CellPress

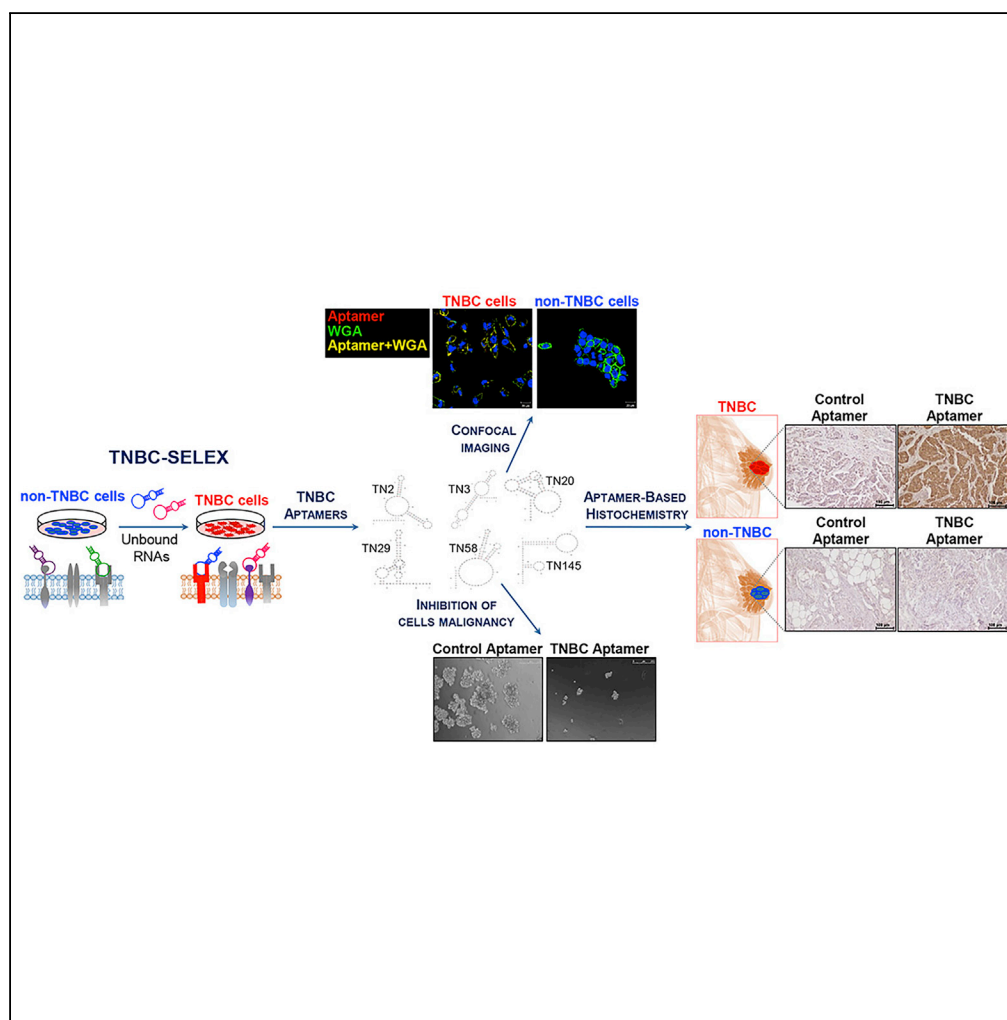**Article**

Novel Aptamers Selected on Living Cells for Specific Recognition of Triple-Negative Breast Cancer

Simona Camorani,
Ilaria Granata,
Francesca
Collina, ..., Monica
Fedele, Mario
Rosario
Guarracino, Laura
Cerchia

cerchia@unina.it

HIGHLIGHTS

Six 2′FPy-RNA aptamers were obtained by TNBC Cell-SELEX/NGS

They distinguish TNBC cells from non-malignant and non-TNBC breast cancer cells

They differentiate TNBC histological specimens by aptamer-based staining

They inhibit TNBC cell lines capacity of growing *in vitro* as mammospheres

Article

Novel Aptamers Selected on Living Cells for Specific Recognition of Triple-Negative Breast Cancer

Simona Camorani,[1,4] Ilaria Granata,[2,4] Francesca Collina,[3] Francesco Leonetti,[1] Monica Cantile,[3] Gerardo Botti,[3] Monica Fedele,[1] Mario Rosario Guarracino,[2] and Laura Cerchia[1,5,*]

SUMMARY

Triple-negative breast cancer (TNBC) is a high heterogeneous group of tumors with a distinctly aggressive nature and high rates of relapse. So far, the lack of any known targetable proteins has not allowed a specific anti-tumor treatment. Therefore, the identification of novel agents for specific TNBC targeting and treatment is desperately needed. Here, by integrating cell-SELEX (Systematic Evolution of Ligands by EXponential enrichment) for the specific recognition of TNBC cells with high-throughput sequencing technology, we identified a panel of 2′-fluoropyrimidine-RNA aptamers binding to TNBC cells and their cisplatin- and doxorubicin-resistant derivatives at low nanomolar affinity. These aptamers distinguish TNBC cells from both non-malignant and non-TNBC breast cancer cells and are able to differentiate TNBC histological specimens. Importantly, they inhibit TNBC cell capacity of growing *in vitro* as mammospheres, indicating they could also act as anti-tumor agents. Therefore, our newly identified aptamers are a valuable tool for selectively dealing with TNBC.

INTRODUCTION

Triple-negative breast cancer (TNBC) accounts for ∼15%–20% of breast cancers, and it annually affects approximately 170,000 patients worldwide (Newman et al., 2015). Compared with other breast cancers, TNBC generally arises at a younger age, is larger, of higher grade, and biologically more aggressive (Dent et al., 2007). Molecularly, TNBC is defined by the absence of estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2), excluding the possibility of using targeted therapies against these proteins. In the last two years, the first two targeted therapies have been approved by the Food and Drug Administration (FDA) for a very limited group of women with advanced TNBC: (1) PARP inhibitors, to treat patients with germline mutations in BRCA1/2 who have previously received chemotherapy (Beniey et al., 2019) and (2) the anti-programmed cell death ligand 1 (PD-L1) immunotherapy, used in combination with chemotherapy (Cortés et al., 2019). However, cytotoxic chemotherapy remains the mainstay treatment for the majority of patients (Gadi and Davidson, 2017), but, unfortunately, many patients with early stages of TNBC do not respond to treatment (Echeverria et al., 2019), and those who initially respond commonly develop metastatic and chemoresistant tumors showing a median overall survival of 13–18 months (Garrido-Castro et al., 2019).

One major obstacle to TNBC treatment is its high degree of heterogeneity, being characterized by different subtypes with unique gene expression profiles and distinct clinical behavior, including response to therapies (Lehmann et al., 2011, 2016; Burstein et al., 2015).

Therefore, it is fundamental to develop novel targeted treatment approaches able to control each individual TNBC subtype, diminish toxicity, and delay the onset of patient resistance to chemotherapy. Highly selective nucleic acids aptamers are emerging as the most promising candidates for targeting tumor cells. They have risen increasing attention for cancer diagnosis and therapy, as a result of their low molecular weight, low/no immunogenicity, and versatility to manipulation for improved stability and targeting efficacy (Zhou and Rossi, 2017; Camorani et al., 2018a).

Aptamers are isolated by the Systematic Evolution of Ligands by EXponential enrichment (SELEX) process (Zhou and Rossi, 2017; Camorani et al., 2018a, 2018b; Keefe et al., 2010).

[1]Institute of Experimental Endocrinology and Oncology "G. Salvatore" (IEOS), National Research Council (CNR), Naples 80131, Italy

[2]Computational and Data Science Laboratory, High Performance Computing and Networking Institute, National Research Council (CNR), Naples 80131, Italy

[3]Pathology Unit, Istituto Nazionale Tumori-IRCCS-Fondazione G. Pascale, Naples 80131, Italy

[4]These authors contributed equally

[5]Lead Contact

*Correspondence: cerchia@unina.it

https://doi.org/10.1016/j.isci. 2020.100979

Although SELEX is typically carried out using purified target proteins, whole living cells are also employable as selection target (cell-SELEX) allowing to obtain aptamers against cell surface proteins in their native conformation. By cell-SELEX, target-binding aptamers are isolated from large libraries of randomized oligonucleotides over several rounds of selection against entire cells in culture, even without the prior knowledge of molecules present at the cell surface. Indeed, by altering target cell selection and off-target cell elimination, aptamers that may identify subtle differences existing among different cells, also belonging to the same tumor type, which drive important tumor cell behaviors, including resistance to therapy, tumorigenicity, stemness and capacity to metastasize, can be isolated (Tang et al., 2007; Camorani et al., 2014; Sefah et al., 2009; Chen et al., 2008; Shangguan et al., 2006; Esposito et al., 2011).

Also, cell-SELEX approaches for targeting of breast cancer cells have been reported to select aptamers for recognition of metastatic cancer cells (Li et al., 2014) and HER2-positive breast cancer cells (Liu et al., 2018). Furthermore, a SELEX method using genetic alphabet expansion allowed to select aptamers on breast cancer cells with some biological activities, such as internalization within the target cells and anti-proliferative activity (Futami et al., 2019).

Here, by integrating a cell-SELEX method for the specific recognition of TNBC cells with high-throughput sequencing technology, we identified a panel of nuclease-resistant RNA aptamers that bind to target cells, and their cisplatin- and doxorubicin-resistant derivatives, with an equilibrium dissociation constant (K_d) in the low nanomolar range, and distinguish TNBC cells from both normal breast cells and non-TNBC breast cancer cells. Importantly, the selected aptamers differentiate TNBC human samples when used for histochemical tissue staining, allowing to cluster them in molecular subtypes. Further, these aptamers interfere with the TNBC cells' capacity of growing *in vitro* as mammospheres, a feature associated with the malignant phenotype, thus indicating they could be employed as important anti-tumor agents.

RESULTS

TNBC Cell-SELEX

We reasoned that TNBC carries characteristic cell surface signatures that distinguish itself from the other breast cancer subtypes, thus allowing to identify, by a differential cell-SELEX screening, aptamers able to specifically bind to TNBC, eventually discriminating among different TNBC subtypes, over non-TNBC breast cancer expressing HER2, PR, and ER. To reach this goal, we started from a library of nuclease-resistant 2'fluoro-pyrimidines (2'F-Py) RNAs and performed a total of 14 consecutive rounds of positive selection on human MDA-MB-231 cells (ER⁻, PR⁻, HER2⁻), with increasing selection stringency (Figure 1A, Table S1). MDA-MB-231 cells represent an established model for aggressive TNBC cells and largely recapitulate the gene expression patterns and mutations found *in vivo* (Lehmann et al., 2011; Nguyen et al., 2014). These cells are characterized by the expression of epithelial-mesenchymal transition markers, highly malignant and invasive phenotype, and a strong tendency to form vasculogenic mimicry (Camorani et al., 2017a; Blick et al., 2008; Betapudi et al., 2006; Han et al., 2008; D'Ippolito et al., 2016; Camorani et al., 2017b; Camorani et al., 2018c). Starting from the second SELEX round (Figure 1A, Table S1), the positive selection was preceded by counterselection steps against the well-characterized BT-474 epithelial breast cancer cell line (ER⁺, PR⁺, HER2 over-expression) (Nowsheen et al., 2012; Dai et al., 2017; Pasleau et al., 1993) to deplete RNA molecules capable of recognizing non-TNBC cells. In order to avoid loss of specific sequences, the counterselection was not included in the first round. Additionally, because MDA-MB-231 cells express abundant levels of epidermal growth factor receptor (EGFR), a receptor frequently overexpressed in TNBC (Nair et al., 2018) and already used as target for aptamer recognition by our and other groups (Camorani et al., 2018a), we chose to include into the selection cycle, starting from the fifth SELEX round, a second counterselection against EGFR-overexpressing epidermoid carcinoma A431 cells (Ullrich et al., 1984) (Figure 1A, Table S1), to avoid that sequences against EGFR could dominate the selection. Importantly, an increase of MDA-MB-231 target cells' recognition was observed as the selection progressed through additional rounds, whereas the enriched libraries did not interact with non-target BT-474 cells (Figure 1B), thus indicating that the counterselection strategy allowed to foster specificity toward TNBC cells.

Identification of TNBC-Specific Candidate Aptamers

We implemented the SELEX approach by using next-generation sequencing (NGS) and bioinformatics analysis to monitor the evolution of the selection procedure along with the screening cycles and to
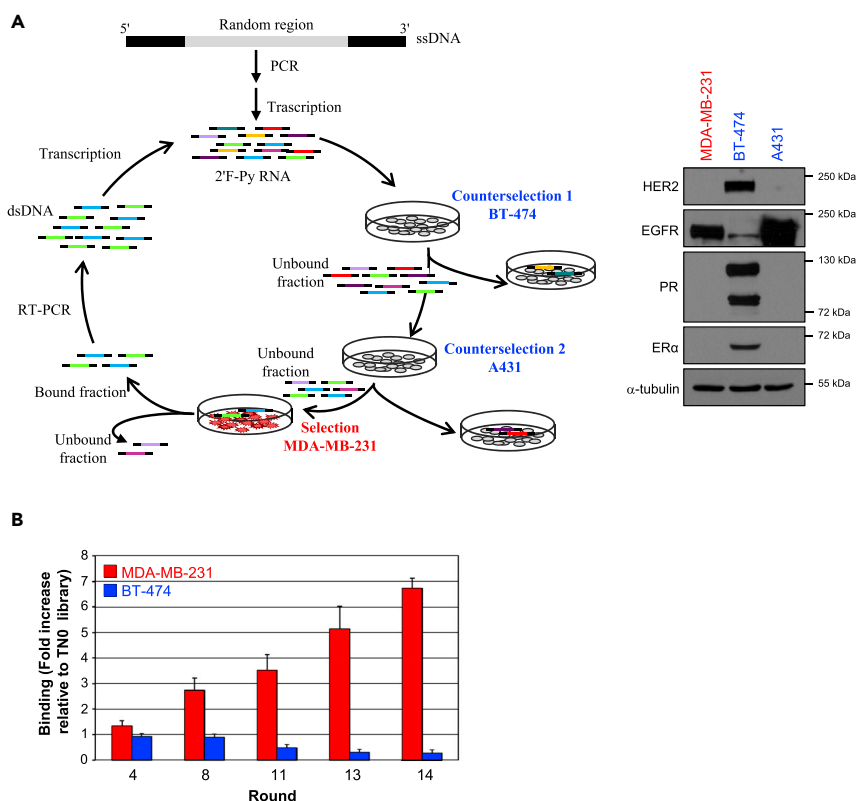
Figure 1. TNBC Cell-SELEX

(A) *Left*: schematic protocol for the selection of TNBC-specific aptamers. A pool of 2'F-Py RNAs (contained a 40-mer random sequence region flanked by two constant sequence regions at 5' and 3' of 21-mer and 23-mer, respectively) was incubated with BT-474 cells (starting from second round) for a first counterselection step. Unbound sequences were recovered and incubated with MDA-MB-231 cells for the selection step (from second to fourth round) or with A431 cells (from fifth round up to the end of SELEX) for a second counterselection step. Unbound sequences from the second counterselection were recovered and incubated with MDA-MB-231 cells for the selection step. The counterselection was not included in the first round. Unbound sequences were discarded by several washings, and bound sequences were recovered by phenol extraction. Sequences enriched by the selection step were amplified by RT-PCR and *in vitro* transcribed before a new cycle of selection. *Right*: characterization of cells used for the SELEX scheme. Lysates from MDA-MB-231, BT-474, and A431 cells were immunoblotted with anti-HER2, anti-EGFR, anti-PR, and anti-ERα antibodies, as indicated. α-tubulin was used as an internal control.

(B) Assessment of the selection progress. The indicated selected pools or the TN0 starting library (200-nM final concentration) were incubated for 15 min at 37°C with MDA-MB-231 and BT-474 cells and binding was detected by RT-qPCR. The results are expressed relative to the background binding detected with the TN0 library. Bars depict mean ± SD of three independent experiments.

support the choice of which aptamers were worth testing in binding assays. Thus, ten selection rounds (0, 3, 5, 7, 9, 10, 11, 12, 13, and 14) were analyzed by Illumina NGS, and data were compared with those obtained by classical cloning of the last selected pool. Figure 2A shows the workflow used in the present study for filtering, clustering, and identifying candidate aptamers to test in cell-binding assays. After quality filtering and primers trimming of the paired ends merged reads, we obtained reads quantity ranging from 2633075 to 7803708. Before proceeding with length filtering, we examined the overall variable region nucleotide length of unique reads along all the rounds (Figure 2B). As expected, the mostly frequent variable region length was 40 nucleotides for all the rounds, with low frequencies also of 38, 39, 41, and 42 nucleotides length. Based on these observations, we included sequences with variable region lengths ranging from 38 to 42 nucleotides in our subsequent analyses. As shown in Figure 2C, the enrichment of the libraries shows an increasing trend throughout the rounds, with the most pronounced changes in percentage enrichment starting from selection round 7, thus indicating the convergence of potential aptamers in the resulting pools.
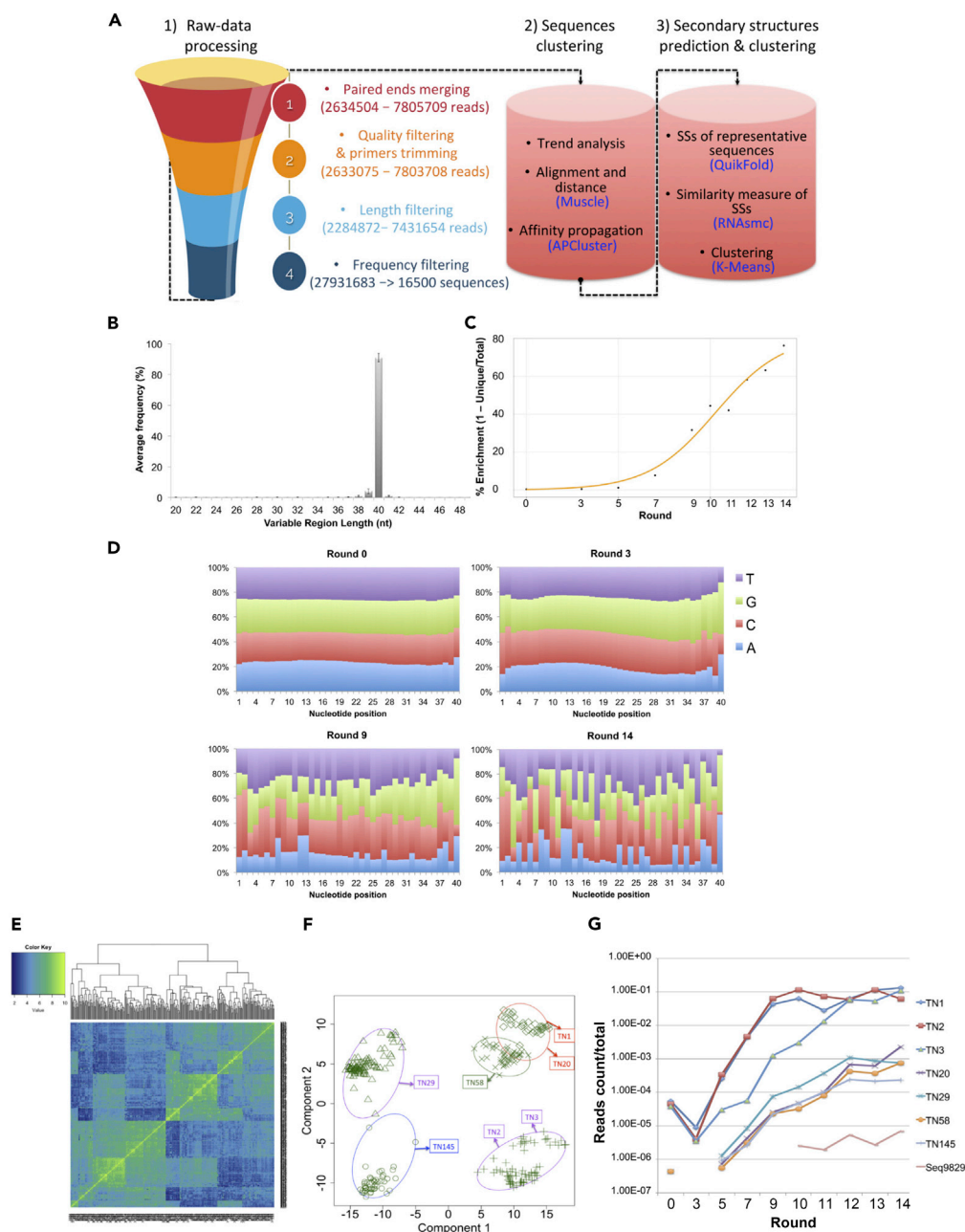
Figure 2. High-Throughput Sequencing and Data Processing

(A) A schematic view of the pipeline used in this study is shown. Briefly, the processing of the raw fastq files involved several steps of filtering after the merging of paired ends, such as quality, length, and frequency filtering. The ranges of the reads number obtained after each step are shown. Finally, 16,500 unique sequences got through the downstream analyses: sequences and secondary structure clustering. The methods used are colored in blue.

(B) The plot shows the average frequency (mean $\pm$ SD) of the variable region length of the RNA sequences in all the rounds. As expected, the most frequent length is 40 nt. Few sequences show a length of 38, 39, 41, and 42 nt.

(C) The plot shows the percentage enrichment at each round (black circle), calculated by: % Enrichment = 1-Unique/Total. A sigmoidal curve fit was added to the plot.

(D) Nucleotide frequency distribution for four rounds (0, 3, 9, 14).

(E) The heatmap shows the matrix of secondary structure similarities computed by RNAsmc R package. The length of the branch and the colors of the heatmap correspond to degree of similarity between the structures predicted.

(F) K-Means clustering of the predicted secondary structures, considering five clusters according to the connectivity parameter. The different clusters are highlighted by the different shape of the points, whereas the color of the ellipses

Figure 2. *Continued*

indicates the density of the cluster. The two components explain 66.27% of the point variability. In the boxes are highlighted the candidate sequences, and the arrows indicate their belonging to the specific clusters.

(G) Trend of the candidate sequences along the selection rounds. The counts are normalized over the total number of sequences at each round. The y scale is log transformed and the points equal to 0 are not displayed. The candidate aptamers TN1, TN2, and TN3 cover the first three positions according to the slope ranking. They are already present at the round 0 and their concentrations continue to increase till round 14 (TN3) or till round 10 and then show a stable concentration till round 14 (TN1 and TN2). At the round 3 all three of them show a concentration decrease. The candidate aptamers TN20, TN29, TN58, and TN145 start to appear at round 5 (except the TN29 and TN58 that have a weak concentration at round 0) and show an increasing trend till round 14. Sequence 9829 (out of the top 5,000 sequences) is reported as an example of a sequence that did not show up as an enriched sequence in our analysis.

The Figure 2D shows that at the beginning of the selection experiment (round 0) the frequency of the four nucleotides is equally distributed, whereas in the subsequent rounds the enrichment of particular sequences is highlighted by the change of this distribution along the sequence positions.

In order to keep only the enriched sequences and ranking them based on the increasing trend along with the rounds, we applied the slope formula to the sequences showing counts equal or greater than 20 in almost one of the last three rounds, thus obtaining 16,500 sequences, from almost 28 million obtained by combining the unique sequences of each round. Then, based on the slope scores, the top 5,000 sequences were selected, which, based on their identity-based distances from alignment, gave rise to 293 clusters (Figure S1). Although aptamers are typically categorized just by the analysis of sequence similarity, we considered the undoubted role of the secondary structures in defining their functionality and added a step of clustering based on structural similarity. From each alignment-based cluster, we selected the most enriched sequences and predicted their secondary structure using the Quikfold software. The heatmap obtained from the 293 secondary structure similarity is shown in Figure 2E. In order to statistically determine the most appropriate number of clusters to our data, we applied the K-Means clustering algorithm, taking in consideration internal measures such as Connectivity, Dunn Index (DI), and Silhouette Width (Handl et al., 2005) (Table S2). The connectivity indicates the degree of connectedness of the clusters. It has a value between 0 and infinity and should be minimized. The Silhouette Width is the average of each observation's Silhouette value. Well-clustered observations show values near 1, and poorly clustered observations show values near -1. The DI is the ratio between the minimum inter-cluster distance and the maximum cluster size. It has a value between 0 and infinity and a higher DI implies better clustering (Handl et al., 2005) (Table S2). We decided to use the best three k settings (5, 6, 7) and found that the most stable discrimination occurred setting five clusters (Figure 2F). The best candidate RNAs were chosen by the following criteria: sequences with an optimum enrichment score that were also cloned from the 14th pool and covering all the different clusters. Accordingly, we extracted seven sequences for testing by cell-binding assays: TN1, TN2, TN3, TN20, TN29, TN58, and TN145 (Figure 2F). The enrichment of the selected sequences is shown in Figure 2G. Accordingly, a sequence out of the top 5,000 sequences, sequence 9829, did not show up as an enriched sequence in our analysis. Noteworthily, four sequences were particular abundant (>50 counts) in the initial, unselected RNA library but, among them, TN1, TN2, and TN3 survived to our filtering criteria and were thus taken into account for further characterization. The theoretical secondary structures of the candidate aptamers are shown in Figure S2.

Characterization of Aptamers' Binding Affinity to MDA-MB-231 Target Cells

The seven candidate aptamers were tested in binding assays to evaluate their ability to specifically bind to MDA-MB-231 target cells with respect to the cells used for the counterselections (BT-474 and A431). To this aim, three different methods were harnessed to detect cell binding of the chosen sequences: reverse transcription-quantitative polymerase chain reaction (RT-qPCR)-based, streptavidin-biotin-based (colorimetric), and flow cytometric assays. The last two approaches, involving the use of 3′-biotinylated and Alexa 647-labeled aptamers, respectively, were applied for binding affinity (K_d values) calculation. Results from RT-qPCR-based assay are shown in Figure S3. The curves of binding to target cells for each aptamer, by both colorimetric and flow cytometric assays, are shown in Figures S4 and S5, respectively, and their K_d for each cell line are listed in Table 1. As shown, except for aptamer TN1 that did not bind to either the target cell line (Figures S3–S5) or the cells of the counterselections (Figures S3 and S4), the other six aptamers bound with high affinity to MDA-MB-231 target cells (Figures S4 and S5) and comparable K_d values, in the low nanomolar range (from 9.84 $\pm$ 1.63 nM to 59.12 $\pm$ 12.93 nM and from 10.77 $\pm$ 2.41 nM to 73.02 $\pm$ 12.08 nM by colorimetric and flow cytometric assays, respectively), were obtained

| Aptamer	Sequence Random Region	MDA-MB-231		BT-474	A431
		K_d, nM Biotin-RNA	K_d, nM Alexa-RNA	K_d, nM Biotin-RNA	K_d, nM Biotin-RNA
TN1	GUGUCAGACGUAAUGUGUCGC ACAUCUUGUCAUGCUACUG	ND	ND	ND	ND
TN2	AAGGCCGACGUAAUGUGUC GGUCGUUACGCGUCGUGCACG	59.12 ± 12.93	73.02 ± 12.08	ND	ND
TN3	CCGAUCUCACGCGCACCUU CUCUUCAGCGCGCGACUGGCA	21.91 ± 3.08	28.93 ± 9.72	ND	62.98 ± 23.61
TN20	CGAUGCGCACCGAUCUCUC UUCUGCACGUCCUUCGGCACA	24.30 ± 4.99	23.92 ± 3.59	ND	ND
TN29	CCUGCCCCAACCAUCGCUU CCUCGACGCGCGUUGUCGGCA	9.84 ± 1.63	10.77 ± 2.41	ND	ND
TN58	GCAACGUUGUGGUCCCGUUUGC ACUUUGUUUACGCGCGCA	17.17 ± 2.87	26.09 ± 3.20	ND	ND
TN145	CCUCAGCGCGCAACUUCCCUCCGU UCCCUGCCACGCGUCA	26.88 ± 3.68	37.36 ± 3.50	ND	ND
Seq9829	GUCUUUUGCUCGUCCUUCCACCG CUCACGUCCUUCCGACA	ND	ND	–	–

Table 1. Dissociation Constant Values for the Candidate Aptamers to Target Cells (MDA-MB-231) and Cells of Counterselections (BT-474 and A431)
The random sequence of each aptamer is reported; for simplicity fixed-primer sequences at 5′ and 3′ are not shown. Kd values were calculated by streptavidin-biotin colorimetric and flow cytometric assays. The mean values of three independent measurements are presented. ND, not detectable; "-", not determined. As expected, sequence 9829, which was filtered out according to lack of enrichment along the cell-SELEX rounds, did not bind to target cells by both the binding assays.

for each sequence by the two applied methods (Table 1). Additionally, the sequence 9829, which we chose among those filtered out according to frequency before the slope calculation, gave no binding to MDA-MB-231 target cells (Figure S3, Table 1), therefore confirming the validity of our approach for the identification of the TNBC-specific candidate aptamers. Importantly, none of the six aptamers bound to BT-474 cells (Figures S3 and S4), demonstrating the successful use of these cells for the counterselection. Further, only TN3 bound to the cell line of the second counterselection, A431, even if at a much lesser extent than MDA-MB-231 cells, indicating that the counterselection step did not completely abrogate its binding to these cells (Figure S4). Noteworthily, TN3 is among the top sequences in the initial library and started to enrich very early on, before the introduction of the counterselection against A431 cells (Figure 2G), possibly explaining why this sequence shows some binding to A431. However, TN3 did not bind to mouse NIH3T3 fibroblasts engineered to overexpress human EGFR (Camorani et al., 2015), thus excluding EGFR as its target (data not shown).

To confirm the validity of the above results, we compared the binding of one of the six candidate aptamers (TN20) with that of the non-binder TN1 aptamer to MDA-MB-231 cells by confocal imaging. As shown, TN20 aptamer specifically co-localized at the membrane level with the Wheat Germ Agglutinin (WGA) surface marker on MDA-MB-231 cells after 10-min incubation at RT (Figure 3A). Conversely, very little to no signal was observed with TN1 sequence (Figure 3A). As a further control experiment, both TN20 and TN1 aptamers were incubated with BT-474 cells, resulting in no binding (Figure 3A). It is well known that many aptamers internalize upon binding to cell surface targets, making them useful targeted delivery vehicles for therapeutic secondary reagents. Preliminary experiments, carried out so far only with TN20, show that it internalizes within the target cells and accumulates in compartments positive for LysoTracker, which stains acidic compartments in live cells, following 30-min incubation at 37°C (Figure 3B).

To test preliminarily whether the targets of the six aptamers are cell surface proteins, we treated MDA-MB-231 target cells with proteinase K and then analyzed cell binding with Alexa-labeled aptamers using flow cytometry. Importantly, the treatment prevented the binding of the six aptamers to MDA-MB-231 cells (Figure 3C), indicating that the target molecules are most likely cytomembrane proteins.
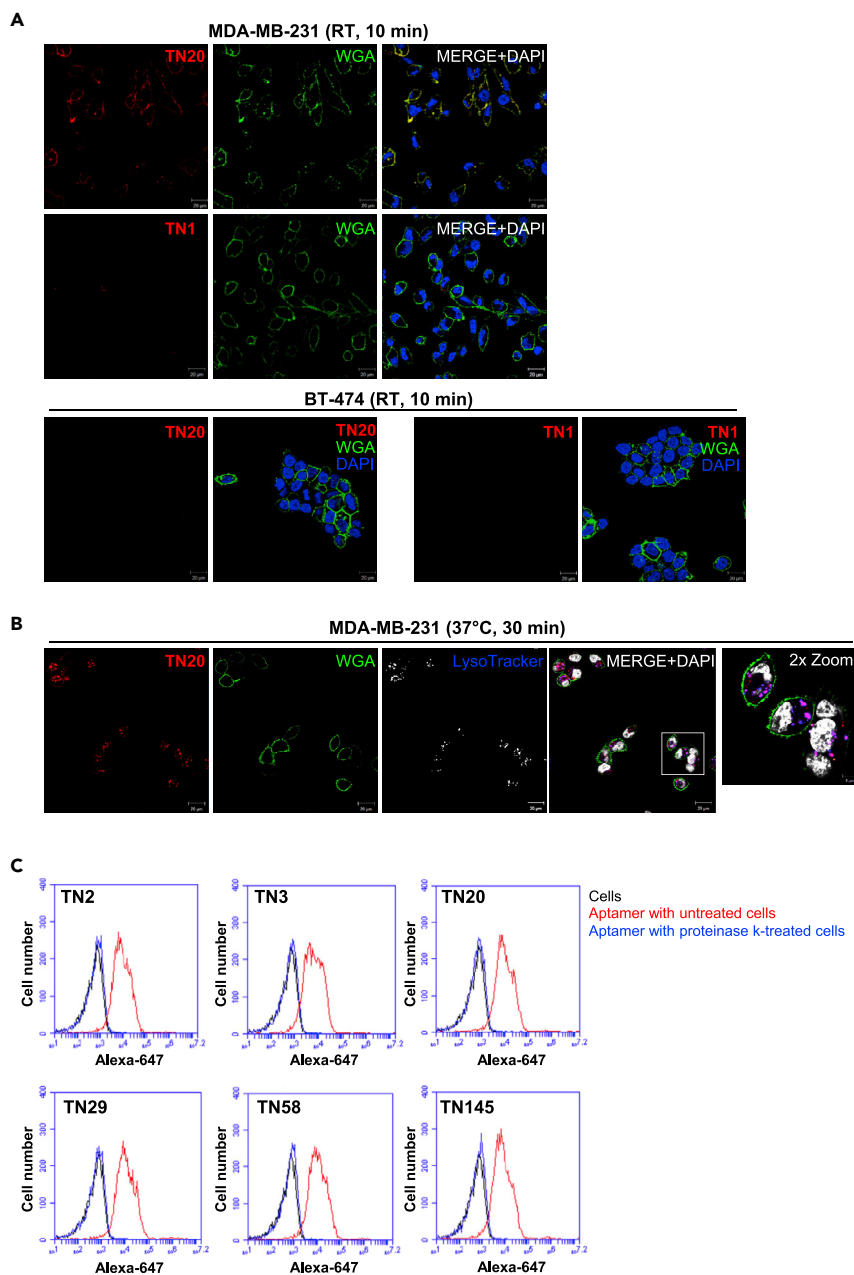
Figure 3. Target-Type Analysis of TNBC Aptamers

(A) Following 10-min incubation at RT with 500 nM Alexa 647-labeled TN20 or TN1 aptamers, MDA-MB-231 and BT-474 cells were fixed and incubated with WGA, visualized by confocal microscopy, and photographed. Alexa-labeled aptamer, WGA, and nuclei are visualized in red, green, and blue, respectively. Co-localization results appear yellow in the merged images. All digital images were captured at the same setting to allow direct comparison of staining patterns.

(B) Internalization of TN20 into MDA-MB-231 cells. Cells were incubated for 30 min at 37°C in the presence of 500 nM Alexa 647-labeled TN20, after the incubation with LysoTracker to detect acidic organelles inside the cells, visualized by confocal microscopy, and photographed. Alexa-labeled aptamer, WGA, LysoTracker, and nuclei are visualized in red, green, blue, and gray, respectively. Co-localization results appear purple in the merged images.

(A and B) Magnification 63×, 0.5× digital zoom, scale bar = 20 μm. Inset: 2× digital zoom, scale bar = 5 μm.

(C) Binding of TNBC aptamers to proteinase K-treated MDA-MB-231 (blue) and untreated (red) cells. The concentration of the aptamers in the binding buffer was 200 nM.

(A–C) At least three independent experiments were performed.

Aptamers' Binding to Chemoresistant MDA-MB-231 Cells

Chemoresistance of TNBC is a major reason for treatment failures and contributes to metastasis, which further decreases the prognosis of patients (Nabholtz et al., 2014). Accumulating evidences indicate that certain cell surface proteins display differential expression in chemoresistant cancer cells compared with chemosensitive cells (O'Reilly et al., 2015; Ziegler et al., 2014). By treating MDA-MB-231 cells with either doxorubicin or cisplatin, both used in clinic for TNBC, over a period of about five months (see Methods), we generated MDA-MB-231 derivative cells (MDA-MB-231/dox and MDA-MB-231/cis) with increased resistance to doxorubicin and cisplatin compared with the parental cells (Figure S6A). Consistent with the acquired resistance, these cells showed a significant enhanced expression of platelet-derived growth factor receptor β (PDGFRβ), integrin αv, and PD-L1 (Figure S6B), three cell surface proteins related to chemoresistance. We then tested whether our novel TNBC aptamers preserve the ability to bind to chemoresistant cells. Importantly, binding curves obtained by flow cytometry analysis show that all the selected aptamers are capable of targeting both MDA-MB-231/dox (Figure S7A) and MDA-MB-231/cis (Figure S7B) cells with K_d values in the low nanomolar range (Figure 4A), thus suggesting that changes in protein expression under therapeutic selection pressure may not limit the effectiveness of targeting. Interestingly, as shown in Figure 4B, the TN58 aptamer displayed a higher affinity for chemoresistant cells with respect to the parental counterpart, thus suggesting that its target is a protein enriched in chemoresistant cells.

Overall, these findings imply that, based on protein expression, personalizing aptamer-based tumor treatments might become possible.
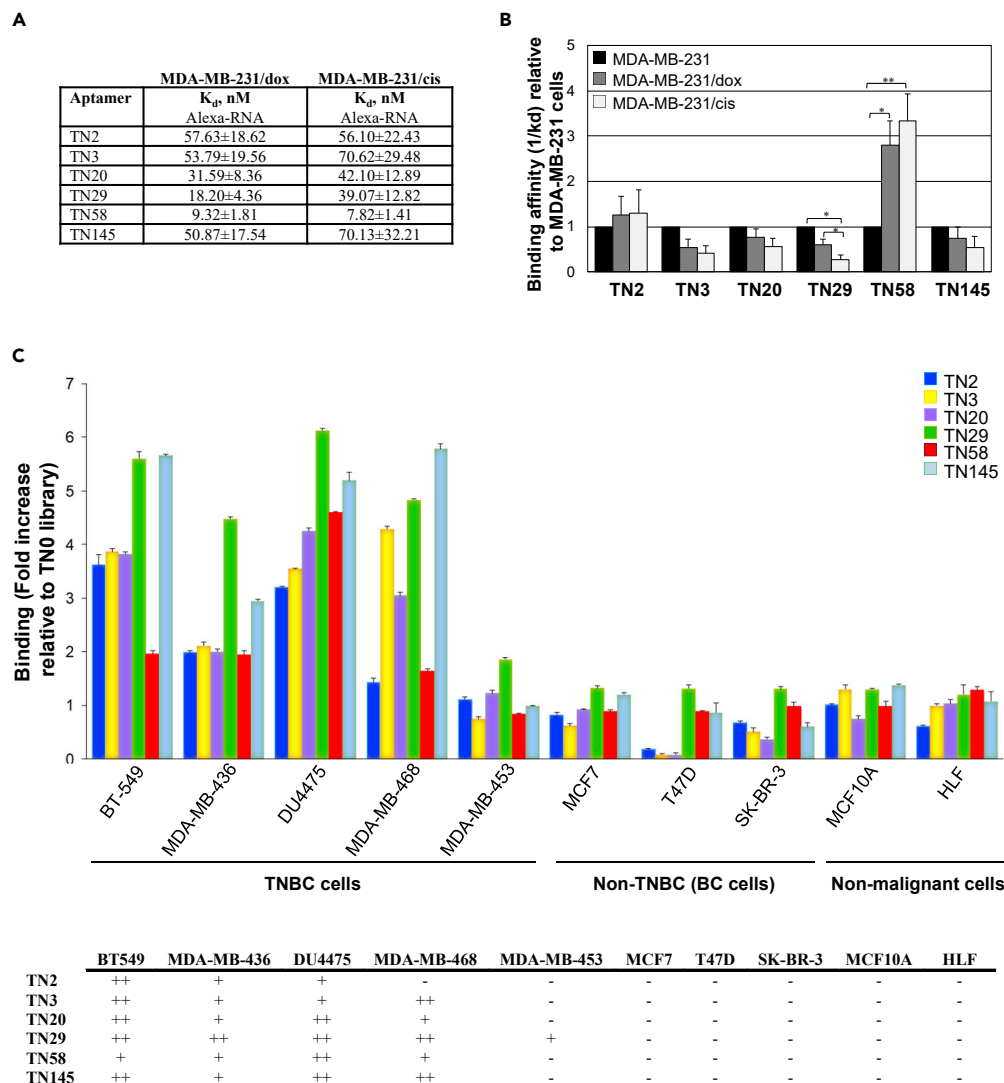
Selected Aptamers Specifically Bind to Different TNBC Cell Lines

Next, in addition to the cells used in the selection (MDA-MB-231) and counterselection (BT-474), binding analyses were extended to other human breast cancer cell lines, TNBC and non-TNBC, to validate binding specificity for the desired cell type. All these cell lines have been extensively characterized and show gene expression profiles similar to the respective tumor subtypes (Lehmann et al., 2011, 2016; Ross and Perou, 2001; Sandberg and Ernberg, 2005). Importantly, tested at the same concentration, none of the selected aptamers bound to non-TNBC breast cancer cell lines representative of luminal A (MCF-7 and T47D, ER^+, PR^+, $HER2^-$) and HER2-positive (SK-BR-3, ER^-, PR^-, $HER2^+$) molecular categories (Figure 4C). Conversely, they bound to most TNBC cells tested (BT-549, MDA-MB-436, DU4475, MDA-MB-468) (Lehmann et al., 2011), except only for TN2 aptamer that gave no appreciable binding on the latter cell line. Noteworthy, MDA-MB-453 cells, which even if classified among TNBC cell lines have been proven to express HER2 (Garnett et al., 2012), were moderately or even not recognized by the six aptamers. Importantly, none of the selected aptamers bound to normal human lung fibroblasts and the human mammary epithelial MCF10A cells, chosen as non-malignant breast cellular model (Figure 4C).

These results clearly demonstrated that the aptamers selected by our cell-SELEX procedure bind to TNBC cells and discriminate them from normal cells or breast cancer cells expressing HER2 and/or ER and PR, suggesting that they may interact with unique surface-binding entities expressed on TNBC cells.

Aptamer-Based Histochemical Staining of Human TNBC Samples

In order to translate our results to the clinics, we evaluated the capacity of the six identified aptamers to be a new tool for the characterization of TNBC tissues. To this aim, we tested *ex vivo* the selected aptamers by setting up the best binding conditions on deparaffinized tissue sections, of biotinylated aptamers used for histochemical staining of 18 human TNBC samples included in a TMA. The main clinicopathological data of the cases are set out in Table S3. As shown in Figure 5, each aptamer showed a distinct pattern of binding on different tumors, which we scored as absent, low, moderate, or high, based on both staining intensity and cell percentage, thus highlighting the TNBC heterogenicity (Figure 5A). Interestingly, it is possible clustering TNBC samples according to a signature of aptamer staining (Figure 5B). For instance, samples #6, #10, and #14 (marked in red) on one side, and #16 and #18 (marked in green) on another side, appear to belong to two different clusters according to the same pattern of binding by all the six selected aptamers (Figure 5B). Further, the staining of four aptamers (TN2, TN3, TN20, and TN145) is common to samples #2 and #3 (marked in purple), and the other aptamers (TN29 and TN58) stained the two tumors at high or moderate intensity (score), respectively, suggesting they may represent another TNBC cluster (Figure 5B). For each aptamer the differences observed in the staining extent likely reflect the relative concentrations of the same target molecules in the different tumors. Importantly, when applied to breast

A

Aptamer	MDA-MB-231/dox Kd, nM Alexa-RNA	MDA-MB-231/cis Kd, nM Alexa-RNA
TN2	57.63±18.62	56.10±22.43
TN3	53.79±19.56	70.62±29.48
TN20	31.59±8.36	42.10±12.89
TN29	18.20±4.36	39.07±12.82
TN58	9.32±1.81	7.82±1.41
TN145	50.87±17.54	70.13±32.21

B

C

	BT549	MDA-MB-436	DU4475	MDA-MB-468	MDA-MB-453	MCF7	T47D	SK-BR-3	MCF10A	HLF
TN2	++	+	+	-	-	-	-	-	-	-
TN3	++	+	+	++	-	-	-	-	-	-
TN20	++	+	++	+	-	-	-	-	-	-
TN29	++	++	++	++	+	-	-	-	-	-
TN58	+	+	++	+	-	-	-	-	-	-
TN145	++	+	++	++	-	-	-	-	-	-

Figure 4. Targeting Specificity of TNBC Aptamers

(A) TNBC aptamers target chemoresistant MDA-MB-231/dox and MDA-MB-231/cis cells with nanomolar K_d values.

(B) Binding affinity ($1/K_d$) of each TNBC aptamer to MDA-MB-231/dox and MDA-MB-231/cis cells is expressed relative to the corresponding binding affinity to MDA-MB-231 cells. Note that the K_d values obtained on parental and chemoresistant cells by using the flow cytometric assay were taken into account for this analysis.

(C) Binding of the indicated aptamers (200-nM final concentration, colorimetric assay) on TNBC cells, covering different subtypes (Lehmann et al., 2011), non-TNBC breast cancer cells, human lung fibroblasts (HLF), and MCF10A cells, chosen as non-malignant breast cellular model. The results are expressed relative to the background binding detected with the TN0 negative control. The binding capacity of the aptamers to the cells is reported as "++" for high binding (more than 3.5-fold), "+" middle binding (between 1.5- and -3.5-fold), and "−"no binding (less than 1.5-fold).

(B and C) Bars depict mean ± SD of three independent experiments. (B) **$p < 0.01$; *$p < 0.05$ (one-way ANOVA followed by Tukey's multiple comparison test).

cancer samples expressing ER, PR, and HER2, the aptamers resulted in negative or low staining (Figures 5B and 5C). Also, accordingly to binding analyses on human cell lines, none of the six aptamers bound to normal samples dissected adjacent to the tumors (Figures 5B and 5C).

Overall, we validated the use of these aptamers in human histologic sections and propose their combined application as a potential tool to allow simple sub-classification of TNBC subtypes.
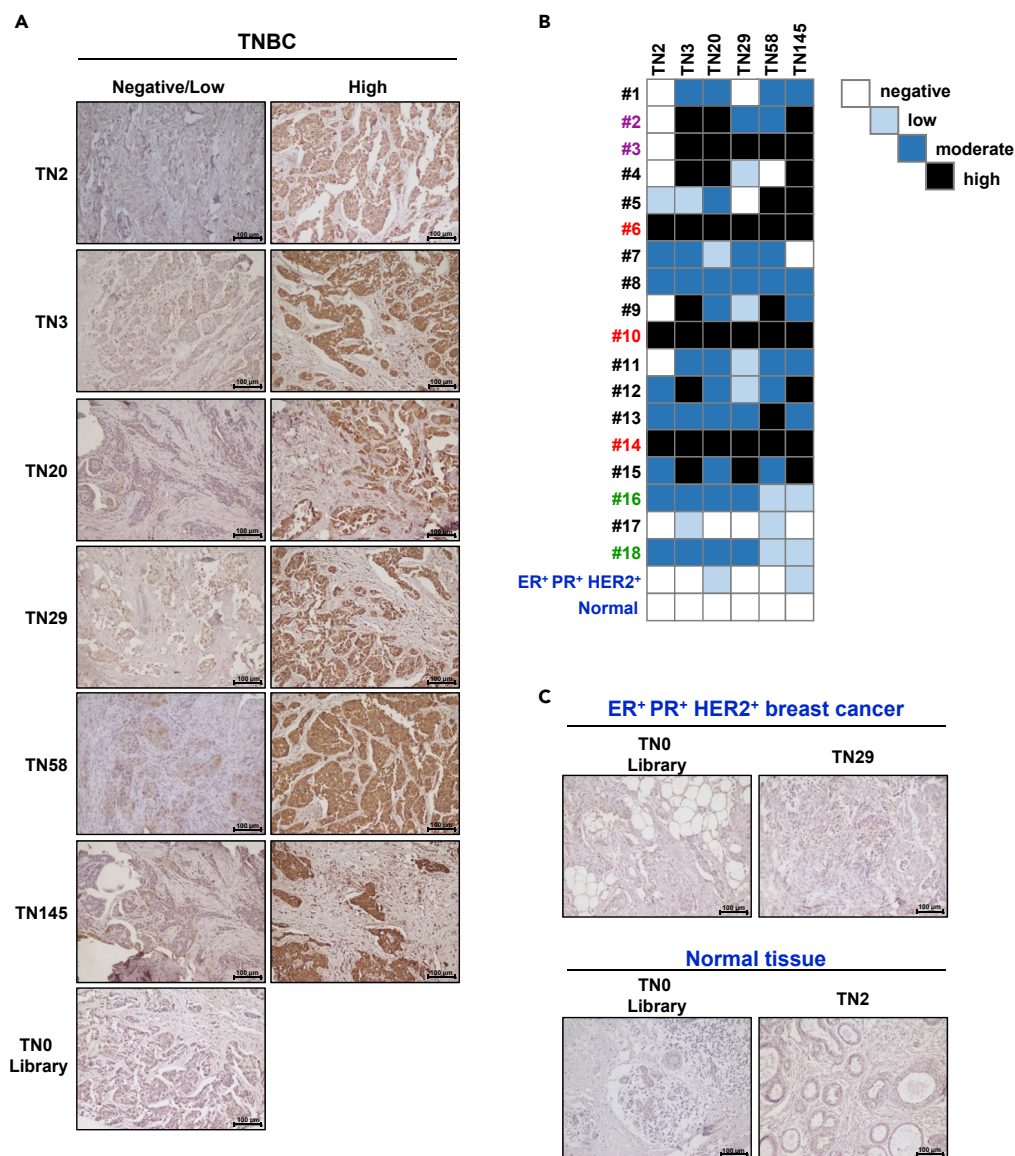
Figure 5. Human Tissue Staining by TNBC Aptamers

(A) Images of representative human TNBC cases stained with the six candidate aptamers or TN0 starting library. The images were specifically chosen to represent a variation in staining pattern for each aptamer. Negative/low (left panels) and high (right panels) staining scores are shown.

(B) Aptamer-staining scores (calculated based on both staining intensity and cell percentage) on the 18 TNBC samples, a representative triple-positive (ER$^+$, PR$^+$, HER2$^+$) breast cancer, and a representative normal breast sample. Different scores are highlighted by different colors: white (negative, score 0), light blue (low, score 1), dark blue (moderate, score 2–4), and black (high, score 6–9). Groups of samples with same/similar pattern of aptamer staining are marked in red (#6, #10, and #14), in green (#16 and #18), or in purple (#2 and #3).

(C) Breast cancer samples expressing ER, PR, and HER2 (upper panels) and breast normal samples dissected adjacent to TNBC cases (lower panels) were stained with TNBC aptamers or TN0 starting library. Images of representative samples of breast cancers and normal tissues stained with a representative TNBC aptamer and TN0 starting library are shown.

(A and C) Magnification 20×, scale bar = 100 μm.

In Vitro Functional Efficacy of TNBC Aptamers

In most cases, by binding to their proper protein target, aptamers strongly interfere with its biological activity, thus exerting significant anti-tumor activity in the case the target is linked to cancer cell growth and progression (Camorani et al., 2018a, 2018b). Therefore, we asked whether the selected TNBC

aptamers affect cancer cells malignancy. Both MDA-MB-231 and BT-549 cells have the propensity to form mammospheres under non-adherent and non-differentiated culture conditions, a cancer-stem-like trait that is associated with aggressive and metastatic nature of TNBC (Oh et al., 2018). Based on this, as readout of their anti-tumor functionality, the TNBC-specific aptamers were tested for inhibition of *in vitro* mammosphere-forming ability of both cell lines (Figure 6). Importantly, the mammosphere formation was drastically impaired in the presence of each of the six aptamers, while resulting indifferent to the starting library negative control (Figures 6A and 6B), as evidenced by a significant reduction in both number and size of mammospheres, after four days of treatment. Moreover, according to the role of CD44 in enhancing self-renewal and mammosphere growth (To et al., 2010), both MDA-MB-231 and BT-549 cells strongly up-regulated the CD44 gene, when grown in stem-permissive conditions, and the aptamers counteracted its induction of expression (Figure 6C).

The presence of 2'F-Py allows extending half-lives in serum of unmodified aptamers that can be as short as few seconds with RNA aptamers (Chen et al., 2017; Dassie et al., 2014). Accordingly, we found that TN2 aptamer, chosen among the six aptamers because it contains the lowest number of 2'F-Py, is almost fully stable up to 8 h in 80% human serum and then slowly degrades.

These findings suggest a potential therapeutic role for the selected aptamers, which likely target important surface TNBC biomarkers.

DISCUSSION

The development of selective TNBC targeting agents, which might ensure therapeutic applications, would greatly advance the development of personalized cancer therapy.

Here, we performed a new TNBC cell-SELEX approach, combined with high-throughput NGS and bioinformatics, which applied stringent criteria for identifying the best candidate aptamers to be tested for targeting efficacy. This strategy allowed us to restrict the efficacy testing to a small panel of potential candidates starting from 16,500 analyzed sequences. Noteworthily, among the seven aptamer candidates, TN1, TN2, and TN3 had unusual high frequencies in the initial RNA library probably due to some biased sequences in the custom library, as previously reported (Takahashi et al., 2016). Importantly, one of these biased sequences was selected out because it was not enriched along the cell-SELEX selection rounds. However, consistently with Takahashi et al. (2016), the three selected top sequences did not limit the identification of several high-affinity aptamers (TN20, TN29, TN58, and TN145), and also, among them, TN2 and TN3 showed optimal performance in all the assays (targeting and functional), whereas TN1 was discarded in binding analyses.

The six aptamers bind to the TNBC MDA-MB-231 target cells with K_d values in the low nanomolar range. In agreement with the principles of the differential cell-SELEX approach, able to provide multiple ligands that recognize molecular differences (e.g. protein expression level or protein conformation) between the target cell population and the cells used for counterselection (Camorani et al., 2018a, 2018b; Keefe et al., 2010), the selected RNA aptamers efficiently discriminate the target cells from triple-positive (HER2[+], ER[+], PR[+]) breast cancer cells (BT-474) used in the counterselection. Furthermore, the six aptamers have unequivocal efficacy in targeting TNBC cells lines, representative of different molecular subtypes, distinguishing them from non-TNBC breast cancer cells and non-malignant cells. Despite the yet limited published body of work regarding aptamer-based histochemistry, aptamers have the potential to revolutionize the field of histopathological diagnostics, and this is one of the most intense research areas by the aptamer scientific community (Bauer et al., 2016). Noteworthy, we succeed in using all the six aptamers in a biotinylated form for histochemical staining of 18 human TNBC samples in a TNBC tissue microarray. Importantly, they are not or poorly recognized samples from triple-positive (ER[+], PR[+], HER2[+]) breast cancers and did not bind to normal adjacent tissues, thus encouraging further analyses by extending the aptamer-based staining to a larger cohort of tumors. Noteworthily, each TNBC sample has its own unique molecular expression signature revealed by the combination of staining intensity of the six aptamers. Therefore, we believe that the aptamers identified in this study may help to advance the needs of precise molecular subtyping of breast cancers. Also, they could provide new biomarkers for detection and delivery of personalized medicine and, hopefully, reducing time and effort required for conventional immunohistochemistry.

Our results suggest that the targets of the TNBC-specific aptamers are cell surface proteins with an important role in cancer cell malignancy. Importantly, aptamer's binding is able to disrupt such function,
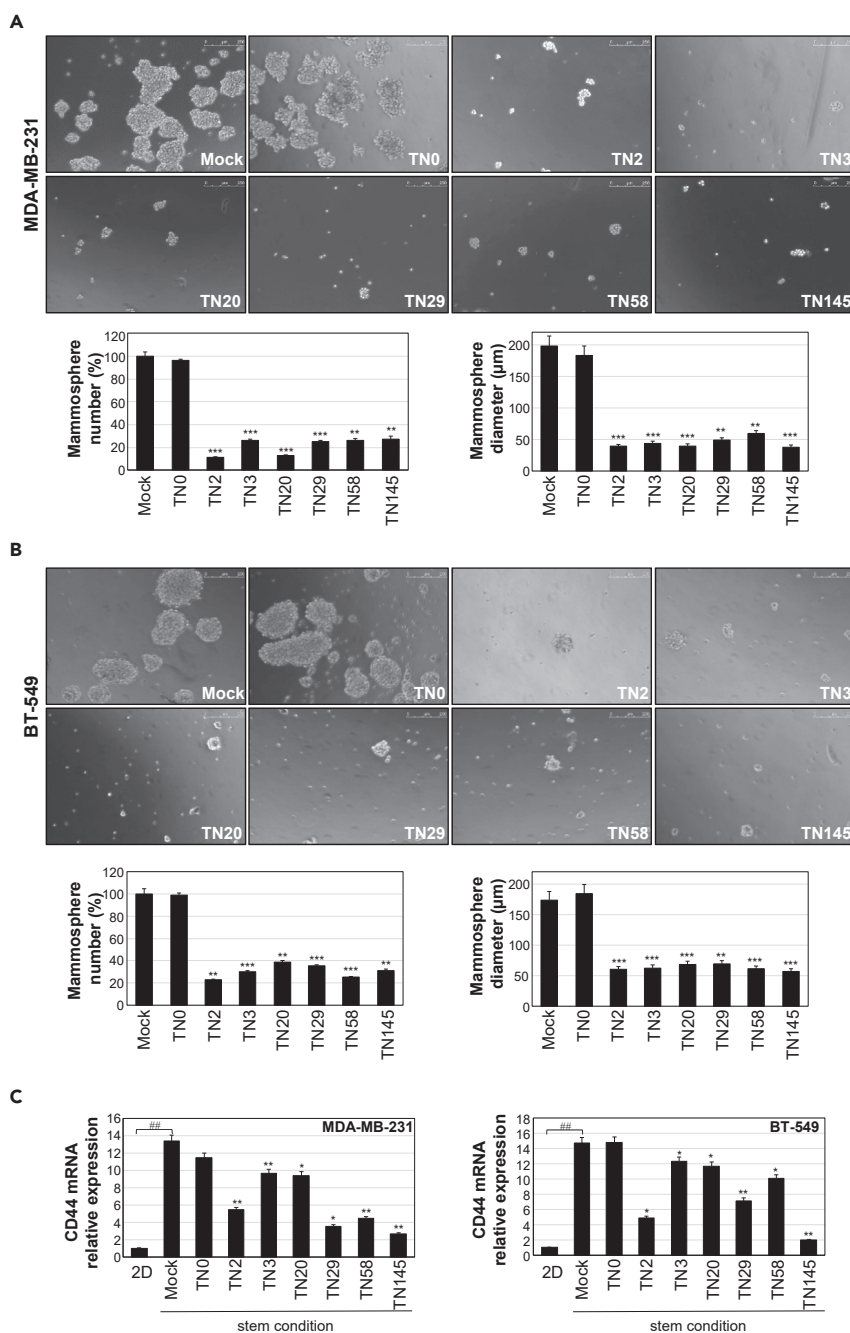
Figure 6. TNBC Aptamers Affect Mammosphere-Forming Efficiency of ML TNBC Cells

(A and B) (A) MDA-MB-231 and (B) BT-549 cells were plated in ultralow attachment 24-multiwell-plates and grown in stem-permissive conditions in the presence of 400 nM indicated aptamers or TN0 starting library for four days. Representative phase-contrast images are shown. Magnification 10×, scale bar = 250 μm. Cell treatment with specific aptamers, but not the TN0, inhibits both the number, expressed as percentage with respect to mock-treated cells (*left*), and diameter (*right*) of mammospheres.

(C) RT-qPCR analysis of CD44 gene expression in MDA-MB-231 (*left*) and BT-549 (*right*) cells grown in adherent two-dimensional condition (2D) and stem condition in the absence (Mock) or presence of indicated aptamers or TN0 library.

(A–C) Bars depict mean ± SD of three independent experiments. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$ relative to mock-treated cells; ##$p < 0.01$ (unpaired *t* test). No statistically significant variations among TN0 and mock treatment were obtained.

as revealed by their action to interfere with TNBC cell growth as mammospheres, opening up a potential therapeutic role for these aptamers, with the possibility of using them in different combinations depending on distinct molecular and/or clinical TNBC phenotypes. It is notable that the six selected aptamers have different sequences and predicted secondary structures, suggesting they likely hit different targets on TNBC cells. Although we are aware that the RNA structural complexity represents a big challenge for computational modeling, in our approach we used the secondary structure prediction to further define similarity among sequences and to select aptamers mostly different from each other. Consistently, each of them has a unique pattern of intensity of staining onto the panel of TNBC tissues analyzed in our study. We are currently investigating the minimized versions for each aptamer that retain the cell targeting properties of the full-length sequence to be used for target-identification studies.

Cytotoxic chemotherapy is the mainstay treatment for patients with TNBC in both the early and advanced stages (Bianchini et al., 2016). However, the treatment is toxic and, after an initial response, a large percentage of tumors commonly relapse causing a death rate in the metastatic setting disproportionately higher than any other breast cancer subtypes (Dent et al., 2007; Abramson et al., 2015). Identification of novel drug targets, especially those that might impact chemotherapy resistance, is an important unmet medical need. Importantly, our selected TNBC aptamers bind at high affinity to TNBC cells resistant to cisplatin and doxorubicin, with TN58 aptamer displaying, approximately, a 3-fold higher affinity for chemoresistant cells with respect to their parental counterpart. Whether the aptamers interfere with aggressiveness of chemo-resistant cells and/or revert their resistance to cisplatin and doxorubicin remains a not yet tested possibility.

Nevertheless, whether TNBC aptamers are able to readily internalize into target cells, as already assessed for TN20 aptamer and other aptamers targeting MDA-MB-231 cells (Futami et al., 2019), they may be explored as agents to deliver novel therapeutics and/or classical drugs to the tumor. This will allow to increase the concentration at the tumor site while dispensing lower absolute doses of the cargo (therefore possibly reducing systemic sideeffects) (Alshaer et al., 2018).

Nanoparticle technology for chemotherapeutic delivery has been approved by the FDA or are in clinical trials to treat patients with breast cancer (Tang et al., 2017), and several preclinical studies demonstrate that aptamer-conjugated nanoparticles can guide drugs to tumor sites with superior tumor penetration than antibodies (Xiao and Farokhzad, 2012). Recently, the anti-nucleolin AS1411 aptamer has been used to deliver the anti-cancer triptolide specifically to MDA-MB-231-derived tumors implanted in mice, which efficiently inhibited tumor growth without affecting healthy organs (He et al., 2020). These results highlight the power of aptamers to transport cytotoxic agents to TNBC overcoming limitation due to high toxicity, poor solubility, and poor bioavailability. In this frame, the identification of our multiple tumor-targeting aptamers opens the way to their use in different combinations depending on distinct molecular and/or clinical TNBC phenotypes.

In addition to therapeutic utility, TNBC aptamers may be useful as specific probes for molecular imaging, thus increasing a repertoire that is currently extremely scarce (Camorani et al., 2018b), as we recently demonstrated for the PDGFRβ aptamer that can act for both specific non-invasive imaging and suppression of lung metastases in mouse models of TNBC (Camorani et al., 2018c).

Limitations of the Study

Although the aptamers we identified in this study are highly selective binders of surface proteins of TNBC target cells and discriminate TNBC cell lines from both non-malignant cells and non-TNBC breast cancer cells, a great challenge is that their molecular targets need to be identified and validated. This aspect of the study is crucial to address the relevance and in-depth mechanism of aptamer binding to its own protein target. Moreover, as we aim to apply the proposed aptamer-based staining approach for advancing molecular subtyping of breast cancers, it is necessary to further extend the analyses to a larger cohort of tumors, correlating the expression of the target protein to clinical features. Finally, although the six selected aptamers show strong potential for TNBC diagnosis and therapy, additional studies are required to assess their efficacy *in vivo* (including stability and lack of immunogenicity). Future studies in animal models will help to determine whether this is a viable strategy for TNBC treatment.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.100979.

AUTHOR CONTRIBUTIONS

S.C. performed the majority of the experiments, analyzed data, and assisted with manuscript preparation. I.G. and M.R.G. performed the bioinformatic analyses of NGS data and participated in editing of the manuscript. F.C. and M.C. performed the aptamer-based TMA staining and evaluated tumor morphology and histochemical staining. F.L. contributed to the cell-based assays experiments. G.B. provided the human TNBC samples. M.F. provided conceptual advice, analyzed the data, and participated in editing of the manuscript. L.C. conceived and designed the study, analyzed the data and coordinated the research, was responsible for funding, and wrote the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Abramson, V.G., Lehmann, B.D., Ballinger, T.J., and Pietenpol, J.A. (2015). Subtyping of triple-negative breast cancer: implications for therapy. Cancer *121*, 8–16.

Alshaer, W., Hillaireau, H., Vergnaud, J., Mura, S., Deloménie, C., Sauvage, F., Ismail, S., and Fattal, E. (2018). Aptamer-guided siRNA-loaded nanomedicines for systemic gene silencing in CD-44 expressing murine triple-negative breast cancer model. J.Control Release *271*, 98–106.

Bauer, M., Macdonald, J., Henri, J., Duan, W., and Shigdar, S. (2016). The application of aptamers for immunohistochemistry. Nucleic Acid Ther. *26*, 120–126.

Beniey, M., Haque, T., and Hassan, S. (2019). Translating the role of PARP inhibitors in triple-negative breast cancer. Oncoscience *6*, 287–288.

Betapudi, V., Licate, L.S., and Egelhoff, T.T. (2006). Distinct roles of nonmuscle myosin II isoforms in the regulation of MDA-MB-231 breast cancer cell spreading and migration. Cancer Res. *66*, 4725–4733.

Bianchini, G., Balko, J.M., Mayer, I.A., Sanders, M.E., and Gianni, L. (2016). Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. Nat. Rev. Clin. Oncol. *13*, 674–690.

Blick, T., Widodo, E., Hugo, H., Waltham, M., Lenburg, M.E., Neve, R.M., and Thompson, E.W. (2008). Epithelial mesenchymal transition traits in human breast cancer cell lines. Clin. Exp. Metastasis *25*, 629–642.

Burstein, M.D., Tsimelzon, A., Poage, G.M., Covington, K.R., Contreras, A., Fuqua, S.A., Savage, M.I., Osborne, C.K., Hilsenbeck, S.G.,

Chang, J.C., et al. (2015). Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. Clin. Cancer Res. *21*, 1688–1698.

Camorani, S., Crescenzi, E., Colecchia, D., Carpentieri, A., Amoresano, A., Fedele, M., Chiariello, M., and Cerchia, L. (2015). Aptamer targeting EGFRvIII mutant hampers its constitutive autophosphorylation and affects migration, invasion and proliferation of glioblastoma cells. Oncotarget *6*, 37570–37587.

Camorani, S., Crescenzi, E., Fedele, M., and Cerchia, L. (2018a). Oligonucleotide aptamers against tyrosine kinase receptors: prospect for anticancer applications. Biochim.Biophys.Acta Rev. Cancer *1869*, 263–277.

Camorani, S., Crescenzi, E., Gramanzini, M., Fedele, M., Zannetti, A., and Cerchia, L. (2017a). Aptamer-mediated impairment of EGFR-integrin αvβ3 complex inhibits vasculogenic mimicry and growth of triple-negative breast cancers. Sci. Rep. *7*, 46659.

Camorani, S., Esposito, C.L., Rienzo, A., Catuogno, S., Iaboni, M., Condorelli, G., de Franciscis, V., and Cerchia, L. (2014). Inhibition of receptor signaling and of glioblastoma-derived tumor growth by a novel PDGFRβ aptamer. Mol. Ther. *22*, 828–841.

Camorani, S., Fedele, M., Zannetti, A., and Cerchia, L. (2018b). TNBC challenge: oligonucleotide aptamers for new imaging and therapy modalities. Pharmaceuticals (Basel). *11*, 123.

Camorani, S., Hill, B.S., Collina, F., Gargiulo, S., Napolitano, M., Cantile, M., Di Bonito, M., Botti, G., Fedele, M., Zannetti, A., et al. (2018c).

Targeted imaging and inhibition of triple-negative breast cancer metastases by a PDGFRβ aptamer. Theranostics *8*, 5178–5199.

Camorani, S., Hill, B.S., Fontanella, R., Greco, A., Gramanzini, M., Auletta, L., Gargiulo, S., Albanese, S., Lucarelli, E., Cerchia, L., et al. (2017b). Inhibition of bone marrow-derived mesenchymal stem cells homing towards triple-negative breast cancer microenvironment using an anti-PDGFRβaptamer. Theranostics *7*, 3595–3607.

Chen, H.W., Medley, C.D., Sefah, K., Shangguan, D., Tang, Z., Meng, L., Smith, J.E., and Tan, W. (2008). Molecular recognition of small-cell lung cancer cells using aptamers. ChemMedChem *3*, 991–1001.

Chen, Z., Liu, H., Jain, A., Zhang, L., Liu, C., and Cheng, K. (2017). Discovery of aptamer ligands for hepatic stellate cells using SELEX. Theranostics *7*, 2982–2995.

Cortés, J., André, F., Gonçalves, A., Kümmel, S., Martín, M., Schmid, P., Schuetz, F., Swain, S.M., Easton, V., Pollex, E., et al. (2019). IMpassion132 Phase III trial: atezolizumab and chemotherapy in early relapsing metastatic triple-negatve breast cancer. Future Oncol. *15*, 1951–1961.

Dai, X., Cheng, X., Bai, Z., and Li, J. (2017). Breast cancer cell line classification and its relevance with breast tumor subtyping. J. Cancer *8*, 3131–3141.

Dassie, J.P., Hernandez, L.I., Thomas, G.S., Long, M.E., Rockey, W.M., Howell, C.A., Chen, Y., Hernandez, F.J., Liu, X.Y., and Wilson, M.E. (2014). Targeted inhibition of prostate cancer metastases with an RNA aptamer to

prostate-specific membrane antigen. Mol. Ther. 22, 1910–1922.

Dent, R., Trudeau, M., Pritchard, K.I., Hanna, W.M., Kahn, H.K., Sawka, C.A., Lickley, L.A., Rawlinson, E., Sun, P., and Narod, S.A. (2007). Triple-negative breast cancer: clinical features and patterns of recurrence. Clin. Cancer Res. 13, 4429–4434.

D'Ippolito, E., Plantamura, I., Bongiovanni, L., Casalini, P., Baroni, S., Piovan, C., Orlandi, R., Gualeni, A.V., Gloghini, A., Rossini, A., et al. (2016). miR-9 and miR-200 regulate PDGFRβ-mediated endothelial differentiation of tumor cells in triple-negative breast cancer. Cancer Res. 76, 5562–5572.

Echeverria, G.V., Ge, Z., Seth, S., Zhang, X., Jeter-Jones, S., Zhou, X., Cai, S., Tu, Y., McCoy, A., Peoples, M., et al. (2019). Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. Sci. Transl Med. 11, eaav0936.

Esposito, C.L., Passaro, D., Longobardo, I., Condorelli, G., Marotta, P., Affuso, A., de Franciscis, V., and Cerchia, L. (2011). A neutralizing RNA aptamer against EGFR causes selective apoptotic cell death. PLoS One 6, e24071.

Futami, K., Kimoto, M., Lim, Y.W.S., and Hirao, I. (2019). Genetic alphabet expansion provides versatile specificities and activities of unnatural-base DNA aptamers targeting cancer cells. Mol. Ther.Nucleic Acids 14, 158–170.

Gadi, V.K., and Davidson, N.E. (2017). Practical approach to triple-negative breast cancer. J. Oncol. Pract. 13, 293–300.

Garnett, M.J., Edelman, E.J., Heidorn, S.J., Greenman, C.D., Dastur, A., Lau, K.W., Greninger, P., Thompson, I.R., Luo, X., Soares, J., et al. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. Nature 483, 570–575.

Garrido-Castro, A.C., Lin, N.U., and Polyak, K. (2019). Insights into molecular classifications of triple-negative breast cancer: improving patient selection for treatment. Cancer Discov. 9, 176–198.

Han, H.J., Russo, J., and Kohwi, Y. (2008). SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. Nature 452, 187–193.

Handl, J., Knowles, J., and Kell, D.B. (2005). Computational cluster validation in post-genomic data analysis. Bioinformatics 21, 3201–3212.

He, J., Peng, T., Peng, Y., Ai, L., Deng, Z., Wang, X.Q., and Tan, W. (2020). Molecularly engineering triptolide with aptamers for high specificity and cytotoxicity for triple-negative breast cancer. J. Am. Chem.Soc. 142, 2699–2703.

Keefe, A.D., Pai, S., and Ellington, A. (2010). Aptamers as therapeutics. Nat. Rev. Drug Discov. 9, 537–550.

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol,

J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J. Clin.Invest 121, 2750–2767.

Lehmann, B.D., Jovanović, B., Chen, X., Estrada, M.V., Johnson, K.N., Shyr, Y., Moses, H.L., Sanders, M.E., and Pietenpol, J.A. (2016). Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. PLoS One 11, e0157368.

Li, X., Zhang, W., Liu, L., Zhu, Z., Ouyang, G., An, Y., Zhao, C., and Yang, C.J. (2014). In vitro selection of DNA aptamers for metastatic breast cancer cell recognition and tissue imaging. Anal. Chem. 86, 6596–6603.

Liu, M., Wang, Z., Tan, T., Chen, Z., Mou, X., Yu, X., Deng, Y., Lu, G., and He, N. (2018). An aptamer-based probe for molecular subtyping of breast cancer. Theranostics 8, 5772–5783.

Nabholtz, J.M., Abrial, C., Mouret-Reynier, M.A., Dauplat, M.M., Weber, B., Gligorov, J., Forest, A.M., Tredan, O., Vanlemmens, L., Petit, T., et al. (2014). Multicentric neoadjuvant phase II study of panitumumab combined with an anthracycline/taxane-based chemotherapy in operable triple-negative breast cancer: identification of biologically defined signatures predicting treatment impact. Ann. Oncol. 25, 1570–1577.

Nair, A., Chung, H.C., Sun, T., Tyagi, S., Dobrolecki, L.E., Dominguez-Vidana, R., Kurley, S.J., Orellana, M., Renwick, A., Henke, D.M., et al. (2018). Combinatorial inhibition of PTPN12-regulated receptors leads to a broadly effective therapeutic strategy in triple-negative breast cancer. Nat.Med. 24, 505–511.

Newman, L.A., Reis-Filho, J.S., Morrow, M., Carey, L.A., and King, T.A. (2015). The 2014 society of surgical oncology Susan G. Komen for the cure symposium: triple-negative breast cancer. Ann. Surg. Oncol. 22, 874–882.

Nguyen, L.V., Cox, C.L., Eirew, P., Knapp, D.J., Pellacani, D., Kannan, N., Carles, A., Moksa, M., Balani, S., Shah, S., et al. (2014). DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. Nat. Commun. 5, 5871.

Nowsheen, S., Cooper, T., Bonner, J.A., Lo Buglio, A.F., and Yang, E.S. (2012). HER2 overexpression renders human breast cancers sensitive to PARP inhibition independently of any defect in homologous recombination DNA repair. Cancer Res. 72, 4796–4806.

Oh, E., Kim, Y.J., An, H., Sung, D., Cho, T.M., Farrand, L., Jang, S., Seo, J.H., and Kim, J.Y. (2018). Flubendazole elicits anti-metastatic effects in triple-negative breast cancer via STAT3 inhibition. Int. J. Cancer 143, 1978–1993.

O'Reilly, E.A., Gubbins, L., Sharma, S., Tully, R., Guang, M.H., Weiner-Gorzel, K., McCaffrey, J., Harrison, M., Furlong, F., Kell, M., et al. (2015). The fate of chemoresistance in triple negative breast cancer (TNBC). BBA Clin. 3, 257–275.

Pasleau, F., Grooteclaes, M., and Gol-Winkler, R. (1993). Expression of the c-erbB2 gene in the BT474 human mammary tumor cell line:

measurement of c-erbB2 mRNA half-life. Oncogene 8, 849–854.

Ross, D.T., and Perou, C.M. (2001). A comparison of gene expression signatures from breast tumors and breast tissue derived cell lines. Dis. Markers 17, 99–109.

Sandberg, R., and Ernberg, I. (2005). Assessment of tumor characteristic gene expression in cell lines using a tissue similarity index (TSI). Proc. Natl. Acad. Sci. U S A. 102, 2052–2057.

Sefah, K., Tang, Z.W., Shangguan, D.H., Chen, H., Lopez-Colon, D., Li, Y., Parekh, P., Martin, J., Meng, L., Phillips, J.A., et al. (2009). Molecular recognition of acute myeloid leukemia using aptamers. Leukemia 23, 235–244.

Shangguan, D., Li, Y., Tang, Z., Cao, Z.C., Chen, H.W., Mallikaratchy, P., Sefah, K., Yang, C.J., and Tan, W. (2006). Aptamers evolved from live cells as effective molecular probes for cancer study. Proc. Natl. Acad. Sci. U S A. 103, 11838–11843.

Takahashi, M., Wu, X., Ho, M., Chomchan, P., Rossi, J.J., Burnett, J.C., and Zhou, J. (2016). High throughput sequencing analysis of RNA libraries reveals the influences of initial library and PCR methods on SELEX efficiency. Sci. Rep. 6, 33697.

Tang, X., Loc, W.S., Dong, C., Matters, G.L., Butler, P.J., Kester, M., Meyers, C., Jiang, Y., and Adair, J.H. (2017). The use of nanoparticulates to treat breast cancer. Nanomedicine (Lond). 12, 2367–2388.

Tang, Z., Shangguan, D., Wang, K., Shi, H., Sefah, K., Mallikratchy, P., Chen, H.W., Li, Y., and Tan, W. (2007). Selection of aptamers for molecular recognition and characterization of cancer cells. AnalChem. 79, 4900–4907.

To, K., Fotovati, A., Reipas, K.M., Law, J.H., Hu, K., Wang, J., Astanehe, A., Davies, A.H., Lee, L., Stratford, A.L., et al. (2010). Y-box binding protein-1 induces the expression of CD44 and CD49f leading to enhanced self-renewal, mammosphere growth, and drug resistance. Cancer Res. 70, 2840–2851.

Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yarden, Y., Libermann, T.A., Schlessinger, J., et al. (1984). Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature 309, 418–425.

Xiao, Z., and Farokhzad, O.C. (2012). Aptamer-functionalized nanoparticles for medical applications: challenges and opportunities. ACS Nano 6, 3670–3676.

Zhou, J., and Rossi, J. (2017). Aptamers as targeted therapeutics: current potential and challenges. Nat. Rev. Drug Discov. 16, 181–202.

Ziegler, Y.S., Moresco, J.J., Tu, P.G., Yates, J.R., 3rd, and Nardulli, A.M. (2014). Plasma membrane proteomics of human breast cancer cell lines identifies potential targets for breast cancer diagnosis and treatment. PLoS One 9, e102341.
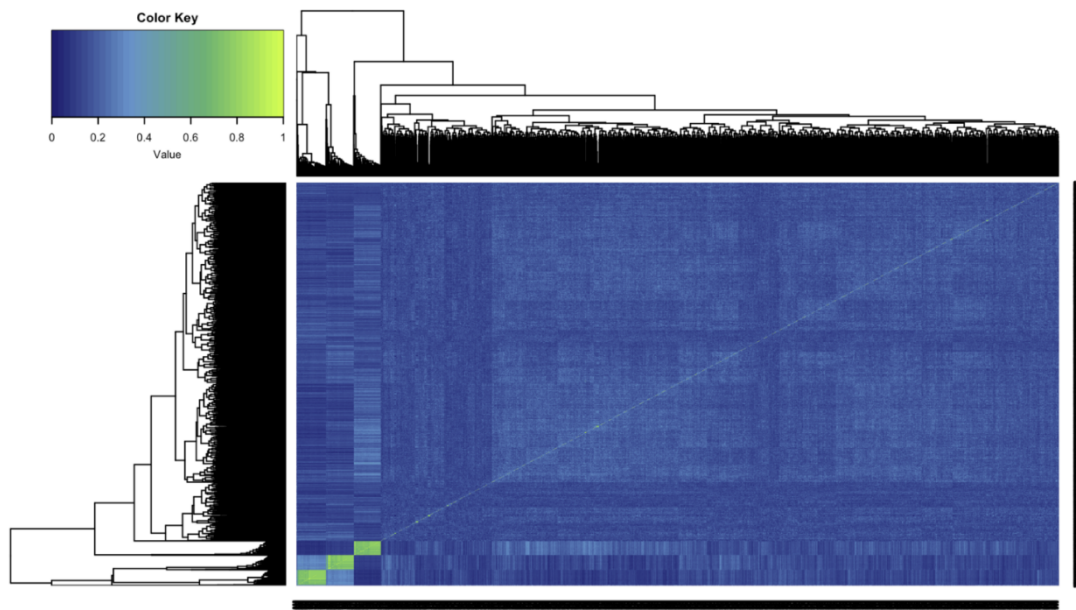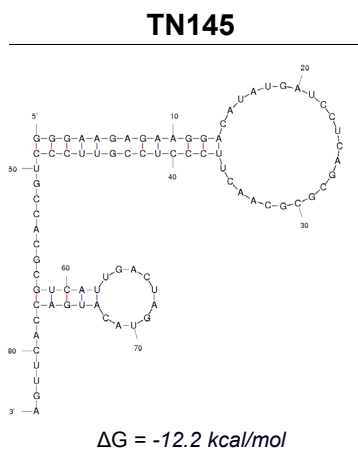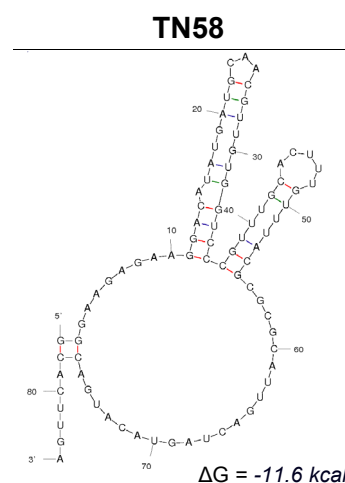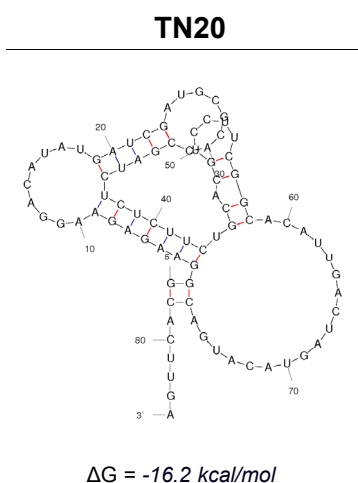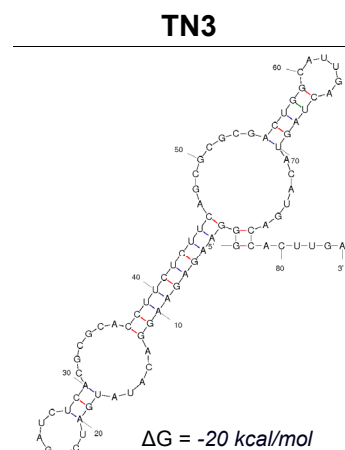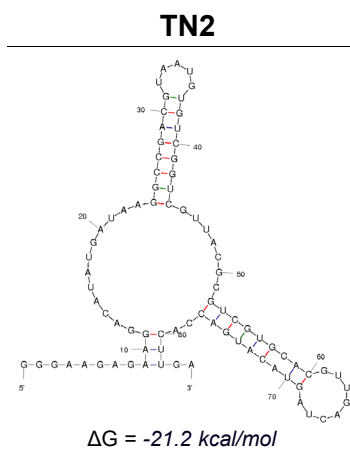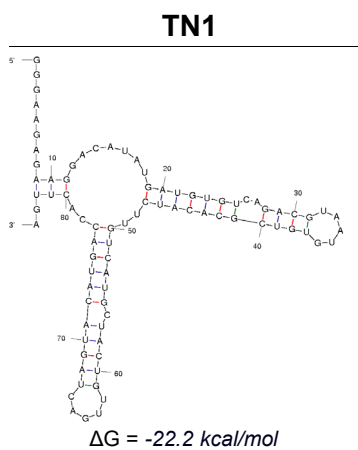
Supplemental Information

Novel Aptamers Selected

on Living Cells for Specific Recognition

of Triple-Negative Breast Cancer

Simona Camorani, Ilaria Granata, Francesca Collina, Francesco Leonetti, Monica Cantile, Gerardo Botti, Monica Fedele, Mario Rosario Guarracino, and Laura Cerchia

Figure S1. Alignment-based similarity heatmap, Related to Figure 2. Only three groups of sequences show a high similarity, as it is highlighted by the dendrogram on the top and on the left, but in general the 5000 sequences show poor identity, indicating a great divergence in the variable region.

TN1
ΔG = -22.2 kcal/mol

TN2
ΔG = -21.2 kcal/mol

TN3
ΔG = -20 kcal/mol

TN20
ΔG = -16.2 kcal/mol

TN29
ΔG = -14.2 kcal/mol

TN58
ΔG = -11.6 kcal/mol

TN145
ΔG = -12.2 kcal/mol
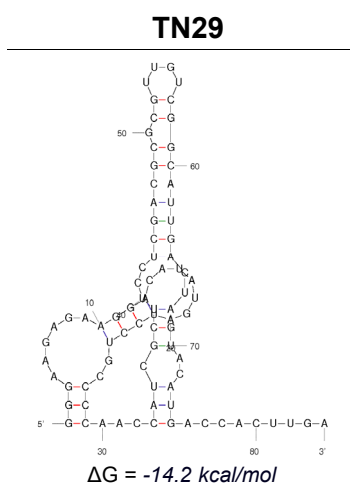
Figure S2. Secondary structure of candidate aptamers, Related to Figure 2. Structures were predicted by using Quikfold software. ΔG values are reported.
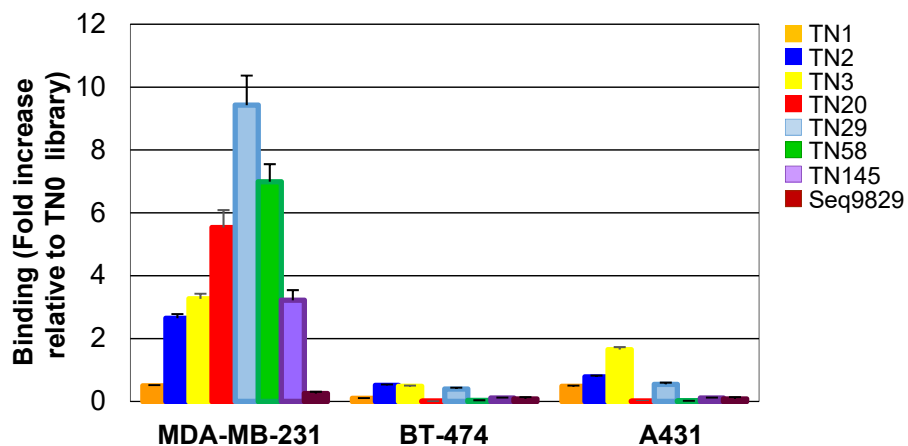
Figure S3. Binding analysis of candidate aptamers with target cells and cells of counterselections by RT-qPCR, Related to Table 1. The indicated aptamers, sequence 9829 or the starting library (200-nM final concentration) were incubated for 15 min at 37°C with MDA-MB-231 (selection), BT-474 (first counterselection) or A431 (second counterselection) cells and binding was detected by RT-qPCR. The results are expressed relative to the background binding detected with the TN0 library. Bars depict mean ± SD of three independent experiments.
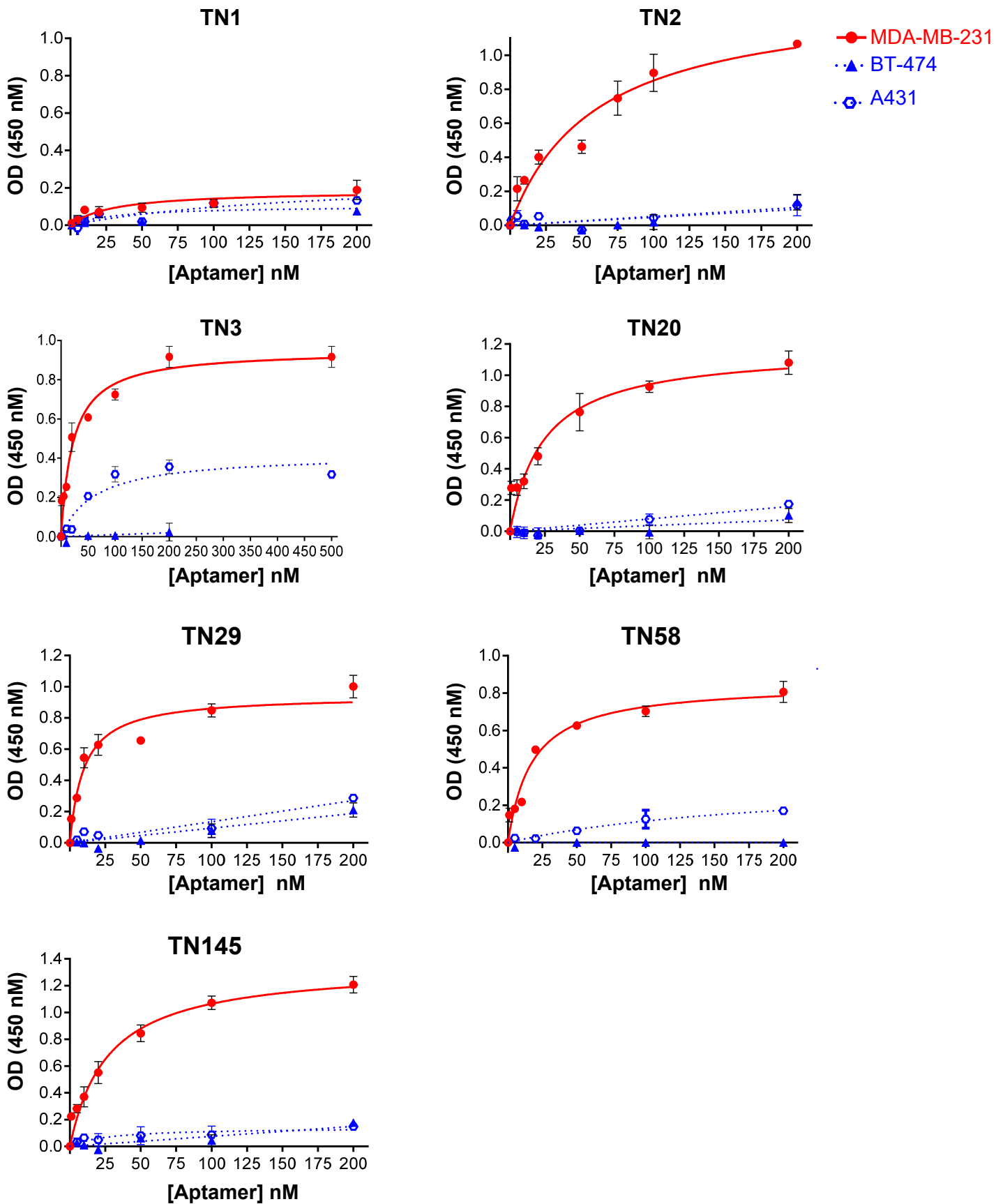
Figure S4. Binding analysis of candidate aptamers with target cells and cells of counterselections by streptavidin-biotin-based assay, Related to Table 1. Binding curves of the indicated 3'-biotinylated aptamers on MDA-MB-231 (selection, red line) and BT-474 or A431 (counterselections, dotted blue lines) cells for calculation of the apparent K_d of aptamer-cell interaction. Binding was analyzed using streptavidin-biotin-based (colorimetric) assay. The background binding value for the TN0 library was subtracted from every data point. Data shown are mean ± SD of three independent experiments.
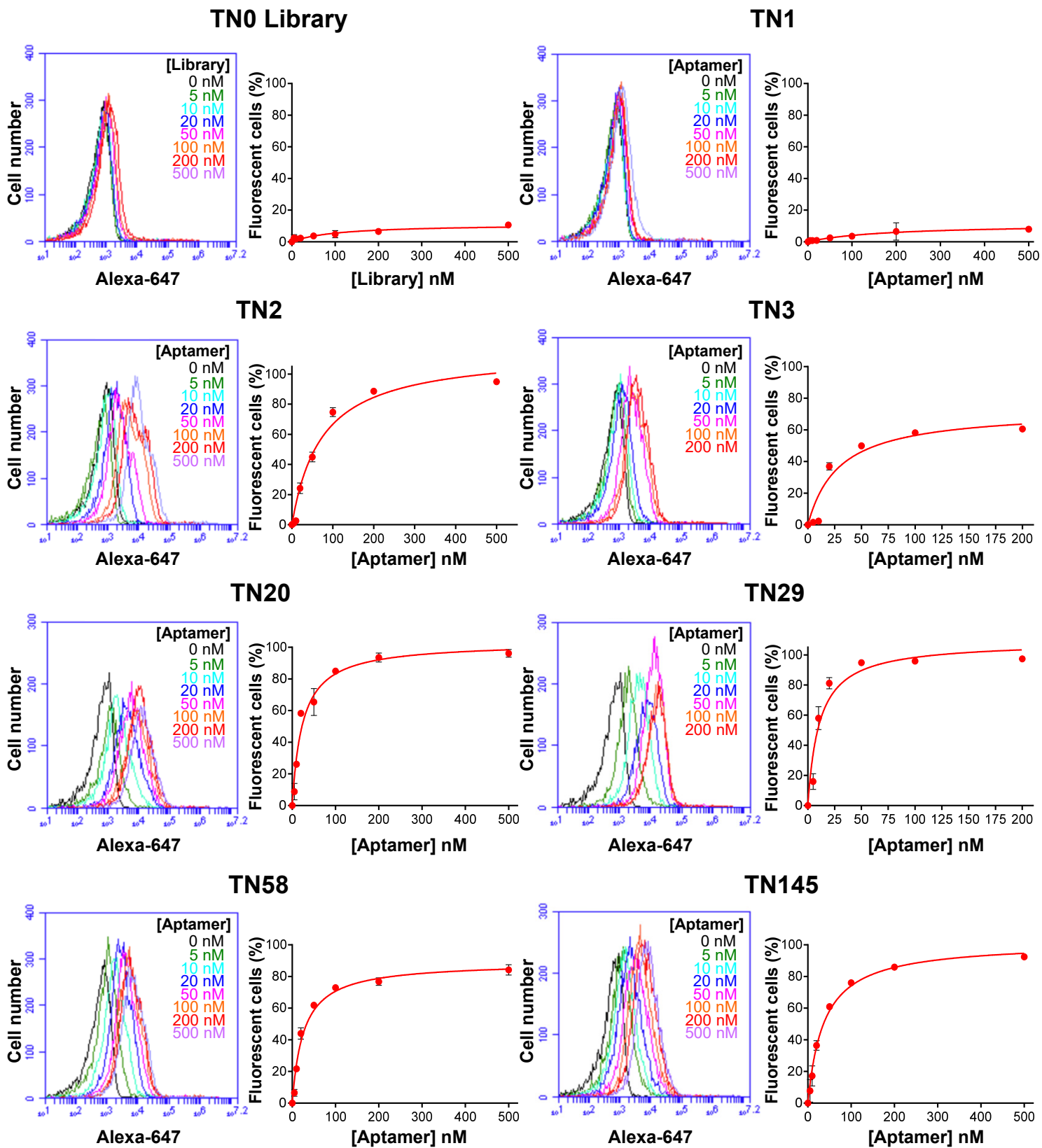
Figure S5. Binding analysis of candidate aptamers with target cells by flow cytometry, Related to Table 1. For each Alexa 647-labeled aptamer and the TN0 library is shown the flow cytometry analysis of increasing concentrations of aptamer incubated with MDA-MB-231 cells (*left*) and the binding curve for calculation of the apparent K_d of aptamer-cell interaction, which was obtained by plotting the percentage of fluorescent cells (reflecting the percentage of cells targeted by the Alexa 647-labeled aptamer), as calculated by BD Accuri C6 Software with respect to mock-treated cells, against varying aptamer concentrations (*right*). Data shown are mean ± SD of three independent experiments.

A

B
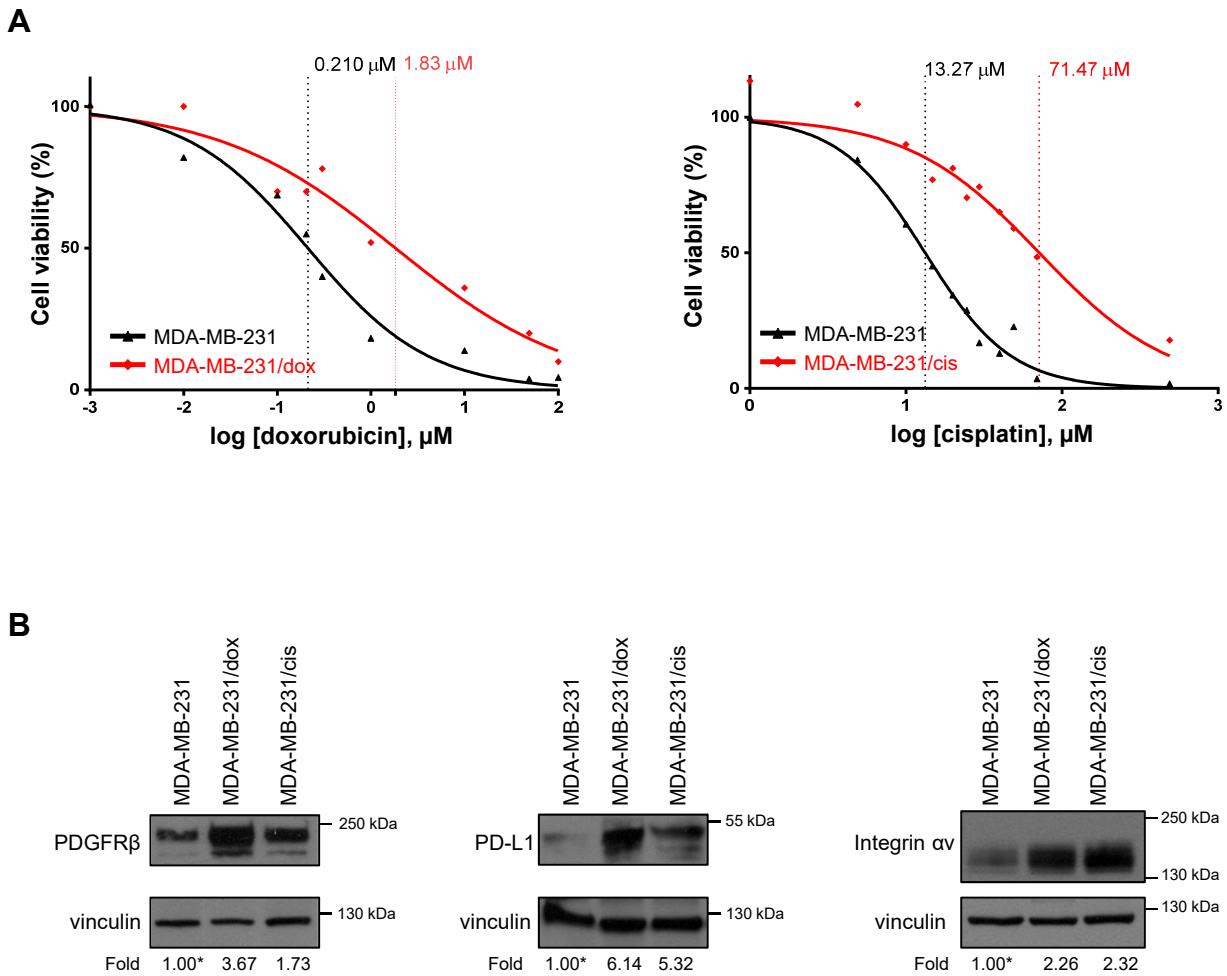
Figure S6. Characterization of chemoresistant MDA-MB-231 cells, Related to Figure 4. (A) MDA-MB-231/dox and MDA-MB-231/cis cells exhibit increased fold changes in IC50 concentrations with respect to parental MDA-MB-231 cells. Following 72 h of treatment of dox or cisplatin cell viability was assessed and the IC50 value for each treatment was calculated as described in Methods. Data are representative of three independent experiments. (B) Lysates from MDA-MB-231, MDA-MB-231/dox and MDA-MB-231/cis cells were immunoblotted with anti-PDGFRβ, anti-PDL1 and anti-integrin αv antibodies, as indicated. Vinculin was used as an internal control. Values below the blot indicate signal levels, normalized to the respective anti-vinculin signal level, and reported as relative to MDA-MB-231 cells, arbitrarily set to 1 (labelled with asterisk).
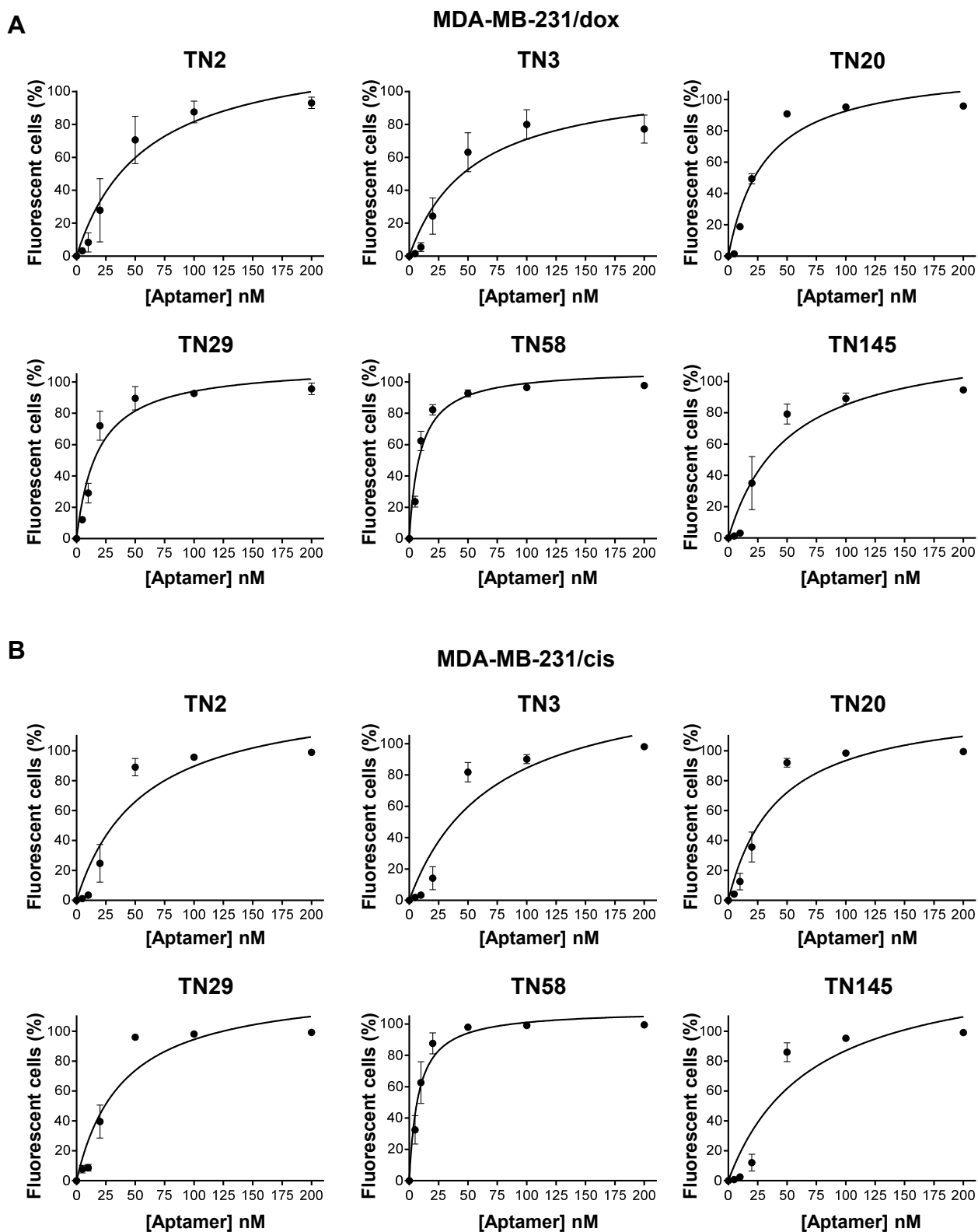
Figure S7. Binding curves of TNBC aptamers on chemoresistant MDA-MB-231 cells, Related to Figure 4.
Binding curves of the indicated Alexa 647-labelled aptamers on MDA-MB-231/dox (A) and MDA-MB-231/cis (B) cells for calculation of the apparent K_d of aptamer-cell interaction. Binding was analysed using Flow cytometric assay. Data shown are mean ± SD of three independent experiments.
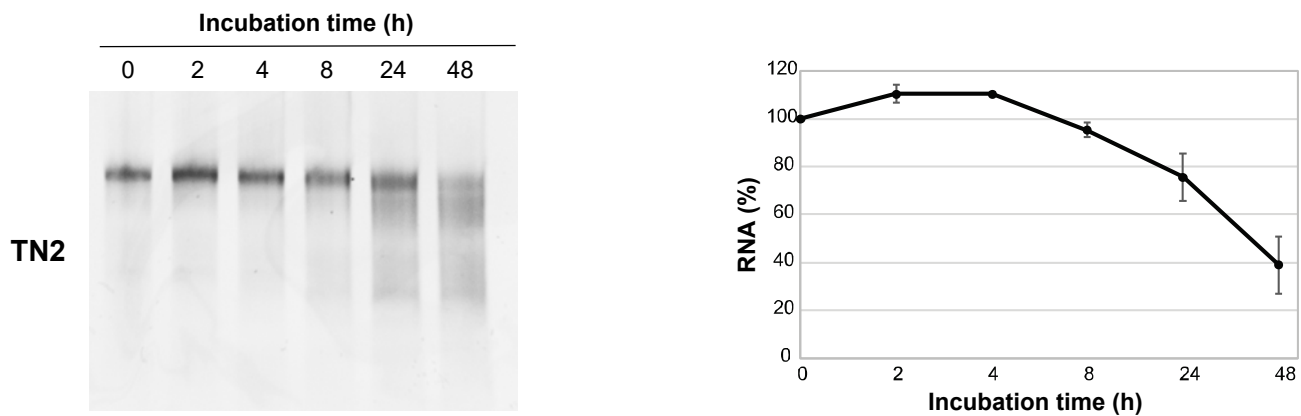
Figure S8. Analyses of the *in vitro* serum stability of TN2 Aptamer, Related to Figure 6. (*Left*) Denaturing PAGE of TN2 following incubation with 80% human serum at the indicated times. Depicted results represent one of three typical experiments performed. TN2 contains 2'F-Py in all the sequence to increase nuclease resistance. (*Right*) Bands intensity was quantified using ImageJ (v1.46r) at each time point and expressed as percentage with respect to time 0 RNA. Note that time zero was taken after 1 h incubation with proteinase K. Data shown are mean ± SD of three independent experiments.

Table S1. Selection pressure used during cell-SELEX, Related to Figure 1.

Round number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
RNA amount (nmoles)	1.5	1.5	1.5	1.2	1.4	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Cells number (x10^6)	7.0	7.0	7.0	7.0	7.0	6.3	6.3	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Number of counter-selections on BT-474	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Number of counter-selections on A431	0	0	0	0	1	1	1	1	1	1	1	1	1	1
Incubation time (min)	30	30	30	30	30	30	30	30	15	15	15	15	15	15
Washing number	1	1	2	2	3	3	4	5	5	5	5	5	5	5

The grey highlighting corresponds to the conditions that changed between two rounds.

Table S2. Clustering validation, Related to Figure 2.

Number of clusters:	5	6	7	8	9	10
Connectivity	**4.04**	4.04	4.04	4.93	18.24	32.11
Dunn	0.5	**0.51**	0.51	0.51	0.33	0.4
Silhouette	0.35	0.36	**0.38**	0.37	0.34	0.33

Internal validation measures for number of clusters ranging from 5 to 10. The internal measures include the Connectivity, Dunn Index and Silhouette Width. In bold the best result for each measure. For the connectivity the smaller is the better, while for the other two measures higher scores correspond to better clustering. Under the same values, the smaller number of clusters is preferred.

Table S3. Clinicopathological features of 18 TNBC patients, Related to Figure 5.

TNBC cases	n.18
Variables	
Age	
<40	2
40-60	10
>60	6
Histotype	
CDI	18
Grading	
1	0
2	8
3	10
LN Metastases	
Negative	11
Positive	6
NA	1
Tumor size (cm)	
<2	10
2-5	6
>5	1
NA	1
Distant Metastases	
Absent	1
Present	3
NA	14

NA: not available. LN: lymph node.

Transparent Methods

Cell lines and culture conditions

All cell lines used in this study were purchased from the American Type Culture Collection (ATCC, Manassas, VA). MDA-MB-231, BT-549, MDA-MB-468, MDA-MB-436, MDA-MB-453, DU4475, and T47D were grown in Roswell Park Memorial Institute-1640 medium (RPMI-1640, Sigma-Aldrich, Milan, Italy) supplemented with 10%, or 20% for DU4475, fetal bovine serum (FBS, Sigma-Aldrich); BT-474 were grown in HybriCare medium (ATCC) supplemented with 10% FBS; MCF7 and SK-BR-3 were grown in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich) supplemented with 10% FBS. Human mammary epithelial MCF10A cells were grown in DMEM/F-12 (Sigma-Aldrich) supplemented with 5% horse serum, 0.5 µg/ml hydrocortisone (Sigma-Aldrich), 10 ng/ml epidermal growth factor (EGF, Sigma-Aldrich) and 10 µg/ml insulin (Sigma-Aldrich). Human normal lung fibroblasts and A431 cells were grown in DMEM supplemented with 20% and 10% FBS, respectively. All cells were maintained in 95% air/5% CO_2 atmosphere at 37°C.

Cell-SELEX

Generation of the initial library

The RNA library was generated starting from the extension of a DNA template oligo containing a 40-mer random sequence region (N40) flanked by two constant sequence regions, 5'-TAGGGAAGAGAAGGACATATGAT(N40)TTGACTAGTACATGACCACTTGA-3' (TriLink Biotechnologies, San Diego, CA), by using the pRv primer, 5'-TCAAGTGGTCATGTACTAGTCAA-3' (anneal protocol: 95°C 5 min, 25°C 20 min; extend protocol: 72°C 30 min, 25°C 10 min). The pRv primer and the pT7 primer, containing the T7 transcription promoter (5'-TTCAGGTAATACGACTCACTATAGGGAAGAGAAGGACATATGAT-3'), were used for PCR amplification. The double-stranded DNA library was *in vitro* transcribed (37°C 16 h) in the presence of 1 mM 2'F-Py (TriLink BioTechnologies) and a mutant form of T7 RNA polymerase (2.5 u/ml T7 R&DNA polymerase, Epicentre Biotechnologies, Madison, WI) was used to improve yields. The RNA was purified by denaturing (7 M urea) polyacrylamide gel (8% acrylamide) electrophoresis (PAGE).

1

The RNA band was detected by ultraviolet (UV) shadowing, excised and eluted from the gel with 0.3 M NaAc buffer containing 2 mM EDTA (Invitrogen) for 2 h at 42°C in a rotator. RNA OD260 and 260/280 ratio were determined by NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA) to assess quantity and purity of RNA library.

SELEX cycle

The SELEX cycle was performed essentially as described (Camorani et al., 2014). Specifically, prior to each round, 2'F-Py RNAs (1500-600 pmol) were heated at 85°C for 5 min in 1.5 ml of RPMI-1640, snap-cooled on ice for 3 min, allowed to warm up to 37°C. Before incubation with the cells, RPMI-1640 was added to RNA to reach a final volume of 15 ml. Cells were plated in 150-mm cell plate the day before the cycle. For the first round, the RNA library was incubated with MDA-MB-231 target cells (7×10^6) for 30 min at 37°C, unbound sequences were eliminated and, after washing of the cells with RPMI-1640, the bound RNAs were recovered by total RNA extraction (TRIzol, Invitrogen, Carlsbad, CA). Starting from round 2, counterselection steps were introduced into the cycle, prior the selection phase with target cells. RNAs were first incubated for 30 min at 37 °C with 7×10^6 BT-474 cells. Unbound sequences in the supernatant were recovered for the selection phase (from round 2 to round 4) or incubated for 30 min at 37°C with 7×10^6 A431 cells, for a second counterselection step (from round 5 to round 14) preceding the incubation with MDA-MB-231 target cells. Sequences enriched by the selection step were amplified by RT-PCR and transcribed, before a new cycle of selection. The nature and growth state of cells, and the binding temperature were rigorously monitored during the entire process. Further, in order to obtain aptamers with high affinity and specificity toward the target cells, the selective pressure from second round was gradually increased (as schematized in **Table S1**) by decreasing the target-cell number (from 7×10^6 up to 4.0×10^6) and the incubation time (from 30 to 15 min), and by increasing the number of washes of target cells (from one up to five) and the counterselection steps.

High-throughput sequencing technology and bioinformatics analysis

Monitoring aptamer selection and identification of TNBC-specific aptamers was performed by Illumina ® NGS on DNA pools coming from 0, 3, 5, 8, 9, 10, 11, 12, 13 and 14 selection rounds. To prepare samples, the RNA pools coming from each selection round were reverse transcribed using Tetro Reverse Transcriptase (Bioline,

London, UK) and specific pRv primer. The reverse transcription protocol was as follows: the RNA pool and the primer were heated at 65°C for 5 min, annealed at 22°C for 5 min and extended at 42°C for 30 min followed by an extension at 50°C for 30 min and enzyme inactivation at 85°C for 5 min. The product from the reverse transcription reaction was purified by using centrifugal filter devices (Amicon Ultra-0.5 ml 10,000 MW-cutoff centrifugal filter, Millipore, Billerica, MA), quantitated using a UV spectrophotometer (OD A260/A280 ratio ≥ 1.8 and A260/230 ≥ 1.9) and submitted for Illumina sequencing (Microgem Laboratory, Naples, Italy).

Primer trimming and quality filtering of sequences

The quality of fastq files was assessed using FastQC software v0.11.2 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The paired ends were merged using Flash software v1.2.11 (https://ccb.jhu.edu/software/FLASH/) (Magoč and Salzberg, 2011). The quality filtering was performed by Fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). In details, sequences that contained more than 50% of bases with a quality score below 30 were removed. Adapters and primers were removed using Cutadapt v1.9.1 (https://cutadapt.readthedocs.io/en/stable/index.html) (Marcel, 2011). Trimmed sequences shorter than 38 nt and longer than 42 nt were also discarded, leaving only the variable region sequences for downstream analysis.

Clustering of sequences in families based on pairwise alignment and secondary structures

The frequency of each sequence in the different libraries was calculated.

Unique sequences were counted for each round, and the overall enrichment through the ten rounds was calculated as 1-(Unique/Total).

Only sequences with counts >20 in at least one of the three last rounds (12, 13, 14) were kept (16500 sequences), to remove those of poor interest.

The frequency of the four nucleotides along the variable region positions in each round was calculated on unique sequences using Fastx toolkit.

The trend of the enrichment throughout the cycles for each sequence was calculated using the slope formula:

$$b = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$$

3

The slope (*b*) represents the rate of change in *y* as *x* changes and is calculated as the covariance of *x* and *y*, divided by the sum of squares (variance) of *x*.

Sequences were ranked according to the slope values, and based on this the first 5000 sequences were selected for further analysis.

Pairwise alignment by Muscle (Edgar, 2004) implementation of msa R package v1.14.0 (Bodenhofer et al., 2015) was performed to calculate identity-based similarity score between each pair of sequence. Affinity propagation was performed, using the R package ApCluster v1.4.7 (Bodenhofer et al., 2011), on the similarity matrix to cluster the sequences in families.

The most enriched sequence for each cluster was identified, the forward and reverse primers sequences added, and the secondary structure of the ~84 nt sequences predicted by Quikfold (http://unafold.rna.albany.edu/?q=DINAMelt/Quickfold). The .ct files were obtained and a similarity matrix between secondary structures was computed through RNAsmc v0.4.0 package (https://CRAN.R-project.org/package=RNAsmc). In order to identify the best number of clusters for our data, we used the function CLvalid of the homonym R package calculating the internal validation measures for k=5:10 for k-means clustering algorithm. Then, the K-Means clustering with the suggested number of clusters was applied through the "Hartigan-Wong" algorithm with maximum 1000 interactions using the stats R package v3.5.0 (https://www.R-project.org/).

The proposed pipeline is shown in **Figure 2A.**

Development of chemoresistant cell lines

To create stable MDA-MB-231 cells chronically resistant to cisplatin (MDA-MB-231/cis) or doxorubicin (MDA-MB-231/dox), MDA-MB-231 cells were exposed to an initial concentration of 6.6 μM (cisplatin) or 0.105 μM (doxorubicin) in RPMI-1640 plus 10% FBS for 24 h. The treatment was followed by the washing of the cells three times with Dulbecco's phosphate-buffered saline (DPBS), trypsinization and splitting. The surviving population of cells was grown to 80% confluence and passaged once a week over a period of at least two weeks to ensure viability. The concentration of drugs the surviving population was exposed to was then sequentially increased to 13 μM, 18 μM, 20 μM and 25 μM (cisplatin) or 200 nM, 300 nM and 400 nM (doxorubicin).

Viability of MDA-MB-231/dox and MDA-MB-231/cis (4.0×10^3 cells/well, 96-well plates) treated for 72 h with increasing concentrations of doxorubicin (1 nM, 10 nM,

4

100 nM, 200 nM, 300 nM, 1 µM, 10 µM, 50 µM and 100 µM, diluted in cell culture medium) and cisplatin (1 µM, 5 µM, 10 µM, 15 µM, 20 µM, 25 µM, 30 µM, 40 µM, 50 µM, 70 µM and 500 µM, diluted in cell culture medium), respectively, was assessed by CellTiter 96 AQueous One Solution Cell Proliferation Assay (Promega BioSciences Inc., San Luis Obispo, CA) according to the manufacturer's instructions. Data about cell viability were plotted in Prism 6 software (GraphPad) to draw dose-response curve and to calculate half maximal inhibitory concentrations (IC50), which were compared to those of parental MDA-MB-231 cells.

Western blot

Cell lysates preparation and Western blot analyses were performed as previously reported (Camorani et al., 2015). Filters were probed with the indicated primary antibodies: anti-EGFR, anti-integrin αv (D2N5H), anti-PDGFRβ, anti-PD-L1 (E1L3N®) (Cell Signaling Technology Inc., Danvers, MA); anti-ERα (clone 1D5), anti-PR (clone PgR 636), anti-HER2 (Dako, Glostrup, Denmark); anti-α-tubulin (TU-02) and anti-vinculin (N-19) (Santa Cruz Biotechnology, Santa Cruz, CA). Densitometric analyses were performed on at least two different expositions to assure the linearity of each acquisition using ImageJ (v1.46r). Blots shown are representative of at least three independent experiments.

Cell Binding analyses

Binding of TNBC selected aptamers to intact cells was assessed by RT-qPCR-based assay, streptavidin-biotin-based (colorimetric) assay and flow cytometric assay. In all the assays, cells were incubated with RNAs diluted to the desired concentration in binding buffer (BB) consisting of BlockAid™ blocking solution (Invitrogen) with 1 mg/ml yeast tRNA and 1 mg/ml ultrapure™ salmon sperm DNA (Invitrogen), as nonspecific competitors.

RT-qPCR-based assay

RT-qPCR based assay to detect RNA binding to MDA-MB-231 cells (selection) and BT-474 and A431 cells (counterselection) was performed as previously described (Camorani et al., 2015), by using 200 nM selected pools (from round 4, 8, 11, 13 and 14), individual TNBC aptamer sequences, or TN0 starting library as a control. Prior the dilution in BB and the use in binding assays, RNAs were unfolded at 85°C for 5 min,

5

snap-cooled on ice for 3 min and allowed to warm up to 37°C. After 15 min-incubation of the cells with RNAs at 37°C, bound RNA was recovered by TRIzol (Invitrogen) and reverse transcribed as described in "High-throughput sequencing technology". The product from the reverse transcription reaction was qPCR amplified with iQ SYBR Green Supermix (Bio-Rad Laboratories, Inc., Hercules, CA) in the presence of pT7 and pRv primers reported in "Generation of the initial library". Gint4.T aptamer, added into TRIzol at 0.5 pM-final concentration, was used as a reference control, and reverse transcribed and amplified by using Gint4.T-specific 5' and 3′ primers (Monaco et al., 2017).

Streptavidin-biotin-based (colorimetric) assay

This assay was conceived to detect the aptamer binding to the cancer cell surface by using (HRP)-conjugated streptavidin. Thus, aptamers were 3'-end biotinylated by terminal transferase and Biotin-ddUTP according to manufacturing instruction (Roche Diagnostics, Indianapolis, IN). Cells were detached from culture plates with 0.02% EDTA, plated (2.0×10^4 cells/well) into clear round bottom 96-well plate (Corning Incorporate, Corning, NY) and mock-treated or incubated with increasing concentrations of TNBC aptamers or TN0 library (1 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM and 500 nM) for 10 min at room temperature (RT). After three washes with DPBS, cells were incubated with HRP-streptavidin (Bio-Rad Laboratories) for 30 min at RT. After three washes with DPBS, cells were incubated with 3,3′,5,5′-Tetramethylbenzidine substrate (Sigma-Aldrich) for 10 min at RT before quenching with an equal volume of 1 N HCl. Absorbance at 450 nm was measured by the multiskan™ FC microplate photometer (Thermo Fisher Scientific, Waltham, Massachusetts).

Flow cytometric assay

For flow cytometric assay, aptamers were internal-labeled with Alexa Fluor 647 fluorescent probe following the provider indications (Invitrogen). Cells were detached from culture plates with 0.02% EDTA and washed twice with 500 µl DPBS. Cells (2.0×10^5) were mock-treated or incubated with increasing concentrations of TNBC aptamers or TN0 library (5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM and 500 nM) for 10 min at RT. After three washes with DPBS, cells were suspended in 500 µl DPBS and analyzed by flow cytometry (BD Accuri™ C6).

6

The Streptavidin-biotin-based colorimetric assay and Flow cytometric assay were applied for binding affinity (K_d values) calculation, by using the TN0 library as a negative control to determine the nonspecific binding. The apparent K_d of aptamer-cell interaction was obtained by fitting the dependence of specific binding on the aptamer concentration to the equation

$Y = Bmax \times X / (K_d + X)$, using GraphPad Prism version 6.00.

Confocal Microscopy

MDA-MB-231 and BT-474 cells (1.0×10^5 cells/well in 24-well), previously seeded on a coverslip for 24 h, were incubated with Alexa Fluor 647-labeled TN20 or TN1 aptamer (500 nM-final concentration in BB). After 10 min incubation, at RT, cells were washed three times with DPBS and then fixed with 4% paraformaldehyde in DPBS for 20 min. Then, cells were incubated with WGA-Alexa Fluor 488 conjugate (WGA-488, Invitrogen) for 30 min at RT. Finally, after three washes with DPBS, cells were incubated with 1.5 μM 4′,6-Diamidino-2-phenylindole (DAPI, D9542, Sigma-Aldrich) and mounted with glycerol/DPBS. For internalization experiments, MDA-MB-231 cells were incubated with LysoTracker Red DND-99 (1:1000, Invitrogen) in RPMI-1640 medium supplemented with 10% FBS. After 1 h at 37°C, cells were incubated with TN20 aptamer for 30 min at 37°C and then washed, fixed and incubated with WGA-488 and DAPI, as described above. Samples were visualized by Zeiss LSM 700 META confocal microscopy equipped with a Plan-Apochromat 63x/1.4 Oil DIC objective.

Target-type analysis

MDA-MB-231 cells (5.0×10^6) were washed with DPBS and then incubated with 0.1 mg/ml proteinase K (Invitrogen) in DPBS at 37 °C for 10 min. FBS was then added to quench the proteinases. After washing with DPBS, the treated cells were used for aptamer-binding assay as described in "Flow-cytometric Assay" by using TNBC aptamers at 200 nM-final concentration.

Histological tissue sections staining with aptamers

TNBC patients selection and tissue microarray (TMA) building

From 2008 to 2009, 18 patients who underwent a tumorectomy, quadrantectomy or mastectomy at the National Cancer Institute IRCCS Fondazione "Giovanni Pascale" of Naples, Italy, were selected from the Pathology Unit and enrolled into this study. All samples were reviewed according to WHO classification criteria, as previously reported (Collina et al., 2016). The cohort included patients with an average age of 55 years, ranging from 33 to 80. The samples were invasive ductal carcinomas, 8 (44.4%) grade II and 10 (55.6%) grade III. Tumor size was lower than 2 cm in 58.8% (10/17) of the samples, between 2 and 5 cm in 35.3% (6/17) and larger than 5 cm in 5.9% (1/17). For 1 case, we were not able to retrieve this information. Metastatic lymph nodes (LNM) were found in 35.3% of patients (6/17) at surgery. This information for 1 patient was lost. Distant metastases were found in 27.5% of patients (3/4). For 14 cases we were unable to recover this information.

TMA was built using five tissue cores for each sample, taken from the periphery to the center of the tumor. Tissue cylinders with a diameter of 1 mm were punched from morphologically representative tissue areas of each 'donor' tissue block and brought into one recipient paraffin block (3 × 2.5 cm) using a semi-automated tissue arrayer (Galileo TMA).

Immunohistochemical staining (IHC) was done on slides from formalin-fixed, paraffin-embedded tissues from a TNBC TMA, to evaluate the expression of ER, PR, HER2 and Ki-67 markers. All details of IHC analysis was previously reported (Collina et al., 2016). The expression of proliferation factor Ki-67 was high (>20%) in 83.3% (15/18) and low (≤20%) in 16.7% of cases (3/18). The main clinicopathological data of the cases are set out in **Table S3**.

Aptamer-based staining

The tissue sections were deparaffinized in xylene and rehydrated through graded alcohols. For antigen retrieval, tissue sections were incubated in 10 mM sodium citrate buffer (pH 6.0) at 95°C for 20 min. After two washes with DPBS containing 0.05% Tween-20 (pH 7.4), tissue sections were incubated in 3% H_2O_2 for 15 min. To mask endogenous biotin binding, sections were treated with biotin-blocking solution (BioLegend, San Diego, CA, USA) following the manufacturer's instructions and then with BlockAid™ blocking solution (Invitrogen) for 40 min and with Ultra V Block Solution (Thermo Fisher Scientific) for 5 min to minimize non-specific antibody

8

binding. Finally, tissue sections were probed with biotin-labeled TNBC aptamer (0.25 pmol/µl) in DPBS, in the presence of an excess (5 pmol/µl) of unlabeled starting library, 65 ng/µL sheared salmon sperm DNA with formamide and 10% BlockAid as blocking agents, for 30 min at 23°C, as detected by air conditioner control in the laboratory. Slides were washed with DPBS and incubated with HRP-conjugated streptavidin, washed with DPBS, incubated with the peroxidase substrate, washed with DPBS, mounted, and analyzed by light microscopy. Cells specifically recognized by the aptamer were stained in brown. The intensity, extent and subcellular distribution of aptamer-staining was evaluated by an expert pathologist. For each sample, at least five fields and >500 cells were analyzed. To classify aptamer-staining we used the combo scores which are obtained by multiplying cytoplasmic/cell surface staining intensity score (based on 4 point system: 3=high, 2=moderate, 1=low, 0=absent staining) with percentage of stained cells (0=negative, 1=1-30%, 2=31-60, 3>60%). We have considered four final scores to stratify the samples: negative (combo score=0), low (combo score=1), moderate (combo score=2-4) and high (combo score=6-9).

Mammosphere formation assay

MDA-MB-231 and BT-549 cells (1.0×10^4 cells/well) were plated in 24-ultralow attachment plates (Corning Incorporate) and grown in DMEM/F-12 serum-free medium (Sigma-Aldrich), supplemented with B27 (1:50, Invitrogen), 20 ng/ml basic fibroblast growth factor (bFGF, Sigma- Aldrich) and 20 ng/ml EGF, with or without aptamer (400 nM-final concentration). Before the treatment, the aptamers were subjected to a short denaturation-renaturation step as reported in "RT-qPCR-based assay". The aptamer treatment was renewed every 24 h. After 4 days, the size and number of spheres were analyzed under a phase-contrast microscopy (Leica DMI3000 B apparatus).

All spheres with a diameter >40 μm were analyzed, and the number and size of spheres were quantitated in at least 10 fields per condition. At the end of treatments, mammospheres were collected by gentle centrifugation (1000 rpm) and RNA was recovered using TRIzol reagent and then processed for RT-qPCR as previously described [24] by using the following primers for CD44 expression: Fwd 5′-TTCCCAAAAAGAGGCTGAGA-3′, Rev 5′-CAATGT TGCAAGGGTTTGTG-3′. β-actin, Fwd 5′-CAAGAGATGGCCACGGCTGCT-3′, Rev 5′-TCCTTCTGCATCCTGTCGGCA-3′ was used for internal control.

9

TN2 aptamer stability in human serum

Human serum stability of TN2 aptamer was evaluated by incubating 5 μM 2'F-Py RNA aptamer in DPBS containing 80% human serum (Sigma-Aldrich) at 37°C up to 48 h. RNA (50 pmol) from all time points was incubated for 1 h at 37°C with 100 μg proteinase K solution (Invitrogen), in order to remove serum proteins that interfere with electrophoretic migration, and then mixed with formamide RNA loading dye and stored at -80°C. All samples were resolved by denaturing PAGE (12% polyacrylamide with 7 M urea). The gel was stained with ethidium bromide and UV exposed to visualize the RNA bands.

Data and Code Availability

The accession number for Illumina Next Generation Sequencing raw data reported in this paper is NCBI SRA: PRJNA558191.

Statistical analysis

All statistical values were defined using GraphPad Prism version 6.00 by unpaired *t*-test (two variables) or one-way ANOVA followed by Tukey's multiple comparison test (more than two variables). *P* value < 0.05 was considered significant for all analyses.

Supplemental References

Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics. 31, 3997-3999.

Bodenhofer, U., Kothmeier, A., and Hochreiter, S. (2011). APCluster: an R package for affinity propagation clustering. Bioinformatics. 27, 2463-2464.

Camorani, S., Crescenzi, E., Colecchia, D., Carpentieri, A., Amoresano, A., Fedele, M., Chiariello, M., and Cerchia, L. (2015). Aptamer targeting EGFRvIII mutant hampers its constitutive autophosphorylation and affects migration, invasion and proliferation of glioblastoma cells. Oncotarget. 6, 37570-37587.

Camorani, S., Esposito, C.L., Rienzo, A., Catuogno, S., Iaboni, M., Condorelli, G., de Franciscis, V., and Cerchia, L. (2014). Inhibition of receptor signaling and of glioblastoma-derived tumor growth by a novel PDGFRβ aptamer. Mol Ther. 22, 828-841.

Collina, F., Cerrone, M., Peluso, V., Laurentiis, M.D., Caputo, R., Cecio, R.D., Liguori, G., Botti, G., Cantile, M., and Bonito, M.D. (2016). Downregulation of androgen receptor is strongly associated with diabetes in triple negative breast cancer patients. Am J Transl Res. 8, 3530-3539.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792-1797.

Magoč, T., and Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 27, 2957-2963.

Marcel, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. 10.14806/ej.17.1.200.