



## Data in Brief

# Transcriptome analysis of basal and luminal tumor-initiating cells in ErbB2-driven breast cancer



Nicholas Borcharding<sup>a,b,c,f</sup>, Nicholas Bormann<sup>f</sup>, David Kusner<sup>a,d</sup>, Ryan Kolb<sup>a,e</sup>, Weizhou Zhang<sup>a,b,c,d,e,f,\*</sup>

<sup>a</sup> Department of Pathology, United States

<sup>b</sup> Medical Science Training Program, United States

<sup>c</sup> Holden Comprehensive Cancer Center, United States

<sup>d</sup> Molecular and Cellular Biology Program, United States

<sup>e</sup> Immunology Program, United States

<sup>f</sup> Carver College of Medicine, University of Iowa, Iowa City, IA 52242-1109, United States

## ARTICLE INFO

## Article history:

Received 31 March 2015

Accepted 2 April 2015

Available online 17 April 2015

## Keywords:

Breast cancer

Tumor-initiating cells

Mammary stem cells

Luminal progenitors

Myoepithelial cells

## ABSTRACT

Breast cancer is the leading cause of cancer-related mortality for females worldwide [1]. Improving early screening strategies and understanding the events that lead to tumor initiation have led to demonstrable improvements in clinical outcome. Our previous work revealed a variance in the tumorigenic capacity between different mammary epithelial cell populations in an *MMTV-ErbB2* mouse model. In order to greater understand how different mammary epithelial cells influence the tumorigenic capacity in ErbB2-induced breast cancer, we transplanted different cell populations from pre-neoplastic *MMTV-ErbB2* female mice into recipient mice for tumorigenic study. We found that different mammary epithelial cells bear different tumorigenic potentials even when induced by the same ErbB2 proto-oncogene. To understand the difference in tumors formed from different epithelial cells, we performed gene expression profiling using these tumors (GSE64487). Several genes were further validated using real-time reverse transcription polymerase chain reaction (RT-PCR). Here we provide further details on the experimental methods and microarray analysis. This data provides a resource to further understanding how different mammary cell populations can initiate ErbB2-driven tumors and the role of these cell populations as putative tumor-initiating cells (TICs).

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## Specifications

Organism/cell line/tissue	<i>Mus musculus</i>
Sex	Female
Sequencer or array type	Illumina Infinium MouseWG-6 v2.0
Data format	Raw: Available in TXT file Processed: Available in series_matrix.TXT
Experimental factors	Mammary tumors derived from basal and luminal mammary compartments.
Experimental features	Gene expression profiling in <i>MMTV-ErbB2</i> tumors derived from wild-type (WT) luminal progenitors (LP), myoepithelial cells (Myo), and mammary stem cells (MSC). Additional expression profiling for tumors derived from LP from <i>MMTV-ErbB2/lkko<sup>AA/AA</sup></i> (AA). The expression levels of selected genes were confirmed using RT-PCR.
Consent	N/A
Sample source location	Department of Pathology, University of Iowa, Iowa City, IA

## Direct link to deposited data

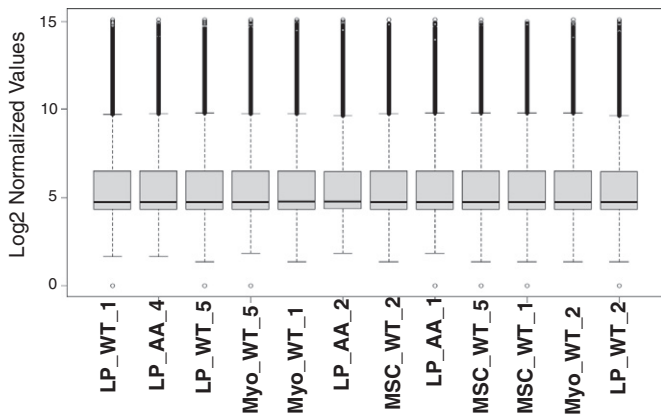
Microarray: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE64487>

## Experimental design, materials and methods

## Tumor samples

Tumor samples were collected from paired primary cell grafts from our previous work [2]. Briefly,  $5 \times 10^4$  sorted mammary epithelial cells from *MMTV-ErbB2* background were orthotopically transplanted into the #2 mammary fat pad of *Rag1<sup>-/-</sup>* female mice in a pairwise fashion. The sorting scheme for individual cell populations consisted of epithelial enrichment using Mammary Epithelial Cell Enrichment Kit (Stemcell, Vancouver, BC, Canada), and magnetic purification against CD45-, CD31-, and Ter119-positive cells. Luminal and basal cell populations were isolated using fluorescence activated cell sorting (FACS) in the following scheme: CD24<sup>med</sup>CD49<sup>phi</sup> mammary stem cells (MSCs), CD24<sup>med</sup>CD49<sup>lo</sup> mature myoepithelial

\* Corresponding author at: Department of Pathology, University of Iowa/Carver College of Medicine, 200 Hawkins Drive, ML 1020, Iowa City, IA 52242, United States.  
E-mail address: [weizhou-zhang@uiowa.edu](mailto:weizhou-zhang@uiowa.edu) (W. Zhang).

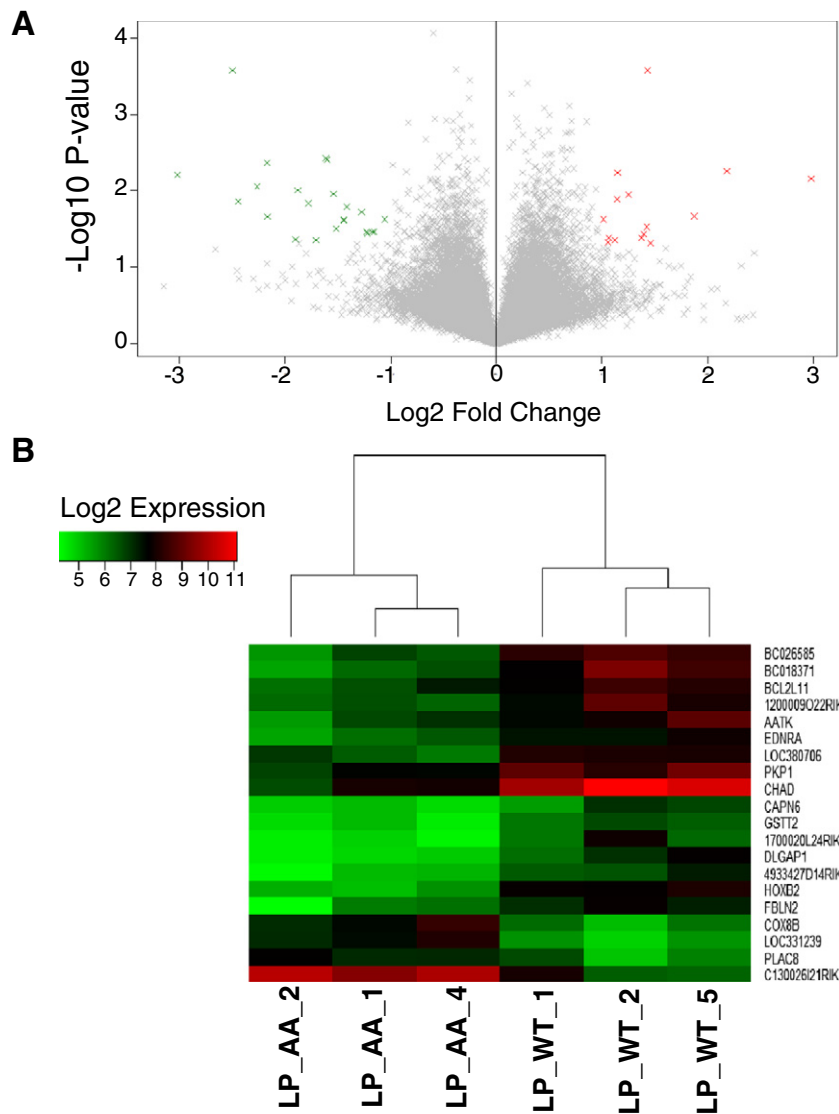


**Fig. 1.** Box plot of the quantile normalized expression level for the 12 microarrays. The line bisecting the boxplot is the mean probe value. Samples appear in the order of series matrix file of GSE64487 dataset, the order of the original blinding of the RNA experiment.

cells (Myo), and CD24<sup>hi</sup>CD49f<sup>lo</sup>CD61<sup>+</sup> luminal progenitors (LP). Tumor growth was monitored over 3 months and tumor samples were harvested before RNA extraction. Sample IDs in GSE64487 correspond to Tumor\_Cell Type\_Allele\_Mouse, i.e. Tumor\_LP\_WT\_1. The final number on the ID corresponds to paired tumors derived from the same mouse.

#### RNA preparation

Total RNA was isolated from 20 mg of tumor samples that were homogenized into RLT buffer (QIAGEN, Venlo, Limburg, Netherlands). RNA was isolated using the RNeasy Plus mini kit (QIAGEN), according to the manufacturer's instruction. RNA extracts were assessed for quality by Agilent 2100 Bioanalyzer, samples with A260/280 ( $2.0 \pm 0.1$ ), A260/230 ( $2.0 \pm 0.1$ ), and RNA integrity number (RIN)  $\geq 8.7$  were used for further experimentation [3]. Twelve tumor samples were selected: 3 WT LP, 3 WT MSC, 3 WT Myo, and 3 *MTV-ErbB2/Ikka*<sup>AA/AA</sup> (AA) LP.



**Fig. 2.** A. Volcano plot displaying the gene expression fold change and *P*-value for LP WT versus LP *Ikka*<sup>AA/AA</sup> (AA). Genes with fold-change  $\geq 1$  and *P*-value  $\leq 0.05$  (15 genes) are highlighted in red and fold-change  $\leq -1$  and *P*-value  $\leq 0.05$  (23 genes) in green. B. Heatmap of the 20 most differentially expressed genes, as measured by absolute fold-change difference. Samples are labeled according to the GSE64487 dataset.

Gene expression microarray

A total of 100 ng of RNA for each sample was submitted to the Iowa Institute of Human Genetics Genomics Division for RNA sample preparation (cDNA synthesis and in vitro transcription). The Genomics Division also performed the subsequent hybridization onto the Illumina Mouse WG-6 v2.0 for 17 h at 58 °C. Chips were stained with streptavidin-Cy3 (GE Healthcare, Piscataway, NJ) and scanned. More detailed methodology of the Iowa Institute of Human Genetics Genomics Division procedure for Illumina Mouse WG-6 v2.0 for this microarray profiling has been described previously in supplemental methods [4].

Microarray analysis

Beadchips were scanned with the Illumina iScan System and data was collected using GenomeStudio software v2011.1. Microarray data was quantile normalized and transformed into log<sub>2</sub> expression by the Iowa Institute of Human Genetics Bioinformatics Division (Fig. 1).

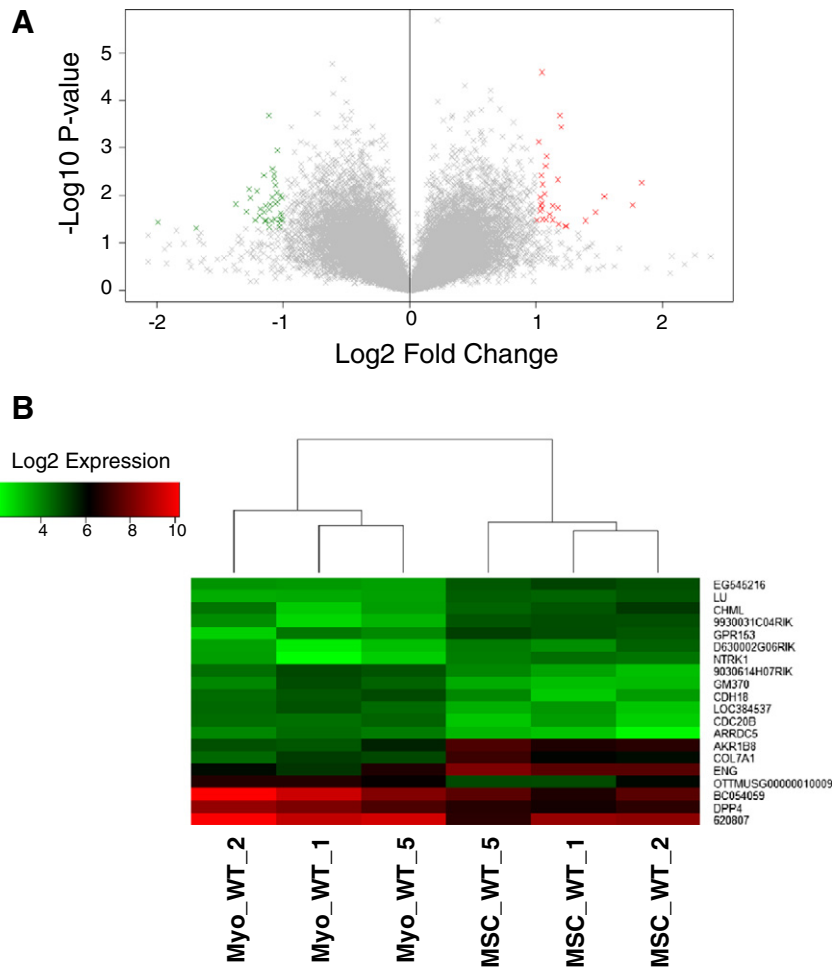
A total of 33,622 coding transcripts were analyzed using the quantile-normalized values. Fold change was calculated by the average log<sub>2</sub> expression differences in the indicated group comparisons. The volcano plots were generated in R with the use of Student's *T* test and generic plot function. Genes highlighted in the volcano plots had a *P*-value < 0.05 and an absolute value of the average log<sub>2</sub> fold change > 1 (red ≥ 1 log<sub>2</sub> fold change and green ≤ -1). The 20 genes

with the greatest fold change between the groups were further displayed in heatmaps, generated in R using the gplots package [5].

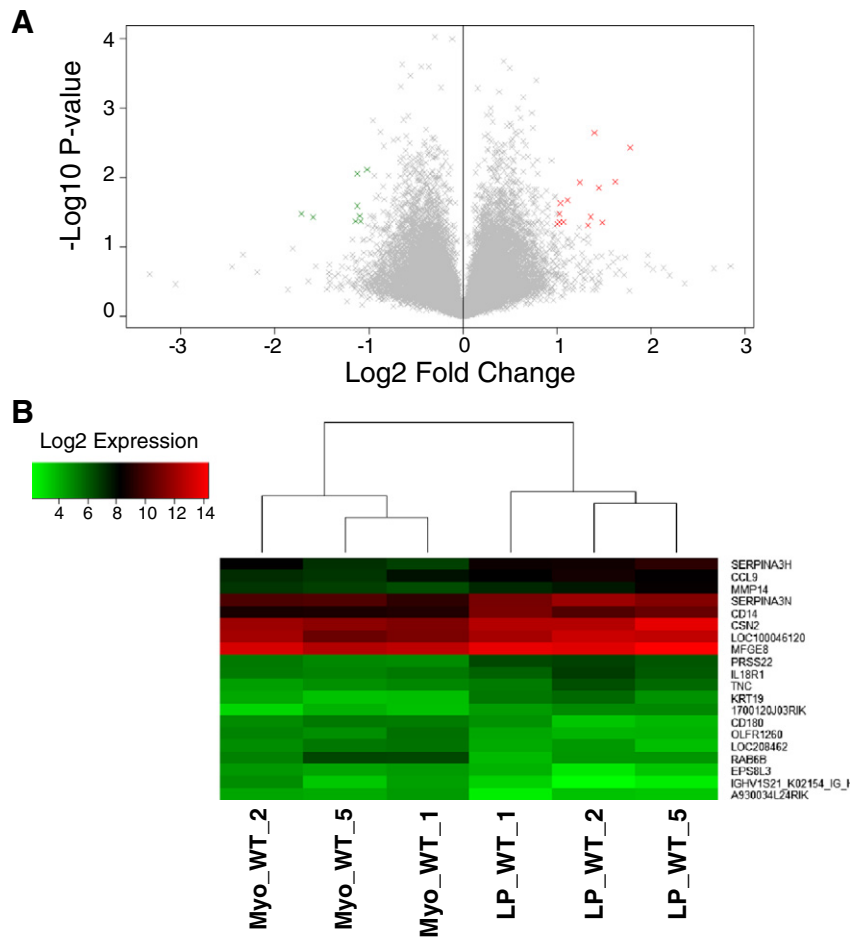
Expression values were compared based on the tumor-initiating cells (TICs) that gave rise to the tumors. Previously, we published the comparison between LP- and MSC-derived tumors, finding that the non-canonical Wnt5a was downregulated in MSC-derived tumors [4]. We also conducted comparisons between the tumors from LP WT versus LP AA to investigate the transcriptional differences with the inactivation of IKK-α (CHUK) in luminal progenitors (Fig. 2A, B). We continued our transcriptional comparisons with the analysis of tumors derived from Myo and MSC, two cell types of the basal compartment that gave rise to ErbB2-driven tumors (Fig. 3A, B). Finally, as another comparison between tumors derived from cells of the basal versus luminal compartments, we compared the transcriptional profiles of the Myo- and LP-derived tumors (Fig. 4A, B).

RT-PCR validation

For each isolated RNA sample, 1 µg was reverse-transcribed with iScript cDNA synthesis Kit (Bio-Rad, Hercules, CA). cDNA samples were diluted (1:12) in nuclease-free water before RT-PCR was performed using 96-Well Optical Reaction Plates (Applied Biosystems, Life Technologies, Carlsbad, CA) and iTaq Universal SYBER Green Supermix (Bio-Rad). A total of 15 µL of total reaction volume was used in a ViiA 7 RT-PCR system (Applied Biosystems). Relative gene expression was quantified using the 2<sup>-ΔΔCT</sup> method [6], using *Cyclophilin A*



**Fig. 3.** A. Volcano plot displaying the gene expression fold change and *P*-value for Myo WT versus MSC WT. Genes with fold-change ≥ 1 and *P*-value ≤ 0.05 (30 genes) are highlighted in red and fold-change ≤ -1 and *P*-value ≤ 0.05 (39 genes) in green. B. Heatmap of the 20 most differentially expressed genes, as measured by absolute fold-change difference. Samples are labeled according to the GSE64487 dataset.



**Fig. 4.** A. Volcano plot displaying the gene expression fold change and *P*-value for Myo WT versus LP WT. Genes with fold-change ≥ 1 and *P*-value ≤ 0.05 (14 genes) are highlighted in red and fold-change ≤ -1 and *P*-value ≤ 0.05 (8 genes) in green. B. Heatmap of the 20 most differentially expressed genes, as measured by absolute fold-change difference. Samples are labeled according to the GSE64487 dataset.

(*Ppia*) (forward 5'-CAGTGCTCAGAGCTCGAAAGT-3' and reverse 5'-GTGTTCTCGACATCACGC-3') as a housekeeping gene. A total of 24 genes were examined based on the differential expression of LP- versus MSC-derived tumors. Results and primer sequences can be found in previously published supplemental material [4].

## Discussion

Herein we describe transcriptional profiling of breast tumors derived from LP, MSC and Myo cell populations in the *MMTV-ErbB2* background. We believe that this dataset could provide insights into the characteristics of ErbB2-driven tumors derived from basal and luminal tumor-initiating cells, as both compartments are able to generate tumors [2]. As *MMTV-ErbB2*-driven tumors are a murine model of the aggressive HER2+ molecular subtype of breast cancer, we believe that this data may assist in further elucidation of the divergence in the 2 clinically defined subclasses of HER2+ breast tumors: HER2-enriched mRNA subclass and luminal-mRNA/HER2+ subclass [7].

## Disclosures

All authors possess no conflicts of interests.

## Acknowledgments

We would like to thank Dr. Kevin Knudtson and Dr. Tom Bair of the Iowa Institute of Human Genetics for their help and insight into

the microarray workflow. This work was supported by NIH grant K99/R00 CA158055 (W.Z.), NIH T32 GM007337 (N.B.), NIH T32 AI007260 (R.K.), a V Scholar award from the V Research Foundation for the Cancer (W.Z.). Additional support for this work was received from the Department Startup Grant and Seed Grant from the Department of Pathology (W.Z.). Lastly, this work benefited from funding through a Breast Cancer Research Grant and an ACS Seed Grant from Holden Comprehensive Cancer Center, University of Iowa Carver College of Medicine (W.Z.).

## References

- [1] American Cancer Society, Global Cancer Facts & Figures. 2nd ed. American Cancer Society, Atlanta, 2011.
- [2] W. Zhang, et al., A NIK-IKK $\alpha$  module expands ErbB2-induced tumor-initiating cells by stimulating nuclear export of p27/Kip1. *Cancer Cell* 23 (2013) 647–659, <http://dx.doi.org/10.1016/j.ccr.2013.03.012>.
- [3] A. Schroeder, et al., The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* 7 (2006) 3, <http://dx.doi.org/10.1186/1471-2199-7-3>.
- [4] N. Borchering, et al., Paracrine WNT5A signaling inhibits expansion of tumor-initiating cells. *Cancer Res.* (2015) <http://dx.doi.org/10.1158/0008-5472.CAN-14-2761>.
- [5] G.R. Warnes, B. Bolker, L. Bonebakker, R. Gentleman, W. Huber, A. Liaw, T. Lumley, M. Maechler, A. Magnusson, S. Moeller, M. Schwartz, B. Venables, gplots: various R programming tools for plotting data. R package version 2 (2009).
- [6] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> method. *Methods* 25 (2001) 402–408, <http://dx.doi.org/10.1006/meth.2001.1262>.
- [7] Cancer Genome Atlas Network, Comprehensive molecular portraits of human breast tumours. *Nature* 490 (2012) 61–70, <http://dx.doi.org/10.1038/nature11412>.