

# Network Analysis Identifies Regulators of Basal-Like Breast Cancer Reprogramming and Endocrine Therapy Vulnerability

Sea R. Choi, Chae Young Hwang, Jonghoon Lee, and Kwang-Hyun Cho



## ABSTRACT

Basal-like breast cancer is the most aggressive breast cancer subtype with the worst prognosis. Despite its high recurrence rate, chemotherapy is the only treatment for basal-like breast cancer, which lacks expression of hormone receptors. In contrast, luminal A tumors express ER $\alpha$  and can undergo endocrine therapy for treatment. Previous studies have tried to develop effective treatments for basal-like patients using various therapeutics but failed due to the complex and dynamic nature of the disease. In this study, we performed a transcriptomic analysis of patients with breast cancer to construct a simplified but essential molecular regulatory network model. Network control analysis identified potential targets and elucidated the underlying mechanisms of reprogramming basal-like cancer cells into luminal A cells. Inhibition of *BCL11A*

and HDAC1/2 effectively drove basal-like cells to transition to luminal A cells and increased ER $\alpha$  expression, leading to increased tamoxifen sensitivity. High expression of *BCL11A* and *HDAC1/2* correlated with poor prognosis in patients with breast cancer. These findings identify mechanisms regulating breast cancer phenotypes and suggest the potential to reprogram basal-like breast cancer cells to enhance their targetability.

**Significance:** A network model enables investigation of mechanisms regulating the basal-to-luminal transition in breast cancer, identifying *BCL11A* and *HDAC1/2* as optimal targets that can induce basal-like breast cancer reprogramming and endocrine therapy sensitivity.

## Introduction

Breast cancer has the highest cancer-related mortality rate among women worldwide (1). It has been categorized into various subtypes, including basal-like (or triple-negative), luminal A, luminal B, and HER2-enriched (2–4). Among these subtypes, luminal A is the least aggressive breast cancer (5). Almost 70% of patients with breast cancer are diagnosed as luminal A and have normal-like features such as expressing estrogen receptor alpha (ER $\alpha$ ; ref 6) by which they are probable in undergoing ER $\alpha$ -targeted therapy as their first line of treatment (7, 8). In contrast, basal-like patients have the worst prognosis with a very high tumor recurrence rate despite its low prevalence ratio of 15% (9). They have highly elevated expression of the epidermal growth factor receptor (EGFR; ref. 10), which is an important characteristic feature of the basal-like subtype. Unfortunately, chemotherapy is the only available treatment option for basal-like patients because they are deficient in hormone receptors, including ER $\alpha$ , and clinical trials of all single anti-EGFR therapy have failed (11). Researchers have speculated that this might be due to interactions between the EGFR signaling pathway and epithelial–mesenchymal transition (EMT) regulators, including SLUG and ZEB1 (12), that are more likely to develop distant metastasis. As a result, several studies

showed that blocking EMT regulators can abolish mesenchymal features of the basal-like subtype (13), or induce the expression of ER $\alpha$  (14), which allows therapeutic responsiveness to anti-ER $\alpha$  drugs. Nevertheless, there still remains a limitation in using conventional ER $\alpha$ -targeted therapies for basal-like patients because the increased expression of ER $\alpha$  is not a sufficient measure to consider these basal-like cells being reprogrammed into luminal A; there are complex interactions and feedback between genes and molecules that retain basal-like characteristics even after the expression of ER $\alpha$  is induced. To overcome this limitation, we investigated the regulatory mechanisms and underlying dynamics of differentially expressed genes (DEG) that are associated with the corresponding signaling cascades, which can possibly determine the responsiveness of anti-ER $\alpha$  therapy at a system level.

A cancer cell can be represented by a dynamical network model composed of various genes, proteins, and small molecules as nodes and their interactions as corresponding edges connecting them (15). A well-defined and validated logical network model can capture differential behaviors of cellular systems to unveil their hidden mechanisms and predict cellular responses when certain perturbations are given (16–18). In this study, we have constructed a molecular regulatory network model that can represent the biological conditions of basal-like and luminal A cells. We then applied a complex network control strategy to identify potential targets that can induce a subtype transition from the most invasive basal-like breast cancer subtype to the terminally differentiated luminal A subtype, which is termed basal-to-luminal A transition (BLT). We also elucidated the underlying mechanisms of BLT through analyzing the logical relationships between the key molecules in our network model. We identified B-cell lymphoma/leukemia 11A (*BCL11A*) and histone deacetylase 1/2 (*HDAC1/2*) in combination as novel targets that can reprogram extremely aggressive basal-like cells into luminal A cells by inducing the activity of luminal A phenotypic markers while reducing the activity of basal-like phenotypic markers. We further validated our simulation results by performing *in vitro* experiments and analyzing

Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea.

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

**Corresponding Author:** Kwang-Hyun Cho, Department of Bio and Brain Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea (South). Phone: 82-42-350-4325; Fax: 82-42-350-4310; E-mail: ckh@kaist.ac.kr

Cancer Res 2022;82:320–33

doi: 10.1158/0008-5472.CAN-21-0621

©2021 American Association for Cancer Research

## BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer

clinical data. Together, our study provides new insights into the subtype transition between breast cancer cells and a development of new treatment for patients with the aggressive breast cancer subtype.

## Materials and Methods

### Boolean network model construction for the BLT using DEG analysis

The Cancer Genome Atlas (TCGA) Breast Invasive Carcinoma (BRCA) RNA sequencing (RNA-seq) expression data were downloaded as Upper Quartile normalized fragments per kilobase of transcript per million mapped reads (FPKM-UQ) values as well as their genomic data from the Genomic Data Commons (GDC) data portal (<https://portal.gdc.cancer.gov/>). All genomic data as well as RNA-seq expression data from the Cancer Cell Lines Encyclopedia (CCLE; ref. 19) breast cancer cell lines were downloaded as reads per kilobase of transcript per million mapped reads (RPKM) values. These were classified according to subtype classification results from PAM50 (20). We then performed DEG analysis of the basal-like and luminal A subtypes of TCGA BRCA RNA-seq expression data using the DEGseq R package (1.38.0).

### Input-output relationships of the BLT model through *in silico* simulation

Boolean network is a logical dynamic model with binary node states, active (ON) as 1 or inactive (OFF) as 0, that does not require detailed kinetic parameters (21, 22). One can adopt a logical Boolean network model to describe a network where each state of nodes is based on their regulatory logical rules by Boolean operators, “OR,” “NOT,” and “AND” (23). We found that every initial state of our network model reaches a stable state within 500-time steps. Hence, we updated our Boolean logical functions for 1,000 times to analyze the activities of network components according to various input settings. To represent the input frequency of each simulation, the input node is set on a cycle that capitulates a desired percentage of ON state. For example, if the input node is set to be 50% ON, it would be placed on a cycle of “01010101” or “10101010” that would remain constant throughout each simulation.

### Attractor landscape and perturbation analysis using the BLT Boolean network model

Our BLT network model has  $2^{28}$  or 268,435,456 possible states that constitute the entire state space of the network with 28 nodes excluding input nodes. We randomly sampled  $2^{14}$  initial states, which was a sufficient number to cover the major attractors, that ultimately converge to either point or cyclic attractor with a basal-like or luminal A phenotype.

We calculate a phenotypic score  $P_A$  to define the phenotype for each attractor  $A$  in the BLT network model as follows:

$$P_A = \frac{N_L}{\text{Total \# of luminal - A marker nodes}} - \frac{N_B}{\text{Total \# of basal - like marker nodes}}$$

where  $N_L$  is defined as luminal A marker nodes that are ON and  $N_B$  as basal-like marker nodes that are ON for each attractor. The larger the  $P_A$  as 1, the more luminal A phenotype an attractor has. The lower the  $P_A$  as  $-1$ , the more basal-like phenotype an attractor has:

$$\begin{aligned} 0 \leq P_A \leq 1: & \text{ luminal A phenotype} \\ -1 \leq P_A < 0: & \text{ basal-like phenotype} \end{aligned}$$

For cyclic attractor, we first define the phenotype of existing states within a cyclic attractor by using  $P_A$ , and then determine the phenotype as follows:

$$\begin{aligned} \text{If } C_L \geq C_B, & \text{ then a cyclic attractor has a luminal A phenotype,} \\ \text{If } C_L < C_B, & \text{ then a cyclic attractor has a basal-like phenotype} \end{aligned}$$

where  $C_L$  is a number of luminal A attractors within a cyclic attractor and  $C_B$  is a number of basal-like attractors within a cyclic attractor. We perform attractor analysis using the BoolNet R package (2.1.9).

### Clinicopathologic analysis

To compare overall survival of patients with breast cancer with basal-like and/or luminal A patient groups, the Kaplan—Meier analysis and log-rank test were used. We downloaded clinical information of TCGA and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) patients with breast cancer from cBioPortal (<http://www.cbioportal.org/>; ref. 24). We first classified the patients into each patient group and then found upper, intermediate, and lower quartiles for each of *BCL11A*, *HDAC1*, and *HDAC2* genes according to their  $Z$ -score normalized mRNA expressions per group. For this, we labeled and compared between high (upper quartile; above the 80th percentile) and low (lower quartile; below the 20th percentile) expressions of these genes across the patient samples in each group. The R package “survival” was used to perform overall survival analysis (25).

### Gene set variation analysis with a list of gene sets

We downloaded the gene expression data of 356 patients with breast cancer with tamoxifen treatment, resistant ( $n = 203$ ) and sensitive ( $n = 153$ ), from the Gene Expression Omnibus (GEO) database (GSE158309). Considering our *in vitro* experimental conditions, we categorized the samples into four groups, BCL11A & HDAC1/2 high, BCL11A low, HDAC1/2 low, and BCL11A & HDAC1/2 low, based on the expression levels of our target genes. For this, the expression data of each gene was  $Z$ -score normalized and labeled as H (upper quartile), L (lower quartile), or I (middle) across the patient samples (lower quartile:  $\leq -0.75$ , middle:  $-0.75 < x < 0.75$ , upper quartile:  $\geq 0.75$ ). The samples were then categorized according to their label as follows:

- BCL11A & HDAC1/2 high group if two or more of these gene expressions of a sample were labeled as H,
- BCL11A low group if *BCL11A* gene expression of a sample was labeled as I or L,
- HDAC1/2 low group if *HDAC1/2* gene expressions of a sample were labeled as I or L,
- BCL11A & HDAC1/2 low group if two or more of these gene expressions of a sample were labeled as L.

We performed gene set variation analysis (GSVA) by using the R package “GSVA” (26) to assess gene set enrichment of these patient samples. A list of gene sets was downloaded from the Molecular Signature Databases (MSigDB; refs. 27, 28) and included in Supplementary Data S5. Additional information regarding any relevant data supporting the findings of this study is available from the authors upon request.

### Cell culture and reagents

Human breast cancer cell lines, BT20, MDA-MB-231, HS578T, MCF7, and T47D, obtained from the Korean Cell Line Bank were cultured in DMEM purchased from WelGENE Inc. containing 10% fetal bovine serum (FBS, WelGENE Inc.) and antibiotics (100 U/mL of penicillin, 100 µg/mL of streptomycin, and 0.25 µg/mL of fungizone)

Choi et al.

purchased from Life Technologies Corp. at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. 4-hydroxytamoxifen or tamoxifen (ER $\alpha$  inhibitor) and romidepsin (HDAC1/2 inhibitor) were purchased from Sigma-Aldrich. Estradiol (E2) and EGF were purchased from Sigma-Aldrich. shBCL11A was purchased from Sigma-Aldrich to specifically target the gene or scrambled shBCL11A-negative control. *Mycoplasma* infection was regularly checked using the e-Myc Mycoplasma PCR Detection Kit from iNtRON Biotechnology Inc.

#### Plasmid generation and virus production for overexpression experiments

We generated lentiviral particles by transfecting HEK293T cells with relevant lentiviral plasmid and packaging DNA mixture (pLP1, pLP2, and pLP/VSVG) using polyethylenimine (PEI; Invitrogen) according to the manufacturer's instructions. To perform overexpression experiments, we amplified the full-length of human *HDAC1* and *HDAC2* through PCR from the human kidney and colon cDNA library (Clontech), respectively. The full-length human *BCL11A* was purchased from Korea Human Gene Bank. The corresponding amplified gene or genes in combination were ligated into the pLentiM1.4 lentiviral vector. We then confirmed that ligation is successfully done by sequencing.

#### RNA extraction and real-time PCR analysis

RNA was extracted from cells using RNA-spin RNA Mini Kit purchased from iNtRON Biotechnology Inc. according to the manufacturer's instructions and treated with RNase-free DNase I (Thermo Fisher Scientific Inc.) to remove genomic DNA. Complementary DNA (cDNA) was synthesized by reverse transcription (RT) using a DiaStar RT kit purchased from Solgent. RT-PCR was performed using the PCR system (Veriti 96-well Thermal Cycler purchased from Applied Biosystems), and sequences of primers are listed in Supplementary File S1. All primers were purchased from Neoprobe. qRT-PCR analysis was performed using the QuantStudio 5 real-time PCR system (Applied Biosystems) with the primers.

#### Crystal violet assay

Cells were seeded at  $1 \times 10^4$  cells/well in a 96-well plate or  $1 \times 10^5$  cells/well in a 24-well plate for 24 hours. Cells were then treated with the indicated drugs and doses, and incubated for 72 hours. Cells were washed with phosphate-buffered saline (PBS) and stained with 0.5% (w/v) crystal violet from Sigma-Aldrich for 30 minutes at room temperature. Finally, plates were washed with tap water, air-dried, and photographed.

#### Western blot analysis

Cells were washed in PBS and lysed in lysis buffer (20 mmol/L of HEPES, 150 mmol/L of NaCl, 0.5% Triton X-100, 10% glycerol, 1  $\mu$ g/mL of aprotinin, 1  $\mu$ g of leupeptin, 1 mmol/L of Na<sub>3</sub>VO<sub>4</sub>, and 1 mmol/L of NaF). For immunoblotting, anti-phospho-ERK1/2, anti-ERK1/2, anti-phospho-EGFR, and anti-EGFR purchased from Cell Signaling Technology Inc., anti- $\alpha$ -actinin and anti-ER $\alpha$  purchased from Santa Cruz Biotechnology Inc. were used. For quantifying intensity of the protein bands, ImageJ software was used (<http://imagej.nih.gov/ij>) and normalized by GAPDH.

#### Hormone dependency and cell viability assay using automated live-cell imaging instrument and analysis

Breast cancer cells were seeded into a 96-well plate at a density of  $1 \times 10^4$  cells/well, and cultured in phenol red-free and serum-free medium overnight. The cells were then exposed to medium containing 10% FBS

plus or minus growth factors, 0.1 nmol/L E2 or 100 ng/mL EGF, or DMSO for control. Then, the cells were collected for RNA/protein extraction after 24 hours, or cell growth was measured at each time point using IncuCyte ZOOM (Essen Biosciences) for 5 days. Images using the software were captured every 3-hour interval from three different regions per well with a 20 $\times$  objective. For testing tamoxifen sensitivity, cells were seeded into a 96-well plate at a density of  $1 \times 10^4$  cells/well, cultured for 24 hours, and then treated with the indicated concentrations of tamoxifen and/or romidepsin for 72 hours. To analyze cell confluency, cells were measured using IncuCyte ZOOM to assess cell viability rates according to the manufacturer's instructions.

## Results

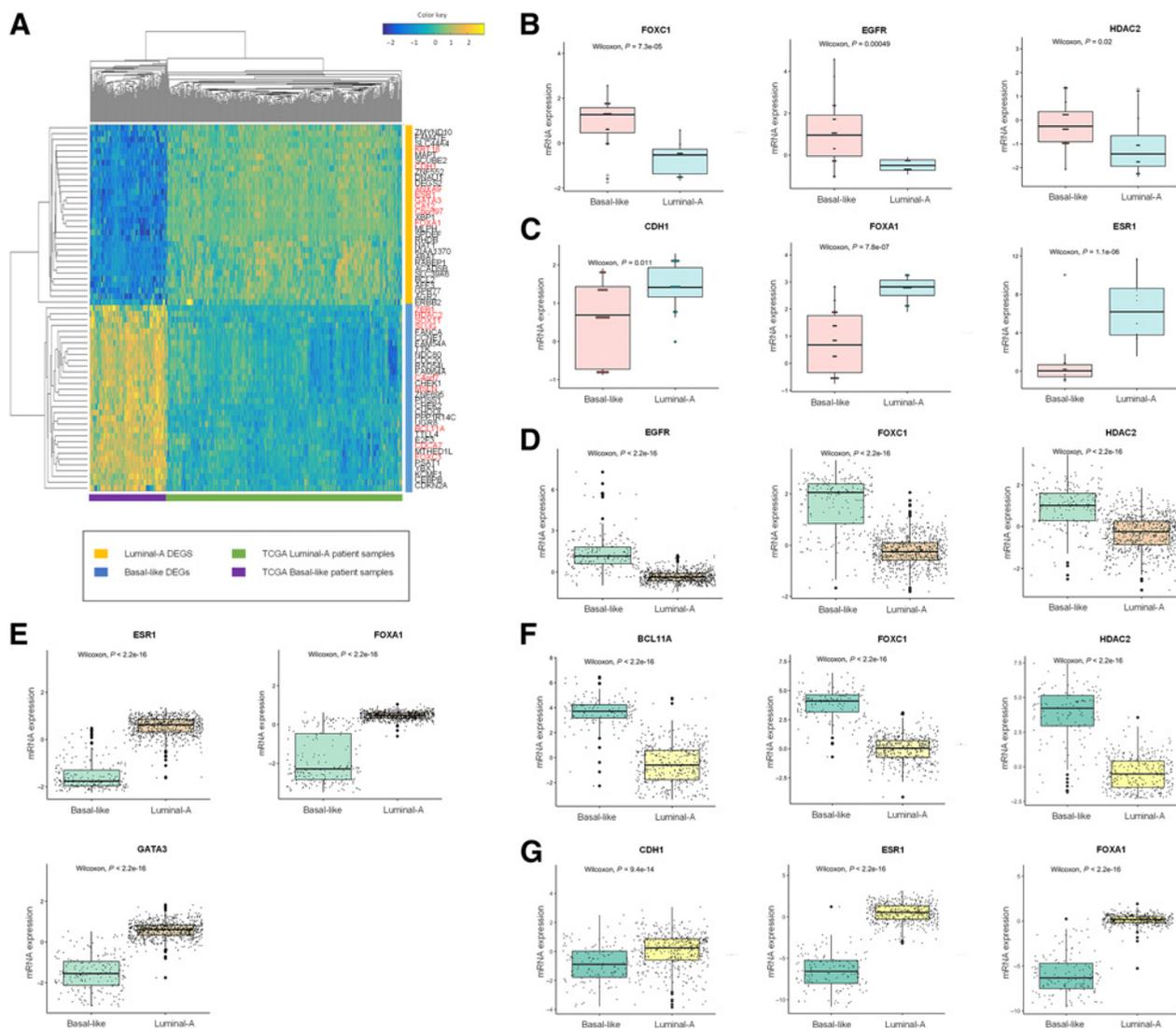
### A mathematical model of the BLT

We summarized the workflow of this study in Supplementary Fig. S1. There are various regulatory feedback and interactions intertwined in cancer cells that the mechanism of driving the BLT may not be easily predictable. To systemically analyze this mechanism, we have constructed a mathematical model that can explain the dynamics between basal-like and luminal A breast cancer cells during the BLT. For this purpose, we have reconstructed our network model based on two major pathways, EGFR and ER $\alpha$  signaling, that govern cell growth, survival, and tumorigenesis in these cell types (29, 30) by using public databases such as Kyoto Encyclopedia of Genes and Genomes (31). We also used the activity of these receptors in the pathways, ER $\alpha$  and EGFR, as a major phenotypic marker node for determining a cellular state for the luminal A and basal-like subtype, respectively. We further identified and integrated DEGs for the corresponding subtypes as additional phenotypic markers in our network model (Supplementary Fig. S2A; Supplementary File S2). For this, we compared the mRNA expression profiles of TCGA basal-like and luminal A breast cancer patients, and we listed the top 10% ranked DEGs (Fig. 1A) that are highly expressed in one of the subtypes but lowly expressed from another for each subtype. The selected nodes and their associated links in our network model are literature-based and supported by experimental evidence obtained from public databases (Supplementary File S3). We also confirmed the expression patterns of these selected DEGs from the corresponding TCGA patient samples by using the CCLE and METABRIC (32) cohorts (Fig. 1B–G; Supplementary Fig. S2B–S2G). As a result, we have constructed a simplified but essential regulatory network model for the BLT, which consists of 30 nodes and 73 links that can present molecular interactions between the ER $\alpha$  and EGFR signaling pathways as well as DEGs of the corresponding subtypes (Fig. 2A).

### Network modification specific to basal-like cell lines

Although cancer cell lines were originated from the same tissue, they have various differences depending on their genomic information, including copy number alteration (CNA), mRNA expression, and somatic mutation. Prior to computational simulation, we have modified our BLT network model to cell line-specific networks by analyzing genomic data of 21 basal-like breast cancer cell lines from the CCLE cohort. We determined the functional results of each genomic change for the nodes in our model (Fig. 2B–D). A result of hierarchical clustering using the functional genomic profile shows that the basal-like cell lines fall into two groups (Fig. 2E). The main difference between the two groups is that most cell lines have PI3K and/or PTEN alterations in Group 1, whereas KRAS alteration is more frequent in Group 2. Thus, we selected BT20 and MDA-MB-231 cell lines with the

## BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer

**Figure 1.**

DEG analysis for constructing the BLT network model. **A**, DEG analysis using TCGA database. TCGA samples of luminal A ( $n = 418$ ) and basal-like ( $n = 138$ ) subtypes are in a green and purple rectangle, respectively. DEGs that were highly expressed from the luminal A subtype or basal-like subtype are in a yellow or blue rectangle, respectively. The top-tier DEGs that were selected for the network construction are labeled in red. **B** and **C**, Some of the selected DEGs that were highly expressed in the basal-like ( $n = 21$ ; **B**) and luminal A ( $n = 12$ ; **C**) subtype from the CCLE cohort. **D–G**, Some of the selected DEGs that were highly expressed in the basal-like ( $n = 801$ ; **D**) and luminal A ( $n = 1103$ ; **E**) subtype from the METABRIC cohort. **F** and **G**, Some of the selected DEGs that were highly expressed in the basal-like ( $n = 138$ ; **F**) and luminal A ( $n = 418$ ; **G**) subtype from the TCGA cohort. Data were analyzed and are presented by two-sample Wilcoxon rank-sum (Mann-Whitney) test.

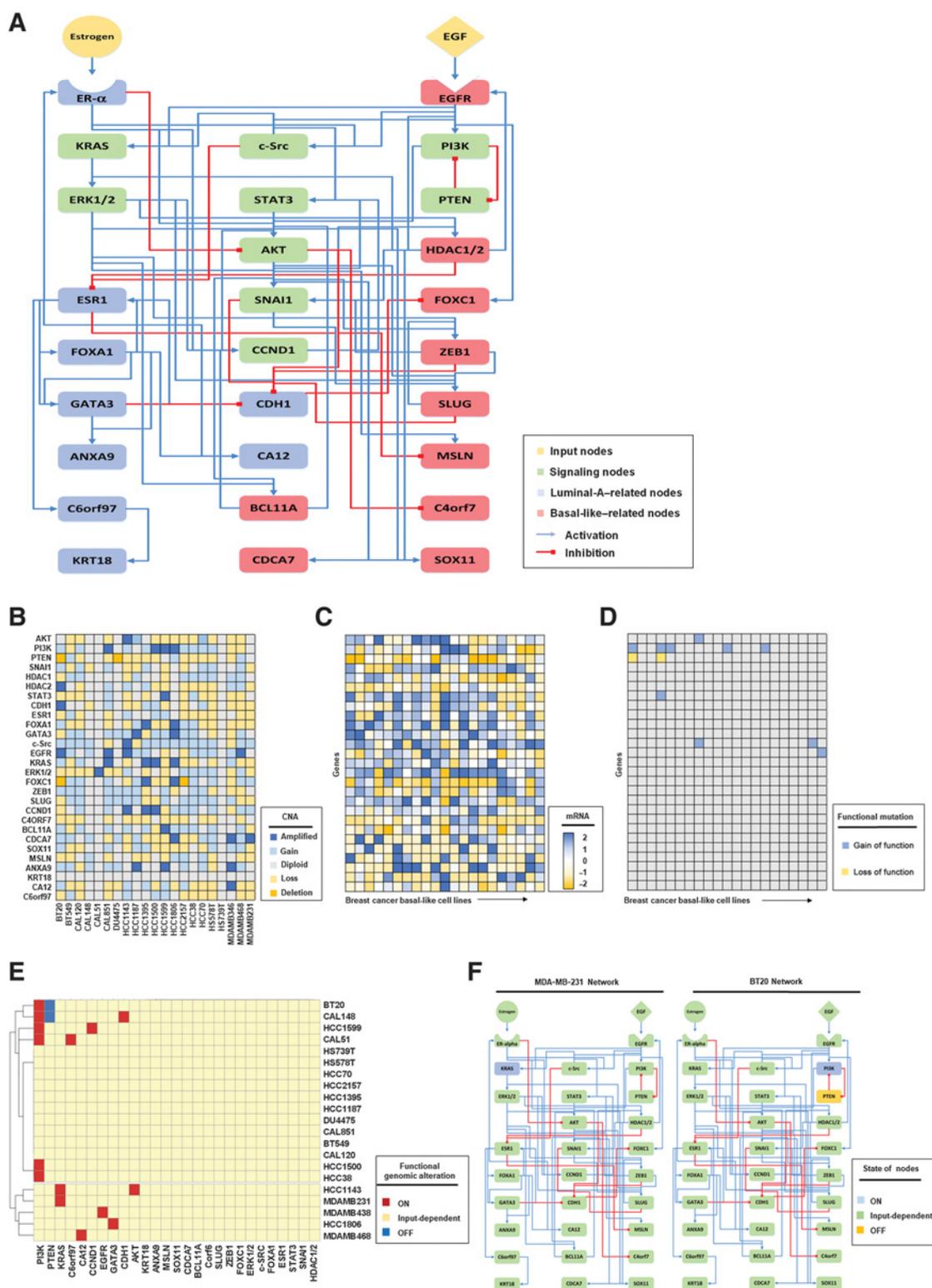
corresponding alterations from each of the two groups (Supplementary Fig. S3). To modify our BLT model to cell line-specific networks, we mapped the functional genomic alterations of these cell lines onto our model to differentially rewire the network model (Fig. 2F; Supplementary File 4).

### Reflecting molecular features for the breast cancer subtypes to the BLT model

To affirm that our BLT network model can reproduce the biological condition of both luminal A and basal-like cells, we have performed *in silico* simulation and observed the activity change of our network components depending on the intensity of each input signal, estrogen (33) and EGF (34). Considering to perform these qualitative

simulations dependent on incoming signals, one can investigate the dynamics of a network model by evaluating the output results based on different concentrations of input signals (35). For detailed information, please see Materials and Methods. We then compared our results with previous experimental data to validate our model. As a result, we were able to reproduce the physiological condition of luminal A cells, of which, the activities of luminal A DEGs (36–38), ER $\alpha$  (39), and ER $\alpha$ -downstream molecules, including c-Src, AKT, and ERK1/2 (40, 41), are dependent on the intensity of estrogen. In contrast, the activities of basal-like DEGs (42, 43), EGFR, and EGFR-downstream molecules, including PI3K, AKT, KRAS, and ERK1/2 (34), are dependent on the intensity of EGF (Supplementary Fig. S4A) in basal-like cells.

Choi et al.



**Figure 2.**

BLT network model and its modification by mapping genomic information of basal-like cancer cell lines. **A**, Our constructed network with the selected DEGs and their molecular interactions by using public databases. **B–D**, Copy number alteration (**B**), mRNA expression (**C**), and functional mutation profile (**D**) of the network for 21 basal-like breast cancer cell lines. **E**, Using the genomic data of these cell lines, the functional genomic profile is created for each cell line. Clustering analysis was performed and one cell line was selected from each cluster, MDA-MB-231 and BT20, to represent the basal-like subtype. **F**, Cell line-specific networks of the selected breast cancer cell lines are shown. The BLT network was modified to cell line-specific networks by mapping the functional genomic profile of corresponding cell lines.

**BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer**

Furthermore, basal-like cancer cells are solely dependent on EGFs to activate the downstream molecules even when estrogens coexist due to ER $\alpha$  being absent on the cell membrane. To validate this physiologic condition of the cells, we performed the same simulation with differential input settings where EGF is increasing its intensity from 0% to 100% by 1% increments with constitutively active estrogen at 0%, 50%, or 100% (Supplementary Fig. S4B). At 0% of active estrogen, the basal-like DEGs, EGFR, and its downstream molecules respond to EGF and continuously increase their activity as the intensity of EGF increases. However, the activity of luminal A DEGs and ER $\alpha$  shows no change because there is no estrogen. At 50% of active estrogen, the activity of basal-like DEGs, EGFR, and EGFR-downstream molecules increases, whereas the activity of luminal A DEGs and ER $\alpha$  decreases. Likewise, at 100% of active estrogen, the activity of basal-like-related nodes is dependent on EGF and increases their activity, whereas the luminal A-related nodes decrease their activity even with constitutively active estrogen at the highest level of intensity. These results illustrate that our model can well reproduce the cellular responses of basal-like cancer cells as well as luminal A cancer cells according to various input settings. Therefore, it is pertinent to note that our model is adequate to investigate the mechanism of the BLT at a system level.

**Identifying possible targets for BLT through a systems biological approach**

Next, we employed a complex network control strategy called logical domain of influence (LDOI)-based target control strategy (44) to identify potential targets that can drive the BLT and uncover its underlying mechanisms in our network model. The LDOI of a node can be defined as nodes and their corresponding state that are influenced and ultimately become stabilized by the node regardless of any initial states of a system. The algorithm of LDOI-based target control strategy retracts back what state of nodes influences a specific target node of interest based on the provided logical rules of a network model. As a result, it revealed nodes and their corresponding states as solution sets that can drive a system to any desired state (44).

Our network model does not have a single node that defines whether an attractor has either luminal A or basal-like phenotype. Thus, we have defined the luminal A and basal-like phenotype according to the activity of phenotypic markers within attractors. For instance, we defined a state as luminal A if the activity of ER $\alpha$  and/or luminal A DEGs are active (ON), and EGFR signaling molecules and/or basal-like DEGs are inactive (OFF). Conversely, we defined a state as basal-like if the activity of EGFR signaling molecules and/or basal-like DEGs are ON, and ER $\alpha$  and/or luminal A DEGs are OFF. Hereafter, we denoted “ $\sim$ ” as OFF or negation of a node. For instance, [ $\sim$ EGFR] corresponds to “EGFR OFF” or “NOT EGFR”. Due to complexity problem, we primarily considered identifying potential targets that can induce the activity of ER $\alpha$  and luminal A DEGs, and simultaneously block the activity of EGFR and its downstream molecules. Among the list of solution sets, we identified the optimal targets that can also predominantly block the activity of basal-like DEGs to prompt the BLT. From the BLT network without any mutation (nominal), we found a total of three potential target solution sets, which are [ $\sim$ KRAS,  $\sim$ HDAC1/2], [ $\sim$ PI3K,  $\sim$ HDAC1/2], and [ $\sim$ HDAC1/2,  $\sim$ BCL11A] (Table 1). For modifying our model to cell line-specific networks, we fixed KRAS as ON for the MDA-MB-231 network according to our functional genomic alteration profile. Likewise, we fixed PI3K as ON and PTEN as OFF for the BT20 network. As a result, we found potential target solution sets as [ $\sim$ PI3K,  $\sim$ HDAC1/2] and [ $\sim$ HDAC1/2,  $\sim$ BCL11A] for the MDA-MB-231 network, and

**Table 1.** LDOI solutions that drive basal-like to luminal A state.

No.	LDOI target solution	Total number of solutions found
<i>Network without any genomic alteration</i>		
1	BCL11A <b>OFF</b> & HDAC1/2 <b>OFF</b>	46
2	PI3K <b>OFF</b> & HDAC1/2 <b>OFF</b>	11
3	KRAS <b>OFF</b> & HDAC1/2 <b>OFF</b>	10
<i>BT20 network</i>		
1	BCL11A <b>OFF</b> & HDAC1/2 <b>OFF</b>	16
2	KRAS <b>OFF</b> & HDAC1/2 <b>OFF</b>	12
<i>MDA-MB-231 network</i>		
1	BCL11A <b>OFF</b> & HDAC1/2 <b>OFF</b>	32
2	PI3K <b>OFF</b> & HDAC1/2 <b>OFF</b>	4

Note: A list of LDOI solution sets that can drive a state to luminal A state with luminal A phenotypic marker nodes ON and basal-like phenotypic marker nodes OFF. The luminal A phenotypic marker nodes include ER $\alpha$ , C6orf97, KRT18, FOXA1, ESR1, CA12, ANXA9, and GATA3. The basal-like phenotypic marker nodes include EGFR, AKT, and ERK1/2.

Desired state:

ON: ER $\alpha$ , C6orf97, KRT18, FOXA1, ESR1, CA12, ANXA9, GATA3.

OFF: EGFR, AKT, ERK1/2.

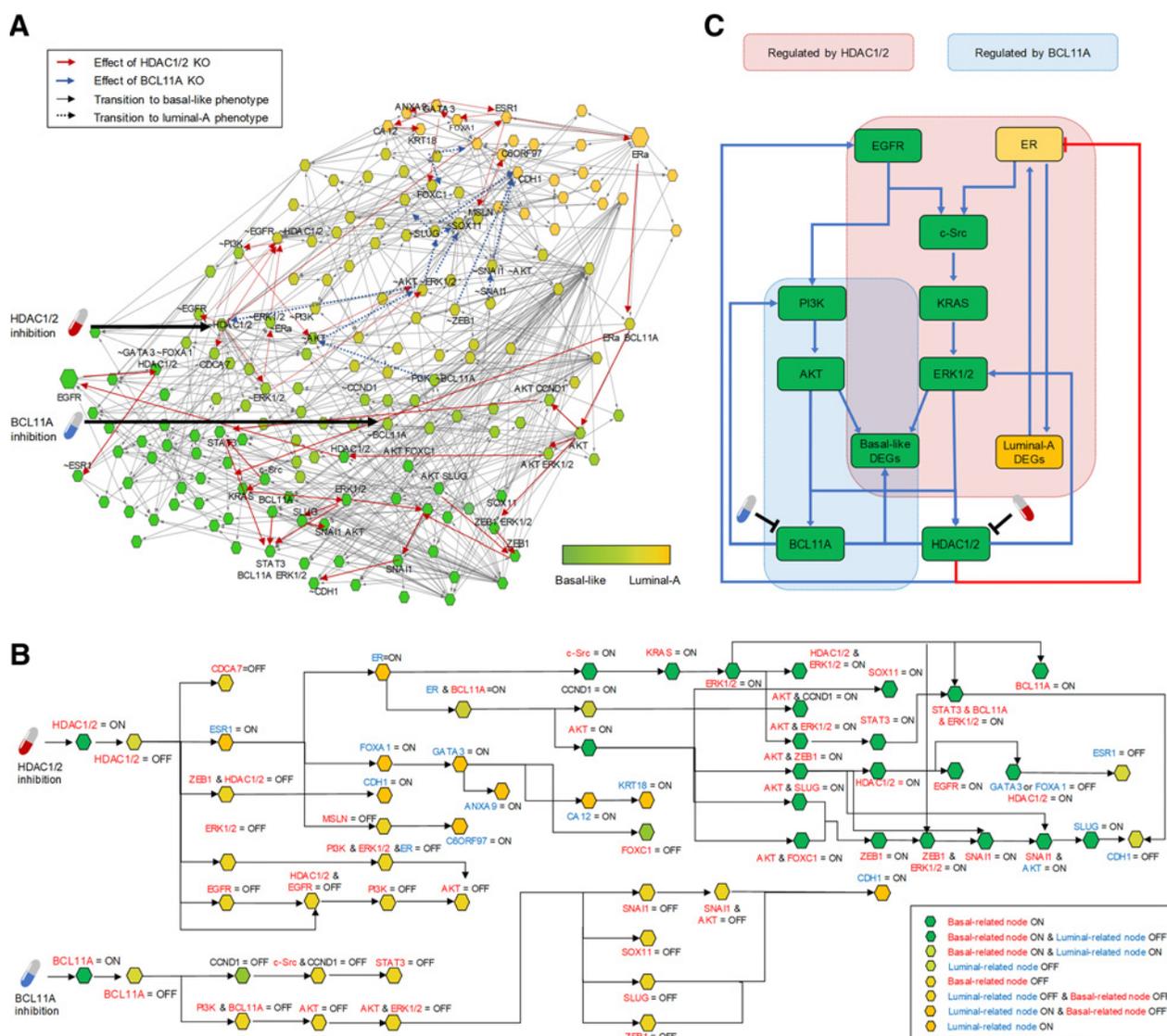
[ $\sim$ KRAS,  $\sim$ HDAC1/2] and [ $\sim$ HDAC1/2,  $\sim$ BCL11A] for the BT20 network. Interestingly, [ $\sim$ HDAC1/2,  $\sim$ BCL11A] was commonly found in all our nominal and cell line-specific network models, and thereby is the optimal combinatorial targets that can sufficiently induce the BLT. These results were in accordance with our quantitative Boolean attractor simulation by perturbing every single and double node(s) in our network models (Supplementary Fig. S5).

**Double KO of BCL11A and HDAC1/2 drives the cell-fate switching to luminal A phenotype**

An expanded network provided by the LDOI-based target control strategy is similar to hypergraph that integrates regulatory interactions and dynamics of a network (44). We employed the expanded network of our BLT model to analyze how BCL11A and HDAC1/2 interact with the other network components. The resulting expanded network of our model encompasses every footprint of the network components, including all possible initial states (basal-like phenotype) to our desired state, luminal A (Fig. 3A). We then analyzed how the BLT is occurring by navigating the transition after each perturbation, HDAC1/2 KO and/or BCL11A KO.

KO of HDAC1/2 induced the activity of [ESR1], which prompted the activity of luminal A DEGs and ER $\alpha$  as well as directly initiated [ $\sim$ EGFR] and [ $\sim$ ERK1/2] (Fig. 3B). As soon as ER $\alpha$  was activated, its direct downstream molecule, [c-Src], was also activated and induced its further downstream molecules, including [AKT] and [ERK1/2]. Therefore, KO of HDAC1/2 initially induced the activity of ER $\alpha$  and luminal A DEGs but ultimately turned OFF their activity by activating the EGFR downstream molecules that are shared with the ER $\alpha$  signaling pathway (Fig. 3B). In contrast, KO of BCL11A directly downregulated [PI3K], [AKT], and some of the basal-like DEGs such as [SOX11], [ZEB1], and [SLUG] (Fig. 3C). It can also partially regulate the activity of HDAC1/2 that can be turned OFF only if both [AKT] and [ERK1/2] are inactive. Together, it is imperative that both KO of HDAC1/2 and BCL11A are required to induce the activity of ER $\alpha$  and luminal A DEGs, and simultaneously block the activity of EGFR signaling molecules and basal-like DEGs.

Choi et al.

**Figure 3.**

The mechanisms of BLT induced by KO of BCL11A and HDAC1/2. **A**, The entire state transitions are shown. Each node represents an individual state of node(s) making a transition from one to another. The commonly shared expanded network of our BLT and cell line-specific models after inhibiting BCL11A and HDAC1/2. The effect after HDAC1/2 KO or BCL11A KO is represented by red or blue arrows, respectively. The state transition that approaches basal-like or luminal A phenotype is represented by solid or dotted arrows, respectively. **B**, The state transitions after KO of HDAC1/2 or BCL11A are shown in a hierarchical order. KO of HDAC1/2 (~HDAC1/2) or BCL11A (~BCL11A) is represented by red or blue drug, respectively. The basal-like-related nodes' names are labeled in red, and the luminal A-related nodes' names are labeled in blue. A state of basal-like-related node ON, or a state of both basal-like-related node ON and luminal A-related node OFF is color-coded in green. A state of luminal A-related node OFF, or a state of both basal-like-related node ON and luminal A-related node ON is color-coded in lighter green. A state of basal-like-related node OFF, or a state of both luminal A-related node OFF and basal-like-related node OFF is color-coded in yellowish green. A state of luminal A-related node ON, or a state of both luminal A-related node ON and basal-like-related node OFF is color-coded in yellow. **C**, Overall schematic representation of the regulation after KO of HDAC1/2 and/or BCL11A. Nodes that are regulated by HDAC1/2 are within the dotted boundary in red. Nodes that are regulated by BCL11A are within the dotted boundary in blue.

### Modification of dynamic stability of the attractors after KO of BCL11A and HDAC1/2

To investigate dynamic stability of the states and to examine effectiveness of the targets for the BLT, we performed attractor landscape analysis by mathematically analyzing the basin of each attractor. An attractor is a steady state within a network model that can be defined by binary activity of molecules (23). The attractor can be either stabilized at a single state, which is named as a point attractor, or

as a limit cycle of multiple states that are recurrently traversed by dynamics of a network, which is named as a cyclic attractor (45). Boolean network with  $n$  nodes has  $2^n$  different initial states as their entire state space. A basin of attractors is a set of initial states that will ultimately converge to a particular attractor determined by the interactions between network components. We can analyze the relative stability of those components by measuring the basin size of attractors (46, 47). For this, we first defined the phenotype of each attractor

**BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer**

by using binary activities of the phenotypic marker nodes within each attractor (see Materials and Methods for details). We then calculated how many initial states converge to the attractors with a particular phenotype over time out of a total number of the initial states in each network before and after the target perturbation. In the MDA-MB-231 network (no perturbation), there were four attractors whose phenotype is basal-like according to their phenotypic score, and the basin of luminal A was 0% (Fig. 4A). After KO of *BCL11A* or *HDAC1/2* from the network, there were distributed attractors, including point and cyclic attractors with repeatedly visited luminal A states, and the basin of luminal A phenotype was increased to 49.342% or 33.956%, respectively. When both *BCL11A* and *HDAC1/2* were KO, the basin of the entire eight attractors was shown to have the luminal A phenotype with 100% basin size. These results were consistent with the BT20 network (Fig. 4B).

**High expressions of *BCL11A* and/or *HDAC1/2* correlated with poor prognosis**

To validate our network analysis and simulation results, we performed Kaplan–Meier analysis of overall survival among basal-like and/or luminal A patients using TCGA (Fig. 4C–K) and METABRIC (Supplementary Fig. S6) cohorts. Among TCGA basal-like patients ( $n = 115$ ), overall survival for *BCL11A* (high = 27, low = 29,  $P = 0.049$ ), *HDAC1* (high = 16, low = 28,  $P = 0.04$ ), and *HDAC2* (high = 17, low = 27,  $P = 0.012$ ) showed that high expressions of these genes were correlated with poor prognosis (Fig. 4C–E). Similar results were shown from TCGA luminal A patients ( $n = 356$ ; Fig. 4F–H). Finally, among basal-like and luminal A patients ( $n = 471$ ), Kaplan–Meier analysis of overall survival for *BCL11A* (high = 76, low = 93,  $P = 0.049$ ), *HDAC1* (high = 51, low = 89,  $P = 0.024$ ), and *HDAC2* (high = 91, low = 94,  $P = 0.011$ ) also showed that high expressions of the corresponding genes were correlated with poor prognosis (Fig. 4I–K). These results were in concordance with the results using the METABRIC dataset (Supplementary Fig. S6).

**BCL11A and HDAC1/2 regulate protein activities and mRNA expression of luminal A and basal-like phenotypic markers**

To further validate our analyses, we performed *in vitro* experiments for short hairpin RNA (shRNA) knockdown (KD) and drug treatment in three of the basal-like breast cancer cell lines, MDA-MB-231 (Fig. 5; Supplementary Fig. S7), BT20 (Fig. 5), and HS578T (Supplementary Fig. S7A and S7B). All experiments were repeated three times. We have observed the BLT through cell culture experiments (mRNA, protein expressions, and cell morphology), drug responsiveness, and growth factor–dependent cell growth. We used shRNA for blocking *BCL11A* gene (Fig. 5A and C), whereas romidepsin (an *HDAC1/2*-specific inhibitor; Fig. 5B and D) that biochemically binds to *HDAC1/2* to prevent their protein activity. We also have used shRNA for blocking *HDAC1/2* genes in the MDA-MB-231 cells to observe the expression of *ESR1* and *ER $\alpha$*  (Supplementary Fig. S7C and S7D). The effect of BLT using shRNA was consistent with romidepsin through the mRNA expression of *ESR1* and protein expression of *ER $\alpha$* .

Blocking *HDAC1/2* remarkably increased the mRNA expression of *ESR1* by 4-fold in MDA-MB-231 cells ( $P = 0.048$ ) and 2-fold in BT20 cells ( $P = 0.049$ ; Fig. 5B and D). KD of *BCL11A* did not show much change for the expression of *ESR1* in MDA-MB-231 cells ( $P = 0.015$ ), whereas it increased the expression of *ESR1* by 1.5-fold in BT20 cells ( $P = 0.035$ ). After both *BCL11A* and *HDAC1/2* were inhibited, the expression of *ESR1* was increased by 10-fold in MDA-MB-231 cells ( $P = 0.005$ ) and 3.5-fold in BT20 cells ( $P = 0.003$ ). These results showed that the fold changes of *ESR1* expression after inhibiting the

target(s) can vary depending on the cell lines with different mutation profiles. However, it was consistent that inhibiting both *BCL11A* and *HDAC1/2* can synergistically increase the expression of *ESR1* to drive the BLT in both MDA-MB-231 and BT20 cells.

Next, we accessed the effect of BLT by measuring the protein expression of *ER $\alpha$* , *EGFR*, and *ERK1/2* and then quantified the results using ImageJ software. Using romidepsin to block *HDAC1/2* considerably increased the expression of *ER $\alpha$*  in BT20 cells. Although the effect was not as substantial in sh*BCL11A*-treated cells, the expression of *ER $\alpha$*  was increased by 15-fold in MDA-MB-231 cells ( $P = 0.031$ ) and 18-fold in BT20 cells ( $P = 0.046$ ) after inhibiting both *BCL11A* and *HDAC1/2*. Moreover, the expression of phosphorylated *EGFR* (p*EGFR*) was decreased by 10-fold in MDA-MB-231 cells ( $P = 0.047$ ) and 100-fold in BT20 cells ( $P = 0.006$ ) after inhibiting both *BCL11A* and *HDAC1/2*. Similarly, the expression of phosphorylated *ERK1/2* (p*ERK1/2*) was reduced by 3-fold in MDAMB231 cells ( $P < 0.001$ ) and 114-fold in BT20 cells ( $P = 0.003$ ; Fig. 5E and F).

**BCL11A and HDAC1/2 regulate tamoxifen sensitivity in basal-like breast cancer cells**

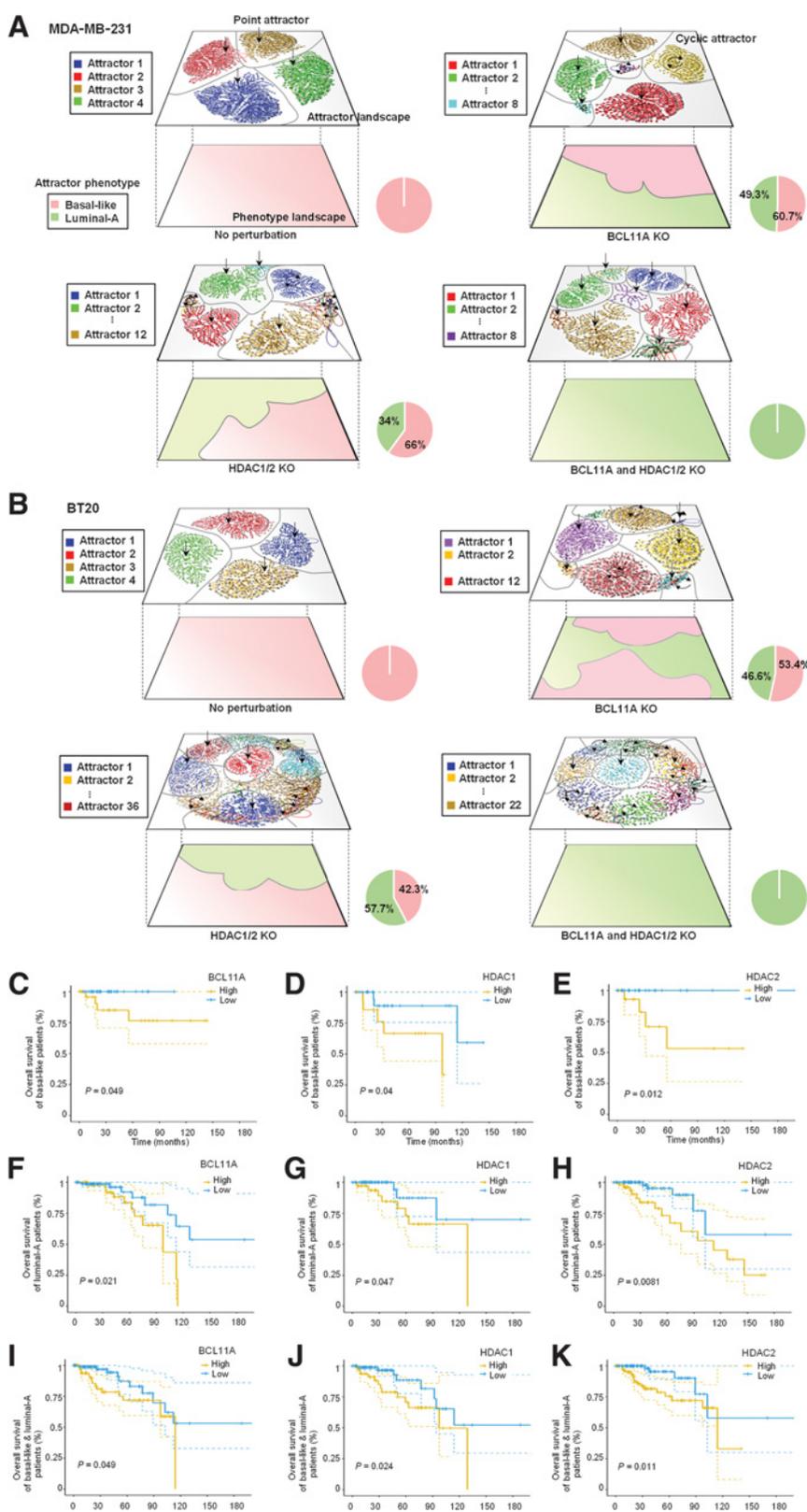
Increased protein expression of *ER $\alpha$*  in basal-like cells after blocking *BCL11A* and *HDAC1/2* suggested that the cells have been reprogrammed into luminal A cells. We hypothesized that these reprogrammed cells would now be sensitive to anti-hormone treatment targeting *ER $\alpha$* . To validate this hypothesis, we used 5  $\mu$ g/mL of 4-OH-tamoxifen (tamoxifen), which is an *ER $\alpha$* -targeted drug for luminal breast cancer patients, to treat MDA-MB-231 and BT20 cells. While the control cells (shScrambled) showed no sensitivity, *BCL11A*-silenced cells (using sh*BCL11A*) or *HDAC1/2*-inhibited cells (using romidepsin) showed some sensitivity to tamoxifen. As expected, the cells that were treated with both sh*BCL11A* and romidepsin, which had the highest protein expression level of *ER $\alpha$* , showed the most sensitivity to tamoxifen in both MDA-MB-231 (Fig. 5G and I) and BT20 (Fig. 5H and J) cells. These results support our hypothesis that depletion of both *BCL11A* and *HDAC1/2* can sufficiently induce the BLT through their increased expression of *ESR1* and *ER $\alpha$* , and tamoxifen sensitivity.

**Overexpression of BCL11A and HDAC1/2 can reactivate EGFR-ERK1/2 signaling in luminal A cells**

To further validate, we conversely performed *in vitro* experiments for *BCL11A* and *HDAC1/2* overexpression (OE), and drug treatment in two of the luminal A breast cancer cell lines, T47D (Fig. 6) and MCF7 (Supplementary Fig. S7E–S7I). Fold difference >1.5-fold increase was considered to be overexpression. *BCL11A* and/or *HDAC1/2* OE did not show dramatic changes of *ESR1* or *EGFR* mRNA expression levels in T47D cells (Fig. 6A–C). After *BCL11A* or *HDAC1/2* OE, however, we observed decreased protein expressions of *ER $\alpha$*  as well as increased expressions of p*EGFR* and p*ERK1/2*. Almost no expression of *ER $\alpha$*  was observed when all these genes were overexpressed in T47D cells (Fig. 6D). Similarly, *BCL11A* and *HDAC1/2* OE synergistically decreased the expression of *ESR1* for 2-fold and increased the *EGFR* expression for 30-fold in MCF7 cells (Supplementary Fig. S7E–S7G) as well as their protein expressions accordingly (Supplementary Fig. S7H). Moreover, *BCL11A* and/or *HDAC1/2* OE also drastically changed the morphology of MCF7 cells towards a more mesenchymal-like phenotype (Supplementary Fig. S7I).

Decreased protein expression of *ER $\alpha$*  in luminal A cells after *BCL11A* and *HDAC1/2* OE suggested that the cells possibly have been reprogrammed into basal-like cells that are not responsive to tamoxifen.

Choi et al.



**Figure 4.**

Dynamic stability of the attractors and patient outcomes depend on the expression of *BCL11A* and *HDAC1/2*. The basin size of attractors is used to analyze the stability of each state. Each attractor is defined as a basal-like or luminal A phenotype in pink or green, respectively. **A**, In the MDA-MB-231 network (no perturbation), there are four major attractors in blue, red, yellow, and green, and all of them have the luminal A phenotype with 0% basin. After *BCL11A* KO, there were eight attractors, including both point and cyclic attractors, and the basin of luminal A phenotype was increased to 49.3%. After *HDAC1/2* KO, there were 12 attractors, including both point and cyclic attractors, and the basin of luminal A phenotype was 34%. After *BCL11A* and *HDAC1/2* KO, the basin of the network was shown to have the luminal A phenotype with 100%. **B**, In the BT20 network (no perturbation), there are four major attractors with 0% basin of the luminal A phenotype. After *BCL11A* KO, there were 12 attractors and the basin of luminal A phenotype was increased to 46.6%. The network has 57.7% or 100% of the luminal A phenotype after *HDAC1/2* KO or both *BCL11A* and *HDAC1/2* KO, respectively. Kaplan-Meier analysis of overall survival among basal-like patients ( $n = 115$ ) with high ( $n = 27$ ) and low ( $n = 29$ ) expression of *BCL11A* (**C**), high ( $n = 16$ ) and low ( $n = 28$ ) expression of *HDAC1* (**D**), and high ( $n = 17$ ) and low ( $n = 27$ ) expression of *HDAC2* (**E**) using TCGA cohort. Overall survival among TCGA luminal A patients ( $n = 356$ ) with high ( $n = 52$ ) and low ( $n = 86$ ) expression of *BCL11A* (**F**), high ( $n = 35$ ) and low ( $n = 62$ ) expression of *HDAC1* (**G**), and high ( $n = 54$ ) and low ( $n = 89$ ) expression of *HDAC2* (**H**). Overall survival among basal-like and luminal A patients ( $n = 471$ ) with high ( $n = 76$ ) and low ( $n = 93$ ) expression of *BCL11A* (**I**), high ( $n = 51$ ) and low ( $n = 89$ ) expression of *HDAC1* (**J**), and high ( $n = 91$ ) and low ( $n = 94$ ) expression of *HDAC2* (**K**).

To validate our hypothesis, we used 10  $\mu\text{g}/\text{mL}$  of tamoxifen to treat the cells. While the control cells responded to tamoxifen, *BCL11A* OE and/or *HDAC1/2* OE cells showed no sensitivity to tamoxifen in T47D cells (Fig. 6E–G). These results further substantiate our hypothesis that both *BCL11A* and *HDAC1/2* OE can conversely induce the reverse-BLT through the upregulated expression of pERK1/2 and pEGFR as well as the downregulated expression of ER $\alpha$ , which ultimately regulate tamoxifen sensitivity.

### BCL11A and HDAC1/2 regulate cell growth dependency in basal-like and luminal A cells

To analyze cell growth dependency on different growth factors, cells were measured using IncuCyte ZOOM to assess cell viability rates. While the shScrambled basal-like cells showed no dependency on estradiol (E2) for growth, sh*BCL11A* or sh*HDAC1/2* cells showed some dependency, and sh*BCL11A* and sh*HDAC1/2* cells became fully dependent on E2 for growth (Supplementary Fig. S8A). When treated with EGF, the shScrambled cells were dependent on EGF for growth, whereas all of *BCL11A*- and/or *HDAC1/2*-silenced cells showed no dependency on EGF for growth (Supplementary Fig. S8B). This is due to increased protein expressions of ER $\alpha$  (Supplementary Fig. S8C) and decreased expressions of pEGFR and pERK1/2 (Supplementary Fig. S8D) after inhibiting *BCL11A* and *HDAC1/2*. Furthermore, E2-triggered basal-like cells showed increased RNA expression patterns of ER $\alpha$  downstream target genes after silencing the target(s) (Supplementary Fig. S8E), whereas the cells showed no change or decreased RNA expression patterns of EGFR downstream target genes even after EGF stimulation (Supplementary Fig. S8F). This indicates that inhibition of *BCL11A* and *HDAC1/2* not only enhances targetability of basal-like cells but also switches their dependency from EGF to E2 for growth.

To further validate, we conversely analyzed cell growth dependency in luminal A cells. The results showed that the control cells had complete dependency on E2 for growth, whereas *BCL11A* and/or *HDAC1/2* OE cells were not dependent on E2 for growth (Supplementary Fig. S8G). In addition, the control cells did not show any cell growth increase after EGF stimulation, whereas *BCL11A* and/or *HDAC1/2* OE cells showed dependency on EGF for growth (Supplementary Fig. S8H). Although the OE cells still showed increased protein expressions of ER $\alpha$  after E2 stimulation (Supplementary Fig. S8I), they also showed increased expression of pEGFR and pERK1/2 (Supplementary Fig. S8J), which possibly had led them to depend on EGF for growth. Furthermore, E2-triggered *BCL11A* and/or *HDAC1/2* OE luminal A cells showed no change in RNA expression patterns of the ER $\alpha$ -downstream target genes in contrast to their control (Supplementary Fig. S8K). However, the OE cells showed increased RNA expression patterns of the EGFR downstream target genes after EGF stimulation (Supplementary Fig. S8L). These results further support our hypothesis that depletion of both *BCL11A* and *HDAC1/2* can sufficiently induce the BLT, whereas *BCL11A* and *HDAC1/2* OE can conversely induce the reverse-BLT through activating the EGFR pathway.

## Discussion

Switching the signal flow from the EGFR- to ER $\alpha$  signaling pathway can make cells rely on the ER $\alpha$  cascade for their growth. This can be partially done by solely blocking *BCL11A* in our network model. The role of *BCL11A* KO is to inactivate the EGFR/AKT signaling pathway by blocking its interaction with PI3K, which is in accord with the previous study (48). To inactivate further downstream molecules such as ZEB1, SLUG, and SOX11, ERK1/2 also needs to be inactivated,

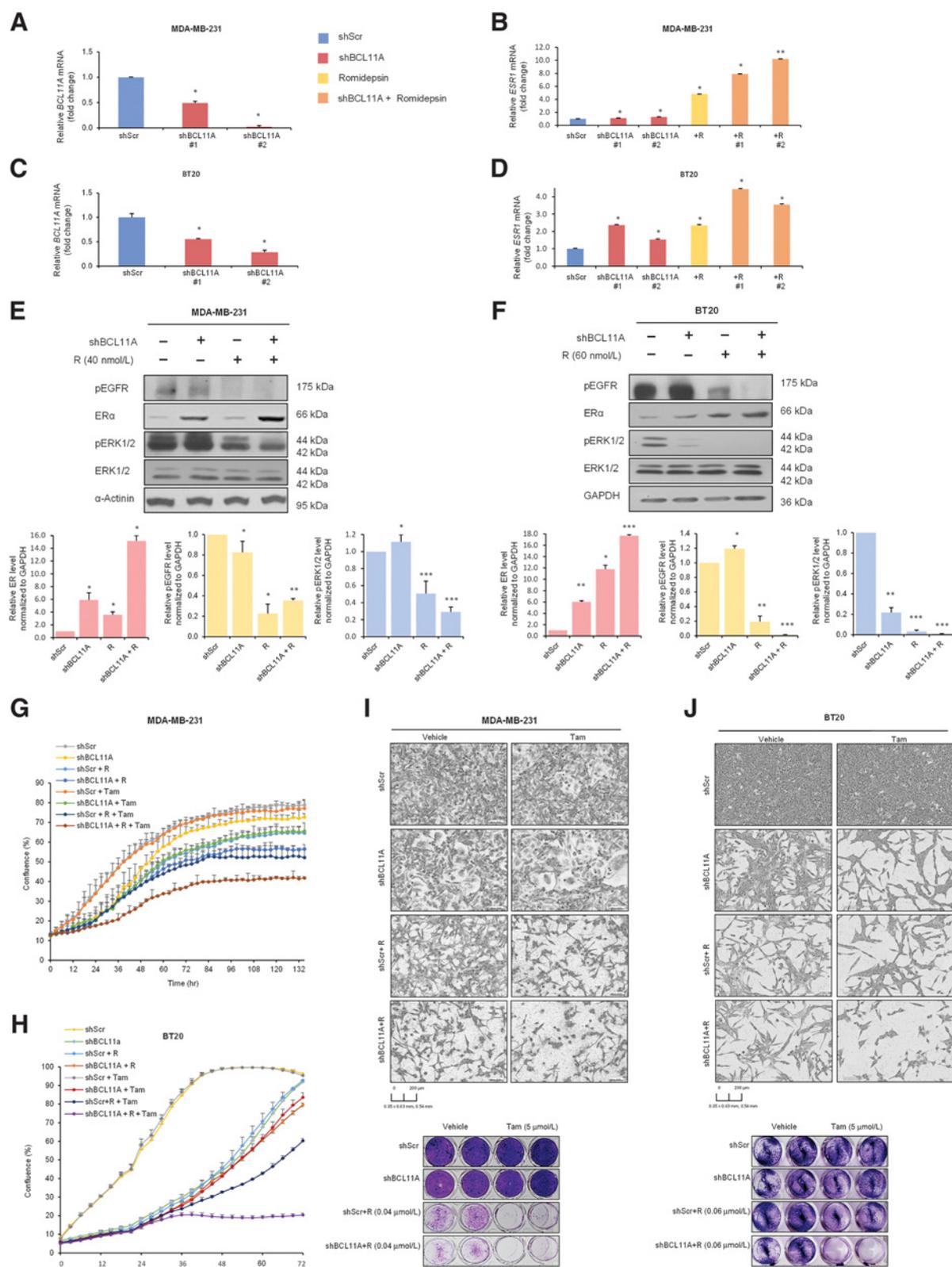
which can be done by blocking *HDAC1/2*. KO of *HDAC1/2* has two opposite roles in our network model. It initially induces the activity of ESR1 and thereby reduces the activity of EGFR, ERK1/2, and the basal-like DEGs, CDCA7 and ZEB1. This is consistent with previous studies that *HDAC1/2* transcriptionally regulates various genes, including *ESR1* (49, 50) and *EGFR* (47, 51). Inactivation of both ERK1/2 and AKT can terminate the activity of entire basal-like DEGs. Moreover, induction of ESR1 activity further activates ER $\alpha$  and most luminal A DEGs. As soon as the activity of ER $\alpha$  is induced, however, it starts to trigger the activity c-SRC and its downstream molecules that are shared between the EGFR and ER $\alpha$  signaling pathways. As a result, the EGFR signaling molecules as well as the basal-like DEGs become activated and, eventually, intervene positive feedback between the luminal A DEGs and ER $\alpha$  signaling pathway. Therefore, both *BCL11A* and *HDAC1/2* are required to be terminated to completely reprogram basal-like cells into luminal A cells by turning OFF the basal-like phenotypic nodes and turning ON the luminal A phenotypic nodes.

To confirm that our results are clinically reasonable, we analyzed the gene expression data of 356 patients with breast cancer with tamoxifen treatment. We first categorized the patient samples into four groups, *BCL11A* & *HDAC1/2* high, *BCL11A* low, *HDAC1/2* low, and *BCL11A* & *HDAC1/2* low, based on the expression of these genes. We then analyzed the relationship between tamoxifen sensitivity and the expression of our target genes according to each conditional subgroup. As a result, a higher percentage of the tamoxifen-resistant group was enriched within the *BCL11A* & *HDAC1/2* high group, whereas the tamoxifen-sensitive group was highly enriched in the *BCL11A* & *HDAC1/2* low group (Supplementary Fig. S9A). Thus, we further affirmed that high expressions of the target genes were enriched in the tamoxifen-resistant patient group and inhibiting them can possibly drive the patients to become sensitive to the drug.

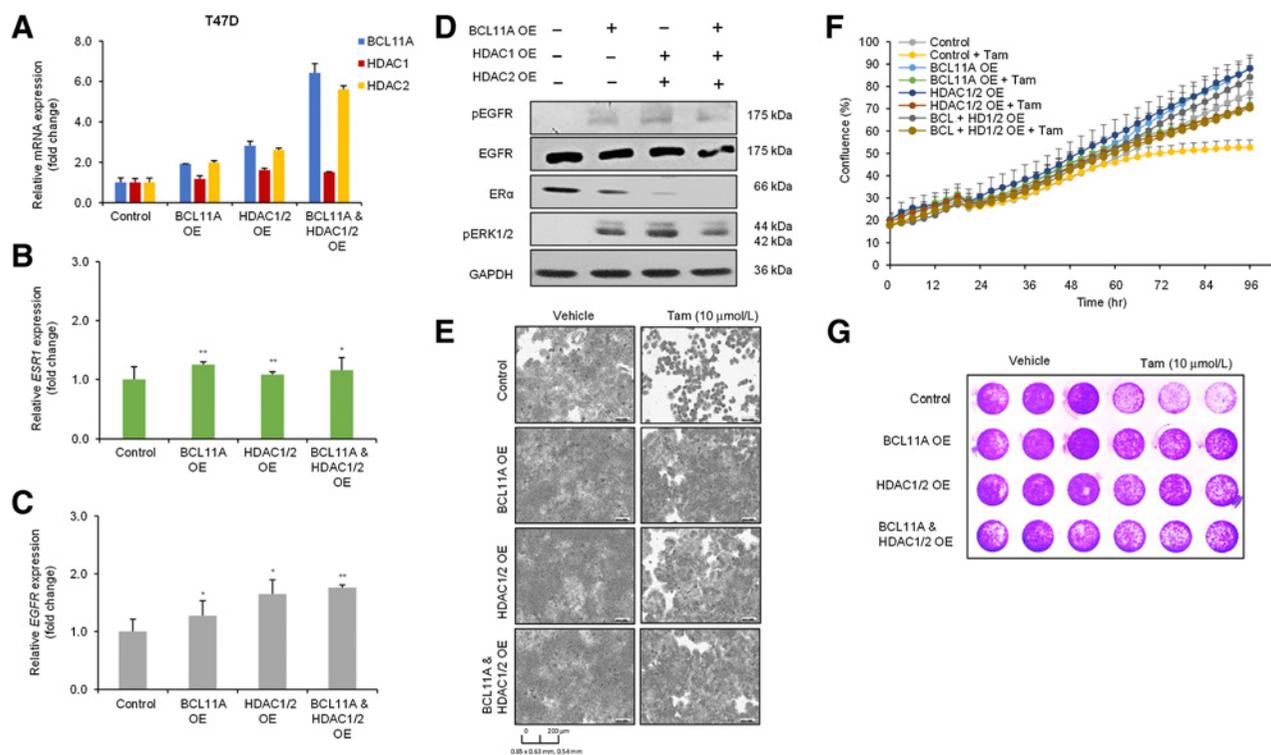
Finally, we performed GSVA to assess gene set enrichment of these patient samples by calculating sample-wise gene set enrichment scores of various gene sets that are associated with the ER $\alpha$  and EGFR signaling pathways. GSVA scores of the gene sets that are upregulated in basal-like or luminal A breast cancer (Supplementary Fig. S9B), and are downstream genes of ER $\alpha$  or EGFR signaling pathway (Supplementary Fig. S9C) for each of the conditional subgroups. These results show that cancer cells with high expression levels of *BCL11A* and *HDAC1/2* are more likely to be tamoxifen-resistant and dependent on the EGFR signaling pathway for growth. Conversely, cancer cells with low expression levels of *BCL11A* and *HDAC1/2* are more likely to be tamoxifen-sensitive and dependent on the ER $\alpha$  signaling pathway for growth. Therefore, blocking *BCL11A* and *HDAC1/2* will drive tamoxifen-resistant cells to activate the ER $\alpha$  signaling pathway, which may provide a vulnerability for endocrine therapy in basal-like patients.

*HDAC1/2* is a well-known epigenetic regulator that governs the expression of ER $\alpha$  (49) by removing acetyl groups that are bound at the estrogen response element (ERE) site of ER $\alpha$  or from the promoter region of *ESR1* (52). In addition, *BCL11A* is a transcription factor that regulates the cell cycle and apoptosis in lymphoma and hematopoietic stem cells (24, 53). *BCL11A* is also known for regulating various chromatin regulators such as histone deacetylases. A previous study showed that *BCL11A* can recruit endogenous SIRT1, the class III histone deacetylase, to the promoter template, which plays an important role in transcriptional regulation (54). In addition, *BCL11A* is a potential regulator of fetal hemoglobin that can ameliorate the major  $\beta$ -hemoglobin disorders by regulating the nucleosome remodeling deacetylase (NuRD) complex, including *HDAC1/2*, in erythroid

Choi et al.



## BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer



**Figure 6.**

Effects of *BCL11A* and *HDAC1/2* overexpression and tamoxifen drug responses. **A**, The mRNA expression of *BCL11A*, *HDAC1*, and *HDAC2* was effectively overexpressed in T47D cells. **B** and **C**, The mRNA expression of *ESRT1* (**B**) and *EGFR* (**C**) after *BCL11A* and/or *HDAC1/2* OE in T47D cells are shown. **D**, Protein expression of ER $\alpha$  and pEGFR and pERK1/2 after overexpressing *BCL11A* and/or *HDAC1/2* are shown. **E**, Left, T47D cells expressing vector alone, or *BCL11A* OE and/or *HDAC1/2*, respectively, are shown. Right, corresponding cells treated with tamoxifen (Tam: 10  $\mu$ mol/L) are shown. **F**, Drug sensitivity to tamoxifen after overexpressing *BCL11A* and/or *HDAC1/2* in T47D cells. **G**, Crystal violet staining of cells treated with tamoxifen. *BCL11A* OE, *BCL11A* overexpression. *HDAC1/2* OE, *HDAC1/2* overexpression. Tam, tamoxifen. Data are presented as mean  $\pm$  SEM (error bars;  $n = 3$ ). \*,  $P < 0.05$ ; \*\*,  $P < 0.01$  by Student  $t$  test.

cells (55, 56). Furthermore, its interaction with histone deacetylase complexes promotes tumorigenesis in triple-negative breast cancer (TNBC; ref. 57). Although *BCL11A* and *HDAC1/2* were previously reported as highly expressed genes in TNBC (49, 58), their role in the cellular functions of breast cancer cells was not previously identified.

Cancer is a highly heterogeneous disease with various genetic and epigenetic modifications. Although we are only considering two subtypes within breast cancer, our analysis has shown that even these relatively homogeneous groups of patients with breast cancer and cell lines could be very heterogeneous with their diversified mutation profiles. For instance, we found out from our *in vitro* experiments that interactive relationships between *BCL11A* and *HDAC1/2* could be different at the gene expression level and independently related to cancer subtypes or cell types with different mutation profiles. In addition, the scope of this study was limited to *in vitro* experiments

to validate the identified targets using the systems biological approach. Even with our effort to overcome these limitations, it is still hard to rationalize that our network can recapitulate the entire basal-like and luminal A breast cancer subtypes with diverse genetic backgrounds. Thus, it is imperative to study the relationship between *BCL11A* and *HDAC1/2* more in-depth using animal models as well as to more precisely represent each of patient samples and cell lines.

Here, we suggest that highly expressed *BCL11A* may recruit more frequently existing *HDAC1/2* to the promoter region of *ESRT1* to block its transcription as well as its target genes. Moreover, we show how double inhibition of *BCL11A* and *HDAC1/2* can regulate the ER $\alpha$  and EGFR signaling cascades as well as the basal-like and luminal DEGs to synergistically reprogram basal-like cells into luminal A cells, and vice versa. Therefore, our study demonstrates that the systems biological approach to identify such targets can be a useful tool for analyzing

**Figure 5.**

Effects of *BCL11A* and *HDAC1/2* silencing on mRNA and protein expressions, and tamoxifen drug response. **A-D**, Effects of mRNA expression after silencing *BCL11A* and *HDAC1/2* in MDA-MB-231 and BT20 cells. The mRNA expression of *BCL11A* was effectively silenced with *BCL11A* shRNA sequences in MDA-MB-231 (**A**) and BT20 (**C**) cells. The mRNA expression of *ESRT1* after inhibiting *BCL11A* and/or *HDAC1/2* in MDA-MB-231 (**B**) and BT20 (**D**) cells. **E** and **F**, Protein expression of ER $\alpha$ , and phosphorylated EGFR (pEGFR) and ERK1/2 (pERK1/2) after inhibiting *BCL11A* and/or *HDAC1/2* in MDA-MB-231 cells (**E**) and BT20 cells (**F**). Bottom, representative Western blot results were quantified by ImageJ and normalized by GAPDH. **G** and **H**, Drug sensitivity to tamoxifen after inhibiting *BCL11A* and/or *HDAC1/2* in MDA-MB-231 cells (**G**) and BT20 cells (**H**). **I** and **J**, Images of MDA-MB-231 cells (**I**) and BT20 cells (**J**) were taken after 3-day treatment with the indicated drugs. Left, MDA-MB-231 cells expressing vector alone, or sh*BCL11A* and/or treated with romidepsin (R: 0.04  $\mu$ mol/L for MDA-MB-231 or 0.06  $\mu$ mol/L for BT20), respectively, are shown. Right, corresponding cells treated with tamoxifen (Tam; 5  $\mu$ mol/L) are shown. Bottom, crystal violet staining of cells treated with the indicated drugs. shScr, scrambled shRNA. sh*BCL11A*, *BCL11A* shRNA (#1, sequence #1; #2, sequence #2). Data are presented as mean  $\pm$  SEM (error bars;  $n = 3$ ). \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$  by Student  $t$  test.

Choi et al.

and governing the repertoire of complex biological networks of biological phenomena.

### Authors' Disclosures

No disclosures were reported.

### Authors' Contributions

**S.R. Choi:** Resources, data curation, formal analysis, validation, investigation, visualization, methodology, writing—original draft. **C.Y. Hwang:** Resources, formal analysis, validation, investigation, visualization, methodology. **J. Lee:** Software, formal analysis, investigation, visualization, methodology. **K.-H. Cho:** Conceptualization, formal analysis, supervision, funding acquisition, validation, investigation, writing—original draft, project administration, writing—review and editing.

### Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korea Government, the Ministry of Science and ICT (2020R1A2B5B03094920), and by Electronics and Telecommunications Research Institute (ETRI) grants (21ZS1100, Core Technology Research for Self-Improving Integrated Artificial Intelligence System). It was also supported by the KAIST Grand Challenge 30 Project. We thank Dongkwan Shin and Jae Il Joo for their analytic support.

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 February 22, 2021; revised July 13, 2021; accepted November 15, 2021; published first November 29, 2021.

### References

- Torre LA, Islami F, Siegel RL, Ward EM, Jemal A. Global cancer in women: Burden and trend. *Cancer Epidemiol Biomarkers Prev* 2017;26:444–57.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Prat A, Perou CM. Mammary development meets cancer genomics. *Nat Med* 2009;15:842–4.
- Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 2009;15:907–13.
- Dai X, Cheng H, Chen X, Li T, Zhang J, Jin G, et al. FOXA1 is prognostic of triple negative breast cancers by transcriptionally suppressing SOD2 and IL6. *Int J Biol Sci* 2019;15:1030–41.
- Esteve FJ, Sahin AA, Cristofanilli M, Arun B, Hortobagyi GN. Molecular prognostic factors for breast cancer metastasis and survival. *Semin Radiat Oncol* 2002;12:319–28.
- Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol* 2010;28:1684–91.
- Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LA, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *J Natl Cancer Inst* 2014;106:dju055.
- Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Dressler LG, et al. Epidemiology of basal-like breast cancer. *Breast Cancer Res Treat* 2008;109:123–39.
- Jiang W, Wang X, Zhang C, Xue L, Yang L. Expression and clinical significance of MAPK and EGFR in triple-negative breast cancer. *Oncol Lett* 2020;19:1842–8.
- Nakai K, Hung MC, Yamaguchi H. A perspective on anti-EGFR therapies targeting triple-negative breast cancer. *Am J Cancer Res* 2016;6:1609–23.
- Majorini MT, Manenti G, Mano M, De Cecco L, Conti A, Pinciroli P, et al. cIAP1 regulates the EGFR/Snai2 axis in triple-negative breast cancer cells. *Cell Death Differ* 2018;25:2147–64.
- Su Y, Hopfinger NR, Nguyen TD, Pogash TJ, Santucci-Pereira J, Russo J. Epigenetic reprogramming of epithelial mesenchymal transition in triple negative breast cancer cells with DNA methyltransferase and histone deacetylase inhibitors. *J Exp Clin Cancer Res* 2018;37:314.
- Zhao H, Agazie YM. Inhibition of SHP2 in basal-like and triple-negative breast cells induces basal-to-luminal transition, hormone dependency, and sensitivity to anti-hormone treatment. *BMC Cancer* 2015;15:109.
- Kwon YK, Cho KH. Boolean dynamics of biological networks with multiple coupled feedback loops. *Biophys J* 2007;92:2975–81.
- Fumia HF, Martins ML. Boolean network model for cancer pathways: predicting carcinogenesis and targeted therapy outcomes. *PLoS One* 2013;8:e69008.
- Shin SY, Kim T, Lee HS, Kang JH, Lee JY, Cho KH, et al. The switching role of beta-adrenergic receptor signalling in cell survival or death decision of cardiomyocytes. *Nat Commun* 2014;5:5777.
- Kim JR, Kim J, Kwon YK, Lee HY, Heslop-Harrison P, Cho KH. Reduction of complex signaling networks to a representative kernel. *Sci Signal* 2011;4:ra35.
- Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 2012;483:603–7.
- Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;27:1160–7.
- Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, Klamt S, et al. Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol Syst Biol* 2009;5:331.
- Kim J, Park SM, Cho KH. Discovery of a kernel for controlling biomolecular regulatory networks. *Sci Rep* 2013;3:2223.
- Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 1969;22:437–67.
- Gao Y, Wu H, He D, Hu X, Li Y. Downregulation of BCL11A by siRNA induces apoptosis in B lymphoma cell lines. *Biomed Rep* 2013;1:47–52.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80.
- Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
- Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 2015;1:417–25.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
- Cheang MCU, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, et al. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin Cancer Res* 2008;14:1368–76.
- Arnal JF, Lenfant F, Metivier R, Flouriot G, Henrion D, Adlanmerini M, et al. Membrane and nuclear estrogen receptor alpha actions: From tissue specificity to medical implications. *Physiol Rev* 2017;97:1045–87.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27–30.
- Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers defines their genomic and transcriptomic landscapes. *Nat Commun* 2016;7:11479.
- Welboren WJ, Stunnenberg HG, Sweep FC, Span PN. Identifying estrogen receptor target genes. *Mol Oncol* 2007;1:138–43.
- Zakaria Z, Zulkifli MF, Wan Hasan WAN, Azlah AK, Raub SHA, Eswaran J, et al. Epidermal growth factor receptor (EGFR) gene alteration and protein overexpression in Malaysian triple-negative breast cancer (TNBC) cohort. *Oncol Targets Ther* 2019;12:7749–56.
- Helikar T, Kochi N, Kowal B, Dimri M, Naramura M, Raja SM, et al. A comprehensive, multi-scale dynamical model of ErbB receptor signal transduction in human mammary epithelial cells. *PLoS One* 2013;8:e61757.
- Chaudhary S, Krishna BM, Mishra SK. A novel FOXA1/ESR1 interacting pathway: A study of Oncomine breast cancer microarrays. *Oncol Lett* 2017;14:1247–64.
- Fang SH, Chen Y, Weigel RJ. GATA-3 as a marker of hormone response in breast cancer. *J Surg Res* 2009;157:290–5.
- Witkiewicz AK, Balaji U, Knudsen ES. Systematically defining single-gene determinants of response to neoadjuvant chemotherapy reveals specific biomarkers. *Clin Cancer Res* 2014;20:4837–48.

**BCL11A-HDAC1/2 Regulate Tamoxifen Response in Breast Cancer**

39. Stoica GE, Franke TF, Moroni M, Mueller S, Morgan E, Iann MC, et al. Effect of estradiol on estrogen receptor- $\alpha$  gene expression and activity can be modulated by the ErbB2/PI 3-K/Akt pathway. *Oncogene* 2003;22:7998–8011.
40. Osborne CK, Shou J, Massarweh S, Schiff R. Crosstalk between estrogen receptor and growth factor receptor pathways as a cause for endocrine therapy resistance in breast cancer. *Clin Cancer Res* 2005;11:865s–70s.
41. Presti D, Quaquarini E. The PI3K/AKT/mTOR and CDK4/6 pathways in endocrine resistant HR+/HER2- metastatic breast cancer: biological mechanisms and new treatments. *Cancers* 2019;11:1242.
42. Citro S, Miccolo C, Meloni L, Chiocca S. PI3K/mTOR mediate mitogen-dependent HDAC1 phosphorylation in breast cancer: a novel regulation of estrogen receptor expression. *J Mol Cell Biol* 2015;7:132–42.
43. Khaled WT, Lee SC, Stingl J, Chen XF, Ali HR, Rueda OM, et al. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat Commun* 2015;6:5987.
44. Yang G, Zanudo JGT, Albert R. Target control in logical models using the domain of influence of nodes. *Front Physiol* 2018;9:454.
45. Huang S, Eichler G, Bar-Yam Y, Ingber DE. Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys Rev Lett* 2005;94:128701.
46. Joo JI, Zhou JX, Huang S, Cho KH. Determining relative dynamic stability of cell states using boolean network model. *Sci Rep* 2018;8:12077.
47. Chou CW, Wu MS, Huang WC, Chen CC. HDAC inhibition decreases the expression of EGFR in colorectal cancer cells. *PLoS One* 2011;6:e18087.
48. Zhang CH, Wang J, Zhang LX, Lu YH, Ji TH, Xu L, et al. Shikonin reduces tamoxifen resistance through long non-coding RNA uc.57. *Oncotarget* 2017;8:88658–69.
49. Muller BM, Jana L, Kasajima A, Lehmann A, Prinzler J, Budczies J, et al. Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer—overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. *BMC Cancer* 2013;13:215.
50. Bicaku E, Marchion DC, Schmitt ML, Munster PN. Selective inhibition of histone deacetylase 2 silences progesterone receptor-mediated signaling. *Cancer Res* 2008;68:1513–9.
51. Liu H, Chen S, Yao X, Li Y, Chen CH, Liu J, et al. Histone deacetylases 1 and 2 regulate the transcriptional programs of nephron progenitors and renal vesicles. *Development* 2018;145:dev153619.
52. Linares A, Dalenc F, Balaguer P, Boule N, Cavailles V. Manipulating protein acetylation in breast cancer: a promising approach in combination with hormonal therapies? *J Biomed Biotechnol* 2011;2011:856985.
53. Luc S, Huang JL, McEldoon JL, Somuncular E, Li D, Rhodes C, et al. Bcl11a deficiency leads to hematopoietic stem cell defects with an aging-like phenotype. *Cell Rep* 2016;16:3181–94.
54. Senawong T, Peterson VJ, Leid M. BCL11A-dependent recruitment of SIRT1 to a promoter template in mammalian cells results in histone deacetylation and transcriptional repression. *Arch Biochem Biophys* 2005;434:316–25.
55. Sankaran VG, Menne TF, Xu J, Akie TE, Lettre G, Van Handel B, et al. Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science* 2008;322:1839–42.
56. Bauer DE, Orkin SH. Hemoglobin switching's surprise: the versatile transcription factor BCL11A is a master repressor of fetal hemoglobin. *Curr Opin Genet Dev* 2015;33:62–70.
57. Moody RR, Lo MC, Meagher JL, Lin CC, Stevers NO, Tinsley SL, et al. Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP4/7 and the transcription factor BCL11A in epigenetic complexes. *J Biol Chem* 2018;293:2125–36.
58. Khaled WT, Lee SC, Stingl J, Chen X, Ali HR, Rueda OM, et al. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat Commun* 2015;6:5987.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## Network Analysis Identifies Regulators of Basal-Like Breast Cancer Reprogramming and Endocrine Therapy Vulnerability

Sea R. Choi, Chae Young Hwang, Jonghoon Lee, et al.

*Cancer Res* 2022;82:320-333. Published OnlineFirst November 29, 2021.

<b>Updated version</b>	Access the most recent version of this article at: doi: <a href="https://doi.org/10.1158/0008-5472.CAN-21-0621">10.1158/0008-5472.CAN-21-0621</a>
<b>Supplementary Material</b>	Access the most recent supplemental material at: <a href="http://cancerres.aacrjournals.org/content/suppl/2021/11/16/0008-5472.CAN-21-0621.DC1">http://cancerres.aacrjournals.org/content/suppl/2021/11/16/0008-5472.CAN-21-0621.DC1</a>

<b>Cited articles</b>	This article cites 58 articles, 12 of which you can access for free at: <a href="http://cancerres.aacrjournals.org/content/82/2/320.full#ref-list-1">http://cancerres.aacrjournals.org/content/82/2/320.full#ref-list-1</a>
<b>Citing articles</b>	This article has been cited by 1 HighWire-hosted articles. Access the articles at: <a href="http://cancerres.aacrjournals.org/content/82/2/320.full#related-urls">http://cancerres.aacrjournals.org/content/82/2/320.full#related-urls</a>

<b>E-mail alerts</b>	<a href="#">Sign up to receive free email-alerts</a> related to this article or journal.
<b>Reprints and Subscriptions</b>	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a> .
<b>Permissions</b>	To request permission to re-use all or part of this article, use this link <a href="http://cancerres.aacrjournals.org/content/82/2/320">http://cancerres.aacrjournals.org/content/82/2/320</a> . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.