Title

Discovery of internalizing antibodies to basal breast cancer cells

Journal

Protein Engineering Design and Selection, 31(1)

ISSN

1741-0126

Authors

Zhou, Yu

Zou, Hao

Yau, Christina

et al.

Publication Date

2018

DOI

10.1093/protein/gzx063

Peer reviewed

Protein Engineering, Design & Selection, 2018, vol. 31 no. 1, pp. 17–28
doi: 10.1093/protein/gzx063
Advance Access Publication Date: 28 December 2017
Original Article

OXFORD

Original Article

Discovery of internalizing antibodies to basal breast cancer cells

Yu Zhou[1,*], Hao Zou[1], Christina Yau[2], Lequn Zhao[1], Steven C. Hall[3], Daryl C. Drummond[4], Shauna Farr-Jones[1], John W. Park[5], Christopher C. Benz[2], and James D. Marks[1,*]

[1]Department of Anesthesia and Perioperative Care, University of California, San Francisco Rm 3C-38, Zuckerberg San Francisco General Hospital and Trauma Center, 1001 Potrero Ave, San Francisco, CA 94110, USA, [2]Buck Institute for Research on Aging, 8001 Redwood Blvd., Novato, CA 94945, USA, [3]Department of Obstetrics, Gynecology & Reproductive Sciences, Sandler-Moore Mass Spectrometry Core Facility, 521 Parnassus Avenue, University of California, San Francisco, San Francisco, CA 94143, USA, [4]Merrimack Pharmaceuticals Inc., One Kendall Square, Suite B7201, Cambridge, MA 02139, USA, and [5]Department of Medicine, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, Box 1710, San Francisco, CA 94143, USA

[*]To whom correspondence should be addressed. E-mail: eunice.zhou@ucsf.edu, jim.marks@ucsf.edu
Edited By: Dr Andrew Bradbury

Received 27 November 2017; Revised 0 0; Editorial Decision 28 November 2017; Accepted 1 December 2017

Abstract

We present a strategy to discover recombinant monoclonal antibodies (mAbs) to specific cancers and demonstrate this approach using basal subtype breast cancers. A phage antibody library was depleted of antibodies to common cell surface molecules by incubation with luminal breast cancer cell lines, and then selected on a single basal-like breast cancer cell line (MDA-MB-231) for binding associated receptor-mediated endocytosis. Additional profiling against two luminal and four basal-like cell lines revealed 61 unique basal-specific mAbs from a pool of 1440 phage antibodies. The unique mAbs were further screened on nine basal and seven luminal cell lines to identify those with the greatest affinity, specificity, and internalizing capability for basal-like breast cancer cells. Among the internalizing basal-specific mAbs were those recognizing four transmembrane receptors (EphA2, CD44, CD73 and EGFR), identified by immunoprecipitation–mass spectrometry and yeast-displayed antigen screening. Basal-like breast cancer expression of these four receptors was confirmed using a bioinformatic approach, and expression microarray data on 683 intrinsically subtyped primary breast tumors. This overall approach, which sequentially employs phage display antibody library selection, antigen identification and bioinformatic confirmation of antigen expression by cancer subtypes, offers efficient production of high-affinity mAbs with diagnostic and therapeutic utility against specific cancer subtypes.

Key words: basaloid breast cancer, cancer target discovery, cell selection, internalizing single chain Fv antibody, phage display

Introduction

Cancer is a heterogeneous disease with multiple different cellular origins and molecular features. This is particularly true for breast cancers, where gene expression profiling (Perou *et al.*, 2000; Sorlie *et al.*, 2001) and comprehensive genomic analyses (Network, 2012) have now subdivided breast cancer into four intrinsic molecular subtypes: luminal A, luminal B, HER2 overexpressing (HER2+) and basal-like breast cancers. Basal-like and HER2+ subtypes are recognized as being more aggressive forms of breast cancer, while the two hormone receptor-positive luminal subtypes are generally less aggressive. Multiple hormone- and HER2-targeted therapies have improved the survival outcome of most breast cancer patients;

however, the 15–20% of patients diagnosed with basal-like breast cancer face a very aggressive disease course for which there are no approved target-specific therapeutics (Foulkes *et al.*, 2010). Therefore, target-specific therapies for the management of this most aggressive subtype of breast cancer are urgently needed.

Basal-like breast cancers were formally identified by their genomic transcript expression profile (Perou *et al.*, 2000) that often includes upregulated expression of CK5/17, epidermal growth factor receptor (EGFR), Caveolin1&2, c-Kit and P-Cadherin (Nielsen *et al.*, 2004; Kreike *et al.*, 2007; Rakha *et al.*, 2007), and are named as such because they tend to express the same proteins as that of normal basal and myoepithelial cells within the human breast (i.e. cytokeratins 5, 6, 14, 15, 17; annexin VIII; P-cadherin/CDH3; vimentin; smooth muscle actin, and EGFR). Except for EGFR, these basal-specific markers are not generally considered druggable targets; and while some basal-like breast cancers overexpress EGFR, EGFR-targeted therapeutics have so far not shown clinical promise for any form of breast cancer (Nakai *et al.*, 2016), underscoring the need for additional basal-like breast cancer targeted therapeutics. The analytical and gene expression approaches used to define the four intrinsic subtypes of breast cancer have also revealed that basal-like breast cancer is one of largest molecular subsets within the broader clinical subcategory of triple-negative (ER-/PR-/HER2-) breast cancers, for which extensive clinical studies have also failed to uncover any promising targeted therapies (Irshad *et al.*, 2011). Fortunately, a number of well-characterized human breast cancer cell lines are now available for in-depth study of the molecular differences between intrinsic breast cancer subtypes, in particular, looking for specific cell surface targets that might distinguish basal-like from luminal and HER2+ breast cancers (Neve *et al.*, 2006; Heiser *et al.*, 2012).

Proteome (Leth-Larsen *et al.*, 2009) and kinome (Hochgrafe *et al.*, 2010) profiling of breast cancer cell lines have revealed breast cancer metastasis-associated membrane proteins (CD73, NDRG1, integrin β1, CD44, CD74 and MHC-II proteins) as well as elevated tyrosine phosphorylation of Met, Lyn, EphA2 and EGFR in basal breast cancer cells. The immunohistochemistry analysis of breast cancer tissue biopsies further identified CD73 as a potential target for metastatic breast cancers independent of HER2 overexpression (Leth-Larsen *et al.*, 2009). Meanwhile, an anti-CD73 therapy in a triple-negative breast cancer model confirmed that CD73 is involved in the progression of this breast cancer subtype (Stagg *et al.*, 2010; Loi *et al.*, 2013).

To identify basal-like breast cancer associated and therapeutically relevant targets, as well as monoclonal antibodies (mAbs) bearing fully human sequences to enable immediate clinical development as therapeutics, we established an approach that combines recombinant antibody library with basal-like breast cancer cell selection. In this approach, we employed a number of human cell lines representing diverse intrinsic subtypes of breast cancer and a human phage display single chain Fv (scFv) antibody library selection. Using this strategy, we isolated target-specific mAbs and identified their cognate cell surface antigens. We also found that these mAbs bind basal-like breast cancer cell surface receptors with high affinity and specificity and induce receptor-mediated endocytosis, presenting the opportunity to use these mAbs to deliver toxic chemotherapy payloads as a new treatment modality against basal-like breast cancers.

We have reported some of the antibodies to EphA2 (2D6, D2-1A7, D2-1B1 and D2-1A9) and CD44 (F2-1A6) previously, including characterization with respect to their specificity and internalization. Here, we report new antibodies targeting basal breast cancer cells (CD73 mAbs), and characterization of *in vitro* biological activities of EphA2 antibody D2-1A7 and 2D6.

Materials and Methods

Cell lines, media, antibodies and full-length cDNA clones

Breast cancer cell lines BT20, BT474, BT594, CAMA1, HCC1950, HCC1954, HCC70, HS578T, JIMT1, MCF7, MDAMB157, MDA-MB-231, MDA-MB-361, MDA-MB-453, MDA-MB-468, MX-1, SKBR3, SUM149PT, SUM159PT, SUM185PE, SUM52PE, T47D, UACC812, ZR75-1 and ZR75-30 were obtained from the ATCC (Manassas, VA) or from collections developed in the laboratories of Drs Steve Ethier (Karmanos Institute, MI, USA; SUM cell lines) and Adi Gazdar (University of Texas, Southwestern Medical Center; HCC cell lines). The cell lines were cultured using conditions described previously (Neve *et al.*, 2006).

Yeast strain EBY100 (GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS2 prb1Δ1.6 R can1 GAL) was grown in YPD medium (Treco and Winston, 2008). EBY100 transfected with expression vector pYD2 (Razai *et al.*, 2005) was selected on SD-CAA medium (Treco and Winston, 2008). The Aga2p antigen fusion was expressed on the yeast surface by induction in SG-CAA medium (identical to SD-CAA medium except that the glucose is replaced by galactose) at 20°C for 24–48 h as described previously (Feldhaus *et al.*, 2003). Bacterial strains *Escherichia coli* DH5α (K12, ΔlacU169 (ϕ80 lacZΔM15), supE44, hsdR17, recA1, endA1, gyrA96, thi-1, relA1) and TG1 (K12, Δ(lac-pro), supE, thi, hsdD5/F' traD36, proA + B +, lacIq, lacZΔM15) were used for the preparation of plasmid DNA and the expression of soluble scFv antibodies respectively.

Commercial anti-EphA2 MAb D7 (Upstate Biotechnology) and mouse Ephrin A1-hFc (R&D Systems) were used in receptor down-regulation and invasion assays, and anti-CD73 (Abcam) was used to detect CD73 in mAb 1A9-immunoprecipitates. SV5 antibody was purified from hybridoma (in house) supernatant using Protein G and directly labeled with Alexa-488 or Alexa-647 using a kit provided by the manufacturer (Invitrogen; Carlsbad, CA) and used to detect proteins displayed on the surface of yeast. Biotin conjugated rabbit anti-fd bacteriophage (Sigma) and Streptavidin Phycoerythrin (PE) (Biosource/Invitrogen) were used to detect phage antibody. Full-length cDNAs were obtained from the ATCC, Origene and Open Biosystems.

Selection by internalization of breast cancer subtype-specific phage antibodies

A multivalent fd phage display library derived from a phagemid display library of naïve human scFv (Sheets *et al.*, 1998) was used for the antibody selection (Huie *et al.*, 2001; O'Connell *et al.*, 2002; Liu *et al.*, 2004). Luminal subtype breast cancer cell lines SUM52PE, T47D and MDA-MB-453 were used to deplete the phage library of non-specific and luminal-specific binders by incubating 10^{12} phage particles with 10^8 cells for 2 h at 4°C. The supernatant containing the depleted phage library was then incubated with 5×10^6 basal subtype breast cancer cells for each cell line (BT20, HCC1954, MDA-MB-231, MDA-MB-468, SUM149PT and SUM159PT) for 1 h at 4°C and subjected to the same selection protocol separately. Cells were washed with cold PBS and incubated with pre-warmed (37°C) medium for 30 min at 37°C to allow receptor-mediated

internalization. Non-internalized phage were removed by washing cells with glycine buffer (50 mM glycine, 150 mM NaCl, 200 mM urea and 2% polyvinylpyrrolidone, pH 2.8), while internalized phage were collected from within the cells and amplified in *E. coli* TG1 as described previously (Becerril *et al.*, 1999; Poul *et al.*, 2000).

Production of scFv and IgG

The scFv genes were subcloned from the phage vector into vector pUC119mycHis (Schier *et al.*, 1995) to express soluble scFv with Myc and (His)6 tags, and vector pSyn-Cys-His6 (Liu *et al.*, 2004) to express scFv with a free cysteine at the COOH terminus for liposome conjugation. ScFv antibodies were purified from the osmotic shock fractions using a Ni-NTA agarose column. IgG proteins were secreted in the media by transfected CHO cells as described previously (Nowakowski *et al.*, 2002), and purified by Protein G affinity chromatography (Amersham Pharmacia, Piscataway, NJ, USA).

Flow cytometry measurement of breast cancer cell surface antibody binding

Breast cancer cells were grown to 80–90% confluence and harvested by trypsinization. Phage (10^{10} colony-forming units total), scFv (100 nM) and IgG (10 nM) antibodies were incubated with 5×10^4 cells overnight at 4°C in flow cytometry buffer (phosphate-buffered saline (PBS) (pH 7.4), 1% FBS) in a total volume of 100 µl. After two washes with 200 µL of flow cytometry buffer, bound phage were detected by the addition of 100 µl (1 µg/mL) of biotinylated anti-fd bacteriophage (Sigma) and streptavidin-PE (Biosource/Invitrogen). Bound scFv was detected by the addition of 100 µl (1 µg/mL) of biotinylated His probe and streptavidin-PE; bound IgG was detected by the addition of 100 µl (1 µg/mL) of PE-labeled anti-human Fc specific $F(ab')_2$ (Jackson Immnuoresearch). After 30 min incubation at 4°C, the cells were washed twice and resuspended in PBS containing 4% paraformaldehyde. Fluorescence was measured by flow cytometry in a fluorescent activated cell sorting (FACS) LSRII (BD Biosciences), and mean fluorescence intensity (MFI) calculated. The MFI value of cell surface antibody binding was normalized, transformed by log 10, and analyzed using Cluster 3.0 and TreeView programs (de Hoon *et al.*, 2004). Equilibrium constants were determined as described (Benedict *et al.*, 1997), except that values were fitted to the equation $MFI = MFI_{min} + MFI_{max}*[Ab]/(K_D + [Ab])$ using KaleidaGraph (Version 4.1.3, Synergy Software, Reading, PA, USA).

Correlation analysis of cell surface staining and gene expression

The normalized and transformed MFI value on a panel of breast cancer cell lines was compared with the gene expression data obtained from Lawrence Berkeley National Laboratory (LBL, http://www.ebi.ac.uk/arrayexpress/experiments/E-TABM-157/) (Neve *et al.*, 2006; Heiser *et al.*, 2012) using Excel Correlation function. The correlation coefficient was generated for each gene probe, sorted from high to low, and used for prediction of target candidates.

Identification of cognate antigen by mass spectrometry

MDA-MB-231 cells were grown to 90% confluence and labeled with 0.1 mg/mL Sulfo-NHS-LC-biotin (Pierce). Total proteins were extracted using lysis buffer (10 mM Tris–HCl (pH 8), 150 mM NaCl, 1% NP40) completed with protease inhibitor cocktail, cleared by incubation with protein A-Sepharose (Sigma), incubated with the target mAbs, and the immunocomplex captured by protein A-Sepharose. After SDS-PAGE separation, the resolved protein was transferred to a PVDF membrane and revealed with horseradish peroxidase (HRP)-conjugated streptavidin (Gibco). An identical gel was stained with Coomassie R250, and the specific protein band was excised with reference to the Western Blot.

The excised gel band was subjected to standard sample preparation for mass spectrometry. The gel band was cut into ~1 mm^3 pieces and were stripped of stain using 25 mM NH_4HCO_3/50% acetonitrile (ACN). After discarding the supernatant, the gel pieces were dehydrated using 100% ACN. Upon shrinking and turning white, the ACN was discarded, and the gel pieces were rehydrated with 10 mM dithiothreitol (DTT) in 25 mM NH_4HCO_3 to reduce protein disulfide bonds. The mixture was kept at 56°C for 45 min, and the supernatant removed followed by addition of 55 mM iodoacetamide. The alkylation reaction was allowed to proceed in the dark at room temperature for 30 min. Next, the supernatant was discarded, and the gel pieces were washed with 25 mM NH_4HCO_3 and dehydrated as described above. The gel pieces were taken to complete dryness in a SpeedVac Concentrator (Savant) followed by rehydration with 12.5 ng/mL trypsin (Promega) in 25 mM NH_4HCO_3 and incubated overnight at 37°C. The supernatant was transferred to an Eppendorf tube. Tryptic peptides were extracted from the remaining gel pieces with aqueous 50% ACN in 0.1% formic acid and combined with the supernatant. The tryptic peptide digest was analyzed by reverse-phase nano HPLC MS/MS using a QSTAR XL mass spectrometer (Applied Biosystems) operating in electrospray ionization mode (Liu *et al.*, 2002). Protein identification was achieved by searching the MS/MS data using the Mascot MSDB database (Matrix Science). Database search parameters included taxonomy: mammals; enzyme: trypsin with three missed cleavages allowed; fixed modifications: Cys-carboxyamidomethylation; variable modifications: oxidized methionine, deamidation of asparagine/glutamine, pyroglutamic acid from N-terminal glutamine and N-terminal acetylation.

Screening antigen-specific phage antibody

Antigen candidates with high-correlation coefficients were chosen for screening against the polyclonal phage antibodies from the second round of cell selection. The antigens were either obtained from commercial sources or displayed on yeast surface by gap repair (Gietz and Schiestl, 1991), and used for antibody screening as described previously (Zhou and Marks, 2012). The resulting antigen-specific phage antibodies were further analyzed by DNA sequencing and FACS analysis.

Microarray determination of breast cancer cell line expression of antigenic receptors

Normalized, log2-scaled gene expression profiles from 51 different breast cancer cell lines, and their corresponding phenotypic data including intrinsic assignment into basal-like (these cell lines further subdivided into BaA and BaB categories) or Luminal (LumA, LumB) subtypes, were obtained from the LBL Breast Cancer Cell Collection site as previously reported (Neve *et al.*, 2006; Heiser *et al.*, 2012). Expression data were mean-centered, annotated using data obtained from the Broad Institute Integrative Genomics Viewer (http://www.broadinstitute.org/igv) (Robinson *et al.*, 2011; Thorvaldsdóttir *et al.*, 2013) and collapsed by gene symbol such that expression of a gene represented by multiple probes was computed as the average across probes. Expression levels of the four candidate genes (EGFR,

EPHA2, CD44 and NT5E/CD73) within the two basal categories (BaA and BaB) were compared with the luminal subtype; and significance of the pair-wise comparisons was assessed using the Student's *t*-test.

Microarray assessment of receptor expression by intrinsic breast cancer subtypes

Adjuvant untreated, node-negative breast cancer cases (447 ER+ and 236 ER−, 683 total) annotated for distant-metastasis-free survival were pooled from four sources (GSE2034, GSE5327, GSE7390 and NKI295). Expression data were analyzed as described above for the cell lines. Data generated on different microarray platforms were then mapped together using gene symbols to yield 10 219 unique genes and combined using distance weighted discrimination (DWD). Each sample within the pooled expression dataset was assigned to one of five intrinsic subtypes: luminal A (LumA), luminal B (LumB), HER2, Basal-like or normal-like (Normal). Specifically, training data used to derive the 'Intrinsic/UNC' signature was obtained from the UNC Microarray Database (https://genome.unc.edu/); 245 of the 305 genes in the signature were mapped to the pooled expression dataset by gene name, and DWD was applied to the training and pooled expression datasets to minimize systemic source biases. Intrinsic subtype centroids were created from the DWD-adjusted training data, and each tumor within the pooled expression dataset was assigned an intrinsic subtype according to its nearest centroid as determined by Spearman correlation. Expression levels of EGFR, EPHA2, CD44 and NT5E/CD73 were compared with the other four subtypes; and significance of the pair-wise comparisons was assessed using the Student's *t*-test.

Liposome internalization assay

Immunoliposomes were prepared as described previously (Nellis *et al.*, 2005; Zhou *et al.*, 2007). For microscopy studies, 150 000 cells were incubated with 50 µM phospholipid of untargeted and basal targeted immunoliposomes labeled with 0.3 mol % DiIC18 (3)-DS in a 12-well plate for 2 h at 37°C followed by washing with PBS and further incubation at 37°C for 2 h. The cells were then analyzed by using an inverted fluorescence microscope (Nikon Eclipse, TE300) with a 540/25 nm bandpass filter for excitation and a long pass filter at 565 nm for emission. The diameter of liposomes was 100–110 nm as determined by photon correlation spectroscopy.

Analysis of cell-surface receptor expression by western blot

After incubation with scFv antibodies, cells were chilled on ice and rinsed twice with ice-cold PBS, incubated with 0.5 mg/ml sulpho-NHS-SS-biotin for 45 min at 4°C. After labeling, cells were washed with PBS containing 50 mM glycine then lysed with lysis buffer containing a protease inhibitor cocktail. The total protein extract was cleared by centrifugation, incubated with streptavidin beads to purify the biotinylated surface protein. The beads were washed, and the bound protein was eluted by heating to 94°C for 8 min in protein-loading buffer containing DTT, and resolved by SDS-PAGE and western blot using receptor-specific mAbs.

Cell invasion assay

Modified Boyden chambers (8-µm pore size, Transwell, BD-Falcon) were used in cell-invasion assays as described previously (Lochter *et al.*, 1997). The inserts were coated with 1:5 diluted Matrigel (Becton Dickinson). After trypsinization, 10^5 cells in 200 µl of defined medium without FBS were incubated with scFv antibodies for 2 h at 37°C, added into the upper chamber, with the lower chamber filled with 300 µl of complete culture medium. After incubation for 20 h at 37°C, the inserts were fixed and stained, cells on the top side of the insets removed by Q-tips. A number of the invaded cells from five randomly selected fields was counted under an inverted microscope (Nikon, Eclipse TS100).

Soft agar colony formation assay

Anchorage-independent growth was performed in 6-well cell culture plate precoated with base medium (DMEM with 10% FBS, 1% penicillin/streptomycin) and 0.7% agarose (SeaPlaque, FMC BioProducts Rockland, ME, USA). Cells (5 000 per well) were mixed with DMEM medium containing 0.35% agarose, and varying amounts of 2B4 IgGs were layered on top of the base medium. Plates were cultured at 37°C in a humidified atmosphere containing 5% CO_2, cultures were given 0.5 ml complete medium containing the 2B4 IgGs twice a week, and colonies were counted under light microscopy after 10 days.

Results

Selection of basal-specific internalizing phage antibodies

A human naïve scFv phage antibody library (Sheets *et al.*, 1998; O'Connell *et al.*, 2002) was selected against well-characterized breast cancer cell lines that represent basal-like (SUM149PT, MDAMB468, BT20 and HCC1954), another basal-like subtype known as claudin-low (MDA-MB-231 and BT549), and luminal (MCF7, MDA-MB-453 and T47D) subtypes. The phage antibody library was depleted on multiple luminal breast cancer cells to remove antibodies to common cell surface molecules, and then selected on each basal-like breast cancer cell line under conditions allowing for receptor-mediated endocytosis (Becerril *et al.*, 1999; Poul *et al.*, 2000; Heitner *et al.*, 2001; Zhou *et al.*, 2010). After two rounds of selection, a fraction of the phage antibodies from each selection bound preferentially to basal-like cells.

Specifically for MDA-MB-231 cell selection, 202 of 1440 (14%) screened phage antibodies bound MDA-MB-231 cells after the second round of selection. Profiling these 202 phage antibodies against both luminal and basal-like/claudin-low lines indicated that 131 (9.1% of screened phage antibodies) demonstrated basal-like/claudin-low specificity. Sixty-one of these phage antibodies were unique as determined by DNA sequencing and were further screened against an expanded panel of nine basal-like/claudin-low and seven luminal breast cancer cell lines to identify antibodies with the greatest specificity for the basal-like cells. Similarly, phage antibodies selected and screened on luminal cells showed only luminal specificity.

The mAbs isolated using basal-like/claudin-low lines were further characterized with respect to the specificity and antigen identity following the strategy illustrated in Fig. 1. This strategy can be generalized for various cancer types.

Cell surface epitope specificity of basal and luminal breast cancer cells

Nine unique phage antibodies (5B1, 5A5, B5, 4F10, 5C8, G1, 12E9, 2B11 and 5E4) isolated from three luminal cell selections, 10 unique phage antibodies (1A9, 2B4, 1G1, 2D5, 5D7, 1H8, 2C12, 2D6, F2-1A6 and D2-1A7) from basal MDA-MB-231 cell selection; the

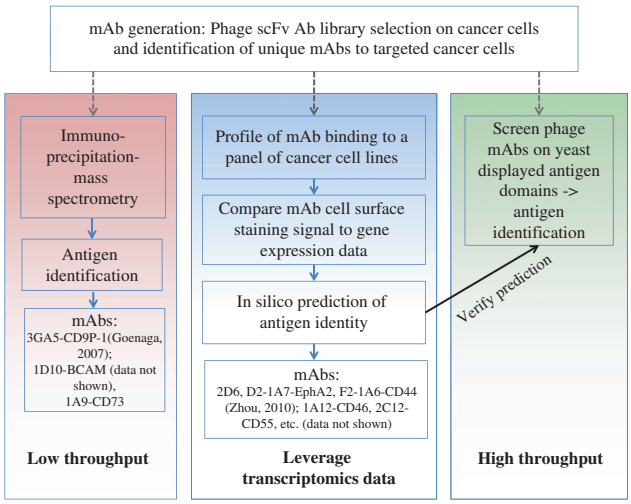

Fig. 1 Schematic of antigen identification strategy. Antigens corresponding to cancer cell specific antibodies were identified using immunoprecipitation–mass spectrometry (IP–MS) at low throughput; by prediction using comparison with transcriptomic databases; or by testing of binding to yeast-displayed candidate antigens. Recombinant mAbs were identified as previously described (Goenaga *et al.*, 2007; Zhou *et al.*, 2010)

EGFR antibody Cetuximab and phage antibody 2224; and the HER2 antibody Trastuzumab were profiled against an expanded panel of nine basal-like and seven luminal breast cancer cell lines using flow cytometry. The mean fluorescent intensity (MFI) values were normalized, transformed, and used for unsupervised hierarchical clustering analysis. The resulting bivariate clustering of antibodies and cell lines indicated that the cell surface molecules bound by MDA-MB-231 specific antibodies were most closely associated with the nine basal-like cell lines, while those bound by the luminal-specific antibodies associated with the seven luminal cell lines (Fig. 2). Interestingly, the EGFR antibody associated with the basal lines more so than the luminal lines, and trastuzumab did not bind either subtype significantly. This unsupervised clustering approach confirmed that the various subtypes of breast cancer cell lines express characteristic surface epitopes presenting the potential for selective targeting by mAbs.

Identification of antibody-targeted basal breast cancer surface antigens

The antigens bound by the selected phage antibodies were identified either by IP and MS (Goenaga *et al.*, 2007) or by yeast display of antigen cDNA and screening for specific binding by phage antibodies (Zhou *et al.*, 2010). The phage antibody 1A9 pulled down one dominant protein from MDA-MB-231 cell lysate, MS identified NT5E/CD73 as the top candidate antigen, and this was confirmed by western blot with an anti-CD73/NT5E mAb (Fig. 3A). Antibodies 1A9 and 2B4 appeared closely associated on the unsupervised cluster heat map (Fig. 2), identifying seven of the nine basal-like cell lines with an identical pattern, suggesting they recognize the same cell surface receptor. This hypothesis was confirmed using a binding competition assay (Fig. 3B) confirming that 2B4 also bound CD73. No other mAbs generated a major band on immunoprecipitation, preventing identification of other basal antigens using this approach.

As an alternative to antigen immunoprecipitation, a correlation analysis for antigen *in silico* prediction was carried out between gene transcription levels and mAb cell staining, using well-characterized breast cancer cell lines and associated microarray gene expression data (Neve *et al.*, 2006). This correlation analysis was performed using the EGFR antibody, Cetuximab (C225), the HER2 antibody, Trastuzumab, and the mAbs 2D6 and 2B4, which were selected on basal breast cancer cells. In the examples of correlation calculations, the paired mAbs (C225 and 2D6) and antigen candidates (EGFR and EphA2) resulted in R values over 0.8 with P value $<$ 0.0001, while the mis-paired mAb (2D6) and antigen (HER2) showed no correlation, suggesting this correlation analysis useful for antigen prediction for a mAb binding to cell surface receptor (Fig. 4). For the EGFR and HER2 mAbs, the expected antigen was in the top six genes, for mAbs 2D6 and 2B4, a number of genes were in the top 10 list including EphA2 for 2D6 and CD73/NT5E for 2B4 (Table I). These results were in agreement with those from immunoprecipitation–mass spectrometry (IP–MS).

To verify the identity of the cognate antigen and screen for additional mAbs, over 20 different antigens were displayed on the surface of yeast cells and used for screening and identification of specific phage antibodies (Zhou *et al.*, 2010; Zhao *et al.*, 2014). Using this approach, we have identified mAbs to EphA2 and CD44 described previously, and mAbs to other targets overexpressed on both luminal and basal subtype breast cancer cells (data not shown). Based on this observation, the approach of yeast display coupled with Ag in silico prediction is therefore complementary to the IP–MS approach. However, the *in silico* prediction was speculative as it worked for a limited number of targets.

In summary, we identified a panel of mAbs specific to EGFR, EphA2, CD44 and CD73 by various means, expressed them as soluble scFv and IgG, and measured their affinities to MDA-MB-231 cells determined using flow cytometry (Table II).

Transcriptional overexpression of the basal selected antigens in breast cancer cell lines and untreated primary breast tumors

To more broadly evaluate the expression pattern and specificity of EGFR, EphA2, CD44 and CD73 for basal-like breast cancers, an expression microarray panel composed of 51 breast cancer cell lines of different subtype origin was first evaluated. This confirmed that mRNA levels for EGFR, EphA2, CD44 and CD73 were all significantly higher ($P < 0.05$) in both subtypes of basal-like cell lines (BaA and BaB) relative to all luminal breast cancer cell lines (Fig. 5A and B, respectively). To extend this comparison to primary breast cancers, we interrogated a pooled expression microarray dataset comprising 683 node-negative and adjuvant chemotherapy naïve breast cancers representing all five intrinsic subtypes. Similar to the cell line results, elevated mRNA levels of the four antibody-targeted receptors were seen across most primary basal-like breast tumors relative to either luminal or HER2+ breast cancer subtypes (Fig. 6). More specifically, in pair-wise statistical comparisons, EGFR, EphA2 and CD73 expression levels were significantly higher in basal-like breast cancers relative to luminal A tumors ($P < 0.05$, Fig. 6A). All four genes showed elevated levels in basal-like relative to luminal B tumors ($P < 0.05$, Fig. 6B); and EGFR, EphA2 and CD44 were significantly overexpressed in basal-like relative to HER2+ tumors ($P < 0.05$, Fig. 6C). In contrast, basal-like expression of EphA2, CD44 and CD73 appeared comparable to normal-like breast cancers ($P < 0.05$, Fig. 6D), a very uncommon breast cancer subtype of questionable clinical significance, with only EGFR expression showing significantly higher expression.

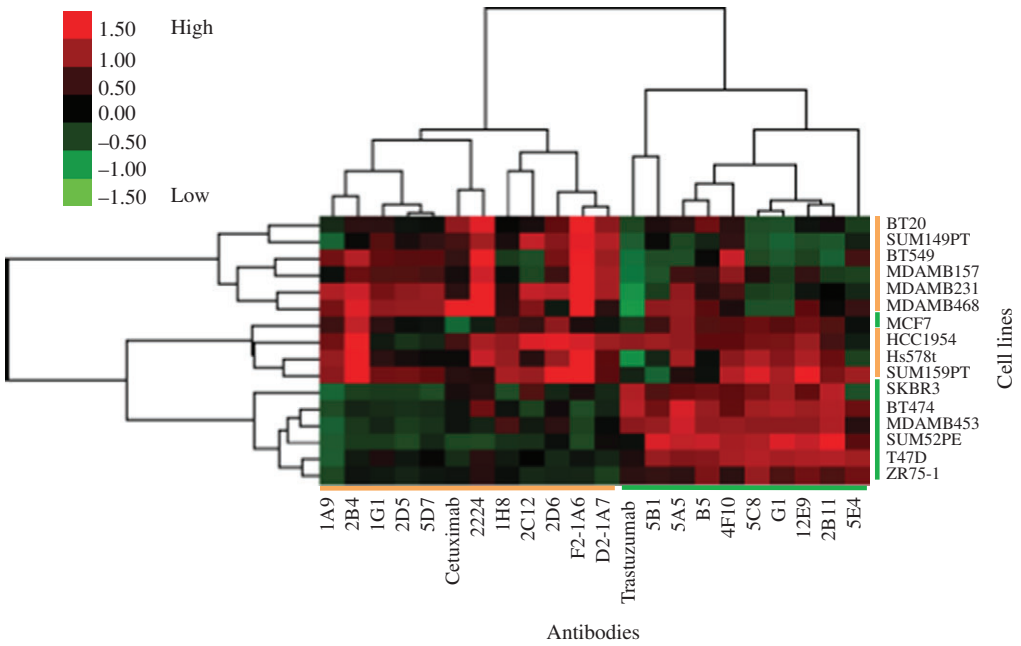

Fig. 2 Hierarchical clustering of mAbs binding to breast cell lines. Antibody binding to cells was measured using flow cytometry. The values of mean fluorescence intensity (MFI) were log-transformed and standardized by the corresponding protein mean and standard deviation followed by classification analysis using Gene Cluster 3.0. Data are presented as a heat map. Breast cell lines (16) comprised seven luminal (indicated by vertical light green line) and nine basal (indicated by vertical orange line) subtypes. Antibodies (22) used comprised those from basal-like cell selection (indicated by horizontal orange line), luminal cell selections (indicated with horizontal light green line) and anti-EGFR antibodies cetuximab (C225), 2224 (Zhou *et al.*, 2007); anti-HER2 antibody trastuzumab; anti-CD44, F2-1A6; anti-EphA2, 2D6 and D2-1A7 (Zhou *et al.*, 2010). We have not yet identified the antigens for the following antibodies: 1G1, 2D5, 5D7, 1H8, 5B1, 5A5, B5, 4F10, 5C8, G1, 12E9, 2B11 and 5E4

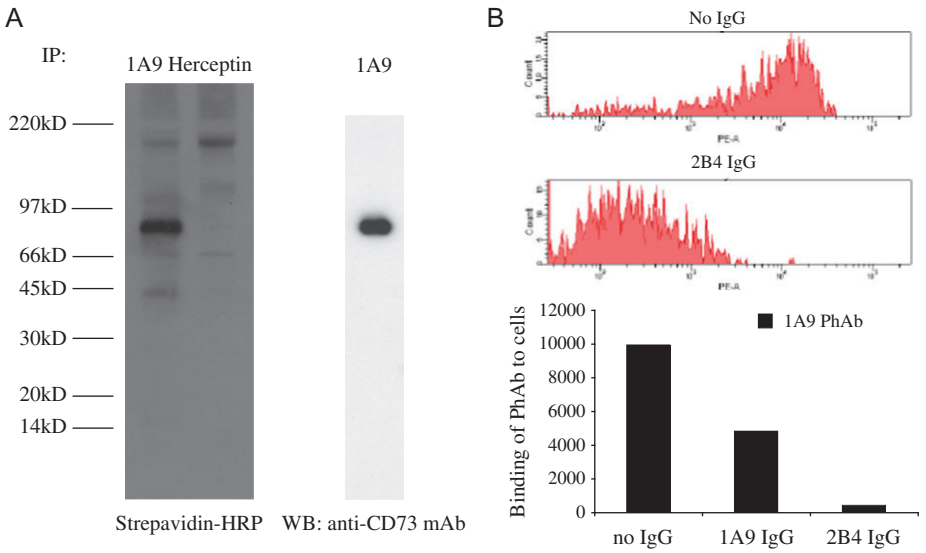

Fig. 3 Identification of cognate antigen to 1A9 and 2B4. (**A**) Western blot of IgG 1A9 immunoprecipitated biotinylated MDA-MB-231 cell lysates stained with Streptavidin-HRP and anti-CD73 mouse mAb. (**B**) 2B4 IgG competed the binding of phage antibody 1A9 to MDA-MB-231 cells

Biologic and therapeutic properties of the basal receptor-targeted antibodies

Phage antibodies targeting EGFR, EphA2, CD73 and CD44 (Table II) were converted to full-length IgG and evaluated for their biological activities. Although the EphA2, CD44 and CD73 antibodies did not inhibit anchorage-dependent cell proliferation, both CD73 IgG 2B4 and EphA2 IgG 2D6 effectively inhibited anchorage-independent cell proliferation as measured by reduction in colony forming units (Fig. 7A), with 2B4 showing greater inhibitory activity than 2D6, possibly due to the differences in target-specific effects and the binding affinities of 2B4 and 2D6 to the cells (data not shown).

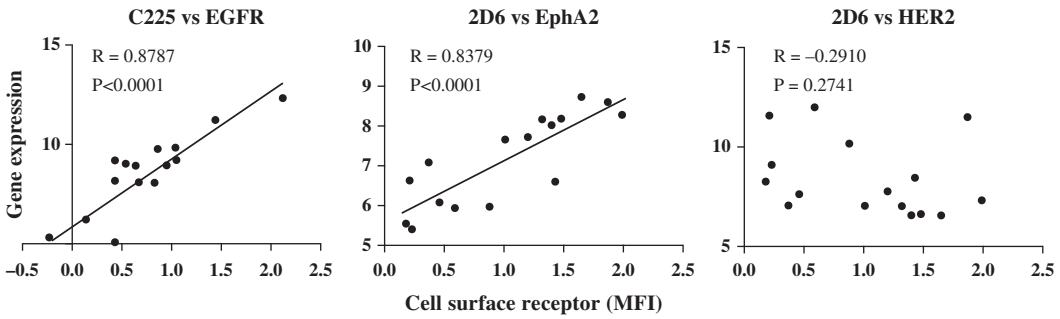

Fig. 4 Correlation analysis of antibody-cell binding profile with gene expression data in breast cancer cell lines. The MFI values of mAb binding to 16 breast cancer cell lines were used to calculate the correlation coefficient (R) comparing with the gene expression data (HG U133A transmembrane-filtered) on the same set of cell lines. Examples of correlation calculations include the paired mAb-antigen candidates (C225/EGFR and 2D6/EphA2), and the mis-paired mAb-antigen (2D6/HER2)

Table I. Correlation coefficient (*R*) of antibody-cell binding value measured by FACS with the gene expression data in breast cancer cell lines[a]

AFFY probe set ID[b]	Gene	Ab C225[c]	AFFY probe set ID	Gene	Ab 2D6[c]	AFFY probe set ID	Gene	Ab 2B4[c]
201983_s_at	EGFR	0.8776	212298_at	NRP1	0.8937	212298_at	NRP1	0.8280
201984_s_at	EGFR	0.8774	208944_at	TGFBR2	0.8763	212203_x_at	IFITM3	0.8196
210984_x_at	EGFR	0.8293	203499_at	EPHA2	0.8377	203939_at	NT5E	0.8071
208091_s_at	DKFZP564K0822	0.8177	210968_s_at	RTN4	0.7800	212063_at	CD44	0.8065
220094_s_at	C6orf79	0.8141	213476_x_at	TUBB3	0.7654	212097_at	CAV1	0.7965
211607_x_at	EGFR	0.7989	210804_x_at	SLC8A1	0.7529	208944_at	TGFBR2	0.7787
203484_at	SEC61G	0.7969	210473_s_at	CDC2L2	0.7501	203499_at	EPHA2	0.7523
204769_s_at	TAP2	0.7840	220663_at	IL1RAPL1	0.7485	203065_s_at	CAV1	0.7376
204581_at	CD22	0.7680	207082_at	CSF1	0.7354	203324_s_at	CAV2	0.7331
220979_s_at	ST6GALNAC5	0.7619	218898_at	FAM57A	0.7346	210968_s_at	RTN4	0.7283
219966_x_at	BANP	0.7472	201920_at	SLC20A1	0.7319	203909_at	SLC9A6	0.7240
219375_at	CEPT1	0.7438	210510_s_at	NRP1	0.7308	209835_x_at	CD44	0.7232
219716_at	APOL6	0.7414	202154_x_at	TUBB3	0.7268	204489_s_at	CD44	0.7223
213137_s_at	PTPN2	0.7325	211945_s_at	ITGB1	0.7261	205875_s_at	TREX1	0.7182
203944_x_at	BTN2A1	0.7279	212097_at	CAV1	0.7249	216080_s_at	C11orf9	0.7114
202727_s_at	IFNGR1	0.7216	214629_x_at	RTN4	0.7227	201601_x_at	IFITM1	0.7105
218084_x_at	FXYD5	0.7203	219131_at	TERE1	0.7201	204490_s_at	CD44	0.7103
219439_at	C1GALT1	0.7134	211509_s_at	RTN4	0.7199	221958_s_at	FLJ23091	0.7096
202638_s_at	ICAM1	0.7069	222206_s_at	NCLN	0.7182	210916_s_at	CD44	0.7085
208780_x_at	VAPA	0.7066	207021_at	ZPBP	0.7131	201798_s_at	FER1L3	0.7037

[a]The gene expression data of 16 breast cell lines determined by Affymetrix probe set HG U133A transmembrane-filtered were retrieved from the breast cell line expression profile database described in Neve *et al.* (2006) (http://www.ebi.ac.uk/arrayexpress/experiments/E-TABM-157/). The MFI value of each antibody binding to these 16 breast cell lines were log-transformed and standardized by the corresponding protein mean and standard deviation. The correlation coefficients (*R* value) between gene expression and antibody binding profile in these 16 cell lines were calculated.

[b]Affymetrix probe set HG U133A identifier, only transmembrane proteins listed.

[c]C225 (Cetuximab) recognizes EGFR; 2D6 binds ephrin type A receptor 2 (EphA2); 2B4 binds CD73.

Table II. Human scFv mAbs binding EGFR, EphA2, CD44 and CD73 on MDA-MB-231 cells

Target specificity	Number of scFvs identified	Affinity range[a] (nM)	Target identification method
EGFR	5	0.94–666	Recombinant EGFR-ECD
EphA2	9	0.6–375	Yeast displayed EphA2-ECD
CD44	5	0.71–482	Yeast displayed CD44-ECD
CD73	2	0.1–5	LC–MS-MS and binding competition

[a]As measured on MDA-MB-231 cells using scFv mAbs by flow cytometry.

Both EGFR and EphA2 are tyrosine kinase receptors activated by their ligands, EGF and Ephrin A1, respectively. The ability of the EphA2 mAbs to compete with Ephrin A1 was assessed by flow cytometry demonstrating that all EphA2 mAbs compete with Ephrin A1 (Zhou *et al.*, 2010). The EphA2 antibody (2D6) competed for the binding of Ephrin A1, and down-regulated cell surface EphA2 in a dose-dependent manner (Fig. 7B). 2D6 also inhibited the invasiveness of MDA-MB-231 cells, as measured using the Boyden Chamber assay (Fig. 7C). Because of its ligand-like functions, 2D6 is agonistic. Likewise, the anti-EGFR antibodies competed with EGF binding and blocked receptor signaling, and thus considered antagonistic antibodies (Zhou *et al.*, 2012) and data not shown.

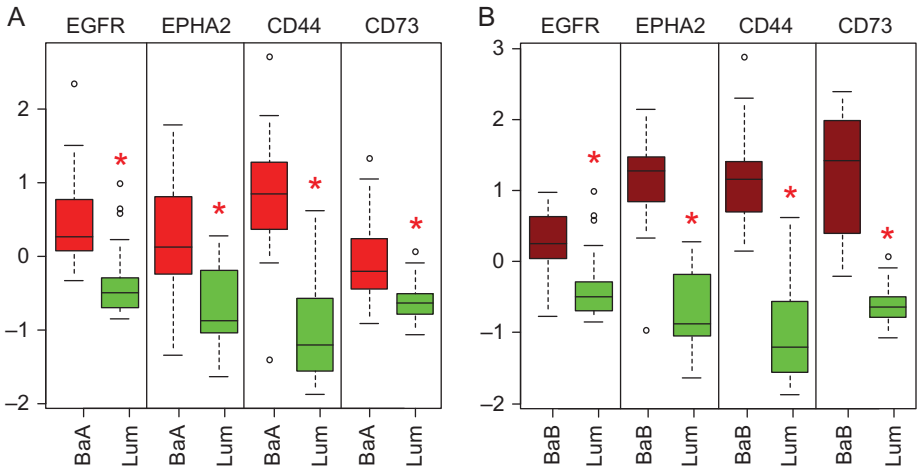

Fig. 5 Comparison of EGFR, EphA2, CD44 and CD73 gene expression levels between basal and luminal breast cancer subtypes in a panel of breast cancer cell lines. Boxplots contrasting the gene expression levels of EGFR, EphA2, CD44 and CD73 between the (**A**) Basal-A vs. Luminal and (**B**) Basal-B vs. Luminal breast cancer cell lines. Red asterisks (*) denote significant difference in mean expression ($P < 0.05$)

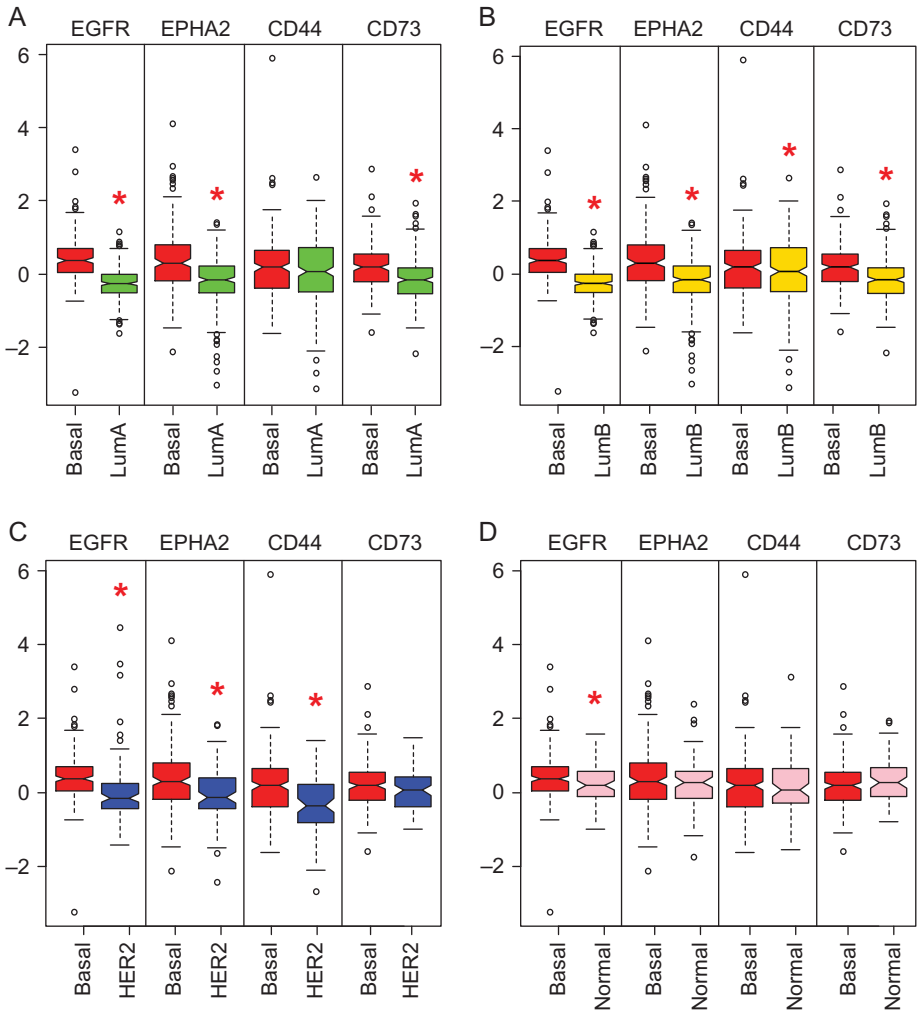

Fig. 6 Comparison of EGFR, EphA2, CD44 and CD73 gene expression levels between basal and other breast cancer subtypes in a pooled set of 683 primary breast cancers. Boxplots contrasting the gene expression levels of EGFR, EphA2, CD44 and CD73 between (**A**) Basal vs. Luminal A, (**B**) Basal vs. Luminal B, (**C**) Basal vs. Her2-enriched and (**D**) Basal vs. normal-like breast cancer intrinsic subtypes. Red asterisks (*) denote significant difference in mean expression ($P < 0.05$)

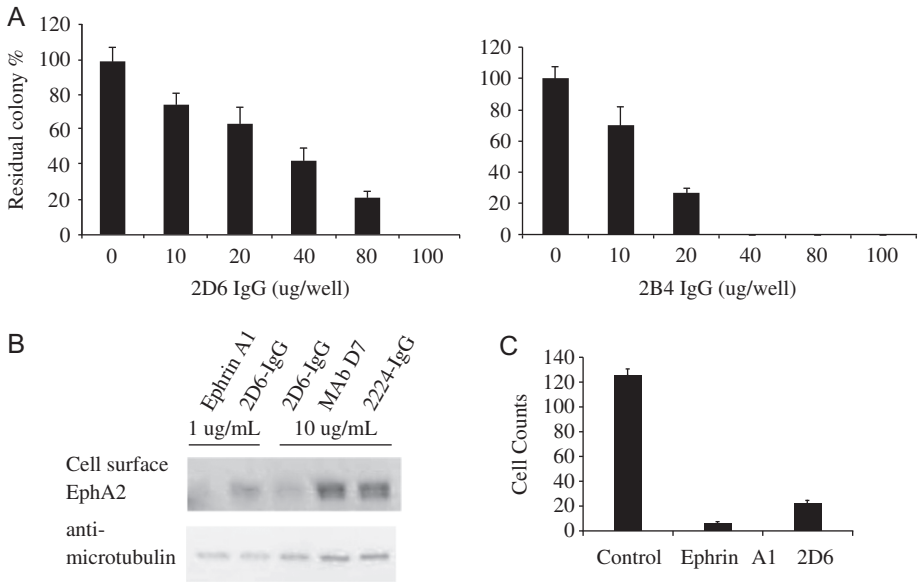

Fig. 7 Characterization of human mAbs targeting EphA2 and CD73. (**A**) 2D6 (EphA2) and 2B4 (CD73) IgG antibodies reduced the colony formation of MDA-MB-231 cells in a soft agar colony forming assay. (**B**) 2D6 reduced the surface level of EphA2 on MDA-MB-231 cells, while the commercial anti-EphA2 MAb D7 and the anti-EGFR 2224-IgG had no effect. (**C**) 2D6 inhibited MDA-MB-231 cells invasion in a modified Boyden Chamber Assay with the ligand Ephrin A1 as a control

We further evaluated the internalization properties of scFv antibodies employing scFv-conjugated and dye (DilC18(3)-DS)-labeled immunoliposomes. Within 3 h of culture exposure at 37°C, MDA-MB-231 cells overexpressing the antigenic targets of mAbs 2B4 (CD73), 2D6 (EphA2) and P2/4 (EGFR) showed specific uptake of the receptor-targeted immunoliposomes while untargeted liposomes were not taken up (Supplementary Fig. S1). Such antibodies could be used to deliver cytotoxic agents into the cell to achieve a therapeutic effect (Park *et al.*, 2001; Junttila *et al.*, 2011).

The EGFR, EphA2, CD73 and CD44 IgG mAbs were also used to quantify the surface density of the four receptors on 25 different breast cancer cell lines (9 luminal, 6 HER2-positive and 10 basal-like). Cell surface overexpression of all four targets was observed primarily on the basal-like cell lines, and was highest in the basal-like/claudin-low subset (Fig. 8A). Also, the EGFR, EphA2 and EGFR/EphA2 dual-targeted liposomes showed predominant uptake into the basal-like cell lines compared to the Luminal and HER2-positive cell lines, with enhanced cellular uptake for the dual-targeted liposomes (Fig. 8B), suggesting therapeutic potential for basal-like breast cancers.

Discussion

To discover cell surface receptors selectively overexpressed in breast and other cancer types for the purpose of developing new tumor-targeting therapeutics, we and others have employed the screening of naive phage Ab libraries against viable cells in culture (Marks *et al.*, 1992; Poul *et al.*, 2000; Heitner *et al.*, 2001; Liu *et al.*, 2004; Zhou *et al.*, 2010; Rust *et al.*, 2013). This approach has previously yielded unique mAbs suitable for targeted drug delivery (Park *et al.*, 2001; Roth *et al.*, 2007; Rust *et al.*, 2013). Here, we report the identification and validation of several internalizing mAbs against over-expressed basal-like breast cancer surface receptors, whose affinity and specificity supports their high potential for development into much-needed biomarkers and therapeutic agents against this most life-threatening form of breast cancer.

Four single transmembrane proteins, EphA2, CD73, CD44 and EGFR, were validated as overexpressed surface markers on basal-like breast cancers, based on the mRNA expression pattern of a broad panel of human breast cancer cell lines as well as a large curated archive of primary human breast cancer transcriptome data representing all intrinsic breast cancer subtypes. The pooled clinical sample analysis was designed to show the therapeutic targeting relevance of these four antigens. While we do not envision the use of these antibodies to classify basal-like from LumA, LumB or HER2+ breast cancers; we believe that this analysis indicates that clinical breast tumors with basal-like characteristics relatively overexpress these targets and, therefore, antibody-based therapies such as immunoliposomes targeting these targets should find greater utility against basal-like breast cancers.

Our findings are also supported by other independent research, as EGFR overexpression in some basal-type breast tumors has been demonstrated (Nielsen *et al.*, 2004; Livasy *et al.*, 2007), and mAbs against CD44 and CD73 capable of associating with basal-like breast cancer were recently developed using a hybridoma immunized with MDA-MB-231 cells and panning of a naïve phage Ab library against these same breast cancer cells (Rust *et al.*, 2013). Our finding of EphA2 receptor overexpression in basal-like breast cancers using the bioinformatics approach, and isolation of mAbs that can bind EphA2 with high affinity and simultaneously induce its rapid internalization are novel and provide evidence that support EphA2 as a therapeutic target (NCT03076372).

Strategies to develop mAbs for tumor-associated antigen receptors include hybridoma followed by humanization, humanized mouse hybridoma, and antibody display technologies. The distinct advantages of direct selection of phage Ab library on live cells include: (i) the antigens are present in their native conformations on the tumor cell surface; (ii) there is no need to produce recombinant antigen, which can be challenging for transmembrane receptors;

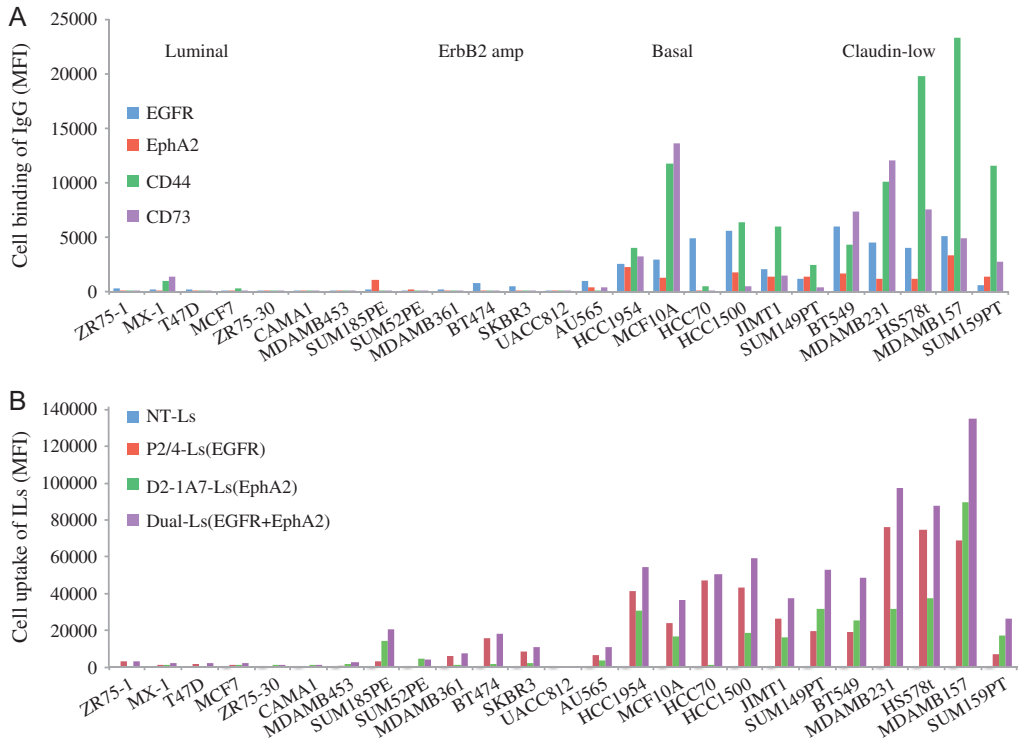

Fig. 8 Evaluation of receptor-mediated internalization of human mAb targeted liposomes in 25 breast cancer cell lines. (**A**) Comparison of cell surface receptor densities of EGFR, EphA2, CD44 and CD73 as shown by staining 25 breast cancer cell lines with corresponding mAbs. (**B**) Quantitative uptake of immunoliposomes with EGFR (P2/4) and EphA2 (D2-1A7) single and double targeting

(iii) the antigen–antibody pairs are identified concurrently; and (iv) internalizing antibodies useful for directing the intracellular drug delivery can be identified using an appropriate selection approach as described here and previously (Poul *et al.*, 2000; Heitner *et al.*, 2001; Liu *et al.*, 2004; Zhou *et al.*, 2010).

A challenge associated with the use of a phage Ab selection approach as described here is identification of the cognate cell surface antigen. In some instances, it is feasible to use mAbs to immunoprecipitate the cognate antigens followed by tandem liquid chromatography (LC–MS-MS) to reveal the antigen identity (CD73 mAbs 1A9 and 2B4 in this report and anti-CD9P1 Ab 3GA5 as previously reported (Goenaga *et al.*, 2007)). However, the immunoprecipitation method requires sufficient quantities of antigen proteins in native conformations, which can be challenging for multiple transmembrane receptors, and this approach is therefore not generalizable for all antigens.

In parallel, we employed an informatics strategy to identify cognate antigens more efficiently. We correlated cell line transcriptional profiles with the level of surface receptor as quantitated by the mAb under evaluation. Candidate antigens thus identified were displayed on yeast, as previously shown (Cochran *et al.*, 2004; Levy *et al.*, 2007; Zhou *et al.*, 2010). In fact, the concordance between antigen mRNA level and protein expression can vary between 46 and 68% (Pascal *et al.*, 2008), suggesting that not all cell surface receptors can be verified by this method. In the pilot study, the antigen probes for EGFR, CD73 and EphA2 were among the top 10 correlated antigens. In particular, the antigen for mAb 2D6 was revealed to be EphA2 using cells and antigen domains displayed on yeast, and binding competition to EphA2 using the Ephrin A1 ligand validated this finding. While this informatics strategy could be generalized for many cell surface receptors to accelerate the discovery of cognate tumor-associated antigens for selected mAbs, there are two limitations: (i) the *in silico* prediction cannot predict the top 10 antigens for all receptors and (ii) some antigens may not be readily displayed on yeast as was the case for CD73 where our mAbs 2B4 and 1A9 failed to bind yeast in flow cytometry analysis. When yeast display fails, mammalian cell overexpression or siRNA knockdown may also be used to complete the antigen identification and validation process (Goenaga *et al.*, 2007; Rust *et al.*, 2013).

Certain IgGs can modulate cell signaling, which is their proposed mechanism of anti-cancer activity (Sato *et al.*, 1983). We have shown that the EphA2 mAbs identified here have agonistic activity and decreased the colony formation of MDA-MB-231 cancer cells in soft agar culture. Similarly, CD73 mAbs also decreased colony formation of MDA-MB-231 cells. This *in vitro* antitumor activity needs to be further investigated using *in vivo* breast tumor models to determine if these mAbs have therapeutic activity comparable to those described by others (Carles-Kinch *et al.*, 2002; Terp *et al.*, 2013). The internalizing properties of our EphA2, EGFR and CD73 mAbs may be exploited to deliver cytotoxic agents into cancer cells, as shown for our HER2 internalizing mAbs (Poul *et al.*, 2000). We generated anti-HER2 immunoliposomes containing doxorubicin, then showed that they could selectively kill HER2+ cancer cells both *in vitro* and *in vivo*. Antitumor potential was confirmed in early phase clinical trials (Noble *et al.*, 2004) (Munster *et al.*, 2014) Following this same model of therapeutic translation, we are similarly coupling our basal-specific mAbs to chemotherapy-loaded immunoliposomes as a new therapeutic modality for the treatment of solid tumors, including basaloid breast cancer (NCT03076372).

Supplementary data

Supplementary data are available at *Protein Engineering, Design and Selection* online.

Author's contributions

Y.Z.: Conceived and designed experiments, developed methodology, analyzed and interpreted data and wrote the article. H.Z.: Performed experiments, acquired, analyzed and interpreted data. C.Y.: Performed the statistical analyses. L.Z.: Performed and analyzed cellular activity assays. S.C.H.: Collected and interpreted mass spectrometry data for identification of cognate antigen. D.C.D.: Provided immunoliposomes and developed methodology, interpreted data and edited article. S.F.J.: Edited the article. J.W.P.: Developed methodology of immunoliposomes, interpreted data and edited the article. C.C.B.: Developed methodology for statistical analysis, interpreted data and edited the article. J.D.M.: Conceived and designed experiments, interpreted data and wrote the article.

Conflict of interest

J.D.M., C.C.B. and J.W. have ownership interests in Merrimack Pharmaceuticals. They are also paid consultants and serve on the scientific advisory board of Merrimack Pharmaceuticals. D.D. is an employee of Merrimack Pharmaceuticals. The other authors disclosed no potential conflicts of interest.

Funding

Research reported in this publication was supported by NIH/NCI (P41 CA196276 to J.D.M., PI), (P50 CA058207 to Laura J. Van't Veer, PI), and (U24 CA143858, R21CA155679, and R01 CA071468 to C.C.B., PI); the Sandler Family Foundation; the Gordon and Betty Moore Foundation; NIH/NCI Center Core Grant (P30 CA082103 to Susan J. Fisher, PI); the Canary Foundation; and NIH National Center for Research Resources Shared Instrumentation Grant (S10RR029446-01 to Ewa Witkowska, PI).

References

Becerril,B., Poul,M.A. and Marks,J.D. (1999) *Biochem. Biophys. Res. Commun.*, **255**, 386–393.

Benedict,C.A., MacKrell,A.J. and Anderson,W.F. (1997) *J. Immunol. Methods*, **201**, 223–231.

Carles-Kinch,K., Kilpatrick,K.E., Stewart,J.C. and Kinch,M.S. (2002) *Cancer Res.*, **62**, 2840–2847.

Cochran,J.R., Kim,Y.-S., Olsen,M.J., Bhandari,R. and Wittrup,K.D. (2004) *J. Immunol. Methods*, **287**, 147–158.

de Hoon,M.J., Imoto,S., Nolan,J. and Miyano,S. (2004) *Bioinformatics*, **20**, 1453–1454.

Feldhaus,M.J., Siegel,R.W., Opresko,L.K. *et al* (2003) *Nat. Biotechnol.*, **21**, 163–170.

Foulkes,W.D., Smith,I.E. and Reis-Filho,J.S. (2010) *N. Engl. J. Med.*, **363**, 1938–1948.

Gietz,R.D. and Schiestl,R.H. (1991) *Yeast*, **7**, 253–263.

Goenaga,A.L., Zhou,Y., Legay,C. *et al* (2007) *Mol. Immunol.*, **44**, 3777–3788.

Heiser,L.M., Sadanandam,A., Kuo,W.-L. *et al* (2012) *Proc. Natl. Acad. Sci. USA.*, **109**, 2724–2729.

Heitner,T., Moor,A., Garrison,J.L., Marks,C., Hasan,T. and Marks,J.D. (2001) *J. Immunol. Methods*, **248**, 17–30.

Hochgrafe,F., Zhang,L., O'Toole,S.A. *et al* (2010) *Cancer Res.*, **70**, 9391–9401. doi:10.1158/0008-5472.CAN-10-0911.

Huie,M.A., Cheung,M.C., Muench,M.O., Becerril,B., Kan,Y.W. and Marks, J.D. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 2682–2687.

Irshad,S., Ellis,P. and Tutt,A. (2011) *Curr. Opin. Oncol.*, **23**, 566–577.

Junttila,T.T., Li,G., Parsons,K., Phillips,G.L. and Sliwkowski,M.X. (2011) *Breast Cancer Res. Treat.*, **128**, 347–356. doi:10.1007/s10549-010-1090-x.

Kreike,B., van Kouwenhove,M., Horlings,H., Weigelt,B., Peterse,H., Bartelink,H. and van de Vijver,M.J. (2007) *Breast Cancer Res.*, **9**, R65.

Leth-Larsen,R., Lund,R., Hansen,H.V., Laenkholm,A.V., Tarin,D., Jensen,O. N. and Ditzel,H.J. (2009) *Mol. Cell Proteomics*, **8**, 1436–1449. doi:10.1074/mcp.M800061-MCP200.

Levy,R., Forsyth,C., LaPorte,S., Geren,I., Smith,L. and Marks,J. (2007) *J. Mol. Biol.*, **365**, 196–210.

Liu,B., Conrad,F., Cooperberg,M.R., Kirpotin,D.B. and Marks,J.D. (2004) *Cancer Res.*, **64**, 704–710.

Liu,B., Huang,L., Sihlbom,C., Burlingame,A. and Marks,J.D. (2002) *J. Mol. Biol.*, **315**, 1063–1073.

Livasy,C.A., Perou,C.M., Karaca,G., Cowan,D.W., Maia,D., Jackson,S., Tse, C.-K., Nyante,S. and Millikan,R.C. (2007) *Hum. Pathol.*, **38**, 197–204.

Lochter,A., Srebrow,A., Sympson,C.J., Terracio,N., Werb,Z. and Bissell,M.J. (1997) *J. Biol. Chem.*, **272**, 5007–5015.

Loi,S., Pommey,S., Haibe-Kains,B., Beavis,P.A., Darcy,P.K., Smyth,M.J. and Stagg,J. (2013) *Proc. Natl. Acad. Sci. USA.*, **110**, 11091–11096. doi:10.1073/pnas.1222251110.

Marks,J.D., Griffiths,A.D., Malmqvist,M., Clackson,T.P., Bye,J.M. and Winter,G. (1992) *Nat. Biotechnol.*, **10**, 779–783.

Munster,P., Krop,I.E., Miller,K. *et al* (2014) *Cancer Res.*, **73**, Abstract P41229. doi:10.1158/0008-5472.

Nakai,K., Hung,M.C. and Yamaguchi,H. (2016) *Am. J. Cancer Res.*, **6**, 1609–1623.

Nellis,D.F., Ekstrom,D.L., Kirpotin,D.B. *et al* (2005) *Biotechnol. Prog.*, **21**, 205–220.

Network,C.G.A. (2012) *Nature*, **490**, 61–70.

Neve,R.M., Chin,K., Fridlyand,J. *et al* (2006) *Cancer Cell*, **10**, 515–527.

Nielsen,T.O., Hsu,F.D., Jensen,K. *et al* (2004) *Clin. Cancer Res.*, **10**, 5367–5374. doi:10.1158/1078-0432.CCR-04-022010/16/5367 [pii].

Noble,C.O., Kirpotin,D.B., Hayes,M.E., Mamot,C., Hong,K., Park,J.W., Benz,C.C., Marks,J.D. and Drummond,D.C. (2004) *Expert. Opin. Ther. Targets*, **8**, 335–353. doi:10.1517/14728222.8.4.335.

Nowakowski,A., Wang,C., Powers,D.B. *et al* (2002) *Proc. Natl. Acad. Sci. USA*, **99**, 11346–11350. doi:10.1073/pnas.172229899.

O'Connell,D., Becerril,B., Roy-Burman,A., Daws,M. and Marks,J.D. (2002) *J. Mol. Biol.*, **321**, 49–56.

Park,J., Kirpotin,D., Hong,K., Shalaby,R., Shao,Y., Nielsen,U., Marks,J., Papahadjopoulos,D. and Benz,C. (2001) *J. Control. Release*, **74**, 95–113.

Pascal,L.E., True,L.D., Campbell,D.S., Deutsch,E.W., Risk,M., Coleman,I. M., Eichner,L.J., Nelson,P.S. and Liu,A.Y. (2008) *BMC Genomics*, **9**, 246.

Perou,C.M., Sørlie,T., Eisen,M.B. *et al* (2000) *Nature*, **406**, 747–752.

Poul,M.A., Becerril,B., Nielsen,U.B., Morisson,P. and Marks,J.D. (2000) *J. Mol. Biol.*, **301**, 1149–1161.

Rakha,E.A., El-Sayed,M.E., Green,A.R., Lee,A.H., Robertson,J.F. and Ellis,I. O. (2007) *Cancer*, **109**, 25–32.

Razai,A., Garcia-Rodriguez,C., Lou,J. *et al* (2005) *J. Mol. Biol.*, **351**, 158–169.

Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and Mesirov,J.P. (2011) *Nat. Biotechnol.*, **29**, 24–26.

Roth,A., Drummond,D.C., Conrad,F., Hayes,M.E., Kirpotin,D.B., Benz,C.C., Marks,J.D. and Liu,B. (2007) *Mol. Cancer Ther.*, **6**, 2737–2746. doi:10.1158/1535-7163.MCT-07-0140.

Rust,S., Guillard,S., Sachsenmeier,K. *et al* (2013) *Mol. Cancer*, **12**, 11. doi:10.1186/1476-4598-12-11.

Sato,J.D., Kawamoto,T., Le,A.D., Mendelsohn,J., Polikoff,J. and Sato,G.H. (1983) *Mol. Biol. Med.*, **1**, 511–529.

Schier,R., Marks,J.D., Wolf,E.J. *et al* (1995) *Immunotechnology*, **1**, 73–81.

Sheets,M.D., Amersdorfer,P., Finnern,R. *et al* (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 6157–6162.

Sorlie,T., Perou,C.M., Tibshirani,R. *et al* (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 10869–10874.

Stagg,J., Divisekera,U., McLaughlin,N., Sharkey,J., Pommey,S., Denoyer,D., Dwyer,K.M. and Smyth,M.J. (2010) *Proc. Natl. Acad. Sci. USA*, **107**, 1547–1552. doi:10.1073/pnas.0908801107.

Terp,M.G., Olesen,K.A., Arnspang,E.C., Lund,R.R., Lagerholm,B.C., Ditzel, H.J. and Leth-Larsen,R. (2013) *J. Immunol.*, **191**, 4165–4173. doi:10.4049/jimmunol.1301274.

Thorvaldsdóttir,H., Robinson,J.T. and Mesirov,J.P. (2013) *Brief Bioinformatics*, **14**, 178–192. doi:10.1093/bib/bbs017.

Treco,D.A. and Winston,F. (2008) *Curr. Protoc. Mol. Biol.*, **Chapter 13**, Unit 13 12. doi:10.1002/0471142727.mb1302s82.

Zhao,L., Qu,L., Zhou,J., Sun,Z., Zou,H., Chen,Y.Y., Marks,J.D. and Zhou, Y. (2014) *PLoS One*, **9**, e111339. doi:10.1371/journal.pone.0111339.

Zhou,Y., Drummond,D.C., Zou,H., Hayes,M.E., Adams,G.P., Kirpotin,D.B. and Marks,J.D. (2007) *J. Mol. Biol.*, **371**, 934–947.

Zhou,Y. and Marks,J.D. (2012) *Methods Enzymol.*, **502**, 43–66. doi:10.1016/B978-0-12-416039-2.00003-3.

Zhou,Y., Zhao,L. and Marks,J.D. (2012) *Arch. Biochem. Biophys.*, **526**, 107–113. doi:10.1016/j.abb.2012.05.007.

Zhou,Y., Zou,H., Zhang,S. and Marks,J.D. (2010) *J. Mol. Biol.*, **404**, 88–99. doi:10.1016/j.jmb.2010.09.006.