

Meta-Analysis of Outlier Expression Profiles in Breast Cancer Identifies AGTR1 Over-Expression as an Opportunity for Targeted Therapy

Daniel R. Rhodes^{1,6}, Scott A. Tomlins¹, Rohit Mehra^{1,5}, Qi Cao¹, Bharathi Laxman¹, Shanker Kalyana-Sundaram¹, Lei Wang¹, Beth E. Helgeson¹, Peter C. Lucas^{1,5}, Celina G. Kleer^{1,5}, Daniel F. Hayes^{1,5}, Sooryanarayana Varambally^{1,5}, and Arul M. Chinnaiyan^{1,2,4,5}
Department of Pathology¹, Bioinformatics², Internal Medicine³, Urology⁴, and the Comprehensive Cancer Center⁵, University of Michigan Medical School, 1301 Catherine Street, Ann Arbor, MI 48109-0602, USA., Compendia Bioscience⁶, 110 Miller Avenue, Floor 2, Ann Arbor, MI 48104, USA.

Abstract

To identify novel cancer targets from microarray gene expression data, we previously developed an analysis strategy that searches for genes with marked overexpression in a subset of tumors. Here, we have extended the approach to include a meta-analysis that validates outlier expression profiles across independent patient cohorts. We analyzed data from 31 breast cancer profiling studies comprising 3,100 tumor expression profiles. We identified genes that showed the most profound overexpression in subsets of tumors, validated across independent datasets. ERBB2 was the highest scoring outlier gene, consistent with the known role that amplification and over-expression of this gene play in breast tumorigenesis. We also identified many novel breast cancer outlier genes not previously identified from expression data, including AGTR1, the angiotensin II receptor type I, which showed marked over-expression in 10–20% of breast cancer cases across several independent patient cohorts. AGTR1-overexpressing tumors were almost exclusively estrogen receptor positive and ERBB2 negative by gene expression. Ectopic over-expression of AGTR1 in mammary epithelial cells combined with angiotensin II stimulation led to a highly invasive phenotype, which was attenuated by treatment with the AGTR1 blocker, losartan. Similarly, breast cancer cells with endogenous over-expression of AGTR1 became invasive upon AT stimulation, again attenuated by losartan. Lastly, we identified copy of number gain of the AGTR1 locus in a subset of AGTR1 overexpressing cases, potentially contributing to AGTR1 over-expression. In summary, our analysis identified and validated genes with profound changes in expression in subsets of tumors, nominating AGTR1 as a novel therapeutic target for a specific sub-population of breast tumors.

Methods

Meta-COPA Analysis

COPA analysis was performed on 31 breast cancer gene expression data sets in OncoPrint (www.oncoPrint.org) as described previously. Genes scoring in the top 1% of COPA scores at any of the three percentile cutoffs (75th, 90th, and 95th) were deemed outliers in their respective datasets. Meta-outliers were defined as genes deemed outliers in a significant fraction ($p < 1E-5$) of datasets as assessed by the binomial distribution.

Quantitative PCR (QPCR)

Quantitative PCR (QPCR) was performed using SYBR Green dye on an Applied Biosystems 7300 Real Time PCR system (Applied Biosystems, Foster City, CA) essentially as described.

AGTR1 Transfection

N-terminal 3XHA-tagged human AGTR1 (Variant 1) in pcDNA 3.1+ (Invitrogen) was obtained from the University of Missouri-Rolla cDNA Resource Center. 3XHA-AGTR1 was amplified by PCR and TOPO-IA cloned into the Gateway entry vector pCR8/GW/TOPO. To generate adenoviral constructs, pCR8/3XHA-AGTR1 was recombined with pAD/CMV/V5 (Invitrogen) using LR Clonase II (Invitrogen). Adenoviruses were generated by the University of Michigan Vector Core. The benign human mammary epithelial cell lines HME and H16N2 were plated in 6-well dishes. 24 h later the cells were infected with 3XHA-AGTR1 adenovirus or LacZ adenovirus. Overexpression was confirmed by QPCR. 48 h post infection, cells were treated with vehicle (ethanol), 1 μ M angiotensin alone, 2 and 5 μ M losartan alone and a combination of 2 and 5 μ M losartan with 1 μ M angiotensin for 24 h prior to the invasion assay.

Cell Invasion Assay

Breast cell lines BT-549, Hs579, H16N2, HCC1500 and prostate carcinoma line DU145 were grown in 100mm tissue culture plates overnight, then transferred to serum free medium. 1 – 2 μ M losartan (Kind gift from Merck, Whitehouse Station, NJ) was added 30 minutes prior to 1 μ M angiotensin II (American Peptide Company, Sunnyvale, CA) treatment. Cell invasion was evaluated using 24-well Matrigel invasion chambers (Becton Dickinson, Franklin Lakes, NJ). Cells were trypsinized and seeded at equal numbers onto the basement membrane matrix present in the insert of a 24 well culture plate. Fetal bovine serum was added to the lower chamber acting as a chemoattractant. After 48 hours of additional incubation, the non-invading cells and EC matrix were removed gently with a cotton swab. The cells that had invaded were present on the lower side of the chamber and were stained, air-dried and photographed. The invaded cells were counted under the microscope assessing six

Conclusions

- AGTR1 is markedly overexpressed in 10–20% of breast cancer. AGTR1 over-expressing tumors were almost exclusively estrogen receptor positive and ERBB2 negative.
- Ectopic over-expression of AGTR1 in mammary epithelial cells combined with angiotensin II stimulation led to an invasive phenotype, which was attenuated by treatment with the AGTR1-blocker, losartan.

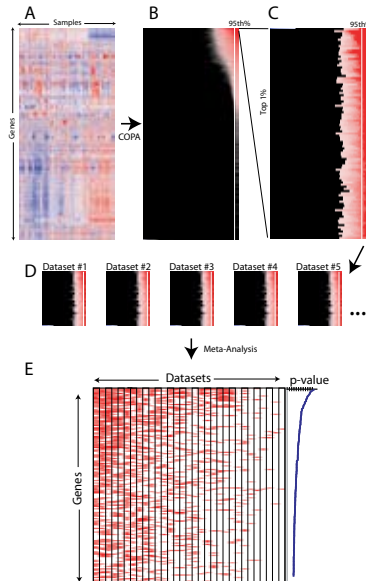


Figure 1.

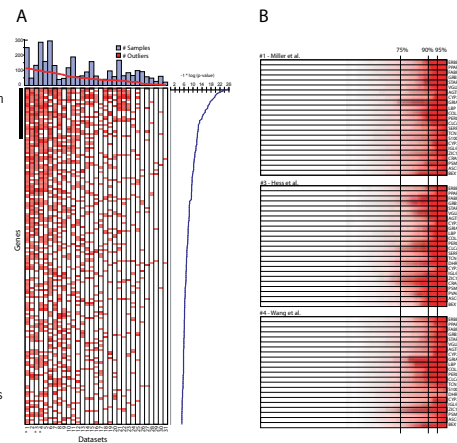
A schematic for the meta-analysis of outlier gene expression profiles (Meta-COPA).

- A cancer gene expression dataset, consisting of thousands of genes measured across tens or hundreds of samples, is normalized and sorted via the COPA method.
- COPA normalizes the median expression per gene to zero and the median absolute deviation to one. A percentile cutoff is selected (e.g., 75th%, 90th% or 95th%) and genes are sorted by their COPA value at the selected percentile. In the COPA map, the intensity of red indicates degree of over-expression and black indicates masked COPA values less than 1.
- The top 1% of genes is deemed to have outlier expression profiles.
- A collection of outliers from independent datasets are compiled and submitted for meta-analysis.
- Multiple datasets of a given cancer type are meta-analyzed to identify genes that are consistently called outliers across independent datasets. Significance is assessed via the binomial distribution.

Figure 2.

Meta-COPA analysis of outlier breast cancer gene expression data.

- Meta-COPA map. Each column in the map represents a breast cancer gene expression dataset. Each row indicates a gene. A red cell indicates that the gene was deemed to have an outlier expression profile in the respective dataset because it scored in the top 1% of COPA values at one of three percentile cutoffs. The line graph along the y-axis indicates the p-value for a gene based on the number of datasets in which the gene was deemed an outlier. 158 genes were called outliers in a significant fraction of datasets ($p < 1E-5$). The bar graph indicates the number of samples in the respective datasets and the contribution of the dataset to the meta-analysis.
- Heatmaps of COPA-normalized values for top scoring meta-outliers across three highly contributory datasets. Genes are ranked by their Meta-COPA p-values. For each gene, samples are



ordered from left to right by their COPA-normalized expression values. Bright Red indicates a COPA normalized value of 6 or greater. White indicates a value of zero or less.

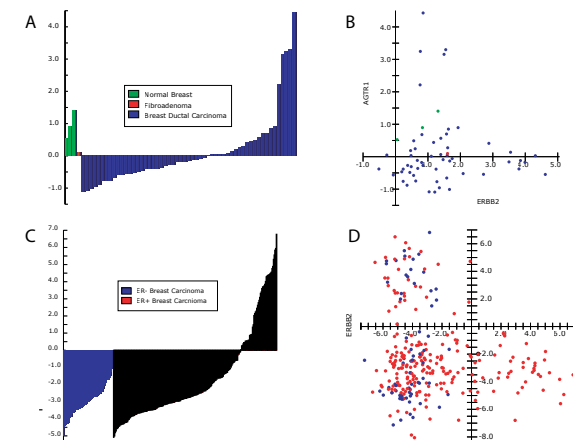


Figure 4.

AGTR1 outlier expression in breast cancer.

- AGTR1 expression profile in the Perou et al. cDNA microarray dataset (n=55).
- In the same dataset, AGTR1 expression vs. ERBB2 expression.

(c) AGTR1 expression profile in the van de Vijver et al. oligonucleotide dataset, segregated by estrogen receptor (ER) status (n=295).

(d) AGTR1 expression vs. ERBB2 expression in the same dataset.

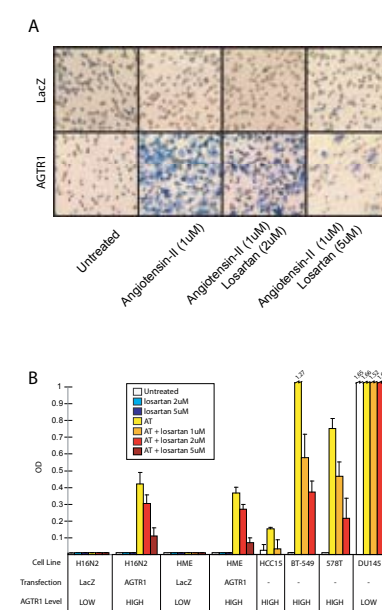


Figure 5.

AGTR1 over-expression and analysis of angiotensin II (AT) and losartan effects on invasion.

- Matrigel invasion assays of H16N2 cells infected with adenovirus expressing AGTR1 or LacZ. Cells were cultured in serum-free media and were pre-treated with and without AT and losartan. Similar results were observed for HME cells.
- Colorimetry read-out of invasion assays. The first four panels represent the adenovirus transfection experiments in which LacZ- or AGTR1-expressing adenovirus was infected into H16N2 and HME immortalized mammary epithelial cells. This experiment set included losartan alone treatments. The last four panels represent invasion assays in three breast cell lines with high endogenous AGTR1 levels and DU145 prostate cancer cells with low AGTR1 levels as measured by quantitative RT-PCR. The optical density (OD) measurements were background subtracted and values below 0.01 were set to 0.01.

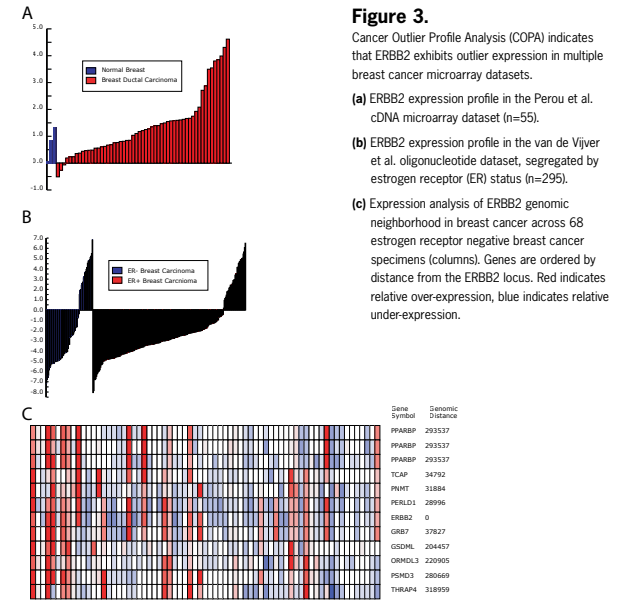


Figure 3.

Cancer Outlier Profile Analysis (COPA) indicates that ERBB2 exhibits outlier expression in multiple breast cancer microarray datasets.

- ERBB2 expression profile in the Perou et al. cDNA microarray dataset (n=55).
- ERBB2 expression profile in the van de Vijver et al. oligonucleotide dataset, segregated by estrogen receptor (ER) status (n=295).
- Expression analysis of ERBB2 genomic neighborhood in breast cancer across 68 estrogen receptor negative breast cancer specimens (columns). Genes are ordered by distance from the ERBB2 locus. Red indicates relative over-expression, blue indicates relative under-expression.

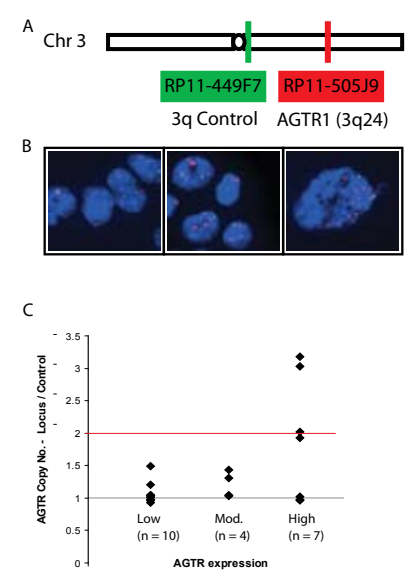


Figure 6.

Copy number analysis of AGTR1 locus.

- A schematic of probes used for FISH analysis.
- Representative image from FISH analysis. The left panel is taken from a representative negative case. The middle and right panels are images from a representative case with definitive copy number gain of AGTR1. Green signal is the AGTR1 locus probe and red signal is the probe near the chromosome 3 centromere.
- Association of AGTR1 over-expression with copy number gain. Three expression bins were defined based on AGTR1 / GAPDH ratios: low (< 1.0), moderate ($1.0 < 2.0$) and high (> 2.0).