Published in final edited form as:

Mol Cancer Res. 2014 August ; 12(8): 1140–1155. doi:10.1158/1541-7786.MCR-13-0637.

Convergent and Divergent Cellular Responses by ErbB4 Isoforms in Mammary Epithelial Cells

Vikram B. Wali^{*,#}, Jonathan W. Haskins[#], Maureen Gilmore-Hebert, James T. Platt, Zongzhi Liu, and David F. Stern

Department of Pathology, Yale School of Medicine, New Haven, Connecticut

Abstract

Associations of ErbB4 (ERBB4/HER4), the fourth member of the EGFR family, with cancer are variable, possibly as a result of structural diversity of this receptor. There are multiple structural isoforms of ERBB4 arising by alternative mRNA splicing, and a subset undergo proteolysis that releases membrane-anchored and soluble isoforms that associate with transcription factors and co-regulators to modulate transcription. In order to compare the differential and common signaling activities of full-length and soluble intracellular isoforms of ERBB4, four JM-a isoforms (full-length and soluble intracellular domain CYT-1 and CYT-2) were expressed in isogenic MCF10A cells and their biological activities were analyzed. Both full-length and intracellular domain CYT-2 promoted cell proliferation and invasion, and CYT-1 suppressed cell growth.

Transcriptional profiling revealed several new and underexplored ERBB4 regulated transcripts including: proteases/protease inhibitors (MMP3, SERPINE2), YAP/Hippo pathway (CTGF, CYR61, SPARC), mevalonate/cholesterol pathway (HMGCR, HMGCS1, LDLR, DHCR7), and cytokines (IL8, CCL20, CXCL1). Many of these transcripts were subsequently validated in a luminal, breast cancer cell line that normally express ERBB4. Furthermore, ChIP-seq experiments identified ADAP1, APOE, SPARC, STMN1, and MXD1 as novel molecular targets of ERBB4. These findings clarify the diverse biological activities of ERBB4 isoforms, and reveal new and divergent functions.

Keywords

ERBB4 isoforms; Breast cancer; Neuregulin-1; ERBB4 gene targets; MCF10A; T47D

Introduction

The four receptor kinases in the Epidermal Growth Factor (EGF) family, EGFR, ERBB2, ERBB3, and ERBB4 regulate developmental processes in the nervous system, cardiovascular system, and in epithelia. EGFR, ERBB2, and ERBB3 are common drivers in human carcinoma and glioblastoma and are targets for US FDA-approved cancer therapeutics. But, ERBB4 has a more ambiguous influence on cancer. *ERBB4* is

*Contact Information of the Corresponding Author: Vikram B. Wali, Yale University School of Medicine, 300 George Street, New Haven, CT 06511, Phone: 203-737-6491; Fax: 203-785-7232, vikram.wali@yale.edu.

#Authors contributed equally

Conflict of interest: None

overexpressed in medulloblastoma, and candidate *ERBB4* activating mutations have been identified in lung cancer, melanoma, and other cancers (1–4). Nonetheless, conflicting reports have been published on *ERBB4* as a prognostic marker, with both positive and negative clinical outcome correlations (5–7).

Inconsistent associations of *ERBB4* with cancer may be explained by the diversity of *ERBB4* regulated signaling processes enabled by mRNA splice variants. JM-a and JM-b isoforms differ in the extracellular juxtamembrane domain (8). JM-b isoforms are conventional receptor tyrosine kinases (RTKs): the ligands, including neuregulin 1 (NRG1), induce receptor phosphorylation and activate subsequent signal transduction. In contrast, JM-a isoforms have a metalloproteinase cleavage site that is clipped by TACE in response to NRG1 binding. This releases the extracellular domain (ECD), leaving the membrane-anchored m80 form. *ERBB4* m80 can then undergo intramembrane cleavage by γ -secretase to release the soluble s80 form comprising the intracellular domain (ICD). s80 relocates to mitochondria and the nucleus (9, 10), where it binds transcriptional co-regulators and transcription factors.

A second alternatively spliced region in the ICD includes (CYT-1) or excludes (CYT-2) an exon that encodes a binding site for the p85 adaptor subunit of phosphatidylinositol (3') kinase, and an overlapping WW domain PPXY binding site. Divergence of signaling processes incited by the four *ERBB4* isoforms may explain the discordance in the *ERBB4* cancer literature: most studies fail to consider these isoforms separately, and the isoform(s) expressed and subcellular localization of *ERBB4* have an impact on prognosis (11, 12).

We previously identified binding of both *ERBB4* ICD isoforms (CYT-1 and CYT-2) with the transcriptional co-repressor KAP1, and identified sixteen other candidate interactors including ubiquitin ligases ITCH and WWP2 (13). The *ERBB4* ICD has been reported by others to associate with transcription factors ER α and Stat5, with transcriptional co-regulators including YAP, WWOX, ETO2, and a TAB2/N-CoR complex, and with ubiquitin ligases Itch and Mdm2 (14–20). In order to better understand the diverse biological outcomes associated with activity of the full-length and truncated *ERBB4* isoforms, we have explored the phenotypic, transcriptional and signaling consequences of introduction and activation of *ERBB4* isoforms, and identified candidate gene target interactions by chromatin immunoprecipitation-sequencing (ChIP-seq).

Materials and methods

Cell culture

MCF10A cells were maintained in DMEM/F12 supplemented with 5% horse serum, 20 ng/ml EGF, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, 10 μ g/ml insulin, 100 units/ml penicillin and 100 μ g/ml streptomycin. MCF10A cells stably expressing full length (FL) JM-a CYT-1-*ERBB4* isoform (CYT-1 MCF10A) or JM-a CYT-2-*ERBB4* isoform (CYT-2 MCF10A) or vector only (V-MCF10A) were generated by lentiviral infection and selection with 10 μ g/ml puromycin and maintained in 1 μ g/ml puromycin. MCF10A cells stably expressing either of the ICD *ERBB4* isoforms: CYT-1 or CYT-2 were produced by lentiviral infection, selection with 10 μ g/ml blastocidin and maintenance in 7 μ g/ml

blastocidin. T47D and MDA-MB-231 cells were cultured in RPMI 1640 with glutamate (Gibco) containing 100 units/ml penicillin, 100 µg/ml streptomycin, and 10% fetal bovine serum (FBS; BioWest). FuGENE 6 (Roche) or Lipofectamine 2000 reagent (Invitrogen Corporation) were used for transfections. T47D cells were transduced with pLKO ERBB4 3'-untranslated region (UTR)-directed shRNA (Sigma, TRCN0000314628) or scrambled control and selected in 1 µg/ml puromycin. These ERBB4 knockdown (KD) T47D stable cell lines were subsequently infected with pInducer20 ERBB4 JM-a CYT-1, CYT-2, or vector control and selected in 400 µg/ml G418. T47D ERBB4 KD, pInducer20 CYT-1 or CYT-2 stable cell lines were maintained in 1 µg/ml puromycin, 200 µg/ml G418, and ERBB4 knockdown and doxycycline(DOX)-inducible ERBB4 isoform re-expression was confirmed by Western blot.

Plasmids

Lentiviral expression plasmids for JM-a FL CYT-1 ERBB4 (EX-A0212-Lv105), CYT-2 ERBB4 (EX-Z4265-Lv105) and negative control vector (EX-EGFP-Lv105) including CMV promoter followed by the ERBB4 coding sequences, puromycin selection cassette, 3'-LTR, poly adenylation sites, ampicillin cassette, pUC Ori, 5' LTR and packaging elements, were obtained from GeneCopoeia (Rockville, MD). ERBB4 plasmids were packaged as lentivirus by co-transfecting 293T cells with pLP/VSV-G, pLP1(Gag/Pol), pLP2(rev), and pcTat (tat) using Lipofectamine 2000 (Invitrogen Corporation, Carlsbad, CA). MCF10A cells were infected with a multiplicity of infection of ~5 in the presence of 8 µg/ml polybrene. Expression of ERBB4 in these MCF10A cells was tested 24h and 72h after infection. Polyclonal stable cell lines were selected with puromycin. The above FL CYT-1 and CYT-2 ERBB4 constructs were also packaged into pInducer20 DOX-inducible expression plasmids that were used to infect ERBB4 knockdown T47D stables to re-express specific JM-a ERBB4 CYT-1 or CYT-2 isoform. The pInducer20 DOX inducible expression plasmid used for cloning was generously provided by Dr. Stephen Elledge (21)

ICD expression cDNAs encoding CYT-1 (amino acids 676-1308) and CYT-2 (amino acids 676-1294) isoforms (GeneCopoeia, Rockville, MD) were cloned into the lentiviral TA cloning vector Lenti6.3-V5 in frame with the 3' V5 epitope tag (Life Sciences Technologies). Stable cell lines for the ICDs and the vector control (Vector) were selected with blasticidin.

Immunoblotting

For NRG1 stimulation, cells were plated at 1×10^6 cells/100 mm plate. The following day, cells were incubated in serum-free OptiMEM medium for 48 h, followed by incubation with 100 ng/ml NRG1. Sample buffer lysates normalized for protein concentration were analyzed by electrophoresis in 4–12% NuPAGE SDS-polyacrylamide midgels (Life Technologies Corporation). For immunoblotting, PVDF membranes were blocked with 2% BSA in 10 mM Tris-HCl, 50 mM NaCl, 0.1% Tween 20, pH 7.4 (TBST) and incubated with anti-phospho-ERBB4 (Tyr 1056) (22), phospho-ERBB4 Tyr 1284 (Cell Signaling Technology #4757), ERBB4 (sc-283), GAPDH (Santa Cruz Biotechnology, CA), phospho-MAPK (Thr202/Tyr204), or phospho-AKT (Ser473; Cell Signaling Technology, MA) diluted 1:5000 to 1:20000 in TBST/2% BSA for 2 h. Membranes were washed five times with

TBST, incubated with horseradish peroxidase-conjugated secondary antibodies in TBST/2% BSA for 1 h, rinsed with TBST, and detected by chemiluminescence (SuperSignal West Pico Chemiluminescent Substrate; Pierce, Rockford, IL).

Cell proliferation assays

Fig. 2A,B,C. Cells were plated at 1000 cells/well in 96-well plates. The next day, four wells per group were fed either serum-free OptiMEM medium or 5% horse serum containing medium, $\pm$ 100 ng/ml NRG1, and incubated for 5 days with refeeding day 2. Proliferation was assayed daily using the ATP-based CellTiter-Glo® Luminescent Cell Viability Assay (Promega). The difference among groups after 5 days was determined by analysis of variance (ANOVA) followed by Newman-Keuls multiple comparison test, with $P < 0.05$ considered to be statistically significant.

Fig. 2E,F. One thousand cells were seeded in 96-well plates in triplicate in complete medium. For scoring, cells were washed three times with PBS and then harvested with 0.25% Trypsin (Invitrogen), stained with trypan blue, and viable cells were enumerated using a cell counter (The Countess, Invitrogen).

Cell invasion

BD BioCoat Matrigel Invasion Chambers (BD Biosciences) with 8 μ m pore PET Matrigel membranes were used. Inserts were hydrated in 500 μ l of OptiMEM for 2 h and transferred to with 5% horse serum-containing medium as chemoattractant. Cells were suspended in OptiMEM with 0.1% horse serum and plated at 50,000 cells/insert in triplicate for both Matrigel and control inserts (lacking Matrigel) in 24 well plates, followed by incubation for 24h. Non-invading cells were removed by scrubbing the upper membrane surface. Invading cells on the lower surface of membrane were stained with Diff-Quick (Invitrogen Corporation) and counted in three microscopic fields per membrane. Percent invasion was calculated as the mean number of cells invading through Matrigel insert membrane/mean number of cells migrating through control insert $\times 100$. The Invasion Index is the percent of invasion of test cells relative to control vector-infected MCF10A cells. Differences among groups were determined by analysis of variance (ANOVA), followed by Newman-Keuls multiple comparison test, with $P < 0.05$ considered to be statistically significant.

Gene expression analysis

FL ERBB4 cell lines and controls were plated at 1×10^6 cells/100 mm plate and incubated in serum-free OptiMem for 48h. The next day, cells were incubated in fresh OptiMem with or without 100ng/ml NRG1 for 2h. RNA was extracted with the RNeasy Plus Mini Kit (Qiagen). For ICD ERBB4 isoforms, RNA was extracted from ICD ERBB4 cell lines maintained in complete medium. RNA samples were analyzed by the Yale Center for Genome Analysis using the Illumina HumanHT-12 v4 Expression BeadChIP (Illumina Inc), with more than 47,000 probes derived from NCBI RefSeq Release 38, and also legacy UniGene content. Both FL and ICD ERBB4 experiments were performed with two biological replicates run in parallel, with each sample analyzed in technical duplicate. The microarray data is available at Gene Expression Omnibus GEO website through accession numbers: GSE57346 (FL ERBB4 experiment) and GSE57339 (for ICD ERBB4).

experiment). The threshold for significant changes in gene expression was set at $p < 0.05$ and fold change (FC) > 1.5 . T47D pLKO ERBB4 KD, pInducer20 ERBB4 expressing stables were similarly serum-starved and treated with NRG1 (100ng/ml) for 2 hrs in the presence of 100ng/ml doxycycline (24hrs). RNA was isolated using the RNeasy Mini Plus kit (Qiagen).

Pathway and Network Analyses

Data from gene expression microarrays were analyzed through the use of IPA (Ingenuity® Systems, www.ingenuity.com). Detailed procedures for each analysis are included in the figure legends.

RNA extraction and real-time PCR

Total RNA was isolated using the RNeasy Plus Mini Kit (Qiagen) and reverse transcribed with the iScript cDNA Synthesis Kit from BioRad using 1 µg of RNA per reaction. Universal TaqMan Master Mix (Applied Biosystems) was used to conduct quantitative real-time PCR (qRT-PCR) analysis. Primers included ADAMTSL4 (Hs00417524_m1), ALDH1A3 (Hs00167476_m1), BTG2 (Hs00988887_m1), CDCA5 (Hs00293564_m1), CDC20 (Hs00426680_m1), CENPF (Hs01118845_m1), CTGF (Hs01026927_g1), CYR61 (Hs00998500_g1), DHCR7 (Hs01023087_m1), DKK1 (Hs00183740_m1), ERBB4 (Hs00171783_m1), FZD2 (Hs00361432_s1), FZD5 (Hs00361869_g1), GAPDH (Hs02758991_m1), HMGCR (Hs00168352_m1), HMGCS1 (Hs00940429_m1), KRT14 (Hs00265033_m1), LDLR (Hs01092524_m1), MMP3 (Hs00968305_m1), MMP9 (Hs00234579_m1), MSBR3 (Hs00827017_m1), MXD4 (Hs01557630_m1), PHLA1 (Hs00378285_g1), PKMYT1 (Hs00993620_m1), SERPINE2 (Hs00385730_m1), SOCS2 (Hs00919620_m1), SPARC (Hs00234160_m1), TP63 (Hs00978343_m1), TPX2 (Hs00234160_m1), TOP2A (Hs01032137_m1), WNT5A (Hs00998537_m1; Applied Biosystems). Relative mRNA expression was determined with the Ct method, with *GAPDH* as reference gene.

Chromatin Immunoprecipitation-sequencing (ChIP-Seq)

Cells were crosslinked with dimethyl 3,3'-dithiobispropionimidate (DTBP; Pierce); chromatin was extracted and sonicated to an average size of 300–500 bp; and individual ChIP assays were performed using antibodies to V5 protein tag and protein G-coupled magnetic beads (23). Sequencing libraries were produced by the Yale Center for Genome Analysis, using 15–18 cycles of amplification, gel-purified from 2% agarose gels, quantified and sequenced on an Illumina Gene Analyser II. Sequence tags (24 bp) were mapped to the human genome (hg19/NCBI Build 37) from the UCSC Genome Browser (<http://genome.ucsc.edu/>) using ELAND (24). Sequence tags were extended to 200 bp and converted to signal map files representing the integer count of mapped tags overlapping at each genomic position. The signal maps were scored using PeakSeq to identify factor binding sites (25). Statistical significance was calculated using a binomial test followed by the Benjamini-Hochberg correction for multiple hypothesis testing to yield a q-value for each candidate region. High confidence bound regions were selected with a q-value cutoff of 0.01, corresponding to an overall FDR of 1%. A q-value cutoff of 0.05 was also used to identify a set of lower confidence regions.

Confirmation of ChIP-Seq Targets

Standard ChIP (26) was performed on 5×10^9 CYT-1 and CYT-2 ICD cells cross-linked with DTBP. Nuclear extracts were divided into three aliquots and precipitated with anti-IgG, anti-histone, (Cell Signaling Technologies) or anti-V5 tag (Invitrogen) antibodies using Protein G magnetic beads. The beads were washed extensively and the antibody complexes were released from the beads. The DNA crosslinks were reversed and DNA was purified using a Qiagen PCR Kit. The DNA from each antibody reaction was used in quantitative PCR. The primers used were:

APOE_F-1, Sequence: GCT ATC TTC CCA TCC GGA AC
 APOE_R-1, Sequence: CAT CTC TGC TGC TGC AGT CT
 SPARC_F-1, Sequence: CAG AGC TCC ACA GAA TGC AG
 SPARC_R-1, Sequence: CAC CCG TCT CTT CTT CTC GA
 STMN1_F-1 Sequence: TCC CAA AGT GCT GGG ATT AG
 STMN1_R-1 Sequence: GCA GGG TGC TGT CTT TGT CT
 ADAP1_F-1 Sequence: AAC ACT ACT GCC CGA TGG TC
 ADAP1_R-1 Sequence: CAG GTG CCA TCT CTT GAG G
 MXD4_F1 Sequence: TTT ACA GCC CAG GAA ACA GG
 MXD4_R1 Sequence: GGC AGG TTC TAG GTC AGT GG

Three biologically independent experiments were done for each ChIP. Binding for each target sequence was calculated as percentage of input binding to that sequence.

Results

ERBB4 has unusually broad signaling potential for a RTK owing to its atypical nuclear functions, and the diversification of ICD isoforms by the CYT-1/CYT-2 splice choice. To compare the functionality of full-length (FL) versus ICD, and CYT-1 versus CYT-2 isoforms, we produced stable cell lines overexpressing different ERBB4 isoforms. MCF10A cells were used first because they express little or no endogenous ERBB4, and because they are a normal-like, non-transformed human mammary cell line.

Expression of Full-length ERBB4 in MCF10A

We first engineered MCF10A cells to express full length CYT-1 and CYT-2 ERBB4 with the cleavable JM-a domain. DNA-mediated gene transfer yielded robust expression of ERBB4 that was stable in T47D cells, which possess endogenous ERBB4. However, ERBB4 expression was lost within 3 days of DNA-mediated gene transfer in MCF10A cells, consistent with a counter-selection against high ERBB4 in MCF10A background (Fig. 1A). Nonetheless, expression of FL ERBB4 was stable in MCF10A cells 3 days after infection with lentivirus (which integrates efficiently) and after puromycin selection. Engineered ERBB4 mRNA expression was higher than endogenous expression in T47D cells (Figure 1B). Although mRNA levels for FL CYT-1 and CYT-2 were comparable (Fig 1B), CYT-2

ERBB4 protein was slightly higher at steady state, and this difference was augmented by stimulation with NRG1 (Figure 1C). This is consistent with the reports that CYT-2 protein is more stable, since it lacks the ubiquitin-ligase binding site present in CYT-1 (27, 28).

NRG1-induced phosphorylation of ERBB4 was greater and more sustained in cells expressing CYT-2 than CYT-1. Vector cells responded with increased MAPK phosphorylation (Thr202/Tyr204) and AKT phosphorylation (Ser473), presumably through activation of endogenous ERBB3. Expression of CYT-1 or CYT-2 enhanced the NRG1 response (Figure 1C) with greater phosphorylation of MAPK in CYT-2. There was little or no difference between control and ERBB4-expressing MCF10A cell lines in relative protein levels of cyclin D1, E-cadherin, vimentin, phospho-YAP(Tyr357), or phospho-YAP(Ser127) over a 24 h period (data not shown).

Biological activities of FL isoforms

Proliferation rates were similar in FL CYT-1, CYT-2, and control cell lines grown in 5% horse serum, with or without addition of NRG1 (Figure 2A). However, in serum-free medium, FL CYT-2 cells grew significantly faster than vector control or FL CYT-1 cells (Figure 2B), both in the absence and presence of NRG1 ($P<0.001$). With NRG1, FL CYT-1 cells grew significantly more slowly than vector cells ($P<0.01$) after 5 days (Figure 2B), so CYT-1 ERBB4 actively reduces growth of these cells. In Boyden chamber Matrigel invasion assays, FL CYT-2 MCF10A cells scored significantly higher ($24.30\pm1.67\%$; $P<0.05$) than both control Vector ($13.1\pm1.70\%$) and FL CYT-1 MCF10A cells ($15.33\pm2.50\%$) (Fig 2C). The invasion index of FL CYT-2 MCF10A cells (1.84) was nearly twice that of vector control cells (1.0), but similar for FL CYT-1 MCF10A and control cells (1.1). Overall, expression of FL CYT-1 reduces proliferation in the absence of serum and the presence of NRG1, whereas FL CYT-2 promotes proliferation and invasion, concordant with somewhat higher relative MAPK signaling.

ERBB4 intracellular domains

The NRG1-activated outputs of the FL JM-a isoforms are a composite of conventional RTK signaling and the activities of the m80 and s80 forms released by cleavage. To evaluate the signaling activities of soluble ICD isoforms, we expressed V5-tagged constructs beginning just beyond the basic residues marking the cytoplasmic face of the transmembrane domain (Fig 2D). These are structurally similar to the forms produced by γ -secretase cleavage of ERBB4, but the actual amino terminus of ERBB4 s80 has not been determined by peptide sequencing. We have reported that these ICDs are Tyr-phosphorylated, concordant with earlier findings (29). Similar to cells expressing FL ERBB4, ICD CYT-2 MCF-10A lines proliferate more rapidly than the Vector or ICD CYT-1 cell lines, whereas ICD CYT-1 cell lines were growth-suppressed compared to the ICD CYT-2 lines or control lines in medium containing 5% horse serum (Fig 2E). The ICD CYT-2 cells invaded through a Matrigel membrane more efficiently ($70\% \pm 1.8$) than the vector or ICD CYT-1 lines, which had low invasive potential.

Transcriptional profiling of FL ERBB4

We used transcriptional profiling to identify the genes commonly and differentially affected by expression of the two FL ERBB4 isoforms. Non-supervised hierarchical clustering of the top genes up-or down-regulated by ERBB4, based on all features with a SD/mean>0.05 in biological replicates, grouped Vector (VEC), CYT-1/CYT-2 without NRG1 and CYT-1/CYT-2 with NRG1 (data not shown). Under serum-free conditions, stimulation with NRG1 (100ng/ml) for 2 hours did not significantly change expression of genes in Vector-MCF10A cells but significantly altered gene expression in FL CYT-1 and FL CYT-2 MCF10A cell lines (Supplementary Figure 2A, left; Supplementary Table 1). Without NRG1, expression of 61 genes in CYT-1 MCF10A and 143 genes in CYT-2 MCF10A cells were significantly altered relative to Vector control cells (adjusted p-value<0.05; fold change (FC)>1.5). Of these, 52 genes were altered in both CYT-1 and CYT-2 cells (Supplementary Figure 2A, center). With NRG1 stimulation, 37 genes in CYT-1 and 57 genes in CYT-2 MCF10A cells were significantly different, out of which 24 were commonly affected in both, in comparison to NRG1-stimulated Vector-MCF10A cells (Supplementary Figure 2A, right). Overall, genes encoding cytokines IL8, CCL20, CXCL1, matrix metalloproteinase MMP3, protease inhibitor SERPINE2, signaling adaptor VAV3, CTGF, and phosphorylation-dependent ubiquitin ligase adaptor FBX032, were the top genes exclusively altered in CYT-2 MCF10A, whereas poly A binding protein PABPC1 and secreted serine protease kallikrein-10 (KLK10) were exclusively altered in CYT-1 MCF10A cells (Supplementary Table 1). qRT-PCR assays validated 10 out of 11 candidate transcriptional changes in this group, including *SERPINE2* and *MMP3* (Supplementary Figure 1A). Some of the genes shared between FL ERBB4 CYT-1 and CYT-2 are listed in Table 1.

Transcriptional profiling of ICD ERBB4

There were many transcriptional differences among MCF10A ICD CYT-1 or CYT-2 and vector control cells (Supplementary Figure 2B). 918 genes were altered in common between the ICD CYT-1 and ICD CYT-2 cells, with 482 ICD CYT-1 specific genes and 168 ICD CYT-2 specific genes (Supplementary Figure 2B). Some of the shared top upregulated (13/20) and downregulated (16/20) genes are listed (Table 2). A subset of genes was validated by qRT-PCR (Supplementary Figure 1B).

FL and ICD ERBB4 differ in ligand-dependence (ICD are constitutively Tyr phosphorylated) (13, 29), subcellular localization, and stability. Overall, NRG1-stimulated cells expressing FL ERBB4 (CYT-1 and CYT-2) altered fewer pathways than the ICD, and with less significance. Endogenous ERBB3-regulated genes were factored out of our analysis by comparison against NRG1-treated controls. In comparison to the respective vector controls, genes encoding transforming growth factor alpha (TGF α), angiopoietin-like 4 (ANGPTL4), plasminogen activator inhibitor SERPINE1, and signaling protein PHLDA1 were among the top commonly affected genes in both FL and ICD, CYT-1 and CYT2 ERBB4 expressing MCF10A cells (Supplementary Tables 1 and 2). HSP70-encoding *HSPA1A* was the only gene that was altered exclusively in CYT-1 isoform in both FL and ICD forms, consistent with a stress response. Similarly, genes including chemokine receptor *CXCL1*, growth-activating *MYC* and *FOSB*, and neural receptor/adhesion protein *CNTNAP2* were altered uniquely in cells expressing FL and ICD CYT-2.

IPA pathway analysis

We used Ingenuity Systems® IPA analysis to evaluate pathways altered by ICD ERBB4 expression. We chose to focus on ERBB4 ICD since the transcriptional changes elicited by the ICD were both stronger and more statistically significant than FL ERBB4, thereby facilitating a more robust analysis. IPA analysis tests for over-representation of genes in a particular annotated process in order to infer altered pathways from patterns of gene expression (Supplementary Figure 3). Interestingly, both CYT-1 and CYT-2 upregulated genes in cholesterol biosynthesis (mevalonate pathway) and ketogenesis. ICD CYT-1 was uniquely linked to increased colorectal cancer metastasis signaling and pancreatic carcinoma signaling and epithelial-to-mesenchymal transition (EMT) associated genes. ERBB4 CYT-2 ICD uniquely affected cyclins/cell cycle regulation and DNA damage-induced 14-3-3 signaling through increases in cyclins and CDK1. ERBB4 ICD CYT-2 and ERBB4 FL CYT-2 both regulated pathways involved in cell cycle and cell cycle regulation (Supplementary Figure 3).

IPA analysis was used to test for overlap of ICD ERBB4-induced genes with genes grouped according to functional annotation. ERBB4 ICD CYT-1 was connected with increases in processes including regulation of microtubules, cytoskeleton, metastasis, neoplasia, cell death, and tumorigenesis, and decreases in skin abnormality, and cell cycle, and ploidy processes (Table 3, left, Supplementary Table 4). ERBB4 ICD CYT-2 was associated with increases in proliferation, angiogenesis, blood vessel development, cardiovascular development, M phase, cytostasis, and malignant tumor growth, and decreases in ploidy. The genes comprising each pathway are listed in Supplementary Tables 2 and 4.

Prediction of transcriptional regulators

Ingenuity® Upstream Regulator Analysis (Table 3, right, and Supplementary Table 3) predicted several transcription factors to be activated in both ERBB4 CYT-1 ICD and CYT-2 ICD HIF1 α , SREBF1, MTPN, E2F1, FOXO1, SREBF2, and NF κ BI. Network analysis of gene expression involving predicted transcriptional regulators suggested activation of YAP/TAZ (Hippo pathway), HIF1 α , and TGF β in both ERBB4 CYT-1 ICD and CYT-2 ICD (Figure 3). ERBB4 CYT-1 ICD was associated with factors including components of the NF κ B pathway, epithelial-to-mesenchymal transition (EMT), and TGF β . ERBB4 CYT-2 ICD was predicted to repress negative regulators of the cell cycle -CDKN2A and RB1 (Figure 3).

ERBB4 ICD ChIP-Seq

ERBB4 is unusual among RTKs in the ability of the soluble ICD to associate with transcription factors and co-regulators and directly participate in gene regulation. Only a small number of nuclear ERBB4 regulatory targets have been identified, so we performed ChIP-seq experiments with ERBB4 ICD CYT-1 to identify the global set of DNA sequences to which ERBB4 binds. Since these binding interactions are likely to be indirect, both protein/protein and protein/nucleic acid crosslinking agents were used prior to shearing and ChIP. Two biological repeats were conducted for the ICD CYT-1. Although there were many fewer reads in the repeat experiment, 236 immunoprecipitated DNA segments were

assigned to the same gene between the two sets, and ninety-four of these were within the same annotated feature (data not shown).

Candidate binding sites based on ChIP-Seq were found in intergenic regions, introns, and promoters. The intergenic regions were most numerous, but they are of uncertain significance since we have not yet attempted to validate them by ChIP. As we were most interested in binding targets with functional impact on gene transcription, initial validation experiments were confined to segments of DNA within 20 kb 5' of the transcriptional start site (TSS) of genes that overlapped with genes altered in RNA profiling (Supplementary Table 5). Of the ten genes that were tested, SPARC, SERPINE 1, and MXD4 were upregulated ICD CYT-1 specific genes, and STMN1 was an ICD CYT-2 specific gene (Supplementary Figure 1). ADAP1/CENTA1, SOCS2 and HS3ST1 were downregulated by both ICD isoforms while APOE, CDKN2AIPNL AND TK1 were upregulated by both. In ChIP validation experiments in cells expressing CYT-1 and CYT-2, no ERBB4 enrichment was seen over the histone and IgG controls for SERPINE1, ANGPTL4, CDKN2AIPNL, HS3ST1 and TK1 (data not shown). Specific ChIP of ERBB4 was detected for each of the other sites analyzed (Figure 4). ERBB4 ChIP enriched for ADAP1/CENTA1 and APOE equivalently from cells expressing CYT-1 ICD and CYT-2 ICD, consistent with the upregulation of APOE and downregulation of ADAP1/CENTA1 in CYT-1 and CYT-2 ICD lines. SPARC and STMN1 are preferentially upregulated in CYT-1 and CYT-2 cells, respectively, but were also evenly enriched by ChIP from cells expressing either of the isoforms. Finally, MXD4 was preferentially enriched by ERBB4 ChIP from CYT1 ICD cells, mirroring CYT-1 preferential transcriptional upregulation.

ERBB4 is not known to be a sequence-specific DNA binding protein, but instead modulates transcription through association with DNA binding proteins and transcriptional coregulators. A search for predicted transcription factor binding sites and transcription factor ChIP-defined sites (Human Genome Build (GRCh37/hg19) within our validated ChIP-Seq targets identified several sites of interest. The ADAP1 ERBB4 binding site includes a CEBPB site found in four of four cell lines tested by CEBPB ChIP-Seq (30). Within the SPARC ERBB4 binding site are BRACH, ATF6, ATF and XBP1 predicted transcription factor binding sites (31). Within the STMN1 ERBB4 site is a TCF4 ChIP-Seq site that was validated in one of the two cell lines tested. APOE ERBB4 site are both HNF4 and ETS1 sites that were found by transcription factor ChIP-Seq in K562 cells, and the MAD ERBB4 binding site includes ChIP-Seq-defined sites for MYC and CTCF in K562 cells. Overall, these results identify five new gene candidates for direct regulation through ERBB4 ICD, and a number of other candidates remain to be tested (Supplementary Table 5). Of these, both ADAP1 and STMN1 have been implicated in intracellular signaling and microtubule regulation, APOE in lipoprotein metabolism, and SPARC, in growth and calcium regulation in the extracellular matrix. *MXD4* encodes a MAD that antagonizes MYC. MXD4 may contribute to growth suppression in CYT-1 ICD cells, where it was preferentially expressed in validation experiments (Supplementary Figure 1B).

ERBB4-dependent mRNA in luminal T47D mammary cancer cells

We initially chose to analyze different ERBB4 isoforms expressed in MCF10 cells, as they do not express significant endogenous ERBB4. These cells have a basal phenotype. We next evaluated ERBB4-induced transcription in a more biologically relevant context, T47D cells, which express endogenous ERBB4 and which have a luminal phenotype. We knocked down endogenous ERBB4 and re-introduced vectors encoding specific ERBB4 isoforms (Figure 5A). Knockdown by 3'-UTR specific shRNA resulted in >50% reduction in ERBB4 protein levels in T47D ERBB4 KD stable cell lines, but NRG1-induced ERBB4 Tyr phosphorylation (B4 sh3, pI20 V) was maintained and possibly slightly increased in a possible compensatory circuit. Introduction of vectors expressing FL ERBB4 CYT-1 (B4 sh3, pI20 C1) or CYT-2 (B4 sh3, pI20 C2) resulted in greater expression of ERBB4 protein (Figure 5A), higher basal ERBB4 phosphorylation, and greater NRG1-induced ERBB4 phosphorylation.

We determined whether ERBB4-regulated genes identified in MCF10A background were similarly regulated in T47D cell lines. The genes most responsive to NRG1 or ERBB4 expression in T47D cells included *CTGF*, *CYR61*, *DKK1*, *LDLR*, *SPARC*, *HMGCR*, *HMGCS1*, *TP63*, and *KRT14*, all of which were upregulated except for *SPARC*, which was reduced with ERBB4 expression (Figure 5B). Most followed the same trend seen in MCF10A cells with the exception of *SPARC*, which was up in MCF10A cells and down in T47D cells, and *TP63* and *SOC2*, which were both reduced in MCF10A cells and up in T47D cells. Additionally, *PHLDA1*, *TOP2A*, *SOC2*, and *SERPINE2* increased with NRG1/ERBB4 in the T47D cells (Supplemental Figure 1C). Of the top genes altered in MCF10A ERBB4 ICD microarray analysis, *CTGF*, *CYR61*, *PHLDA1*, *SERPINE2*, and *TOP2A* all validated, but isoform-specific effects were not as strong as with the ICD. The ERBB4-altered genes overlapping between MCF10A and T47D cells mainly grouped into the mevalonate/cholesterol pathway (*HMGCR*, *HMGCS1*, *LDLR*, *DHCR7*) and the YAP/Hippo pathway (*CTGF*, *CYR61*, *SPARC*) in addition to luminal/basal markers (*KRT14*, *TP63*), *DKK1* (Wnt negative regulator), and *PHLDA1*, an important negative regulator and effector of Aurora A kinase in breast cancer. (Figure 5B).

Since the cholesterol pathway genes *HMGCR* and *LDLR* were among the novel and prominent genes altered in both MCF10A and T47D cells by ERBB4 and its ligand, NRG1, we investigated their expression in pregnant and lactating mouse mammary glands using existing gene expression data (32). *ErbB4* and its ligands are highly expressed and necessary for normal mammary development (33, 34). Intriguingly, *Hmgcr* and *Ldlr* expression is elevated during late pregnancy and early lactation along with *ErbB4* and several of its ligands (Figure 5C).

Discussion

We report here the first direct comparison of the four JM-a FL and artificially truncated CYT-1 and CYT-2 ERBB4 isoforms in an isogenic background. In MCF10A cells, which lack endogenous ERBB4, CYT-1 ERBB4 suppresses growth whereas CYT-2 ERBB4 increases cell proliferation and promotes invasion. Transcriptional profiling revealed genes that are commonly altered by multiple ERBB4 isoforms and also genes that are uniquely

affected by each ERBB4 isoform. ERBB4 was knocked down in T47D cells, and ERBB4 isoforms re-expressed to confirm functionally relevant genes in a luminal-like cell background where ERBB4 is normally expressed. Novel ERBB4 ICD DNA binding regions and candidate ERBB4 target genes were identified by ChIP-Seq.

Our findings of growth inhibitory potential of CYT-1 and growth promoting potential of CYT-2 are consistent with studies in HC11 mouse mammary epithelial cells expressing s80 (35), and with comparisons of MCF7 cells expressing FL JM-a CYT-1 and CYT-2 (36) and MCF10A cells expressing CYT1 s80 (37). Moreover, NR6 mouse fibroblasts expressing JM-a CYT-2 ErbB4 showed enhanced growth compared to JM-b CYT-2, consistent with a potent role for the intracellular cleaved forms in growth promotion by ERBB4 (38). Although ERBB4 is involved in mammary differentiation, (33, 34, 39), we did not find that NRG1-activated FL ERBB4 promotes differentiation or EMT. This may be because MCF10A cells express little or no prolactin receptor or Stat5, both of which regulate mammary differentiation cooperatively with ERBB4 (33, 40). Similarities in biological activities of FL and ICD isoforms suggest that the dominant signaling output of ERBB4 JM-a is mediated by the s80 cleavage products, and implies that ERBB4 JM-a will have very different signaling qualities from JM-b, which is not cleaved.

Major categories of genes regulated by ERBB4 encode proteases/protease inhibitors (MMP3, SERPINE2, KLK-10), YAP/Hippo pathway targets (CTGF, CYR61, SPARC), mevalonate/cholesterol pathway genes (HMGCR, HMGCS1, LDLR, DHCR7), and cytokines (IL8, CCL20, CXCL1). In FL ERBB4 cell lines, a cleaved ERBB4 band was detected even in the absence of NRG1, indicating some basal ERBB4 activity. In comparison of nonstimulated ERBB4 MCF10A versus Vector-MCF10A, genes encoding Rho-GEF VAV3, and protease inhibitor SERPINE2 were upregulated by CYT-2. VAV3, along with VAV2, control a lung metastasis-specific transcriptional program in breast cancer (41), and VAV2 is regulated in hippocampal neurons by NRG1-ERBB4 signaling (42). SERPINE2 is upregulated by oncogenic activation of the MAPK pathway and has been proposed as a therapeutic target in colorectal cancer. Top NRG1-dependent genes unique to CYT-2 cells include *MMP3*, *SERPINE2*, *IL8*, *CCL20* (upregulated) and *FBX032* (downregulated). The invasive properties of MMP3, upregulation of SERPINE2 by MAPK, and pro-angiogenic disposition of VAV-3 agree with the higher MAPK activity, pro-proliferative, and invasive behavior of CYT-2 cells. Activation of ERBB4 by cytokines is well-documented (43, 44). We report a converse upregulation of cytokines IL8 and CCL20 by CYT-2. FBX032 is a novel TGF- β /SMAD4 target gene and a tumor suppressor. Thus, modulation of proteases, cytokines and TGF- β pathways by CYT-2 ERBB4 may contribute to the highly proliferative and invasive phenotype of these cells.

We found that the free ERBB4 ICD is a much more potent transcriptional regulator than FL ERBB4, which may be associated with its constitutive Tyr phosphorylation and chronic signaling activity. There was overlap of predicted pathways in ERBB4 CYT-2 ICD and FL CYT-2 but no overlap for unique CYT-1 ERBB4 ICD and FL. CYT-1 contains a binding site p85 PI3K, and the PI(3') kinase requires membrane localization for signaling, so CYT-1 may have unique signaling roles at the membrane that are bypassed following release of the

s80 ICD. The overlapping ERBB4 CYT-2 pathways were cell cycle related, consistent with the fact that ERBB4 CYT-2 increases proliferation in MCF10A cells.

The predicted upstream transcription factors activated or repressed by ERBB4 ICD (Figure 3) include YAP and HIF1- α both known to bind ERBB4 (45, 46). TWIST1 and SNAI1 were predicted as active only in ERBB4 CYT-1 ICD, raising the possibility that CYT1 has a role in cancer progression through EMT, consistent with the pathway analysis (Table 3). In contrast, many of the predicted CYT-2 ICD pathways were connected with cell cycle and proliferation (Table 3), and ERBB4 CYT-2 ICD gene expression predicted repression of CDKN2A and RB1 (Figure 3), critical negative regulators of cell cycle progression. This prediction is reinforced by the increase in proliferation and M phase associated genes (Table 3) as well as the finding that ERBB4 ICD CYT-2 increases proliferation of MCF10A cells (Fig 2). These findings identify candidate mediators for the differential growth regulation by CYT-1 versus CYT-2, which is not explained by major differences in MAPK or PI3K/AKT pathway signaling (Fig. 1C).

We were surprised to find that ERBB4 upregulated several enzymes in the mevalonate/cholesterol pathway, as this is a novel function of ERBB4. The upregulation of HMG-CoA reductase (HMGCR), HMG-CoA synthase (HMGCS1), 7-dehydrocholesterol reductase (DHCR7), and low-density lipoprotein receptor (LDLR) could result from either direct or indirect transcriptional regulation by ERBB4. Interestingly, sterol regulatory element-binding protein 1 and 2 (SREBF1 and SREBF2), major transcriptional regulators of enzymes critical for sterol biosynthesis including LDLR, HMGCS1, and HMGCR were two of the top predicted activated transcription factors in the IPA analysis. ChIP-seq identified ERBB4 binding very close to the TSS (62-776 bp) of SREBF1 and SREBF2, although expression of these genes was not significantly altered in the MCF10A ERBB4 microarrays. DHCR7 was also identified in the CYT-1 ERBB4 ChIP-seq experiment at ~70kb from the putative TSS. Additionally, APOE, another component of cholesterol metabolism, was transcriptionally upregulated in MCF10A ERBB4 ICD cells and was validated from the ERBB4-ChIP. HMGCR and HMGCS1 are regulated by PPAR α (47), whose activity is modulated by NCOR1, known to form a nuclear complex with ERBB4 (17, 48). Collectively, these data lead us to hypothesize that ERBB4 interacts with SREBP1/2 to directly regulate expression of mevalonate/cholesterol genes and we are currently exploring this mechanism.

In addition to expressing ERBB4 isoforms in MCF10A cells which otherwise lack ErbB4 expression, we also reduced endogenous ERBB4 in luminal T47D cells and then re-expressed specific FL ERBB4 CYT-1 or CYT-2 isoforms. Expression analysis revealed that, similar to MCF10A cells, genes regulating cholesterol and Hippo pathway genes are also significantly altered by ERBB4 expression in T47D cells. Notable Hippo pathway genes include *CTGF*, *CYR61*, and *SPARC*, all of which are YAP/TEAD regulated genes (49). We are currently investigating the mechanism and biological consequences of the ERBB4-YAP biochemical interaction reported previously (46). In T47D cells, ERBB4 also upregulated *HMGCR*, *HMGCS1*, *LDLR* genes, which are chief regulators of the cholesterol pathway. As ERBB4 plays a critical role in mammary gland development during pregnancy and lactation (33, 34), and the temporal pattern of *Hmgcr* and *Ldlr* expression is similar to that of *ErbB4*

and its ligands in mouse mammary glands, there is a potential functional relationship between ERBB4 and the mevalonate pathway. ERBB4 regulates milk proteins through activation of mammary differentiation factor STAT5, so we hypothesize based on these findings that ERBB4 coordinately regulates cholesterol synthesis as another nutritional component of milk. Indeed, cholesterol is synthesized locally in mammary glands in addition to the liver (50). These findings in normal basal-like MCF10A and cancerous luminal T47D cells are notable in light of the recent finding that cholesterol promotes breast cancer growth and metastasis (51), and we further speculate that ERBB4 regulation of cholesterol metabolism, normally occurring in mammary development, could be hijacked during tumorigenesis or cancer progression. It was recently shown that YAP/TAZ activity can be controlled by the SREBP/mevalonate pathway (52), raising the possibility that ERBB4 could be a RTK facilitating mevalonate/Hippo signaling.

A major goal of this work was to identify common and differential cellular responses associated with ERBB4 signaling in mammary background, and to find candidate pathways and protein mediators of these responses. This work underscores the diverging phenotypes of CYT-1 and CYT-2 isoforms despite related responses linked to activation of Hippo and HIF1- α pathway genes. Differential ERBB4 isoform-dependent changes implicate cytokines, growth factors/mitotic cell cycle regulators, and extracellular matrix mediators. The linkage of ERBB4 with cholesterol metabolism, intracellular cytoskeletal regulators, and novel candidate target genes detected by ChIP, as well as TGF β and NF κ B pathways significantly extends the universe of potential processes connected with nuclear signaling by ERBB4. Our findings implicate ERBB4 as a coordinate regulator of growth through the hippo pathway. The possibility that ERBB4 regulates cholesterol metabolism may have important implications for milk production and, more generally, anabolic processes in normal epithelia and cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by USPHS grant R01 CA80065 from the National Cancer Institute and JWH by NIH training grant T32GM07223.

References

1. Prickett TD, Agrawal NS, Wei X, Yates KE, Lin JC, Wunderlich JR, et al. Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. *Nat Genet.* 2009; 41:1127–32. [PubMed: 19718025]
2. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature.* 2008; 455:1069–75. [PubMed: 18948947]
3. Gilbertson R, Hernan R, Pietsch T, Pinto L, Scotting P, Allibone R, et al. Novel ERBB4 juxtamembrane splice variants are frequently expressed in childhood medulloblastoma. *Genes Chromosomes Cancer.* 2001; 31:288–94. [PubMed: 11391800]
4. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013; 502:333–9. [PubMed: 24132290]

5. Hollmen M, Elenius K. Potential of ErbB4 antibodies for cancer therapy. *Future Oncol.* 2010; 6:37–53. [PubMed: 20021208]
6. Knowlden JM, Gee JM, Seery LT, Farrow L, Gullick WJ, Ellis IO, et al. c-erbB3 and c-erbB4 expression is a feature of the endocrine responsive phenotype in clinical breast cancer. *Oncogene.* 1998; 17:1949–57. [PubMed: 9788438]
7. Barnes NL, Khavari S, Boland GP, Cramer A, Knox WF, Bundred NJ. Absence of HER4 expression predicts recurrence of ductal carcinoma in situ of the breast. *Clin Cancer Res.* 2005; 11:2163–8. [PubMed: 15788662]
8. Elenius K, Corfas G, Paul S, Choi CJ, Rio C, Plowman GD, et al. A novel juxtamembrane domain isoform of HER4/ErbB4. Isoform-specific tissue distribution and differential processing in response to phorbol ester. *J Biol Chem.* 1997; 272:26761–8. [PubMed: 9334263]
9. Ni CY, Murphy MP, Golde TE, Carpenter G. gamma -Secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. *Science.* 2001; 294:2179–81. [PubMed: 11679632]
10. Naresh A, Long W, Vidal GA, Wimley WC, Marrero L, Sartor CI, et al. The ERBB4/HER4 intracellular domain 4ICD is a BH3-only protein promoting apoptosis of breast cancer cells. *Cancer Res.* 2006; 66:6412–20. [PubMed: 16778220]
11. Machleidt A, Buchholz S, Diermeier-Daucher S, Zeman F, Ortmann O, Brockhoff G. The prognostic value of Her4 receptor isoform expression in triple-negative and Her2 positive breast cancer patients. *BMC Cancer.* 2013; 13:437. [PubMed: 24063248]
12. Veikkolainen V, Vaparanta K, Halkilahti K, Iljin K, Sundvall M, Elenius K. Function of ERBB4 is determined by alternative splicing. *Cell Cycle.* 2011; 10:2647–57. [PubMed: 21811097]
13. Gilmore-Hebert M, Ramabhadran R, Stern DF. Interactions of ErbB4 and Kap1 connect the growth factor and DNA damage response pathways. *Mol Cancer Res.* 2010; 8:1388–98. [PubMed: 20858735]
14. Wong J, Weickert CS. Transcriptional interaction of an estrogen receptor splice variant and ErbB4 suggests convergence in gene susceptibility pathways in schizophrenia. *J Biol Chem.* 2009; 284:18824–32. [PubMed: 19439407]
15. Williams CC, Allison JG, Vidal GA, Burow ME, Beckman BS, Marrero L, et al. The ERBB4/HER4 receptor tyrosine kinase regulates gene expression by functioning as a STAT5A nuclear chaperone. *J Cell Biol.* 2004; 167:469–78. [PubMed: 15534001]
16. Aqeilan RI, Donati V, Palamarchuk A, Trapasso F, Kaou M, Pekarsky Y, et al. WW domain-containing proteins, WWOX and YAP, compete for interaction with ErbB-4 and modulate its transcriptional function. *Cancer Res.* 2005; 65:6764–72. [PubMed: 16061658]
17. Sardi SP, Murtie J, Koirala S, Patten BA, Corfas G. Presenilin-dependent ErbB4 nuclear signaling regulates the timing of astrogenesis in the developing brain. *Cell.* 2006; 127:185–97. [PubMed: 17018285]
18. Linggi B, Carpenter G. ErbB-4 s80 intracellular domain abrogates ETO2-dependent transcriptional repression. *J Biol Chem.* 2006; 281:25373–80. [PubMed: 16815842]
19. Omerovic J, Santangelo L, Puggioni EM, Marrocco J, Dall'Armi C, Palumbo C, et al. The E3 ligase Aip4/Itch ubiquitinates and targets ErbB-4 for degradation. *FASEB J.* 2007; 21:2849–62. [PubMed: 17463226]
20. Arasada RR, Carpenter G. Secretase-dependent tyrosine phosphorylation of Mdm2 by the ErbB-4 intracellular domain fragment. *J Biol Chem.* 2005; 280:30783–7. [PubMed: 15985438]
21. Meerbrey KL, Hu G, Kessler JD, Roarty K, Li MZ, Fang JE, et al. The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc Natl Acad Sci U S A.* 2011; 108:3665–70. [PubMed: 21307310]
22. Gallo RM, Riese DJ 2nd. The antibody sc-33040-R fails to specifically recognize phosphorylation of ErbB4 on tyrosine1056. *Growth Factors.* 2007; 25:329–33. [PubMed: 18236211]
23. O'Geen H, Fietze S, Farnham PJ. Using ChIP-seq technology to identify targets of zinc finger transcription factors. *Methods Mol Biol.* 2010; 649:437–55. [PubMed: 20680851]
24. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res.* 2002; 12:996–1006. [PubMed: 12045153]

25. Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, et al. PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat Biotechnol.* 2009; 27:66–75. [PubMed: 19122651]
26. O'Geen H, Nicolet CM, Blahnik K, Green R, Farnham PJ. Comparison of sample preparation methods for ChIP-chip assays. *Biotechniques.* 2006; 41:577–80. [PubMed: 17140114]
27. Sundvall M, Peri L, Maatta JA, Tvorogov D, Paatero I, Savisalo M, et al. Differential nuclear localization and kinase activity of alternative ErbB4 intracellular domains. *Oncogene.* 2007; 26:6905–14. [PubMed: 17486069]
28. Meijer IM, van Rotterdam W, van Zoelen EJ, van Leeuwen JE. Cbl and Itch binding sites in ERBB4 CYT-1 and CYT-2 mediate K48- and K63-polyubiquitination, respectively. *Cell Signal.* 2013; 25:470–8. [PubMed: 23153581]
29. Linggi B, Cheng QC, Rao AR, Carpenter G. The ErbB-4 s80 intracellular domain is a constitutively active tyrosine kinase. *Oncogene.* 2006; 25:160–3. [PubMed: 16170367]
30. Rosenbloom KR, Dreszer TR, Long JC, Malladi VS, Sloan CA, Raney BJ, et al. ENCODE whole-genome data in the UCSC Genome Browser: update 2012. *Nucleic Acids Res.* 2012; 40:D912–7. [PubMed: 22075998]
31. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, et al. The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res.* 2013; 41:D64–9. [PubMed: 23155063]
32. Anderson SM, Rudolph MC, McManaman JL, Neville MC. Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Res.* 2007; 9:204. [PubMed: 17338830]
33. Jones FE, Welte T, Fu XY, Stern DF. ErbB4 signaling in the mammary gland is required for lobuloalveolar development and Stat5 activation during lactation. *J Cell Biol.* 1999; 147:77–88. [PubMed: 10508857]
34. Tidcombe H, Jackson-Fisher A, Mathers K, Stern DF, Gassmann M, Golding JP. Neural and mammary gland defects in ErbB4 knockout mice genetically rescued from embryonic lethality. *Proc Natl Acad Sci U S A.* 2003; 100:8281–6. [PubMed: 12824469]
35. Muraoka-Cook RS, Sandahl MA, Strunk KE, Miraglia LC, Husted C, Hunter DM, et al. ErbB4 splice variants Cyt1 and Cyt2 differ by 16 amino acids and exert opposing effects on the mammary epithelium in vivo. *Mol Cell Biol.* 2009; 29:4935–48. [PubMed: 19596786]
36. Maatta JA, Sundvall M, Junttila TT, Peri L, Laine VJ, Isola J, et al. Proteolytic cleavage and phosphorylation of a tumor-associated ErbB4 isoform promote ligand-independent survival and cancer cell growth. *Mol Biol Cell.* 2006; 17:67–79. [PubMed: 16251361]
37. Feng SM, Sartor CI, Hunter D, Zhou H, Yang X, Caskey LS, et al. The HER4 cytoplasmic domain, but not its C terminus, inhibits mammary cell proliferation. *Mol Endocrinol.* 2007; 21:1861–76. [PubMed: 17505063]
38. Sundvall M, Veikkolainen V, Kurppa K, Salah Z, Tvorogov D, van Zoelen EJ, et al. Cell death or survival promoted by alternative isoforms of ErbB4. *Mol Biol Cell.* 2010; 21:4275–86. [PubMed: 20943952]
39. Muraoka-Cook RS, Sandahl M, Husted C, Hunter D, Miraglia L, Feng SM, et al. The intracellular domain of ErbB4 induces differentiation of mammary epithelial cells. *Mol Biol Cell.* 2006; 17:4118–29. [PubMed: 16837552]
40. Muraoka-Cook RS, Sandahl M, Hunter D, Miraglia L, Earp HS 3rd. Prolactin and ErbB4/HER4 signaling interact via Janus kinase 2 to induce mammary epithelial cell gene expression differentiation. *Mol Endocrinol.* 2008; 22:2307–21. [PubMed: 18653779]
41. Citterio C, Menacho-Marquez M, Garcia-Escudero R, Larive RM, Barreiro O, Sanchez-Madrid F, et al. The rho exchange factors vav2 and vav3 control a lung metastasis-specific transcriptional program in breast cancer cells. *Sci Signal.* 2012; 5:ra71. [PubMed: 23033540]
42. Allison JG, Das PM, Ma J, Inglis FM, Jones FE. The ERBB4 intracellular domain (4ICD) regulates NRG1-induced gene expression in hippocampal neurons. *Neurosci Res.* 2011; 70:155–63. [PubMed: 21352860]

43. Frey MR, Edelblum KL, Mullane MT, Liang D, Polk DB. The ErbB4 growth factor receptor is required for colon epithelial cell survival in the presence of TNF. *Gastroenterology*. 2009; 136:217–26. [PubMed: 18973758]
44. Hilliard VC, Frey MR, Dempsey PJ, Peek RM Jr, Polk DB. TNF-alpha converting enzyme-mediated ErbB4 transactivation by TNF promotes colonic epithelial cell survival. *Am J Physiol Gastrointest Liver Physiol*. 2011; 301:G338–46. [PubMed: 21617117]
45. Paatero I, Jokilampi A, Heikkinen PT, Iljin K, Kallioniemi OP, Jones FE, et al. Interaction with ErbB4 promotes hypoxia-inducible factor-1alpha signaling. *J Biol Chem*. 2012; 287:9659–71. [PubMed: 22308027]
46. Komuro A, Nagai M, Navin NE, Sudol M. WW domain-containing protein YAP associates with ErbB-4 and acts as a co-transcriptional activator for the carboxyl-terminal fragment of ErbB-4 that translocates to the nucleus. *J Biol Chem*. 2003; 278:33334–41. [PubMed: 12807903]
47. van der Meer DL, Degenhardt T, Vaisanen S, de Groot PJ, Heinaniemi M, de Vries SC, et al. Profiling of promoter occupancy by PPARalpha in human hepatoma cells via ChIP-chip analysis. *Nucleic Acids Res*. 2010; 38:2839–50. [PubMed: 20110263]
48. Battaglia S, Maguire O, Thorne JL, Hornung LB, Doig CL, Liu S, et al. Elevated NCOR1 disrupts PPARalpha/gamma signaling in prostate cancer and forms a targetable epigenetic lesion. *Carcinogenesis*. 2010; 31:1650–60. [PubMed: 20466759]
49. Zhang H, Liu CY, Zha ZY, Zhao B, Yao J, Zhao S, et al. TEAD transcription factors mediate the function of TAZ in cell growth and epithelial-mesenchymal transition. *J Biol Chem*. 2009; 284:13355–62. [PubMed: 19324877]
50. Ontsouka EC, Albrecht C. Cholesterol transport and regulation in the mammary gland. *J Mammary Gland Biol Neoplasia*. 2014; 19:43–58. [PubMed: 24510467]
51. Kaiser J. Cancer. Cholesterol forges link between obesity and breast cancer. *Science*. 2013; 342:1028. [PubMed: 24288308]
52. Sorrentino G, Ruggeri N, Specchia V, Cordenonsi M, Mano M, Dupont S, et al. Metabolic control of YAP and TAZ by the mevalonate pathway. *Nat Cell Biol*. 2014; 16:357–66. [PubMed: 24658687]

Implications

ErbB4 as a regulator of hippo and mevalonate pathways provides new insight into milk production and anabolic processes in normal mammary epithelia and cancer.

Figure 1. Full Length (FL) ERBB4 expression in T47D and MCF10A cells

T47D cells were transfected with vector (V), FL CYT-1 ERBB4 (1) or FL CYT-2-ERBB4 (2) plasmids, and next day, a fraction of cells was collected to extract protein while the rest were re-plated in presence (+NRG1) or absence (-NRG1) of 100ng/ml NRG1 and collected on day 3. Similarly, MCF10A cells transfected by vector (V), CYT-1-ERBB4 (1) or CYT-2-ERBB4 (2) plasmids were collected at day 1 and day 3 and protein whole cell lysates were prepared. Relative levels of FL ERBB4 and GAPDH were determined by immunoblotting (A). MCF10A cells were infected with viruses containing above ERBB4 constructs, and cells were similarly collected at day 1 and day 3, and probed for ERBB4 and GAPDH. Stably infected cells which strongly expressed ErbB4 were produced by selection with 10µg/ml puromycin for 2 weeks, and were maintained in 1µg/ml puromycin supplemented media in culture (A, Right). These MCF10A stable cell lines were probed for ERBB4 mRNA by qRT-PCR using T47D cells as positive control for ERBB4 expression (B). Effect of NRG1 (100ng/ml) on the relative levels of ERBB4, phosphorylated ERBB4 Tyr1056 (P-ERBB4), AKT Ser473 (P-AKT) and MAPK Thr203/Tyr204 (P-MAPK) in MCF10A stables over 24h was determined by immunoblotting (C). The phospho-ERBB4 antibody used here nominally detects Tyr1056, but it has also been reported to detect phosphorylated sites in region 1032 to 1040 present in both CYT-1 and CYT-2 [21].

Figure 2. Proliferation and invasion of FL and ICD ERBB4 expressing MCF10A cells

MCF10A cells stably infected with vector backbone (Vector), FL CYT-1 (FL CYT-1) or CYT-2 (FL CYT-2) ERBB4 were seeded in 96-well plates at 1000 cells/well (4 wells/group) and allowed to grow in 5% horse serum containing media in the absence or presence of 100ng/ml NRG1 (A). Similarly, these cells were also allowed to grow in serum-free media in the absence or presence of 100ng/ml NRG1 (B). Cell proliferation was assessed by the CellTiter-Glo® luminescent cell viability assay every day over 5 day-culture period. Data points indicate the mean luminescence/well $\pm$ SEM in each group. All pairwise group comparisons in B showed significant differences, except CYT-1 vs Vector or CYT-1+NRG1 vs Vector. $***P < 0.0001$ by one-way ANOVA followed by Newman-Keuls multiple comparison test. Vector-MCF10A, FL CYT-1 MCF10A, FL CYT-2-MCF10A and invasive MDA-MB-231 cells were seeded at 50,000 cells/well in 24-well plates with 8- μ m matrigel or control inserts (3 wells/group for matrigel and control inserts each) in OptiMEM with 0.1% horse serum and 100ng/ml NRG1. 5% horse serum and 100ng/ml NRG1 containing media was used as a chemoattractant. Invasive cells were stained after 24h, and percent invasion and invasive potential was determined. Vertical bars indicate mean cell count $\pm$ SEM for three replicates in each group. All group comparisons showed significant differences, except Vector-MCF10A versus FL CYT-1 MCF10A (C). Protein expression of V5- tagged ICD ERBB4 isoforms from 2 MCF10A biological sets: vector (Va, Vb), CYT-1 (1a, 1b), CYT-2 (2a, 2b) was detected by Western blot using GAPDH as internal control (D). Similar to A and B, ICD CYT-1 and CYT-2 cells were plated in 5% containing medium. Proliferation was assessed by counting cells after trypan blue staining over 5-days, data points indicate mean viable count/well $\pm$ SEM in each group. Percent invasion of ICD CYT-1 and ICD CYT-2 MCF10A was measured with 5% horse serum containing medium as chemoattractant (F). $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$ by one-way ANOVA followed by Newman-Keuls multiple comparison test.

Figure 3. Network analyses of predicted upstream transcriptional regulators in cells overexpressing ERBB4 ICD CYT-1 or CYT-2

A) Gene expression from MCF10A stable cells expressing ERBB4 ICD CYT-1 was compared against empty vector MCF10A cells using the Limma statistical analysis package. All genes with adjusted p -value < 0.05 were entered into IPA (Ingenuity Systems) and differential gene expression was analyzed with a 2 fold-change cut-off in either direction compared with empty vector. Networks were generated by manually selecting one or more transcription factors with predicted activation/inhibition by IPA analysis (see Table 3). Transcription factors with known pathway interactions (e.g., SMAD2 and SMAD3) were included in the same network analysis and labeled with a common pathway name (e.g., TGF- β). The genes associated with each transcription factor make up the circumference of the network. Genes in red indicate increased expression in the Limma data set while genes in green indicate decreased expression. Transcription factors in the center of the networks are orange when predicted as activated and blue when predicted as inhibited. Lines connect transcription factors to known regulated genes. Orange lines represent known gene regulation that leads to activation of the associated gene while blue lines represent inhibition of the gene. Yellow lines are associations where the direction of gene regulation from the Limma analysis is inconsistent with the predicted activation state of the transcription factor. Gray lines are published interactions between a transcription factor and a gene where an effect on gene expression is not predicted (e.g., the transcription factor might bind the protein but not regulate gene expression). TWIST1 was included with the authors' discretion even though it had a z -score (1.96) that did not meet the initial criteria of $abs(z\text{-score}) \geq 2.0$ due to its biological importance. B) Same analysis as (A) but comparing gene expression of

MCF10A stable cells expressing ERBB4 ICD CYT-2 with empty vector MCF10A cells. NOTE: For both (A) and (B) any of the transcriptional regulators in Table 3 listed as predicted upstream regulators could have been used to generate these network analyses. Readers should exercise caution in interpreting these networks because they are not comprehensive and represent computationally predicted activation. Presence or absence of a network in the ERBB4 CYT-1 ICD compared with ERBB4 CYT-2 ICD should not be interpreted as a network exclusive to a particular isoform. However, the significance (z-score) of activation can be compared between isoforms in Table 3. These figures are meant to give a sampling of major biological pathways that are significantly over-represented in the IPA analysis when transcription factors were evaluated as potential upstream regulators.

Figure 4. Quantitative Real Time PCR Validation of ChIP-Seq Targets

Sequences were selected for validation that matched ICD transcripts and were no more than 20kb from the promoters. Primers were designed using Primer 3 (ref). ChIP was performed on Cyt1 and Cyt2 ICD nuclear extracts using control rabbit IgG, Histone H3, and V5 antibodies. Quantitative PCR was performed on each set with probes for the following; SPARC, MXD4, STMN1, ADAP1, and APOE. The ChIPs were done for 3 biological repeats. Results are given as percent input. Standard Deviation was determined for the biological repeats.

Figure 5. ERBB4 regulated genes in T47D cells

A) Western blot showing ERBB4 knockdown in T47D cells, and re-expression of CYT-1 or CYT-2 ERBB4 isoform in KD stables using pI20 DOX-inducible expression plasmid. Cells were serum-starved and treated with DOX (100ng/ml) for 24h and NRG1 (100ng/ml) for 2h, and protein whole cell lysates were prepared to measure relative levels of phosphorylated and total ERBB4, CTGF and GAPDH by immunoblotting. B4=ERBB4; pI20=pInducer20; B4 sh3=pLKO ERBB4 3'-UTR shRNA; C1=CYT-1 ERBB4 JM-a; C2=CYT-2 ERBB4 JM-a. B) RT-PCR validating genes identified in the MCF10A ERBB4 microarray in T47D cells. Top row are genes involved in the cholesterol/mevalonate pathway. Middle row are Hippo (YAP/TEAD) regulated genes. Bottom row includes a Wnt negative regulator (DKK1) and basal breast markers (TP63/KRT14). CTGF, SPARC, and KRT14 are plotted as the average of 3 biological replicates. The remaining genes are plotted as technical triplicates of a single experiment with subsequent validation in two additional biological replicates. scr=pLKO scramble; V=pInducer20 vector; B4 sh3=pLKO ERBB4 3'-UTR shRNA; CYT-1=pInducer20 ERBB4 JM-a, CYT-1; CYT-2=pInducer20 ERBB4 JM-a, CYT-2; +NRG1= neuregulin (100ng/ml) for 2h. C) Gene expression during pregnancy and lactation in the mouse mammary gland from Anderson et al. 2007. Data were downloaded from GEO (GSE8191). Preg=pregnancy; Lac=lactation; Invo=involution; D=day.

Table 1
Top gene lists for up- and down-regulated genes with NRG1 stimulation in full-length ERBB4 CYT-1 vs. empty vector or full-length ERBB4 CYT-2 vs. empty vector

Gene expression from stable MCF10A cells expressing full-length ERBB4 CYT-1 or CYT-2 stimulated with NRG1 (100ng/ml, 2h) following starvation was compared against empty vector stable MCF10A cells stimulated with NRG1 (100ng/ml, 2h) following starvation using the Limma statistical package. Genes with adjusted p-value<0.05 were sorted by fold-change (FC) expression over empty vector. Any gene listed more than once in top genes was edited to only include the probe with the highest FC expression. Genes highlighted in blue are unique to the top 20 genes for full-length ERBB4 CYT-1. Genes highlighted in purple are unique to the top 20 genes for full-length ERBB4 CYT-2. Genes in bold are shared between full-length ERBB4 CYT-1 and full-length CYT-2, with the same direction of regulation. NOTE: The top upregulated full-length ERBB4 CYT-1 list was expanded to 21 genes to include TGFA, since it is known that this ligand is important in ERBB family regulation.

	TOP UPREGULATED FL pCYT-1 vs pVEC			
	GENE	DESCRIPTION	FC	adj.P.Val
1	S100A9	Homo sapiens S100 calcium binding protein A9 (calgranulin B) (S100A9), mRNA.	1.6	4.67E-06
2	LOC100132240	PREDICTED: Homo sapiens misc_RNA(LOC100132240), miscRNA.	1.6	2.78E-05
3	LOC441019	PREDICTED Homo sapiens hypothetical LOC441019 (LOC441019), mRNA.	1.6	1.68E-08
4	PABPC1	Homo sapiens poly(A) binding protein, cytoplasmic 1 (PABPC1), mRNA.	1.5	2.01E-02
5	LOC645691	PREDICTED: Homo sapiens similar to heterogeneous nuclear ribonucleoprotein in A1 (LOC645691), mRNA.	1.5	2.25E-02
6	CITED4	Homo sapiens Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 4 (CITED4), mRNA.	1.5	3.85E-05
7	HMGA1	Homo sapiens high mobility group AT-hook 1 (HMGA1), transcript variant 1, mRNA.	1.5	2.73E-02
8	TNFRSF6B	Homo sapiens tumor necrosis factor receptor superfamily, member 6b, decoy (TNFRSF6B), transcript variant M68C, mRNA.	1.5	5.90E-05
9	NT5E	Homo sapiens 5'-nucleotidase, ecto (CD73) (NT5E), mRNA.	1.4	6.45E-03
10	KIR2DL3	Homo sapiens killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3 (KIR2DL3), transcript variant 1, mRNA.	1.4	4.53E-06
11	LRRFIP1	Homo sapiens leucine rich repeat (in FLII) interacting protein 1 (LRRFIP1), mRNA.	1.4	1.37E-02
12	ZNF503	Homo sapiens zinc finger protein 503 (ZNF503), mRNA.	1.4	2.41E-02
13	TDP1	Homo sapiens tyrosyl-DNA phosphodiesterase 1 (TDP1), transcript variant 1, mRNA.	1.4	2.25E-02
14	PHLDA1	Homo sapiens pleckstrin homology-like domain, family A, member 1 (PHLDA1), mRNA.	1.4	3.44E-03
15	C3orf34	Homo sapiens chromosome 3 open reading frame 34 (C3orf34), mRNA.	1.4	2.58E-02
16	LOC641768	PREDICTED: Homo sapiens similar to ribosomal protein S26, transcript variant 2 (LOC641768), mRNA.	1.4	2.28E-02
17	TATDN1	Homo sapiens TatD DNase domain containing 1 (TATDN1), mRNA.	1.4	2.91E-02
18	IL7R	PREDICTED: Homo sapiens interleukin 7 receptor (IL7R), mRNA.	1.4	2.78E-05
19	ANGPTL4	Homo sapiens angiopoietin-like 4 (ANGPTL4), transcript variant 1, mRNA.	1.4	4.66E-04
20	LYAR	Homo sapiens Ly1 antibody reactive homolog (mouse) (LYAR), mRNA.	1.4	1.69E-05

TOP DOWN REGULATED FL pCYT-1 vs pVEC				
	GENE	DESCRIPTION	FC	adj.P.Val
1	EGR1	Homo sapiens early growth response 1 (EGR1), mRNA.	-2.4	7.64E-05
2	KRT1	Homo sapiens keratin 1 (KRT1), mRNA.	-2.2	1.68E-08
3	KRT4	Homo sapiens keratin 4 (KRT4), mRNA.	-2.2	2.62E-08
4	IGFL1	Homo sapiens IGF-like family member 1 (IGFL1), mRNA.	-2.0	3.78E-08
5	LOC100132761	PREDICTED: Homo sapiens hypothetical protein LOC100132761 (LOC 100132761), mRNA.	-2.0	2.52E-02
6	FABP4	Homo sapiens fatty acid binding protein 4, adipocyte (FABP4), mRNA.	-1.9	3.66E-03
7	KLK7	Homo sapiens kallikrein-related peptidase 7 (KLK7), transcript variant 1, mRNA.	-1.9	2.04E-09
8	ATP6V1B1	Homo sapiens ATPase, H+transporting, lysosomal 56/58kDa, V1 subunit B1 (ATP6V1B1), mRNA.	-1.9	3.50E-09
9	KRT10	Homo sapiens keratin 10 (epidermolytic hyperkeratosis; keratosis palmaris et plantaris) (KRT10), mRNA.	-1.8	1.27E-06
10	KLK5	Homo sapiens kallikrein-related peptidase 5 (KLK5), transcript variant 2, mRNA.	-1.8	2.07E-06
11	CRABP2	Homo sapiens cellular retinoic acid binding protein 2 (CRABP2), mRNA.	-1.8	1.01E-06
12	GPX2	Homo sapiens glutathione peroxidase 2 (gastrointestinal) (GPX2), mRNA.	-1.8	1.15E-06
13	LOC642956	PREDICTED: Homo sapiens hypothetical LOC642956 (LOC642956), mRNA.	-1.7	7.45E-06
14	FOLR1	Homo sapiens folate receptor (adult) (FOLR1), transcript variant8, mRNA.	-1.7	2.13E-04
15	KANK4	Homo sapiens KN motif and ankyrin repeat domains 4 (KANK4), mRNA.	-1.7	1.15E-06
16	NUCKS1	Homo sapiens nuclear casein kinase and cyclin-dependent kinase substrate 1 (NUCKS1), mRNA.	-1.7	1.83E-02
17	TXNIP	Homo sapiens thioredoxin interacting protein (TXNIP), mRNA.	-1.6	1.52E-07
18	MGC102966	PREDICTED: Homo sapiens similar to Keratin, type I cytoskeletal 16 (Cytokeratin-16) (CK-16) (Keratin-16) (K16) (MGC102966), misc RNA.	-1.6	5.44E-05
19	CNTNAP2	Homo sapiens contactin associated protein-like 2 (CNTNAP2), mRNA.	-1.6	4.60E-05
20	TCN1	Homo sapiens transcobalamin I (vitamin B12 binding protein, R binder family (TCN1), mRNA.	-1.6	1.24E-06

BOLD = Shared gene between CYT-1 and CYT-2				
	GENE	DESCRIPTION	FC	adj.P.Val
1	MMP3	Homo sapiens matrix metalloproteinase 3 (stromelysin 1, progelatinase) (MMP3), mRNA.	2.2	3.60E-14
2	SERPINE2	Homo sapiens serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2 (SERPINE2), mRNA.	1.9	1.25E-10
3	DKK1	Homo sapiens dickkopf homolog 1 (Xenopus laevis) (DKK1), mRNA.	1.7	1.41E-07
4	TNFRSF6B	Homo sapiens tumor necrosis factor receptor super family, member 6b, decoy (TNF RSF6B), transcript variant M68C, mRNA.	1.7	8.79E-08
5	KIR2DL3	Homo sapiens killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3 (KIR2DL3), tran	1.7	2.75E-10
6	LOC100132240	PREDICTED: Homo sapiens misc_RNA (LOC100132240), miscRNA.	1.7	1.48E-06
7	IL7R	PREDICTED: Homo sapiens interleukin 7 receptor (IL7R), mRNA.	1.7	4.19E-10
8	ANGPTL4	Homo sapiens angiopoietin-like 4 (ANGPTL4), transcript variant 1, mRNA.	1.6	4.51E-08
9	S100A9	Homo sapiens S100 calcium binding protein A9 (calgranulin B) (S100A9), mRNA.	1.6	6.00E-07
10	IL33	Homo sapiens interleukin 33 (IL33), mRNA.	1.6	2.39E-09
11	FST	Homo sapiens follistatin (FST), transcript variant FST344, mRNA.	1.6	5.61E-04
12	NTSE	Homo sapiens 5'-nucleotidase, ecto (CD73) (NTSE), mRNA.	1.6	1.02E-04
13	ACOX2	Homo sapiens acyl-Coenzyme A oxidase 2, branched chain (ACOX2), mRNA.	1.6	8.69E-07

	BOLD = Shared gene between CYT-1 and CYT-2			
	GENE	DESCRIPTION	FC	adj.P.Val
14	CCNA1	Homo sapiens cyclin A1 (CCNA1), mRNA.	1.5	8.73E-09
15	EOMES	Homo sapiens eomesodermin homolog (Xenopus laevis) (EOMES), mRNA.	1.5	2.49E-05
16	CYP27B1	Homo sapiens cytochrome P450, family27, subfamily B, polypeptide 1 (CYP27B1), nuclear gene encoding mitochondrial protein, mRNA.	1.5	8.37E-05
17	ETS1	Homo sapiens v-ets erythroblastosis virus E26 oncogene homolog 1 (avian) (ETS1), mRNA.	1.5	1.07E-04
18	KIR2DL4	Homo sapiens killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4 (KIR2DL4), mRNA.	1.5	3.27E-08
19	MSTP131 (NRG)	Homo sapiens MSTP131 (MST131) mRNA, complete cds	1.5	1.12E-06
20	GJB2	Homo sapiens gap junction protein, beta 2,26kDa (GJB2), mRNA.	1.5	7.71E-07

	TOP DOWNREGULATED FLpCYT-2 vs pVEC			
	GENE	DESCRIPTION	FC	adj.P.Val
1	EGR1	Homo sapiens early growth response 1 (EGR1), mRNA.	-2.7	5.10E-06
2	KRT4	Homo sapiens keratin 4 (KRT4), mRNA.	-2.6	1.01E-10
3	KRT1	Homo sapiens keratin 1 (KRT1), mRNA.	-2.5	2.33E-10
4	KLK7	Homo sapiens kallikrein-related peptidase 7 (KLK7), transcript variant 1, mRNA.	-2.2	1.48E-12
5	KLK5	Homo sapiens kallikrein-related peptidase 5 (KLK5), transcript variant 1, mRNA.	-2.2	6.52E-11
6	KRT10	Homo sapiens keratin 10 (epidermolytic hyperkeratosis: keratosis palmaris et plantaris) (KRT10), mRNA.	-2.2	1.55E-09
7	IGFL1	Homo sapiens IGF-like family member 1 (IGFL1), mRNA.	-2.1	1.01E-09
8	ATP6V1B1	Homo sapiens ATPase, H+transporting, lysosomal 56/58kDa, V1 subunit B1 (ATP6V1B1), mRNA.	-2.0	3.90E-11
9	FABP4	Homo sapiens fatty acid binding protein 4, adipocyte (FABP4), mRNA.	-2.0	1.09E-03
10	FOLR1	Homo sapiens folate receptor 1 (adult) (FOLR1), transcript variant8, mRNA.	-2.0	4.47E-07
11	CRABP2	Homo sapiens cellular retinoic acid binding protein 2 (CRABP2), mRNA.	-1.9	1.46E-08
12	GNPMB	Homo sapiens glycoprotein (transmembrane) nmb (GNPMB), transcript variant 1, mRNA.	-1.9	2.15E-05
13	GPX2	Homo sapiens glutathione peroxidase 2 (gastro intestinal) (GPX2), mRNA.	-1.8	1.19E-07
14	KRT13	Homo sapiens keratin 13 (KRT13), transcript variant 2, mRNA.	-1.8	5.76E-12
15	PSG4	Homo sapiens pregnancy specific beta-1-glycoprotein 4 (PSG4), transcript variant 1, mRNA.	-1.7	4.76E-10
16	FBXO32	Homo sapiens F-box protein 32 (FBXO32), transcript variant 2, mRNA.	-1.7	4.91E-08
17	MATN2	Homo sapiens matrilin 2 (MATN2), transcript variant 1, mRNA.	-1.7	6.58E-05
18	TCN1	Homo sapiens transcobalamin I (vitamin B12 binding protein, R binder family) (TCN1), mRNA.	-1.7	2.10E-08
19	TXNIP	Homo sapiens thioredoxin interacting protein (TXNIP), mRNA.	-1.7	4.83E-09
20	CLDN7	Homo sapiens Claudin 7 (CLDN7), mRNA.	-1.7	3.94E-10

= Unique to CYT-1 top 20 genes

= Unique to CYT-2 top 20 genes

BOLD = Shared gene between CYT-1 and CYT-2

Table 2

Top gene lists for up- and down-regulated genes in ERBB4 ICD CYT-1 vs. empty vector or ERBB4 ICD CYT-2 vs. empty vector. Gene expression from stable MCF10A cells expressing ERBB4 ICD CYT-1 or CYT-2 was compared against empty vector stable MCF10A cells using the Limma statistical package. Genes with adjusted p-value < 0.05 were sorted by fold-change (FC) expression over empty vector. Any gene listed more than once in top genes was edited to only include the probe with the highest FC expression. Genes highlighted in blue are unique to the top 20 genes for ERBB4 CYT-1 ICD. Genes highlighted in purple are unique to the top 20 genes for ERBB4 CYT-2 ICD. Genes in bold are shared between ERBB4 CYT-1 ICD and CYT-2 ICD, with the same direction of regulation.

TOP UPREGULATED CYT-1 ICD				
	GENE	DESCRIPTION	FC	adj.P.Val
1	DCN	Homo sapiens decorin (DCN), transcript variant C, mRNA.	16.3	3.98E-23
2	SRGN	Homo sapiens serglycin (SRGN), mRNA.	13.7	1.03E-21
3	FBXO32	Homo sapiens F-box protein 32 (FBXO32), transcript variant 2, mRNA.	13.5	5.00E-22
4	TAGLN	Homo sapiens transgelin (TAGLN), transcript variant 2, mRNA.	11.0	4.17E-20
5	CRYAB	Homo sapiens crystallin, alpha B (CRYAB), mRNA.	8.3	1.35E-17
6	CTGF	Homo sapiens connective tissue growth factor (CTGF), mRNA.	8.2	5.64E-19
7	DARC	Homo sapiens Duffy blood group, chemokine receptor (DARC), transcript variant 2, mRNA.	7.4	1.16E-18
8	MGP	Homo sapiens matrix Gla protein (MGP), mRNA.	7.0	2.85E-19
9	COL8A1	Homo sapiens collagen, type VIII, alpha 1 (COL8A1), transcript variant 2, mRNA.	6.8	5.24E-19
10	NNMT	Homo sapiens nicotinamide N-methyltransferase (NNMT), mRNA.	6.3	2.21E-19
11	IFI6	Homo sapiens interferon, alpha-inducible protein 6 (IFI6), transcript variant 2, mRNA.	5.8	1.44E-18
12	LOX	Homo sapiens lysyl oxidase (LOX), mRNA.	5.5	2.26E-17
13	SLC46A3	Homo sapiens solute carrier family 46, member 3 (SLC46A3), mRNA.	5.3	8.34E-19
14	CYR61	Homo sapiens cysteine-rich, angiogenic inducer, 61 (CYR61), mRNA.	5.2	1.16E-18
15	SRPX	Homo sapiens sushi-repeat-containing protein, X-linked (SRPX), mRNA.	5.0	8.98E-17
16	PLAC8	Homo sapiens placenta-specific 8 (PLAC8), mRNA.	5.0	2.00E-16
17	IFI44L	Homo sapiens interferon-induced protein 44-like (IFI44L), mRNA.	4.7	1.32E-17
18	ANKRD1	Homo sapiens ankyrin repeat domain 1 (cardiac muscle) (ANKRD1), mRNA.	4.5	2.29E-16
19	PTTG1	Homo sapiens pituitary tumor-transforming 1 (PTTG1), mRNA.	4.5	3.69E-17
20	FAP	Homo sapiens fibroblast activation protein, alpha (FAP), mRNA.	4.5	1.41E-16

TOP DOWNREGULATED CYT-1 ICD				
	GENE	DESCRIPTION	FC	adj.P.Val
1	SPRR2A	Homo sapiens small proline-rich protein 2A (SPRR2A), mRNA.	-6.3	3.40E-18
2	PRSS3	Homo sapiens protease, serine, 3 (mesotrypsin) (PRSS3), mRNA.	-4.5	2.06E-15
3	SERPINB2	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2), mRNA.	-4.5	2.26E-17
4	SPRR2F	Homo sapiens small proline-rich protein 2F (SPRR2F), mRNA.	-4.2	5.46E-14
5	TNFRSF6B	Homo sapiens tumor necrosis factor receptor superfamily, member 6b, decoy (TNFRSF6B), transcript variant M68C, mRNA.	-4.2	7.41E-16
6	CENTA1	Homo sapiens centaurin, alpha 1 (CENTA1), mRNA.	-4.1	1.04E-16

TOP DOWNREGULATED CYT-1 ICD				
	GENE	DESCRIPTION	FC	adj.P.Val
7	TMEM16A	Homo sapiens transmembrane protein 16A (TMEM 16A), mRNA.	-3.3	4.91E-16
8	FLRT3	Homo sapiens fibronectin leucine rich transmembrane protein 3 (FLRT3), transcript variant 2, mRNA.	-3.2	3.71E-15
9	SPRR1B	Homo sapiens small proline-rich protein 1B (comifin) (SPRR1B), mRNA.	-3.0	8.20E-15
10	SOCS2	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), mRNA.	-3.0	1.74E-13
11	HS3ST1	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 1 (HS3ST1), mRNA.	-3.0	4.70E-13
12	FOXQ1	Homo sapiens forkhead box Q1 (FOXQ1), mRNA.	-3.0	1.61E-14
13	EGFR	Homo sapiens epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian) (EGFR), transcript variant 1, mRNA.	-3.0	7.48E-16
14	MMP7	Homo sapiens matrix metalloproteinase 7 (matrilysin, uterine) (MMP7), mRNA.	-2.8	7.20E-14
15	LCN2	Homo sapiens lipocalin 2 (LCN2), mRNA.	-2.8	8.20E-15
16	ITGA6	Homo sapiens integrin, alpha 6 (ITGA6), transcript variant 2, mRNA.	-2.7	8.33E-11
17	AGPAT9	Homo sapiens 1-acylglycerol-3-phosphate O-acyltransferase 9 (AGPAT9), mRNA.	-2.7	1.99E-14
18	UPP1	Homo sapiens uridine phosphorylase 1 (UPP1), transcript variant 1, mRNA.	-2.7	2.44E-13
19	GPR110	Homo sapiens G protein-coupled receptor 110 (GPR110), transcript variant 1, mRNA.	-2.6	3.40E-13
20	ANTXR2	Homo sapiens anthrax toxin receptor 2 (ANTXR2), mRNA.	-2.6	2.50E-14

TOP 20 UPREGULATED CYT-2 ICD				
	GENE	DESCRIPTION	FC	adj.P.Val
1	SRGN	Homo sapiens serglycin (SRGN), mRNA.	9.2	4.24E-20
2	FBXO32	Homo sapiens F-box protein 32 (FBXO32), transcript variant 2, mRNA.	9.1	2.20E-20
3	DCN	Homo sapiens decorin (DCN), transcript variant C, mRNA.	8.2	2.20E-20
4	CTGF	Homo sapiens connective tissue growth factor (CTGF), mRNA.	6.9	4.59E-18
5	MGP	Homo sapiens matrix Gla protein (MGP), mRNA.	6.0	1.45E-18
6	TAGLN	Homo sapiens transgelin (TAGLN), transcript variant 2, mRNA.	5.9	1.25E-17
7	PTTG1	Homo sapiens pituitary tumor-transforming 1 (PTTG1), mRNA.	5.7	4.57E-18
8	TOP2A	Homo sapiens topoisomerase (DNA) II alpha 170kDa (TOP2A), mRNA.	5.7	3.08E-16
9	UBE2C	Homo sapiens ubiquitin-conjugating enzyme E2C (UBE2C), transcript variant 3, mRNA.	5.6	4.57E-18
10	NNMT	Homo sapiens nicotinamide N-methyltransferase (NNMT), mRNA.	5.5	1.45E-18
11	KIF 20A	Homo sapiens kinesin family member 20A (KIF 20A), mRNA.	5.0	6.58E-18
12	DARC	Homo sapiens Duffy blood group, chemokine receptor (DARC), transcript variant 2, mRNA.	4.8	1.65E-16
13	COL8A1	Homo sapiens collagen, type VIII, alpha 1 (COL8A1), transcript variant 2, mRNA.	4.7	3.77E-17
14	AURKB	Homo sapiens aurora kinase B (AURKB), mRNA.	4.5	1.44E-15
15	PTTG3P	Homo sapiens pituitary tumor-transforming 3 (pseudogene) (PTTG3P), non-coding RNA.	4.3	7.72E-16
16	PLAC8	Homo sapiens placenta-specific 8 (PLAC8), mRNA.	4.3	1.35E-15
17	CYR61	Homo sapiens cysteine-rich, angiogenic inducer, 61 (CYR61), mRNA.	4.3	2.46E-17
18	CRYAB	Homo sapiens crystallin, alpha B (CRYAB), mRNA.	4.2	1.93E-14
19	NUSAP1	Homo sapiens nucleolar and spindle associated protein 1 (NUSAP1), transcript variant 2, mRNA.	4.2	3.77E-17
20	SERPINA3	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 anti proteinase, antitrypsin), member 3 (SERPINA3), mRNA.	4.1	3.71E-16

TOP 20 DOWNREGULATED CYT-2 ICD				
	GENE	DESCRIPTION	FC	adj.P.Val
1	SERPINB2	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2), mRNA.	-4.8	1.53E-17
2	FLRT3	Homo sapiens fibronectin leucine rich transmembrane protein 3 (FLRT3), transcript variant 2, mRNA.	-3.9	2.41E-16
3	SPRR2A	Homo sapiens small proline-rich protein 2A (SPRR2A), mRNA.	-3.9	1.44E-15
4	MMP7	Homo sapiens matrix metalloproteinase 7 (matrilysin, uterine) (MMP7), mRNA.	-3.4	2.03E-16
5	PRSS3	Homo sapiens protease, serine, 3 (mesotrypsin) (PRSS3), mRNA.	-3.4	1.03E-13
6	TNFRSF6B	Homo sapiens tumor necrosis factor receptor superfamily, member 6b, decoy (TNFRSF6B), transcript variant M68C, mRNA.	-3.1	2.93E-15
7	SPRR1B	Homo sapiens small proline-rich protein 1B (cornifin) (SPRR1B), mRNA.	-3.0	9.58E-15
8	DHRS9	Homo sapiens dehydrogenase/reductase (SDR family) member 9 (DHRS9), transcript variant 1, mRNA.	-2.9	9.35E-15
9	SOCS2	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), mRNA.	-2.8	3.29E-14
10	LEPREL1	Homo sapiens leprecan-like 1 (LEPREL1), mRNA.	-2.6	5.79E-13
11	CENTA1	Homo sapiens centaurin, alpha 1 (CENTA1), mRNA.	-2.6	2.06E-13
12	HS3ST1	Homo sapiens heparan sulfate (glucosamine) 3-O-sulfotransferase 1 (HS3ST1), mRNA.	-2.6	8.36E-12
13	ITGA6	Homo sapiens integrin, alpha 6 (ITGA6), transcript variant 2, mRNA.	-2.6	3.49E-10
14	SPRR2F	Homo sapiens small proline-rich protein 2F (SPRR2F), mRNA.	-2.5	2.25E-10
15	TMEM16A	Homo sapiens transmembrane protein 16A (TMEM 16A), mRNA.	-2.5	8.13E-14
16	LCN2	Homo sapiens lipocalin 2 (LCN2), mRNA.	-2.5	1.09E-13
17	FOXQ1	Homo sapiens forkhead box Q1 (FOXQ1), mRNA.	-2.4	9.63E-13
18	SPRY2	Homo sapiens sprouty homolog 2 (Drosophila) (SPRY2), mRNA.	-2.4	1.36E-12
19	FLJ12684	Homo sapiens hypothetical protein FLJ12684 (FLJ12684), mRNA. XR_001254	-2.4	1.62E-11
20	ANTXR2	Homo sapiens anthrax toxin receptor 2 (ANTXR2), mRNA.	-2.4	3.13E-13

= Unique to CYT-1 top 20 genes

= Unique to CYT-2 top 20 genes

BOLD = Shared gene between CYT-1 and CYT-2

Table 3

IPA Predicted Activation of Pathways and Transcriptional Regulators of ICD ERBB4ERBB4 ICD predicted function activation/inhibition (left): Gene expression from MCF10A stable cells expressing ERBB4 ICD CYT-1 or ICD CYT-2 was compared against empty vector MCF10A cells using the Limma statistical analysis package. All genes with adjusted p-value<0.05 were entered into IPA (Ingenuity Systems) and differential gene expression was analyzed with a 2 fold-change cut-off in either direction compared with empty vector. IPA functional analysis was used to identify biological trends within functional categories and to predict the effect of ERBB4 induced gene expression changes on biological processes. The IPA Downstream Effects Analysis uses information from literature compiled in the Ingenuity® Knowledge Base to evaluate the expected consequences of altered gene expression in a dataset on biological functions. This analysis identifies genes in the dataset that are known to affect certain functions and cross-references the directionality of gene expression with expected changes based on previous knowledge. IPA calculates a regulation z-score relative to control samples that predicts increased activation if the direction of gene changes is consistent with the literature across most genes associated with that function. A decrease in activation is predicted if the direction of gene changes associated with a particular function are mostly inconsistent with the literature. If there is no clear pattern of up- or down-regulation of a gene function set IPA does not give a prediction on activation. The “Category” column label describe the high-level functional category, while the “Functions Annotation” column label refers to a specific function that is significantly altered. The p-value refers to the significance of overlap of the Limma dataset with the related function and was calculated by the Fisher’s Exact Test. In the “Predicted Activation” column, a function is predicted as increased when the z-score ≥ 2 or as decreased when the z-score ≤ -2 . The number of genes associated with each function used in the IPA algorithms are listed in the # molecules column. The full list of gene names can be found in Supplementary Table 4. Functions shaded in red are predicted to have increased activity in ERBB4 ICD expressing cells and functions shaded in green are predicted to have decreased activity in ERBB4 ICD expressing cells. Predicted activation of upstream transcriptional regulators in MCF10A cells overexpressing ERBB4 ICD CYT-1 or CYT-2 (right): Gene expression from MCF10A stable cells expressing ERBB4 ICD CYT-1 or CYT-2 was compared against empty vector MCF10A cells using the Limma statistical analysis package. All genes with adjusted p-value<0.05 were entered into IPA (Ingenuity Systems) and differential gene expression was analyzed with a 2 fold-change cut-off in either direction compared with empty vector. Ingenuity Upstream Regulator Analysis was used to identify transcriptional regulators that can explain differential gene expression between ERBB4 expressing cells and empty vector cells. This analysis examines the overlap between known targets of each transcriptional regulator and genes from the Limma analysis. The algorithm uses curated information in the Ingenuity® Knowledge Base, comprised of published observed effects of regulators on target genes and direction of regulation (up/down/neutral), to generate a prediction of upstream regulator activity. In this analysis, two statistical measures are used—an overlap p-value and an activation z-score. The overlap p-value (calculated by Fisher’s Exact Test, $p < 0.01$) describes regulators that have significant overlap between genes from the ERBB4 Limma dataset and known targets of a particular transcriptional regulator. The z-score predicts the likely direction of regulation by comparing data against a model that randomly assigns directionality. Transcription factors with $abs(z\text{-score}) \geq 2.0$ were sorted from high to low based on CYT-1 ICD scores. CYT-2 ICD z-scores were first arranged around CYT-1 scores to fill in all commonly regulated transcription factors, then any factors unique

to CYT-1 were added to the list. Red represents strong predicted activation, green represents predicted repression, and gray represents factors not predicted.

Table 1
ICD-CYT-1's VEC FC2

Category	Functions Annotation	p-Value	Predicted Activation	Activation z-score	# Molecules
Cellular Assembly and Organization	microtubule dynamics	7.09E-07	Increased	3.1	46
Cellular Function and Maintenance	microtubule dynamics	7.09E-07	Increased	3.1	46
Cellular Assembly and Organization	organization of cytoskeleton	1.01E-06	Increased	3.0	51
Cellular Function and Maintenance	organization of cytoskeleton	1.01E-06	Increased	3.0	51
Cancer	neoplasia of tumor cell lines	5.42E-07	Increased	2.6	17
Cellular Development	neoplasia of tumor cell lines	5.42E-07	Increased	2.6	17
Cellular Growth and Proliferation	neoplasia of tumor cell lines	5.42E-07	Increased	2.6	17
Cellular Assembly and Organization	development of cytoplasm	8.99E-06	Increased	2.5	23
Cancer	neoplasia of cells	1.43E-07	Increased	2.4	23
Cancer	metastasis	1.04E-21	Increased	2.3	55
Cellular Assembly and Organization	formation of cytoskeleton	1.36E-04	Increased	2.3	18
Cancer	metastasis of cells	1.55E-05	Increased	2.2	12
Cell Death and Survival	cell death of brain cells	4.72E-04	Increased	2.2	14
Cell Death and Survival	cell death of central nervous system cells	1.01E-04	Increased	2.1	16
Cellular Development	branching of cells	4.27E-05	Increased	2.1	11
Cancer	metastasis of tumor cell lines	2.62E-04	Increased	2.1	8
Cellular Development	metastasis of tumor cell lines	2.62E-04	Increased	2.1	8
Cellular Growth and Proliferation	metastasis of tumor cell lines	2.62E-04	Increased	2.1	8
Cancer	tumorigenesis of cells	2.63E-04	Increased	2.0	13
Cell Death and Survival	cell death of cerebral cortex cells	1.89E-04	Increased	2.0	13
Cancer	metastasis of carcinoma cell lines	1.62E-04	Increased	2.0	4
Cellular Development	metastasis of carcinoma cell lines	1.62E-04	Increased	2.0	4
Cellular Growth and Proliferation	metastasis of carcinoma cell lines	1.62E-04	Increased	2.0	4
Dermatological Diseases and Conditions	skin abnormality	3.33E-05	Decreased	-2.0	14
Organismal Injury and Abnormalities	skin abnormality	3.33E-05	Decreased	-2.0	14
Cellular Movement	invasion of fibroblast cell lines	4.38E-06	Decreased	-2.1	8
Cell Cycle	cell cycle progression of tumor cell lines	9.50E-05	Decreased	-2.4	14
Cell Cycle	ploidy	4.77E-04	Decreased	-2.7	10

ICD CYT-2 vs VEC FC2					
Category	Functions Annotation	p-Value	Predicted Activation	Activation z-score	# Molecules
Cellular Growth and Proliferation	proliferation of cells	7.23E-14	Increased	3.1	108
Cellular Growth and Proliferation	proliferation of tumor cell lines	5.97E-11	Increased	3.0	57
Cellular Development	proliferation of tumor cell lines	5.97E-11	Increased	3.0	57
Cardiovascular System Development and Function	angiogenesis	5.22E-04	Increased	2.4	22
Cell Cycle	M phase of tumor cell lines	3.86E-13	Increased	2.4	15
Cancer	growth of malignant tumor	2.36E-06	Increased	2.2	10
Cardiovascular System Development and Function	development of blood vessel	1.04E-05	Increased	2.2	29
Organismal Development	development of blood vessel	1.04E-05	Increased	2.2	29
Cell Death and Survival	cell death of cerebral cortex cells	2.35E-04	Increased	2.2	11
Cardiovascular System Development and Function	development of cardiovascular system	4.64E-05	Increased	2.1	33
Cardiovascular System Development and Function	vasculogenesis	2.79E-06	Increased	2.0	28
Organismal Development	vasculogenesis	2.79E-06	Increased	2.0	28
Cellular Growth and Proliferation	cytostasis	1.76E-04	Increased	2.0	13
Cell Cycle	ploidy	2.85E-14	Decreased	-2.2	20
Cellular Movement	invasion of fibroblast cell lines	7.61E-04	Decreased	-2.2	5
Cell Cycle	ploidy of cells	1.26E-11	Decreased	-2.4	16

Transcription Factor	Activation Z-score	
	ICD CYT-1	ICD CYT-2
HIF1A	3.6	2.6
SREBF1	3.5	3.1
MTPN	3.3	2.4
E2F1	3.2	2.1
FOXO1	3.1	4.1
SREBF2	3.1	3.0
NFKB1A	3.1	2.6
SP1	3.0	0.0

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Transcription Factor	Activation Z-score	
	ICD CYT-1	ICD CYT-2
FOXM1	2.9	3.5
STAT4	2.8	0.0
TBX2	2.6	3.7
EGR2	2.6	2.2
SRF	2.6	2.6
Nfat (family)	2.4	0.0
SMAD2	2.4	0.0
SMARCA4	2.4	0.0
MYOD1	2.4	0.0
STAT3	2.4	2.4
KLF5	2.4	0.0
MKL1	2.4	0.0
RELA	2.2	0.0
PPARGC1B	2.2	0.0
SNAIL1	2.2	0.0
Notch	2.2	0.0
SMAD4	2.1	2.1
SMAD3	2.1	0.0
FOXO4	2.1	0.0
IRF7	2.1	2.6

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Transcription Factor	Activation Z-score	
	ICD CYT-1	ICD CYT-2
YAPI	2.0	2.0
MYBL2	2.0	2.2
NFYA	2.0	0.0
SIRT2	2.0	2.0
RUVBL1	2.0	0.0
NR1H3	0.0	2.2
E2F2	0.0	2.0
MYCN	-2.2	0.0
HOXA10	-2.5	0.0
WT1	-2.6	-2.4
KDM5B	-2.7	-3.2
SMARCE1	0.0	-2.0
HMGAI	0.0	-2.0
ATF3	0.0	-2.0
TCF3	0.0	-2.5
RBI	0.0	-2.6
CDKN2A	0.0	-2.7
NUPR1	0.0	-3.2