International Journal of Scientific Reports
Gaudeau A et al. Int J Sci Rep. 2021 Apr;7(4):196-206
http://www.sci-rep.com
pISSN 2454-2156 | eISSN 2454-2164
DOI: https://dx.doi.org/10.18203/issn.2454-2156.IntJSciRep20211035

Original Research Article

Cell-based siRNA screens highlight triple-negative breast cancer cell epigenetic vulnerability

Albane Gaudeau[1], Coralie Clua Provost[1], Thierry Dorval[1], Andrew Walsh[2], Michael Hannus[2], Franck Perez[3], Jacques Camonis[4], Elaine Del Nery[3,4], Jean-Philippe Stephan[1]*

[1]Center of Excellence Pharmacological Screening, Compound Management and Biobanking, Institut de Recherches Servier, Croissy-sur-Seine, France
[2]siTOOLs Biotech, Planegg/Martinsried, Germany
[3]BioPhenics Laboratory, Translational Research Department, Institut Curie
[4]Institut Curie, Centre de Recherche, Paris Sciences et Lettres Research University, Paris, France

Received: 01 November 2020
Revised: 07 February 2021
Accepted: 08 February 2021

***Correspondence:**
Dr. Jean-Philippe Stephan,
E-mail: jean-philippe.stephan@servier.com

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: Triple-negative breast cancer (TNBC) is a heterogeneous disease defined by ER-, PR- and HER2-negative phenotype and in most cases, a relatively aggressive clinical behaviour. The lack of specific targeted therapies and low efficiency of currently available chemotherapies spurred several clinical trials in the last few years. Despite encouraging results, TNBC still remains a major unmet medical need that prompted us to explore the role of 863 epigenetic modulators in TNBC cell survival.
Methods: A comprehensive siRNA library was screened to explore the role of known epigenetic modulators in TNBC cell viability and growth. The knock-down effect was evaluated for 863 epigenetic genes using 4 siRNAs/gene in two TNBC and a non-TNBC cell lines using ATP-based luminescence and nuclei count image-based assays. Considering siRNA off-target effects, four analysis methods including a classical threshold-based analysis and three ranking methods were applied to determine on-target hits for each screen readout. Hit genes common to both phenotypic readouts highlighted strong epigenetic players involved in TNBC cell survival.
Results: Overall, knock-down of many epigenetic modulator genes mitigates cell survival in TNBC and a non-TNBC cell lines depicted from both phenotypic readouts. Interestingly, ranking-based analysis confirmed hit genes identified in threshold-based analysis and also revealed additional hits enabling us to confirm CDK1 and KMT5A as important regulators in TNBC cell viability and growth. Surprisingly, CHAF1A appeared as a new candidate gene involved in TNBC cell survival.
Conclusions: Taken together, siRNA epigenetic screening results identified CHAF1A as a novel regulator of TNBC cell survival.

Keywords: TNBC, Epigenetics, siRNA, High-throughput screening

INTRODUCTION

Breast cancer is the most prevalent malignancy in women worldwide. Therapeutic progresses enable 70-80% of patients with early-stage, non-metastatic disease to be successfully treated today.[1] However, advanced cancers with distant organ metastases and specific forms of breast cancers still display poor prognosis. Although luminal (Estrogen receptor (ER) and/or progesterone receptor (PR) positive) and human epidermal growth factor receptor 2 (HER2) overexpressing tumors benefit from a rather favourable prognostic with hormone or targeted

therapies, respectively, TNBC representing 15-20% of all breast cancers, remains a significant unmet medical need.

Aggressiveness of TNBC relates to highest rates of metastasis with aberrant epithelial-to-mesenchymal transition (EMT) and stem cell signalling pathway activities characterized by aberrant activity of epigenetic mechanisms, which in turn triggered strong interest in the discovery of small molecules targeting specific epigenetic modulators.[2-5] FDA-approved drugs already target DNA methyltransferases (DNMT), histone acetyltransferases (HAT), histone deacetyltransferase (HDAC) and protein methyltransferases (PMT).[6] Among ongoing clinical trials evaluating epigenetic drug candidates, valproic acid (VPA) was described as a potent HDAC inhibitor able to reprogram both ER-positive and-negative malignant mammary epithelial cells to a more differentiated and physiologic phenotype that may improve the sensitivity to endocrine therapy and/or chemotherapy in breast cancer patients.[7] More recently, Entinostat, another HDAC inhibitor initially shown to be able to reverse the EMT phenotype through reversal of epigenetic repression of E-cadherin, was demonstrated to reduce the percentage of tumor-initiating cells from TNBC cells.[8,9] Given that TNBC tumors are enriched in CSCs, and epigenetic processes control multiple critical aspects of TNBC cell biology, we investigated a large collection of siRNA targeting epigenetic modulators on TNBC and non-TNBC cell viability aiming at identifying epigenetic modulators mitigating TNBC cell survival.[10-12] We used two different TNBC cell lines, MDA-MB-468 and MDA-MB-231, and the HCT116 colorectal cancer cell line on two distinct cell viability assays: (i) the ATP-based luminescence cell viability readout and (ii) the image-based nuclei count assay upon nuclear staining. Finally, we undertook a comprehensive data analysis pipeline using four different methods: threshold-based, and three ranking-based algorithms-Redundant siRNA Activity (RSA), double-hit, and triple-hit methods-specially designed for siRNAs screening analysis, ultimately only considering hits identified through at least three of them. Our investigations confirmed several epigenetic modulators previously identified as targets of interest for TNBC and identified new candidate gene involved in TNBC cell survival.

METHODS

Cell lines

MDA-MB-468 (HTB-132, ATCC) and MDA-MB-231 (HTB-26, ATCC) TNBC cell lines and HCT116 (CCL-247, ATCC) colorectal cancer cell line were cultured in RPMI 1640 media supplemented with GlutaMAX (61870, Life technologies), 10% fetal Bovine serum (Lonza) and 1% HEPES 1M (H0887, ThermoFisher), in a humidified atmosphere containing 5% CO2 at 37°C. In these culture conditions, MDA-MB-468, MDA-MB-231 and HCT116 cell lines have doubling times equal to roughly 28, 36 and 18 hours (h), respectively. Each cell line was validated through short tandem repeat (STR) profiling (Table 1).

Table 1: Short tandem repeat (STR) profiling for MDA-MB-231, MDA-MB-468 and HCT116 cell lines.

Short tandem repeat	MDA-MB-231	MDA-MB-468	HCT116
Amelogenin	X, X	X, X	X, Y
CSF1PO	12, 13	12, 12	7, 10
D12S391	17, 18	18, 18	NA
D13S317	13, 13	12, 12	10, 12
D16S539	12, 12	9, 9	11, 13
D18S51	11, 16	17, 17	17
D19S433	11, 14	12, 13	NA
D1S1656	15, 17	11, 14	NA
D21S11	30, 33.2	27, 28	29, 30
D2S1338	20, 21	17, 25	NA
D3S1358	16, 16	15, 15	12, 19
D5S818	12, 12	12, 12	10, 11
D6S1043	18, 18	13, 23	NA
D7S820	8, 9	8, 8	11, 12
D8S1179	9, 13	13, 13	12, 14
FGA	22, 23	23, 23	18, 23
Penta D	11, 14	8, 10	9, 13
Penta E	11, 11	5, 5	13, 14
TH01	7, 9.3	7, 7	8, 9
TPOX	8, 9	8, 9	8
vWA	15, 18	18, 18	17, 22

Epigenetic siRNA library

The siRNA library was purchased from Dharmacon (GU-006107-E2-01 further completed with siRNA reagents targeting additional genes). The library includes 863 epigenetic modulators, each one being targeted by four individual siRNAs. The 3,452 siRNAs were split into 13 384-well plates. 7 PLK1 siRNAs and 7 non-targeting control siRNAs were used as controls distributed in all plates.

siRNAs were screened through reverse transfection at 12.5 nM final concentration in 80 µL final volume per well. 200 nL of each siRNA were spotted in 384-well screening plates (6057300, Perkin Elmer) before adding the transfection agent and the cells. To avoid position effects, for each gene, the 4 siRNAs were randomly scattered across different plates at different well positions.

siRNA transfection

Real time cell confluence monitoring

Cell seeding density was optimized to reach 60 to 80% confluence after 96h incubation. Real-time cell confluence monitoring was performed using the Incucyte platform (Essen BioScience) acquiring cell images every 2 h for 96 h. Cell confluence was determined at every

time point for every well on the screening plates, using IncucyteZOOM analysis software.

CellTiterGLO assay

Transfection conditions were optimized using an siRNA targeting the gene polo-like kinase 1 (PLK1) (D-003290-08-0005, Dharmacon) and the non-targeting control siRNA (D-001210-02-05, Dharmacon) designed to target no known genes in human, mouse or rat. The optimum transfection conditions were identified based on the maximal and minimal impact of PLK1 siRNA (positive control) and non-targeting control (negative control) siRNA on cell viability, respectively, compared to the non-transfected cells condition. Cell viability was assessed by quantifying ATP levels after 96 h treatment using the CellTiterGLO (CTG) assay (Promega). Luminescence signal was acquired on PheraSTAR (Cysbio) reader, resulting in relative luminescence units (RLU) value per well. The different cell lines were transfected by lipofection using the DharmaFECT 4 lipid (T-2004-02, Dharmacon). MDA-MB-468, MDA-MB-231 and HCT116 cell lines were transfected using 0.04, 0.05 and 0.03 µL of DharmaFECT 4, diluted in 20 µL of Opti-MEM (11058-021, Gibco) and added to 200 nL of siRNAs, respectively. After 15 minutes (mn) incubation, 60 µL of cell suspension containing 2000 MDA-MB-468 or MDA-MB-231 or 500 HCT116 cells were added per well. Cells were then incubated in a humidified atmosphere containing 5% CO2 at 37°C for 96 h, replacing 40 µL of medium with fresh medium 24 h after transfection.

Fluorescence-based nuclei count

Cell count per well was determined 96h following siRNA transfection, using an automated fluorescence-based nuclear staining assay followed by image acquisition and automated cell count analysis. Cells were washed one time with PBS (14190, Gibco) and incubated in 40 µL 4% PFA (18814, Tebu-bio) solution for 15 mn. After one wash, 40 µL NH4Cl 1M (A9434, sigma-Aldrich) per well were used to quench the PFA. Cells were then washed one more time and incubated with 40 µL 1% BSA (A7979, sigma-Aldrich)/PBS solution for 15 mn. Cells were then incubated for 1h in 30 µL 1/1000 Hoechst (H3570, Invitrogen)/ 1% BSA/PBS. After three final washes, cells were then stored in 40 µL 1% PFA/PBS solution at 4°C away from light. Cell images were then acquired on the high-content platform Opera Phenix (Perkin Elmer) at 10X magnification. Following image acquisition, images were exported in Columbus (Perkin Elmer) to detect and count nuclei.

Data analyses

RLU and nuclei count values were analyzed using the open-source software HCS Analyzer.[13] First, a quality-control was applied to each plate to validate the controls before analyzing siRNA values. This step consisted in calculating the z' factor between the negative (non-targeting control siRNA) and positive (PLK1 siRNA) control values as follows:

$$Z' = 1 - \frac{3\,\sigma\,\mathrm{NC} + 3\,\sigma\,\mathrm{PC}}{|\mu\,\mathrm{NC} - \mu\,\mathrm{PC}|}$$

Where σ means the standard deviation, NC means negative controls, PC means positive controls, and μ means arithmetic mean.

Plates with z'<0 were removed from the subsequent data analysis.

Subsequently, raw data were normalized into z-score for each plate using samples and calculated according to the following formula:

$$z_{score} = \frac{Xi - Xp}{\sigma p}$$

Where X means the value of the sample i, or the mean of all plate sample values p, excluding the controls.

Threshold-based and Ranking-based approaches were then used for hits identification.

Threshold-based approach

A threshold was determined as follows:

$$T = \mu\,PC + 3\,\sigma\,PC$$

If z score<T, the siRNA sample was considered as a potential hit. Then, genes were grouped depending on the number of siRNA hits. A gene is identified as a hit if 2 or more out of 4 siRNAs were siRNA hits.

Ranking approaches

Ranking approaches consisted in ranking genes by order in which they are on-target hits depending on the algorithm used. For each of the six screening data sets, we worked with 3 different algorithms resulting in 3 distinct rankings.

The first ranking approach is the redundant siRNA activity (RSA) analysis based on the hypergeometric distribution. First, all z scores values from the 3452 siRNAs assayed are ranked in ascending order (low is good). In this way, each gene is represented at 4 different positions in the list. Then, at each position, RSA tests the probability of finding the given on-target siRNAs (from 1 to 4) in the current list. Therefore, a p-value is assigned to each siRNA, so that there are 4 p values per gene. Finally, the lowest of the 4 p values is assigned to the gene, enabling to rank all 863 genes by ascending order. Here, RSA enables to rank genes by enrichment of their siRNAs towards the top of the ranked list.

The other two algorithms, developed by siTOOLs biotech, are based on a multi-hit analysis, ranking genes by order in which they become double- or triple-hits. First, similarly to RSA, all z score values from the 3452 siRNAs are ranked by ascending order. For the double-hit ranking, once a second siRNA appears in the list, the corresponding ranking position is assigned to the targeted gene. This way, each of the 863 genes have a value, which is used to rank genes by ascending order in which they become double-hits. For the triple-hit ranking, the third siRNA ranking position was considered.

Considering that ranking-based siRNA analysis regards the top 30 genes as true-positive, we focused on the top 30 genes for each of the 3 ranking methods.

On-target hits selection

For each screen, 4 hit lists were obtained: 1 list composed of threshold-based hit genes and 3 lists composed of the top 30 genes ranked by the 3 different ranking-based algorithms. For each gene in each list, hit genes were ultimately considered if they were identified in at least 2 independent hit lists.

TNBC-specific genes selection

The final analysis step focused on removing non-specific hit genes involved in broader cell toxicity. Our final hit list includes epigenetic modulators mitigating both TNBC cells, MDA-MB-468 and MDA-MB-231, but not HCT116 cell line.

RESULTS

siRNA screening datasets

Experimental conditions were first optimized for MDA-MB-468, MDA-MB-231 and HCT116 cell lines (Figure 1). Our extended epigenetic siRNA library, targeting 863 epigenetic modulators using 4 different siRNAs per gene, was used to screen MDA-MB-468 (Figure 2: A, B and C), MDA-MB-231 (Figure 2: D, E and F) and HCT116 (Figure 2: G, H and I) cell lines for two independent readouts, CellTiterGLO and nuclei count. Ultimately, results from the six screens were processed according to the z-score normalization method. The distribution of normalized z-scores for each siRNA, including controls, are presented Figure 2 for each readout.

These results suggest major differences between the two readouts, and between cell lines when using Nuclei count as a viability readout compared to CellTiterGLO readout. Indeed, CellTiterGLO and Nuclei count are two different cell viability assays, one measuring a homogeneous signal, and the other one detecting signal from individual cells, resulting in more homogeneous and heterogeneous results, respectively (Figure 2).

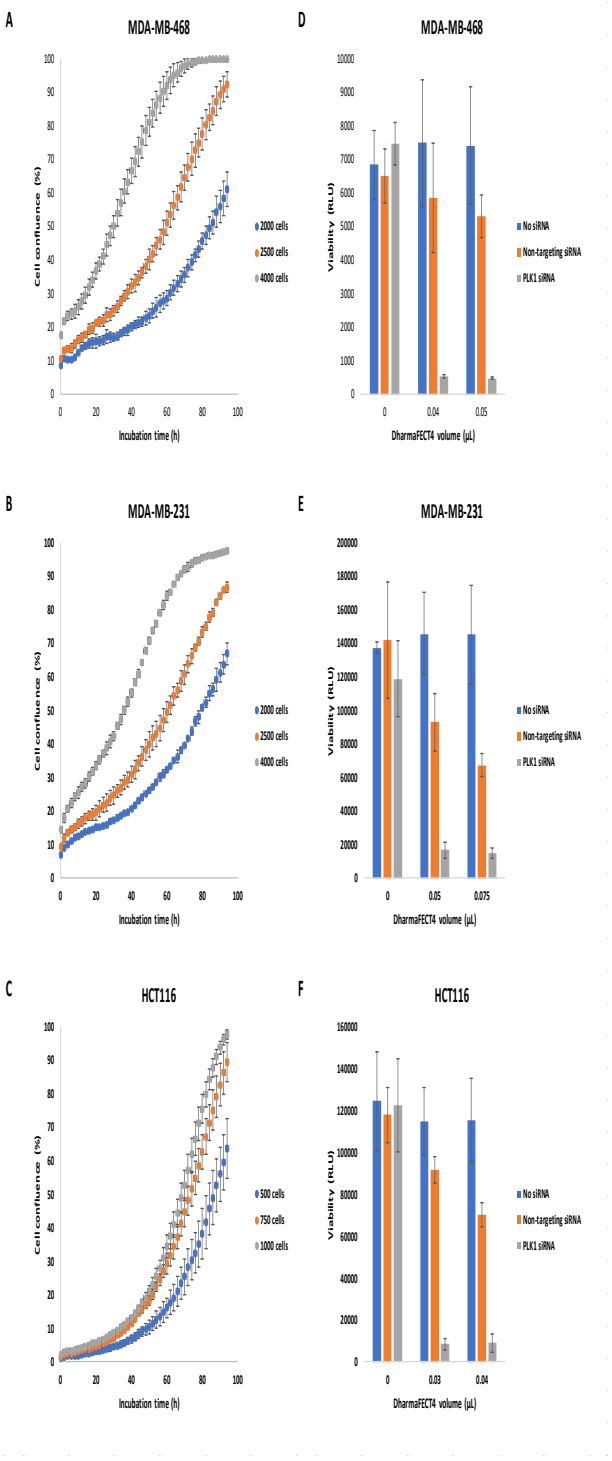

Figure 1: siRNA transfection optimization for the different cell lines. Cell confluence over time for MDA-MB-468 (A), MDA-MB-231 (B) and HCT116 (C) cell lines seeded at different cell density in 384-well plate. Cell viability for MDA-MB-468 (D), MDA-MB-231 (E) and HCT116 (F) cell lines after 96 h treatment with various amounts of DharmaFECT4. Cells were exposed to non-targeting control siRNA, PLK1 siRNA or media alone as non-transfected condition. Cell viability was expressed as relative luminescence units (RLU) based on the CellTiterGLO readout.

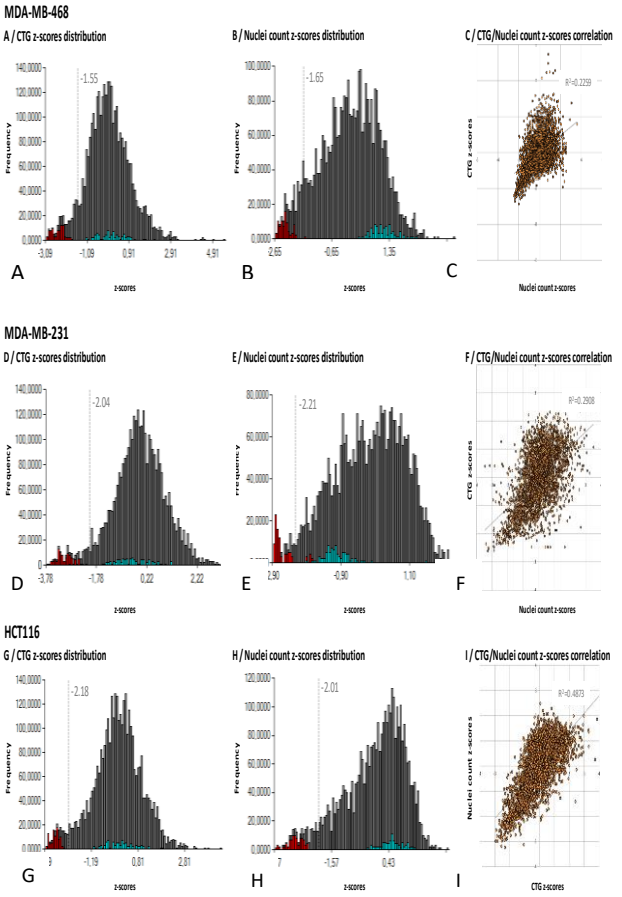

Figure 2: Epigenetic siRNA library screening results using two independent cell viability readouts. Screening data for the MDA-MB-468 (A, B and C), MDA-MB-231 (D, E and F) and HCT116 (G, H and I) cell lines using CellTiterGLO readout (A, D and G) or Nuclei count based on nuclear fluorescent staining readout (B, E and H). Correlation between the two readouts for MDA-MB-468 (C), MDA-MB-231 (F) and HCT116 (I). Z score results for siRNAs targeting epigenetic modulator are shown in dark grey, while z score results for positive and negative controls are shown in red and blue, respectively. Hits selection thresholds are visualized by a bar for each cell line and readout.

siRNA hits identification

Threshold analysis

Threshold values were calculated for each cell line and readout (Figure 3). siRNAs with normalized z-scores lower than the threshold were considered as siRNA hits. Only siRNAs targeting the intended genes by at least 2 out of the 4 siRNAs were considered for the subsequent data analysis. Figure 3 shows the number of genes in which 2 to 4 siRNAs were identified as hits using the threshold method. Ultimately, 5 to 27 genes were identified in all screens, while MDA-MB-231 Nuclei count screening did not reveal any 2 out of 4 siRNAs hit genes using the threshold analysis.

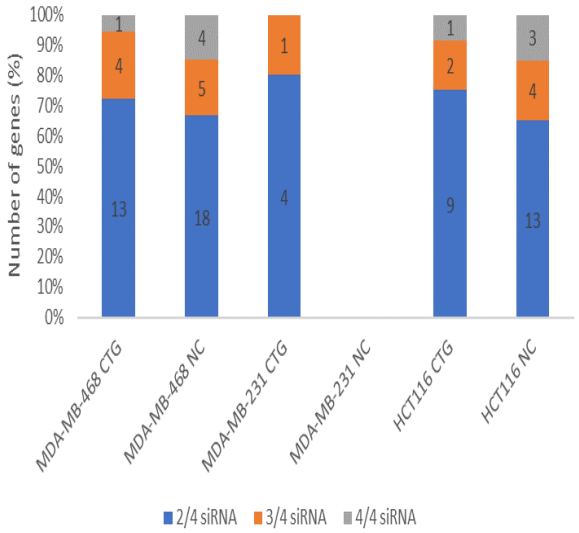

Figure 3: Number of genes identified based on the number of siRNA hits using the threshold method for each cell line screen. For the MDA-MB-231 nuclei count readout, no genes were identified with at least 2 siRNA hits. CTG: CellTiterGLO; NC: nuclei count.

RSA analysis

In this approach, the correlation between normalized z scores for siRNAs randomly paired by genes was investigated for viability screens performed on MDA-MB-468, MDA-MB-231 and HCT116 either using the CellTiterGLO or the nuclei count readouts (Figure 4). CellTiterGLO or nuclei count normalized z scores for randomly paired siRNAs targeting the same genes are overall not correlated, highlighting the impact of the off-target effect previously described for siRNA.

Likewise, scrambled data p-values were also calculated based on a random pairing of the siRNA sequences to z-score values. Finally, all the gene p values were then ranked by ascending order (Figure 5) for both real and scrambled screening data sets, for CellTiterGLO (Figure 5A) and Nuclei count (Figure 5B) readouts performed in MDA-MB-468, MDA-MB-231, and HCT116 cells. Based on these rankings, the real and scrambled p value screening data sets diverge at different ranking positions depending on the cell line and the readouts. This representation demonstrates that genes with the lowest real p values, inferior to scrambled p-values, are truly active in our screens. In these conditions, the top 30 RSA genes were analyzed based on their CellTiterGLO (Figure 6A) and nuclei count (Figure 6B) RSA p values and individual z score values for each of the four siRNAs targeting these genes in the MDA-MB-468, MDA-MB-231 and HCT116 cell lines. As expected, genes with the lowest RSA p values, such as PRPF31 for MDA-MB-468, also have the lowest CellTiterGLO and Nuclei count z scores and the tightest data distribution between the four siRNAs per gene.

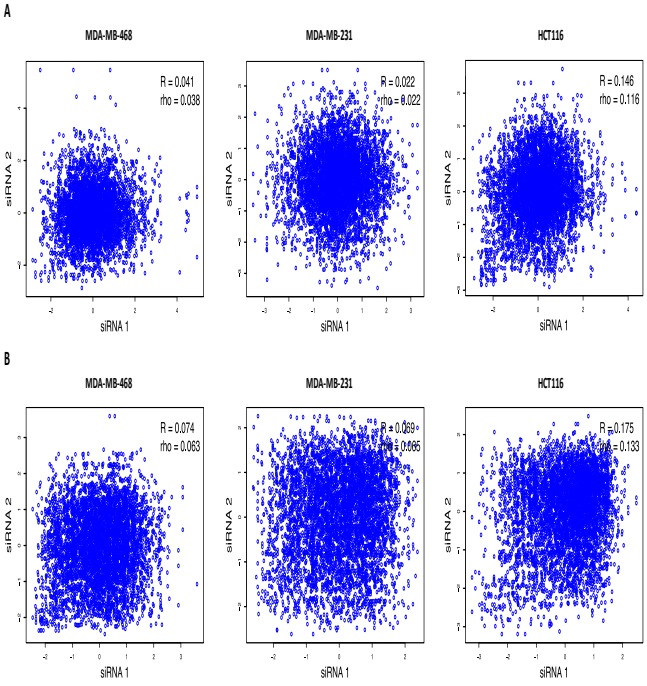

Figure 4: Data correlation for randomly paired siRNAs targeting the same genes in MDA-MB-468, MDA-MB-231 and HCT116. (A) Correlation between CellTiterGLO z-score values for randomly paired siRNAs targeting the same genes. (B) Correlation between Nuclei count z score values for randomly paired siRNAs targeting the same genes.

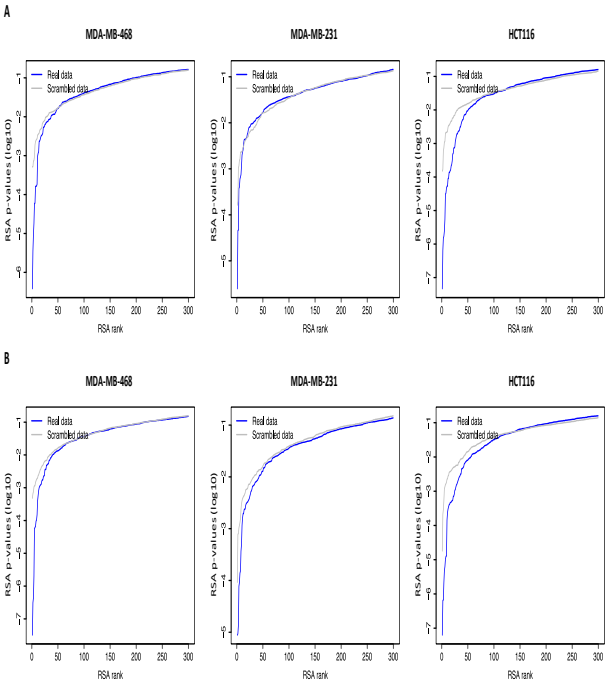

Figure 5: Top 300 RSA real gene p values ranked by ascending order compared to scrambled p values calculated for CellTiterGLO (A) and nuclei count (B) readouts performed in MDA-MB-468, MDA-MB-231, and HCT116 cells.

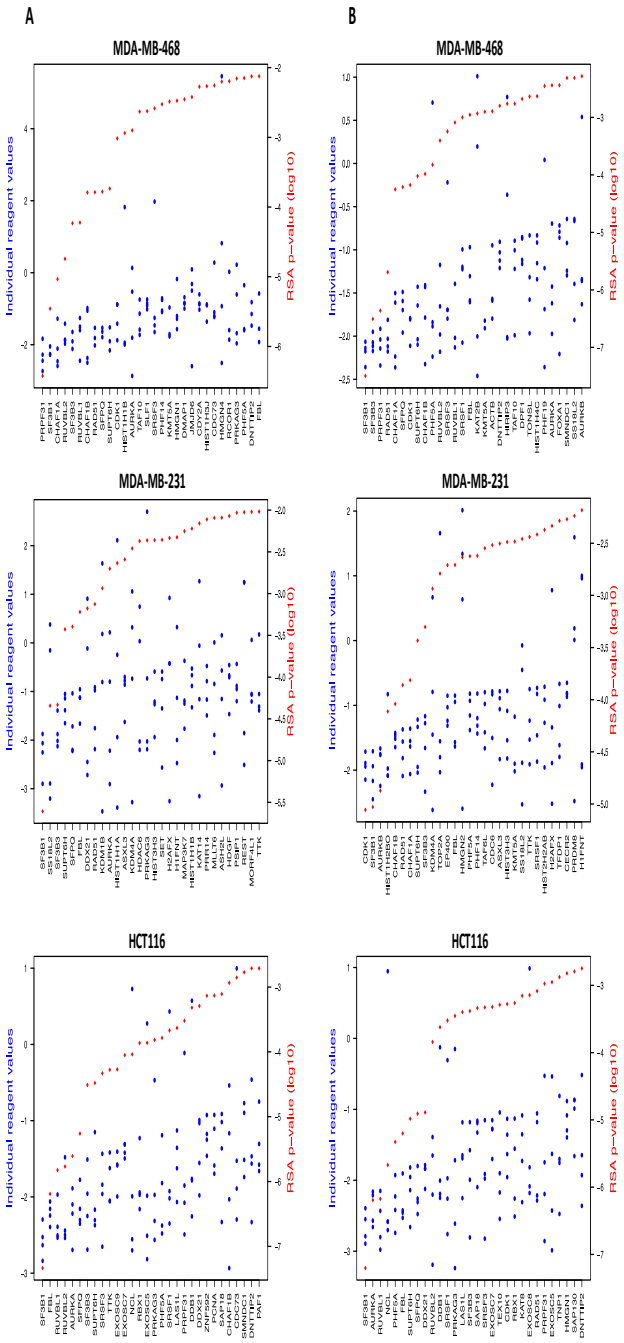

Figure 6: Hit's identification based on RSA ranking. Top 30 genes ranked by RSA p values (log10) for MDA-MB-468, MDA-MB-231 and HCT116 screens using CellTiterGLO (A) and nuclei staining (B) readouts, including individual siRNA z score values for the four siRNAs.

Multi-hit analysis

As for the RSA ranking, CellTiterGLO and Nuclei count z-scores were ranked in ascending order for the multi-hit analyses. Ranking based on real data was then compared to the one obtained when the siRNAs were randomly paired to z score values (scrambled data). The cumulative number of double-hit genes at every position for the top

500 siRNAs compared to scrambled data is presented for the CellTiterGLO and nuclei count readouts Figure 7 (A and B), respectively. As for the RSA analysis, the HCT116 cell line displays the largest separation between real and scrambled data double-hits curves for both CellTiterGLO and Nuclei count readouts. Real and scrambled data sets are slightly less separated for MDA-MB-468, merging near the top 350 and 500 siRNAs for CellTiterGLO and Nuclei count data sets, respectively. For the MDA-MB-231, the real and scrambled data double-hits curves almost overlap for both readouts, with an enrichment in real data further in the siRNA ranking, for the nuclei count readout.

Relevant genes identification based on the different screening readouts and data analysis methods

To identify epigenetic modulators critical for TNBC cell lines viability, hit genes were compared between the two implemented readouts and the four methods considered to analyze the data. The latter were first compared to determine hit genes that were identified by at least three analysis methods, considering hit genes defined as at least 2 out 4 siRNA hits for the Threshold approach and being in the top 30 genes for the ranking algorithms (Figure 8).

Interestingly, the RSA ranking method identified the same hit genes than the double-hit ranking approach regardless of the cell line, for both readouts. Globally, most hit genes identified were commonly identified across the data analysis methods employed. This is the case for 12 hit genes in MDA-MB-468 using CellTiterGLO and 14 hit genes using the nuclei count readout. In MDA-MB-231, 3 hit genes were common to the four methods using CellTiterGLO readout while all methods, except for the threshold analysis, generated 13 hit genes using the nuclei count readout. Twelve hit genes were common to the four methods in the HCT116 cell line using CellTiterGLO and 17 hit genes using the nuclei count readout. Only one gene, SF3B1, was identified through the four methods for all screens regardless of the cell line or the readout, except in the MDA-MB-231 nuclei count that did not generate any hit genes through the Threshold approach. Interestingly, SF3B3 was identified as a hit gene by threshold in MDA-MB-468 using both readouts, and in HCT116 using CellTiterGLO, while it was identified by all ranking methods in all screens. These results indicate that ranking approaches confirm and further complement data from the threshold analysis.

The ranking analysis systematically identified more common hit genes than the threshold method, regardless of the readout or the cell line. As expected, the triple-hit ranking method identified slightly fewer hit genes than the double-hit ranking method regardless of the readout and the cell line, except for the HCT116 cell line with the CellTiterGLO where the two methods performed similarly. Ultimately, between 9 and 22 genes were identified as potential hits depending on the cell line and the readout considered.

Subsequently, our analysis consisted in highlighting common hit genes across the two readouts for each cell line, and ultimately, identifying common hit genes across the three cell lines. As shown in Figure 9, genes identified by threshold or ranking (Figure 9A) are either common to both cell viability readouts, either specific to CellTiterGLO or nuclei count. Globally, numerous hit genes are common to the two readouts, except for MDA-MB-231, for which only 3 genes were common to both readouts. As MDA-MB-231 cells display the slowest doubling time (36 h) among the three cell types, cell incubation for 96 h was probably not sufficient to allow complete biological modulation in response to siRNA and identification of key genes in cell viability. Moreover, for all three cell lines, number of hit genes are slightly higher in nuclei count compared to CellTiterGLO assay, suggesting that more hits are cytostatic. Among hit genes common to both readouts, genes common to both threshold and ranking analyses were identified (Figure 9B). Six are common to the MDA-MB-468 and HCT116 cell lines, five are specific to HCT116, and six are specific of the MDA-MB-468 cell line. Identified through both readout assays and both data analysis approaches, these genes are supposed to be strong candidates involved in cell viability.

Finally, hit genes were analyzed across the different cell lines (Figure 9C), regardless of the readouts or the data analysis methods used to identify them. A total of 46 genes were identified as impacting cell viability of the different cell lines investigated. Amongst these genes, 10, 8 and 11 were specific for the MDA-MB-468, MDA-MB-231 and HCT116 cell lines, respectively. Ultimately, 3 hit genes were identified as being specific for the two TNBC cell lines, MDA-MB-468 and MDA-MB-231: CDK1, CHAF1A and KMT5A.

The role of CDK1 in the TNBC cell viability was highlighted in both readouts for the MDA-MB-468 cell line. Indeed, CDK1 was identified by at least 3 analysis methods using CellTiterGLO. Using the Nuclei count, it is a 3 out of 4 siRNAs hit using the Threshold method and, even though it was not confirmed by at least two ranking approaches, CDK1 was identified in the top 30 double-hit analysis and at the position 32 in the RSA rank. In the MDA-MB-231, CDK1 was identified by the three ranking methods using the nuclei count readout. For CHAF1A and KMT5A, both genes were identified with the four analysis methods in the MDA-MB-468 cell line using both readouts. Both hit genes were also identified with the three ranking approaches in the MBA-MB-231 cell line using the nuclei count readout and as 3 out of 4 and 2 out of 4 siRNAs hit for CHAF1A and 2 out of 4 and 3 out of 4 siRNAs hit for KMT5A with the threshold analysis using the CellTiterGLO and the nuclei count readouts, respectively.

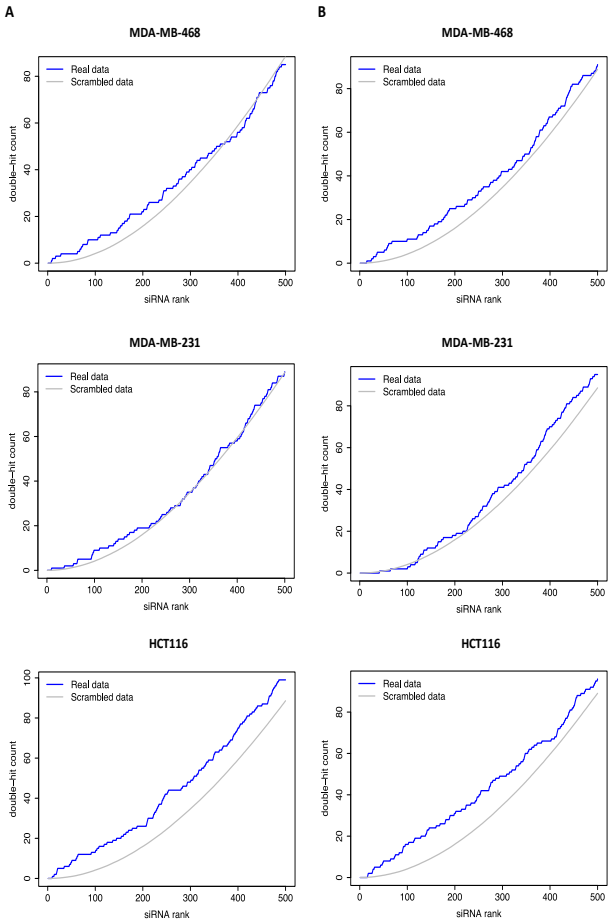

Figure 7: Double-hit enrichment at every siRNA ranking position based on ascending CellTiterGLO (A) and nuclei count (B) z scores or scrambled data for MDA-MB-468, MDA-MB-231 and HCT116 cell lines.

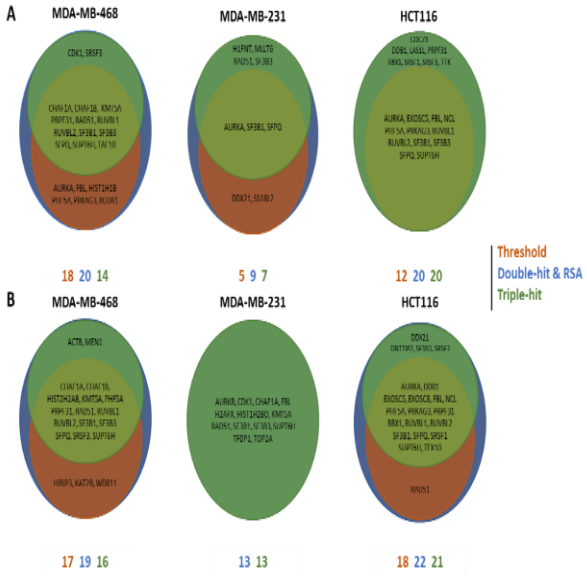

Figure 8: Hit genes identified through each analysis method across each of the three cell lines using either CellTiterGLO (A) or nuclei count (B) readouts.

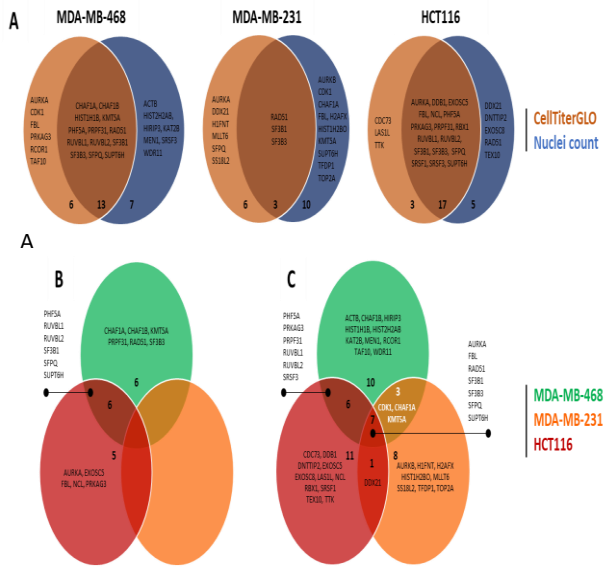

Figure 9: Genes identified in each cell line common to both CellTiterGLO and nuclei count readouts (A) and to both threshold and ranking analyses (B), and finally hit genes identified for each cell line regardless of the readout (C).

DISCUSSION

Among breast cancers, the triple-negative form displays the poorest prognosis and remains the most aggressive and heterogeneous disease, urgently requiring novel therapeutic modalities. Although TNBC is more sensitive to chemotherapy compared to non-TNBC breast cancers about 70% of patients display recurrent disease and all TNBC-diagnosed patients will die from the disease.[14,15] Consequently, numerous clinical trials are ongoing using various emerging therapeutics including PARP inhibitors, PI3K/mTOR signaling inhibitors and HDAC inhibitors.[16] Although the latter provide encouraging results, driven by the fact that TNBC progression like many other cancers is driven by gene expression deregulation in key signaling pathways, we undertook screening a comprehensive epigenetic modulator siRNAs library.[17,18] Ultimately, our results highlight the role of three epigenetic genes in regulating TNBC cell viability and growth: CDK1, CHAF1A, and KMT5A.

Recent studies highlighted the link between the overexpression of the transcription factor Myc and the sensitivity to CDK1 inhibition in TNBC. Horiuchi et al demonstrated that mice bearing overexpressing c-Myc TNBC cells were significantly more sensitive to Dinaciclib, a small molecule inhibitor of CDK 1, 2, 5 and 9, whereas most of the receptor-positive cell lines showed resistance to such treatment.[19] Another study performed by Liu et al obtained similar results through delivery of a CDK1 siRNA with a nanoparticule carrier.[20] Finally, CDK1 inhibition was also investigated through various combination treatments, including MYC and PARP inhibition.[21,22] Studies also highlighted the role of

KMT5A-mediated H4K20me1 promoting the EMT in triple-negative breast cancer. As an example, Yang et al demonstrated that, in addition to physically interact with TWIST, a key player in EMT, KMT5A is also recruited on the promoters of two TWIST target genes, N-cadherin and E-cadherin, in MDA-MB-231 TNBC cells.[23] KMT5A acts as a coactivator and a corepressor in mesenchymal N-cadherin and epithelial E-cadherin markers transcription, respectively, suggesting that KMT5A could be a potential therapeutic target to prevent EMT and metastasis in non-or triple-negative breast cancer. The results of our screen are perfectly aligned with previous published data demonstrating the role of CDK1 and KMT5A in the viability and proliferation of cancers cells and reinforce the rationale suggesting the interest of this substrates phosphorylase and this H4K20 methyltransferase, respectively, specifically in the context of TNBC tumors.

Although CDK1 and KMT5A were previously described as being involved in cancers including TNBC, hence strengthening our screening results, the role of CHAF1A was never highlighted before in the context of TNBC.

As a core component of CAF-1, CHAF1A (CAF p150) plays an essential role in many biological processes, such as chromatin assembly during DNA replication or after DNA repair.[24-26] CHAF1A epigenetically regulates gene expression through interaction with heterochromatin protein 1 (HP1) or with methyl-CpG binding domain protein 1 (MDB1).[27,28] Recent studies associated CHAF1A with the development and progression of solid tumors.[29-31] The importance of CHAF1A as a cancer driver gene is supported by the finding that its overexpression has been linked to tumor progression, cancer susceptibility, and more recently, epigenetic silencing.[32-34] CHAF1A was found overexpressed in gastric cancer cell lines and tissue samples and its high expression was predictive of poor outcome.[35] Functional in vitro studies demonstrated that CHAF1A expression promoted gastric cancer cell proliferation by enhancing transcriptional activation of c-MYC and CCND1 genes, through direct binding to their promoter regions. This activation was achieved in concert with TCF4, a mediator of Wnt signalling pathway, suggesting that CHAF1A may act as a co-activator of this important pathway. CHAF1A was also found to be involved in epithelial ovarian cancer (EOC) cell proliferation and cell apoptosis inhibition. Xia et al found that the positive staining of CHAF1A in EOC was higher than that in normal tissues and overexpression of CHAF1A was strongly associated with cancer stage and lymph node metastasis.[36] Knock down of CHAF1A by siRNA in EOC inhibited cell proliferation, reduced colony formation, caused G0/G1 phase arrest and promoted cell apoptosis. Taken together, these studies correlate with our screening data supporting the role for CHAF1A in cell proliferation and survival. The role of CHAF1A was also investigated in non-small cell lung cancer (NSCLC). Cai et al demonstrated that CHAF1A is a direct target of miR-520b and decreased expression of CHAF1A resulting from the up-regulation of miR-520b could decelerate the proliferation, invasion and migration of lung cancer cells.[37] Interestingly, miR520b is among the miRNAs that have been shown to be down-regulated in triple-negative breast cancer.[38] Interestingly, Montes de Oca et al performed unsupervised hierarchical clustering using some published transcriptome dataset of 1127 breast cancer patients and found several chromatin regulators, including two CAF1 subunits CHAF1A (p150) and p60, to be part of a gene cluster with high expression in triple negative, HER2+ and luminal B breast cancer subtypes and low expression in the luminal A subtype.[39] Ultimately, the specific role of CHAF1A in breast cancer remains to be uncovered but its role in the regulation of H3K9 trimethylation of key target genes regulating proliferation, survival, and differentiation might be compatible with a specific role in TNBC as suggested by our siRNA screening results. Our data allow to drift from the previous correlation between CHAF1A expression and cancer to a more causative importance for CHAF1A to support cancer cell growth/viability.

Interestingly, the expression of the three epigenetic genes CDK1, CHAF1A and KMT5A identified here as key genes in specific TNBC cell viability is correlated to c-Myc expression, a Wnt/β-catenin signalling pathway target gene. Given that aberrant activity of Wnt/β-catenin signaling pathway occurs in cancer stem cells (CSC) and that TNBC MDA-MB-231 cell line expresses elevated c-Myc levels, these findings are of interest since TNBC tumors are enriched in CSC, conferring chemo-resistance, EMT and metastasis. Our data support the rationale of targeting epigenetics in TNBC to eradicate tumors, sensitize cells and prevent metastases. In our study, MDA-MB-468 and MDA-MB-231 were selected as TNBC cell lines of interest. Among TNBC intrinsic molecular subtypes, MDA-MB-468 belongs to basal-like 1 and MDA-MB-231 to Mesenchymal. Further investigations will be necessary to uncover CDK1, CHAF1A, and KMT5A roles in additional TNBC subtypes such as immunomodulatory and luminal androgen receptor (LAR), supposing that their inhibition could potentially overcome the TNBC heterogeneity.

CONCLUSION

Our investigation of a large collection of siRNA targeting epigenetic modulators on TNBC and non-TNBC cell lines using two distinct cell viability assays combined with a comprehensive data analysis pipeline using four different methods enabled us to confirm CDK1 and KMT5A as important regulators in TNBC cell viability and growth and to identify CHAF1A as a novel regulator of TNBC cell survival.

ACKNOWLEDGEMENTS

Authors would like to thank Amandine Ferrand for her valuable technical help in the imaging protocols, Xavier

Scerri for implementing all the automation processes for the screenings, and Clémence Dupré's team and Chantal Bourrier's team for their kind supports in cell banking and cell line characterization.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the institutional ethics committee

REFERENCES

1. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P et al. Breast cancer. Nat Rev Dis Primer. 2019;5:66.
2. Thiery JP. Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol. 2003;15:740-6.
3. Liu S, Dontu G, Wicha MS. Mammary stem cells, self-renewal pathways, and carcinogenesis. Breast Cancer Res. 2005;7:86-95.
4. Toh TB, Lim JJ, Chow EK-H. Epigenetics in cancer stem cells. Mol Cancer. 2017;16:29.
5. Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of Epigenetic Drugs in Disease: An Overview. Genet Epigenetics. 2014;6:GEG:S12270.
6. Ar P, Fernandes GF DS, Jl DS. Clinical Pharmacology: Epigenetic Drugs at a Glance. Biochem Pharmacol. 2018;7:2.
7. Travaglini L, Vian L, Billi M, Grignani F, Nervi C. Epigenetic reprogramming of breast cancer cells by valproic acid occurs regardless of estrogen receptor status. Int J Biochem Cell Biol. 2009;41:225-34.
8. Shah P, Gau Y, Sabnis G. Histone deacetylase inhibitor entinostat reverses epithelial to mesenchymal transition of breast cancer cells by reversing the repression of E-cadherin. Breast Cancer Res Treat. 2014;143:99-111.
9. Schech A, Kazi A, Yu S, Shah P, Sabnis G. Histone Deacetylase Inhibitor Entinostat Inhibits Tumor-Initiating Cells in Triple-Negative Breast Cancer Cells. Mol Cancer Ther. 2015;14:1848-57.
10. Ma F, Li H, Wang H, Shi X, Fan Y, Ding X et al. Enriched CD44+/CD24-population drives the aggressive phenotypes presented in triple-negative breast cancer (TNBC). Cancer Lett. 2014;353:153-9.
11. Idowu MO, Kmieciak M, Dumur C, Burton RS, Grimes MM, Powers CN et al. CD44+/CD24-/low cancer stem/progenitor cells are more abundant in triple-negative invasive breast carcinoma phenotype and are associated with poor outcome. Hum Pathol. 2012;43:364-73.
12. Garmpis N, Christos D, Garmpi A, Kalampokas E, Kalampokas T, Spartalis E et al. Histone Deacetylases as New Therapeutic Targets in Triple-negative Breast Cancer: Progress and Promises. Cancer Genomics Proteomics. 2017;14:299-313.
13. Ogier A, Dorval T. HCS-Analyzer: open-source software for high-content screening data correction and analysis. Bioinformatics. 2012;28:1945-6.
14. Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F et al. The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes. Clin Cancer Res. 2007;13:2329-34.
15. Bonotto M, Gerratana L, Poletto E, Driol P, Giangreco M, Russo S et al. Measures of Outcome in Metastatic Breast Cancer: Insights From a Real-World Scenario. The Oncologist. 2014;19:608-15.
16. Lee A, Djamgoz MBA. Triple negative breast cancer: Emerging therapeutic modalities and novel combination therapies. Cancer Treat Rev. 2018;62:110-22.
17. Jones PA, Baylin SB. The Epigenomics of Cancer. Cell. 2007;128:683-92.
18. Rodríguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. Nat Med. Nature Publishing Group; 2011;17:330-9.
19. Horiuchi D, Kusdra L, Huskey NE, Chandriani S, Lenburg ME, Gonzalez-Angulo AM et al. MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. J Exp Med. 2012;209:679-96.
20. Liu Y, Zhu Y-H, Mao C-Q, Dou S, Shen S, Tan Z-B et al. Triple negative breast cancer therapy with CDK1 siRNA delivered by cationic lipid assisted PEG-PLA nanoparticles. J Controlled Release. 2014;192:114-21.
21. Klauber-DeMore N, Schulte BA, Wang GY. Targeting MYC for triple-negative breast cancer treatment. Oncoscience. 2018;5:120-1.
22. Xia Q, Cai Y, Peng R, Wu G, Shi Y, Jiang W. The CDK1 inhibitor RO3306 improves the response of BRCA-proficient breast cancer cells to PARP inhibition. Int J Oncol. Spandidos Publications; 2014;44:735-44.
23. Yang F, Sun L, Li Q, Han X, Lei L, Zhang H et al. SET8 promotes epithelial-mesenchymal transition and confers TWIST dual transcriptional activities: SET8 promotes TWIST-induced EMT. EMBO J. 2012;31:110-23.
24. Verreault A, Kaufman PD, Kobayashi R, Stillman B. Nucleosome Assembly by a Complex of CAF-1 and Acetylated Histones H3/H4. Cell. 1996;87:95-104.
25. Kadyrova LY, Blanko ER, Kadyrov FA. CAF-I-dependent control of degradation of the discontinuous strands during mismatch repair. Proc Natl Acad Sci U S A. 2011;108:2753-8.
26. Smith CL, Matheson TD, Trombly DJ, Sun X, Campeau E, Han X et al. A separable domain of the p150 subunit of human chromatin assembly factor-1 promotes protein and chromosome associations with nucleoli. Mol Biol Cell. Am Society Cell Biol: 2014;25:2866-81.
27. Murzina N, Verreault A, Laue E, Stillman B. Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. Mol Cell. 1999;4:529-40.

28. Reese BE, Bachman KE, Baylin SB, Rountree MR. The Methyl-CpG Binding Protein MBD1 Interacts with the p150 Subunit of Chromatin Assembly Factor 1. Mol Cell Biol. 2003;23:3226-36.
29. Xu M, Jia Y, Liu Z, Ding L, Tian R, Gu H et al. Chromatin assembly factor 1, subunit A (P150) facilitates cell proliferation in human hepatocellular carcinoma. OncoTargets Ther. 2016;9:4023-35.
30. Peng H, Du B, Jiang H, Gao J. Over-expression of CHAF1A promotes cell proliferation and apoptosis resistance in glioblastoma cells via AKT/FOXO3a/Bim pathway. Biochem Biophys Res Commun. 2016;469:1111-6.
31. Barbieri E, Preter KD, Capasso M, Chen Z, Hsu DM, Tonini GP et al. Histone Chaperone CHAF1A Inhibits Differentiation and Promotes Aggressive Neuroblastoma. Cancer Res. American Association for Cancer Research; 2014;74:765-74.
32. Glinsky GV. Genomic models of metastatic cancer: functional analysis of death-from-cancer signature genes reveals aneuploid, anoikis-resistant, metastasis-enabling phenotype with altered cell cycle control and activated Polycomb Group (PcG) protein chromatin silencing pathway. Cell Cycle Georget Tex. 2006;5:1208-16.
33. Jiao R, Bachrati CZ, Pedrazzi G, Kuster P, Petkovic M, Li J-L et al. Physical and Functional Interaction between the Bloom's Syndrome Gene Product and the Largest Subunit of Chromatin Assembly Factor 1. Mol Cell Biol. 2004;24:4710-9.
34. Poleshko A, Einarson MB, Shalginskikh N, Zhang R, Adams PD, Skalka AM et al. Identification of a Functional Network of Human Epigenetic Silencing Factors. J Biol Chem. 2010;285:422-33.
35. Zheng L, Liang X, Li S, Li T, Shang W, Ma L et al. CHAF1A interacts with TCF4 to promote gastric carcinogenesis via upregulation of c-MYC and CCND1 expression. EBio Medicine. 2018;38:69-78.
36. Xia D, Yang X, Liu W, Shen F, Pan J, Lin Y et al. Over-expression of CHAF1A in Epithelial Ovarian Cancer can promote cell proliferation and inhibit cell apoptosis. Biochem Biophys Res Commun. 2017;486:191-7.
37. Cai Y, Dong ZY, Wang JY. MiR-520b inhibited metastasis and proliferation of non-small cell lung cancer by targeting CHAF1A. Eur Rev Med Pharmacol Sci. 2018;22(22):7742-9.
38. Tang Q, Ouyang H, He D, Yu C, Tang G. MicroRNA-based potential diagnostic, prognostic and therapeutic applications in triple-negative breast cancer. Artif Cells Nanomedicine Biotechnol. 2019;47:2800-9.
39. Montes de Oca R, Gurard-Levin ZA, Berger F, Rehman H, Martel E, Corpet A et al. The histone chaperone HJURP is a new independent prognostic marker for luminal A breast carcinoma. Mol Oncol. 2015;9:657-74.

Cite this article as: Gaudeau A, Clua Provost C, Dorval T, Walsh A, Hannus M, Perez F, Camonis J, Del Nery E, Stephan JP. Cell-based siRNA screens highlight triple-negative breast cancer cell epigenetic vulnerability. Int J Sci Rep 2021;7(4):196-206.