvironment, loss of *Pten* in stromal fibroblasts resulted in activation of the Ras, JNK and Akt pathways. Western blot analysis using protein lysates derived from *Pten*-deleted stromal fibroblasts demonstrated an increase in the phospho-specific forms of Akt (Akt<sup>T308/S473</sup>) and JNK (JNK<sup>T183/Y185</sup>) (Fig. 2f and Supplementary Fig. 8b). Immunohistochemical assays confirmed the activation of Akt and JNK in stromal fibroblasts, and, interestingly, revealed a profound activation of these two pathways in ductal epithelial cells adjacent to the *Pten*-deleted stroma (Fig. 2g and Supplementary Fig. 8b). This analysis also showed increased levels of phospho-Erk1/2 in *Pten*-deleted stromal fibroblasts; however, this increase could not be detected in primary cultured fibroblasts (Fig. 2f, g), presumably owing to the constitutive Pten-independent activation of Erk1/2 by serum-stimulation<sup>23</sup>.

Among the many expression changes observed in *Pten*-deleted stromal fibroblasts we noted that there was a significant increase of *Ets2* messenger RNA (mRNA) levels (2.8-fold,  $P < 0.001$ ). This induction is notable because the *Ets2* transcription factor is known to be transcriptionally induced by MAPK<sup>24,25</sup> activation and its function to be post-translationally enhanced by the Akt- and JNK-mediated phosphorylation of its pointed domain at threonine 72 (*Ets2*<sup>T72</sup>)<sup>23,26</sup>. We confirmed the higher levels of *Ets2* mRNA and protein in *Pten*-deleted fibroblasts (approximately threefold,  $P < 0.001$ ; Supplementary Fig. 9a, b) and, consistent with the activation of Akt and JNK in these mammary glands, there was a marked increase of phospho-*Ets2*<sup>T72</sup> in stromal fibroblasts and adjacent epithelial ducts (Fig. 2h, i). Loss of *Pten* in stromal fibroblasts resulted in the induction of several genes involved in ECM remodelling and macrophage recruitment, two of which, *Mmp9* and *Ccl3*, are known to be direct transcriptional targets of *Ets2*<sup>27,28</sup> (Supplementary Figs 6c and 9c). The increase of *Mmp9* expression appears to be of pathological relevance because *in situ* zymography<sup>29</sup> showed robust *Mmp9* activity in tumour samples (Supplementary Fig. 9d). Chromatin immunoprecipitation (ChIP) assays showed an increase in the loading of *Ets2* onto the *Mmp9* and *Ccl3* promoters in *Pten*-deleted mammary fibroblasts (Supplementary Fig. 9e), suggesting a direct role for *Ets2* in the transcriptional regulation of these two target genes *in vivo*. Together, these data illustrate the extensive molecular reprogramming that takes place in the tumour and its microenvironment in response to ablation of *Pten* in stromal fibroblasts.

### Stromal *Ets2* promotes mammary tumorigenesis

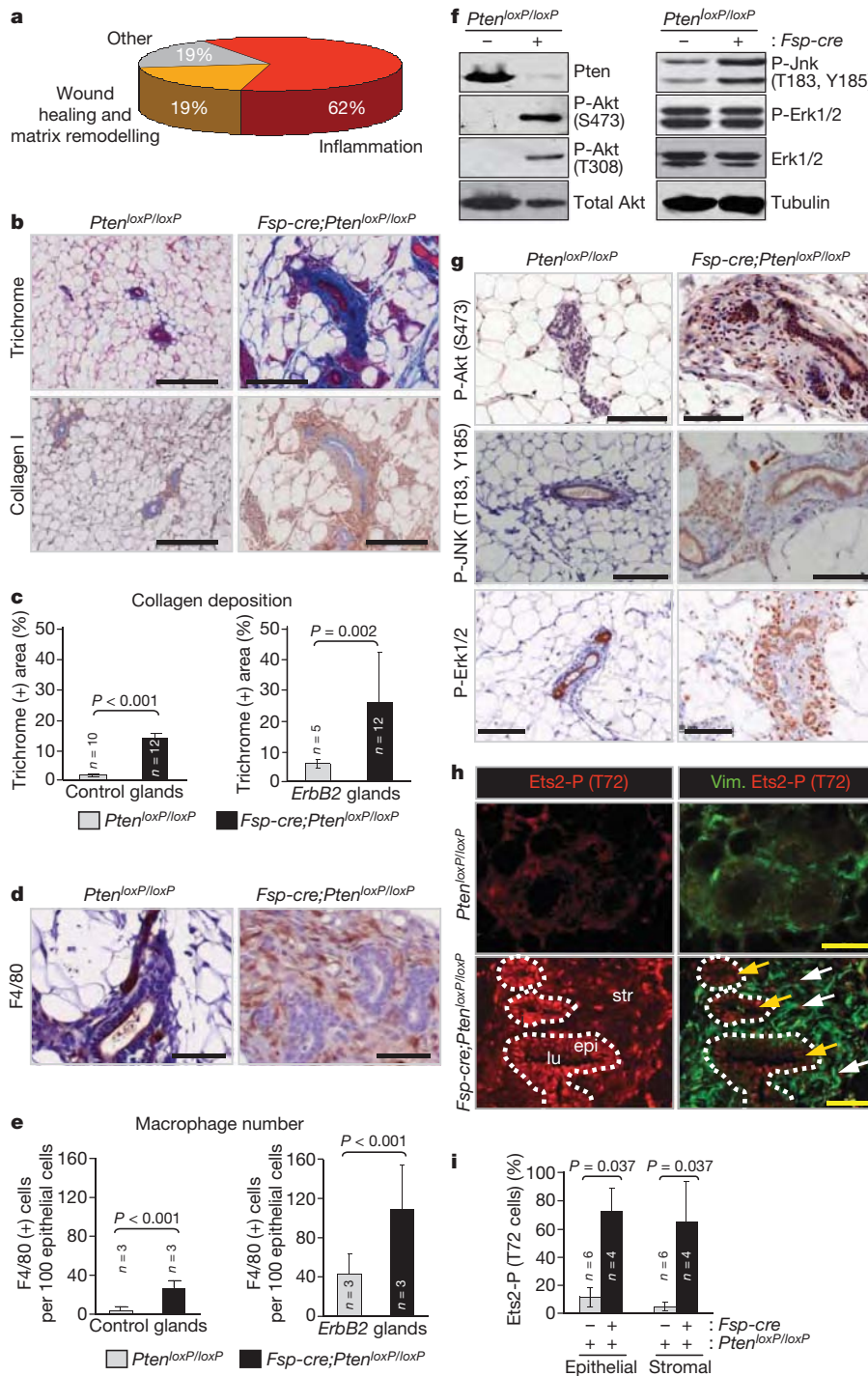
To determine whether *Ets2* promotes a microenvironment conducive to tumour growth, we analysed the consequences of ablating a conditional allele of *Ets2* (*Ets2*<sup>loxP</sup>)<sup>30</sup> in mammary stromal fibroblasts of a well-characterized mouse model of breast cancer, *MMTV-PyMT* (*PyMT*)<sup>31</sup>. The *PyMT* oncogene initiates the rapid onset and progression of mammary tumours and thus represents an appropriate model for evaluating any potential delay that loss of *Ets2* might have on tumorigenesis. The efficient *Fsp-cre*-mediated ablation of *Ets2* in stromal fibroblasts was facilitated by using mice carrying conventional and conditional knockout alleles of *Ets2* (DNA binding domain *Ets2*<sup>db/loxP</sup>)<sup>32</sup> (Fig. 3a and Supplementary Fig. 10a, b). Ablation of *Ets2* in these cells had no detectable physiological consequence on the development of mammary glands, either during puberty or pregnancy (M.C.O., unpublished observations). The evaluation of *PyMT*;*Fsp-cre*;*Ets2*<sup>db/loxP</sup> and control *PyMT*;*Ets2*<sup>db/loxP</sup> mice over a period of three months showed that ablation of *Ets2* in mammary fibroblasts significantly reduced the tumour load (Fig. 3b) and slowed progression to adenoma and early carcinoma (Fig. 3c). Quantitative RT-PCR showed high levels of *Mmp9* expression in tumour-associated fibroblasts containing *Ets2* and low levels in *Ets2*-deleted fibroblasts (Fig. 3e). Because *Mmp9* activity is known to mediate the release of matrix-bound VEGF-A to its active isoforms, including VEGF<sub>164</sub> (ref. 33), we visualized the spatial distribution of VEGF<sub>164</sub> and *Mmp9* activity by immunofluorescence. These assays showed that the accumulation of VEGF<sub>164</sub>, which was particularly acute within collagen-1A-rich stromal locations overlapping *Mmp9* activity, was significantly decreased in stromal-deleted *Ets2* tumours (Fig. 3d, f). Given that VEGF<sub>164</sub> is a specific ligand for VEGF Receptor 2 (VEGFR2; FLK-1; KDR), one of the most potent mediators of VEGF-induced endothelial signalling and angiogenesis<sup>34</sup>, we also evaluated VEGFR2 status by immunostaining tumour sections with antibodies specific for CD31 and the phospho-activated form of the murine VEGF receptor (VEGFR2<sup>Y1173</sup>)<sup>35</sup>. This analysis revealed a fourfold decrease in the number of CD31/VEGFR2<sup>Y1173</sup> double-positive cells in *Ets2*-deleted versus non-deleted tumour samples (Fig. 3f, g). Together, these data show that loss of *Ets2* in stromal fibroblasts results in decreased *Mmp9* activity in the tumour ECM and reduced VEGFR2<sup>Y1173</sup> activation in the tumour vasculature.

### *Ets2* is a key component of the stromal *Pten* axis

We then entertained the hypothesis that *Ets2* may be contributing to the remodelling of the tumour microenvironment caused by stromal deletion of *Pten*. To test this possibility directly, we compared tumour incidence in *Pten*<sup>loxP/loxP</sup>, *Fsp-cre*;*Pten*<sup>loxP/loxP</sup> and *Fsp-cre*;*Pten*<sup>loxP/loxP</sup>;*Ets2*<sup>db/loxP</sup> mammary glands that were orthotopically injected with an established *ErbB2*-initiated mammary tumour cell line (NT2.5)<sup>36</sup>. This orthotopic model recapitulated the consequences of deleting *Pten* in the mammary stroma that were observed in the genetically engineered *ErbB2*-mouse model described earlier in this study. Indeed, tumour incidence and tumour load in injected *Fsp-cre*;*Pten*<sup>loxP/loxP</sup> females was markedly higher than in control *Pten*<sup>loxP/loxP</sup> females (Fig. 4a, b). Importantly, mammary glands doubly deleted for stromal *Pten* and *Ets2* had fewer and smaller tumours than glands deleted for *Pten* only. These mammary tumours had decreased numbers of macrophages and recruitment of new vasculature (Fig. 4c–f). Loss of *Pten* and *Ets2*, however, failed to reduce the tumour load and collagen deposition fully to control levels (Fig. 4b and Supplementary Fig. 11a, b), suggesting that additional effectors must contribute towards *Pten*'s tumour suppressor functions. From these data, we conclude that *Ets2* is a major component of the *Pten* tumour-suppressive axis that acts in the stromal fibroblast compartment of mammary glands.

### Human breast tumour stroma contains a *Pten* signature

To determine the relevance of these findings to human breast cancer, we compared the mouse stromal fibroblast *Pten* expression signature

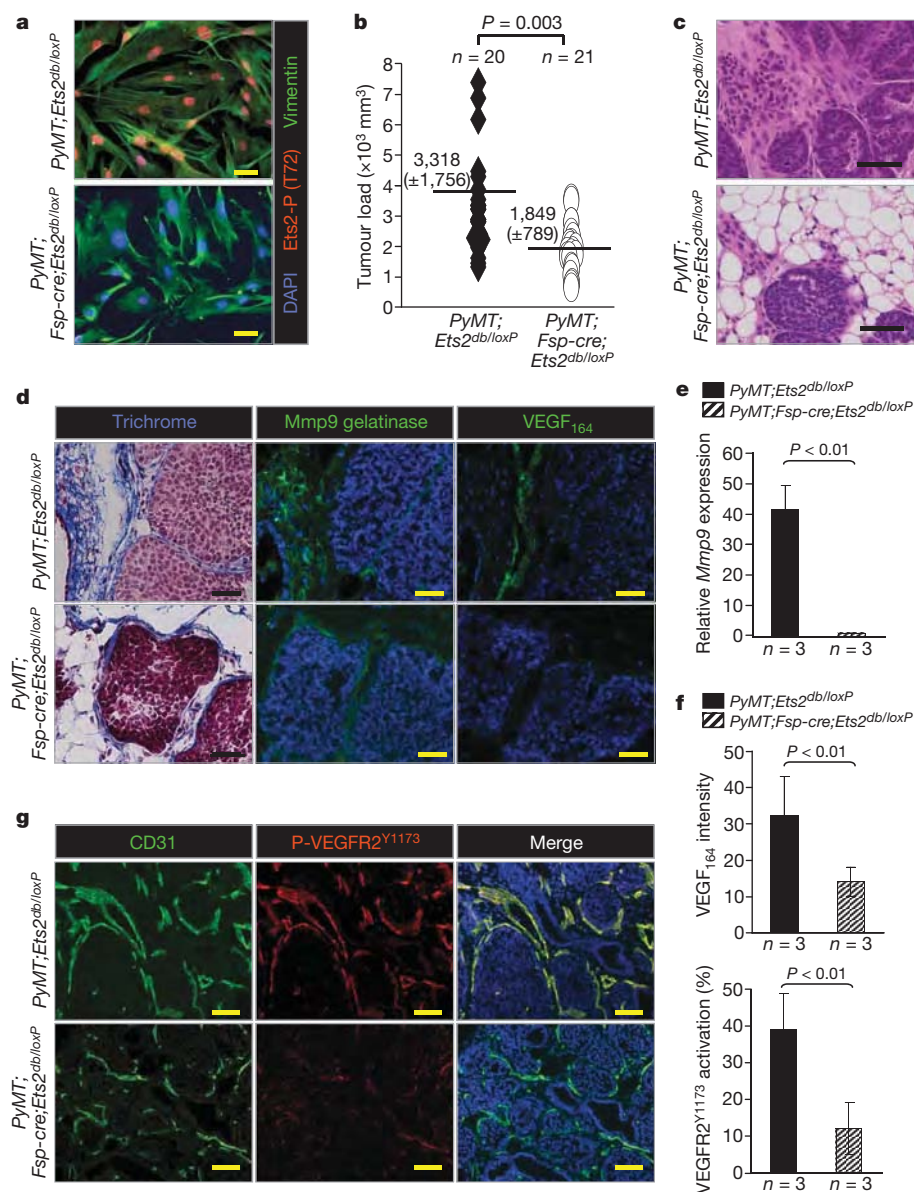


**Figure 2 | Characterization of ECM deposition and immune cell infiltration.**  
**a**, Schematic representation of the biological processes affected by differentially expressed genes (more than fourfold) in *Pten*-deleted stromal fibroblasts. **b**, Mammary gland paraffin sections stained with Masson's trichrome and collagen-I-specific antibodies, respectively; scale bars, 200  $\mu$ m. **c**, Trichrome-stained sections were quantified for collagen deposition; mammary glands in the absence (left) or presence (right) of *ErbB2* were analysed, respectively. Values are means  $\pm$  s.d.; the Wilcoxon rank sum test was used to compare groups. **d**, Mammary gland paraffin sections stained with the macrophage-specific marker F4/80; scale bars, 50  $\mu$ m. **e**, Quantification of F4/80-positive stained stromal cells in mammary glands in the absence (left) or presence (right) of *ErbB2*, respectively. Values are means  $\pm$  s.d.; the Wilcoxon rank sum test was used to compare groups.

**f**, Western blot analysis of whole-cell lysates derived from mammary stromal fibroblasts. **g**, Mammary gland paraffin sections stained with phospho-Akt<sup>S473</sup>-, phospho-JNK<sup>T183/Y185</sup>- and phospho-Erk1/2-specific antibodies; scale bars, 100  $\mu$ m. All analyses were performed using tissue or cells from 8-week-old females. **h**, Frozen mammary tissue sections stained with vimentin (green) and phospho-Ets2<sup>T72</sup>-specific (red) antibodies. Note that loss of *Pten* in the mammary stroma increased Ets2 phosphorylation in both the stromal and epithelial compartments. Dotted white line indicates the stromal-epithelial boundary. lu, Lumen; epi, epithelium; str, stroma; scale bars, 50  $\mu$ m. **i**, Quantification of mammary epithelial and stromal cells that stained positive for nuclear phospho-Ets2<sup>T72</sup>. Values are means  $\pm$  s.d.; the Wilcoxon rank sum test was used to compare groups.

to the expression signatures derived from laser-captured tumour stroma (49 samples) and adjacent normal stroma (52 samples) in patients with breast cancer<sup>37</sup>. This analysis identified 137 human

orthologues from the 150 differentially expressed mouse genes detected by the Affymetrix oligo-arrays shown in Supplementary Fig. 6a, b. Of these 137 orthologues, 129 genes were represented in



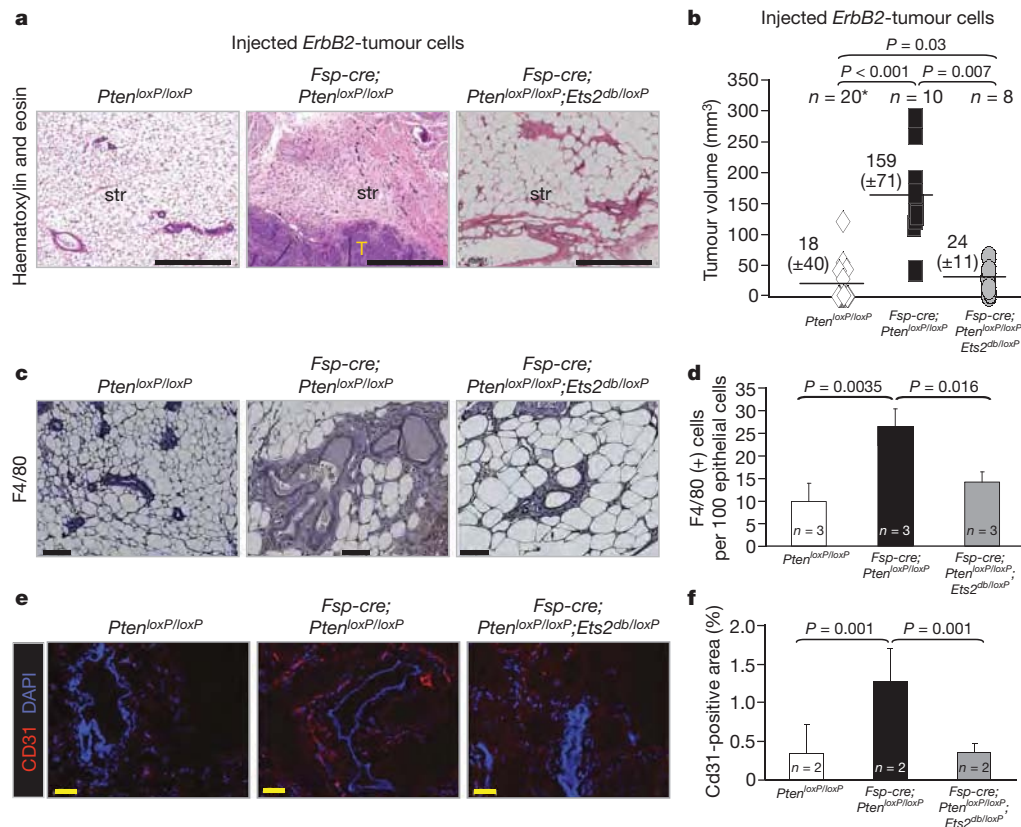
**Figure 3 | Ets2 ablation in stromal fibroblasts restricts mammary tumorigenesis.** **a**, Immunofluorescence staining of cultured mammary fibroblasts with vimentin (green), phospho-Ets2<sup>T72</sup> (red) antibodies and counterstained with 4,6-diamidino-2-phenylindole (DAPI) (blue). **b**, Total mammary tumour volume of PyMT;Ets2<sup>db/loxP</sup> (n = 20) and PyMT;Fsp-cre;Ets2<sup>db/loxP</sup> (n = 21) mice collected 30 days after tumour initiation. Values represent the mean ± s.d. **c**, H&E staining of tumours harvested from PyMT;Ets2<sup>db/loxP</sup> or PyMT;Fsp-cre;Ets2<sup>db/loxP</sup> mice. **d**, Consecutive sections stained for (left to right): trichrome, Mmp9 gelatinase activity and VEGF<sub>164</sub>, and counterstained with DAPI from PyMT;Ets2<sup>db/loxP</sup> and PyMT;Fsp-cre;Ets2<sup>db/loxP</sup> mammary tumours. **e**, Quantification of Mmp9 mRNA

the expression platform used (Agilent) for the analysis of human patient stroma samples<sup>37</sup>. Only 70 of these 129 genes had highly variable gene expression across all human stromal samples (a variance cutoff of greater than 0.5). The heat map generated for the human stroma data set showed that this 70 gene subset derived from the mouse Pten-signature was sufficient to distinguish normal from tumour stroma in all patients (Fig. 5a;  $P = 8 \times 10^{-5}$  as determined by a permutation test). Principal component analysis also discriminated normal from tumour stroma perfectly ( $P < 1 \times 10^{-10}$ ; Supplementary Fig. 12). Interestingly, 12 of the 70 human orthologues identified by the Pten-signature (Fig. 5a: gene names highlighted in red; Supplementary Fig. 13) were previously shown to be differentially expressed in the tumour stroma of patients with breast cancer and to

expression by quantitative RT-PCR. **f**, Quantification of VEGF<sub>164</sub> immunofluorescence staining in tumour stroma (top) and tumour endothelial cells co-expressing CD31 and phospho-VEGFR2<sup>Y1173</sup> (bottom). **g**, Tumour vascular endothelial cells visualized by immunofluorescence double staining with CD31 (green) and p-VEGFR2<sup>Y1173</sup> (red), and counterstained with DAPI (blue) in mammary tumours collected 1 week after tumour initiation in PyMT;Ets2<sup>db/loxP</sup> and PyMT;Fsp-cre;Ets2<sup>db/loxP</sup> mice. All analyses used tissue or cells from 9- to 10-week-old females; scale bars, 50 μm. In **e** and **f**, bars represent means and error bars are s.d. Student's *t*-test was used for all the statistical comparisons between groups.

be associated with recurrence<sup>37</sup> (Fig. 5b;  $P = 2.5 \times 10^{-8}$ , Fisher's exact analysis). These analyses suggest that the fibroblast Pten-expression signature identified by our stromal mouse model represents a significant subset of the total gene signature expressed in the stroma of human breast cancer. We interpret these results to mean that a portion of the transcriptome regulated by Pten in mammary stromal fibroblasts is dysregulated in the tumour stroma of patients with breast cancer.

We also evaluated the relevance of the Pten–Ets2 relationship in human cancer by immunohistochemical staining of breast-cancer tissue microarrays (TMAs) with antibodies specific for Pten, P-Ets2<sup>T72</sup> and P-Akt<sup>S473</sup> (Fig. 5c). From the analysis of 99 patient samples with invasive carcinoma, PTEN expression was scored as



**Figure 4 | Loss of *Ets2* in stromal fibroblasts diminishes tumour growth in stromal *Pten*-deleted mammary glands.** **a**, H&E sections of mammary glands after orthotopic injection of the *ErbB2*-expressing tumour cell line NT2.5. T, tumour; str, stroma; scale bars, 500  $\mu$ m. **b**, Volumes of tumours collected 21 days after injection. *Pten*<sup>loxP/loxP</sup>;*Ets2*<sup>db/loxP</sup> (*n* = 10) and *Pten*<sup>loxP/loxP</sup> (*n* = 10) control groups were combined (\*, *n* = 20) after it was determined that there was no statistical difference in tumour incidence or load between these two control groups. Values are means  $\pm$  s.d. **c**, Sections from mammary glands with the indicated genotypes stained with the

macrophage-specific marker F4/80. Scale bars, 100  $\mu$ m. **d**, Quantification of stromal cells positive for F4/80 in mammary glands. **e**, Frozen mammary gland sections stained with the endothelial-specific antibody, CD31. Scale bars, 50  $\mu$ m. **f**, Quantification of CD31-positive staining. In **d** and **f**, bars represent means and error bars are s.d. For all the statistical analyses, an analysis of variance model with Bonferroni adjustment was used. Pairwise comparisons shown have a significant difference between marked genetic groups.

absent or low in approximately 50% of samples. Importantly, PTEN staining in the TMAs was negatively correlated with P-AKT<sup>S473</sup> and nuclear P-ETS2<sup>T72</sup>, whereas P-AKT<sup>S473</sup> and nuclear P-ETS2<sup>T72</sup> showed a positive association (Fig. 5d). These results suggest that activation of P-ETS2<sup>T72</sup> in human breast-cancer stroma is a pathological event that is favoured by a reduction in PTEN expression.

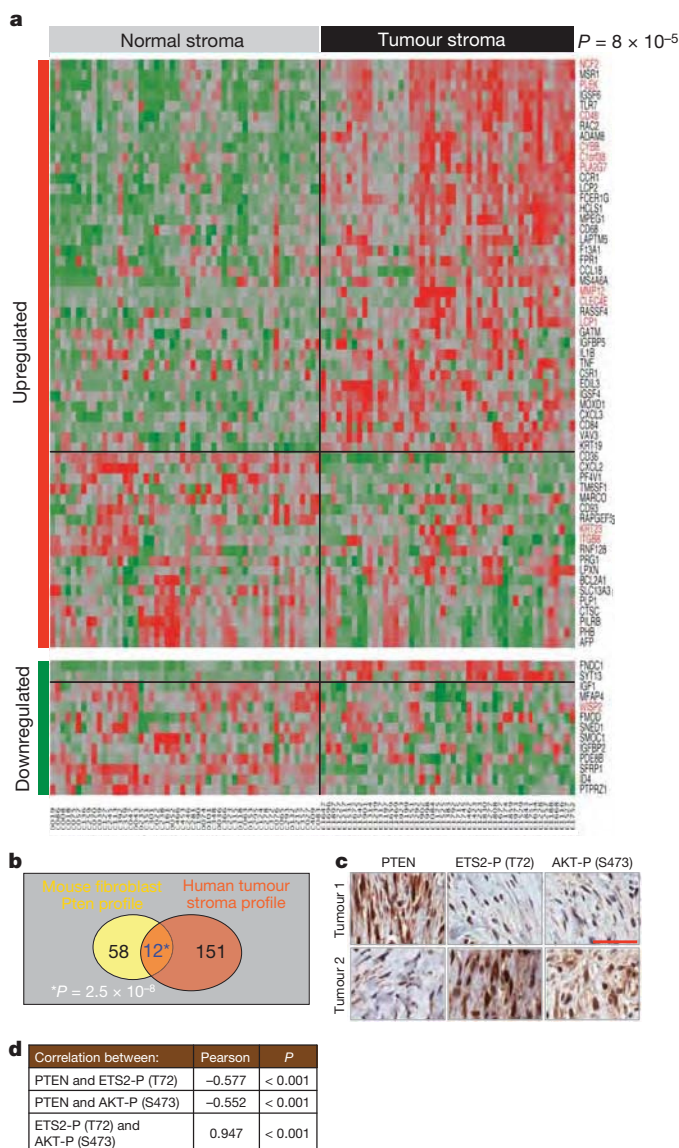
**Discussion**

Histopathology and molecular studies suggest that malignant tumours consist of a complex cellular system that is dependent on reciprocal signalling between tumour cells and the adjacent stroma. However, the signalling pathways that mediate the communication between the various cell types in the tumour remain virtually unknown. We recently developed a mesenchymal-specific *cre* mouse<sup>15</sup> and used it here to examine the consequences of inactivating *Pten* in mammary stromal fibroblasts. Using this system we show, for the first time to our knowledge, that *Pten* in stromal fibroblasts has a critical role in the suppression of epithelial mammary tumours that is, in part, mediated through an *Ets2*-regulated transcriptional program.

The tumour suppressor functions of PTEN have been extensively studied in the tumour cell<sup>38–40</sup>. We show here that genetic ablation of *Pten* in mammary stromal fibroblasts of mice alters the expression profile of these cells to increase ECM, chemokine and cytokine production in the tumour microenvironment. As a result, *Pten* stromal-deleted tumours exhibit high levels of collagen, macrophage recruitment and vascular networks, which together favour the initiation and progression of mammary epithelial tumours. Remarkably, side-by-side evaluation of histopathology by independent pathologists could not distinguish

tumours between *Pten* stromal-deleted mice and human patients with breast cancer, highlighting the importance of modelling stromal cell compartments of the tumour microenvironment. The mechanism by which *Pten* in the stroma likely exerts its tumour suppressor role is through the control of multiple signalling pathways, including components of the Ras, Akt and JNK networks, which together culminate in the regulation of *Ets2* transcriptional activity. The fact that loss of *Ets2* in mammary stromal fibroblasts diminished the oncogenic consequences of deleting *Pten* in these cells underscores the importance of the stromal *Pten*–*Ets2* axis in stromal fibroblasts during tumour suppression. These observations are consistent with previous work from Oshima and colleagues that showed a critical cell non-autonomous role for *Ets2* in the growth of mammary tumours in mice<sup>41</sup> and with the identification of *Ets2* activation as a key event associated with breast cancer in human patients having poor prognosis<sup>42–44</sup>. The relevance of this mouse *Pten*–*Ets2* tumour suppressor axis to breast cancer is underlined by the high correspondence between the mouse and human stromal expression signatures. The observation that the dire consequences of targeting this *Ets2*-driven stromal program are tumour-specific, sparing normal mammary development, emphasizes the potential use of stromal-specific strategies for therapeutic intervention in human breast cancer.

In summary, this work identifies *Pten*–*Ets2* as a key regulatory axis in stromal fibroblasts that suppresses mammary epithelial tumours by profoundly attenuating some of the most malignant characteristics of the tumour microenvironment. This novel function of *Pten* may be relevant in the suppression of epithelial tumours of other organs, but may also extend beyond cancer, to conditions where the microenvironment may impact disease manifestation, such as in autoimmune



**Figure 5 | Pten-signature is represented in breast cancer stroma.** **a**, Heat map displaying the expression of the human orthologues of the 70-gene subset in normal- and tumour-stroma from human patients with breast cancer. The 70-gene subset derived from the mouse Pten-signature includes 57 genes upregulated (red bar on y-axis) and 13 genes downregulated (green bar on y-axis). Red and green regions inside the heat map indicate relative gene expression levels (red, increased; green, decreased). The P value indicates the ability of the mouse 70-gene signature to partition normal and tumour stroma in patients with breast cancer (see Statistical Methods). The coded patient identities are listed at the bottom. **b**, Venn diagram depicting the overlap between the mouse Pten-deleted fibroblast and human stroma microarray data sets. The 12 genes highlighted in red (**a**, right margin) are common between the mouse Pten 70-gene signature and the human 163-gene signature that has been shown to associate with recurrence<sup>38</sup>. This overlap is highly significant ( $P = 2.5 \times 10^{-8}$ ; Fisher's exact test). **c**, Representative PTEN, P-ETS2<sup>T72</sup> and P-AKT<sup>S473</sup> IHC staining in human breast carcinoma samples from the tissue microarray. Scale bar, 50  $\mu\text{m}$ . **d**, Pearson correlations between PTEN, P-ETS2<sup>T72</sup> and P-AKT<sup>S473</sup> expression based on Allred scores of a tissue microarray (**c**) containing 99 patients with advanced breast carcinoma.

syndromes<sup>45</sup>, lung fibrosis<sup>46</sup> and neurodegeneration<sup>47</sup>. Interestingly, the stromal Pten expression signature identified here includes genes that have been causally linked to ECM deposition and inflammation in rheumatoid arthritis, lung fibrosis and neurodegeneration (Supplementary Tables 1 and 2). These data offer a molecular basis for how altered Pten signalling in the tumour stroma may elicit broad

responses in a variety of cells in the tumour microenvironment that contribute to disease manifestation.

## METHODS SUMMARY

**Transgenic mice.** Generation of *Fsp-cre* mice has been described<sup>15</sup>. *Pten*<sup>loxP</sup> mice were created following the strategy described in Supplementary Fig. 1. *Ets2*<sup>loxP</sup> mice were generated by standard techniques<sup>30</sup>. Animals were maintained and killed following institutional guidelines. Tenth-generation congenic (N10) FVB/N animals were used for transplantation and orthotopic injection studies.

**Tissue processing, histology.** Tissues were either embedded in OCT or fixed (4% paraformaldehyde or formalin) and embedded in paraffin. Frozen sections were used for X-gal staining as previously described<sup>15</sup>. IHC on tissue microarray sections from 99 patients with breast carcinoma was scored using Allred score's system<sup>48</sup>.

**Isolation of primary mammary fibroblasts.** Primary mammary fibroblasts were purified following the protocol published previously with minor modifications<sup>49</sup>. Mammary glands were dissected from 8-week-old female mice, minced and digested with collagenase I, 160 U ml<sup>-1</sup> hyaluronidase, 1  $\mu\text{g ml}^{-1}$  hydrocortisone and 10  $\mu\text{g ml}^{-1}$  insulin with penicillin and streptomycin in a 5% CO<sub>2</sub> incubator overnight at 37 °C. Collagenase was neutralized with 10% FBS-DMEM medium. Digested tissue was resuspended in medium and subjected to gravity for 12–15 min. Pellets were washed three times to collect epithelial organoids, and supernatants were subjected four more times to gravity sedimentation and then cultured.

**RNA and microarray analysis.** RNA was harvested with Trizol according to the manufacturer's instructions (Invitrogen). RNA quality and concentration were assessed with Bioanalyser and Nanodrop RNA 6000 nano-assays. RNA samples were hybridized to Affymetrix GeneChip Mouse genome 430 2.0 platform at the Microarray Shared Resource Facility, Ohio State University Comprehensive Cancer Center. The microarray data were deposited with Gene Expression Omnibus (GEO) and can be viewed at <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=zhmbrcoqaacyite&acc=GSE16073>.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 29 May; accepted 7 September 2009.

- Wiseman, B. S. & Werb, Z. Stromal effects on mammary gland development and breast cancer. *Science* **296**, 1046–1049 (2002).
- Nelson, C. M. & Bissell, M. J. Of extracellular matrix, scaffolds, and signaling: tissue architecture regulates development, homeostasis, and cancer. *Annu. Rev. Cell Dev. Biol.* **22**, 287–309 (2006).
- Mueller, M. M. & Fusenig, N. E. Friends or foes – bipolar effects of the tumour stroma in cancer. *Nature Rev. Cancer* **4**, 839–849 (2004).
- Schedin, P. Pregnancy-associated breast cancer and metastasis. *Nature Rev. Cancer* **6**, 281–291 (2006).
- Littlepage, L. E., Egeblad, M. & Werb, Z. Coevolution of cancer and stromal cellular responses. *Cancer Cell* **7**, 499–500 (2005).
- Bhowmick, N. A., Neilson, E. G. & Moses, H. L. Stromal fibroblasts in cancer initiation and progression. *Nature* **432**, 332–337 (2004).
- Kalluri, R. & Zeisberg, M. Fibroblasts in cancer. *Nature Rev. Cancer* **6**, 392–401 (2006).
- Cully, M., You, H., Levine, A. J. & Mak, T. W. Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. *Nature Rev. Cancer* **6**, 184–192 (2006).
- Bergamaschi, A. *et al.* Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. *J. Pathol.* **214**, 357–367 (2008).
- Myers, M. P. *et al.* The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *Proc. Natl Acad. Sci. USA* **95**, 13513–13518 (1998).
- Stambolic, V. *et al.* Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *Cell* **95**, 29–39 (1998).
- Salmela, L., Carracedo, A. & Pandolfi, P. P. Tenets of PTEN tumor suppression. *Cell* **133**, 403–414 (2008).
- Knobbe, C. B., Lapin, V., Suzuki, A. & Mak, T. W. The roles of PTEN in development, physiology and tumorigenesis in mouse models: a tissue-by-tissue survey. *Oncogene* **27**, 5398–5415 (2008).
- Di Cristofano, A., Pesce, B., Cordon-Cardo, C. & Pandolfi, P. P. Pten is essential for embryonic development and tumour suppression. *Nature Genet.* **19**, 348–355 (1998).
- Trimboli, A. J. *et al.* Direct evidence for epithelial–mesenchymal transitions in breast cancer. *Cancer Res.* **68**, 937–945 (2008).
- Guy, C. T. *et al.* Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc. Natl Acad. Sci. USA* **89**, 10578–10582 (1992).
- Cases, S. *et al.* Development of the mammary gland requires DGAT1 expression in stromal and epithelial tissues. *Development* **131**, 3047–3055 (2004).

18. Andrechek, E. R. *et al.* Amplification of the neu/erbB-2 oncogene in a mouse model of mammary tumorigenesis. *Proc. Natl Acad. Sci. USA* **97**, 3444–3449 (2000).
19. Desai, K. V. *et al.* Initiating oncogenic event determines gene-expression patterns of human breast cancer models. *Proc. Natl Acad. Sci. USA* **99**, 6967–6972 (2002).
20. Mootha, V. K. *et al.* PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genet.* **34**, 267–273 (2003).
21. Dennis, G., Jr *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, 3 (2003).
22. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57 (2009).
23. Weng, L. P., Brown, J. L., Baker, K. M., Ostrowski, M. C. & Eng, C. PTEN blocks insulin-mediated ETS-2 phosphorylation through MAP kinase, independently of the phosphoinositide 3-kinase pathway. *Hum. Mol. Genet.* **11**, 1687–1696 (2002).
24. Fowles, L. F. *et al.* Persistent activation of mitogen-activated protein kinases p42 and p44 and ets-2 phosphorylation in response to colony-stimulating factor 1/c-fms signaling. *Mol. Cell. Biol.* **18**, 5148–5156 (1998).
25. McCarthy, S. A. *et al.* Rapid phosphorylation of Ets-2 accompanies mitogen-activated protein kinase activation and the induction of heparin-binding epidermal growth factor gene expression by oncogenic Raf-1. *Mol. Cell. Biol.* **17**, 2401–2412 (1997).
26. Smith, J. L. *et al.* ets-2 is a target for an akt (Protein kinase B)/jun N-terminal kinase signaling pathway in macrophages of motheaten-viable mutant mice. *Mol. Cell. Biol.* **20**, 8026–8034 (2000).
27. Watabe, T. *et al.* The Ets-1 and Ets-2 transcription factors activate the promoters for invasion-associated urokinase and collagenase genes in response to epidermal growth factor. *Int. J. Cancer* **77**, 128–137 (1998).
28. Wei, G. *et al.* Activated Ets2 is required for persistent inflammatory responses in the motheaten viable model. *J. Immunol.* **173**, 1374–1379 (2004).
29. Ludwig, T. Local proteolytic activity in tumor cell invasion and metastasis. *Bioessays* **27**, 1181–1191 (2005).
30. Wei, G. *et al.* Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* **114**, 1123–1130 (2009).
31. Lin, E. Y. *et al.* Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am. J. Pathol.* **163**, 2113–2126 (2003).
32. Yamamoto, H. *et al.* Defective trophoblast function in mice with a targeted mutation of Ets2. *Genes Dev.* **12**, 1315–1326 (1998).
33. Lee, S., Jilani, S. M., Nikolova, G. V., Carpizo, D. & Iruela-Arispe, M. L. Processing of VEGF-A by matrix metalloproteinases regulates bioavailability and vascular patterning in tumors. *J. Cell Biol.* **169**, 681–691 (2005).
34. Millauer, B. *et al.* High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell* **72**, 835–846 (1993).
35. Sakurai, Y., Ohgimoto, K., Kataoka, Y., Yoshida, N. & Shibuya, M. Essential role of Flk-1 (VEGF receptor 2) tyrosine residue 1173 in vasculogenesis in mice. *Proc. Natl Acad. Sci. USA* **102**, 1076–1081 (2005).
36. Dakappagari, N. K. *et al.* Conformational HER-2/neu B-cell epitope peptide vaccine designed to incorporate two native disulfide bonds enhances tumor cell binding and antitumor activities. *J. Biol. Chem.* **280**, 54–63 (2005).
37. Finak, G. *et al.* Stromal gene expression predicts clinical outcome in breast cancer. *Nature Med.* **14**, 518–527 (2008).
38. Berns, K. *et al.* A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* **12**, 395–402 (2007).
39. Saal, L. H. *et al.* Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair. *Nature Genet.* **40**, 102–107 (2008).
40. Saal, L. H. *et al.* Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. *Proc. Natl Acad. Sci. USA* **104**, 7564–7569 (2007).
41. Tynan, J. A., Wen, F., Muller, W. J. & Oshima, R. G. Ets2-dependent microenvironmental support of mouse mammary tumors. *Oncogene* **24**, 6870–6876 (2005).
42. Buggy, Y. *et al.* Ets2 transcription factor in normal and neoplastic human breast tissue. *Eur. J. Cancer* **42**, 485–491 (2006).
43. Park, E. S. *et al.* Heterologous tissue culture expression signature predicts human breast cancer prognosis. *PLoS One* **2**, e145 (2007).
44. Svensson, S. *et al.* ERK phosphorylation is linked to VEGFR2 expression and Ets-2 phosphorylation in breast cancer and is associated with tamoxifen treatment resistance and small tumours with good prognosis. *Oncogene* **24**, 4370–4379 (2005).
45. Pap, T. *et al.* Activation of synovial fibroblasts in rheumatoid arthritis: lack of Expression of the tumour suppressor PTEN at sites of invasive growth and destruction. *Arthritis Res.* **2**, 59–64 (2000).
46. White, E. S. *et al.* Negative regulation of myofibroblast differentiation by PTEN (phosphatase and tensin homolog deleted on chromosome 10). *Am. J. Respir. Crit. Care Med.* **173**, 112–121 (2006).
47. Gibson, G. E. & Huang, H. M. Oxidative processes in the brain and non-neuronal tissues as biomarkers of Alzheimer's disease. *Front. Biosci.* **7**, d1007–d1015 (2002).
48. Allred, D. C., Harvey, J. M., Berardo, M. & Clark, G. M. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod. Pathol.* **11**, 155–168 (1998).
49. Soule, H. D. & McGrath, C. M. A simplified method for passage and long-term growth of human mammary epithelial cells. *In Vitro Cell. Dev. Biol.* **22**, 6–12 (1986).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank M. Rawahneh and J. Moffitt for histotechnical assistance, K. Kornacker, S. Cory and I. Vasudeva Murthy for bioinformatics assistance, P. Gulati for statistics assistance, the Ohio State University Human Tissue Resource Network and the Ohio State University Comprehensive Cancer Center Microarray, Nucleic Acids, Transgenics and Flow Cytometry Shared Facilities for technical assistance. *MMTV-ErbB2* mice were provided by W. Muller. This work was funded by the National Institutes of Health to G.L. (R01CA85619, R01HD47470, P01CA097189) and to M.C.O. (R01CA053271, P01CA097189), by the Komen Breast Cancer Foundation and Evelyn Simmers Charitable Trust to M.C.O., by the Terry Fox New Frontiers Group Grant to M.P., and by the Natural Science and Engineering Research Council of Canada Discovery Grants Program grant to M.H. F.L. and F.P. were funded by Department of Defense Pre-doctoral Fellowships and J.-L.C. was funded by a Department of Defense Postdoctoral Fellowship. G.L. is the recipient of the Pew Charitable Trusts Scholar Award and the Leukemia and Lymphoma Society Scholar Award. M.P. holds the Diane and Sal Guerrero Chair in Cancer Genetics at McGill University.

**Author Contributions** G.L. and M.C.O. designed and supervised this study, analysed data, and helped write and edit the manuscript. A.J.T., C.Z.C., F.L. and J.A.W. designed and performed experiments, collected and analysed data, and co-wrote the paper. N.C., J.C.T., H.W., J.-L.C., S.M.S. and M.N.G. technically assisted with experiments, and collected and analysed data. G.W., A.J.T., M.L.R. and M.W. performed experiments in initial stages of the project, particularly in designing and characterizing the mouse models. S.N., P.S. and T.J.R. contributed to the histopathological analysis of the mouse mammary tumour models and writing the manuscript. S.H.B. and L.Y. contributed to the histopathological analysis of human samples and writing the manuscript. S.A.F. and J.A.S. contributed to the statistical analyses of data and writing the manuscript. A.M., F.P., J.A.W., E.C., M.H. and M.P. contributed to the analysis and comparison of mouse and human microarray data and writing the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to M.C.O. ([michael.ostrowski@osumc.edu](mailto:michael.ostrowski@osumc.edu)) or G.L. ([gustavo.leone@osumc.edu](mailto:gustavo.leone@osumc.edu)).



## METHODS

**Transgenic mice.** Animals were housed according to federal and University Laboratory Animal Resources at Ohio State University regulations.

*Fsp-cre*: the generation of the *Fsp-cre* mouse line has been previously described<sup>15</sup>.  
*Ets2<sup>loxP</sup>*: the *Ets2* conditional transgenic mouse line was generated such that the *Ets2* pointed domain is 'floxed'<sup>28</sup>. The pointed domain is encoded by exons 3–5 and important for the protein–protein interaction and signal transduction.

*Ets2<sup>Δb</sup>*: the *Ets2<sup>Δb</sup>* mouse was a gift from R. G. Oshima<sup>32</sup>.

*Pten<sup>loxP</sup>*: the *Pten<sup>loxP</sup>* mice were generated in our laboratory. *LoxP* sites were introduced into two *HpaI* sites within introns 3 and 5 of the *Pten* gene, respectively, to flank exons 4 and 5. Exon 5 encodes the lipid phosphatase domain. Tissue-specific expression of *cre* will excise exons 4 and 5, generating a loss-of-function *Pten<sup>Δ</sup>* allele (Supplementary Fig. 1).

*Rosa-LacZ*: the generation of the conditional *Rosa-LacZ* mouse line has been previously described<sup>50</sup>. A list of PCR primers can be found in Supplementary Table 3.

**Mammary tissue transplantation.** Transplant procedure was based on a previously published method<sup>17</sup>. The day before surgery, recipient mice were anaesthetized with isoflurane (Abbott Laboratories) and a square area along the scapular region was shaved (roughly 25 mm × 25 mm). At the time of surgery, inguinal and groin mammary tissue (approximately 5 mm × 5 mm in size) was removed from 8- to 9-week-old donor females and placed subcutaneously into the scapular region of wild-type hosts (littermates or FVB/N purchased from Taconic) through two 5- to 10-mm incisions on the left and right side under aseptic conditions. Only tissue from the abdominal (#4 and #9) and groin (#5 and #10) mammary glands was used for transplant after the lymph nodes were removed. The small incisions were closed using a 9-mm wound clip. Animals were monitored twice a week until tumour onset. Mice were killed either at specific time points, when the tumour was about 2 cm in size, or it presented a health problem to the animal such as exterior ulceration at the site of the tumour.

**Orthotopic mammary gland injection.** The neu-expressing mouse mammary carcinoma cell line NT2.5 was provided by P. T. Kaumaya and was maintained as described<sup>36</sup>. Eight-week-old female mice of each genotype were anaesthetized and injected with 5 × 10<sup>5</sup> NT2.5 cells at both inguinal mammary glands. Tumour initiation was monitored by palpating twice a week. All the mice were killed three weeks after injection. Tumour volume was calculated by the formula  $V = \frac{1}{2} \times \text{length} \times (\text{width})^2$ .

**Tissue processing and X-gal staining.** Large individual tumours (typically ~1–2 cm) or the remainder of transplanted tissues were removed, divided and either fixed in 4% paraformaldehyde for 24–48 h or embedded directly in OCT (Sakura). Fixed tissue samples were embedded in paraffin and cut into 5-μm sections for H&E, IHC or immunofluorescence staining. For each sample collected, two sets of sections were obtained at 25-μm intervals for analysis. Corresponding OCT-embedded tissue was sectioned (7 μm) in a similar manner for X-gal or immunofluorescence staining. For X-gal staining, frozen tissue sections were dried 15 min at room temperature before fixing in a glutaraldehyde solution (0.2% glutaraldehyde, 1.25 mM EGTA, pH 7.3 and 2 mM magnesium chloride in 1 × PBS) for 30 min. The sections were washed with LacZ wash buffer (2 mM magnesium chloride, 0.01% sodium deoxycholate, 0.02% IGEPAL CA-630 (Sigma) in PBS) for 5 min three times and then stained in LacZ staining solution (4 mM potassium ferricyanide, 4 mM potassium ferrocyanide, 1 mg ml<sup>-1</sup> X-gal in LacZ wash buffer) at 37 °C overnight (~18 h) and protected from light. Stained sections were washed in PBS for 5 min three times and then rinsed with water for 2 min before counter-staining with nuclear fast red. Collagen deposition was visualized using standard Masson's trichrome staining. Whole mammary gland sections were stained and imaged using an Aperio Scanscope CS whole-slide scanner. Areas of skin and muscle, which also stain positive, were manually encircled and excluded from the analysis.

**IHC and immunofluorescence.** IHC and immunofluorescence were performed using paraffin sections with the following antibodies: Pten (1:100, Cell Signaling, overnight), collagen I (1:100, Abcam, 30 min), F4/80 (1:50, Caltag, overnight), phospho-Akt<sup>S473</sup> (1:50, Cell Signaling, overnight), phospho-JNK<sup>T183/Y185</sup> (1:50, Cell Signaling, overnight), phospho-Erk1/2 (1:100, Cell Signaling, overnight), phospho-Ets2<sup>T72</sup> (1:25, M.C.O. laboratory, overnight), Cytokeratin 8/18 (1:300; Research Diagnostics, 30 min), E-cadherin (1:700, BD-Pharmingen, 30 min), mouse α-SMA (1:200, Sigma, 30 min) and collagen IV (1:100, Chemicon, 30 min). In general, paraffin sections were deparaffinized and the antigen retrieval accomplished by incubation in antigen retrieval solution (DAKO) at >95 °C (30 min). For IHC, staining was developed using the biotin/avidin/horseradish peroxidase system from Vector Laboratories according to the manufacturer's instructions. All IHC slides were counterstained with haematoxylin and images obtained using an Eclipse 50i microscope (Nikon) and an AxioCam HRC camera (Zeiss). For IHC quantification, cells were counted manually from

ten random fields and reported as a percentage of positive cells from the total cell population.

Frozen sections of mammary glands for immunofluorescence were fixed at 4 °C in either 4% paraformaldehyde in PBS or acetone. Samples were treated with phospho-Ets2<sup>T72</sup> (1:125; M.C.O. laboratory, overnight), vimentin (1:50; Santa Cruz Biotech, overnight), Pten (1:100, Cell Signaling, overnight), VEGF<sub>164</sub> (1:100, R&D Systems, overnight), phospho-VEGFR2<sup>Y1173</sup> (1:100, Cell Signaling, overnight) and CD31 (1:50, BD Biosciences, 30 min) primary antibodies. Fluorescent staining was developed using secondary antibodies conjugated to AlexaFluor dyes following standard protocol (Invitrogen; Molecular Probes). For Pten immunofluorescence on paraffin sections, the signal was amplified using a biotinylated secondary antibody and Texas red conjugated to streptavidin. All immunofluorescence sections were counterstained with DAPI and images obtained using an AxioScope 40 microscope (Zeiss) equipped with an AxioCam HRC camera (Zeiss). Image analysis for quantifying CD31 and phospho-VEGFR2<sup>Y1173</sup> areas was performed using standard image processing techniques implemented in Matlab. VEGF<sub>164</sub> images were quantified with ImageJ software.

**Gelatinase *in situ* zymography.** This was performed as described, with minor modification<sup>51</sup>. Briefly, frozen sections (10 μm) were quickly fixed with cold acetone, rehydrated with PBS and then incubated with 40 μg ml<sup>-1</sup> DQ-gelatin fluorescein conjugate (Molecular Probes) in reaction buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM CaCl<sub>2</sub> and 0.2 mM NaN<sub>3</sub>, pH 7.6) for 10 h. The reaction was quenched with 10 mM EDTA-PBS wash. Nuclei were counterstained with DAPI. A consecutive slide was stained with H&E to localize the MMP9 activity. Fluorescent images were acquired with AxioScope40 microscope (Zeiss) and AxioCam HRC camera (Zeiss).

**Western blot.** One million to two million cells were lysed with radio immunoprecipitation assay (RIPA) buffer (50 mM pH 7.4 Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate and 0.1% SDS) containing protease and phosphatase inhibitors (Roche). Primary antibodies for Pten (Cell Signaling), Akt (Cell Signaling), phospho-Akt<sup>S473</sup> and phospho-Akt<sup>T308</sup> (Cell Signaling), phospho-JNK<sup>T183/Y185</sup> (Cell Signaling), phospho-Erk1/2 (Cell Signaling), Erk1/2 (Santa Cruz), Ets2 (M.C.O. laboratory) and Collagen 1 (Abcam) were incubated overnight at a dilution of 1:1,000, whereas tubulin (Sigma) was used at 1:5,000 dilution. Washed membranes were blotted with either horseradish-peroxidase-conjugated anti mouse IgG or anti rabbit IgG antibodies and developed with enhanced chemiluminescent substrate (Thermo Scientific).

**Microarray analysis.** The data were analysed using WEDGE++ expression analysis<sup>52</sup>. Heat map representation (Supplementary Fig. 6) was performed using the TIGR Multiexperiment Viewer program MeV version 4.1.

**Quantitative RT-PCR.** Quantitative gene expression was performed using 50 ng complementary DNA per reaction. Taqman Roche Universal Probe Library system probe and primers (Roche) following the manufacturer's instructions. Reactions were performed on the I-cycler iQ Real-Time machine (Bio-Rad). A list of quantitative RT-PCR primers can be found in Supplementary Table 3. The reference gene used for all quantitative RT-PCR assays is *Rpl4*.

**ChIP and quantitative RT-PCR.** ChIP assays were performed as described by Hu *et al.*<sup>53</sup>. Primary fibroblasts were cross-linked with 1% formaldehyde and soluble chromatin was prepared with sonication to an average DNA length of 200–1,000 base pairs. Sheared soluble chromatin was pre-cleared with transfer-RNA-blocked Protein G Agarose, and 10% of the pre-cleared chromatin was set aside as input control. Immunoprecipitation was performed with 5 μg of Ets2 antibody or rabbit IgG overnight at 4 °C. Immune complexes were pulled down with Protein G-Agarose, washed, and eluted with elution buffer (0.1 M NaHCO<sub>3</sub>, 1% SDS), and cross-links removed by incubating with 200 mM NaCl containing 50 μg ml<sup>-1</sup> of RNase A (Sigma) at 65 °C overnight. DNA was purified with the Qiagen PCR Purification Kit after proteinase K treatment according to the manufacturer's instructions. Samples were analysed by quantitative RT-PCR as indicated above. The threshold for the promoter being studied was adjusted by that of input values and represented as relative abundance. All quantitative RT-PCR reactions were analysed by melt curve analysis and agarose gels to confirm the specificity of the reaction. A list of ChIP primers can be found in Supplementary Table 3.

**Generating the human stroma heat map with genes identified in *Pten* null fibroblasts.** Analysis of the *Pten* null mouse microarray data using WEDGE++ software<sup>52</sup> led to the identification of 195 differentially expressed probe sets matching to 150 unique mouse genes. A search for human orthologues using Ensembl and MGI databases yielded a list of 137 genes. These genes were queried against the McGill Cancer Center's Breast Stroma Microarray data (GSE9014 and GSE4823). Of the 137 genes, 129 were represented on the Agilent Custom Array used in the McGill study. A heat map was generated for the human stroma data set (52 normal stroma and 49 tumour stroma samples). To achieve better resolution

on the heat map, and to identify only those genes that had highly variable gene expression across all samples, a variance cutoff of  $>0.5$  was used to generate a subset of 70 genes. The heat map (Fig. 5a) shows the ability of these 70 genes to separate the normal and tumour stroma samples based solely on their gene expression profiles. This partitioning is highly significant ( $P = 3.9 \times 10^{-15}$ ), as determined by Wilcoxon's test on the average expression of *Pten* null-signature in all samples.

Interestingly, comparison of 137 human orthologues of the *Pten* null list with the 163 genes associated with recurrence in tumour stroma from the McGill study showed that there are 12 genes (highlighted in red; Fig. 5a) present on both gene lists. This overlap is again highly significant ( $P = 2.5 \times 10^{-8}$ ; Fisher's exact analysis).

**TMA.** TMA slides containing paraffin-embedded breast carcinoma tissues were processed at the Pathology Core Facility and Tissue Archives Human Tissue Resource Network at Ohio State University. IHC on these slides was performed with phospho-Ets2<sup>T72</sup> (1:25, M.C.O. laboratory), Pten (1:100, Cell Signaling) and phospho-Akt<sup>S473</sup> (1:50, Cell Signaling) following standard procedures. Level of PTEN and phospho-ETS2 and phospho-AKT expression in stromal cells were scored with Allred score method<sup>48</sup>.

**Statistical analysis.** The number of animals and experiments are indicated in the figures. All the statistical tests were two-sided. The Wilcoxon rank test was used for some of the statistical analyses when the outcome variables were skewed and could not be transformed to achieve normality (Figs 1f and 2c, e, i). General linear models (analysis of variance) were used to study differences in continuous outcome variables among groups. *Ad hoc* pairwise comparisons were performed if the overall test was significant (Fig. 4b, d, f and Supplementary Figs 9c and 11b). Fisher's exact test was used to compare counts in Fig. 1e because low counts (below 5) were observed in one of the groups. Fisher's exact test was also used to compare gene overlap in Fig. 5b. Multiplicity adjustment methods (Bonferroni method) were used when multiple pairwise comparisons were performed. A two-sample (unpaired) Student's *t*-test was used for tumour load and the fluorescent microscopy images of VEGF<sub>164</sub>, CD31 and pVEGFR<sup>Y1173</sup> staining in Fig. 3b, e, f and Supplementary Fig. 9a, e. For human TMA, a Pearson correlation

was used to evaluate the association of Allred scores between PTEN/P-ETS2<sup>T72</sup>, PTEN/P-AKT<sup>S473</sup> and P-ETS2<sup>T72</sup>/P-AKT<sup>S473</sup>.

We show the expression data comparing normal and tumour stroma in Fig. 5a as a heat map because it highlights the 'direction' of the change in expression. Moreover, a permutation test strategy was used to study whether the 70-gene subset of the mouse *Pten* signature discriminated between human tumour and normal stroma samples<sup>54</sup> (Fig. 5a). A total of 100,000 random permuted samples was used to compute the empirical *P* value of the 70-gene signature. A Wilcoxon rank sum value for each random 70-gene permutation was computed from the average log expression difference (between normal and tumour stroma) of the 70 genes. For computing the average log expression difference of the mouse *Pten* 70-gene signature, we assigned negative values to the 13 downregulated genes. The empirical *P* value obtained using this method ( $8 \times 10^{-5}$ ) represents the proportion of random permuted samples that had a Wilcoxon rank sum value larger than the Wilcoxon rank sum value of the *Pten* 70-gene unpermuted signature. In addition, principal component analysis was used to explore whether the *Pten* 70-gene signature discriminated normal versus tumour stroma; this principal component analysis, like the heatmap presented in Fig. 5a, showed perfect separation of normal and tumour samples (data not shown), with the first component being sufficient for this separation. A Wilcoxon rank sum test yielded a *P* value of less than  $1 \times 10^{-10}$ .

50. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genet.* **21**, 70–71 (1999).
51. Mook, O. R., Van Overbeek, C., Ackema, E. G., Van Maldegem, F. & Frederiks, W. M. *In situ* localization of gelatinolytic activity in the extracellular matrix of metastases of colon cancer in rat liver using quenched fluorogenic DQ-gelatin. *J. Histochem. Cytochem.* **51**, 821–829 (2003).
52. Auer, H. *et al.* Gene-resolution analysis of DNA copy number variation using oligonucleotide expression microarrays. *BMC Genomics* **8**, 111 (2007).
53. Hu, R. *et al.* Eos, MITF, and PU.1 recruit corepressors to osteoclast-specific genes in committed myeloid progenitors. *Mol. Cell. Biol.* **27**, 4018–4027 (2007).
54. Westall, P. H. & Young, S. S. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment* (John Wiley, 1993).

# Allele-specific tumor spectrum in *Pten* knockin mice

Hui Wang<sup>a,b,c</sup>, Matt Karikomi<sup>a,b,c,1</sup>, Shan Naidu<sup>a,b,c,d,1</sup>, Ravi Rajmohan<sup>a,b,c,1</sup>, Enrico Caserta<sup>a,b,c</sup>, Hui-Zi Chen<sup>a,b,c,e</sup>, Maysoun Rawahneh<sup>a,b,c</sup>, Julie Moffitt<sup>a,b,c</sup>, Julie A. Stephens<sup>f</sup>, Soledad A. Fernandez<sup>f</sup>, Michael Weinstein<sup>a,b,c</sup>, Danxin Wang<sup>g</sup>, Wolfgang Sadee<sup>g</sup>, Krista La Perle<sup>d</sup>, Paul Stromberg<sup>d</sup>, Thomas J. Rosol<sup>d</sup>, Charis Eng<sup>h</sup>, Michael C. Ostrowski<sup>i,j,2</sup>, and Gustavo Leone<sup>a,b,c,e,j,2</sup>

<sup>a</sup>Department of Molecular Genetics, College of Biological Sciences, <sup>b</sup>Human Cancer Genetics Program, Comprehensive Cancer Center, College of Medicine, <sup>c</sup>Department of Molecular Virology, Immunology, and Medical Genetics, College of Medicine, <sup>d</sup>Department of Veterinary Biosciences, College of Veterinary Medicine, <sup>e</sup>Medical Scientist Program, College of Medicine, <sup>f</sup>Center for Biostatistics, College of Medicine, <sup>g</sup>Department of Pharmacology, College of Medicine, <sup>h</sup>Department of Molecular and Cellular Biochemistry, College of Medicine, and <sup>i</sup>Tumor Microenvironment Program, Comprehensive Cancer Center, College of Medicine, Ohio State University, Columbus, OH 43210; and <sup>j</sup>Genomic Medicine Institute, Lerner Research Institute and Taussing Cancer Institute, Cleveland Clinic, Cleveland, OH 44195

Edited\* by Albert de la Chapelle, Ohio State University Comprehensive Cancer Center, Columbus, OH, and approved January 29, 2010 (received for review November 2, 2009)

**Germline mutations in the tumor suppressor gene *PTEN* (phosphatase and tensin homology deleted on chromosome 10) cause Cowden and Bannayan–Riley–Ruvalcaba (BRR) syndromes, two dominantly inherited disorders characterized by mental retardation, multiple hamartomas, and variable cancer risk. Here, we modeled three sentinel mutant alleles of *PTEN* identified in patients with Cowden syndrome and show that the nonsense *Pten*<sup>Δ4–5</sup> and missense *Pten*<sup>C124R</sup> and *Pten*<sup>G129E</sup> alleles lacking lipid phosphatase activity cause similar developmental abnormalities but distinct tumor spectra with varying severity and age of onset. Allele-specific differences may be accounted for by loss of function for *Pten*<sup>Δ4–5</sup>, hypomorphic function for *Pten*<sup>C124R</sup>, and gain of function for *Pten*<sup>G129E</sup>. These data demonstrate that the variable tumor phenotypes observed in patients with Cowden and BRR syndromes can be attributed to specific mutations in *PTEN* that alter protein function through distinct mechanisms.**

cancer genetics | Cowden syndrome

The tumor suppressor gene *PTEN* (phosphatase and tensin homology deleted on chromosome ten) encodes a product with both protein and lipid phosphatase activity (1–3). The lipid phosphatase activity negatively regulates phosphoinositide-3 kinase (PI-3K) and the downstream Akt and mammalian target of rapamycin pathway components (4). Consistent with a requirement for lipid phosphatase activity in tumor suppression, the mutant *PTEN*<sup>G129E</sup> allele, originally identified in patients with Cowden syndrome, selectively lacks lipid phosphatase activity (5). However, *PTEN* also has protein phosphatase activity, and although controversial, there is a body of evidence suggesting that *PTEN*'s protein phosphatase activity may also contribute to tumor suppression (6). The Cowden syndrome *PTEN*<sup>C124R</sup> allele, which encodes a protein product lacking lipid and protein phosphatase activity (7), provides genetic support for an involvement of dual *PTEN* phosphatase activities in tumor suppression. In addition, recent data suggest a role for nuclear *PTEN* in genomic stability that may be independent of its lipid phosphatase and PI-3K signaling activities (8). Despite intensive efforts in the past decade to understand the biochemical functions of *PTEN*, it remains unclear which of its many functions endow this gene with tumor suppressor status and whether the different functions contribute to tumor suppression selectively in an organ-specific and/or signaling pathway-specific manner.

Individuals with Cowden syndrome show a wide variation in disease manifestation, including cancer predisposition (9), that is thought to be driven by a combination of genotype as well as polymorphic heterogeneity in the population (10). The reason why polymorphic heterogeneity has been invoked is because identical mutations result in disparate phenotypes, ranging from mild developmental disorders in Cowden syndrome to very severe disorders in Bannayan–Riley–Ruvalcaba (BRR) syndrome (9). Here, we used homologous recombination in mice to model three differ-

ent mutant alleles of *PTEN* originally identified in Cowden syndrome. We found that each mutant allele displayed distinct tumor phenotypes in organs targeted by Cowden syndrome. These functional differences are not manifested during embryonic development and do not correspond strictly to the level of Akt activation. Rather, the variable phenotypes can be attributed to functions beyond PI3K–Akt activation that include gain of function for the *Pten*<sup>G129E</sup> allele and loss of function for the *Pten*<sup>Δ4–5</sup> and *Pten*<sup>C124R</sup> alleles. Together, these results demonstrate that specific germline mutations have a strong influence in the variable predisposition to cancer of patients who have Cowden syndrome.

## Results

**Abnormal Embryonic Development in *Pten*<sup>Δ4–5</sup>, *Pten*<sup>G129E</sup>, and *Pten*<sup>C124R</sup> Knockin Mice.** To explore the physiological consequence of inactivating distinct arms of *PTEN*'s tumor suppressor arsenal, we generated, characterized, and compared mouse models that have a frameshift (*Pten*<sup>Δ4–5</sup>) (11) leading to the premature stop and missense mutations in *Pten* (*Pten*<sup>G129E</sup> or *Pten*<sup>C124R</sup>) that can be found in patients with Cowden syndrome (12) (Fig. 1*A*). Targeting of the three *Pten* mutant alleles was verified by Southern blot analysis, PCR, and direct sequencing of genomic DNA (11) (Fig. 1*B* and *C* and Fig. S1*A*, *Left*). Allelic expression imbalance analysis showed that the two missense mutant alleles lacking the neomycin (neo) cassette (neo–) were expressed at a similar level as the WT allele (Fig. 1*D*), whereas the same mutant alleles containing the neo cassette (neo+) were expressed, as expected, at lower levels (Fig. S1*A*, *Right*).

We first evaluated the effect of *Pten*<sup>Δ4–5</sup>, *Pten*<sup>G129E</sup>, and *Pten*<sup>C124R</sup> on embryonic development by examining offspring derived from intercrosses between heterozygous knockin mice (Fig. 2*A*). The intercrosses yielded no live homozygous progeny at birth for any of the three mutant alleles; however, mutant embryos were recovered at embryo day (E) 8.5 and E9.5, albeit at a lower than expected frequency. Development of the few mutant embryos that were obtained was severely compromised, with defects in anterior-posterior patterning, failure of axial rotation, and absence of overt tissue differentiation leading to resorption by E9.5 (Fig. 2*A* and *B* and Fig. S1*B*). Heterozygous *Pten*<sup>Δ4–5/+</sup>,

Author contributions: H.W., M.C.O., and G.L. designed research; H.W., M.K., S.N., R.R., E.C., H.-Z.C., M.R., and J.M. performed research; D.W. and W.S. contributed new reagents/analytic tools; H.W., H.-Z.C., J.A.S., S.A.F., M.W., D.W., W.S., K.L.P., P.S., T.J.R., C.E., M.C.O., and G.L. analyzed data; and H.W., M.C.O., and G.L. wrote the paper.

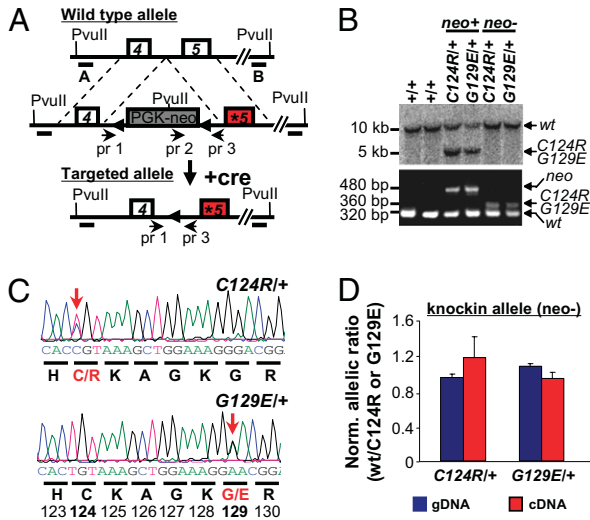
The authors declare no conflict of interest.

\*This Direct Submission article had a prearranged editor.

<sup>1</sup>M.K., S.N., and R.R. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. E-mail: michael.ostrowski@osumc.edu or gustavo.leone@osumc.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0912524107/DCSupplemental](http://www.pnas.org/cgi/content/full/0912524107/DCSupplemental).



**Fig. 1.** Generation, targeting, and verification of point mutations of *Pten* alleles. (A) Endogenous WT *Pten* allele (Top), the two targeted missense mutations in exon 5 (\*) of *Pten* allele containing the selectable phosphoglycerate kinase promoter (PGK)-neo cassette (flanked by *LoxP* sites, triangles; Middle), and the two targeted mutant *Pten* knockin alleles (*Pten*<sup>C124R</sup> and *Pten*<sup>G129E</sup>) lacking the PGK-neo cassette (after mating with *EIIA-cre*-expressing mice; Bottom). A and B, DNA probes used for Southern blot analysis; pr, primers used for PCR genotyping. (B) Southern blot analysis (Upper) and genotyping (Lower) of tail DNA with the indicated genotypes. Genomic DNA was digested with *PvuII* and probed with probe A, and expected band sizes are indicated for each allele. PCR amplification using primer pairs 1–3 and 2–3 (primer information provided in Table S3) yielded specific fragment size for different alleles as indicated. (C) Sequence analysis of tail DNA isolated from *Pten*<sup>C124R/+</sup> and *Pten*<sup>G129E/+</sup> mice. Chromatograms demonstrating the successful targeting of the *Pten* locus and translated amino acids are shown below the codons. Red letters and bold numbers denote the two targeted amino acids (C124R and G129E). Arrows point to targeted nucleotides. (D) Allelic expression imbalance analysis of allele-specific expression. The graph shows the proportion of mRNA expressed from the WT allele over the indicated mutant allele in lungs of *Pten*<sup>C124R/+</sup> and *Pten*<sup>G129E/+</sup> mice. neo<sup>-</sup>, mice lacking the PGK-neo cassette. Genomic DNA (gDNA) was used as an internal control.

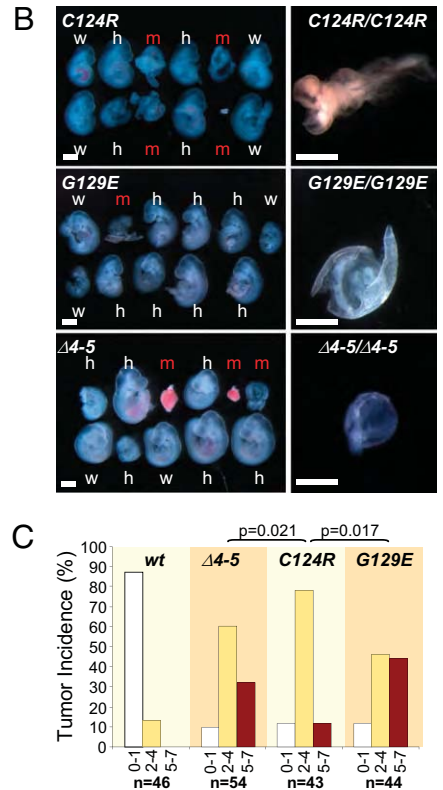
*Pten*<sup>G129E/+</sup>, and *Pten*<sup>C124R/+</sup> embryos were morphologically indistinguishable from WT littermates (Fig. 2A and B and Fig. S1C). Thus, overall, there were no obvious differences among homozygous knockin embryos of the three *Pten* mutant genotypes.

**Organ-Selective Cancer Predisposition in *Pten*<sup>Δ4-5/+</sup>, *Pten*<sup>G129E/+</sup>, and *Pten*<sup>C124R/+</sup> Mice.** The tumor spectrum in animals harboring one copy of each of the three mutant *Pten* alleles was determined by analyzing groups of >40 adult male and female mice. Similar to previous studies that analyzed heterozygous mice harboring a *Pten* knockout allele (13–17), mice carrying one copy of each of the three *Pten* mutant alleles exhibited neoplasms in multiple organs that are characteristically involved in patients who have Cowden syndrome, confirming these are causal mutations that drive this syndrome. Interestingly, pathological examination of 9-month-old mice revealed differences in the frequency and severity of the individual proliferative lesions found in the three genetic groups (Table S1). The intergroup differences were amplified when the propensity of an individual animal to develop tumors at two or more organ sites was considered (Fig. 2C). A significant bias for lesions also occurred in specific organ pairs for each of the three genetic cohorts. Mammary gland-lymph node ( $P < 0.001$ ), uterus-lymph node ( $P = 0.032$ ), thyroid-lymph node ( $P = 0.003$ ), and prostate-stomach ( $P = 0.003$ ) lesions were frequently present in *Pten*<sup>Δ4-5/+</sup> mice; mammary gland-lymph node ( $P = 0.050$ ) lesions were present in *Pten*<sup>G129E/+</sup> mice;

**A Table 1. Embryo viability**

E9.5	+/+	m/+	m/m	Total
<i>C124R</i>	observed	19	37	0(14)* 70
	expected	18	35	18
<i>G129E</i>	observed	8	35	0(3)* 46
	expected	12	24	12
$\Delta 4-5$	observed	6	16	0(3)* 25
	expected	6	13	6
E8.5	+/+	m/+	m/m	Total
<i>C124R</i>	observed	14	46	0(11)* 71
	expected	18	36	18
<i>G129E</i>	observed	20	43	0(3)* 66
	expected	17	33	17
$\Delta 4-5$	observed	14	45	0(6)* 65
	expected	16	33	16

\* $p < 0.001$ , Chi-square test or Fisher's exact test



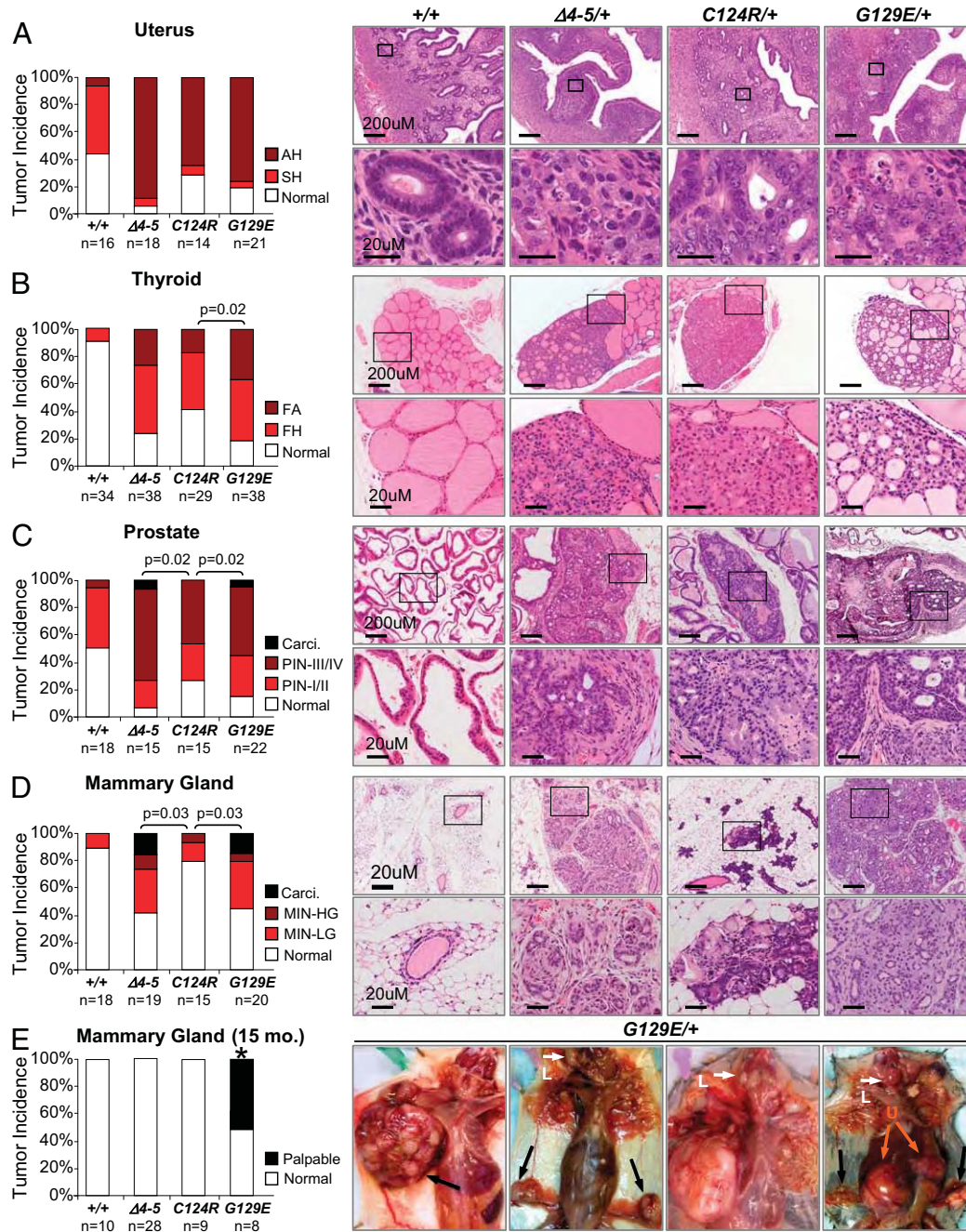
**Fig. 2.** *Pten*<sup>Δ4-5</sup>, *Pten*<sup>C124R</sup>, and *Pten*<sup>G129E</sup> mutations cause early embryonic lethality and exhibit allele-specific tumor syndromes. (A) Offspring from *Pten* (*Pten*<sup>C124R</sup>, *Pten*<sup>G129E</sup>, or *Pten*<sup>Δ4-5</sup>) heterozygous intercrosses were examined. The number of observed and expected E8.5 and E9.5 embryos is indicated. Fisher's exact or  $\chi^2$  tests were used to compare differences between observed and expected homozygous embryos. The number of dead embryos is shown in parentheses. m, mutant allele (*Pten*<sup>C124R</sup>, *Pten*<sup>G129E</sup>, or *Pten*<sup>Δ4-5</sup>). (B) (Left) Stereomicroscopic images of E9.5 WT (w), heterozygous (h), and homozygous (m) *Pten*-mutant littermate embryos. (Right) Higher magnification images of severely affected homozygous mutant embryos. (Scale bar: 1 mm.) (C) Organ distribution of tumor lesions in animals harboring the indicated *Pten* alleles. To simplify the graphical representation of data, mice were grouped based on the number of organs (0–1, white bars; 2–4, yellow bars; 5–7, red bars) afflicted with tumors. All uncategorized data were analyzed by Poisson regression methods, and significant differences are shown; all *Pten* mutant-WT allele comparisons were significant ( $P < 0.0001$ ; not indicated).

and mammary gland-intestine ( $P = 0.007$ ), mammary gland-adrenal ( $P = 0.049$ ), and prostate-stomach ( $P = 0.002$ ) lesions were present in *Pten*<sup>C124R/+</sup> mice (Table S2). Consistent with observations made in human patients with Cowden syndrome, there was a gender bias in the organ-specific manifestation of lesions, with lymph node lesions

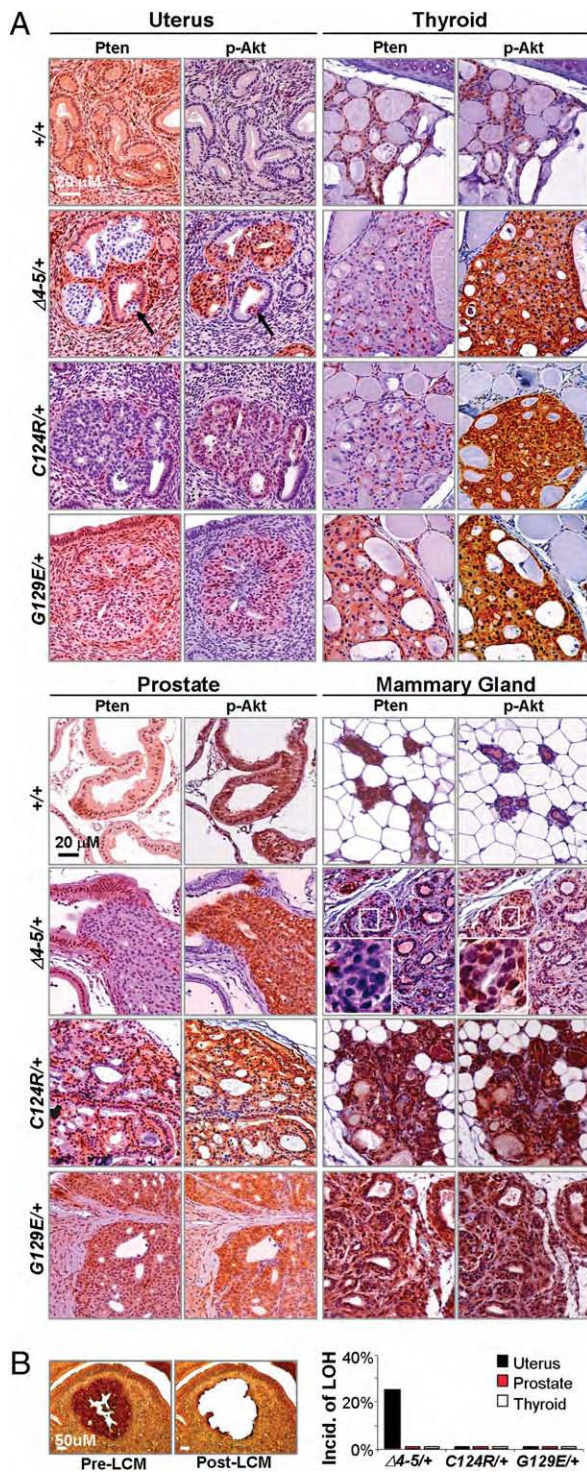
more frequently found in female mice (for *Pten* <sup>$\Delta$ 4-5</sup>,  $P < 0.001$ ) and adrenal and stomach lesions more commonly found in male mice (for *Pten* <sup>$\Delta$ 4-5</sup>,  $P = 0.008$  and  $P = 0.012$ , respectively; for *Pten*<sup>C124R</sup>,  $P = 0.034$ ) (Fig. S2). These findings challenge the current view of polymorphic heterogeneity in the population as the principal reason for the variable penetrance observed in patients with Cowden syn-

drome. We suggest that this variability may instead be attributed, at least in part, to allele-specific effects of *PTEEN* mutations.

Based on these initial observations, we performed a thorough histopathological analysis of neoplastic lesions in organ sites frequently involved in Cowden syndrome, including the uterus, thyroid, mammary gland (2, 18–20) and the prostate (21). Detailed



**Fig. 3.** Allele-specific and organ-specific tumor development in *Pten* <sup>$\Delta$ 4-5/+</sup>, *Pten*<sup>C124R/+</sup>, and *Pten*<sup>G129E/+</sup> mice. Histopathological grades of lesions were compared between WT, *Pten* <sup>$\Delta$ 4-5/+</sup>, *Pten*<sup>C124R/+</sup>, and *Pten*<sup>G129E/+</sup> 9-month-old mice in the uterus (A), thyroid (B), prostate (C), and mammary gland (D). Note the significant difference in the severity of lesions among the various animals with mutant *Pten* alleles. (Right) Histology with the highest grade lesions found in each genetic group. (Lower) Magnified view of the boxed region in the upper panels. Detailed histopathological criteria used to grade the lesions are included in Fig. S3. AH, atypical hyperplasia; SH, simple hyperplasia; PIN-I/II, prostatic intraepithelial neoplasia grade I or II; PIN-III/IV, prostatic intraepithelial neoplasia grade III or IV; FA, follicular adenoma; FH, follicular hyperplasia; Carci., carcinoma; MIN, mammary intraepithelial neoplasia; MIN, low grade; HG, high grade; Normal, no gross or microscopic tumor. Comparisons between genetic groups in A–D were analyzed by  $\chi^2$  or Fisher's exact tests and adjusted by Holm's method. (E) Large palpable tumor masses in 12–15-month-old *Pten*<sup>G129E/+</sup> female mice. Black and orange arrows point to mammary gland and uterine (U) tumors, respectively. White arrows point to enlarged lymph nodes (L). Normal, no gross mammary gland tumors; Palpable, palpable mammary gland tumors. Comparisons between genetic groups in E were analyzed by a binomial exact test. All comparisons between *Pten*<sup>G129E/+</sup> and other genetic groups were found to be significant (\* $P < 0.01$ ).



**Fig. 4.** Expression of Pten and p-Akt in proliferative lesions of various organs. (A) Tissue sections from the uterus and thyroid (Upper) and prostate and mammary gland (Lower) of mice with the indicated genotypes were stained with Pten- and p-Akt-specific antibodies. Complete loss of Pten occurred in  $Pten^{\Delta4-5/+}$  lesions of the uterus, thyroid, and prostate, whereas decreased levels of Pten were observed in  $Pten^{C124R/+}$  lesions of the uterus and thyroid. Pten protein expression was lost in a mosaic fashion in advanced lesions of  $Pten^{\Delta4-5/+}$  mice but persisted in  $Pten^{C124R/+}$  MIN and  $Pten^{G129E/+}$  mammary gland carcinomas. (Right) Magnified images of the boxed region. (B) LOH analysis of WT *Pten* allele detected by PCR for laser capture microdissected (LCM) lesions of the uterus, prostate, and thyroid of 9-month-old heterozygotes. (Left) Representative images showing pre- and post-LCM tissue of "quick" p-Akt immunostained sections. (Right) Incidence

criteria for grading lesions are presented in Fig. S3. In the uterus, proliferative lesions of the endometrium (simple hyperplasia and atypical hyperplasia) were encountered in  $Pten^{\Delta4-5/+}$ ,  $Pten^{C124R/+}$ , and  $Pten^{G129E/+}$  female mice with equal frequency (Fig. 3A), similar to what has been reported in patients with Cowden syndrome. Advanced thyroid follicular lesions, however, were more frequently observed in  $Pten^{G129E/+}$  mice compared with  $Pten^{\Delta4-5/+}$  or  $Pten^{C124R/+}$  mice (Fig. 3B and Fig. S4A). In the prostate, both  $Pten^{\Delta4-5/+}$  and  $Pten^{G129E/+}$  mice were more likely to have lesions of higher histological grade, which included prostatic adenocarcinoma, than  $Pten^{C124R/+}$  mice (Fig. 3C). In the mammary gland, advanced lesions such as mammary intraepithelial neoplasia (MIN), invasive carcinoma, and adenocarcinoma were observed in both  $Pten^{\Delta4-5/+}$  and  $Pten^{G129E/+}$  female mice but rarely in  $Pten^{C124R/+}$  female mice (Fig. 3D;  $P = 0.03$ ). Strikingly, only lesions in  $Pten^{G129E/+}$  female mice ( $n = 8$ ) advanced to form large palpable tumors with unique adenosquamous morphology and a prominent stromal component (Fig. 3E and Fig. S4B), which are classic characteristics of breast carcinoma seen in individuals with Cowden syndrome (22). Despite small sample sizes in our database comprising >400 *PTEN* mutation-positive individuals, all three *PTEN*<sup>G129E</sup>-affected women had breast cancer, whereas two *PTEN*<sup>C124R</sup>-affected individuals only had fibrocystic disease of the breast. Moreover, we also observed high-grade MIN in  $Pten^{G129E/+}$  male mice but not in males of other genotypes (Fig. S4C), consistent with an increased incidence of breast cancer observed in male patients with Cowden syndrome with this particular *PTEN* mutation (5, 23). The development of more severe lesions in select organs of  $Pten^{G129E/+}$  mice when compared with  $Pten^{\Delta4-5/+}$  mice suggests that the  $Pten^{G129E}$  allele, in addition to lacking lipid phosphatase activity, may have gained protumorigenic functions. Together, these findings demonstrate the critical role of specific *PTEN* mutant alleles in driving the broad clinical manifestations of Cowden syndrome that often present in an organ-selective fashion and with variable penetrance and expressivity.

**Organ-Specific *Pten* Inactivation During Tumor Initiation and Progression.** Molecular analysis of embryos and tumors with each of the three different *Pten* mutant genotypes revealed intriguing tissue-specific mechanisms of how *Pten* mutant alleles differentially contribute to tumor development and progression. As might have been expected, there was focal loss of WT Pten protein expression in  $Pten^{\Delta4-5/+}$  epithelial lesions of the endometrium, thyroid, and prostate (Fig. 4A). In the endometrial lesions analyzed, there was loss of heterozygosity (LOH) in 25% of cases (Fig. 4B). In the remaining endometrial lesions, as well as in the limited number of thyroid and prostate  $Pten^{\Delta4-5/+}$  tumors analyzed, LOH was not observed (Fig. 4B), consistent with the low percentage of LOH (11%) observed in Cowden syndrome-related nonmalignant neoplasias (24). These results suggest that the frequency of LOH in neoplastic lesions of all three mutant heterozygotes was low and that loss of expression from the WT *Pten* allele in the majority of  $Pten^{\Delta4-5/+}$  tumors was mediated by a posttranscriptional mechanism.

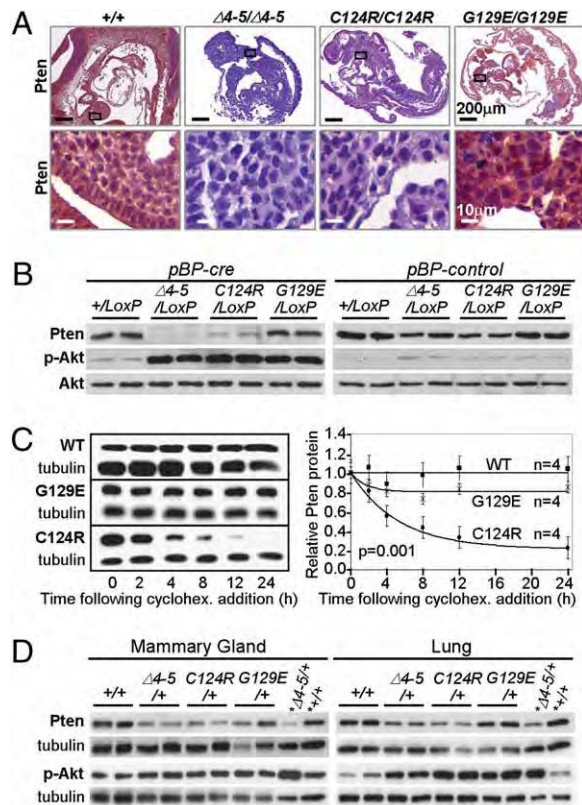
Interestingly, Pten protein in lesions of the uterus, thyroid, and prostate was frequently absent in hyperplastic clonal areas containing as few as three to five cells (Fig. 4A and Fig. S5 A–C), suggesting that loss of Pten expression in these organs was an early or initiating event during tumor development. In mammary preneoplastic lesions, however,  $Pten^{\Delta4-5/+}$  female mice retained normal amounts of Pten protein, and its loss, which was manifested

of LOH in lesions from the indicated organs. For LOH in the uterus, 25  $Pten^{\Delta4-5/+}$ , 5  $Pten^{C124R/+}$ , and 5  $Pten^{G129E/+}$  mice were examined. LOH in other organs was sampled from 5 mice per genotype group.

in a mosaic pattern, was only detected in advanced adenocarcinomas (Fig. 4A and Fig. S5D). When we examined epithelial cells from three adenocarcinomas in *Pten*<sup>Δ4-5/+</sup> and *Pten*<sup>G129E/+</sup> female mice, no LOH was detected in any of these samples. From these data, we conclude that the level of WT Pten protein that accumulates in *Pten*<sup>Δ4-5/+</sup> mammary glands is haploinsufficient to suppress tumor initiation, and that loss of expression from the WT allele is a late event that may be required for malignant progression of mammary cancers. Together, these observations support a tissue-selective requirement for loss of Pten protein expression.

#### Allele-Specific Mechanism of Pten-Mediated Tumor Suppression.

Inactivation of tumor suppression by the two missense mutations (*Pten*<sup>G129E</sup> and *Pten*<sup>C124R</sup>) is thought to result from the ablation of enzymatic phosphatase activity (1, 25, 26). It was therefore surprising to find that lesions in the uterus, thyroid, and prostate of *Pten*<sup>C124R/+</sup> mice had little (if any) detectable Pten protein (Fig. 4A). This was specific to this cohort of mice, because Pten protein in the corresponding lesions of *Pten*<sup>G129E/+</sup> mice was clearly



**Fig. 5.** Decreased *Pten*<sup>C124R</sup> protein stability. (A) Pten IHC staining of homozygous embryos (E9.5) with the indicated genotypes. (Lower) High magnification of boxed areas in the upper panels. (B) Western blots of control- (*pBP-control*) or *cre*-retrovirus- (*pBP-cre*) infected MEFs of the indicated genotypes were probed with antibodies specific for Pten, p-Akt S473, and total Akt. Complete *cre*-mediated deletion of exons 4 and 5 was confirmed by PCR (Fig. S6B). (C) (Left) Western blots of cycloheximide- (cyclohex.) treated MEFs treated as in B and expressing WT Pten (WT), *Pten*<sup>C124R</sup> (C124R), and *Pten*<sup>G129E</sup> (G129E). (Right) Protein half-lives of WT Pten (WT), *Pten*<sup>C124R</sup> (C124R), and *Pten*<sup>G129E</sup> (G129E) were calculated as described in *SI Text* and plotted. n, number of independent experiments. ANOVA with a repeated-measures model was used to evaluate differences across the various time points and between genotypes. (D) Western blots of lysates derived from mammary glands and lungs of two 9-month-old female mice with the indicated genotypes probed with Pten- and p-Akt (S473)-specific antibodies as indicated. \*Lysates derived from MEFs with the indicated genotypes were included in these blots. Tubulin protein levels were monitored as a loading control.

present. Interestingly, the levels of Pten protein in preneoplastic mammary gland lesions in *Pten*<sup>C124R/+</sup> mice were uniformly high, suggesting that, as in *Pten*<sup>Δ4-5/+</sup> mice, tumor initiation in mammary glands did not require LOH (note that carcinoma was never detected in this genetic group) (Fig. 4A). These results suggest that the two missense mutations (*Pten*<sup>G129E</sup> and *Pten*<sup>C124R</sup>) contribute to tumorigenesis differently, with the *Pten*<sup>C124R</sup> allele producing a protein product that is particularly labile.

The unexpectedly low levels of *Pten*<sup>C124R</sup> protein in tumors prompted us to reexamine its expression in normal tissues. We used immunohistochemistry (IHC) to compare Pten protein levels in E9.5 *Pten*<sup>+/+</sup>, *Pten*<sup>Δ4-5/Δ4-5</sup>, *Pten*<sup>G129E/G129E</sup>, and *Pten*<sup>C124R/C124R</sup> embryos. Remarkably, *Pten*<sup>C124R</sup> protein was almost undetectable in most tissues of homozygous embryos (Fig. 5A), even though its mRNA levels were not altered. In contrast, there were normal amounts of Pten protein expressed from the *Pten*<sup>G129E</sup> allele. To determine whether the absence of *Pten*<sup>C124R</sup> protein could be attributable to reduced protein stability, we generated and compared mouse embryonic fibroblasts (MEFs) containing a conditional *Pten*<sup>LoxP</sup> null allele (with *LoxP* sites flanking exons 4–5) and either a WT allele or each of the three mutant knockin alleles (*Pten*<sup>+/LoxP</sup>, *Pten*<sup>Δ4-5/LoxP</sup>, *Pten*<sup>G129E/LoxP</sup>, and *Pten*<sup>C124R/LoxP</sup>; Fig. S6A). Retroviral transduction of *cre* recombinase (*cre*) in the MEF cultures resulted in efficient deletion of *Pten*<sup>LoxP</sup> (Fig. S6B), generating cells that express *Pten* solely from the remaining WT allele or from each of the three knockin alleles (Fig. S6B). Western blot analysis confirmed normal steady-state levels of Pten protein in *cre-Pten*<sup>+/LoxP</sup> and *cre-Pten*<sup>G129E/LoxP</sup> MEFs (compared with *control-Pten*<sup>Δ4-5/LoxP</sup> MEFs), undetectable Pten in *cre-Pten*<sup>Δ4-5/LoxP</sup> MEFs, and decreased Pten in *cre-Pten*<sup>C124R/LoxP</sup> MEFs (Fig. 5B). Quantitative RT-PCR analysis showed equivalent amounts of *Pten* mRNA in MEFs of all genotypes (Fig. S6C). Pulse-chase experiments using cycloheximide demonstrated a significant reduction in the half-life of *Pten*<sup>C124R</sup> protein when compared with either WT or *Pten*<sup>G129E</sup> protein ( $P = 0.001$ ;  $\sim 3.7$  vs.  $>24$  h; Fig. 5C). From these results, we suggest that one mechanism by which the *Pten*<sup>C124R</sup> allele promotes Cowden syndrome and sporadic cancers is through the inherent instability of its protein product.

The activation of Akt, through its PI-3K-mediated phosphorylation at either residue 308 or 473, is thought to be attributable to the inactivation of Pten (1). Indeed, the level of phosphorylated Akt (p-Akt) was increased in *cre-Pten*<sup>Δ4-5/LoxP</sup> MEFs relative to either *cre-Pten*<sup>+/LoxP</sup> or *control-Pten*<sup>Δ4-5/LoxP</sup> MEFs. Interestingly, p-Akt levels were similarly increased in MEFs expressing *Pten*<sup>G129E</sup> or *Pten*<sup>C124R</sup> protein (Fig. 5B). Moreover, IHC revealed that E9.5 *Pten*<sup>Δ4-5/Δ4-5</sup>, *Pten*<sup>G129E/G129E</sup>, and *Pten*<sup>C124R/C124R</sup> embryos had equivalent levels of p-Akt, indicating that signaling downstream of PI-3K in embryos of all three mutant genotypes was similarly engaged (Fig. S1B). Because *Pten*<sup>Δ4-5/+</sup>, *Pten*<sup>G129E/+</sup>, and *Pten*<sup>C124R/+</sup> mice had different tumor and gender biases characteristic of Cowden syndrome with distinct pathological findings and frequencies, we investigated the extent of Akt phosphorylation in tumor samples. Again, IHC of all tumor types and Western blot analysis of lysates derived from mammary glands, lungs, and MEFs failed to reveal any difference in the amounts of p-Akt between the three genotypes (Figs. 4A and 5B and D). Therefore, the severity of cancer phenotypes attributed to *Pten*<sup>C124R</sup>, *Pten*<sup>G129E</sup>, and *Pten*<sup>Δ4-5</sup> alleles did not correspond to Akt activation (8, 27, 28). These data indicate that unique functions, beyond Akt activation, distinguish the three mutant alleles from one another in their ability to instigate cancer development and progression.

#### Discussion

Germline *PTEN* mutations cause a broad array of seemingly unrelated syndromes, which are collectively called the PTEN hamartoma tumor syndrome and include BRR and Cowden syndromes (9). Puzzlingly, certain identical mutations, such as

R130X and R233X, can result in either Cowden or BRR syndrome (29), suggesting that polymorphic heterogeneity strongly influences disease outcome. Even individuals with Cowden syndrome have a wide variation in disease manifestation and in the penetrance of phenotypes, including cancer predisposition (2, 12, 30). Whether this variability is attributable to gene mutations that differentially alter the function of Pten protein directly or to polymorphic variation and/or mutations in genetic modifiers of *PTEN* has been a subject of controversy (16).

Mouse modeling of three important mutant alleles of *PTEN* identified in patients with Cowden syndrome demonstrates that nonsense (*Pten*<sup>Δ4-5</sup>) and missense (*Pten*<sup>G129E</sup> and *Pten*<sup>C124R</sup>) mutations in *Pten* display distinct tumor phenotypes in organs targeted by Cowden syndrome. These functional differences are not manifested during embryonic development and do not appear to depend on differential PI-3K-Akt signaling. We have not formally ruled out, however, the possibility that the timing of Akt activation during tumor initiation/progression may differ between the three *Pten* knockin genetic groups. Nonetheless, it is interesting that phenotypic differences in human patients with Cowden syndrome also do not strictly correspond to the level of Akt activation. Rather, the striking differences in tumor spectrum and severity, with phenotypes in *Pten*<sup>G129E/+</sup> > *Pten*<sup>Δ4-5/+</sup> >> *Pten*<sup>C124R/+</sup> mice, may be partially accounted for by a gain of function in *Pten*<sup>G129E/+</sup> individuals and by reduced protein stability in *Pten*<sup>C124R/+</sup> individuals. The gain-of-function phenotypes manifested in *Pten*<sup>G129E/+</sup> mice may be attributable to the remaining protein phosphatase activity of mutant protein or to interference of function by mutant protein in protein complexes. Although further experimentation will be required to define fully the mechanisms underlying our observations, the *in vivo* models we describe provide the physiological and genetic foun-

ation necessary to unravel this puzzle. In summary, the results demonstrate that specific germline mutations disrupt *PTEN* function by distinct mechanisms that may explain the organ-selective variability associated with cancer predisposition in Cowden syndrome. These findings raise the possibility that specific somatic alterations, whether genetic or epigenetic, in *PTEN* will also have a differential impact on sporadic cancers that develop in patients who do not have Cowden syndrome. We believe that the genotype-phenotype relations revealed here may be used to guide more effective and personalized targeted therapies for cancer patients with *PTEN* mutations.

## Methods

*Pten*<sup>C124R/+</sup> and *Pten*<sup>G129E/+</sup> knockin mice were generated using a standard homologous recombination approach. Isolation of primary MEFs from E13.5 embryos, retroviral infections, and cell culture conditions were done using standard methods. *Pten*<sup>LoxP</sup> deletion by *cre* was confirmed by genotyping of MEF DNA and Western blot analysis using *PTEN* antibody. Statistics analyses were performed for all the data. Detailed methods can be found in *SI Text*.

**ACKNOWLEDGMENTS.** The authors thank John C. Thompson for excellent technical assistance, the Ohio State University Comprehensive Cancer Center Nucleic Acids and Transgenics Shared Facilities for technical assistance, and Dr. Philip Popovich and Wenmin Lai for guidance in the laser capture microdissection technique. This work was funded by National Institutes of Health Grants R01CA85619, R01HD47470, and P01CA097189 (to G.L.) and R01CA053271 and P01CA097189 (to M.C.O.). H.W. is supported by an American Cancer Society Postdoctoral Fellowship. M.C.O. received awards from both the Susan Komen Foundation and the Evelyn Simmers Foundation, and G.L. received an award from the Susan Komen Foundation and US Department of Defense. G.L. is the recipient of the Pew Charitable Trusts Scholar Award and the Leukemia and Lymphoma Society Scholar Award.

- Maehama T, Dixon JE (1998) The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem* 273: 13375–13378.
- Li J, et al. (1997) PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* 275:1943–1947.
- Myers MP, et al. (1997) P-TEN, the tumor suppressor from human chromosome 10q23, is a dual-specificity phosphatase. *Proc Natl Acad Sci USA* 94:9052–9057.
- Manning BD, Cantley LC (2007) AKT/PKB signaling: Navigating downstream. *Cell* 129: 1261–1274.
- Liaw D, et al. (1997) Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet* 16:64–67.
- Leslie NR, Maccario H, Spinelli L, Davidson L (2009) The significance of PTEN's protein phosphatase activity. *Adv Enzyme Regul* 49:190–196.
- Nelen MR, et al. (1997) Germline mutations in the PTEN/MMAC1 gene in patients with Cowden disease. *Hum Mol Genet* 6:1383–1387.
- Shen WH, et al. (2007) Essential role for nuclear PTEN in maintaining chromosomal integrity. *Cell* 128:157–170.
- Gustafson S, Zbuk KM, Scacheri C, Eng C (2007) Cowden syndrome. *Semin Oncol* 34: 428–434.
- Pezzolesi MG, Platzer P, Waite KA, Eng C (2008) Differential expression of PTEN-targeting microRNAs miR-19a and miR-21 in Cowden syndrome. *Am J Hum Genet* 82: 1141–1149.
- Trimboli AJ, et al. (2009) Pten in stromal fibroblasts suppresses mammary epithelial tumours. *Nature* 461:1084–1091.
- Marsh DJ, et al. (1999) PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. *Hum Mol Genet* 8:1461–1472.
- Suzuki A, et al. (1998) High cancer susceptibility and embryonic lethality associated with mutation of the PTEN tumor suppressor gene in mice. *Curr Biol* 8:1169–1178.
- Di Cristofano A, Pesce B, Cordon-Cardo C, Pandolfi PP (1998) Pten is essential for embryonic development and tumour suppression. *Nat Genet* 19:348–355.
- Podsypanina K, et al. (1999) Mutation of Pten/Mmac1 in mice causes neoplasia in multiple organ systems. *Proc Natl Acad Sci USA* 96:1563–1568.
- Freeman D, et al. (2006) Genetic background controls tumor development in PTEN-deficient mice. *Cancer Res* 66:6492–6496.
- Stambolic V, et al. (2000) High incidence of breast and endometrial neoplasia resembling human Cowden syndrome in pten<sup>+/-</sup> mice. *Cancer Res* 60:3605–3611.
- Risinger JI, Hayes AK, Berchuck A, Barrett JC (1997) PTEN/MMAC1 mutations in endometrial cancers. *Cancer Res* 57:4736–4738.
- Dahia PL, et al. (1997) Somatic deletions and mutations in the Cowden disease gene, PTEN, in sporadic thyroid tumors. *Cancer Res* 57:4710–4713.
- Steck PA, et al. (1997) Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet* 15: 356–362.
- Cairns P, et al. (1997) Frequent inactivation of PTEN/MMAC1 in primary prostate cancer. *Cancer Res* 57:4997–5000.
- Schrager CA, Schneider D, Gruener AC, Tsou HC, Peacocke M (1998) Clinical and pathological features of breast disease in Cowden's syndrome: An underrecognized syndrome with an increased risk of breast cancer. *Hum Pathol* 29:47–53.
- Fackenthal JD, et al. (2001) Male breast cancer in Cowden syndrome patients with germline PTEN mutations. *J Med Genet* 38:159–164.
- Marsh DJ, et al. (1998) Germline PTEN mutations in Cowden syndrome-like families. *J Med Genet* 35:881–885.
- Myers MP, et al. (1998) The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *Proc Natl Acad Sci USA* 95:13513–13518.
- Ramaswamy S, et al. (1999) Regulation of G1 progression by the PTEN tumor suppressor protein is linked to inhibition of the phosphatidylinositol 3-kinase/Akt pathway. *Proc Natl Acad Sci USA* 96:2110–2115.
- Vivanco I, et al. (2007) Identification of the JNK signaling pathway as a functional target of the tumor suppressor PTEN. *Cancer Cell* 11:555–569.
- Chang CJ, et al. (2008) PTEN nuclear localization is regulated by oxidative stress and mediates p53-dependent tumor suppression. *Mol Cell Biol* 28:3281–3289.
- Eng C (2003) PTEN: One gene, many syndromes. *Hum Mutat* 22:183–198.
- Eng C, Peacocke M (1998) PTEN and inherited hamartoma-cancer syndromes. *Nat Genet* 19:223.





## Tensor classification of $N$ -point correlation function features for histology tissue segmentation

Kishore Mosaliganti<sup>a,\*</sup>, Firdaus Janoos<sup>a</sup>, Okan Irfanoglu<sup>a</sup>, Randall Ridgway<sup>a</sup>, Raghu Machiraju<sup>a</sup>, Kun Huang<sup>b</sup>, Joel Saltz<sup>b</sup>, Gustavo Leone<sup>c</sup>, Michael Ostrowski<sup>c</sup>

<sup>a</sup>Department of Computer Science and Engineering, The Ohio State University, Columbus, OH 43210, USA

<sup>b</sup>Department of Biomedical Informatics, The Ohio State University, Columbus, OH 43210, USA

<sup>c</sup>Department of Human Cancer Genetics, The Ohio State University, Columbus, OH 43210, USA

### ARTICLE INFO

#### Article history:

Received 16 April 2007

Received in revised form 24 April 2008

Accepted 23 June 2008

Available online 25 July 2008

#### Keywords:

$N$ -point correlation functions

Phenotyping

Image segmentation

Microstructure

### ABSTRACT

In this paper, we utilize the  $N$ -point correlation functions ( $N$ -pcfs) to construct an appropriate feature space for achieving tissue segmentation in histology-stained microscopic images. The  $N$ -pcfs estimate microstructural constituent packing densities and their spatial distribution in a tissue sample. We represent the multi-phase properties estimated by the  $N$ -pcfs in a tensor structure. Using a variant of higher-order singular value decomposition (HOSVD) algorithm, we realize a robust classifier that provides a multi-linear description of the tensor feature space. Validated results of the segmentation are presented in a case-study that focuses on understanding the genetic phenotyping differences in mouse placenta.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

Developments in microscopy imaging technologies (Levenson and Hoyt, 2000, 2003) have generated large high-resolution datasets that have spurred medical researchers to conduct investigations into tissue organization, interfaces and internal cellular arrangements (Wu et al., 2003, 2007). Deeper insights into mechanisms of onset and growth of disease including cancer are now possible. For example, existing imaging modalities and imaging agents allow the microstructure characterization of the tumor microenvironment (Sloane et al., 2006). The tumor microstructure is best described by the composition and three-dimensional arrangement of cellular matrices, vasculature, and ducts that are embedded into a salient tissue structure. In this work, we focus on achieving tissue segmentation in histology images obtained from light microscopy by exploiting the properties of the microstructure. Segmentation of tissue layers when combined with other information (genetic and molecular expressions) will further the understanding of disease mechanisms (Ohtake et al., 2001).

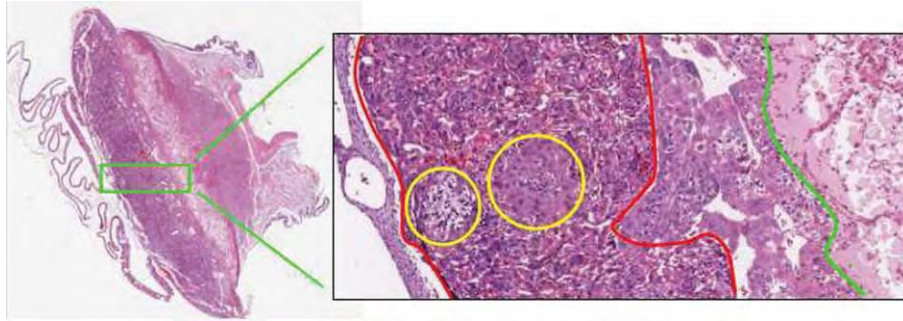
Tissue layers differ mainly in the spatial distributions and packing of microstructure components such as the red blood cells (RBCs), nuclei, extracellular matrix and background material. We algorithmically process high-resolution datasets to determine these distributions. These large images are generated from serial-section stacks that were digitized using light microscopy. Robust segmentation involves the discovery of feature spaces that estimate and spatially delineate component distributions, wherein the tissue layers naturally appear as salient clusters. The clusters can then be suitably classified.

Fig. 1 shows a typical mouse placenta section that we encounter in our work. The interface between two tissue types at full resolution is marked out using a red boundary. Note the clear lack of a well-defined boundary between the two tissue types. Instead, the change is better perceived by noting the subtle changes in microstructure properties (texture) that manifests within each tissue region. Fig. 2 shows another histology-stained section of the mouse mammary tissue. The epithelial cell lining surrounding a duct has a characteristic packing arrangement. The material surrounding the duct consists of fat cells arranged in honeycomb-like cellular matrices. Our examples serve to illustrate that the tissues are best identified by the relative packing densities and spatial distributions of nuclei, RBCs, extracellular material and background components.

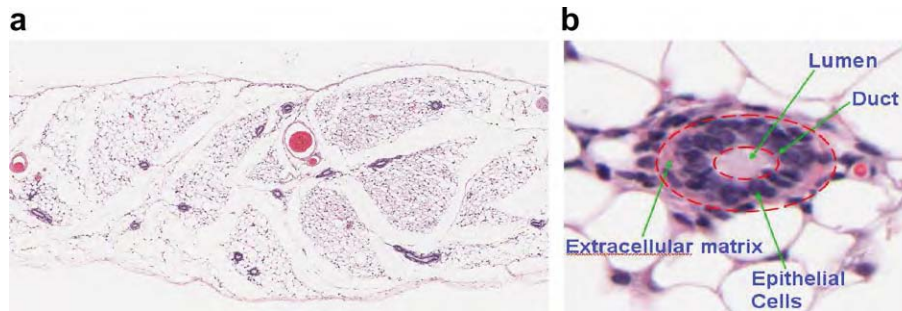
In Section 3, we propose the use of  $N$ -point correlation functions ( $N$ -pcfs) borrowed from the material science literature (Torquato, 2004) for image segmentation. These functions efficiently

\* Corresponding author. Tel.: +1 614 292 4029; fax: +1 614 292 2911.

E-mail addresses: [mosaligk@cse.ohio-state.edu](mailto:mosaligk@cse.ohio-state.edu) (K. Mosaliganti), [janoos@cse.ohio-state.edu](mailto:janoos@cse.ohio-state.edu) (F. Janoos), [irfanogl@cse.ohio-state.edu](mailto:irfanogl@cse.ohio-state.edu) (O. Irfanoglu), [ridgwayr@cse.ohio-state.edu](mailto:ridgwayr@cse.ohio-state.edu) (R. Ridgway), [raghu@cse.ohio-state.edu](mailto:raghu@cse.ohio-state.edu) (R. Machiraju), [khuang@bmi.ohio-state.edu](mailto:khuang@bmi.ohio-state.edu) (K. Huang), [joel.saltz@osumc.edu](mailto:joel.saltz@osumc.edu) (J. Saltz), [gustavo.leone@osumc.edu](mailto:gustavo.leone@osumc.edu) (G. Leone), [michael.ostrowski@osumc.edu](mailto:michael.ostrowski@osumc.edu) (M. Ostrowski).



**Fig. 1.** A slice of the placenta tissue showing the labyrinth–spongiotrophoblast interface (red). The yellow circles show glycogen tissue embedded within the labyrinth. The green boundary is the interface between the spongiotrophoblast and the remaining maternal layers. Note the subtle change in tissue microstructure across the boundaries.



**Fig. 2.** A histology section of the mouse mammary gland showing several duct cross-sections. A zoomed duct section reveals the characteristic packing densities of the epithelial cell lining.

characterize the microstructure in a heterogeneous substrate (Gokhale, 2004, 2005, 2006). The  $N$ -pcfs estimate a feature tensor at each pixel location. The tensor encodes information relevant to the spatial proximity of different components relative to each other. In order to classify the tensors into salient tissue classes, we invoke higher-order singular value decomposition (de Lathauwer et al., 2000) methods. This allows us to identify fundamental descriptors that best explain the unique distributions spanned by a salient tissue in the tensor space.

Finally, in Section 4, we apply these methods to a genetic phenotyping study and provide rigorous validation proof. Using manual ground-truth, we compare the performance of the tensor classification framework with the  $k$ -nearest neighbor method. The segmentations from using the  $N$ -pcf measures are compared with the Haralick and Gabor measures. Additionally, the automated segmentations on the entire image stack are compared against manual segmentations. Thus, we obtain effective high-resolution tissue segmentations as shown in Fig. 14 (placental labyrinth) and Fig. 15.

## 2. Related work

Medical research has focussed on trying to understand the cellular and molecular mechanisms of transformation especially in cytogenetics (Streicher et al., 2000) and cancer pathology (Braumann et al., 2005, 2006) by utilizing high-resolution microscopy. Imaging studies have had a high impact in terms of providing automated and objective methods for comparative analysis (Wenzel et al., 2007). In this context, pertinent research in serial-section visualization focusses on (i) extracting/ segmenting relevant 2D features (neurons, vasculature, ductal profiles, etc.) and regions (labyrinth layer, mammary epithelial tissue, etc.) (ii) performing automated serial-section alignment (Fiala and Harris, 2002, 2005, 2006), and (iii) evaluating 3D structure from 2D profiles

(Jeong and Radke, 2006, 2005, 2007). The work presented in (Oh-take et al., 2001, 2004) follows this mantra successfully to get 3D tangible reconstructions of duct/vasculature anatomy. In this work, we address 2D segmentation challenges that these tasks require as a precursor to 3D reconstructions.

Direct volume rendering using transfer functions has a limited scope in the visualization of the raw image stacks since it requires the presence of 3D image gradients and color variations. In serial-section microscopy, information is available on a per image basis with little or no correspondence of structures in subsequent slides owing to thick sections (Fiala and Harris, 2002) that are often acquired. Our data had in-plane and out-plane resolutions of 0.5  $\mu\text{m}$  and 5  $\mu\text{m}$ , respectively. Furthermore, all the tissue layers resemble a multi-phase sample having the same microstructural components. They vary in the presentation of ensemble properties. Hence, a segmentation of the various regions (Fiala, 2005) is first required before meaningful visualizations are obtained.

Segmentation techniques reported for histology image segmentation in the literature may be categorized into two specific classes: color-based and image-based. Color space methods target features of interest that are easily discriminated and extracted solely on the basis of color at every pixel or local region. For example, in our case, the nuclei (dark), red blood cells/blood vessels (scarlet), and extracellular material/neoplasm (mauve) appear with distinctive hues in the standard H&E histological staining protocol. In an earlier effort, Pan and Huang (2005) devised a Bayesian supervised segmentation method incorporating image features such as windowed color and gradient histograms into a long feature vector for classification. This approach did not exploit the spatial arrangement patterns of the microstructural components.

The image-based methods seek to outline salient regions in the image. Often, their goal is to detect boundaries or interfaces between different regions. Hence, they come under the broader class of region segmentation algorithms. The level-set methods (Malladi

et al., 1995, 2004), active contour models (Caselles et al., 1997), Gibbs models (Chen and Metaxas, 2005), watershed methods (Beucher, 1991), and texture analysis methods (Haralick et al., 1973) are prime examples of this approach in the medical imaging literature. They have found popular use in separating overlapping nuclei clusters, cellular constellations, ductal and vasculature pathways. In (Huang and Murphy, 2004), the authors use an extended Haralick feature set to locate sub-cellular patterns in fluorescence microscopy images. Our images do not present easily discernible boundaries (Figs. 1 and 2). Salient regions are identified by the subtle changes in ensemble properties. Therefore, we adapt methods from material science literature for obtaining region segmentations.

The quantitative characterization of spatial distributions of finite-size objects or points in multidimensional spaces has been studied in several disciplines such as spatial statistics (Torquato, 2004, 1985), materials science (Gokhale, 2004, 2005), signal processing (Aste and Weaire, 2000), biology (Zou and Wu, 1995), physics (Chandrasekhar, 1996), and astronomy (Babu and Feigelson, 1943). Microstructure irrespective of its origin (material science or biological sections) may be defined as a collection (ensemble) of points, lines, internal surfaces, and volumes (Gokhale et al., 2005). Each microstructural feature is associated with size, shape, volume, surface area, length, curvature attributes, etc., morphological orientation, and location. Statistical distributions of such geometric attributes of ensembles of microstructural features collectively specify the geometric state of a microstructure. Mathematically, these properties of the microstructure are formalized by the statistical  $N$ -point correlation functions ( $N$ -pcfs) (Stoyan et al., 1985). There are fundamental geometric constraints that are enforced while realizing  $N$ -pcfs. These constraints, therefore, provide useful user-input in choosing a certain functional form that is best representative of the microstructure. A formal introduction to  $N$ -pcfs with applications in material characterization studies can be found in Torquato (Torquato, 2004). In practice, the 2-pcf is most useful for microstructure representation (Saheli et al., 2004, 2005). Recently, a digital image analysis based technique has been developed for the realistic computer simulation or reproduction of microstructure from measurements of the two-point correlation functions on 2D sections (Singh et al., 2006).

The  $N$ -pcf features are similar to the co-occurrence matrices of Haralick et al. (1973) in some sense. However, significant conceptual as well implementation differences exist. For example, the computation of the  $N$ -pcfs is randomized and does not exploit the pixel grid structure. These functions are used to compute other physical properties of the tissue substrate (such as porosity, etc.). The functions capture both the geometry and statistical nature of textural regions. Given the inherent generality of the  $N$ -pcfs, we explore their use in segmenting light microscopy images.

The work describe in this paper is a culmination of several related efforts. Ridgway et al. (2006) initially introduced the  $N$ -pcfs for histology image segmentation in conjunction with the HOSVD classifier. Some initial promising results were observed on individual 2D images. The segmentation results were however

validated on a single marked image. Janoos et al. (2007) extended the above approach to incorporate multi-resolution scale-space strategies to achieve better time performance. They note that the coarse segmentation results at lower resolutions as a trade-off against better time performance. In this work, we provide a consolidated and rigorously validated solution on large serial-section stacks.

### 3. Segmentation algorithm

We describe our segmentation framework as a consequence of three processing stages. Please refer to Fig. 3 for a flowchart.

- (i) *Identifying homogeneous phase components*: At the outset, we identify the microstructural components namely, the nuclei, RBC, extracellular material and background from color images using a standard Gaussian maximum likelihood estimation (MLE) framework (Section 3.1).
- (ii) *Estimating component distributions*: We treat a slice as a multi-phase material wherein each tissue layer can be independently analyzed for the ensemble properties. The tissue regions present salient packing and spatial distributions of the components that are measured by the  $N$ -pcfs. In Section 3.2, the  $N$ -pcfs are estimated using a sliding window strategy that is applied throughout the image to yield a feature tensor at each pixel location.
- (iii) *Tensor classification*: The  $N$ -pcf features of a tissue sample are naturally expressed as  $N + 1$  mode (order  $N + 1$ ) tensors. In Section 3.3, we decompose the tensor feature space for a given tissue type by using the higher-order singular value decomposition method on training data. Novel tissue regions are then projected onto a lower-dimensional tensor space and classified.

#### 3.1. Color segmentation

The RGB pixel data is color-classified to determine the individual microstructural components in the image (denoted as  $I$ ). We use a standard Gaussian MLE classifier to label each pixel  $p \in I$  as belonging to one of the four classes, namely, (i) nuclei, (ii) RBC, (iii) extracellular material, and (iv) white background. The MLE algorithm assumes that the histograms of the bands of data have normal distributions (shown in Fig. 4a).

The *a priori* information related to the four classes is learnt from training data. Pixel wise labeled data was generated using a random sampling of the given image. A custom-built application randomly displays patches of the training image, and highlights the center pixel. The user then chooses between red blood cells, cytoplasm, background, nuclei, and pass. The spatial location, RGB triplet values, and the user chosen class are used as class attributes. Covariance matrices ( $\Sigma_i$ ), mean ( $\mu_i$ ), and prior probability weights ( $a_i$ ) are then calculated for each individual class.

$$P_i(p) = \ln(a_i) - [0.5 \ln(|\Sigma_i|)] - [0.5(I(p) - \mu_i)^T \Sigma_i^{-1} (I(p) - \mu_i)], \quad (1)$$

$$\pi(p) = \arg \max_i P_i(p).$$

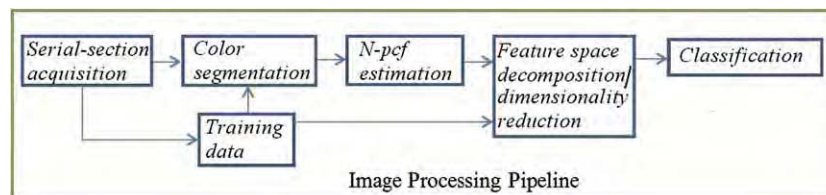
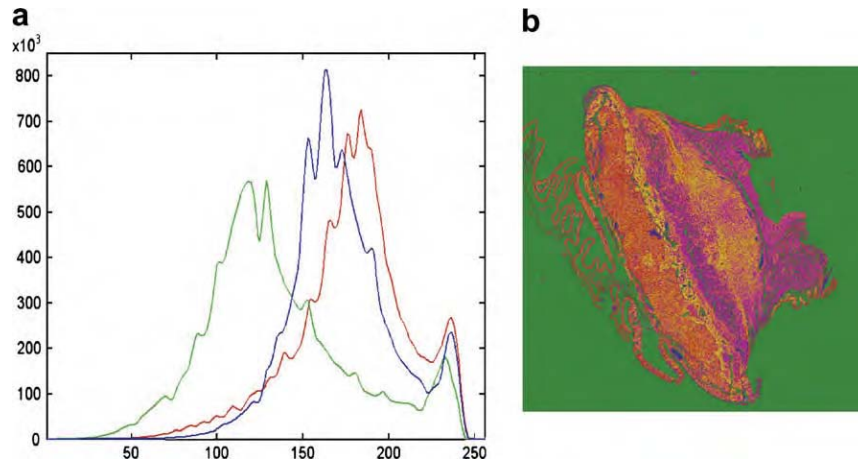


Fig. 3. Segmentation pipeline for serial-section stacks.



**Fig. 4.** (a) RGB histograms of the placenta section are shown to be approximately having a normal distribution. (b) Component labeled image-C of the original RGB placenta image-I shown in Fig. 1.

Eq. 1 determines the probability ( $P_i(p)$ ) associated with a pixel  $p$  towards a class  $i$ . The maximum logarithmic probability rule is invoked to determine the final class membership  $\pi(p)$ . The result of this classification is a material component labeled image that is shown in Fig. 4b. Having identified the microstructural components in the image, we would like to measure their relative packing and spatial distributions. The functions defined in the next section describe an elegant way to do so.

### 3.2. $N$ -point correlation functions

The histology images normally are composed of four component phases as explained earlier. However, to simplify the presentation here, assume the presence of only two phases in an image  $I$ , namely, 0 (black) and 1 (white). Examples of such image textures are shown in Fig. 6. Consider placing a  $N$ -sided regular polyhedron with edge length  $\ell$  in  $I$ . The probability that all the  $N$ -vertices lie in phase 0 is defined as an  $N$ -point correlation function ( $N$ -pcf),  $P_{i_1 i_2 \dots i_N}^\ell$ , where  $i_m = 1$ . The subscript  $i_m$  denotes the phase of the  $m$ th polyhedron vertex. The  $N$ -pcf for a regular polyhedron of edge length  $\ell$  depends on its orientation  $(\theta, \phi)$  and location in the 3D space of the microstructure. The orientation averaged  $N$ -pcf  $P_{ij}^\ell$  can be computed from the corresponding direction-dependent functions  $\tilde{P}_{ij}^\ell(\theta, \phi)$  as follows:

$$\langle P_{i_1 i_2 \dots i_N}^\ell \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \tilde{P}_{i_1 i_2 \dots i_N}^\ell(\theta, \phi) d\theta d\phi. \quad (2)$$

We now provide some insight into the probability measures captured by these functions. Consider the simple case of a 1-pcf, say  $P_0$ . It represents the probability that a point  $p$  is in phase 0. This quantity measures the volume fraction of phase 0 in the microstructure. Similarly,  $P_1$  is the volume fraction of phase 1 and we have:

$$P_0 + P_1 = 1. \quad (3)$$

A 2-pcf is the probability of a straight line segment of length  $\ell$  randomly placed in the microstructure such that one end is in phase  $i_1 \in \{0, 1\}$  and the other end is in phase  $i_2 \in \{0, 1\}$ . For a two-phase microstructure, there are four possible 2-pcfs namely  $P_{00}^\ell$ ,  $P_{01}^\ell$ ,  $P_{10}^\ell$  and  $P_{11}^\ell$  and

$$P_{00}^\ell + P_{01}^\ell + P_{10}^\ell + P_{11}^\ell = 1; \quad P_{01}^\ell = P_{10}^\ell; \\ P_{00}^\ell + P_{00}^\ell = f_0; \quad P_{10}^\ell + P_{11}^\ell = f_1.$$

Parameters  $f_0$  and  $f_1$  represent the volume fractions of the individual phases. Similarly, the 3-pcf descriptor of the material ensemble

is  $P_{i_1 i_2 i_3}^\ell$ , where  $i_1, i_2, i_3 \in \{0, 1\}$  are the phase indices of the three points, and  $\ell$  is the separation distance between them (the three points describe an equilateral triangle whose side has length  $\ell$ ) and

$$\sum_{p=0}^2 \sum_{q=0}^2 \sum_{r=0}^2 P_{pqr}^\ell = 1; \quad P_{pqr}^\ell = P_{rpq}^\ell; \quad \sum_{q=0}^2 \sum_{r=0}^2 P_{pqr}^\ell = f_p. \quad (5)$$

Note that each individual texture class in an image may provide a unique or characteristic  $N$ -pcf feature measure for a certain value of the separation distance  $\ell$ . The presentation of these characteristic values makes the texture class to be easily identified. It is not known *a priori* what these values of  $\ell$  are for a given image. Hence, in practice, a range of values need to be explored while estimating these functions for the given image. The set of possible integral values that  $\ell$  may assume is represented by the discrete set  $K \subset \mathbb{Z}$  and the set of all component phases by  $Q \subset \mathbb{Z}$ . The  $N$ -pcf feature descriptor for a tissue region represented by  $P_{i_1 i_2 \dots i_N}^{(k)} \in \mathfrak{R}^{K \times Q^N}$  is an  $N + 1$  mode tensor.

#### 3.2.1. Algorithms for the 2-pcf and 3-pcf

Essentially, a  $N$ -pcf is a multivariate distribution function. To estimate this function, we resort to using samples. Conceptually, for a given separation length  $\ell$ , one needs to create auto- and cross-histograms at every point in the discrete image plane. Therefore, Monte Carlo methods were used to sample the distributions of material components using a sliding window. To estimate the functions at a given point in the image, the number of samples ( $S$ ) and window sizes ( $\Omega$ ) need to be specified. The minimum window size is proportional to the maximum distance that the functions will be evaluated for, and the sample size is chosen to keep the variance of the measured result to a desired range.

To evaluate the 2-pcf for separation distance  $\ell$  and phases  $i, j$  in a region containing  $m$ -phases, we place a number of randomly oriented and positioned lines. We then count the fraction of line segments that have an end point in phase  $i$  and the other in  $j$  to give an estimate of the 2-pcf  $P_{ij}^\ell$ . A similar procedure is used to evaluate the 3-pcf  $P_{ijl}^\ell$  among phases  $i, j$  and  $l$ .

#### 3.3. HOSVD-based tensor decomposition

We briefly describe the HOSVD decomposition. More details can be found in (de Lathauwer et al., 2000). Consider the  $M$ -mode tensor  $\mathbb{A} \in \mathfrak{R}^{I_1 \times I_2 \times \dots \times I_M}$  where the dimension of the mode  $i$  is  $I_i$ . The column vectors are referred to as mode-1 vectors and row vectors as mode-2. The mode- $n$  vectors are the column vectors of matrix

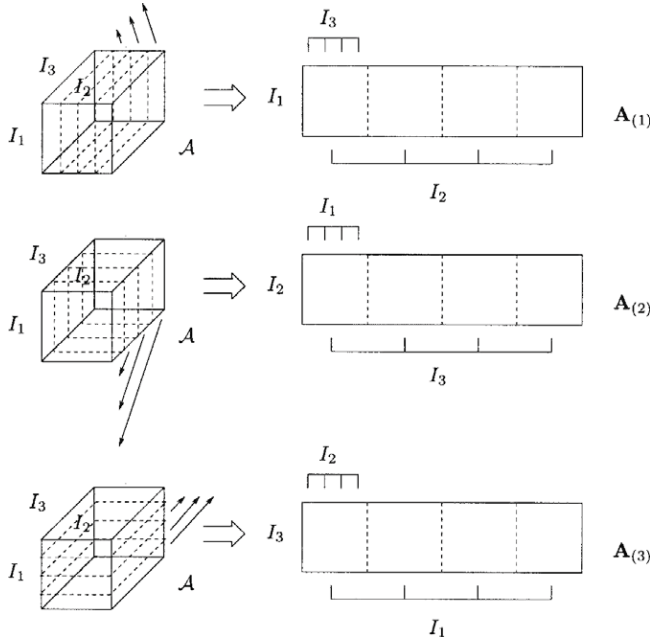


Fig. 5. Unfolding the 3-mode tensor  $\mathbb{A} \in \mathfrak{R}^{I_1 \times I_2 \times I_3}$  to the matrices  $\mathbf{A}_{(1)} \in \mathfrak{R}^{I_1 \times (I_2 \times I_3)}$ ,  $\mathbf{A}_{(2)} \in \mathfrak{R}^{I_2 \times (I_1 \times I_3)}$  and  $\mathbf{A}_{(3)} \in \mathfrak{R}^{I_3 \times (I_1 \times I_2)}$  (from de Lathauwer et al. (2000)).

$\mathbf{A}_{(n)} \in \mathfrak{R}^{I_n \times (I_1 \times I_2 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_M)}$  obtained by flattening  $\mathbb{T}$  along the  $n$ th mode (see Fig. 5).

The mode- $n$  product of a matrix  $\mathbf{X} \in \mathbb{R}^{n \times n}$  with a tensor  $\mathbb{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_n \times \dots \times I_M}$  is denoted by  $\mathbb{A} \times_n \mathbf{X}$  and results in the tensor  $\mathbb{B} \in \mathfrak{R}^{I_1 \times I_2 \times \dots \times I_{n-1} \times I_n \times I_{n+1} \times \dots \times I_M}$ . Its entries are  $b_{i_1 \dots i_{n-1} j_n i_{n+1} \dots i_M} = \sum_{i_n} a_{i_1 \dots i_{n-1} i_n i_{n+1} \dots i_M} x_{j_n i_n}$ . In terms of the flattened matrices,  $\mathbf{B}_{(n)} = \mathbf{X} \mathbf{A}_{(n)}$ .

Singular value decomposition (SVD) is a 2-mode tool commonly used in signal processing to reduce the dimensionality of the space and thereby reduce noise. SVD decomposes a matrix  $\mathbf{A}$  into three other matrices,  $\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$ , where the columns of  $\mathbf{U}$  spans the row space of  $\mathbf{A}$  and  $\mathbf{V}$  spans the column space of  $\mathbf{A}$ , while  $\mathbf{S}$  is a diagonal matrix of singular values. By a similar extension, the  $N$ -mode SVD or higher-order SVD (HOSVD) decomposes the multi-linear space spanned by tensor  $\mathbb{A}$  yielding a core tensor and orthonormal matrices spanning the vector spaces of each mode of the tensor, i.e.,

$$\mathbb{A} = \mathbb{S} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times \dots \times_M \mathbf{U}_M \quad (6)$$

The core tensor  $\mathbb{S}$  is analogous to the diagonal singular value matrix from traditional SVD and coordinates the interaction of matrices to produce the original tensor. Additionally, the Frobenius-norm of the sub-tensors of  $\mathbb{S}$  estimates the variance of the corresponding part of the original tensor. This could be used to reduce the feature space.

**Dimensionality reduction:** Matrices  $\mathbf{U}_i$  are orthonormal and their columns span the space of the flattened tensor  $\mathbf{T}_{(i)}$ . The row vectors of  $\mathbf{U}_i$  are the coefficients describing each dimension along mode  $i$ .<sup>1</sup> After conducting the HOSVD on a given image data, the dimensionality with respect to any mode can be reduced independently, unlike the PCA where dimensionality reduction is only based on the variance. By reducing the number of dimensions in one mode, we can selectively control how that mode explains the original space, and eliminate noise from each mode separately. The dimensions also are reduced by removing the last  $m$ -column vectors from the tensor flattened along the desired mode. The error after dimensionality

<sup>1</sup> They are similar to the coefficients extracted from PCA but there exist different sets of coefficients for each mode in a typical HOSVD analysis. Please refer to Vasilescu and Terzopoulos (2003) for details.

reduction is bounded by the Frobenius norm of the corresponding hyper-planes in the core tensor.

### 3.4. Classification

Let  $C$ ,  $L$  and  $K$  be discrete sets of tissue classes, training samples for each class and separation distances used in the computation of the 2-pcfs. As described in the preceding sections, for each distance  $\ell \in K$  used in the 2-pcf computation, we extract a  $Q \times Q$  correlation matrix, where  $Q = 4$  is the number of material phases in the tissue.<sup>2</sup> We pack all these features, into a 5-mode tensor  $\mathbb{A} \in \mathfrak{R}^{C \times L \times K \times Q \times Q}$ . The first mode corresponds to the tissue classes, the second mode to the training instances, third mode to the separation distances, and fourth and fifth mode to vertex phases in the 2-pcf. This tensor is decomposed as

$$\mathbb{A} = \mathbb{S} \times_1 \mathbf{U}_{C \times C}^{\text{classes}} \times_2 \mathbf{U}_{N \times N}^{\text{trg}} \times_3 \mathbf{U}_{K \times K}^{\text{dist}} \times_4 \mathbf{U}_{Q \times Q}^{\text{pt}_1} \times_5 \mathbf{U}_{Q \times Q}^{\text{pt}_2}. \quad (7)$$

The row vectors of  $\mathbf{U}_{C \times C}^{\text{classes}}$  are the coefficients describing each class and are used in classification.

Let  $\mathbb{X}_{K \times Q \times Q}$  be the 2-pcf feature tensor for one training instance, with class  $c \in C$  and training instance number  $l \in L$ . Let  $\mathbf{u}_{1 \times C}^{\text{classes}}(c)$  be the  $j$ th row in the class matrix  $\mathbf{U}_{1 \times C}^{\text{classes}}$  (row =  $j$ ) describing class  $j$  and  $\mathbf{u}_{1 \times L}^{\text{trg}}(l)$  be the  $l$ th row vector of the instance matrix,  $\mathbf{U}_{1 \times L}^{\text{trg}}$  (row =  $l$ ), describing instance  $l$  in the training data. Then by projecting into the decomposed space in Eq. (7), we are able to reconstruct

$$\mathbb{X}_{K \times Q \times Q} = \left[ \left( \mathbb{S} \times_3 \mathbf{U}_{K \times K}^{\text{dist}} \times_4 \mathbf{U}_{Q \times Q}^{\text{pt}_1} \times_5 \mathbf{U}_{Q \times Q}^{\text{pt}_2} \right) \times_2 \mathbf{u}_{1 \times L}^{\text{trg}}(l) \right] \times_1 \mathbf{u}_{1 \times C}^{\text{classes}}(c).$$

For a test feature tensor  $\mathbb{Z}_{K \times Q \times Q}$ , the goal is now to find the class coefficients  $\mathbf{z}_{1 \times C}$  that will minimize its reconstruction error. Then, the reconstructed feature tensor is

$$\widehat{\mathbb{Z}}_{K \times Q \times Q} = \left( \mathbb{S} \times_3 \mathbf{U}_{K \times K}^{\text{dist}} \times_4 \mathbf{U}_{Q \times Q}^{\text{pt}_1} \times_5 \mathbf{U}_{Q \times Q}^{\text{pt}_2} \times_2 \mathbf{u}_{1 \times L}^{\text{trg}}(l) \right) \times_1 \mathbf{z}_{1 \times C}^{\text{classes}}, \quad (8)$$

$$\text{and let } \mathbb{M}_{C \times 1 \times K \times Q \times Q}^{(l)} = \left( \mathbb{S} \times_3 \mathbf{U}_{K \times K}^{\text{dist}} \times_4 \mathbf{U}_{Q \times Q}^{\text{pt}_1} \times_5 \mathbf{U}_{Q \times Q}^{\text{pt}_2} \times_2 \mathbf{u}_{1 \times L}^{\text{trg}}(l) \right), \quad (9)$$

$$\text{therefore, } \widehat{\mathbb{Z}}_{K \times Q \times Q} = \mathbb{M}_{C \times 1 \times K \times Q \times Q}^{(l)} \times_1 \mathbf{z}_{1 \times C}^{\text{classes}}. \quad (10)$$

The best class coefficients  $\mathbf{z}_{1 \times C}^{\text{classes}}$  are obtained by minimizing the following objective function

$$\varepsilon^{(l)} = \sum_i \sum_j \sum_k \left( Z_{ijk} - \widehat{Z}_{ijk} \right)^2 \quad \forall l \in [1, L]. \quad (11)$$

The solution to this optimization problem turns out to be the solution to the linear system:  $\mathbf{A}^{(l)} \mathbf{x}^{(l)} = \mathbf{b}^{(l)}$ , where

$$\mathbf{A}_{pq}^{(l)} = \sum_i \sum_j \sum_k \left( \mathbb{M}_{p1ijk}^{(l)} \times \mathbb{M}_{q1ijk}^{(l)} \right), \quad (12)$$

$$\text{and } b_p^n = \sum_i \sum_j \sum_k \left( \mathbb{M}_{p1ijk}^{(l)} \times Z_{ijk} \right) \quad \text{where } p, q \in [1, C]. \quad (13)$$

The best coefficient set  $\mathbf{z}_{1 \times C}$  is obtained by first selecting the training instance  $l$  that minimizes the error  $\varepsilon^{(l)}$ , and then computing the corresponding  $\mathbf{z}_{1 \times C}$ .

## 4. Results

### 4.1. $k$ -Nearest neighbor classification using the 2,3-pcf features

Here, we illustrate the working of the  $N$ -pcf features on very simple images with binary components. The  $N$ -pcfs were evaluated using three synthetic images with well-defined textures shown in Fig. 6. The images consisted of two phases represented as black (phase 0) and white (phase 1) regions. The 2-pcfs were evaluated

<sup>2</sup> Nuclei, RBCs, cytoplasm, and background.

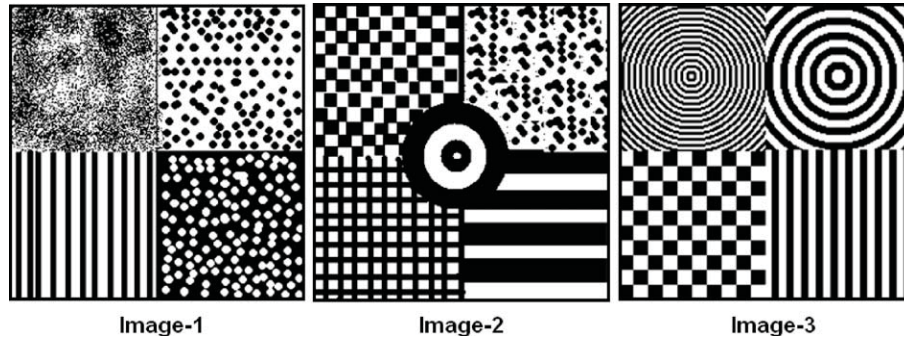


Fig. 6. Synthetic images constructed two phases namely the black and white areas. Each image is composed of different texture regions that depend on the spatial distribution of these two phases.

using line segments having varying lengths  $\ell \in [1, 25]$  pixel units. Similarly, the 3-pcfs were evaluated using equilateral triangles with edge lengths  $\ell \in [1, 25]$  pixel units. The random sampling parameter was set as  $S = 500$  with image windows  $\Omega$  of size  $40 \times 40$ .

A  $k$ -nearest neighbor classifier ( $k = 5$ ) was used to identify the different texture types based on the observed 2-pcf and 3-pcf evaluations. This relatively simple classifier to study the feature space spanned by the  $N$ -pcfs. The training set for the classifier consisted of hundred regions randomly chosen from each texture type. For example, image-1 presented four different textures. Therefore,

the training set consisted of the 2-pcf and 3-pcf output on 400 regions. Classification was performed by considering the 2-pcf and the 3-pcf features independently. We present our results in Figs. 7–9. The algorithm generated region labels (along the column) are tabulated against the ground-truth labels (along the rows).

The accuracies for each texture class in image-1 (Fig. 7) was over 95% for both the 2-pcf and the 3-pcf features. At the same time, the rate of false identifications was low (<5%). The only notable exception was observed in class 3 (the vertical bars). The vertical bars fared poorly when evaluated using the 2-pcf feature set having a false positive rate of 9.1%. Most of the inaccuracies were

 2-pcf		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Acc</b>
	<b>Class 1</b>	15699	250	273	34	96.5%
	<b>Class 2</b>	233	14801	697	16	93.9
	<b>Class 3</b>	171	345	15748	120	96.1%
	<b>Class 4</b>	0	153	608	15623	95.3
	<b>FP</b>	2.5%	4.8%	9.1%	1%	
 3-pcf		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Acc</b>
	<b>Class 1</b>	15865	64	327	0	97.5%
	<b>Class 2</b>	87	15602	53	5	99%
	<b>Class 3</b>	119	199	15993	73	97.6%
	<b>Class 4</b>	0	209	378	15797	96.4%
	<b>FP</b>	1.2%	2.9%	4.5%	0.4%	

Fig. 7. Classification results tabulated against the ground-truth on synthetic image-1.

 2-pcf		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Class 5</b>	<b>Acc</b>
	<b>Class 1</b>	13775	212	474	0	164	94.1%
	<b>Class 2</b>	104	13783	534	114	253	93.2%
	<b>Class 3</b>	317	277	13197	35	29	95.2%
	<b>Class 4</b>	3	73	149	11025	3328	75.6%
	<b>Class 5</b>	0	2	1	261	5476	95.4%
	<b>FP</b>	2.9%	3.9%	8.0%	3.5%	40.8%	
 3-pcf		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Class 5</b>	<b>Acc</b>
	<b>Class 1</b>	13963	316	163	13	170	95.4%
	<b>Class 2</b>	1310	13150	28	163	137	88.9%
	<b>Class 3</b>	358	92	13393	11	1	96.6%
	<b>Class 4</b>	272	145	3	10284	3874	70.5%
	<b>Class 5</b>	23	24	0	439	5254	91.5%
	<b>FP</b>	12.3%	4.2%	1.4%	5.7%	4.4%	

Fig. 8. Classification results tabulated against the ground-truth on synthetic image-2.



 <b>2-pcf</b>		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Acc</b>
	<b>Class 1</b>	9594	81	0	27	98.8%
	<b>Class 2</b>	42	7988	38	1632	82.3%
	<b>Class 3</b>	109	60	9453	278	95.4%
	<b>Class 4</b>	0	895	13	9192	91.0%
	<b>FP</b>	1.5%	11.4%	0.5%	17.4%	
 <b>3-pcf</b>		<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Acc</b>
	<b>Class 1</b>	9328	273	95	6	96.1%
	<b>Class 2</b>	2	8553	583	562	88.1%
	<b>Class 3</b>	60	96	9600	144	96.9%
	<b>Class 4</b>	0	226	110	9764	96.6%
	<b>FP</b>	0.6%	6.5%	7.5%	6.7%	

Fig. 9. Classification results tabulated against the ground-truth on synthetic image-3.

confined to the boundaries since textures are not well-defined locally. Window patches at texture boundaries contain material from different tissues. Hence, the microstructure properties as captured by the  $N$ -pcf are representative of an intermediate material.

Image-2 suffered from higher inaccuracies but still maintained acceptable performance (>82%) for four of the five textures (Fig. 8). The performance was relatively poor on the portion of the image with wide horizontal bars. The classification of the bars was often confused with that of the circles in the center of the image. The ambiguity occurred primarily at the boundaries. This was mainly a result of the features not being sufficiently discriminating at the scale of distances ( $\ell$ ) considered in the  $N$ -pcf features. As a consequence, this resulted in a classification accuracy ranging from 70% to 75% for the horizontal bars and false positive rates of 3.5–6%. It should be noted that the false positive rates for the circles in the center were higher than the normal 4–50%. The 3-pcf features performed much better than the 2-pcf. The high false positive rates are attributed to the same reasons of scale that was discussed above. This experiment shows that incorporating the right scale of  $N$ -pcf features is important in obtaining good accuracies while maintaining a low false positive rate.

Classification accuracy in image-3 was within the range of 82.3–98.8% (Fig. 9). Significantly, both the best and worst performance occurred while using the 2-pcf features. The thin concentric circles was identified with 98.8% accuracy while the wider concentric circles were identified with an accuracy of 82.3%. The lack of accuracy in the latter case was the result of identifying the wide concentric circles near the image boundary as vertical bars. This is well illustrated in Fig. 10 where circles at large radii are similar to vertical bars. False positives were limited to the range of 0.5% and 21.6%. From this experiment, we observe that the 2-pcf features performed better as compared to the 3-pcf features which is a change from the trend observed in image-2 wherein the 3-pcf features fared better.

We now make the following observations about the performance of  $N$ -pcfs:



Fig. 10. Window patches from (a) texture class 2 and (b) texture class 4. Note the similarity in organization observed locally.

- (1) The textures regions are well-identified in the interiors. Ambiguity arise at the boundaries where the texture is not well-defined locally. The resolution depends on the window sizes  $\Omega$  considered.
- (2) The correct scale of separation distances need to be considered. A wrong choice can lead to ambiguity across different texture classes.
- (3) While the 2-pcf and the 3-pcf output trends are correlated in most textures, some cases cause one of the measures to perform relatively better. Hence, it is useful to consider both the measures in any given image.

We now apply our segmentation framework to datasets acquired from phenotyping studies.

#### 4.2. Mouse placenta phenotyping

A central issue in human cancer genetics requires the understanding of how the genotype change (e.g., gene mutations and knockout) affects phenotype (e.g., tissue morphology or animal behavior) and is valuable in the development of therapies to treat the diseases such as cancers. In common experimental conditions, the genotype change can be well controlled. Therefore, it is useful to quantitatively assess the corresponding phenotype change in the organism. Serial-section imaging (Fiala and Harris, 2002, 2003, 2005, 2007) is one of the widely used phenotyping tools for making large-scale objective and automated analysis. We have applied our methods to two phenotyping studies that required a significant image analysis component. Image analysis is capable of providing viable biomarkers that can be used towards hypothesis generation.

We implemented our framework using the National Library of Medicine's (NIH/ NLM) Insight Segmentation and Registration Toolkit (ITK) (Yoo et al., 2002) and the Visualization Toolkit (VTK) from Kitware Inc. The classified volumetric datasets are loaded into Kitware's VolView volume visualization software to render the surface appropriately in both the case-studies. All our tasks were conducted on a 2.5 GHz Pentium machines running Linux with 1 GB main memory. The segmentation software processed a placenta slice completely in under 15 min. The mammary sections required 3–5 min owing to the use of a single 2-pcf feature.

Mouse placentas are composed of three distinct layers: the labyrinth, spongiotrophoblast, and glycogen layers as shown in Fig. 1. The phenotyping experiments (Wu et al., 2003, 2007) required the quantification of morphological parameters related to the surface area, volume of the labyrinth and its interface with the spongio-

trophoblast layer. Hence, segmentation was deemed necessary and important.

The microscopic images are obtained from the standard histologically stained slides of both a wild-type and a mutated ( $Rb^-$ ) mouse placenta. The slides are collected by sectioning the wax fixed sample (placenta) at  $3\mu\text{m}$  thickness. They are then digitized using a an Aperio Scanscope digitizer using a  $20\times$  magnification objective lens, producing effective magnification of  $200\times$  under which each image is of size between 500 MB and 1 GB. A total of more than 2000 images are obtained (1278 images for the wild-type and 786 images for the  $Rb^-$ ) with total file size of approximately 1.7 TB.

4.3. Segmenting the placental tissue layers

The first step consisted of identifying the tissue components. These were identified by modeling each component as a Gaussian distribution in the RGB color space. The modeling was performed by using the pixel training data gathered by the method described in Section 3. The 2 and 3-pcfs were evaluated at distances  $k \in [1, 24]$  pixel units. A window size  $\Omega$  of  $51 \times 51$  dimensions was considered with sample size  $S = 1000$ . Since we were interested in comparing the utility of the  $N$ -pcf features with those of Haralick and Gabor features, classification was accomplished using the  $k$ -NN scheme ( $k = 5$ ). Training data for the  $k$ -NN method was obtained as follows:

*Training data:* Image tissue patches of size  $20 \times 20$  were generated and labeled as either labyrinth (Lab), spongiotrophoblast (SP), Glycogen (Glyc), or background (BG). A total of 2200 regions were selected from one image slide (800 for labyrinth, 800 for spongiotrophoblast, and 600 for the background). A total of 150 of these regions were used in training (50 for each region) and the rest is used for testing.

In Fig. 11, we observe that the  $N$ -pcfs identified and labeled the labyrinth layer with an acceptable level of accuracy ( $\approx 96\%$ ) and a low false positive rate of 8.5–10.5%. The spongiotrophoblast was also classified with an accuracy of  $\approx 96\%$  and a false positive rate of 6–8%. The background material is identified with an accuracy of 99–100%. Glycogen was identified with an accuracy of only  $\approx 50\%$ . This was because the glycogen regions are smaller than the window size. They have almost the same microstructure characteristics as the labyrinth layer (refer to Fig. 1). They are often difficult to perceive even to the human eye.

The Haralick features fared well on the labyrinth layer (95%) and the spongiotrophoblast (89%) but still presented slightly lower values than the  $N$ -pcf features. The Gabor features performed poorly only delivering a classification accuracy of 53% for the labyrinth and 80% for the spongiotrophoblast. The correlation func-

tions meanwhile had the lowest rate of false positives of all of the features.

The better performance observed with the material science measures is a result of the component hierarchy used in the technique. While the Gabor and Haralick features are concerned with the luminance and co-occurrence of the image features, the newer measures are able to leverage the knowledge of the components and spatial distributions for better segmentations.

4.4. Comparing the  $k$ -nearest neighbor and the HOSVD classification

In this section, we compare the performance of the simple  $k$ -NN and HOSVD classification scheme using the same labeled data that was prepared for the experiments in Section 4.3.

The resulting confusion matrix while using the  $k$ -NN classification on the  $N$ -pcf features is shown in Table 1. The  $k$ -NN classifier achieves above 90% in the labyrinth and the spongiotrophoblast regions. Notably, the  $k$ -NN algorithm performs very well in the spongiotrophoblast regions and achieves a classification accuracy of 96.8% but provides a high false positive rate of detection. Meanwhile, the HOSVD classifier without any dimensional reduction performs comparably to the  $k$ -NN classifier in Table 2. HOSVD improves the results to 94% in the labyrinth region but the accuracy decreases to 93.9% in the spongiotrophoblast areas. Nevertheless, the false positive rate of detection is consistently lower. The classi-

Table 1  
Confusion matrix entries for  $k$ -NN

	Lab	SP	BG	Acc (%)
Lab	684	66	0	91.2
SP	61	726	0	96.8
BG	0	0	550	100
FP (%)	8.1	8.3	0	–

Acc, accuracy; FP, false positives; Lab, labyrinth layer; ST, spongiotrophoblast; BG, background.

Table 2  
Confusion matrix entries for HOSVD classification with no dimensional reduction/35 dimensions/9 dimensions, respectively

	Lab	SP	BG	Acc (%)
Lab	705 717 617	45 33 47	0	94 95.6 92.3
SP	46 53 30	704 697 720	0	93.9 92.9 96.3
BG	0	0	550	100
FP (%)	6 6.8 4.6	6 4 6.1	0	–

Acc, accuracy; FP, false positives; Lab, labyrinth layer; ST, spongiotrophoblast; BG, background.

<b>2-Point</b>	Lab	ST	Glyc	BG	Acc	<b>3-Point</b>	Lab	ST	Glyc	BG	Acc
<b>Lab</b>	6876	215	106	0	95.5%	<b>Lab</b>	6921	184	92	0	96.1%
<b>ST</b>	413	6529	256	0	90.7%	<b>ST</b>	397	6589	212	0	91.5%
<b>Glyc</b>	239	385	726	0	53.7%	<b>Glyc</b>	432	252	666	0	49.3%
<b>BG</b>	0	0	0	5396	100%	<b>BG</b>	0	0	0	5396	100%
<b>FP</b>	8.6%	8.4%	33.2%	0%		<b>FP</b>	10.6%	6.2%	31.3%	0%	
<hr/>											
<b>Haralick</b>	Lab	ST	Glyc	BG	Acc	<b>Gabor</b>	Lab	ST	Glyc	BG	Acc
<b>Lab</b>	6853	189	155	0	95.2%	<b>Lab</b>	3843	2878	376	103	53.3%
<b>ST</b>	731	6432	35	0	89.3%	<b>ST</b>	802	5822	168	408	80.8%
<b>Glyc</b>	865	68	417	0	30.8%	<b>Glyc</b>	711	457	152	30	11.2%
<b>BG</b>	4	15	0	5377	99.6%	<b>BG</b>	32	147	10	5211	96.5%
<b>FP</b>	18.9%	4.0%	31.3%	0%		<b>FP</b>	28.6%	37.4%	78.4%	9.4%	

Fig. 11. Comparison of the placenta segmentation output using four different feature sets. Acc, accuracy; FP, false positives; Lab, labyrinth layer; ST, spongiotrophoblast; Glyc, glycogen; BG, background.



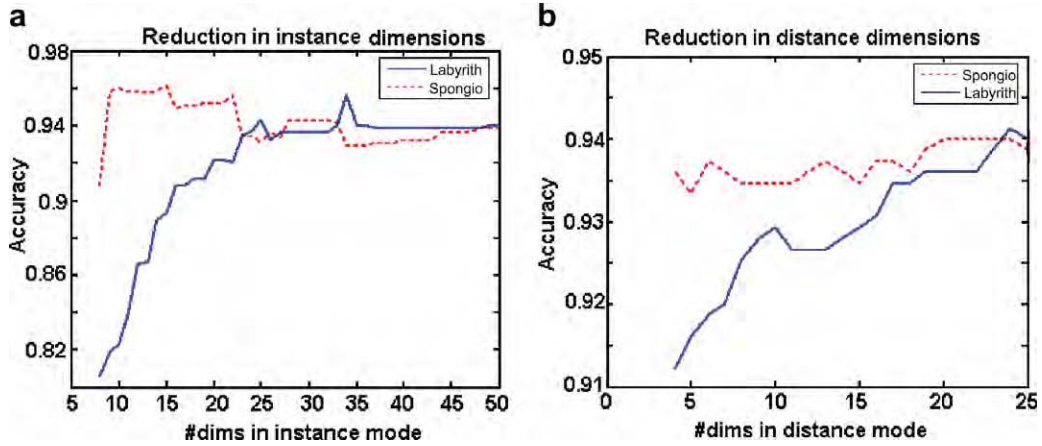


Fig. 12. Classification accuracies observed after dimensionality changes in (a) instances mode alone and (b) distance mode alone.

fication accuracy for the background is at 100% with 0% false positive rate for both the classifiers.

Fig. 12a shows the classification accuracies observed for all the four classes when the dimension is reduced in the instance mode alone. The accuracy of detection sharply rises for the labyrinth

**Table 3**  
Confusion matrix for combined dimensional reduction using the HOSVD scheme

	Lab	SP	BG	Acc (%)
Lab	692	58	0	92.30
SP	35	715	0	95.30
BG	0	0	550	100
FP	5.0	8.11	0	

Acc, accuracy; FP, false positives; Lab, labyrinth layer; ST, spongiotrophoblast; Glyc, Glycogen; BG, background.

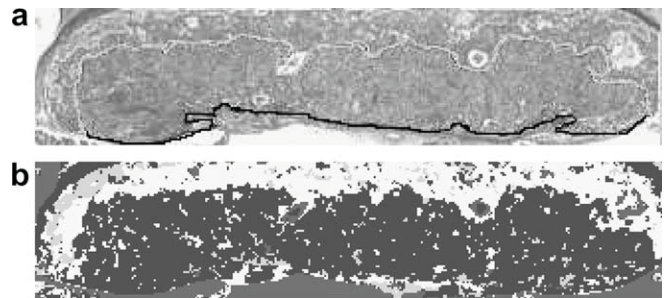


Fig. 14. (a) Cropped placenta image with the boundary between labyrinth and maternal layer outlined in black, the interface between labyrinth and spongiotrophoblast marked as white, and glycogen marked as gray. (b) Segmentation using *N*-pcfs with the labyrinth as dark gray, glycogen as medium gray and spongiotrophoblast as light gray.

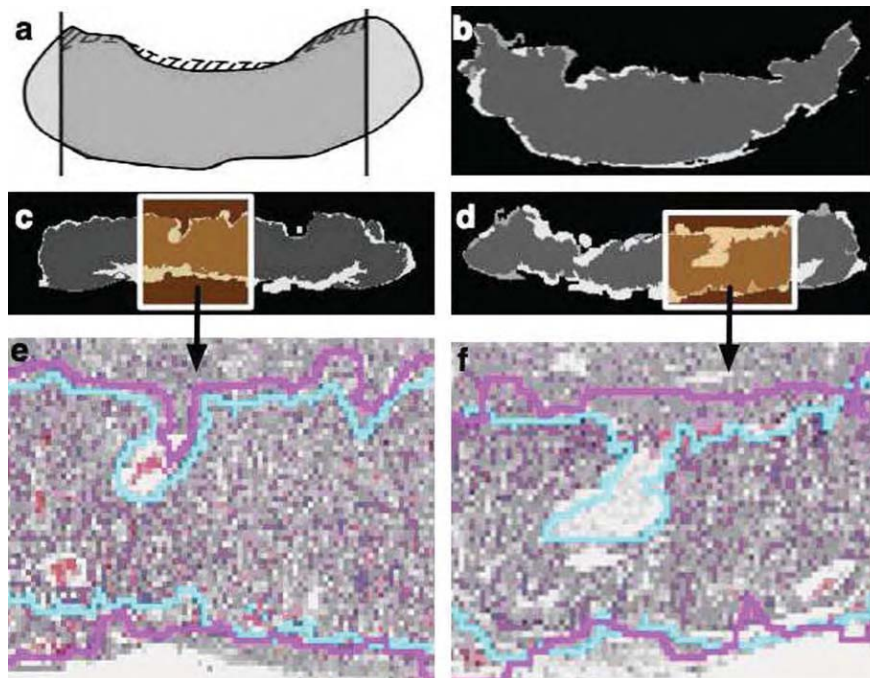


Fig. 13. Evaluation of the automatic segmentation algorithm. (a) The solid line is the manually marked boundary and the dashed line is the automatic segmentation result. (b–d) Examples of images with boundary estimation errors being 2.5%, 8.4%, and 16.5%. (e and f) A larger view of the difference between manual segmentation (cyan) and automatic segmentation (purple).

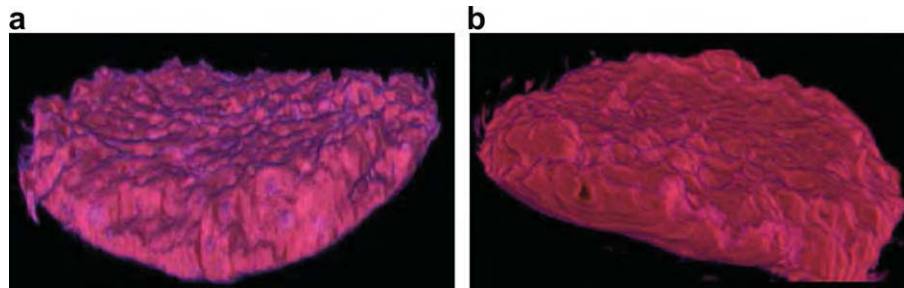


Fig. 15. Volume rendering of the binary labyrinth segmentation masks for (a) wild-type and (b) mutant placenta.

layer when 35 instances per class are considered. Meanwhile, the spongiotrophoblast detection is better while using around nine samples. These numbers indicate that the variance of features describing spongiotrophoblast is mostly due to the noise, whereas the variance for the labyrinth is due to the variation in the data. We present the confusion matrices for both the cases in Table 2 (35 and 9, respectively).

Fig. 12b shows the classification accuracies observed for all the four classes when the dimension is reduced in the distance mode alone. The classification accuracy for the labyrinth layer decreases monotonically while the spongiotrophoblast remains unaffected. Finally, a choice that would serve both the labyrinth and spongiotrophoblast regions equally may be obtained by reducing the dimensionality by 30 in instances mode and by two in distance mode. This setup produces the confusion matrix shown in Table 3.

To summarize our results:

- (1) The high accuracy in detection of the  $k$ -NN classifier are offset by the presence of high false positive rate.
- (2) The HOSVD scheme with dimensional reduction eliminates the modes that do not explain the tensor feature space and retains the significant ones. Hence, it helps in eliminating noise and cause better data separation.
- (3) Our results indicate that dimensionality reduction helps in providing the user more control in obtaining better segmentations on a particular region-of-interest as a trade-off against other regions.

In addition to the standard validation, we also measure the efficacy of the framework by inspecting the labyrinth–spongiotrophoblast interface.

#### 4.5. Labyrinth–spongiotrophoblast interface validation

Our framework processed four pairs (eight in total) placenta datasets. Validation was carried out on three placenta datasets. For each placenta, manual segmentation of the labyrinth layer was carried out on ten images that are evenly spaced throughout the image stack. In Fig. 13, the automatically segmented labyrinth is overlaid on the manually segmented labyrinth tissue. For all the manually segmented images, the error is measured as the ratio of area difference between the two segmentations to the area of the manual labyrinth segmentation. Formally, let  $I_m$  and  $I_a$  represent the manually and automatically segmented images. The boundary estimation error is defined as  $\frac{\text{Area}(|I_m - I_a|)}{\text{Area}(I_m)}$ . For the three samples, the boundary estimation errors are 6.61.6%, 5.33.3%, and 16.77.4%. The errors are quite low for two of the three placentas given the fact that the validation has been conducted across the serial-section stack. As shown in Fig. 13e and f, the error can be attributed to two major reasons: (i) large window sizes in the  $N$ -pcf algorithm

leads to a boundaries that suffer from a stair-case effect and (ii) discrepancy in assigning the large white areas on the boundary. The latter reason also explains why the automatic segmentation works less reliably in the case of the third placenta with the largest error (16.77.4%). The other two placenta (one wild-type and one mutant) with mean error less than 7% were then used for visualization. Fig. 14 shows a placenta with marked up boundaries and the final segmentation result. Fig. 15 shows the pair of selected wild-type and mutant placentas whose labyrinth layer is rendered as a binary volume using a simple transfer function.

## 5. Summary

In this paper, we described a tissue segmentation algorithm that is applicable in histology images obtained from serial-section microscopy. We estimated the packing and material component distribution locally using the  $N$ -point correlation functions. These functions are realized using suitable windowing and sampling strategies to provide tensor feature representations. Multi-linear properties of the tensor feature space are then extracted using a variant of the HOSVD algorithm. The algorithm allows the reduction of the dimensionality of the tensor feature space with respect to the significant modes thereby achieving robust classification. Our methods have been applied in a mouse model phenotyping study requiring the segmentation of the labyrinth tissue layer in the placenta. The classifier performance and the final segmentation output was validated extensively using manual ground-truth. In future, we would like to extend the feature set with additional spatial proximity measures that efficiently represent the microstructure. These measures would address some of the challenges in the segmentation of large histology images. Additionally, we would like to investigate possible enhancements in performance that may be obtained from utilizing efficient structures such as the quadtrees and  $kd$ -trees on large images.

## References

- Aste, T., Weaire, D., 2000. The Pursuit of Perfect Packing. Institute of Physics Publishing.
- Babu, G., Feigelson, E., 1943. Stochastic problems in physics and astronomy. *Reviews of Modern Physics* 15, 1–89.
- Beucher, S., 1991. The watershed transformation applied to image segmentation. In: *Conference on Signal and Image Processing in Microscopy and Microanalysis*, pp. 299–314.
- Braumann, U., Kuska, J., Einkenkel, J., Horn, L., Löffler, M., Höckel, M., 2005. Three-dimensional reconstruction and quantification of cervical carcinoma invasion fronts from histological serial sections. *IEEE Transactions in Medical Imaging* 24 (10), 5–7.
- Caselles, V., Kimmel, R., Sapiro, G., 1997. Geodesic active contours. *International Journal on Computer Vision* 22 (1), 61–97.
- Chandrasekhar, S., 1996. Spatial point processes in astronomy. *Journal of Statistical Planning and Inference* 50 (3), 311–326.
- Chen, T., Metaxas, D., 2005. A hybrid framework for 3D medical image segmentation. *Medical Image Analysis* 9, 547–565.
- Cooper, L., Huang, K., Sharma, A., Mosaliganti, K., Pan, T., 2006. Registration vs. reconstruction: building 3-D models from 2-D microscopy images. In:

- Proceedings of Workshop on Multiscale Biological Imaging, Data Mining and Informatics, pp. 57–58.
- de Lathauwer, L., de Moor, B., Vandewalle, J., 2000. A multilinear singular value decomposition. *SIAM Journal of Matrix Analysis and Applications* 21 (4), 1253–1278.
- Fernandez-Gonzalez, R., Deschamps, T., Idica, A., Malladi, R., de Solorzano, C., 2004. Automatic segmentation of histological structures in mammary gland tissue sections. *Biomedical Optics* 9 (3), 444–453.
- Fiala, J., 2005. Reconstruct: a free editor for serial section microscopy. *Journal of Microscopy* 218, 52–61.
- Fiala, J., Harris, K., 2002. Computer-based alignment and reconstruction of serial sections. *Microscopy and Analysis* 52, 5–7.
- Fridman, Y., Pizer, S., Aylward, S., Bullitt, E., 2004. Extracting branching tubular object geometry via cores. *Medical Image Analysis* 8, 169–176.
- Gokhale, A., 2004. Experimental measurements and interpretation of microstructural  $n$ -point correlation functions. *Microscopy and Microanalysis* 10, 736–737.
- Gokhale, A., Tewari, A., Garmestani, H., 2005. Constraints on microstructural two-point correlation functions. *Scripta Materialia* 53, 989–993.
- Haralick, R., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 610–621.
- Huang, K., Murphy, R.F., 2004. Automated classification of subcellular patterns in multicell images without segmentation into single cells. In: *IEEE International Symposium of Biomedical Imaging (ISBI 2004)*, pp. 1139–1142.
- Huang, A., Nielson, G., Razdan, A., Farin, G., Baluch, D., Capco, D., 2006. Thin structure segmentation and visualization in three-dimensional biomedical images: a shape-based approach. *IEEE Transactions in Visualization and Computer Graphics* 12 (1), 93–102.
- Janoos, F., Irfanoglu, O., Mosaliganti, K., Machiraju, R., Huang, K., Wenzel, P., deBruin, A., Leone, G., 2007. Multi-resolution image segmentation using the 2-point correlation functions. In: *IEEE International Symposium of Biomedical Imaging*.
- Jeong, Y., Radke, R., 2006. Reslicing axially sampled 3D shapes using elliptic fourier descriptors. *Medical Image Analysis* 11, 197–206.
- Levenson, R., Hoyt, C., 2000. Spectral imaging and microscopy. *American Laboratory* 32 (1), 26–34.
- Malladi, R., Sethian, J.A., Vermuri, B.C., 1995. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (2), 158–174.
- Mosaliganti, K., Machiraju, R., Heverhagen, J., Saltz, J., Knopp, M., 2005. Exploratory segmentation using geometric tessellations. In: *Proceedings of 10th International Fall Workshop on Vision, Modeling and Visualization*, pp. 1–8.
- Ohtake, T., Kimijima, I., Fukushima, T., Yasuda, H., Sekikawa, M., Takenoshita, S., Abe, R., 2001. Computer-assisted complete three-dimensional reconstruction of the mammary ductal/lobular systems. *Cancer* 91 (12), 2263–2314.
- Pan, T., Huang, K., 2005. Virtual mouse placenta: tissue layer segmentation. In: *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Ridgway, R., Irfanoglu, O., Machiraju, R., Huang, K., 2006. Image segmentation with tensor-based classification of  $N$ -point correlation functions. In: *MICCAI Workshop on Medical Image Analysis with Applications in Biology*.
- Saheli, G., Garmestani, H., Adams, B., 2004. Microstructure design of a two phase composite using two-point correlation functions. *Computer-Aided Materials Design* 11 (2–3), 103–115.
- Sharp, R., Ridgway, R., Mosaliganti, K., Wenzel, P., Pan, T., Bruin, A., Machiraju, R., Huang, K., Leone, G., Saltz, J., 2007. Volume rendering phenotype differences in mouse placenta microscopy data. *Special Issue on Anatomic Rendering and Visualization, Computing in Science and Engineering* 9 (1), 38–47.
- Singh, H., Mao, Y., Sreeranganathan, A., Gokhale, A., 2006. Application of digital image processing for implementation of complex realistic particle shapes/morphologies in computer simulated heterogeneous microstructures. *Modeling and Simulation in Material Science and Engineering* 14, 351–363.
- Sloane, B., Gillies, R., Mohla, S., Sogn, J., Menkens, A., Sullivan, D., 2006. I2 imaging: cancer biology and the tumor microenvironment. *Cancer Research* 66 (23), 11097–11101.
- Stephens, D., Allan, V., 2003. Light microscopy techniques for live cell imaging. *Science* 300 (5616), 82–86.
- Stoyan, D., Kendall, S., Mecke, J., 1985. *Stochastic Geometry and its Applications*. John Wiley and Sons, New York.
- Streicher, J., Donat, M., Strauss, B., Spörle, R., Schughart, K., Müller, G., 2000. Computer-based three-dimensional visualization of developmental gene expression. *Nature Genetics* 25, 147–152.
- Sundararaghavan, V., Zabaras, N., 2005. Classification and reconstruction of threedimensional microstructures using support vector machines. *Computational Materials Science* 32, 223–239.
- Torquato, S., 2004. *Random Heterogenous Material*. Springer Verlag.
- Vasilescu, M.A.O., Terzopoulos, D., 2003. Multilinear subspace analysis for image ensembles. In: *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR'03)*, vol. 2, pp. 93–99.
- Wenzel, P., Wu, L., deBruin, A., Chen, W., Dureska, G., Sites, E., Pan, T., Sharma, A., Huang, K., Ridgway, R., Mosaliganti, K., Sharp, R., Machiraju, R., Saltz, J., Yamamoto, H., Cross, J., Robinson, M., Leone, G., 2007. Rb is critical in a mammalian tissue stem cell population. *Genes and Development* 21 (1), 85–97.
- Wu, L., de Bruin, A., Saavedra, H.I., Starovic, M., Trimboli, A., Yang, Y., Opavska, J., Wilson, P., Thompson, J., Ostrowski, M., Rosol, T., Woollett, L., Weinstein, M., Cross, J., Robinson, M., Leone, G., 2003. Extra-embryonic function of Rb is essential for embryonic development and viability. *Journal of Nature* 421, 942–947.
- Yoo, T., Ackerman, M., Lorenson, W., Schroeder, W., Chalana, V., Aylward, S., Metaxas, D., Whitaker, R., 2002. Engineering and algorithm design for an image processing api: a technical report on itk—the insight toolkit. In: *Proceedings of the 10th Annual Medicine Meets Virtual Reality (MMVR)*, pp. 586–592.
- Zou, G., Wu, H., 1995. Nearest-neighbor distribution of interacting biological entities. *Theoretical Biology* 172 (4), 347–353.



# Cancer Research

## An *Ets2*-Driven Transcriptional Program in Tumor-Associated Macrophages Promotes Tumor Metastasis

Tahera Zabuawala, David A. Taffany, Sudarshana M. Sharma, et al.

*Cancer Res* 2010;70:1323-1333. Published OnlineFirst February 9, 2010.

### Updated Version

Access the most recent version of this article at:  
doi:[10.1158/0008-5472.CAN-09-1474](https://doi.org/10.1158/0008-5472.CAN-09-1474)

### Supplementary Material

Access the most recent supplemental material at:  
<http://cancerres.aacrjournals.org/content/suppl/2010/02/12/0008-5472.CAN-09-1474.DC1.html>

### Cited Articles

This article cites 42 articles, 18 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/70/4/1323.full.html#ref-list-1>

### E-mail alerts

[Sign up to receive free email-alerts](#) related to this article or journal.

### Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

### Permissions

To request permission to re-use all or part of this article, contact the AACR Publications Department at [permissions@aacr.org](mailto:permissions@aacr.org).

## An *Ets2*-Driven Transcriptional Program in Tumor-Associated Macrophages Promotes Tumor Metastasis

Tahera Zabuawala<sup>1,2,7</sup>, David A. Taffany<sup>1,7</sup>, Sudarshana M. Sharma<sup>1,7</sup>, Anand Merchant<sup>1,7</sup>, Brett Adair<sup>2,7</sup>, Ruchika Srinivasan<sup>1,7</sup>, Thomas J. Rosol<sup>3,7</sup>, Soledad Fernandez<sup>6,7</sup>, Kun Huang<sup>4,7</sup>, Gustavo Leone<sup>2,5,7</sup>, and Michael C. Ostrowski<sup>1,2,7</sup>

### Abstract

Tumor-associated macrophages (TAM) are implicated in breast cancer metastasis, but relatively little is known about the underlying genes and pathways that are involved. The transcription factor *Ets2* is a direct target of signaling pathways involved in regulating macrophage functions during inflammation. We conditionally deleted *Ets* in TAMs to determine its function at this level on mouse mammary tumor growth and metastasis. *Ets2* deletion in TAMs decreased the frequency and size of lung metastases in three different mouse models of breast cancer metastasis. Expression profiling and chromatin immunoprecipitation assays in isolated TAMs established that *Ets2* repressed a gene program that included several well-characterized inhibitors of angiogenesis. Consistent with these results, *Ets2* ablation in TAMs led to decreased angiogenesis and decreased growth of tumors. An *Ets2*-TAM expression signature consisting of 133 genes was identified within human breast cancer expression data which could retrospectively predict overall survival of patients with breast cancer in two independent data sets. In summary, we identified *Ets2* as a central driver of a transcriptional program in TAMs that acts to promote lung metastasis of breast tumors. *Cancer Res*; 70(4): 1323–33. ©2010 AACR.

### Introduction

Sporadic human cancer results from somatic gene mutations that lead to aberrant growth, survival, genetic instability, and increased motility of tumor cells (1). In addition to genetic complexity, it is increasingly apparent that cellular complexity inherent in the tumor stroma plays an active role in promoting all stages of tumor progression (2). Among the many cell types in the tumor stroma, the tumor-associated macrophage (TAM) is a broadly defined myeloid cell type that has been implicated in tumor progression (2). TAMs are thought to be a polarized M2 subtype of macrophage that promotes tumor growth, invasion, and angiogenesis (3). Alternatively, the pleiotropic effects of macrophages within the tumor microenvironment may be mediated by distinct

subpopulations of TAMs that can selectively affect distinct processes such as tumor angiogenesis or invasion (4, 5).

The link between TAMs and tumor progression is especially well-established in breast cancer. Human clinical studies have shown that a high focal infiltration of TAMs in primary human breast tumors correlates directly with tumor cell invasion, increased vascularization, axillary lymph node involvement, and reduced relapse-free survival of patients (6–9). In a mouse mammary tumor model, genetic ablation of colony-stimulating factor-1 (*Csf-1*), a growth factor critical for macrophage growth, differentiation, and survival, results in a reduction in mammary TAMs and a lower incidence of lung metastasis (10).

*Ets2*, a member of the *Ets* family of transcription factors, is a direct effector of CSF-1 signaling pathways that modulates macrophage functions and survival during inflammation (11, 12). *ETS2* activates or represses the transcription of target genes in a context-dependent manner (13, 14). Elevated expression of *ETS2* has been correlated with human breast cancer (15). However, in mouse mammary tumor models, *Ets2* promotes tumor progression from the stroma and not the tumor epithelial cell (16).

In the current study, a genetic approach was used to define the action of *Ets2* in mouse mammary TAMs. The results show that *Ets2* in TAMs decreased the growth rate of the primary tumor and tumor metastases as well as the mechanism involved repressing genes that are inhibitors of angiogenesis. One hundred and thirty-three human genes orthologous to the *Ets2*-TAM profile could retrospectively predict disease-free survival among patients present in two human breast cancer microarray data sets (17, 18). These results identify

**Authors' Affiliations:** Departments of <sup>1</sup>Molecular and Cellular Biochemistry, <sup>2</sup>Molecular Genetics, <sup>3</sup>Veterinary Biosciences, <sup>4</sup>Biomedical Informatics, and <sup>5</sup>Molecular Virology, Immunology, and Medical Genetics, <sup>6</sup>Center for Biostatistics, College of Public Health, and <sup>7</sup>Tumor Microenvironment Program, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Microarray data has been deposited with the Gene Expression Omnibus and is available through this link: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12863>.

**Corresponding Author:** Michael C. Ostrowski, The Ohio State University, 460 West 12th Avenue, BRT 810, Columbus, OH 43210. Phone: 614-688-3824; Fax: 614-688-4141; E-mail: [ostrowski.4@osu.edu](mailto:ostrowski.4@osu.edu).

doi: 10.1158/0008-5472.CAN-09-1474

©2010 American Association for Cancer Research.

an *Ets2*-regulated transcriptional program in TAMs that regulates the growth and spread of breast tumors.

## Materials and Methods

**Mice.** The *Ets2*<sup>LoxP</sup> allele, *Ets2*<sup>db</sup> knockout allele, *MMTV-PyMT* transgenic mice, and *Lys-Cre* knockin allele have been previously described (19–22). The *c-fms-YFP* construct was identical to the published *c-fms-EGFP* construct except for the substitution of YFP (23). Transgenic mice were produced by standard DNA microinjection procedures. All alleles used were >10 generations FVB/N background. Use and care of mice were approved by the Ohio State University Institutional Animal Care and Use Committee.

**Orthotopic and tail vein injection assays.** Two breast cancer cell lines, Met-1 (*MMTV-PyMT*) and MVT-1 (*MMTV-c-Myc;MMTV-VEGF*) were used (24, 25). The cell lines were cultured in DMEM containing 10% fetal bovine serum at 37°C in a 5% CO<sub>2</sub> incubator. Cultured tumor cells were harvested at 80% to 90% confluence and suspended in filtered cold 0.9% NaCl. Three million Met-1 cells or 200,000 MVT-1 cells were injected into the tail vein or mammary gland, respectively. Tail vein and orthotopically injected animals were dissected 18 and 35 d postinjection, respectively.

**Isolation of TAMs.** Minced mammary glands or lungs with metastatic tumors were digested with 20 mg of collagenase type 2 (Worthington), 480 units of DNaseI (Boehringer), and 1 mmol/L of MgCl<sub>2</sub> at 37°C and stroma was enriched by gravity separation (26). The YFP-positive cell population was sorted using fluorescent activated cell sorting with BD FACSAria.

**RNA extraction and quantitative real-time PCR.** RNA extraction and cDNA preparation were done as described previously (27). For samples used in microarray analysis, RNA was extracted with the RNeasy Stratagene micro-prep column (Stratagene) according to the instructions of the manufacturer. Two independent sets of RNA isolated from different TAMs/mice other than RNA used for the microarrays were used for verification.

Real-time quantitative reverse transcription-PCR was conducted using the Roche Universal Probe Library system (Roche Diagnostics) in an iCycler iQ Real-time Detection System (Bio-Rad) as described previously (27). Primer-probe combinations are available upon request.

**Histology and immunohistochemistry.** Tumor tissues were fixed in formalin overnight, processed, paraffin-embedded, and 5 μmol/L sections were prepared. For immunostaining, rat α-mouse F4/80 (1:40 dilution; Caltag Labs), rat α-mouse CD31 (1:50 dilution; Abcam), mouse α-human THBS1 (1:50 dilution; Abcam), mouse α-mouse THBS2 (1:50 dilution; BD Biosciences), goat α-mouse SPARC (1:100 dilution; BD Biosciences), and mouse α-mouse bromodeoxyuridine (1:50 dilution; DAKO) primary antibodies were used. Biotinylated goat α-rat, goat α-mouse, or donkey α-goat (BD Biosciences) were the secondary antibodies used for immunohistochemical analyses. Images were acquired using an AxioScope 40 microscope (Zeiss) equipped with an Axio-cam HRc camera (Zeiss). Immunohistochemical data was

quantified by calculating the area of antibody staining per unit area of tumor using Metamorph 6.0 software. Whole mount hematoxylin staining of lungs was performed as described (28).

For colocalization studies, frozen sections of mammary tumors fixed in 4% paraformaldehyde were double-immunostained with α-F4/80 antibody (Alexa-594 secondary antibody; Invitrogen) and either α-THBS2, α-THBS1, or α-SPARC antibody (Alexa-488 secondary antibody; Invitrogen). Nuclei were stained with DRAQ5. Images of stained mammary tumors were acquired using a Zeiss 510 META laser scanning confocal microscope. Results are presented as the percentage of F4/80-positive or -negative cells that had colocalized staining in or around (extracellular space) for α-THBS2, α-THBS1, or α-SPARC, respectively.

**Chromatin immunoprecipitation assays.** Chromatin immunoprecipitation (ChIP) assays were performed as described (27). Immunoprecipitation was carried out with 2.5 μg of antibodies. The ETS2 antibody has been previously described (19). Rabbit α-mouse HDAC1 and rabbit-IgG were purchased from Santa Cruz Biotechnology and Upstate, respectively. For lung TAMs, the immunoprecipitated chromatin was amplified using an unbiased genome amplification kit (Sigma Aldrich). Samples were analyzed by real-time PCR using the Roche Universal Probe Library (Roche Diagnostics) and the FastStart TaqMan master kit (Roche Diagnostics).

**Microarray and survival analysis.** Microarrays were performed on the Mouse Affymetrix 130A.2 platform. The primary data was analyzed by a modified robust multi-array average method to yield an average gene expression value (29, 30). The detailed description of the experiment and subsequent data analysis is presented in Supplementary Table S1A.

A high-confidence 142 probe set ( $P < 0.05$ ) human *Ets2*-TAM signature was generated by comparing 407 mouse probe sets (357 genes, absolute INT > 1.5) to the 98 lymph node-negative Rosetta cohort (divided into two groups based on lymphocyte/leukocyte infiltration status; ref. 31).<sup>8</sup> For survival analysis, the 142 probe set *Ets2*-TAM signature was used as a query to retrieve gene expression data from Stockholm (GSE1456) breast cancer microarrays (downloaded from the National Center for Biotechnology Information-Gene Expression Omnibus web page). Similarly, gene expression data was also extracted from total and lymph node-negative Rosetta microarrays. The resultant data sets were loaded onto BRB Array Tools as described in Supplementary Table S3. Briefly, unsupervised K-means clustering of each data set was performed by using Cluster 3.0 (32), and samples were assigned into two groups. Kaplan-Meier survival analysis was performed by using the Survival Analysis module of the software package StatsDirect (StatsDirect, Ltd.). Significance of survival analyses was performed by using the log rank (Peto) test.

**Statistical analysis.** For lung metastases data, a nonparametric Kruskal-Wallis test with no multiplicity adjustment was used to compare medians between experimental and

<sup>8</sup> <http://www.rii.com/publications/2002>

control groups. A repeated measures ANOVA model was used to analyze mammary tumor progression between the genetic groups over a period of 42 d. This approach takes into consideration longitudinal data, and the following terms were included in the model: genetic group, time and interaction (genotype \* time). For the statistical analysis of imaging data, an unpaired Student's *t* test was used. All the tests were two-sided.

## Results

**Deletion of Ets2 in TAMs decreases lung metastasis in spontaneous and orthotopic breast tumor models.** Cre/*LoxP* technology was used to conditionally delete *Ets2* in TAMs in the *PyMT* model, a penetrant breast cancer model with a high frequency of lung metastasis (21). The conditional *Ets2<sup>LoxP</sup>* allele used for this study contained *LoxP* sites flanking exon3-exon5 so that Cre-mediated recombination of the region resulted in the generation of a null allele (19). The well-characterized *Lys-Cre* knockin allele was used to delete *Ets2* specifically in the macrophage compartment (22). However, initial studies revealed that Cre recombination in *Lys-Cre;Ets2<sup>LoxP/LoxP</sup>* mice was only 30% to 50% efficient (data not shown). To circumvent this problem, we adopted a strategy whereby mice contained one conditional *Ets2<sup>LoxP</sup>* allele and one conventional knockout allele, *Ets2<sup>db</sup>* (20). In the final cross, *PyMT;Lys-Cre;Ets2<sup>db/+</sup>* males were crossed with *Ets2<sup>LoxP/LoxP</sup>* females to generate both the experimental genotype, *PyMT;Lys-Cre;Ets2<sup>LoxP/db</sup>*, and the control genotype, *PyMT;Ets2<sup>LoxP/db</sup>* (Supplementary Fig. S1A). The frequency of *Ets2* rearrangement in isolated mammary tumor macrophages varied between 70% and 90% with this allele configuration (Supplementary Fig. S1B).

Tumor progression was monitored in females of the two genotypes. Tumor initiation was identical between experimental and control mice (data not shown). A small, but statistically significant, decrease in overall tumor growth was observed in the experimental group (Supplementary Fig. S1C). This difference in tumor growth was not significant in the early carcinoma stage of progression, but was more pronounced during the late carcinoma stage (days 21–35 postinitiation; Supplementary Fig. S1C). However, the final tumor burden and tumor volume were similar in both *PyMT;Lys-Cre;Ets2<sup>LoxP/db</sup>* and *PyMT;Ets2<sup>LoxP/db</sup>* mice (Supplementary Fig. S1D).

Lung metastasis in both genetic groups was studied by whole mount analysis (Fig. 1A; Supplementary Fig. S1E). After image acquisition, the size of the tumors relative to total lung area and the total number of metastases in *PyMT;Lys-Cre;Ets2<sup>LoxP/db</sup>* versus *PyMT;Ets2<sup>LoxP/db</sup>* mice were quantified. The results showed that both the size and number of lung metastases were significantly reduced in *PyMT;Lys-Cre;Ets2<sup>LoxP/db</sup>* mice compared with controls (Fig. 1A, right; size decreased 3- to 4-fold; *P* = 0.001; Supplementary Fig. S2A, number decreased 2-fold; *P* = 0.02).

To confirm and extend the results obtained in the genetic *PyMT* model, a syngeneic model was used. The highly metastatic cell line, MVT-1, derived from mice doubly transgenic

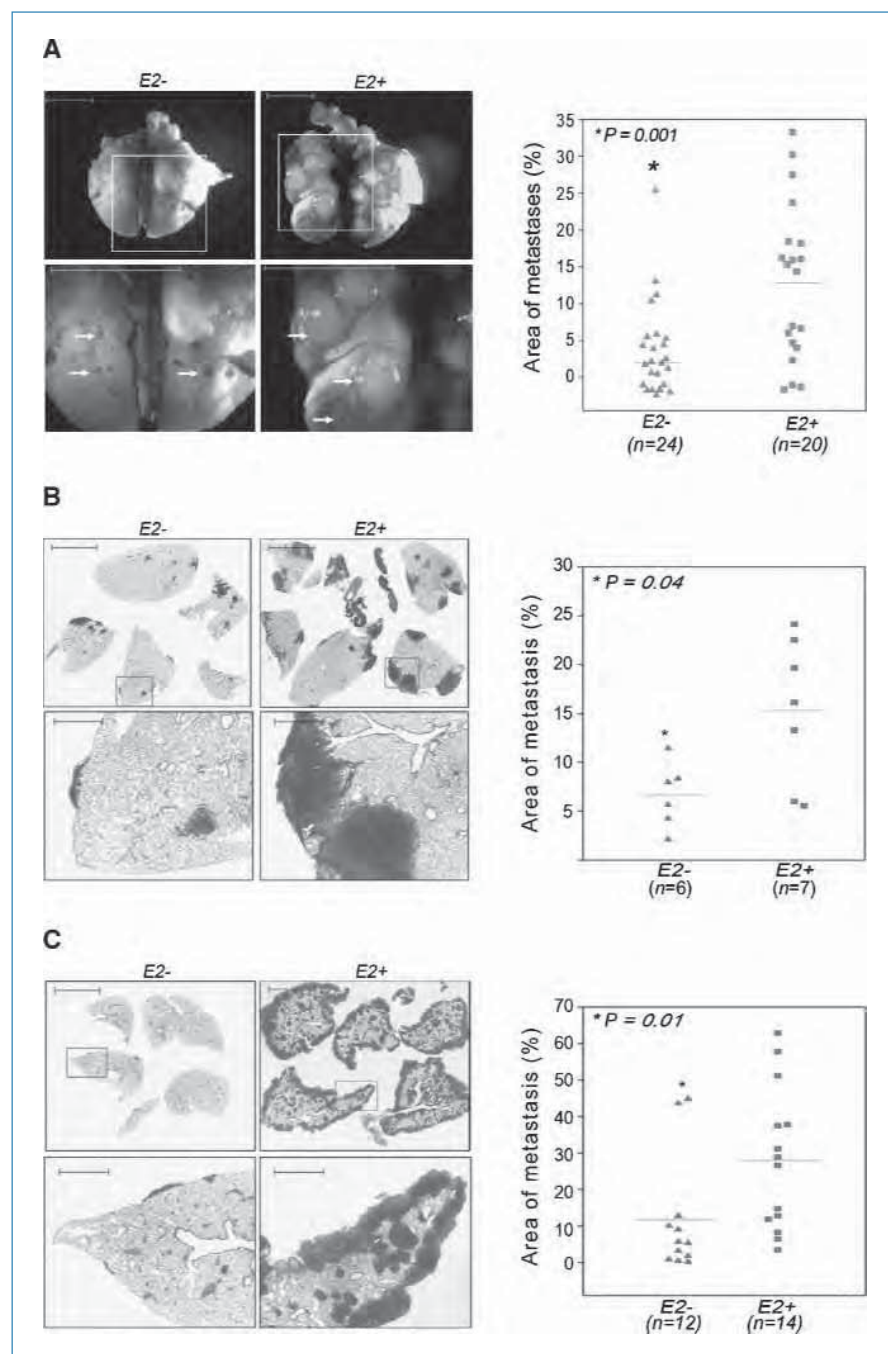
for *MMTV-c-Myc* and *MMTV-VEGF* (25) was injected into the mammary fat pads of *Lys-Cre;Ets2<sup>LoxP/db</sup>* and control *Ets2<sup>LoxP/db</sup>* female mice. After 35 days, mice were euthanized and examined. Although there were no differences in the final tumor burden for the primary tumors (data not shown), the size of metastases per total lung area was reduced 3-fold in the experimental *Lys-Cre;Ets2<sup>LoxP/db</sup>* group compared with the control group (Fig. 1B). These results indicate that the effect of *Ets2* is independent of the *PyMT* oncogene and also show that haploinsufficiency of *Ets2* in the *PyMT* model is not a confounding factor.

**Ets2 in lung macrophages is required for breast tumor metastasis.** To firmly establish that the effect of *Ets2* in TAMs on metastasis was independent of effects at the primary mammary tumors, a tail vein injection model was used. A metastatic *PyMT* cell line, Met-1 (24), was injected into the circulation via the tail vein in the same two genetic groups as above. After 18 days, mice were euthanized and metastases to lungs were quantified in H&E-stained sections (Fig. 1C). The results showed that the size of lung metastases were significantly reduced more than 3-fold in the *Lys-Cre;Ets2<sup>LoxP/db</sup>* mice compared with controls.

A potential explanation for the lower levels of metastasis observed in all three models might be that *Ets2*-regulated genes were required for macrophage survival and/or motility (11, 12). Immunostaining of tumor sections with F4/80 antibody, a marker for mature macrophages, revealed that *Ets2* deletion did not result in a decrease in F4/80-positive macrophages associated with either primary or metastatic tumors (Supplementary Fig. S2B–C, respectively).

**Identification of Ets2 target genes in TAMs.** To address the mechanism of *Ets2* function in TAMs, mammary TAMs were isolated and subjected to gene expression profiling using the Affymetrix platform. To accomplish this, mammary TAMs were tagged using a *c-fms-YFP* transgene (ref. 23; Supplementary Fig. S3A). This transgene was incorporated into the breeding scheme outlined above to produce experimental *PyMT;Lys-Cre;Ets2<sup>LoxP/db</sup>;c-fms-YFP* and control *PyMT;Ets2<sup>LoxP/db</sup>;c-fms-YFP* mice. YFP-positive cells isolated from collagenase-digested tissue by high-speed digital fluorescence-activated cell sorting represented ~10% to 15% of the total cells from the primary mammary tumor site (Supplementary Fig. S3B). Greater than 90% of these YFP-positive cells coexpressed macrophage markers such as F4/80 (Supplementary Fig. S3C). Typically,  $3 \times 10^5$  to  $5 \times 10^5$  YFP-positive TAMs could be isolated from a single mouse.

YFP-positive TAMs were isolated from both genetic groups at the stage when early carcinoma was initially detected in the *PyMT* model (21). The percentage of YFP-positive cells per mammary gland isolated by fluorescence-activated cell sorting was similar in both genetic groups, supporting the conclusion that a reduction in tumor macrophages was not responsible for the observed effects (Supplementary Fig. S3B). Because macrophages have also been shown to play a central role in tissue remodeling during mammary gland development (33), YFP-positive macrophages were extracted from the mammary gland of *Lys-Cre;Ets2<sup>LoxP/db</sup>;c-fms-YFP* and *Ets2<sup>LoxP/db</sup>;c-fms-YFP* females ~14 days after the onset



**Figure 1.** Deletion of *Ets2* in TAMs decreases lung metastasis in spontaneous, orthotopic, and tail vein injection breast cancer models. A, whole mount images of lungs obtained from *PyMT*;*Lys-Cre*;*Ets2*<sup>LoxP/dB</sup> (*E2-*, left) and *PyMT*;*Ets2*<sup>LoxP/dB</sup> (*E2+*, right) mice at late carcinoma stage. B, analysis of metastatic tumor burden in H&E-stained lung sections obtained from *Lys-Cre*;*Ets2*<sup>LoxP/dB</sup> (*E2-*, left) and *Ets2*<sup>LoxP/dB</sup> (*E2+*, right) mice in the MVT-1 orthotopic model. C, analysis of metastatic tumor burden in H&E-stained lung sections obtained from *E2-* (left) and *E2+* (right) mice in the Met-1 tail vein injection model. Bottom panels, high-magnification images of insets from the respective top panels (A, B, and C). Lung metastases are indicated by white arrows (A) or outlined with the dotted red line (B and C). Bars, 5 mm. Right, scatter plots indicate the size of the metastatic tumors in mice with the indicated genotype. Data is presented as the tumor metastases area per unit area. The mean size in each genotype is indicated by the horizontal line. *n*, number of mice per genetic group. Statistical significance (*P* value evaluated by nonparametric Kruskal-Wallis test) is shown.

of puberty. We reasoned that the role of macrophages in tissue remodeling during mammary gland development would provide a useful comparison to unmask the tumor-specific effects of *Ets2*.

Expression profiling was performed on the resulting four sets of RNA samples. Comparisons between all four sets of expression data were used to identify 357 genes (407 probe sets), the expressions of which depended on both loss of *Ets2* and the presence of tumor (see Supplementary Table S1 for

details). Approximately 25% of these genes were negatively regulated in the tumor microenvironment and the expression of these genes increased when *Ets2* was deleted in TAMs. Gene ontology indicated that genes encoding extracellular components were principally affected by *Ets2* deletion (Fig. 2A). The major biological process represented was angiogenesis, with 34% of the genes annotated as having a role in this process (Fig. 2A). Many of the genes in the angiogenesis class were classified as inhibitors of angiogenesis.



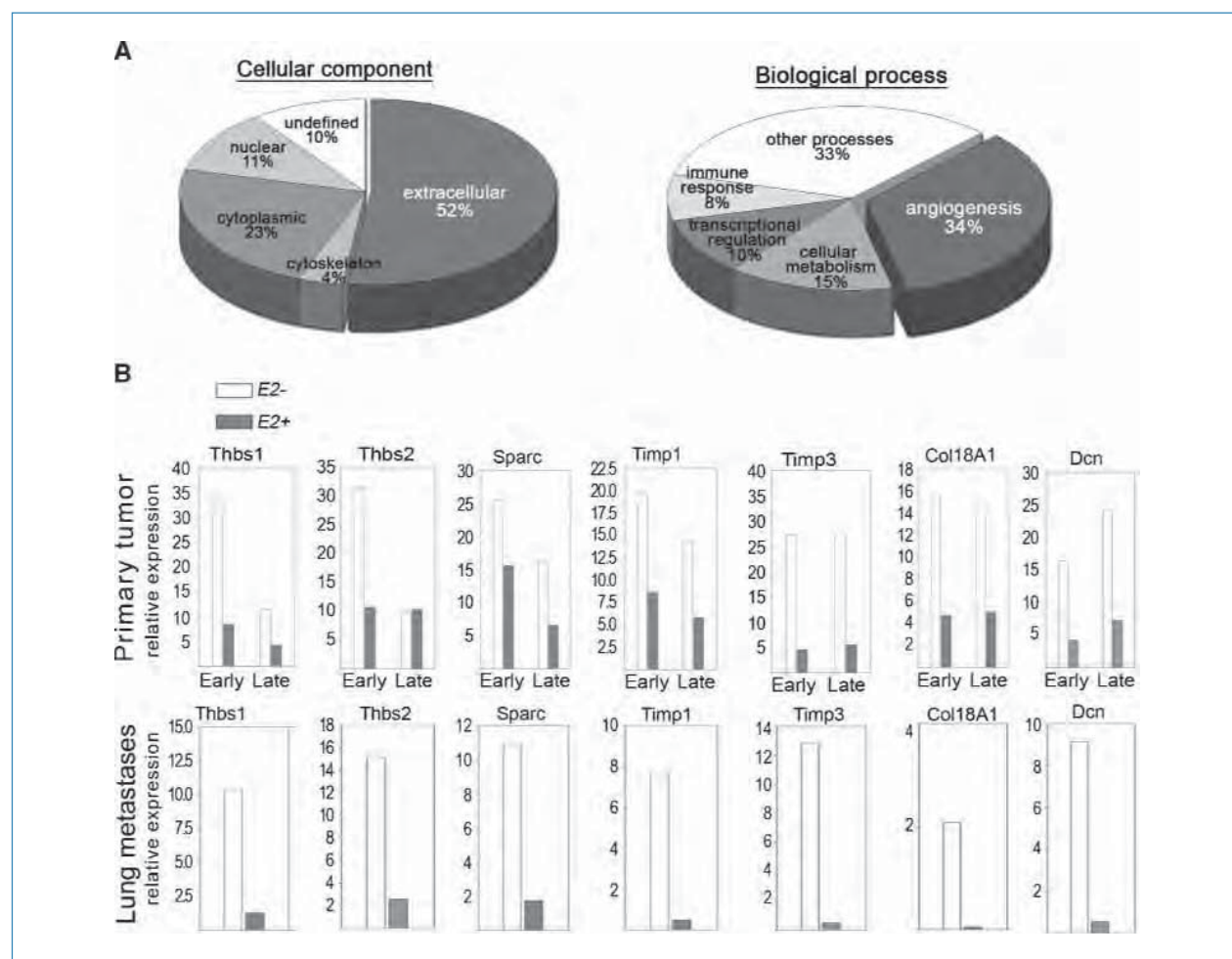
Quantitative reverse transcription-PCR using RNA from independently isolated mammary TAMs representing early (first palpable tumor) and late (6 weeks after tumor initiation) carcinoma stages were used to verify the microarray results (Fig. 2B). Of 31 genes tested, 25 were confirmed to be differentially expressed in TAMs with or without *Ets2* (Supplementary Table S2). Data for 14 of the genes classified as encoding inhibitors of angiogenesis are shown (Fig. 2B; Supplementary Fig. S3D). Expression of these genes in both early and late tumors was increased when *Ets2* was deleted. In contrast, potential ETS2 targets known to be involved in inflammation such as *Mmp9* and *Tnfa* (12), and other genes associated with inflammation such as *Il6*, were not significantly affected by *Ets2* deletion in TAMs, emphasizing that the analysis identified tumor-specific targets of ETS2 (Supplementary Fig. S3D).

The same 31 genes were also studied in lung TAMs isolated following tail vein injection of Met-1 cells (bottom, Fig. 2B;

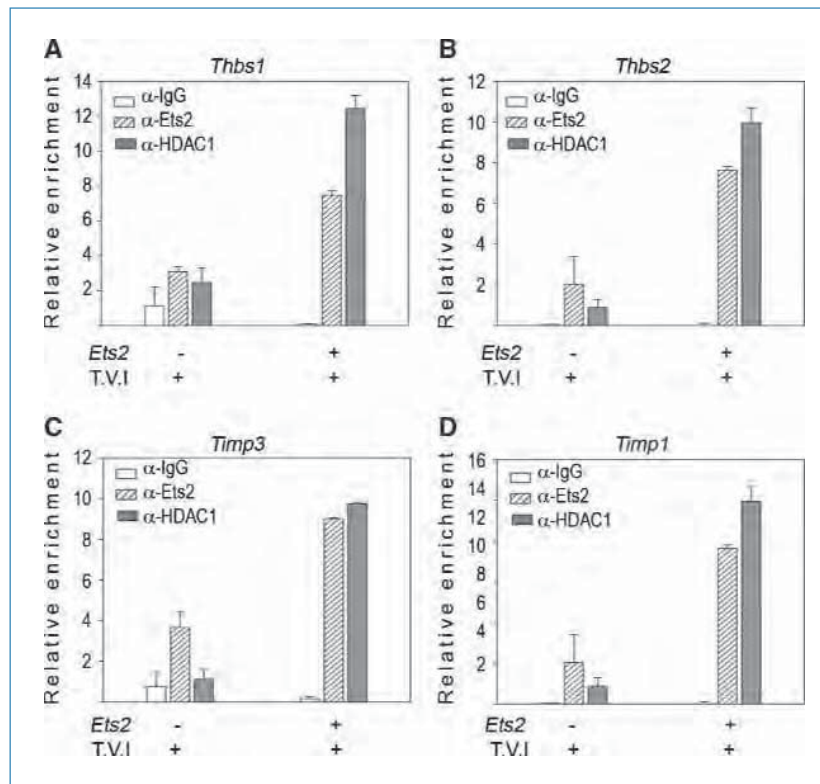
Supplementary Table S2). In these TAMs, 25 of 31 genes were differentially expressed when *Ets2* was deleted, including the angiogenic gene set, indicating the *Ets2* targets were similar in mammary or lung TAMs.

**ETS2 directly regulates antiangiogenic genes in isolated TAMs.** Examination of 1 kb of the proximal promoter regions of four candidate genes not previously reported as ETS2 targets (*Thbs1*, *Thbs2*, *Timp1*, and *Timp3*) revealed conserved ETS binding motifs in their proximal promoter regions (Supplementary Fig. S4A). Based on these conserved sequences, ChIP experiments were performed on lung TAMs from mice with or without *Ets2*. For the experiments, ~50,000 YFP-tagged, F4/80-positive cells were isolated from lungs containing metastases following tail vein injection of Met-1 cells. Antibodies against ETS2 and its corepressor HDAC1 (14) were used in the ChIP assays (Fig. 3).

The ChIP experiments revealed that in wild-type cells, ETS2 and HDAC1 were both enriched at all four of these promoter



**Figure 2.** *Ets2* represses the expression of extracellular matrix-modifying genes in TAMs. A, gene ontology based on cellular localization (left pie chart) and biological process (right pie chart) of the genes differentially regulated in TAMs with or without *Ets2*. B, confirmation of genes identified by the microarray analysis using real-time quantitative PCR. RNA was extracted from independently isolated sets of TAMs derived from mammary glands (top) or from lungs of mice injected with the Met-1 cell line (bottom). Genotypes and stage of tumor development are indicated. Data is represented as average fold induction in samples analyzed in duplicates.



**Figure 3.** *Ets2* represses the expression of antiangiogenic genes in TAMs. ChIP analysis of the *Ets2*-binding sites in *Thbs1* (A, distal *Ets2* site), *Thbs2* (B), *Timp3* (C), and *Timp1* (D) promoters from YFP+ cells extracted from the lung following tail vein injection of Met-1 cells. ChIP was performed with  $\alpha$ -ETS2 and  $\alpha$ -HDAC1 antibodies and rabbit IgG control, as indicated. Subsequently, quantitative PCR was performed on the immunoprecipitated chromatin. Columns, mean of the relative enrichment of amplified chromatin from two independent experiments; bars, SEM.

sequences (Fig. 3). In contrast, when *Ets2* was conditionally deleted, both the levels of ETS2 and HDAC1 were significantly reduced at each of the four promoters. Similar results were obtained for the *Thbs1* promoter in TAMs isolated from the primary mammary tumor (Supplementary Fig. S4B).

**Expression of angiogenesis inhibitors in TAMs lacking *Ets2* correlates with reduced tumor angiogenesis and proliferation.** To verify the expression of ETS2 targets *in situ*, we performed immunohistochemical staining on paraffin-embedded samples prepared from metastatic lung tumors using commercially available antibodies. This analysis showed robust expression of THBS2, THBS1, and SPARC within tumors from mice with *Ets2* deletions in TAMs compared with *Ets2*+ controls (Fig. 4A; Supplementary Fig. S5A–B).

To confirm that the tumor macrophages were expressing these proteins, frozen mammary tumor sections were analyzed by double immunofluorescent staining using F4/80 to identify TAMs. The MVT-1 orthotopic mammary fat pad injection model was used for this analysis. Staining with  $\alpha$ -F4/80 and  $\alpha$ -THBS2 showed extensive overlap between the two proteins in sections obtained from tumors with *Ets2* deletions (Fig. 4B, top; Supplementary Movie 1). Because THBS2 is an extracellular protein, expression was found both intracellularly and in the adjacent extracellular space in ~75% of F4/80-positive cells, as clearly evident in confocal reconstructions of 15- $\mu$ m sections (see Supplementary Movie 1). In contrast, coexpression of THBS2 in F4/80-positive cells was 10-fold lower in tumors with *Ets2* (Fig. 4B, bottom). Importantly, expression of THBS2 in F4/80-negative cells was not

affected by deletion of *Ets2* in TAMs (Fig. 4B, bottom bar graph). Identical results were obtained for THBS1 and SPARC (Supplementary Fig. S6A–B and Supplementary Movies 2–3, respectively).

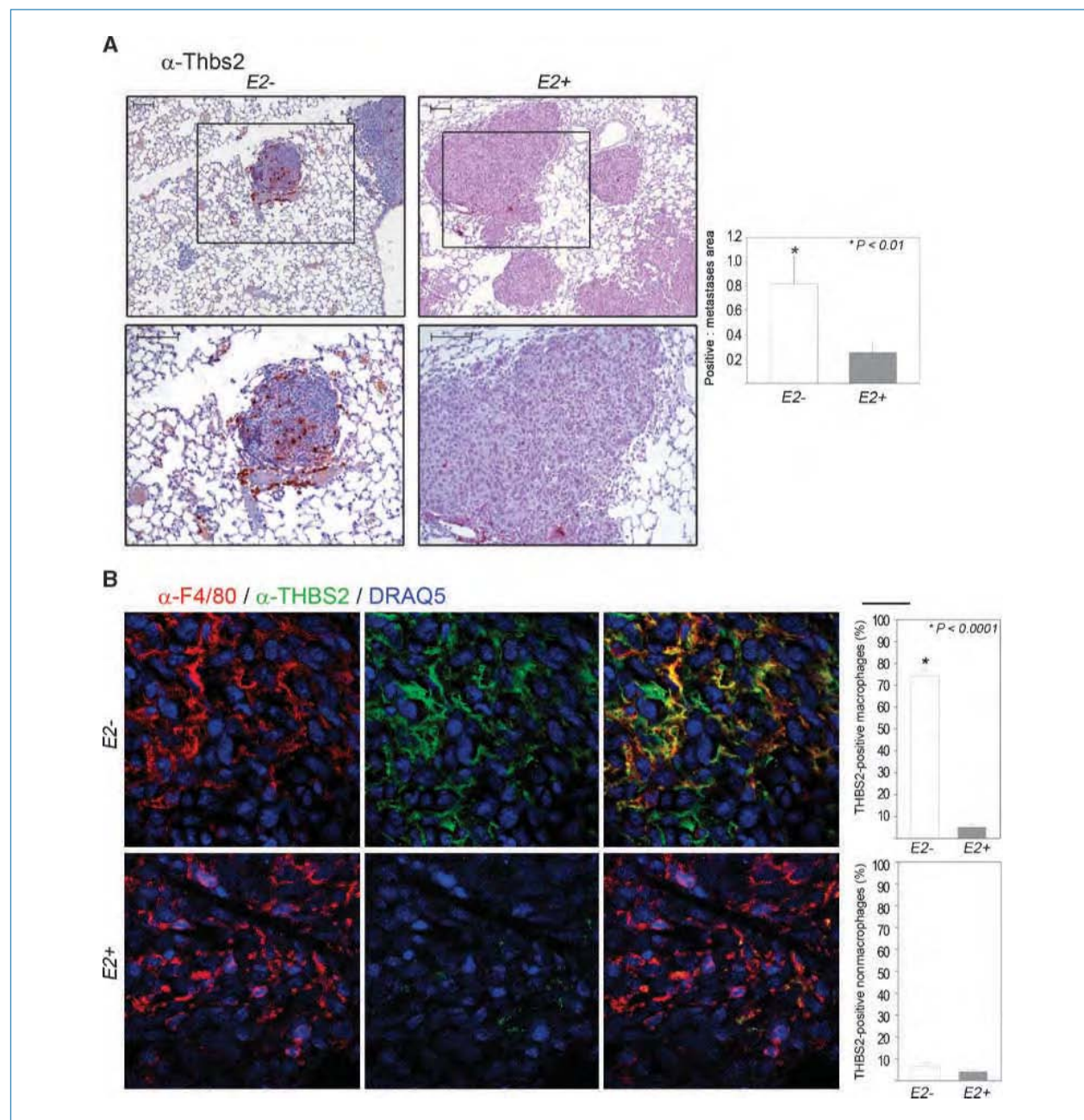
Because many of the tumor-specific *Ets2* targets detected, including THBS1, THBS2, and SPARC have been implicated in angiogenesis, blood vessel density was analyzed in experimental and control mice using  $\alpha$ -CD31 immunostaining of paraffin-embedded tumor sections. For these experiments, both primary MVT-1 tumors and lung tumors formed by tail vein injection of Met-1 cells were studied (Fig. 5A). A significant 2- to 3-fold reduction in tumor vasculature was observed in both primary mammary tumors and lung metastases (Fig. 5A).

Bromodeoxyuridine incorporation was used to measure cell proliferation in lung metastases in the Met-1 tail vein injection model (Fig. 5B). The analysis showed a significant 2.5-fold decrease in bromodeoxyuridine-labeled tumor cells in mice with *Ets2*-deficient TAMs compared with controls. Tumor cell apoptosis, measured by staining with activated caspase-3 antibody, was not significantly affected by *Ets2* deletion (Supplementary Fig. S2D).

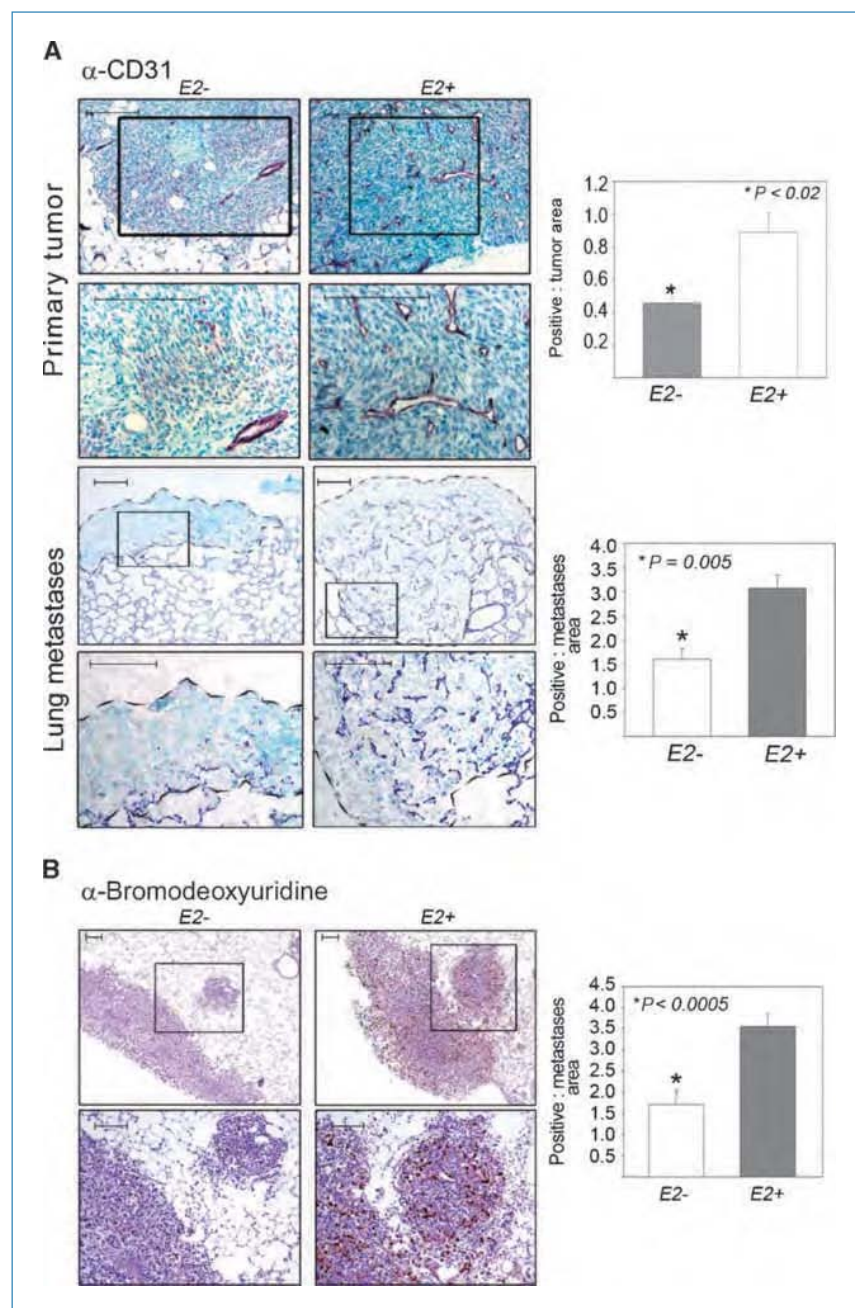
**The *Ets2*-TAM gene expression signature predicts survival of patients with breast cancer.** To determine if the mouse genetic studies were relevant to human disease, the mouse expression data was compared with the Rosetta human breast cancer data set (31). Initially, 407 mouse probe sets that were differentially expressed in mouse TAMs with or without *Ets2* were compared with the Rosetta array platform and 341 homologous human probe sets were identified (see

Supplementary Table S3 for details). These 341 probe sets were compared with 2,856 probe sets that represented genes differentially expressed in 117 human samples annotated as with or without lymphocyte/leukocyte infiltration (31). This comparison

showed that 142 of the mouse *Ets2*-TAM probe sets, representing 133 genes, were significantly differentially expressed in lymphocyte/leukocyte infiltration-positive versus -negative human breast cancers ( $P < 0.05$ , see Supplementary Table S3



**Figure 4.** Increased expression of thrombospondin-2 in TAMs that lack *Ets2*. **A**, lung sections from mice injected with the Met-1 cell line in the indicated genotypes, immunostained with  $\alpha$ -THBS2. Bottom panels, high-magnification images of insets from the respective top panels (**B**). Bars, 100  $\mu$ m. Quantification of antibody staining is presented as the average area of staining per tumor area (graphs, bottom). Five different tumor areas from five different mice in each group were analyzed. **B**, images of mammary tumor sections from mice injected with the MVT-1 cell line harvested 1 wk postinjection. Double-immunostained with  $\alpha$ -F4/80 (red, left) and  $\alpha$ -THBS2 (green, middle), and merged F4/80-THBS2 images (yellow, right). Quantification of antibody staining is presented as the average percentage of F4/80-positive cells that are also positive for THBS2 in the mammary tumors (graph at right). Bar at the top right corner, 20  $\mu$ m. Five different tumor areas from four different tumors in each group were analyzed. Statistical significance ( $P$  value evaluated by unpaired Student's  $t$  test) is shown.



**Figure 5.** Impaired tumor angiogenesis and proliferation with *Ets2* deletion in TAMs. A, images of mammary (top) and lung (bottom) sections from mice injected with the MVT-1 and Met-1 cells, respectively, genotypes as indicated. Immunostained with  $\alpha$ -CD31. B, micrographs of lung sections from mice injected with the Met-1 cell line in the indicated genotypes, immunostained with  $\alpha$ -bromodeoxyuridine. Panels 2 and 4 (A) and bottom panels (B), high-magnification images of insets from the respective top panels. Bar, 100  $\mu$ m. Quantification of antibody staining is presented as the average area of staining per tumor area (graphs, right). Five different tumor areas from five different mice in each group were analyzed. Statistical significance ( $P$  value evaluated by unpaired Student's  $t$  test) is shown.

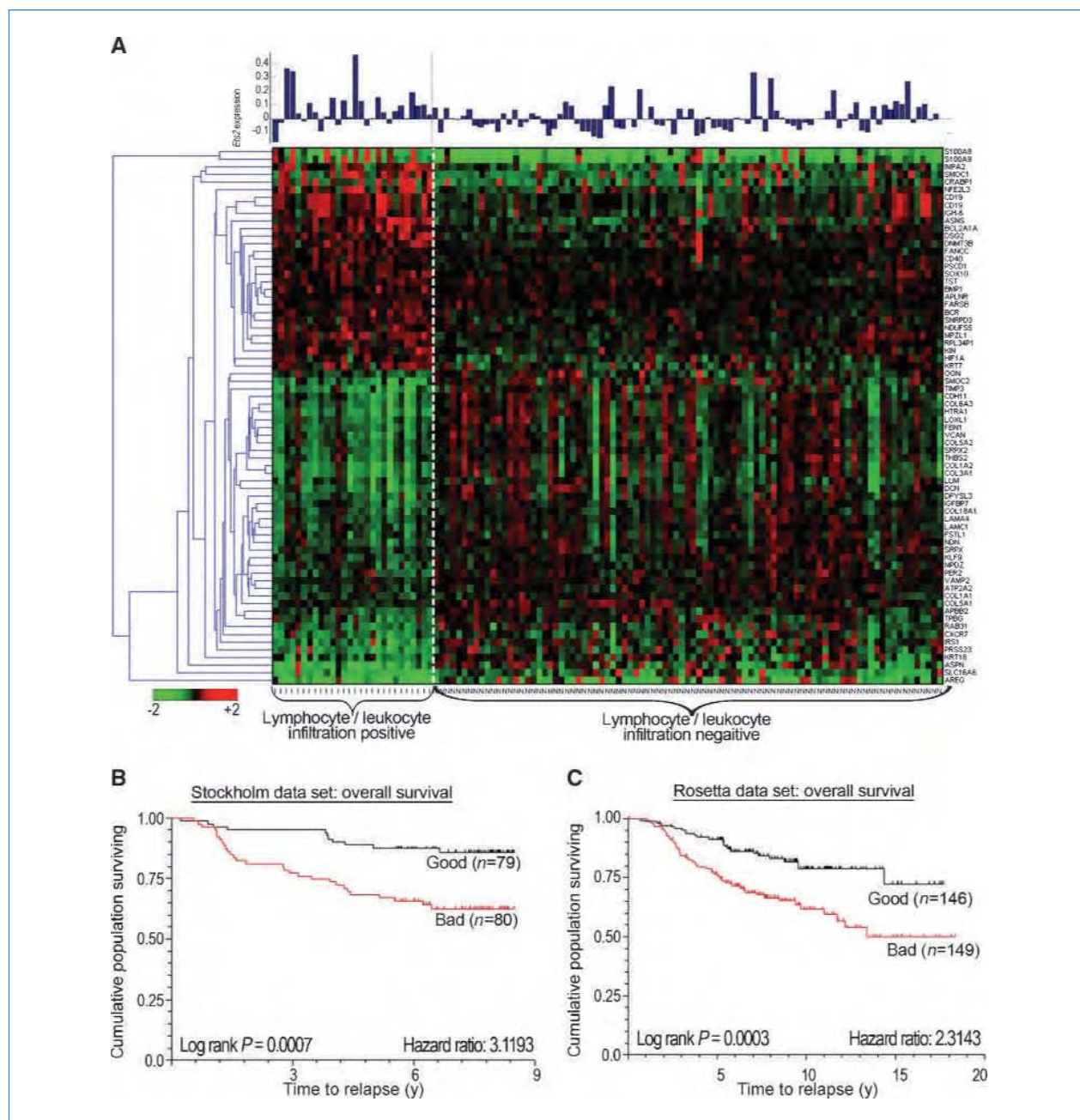
and Supplementary Fig. S7A). Gene ontology analysis of these human genes showed that extracellular matrix components and angiogenesis were predominantly affected, just as for the mouse *Ets2*-TAMs genes (Supplementary Fig. S7B). A subset of 70 genes differentially expressed with high significance ( $P < 0.001$ ) is represented in the heat map presented in Fig. 6A. Interestingly, *Ets2* expression itself was on average 8-fold higher in lymphocyte/leukocyte infiltration-positive patients when compared with the negative group (Fig. 6A, bar graph;  $P = 0.0002$ ).

To determine if the TAM gene signature correlated with the clinical outcome of patients, the 133 human *Ets2*-TAM gene signature was used for unsupervised clustering of expression data obtained from 159 patients with sporadic breast cancer in the Stockholm data set (ref. 18; see Supplementary Table S3). Expression of the *Ets2*-TAM signature predicted overall survival in this group with high confidence (Fig. 6B;  $P = 0.0007$ ; hazard ratio, 3.1). Similar results were obtained with the entire 295-patient Rosetta sample set (Fig. 6C;  $P = 0.00003$ ; hazard ratio, 2.31).

## Discussion

The influence of the microenvironment, particularly macrophages, on tumor growth and metastasis have long been recognized, but relatively little is known of the gene pathways and mechanisms macrophages use to promote tumor malignancy (34). The results presented here show that in mouse models, *Ets2* in tumor macrophages promotes angio-

genesis and growth of both primary tumors and lung metastases. The mechanism of action of ETS2 in TAMs involved direct repression of genes encoding predominantly extracellular products, including well-characterized inhibitors of angiogenesis. Recently, an independent report of global gene profiling in TAMs also observed the expression of several antiangiogenic genes along with well-known positive regulators such as *Vegf-a*, results consistent with our data (35).



**Figure 6.** The *Ets2*-TAM gene signature predicts survival in human breast cancer patients. A, heat map of differential expression of the *Ets2*-TAM 70 gene profile ( $P < 0.001$ , see text) in 98 breast cancer samples distinguished by lymphocyte/leukocyte infiltration. Bar graph (top), the level of *Ets2* expression in each of the 98 cancer samples. B and C, Kaplan-Meier analysis of overall survival in Stockholm and Rosetta breast tumor cohorts, respectively.

However, the antiangiogenic effect of TAMs lacking *Ets2* is dominant even in the context of MVT-1 tumor cells that overexpress *VEGF-A*. Additionally, the presumed role of VEGF-A produced by TAMs in triggering the angiogenic switch have been challenged by recent findings showing that deletion of VEGF-A in TAMs actually results in increased tumor growth (36, 37). Thus, *Ets2* has a previously unappreciated role in TAMs in controlling the balance between positive and negative regulators of angiogenesis necessary for tumor metastasis.

*Ets2* in TAMs increased the growth of primary and metastatic tumors. *Ets2* could indirectly affect tumor growth by modulating angiogenesis, or directly through paracrine mechanisms. The *Ets2* targets identified would favor the former possibility, as obvious paracrine candidates such as *Il6* or *Egf* were not differentially expressed. In either case, the results are consistent with the *Ets2* pathway playing a role in some activities associated with the alternatively activated M2 macrophage population (2, 3). M2 macrophages are believed to modulate inflammatory response and to promote tissue remodeling and angiogenesis; in the context of tumor progression, M2-like cells are believed to promote immune suppression as well as tumor angiogenesis, invasion, and metastasis (2, 3). Extracellular function and angiogenesis are the major *Ets2* targets identified in our studies, providing a molecular mechanism by which M2-like tumor macrophages modulate the extracellular microenvironment to promote tumor growth and angiogenesis at both primary and tumor sites.

A key finding is that a portion of the mouse *Ets2*-TAM gene expression signature was present in human breast cancer expression data and that it could retrospectively predict overall survival in two independent cohorts of patients with sporadic breast cancer. This 133-gene signature is independent of other breast tumor signatures capable of predicting patient outcome, including stromal gene signatures (38, 39). Although further efforts will be required to fully implement these findings and determine their significance to human disease, the results validate the relevance of our hypothe-

sis-driven mouse modeling approach for dissecting TAM functions in tumor growth and metastasis.

Dispersed tumor cells are present in many patients with breast cancer and may be the mediators of tumor recurrence (40). Breast tumor micrometastases are genetically distinct from the primary tumor indicating that they are disseminated early in tumor progression (41, 42). Results obtained in the *PyMT* and *Her2/Neu* mouse models show an early spread of mammary epithelial cells before the carcinoma stage, providing experimental verification of the human data (42). Thus, understanding how dispersed dormant cells progress to growing metastases is a problem with considerable clinical relevance. Further studies on *Ets2* and its downstream targets could provide unique insights into understanding how the microenvironment modulates the growth of tumor cells at metastatic sites.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Acknowledgments

We thank Alexander Borowsky and Michael Johnson for the Met-1 and MVT-1 cell lines, respectively, Robert Oshima for the *Ets2<sup>fl</sup>* mice, Karl Kornacker for microarray data analysis, Kartic Krishnamurthy for image analysis software, and Lisa Rawahneh for histology support. We also acknowledge the Cancer Center Microscopy/Imaging, Microarray, Genomic, Transgenic/Knockout, Histology, and Flow Cytometry Shared Resources.

### Grant Support

DOD Predoctoral Fellowship (T. Zabuawala). This work was supported by National Cancer Institute grants P01 CA097189 (M.C. Ostrowski, G. Leone, T.J. Rosol) and R01 CA053271 (M.C. Ostrowski), and by the Evelyn Simmer's Charitable Trust (M.C. Ostrowski).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 4/27/09; revised 11/15/09; accepted 12/1/09; published OnlineFirst 2/9/10.

### References

- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
- Tlsty TD, Coussens LM. Tumor stroma and regulation of cancer development. *Annu Rev Pathol* 2006;1:119–50.
- Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature* 2008;454:436–44.
- Lewis CE, Pollard JW. Distinct role of macrophages in different tumor microenvironments. *Cancer Res* 2006;66:605–12.
- Wyckoff JB, Wang Y, Lin EY, et al. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res* 2007;67:2649–56.
- Leek RD, Lewis CE, Whitehouse R, Greenall M, Clarke J, Harris AL. Association of macrophage infiltration with angiogenesis and prognosis in invasive breast carcinoma. *Cancer Res* 1996;56:4625–9.
- Bolat F, Kayaselcuk F, Nursal TZ, Yagmurdu MC, Bal N, Demirhan B. Microvessel density, VEGF expression, and tumor-associated macrophages in breast tumors: correlations with prognostic parameters. *J Exp Clin Cancer Res* 2006;25:365–72.
- Tsutsui S, Yasuda K, Suzuki K, Tahara K, Higashi H, Era S. Macrophage infiltration and its prognostic implications in breast cancer: the relationship with VEGF expression and microvessel density. *Oncol Rep* 2005;14:425–31.
- Valkovic T, Dobrila F, Melato M, Sasso F, Rizzardi C, Jonjic N. Correlation between vascular endothelial growth factor, angiogenesis, and tumor-associated macrophages in invasive ductal breast carcinoma. *Virchows Arch* 2002;440:583–8.
- Lin EY, Nguyen AV, Russell RG, Pollard JW. Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. *J Exp Med* 2001;193:727–40.
- Sevilla L, Aperlo C, Dulic V, et al. The *Ets2* transcription factor inhibits apoptosis induced by colony-stimulating factor 1 deprivation of macrophages through a Bcl-xL-dependent mechanism. *Mol Cell Biol* 1999;19:2624–34.
- Wei G, Guo J, Doseff AI, et al. Activated *Ets2* is required for persistent inflammatory responses in the motheaten viable model. *J Immunol* 2004;173:1374–9.
- Baker KM, Wei G, Schaffner AE, Ostrowski MC. *Ets-2* and components of mammalian SWI/SNF form a repressor complex that negatively regulates the BRCA1 promoter. *J Biol Chem* 2003;278:17876–84.

14. Wei G, Schaffner AE, Baker KM, Mansky KC, Ostrowski MC. Ets-2 interacts with co-repressor BS69 to repress target gene expression. *Anticancer Res* 2003;23:2173–8.
15. Turner DP, Findlay VJ, Moussa O, Watson DK. Defining ETS transcription regulatory networks and their contribution to breast cancer progression. *J Cell Biochem* 2007;102:549–59.
16. Man AK, Young LJ, Tynan JA, et al. Ets2-dependent stromal regulation of mouse mammary tumors. *Mol Cell Biol* 2003;23:8614–25.
17. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009.
18. Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;7:R953–64.
19. Wei G, Srinivasan R, Cantemir-Stone C, et al. Ets1 and Ets2 are required for endothelial cell survival during embryonic angiogenesis. *Blood* 2009;114:1123–30.
20. Yamamoto H, Flannery ML, Kupriyanov S, et al. Defective trophoblast function in mice with a targeted mutation of Ets2. *Genes Dev* 1998;12:1315–26.
21. Lin EY, Jones JG, Li P, et al. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol* 2003;163:2113–26.
22. Clausen BE, Burkhardt C, Reith W, Renkawitz R, Forster I. Conditional gene targeting in macrophages and granulocytes using LysMcre mice. *Transgenic Res* 1999;8:265–77.
23. Sasmono RT, Oceandy D, Pollard JW, et al. A macrophage colony-stimulating factor receptor-green fluorescent protein transgene is expressed throughout the mononuclear phagocyte system of the mouse. *Blood* 2003;101:1155–63.
24. Borowsky AD, Namba R, Young LJ, et al. Syngeneic mouse mammary carcinoma cell lines: two closely related cell lines with divergent metastatic behavior. *Clin Exp Metastasis* 2005;22:47–59.
25. Pei XF, Noble MS, Davoli MA, et al. Explant-cell culture of primary mammary tumors from MMTV-c-Myc transgenic mice. *In Vitro Cell Dev Biol Anim* 2004;40:14–21.
26. Soule H, McGrath C. A simplified method for passage and long-term growth of human mammary epithelial cells. *In Vitro Cell Dev Biol* 1986;22:6–12.
27. Hu R, Sharma SM, Bronisz A, Srinivasan R, Sankar U, Ostrowski MC. Eos MITF, PU.1 recruit corepressors to osteoclast-specific genes in committed myeloid progenitors. *Mol Cell Biol* 2007;27:4018–27.
28. Jessen KA, Liu SY, Tepper CG, et al. Molecular analysis of metastasis in a polyomavirus middle T mouse model: the role of osteopontin. *Breast Cancer Res* 2004;6:R157–69.
29. Auer H, Newsom D, Nowak N, et al. Gene-resolution analysis of DNA copy number variation using oligonucleotide expression microarrays. *BMC Genomics* 2007;8:111.
30. Irizarry R, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4:249–64.
31. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
32. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* 2004;20:1453–4.
33. Gouon-Evans V, Rothenberg ME, Pollard JW. Postnatal mammary gland development requires macrophages and eosinophils. *Development* 2000;127:2269–82.
34. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell* 2006;124:263–6.
35. Ojalvo L, King W, Cox D, Pollard J. High-density gene expression analysis of tumor-associated macrophages from mouse mammary tumors. *Am J Pathol* 2009;174:1048–64.
36. Lin EY, Li JF, Bricard G, et al. VEGF restores delayed tumor progression in tumors depleted of macrophages. *Mol Oncol* 2007;1:288–302.
37. Stockmann C, Doedens A, Weidemann A, et al. Deletion of vascular endothelial growth factor in myeloid cells accelerates tumorigenesis. *Nature* 2008;456:814–8.
38. Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat Med* 2008;14:518–27.
39. Ma X, Dahiya S, Richardson E, Erlander M, Sgroi D. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res* 2009;11:R7.
40. Braun S, Vogl FD, Naume B, et al. A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med* 2005;353:793–802.
41. Schmidt-Kittler O, Ragg T, Daskalakis A, et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* 2003;100:7737–42.
42. Hüsemann Y, Geigl J, Schubert F, et al. Systemic spread is an early step in breast cancer. *Cancer Cell* 2008;13:58–68.



## An imaging workflow for characterizing phenotypical change in large histological mouse model datasets

Kishore Mosaliganti<sup>a,b</sup>, Tony Pan<sup>a</sup>, Randall Ridgway<sup>b</sup>, Richard Sharp<sup>b</sup>, Lee Cooper<sup>a</sup>, Alex Gulacy<sup>a</sup>, Ashish Sharma<sup>a</sup>, Okan Irfanoglu<sup>b</sup>, Raghu Machiraju<sup>a,b</sup>, Tahsin Kurc<sup>a</sup>, Alain de Bruin<sup>c</sup>, Pamela Wenzel<sup>c</sup>, Gustavo Leone<sup>c</sup>, Joel Saltz<sup>a,b</sup>, Kun Huang<sup>a,b,d,\*</sup>

<sup>a</sup> Department of Biomedical Informatics, The Ohio State University, 3190 Graves Hall, 333 West 10th Avenue, Columbus, OH 43210, USA

<sup>b</sup> Department of Computer Science and Engineering, 395 Dreese Laboratories, 2015 Neil Avenue, Columbus, OH 43210, USA

<sup>c</sup> Department of Molecular Virology, Immunology, and Medical Genetics, Biomedical Research Tower, 460 West 12th Avenue, Columbus, OH 43210, USA

<sup>d</sup> The Biomedical Informatics Shared Resources, The Ohio State University Comprehensive Cancer Center, 204 Biomedical Research Tower, 460 West 12th Avenue, Columbus, OH 43210, USA

### ARTICLE INFO

#### Article history:

Received 29 January 2007

Available online 7 April 2008

#### Keywords:

Light microscopy  
Histology staining  
Genetic phenotyping  
Mutation  
Morphometrics  
Image analysis  
Image registration  
Segmentation  
Visualization  
Imaging workflow

### ABSTRACT

**Motivation:** This paper presents a workflow designed to quantitatively characterize the 3D structural attributes of macroscopic tissue specimens acquired at a micron level resolution using light microscopy. The specific application is a study of the morphological change in a mouse placenta induced by knocking out the *retinoblastoma* gene.

**Result:** This workflow includes four major components: (i) serial section image acquisition, (ii) image pre-processing, (iii) image analysis involving 2D pair-wise registration, 2D segmentation and 3D reconstruction, and (iv) visualization and quantification of phenotyping parameters. Several new algorithms have been developed within each workflow component. The results confirm the hypotheses that (i) the volume of labyrinth tissue decreases in mutant mice with the *retinoblastoma* (*Rb*) gene knockout and (ii) there is more interdigitation at the surface between the labyrinth and spongiotrophoblast tissues in mutant placenta. Additional confidence stem from agreement in the 3D visualization and the quantitative results generated.

**Availability:** The source code is available upon request.

© 2008 Elsevier Inc. All rights reserved.

### 1. Introduction

This paper presents an imaging workflow designed to quantitatively characterize 3D structural attributes of microscopic tissue specimens at micron level resolution using light microscopy. The quantification and visualization of structural phenotypes in tissue plays a crucial role in understanding how genetic and epigenetic differences ultimately affect the structure and function of multicellular organisms [1–5].

The motivation for developing this imaging workflow is derived from an experimental study of a mouse placenta model system wherein the morphological effects of inactivating the *retinoblastoma* (*Rb*) tumor suppressor gene are studied. The *Rb* tumor suppressor gene was identified over two decades ago as the gene responsible for causing retinal cancer (retinoblastoma) but has also been found to be mutated in numerous other human cancers.

Homozygous deletion of *Rb* in mice results in severe fetal and placental abnormalities that lead to lethality by prenatal day 15.5 [6–8]. Recent studies suggest that *Rb* plays a critical role in regulating development of the placenta and *Rb*<sup>-/-</sup> placental lineages have many fetal abnormalities [8–10].

Our previous work suggested that deletion of *Rb* leads to extensive morphological changes in the mouse placenta including possible reduction of total volume and vasculature of the placental labyrinth, increased infiltration from the spongiotrophoblast layer to the labyrinth layer, and clustering of labyrinthic trophoblasts [8]. However, these observations are based solely on the qualitative inspection of a small number of histological slices from each specimen alone. In order to fully and objectively evaluate the role of *Rb* deletion, a detailed characterization of the mouse placenta morphology at cellular and tissue scales is required. This permits the correlation of cellular and tissue phenotype with *Rb*<sup>-/-</sup> genotype. Hence, we develop a microscopy image processing workflow to acquire, reconstruct, and quantitatively analyze large serial sections obtained from a mouse placenta. In addition, this workflow has a strong visualization component that enables exploration of complicated 3D structures at cellular/tissue levels.

\* Corresponding author. Address: Department of Biomedical Informatics, The Ohio State University, 3190 Graves Hall, 333 West 10th Avenue, Columbus, OH 43210, USA. Fax: +1 614 688 6600.

E-mail address: [khuang@bmi.osu.edu](mailto:khuang@bmi.osu.edu) (K. Huang).



Using the proposed workflow, we analyzed six placenta samples which included three normal controls and three mutant ( $Rb^{-/-}$ ) samples. A mouse placenta contains a maternally derived decidua layer and two major extra-embryonic cell derivatives namely, labyrinth trophoblasts and spongiotrophoblasts (Fig. 1). Placental vasculature that lays embedded within the labyrinth layer is the main site of nutrient-waste exchange between mother and fetus and consists of a network of maternal sinusoids interwoven with fetal blood vessels. The quantitative analysis of placenta samples validates observations published in [10] that  $Rb$ -deficient placenta suffer from a global disruption of architecture marked by increased trophoblast proliferation, a decrease in labyrinth and vascular volumes, and disorganization of the labyrinth-spongio-trophoblast interface.

To summarize, in this paper, we report the architecture and implementation of a complete microscopic image processing workflow as a novel universal 3D phenotyping system. The resulting 3D structure and quantitative measurements on the specimen enable further modeling in systems biology study. While some of the algorithms presented here are optimized for characterizing phenotypical changes in the mouse placenta in gene knockout experiments, the architecture of the workflow enables the system to be easily adapted to countless biomedical applications including our exploration of the organization of tumor microenvironment [16].

### 1.1. Related work

The quantitative assessment of morphological features in biomedical samples is an important topic in microscopic imaging. Techniques such as stereology have been used to assess 3D attributes by sampling a small number of images [17]. Using statistical sample theory, stereological methods allow the researcher to gain insights on important morphological parameters such as cell density and size [18,19]. However, an important limitation of stereology is that it is not useful for large scale 3D visualization and tissue segmentation, both of which are potentially critical for biological discovery. Therefore, we need new algorithms to enable objective large scale image analysis. Since our work involves multiple areas of image analysis research, we delegate algorithmic literature review to the corresponding subsections in Section 2.

There has been some work focusing on acquiring the capability for analyzing large microscopic image sets. Most of these efforts involve developing 3D anatomical atlases for modeling animal systems. For instance, in [20], the authors developed a 3D atlas for the brain of honeybees using stacks of confocal microscopic images. They focus on developing a consensus 3D model for all key functional modules of the brain of the bees. In the Edinburgh Mouse Atlas Project (EMAP), 2D and 3D image registration algorithms have been developed to map the histological images with 3D optical tomography images of the mouse embryo [21]. Apart from atlas related work, 3D reconstruction has also been used in clinical settings. In [1], the authors build 3D models for human cervical cancer samples using stacks of histological images. The goal was to develop an effective non-rigid registration technique and identify the key morphological parameter for characterizing the surface of the tumor mass. In this paper, instead of focusing on a single technique, we present the entire workflow with a comprehensive description of its components (Fig. 2).

## 2. Components and algorithms of the workflow

In this section, we describe the components of the workflow and the related image processing algorithms. Please refer to Fig. 2 for a schematic representation of the three stages.

1. In the first stage, large sets of histological slides are produced and digitized. The preprocessing of the images includes color correction to compensate for intensity inconsistency across slides due to staining variations and pixel-based color classification for segmenting the image components such as cell nuclei, white spaces (including purported vasculature spaces), cytoplasm, and red blood cells. These standard preprocessing steps build the foundation for the next two stages of investigation.
2. The second (middle) stage consists of *image registration and segmentation*. The *registration* process aligns 2D images in a pair-wise manner across the stack. Pair-wise alignments provide 3D coordinate transforms to assemble a 3D volume of the mouse placenta. The *segmentation* process identifies regions corresponding to different tissue structures such as the labyrinth and spongiotrophoblast layers. In our current realization, the *image registration* and *segmentation* process do not directly interact with each other. However, in other applications, results from image segmentation provide the landmarks that may be used in image registration [16].
3. The final stage (bottom) of the workflow supports user-interaction, exploration via visualization and quantification. For this project, the quantification is focused on testing three hypotheses about the effects of  $Rb$  deletion in placental morphology. We provide the hypotheses specifics later in Section 2.6. The quantification step in our workflow provides measurements of morphological attributes relevant to the hypothesis. The visualization step allows the researcher to further study the 3D structures in detail. Volumetric rendering techniques are developed because we are interested in visualization of multiple interleaving types of tissue that will further confirm the quantifications.

The details in the three levels of the workflow are given in Sections 2.1–2.6. Please note that in stage 2, we adopt a multiresolution strategy. For example, image registration/segmentation is carried out at lower resolutions in order to reduce computational costs. Furthermore, we note that the performance of a segmentation algorithm is dependent on the resolution scale. Later stages often process segmented images at different resolutions. Hence, multiple algorithms have been developed for the same technical component.

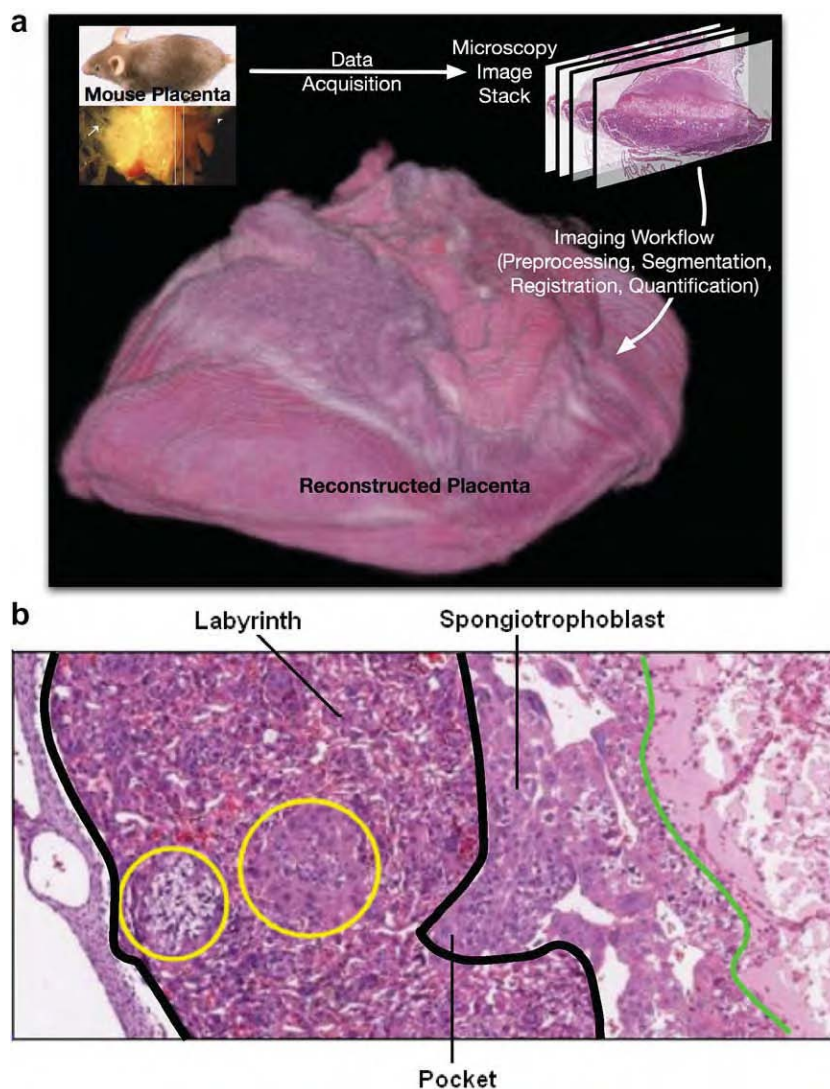
### 2.1. Data acquisition

#### 2.1.1. Image acquisition and stitching

Six mouse placenta samples, three wild-type and three  $Rb^{-/-}$ , were collected at embryonic day 13.5. The samples were fixed in formalin, paraffin-embedded, sectioned at 5- $\mu$ m intervals and stained using standard haematoxylin and eosin (H&E) protocols. We obtained 500–1200 slides approximately for each placenta specimen that were digitized using a Aperio ScanScope slide scanner with 20 $\times$  objective length and image resolution of 0.46  $\mu$ m/pixel. Digitized whole slides were acquired as uncompressed stripes due to the constrained field-of-view of the sensor. The digitization process also produces a metadata file that contains global coordinates of the stripes and describes the extent of any overlap with adjacent stripes. This file is used to reconstruct the digital file of the whole slide from the stripes using a custom Java application that we developed for this purpose.

#### 2.1.2. Image re-sampling

Each serial section produces a digitized RGB format image with dimensions approximately 16 K  $\times$  16 K pixel units. An entire set of the placenta image stacks (each containing approximately 500–1200 images) occupies more than three Terabytes (Tb) of data storage. The processing of such large datasets is beyond the



**Fig. 1.** (a) A mouse placenta reconstructed in 3D with the described imaging workflow. (b) Zoomed placenta image showing the different tissue layers. The tissue between the two thick black boundaries is the labyrinth tissue. The pocket area is an example of the infiltration (interdigitation) from the spongiotrophoblast layer to the labyrinth layer. The cells in the left circle are glycogen cells.

computational capability of most workstations, especially since most imaging algorithms require the full image to be loaded into memory. For certain tasks, it is convenient to down-sample images by a factor of 2–10 depending on the algorithm and performance. The down-sampling process employs linear interpolation to maintain continuity of the features.

## 2.2. Image preprocessing

### 2.2.1. Color correction

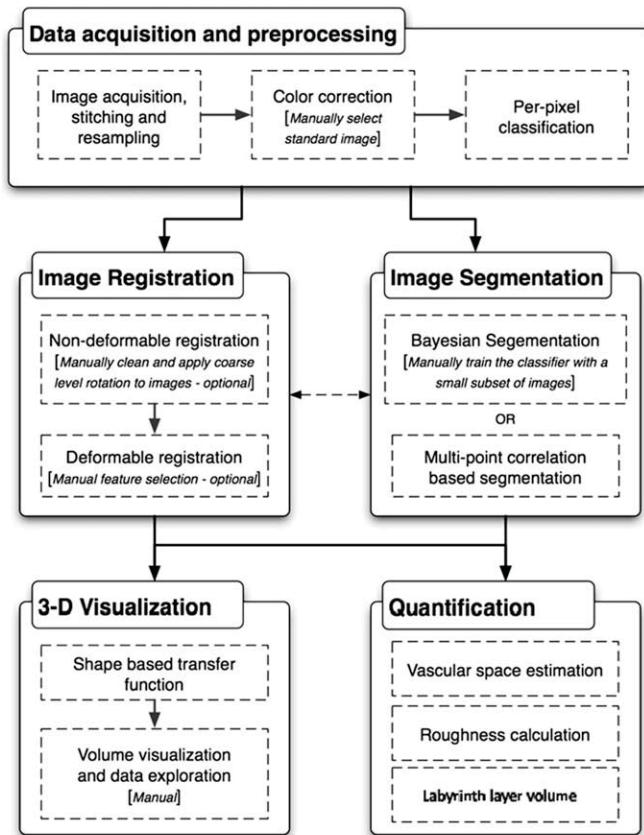
Digitized images of sectioned specimens usually exhibit large staining variations across the stack. This occurs due to idiosyncrasies in the slide preparation process, including section thickness, staining reagents and reagent application time. The process of color correction seeks to provide similar color distributions (histograms) in images from the same specimen. This process greatly facilitates later processing steps, because consistent color profiles narrow the range of parameter settings in algorithms. Color correction is accomplished by normalizing all images in a specimen to a standard color histogram profile. The standard histogram is computed from a manually pre-selected image with a color profile that is representative of the whole image stack.

The color profiles are normalized using MATLAB's Image Toolbox histogram equalization function [22]. We ensure that pixels representing foreground tissue alone participate in the color normalization process. We developed an algorithm to identify foreground tissue pixels from background by thresholding the image in HSL (hue, saturation, and luminance) color space. The HSL color space is less sensitive to intensity gradients within a single image that result from light leakage near edges of glass slides.

### 2.2.2. Pixel-based color segmentation

Pixels in an H&E-stained image correspond to biologically salient structures, such as placental trophoblast, cytoplasm, nuclei, and red blood cells. These different cellular components can be differentiated based on color in each specimen, and the per-pixel classification result is used in image registration and segmentation.

A maximum likelihood estimation (MLE) algorithm is implemented to classify the pixels into four classes in the RGB color space: red blood cells, cytoplasm, nuclei, and background [15]. For simplicity, we assume that the histograms of the bands of data have normal distributions. The *a priori* information related to the four classes is learnt via the following training process. For the image dataset of each placenta specimen (usually contains



**Fig. 2.** The imaging workflow for characterizing phenotypical changes in microscopy data. Components that involve manual intervention are identified.

500–1200 images), one representative image is selected as training image (the same one used to normalize the color profile). A custom-built application randomly selects pixels from the images, displays patches of the training image centering at the selected pixel and highlights the center pixels. The user then chooses among the four classes and a pass option. This procedure provides the training samples and their classifications from manual input. The spatial locations and RGB triplet values are used as attributes for these randomly selected pixels. The covariance matrices, mean values and prior probability weights are then calculated for each individual class. The maximum logarithmic probability rule is invoked to determine the final class membership. Here, the pixels classified as background are from three possible sources. One source is the white background of the images. In each image, the foreground (the region corresponding to the specimen) is surrounded by a large region of white background space. Therefore, pixels in the largest region of background can be easily removed. Another source of background pixels is the white space in the blood vessels. Since most red blood cells are removed from the blood vessel during the preparation of the slides, the regions corresponding to cross-sections of blood vessels usually appear in the form of small white areas with a small number of red pixel clusters (red blood cells). The pixels corresponding to the blood vessels are important in determining the area of vasculature space in the images. The third source of white pixels is the cytoplasm areas for large cells such as giant cells in the spongiotrophoblast layer and the glycogen cell clusters. An example of the pixel classification result is shown in Fig. 3. The classification results are used in the subsequent stages based on requirements in classification granularity.

### 2.3. Image registration

During the slide preparation process, a tissue section is mounted with a random orientation on the glass slide. The section remains displaced in orientation and offset from the previous sliced section. The nature of physical slicing causes deformation and non-linear shearing in the soft tissue. Image registration seeks to compensate for the misalignment and deformation by aligning pair-wise images optimally under pre-specified criteria. Hence, image registration allows us to assemble a 3D volume from a stack of images. In our study, we employ rigid and non-rigid registration algorithms successively. While rigid registration provides the rotation and translation needed to align adjacent images in a global context, it also provides an excellent initialization for the deformable registration algorithm [1]. Non-rigid registration compensates for local distortions in an image caused by tissue stretching, bending and shearing [22,24,26–29].

#### 2.3.1. Rigid registration algorithms

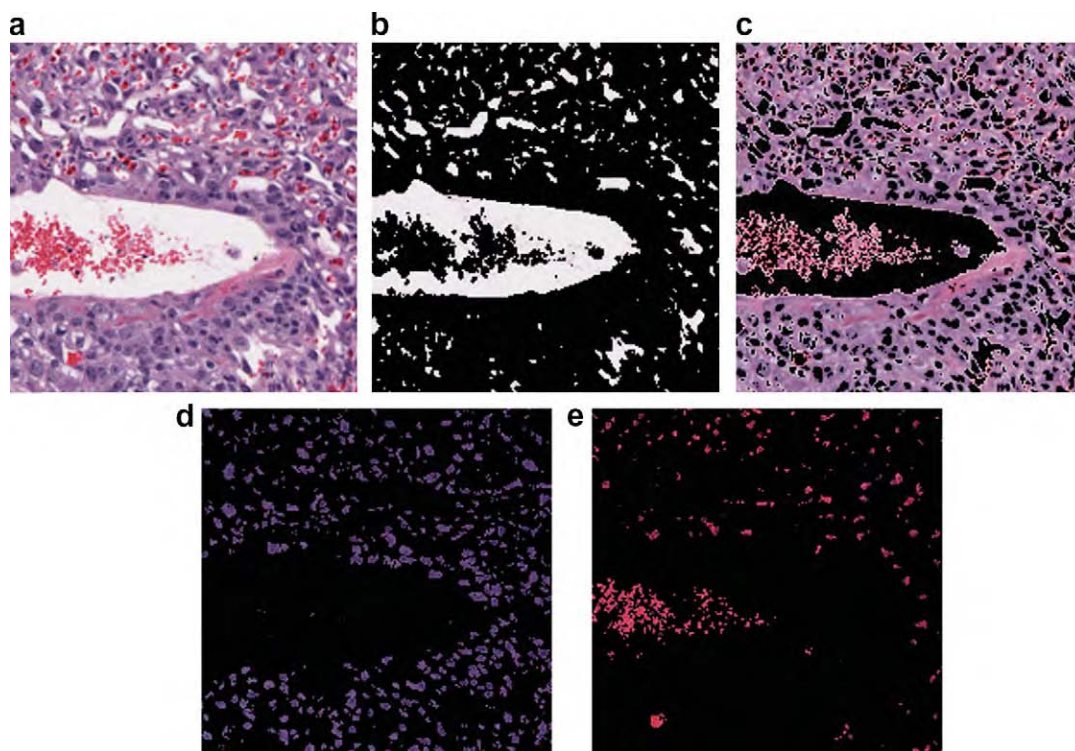
Rigid registration methods involve the selection of three components: the image similarity metric (cost function), the transformation space (domain), and the search strategy (optimization) for an optimal transform. We present two algorithms for rigid registration. The first algorithm is used for reconstructing low-resolution mouse placenta images. The second algorithm is optimized for higher resolution images.

**2.3.1.1. Rigid registration via maximization of mutual information.** This algorithm exploits the fact that the placenta tissue has an elongated oval shape. We carry out a principal component analysis of the foreground region to estimate the orientation of the placenta tissue. This orientation information is used to initialize an estimate of the rotation angle and centroid translation. After the images are transformed into a common coordinate reference frame, a maximum mutual information based registration algorithm is carried out to refine the matching [12,23]. The algorithm searches through the space of all possible rotation and translations to maximize the mutual information between the two images.

MI based methods are effective in registering multi-modal images where pixel intensities between images are not linearly correlated. While the placenta images are acquired using the same protocol, they have multimodal characteristics due to staining variations and the occasional luminance gradients. Rigid body registration techniques requiring intrinsic point or surface-based landmarks [41] and intramodal registration methods [42] that relying on linear correlation of pixel values are inadequate under these conditions.

It has been shown [43] that MI registration with multiresolution strategies can achieve similar robustness compared to direct registration. Studholme and Hill [44] reported no loss in registration precision and significant computational speed-up when comparing different multiresolution strategies. We adopt the multiresolution approach, using 3-level image pyramids. The image magnifications used were 10 $\times$ , 20 $\times$ , and 50 $\times$ . Optimal transforms obtained from a lower magnification are scaled and used as initialization for registration of the next higher magnification. Registration is then performed on the images, potentially with different optimizer parameters, to refine the transforms. The process is repeated for each magnification level to obtain the final transforms. We note that at magnifications higher than 50 $\times$ , the computation cost for registration outweighs the improvements in accuracy. The details of the implementation can be found in [12].

**2.3.1.2. Fast rigid registration using high-level features.** This algorithm segments out simple high-level features that correspond to anatomical structures such as blood vessels using the color-based



**Fig. 3.** An example of the color segmentation. (a) A 200-by-200 pixels patch of the original image (down-sampled by four times for visualization purposes). (b) Segmented background region. Most of the white background regions correspond to blood vessels. A small fraction of them (in the bottom left corner of the image) correspond to cytoplasm regions for the large cells in the spongiotrophoblast layer. (c) Segmented cytoplasm region. (d) Segmented cell nuclei region. (e) Segmented red pixels corresponding to the remaining red blood cells in the blood vessels.

segmentation results in both images. Next, it matches the segmented features across the two images based on similarity in areas and shapes. Any two pairs of matched features can potentially be used to compute rigid transformation between the two images. The mismatched features are removed with a voting process, which selects the most commonly derived rigid transformation (rotational and translational) from the pairs of matched features. This algorithm was implemented to register large images with high speed [11].

### 2.3.2. Non-rigid registration

In our workflow, the rigidly registered image stack serves as input for further refinement using non-rigid methods. In order to visualize a small localized tissue microenvironment, non-rigid registration was conducted by manually selecting point features in each slice of the microenvironment. While we obtained good quality visualizations, repeating this procedure is cumbersome and forced us to consider automated techniques.

There are many previous studies on automatic non-rigid registration [45–48]. Johnson and Christensen present a hybrid landmark/intensity-based technique [45]. Arganda-Carreras et al. present a method for automatic registration of histology sections using Sobel transforms and segmentation contours [47]. Leung and Malik et al. use the powerful cue of contour continuity to provide curvilinear groupings into region-based image segmentation [48]. Our data does not, however, have well defined contours on a slice by slice basis. Thus, contour based registration techniques fail on our dataset.

In our approach, automated pair-wise non-rigid registration is conducted by first identifying a series of matching points between images. These points are used to derive a transformation by fitting a non-linear function such as a thin-plate spline [26] or polynomial functions [25,28]. We have developed an automatic procedure for

selecting matching points by searching for those with the maximum cross correlation of pixel neighborhoods around the feature points [11].

Normally, feature points in an image are selected based on their prominence. Our approach differs with the previous ones in that we select points uniformly. For instance, we choose points that are 200 pixels apart both vertically and horizontally. The variation in a  $31 \times 31$  pixels neighborhood centered at each sampled point is analyzed. The selection of the neighborhood window size depends on the resolution of the image so that a reasonable number of cells/biological features are captured. Please note that we only retain feature points belonging to the foreground tissue region. The neighborhood window is transformed into the grayscale color space and its variance is computed. We retain the selected point as a feature point only when the variance of the neighboring window pixel intensity value is large enough (which implies a complex neighborhood). The unique correspondence of a complex neighborhood with a novel region in the next image is easy to determine. On the other hand, regions with small intensity variance tend to generate many matches and prone to false-positives. For example, consider an extreme example in which a block of white space can be matched to many other blocks of white spaces without knowing the correct match. This step usually yields about 200 feature points that are uniformly distributed across the foreground of each image.

In the second step, we rotate the window around the feature point by the angle that is already computed in the rigid registration procedure. This gives a template patch for initialization in the next image. In the next image, a much larger neighborhood (e.g.,  $100 \times 100$  pixels) is considered at the same location. A patch in this larger neighborhood with the largest cross correlation with the template patch from the first image is selected. The center of this patch is designated as the matching feature point. The two steps

together usually generate more than 100 matched feature points between the two images. These points are then used as control points to compute the non-linear transformation using the thin-plate splines or polynomial transformations [25,28]. In this project, we tested both six-degree polynomial transformations and piecewise affine transformations. The 3D reconstructions are similar in both schemes while the piecewise affine transformation is easier to compute and propagate across a stack of images. Fig. 4 shows renderings of the placenta that were reconstructed using the rigid and deformable registration algorithms. This approach is used to generate high-resolution 3D reconstructions of the samples.

#### 2.4. Image segmentation

In processing biological images, a common task is to segment the images into regions corresponding to different tissue types. For analysis of the mouse placenta, we segmented images into three tissue types, labyrinth trophoblast, spongiotrophoblast, and glycogen cells (a specialized derivative of the spongiotrophoblast lineage). Each H&E-stained tissue type can be classified by distinctive texture and color characteristics of cell nuclei and cytoplasm and by presence of vacuoles and red blood cells. The

segmentation algorithm, therefore, is based on object texture, color, and shape.

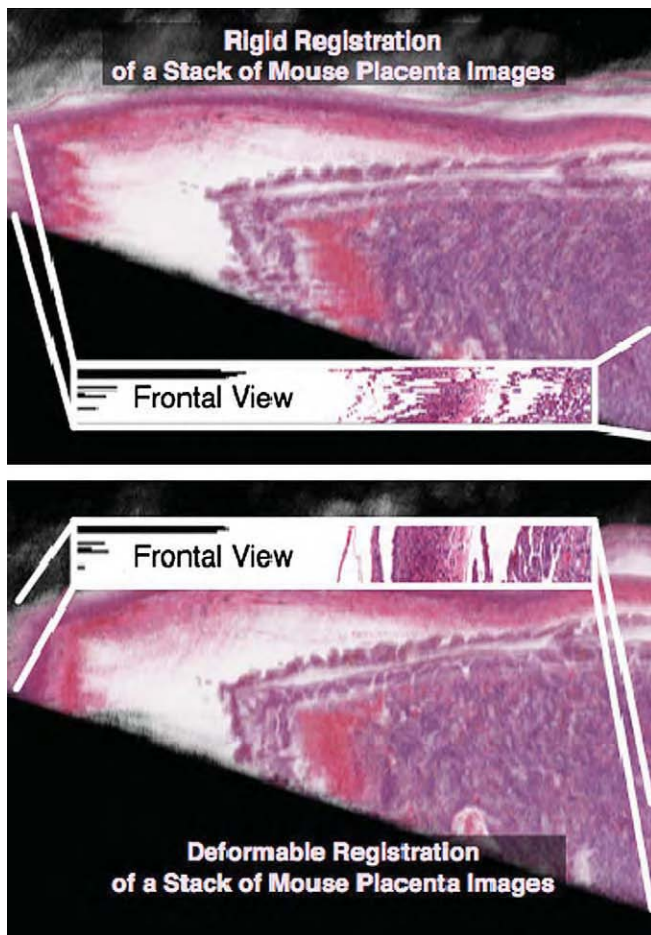
The automatic segmentation of natural images based on texture and color has been widely studied in computer vision [30–32]. Most segmentation algorithms contain two major components: the image features and the classifier (or clustering method). Image features include pixel intensity, color, shape, and spatial statistical features for textures such as Haralick features and Gabor filters [33,34]. A good set of image features can substantially ease the design of the classifier. Supervised classifiers are used when training samples are available. Examples of such classifiers include Bayesian classifier, K-nearest neighbor (KNN), and support vector machine (SVM). If no training example is available, unsupervised clustering algorithms are needed. Examples of such algorithms are K-means, generalized principal component analysis (GPCA) [32], hierarchical clustering, and self-organizing maps (SOM). Active contour algorithms, such as the level-set based ones [35,37,39], can also be considered as an unsupervised method.

In our project, both manual and automatic segmentation procedures have been conducted on the image sets. For each placenta, manual segmentation of the labyrinth layer was carried out on ten images that are evenly spaced throughout the image stack. These manually segmented images are used as the ground-truth for training and testing the automatic segmentation algorithms. In addition, manual segmentation allows for higher level of accuracy in the estimation of area of the labyrinth layer, which also translates to more accurate volume estimates. However, manual segmentations are not feasible for the purpose of visualizing the boundary between the labyrinth and the spongiotrophoblast layers since it is impractical to manually segment all the images. Instead, we adopted automatic segmentation for this purpose.

##### 2.4.1. New features for histological images

In histology-based microscopy images, there has been little work on the automatic segmentation of different types of tissues or cell clusters in histological images. Due to the complicated tissue structure and large variance in biological samples, none of the commonly used image segmentation algorithms that we have tested can successfully distinguish the biological patterns in microstructure and organization [13]. To solve this problem, we designed new segmentation algorithms. The idea was to treat each tissue type as one type of heterogeneous biomaterial composed of homogeneous microstructural components such as the red blood cells, nuclei, white background and cytoplasm. The distribution and organization of these components determine the tissue type. For such biomaterials, quantities such as multiple-point correlation functions (especially the two-point correlation function) can effectively characterize their statistical properties [36] and thus serve as effective image features.

The two-point correlation function (TPCF) for a heterogeneous material composed of two components is defined as the probability that the end points of a random line with length  $l$  belong to the same component. TPCF has been used in analyzing microstructures of materials and large images in astrophysics. However, our study marks the first time that TPCF is introduced in characterizing tissue structures in histological images. For materials with more than two components, a feature vector replaces the probability with each entry being the correlation function for that component. In our work, the four components are cell nuclei, cytoplasm, background, and red blood cells, which are obtained through pixel classification in the preprocessing stage. In addition to the two-point correlation function, three-point correlation function and lineal-path function can also be similarly defined. These functions form an excellent set of statistical features for the images, as demonstrated in Section 3.



**Fig. 4.** Comparison of rigid and deformable registration algorithms. A stack of 25 images were registered using rigid registration algorithm (top) and non-rigid registration algorithm (bottom) and the 3D reconstruction results are rendered. The frontal views show the cross-sections of the reconstructed model. The benefits of using deformable registration algorithms are clearly visible in the frontal view of the image stack cross-section. In the top frontal view which is the cross-section of the rigid registered images, the structures are jaggy and discontinuous. In the bottom frontal view, the results from the non-rigid (deformable) registration algorithm display smooth and continuous structures.

#### 2.4.2. Supervised classification

In addition to feature selection, another aspect of the segmentation problem is to select the classification procedure. In our project, we selected the K-nearest neighbor (KNN) due to both its effectiveness and easy implementation [14]. For each placenta specimen, about 500–1200 serial images are generated. Due to the large variation in morphology, intensity and cell distributions across the different placenta datasets, the KNN classifier is trained on a per placenta dataset basis prior to segmenting all the images.

Within each placenta dataset, 10 evenly spaced images were selected from the stack. These 10 images were then manually segmented by the pathologist. A representative image of the 2D morphology for this placenta specimen was selected by the pathologist as the training sample from the set of 10 images. Image patches of size 20-by-20 pixels were randomly generated and labeled as labyrinth, spongiotrophoblast, glycogen cells or background. A patch lying on the boundaries remained ambiguous and was not chosen into the training dataset. A total of 2200 regions were selected from the image slide (800 for labyrinth, 800 for spongiotrophoblast, and 600 for the background) for training. Please note that the color correction of the serial section stacks (Section 2.1.2) allowed the tissue components to share similar color distributions across the images and hence training based on a representative slide was applicable throughout. The remaining nine images were used for validation purposes as ground-truth.

#### 2.4.3. Evaluation of the automatic segmentation algorithm

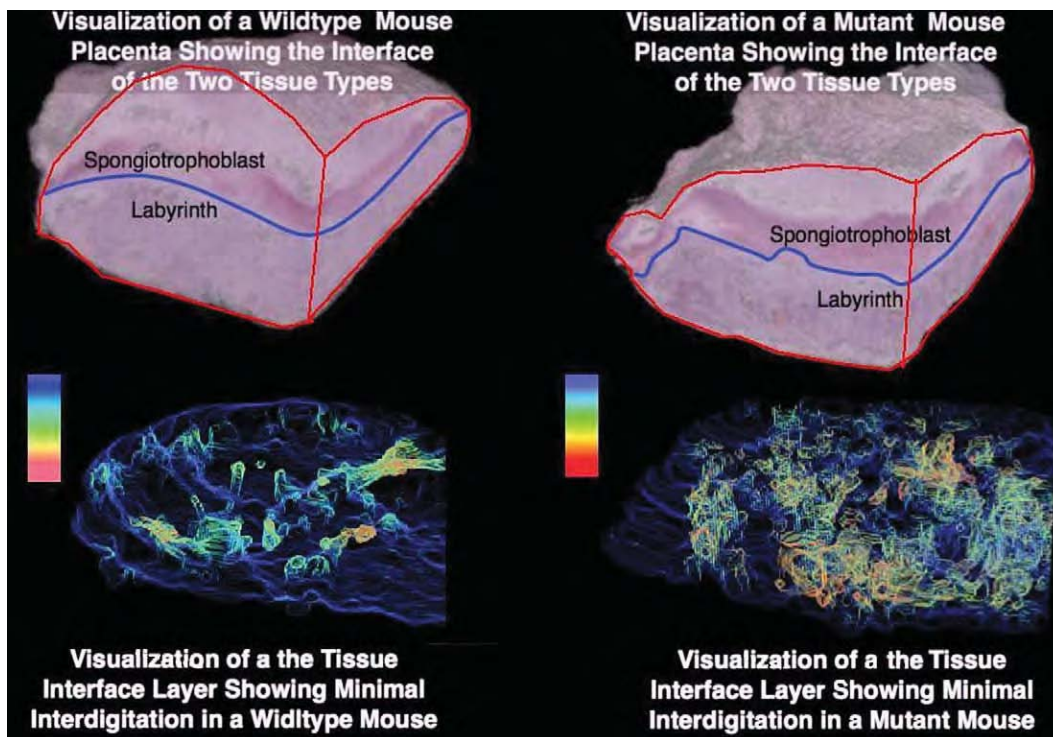
In our study, we found that automatic segmentation tends to generate relatively large error in images obtained from the end regions of the placenta slice sequence, which can bias the volume estimation. However, for the mid-section of the sequence, automated segmentation provided a visually satisfactory boundary between the two layers of tissues. These tests were carried out in three placentae with one control and two mutants. The observation is further confirmed by a quantitative evaluation process as shown in Fig. 6. In the figure, the automatically segmented labyrinth is overlaid on the manually segmented labyrinth tissue. For all the manually segmented images, the error is measured as the ratio between the area encircled by the two tissue boundaries (manually and automatically generated boundaries) and the manually segmented labyrinth area. For the three samples, the mean errors are  $6.6 \pm 1.6\%$ ,  $5.3 \pm 3.3\%$ , and  $16.7 \pm 7.4\%$ . The two samples (one control and one mutant) with mean error less than 8% are then used for visualization. As shown in Fig. 6e and f, the discrepancy between the two segmentation methods can be attributed to two major factors: the use of a large sliding window in automatic segmentation which leads to the “dilation effect”, and the discrepancy in assigning the large white areas on the boundary. This white region is actually the cross-section of a blood vessel at the boundary of the labyrinth tissue layer and the spongiotrophoblast tissue layer. The designation of such regions usually requires post-processing based on explicit anatomical knowledge which is not incorporated in the current version of the automatic segmentation algorithm.

rinth is overlaid on the manually segmented labyrinth tissue. For all the manually segmented images, the error is measured as the ratio between the area encircled by the two tissue boundaries (manually and automatically generated boundaries) and the manually segmented labyrinth area. For the three samples, the mean errors are  $6.6 \pm 1.6\%$ ,  $5.3 \pm 3.3\%$ , and  $16.7 \pm 7.4\%$ . The two samples (one control and one mutant) with mean error less than 8% are then used for visualization. As shown in Fig. 6e and f, the discrepancy between the two segmentation methods can be attributed to two major factors: the use of a large sliding window in automatic segmentation which leads to the “dilation effect”, and the discrepancy in assigning the large white areas on the boundary. This white region is actually the cross-section of a blood vessel at the boundary of the labyrinth tissue layer and the spongiotrophoblast tissue layer. The designation of such regions usually requires post-processing based on explicit anatomical knowledge which is not incorporated in the current version of the automatic segmentation algorithm.

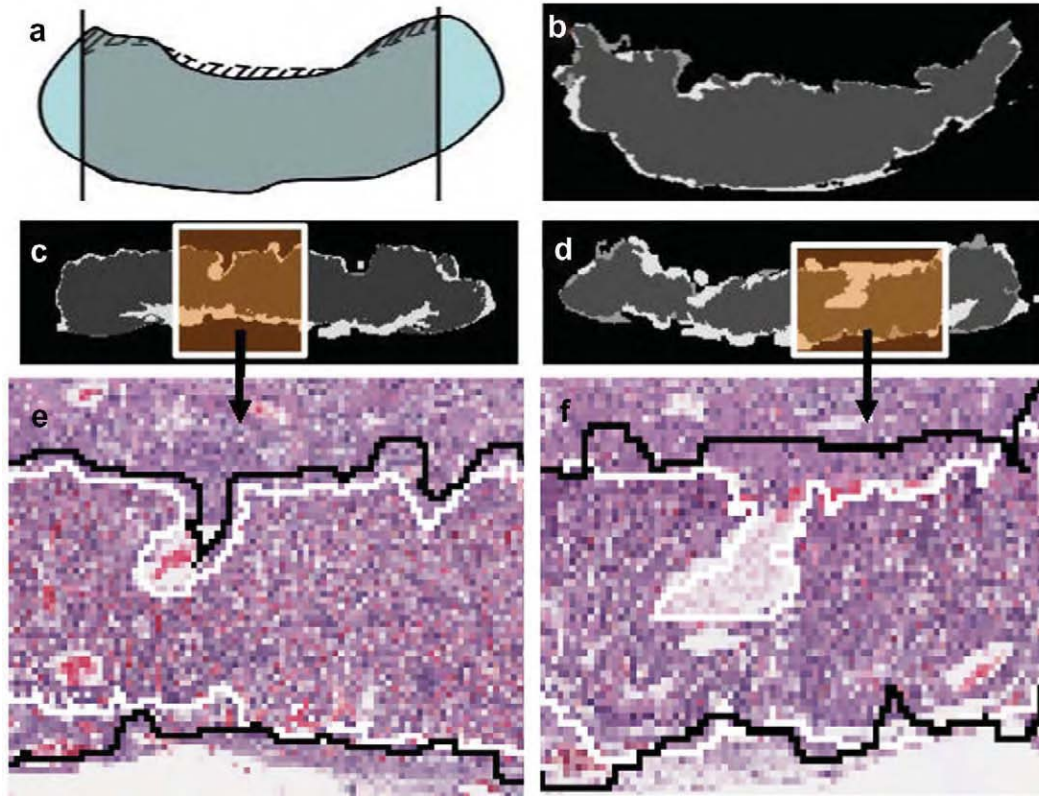
#### 2.5. Visualization in the 3D space

We are interested in quantifying the 3D finger-like infiltration (referred as pockets) that occurs on the labyrinth–spongiotrophoblast tissue interface of the mouse placenta (Fig. 5). The presence of pockets has a direct correlation with surface morphological parameters such as interface surface-area, convolutedness, and the extent of tissue infiltration.

The registered stack of images is treated as volume data and visualized using volumetric rendering techniques. In volumetric rendering, a transfer function maps the feature value (e.g., pixel intensity) to the rendered color and opacity values. It allows the user to highlight or suppress certain values by adjusting the transfer function. In our approach, we evolve a front in the close vicinity of the target surface. The front initially represents a global shape of the surface without pockets. As the front progresses towards the target surface, it acquires the features on the surface and finally



**Fig. 5.** Visualizing the interdigitation at the interface of the labyrinth and the spongiotrophoblast tissue layers in control (left) and mutant (right) mouse placenta. The detected pockets are colored using a heat map. Red regions indicate large pockets and yellow regions indicate shallow pockets.



**Fig. 6.** Evaluation of the automatic segmentation algorithm. (a) The solid line is the manually marked boundary and the dashed line is the automatic segmentation result. The boundary estimation error is defined as the ratio between the shaded area and the gray area. (b–d) Examples of images with boundary estimation errors being 2.5%, 8.4%, and 16.5%. The boundary is in the top portion of the image. The dark gray area is the manual segmentation result, and the light gray area is the automatic segmentation result. (e and f) A larger view of the difference between manual segmentation (black) and automatic segmentation (white).

converges to it. This leads to a natural definition of feature size at a point on the contour as the distance traveled by it from the initial front to the target surface. Surface pockets have larger feature sizes compared to the flat regions owing to the larger distances traversed. Hence, they are suitably extracted. Fig. 5 shows the resultant visualizations from a transfer function that highlights high feature values implemented using the Visualization Toolkit (VTK) from Kitware Inc. [38]. The details of the implementation can be found in [49].

## 2.6. Quantification

Our application requires the quantitative testing of three hypotheses regarding the morphological changes in mouse placenta induced by the deletion of *Rb*. These hypothesized changes include the increased surface complexity between the labyrinth layer and the spongiotrophoblast layer, the reduced volume of the labyrinth layer, and reduced vasculature space in the labyrinth layer. Here, we describe the quantification processes for measuring the three morphological parameters.

### 2.6.1. Characterizing the complexity of the tissue layer interface

*Rb* mutation increases the number of shallow interdigitations at the interface of the spongiotrophoblast and the labyrinth tissue layers. In order to quantify the increased interdigitation, we calculate the number of pixels at the interface and the roughness of the interfacial area between the two layers, based on the assumption that increased interdigitation is manifested as increased area of the interface and greater roughness. The number of pixels at the interface is computed based on the image segmentation results. In addition, given the fractal nature of the surface-area between

the two tissue layers, the boundary roughness is quantified by calculating the Hausdorff dimension, a technique that is well-known and commonly used in geological and material sciences for describing the fractal complexity of the boundary [40]. Typically, the higher the Hausdorff dimension, the more rough the boundary. In order to calculate the Hausdorff dimension, we take the 2D segmented image and overlay a series of uniform grids with cell size ranging from 64 to 2 pixels. Next, we count the number of grid cells that lie at the interface of the two tissue layers. If we denote the cell size of the grids as  $\epsilon$  and the number of grid cells used to cover the boundary as  $N(\epsilon)$ . Then the Hausdorff dimension  $d$  can be computed as

$$d = -\lim_{\epsilon \rightarrow 0} (\ln(N(\epsilon)) / \ln(\epsilon)).$$

In practice,  $d$  is estimated as the negated slope of the log–log curve for  $N(\epsilon)$  versus  $\epsilon$ .

### 2.6.2. Estimating the volumes of the labyrinth tissue layer in mouse placentae

The volume of the labyrinth is estimated using an approach analogous to the Riemann Sum approximation for integration in calculus. The labyrinth volume for a slice is computed from the pixel count of the labyrinth mask obtained from the 2D segmentation, the 2D pixel dimensions, and section thickness. The labyrinth volume is accumulated across all serial sections in a dataset to obtain an approximation of the total labyrinth volume.

### 2.6.3. Estimating the vascularity in the labyrinth tissue layer

The vascularity of the labyrinth is estimated by the ratio of total blood space volume to total labyrinth volume, which is referred to as intravascular space fraction. The estimation of total labyrinth

volume is described in the above section. The total blood space is calculated by counting all pixels previously classified as red blood cell pixels or as background pixels within the labyrinth tissue. The labyrinth mask generated by the segmentation step is used to identify the background pixel in the second case. The intravascular space fraction is then computed.

### 3. Results: a case study on the effects of Rb deletion in placental morphology

#### 3.1. Manual and automated stages

Whole slide imaging for histology and cytology usually involves a large amount of data and is typically not suitable for manual annotation. Three-dimensional processing of serial sections further motivates the need for automation of different stages in the workflow. However, biological systems are characterized by a high incidence of exceptions, and these are especially evident in systems with high-level of detail such as microscopic imaging. Human intervention and semi-automated approaches are often essential components in image analysis workflows. The manual components are identified in the schematic representation shown in Fig. 2.

#### 3.2. Results

The last stage of the workflow discussed in Section 2 generates results for the application—namely quantified parameters and visualizations. For this project, the quantification is focused on testing the three hypotheses about the effects of *Rb* deletion in placental morphology, namely reduced volumes of the placental labyrinth layer (Section 3.2.3), decreased vasculature space in the labyrinth layer (Section 3.2.4), and increased roughness of the boundary between the labyrinth and spongiotrophoblast layers (Sections 3.2.1 and 3.2.2).

##### 3.2.1. Reconstruction and visualization in 3D

Fig. 5 shows the final reconstructed mouse placenta using rigid registration results. Different tissues are highlighted by incorporating the segmentation results in the transfer function adjustment during volumetric rendering. Earlier, in Section 2.5, we mentioned about the 3D finger-like infiltration that occurs on the labyrinth–spongiotrophoblast tissue interface of the mouse placenta. The presence of pockets has a direct correlation with surface morphological parameters such as interface surface-area, convolutedness and the extent of tissue infiltration. We automatically detect pockets using a level-set based pocket detection approach to determine a pocket size feature measure along the interface [16]. The bottom section of the figure shows the infiltration structure in detail by using these feature measurements in the transfer function. The resulting visualization reveals extensive shallow interdigitation in mutant placenta in contrast with fewer but larger interdigitations in the control specimen. These observations are quantitatively verified by calculating the fractal dimension.

##### 3.2.2. Quantifying complexity of the tissue interface

We first computed the number of pixels at the interface between the two tissue layers in littermates. The number of interface pixels for the controls are 1738 and 2374 (in the images down-sampled by 20 times to save computational cost for the image segmentation algorithm) while the interface pixels for the corresponding mutants are 3413 and 4210, respectively. Therefore in both cases, the numbers of interface pixels are almost doubled in mutants than in controls. However, the result for computing the Hausdorff dimension is not as significant. Among the three pairs of littermates, the increase in the Hausdorff dimensions in mutants comparing to the controls are only 3%, 2.5%, and 0.5% when the grid

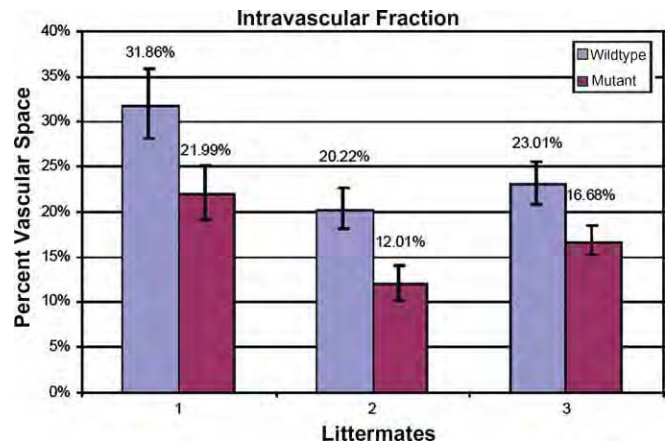


Fig. 7. Intravascular space fraction estimation. The intravascular space fraction is measured for each sample in ten manually segmented images. The mean and standard deviation of the measurement are presented here.

cell sizes between 2 and 64 pixels are used. However, in the mutant placenta, the number of grid cells of size no more than 8 pixels that lie on the interface layer is significantly increased. This suggests that most of the disruption at the interface is due to small shallow interdigitations which are difficult to be characterized using fractal dimensions. This observation supports our result determined in Section 3.2.1 above on surface pockets. Available work in the literature have also reported difficulty in computing fractal dimensions [1].

##### 3.2.3. Volume of labyrinth tissue layer estimation

The volume of the labyrinth tissue layer for each specimen was estimated by summing the areas of the labyrinth layer in each of the ten manually segmented images then multiplying by the distance between consecutive images. This method gives a first order approximation of the labyrinth layer volume. The estimated volumes of the labyrinth layer for the three control mice are 11.0, 9.0, and 12.8 mm<sup>3</sup>. While the measurements for their corresponding littermates are 7.9, 8.2, and 9.3 mm<sup>3</sup>. A consistent reduction of labyrinth layer volume in the range of 9–28% is, therefore, observed for the three pairs of littermates.

##### 3.2.4. Intravascular space fraction estimation

The intravascular space fraction is estimated by combining the color segmentation and image segmentation results. We compute the percentage of white and red pixels in the segmented labyrinth layers. As shown in Fig. 7, for all three pairs of mutant and control samples, significant decrease in intravascular space fraction is observed.

The reduction in the volume and the intravascular space of the labyrinth layer in the mouse placenta is consistent with our hypothesis that *Rb* deletion causes significant morphological disruption in mouse placenta which negatively affects fetal development.

## 4. Conclusion and discussion

In this paper, we presented an imaging workflow for reconstructing and analyzing large sets of microscopy images in the 3D space. The goal of this work is to develop a new phenotyping tool for quantitatively studying sample morphology at tissue and cell level. We developed a set of algorithms that include the major components of the workflow using a mouse placenta morphology study as a driving application. This workflow is designed to acquire, reconstruct, analyze, and visualize high-resolution light



microscopy data obtained from a whole mouse placenta. It allows the researchers to quantitatively assess important morphological parameters such as tissue volume and surface complexity on a microscopic scale. In addition, it has a strong visualization component that allows the researcher to explore complicated 3D structures at cellular and tissue levels. Using the workflow, we analyzed six placenta including three controls and three *Rb*<sup>-/-</sup> mutant with gene knockout and quantitatively validated the hypotheses relating to *Rb* in placenta development [10]. Analysis of placenta indicated that *Rb* mutant placenta exhibit global disruption of architecture, marked by an increase in trophoblast proliferation, a decrease in labyrinth and vascular volumes, and disorganization of the labyrinth–spongiotrophoblast interface. The analytical results are consistent with previously observed impairment in placental transport function [8,10]. These observations include an increase in shallow finger-like interdigitations of spongiotrophoblast that fail to properly invade the labyrinth and clustering of labyrinth trophoblasts that was confirmed with the 3D visualization. Due to the intricacy of carrying out experiments with transgenic animals, we had only a small number of placenta samples which just satisfied the basic statistical requirement. However, the consistent changes in placental morphology we have obtained from large scale image analysis and visualization provide strong evidence to support our hypothesis.

One of the major challenges we faced in the process of workflow development was to strike a good balance between automation and manual work. On one hand, large data size forced us to develop automatic methods to batch process the images. On the other hand, large variations in the images required us to take several manual steps to circumvent the technical difficulties and achieve more flexibility. While this work was largely driven by the mouse placenta study, it is subsequently applied to process other data sets including our ongoing work in phenotyping the mouse breast tumor microenvironment. Other directions include developing a parallel processing framework for handling images in their original high-resolution and a middleware system to support the execution of the workflow on multiple platforms, improving the accuracy of the image segmentation algorithm to obtain higher accuracy and better time performance, and extending the image registration algorithm to deal with images from slides stained with different staining techniques (e.g., H&E versus immunohistochemical staining) so that we can map molecular expression to different types of cells.

## Acknowledgments

This work was supported in part by funding from the National Institutes of Health NIBIB BISTI P20-EB000591, and The National Science Foundation (Grants CNS-0615155 and CNS-0509326).

## References

- Braumann UD, Kuska JP, Einkenkel J, Horn LC, Löffler M, Höckel M. Three-dimensional reconstruction and quantification of cervical carcinoma invasion fronts from histological serial sections. *IEEE Trans Med Imaging* 2005;24(10):1286–307.
- Chen W, Reiss M, Foran D. A prototype for unsupervised analysis of tissue microarrays for cancer research and diagnostics. *IEEE Trans Inf Technol Biomed* 2004;8(2):89–96.
- Chen X, Zhou X, Wong S. Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE Trans Biomed Eng* 2006;53(4):762–6.
- Price DL, Chow SK, Maclean NA, Hakozaki H, Peltier S, Martone ME, et al. High-resolution large-scale mosaic imaging using multiphoton microscopy to characterize transgenic mouse models of human neurological disorders. *Neuroinformatics* 2006;4(1):65–80.
- Sarma S, Kerwin J, Puelles L, Scott M, Strachan T, Feng G, et al. 3d modelling, gene expression mapping and post-mapping image analysis in the developing human brain. *Brain Res Bull* 2005;66(4–6):449–53.
- Jacks T, Fazeli A, Schmitt EM, Bronson RT, Goodell MA, Weinberg RA. Effects of an *Rb* mutation in the mouse. *Nature* 1992;359(6393):295.
- Lee EYHP, Chang C-Y, Hu N, Wang Y-CJ, Lai C-C, Herrup K, et al. Mice deficient for *Rb* are nonviable and show defects in neurogenesis and haematopoiesis. *Nature* 1992;359(6393):288.
- Wu L, de Bruin A, Saavedra HI, Starovic M, Trimboli A, Yang Y, et al. Extra-embryonic function of *Rb* is essential for embryonic development and viability. *Nature* 2003;421(6926):942.
- de Bruin A, Wu L, Saavedra HI, Wilson P, Yang Y, Rosol TJ, et al. *Rb* function in extraembryonic lineages suppresses apoptosis in the Cns of *Rb*-deficient mice. *PNAS* 2003;100(11):6546–51.
- Wenzel P, Wu L, de Bruin A, Chong J, Chen W, Dureska G, et al. *Rb* is critical in a mammalian tissue stem cell population. *Genes Dev* 2007;21(1):85–97.
- Huang KLC, Sharma A, Pan T. Fast Automatic registration algorithm for large microscopy images. In: Wong, S, editor. *IEEE/NLM life science systems & applications workshop*. Bethesda, MD; 2006.
- Mosaliganti R, Pan T, Sharp R, Ridgway R, Iyengar S, Gulacy A et al. Registration and 3d visualization of large microscopy images. In: *Proceedings of the SPIE annual medical imaging meetings*; 2006.
- Pan T, Huang K. Virtual mouse placenta: tissue layer segmentation. In: *International conference of the IEEE engineering in medicine and biology society*. Shanghai, China: IEEE Publishing; 2005.
- Ridgeway R, Irfanoglu O, Machiraju R, Huang K. *Image segmentation with tensor-based classification of N-point correlation functions*. In: *Microscopic image analysis with applications in biology (MIAAB) workshop in MICCAI*. Copenhagen, Denmark; 1979.
- Sharp R, Ridgway R, Mosaliganti K, Wenzel P, Pan T, de Bruin A, et al. Volume rendering phenotype differences in mouse placenta microscopy data. *Comput Sci Eng* 2007;38–47.
- Cooper L, Huang K, Sharma A, Mosaliganti R, Pan T. Registration vs. reconstruction: building 3-D models from 2-D microscopy images. In: Auer M, Peng H, et al., editors. *Workshop on multiscale biological imaging, data mining and informatics*. Santa Barbara, CA: University of California; 2006. p. 57–8.
- Gundersen H, Bagger P, Bendtsen T, Evans S, Korbo L, Marcussen N, et al. The new stereological tools: disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. *APMIS* 1988;96:379–94.
- Mouton P, Gokhale A, Ward N, West M. Stereological length estimation using spherical probes. *J Microsc* 2002;206:54–64.
- West M. Stereological methods for estimating the total number of neurons and synapses: issues of precision and bias. *Trends Neurosci* 1999;22:51–61.
- Brandt R, Rohlfing T, Rybak J, Kroczyk S, Maye A, Westerhoff M, et al. A three-dimensional average-shape atlas of the honeybee brain and its applications. *J Comp Neurol* 2005;492(1):1–19.
- Hill W, Baldock R. The constrained distance transform: interactive atlas registration with large deformations through constrained distances. In: *Workshop on image registration in deformable environments*. Edinburgh, UK; 2006.
- Gonzalez R, Woods R, Eddins S. *Digital image processing using matlab*. Prentice Hall; 2004.
- Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 1997;16(2):187–98.
- Goshtasby A. 2-D and 3-D image registration: for medical, remote sensing, and industrial applications. Wiley-Interscience; 2005.
- Hajnal J, Derek H, Hawkes D. *Medical image registration*. CRC; 2001.
- Bookstein F. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans Pattern Anal Mach Intell* 1989;11(6):567–85.
- Crum W, Hartkens T, Hill D. Non-rigid image registration: theory and practice. *Br J Radiol* 2004;77:S140–53.
- Rohr K. *Landmark-based image analysis: using geometric and intensity models*. Springer; 2007.
- Bajcsy R, Kovacic S. Multiresolution Elastic Matching. *Comput Vis Graph Image Process* 1989;46:1–21.
- Belongie S, Carson C, Greenspan H, Malik J. Color- and texture-based image segmentation using Em and its application to content-based image retrieval. In: *IEEE international conference on computer vision*. IEEE Publisher; 1998.
- Chen Y, Li J, Wang J. *Machine learning and statistical modeling approaches to image retrieval*. Kluwer Academic Publishers; 2004.
- Hong W, Wright J, Huang K, Ma Y. Multiscale hybrid linear models for lossy image representation. *IEEE Trans Image Process* 2006;15(12):3655–71.
- Haralick R, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;SMC-3(6):610–21.
- Saito T, Kudo H, Suzuki S. Texture image segmentation by optimal gabor filters. In: *The third IEEE international conference on signal processing*. IEEE Publishing; 1996.
- Chan T, Vese L. Active contours without edges. *IEEE Trans Image Process* 2001;10(2):266–77.
- Torquato S. *Random heterogeneous materials: microstructure and macroscopic properties*. Interdisciplinary applied mathematics, vol. 16. New York: Springer; 2002 [xxi, 701p].
- Yoo T. *Insight into images: principles and practice for segmentation, registration, and image analysis*. AK Peters; 2004.
- Kitware. *VTK User's Guide Version 5*. 5th ed: Kitware.
- Caselles V, Kimmel R, Sapiro G. Geodesic active contours. *Int J Comput Vis* 1997;22(1):61–79.

- [40] Falconer K. Fractal geometry: mathematical foundations and applications. 2nd ed. Wiley; 2003.
- [41] Maintz JAB. Retrospective registration of tomographic brain images, Ph.D. dissertation. Utrecht, The Netherlands: University of Utrecht; 1996.
- [42] Maes F, Vandermeulen D, Suetens P. Medical image registration using mutual information. *Proc IEEE* 1996;91(10):1699–722.
- [43] Maes F, Vandermeulen D, Suetens P. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Med Image Anal* 1999;3(4):373–86.
- [44] Studholme C, Hill DLG. Automated 3-D registration of MR and CT images of the head. *Med Image Anal* 1996;1(2):163–75.
- [45] Johnson H, Christensen G. Consistent landmark and intensity based image registration. *IEEE Trans Med Imaging* 2002;21:450–61.
- [46] Chui H, Rambo J, Duncan JS, Schultz R, Rangarajan A. Registration of cortical anatomical structures via robust 3d point matching. In: *IPMI '99: Proceedings of the 16th international conference on information processing in medical imaging*; 1999. p. 168–81.
- [47] Carreras A, Fernandez-Gonzalez R, Ortiz de Solorzano C. Automatic registration of serial mammary gland sections. In: *Proceedings of the 26th annual international conference of the IEEE EMBS*; 2004.
- [48] Leung T, Malik J. Contour continuity in region based image segmentation. In: *Lecture notes in computer science*, vol. 1406; 1998. p. 544–59.
- [49] Mosaliganti K, Janoos F, Sharp R, Ridgway R, Machiraju R, Huang K, et al. Detection and visualization of surface-pockets to enable phenotyping studies. *IEEE Trans Med Imaging* 2007;26(9):1283–90.

## Protocol Identification

**Protocol Title:** Animal Models of Cancers and Development (Renewal 1)  
**Principal Investigator:** Gustavo Leone  
**Protocol Type:** Research  
**Academic Unit:** Molecular Genetics

## Protocol Aims

**Describe the overall goals of the proposed project and detail the scientific problem/question being addressed.**

The main goal of this study is to understand, mechanistically, the function of transcription factors, tumor suppressor genes and oncogenes in cancer and development. Mammalian organs and tumors are both complex mixture of different cells that require the interaction and communication between them, in order for their respective development and function. We plan to use several mouse genetic models to test these ideas in several organs. Using mouse genetics, we can eliminate, reduce or increase gene function to determine their effect on organ development and tumor initiation, growth and metastasis.

**Describe the relevance the proposed project will have to human and/or animal health, the advancement of knowledge, and/or the benefit for society.**

If successful, this type of research could identify the genes necessary for both cell types to communicate to each other. Therefore, the basic knowledge gained from our studies could directly and indirectly drive the development of cancer therapies for the treatment of cancer in humans.

## Rationale for Using Animals

**Provide rationale for the use of live vertebrate animals and/or vertebrate animal tissues.**  
The rationale should include reasons why non-animal models cannot be used.

There are no in vivo tissue culture systems or computer model systems in which we could accurately model the complex genetic and cell interactions we want to study. Animals are needed because cancer is a complex disease involving many different cell types in a unique context. Thus, to effectively study cancer, we require that the studies be performed in the most physiological setting possible. Current mouse model systems give us the ability to dissect the action of specific genes in specific cell types, which is perfectly suited for our studies.

## Unnecessary / Duplicative Research

**To the best of your knowledge, does the work proposed in this protocol duplicate previously published research?**

Yes  No

## Protocol Groups

**Define the group(s) of animals to be used in this protocol.**

Protocol groups are used to link activities, housing, and number justification within a study. Properly defining your animal groups is fundamental to the development of your protocol in this system.

- Each group may include only one species.
- All animals of a single species may be included in one protocol group if they will undergo basically the same activities.
- Do not use extra protocol groups unless the animals will experience substantially different protocol activities for a species. Keeping groups to the minimum required will streamline data entry and minimize the chance for errors.

Name each group with a short descriptive label that will readily identify that animal group.

### Protocol Group(s)/Species:

1. Cellular Growth and Development: the role of E2F's and tumor suppressors/Mouse, Standard (Mus Musculus).
2. Cancer and Inflammation: the role of E2F's and tumor suppressors/ Mouse, Standard (Mus Musculus).

## Rationale for Species

**Provide a rationale for the appropriateness of each species to be used in regards to addressing the scientific problem/question.**

The mouse has been an important model for the scientific study of cancer for nearly forty years. The advent of transgenic technologies allows the genome of the mouse to be specifically altered. This model is perfectly suited for the genetic studies we propose here. We can examine the action of an oncogene within tumor cells in the context of other gene alterations in the surrounding environment, experiments that can't be accomplished in a mammalian cell culture model system.

## Transgenic Animal Use

**Does this research involve transgenic species other than rodents?**  Yes  No

**If yes, will you be creating the transgenic rodents?** The term "creating" *does not include*: breeding, purchasing, transferring from another protocol, or transferring from another investigator.  Yes  No

## Animal Sources

**Add the sources from which animals in this protocol will be obtained.**

- You must identify at least one source for each protocol group.
- You may apply one source to multiple protocol groups.

Source            Protocol Group(s) / Species

### **Obtained Externally: From Vendor-**

1. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### **Obtained Externally: Donation-**

1. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### **Obtained Internally: Breeding-**

1. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

## Breeding Plan:

Breeders will be grouped as one male and one or two females. If it becomes necessary to group three females with one male, females will be monitored for plugs or pregnancies and separated from the group. Breeders will be retired and euthanized after 10 months or failure to produce litters. We are expecting approximately 8 pups per litter, however smaller litter sizes may occur due to certain crosses, although this is unknown.

The following formula is an example how we arrive at the numbers required to generate and maintain our genetic lines. At a 1 to 2 (1:2) ratio during the life of the protocol: 30 breeding pair/strain (=10 pair per year x 3 years) x 2 litters/pair x 8 pups/litter = 480 mice total;  $480 \times (1/2) = 240$  mice with useful genetics, the remainder to be euthanized or potentially transferred. Initial breeding stock is on hand, any deficiencies will be made up through one of three different methods: 1) transferred from another (internal) protocol, 2) purchased externally from an approved vendor or 3) recovery through in vitro fertilization from archived frozen semen.

Genotyping will be performed at <21 days of age via tail snips of <5mm. Only offspring of the appropriate genotype (and sex) can be used for future breeding or experiments. Unused offspring will be euthanized, or, if appropriate, will be transferred to approved-protocols for experimental use.

Experimental lines: There are 10 main projects, which all together will require 140890 mice over the next 3 years. The main reason for the large number of animals needed for our studies is that in all the projects, the experimental animals contain at least two transgenes, but in most cases harbor 4 or more of them. To incorporate such a large number of transgenes within one animal, a significant amount of breeding is needed to obtain the desired collection of genes. There is also the possibility that some combinations (particularly mice possessing homozygous transgenic alleles) may have deleterious developmental consequences (leading to infertility or death etc.), permitting only heterozygous crosses.

The number of animals in the Breeder category were determined based on the number of experimental animals required, the expected fertility of the breeder mice (which varies depending on the biology of the particular knock out mouse), and the number of transgenes the breeder mouse carries. The number of animals that will be used in the experimental design also varies between genetic groups. This variation is based on statistical considerations (expected results), the availability of existing data (published or unpublished), whether the genetic groups represent controls and thus can be used in multiple experimental set-ups, and the biology of the particular genetic group.

For example: 25 mice are required for one particular group that requires 5 transgenes in one mouse. The best breeding combination that can be provided will generate 1 test mouse for every 16 mice born (i.e., 1:16). However, only females are required, which now doubles the statistical chance to 1:32; meaning that it will take 800 births to generate the 25 required mice. At 8 births per litter, it would require 100 breeding pairs producing 1 litter each, but it is decided to use 25 pairs producing 4 litters each.

## Housing Locations

Please add the locations where animals will be housed for greater than 12 consecutive hours.

- You must identify at least one housing location for each protocol group.
- You may apply one housing location to multiple groups.

**Do not provide locations where animals will be used only (and not housed).** You will be asked to supply the locations where animals will be used in the "Animal Use Locations" section of this submission.

Housing Location- ULAR Vivarium / ULAR Animal Facility

Protocol Group(s)/Species:

1. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

May receive environmental enrichment: Yes

## Protocol Activities

### I. Non-Survival Surgery - Perfusion

#### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

#### Agent Administration

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
ketamine	PBS	100mg/kg	Intraperitoneal	100-400uL	once	Therapeutic
paraformaldehyde	n/a	to effect	Intravenous	as needed for perfusion	x1	Experimental
PBS	n/a	to effect	Intravenous	as needed for perfusion	x1	Experimental
xylazine	PBS	10 mg/kg	Intraperitoneal	100-400	once	Therapeutic

#### Non-Survival Surgery

**Detail all non-survival surgical procedures to be performed addressing all of the following applicable topics:**

- Pre-Operative Care
- Intra-Operative Monitoring Parameters
- Anesthetic Plan
- Surgical Procedure(s)
- Duration of Each Surgical Procedure



## Experimental Procedures:

Procedure: Mice will be anesthetized with ketamine (100mg/kg, i.p.) plus xylazine (10 mg/kg i.p.) using a syringe with a 27G x ½ needle. An effective state of anesthesia will be assessed by the lack of a toe pinch reflex, at which time the chest cavity will be opened to begin perfusion through the heart. Phosphate buffered saline (PBS) will be used to remove blood from the vasculature and continue until the liver becomes pale. At this point tissues could be removed for cell isolation. For most other procedures, the PBS would then be replaced with a fixative (paraformaldehyde), which will be passed through the vasculature for 10 mins. Following perfusion with the fixative, tissue(s) of interest will be removed and processed accordingly.

Reason: Certain analyses are adversely affected by the presence of blood cells and serum proteins in the tissue, such as cell isolation, protein-binding assays, immunohistochemistry and fluorescent imaging. Profusion of the animals prior to removal and processing of tissues will improve the information gathered from such analyses.

## Major Non-Survival Surgery

**\* Does the surgery penetrate and expose a body cavity or produce substantial impairment of physical or physiological function?**

Yes  No

## II. Survival Surgery- Mammary Tissue Transplant

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### Agent Administration

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
Ibuprofen	water	50-80mg/kg	Intraperitoneal	100-200uL	Once a day	Therapeutic
Ibuprofen	water	0.4mg/mL (in water bottle)	Oral	Variable	ad libitum	Therapeutic
Isoflorane	Oxygen/air mix (2L/min oxygen flow)	5% initially then maintained at 1.2 to 2%	Inhalation	N/A	Duration of the procedure	Therapeutic

### Survival Surgery Assurances

**Survival Surgery:** In absence of complications, the animal is expected to recover from anesthesia following the procedure.

For further information, please see the IACUC guidelines on [Post-Operative Care](#) and [Survival Surgery in Rodents](#).

**\*Post-operative analgesics will be administered and have been listed in the "Agent Administration" section above.**  Yes  No

**Post-operative monitoring will be conducted by the study team.**

Yes  No

## Survival Surgery

**Detail all surgical procedures to be performed addressing all of the following applicable topics:**

- Intra-Operative Monitoring Parameters
- Anesthetic Plan
- Analgesics
- Surgical Procedures
- Duration of Each Surgical Procedure
- Post-Operative Care
- Frequency of Observation
- Long Term Care

Experimental Procedures:

Anesthesia and monitoring parameters: The surgeon will pay close attention to animal's level of anesthesia. Respiratory rate, muscle tone, heart rate & toe pinch reflexes can help assess the anesthesia level and additional isoflurane will be administered as needed to maintain a surgical plane.

Analgesics: Ibuprofen as determined by criteria below. Previously the veterinary and research staff have not noted signs of pain due to the quick surgery and small incision size. Mice typically are up and moving about the cage normally within 5 minutes. If any signs of pain/distress are noted during this immediate post-op, an analgesic will be provided.

Pre-operative Considerations and Animal Preparation:

Transplantation of mammary tissue will be performed under aseptic conditions using mice between eight and ten weeks of age. The day prior to surgery, the scapular region of recipient mice are shaved while under anesthesia using Isoflurane.

Surgical Procedures:

On the day of surgery the donor animals are sacrificed by cervical dislocation. The dead mouse is then completely immersed in 70% ethanol solution to sterilize it, transferred to a laminar flow hood and the mammary tissue is extracted after removing any lymph nodes. The excised tissue is kept hydrated by placing it in a sterile PBS solution. Recipient mice are anesthetized with isoflurane in a laminar flow hood and the shaved surgical site is prepped with Nolvasan surgical scrub and 70% isopropyl alcohol solution 3 times each, alternating between first soap and then the alcohol solution. Next a 4-5mm opening is made in the skin near one shoulder blade using forceps and a pair of small straight bladed scissors. A small pair of curved scissors are then inserted into the opening; away from the shoulder blade and along the side of the rib cage to one side in order to separate the skin from the rib cage. The scissors are then partially opened to create a wider pocket for the donated tissue to rest and then carefully removed. A pair of forceps are used to hold the incision open and a second pair of forceps (curved) are used to place the donated tissue into the newly formed sub-dermal pocket. After the donated mammary tissue is placed inside and the forceps removed, the incision is held together with the pair of curved

forceps and then stapled closed using a single 9mm wound clip. The tools are rinsed thoroughly in a 70% ethanol solution and patted dry with sterile gauze to repeat this procedure on the other side of the animal.

**Duration:**

The entire procedure will last approximately 20min, with the recipient mouse being under anesthesia between 10 and 15 minutes.

**Post-operative Monitoring/Observations/Treatment:**

After removing the mouse from anesthesia it is placed in a clean cage on a paper towel with supplemental heat via warming lamp. Once the mice are observed to move about and eat & drink, they are placed back into their animal facility room of origin. The animals will be observed daily for one week by the PI group. Thereafter, between 7 and 14 days post-operation, the incision-closing staples can be removed and animal will no longer be a need for the post-surgery wounds to be so closely monitored. They can be checked during the subsequent monitoring for tumor growth.

**Post-surgery:**

Hyper or hypo-responsiveness to manipulation, lack of normal ambulation, excessive scratching of incision site would require administration of analgesics (Ibuprofen) either orally or by injection depending on the severity of the mouse's condition, or removal from the study (euthanasia).

## **Major Non-Survival Surgery**

**\* Does the surgery penetrate and expose a body cavity or produce substantial impairment of physical or physiological function?**

Yes  No

### III. Survival Surgery- **Hepatectomy**

#### **Protocol Groups**

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

#### **Agent Administration**

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
Buprenorphine	PBS	0.05-0.1mg/kg	Subcutaneous	100-200uL	2 times a day	Therapeutic
Ibuprofen	water (via bottle)	0.8mg/mL	Oral	N/A	N/A	Therapeutic
Isoflurane	Oxygen/air mix (at 2L/min oxygen flow)	5% initially and then 2 to 1.5% to maintain	Inhalation	N/A	For the duration of the procedure	Therapeutic
NaCl solution (saline)	water	0.9% wt/vol	Subcutaneous	1-2mL	As need to alleviate dehydration	Therapeutic
Puralube (eye lubricant)	N/A	N/A	Topical	N/A- enough to coat each eye thoroughly	As needed during surgery	Therapeutic

## Survival Surgery Assurances

**Survival Sugery:** In absence of complications, the animal is expected to recover from anesthesia following the procedure.

For further information, please see the IACUC guidelines on [Post-Operative Care](#) and [Survival Surgery in Rodents](#).

**\*Post-operative analgesics will be administered and have been listed in the "Agent Administration" section above.**  Yes  No

**Post-operative monitoring will be conducted by the study team.**

Yes  No

## Survival Surgery

**Detail all surgical procedures to be performed addressing all of the following applicable topics:**

- Intra-Operative Monitoring Parameters
- Anesthetic Plan
- Analgesics
- Surgical Procedures
- Duration of Each Surgical Procedure
- Post-Operative Care
- Frequency of Observation
- Long Term Care

Experimental Procedures:

Rationale: Hepatocytes undergo a specialized type of cell cycle known as the endocycle. Cells that utilize the endocycle obtain polyploid genomes because cells continue rounds of DNA synthetic (S) phase without an intervening mitotic (M) phase. We are looking to determine the role of two E2F family members known to impact endocycle progression in hepatocyte proliferation and hence liver regeneration, a process that when disrupted contributes to a number of human pathological conditions including viral and toxin-induced liver cirrhosis and cancer.

Surgery: Anesthesia of mice will be induced using 2% isoflurane (at 2L/min oxygen flow), 2-3 minutes; anesthesia will be maintained during surgery by 1.5% isoflurane (1 L/min oxygen flow) inhalation through a mouthpiece.

Animal Prep: Hair at and around the site of incision will be removed by electric clippers while the animal is anesthetized. Puralube<sup>®</sup> will be applied around the rim of the eyes to prevent drying. Animal will then be moved to the location where surgery will be performed, and the incision site will be sterilized 3x with alternating Chlorascrub and alcohol swabs.

Next a horizontal 2-cm midline incision through the abdominal skin, muscle and peritoneum will be made a few mm below the tip of the xiphoid. A sterile 2x2 gauze will be rolled up and inserted underneath the animal to elevate the site of incision and expose the liver.

The falciform ligament will be cut using curved microsurgery scissors, while gently pulling down the median lobe with a saline-moistened cotton tip. The median and left lateral lobe will be held against the diaphragm/thorax with a moistened cotton tip, as the stomach is pulled down with the micro-dissecting forceps to visualize and section the membrane that links the caudate and the left lateral lobe.

A piece of 4-0 silk thread will be looped around the base of the left lateral lobe and a slipknot made as close to the base of the lobe as possible. While gently pulling upward on the loose ends of the silk thread, cauterization will be performed using a high temperature cautery pen at the base of the lobe.

This same course of procedure will be repeated for the removal of the right and caudate lobes, leaving only the median lobe (which is intimate to the gallbladder) intact, thus completing the 2/3 partial hepatectomy.

The peritoneum will be sutured with a 5-0 absorbable suture and then the skin with either wound clips or 4-0 prolonged absorbable (PDS, vicryl) or non-absorbable suture (nylon).

Post-operative care:

- During the immediate post-op period, the animal will be observed closely until full recovery from anesthesia. The animal will be placed on a warming pad (Deltaphase isothermal pad, Braintree Scientific) to prevent hypothermia.
- Analgesics will be administered upon completion of surgery (Buprenorphine, 0.05-0.1 mg/kg BW SC twice a day) for that day and thereafter as needed. Ibuprofen (Motrin, 7.5 mg/kg BW PO) will be maintained in the water bottle on recovery and through the next three days.
- Warm saline will be provided (1-2mL per 100g BW) to the animals based on the criteria of blood loss intra-operatively.
- Animals will be housed individually to prevent cannibalism and will not be returned to the general colony population post-surgery until they appear fully stable and. They will be kept in the surgery room (BRT procedure room) and post-operative records will be kept and posted in this room where the mice are housed.

Long term care:

- External wound clips or sutures will be removed 10-14 days after surgery.
- Post-surgical animals will be monitored on a daily basis until all wounds have healed and sutures or wound clips are removed.

## **Major Survival Surgery**

**\* Does the surgery penetrate and expose a body cavity or produce substantial impairment of physical or physiological function?**

Yes  No

## **Multiple Major Survival Surgeries**

**\* Will any individual animal be used in multiple survival surgical procedures?**

Yes  No

## IV. Survival Surgery- Wound Healing

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### Agent Administration

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
Ibuprofen	water	30-50mg/kg	Intraperitoneal	100-200uL	Once a day	Therapeutic
Ibuprofen	water	0.4mg/mL (in water bottle)	Oral	Variable	ad libitum	Therapeutic
Isoflorane	Oxygen/air mix (2L/min oxygen flow)	5% initially then maintained at 1.2 to 2%	Inhalation	N/A	Duration of the procedure	Therapeutic

### Survival Surgery Assurances

**Survival Surgery:** In absence of complications, the animal is expected to recover from anesthesia following the procedure.

For further information, please see the IACUC guidelines on [Post-Operative Care](#) and [Survival Surgery in Rodents](#).

**\*Post-operative analgesics will be administered and have been listed in the "Agent Administration" section above.**  Yes  No

**Post-operative monitoring will be conducted by the study team.**

Yes  No

### Survival Surgery



**Detail all surgical procedures to be performed addressing all of the following applicable topics:**

- Intra-Operative Monitoring Parameters
- Anesthetic Plan
- Analgesics
- Surgical Procedures
- Duration of Each Surgical Procedure
- Post-Operative Care
- Frequency of Observation
- Long Term Care

**Experimental Procedures:**

Reason: Our lab will use this procedure as an in vivo model to study the effect of tumor suppressor genes function during wound healing and cell migration related to tumor metastasis.

Anesthesia and monitoring parameters: The surgeon will pay close attention to animal's level of anesthesia. Respiratory rate, muscle tone, heart rate & toe pinch reflexes can help assess the anesthesia level and additional isoflurane will be administered as needed to maintain a surgical plane.

**Procedure:**

All mice will be anesthetized by isoflurane inhalation. The surgical site will be clipped of hair and then scrubbed with betadine in a circular, outward expanding pattern. The betadine will be followed by 70% alcohol. This procedure will be repeated three times. When no longer responding to a toe pinch, two full-thickness excisional 8mm wounds will be made through the skin and panniculus in the inguinal regions using a sterile biopsy punch. The biopsy punches will be made between the most caudal two sets of nipples on each side. Sharp scissors will be used to complete the biopsy if needed. The mammary gland underlying the biopsy punches will be exteriorized with forceps and a small section (less than 8mm) of the mammary gland will be excised.

**Duration:**

The procedure will last no more than 10min per mouse

**Post-operative care:**

After the procedure, the mice will be allowed to recover in a warm environment and will be administered 0.05mg/kg buprenorphine SQ. The mice will be housed individually for 24 hours at which time they will be re-evaluated and group housed if there is no trauma to the surgical sites. At this time their pain will also be re-evaluated and additional doses buprenorphine administered if needed (0.05mg/kg SQ).

Ibuprofen (30 mg/kg) will be provided in the drinking water for 5–7 days post-operatively. The mice will be monitored daily during the first 7 days after wounding. If any moisture or discharge is present at the surgical sites, they will be gently cleaned with betadine or chlorhexidine.

Mice will also be weighed prior to the surgical procedure and once daily for five days post-operatively. Mash (feed pellets plus water) will be provided on the floor of the cage for five days post-operatively as well.

Long term:

After the first week the mice will be weighed and monitored once weekly, until euthanized for tissue at determined endpoints.

### **Major Survival Surgery**

**\* Does the surgery penetrate and expose a body cavity or produce substantial impairment of physical or physiological function?**

Yes  No

### **Multiple Major Survival Surgeries**

**\* Will any individual animal be used in multiple survival surgical procedures?**

Yes  No

## V. Survival Surgery- Pulmonary Fibrosis

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### Agent Administration

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
Bleomycin	saline solution	1 µg/µl	intertracheal- injected into the tracheal lumen	50 µl	Once	Experimental
Isoflurane	Oxygen/air mix (2L/min oxygen flow)	5% initially then maintained at 1.2 to 2%	Inhalation	N/A	Duration of the procedure	Therapeutic

### Survival Surgery Assurances

**Survival Surgery:** In absence of complications, the animal is expected to recover from anesthesia following the procedure.

For further information, please see the IACUC guidelines on [Post-Operative Care](#) and [Survival Surgery in Rodents](#).

**\*Post-operative analgesics will be administered and have been listed in the "Agent Administration" section above.**  Yes  No

**Post-operative monitoring will be conducted by the study team.**

Yes  No

## Survival Surgery

Detail all surgical procedures to be performed addressing all of the following applicable topics:

- Intra-Operative Monitoring Parameters
- Anesthetic Plan
- Analgesics
- Surgical Procedures
- Duration of Each Surgical Procedure
- Post-Operative Care
- Frequency of Observation
- Long Term Care

### Experimental Procedures:

This is a new line of work in our lab to study the effect of various mutations of PTEN in pulmonary fibrosis. This work would build on our observations that loss of PTEN is associated with mild to moderate fibrosis in mouse mammary glands. Bleomycin is a chemical agent known to induce lung fibrosis upon intratracheal administration. The following procedure will be used to determine if specific genetic alterations in the mice change this fibrotic response in the mouse lungs after chemical insult.

Anesthesia and monitoring parameters: The surgeon will pay close attention to animal's level of anesthesia. Respiratory rate, muscle tone, heart rate & toe pinch reflexes can help assess the anesthesia level and additional isoflurane will be administered as needed to maintain a surgical plane.

Analgesics: None

Procedure: All animals will be aseptically prepared according to IACUC rodent surgery guidelines. Mice will be anesthetized by isoflurane inhalation and the trachea will be exposed using sterile technique. A single dose of bleomycin (50 $\mu$ g in 50  $\mu$ l saline solution) will be injected into the tracheal lumen using a 26-gauge needle. After the inoculation, the incision will be closed using surgical staple(s).

Post-operative care: After removing the mouse from anesthesia it is placed in a clean cage with supplemental heat via warming lamp. Once the mice are observed to move about and eat & drink, they are placed back into their animal facility room of origin.

Long term care and observations: The mice will be euthanized at predetermined time points after bleomycin injection according to IACUC rodent euthanasia guidelines. The lungs will be harvested histological and biochemical analysis.

## Major Non-Survival Surgery

\* Does the surgery penetrate and expose a body cavity or produce substantial impairment of physical or physiological function?

Yes  No

## VI. Agent Administration (Not Part of Any Other Activity Listed)

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### Agent Administration

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

#### 1. Bromodeoxyuridine

**Agent Name:** Bromodeoxyuridine (5-bromo-2-deoxyuridine, BrdU)

**Vehicle:** PBS (phosphate buffered saline)

**Dose:** 50mg/kg

**Route:** Intraperitoneal

**Volume:** 100-400uL

**Frequency:** 1x 1 to 3 hours prior to euthanasia

**Will this agent be used for therapeutic or experimental purposes?** Experimental

**Is this agent human source material?**  Yes  No

**Is this agent infectious?**  Yes  No

**Is this agent radioactive?**  Yes  No

**Is this agent a chemical hazard?** [carcinogens, mutagens, reproductive hazards, biological toxins, or chemicals with LD50 in rodents of <500 mg/kg, corrosives, or chemicals with known effects on major organ systems (i.e., liver, kidney, central nervous system, hematopoietic system)]

Yes  No

## 2. Polyinosinic–polycytidylic acid

**Agent Name:** Polyinosinic–polycytidylic acid

**Vehicle:** PBS (phosphate buffered saline)

**Dose:** 10mg/kg

**Route:** Intraperitoneal

**Volume:** 200-300uL

**Frequency:** one injection every other day for 5 days.

**Will this agent be used for therapeutic or experimental purposes?** Experimental

**Is this agent human source material?**  Yes  No

**Is this agent infectious?**  Yes  No

**Is this agent radioactive?**  Yes  No

**Is this agent a chemical hazard?** [carcinogens, mutagens, reproductive hazards, biological toxins, or chemicals with LD50 in rodents of <500 mg/kg, corrosives, or chemicals with known effects on major organ systems (i.e., liver, kidney, central nervous system, hematopoietic system)]

Yes  No

## 3. Tamoxifen

**Agent Name:** Tamoxifen

**Vehicle:** corn oil

**Dose:** 4 to 40mg/kg body weight

**Route:** Intraperitoneal

**Volume:** 200-500uL

**Frequency:** one injection every day for 5 days.

**Will this agent be used for therapeutic or experimental purposes?** Experimental

Is this agent human source material?  Yes  No

Is this agent infectious?  Yes  No

Is this agent radioactive?  Yes  No

**Is this agent a chemical hazard?** [carcinogens, mutagens, reproductive hazards, biological toxins, or chemicals with LD50 in rodents of <500 mg/kg, corrosives, or chemicals with known effects on major organ systems (i.e., liver, kidney, central nervous system, hematopoietic system)]

Yes  No

#### 4. $\beta$ -naphthoflavone

**Agent Name:**  $\beta$ -naphthoflavone

**Vehicle:** corn oil

**Dose:** 80 mg/kg body weight

**Route:** Intraperitoneal

**Volume:** 200-500uL

**Frequency:** three injections within 24 hours.

**Will this agent be used for therapeutic or experimental purposes?** Experimental

Is this agent human source material?  Yes  No

Is this agent infectious?  Yes  No

Is this agent radioactive?  Yes  No

**Is this agent a chemical hazard?** [carcinogens, mutagens, reproductive hazards, biological toxins, or chemicals with LD50 in rodents of <500 mg/kg, corrosives, or chemicals with known effects on major organ systems (i.e., liver, kidney, central nervous system, hematopoietic system)]

Yes  No



## VII. Food and Water Modification

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### Agent Administration

**Add all agents (including all anesthetics and analgesics) to be administered/used in this activity.**

For further information, please see the IACUC guideline on Therapeutic Agents (Anesthesia/Analgesia).

Agent Name	Vehicle	Dose	Route	Volume	Frequency	Therapeutic or Experimental
doxy	drinking water or chow	200mg - 3gm/kg	Oral	ad lib	daily x 730d determined by group	Experimental
tamoxifen	water or chow	400 mg/kg	Oral	ad lib	daily for life	Experimental

### Food and/or Water Modification

**Describe the feeding and/or watering schedule to be used and provide scientific justification for why modification is required.** Include the level of modification, the duration of the modification, and the measures that will be taken to prevent adverse effects on the health of the animals. **Note that any pre-surgical food restriction should be detailed in the surgery activities section and does not require completion of this section nor specific justification.**

For further information, please see the IACUC guideline on [Food and Fluid Restriction](#).

Add info for doxy water administration.

1. Commercially prepared chow containing either doxycycline or tamoxifen will be fed to assigned experimental groups in a routine fashion by research personnel. However, ULAR animal caretakers have access to this chow and can add food to the animals cages if necessary during routine, daily care.

2. Water bottles containing doxycycline will be used either as an alternative to the chow containing it (mentioned above), or in addition to it, depending upon the level of expression required. The level of expression is determined from examination of tissues from sacrificed animals. Water bottles will be changed on a weekly basis and water levels every 2-3 days, and daily observations performed by ULAR vivarium personnel, who will report all care issues to vet and research personnel.

GEM animals which are conditional knockouts/ins will receive the special diet/water as per their genotype.

## **Adverse Effects / Early Removal Criteria**

**Add the anticipated adverse effects and early removal criteria for the animals to be used on this protocol.**

- You must identify adverse effects and early removal criteria for each protocol group.
- You may apply one set of adverse effects and early removal criteria to multiple protocol groups.

### **Protocol Groups**

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

### **Adverse Effects**

**Provide a description of any anticipated effects that would negatively impact the health and well-being of the animal (neurologic deficit, tumor burden, disease state, etc.) as a result of the experimental manipulations.**

Potential adverse effects for all studies (both development of new strains and experimental animals) may present as the following (to varying degrees) for both:

1. Hair coat (slickness, greasiness, separated or well groomed)
2. Diminished alertness and mobility
3. Decreased ability to feed and drink (weight loss)
4. Body distress associated with tumor burden (such as loss of body condition, dyspnea, lethargy, rough hair coat, mass location or size preventing mobility)

Potential surgical complications for mammary surgery, wound healing, bleomycin, and hepatectomy surgeries may also present as those listed above, and in addition may include:

1. Inflammation/infection of surgical site
2. Self mutilation (opening of surgical closure/removal of wound clips)
3. Dehydration, blood loss
4. Hepatectomy: surgical dehiscence; severe blood loss
5. Bleomycin (fibrosis): difficult breathing

## Early Removal Criteria

**Provide the criteria (e.g., tumor size, percentage body weight gain or loss, inability to eat or drink, behavioral abnormalities, clinical symptomatology, signs of toxicity) that will be used to determine either the endpoint of the study or when animals will be removed prior to a planned endpoint.** For further information, see the IACUC guideline on Early Removal Criteria and Non-CO<sub>2</sub> Euthanasia.

Overall health/distress of each animal will be evaluated on an individual basis. Determination for removal from a study will be made by the presence of one or more criteria as determined by research personnel. ULAR veterinary personnel will also provide oversight and input regarding condition of animals.

Early Removal Criteria for all studies:

1. Ulceration of tumors\*
2. Unresponsive or immobile
3. Inability to feed and drink
4. Body condition score of less than 2/5 indicating muscle wasting; dyspnea, lethargy, rough hair coat, mass location or size preventing mobility
5. Surgical complications, although not anticipated, will be addressed on an individual basis, such as providing replacement fluids. (Post op treatment specified in protocol activities.) If there is a dehiscence of the surgical site, the animal will be euthanized.

\*Specific to tumor studies-

Severity of ulcerations as assessed by depth (no deeper than 2mm will be allowed), size, signs of pain – i.e. self-mutilation, guarding, etc.) Treatment of ulcerations is not an option as the studies would be significantly biased and difficult to properly control.

NOTE: Animals displaying any of the early removal criteria above will be evaluated by research personnel &/or ULAR vet staff. Treatment or euthanasia will be initiated as appropriate per the veterinary staff.

## Euthanasia Methods

Provide the methods of euthanasia to be used, either as a component of the proposed project or as an emergency procedure if necessary.

Euthanasia must meet the guidelines of the 2007 AVMA Guidelines on Euthanasia. For further information, see the IACUC guidelines on Euthanasia/Death as an Endpoint and CO2 Euthanasia of Rodents including Guinea Pigs.

### Protocol Groups

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus).
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus).

### Euthanasia Method

**Does this euthanasia method apply to:** All

*Note: Euthanasia methods must be defined for all animals being euthanized.*

Select the euthanasia method to be used for each group you have marked above.

Carbon Dioxide (CO2)  
Cervical Dislocation w/ anesthesia  
Decapitation in awake animal  
Exsanguination w/ anesthesia

**Agent for Euthanasia:** CO2  
(May be redundant with method indicated above.)

**Dose:** 70% to effect

**Route:** Inhalation

**Vehicle:** N/A

### Euthanasia Detail

**Detail the method/procedure being used.**

Animals will be placed in CO2 chamber and once breathing ceases, a secondary method of euthanasia will be used as indicated above. For decapitation, a sharp instrument (scissors, scalpel) will be used. For cervical dislocation with anesthesia, once animals are fully anesthetized with isoflurane, the mouse will be cervically dislocated by placing the fingers behind the skull and rapidly pulling on the base of the tail to dislocate the spinal cord from the brain.

For the exsanguination under anesthesia, see the protocol activity "non-survival surgery."

**Confirmation of Euthanasia**

**Regardless of euthanasia method, indicate how death will be confirmed.**

Cervical Dislocation

Decapitation

Removing Vital Organ(s)

## Animal Use Locations- Location 1

**Building:** Biomedical Research Tower

**Room:** Suite 67

If occurring within a ULAR vivarium specify "ULAR" for the room.

**Select the protocol groups to be used at the location selected above.**

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

**Select the activities to be performed at the location selected above.** Be certain that the groups selected in the previous question are included in the activities you are selecting here.

Activity	Protocol Groups
Agent Administration (Not Part of Any Other Activity Listed)	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Food and/or Water Modification	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Surgery: Non Survival	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Surgery: Survival	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Surgery: Survival	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Surgery: Survival	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors
Surgery: Survival	1. Cellular Growth and Development: the role of E2F's and tumor suppressors 2. Cancer and Inflammation: the role of E2F's and tumor suppressors

## Animal Use Locations- Location 2

**Building:** Biomedical Research Tower

**Room:** 855

If occurring within a ULAR vivarium specify "ULAR" for the room.

**Select the protocol groups to be used at the location selected above.**

1. Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)
2. Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)

**Select the activities to be performed at the location selected above.** Be certain that the groups selected in the previous question are included in the activities you are selecting here.

Activity	Protocol Groups
Agent Administration (Not Part of Any Other Activity Listed)	<ol style="list-style-type: none"><li>1. Cellular Growth and Development: the role of E2F's and tumor suppressors</li><li>2. Cancer and Inflammation: the role of E2F's and tumor suppressors</li></ol>
Surgery: Non Survival	<ol style="list-style-type: none"><li>1. Cellular Growth and Development: the role of E2F's and tumor suppressors</li><li>2. Cancer and Inflammation: the role of E2F's and tumor suppressors</li></ol>



## Justification for Number of Animals Requested

### Breeding Animals Requested

**Add only those animals involved in the generation of offspring.** This number should be the number that will be used during the three year life of the protocol.

Protocol Group(s)/Species	Male	Female	Number of Expected Offspring	Number of Offspring to Be Used	Total
Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)	1295	1295	37040	8920	39630
Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)	2490	2490	96280	18960	101260

**Total Breeding Animals Requested: 140890**

**Briefly list/describe each experimental group identified in the table above and justify the number of animals per group.** If all groups have the same number of animals, this should be noted and the justification need only be done once.

- To justify the number of animals per group, provide either a reference to statistical assessment (such as a power analysis) or a justification based on other relevant criteria, such as current standards in the literature or specialized experimental requirements.
- Any anticipated animal losses should be clearly described.
- The number of animals requested should be the minimum required for statistically valid results.
- The presentation of the requested animal numbers in a summary table or flow chart may be beneficial (add this below).

Breeding activities will generate mice for most experiments. Some mice will be purchased for breeding stock, experimental purposes and control groups.

The number of actual crosses required for the experiments will be determined by genotyping, testing of tissue from sacrificed animals and statistical analysis. In order to obtain the number of offspring for testing that are the appropriate sex and genotype, the number of breeding pairs must be sufficient to obtain the desired strain.

GEM breeding scheme example: 5 breeding pairs x 5 subgroups x 10 mice/subgroup x 2 litters per pair x 8 pups/litter = 4,000 offspring. However, breeding will be optimized to produce as few unneeded offspring as possible.

All together we will require 140890 mice over the next 3 years. The main reason for the large number of animals needed for our studies is that in all the projects, the experimental animals contain at least two transgenes, but in most cases harbor 4 or more of them. To incorporate such a large number of transgenes within one animal, a significant amount of breeding is needed to obtain the desired collection of genes. There is also the possibility that some combinations (particularly mice possessing homozygous transgenic alleles) may have deleterious developmental consequences (leading to infertility or death etc.), permitting only heterozygous crosses.

Attach is an Excel spreadsheet which describes the GEM breeding needs. The final tab "Sheet2" contain the summary of the various GEMs. The above numbers will generate the GEM (genetically engineered mice) that will then be used for the studies as outlined in the timeline groups. Specifically, a subset of all the GEM strains will undergo hepatectomy, mammary transplantation, wound healing, and fibrosis. Based on our published literature, a group size of 6 - 20 mice (depending upon genetic background) is required for the specified surgical interventions on each genetic background to be studied for its effect on inflammation; tumor cellular development/pathogenesis in conjunction with the surgical intervention. Based on this, 20 mice/group x 4 surgical interventions(wound healing, hepatectomy, fibrosis, mammary transplantation) x 40 GEM strains = 3,200 mice. The remaining mice will be followed for the development of spontaneous tumors/ tissue harvesting due to the mutations. Estimates based on approximately 308 various combinations of oncogenes, tumor suppressors and E2F transcription factors with the various tissue specific deleter -cre transgenes at 4 time points and 20 mice per group, gives us the remainder of the total mice requested, 24640 animals. Please note that the GEM breeding numbers for the various genotypes in the attached excel spreadsheet will vary. For instance, those GEM phenotypes which have more global expression will be used for all the projects versus those which have a hepatic phenotype will be used for only those studies analyzing liver cancer pathogenesis/inflammation.

## Non-Breeding Animals Being Requested

**Add all animals NOT involved in the generation of offspring.** This number should be the number that will be used during the three year life of the protocol.

Protocol Group(s)/Species	Number of Experiments	Number of Experimental Groups	Number of Animals per Experimental Group	Number of Animals Requested
Cancer and Inflammation: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)				300
Cellular Growth and Development: the role of E2F's and tumor suppressors - Mouse, Standard (Mus Musculus)				100
Total Number of Animals Requested:				400

## Genotyping Methods

**Indicate the genotyping method, weaning age, and animal identification method to be employed for each animal type.** For further information, please see the IACUC guidelines on Methods of Animal Identification, Genotyping of Rodents, and Mouse Breeding Cages.

Genotyping Method	Weaning Age	Method of ID	Protocol Groups
Tail Snips <21 days of age	28	Ear Punch/Ear Notch	Cellular Growth and Development: the role of E2F's and tumor suppressors Cancer and Inflammation: the role of E2F's and tumor suppressors
Tail Snips >21 days of age [Requires General Anesthesia]	28	Ear Punch/Ear Notch	Cellular Growth and Development: the role of E2F's and tumor suppressors Cancer and Inflammation: the role of E2F's and tumor suppressors