ERBB2b mRNA isoform encodes a nuclear variant of the ERBB2 oncogene in breast cancer

Guoqiang Hua, Aurélie Bergon, Pierre Cauchy, Brigitte Kahn-perlès, François Bertucci, Daniel Birnbaum, Nadia Benkirane-jessel, Jean Imbert

ERBB2b mRNA isoform encodes a nuclear variant of the *ERBB2* oncogene in breast cancer

Running title: ERBB2b nuclear isoform of ERBB2

Guoqiang Hua[1,3,4,*], Aurélie Bergon[1], Pierre Cauchy[1,5], Brigitte Kahn-Perlès[1], François Bertucci[2], Daniel Birnbaum[2], Nadia Benkirane-Jessel[3,4] and Jean Imbert[1,*]

[1]Inserm UMR1090 TAGC, Aix-Marseille University, 163 Avenue de Luminy, 13288, Marseille, cedex 9, France.

[2]Laboratoire d'Oncologie Prédictive, CRCM, CNRS UMR 7258, Inserm U1068, Institut Paoli-Calmettes, Aix-Marseille University, 232 Boulevard de Sainte-Marguerite, 13009, Marseille, France

[3]Inserm UMR1260, RNM, FMTS, 11 rue Humann, 67000, Strasbourg, France

[4]Université de Strasbourg, Faculté de Chirurgie Dentaire de Strasbourg, 8 rue Sainte-Elisabeth, 67000, Strasbourg, France

[5]Max Planck Institute of Immunobiology and Epigenetics, Stübeweg 51, 79108 Freiburg im Breisgau, Germany

*Correspondence: Guoqiang Hua, g.hua@unistra.fr; Jean Imbert, jean.imbert@inserm.fr

Abstract

The presence of nuclear ERBB2 receptor-type tyrosine kinase is one of the causes of the resistance to membrane ERBB2-targeted therapy in breast cancers. It has been previously reported that this nuclear location arises through at least two different mechanisms: proteolytic shedding of the extracellular domain of the full-length receptor and translation of the mRNA encoding *ERBB2* from internal initiation codons. Here, we report a new mechanism and function where a significant portion of nuclear ERBB2 results from the translation of the variant *ERBB2* mRNA under the transcriptional control of a distal promoter which is actively used in breast cancer cells. We show that both membrane ERBB2a and nuclear ERBB2b isoforms are prevalently expressed in breast cancer cell lines and carcinoma samples. The ERBB2b isoform, which is translated from mRNA variant 2, can directly translocate into the nucleus due to the lack of the signal peptide which is required for an intermediate membrane location. siRNA-mediated gene silencing showed that ERBB2b can repress ERBB2a expression, encoded by variant 1, whereas ERBB2a activates ERBB2b. Nuclear ERBB2 binding to its own promoter was revealed by chromatin immunoprecipitation assay. Altogether, our results provide new insights into the origin and function of nuclear ERBB2 where it can participate at the same time in a positive or a negative feedback autoregulatory loop, dependent on which of its promoters this bona fide transcription factor is acting. They also provide a new understanding for the resistance to therapies targeting the membrane-anchored ERBB2 in breast cancer.

Key words: ERBB2, isoform, nuclear translocation, breast cancer, resistance

1 INTRODUCTION

The ERBB2/HER2 oncogene encodes a 185 kDa receptor tyrosine kinase (RTK) [Akiyama et al., 1986; Yarden and Sliwkowski, 2001] that belongs to the epithelial growth factor receptor (EGFR) family, including epidermal growth factor receptor (EGFR)/ERBB1 [Ullrich et al., 1984], ERBB3 [Kraus et al., 1989], and ERBB4 [Plowman et al., 1993]. These receptors are composed of an extracellular domain, a transmembrane domain, and an intracellular domain. ERBB2 has no known ligand binding to its extracellular domain [Klapper et al., 1999]; however, it becomes the preferred heterodimeric partner of the other three ERBB members [Graus-Porta et al., 1997]. Ligand binding to these high-affinity complexes stimulates the kinase activity, phosphorylates tyrosine residues and activates diverse downstream signaling cascades, such as mitogen-activated protein kinase (MAPK), signal transducers and activators of transcription (STAT), and the phosphatidylinositol 3-OH kinase (PI3K) pathway [Citri and Yarden, 2006; Egan and Weinberg, 1993].

Amplification and overexpression of the ERBB2/HER2 receptor are observed in ~20% of breast cancers, and in other types of cancer with variable proportions, and are associated with poor prognosis and higher metastatic potential [Hurst, 2001; Kim and Scott, 2017; Vernimmen et al., 2003; Yarden and Sliwkowski, 2001]. Major clinical treatments of ERBB2-positive breast cancers include humanized monoclonal antibodies (trastuzumab and pertuzumab), which bind to the extracellular domain of ERBB2, disrupting the dimerization of the receptor and the downstream signaling [Franklin et al., 2004; Hudis, 2007], and tyrosine kinase inhibitors (TKI, lapatinib and neratinib), which antagonize the kinase activity of the homo- or heterodimerized receptors, inhibiting phosphorylation of their substrates and downstream signaling [Barok et al., 2014; Burris, 2004; Burstein et al., 2010; Diéras et al., 2017; Xia et al., 2005]. These therapies have greatly improved the overall survival of ERBB2-positive breast cancer patients [Slamon et al., 2001]. However, resistances to monoclonal antibody and TKI treatments are observed in the majority of patients [Blackwell et al., 2009; Esteva et al., 2002; Nahta et al., 2006]. Although different mechanisms have been proposed [Esteva et al., 2010; Spector and Blackwell, 2009], the resistance to anti-ERBB2 treatment remains a major clinical issue.

The presence of nuclear forms of RTKs, including the four members of the EGFR family, has been revealed since the last two decades [Carpenter, 2003; Lin et al., 2001; Ni et al., 2001; Offterdinger et al., 2002; Wang et al., 2004] and several mechanisms for the nuclear translocation and the function of those nuclear RTKs have been reported [Giri et al., 2005; Lin et al., 2001; Lo et al., 2006; Ni et al., 2001; Offterdinger et al., 2002; Rio et al., 2000; Wang et al., 2004; Williams et al., 2004]. Interestingly, recent studies shed new light on the role played by the nuclear form of ERBB2 in the resistance to anti-ERBB2 treatments [Cordo Russo et al., 2015, 2019; Li et al., 2011; Schillaci et al., 2012; Venturutti et al., 2016].

The *ERBB2* gene can be transcribed as at least 17 different transcripts generated by alternative

splicing and alternative promoter usage [Nezu et al., 1999]. All transcript variants are apparently under the control of two major promoters separated by 12 Kb (Fig. 1). The proximal promoter controls the transcription of *ERBB2* transcript variant 1 (NM_004448) which encodes the canonical form of the receptor ERBB2a, generally named ERBB2, while the 5'-end distal promoter controls the transcription of *ERBB2* transcript variant 2 (NM_001005862.2). *ERBB2* transcript variant 1 is composed of 27 exons. Its first exon codes for 30 amino acids, where the first 22 amino acids constitute a signal peptide which is cleaved during the migration of ERBB2 mature form to the cell membrane. *ERBB2* transcript variant 2 is composed of 30 exons where the first 4 exons are non-coding. Consequently, the translation of the ERBB2b isoform starts at an alternative ATG within its fifth exon which is identical to the second exon of variant 1 (Fig. 1). Therefore, sequences coding for the signal peptide in ERBB2a exon 1 are missing in the ERBB2b isoform. Because this signal peptide is cleaved once it transports the protein to the cell membrane, the sequence difference of mature proteins between ERBB2a and ERBB2b isoforms is only eight N-terminal amino acids. Beside its canonical location at cell membrane, alternative forms of the ERBB2/HER2 receptor have been observed into cell nuclei where they can escape current therapeutic approaches. These nuclear forms have been described as the result of at least two different mechanisms: proteolytic shedding of the extracellular domain of the full-length receptor and translation of the mRNA encoding ERBB2 from internal initiation codons [Anido et al., 2006; Christianson et al., 1998].

Here, we report a new mechanism for ERBB2 nuclear expression, where the ERBB2b isoform can directly translocate into the nucleus without requirement of an intermediate membrane location. siRNA analyses showed that ERBB2b isoform can repress the ERBB2a isoform, whereas ERBB2a isoform can activate the ERBB2b isoform. In addition, nuclear ERBB2 binding to its own promoter was revealed by chromatin immunoprecipitation (ChIP) assay. Altogether, our results provide new insights into the function of ERBB2 and provide an additional explanation for the resistance to therapies targeting the membrane-anchored ERBB2 in breast cancer.

2 MATERIALS AND METHODS

2.1 Cell lines and antibodies

The non-oncogenic human primary mammary epithelium cells HME-1 (Clontech, Mountain View, CA, USA) and the breast cancer cell lines BT474, BT-483, HCC-202, HCC-1806, HCC-1954, HCC-1569, MDA-MB-175, MDA-MB-361, MDA-MB-453, SK-BR-3, UACC-812, ZR-75-30, Br-Ca-Mz-01 (http://www.atcc.org) and SUM-185, SUM-190, SUM-206 (http://www.cancer.med.umich.edu/breast_cell/production) were grown according to the recommendations of the supplier. Twenty-eight breast carcinoma samples were obtained from patients treated at the Institut Paoli-Calmettes (Marseille, France), and corresponded to surgical specimens in case of early-stage disease and/or diagnostic biopsies in case of advanced stage. Their characteristics are summarized in Table 1 and detailed in Table 2. Twenty cases were ERBB2+, whereas 8 were ERBB2-. Each patient gave written informed consent and the study was approved by our institutional review board. The following antibodies were purchased from a commercial source: ERBB2 (Santa Cruz, sc-284), Sp1 (Upstate, 07-645), anti-β-tubulin (Abcam, ab11307), IgG Rabbit (ab171870).

2.2 ChIP-seq data processing

Reads were aligned to the hg19 genome using bowtie 1.1.1 [Langmead et al., 2009], with the parameters --all --best --strata -v 2 -m 1 -S. Peak detection was performed using MACS 1.4.2 [Zhang et al., 2008] using -g hs --keep-dup=auto -w -S. Annotation of peaks was performed using Homer annotatePeaks (version 4.7.2) [Heinz et al., 2010]. Public BT474 rabbit IgG and RNAPII HCC-1806 data were recovered from the Gene Expression Omnibus (GEO) from accessions GSM1145935 [Heinonen et al., 2015] and GSM3317507 [Perreault et al., 2019], respectively. Intersection of peaks was carried out using bedtools intersect 2.19.0 [Quinlan and Hall, 2010]. Annotation to the closest gene was performed using bedtools closest 2.19.0 using -t first as a parameter. Read depth coverage were normalized to BT474 Pol II ChIP-seq data.

2.3 siRNA Assay

Three 25-mer duplex small interfering RNAs (siRNAs) to specifically target either ERBB2a or b mRNA variants were designed and ordered (Select Stealth™ RNAi, Invitrogen, Table 3). A mixed of each three siRNA duplexes (10 nM) and a mixed Stealth™ RNAi Negative Control Duplexes were transiently transfected by Lipofectamine™ RNAiMAX (Invitrogen) for 48 h at 37°C in a CO_2 incubator according to the manufacturer instruction. Gene knockdown was confirmed by quantitative reverse transcription-PCR and immunoblotting.

2.4 ERBB2a- / ERBB2b-GFP construction

The full-length cDNAs of the two isoforms ERBB2a and b were obtained by reverse transcription (Super Script ™ First-Strand, Invitrogen) of mRNA extracted from BT474 cells (RNeasy® Micro Kit, Qiagen) and amplified by PCR (Platinum® TaqDNA Polymerase High Fidelity, Invitrogen) using Gateway Specific Primers: ERBB2a-GW-F 5'-AttB2-GGGGACAAGTTTGTACAAAAAAGCAGGCTAGCCGCAGTGAGCACCAT-3'; ERBB2b-GW-F 5'-AttB2-GGGGACAAGTTTGTACAAAAAAGCAGGCTTGTGCACCGGCACAGACAT-3'; ERBB2-GW-R 5'-AttB1-GGGGACCACTTTGTACAAGAAAGCTGGGTGACTTGGCCTTCTGGTTCA-3'. They were subsequently cloned into the expression vectors pEGFP, according to the Gateway protocol (Invitrogen).

2.5 Transient transfection

Cells were transfected using FuGENE 6 reagent (Roche Diagnostics). $3x10^5$ cells were seeded in 6-well plates (Falcon 3046, BD Biosciences, Lincoln Park, NJ) and then transfected the next day using the FuGENE reagent (FuGENE:DNA ratio of 3 µl: 2 µg) for 48 h in complete medium.

2.6 RNA extraction and Reverse-Transcription PCR

Total RNA was extracted from cells using RNeasy® Micro Kit (QIAGEN) according to the manufacturer's protocol. Total RNA was quantified using a Nanodrop 1000 Spectrophotometer device. RT-PCR was performed using SuperScript™ II Reverse Transcriptase (Invitrogen) and random hexamer primers.

2.7 Subcellular fractionation

The cytosolic, membrane, and nuclear fractions were prepared from fractions 1, 2 and 3 using ProteoExtract® Subcellular Proteome Extraction (Calbiochem, Merck). The fractionation efficiency was assessed by antibodies against Sp1 and β-tubulin.

2.8 Immunofluorescence

Cultured cells were fixed with 4% paraformaldehyde for 20 min, permeabilized with 0.1% Tween-20 for 10 min, and then immunostained with primary antibodies against ERBB2 (1:200 dilution in PBS with 1% BSA) for 1 hr at room temperature. After three washes with PBS, the Alexa 488-conjugated secondary antibody was then applied for 45 min at room temperature. The nucleus was stained with DAPI before mounting. For immunofluorescence microscopy, the images were captured with a Zeiss 510 Meta UV confocal microscope.

2.9 Western blot

10^5 cultured cells were lysed in 10 µl of lysis buffer (25 mM Tris pH 7.9, 1% w/v SDS, 1 mM dithiothreitol) and fractionated on a 7.5% SDS-polyacrylamide gel, electrotransferred to a PVDF membrane. Specific polypeptides were revealed with the rabbit polyclonal anti-ERBB2 (C18, sc-284), rabbit polyclonal anti-Sp1 (Upstate, 07-645), and mouse monoclonal anti-β-tubulin (Abcam, ab11307) at a proper dilution in TBS-Tween containing 5% non-fat milk or BSA according to the recommendations of the supplier, and then reacted with secondary antibodies at a dilution of $1:10^4$ in the same buffer.

2.10 Chromatin Immunoprecipitation Assay

BT474 cells were fixed with 1% formaldehyde for 15 min, and the cross-linking reaction was stopped by the addition of 125 mM glycine incubating 5 min. Cells were washed and lysed in three Lysis Buffers containing 1x Complete Protease Inhibitors Cocktail (Roche) respectively. 1% Triton X-100 was added to the lysate and all lysates were sonicated at 4°C to generate 100-300 base pair chromatin fragments. Pretreated mixture of 3 µg of antibody (rabbit polyclonal anti-ERBB2, C18 sc-284 or rabbit polyclonal anti-PolII, N20 sc-899), and protein G-conjugated magnetic Dynabeads (Dynal Biotech, Invitrogen) was added to the supernatant and incubated on a rotating wheel at 4°C for overnight. The immunocomplexes were then washed with RIPA wash buffer (S1) for 6 times and once with TE containing 50 mM NaCl. The Ab–protein–DNA complexes were eluted twice in 100 µl elution buffer at 65°C for 15 min, briefly vortexed every 2 min. The combined eluates as well as the input sample (1% of the amount used in the IP procedure) were treated to reverse the crosslink by incubating overnight at 65°C. After Rnase and proteinase K digestion, the DNA fragments were extracted using phenol-chloroform and purified using a Qiagen PCR purification kit.

2.11 Quantitative PCR and normalization

DNA was quantified using LC FastStart DNA Master SYBR Green I and read with a Light Cycler® 2 instrument (Roche Diagnostics) according to the manufacturer. For cDNA analysis, threshold cycles (Ct) were determined for quantification of the input target number. For comparison between different biological samples, the target gene expression level in each breast cancer cell line and breast carcinoma sample was normalized using a calibrator RNA sample extracted from the HME-1 cell line. The normalized gene expression (NGE) of the target gene was expressed as the N_{ratio}, where each normalized gene expression gene value was divided by a calibrator normalized gene value (*TBP*): $N_{target} = (Ct_{target,HME\text{-}1}\text{-}Ct_{TBP,HME\text{-}1}) - (Ct_{target,sample}\text{-}Ct_{TBP,sample})$. Primers used in this study are listed in Table 4.

For RNA interference assay, the relative gene expression (RGE) was calculated as a percentage of the measured mRNA level in cells transfected with target siRNA compared to cells transfected with matched control siRNA. Data were normalized against the *TBP* signal and the $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression level of a given gene [Livak and Schmittgen, 2001].

For chromatin immunoprecipitation (ChIP) assay, the amount of target and endogenous reference DNA was determined from a standard curve. The standard curve was constructed with four-fold serial dilutions of input DNA. The results from specific antibodies were reported to that observed in the corresponding IgG samples. A region between the *GAPDH* and *CNAP1* genes was used as the negative control to standardize the results from COX2 and ERBB2 HAS regions (ChIP-It protocol, Active Motif). Primers used in ChIP assays are: ERBB2 F 5'- TCCTTTCGATGTGACTGTCTCC – 3', ERBB2 R 5'- CTAAATGCAGAGGCTGGTGACT – 3'; COX2 F 5'- AACATGGCTTCTAACCCAAAC – 3', COX2 R 5'- AGGAAGCCTTTCTCCTCCTC – 3'; Neg F 5'- ATGGTTGCCACTGGGGATCT -3', Neg R 5'- TGCCAAAGCCTAGGGGAAGA- 3'.

2.12 Motif discovery and average profiles

A custom position weight matrix was created for HAS motifs [Wang et al., 2004]. Mapping of motifs to specific and shared peaks was performed via Homer annotatePeaks using -hist 10 -m <motif> -size 2000 as parameters. Average profiles were computed for each peak population. Motif discovery was performed using Homer findMotifsGenome using hg19 as a parameter.

2.13 Gene Set Enrichment Analyses

Publicly available gene expression data for BT474 and HCC-1806 cells were retrieved from GEO accessions GSM1230347 and GSM1230072 [Dezső et al., 2014; Milosevic et al., 2013] and were processed in R using the affy package. Gene Set Enrichment Analyses (GSEAs) were computed using GSEA 2.2.4 [Subramanian et al., 2005], using the top RNAPII 500 peaks in summit height for BT474 and HCC-1806 as the genesets, the GENE_SYMBOL.chip chip platform, gene_set permutation for p-value computation, $\log^2$ ratio of classes as the ranking metric, and weighted statistic.

2.14 Statistical Analysis

Statistical analyses were done using the SPSS 16.0 program for Windows. The Spearman's correlation coefficient (ρ) was used to compare the correlation between two target gene expression levels in breast cancer cell lines and tumor samples. A p-value < 0.05 was set as the criteria for statistical significance, while a p-value < 0.01 was considered highly significant. For p-values pertaining to motif densities, paired, two-tailed student t-tests were performed.

3 RESULTS

3.1 Both proximal and distal ERBB2 gene promoters are actively used in breast cancer cells

To confirm that both proximal and distal ERBB2 gene promoters are actively used in breast cancer cells, we performed a PolII ChIP-seq assay using BT474 cells and extracted data pertinent to the *ERBB2* gene locus. As shown in Figure 1A, PolII was significantly recruited at both promoters with predominance for the well characterized proximal promoter that controls the expression of the canonical transmembrane form of this RTK. Pol II signal was indeed enriched at both promoters as compared to a publicly available IgG ChIP-Seq control dataset performed in BT474 cells [Heinonen et al., 2015]. The recruitment of PolII at the distal promoter clearly indicates that both promoters might contribute to ERBB2 overexpression in breast cancer cells.

3.2 Prevalent presence of ERBB2b isoform in breast cancer

To determine the relative expression of the *ERBB2a* and *ERBB2b* mRNA isoforms in human breast cancer, mRNAs were extracted from 16 breast cancer cell lines and 28 clinical tumour samples (see Table 1 and Table 2). The human normal epithelial breast cell line (HME-1) was used as a reference to measure the relative gene expression (RGE). Quantitative-RT-PCRs using primers specific for either *ERBB2a* or *ERBB2b* isoform showed that both *ERBB2a* and *ERBB2b* isoforms were expressed in 15 breast cancer cell lines and all carcinoma samples (Fig. 2A-B), whereas the *ERBB2b* isoform was not expressed in the HCC-1569 cell line. Importantly, *ERBB2b* expression was highly correlated to that of *ERBB2a* with a Spearman rank correlation coefficient equal to 0.912 and 0.882 for cell lines (without HCC-1569) and carcinoma samples, respectively (Fig. 2C-D). Because the amplification of the chromosome 17q12 region is often observed in various solid tumours [Kao and Pollack, 2006; Maqani et al., 2006], three genes adjacent to the *ERBB2* locus within this chromosome region, *PERLD1* (*PGAP,* NM_033419.5)*, C17orf37* (*MIEN1*, NM_032339.5), and *GRB7* (NM_005310.5) were also analysed for their RGEs in this series of breast cancer samples. Highly positive correlations were observed among all five transcripts. The highest correlation was observed around the *ERBB2* locus (Fig. 2E). These results indicated the prevalent presence of *ERBB2b* isoform in breast cancer.

3.3 ERBB2b isoform is a nuclear form of ERBB2

To study the protein expression of the two ERBB2 isoforms in different cell compartments, four breast cancer cell lines were selected based on the RGE of the two *ERBB2* isoforms: BT474 and HCC-1954 overexpress both isoforms, whereas HCC-1569 overexpresses only *ERBB2a*, and Br-Ca-Mz-01 weakly expresses *ERBB2b* but not *ERBB2a* (Fig. 2A). Cytosolic, membrane, and nuclear fractions were

separated for each of these four cell lines. ERBB2 protein expression was observed mainly in the membrane fraction, but also in the cytosolic and nuclear fractions in BT474 and HCC-1954 cells, which is expected for a transmembrane receptor (Fig. 3A). Interestingly, no ERBB2 nuclear expression was revealed in HCC-1569 cells (Fig. 3A). However, no detectable expression of ERBB2 by Western-blot was observed for Br-Ca-Mz-01 cell lines probably due to its very weak expression level (Fig. 3A).

Immunofluorescence assays were used to precisely determine the expression location of ERBB2a and ERBB2b in single breast cancer cells. Confocal microscopy analyses revealed that clear membrane ERBB2 expression was observed in BT474, HCC-1954, and HCC-1569 cells, but not in Br-Ca-Mz-01 cells, which is in agreement with the lack of *ERBB2a* isoform expression (Fig. 3B). However, nuclear ERBB2 expression was observed in all the four tested cell lines, including HCC-1569 in which *ERBB2b* mRNA was not detected by transcription analysis (Fig. 3B). Altogether these results demonstrated that ERBB2b isoform does not contribute to a detectable membrane expression of ERBB2.

To confirm that ERBB2b is a *bona fide* nuclear form of ERBB2, we constructed two expression vectors that contain either *ERBB2a* or *ERBB2b* cDNA fused to green fluorescent protein (GFP) using the Gateway cloning system (Invitrogen Tm). These two vectors were then transiently transfected into BT474 (high ERBB2 expression) and HCC-1806 (low ERBB2 expression) cell lines. ERBB2-GFP expression was observed in whole cells (Fig. 3C left panels) with the transfection of *ERBB2a*, whereas ERBB2-GFP expression was specifically located in cell nucleus upon transfection of the *ERBB2b* isoform (Fig. 3C right panels). These results clearly demonstrated that *ERBB2* mRNA variant 2 (*ERBB2b*) encodes a *bona fide* nuclear isoform of ERBB2.

3.4 ERBB2a isoform upregulates ERBB2b whereas ERBB2b isoform inhibits ERBB2a

To study putative functional relations between ERBB2a and ERBB2b isoforms, RNA interference assays were carried out in BT474 and HCC-1954 cell lines, in which both isoforms are overexpressed. Three specific custom Stealth™ siRNAs (Invitrogen) were designed against each isoform and transiently transfected into two cell lines (Fig. 4A). Control siRNAs were simultaneously transfected into cells and resulting RGEs were defined as 100%. qRT-PCR showed that ERBB2a siRNAs inhibited 30-40% of *ERBB2a* isoform expression, and 50-70% of *ERBB2b* isoform expression in both cell lines (Fig. 4B). However, in marked contrast, ERBB2b siRNAs induced a significant increase of *ERBB2a* expression, while efficiently inhibited *ERBB2b* expression (Fig. 4B). These results suggested that the ERBB2a isoform upregulates the expression of the *ERBB2b* isoform, whereas ERBB2b inhibits the expression of *ERBB2a*.

Immunofluorescence assay were further carried out in HCC-1954 and Br-Ca-Mz-01 cell lines treated with siRNAs against either ERBB2a or ERBB2b isoforms. In HCC-1954 cells, the diffusion of

ERBB2 protein expression was observed upon transfection of ERBB2a siRNAs: less ERBB2 membrane protein was observed and nuclear proteins were present (Fig. 4C mid panels). In marked contrast, ERBB2 protein expression was strongly increased in membranes, and reduced in nuclei when cells were transfected with ERBB2b siRNAs (Fig. 4C lower panels). In Br-Ca-Mz-01 cells, the already weak cytoplasmic and nuclear ERBB2 expressions were further reduced by both ERBB2a and ERBB2b siRNAs (Fig. 4D). However, no convincing ERBB2 membrane accumulation was observed when ERBB2b expression was inhibited, due to the low expression of ERBB2a (Fig. 4D). These observations, together with the siRNA assays, strongly suggest a 2-stage feedback interaction between ERBB2a and ERBB2b where the ERBB2b isoform upregulated by ERBB2a can subsequently downregulates ERBB2a expression.

3.5 Binding of nuclear ERBB2 to a specific HAS motif of the *ERBB2* gene proximal promoter

Nuclear ERBB2 was reported to bind to the *COX2* gene promoter *via* an HER2/ERBB2-associated sequence (HAS), a specific AT-rich DNA motif [Wang et al., 2004]. By searching a similar HAS regulatory element in the vicinity of the *ERBB2* gene promoter, a potential ERBB2-binding sequence was found 500 bps upstream of the transcriptional start site of *ERBB2a* isoform (NM_004448) (Fig. 5A). Chromatin immunoprecipitation assays in BT474 cells confirmed the binding of nuclear ERBB2 to the *COX2* gene promoter *via* HAS (Fig. 5B). Most importantly, our ChIP assay also revealed the binding of nuclear ERBB2 to *ERBB2* proximal promoter region (Fig. 5B).

3.6 Regulatory regions are enriched in HAS motifs and correlate with increased gene expression in BT474 cells

We next sought to determine whether HAS motifs were enriched in other regulatory regions in BT474 cells. To this end, we used the RNAPII ChIP-seq that we performed above (Fig. 1A) and performed peak detection. This analysis identified 13,897 peaks, which mostly corresponded to distal regions, with ~1/5 of peaks annotated to promoter and 5' UTR regions (Fig. 6A). We hypothesized that if ERBB2b binds to certain regulatory regions in BT474 cells, those regions would not be present in ERBB2b low-expressing cells. We thus compared our genome-wide mapping of RNAPII with that performed in HCC-1806 [Perreault et al., 2019], a triple-negative, low- *ERBB2a* and *ERBB2b* expressing breast cancer cell line (Fig. 2A and Fig. 6B). By mapping HAS motifs to BT474-, HCC-1806- specific and shared peaks, we observed that BT474-specific peaks were significantly more enriched in HAS motifs than shared or HCC-1806-specific peaks (Fig. 6C). Further, unsupervised motif discovery revealed a compatible motif enriched in BT474-specific RNAPII peaks (Fig. 6D), consistent with putative ERBB2b binding in those regions. We next aimed to determine whether the elements identified via genome-wide mapping of RNAPII targets corresponded to putative enhancers regions. Importantly, gene set enrichment analysis

(GSEAs) [Subramanian et al., 2005] done using previously published BT474 and HCC-1806 gene expression data [Dezső et al., 2014; Milosevic et al., 2013] showed positive, significant correlations between RNAPII peaks and gene expression fold change in BT474 as well as in HCC-1806 cells, consistent with a trans-activating roles for those regions (Fig. 6E).

4 DISCUSSION

Nuclear translocation of receptor tyrosine kinases (RTKs) has been described for twenty years. Here, we report a new mechanism for a constant and ubiquitous presence of a nuclear form of ERBB2. The ERBB2b isoform, lacking a signal peptide as compared to the canonical ERBB2a isoform (see Fig. 1), most probably cannot be expressed at the cell membrane. This was confirmed by immunofluorescence analyses of Br-Ca-Mz-01 cells, which do not express the ERBB2a isoform (Fig. 3B). Our results also indicate that membrane ERBB2 is mostly if not exclusively expressed by the *ERBB2a* isoform. Importantly, transient transfection of an *ERBB2b* expression vector coupled with a GFP reporter gene clearly demonstrated that ERBB2b is a nuclear form of ERBB2 (Fig. 3C). It was reported that membrane ERBB2 could contribute to the nuclear translocation as a consequence of endocytic internalization [Giri et al., 2005], which could explain the nuclear expression of ERBB2 observed in HCC-1569 cells which express only ERBB2a isoform, but not ERBB2b isoform (Fig. 3B). Nuclear ERBB2 has been reported to regulate target gene expression through binding to HASs or tethered binding to other transcription factors [Béguelin et al., 2010; Diaz Flaqué et al., 2013; Li et al., 2011; Proietti et al., 2009; Venturutti et al., 2016; Wang et al., 2004]. Accordingly, we observed that nuclear ERBB2 binds to a HAS located in the promoter region of the *ERBB2* gene (Fig. 5). However, we could not exclude that nuclear ERBB2 binds to this sequence through tethered binding to other transcription factors, since no DNA-binding domain has been yet identified in the ERBB2 protein.

Transcriptional analyses from both breast cancer cell lines and clinical samples revealed a positive correlation between the expression of *ERBB2a a*nd *ERBB2b* mRNAs (Fig. 2A-D). Using siRNAs targeting specifically and independently both mRNA species, we have established a functional link between these two isoforms, where ERBB2a could upregulate the ERBB2b, in agreement with the positive correlation observed between these two isoforms (Fig. 4). However, our siRNA assays provided an unexpected result when overexpressed ERBB2b inhibited ERBB2a mRNA expression in breast cancer cell lines. This apparent discrepancy could be explained when considering that the positive correlation between these two isoforms were observed in breast cancer cell lines and carcinomas samples under near physiological condition, whereas siRNA assays knock-out target gene expression and alter its downstream gene expression which subsequently break the physiological balance between these two ERBB2 isoforms. In addition and perhaps more importantly, the well-

known amplification of the ERBB2 locus in chromosome region 17q12 in breast cancers may also contribute to this positive correlation between the two isoforms by overcoming ERBB2b repressive activity as previously hypothesized for several other transcriptional repressors of the *ERBB2* gene such as FOXP3, GATA4, MYB, PAX3 [Birnbaum et al., 2009]. Further *in vivo* experiments using transgenic mouse models in which the expression of either isoform would be knocked-out or knocked-in should help to further decipher the respective functional role played by each ERBB2 isoform.

Nuclear ERBB2 has been reported to be involved in tumour growth and metastasis as well as resistance to anti-ERBB2 treatment in breast cancers [Béguelin et al., 2010; Cordo Russo et al., 2019; Diaz Flaqué et al., 2013; Díaz Flaqué et al., 2013; Venturutti et al., 2016; Wang et al., 2004]. C-terminal fragments of ERBB2 generated by alternative initiation from the internal AUG codon have also been shown to have tumorigenic property [Anido et al., 2006; Arribas et al., 2011; Vernieri et al., 2019]. In this study, we provide evidence that nuclear ERBB2, in addition to the two known origins (proteolytic shedding of the extracellular domain of the full-length receptor, translation of the mRNA encoding *ERBB2/HER2* from internal initiation codons) also result from the translation of the *ERBB2* mRNA variant 2 lacking the exons encoding the first 30 amino acids that are present in the ERBB2a full-length isoform. Our results together with published observations strongly indicate that the ERBB2a and ERBB2b isoforms may have different functions both in normal and breast cancer cells. Thus, a specific antibody against the small peptide encoded by amino acids 23-30 of the ERBB2a isoform might help to further study the role of nuclear ERBB2. With such discriminative tool, the composition of ERBB2 nuclear fraction would be further characterize, and specifically study each ERBB2 isoform both in normal and breast cancer cells, which could contribute to the screen and identification of new potential treatments in HER2-positive breast cancer.

Receptor-type tyrosine kinases (RTK) are major regulators of cellular processes that are often mutated resulting in their constitutive activation in human cancers. In contrast to many oncogenic RTKs oncogenes, some others such as ERBB2 are amplified without mutation or rearrangement. We have previously hypothesized that ERBB2 gene amplification is used to overcome repression of its expression by several sequence-specific transcription factors [Birnbaum et al., 2009]. The present characterization of a new negative feedback autoregulatory loop where ERBB2b can repress ERBB2a fits with this model, adding one more member in the list of the transcription factors known to repress *ERBB2* gene expression [Hua et al., 2009].

ACKNOWLEDGEMENTS

We thank Karim Ben Salah, Hélène Holota (TAGC, Marseille) and José Adélaïde (CRCM, Marseille) for excellent technique support. We thank Keiji Zhao (NHLBI/NIH, Bethesda) for his invaluable assistance

with the ChiP-seq assays. This work was supported by Association pour la recherche sur le cancer (ARC), Fondation Groupe EDF, Institut national du cancer (INCA), Ligue Nationale Contre le Cancer (EL2016 Birnbaum and EL2019 Bertucci), the European Molecular Biology Organization (EMBO Short Term Fellowship ASTF 25.00-2009 HUA) and Boehringer Ingelheim Fonds (Travel Grant 2008 HUA).

CONFLICT OF INTEREST

The authors here declare that no conflicts of interest exist during the completion of the paper.

AUTHOR CONTRIBUTIONS

G.H and J.I designed the research. G.H, A.B, P.C and B.K-P performed the experiments. G.H, A.B, P.C, B.K-P and J.I analysed the data. G.H, P.C and J.I wrote the manuscript. F.B, D.B and N.B-J revised the manuscript. The paper was approved by all authors.

REFERENCES

Akiyama T, Sudo C, Ogawara H, Toyoshima K, Yamamoto T. 1986. The product of the human c-erbB-2 gene: a 185-kilodalton glycoprotein with tyrosine kinase activity. Science 232(4758):1644–1646.

Anido J, Scaltriti M, Bech Serra JJ, Santiago Josefat B, Todo FR, Baselga J, Arribas J. 2006. Biosynthesis of tumorigenic HER2 C-terminal fragments by alternative initiation of translation. EMBO J 25(13):3234–3244.

Arribas J, Baselga J, Pedersen K, Parra-Palau JL. 2011. p95HER2 and breast cancer. Cancer Res 71(5):1515–1519.

Barok M, Joensuu H, Isola J. 2014. Trastuzumab emtansine: mechanisms of action and drug resistance. Breast Cancer Res 16(2):209.

Béguelin W, Díaz Flaqué MC, Proietti CJ, Cayrol F, Rivas MA, Tkach M, Rosemblit C, Tocci JM, Charreau EH, Schillaci R, Elizalde PV. 2010. Progesterone Receptor Induces ErbB-2 Nuclear Translocation To Promote Breast Cancer Growth via a Novel Transcriptional Effect: ErbB-2 Function as a Coactivator of Stat3. Mol Cell Biol 30(23):5456–5472.

Birnbaum D, Sircoulomb F, Imbert J. 2009. A reason why the ERBB2 gene is amplified and not mutated in breast cancer. Cancer Cell Int 9:5.

Blackwell KL, Pegram MD, Tan-Chiu E, Schwartzberg LS, Arbushites MC, Maltzman JD, Forster JK, Rubin SD, Stein SH, Burstein HJ. 2009. Single-agent lapatinib for HER2-overexpressing advanced or metastatic breast cancer that progressed on first- or second-line trastuzumab-containing regimens. Ann Oncol 20(6):1026–1031.

Burris HA. 2004. Dual kinase inhibition in the treatment of breast cancer: initial experience with the EGFR/ErbB-2 inhibitor lapatinib. Oncologist 9 Suppl 3:10–15.

Burstein HJ, Sun Y, Dirix LY, Jiang Z, Paridaens R, Tan AR, Awada A, Ranade A, Jiao S, Schwartz G, Abbas R, Powell C, Turnbull K, Vermette J, Zacharchuk C, Badwe R. 2010. Neratinib, an irreversible

ErbB receptor tyrosine kinase inhibitor, in patients with advanced ErbB2-positive breast cancer. J Clin Oncol 28(8):1301–1307.

Carpenter G. 2003. Nuclear localization and possible functions of receptor tyrosine kinases. Curr Opin Cell Biol 15(2):143–148.

Christianson TA, Doherty JK, Lin YJ, Ramsey EE, Holmes R, Keenan EJ, Clinton GM. 1998. NH2-terminally truncated HER-2/neu protein: relationship with shedding of the extracellular domain and with prognostic factors in breast cancer. Cancer Res 58(22):5123–5129.

Citri A, Yarden Y. 2006. EGF-ERBB signalling: towards the systems level. Nat Rev Mol Cell Biol 7(7):505–516.

Cordo Russo RI, Béguelin W, Díaz Flaqué MC, Proietti CJ, Venturutti L, Galigniana N, Tkach M, Guzmán P, Roa JC, O'Brien NA, Charreau EH, Schillaci R, Elizalde PV. 2015. Targeting ErbB-2 nuclear localization and function inhibits breast cancer growth and overcomes trastuzumab resistance. Oncogene 34(26):3413–3428.

Cordo Russo RI, Chervo MF, Madera S, Charreau EH, Elizalde PV. 2019. Nuclear ErbB-2: a Novel Therapeutic Target in ErbB-2-Positive Breast Cancer? Horm Cancer 10(2–3):64–70.

Dezső Z, Oestreicher J, Weaver A, Santiago S, Agoulnik S, Chow J, Oda Y, Funahashi Y. 2014. Gene expression profiling reveals epithelial mesenchymal transition (EMT) genes can selectively differentiate eribulin sensitive breast cancer cells. PLoS ONE 9(8):e106131.

Díaz Flaqué MC, Galigniana NM, Béguelin W, Vicario R, Proietti CJ, Russo R, Rivas MA, Tkach M, Guzmán P, Roa JC, Maronna E, Pineda V, Muñoz S, Mercogliano M, Charreau EH, Yankilevich P, Schillaci R, Elizalde PV. 2013. Progesterone receptor assembly of a transcriptional complex along with activator protein 1, signal transducer and activator of transcription 3 and ErbB-2 governs breast cancer growth and predicts response to endocrine therapy. Breast Cancer Res 15(6):R118.

Diaz Flaqué MC, Vicario R, Proietti CJ, Izzo F, Schillaci R, Elizalde PV. 2013. Progestin drives breast cancer growth by inducing p21(CIP1) expression through the assembly of a transcriptional complex among Stat3, progesterone receptor and ErbB-2. Steroids 78(6):559–567.

Diéras V, Miles D, Verma S, Pegram M, Welslau M, Baselga J, Krop IE, Blackwell K, Hoersch S, Xu J, Green M, Gianni L. 2017. Trastuzumab emtansine versus capecitabine plus lapatinib in patients with previously treated HER2-positive advanced breast cancer (EMILIA): a descriptive analysis of final overall survival results from a randomised, open-label, phase 3 trial. Lancet Oncol 18(6):732–742.

Egan SE, Weinberg RA. 1993. The pathway to signal achievement. Nature 365(6449):781–783.

Esteva FJ, Valero V, Booser D, Guerra LT, Murray JL, Pusztai L, Cristofanilli M, Arun B, Esmaeli B, Fritsche HA, Sneige N, Smith TL, Hortobagyi GN. 2002. Phase II study of weekly docetaxel and trastuzumab for patients with HER-2-overexpressing metastatic breast cancer. J Clin Oncol 20(7):1800–1808.

Esteva FJ, Yu D, Hung M-C, Hortobagyi GN. 2010. Molecular predictors of response to trastuzumab and lapatinib in breast cancer. Nat Rev Clin Oncol 7(2):98–107.

Franklin MC, Carey KD, Vajdos FF, Leahy DJ, Vos AM de, Sliwkowski MX. 2004. Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell 5(4):317–328.

Giri DK, Ali-Seyed M, Li L-Y, Lee D-F, Ling P, Bartholomeusz G, Wang S-C, Hung M-C. 2005. Endosomal transport of ErbB-2: mechanism for nuclear entry of the cell surface receptor. Mol Cell Biol 25(24):11005–11018.

Graus-Porta D, Beerli RR, Daly JM, Hynes NE. 1997. ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. EMBO J 16(7):1647–1655.

Heinonen H, Lepikhova T, Sahu B, Pehkonen H, Pihlajamaa P, Louhimo R, Gao P, Wei G-H, Hautaniemi S, Jänne OA, Monni O. 2015. Identification of several potential chromatin binding sites of HOXB7 and its downstream target genes in breast cancer. Int J Cancer 137(10):2374–2383.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38(4):576–589.

Hua G, Zhu B, Rosa F, Deblon N, Adélaïde J, Kahn-Perlès B, Birnbaum D, Imbert J. 2009. A Negative Feedback Regulatory Loop Associates the Tyrosine Kinase Receptor ERBB2 and the Transcription Factor GATA4 in Breast Cancer Cells. Mol Cancer Res 7(3):402–414.

Hudis CA. 2007. Trastuzumab--mechanism of action and use in clinical practice. N Engl J Med 357(1):39–51.

Hurst HC. 2001. Update on HER-2 as a target for cancer therapy: the ERBB2 promoter and its exploitation for cancer treatment. Breast Cancer Res 3(6):395–398.

Kao J, Pollack JR. 2006. RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes. Genes Chromosomes Cancer 45(8):761–769.

Kim ES, Scott LJ. 2017. Palbociclib: A Review in HR-Positive, HER2-Negative, Advanced or Metastatic Breast Cancer. Target Oncol 12(3):373–383.

Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, Yarden Y. 1999. The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. Proc Natl Acad Sci USA 96(9):4995–5000.

Kraus MH, Issing W, Miki T, Popescu NC, Aaronson SA. 1989. Isolation and characterization of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: evidence for overexpression in a subset of human mammary tumors. Proc Natl Acad Sci USA 86(23):9193–9197.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25.

Li L-Y, Chen H, Hsieh Y-H, Wang Y-N, Chu H-J, Chen Y-H, Chen H-Y, Chien P-J, Ma H-T, Tsai H-C, Lai C-C, Sher Y-P, Lien H-C, Tsai C-H, Hung M-C. 2011. Nuclear ErbB2 enhances translation and cell growth by activating transcription of ribosomal RNA genes. Cancer Res 71(12):4269–4279.

Lin SY, Makino K, Xia W, Matin A, Wen Y, Kwong KY, Bourguignon L, Hung MC. 2001. Nuclear localization of EGF receptor and its potential new role as a transcription factor. Nat Cell Biol 3(9):802–808.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402–408.

Lo H-W, Ali-Seyed M, Wu Y, Bartholomeusz G, Hsu S-C, Hung M-C. 2006. Nuclear-cytoplasmic transport of EGFR involves receptor endocytosis, importin beta1 and CRM1. J Cell Biochem 98(6):1570–1583.

Maqani N, Belkhiri A, Moskaluk C, Knuutila S, Dar AA, El-Rifai W. 2006. Molecular dissection of 17q12 amplicon in upper gastrointestinal adenocarcinomas. Mol Cancer Res 4(7):449–455.

Milosevic J, Klinge J, Borg A-L, Foukakis T, Bergh J, Tobin NP. 2013. Clinical instability of breast cancer markers is reflected in long-term in vitro estrogen deprivation studies. BMC Cancer 13:473.

Nahta R, Yu D, Hung M-C, Hortobagyi GN, Esteva FJ. 2006. Mechanisms of disease: understanding resistance to HER2-targeted therapy in human breast cancer. Nat Clin Pract Oncol 3(5):269–280.

Nezu M, Sasaki H, Kuwahara Y, Ochiya T, Yamada Y, Sakamoto H, Tashiro H, Yamazaki M, Ikeuchi T, Saito Y, Terada M. 1999. Identification of a novel promoter and exons of the c-ERBB-2 gene. Biochem Biophys Res Commun 258(3):499–505.

Ni CY, Murphy MP, Golde TE, Carpenter G. 2001. gamma -Secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. Science 294(5549):2179–2181.

Offterdinger M, Schöfer C, Weipoltshammer K, Grunt TW. 2002. c-erbB-3: a nuclear protein in mammary epithelial cells. J Cell Biol 157(6):929–939.

Perreault AA, Sprunger DM, Venters BJ. 2019. Epigenetic and transcriptional profiling of triple negative breast cancer. Sci Data 6:190033.

Plowman GD, Culouscou JM, Whitney GS, Green JM, Carlton GW, Foy L, Neubauer MG, Shoyab M. 1993. Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal growth factor receptor family. Proc Natl Acad Sci USA 90(5):1746–1750.

Proietti CJ, Rosemblit C, Beguelin W, Rivas MA, Díaz Flaqué MC, Charreau EH, Schillaci R, Elizalde PV. 2009. Activation of Stat3 by heregulin/ErbB-2 through the co-option of progesterone receptor signaling drives breast cancer growth. Mol Cell Biol 29(5):1249–1265.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842.

Rio C, Buxbaum JD, Peschon JJ, Corfas G. 2000. Tumor necrosis factor-alpha-converting enzyme is required for cleavage of erbB4/HER4. J Biol Chem 275(14):10379–10387.

Schillaci R, Guzmán P, Cayrol F, Beguelin W, Díaz Flaqué MC, Proietti CJ, Pineda V, Palazzi J, Frahm I, Charreau EH, Maronna E, Roa JC, Elizalde PV. 2012. Clinical relevance of ErbB-2/HER2 nuclear expression in breast cancer. BMC Cancer 12:74.

Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L. 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 344(11):783–792.

Spector NL, Blackwell KL. 2009. Understanding the mechanisms behind trastuzumab therapy for human epidermal growth factor receptor 2-positive breast cancer. J Clin Oncol 27(34):5838–5847.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102(43):15545–15550.

Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J. 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature 309(5967):418–425.

Venturutti L, Romero LV, Urtreger AJ, Chervo MF, Cordo Russo RI, Mercogliano MF, Inurrigarro G, Pereyra MG, Proietti CJ, Izzo F, Díaz Flaqué MC, Sundblad V, Roa JC, Guzmán P, Bal de Kier Joffé ED, Charreau EH, Schillaci R, Elizalde PV. 2016. Stat3 regulates ErbB-2 expression and co-opts ErbB-2 nuclear function to induce miR-21 expression, PDCD4 downregulation and breast cancer metastasis. Oncogene 35(17):2208–2222.

Vernieri C, Milano M, Brambilla M, Mennitto A, Maggi C, Cona MS, Prisciandaro M, Fabbroni C, Celio L, Mariani G, Bianchi GV, Capri G, Braud F de. 2019. Resistance mechanisms to anti-HER2 therapies in HER2-positive breast cancer: Current knowledge, new research directions and therapeutic perspectives. Crit Rev Oncol Hematol 139:53–66.

Vernimmen D, Gueders M, Pisvin S, Delvenne P, Winkler R. 2003. Different mechanisms are implicated in ERBB2 gene overexpression in breast and in other cancers. Br J Cancer 89(5):899–906.

Wang S-C, Lien H-C, Xia W, Chen I-F, Lo H-W, Wang Z, Ali-Seyed M, Lee D-F, Bartholomeusz G, Ou-Yang F, Giri DK, Hung M-C. 2004. Binding at and transactivation of the COX-2 promoter by nuclear tyrosine kinase receptor ErbB-2. Cancer Cell 6(3):251–261.

Williams CC, Allison JG, Vidal GA, Burow ME, Beckman BS, Marrero L, Jones FE. 2004. The ERBB4/HER4 receptor tyrosine kinase regulates gene expression by functioning as a STAT5A nuclear chaperone. J Cell Biol 167(3):469–478.

Xia W, Gerard CM, Liu L, Baudson NM, Ory TL, Spector NL. 2005. Combining lapatinib (GW572016), a small molecule inhibitor of ErbB1 and ErbB2 tyrosine kinases, with therapeutic anti-ErbB2 antibodies enhances apoptosis of ErbB2-overexpressing breast cancer cells. Oncogene 24(41):6213–6221.

Yarden Y, Sliwkowski MX. 2001. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2(2):127–137.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol 9(9):R137.

Figure legends

FIGURE 1 ERBB2 gene proximal and distal promoter usage and schematic representation of its two major mRNA isoforms. A, UCSC genome browser screenshot of RNAPII and IgG read depth (2 upper black profiles) at the *ERBB2* gene locus in BT474 cells. Vertical arrows indicate the genomic location of both distal and proximal ERBB2 promoters. Known isoforms are annotated underneath, with mammal conservation at the bottom (blue). B, Domain structure of ERBB2a isoform: signal peptide in grey; transmembrane domain in black which separates the extracellular domain from intracellular domain; tyrosine kinase domain in red. C, Exon map for ERBB2a isoform, the Exon 1 containing sequences coding the signal peptide in red. D, Domain structure of ERBB2b isoform with transmembrane domain in black and tyrosine kinase domain in red. E, Exon map for ERBB2b isoform, the first four non-coding exons in brown.

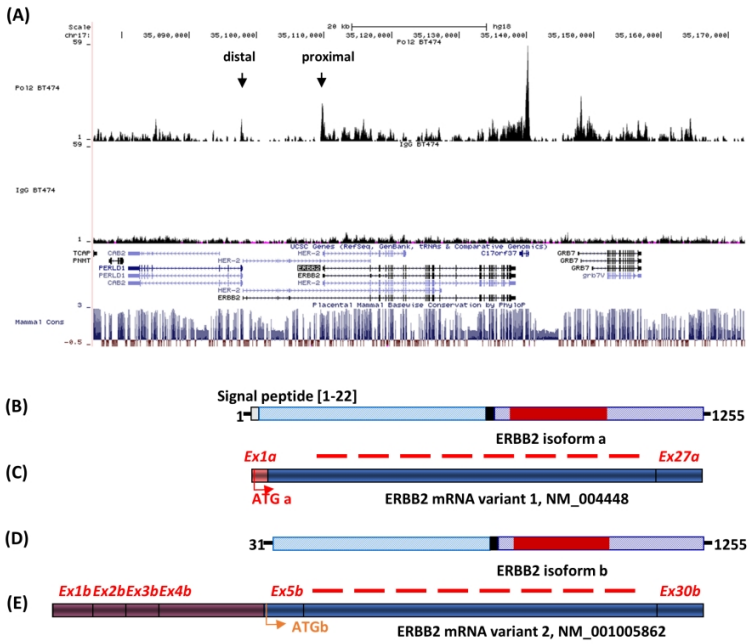
FIGURE 2 The expression of *ERBB2a* and *ERBB2b* isoforms are positively correlated in breast cancer. Quantitative-RT-PCR was performed from breast cancer cell lines and tumor samples. Relative gene expressions (RGE) of *ERBB2a* and *ERBB2b* were presented: as $2^{\Delta\Delta Ct}$ in breast cancer cell lines (A) and carcinomas samples (B); or as ΔΔCt in breast cancer cell lines (C) and tumor samples (D). E, The correlation coefficient and significant rate were calculated using ΔΔCt of relative gene expressions of three genes co-amplified at the 17q12 locus *PERLD1*, *C17orf37* and *GRB7* in breast cancer carcinomas samples.

FIGURE 3 ERBB2b isoform is a nuclear form of ERBB2. A, Subcellular fractionations were performed in BT474, HCC-1954, HCC-1569 and Br-Ca-Mz-01 cell lines to separate cytosolic, membrane and nuclear fraction. Western blot analyses were followed to determine ERBB2 localization in these two cell lines, Sp1 and β-tubulin were chosen as nuclear and cytosolic marker, respectively. B, Immunofluorescence assay coupled with confocal microscopy were carried out with an ERBB2 antibody against its carboxyl terminus. Nuclei were stained with DAPI. C, Expression of pEGFP-ERBB2a or pEGFR-ERBB2b in BT474 and HCC1806 cells. Upper panels: plain fluorescent microscope fields. Lower panels: GFP-positive cell magnification. These representative examples are extracted from 6 to 10 randomly selected confocal microscope fields per conditions.

FIGURE 4 ERBB2a activates ERBB2b whereas ERBB2b inhibits ERBB2a. A, Location of specific siRNAs targeting either ERBB2a or 2b isoform. B, Transcription analyses upon transient transfection of siRNA in breast cancer cell line BT474 or HCC-1954. C, Immunofluorescence analysis was performed upon transient transfection of siRNA in HCC-1954 cells. D, Immunofluorescence analysis was performed upon transient transfection of siRNA in Br-Ca-Mz-01 cells.

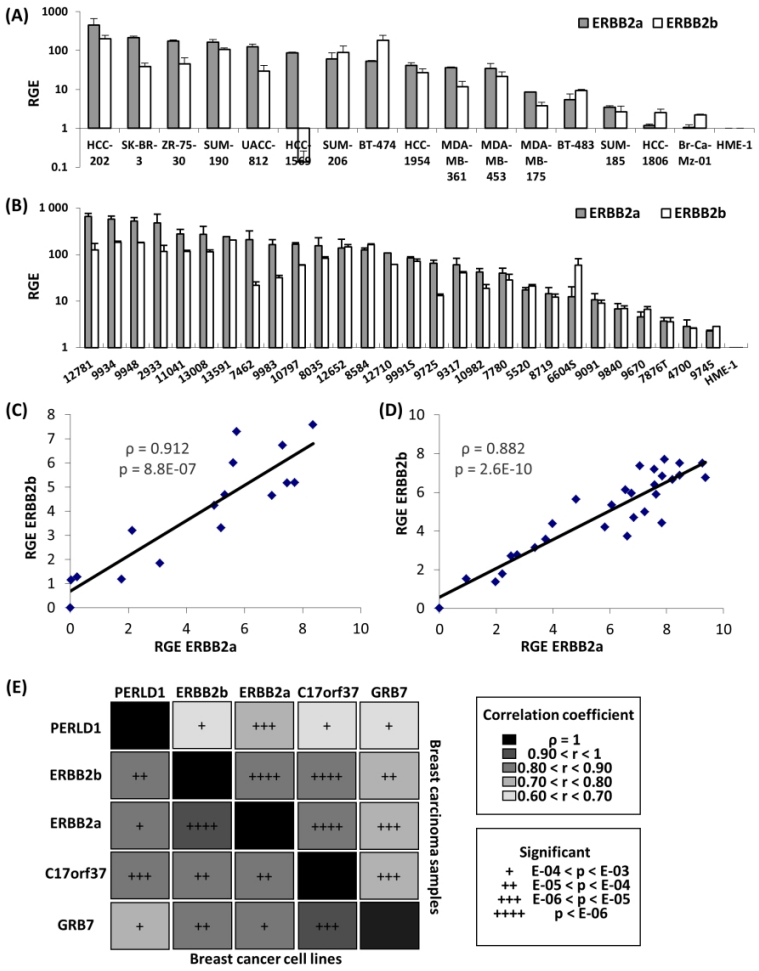
FIGURE 5 Binding of nuclear ERBB2 to a specific DNA sequence of the *ERBB2* promoter. A, Representation of the 5'-end flanking sequence of human *ERBB2* gene showing a putative ERBB2 binding sequence in black box. B, Chromatin immunoprecipitation assays carried out in BT474 cell line using ERBB2 antibody or IgG as a negative control. DNA binding were measured for COX2 and ERBB2 gene by the percent of input of DNA either pulled down by ERBB2 or IgG.

FIGURE 6 RNAPII peaks in BT474 cells are enriched in HAS motifs and correlate with increased gene expression. A, Pie chart showing annotation of BT474 RNAPII peaks to genomic features. B, Venn diagram illustrating intersection of RNAPII peaks in BT474 and HCC-1806 cells. C, Average profiles of ERBB2/HER2-associated sequence motif (HAS) densities in BT474-specific (red), shared (light blue), HCC-1806-specific RNAPII peaks (green). D, Motif discovery result in BT474-specific RNAPII peaks showing motif #8, with a sequence compatible with HAS motifs. E, Gene set enrichment analysis (GSEA) showing correlations between presences of BT474 (left) and HCC-1806 (right) RNAPII peaks with BT474/HCC-1806 mRNA expression fold change.



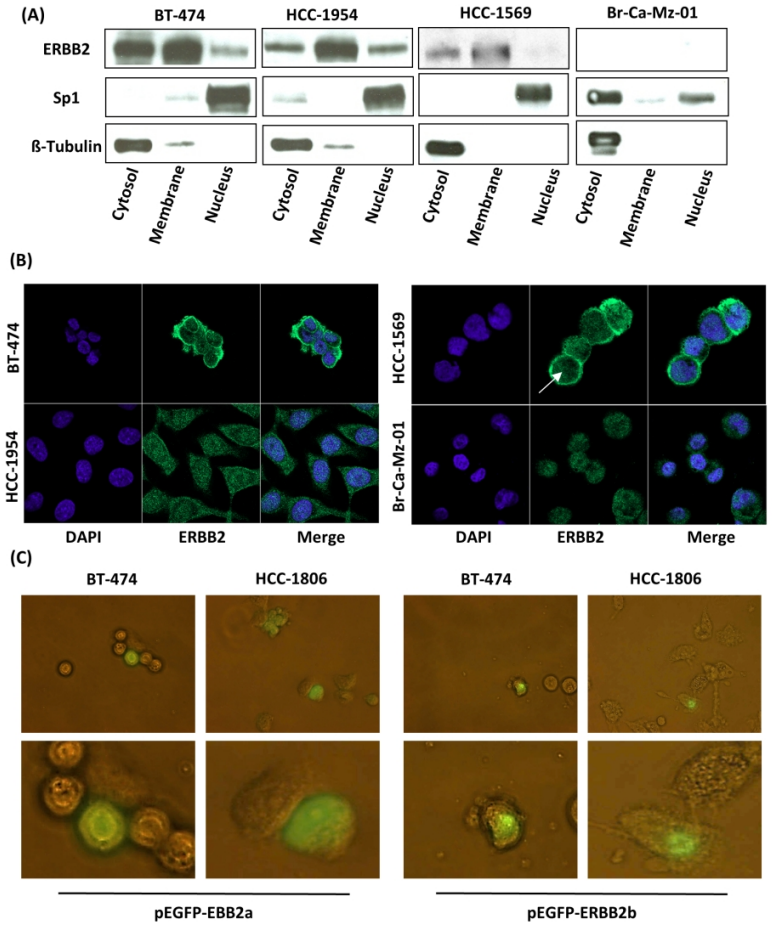
ERBB2 gene proximal and distal promoter usage and schematic representation of its two major mRNA isoforms.

190x254mm (307 x 307 DPI)



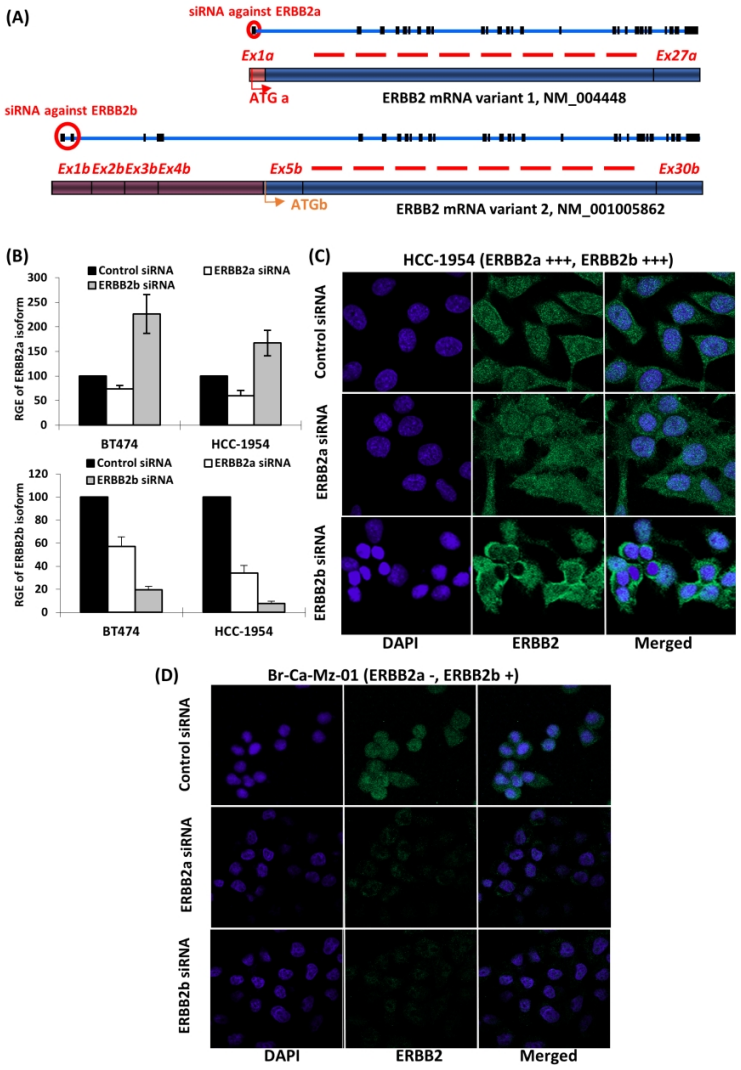
The expression of ERBB2a and ERBB2b isoforms are positively correlated in breast cancer.

190x254mm (307 x 307 DPI)



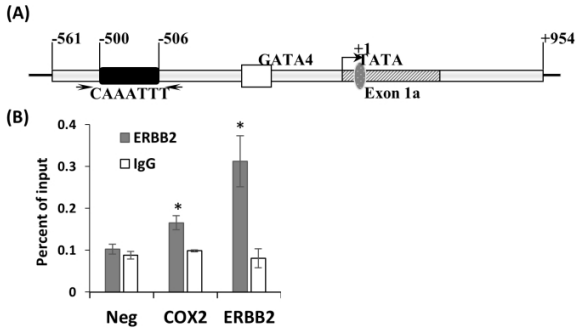
ERBB2b isoform is a nuclear form of ERBB2.

190x254mm (307 x 307 DPI)



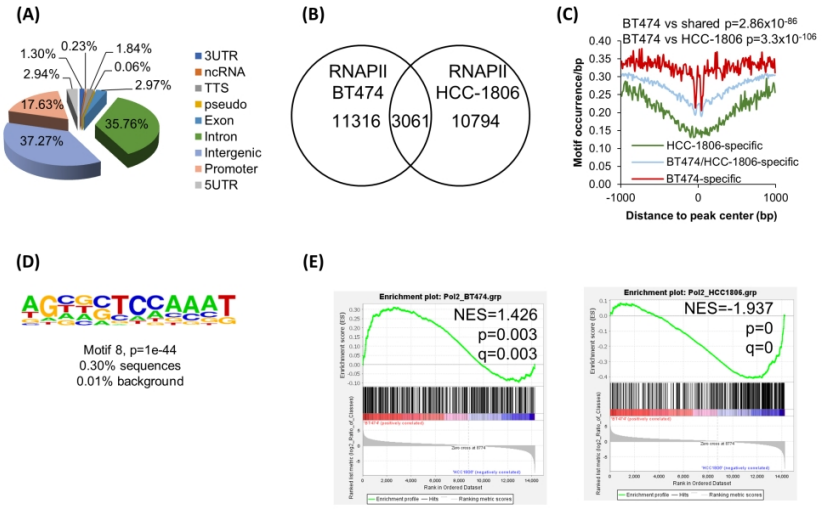
ERBB2a activates ERBB2b whereas ERBB2b inhibits ERBB2a.

190x254mm (307 x 307 DPI)



Binding of nuclear ERBB2 to a specific DNA sequence of the ERBB2 promoter.

190x254mm (307 x 307 DPI)



RNAPII peaks in BT474 cells are enriched in HAS motifs and correlate with increased gene expression.

190x254mm (307 x 307 DPI)

Table 1: Clinicopathological characteristics of 28 breast cancer samples

Characteristics		N (%)
Age at diagnosis	<=50	11 (39%)
	>50	17 (61%)
Menopausal status	no	11 (39%)
	yes	17 (61%)
AJCC stage	1	10 (36%)
	2	12 (43%)
	3	5 (18%)
	4	1 (4%)
Pathological tumor type	ductal	22 (79%)
	lobular	2 (7%)
	other	4 (14%)
Pathological tumor grade	1	5 (19%)
	2	6 (22%)
	3	16 (59%)
Pathological tumor size	pT1	10 (40%)
	pT2	8 (32%)
	pT3	7 (28%)
Pathological axillary lymph node status	negative	11 (42%)
	positive	15 (58%)
ER status*	negative	18 (67%)
	positive	9 (33%)
PR status*	negative	20 (74%)
	positive	7 (26%)
ERBB2 status**	negative	8 (29%)
	positive	20 (71%)
Molecular subtype	HR+/ERBB2-	4 (14%)
	ERBB2+	20 (71%)
	TN	4 (14%)

* IHC status with 10% positivity cut-off; **, status defined using IHC, then array-CGH, HR: hormone receptor; TN: triple-negative

Table 3. Targeted sequences used for siRNA Assay.

RNA	Target Name	Sequence
ERBB2 isoform a (NM_004448)	NM_004448_stealth_14 sense strand	GAGGAGGGCUGCUUGAGGAAGUAUA
	NM_004448_stealth_14 antisense strand	UAUACUUCCUCAAGCAGCCCUCCUC
	NM_004448_stealth_15 sense strand	AGGAGGGCUGCUUGAGGAAGUAUAA
	NM_004448_stealth_15 antisense strand	UUAUACUUCCUCAAGCAGCCCUCCU
	NM_004448_stealth_17 sense strand	GAGGGCUGCUUGAGGAAGUAUAAGA
	NM_004448_stealth_17 antisense strand	UCUUAUACUUCCUCAAGCAGCCCUC
ERBB2 isoform b (NM_001005862)	NM_001005862_stealth_90 sense strand	CCUGAUGGGUUAAUGAGCAAACUGA
	NM_001005862_stealth_90 antisense strand	UCAGUUUGCUCAUUAACCCAUCAGG
	NM_001005862_stealth_97 sense strand	GGUUAAUGAGCAAACUGAAGUGUUU
	NM_001005862_stealth_97 antisense strand	AAACACUUCAGUUUGCUCAUUAACC
	NM_001005862_stealth_187 sense strand	GGUAGAACCUUUGCUGUCCUGUUCA
	NM_001005862_stealth_187 antisense strand	UGAACAGGACAGCAAAGGUUCUACC

Table 4. Sequences of the oligonucleotide primers used for qRT-PCR.

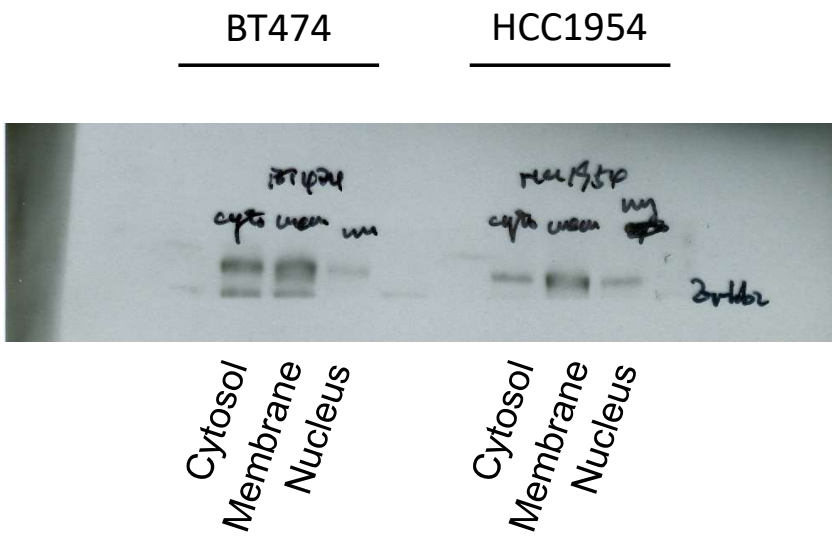
Gene	Oligonucleotide	Sequence	PCR product size (bp)
ERBB2a mRNA	Forward Primer	5'- CTCCTCCTCGCCCTCTTG – 3'	253
	Reverse Primer	5'- GACCTGCCTCACTTGGTTGT – 3'	
ERBB2b mRNA	Forward Primer	5'- ACGCCTGATGGGTTAATGAG – 3'	154
	Reverse Primer	5'- CCAAATTCTGTGCTGGAGGT – 3'	
TBP mRNA	Forward Primer	5'- CTTGTGCTCACCCACCAAC – 3'	228
	Reverse Primer	5'- GGAGGCAAGGGTACATGAGA – 3'	
PERLD1 mRNA	Forward Primer	5'- AGTTTTCCACACCAGGGACA – 3'	203
	Reverse Primer	5'- CATAGTCGAAGCGGATGAGG – 3'	
C17orf37 mRNA	Forward Primer	5'- TTCGAGGCGACCTACCTG – 3'	225
	Reverse Primer	5'- ACGGCTGTTGGTGATCTTTT – 3'	
GRB7 mRNA	Forward Primer	5'- CGTGTGTGAAATGCTGGTG – 3'	225
	Reverse Primer	5'- TCTGGGAACAGGGAGTGTG – 3'	

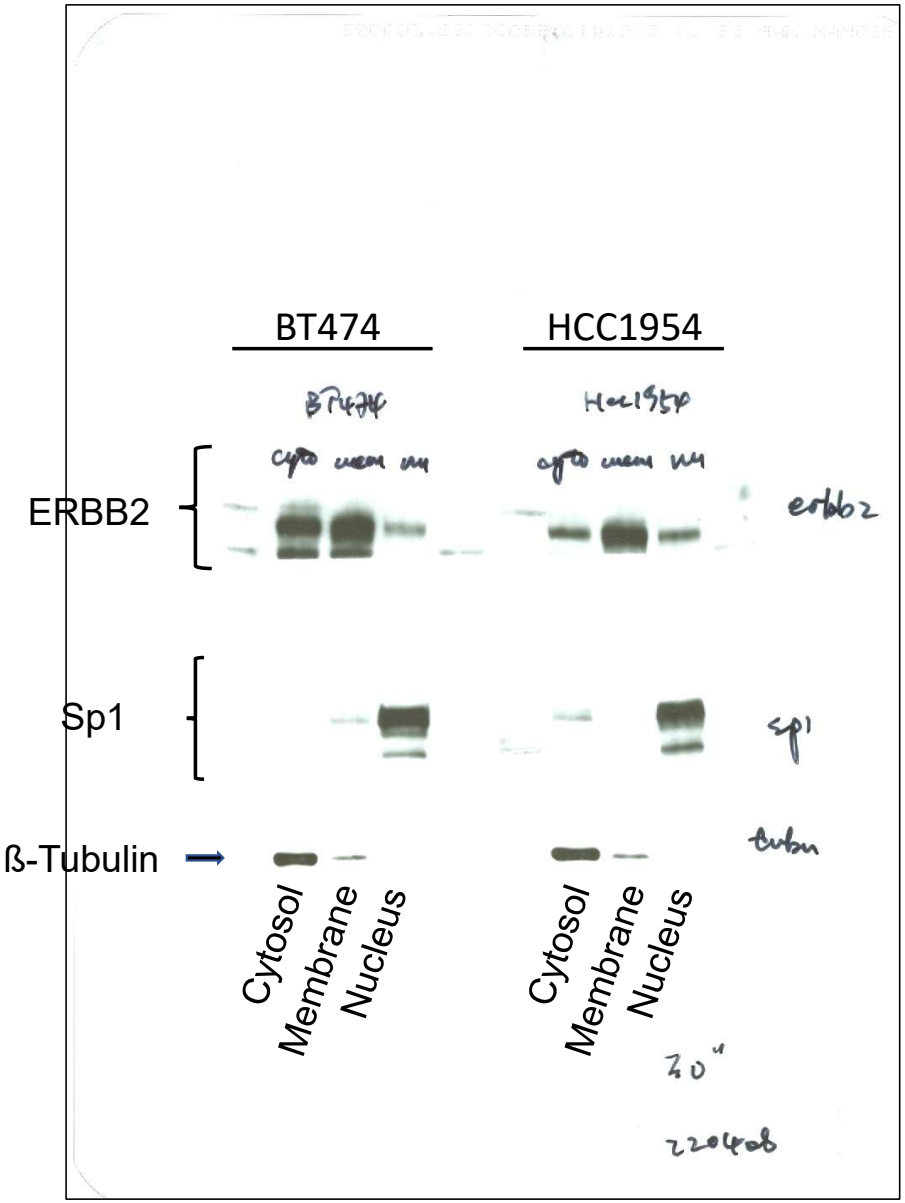
Table 2 Detailed clinicopathological characteristics of 28 breast cancer samples

ID	Age at diagnosis	Menopausal status	AJCC stage	Pathological tumor type	Pathological tumor grade	Pathological tumor size	Pathological axillary lymph node status	ER status*	PR status*	ERBB2 status**	Molecular subtype
12781	<=50	no	1	other	NA	pT1	neg			pos	ERBB2+
9934	>50	yes	3	ductal	3			neg	neg	pos	ERBB2+
9948	>50	yes	3	ductal	3	pT1		pos	pos	pos	ERBB2+
2933	>50	yes	2	ductal	3	pT1	pos	neg	neg	pos	ERBB2+
11041	>50	yes	2	ductal	3	pT2	pos	neg	neg	pos	ERBB2+
13008	<=50	no	2	ductal	3	pT3	pos	pos	pos	pos	ERBB2+
13591	>50	yes	4	ductal	3	pT3	pos	neg	neg	pos	ERBB2+
7462	<=50	no	2	ductal	3	pT3	pos	neg	neg	pos	ERBB2+
9983	<=50	no	2	ductal	2	pT1	neg	neg	neg	pos	ERBB2+
10797	>50	yes	1	ductal	3	pT2	pos	neg	neg	pos	ERBB2+
8035	<=50	no	2	ductal	2	pT3	pos	neg	neg	pos	ERBB2+
12652	<=50	no	3	ductal	3		pos	pos	pos	pos	ERBB2+
8584	>50	yes	2	other	2	pT2	neg	pos	pos	pos	ERBB2+
12710	<=50	no	3	ductal	3		neg	neg	neg	pos	ERBB2+
9991	>50	yes	1	ductal	2	pT2	neg	neg	neg	pos	ERBB2+
9725	<=50	no	1	ductal	3	pT3	pos	neg	neg	pos	ERBB2+
9317	>50	yes	2	ductal	3	pT2	pos	pos	neg	pos	ERBB2+
10982	>50	yes	2	ductal	3	pT3	pos	neg	neg	pos	ERBB2+
7780	<=50	no	3	ductal	3	pT2	pos	neg	neg	pos	ERBB2+
5520	>50	yes	1	ductal	3	pT1	neg	neg	neg	pos	ERBB2+
8719	<=50	no	1	other	2	pT1	neg	pos	pos	neg	HR+/ERBB2-
6604S	>50	yes	2	ductal	1	pT1	pos	pos	pos	neg	HR+/ERBB2-
9091	>50	yes	1	ductal	1	pT1	neg	pos	neg	neg	HR+/ERBB2-
9840	>50	yes	1	ductal	1	pT3	neg	neg	neg	neg	TN
9670	<=50	no	1	ductal	1	pT2	neg	pos	pos	neg	HR+/ERBB2-
7876T	>50	yes	2	lobular	1	pT1	neg	neg	neg	neg	TN
4700	>50	yes	2	other	3	pT2	pos	neg	neg	neg	TN
9745	>50	yes	1	lobular	2	pT1	pos	neg	neg	neg	TN

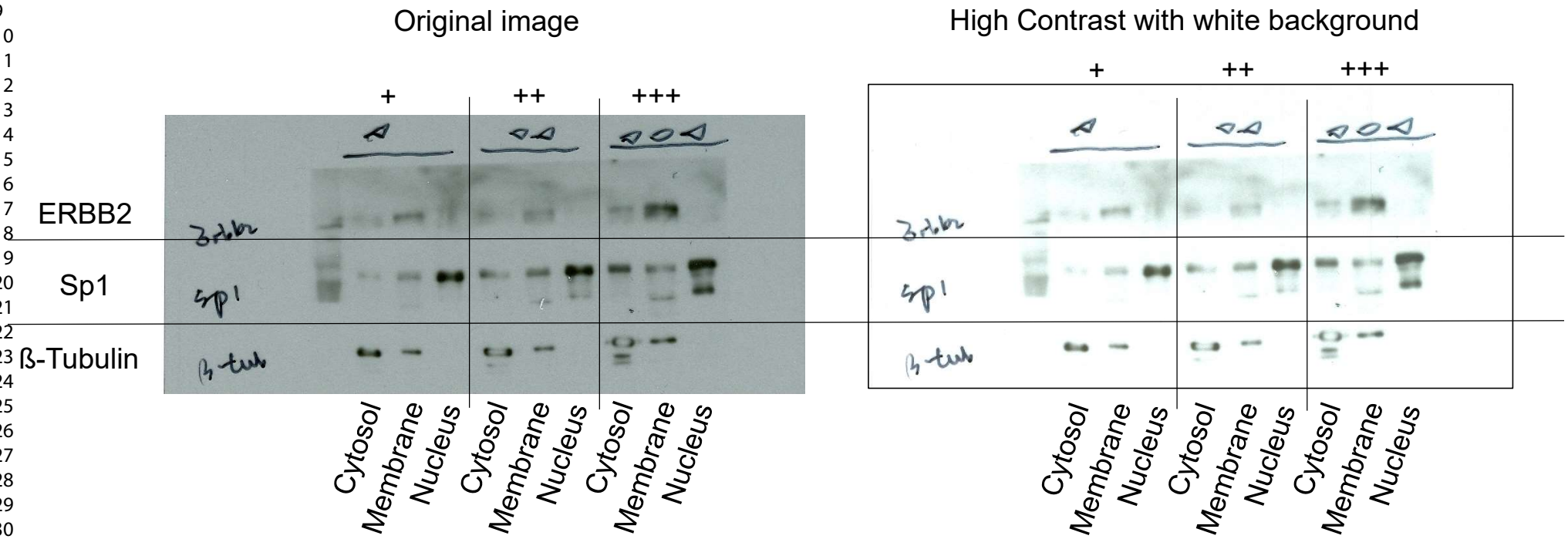
*, IHC status with 10% positivity cut-off; **, status defined using IHC, then array-CGH

HR: hormone receptor; TN: triple-negative





Cytosolic, membrane, and nuclear fractions were separated for BT474 and HCC1954 cells, Western blot analyses were performed with ERBB2, SP1 and ß-Tubulin antibodies.



Cytosolic, membrane, and nuclear fractions were separated for HCC1569 cells, Western blot analyses were performed with ERBB2, SP1 and ß-Tubulin antibodies.

Cytosolic, membrane, and nuclear fractions were separated for Br-Ca-Mz-01 and HCC1569 cells, Western blot analyses were performed with ERBB2, SP1 and ß-Tubulin antibodies.