

## RESEARCH ARTICLE

## Open Access



# Screening and characterization of novel specific peptides targeting MDA-MB-231 claudin-low breast carcinoma by computer-aided phage display methodologies

Franklin L. Nobrega<sup>1</sup>, Débora Ferreira<sup>1</sup>, Ivone M. Martins<sup>1</sup>, Maria Suarez-Diez<sup>2</sup>, Joana Azeredo<sup>1</sup>, Leon D. Kluskens<sup>1^</sup> and Lígia R. Rodrigues<sup>1\*</sup>

## Abstract

**Background:** Claudin-low breast carcinoma represents 19% of all breast cancer cases and is characterized by an aggressive progression with metastatic nature and high rates of relapse. Due to a lack of known specific molecular biomarkers for this breast cancer subtype, there are no targeted therapies available, which results in the worst prognosis of all breast cancer subtypes. Hence, the identification of novel biomarkers for this type of breast cancer is highly relevant for an early diagnosis. Additionally, claudin-low breast carcinoma peptide ligands can be used to design powerful drug delivery systems that specifically target this type of breast cancer.

**Methods:** In this work, we propose the identification of peptides for the specific recognition of MDA-MB-231, a cell line representative of claudin-low breast cancers, using phage display (both conventional panning and BRASIL). Binding assays, such as phage forming units and ELISA, were performed to select the most interesting peptides (i.e., specific to the target cells) and bioinformatics approaches were applied to putatively identify the biomarkers to which these peptides bind.

**Results:** Two peptides were selected using this methodology specifically targeting MDA-MB-231 cells, as demonstrated by a 4 to 9 log higher affinity as compared to control cells. The use of bioinformatics approaches provided relevant insights into possible cell surface targets for each peptide identified.

**Conclusions:** The peptides herein identified may contribute to an earlier detection of claudin-low breast carcinomas and possibly to develop more individualized therapies.

**Keywords:** Claudin-low breast cancer, Phage display, MDA-MB-231, PRWAVSP, DTFNSFGRVRIE

## Background

Breast cancer is the most frequent cancer among women, representing 25% of all cancer cases, and the most frequent cause of cancer death in less developed countries and the second in developed regions [1].

Breast cancer has long been recognized as a heterogeneous disease [2], challenging an effective detection, diagnosis and treatment. Initially based on morphological

observations, this heterogeneity has been confirmed by high-throughput methods such as molecular profiling with microarrays. These have allowed the identification of specific biomarkers whose presence or absence enable distinguishing breast cancers into different subtypes. The currently accepted biomarkers include the estrogen (ER), progesterone (PR) and human epidermal growth factor 2 (HER2) receptors [3], dividing breast cancer into the following subtypes: luminal A (ER<sup>+</sup>, PR<sup>+/-</sup>, HER2<sup>-</sup>), luminal B (ER<sup>+</sup>, PR<sup>+/-</sup>, HER2<sup>+</sup>), HER2 (ER<sup>-</sup>, PR<sup>-</sup>, HER2<sup>+</sup>), basal-like (ER<sup>-</sup>, PR<sup>-</sup>, HER2<sup>-</sup>) and claudin-low (ER<sup>-</sup>, PR<sup>-</sup>, HER2<sup>-</sup>) [4, 5]. The claudin-low subtype was initially clustered together with the

\* Correspondence: [lmr@deb.uminho.pt](mailto:lmr@deb.uminho.pt)

<sup>^</sup>Deceased on April 1st 2016

<sup>1</sup>Centre of Biological Engineering (CEB), University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Full list of author information is available at the end of the article



basal-like but the presence of unique features (e.g., downregulation of claudin-3 and claudin-4 and the low expression of proliferation marker Ki67) led to its own subtype [6, 7].

Each cancer subtype has a different prognosis and treatment response [4]. Luminal A and luminal B subtypes, characterized by the presence of ER, are commonly treated with hormone therapy with a good overall outcome; HER2 subtype, with the presence of HER2 can be treated with anti-HER2 monoclonal antibody therapy; but the basal-like and claudin-low subtypes, due to the absence of expression of a recognizable therapeutic target, lack targeted therapeutic options [8, 9]. Unfortunately, these represent about 19% of all breast cancer cases and include those with worst prognosis due to its aggressive and metastatic nature and high rates of relapse [10]. The identification of specific molecular biomarkers for these subtypes would be a valuable contribution to a more precise diagnosis and to the development of individualized therapies to different molecular subgroups.

However, the quest for molecular biomarkers specific for cancer cells remains a challenge due to the lack of affinity reagents that can specifically bind to unique molecular targets on the surface of these cells. The isolation and identification of such reagents is vital for clinical applications in cancer diagnosis and therapy [11]. Evolutionary screening techniques, such as phage display [12], have demonstrated incredible capacity to identify affinity reagents for a wide variety of targets (proteins, nucleic acids, inorganic materials, cells, among others) [13, 14]. In fact, phage display has already been used to generate recombinant antibody fragments that specifically recognize breast cancer subpopulations [15], as well as cell-targeting peptides for SK-BR-3 breast cancer cells [16]. In addition, phage display does not require prior knowledge of the cell surface, has low costs, and the cell-specific peptides identified typically present low immunogenicity [17, 18].

In this work, we used phage display to identify peptides specifically recognizing the claudin-low breast cancer cell line MDA-MD-231. The identification of such peptides could open new perspectives for the development of targeted therapies against this specific breast cancer subtype. Binding assays were performed to select the most specific peptides and a bioinformatics analysis was implemented to evaluate their potential targets on the cell surface.

## Methods

### Library diversity and preparation

The M13KE phage and its host, *Escherichia coli* ER2387, were obtained from New England Biolabs (NEB). Two different libraries of M13KE were used, namely a homemade 7-mer library and a commercial 12-mer library

from NEB (E8110S). The construction of the 7-mer library was performed as described in [19], using primers 5'-CATGCCCGGGTACCTTTCTATTCTC-3' and 5'- (NNN)<sub>7</sub>AGAGTGAGAATAGAAAGGTACCCG GG-3' and digested as in the protocol for M13KE DNA insertion (7.2 kb).

### Cell line and culture

The human cancer cell lines MDA-MB-231 (claudin-low subtype), SK-BR-3 (HER2 subtype), Hs 578 T (basal-like subtype) and MDA-MB-435 (melanoma [20]) were kindly provided by the Institute of Molecular Pathology and Immunology at the University of Porto (IPA-TIMUP). The human mammalian cell line MCF-10-2A (ATCC CRL-10781) is non-tumorigenic and was used as a control. MDA-MB-231, SK-BR-3, Hs 578 T, and MDA-MB-435 cells were routinely cultured in Dulbecco's Modified Eagle Medium (DMEM, Biochrom) supplemented with 10% (v/v) fetal bovine serum (FBS, Biochrom) and 1% (v/v) penicillin-streptomycin (Biochrom). MCF-10-2A cells were grown in a 1:1 solution of DMEM and HAM's F-12 medium supplemented with 5% horse serum (Merck Millipore), 20 ng.mL<sup>-1</sup> epidermal growth factor (Merck Millipore), 100 ng.mL<sup>-1</sup> cholera toxin (Sigma-Aldrich), 0.01 mg.mL<sup>-1</sup> insulin (Sigma-Aldrich), 500 ng.mL<sup>-1</sup> hydrocortisone, 95% (Sigma-Aldrich) and 1% penicillin-streptomycin. All cell lines were cultured at 37 °C and 5% CO<sub>2</sub>. Subculturing was performed at 80% confluence, by washing the monolayer with sterile phosphate buffered-saline (PBS), pH 7.4, without Ca<sup>2+</sup> and Mg<sup>2+</sup>, and detaching the cells with Trypsin/EDTA solution 0.05%/0.2% (w/v) (Biochrom). The cell suspension was centrifuged at 250 × g for 7–10 min and the cell pellet was resuspended on fresh growth medium, counted and split according to the experimental needs.

### Panning experiments – conventional selection versus BRASIL

Both conventional phage display and BRASIL [21] methods were used to compare their performance in the selection of a peptide specific to the MDA-MD-231 cells. The BRASIL method is in principle faster than the conventional panning and by using counter-selection it reduces the number of false positives. However, this methodology uses cells in suspension, which may hide surface receptors that are only available in the adherent state. The panning experiments with both methodologies were performed equally for the 7-mer and the 12-mer libraries. The experimental setting can be seen in Additional file 1: Table S1.

### Conventional selection (surface panning procedure – direct target coating)

One mL of MDA-MB-231 cell suspension at a concentration of 10<sup>6</sup> cells.mL<sup>-1</sup> was added to a 6-well

microtiter plate and incubated overnight at 37 °C in a 5% CO<sub>2</sub> humidified incubator. The medium was then removed and the wells completely filled with blocking buffer (0.1 M NaHCO<sub>3</sub> (pH 8.6, Sigma), 5 mg.mL<sup>-1</sup> Bovine Serum Albumin (BSA) (Sigma) solution IgG-free, low endotoxin suitable for cell culture (Sigma). After an incubation of 1 h at 4 °C, the blocking solution was discarded and the wells washed 6 times with Tris Buffered Saline with Tween-20 (TBST, TBS with 0.1% (v/v) Tween-20) (Sigma-Aldrich). One mL of a 100-fold dilution in TBST of the library (7-mer or 12-mer) (1x10<sup>11</sup> for a library with 2x10<sup>9</sup> clones) was added to the coated wells and rocked gently for 60 min at 4 °C (to limit phage internalization). The non-binding phage was discarded and the wells were washed 10 times with TBST. The bound phage was then eluted with 750 µL of PBS 1x (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub> and 1.8 mM KH<sub>2</sub>PO<sub>4</sub>), and rocked gently for 60 min at 4 °C. The eluate was transferred to a microcentrifuge tube and the titer was determined using the double layer agar technique [22] in LB plates containing 100 µM IPTG and 20 µg.mL<sup>-1</sup> X-gal, counting the blue colonies. The remaining eluate was amplified by adding the eluate to 20 mL early-log ER2738 culture and incubating with vigorous shaking for 4.5 h at 37 °C. The culture was spun at 12,000 × g for 10 min at 4 °C, and the supernatant was transferred to a fresh tube and re-spun. The upper 80% of the supernatant was transferred to a new tube and the phage was precipitated with 1/6 volume of 20% polyethylene glycol (PEG) 8000/2.5 M NaCl for at least 2 h at 4 °C. This solution was centrifuged at 12,000 × g for 15 min at 4 °C, the supernatant was discarded and the phage pellet was suspended in 1 mL TBS. PEG/NaCl precipitation was repeated and the final pellet suspended in 200 µL TBS. The titer was determined as previously described. The whole process was repeated for a total of 8 rounds of panning.

A control panning experiment was carried out using streptavidin as the target, including 0.1 µg.mL<sup>-1</sup> streptavidin in the blocking solution. The bound phage was eluted with 0.1 mM biotin in TBS for at least 30 min. After 3 rounds of enrichment/amplification, the consensus sequence for streptavidin-binding peptides was assessed to confirm the inclusion of the motif His-Pro-Gln.

#### **BRASIL**

A biopanning protocol was used as described in [21]. Briefly, MDA-MB-231 cells.mL<sup>-1</sup> were collected, centrifuged (250 × g, 10 min) and the pellet suspended in 1 mL of complete DMEM medium, containing 1% (w/v) BSA. The solution was centrifuged and this step repeated 3 times; the cells were re-suspended in complete growth medium containing 3% (w/v) BSA solution and kept on ice. Ten µL of the phage library (7-mer or 12-

mer) were added to the previous cell suspension and incubated on ice for 4 h. A bubble of 300 µL PBS was formed on a non-miscible organic phase (cyclohexane:dibutyl phthalate (1:9, v/v, Sigma)), and 200 µL of the cell suspension incubated with the phage library were gently inserted into the bubble. After centrifuging at 10,000 × g for 10 min, the pellet was recovered and washed with 50 µL Tris-HCl (10 mM, pH 9.5). Eluted phages were amplified between rounds using *E. coli* ER2738, purified and concentrated with 20% PEG 8000/2.5 M NaCl. Phage titer was determined as described above. The amplified phages were used for additional rounds of biopanning in a total of eight. A final round of counter-selection with MCF-10-2A cells (non-tumorigenic) was performed, differing from the previous rounds in the fraction collected, which in this case was the aqueous phase containing the phages that did not bind to the control cells.

#### **Preliminary analysis of the specificity and selectivity of a phage pool**

##### **Flow cytometry analysis**

To characterize pool specificity and selectivity, the last round of the 12-mer phage pool from conventional panning was conjugated with Alexa 488 and analyzed using flow cytometry to evaluate the binding to MCF-10-2A (control, non-tumorigenic cells), MDA-MB-231, MDA-MB-435, SK-BR-3 and Hs 578 T cell lines. Briefly, 1x10<sup>5</sup> cells were harvested, washed in PBS and blocked using PBS with 3% BSA at 4 °C for 1 h. Subsequently, the cells were washed with PBST 1x (PBS with 0.1% (v/v) Tween-20) and were incubated with 100 µL of fluorescent phage particles. The cells were rinsed again with PBST 1x and finally resuspended in 200 µL of PBS for flow cytometry analysis using a EC800™ flow cytometer analyzer (Sony Biotechnology Inc.) counting 20,000 events.

##### **Tissue section analysis**

For immunohistochemical analysis, serial sections of paraffin-embedded 231 mammary cancer tissue sections, kindly provided by Dr. João Nuno Moreira (CNC, Coimbra, Portugal), were treated as described in [23]. To maximize antibody binding, antigen retrieval was performed by heating the slides in 10 mM sodium citrate buffer (pH 6.0) at 95 °C for 20 min and the slow cooling at room temperature in the same buffer for about 20 min. Tissues were maintained humid at all time. Tissue sections were blocked using a 5% BSA solution and were incubated at room temperature for 30 min. Immunostaining was performed by adding 100 µL of the last round of the 12-mer phage pool (10<sup>9</sup> PFUs.mL<sup>-1</sup>) to the tissue overnight at 4 °C [24, 25]. Sections were washed 4 times in TBST 1x for 5 min and 100 µL of the primary antibody rabbit anti-fd bacteriophage (working dilution

of 1:5000 in BSA 1%), was added and incubated at 4 °C overnight. Sections were washed several times with TBST 1x and were challenged with the fluorescein isothiocyanate (FITC)-labelled goat anti-rabbit IgG secondary antibody (working dilution of 1:40 in 1% BSA) for 2 h at room temperature. After additional washing of the sections with TBST buffer, sections were counterstained with 4', 6 - diamidino-2-phenylindole (DAPI, Vector Laboratories) for nuclear labelling and were mounted with Vectashield® mounting medium (Vector Laboratories). The tissue sections were allowed to dry for 1 h at room temperature in the dark and were sealed with nail polish. Images of the slides were captured using an Olympus BX51 microscope incorporated with a high-sensitivity camera Olympus DP71 with 60× magnification.

### Selection and screening of cell-specific peptides

#### Preparation of individual clones for peptide analysis

Single-stranded DNA (ssDNA) was prepared according to the standard protocol described in [19], using iodide buffer (10 mM Tris-HCl, 1 mM EDTA and 4 M NaI (Sigma-Aldrich), pH 8.0) and ethanol precipitation. The DNA pellet was suspended in 30 µL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), quantified using Nanodrop 1000 and confirmed by 2% gel electrophoresis in SGTB (GRISP) buffer 1× at 200 V for 30 min.

#### PCR and confirmation electrophoresis

The insert sizes of the individual clones, as well as of the complete library were assessed by PCR using the forward primer 5'-TTAACTCCCTGCAAGCCTCA-3' and the reverse primer 5'-CCCTCATAGTTAGCGTAACG -3'. PCR reactions were carried out using KAPA *Taq* polymerase in 20 µL reaction volume, containing 2 µL of phage DNA. The PCR conditions were the following: 25 cycles of denaturation at 95 °C for 30 s; annealing in the temperatures range from 45 to 70 °C, for 30 s; and extension at 72 °C for 30 s. Amplification was confirmed by 2% gel electrophoresis in SGTB buffer 1× at 200 V for 30 min.

#### DNA sequencing and insert analysis

The DNA products obtained were prepared for sequencing using Illustra ExoProStar 1-Step (GE Healthcare) and sent to Macrogen Inc. service using the M13-P111 sequencing primer 5'- TTAAGTCCCTGCAAGCCTCA-3', provided with the Ph.D.12-mer library kit for forward reading and the primer 5' -CCCTCATAGTTAGCGTAACG-3' for reverse reading. The Vector NTI Advance 11.5.0 software (Invitrogen - Life Technologies) was used for the analysis of correct insertion of the peptides taking into account that the displayed peptides are expressed at the

N-terminus of pIII, followed by a short spacer (Gly-Gly-Gly-Ser) and then the wild-type pIII sequence.

### Binding assays

#### Binding assay with counting of blue colony forming units (pfu)

The binding of the peptides displayed on M13KE phage was evaluated following a procedure similar to the conventional panning. First, the individual clones were amplified, centrifuged at 12,000 × *g* for 10 min at 4 °C, and the supernatant used for phage concentration with 20% PEG 8000/2.5 M NaCl. Phages were suspended in 50 µL TBS and the titer was determined using the double layer agar technique. Then, 1 mL of MDA-MB-231 cells at a concentration of 10<sup>6</sup> cells.mL<sup>-1</sup> was added to a 6-well microtiter plate and incubated overnight at 37 °C and 5% CO<sub>2</sub>. MDA-MB-435 cells were used as a negative control in the same conditions. The cell medium was removed and the wells were washed 6 times with TBST. Then, 1 mL of each M13KE-peptide suspension, at a concentration of 1×10<sup>11</sup> PFU.mL<sup>-1</sup> was added to the wells and incubated for 60 min at 4 °C. The non-binding phage was discarded and the wells were washed 10 times with TBST. The bound phages were then eluted with 750 µL of PBS 1x and rocked gently for 60 min at 4 °C. The eluate was collected and the titer was determined using the double layer agar technique in IPTG/X-gal plates.

#### ELISA with direct target coating

ELISA was performed to rapidly determine whether a selected phage clone binds the target, using the protocol described in the NEB Phage Display manual [19]. For each clone to be characterized, one row of coated (with target cells) and uncoated wells were used. Plates were read at 405 to 415 nm (Promega Glomax 20/20 luminometer) and the signals (RLUs) obtained with and without target protein (cells) were compared.

### Bioinformatics analysis

#### Library analysis

Sequence similarities between the peptides obtained in this work and peptides reported in the literature targeting cancer cells (see Additional file 2: Table S3) were scored using Blossum45 matrices and the Needleman-Wunsch algorithm as implemented by the pairwise alignment function from the R *Biostrings* package version 2.38.2 [26]. The symmetric matrix containing the scores for the pairwise sequence alignments,  $SC(i,j)$ , was converted into a similarity matrix taking into account the background values for each sequence following a procedure similar to the Context Likelihood of Relatedness (CLR) algorithm used to detect spurious association in transcriptional or metabolite association networks [27, 28]. Briefly, the



likelihood of  $SC(i,j)$  is estimated using a null model given by considering all the alignment scores involving independently sequences  $i$  and  $j$ ,  $SC_i$  and  $SC_j$ , respectively. The background score is approximated as a joint normal distribution with  $SC_i$  and  $SC_j$  treated as independent variables. The final form of the likelihood estimate is:

$$f(z_i, z_j) = \sqrt{z_i^2 + z_j^2} \tag{1}$$

where

$$z_i = \max\left(0, \frac{SC(i, k) - \mu_i}{\sigma_i}\right) \tag{2}$$

and  $\mu_i$  and  $\sigma_i$  are, respectively, the mean and the standard deviation of the empirical distribution of  $SC(i, k)$  with  $k = 1, \dots, n$ , and  $n$  the total number of considered sequences. The similarity estimate is then a matrix with entries  $f(z_i, z_j)$ . The similarity estimate was normalized, through dividing by its highest values, to use in Multidimensional scaling (MDS) plots, clustering and heatmap reconstruction using the R *gplots* library [29].

**Docking studies**

Known biomarkers of breast cancer were selected from a literature and databases search (see Additional file 3: Table S4). The biomarkers found were retrieved through the Kyoto Encyclopedia of Genes and Genomes (KEGG) for pathways and function analysis of biomarkers, UniProt for protein characterization and amino acid sequences, GenBank for gene sequences, and Protein Data Bank (PDB) for tri-dimensional protein structures [30]. When protein structures were not available, they were predicted using the PHYRE2 software [31] and the peptide structures were predicted using PEPstrMOD [32, 33]. The resulting pdb files were used in a protein-peptide analysis performed using ClusPro 2.0 [34, 35] in all available models, by the peptide sequences identified by phage display against the tri-dimensional structures of the breast cancer biomarkers. Weighted score ( $E$ ) was obtained by:

$$E = 0.40E_{rep} + -0.40E_{att} + 600E_{elec} + 1.00E_{DARS} \tag{3}$$

(Balanced coefficients)

$$E = 0.40E_{rep} + -0.40E_{att} + 1200E_{elec} + 1.00E_{DARS} \tag{4}$$

(Electrostatic-favored coefficients)

$$E = 0.40E_{rep} + -0.40E_{att} + 600E_{elec} + 2.00E_{DARS} \tag{5}$$

(Hydrophobic-favored coefficients)

$$E = 0.40E_{rep} + -0.10E_{att} + 600E_{elec} + 0.00E_{DARS} \tag{6}$$

(Vand der Waals and Electrostatic coefficients)

where the lowest energy state represents the highest binding. The tri-dimensional model structures obtained were visualized using UCSF Chimera version 1.10.2 [36]. Alignments were scored using Blosum45, 50 and 62 matrices.

**Statistical analysis**

GraphPad Prism 5.03 (GraphPad Software, Inc.) was used for statistical analysis of the data. The significance of differences was evaluated using the One-way ANOVA with Tukey's Multiple Comparison Test, considering a significance level of 95%.

**Results**

**Identification by phage display of a peptide that recognizes the breast cancer cell line MDA-MB-231**

Phage display search of ligands specific for breast cancer cell surface receptors, as any other variety of targets, is a balance between the affinity to the target and its frequency on the library pool. Therefore, the library heterogeneity is a critical step for the success of panning experiments. In this study, we initially used a commercial 12-mer library aiming to isolate highly specific peptides directed against potential biomarkers present on the cell surface of MDA-MB-231 cells. For this purpose, we used the conventional phage display methodology. The phage pool of the last round of phage display was subjected to preliminary assays against several cell lines (MDA-MB-231, MCF-10-2A, SK-BR-3, Hs 578 T and MDA-MB-435) using flow cytometry to evaluate its specificity for MDA-MB-231 cells. The flow cytometry results, presented in Additional file 4: Figure S1, clearly indicate the selectivity of the phage pool towards MDA-MB-231 cells, with statistical significance as compared to the remaining cell lines evaluated. Then, this preliminary analysis was refined to study the interaction of the phage pool with the MDA-MB-231 cells by immunohistochemistry. Additional file 5: Figure S2 demonstrates binding of the phage pool to MDA-MB-231 tissue sections (identified by green fluorescence in Additional file 5: Figure S2B), in contrast to the wild type M13KE phages, which exhibit no staining (Additional file 5: Figure S2A), thus clearly suggesting the capacity of the peptides selected by phage display techniques to interact with the target cells.

With these initial results we could confirm the possibility of obtaining specific and selective peptides for the MDA-MB-231 cell line and so, we enlarged the study using an additional phage display library, containing only

7 amino acids (7-mer library), as well as a more recently developed phage display methodology, Biopanning and rapid analysis of selective interactive ligands (BRASIL). A library of smaller peptides may offer an advantage over the 12-mer library on the strength of binding of the peptides selected. Additionally, BRASIL presents the advantages of being faster and using counter-selection (to remove peptides that bind to targets present on other cells), but can be limited by the use of suspended cells, potentially hiding surface receptors only present in the adherent state.

For each phage display methodology and library, eight rounds of panning were performed and the peptides obtained from the last panning round of each experimental set (details provided in Additional file 1: Table S1) are presented in Table 1. Also, the consensus sequence with the respective overall percentage was determined (Table 1).

Conventional phage display and BRASIL methodologies resulted in similar consensus sequences. In fact, for the 7-mer library the sequence is identical (PRLNVSP), and for the 12-mer library only the first two amino acids are different (TTFNSFGRVRIE for the conventional method and WWFNSFGRVRIE for BRASIL). On the other hand, comparing the two libraries herein used, the consensus peptides obtained are very different. Furthermore, the overall percentage of consensus is higher for the commercial 12-mer library (86 %, 87%) than for the home-made 7-mer library (70 %, 60%).

#### Binding assays

Peptides 1.3(7/52) (PRWAVSP), 5.3(14/45) (WWFNSFGRVRIE), 5.3(19/45) (WWFFSFGRVRIE), 6.2(8/17) (TTEYSFGRTSTL) and 6.2(9/17) (DTFNSFGRVRIE) were

**Table 1** Conventional and modified BRASIL enrichment results based on binding affinity to the breast cancer MDA-MB-231 cell line and a counter selection with the non-tumorigenic MCF-10-2A cell line with the 7-mer and 12-mer libraries

Method	7-mer library		12-mer library	
	Phage clone	Sequence	Phage clone	Sequence
Conventional	3.1 (5/27)	PRLNTSP	6.2 (1/17)	TFNSFGRVAIE
	3.1 (6/27)	PTLDVAP	6.2 (2/17)	TFCSFGRVRIE
	3.1 (8/27)	PQLNVSP	6.2 (3/17)	TFNSFGRVRIE
	3.1 (9/27)	PGAQVSP	6.2 (4/17)	TFNSFGRVHWE
	3.1 (10/27)	PRTNVAP	6.2 (5/17)	TFNSFGKVRIE
	3.1 (11/27)	PRKTVSP	6.2 (7/17)	TYNSFGRVRIE
	3.1 (12/27)	PAMNVSP	6.2 (8/17)	TTEYSFGRTSTL
	3.1 (13/27)	PQENASP	6.2 (9/17)	DTFNSFGRVRIE
	3.1 (19/27)	PSLNVSP	6.2 (10/17)	TFNSFGRVRIQ
	3.1 (20/27)	ARLNVAP	6.2 (12/17)	TFNSFGRVRIE
	3.1 (21/27)	TMLMVRP	6.2 (13/17)	TNFNDFKVRGE
	3.1 (22/27)	ARLNTQP	6.2 (14/17)	TFSSFGRVRIG
	3.1 (23/27)	PMMAVAP	6.2 (15/17)	TFNSNGRWIE
	3.1 (26/27)	PRLNVSP	6.2 (16/17)	TFNSFGRVRRG
	3.1 (27/27)	PRLNVSP	6.2 (17/17)	TFNSFYRVRIE
		<i>Consensus (overall %)</i>	<b>PRLNVSP (70%)</b>	<i>Consensus (overall %)</i>
BRASIL	1.3 (2/52)	PRQNQSP	5.3 (1/45)	WWFNSFGRVRIE
	1.3 (3/52)	QLKNVTP	5.3 (3/45)	WAFNSFGRVRIE
	1.3 (5/52)	PRLNVAT	5.3 (5/45)	WWFNKFGKVRIE
	1.3 (6/52)	PRLNVTT	5.3 (14/45)	WWFNSFGRVRIE
	1.3 (7/52)	PRWAVSP	5.3 (16/45)	IWFNSFFRVRIE
	1.3 (8/52)	PRLNHSP	5.3 (19/45)	WWFFSFGRVRIE
	1.3 (10/52)	PQMTAMP	5.3 (20/45)	GWFNSFGRWSWL
	1.3 (11/52)	MFLNGAP	5.3 (23/45)	WWFNSFGRVRIE
	1.3 (15/52)	TRLQVSP	5.3 (43/45)	WWFFSFGRWSWL
	1.3 (19/52)	WRMAHSP	5.3 (45/45)	WWFNSFGRVRIE
		<i>Consensus (overall %)</i>	<b>PRLNVSP (60%)</b>	<i>Consensus (overall %)</i>

Bold words represent the consensus (overall %) for the two methods

selected among those identified by phage display to assess in vitro for their binding affinity to MDA-MB-231 (claudin-low breast cancer subtype), by incubation of the cells with M13KE phages containing each peptide in analysis. The melanoma MDA-MB-435 cells were used as a negative control to evaluate the specificity of the peptides for the breast cancer MDA-MB-231 cells. The results, presented as the ratio between the concentration of phages bound to each cell line and the initial phage concentration used, are shown in Fig. 1.

The phages displaying the selected peptides have a higher binding affinity to MDA-MB-231 cells than to MDA-MB-435 cells, with the differences ranging from 0.55 (corresponding to 6 logs, for peptides 5.3(14/45), sequence WWFNSFGRVRIE and 5.3(19/45), sequence WWFFSFGRVRIE) to 0.80 (9 logs, for peptides 1.3(7/52), sequence PRWAVSP and peptide 6.2 (9/17) sequence DTFNSFGRVRIE), with the latter two demonstrating the most promising results in terms of specificity and binding strength.

Enzyme-linked immunosorbent assays (ELISA) were performed with the selected peptides against MDA-MB-231 cells. MDA-MB-435 cells were used as a negative control and streptavidin as a positive control (using an M13KE phage displaying affinity peptides towards streptavidin). Results were read in a luminometer and the relative light units (RLUs) obtained are shown in Fig. 2.

The ELISA assays are in good agreement with the results obtained from the binding assays described above (Fig. 1), with all peptides showing higher affinity to the MDA-MB-231 cells than to the MDA-MB-435 cells. The

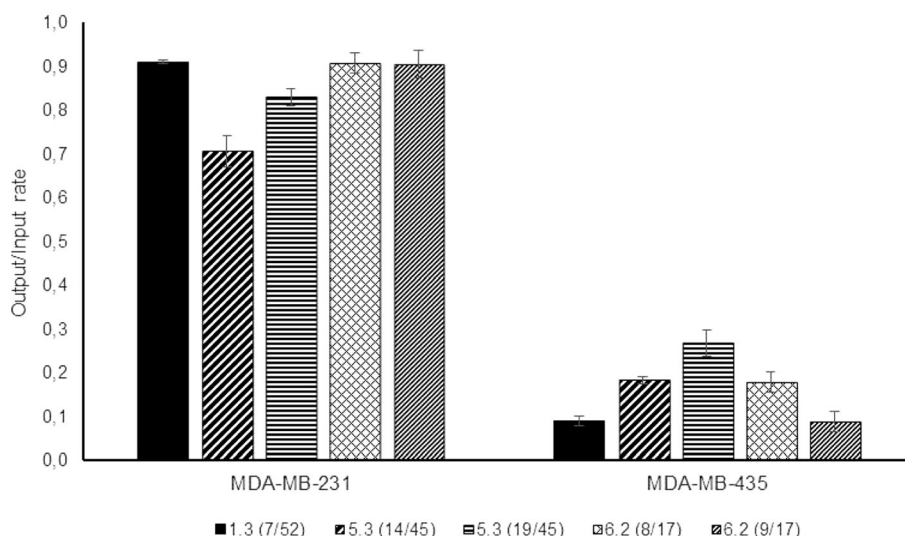
differences observed between the two cell lines range from 3 to 4 logs.

**Library analysis**

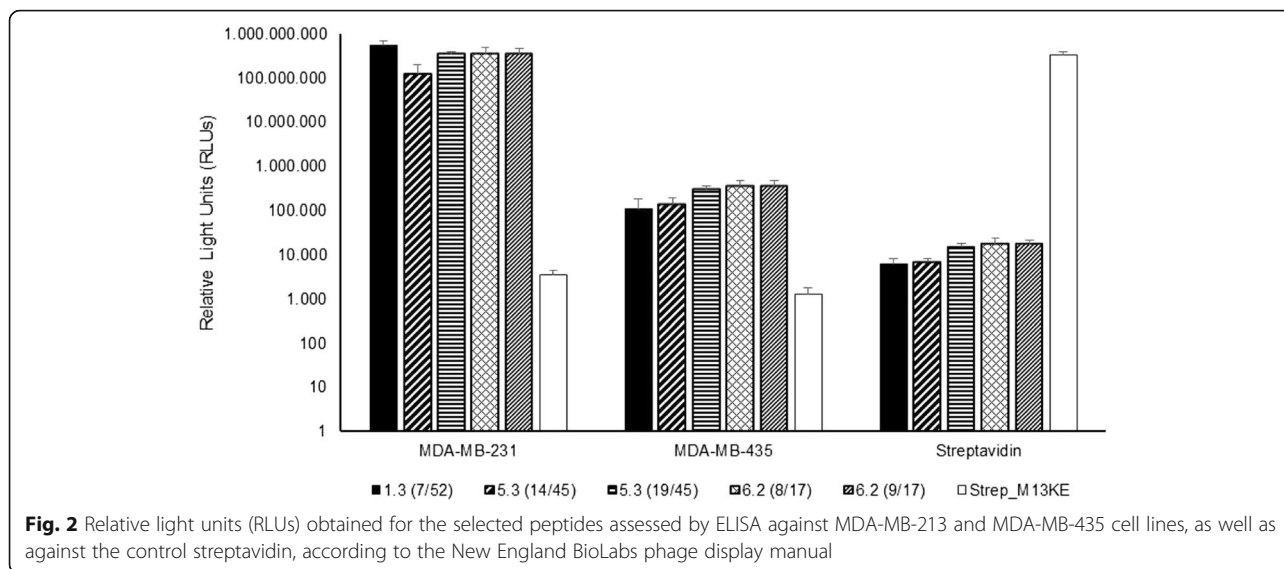
The peptides obtained in this work were compared to previously reported peptides (specific to breast cancer cells) to assess possible similarities. This was performed using pair-wise sequence alignments to prevent the bias towards the discovery of consensus sequence obtained when using multiple sequence alignments. Blosum45, 50 and 62 were used and compared, with the Blosum45 matrix being chosen to score the alignments since it is more adequate to score divergent sequences. An initial analysis demonstrated a high impact of sequence length on the similarity computation (see Additional file 6: Figure S3). Therefore, to consider the local background of each sequence regarding the alignment score, the CLR algorithm was adapted to this context. This algorithm has been used successfully to take the local background into account when assessing similarities between gene expression profiles or metabolite concentrations [27, 28].

The multidimensional scaling of the peptides can be seen in Fig. 3, where graphical distances between the item represents the (dis)similarities between the sequences. The algorithm places the newly identified sequences in the outskirts of the figure, indicating an average low similarity shared with previously identified peptides.

The similarities between all sequences were illustrated in a heatmap (Additional file 6: Figure S3). Even though the local context of each sequence has been considered, there was still a prevalence of association between



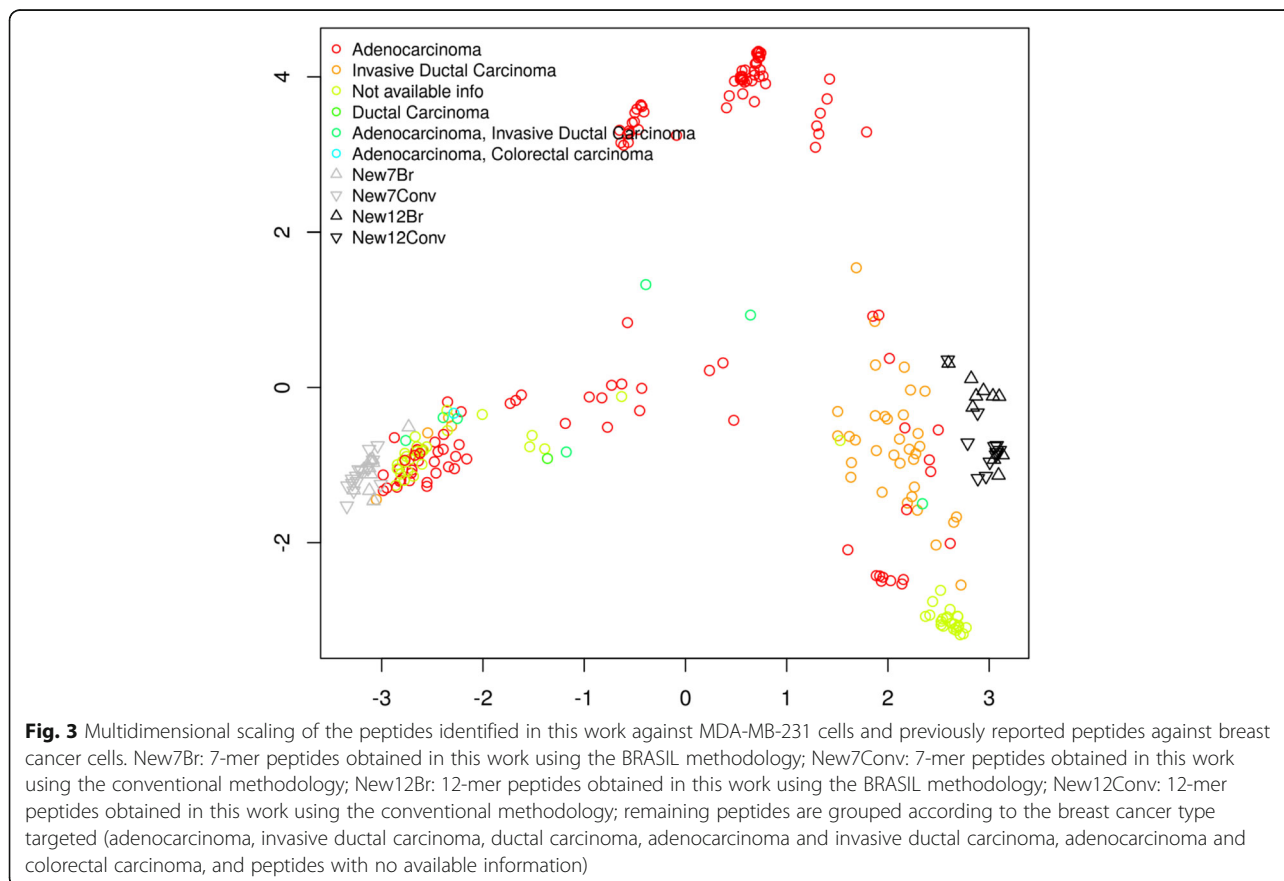
**Fig. 1** Binding assays of selected peptides against MDA-MB-231 and MDA-MB-435 cell lines, using an initial phage concentration of  $1 \times 10^{11}$  PFUs.mL<sup>-1</sup>. Results are expressed as the ratio between the concentration of phages bound to the cells (output) and the initial phage concentration used (input)



sequences of similar length. Therefore, to fully consider the effect of this bias, separated heatmaps for the 7-mer (Fig. 4) and 12-mer peptides (Fig. 5) were built only considering peptides of the same length. Results show that indeed the newly identified peptides are far (in sequence space) from those previously reported.

**Docking studies**

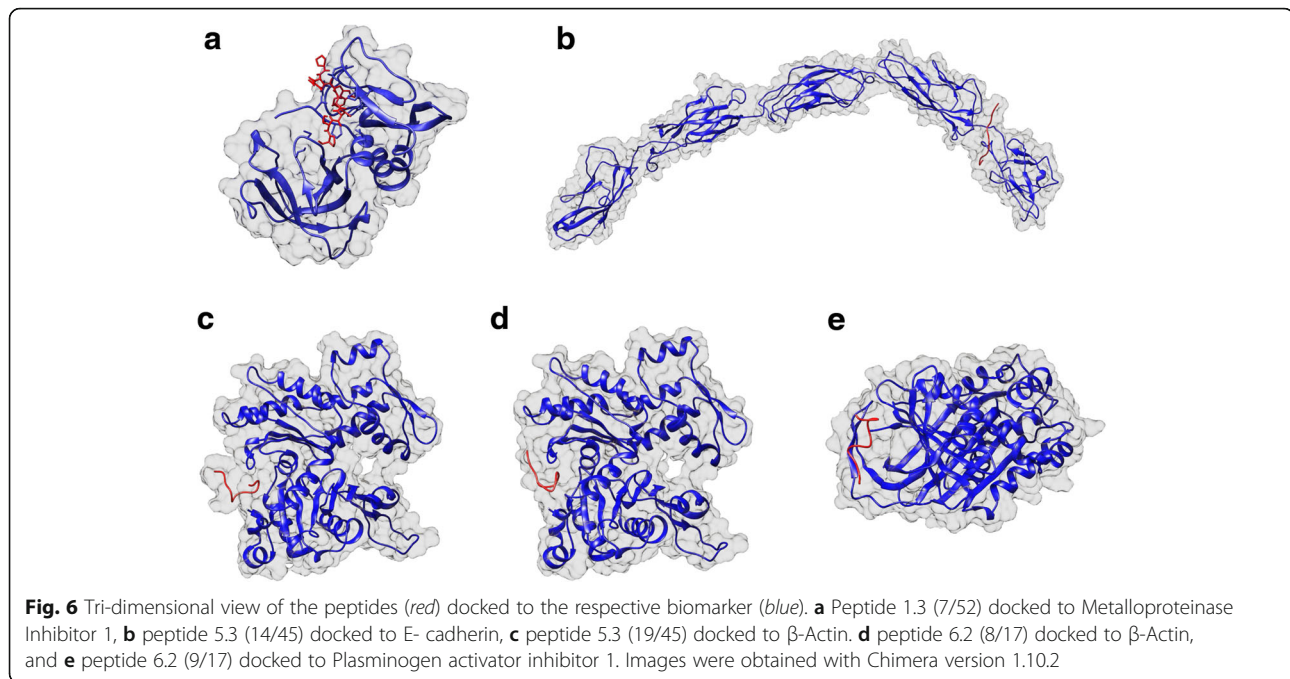
A structural bioinformatics approach was implemented to identify potential targets of the peptides in the MDA-MB-231 cells. For this purpose, established biomarkers present in breast cancer cells were retrieved from the literature using search engines such as PubMed (with











were used in counter-selection. Afterwards, immunohistochemistry analysis of the phage pool against tissue sections of the target cells (MDA-MB-231) was carried out (Additional file 5: Figure S2), demonstrating the binding affinity of the pool for the target. These initial results proved the feasibility of obtaining a specific peptide targeting the MDA-MB-231 cells using phage display approaches. To increase the possibility of identifying peptides with strong binding affinities, an additional phage display methodology (BRASIL) and a home-made 7-mer M13KE library were included in this study.

The 7-mer and 12-mer libraries led to different consensus sequences (Table 1). This indicates a strong influence of the library on the phage display results, which was expected due to the difference in length of the peptides (7 or 12 amino acids). Indeed, although the 7-mer library is adequate for a biopanning strategy, it is more useful for targets requiring binding elements concentrated in a short sequence of amino acids [38, 39]. In turn, the 12-mer library may have an advantage if the binding amino acids (which most of the times are less than 12) are spread out over the peptide sequence [38, 40]. Moreover, the 12-mer library may also increase the effective peptide diversity since each 12-mer peptide contains 7-mer peptides with different flanking sequences [39]. However, due to the increased length of the 12-mer peptides, it is possible that sequences with multiple weak binding are selected instead of sequences with few strong bindings [38]. Nevertheless, libraries of

both lengths have been successfully used for biopanning experiments, as also observed herein.

Comparing the phage display methodologies, for the 7-mer library, both BRASIL and conventional phage display resulted in the same consensus sequence. However, for the 12-mer library the consensus sequence differed between the two methodologies in the first two amino acids, perhaps because as explained above, these two amino acids may not be relevant for the strength of binding of the 12-mer peptide. However, these methods - conventional panning and BRASIL - do not display a significant difference. Since BRASIL is simple and faster, it is possible to say that this method is preferable for practical purposes [21].

Five peptides from the phage pools were selected to evaluate their binding affinities through different experimental assays. Both phage forming units (PFUs, Fig. 1) and ELISA (Fig. 2) assays suggest that all peptides exhibit specificity to the MDA-MB-231 cells with a lower binding capacity to MDA-MB-435 cells, as expected since this latter cell line was used in the counter-selection. These differences were of up to 80% (corresponding to 9 logs) in the PFUs assays (Fig. 1) and 4 logs in the ELISA assays (Fig. 2). These differences of peptide affinity between target and non-target cells are in accordance with values previously reported for phage display studies identifying peptides specific for other cancer cells, e.g., renal carcinoma A498 cells [41], breast cancer SKBR3 cells [42] and ovarian cancer HO8910 cells [43]. Among the peptides, 1.3(7/52) (PRWAVSP) and 6.2(9/17) (DTFNSFGRVRIE)

present the most promising results for targeted therapies of claudin-low cancer subtype due to the higher binding strengths and selectivity for the MDA-MB-231 cells.

Bioinformatics analysis (Table 2) indicate that peptides 1.3(7/52) (PRWAVSP) and 6.2(9/17) (DTFNSFGRVRIE) specifically target TIMP-1 and PAI1, respectively, with both biomarkers being related to breast cancer and to MDA-MB-231 cells. TIMP-1 is often overexpressed in many malignancies and is associated with increased histological grade, lymph-node and distant metastasis and decreased survival in breast cancer [44]. It is present and overexpressed in the MDA-MB-231 cells but also in Hs 578 T cells, with no expression on the remaining cell lines evaluated [45]. Although TIMP-1 has been considered a potential target for prognosis and therapeutic purposes, to our knowledge no specific peptide, antibody, aptamer or other molecule has been identified against this protein. Hence, the peptide herein identified represents a promising development for the establishment of prognosis tools, as well as targeted therapies. On the other hand, PAI1 is considered a prognostic marker due to a strong correlation with tumor aggressiveness and poor clinical outcome in breast cancer [46]. It is highly expressed in MDA-MB-231 cells, but also on MCF-10-2A and SK-BR-3 cells, exhibiting low levels of expression on Hs 578 T and MDA-MB-435 cells, which is somewhat in accordance with the preliminary flow cytometry results (Additional file 4: Figure S1) obtained for the phage pool from which the peptide 6.2(9/17) was obtained. Several aptamers have been developed that bind and inhibit PAI1, exhibiting potential therapeutic applications as anti-metastatic agents [47, 48]. The peptide here identified represents another alternative also with therapeutic and prognosis potential.

From the bioinformatics analysis, it is also interesting to note that peptides with different sequences (5.3 (19/45) sequence WWFFSFGRVRIE, and 6.2 (8/17) sequence TTEYSFGRTSTL) can exhibit affinities towards the same target (in this case the MDA-MB-231 biomarker  $\beta$ -actin, associated with breast cancer metastasis [37]), while similar peptides differing only in one amino acid have a different target (e.g., 5.3 (14/45), sequence WWFNSFGRVRIE targeting E-cadherin). This indicates that not only the amino acid sequence, but also the tri-dimensional conformation of the peptides influence the peptide interactions with the cells.

Finally, the peptide 5.3 (14/45) (WWFNSFGRVRIE) exhibited the highest affinity for a biomarker that is not present in the MDA-MB-231 cells (E-cadherin) [49], although showing a similar affinity for a MDA-MB-231 biomarker ( $\alpha$ -1-antichymotrypsin) (see Additional file 7: Table S2). This might explain the lowest selective affinity of this peptide for MDA-MB-231 cells as demonstrated in the binding studies.

All the peptides identified in this study were compared with reported breast cancer specific peptides and no similarities have been found (Figs. 3, 4, and 5), thus strongly supporting their novelty.

## Conclusions

In this work we identified new peptides specific for the MDA-MB-231 cells, which is representative of the claudin-low subtype of breast carcinomas, using phage display aided by bioinformatics tools. The methodology used together with the interpretation of phage display results (peptide sequences) being aided by bioinformatics approaches can be very useful to predict the potential cell targets (biomarkers) and to isolate peptides that are specific for the desired cells from those binding to other cancer subtypes. The selected peptides, PRWAVSP and DTFNSFGRVRIE, exhibit a strong binding to the MDA-MB-231 cells and a good specificity as demonstrated by the low binding to the MDA-MB-435 cells. Such peptides can be a valuable contribute towards future clinical applications through the development of more specific and targeted therapeutic solutions against the claudin-low breast cancer subtype.

## Additional files

**Additional file 1: Table S1.** Experimental setting of the phage display experiments, with BRASIL (B) and Conventional (C) methodologies, using the 7-mer and the 12-mer libraries. (DOCX 27 kb)

**Additional file 2: Table S3.** Breast cancer specific peptides reported in the literature, with amino acid sequence, sequence size, breast cancer stage, cell line targeted, cancer histological subtype and PubMed unique identifier number (PMID). (DOCX 58 kb)

**Additional file 3: Table S4.** Data collected for potential biomarkers of breast cancer cells, retrieved from Kyoto Encyclopedia of Genes and Genomes (KEGG), Uniprot, GenBank, and Protein Data Bank (PDB), with those present in MDA-MB-231 represented in bold. (DOCX 27 kb)

**Additional file 4: Figure S1.** Flow cytometry results, in terms of percentage of binding, of the phage pool from the last round of 12-mer conventional panning against normal breast cell line MCF-10-2A, breast cancer cell lines MDA-MB-231, SK-BR-3, Hs 578 T and MDA-MB-435 cell line [19]. Statistically significant (*P*) differences are represented by \*\*\*. (DOCX 43 kb)

**Additional file 5: Figure S2.** Immunofluorescence staining of MDA-MB-231 tissue sections. Sections were incubated with (A) wild-type M13KE phage particles and (B) M13KE phage particles of the phage pool from the last round of conventional phage display. Images were acquired with blue filter (1), green filter (2) and filter overlapping (3). Peptide affinity was detected using a primary anti-M13 antibody and a secondary goat anti-rabbit FITC conjugate antibody. Images were acquired using an inverted LEICA DMI 3000B (Leica Microsystems) with incorporated camera (Model DFC 450C). Scale bar of 10  $\mu$ m. (DOCX 1254 kb)

**Additional file 6: Figure S3.** Heatmap representation of the similarities between all peptides identified in this work with those previously reported. New12Br: 12-mer peptides obtained in this work using the BRASIL methodology; New12Conv: 12-mer peptides obtained in this work using the conventional methodology; Previous: 12-mer peptides reported in previous studies. Legend bar on the right represents the peptides of x-mer in different colours. (DOCX 398 kb)



**Additional file 7: Table S2.** Data from docking analysis with the selected 7-mer and 12-mer phage display peptides against breast cancer biomarkers retrieved from the literature: lowest energy weighted score (*E*), cluster members (*CM*) and type of interaction (*I*). Biomarkers present in MDA-MB-231 cells and best model scores are given in bold. (DOCX 41 kb)

### Abbreviations

BRASIL: Biopanning and rapid analysis of selective interactive ligands; BSA: Bovine serum albumin; CLR: Context likelihood of relatedness; DAPI: 4', 6-diamidino-2-phenylindole; DMEM: Dulbecco's Modified Eagle Medium; EDTA: Ethylenediaminetetraacetic acid; ELISA: Enzyme-linked immunosorbent assay; ER: Estrogen receptor; FBS: Fetal bovine serum; FITC: Fluorescein isothiocyanate; HER2: Human epidermal growth factor 2 receptor; IPATIMUP: Institute of Molecular Pathology and Immunology at the University of Porto; KEGG: Kyoto Encyclopedia of Genes and Genomes; PAI1: Plasminogen activator inhibitor 1 precursor; PBS: Phosphate buffered-saline; PBST: Phosphate buffered-saline with Tween-20; PDB: Protein data bank; PEG: Polyethylene glycol; PFU: Plaque forming unit; PR: Progesterone receptor; RLU: Relative light unit; ssDNA: Single-stranded DNA; TBS: Tris buffered saline; TBST: Tris buffered saline with Tween-20; TE: Tris-EDTA; TIMP-1: Metalloproteinase inhibitor 1

### Acknowledgments

This study was supported by the Portuguese Foundation for Science and Technology (FCT) and the European Community fund FEDER, through Program COMPETE, under the scope of the Projects FCOMP-01-0124-FEDER-021053 (PTDC/SAU-BMA/121028/2010), RECI/BBB-EBI/0179/2012 (FCOMP-01-0124-FEDER-027462), the strategic funding of UID/BIO/04469/2013 unit, and the Projects "BioHealth – Biotechnology and Bioengineering approaches to improve health quality", REF. NORTE-07-0124-FEDER-000027, and "BioInd – Biotechnology and Bioengineering for improved Industrial and Agro-Food processes", REF. NORTE-07-0124-FEDER-000028, co-funded by the Programa Operacional Regional do Norte (ON.2 – O Novo Norte), QREN, FEDER. Franklin L. Nóbrega acknowledges FCT for the grant SFRH/BD/86462/2012.

### Availability of data and materials

All relevant data has been presented in the main manuscript or in the Additional files 1, 2, 3, 4, 5, 6 and 7.

### Authors' contributions

FRK performed the phage display experiments, binding assays and bioinformatics analysis (docking studies); DF performed the flow cytometry and immunohistochemistry assays; IMM contributed to the phage display experiments; MSD performed the bioinformatics analysis (library analysis); JA, LDK and LRR designed and supervised the study; all authors contributed to the writing and revision of the manuscript. All authors read and approved the final manuscript.

### Competing interests

The authors declare they have no competing interests or other interests that might be perceived to influence the results and discussion herein reported.

### Ethics approval

The current study complies with all the ethics requirements and it does not involve human subjects (material or data).

### Author details

<sup>1</sup>Centre of Biological Engineering (CEB), University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. <sup>2</sup>Laboratory of Systems and Synthetic Biology, Wageningen University and Research Centre, Stippeneng 4, 6708WE Wageningen, The Netherlands.

Received: 26 April 2016 Accepted: 25 October 2016

Published online: 14 November 2016

### References

- GLOBOCAN. Cancer Incidence and Mortality Worldwide: IARC CancerBase No.11. 2012. [http://globocan.iarc.fr]. Accessed 9 Mar 2016.
- Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta*. 2010;1805(1):105.
- Weigel MT, Dowsett M. Current and emerging biomarkers in breast cancer: prognosis and prediction. *Endocr Relat Cancer*. 2010;17(4):R245–62.
- Sørbye T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*. 2001;98(19):10869–74.
- Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics*. 2006;7:96.
- Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*. 2010;12(5):R68.
- Mendes TFS, Kluskens LD, Rodrigues LR. Triple Negative Breast Cancer: Nanosolutions for a Big Challenge. *Advanced Science*. 2015;2(11):1–14.
- Holliday D, Speirs V. Choosing the right cell line for breast cancer research. *Breast Cancer Res*. 2011;13(4):215.
- Bertos NR, Park M. Breast cancer — one term, many entities? *J Clin Invest*. 2011;121(10):3789–96.
- Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA. Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clin Cancer Res*. 2007;13(15):4429–34.
- Li X, Mao C. Using Phage as a Platform to Select Cancer Cell-Targeting Peptides. *Methods Mol Biol*. 2014;1108:57–68.
- Fu B, Zhang Y, Long W, Zhang A, Zhang Y, An Y, Miao F, Nie F, Li M, He Y, Zhang J, Zhang G, Teng G. Identification and characterization of a novel phage display-derived peptide with affinity for human brain metastatic breast cancer. *Biotechnol Lett*. 2014;36:2291–301.
- Wölcke J, Weinhold E. A DNA-binding peptide from a phage display library. *Nucleosides, Nucleotides and Nucleic Acids*. 2001;20(4–7):1239–41.
- Shadidi M, Sioud M. Identification of novel carrier peptides for the specific delivery of therapeutics into cancer cells. *The FASEB Journal*. 2003;17(2):256–8.
- Larsen SA, Meldgaard T, Fridriksdottir AJ, Lykkemark S, Poulsen PC, Overgaard LF, Petersen HB, Petersen OW, Kristensen P. Selection of a breast cancer subpopulation-specific antibody using phage display on tissue sections. *Immunol Res*. 2015;62(3):263–72.
- Abbineni G, Modali S, Safiejko-Mroccka B, Petrenko VA, Mao C. Evolutionary selection of new breast cancer cell-targeting peptides and phages with the cell-targeting peptides fully displayed on the major coat and their effects on actin dynamics during cell internalization. *Mol Pharm*. 2010;7(5):1629–42.
- Nilsson F, Tarli L, Viti F, Neri D. The use of phage display for the development of tumour targeting agents. *Adv Drug Deliv Rev*. 2000;43(2–3):165–96.
- Molek P, Strukelj B, Bratkovic T. Peptide Phage Display as a Tool for Drug Discovery: Targeting Membrane Receptors. *Molecules*. 2011;16:857–87.
- NEB: Ph.D.TM Phage Display Libraries Manual.
- Rae JM, Creighton CJ, Meck JM, Haddad BR, Johnson MD. MDA-MB-435 cells are derived from M14 Melanoma cells—a loss for breast cancer, but a boon for melanoma research. *Breast Cancer Res Treat*. 2006;104(1):13–9.
- Giordano RJ, Cardo-Vila M, Lahdenranta J, Pasqualini R, Arap W. Biopanning and rapid analysis of selective interactive ligands. *Nat Med*. 2001;7(11):1249–53.
- Kropinski AM, Mazzocco A, Waddell TE, Lingohr E, Johnson RP: Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay. In: *Bacteriophages: Methods and Protocols*. New York: Edited by Clokie MR, Kropinski AM, vol. 501: Humana Press; 2009:69–76.
- IHC-Paraffin protocol (IHC-P). [http://www.abcam.com/ps/pdf/protocols/ihc\_p.pdf]. Accessed 21 Oct 2015.
- Bar H, Yacoby I, Benhar I. Killing cancer cells by targeted drug-carrying phage nanomedicines. *BMC Biotechnol*. 2008;8(1):1–14.
- Zhang B, Zhang Y, Wang J, Zhang Y, Chen J, Pan Y, Ren L, Hu Z, Zhao J, Liao M, et al. Screening and Identification of a Targeting Peptide to Hepatocarcinoma from a Phage Display Peptide Library. *Mol Med*. 2007;13(5–6):246–54.
- Pagès H, Aboyoun P, Gentleman R, DebRoy S. Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.42.0. 2016.
- Suarez-Diez M, Saccenti E. Effects of Sample Size and Dimensionality on the Performance of Four Algorithms for Inference of Association Networks in Metabonomics. *J Proteome Res*. 2015;14(12):5119–30.
- Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, Gardner TS. Large-Scale Mapping and Validation of Escherichia coli Transcriptional Regulation from a Compendium of Expression Profiles. *PLoS Biol*. 2007;5(1):e8.

29. Warnes G, Bolker B, Lumley T: gplots: Various R programming tools for plotting data. R package version 2.6.0.
30. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 2000;28(1):235–42.
31. Kelley LA, Sternberg MJE. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protocols.* 2009;4(3):363–71.
32. Kaur H, Garg A, Raghava G. PEPstr: A de novo method for tertiary structure prediction of small bioactive peptides. *Protein Pept Lett.* 2007;14:626–30.
33. Singh S, Singh H, Tuknait A, Chaudhary K, Singh B, Kumaran S, Raghava GPS. PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues. *Biol Direct.* 2015;10:73.
34. Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, Vajda S. How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics.* 2013;81(12):2159–66.
35. Comeau S, Gatchell D, Vajda S, Camacho C. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 2004;20(1):45–50.
36. Pettersen E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T. UCSF Chimera - a visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25(13):1605–12.
37. Morse DL, Carroll D, Weberg L, Borgstrom MC, Ranger-Moore J, Gillies RJ. Determining suitable internal standards for mRNA quantification of increasing cancer progression in human breast cells by real-time reverse transcriptase polymerase chain reaction. *Anal Biochem.* 2005;342(1):69–77.
38. Phage library choice. [<https://www.neb.com/faqs/2013/09/03/phage-library-choice>]. Accessed 10 Feb 2015.
39. Work L, Nicklin S, White S, Baker A: Use of phage display to identify novel peptides for targeted gene therapy. In: *Gene Therapy Methods*. Cambridge: Edited by Phillips M, vol. 346: Academic Press. 2002;157–17.
40. Lamichhane A. Identification of *Pseudomonas aeruginosa* ribosome assembly inhibitors. In: *Identification of drug targets and drug leads in Pseudomonas aeruginosa*. Michigan: ProQuest Dissertations Publishing; 2008.
41. Tu X, Zhuang J, Wang W, Zhao L, Zhao L, Zhao J, Deng C, Qiu S, Zhang Y. Screening and identification of a renal carcinoma specific peptide from a phage display peptide library. *J Exp Clin Cancer Res.* 2011;30(1):1–6.
42. Shukla GS, Krag DN. Phage display selection for cell-specific ligands: Development of a screening procedure suitable for small tumor specimens. *J Drug Target.* 2005;13(1):7–18.
43. Zhou C, Kang J, Wang X, Wei W, Jiang W. Phage display screening identifies a novel peptide to suppress ovarian cancer cells in vitro and in vivo in mouse models. *BMC Cancer.* 2015;15(1):1–12.
44. Bigelow R, Williams B, Carroll J, Daves L, Cardelli J. TIMP-1 overexpression promotes tumorigenesis of MDA-MB-231 breast cancer cells and alters expression of a subset of cancer promoting genes in vivo distinct from those observed in vitro. *Breast Cancer Res Treat.* 2008;117(1):31–44.
45. Lacroix M, Leclercq G: An updated view of cell lines as in vitro models for breast tumors. In: *Focus on Breast Cancer Research*. Edited by Yao A. New York: Nova Science Publishers; 2004.
46. Annecke K, Schmitt M, Euler U, Zerm M, Paepke D, Paepke S, von Minckwitz G, Thomssen C, Harbeck N. uPA and PAI-1 in breast cancer: review of their clinical utility and current validation in the prospective NNBC-3 trial. *Adv Clin Chem.* 2008;45:31–45.
47. Damare J, Brandal S, Fortenberry YM. Inhibition of PAI-1 Antiproteolytic Activity Against tPA by RNA Aptamers. *Nucleic Acid Therapeutics.* 2014;24(4):239–49.
48. Trelle MB, Dupont DM, Madsen JB, Andreassen PA, Jørgensen TJD. Dissecting the Effect of RNA Aptamer Binding on the Dynamics of Plasminogen Activator Inhibitor 1 Using Hydrogen/Deuterium Exchange Mass Spectrometry. *ACS Chem Biol.* 2014;9(1):174–82.
49. Lombaerts M, van Wezel T, Philippo K, Dierssen JWF, Zimmerman RME, Oosting J, van Eijk R, Eilers PH, van de Water B, Cornelisse CJ, et al. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. *Br J Cancer.* 2006;94(5):661–71.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

