Universiteit
Leiden
The Netherlands

RNA splicing in breast cancer progression

Koedoot, E.

Citation
Koedoot, E. (2019, December 17). *RNA splicing in breast cancer progression*. Retrieved from https://hdl.handle.net/1887/81820

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/81820

Note: To cite this publication please use the final published version (if applicable).

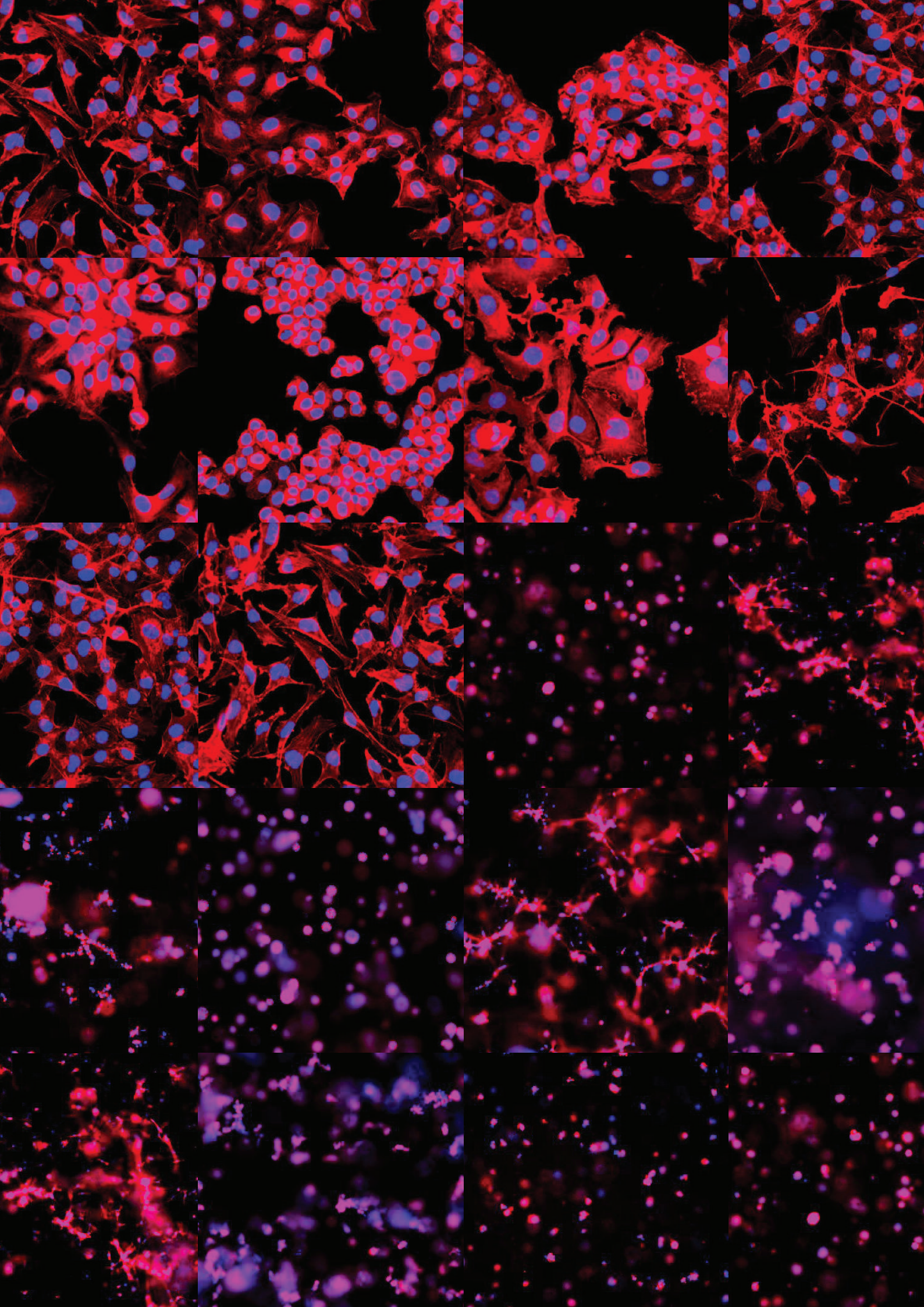
Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/81820 holds various files of this Leiden University dissertation.

Author: Koedoot, E.
Title: RNA splicing in breast cancer progression
Issue Date: 2019-12-17



Chapter 7

Differential reprogramming of breast cancer subtypes in 3D cultures and implications for sensitivity to targeted therapy

Esmee Koedoot, Liesanne Wolters, Marcel Smid, Peter Stoilov, Gerhard A. Burger, Bram Herpers, Kuan Yan, Leo S. Price, John A. Foekens, John W.M. Martens, Sylvia E. Le Dévédec, Bob van de Water

Submitted

Highlights

- A systematic comparison of the transcriptome of breast cancer cell lines in 2D and 3D cultures demonstrated increased mitochondrial metabolism and downregulated cell cycle programs in 3D cultures
- Invasive breast cancer cell lines demonstrated increased expression of extracellular matrix interaction genes in 3D culutures
- 3D cell cultures showed decreased levels of metastatic genes and are better correlated to patient transcriptomes

◄ IN THE PICTURE

Microscopic view of breast cancer cells cultured in 2D (top) and 3D (bottom) conditions. Cells were stained for the actin cytoskeleton (red) and nuclei (blue).

◄ IN BEELD

Microscopisch beeld van borstkankercellen gekweekt in 2D (boven) en 3D (onder) condities. Cellen zijn gekleurd voor het cytoskelet (rood) en celkernen (blauw).

Chapter 7

Screening for effective candidate drugs for breast cancer has shifted from two-dimensional (2D) to three-dimensional (3D) cultures. Here we systematically compared the cellular physiology of these different culture conditions by RNAseq of 14 BC cell lines cultured in both 2D and 3D conditions. All 3D BC cell cultures demonstrated increased mitochondrial metabolism and downregulated cell cycle programs. Luminal BC cells in 3D demonstrated overall limited reprogramming. 3D basal B BC cells showed increased expression of extracellular matrix (ECM) interaction genes, which coincides with an invasive phenotype not observed in other BC cells. Genes downregulated in 3D were associated with metastatic disease progression in BC patients, including cyclin dependent kinases and aurora kinases. Furthermore, the overall correlation of the cell line transcriptome to the BC patient transcriptome was increased in 3D cultures for all TNBC cell lines. To define the most optimal culture conditions to study the oncogenic pathway of interest, an open source bioinformatics strategy was established.

Introduction

Breast cancer is the most prevalent cancer and the second leading cause of cancer death in women with an estimated 40,610 deaths in the United States in 2017[367]. Based on levels of the estrogen, progesterone and HER2 receptors, breast cancer can be divided in different subtypes. The triple-negative subtype (TNBC) lacking the expression of these three hormone receptors accounts for 15-20% of all tumors[368] and is the most aggressive subtype, often leading to metastases[72,73]. Despite the efforts, there is still no targeted therapy for TNBC available[369]. A major reason for this lack in clinical translation may be the use of two-dimensional in vitro experiments that do poorly represent the three-dimensional (3D) tissue physiology observed in human cancer patients. To increase translation from in vitro findings to a clinical setting, different 3D culture systems are now explored, such as organoid cultures, patient-derived xenograft models, reprogrammed stem cell like models, tumor-on-a-chip and 3D cultures of immortalized breast cancer cell lines[370]. While the majority of breast cancer drug screening studies in the last decade have still been performed in 2D[371–376], there is an increasing number of drug screens performed in more complex models such as patient-derived organoids[377,378], tumor-on-a-chip[379] and patient-derived xenograft[380] models. Although these complex models better represent human physiology and should increase clinical translation[381–383], drawbacks of these models include reduced reproducibility[381,384,385], increasing costs, inconvenient maintenance, difficulties in expanding them and generating genetic modifications, making these models less suitable for high-throughput screening[386]. Next to the already widely studied phenotypic changes between different culturing models[380,387] and phenotypic classification of different tumor subtypes[370], transcriptomic analysis can contribute to the understanding of the differences between established in vitro models and help to determine the most suitable model in terms of both clinical translation, costs and efficiency.

Here, we performed RNA-sequencing of 14 breast cancer cell lines cultured on a 2D plastic substrate as well as in a 3D matrigel-collagen environment. In this 3D model, cells spontaneously form spheroid-like structures exhibiting cell-cell as well as cell-extracellular matrix interactions, thereby changing their cell polarity and shape. We unraveled the

transcriptomic differences linked to the invasive phenotype of basal B TNBC compared to basal A and luminal breast cancer and uncovered a spectrum of genes higher expressed in 2D cultures that were related to metastatic progression in breast cancer patients, including CDKs and aurora kinases. Interestingly, these differential expression levels did not change target-specific drug sensitivity. Since the transcriptomic correlation of in vitro cultured cell models to patient tumor tissue was highly subtype and pathway dependent, we established a bioinformatics tool that can be used in future studies to select the most suitable cell type and culture conditions for the pathway of interest. Altogether, this study unraveled the transcriptomic variance between different breast cancer in vitro models and provides an important database that can contribute to selection of the most effective and relevant drug candidates for the treatment of TNBC.

Materials and methods

Cell culture

All cell lines were purchased from ATCC. Cells were grown in RPMI-1640 medium (Gibco, ThermoFisher Scientific, Breda, The Netherlands) supplemented with 10% FBS (GE Healthcare, Landsmeer, The Netherlands), 25 IU/ml penicillin and 25 µg/ml streptomycin (full RPMI) (ThermoFisher Scientific) at 37 °C in a humidified 5% CO_2 incubator. For 2D cultures, cells were grown on plastic tissue culture plates (Corning). For 3D cultures, cells were grown in a gel-mixture containing 6 mg/ml matrigel (Corning, Lot# 5061003), 0.5 mg/ml collagen (Corning, Lot# 5092001), 3.7 g/L $NaHCO_3$ and 0.05 M HEPES.

RNA isolation

For 2D cultures, cells were plated in a 6-well tissue culture plate (Corning, 3516) in full RPMI. RNA was collected 2 days after plating at 70 % confluency using RNeasy plus mini kit (Qiagen) according to the manufacturer's protocol.

For 3D cultures, cells were plated in a 24-wells tissue culture plate (Corning, 3524). Wells were pre-coated with 200 µl matrigel (Corning, Lot# 5061003, 9.4mg/ml) and cells were plated in a total volume of 400 µl. The following densities were used for the different cell lines: MDA-MB-231 (200 cells/µl), Hs578T (200 cells/µl), BT549 (200 cells/µl), SUM149PT (250 cells/µl), HCC1143 (250 cells/µl), HCC1806 (250 cells/µl), HCC1954 (250 cells/µl) and MDA-MB-468 (250 cells/µl). After 30 minutes incubation at 37 °C, 1200 µl full RPMI was added. RNA was collected 7 days after plating. RPMI was removed, 1.2 ml QIAzol (Qiagen) was added per well and cells were lysed by repetitive pipetting. Cells were incubated for 5 minutes at room temperature and 240 µl chloroform was added, followed by 2-3 minutes incubation at room temperature. Samples were centrifuged for 15 minutes at 4°C (12,000 x g), the upper aqueous phase was collected and RNA was precipitated by mixing with 600 µl isopropanol. Samples were incubated for 10 minutes at room temperature, followed by 15 minutes centrifugation at 4°C (12,000 x g) and washing with 1.2 ml 75% v/v ethanol. Samples were centrifuged for 5

minutes at 4°C (7,500 x g). Then the supernatant was removed and the pellet was air-dried and dissolved in 30 µl RNase free water.

RNA sequencing and analysis

DNA libraries were prepared with the TruSeq Stranded mRNA Library Prep Kit and sequenced according to the Illumina TruSeq v3 protocol on an Illumina HiSeq2500 sequencer. 100 base pair paired-end reads were generated and alignment was performed using the HiSat2 aligner (version 2.2.0.4) against the human GRCh38 reference genome. Gene expression was quantified using the HTseq-count software (version 0.6.1) based on the ENSEMBL gene annotation for GRCH38 (release 84). Count data was normalized using the DESeq2 package[92] and DEGs were selected based on log2 fold change > 1 or < -1. Paired t-test were performed on the normalized RNA-seq data using BRB-ArrayTools (v4.4.1) developed by Dr. Richard Simon and the BRB-ArrayTools Development Team and R (v3.2.2). Genes with p-value < 0.01 were considered significant. Upon acceptance, RNA sequencing data will be available in the Sequence Read Archive.

Drug screening

For 2D drug screening, cells were plated in a 96-well tissue culture plate (Corning, 3599) using the following cell densities: HCC1143 (8,000 cells/well), HCC1806 (8,000 cells/well), HCC1954 (8,000 cells/well), MDA-MB-468 (10,000 cells/well), Hs578T (4,000 cells/well), MDA-MB-231 (5,000 cells/well), SUM149PT (8,000 cells/well) and BT549 (8,000 cells/well). Drug treatment was performed 16 hours after plating and 4 days after treatment, cells were fixed for sulforhodamine B (SRB) colorimetric proliferation assay[343] previously described by our group[388].
For 3D drug screening, cells were plated in a 384-well µclear plate (Greiner, 781091) using a CyBi Selma 96/60 robotic liquid dispenser (Analytik Jena AG, Jena, Germany). The same gel composition and cell densities were used as in the RNA isolation procedure. Cells were plated in a total volume of 14.5 µl matrigel-collagen mix. After 30 minutes incubation at 37°C, 44.5 µl full RPMI was added. Drug treatment was performed 3 days after plating and 4 days after treatment, cells were fixed and stained in 0.05 µM Phalloidin Rhodamin (Sigma Aldrich), 0.4 µg/ml Hoechst 33258 (Fisher Biotech), 0.1% Triton X-100 Triton (Sigma Aldrich) and 3.7% formaldehyde (Sigma Aldrich) in PBS O/N at 4°C. Plates were washed with PBS and imaged using ImageXpress Micro XLS system (Molecular Devices, CA, USA).
Both in 2D and 3D cultures, cells were treated with 6 concentrations of the following inhibitors: MG132, bortezomib, flavopiridol HCl, dinaciclib, palbociclib, milciclib, aurora kinase inhibitor I, TAK-901, gefitinib and erlotinib (all Selleck Chem). DMSO was used as a negative control.

Drug screen normalization and analysis

Proliferation was normalized to the number of cells before treatment (set at 0 %) and DMSO treatment (set at 100%). A nonlinear sigmoidal dose–response (with variable slope) curve was used to fit the data and retrieve the IC50 values using Graphpad Prism 7.00.
To analyze the 3D cultures the acquired image stacks (4x objective, 40 sections per well, 50 µm step size) were processed in OcellO Ominer 3D image analysis platform. The software

processes the DAPI channel to measure the number of nuclei and the shapes of the nuclei in each image stack and integrates this information with the shape and number of tumor cell clusters detected in the TRITC channel.

Immunofluorescence

For 2D immunofluorescence staining, cells were plated in a 96-well µclear plate (Greiner). Cells were fixed and permeabilized 72 hours after plating by incubation with 1% formaldehyde and 0.1% Triton X-100 in PBS and blocked with 0.5% w/v BSA in PBS. Cells were incubated with 1:10,000 Hoechst 33258 combined with 1:2000 rhodamine-phalloidin (Molecular Probes, R415) for 1 hour at room temperature. Cells were imaged with a Nikon Eclipse Ti microscope and 20x oil objective.

Hierarchical clustering and pathway analysis

Over-representation analysis was performed using Consensus PathDB[94] using KEGG and Reactome databases. Ranked gene set enrichment analysis (GSEA) was performed on the full ranked gene lists[93]. T-sne plots were generated using the tsne R package. Hierarchical clusterings were generated using the pheatmap R package, based on complete linkage and Euclidean distance or correlations as described in the figure legend.

Data retrieval and normalization

Normalized total gene expression data from The Cancer Genome Atlas (TCGA) were obtained by using the TCGA Assembler R package[281] after the new release in January 2017. Normalized reads were log2 transformed before further analyses were performed. For calculating spearman correlations between the TCGA and cell line RNA sequencing data, log2 normalized data were used for both datasets. Microarray data of primary tumors of 867 untreated patients (MA-867 dataset) used to calculate hazard ratios was previously published and publicly available (GSE2034, GSE5327, GSE2990, GSE7390 and GSE11121). Raw .cel files were downloaded, processed with fRMA and batch effects were corrected using ComBat.

Results

mRNA profiling of breast cancer cells cultured in 3D revealed downregulation of cell cycle-related genes and upregulation of mitochondrial genes

To understand how cell culture systems affect the transcriptome of breast cancer (BC) cells, we performed RNA sequencing of 52 human breast cancer cell lines cultured on 2D tissue culture plastic and 14 cell lines cultured in a 3D matrigel-collagen environment (Fig. 1A, Suppl. Table 1). The selection of the 14 cell lines was based on previously defined subtype classifications[42,389–392] with selected cell lines representing the different BC subtypes. These cell line subtypes were validated in our RNA sequencing dataset; hierarchical clustering based on RNA expression of previously published luminal and basal markers clearly separated the different subtypes (Fig. 1B and Suppl. Fig. 1, Suppl. Table 2)[389,393]. Moreover, cell lines could

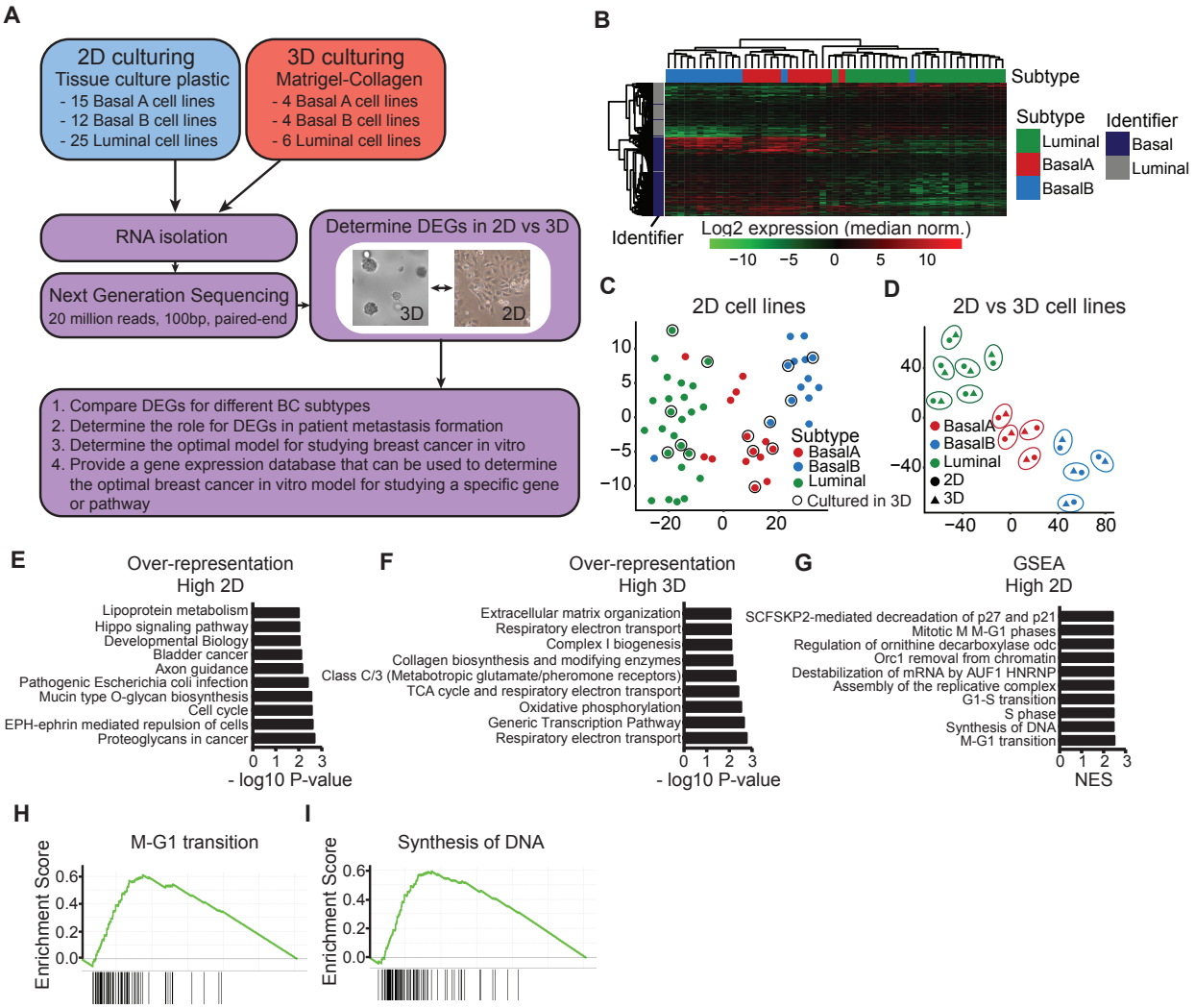


Figure 1. **RNA sequencing of breast cancer cell lines cultured in 2D and 3D environments.** (A) Experimental overview. RNA was isolated from cell lines cultured in 2D and 3D conditions (see Methods section for more details), upon which next generation sequencing was performed. Cell line specific DEGs were selected based on log2 fold change >1 or <-1. Subtype specific DEGs were determined using a paired cell line analysis and selected if the parametric p-value < 0.01. See Methods section for further details. (B) Hierarchical clustering (complete linkage, correlation) of log2 expression levels of previously defined basal and luminal identifiers[393] in 2D cultured cell lines. For every gene, expression levels were normalized to the median cell line expression. (C) T-sne of all cell lines cultured in 2D based on log2 normalized reads from all detected genes. Encircled cell lines were also cultured in 3D in this study. (D) T-sne of cell lines cultured in both 2D and 3D conditions based on log2 normalized reads from all detected genes. 2D and 3D conditions of the same cell line are encircled. (E) Top 10 over-represented pathways in DEGs higher expressed in 2D cultures. DEGs were determined using paired analysis for all cell lines, $p < 0.01$. (F) Top 10 over-represented pathways in DEGs higher expressed in 3D cultures. DEGs were determined using paired analysis for all cell lines, $p < 0.01$. (G) Top 10 enriched pathways higher expressed in 2D cultures using a ranked gene set enrichment analysis (GSEA). (H) and (I) Examples of GSEA of pathways enriched in 2D cultures.

also be assigned to these subtypes based on whole transcriptomic differences (Fig. 1C). Three cell lines (SUM149PT, SUM225CWN and HCC1500) did not cluster with their assigned subtype (Fig. 1B), which was in agreement with the ambiguous subtype classification described in the

literature (Suppl. Table 3)[42,389–392]. To visualize the global transcriptomic differences between 2D and 3D culturing systems, we created a T-sne plot using RNA expression levels from the cell lines that were cultured both in 2D and 3D culture conditions (Fig. 1D). Cell lines separated consistently based on their subtype and, interestingly, the differences between cell lines were much larger than the differences between 2D/3D culture conditions; 2D and 3D samples of the same cell line clustered together (Figure 1D, encircled). However, these 2D and 3D samples were still clearly separated, revealing significant changes in gene expression patterns between the different culture conditions. To investigate the altered pathways comparing 2D to 3D culturing conditions, we calculated the significantly differentially expressed genes (DEGs) pairing 2D and 3D samples for every cell line, followed by over-representation analysis (Figure 1E/1F) and ranked gene set enrichment analysis (GSEA) (Figure 1G). In total, we identified 452 genes higher expressed in 2D and 1762 genes higher expressed in 3D (Suppl. Table 4). Over-represented and enriched pathways in 2D cultures include mainly cell cycle and cancer related pathways (Figure 1E/G/H/I), while pathways related to oxidative phosphorylation and transcription were higher expressed in 3D systems.

To ensure that the identified pathways were not dominated by a subset of genes and generally affected among all cell line, we plotted the log2 fold changes comparing 2D to 3D culturing methods for top-ranked pathways in every single cell line (Suppl. Fig. 2). For the cell cycle-related pathways, we observed a right-shifted normal distribution in basal A and basal B cell lines (but not in luminal cell lines, Suppl. Fig. 2A/C), confirming that the majority of the cell cycle-related genes was affected. Interestingly, for the respiratory electron transport pathway a bimodal distribution was observed in both basal A and B lines, with the majority of genes slightly higher expressed in 2D, while a small proportion was significantly higher expressed in 3D cultures (Suppl. Fig. 2B). Zooming in on this pathway revealed that in particular the mitochondrial genes were up-regulated in 3D (Suppl. Fig. 3).

Altogether, we demonstrated that culturing cell lines in a 3D matrigel-collagen environment mainly reduced expression of cell cycle related genes, but also resulted in elevated expression of mitochondrial genes.

3D culturing induced expression of extracellular matrix organization genes in Basal B cell lines

In Suppl. Fig. 2, we showed that gene expression differences between 2D and 3D cultures can be subtype and cell line specific. Therefore, we now examined the DEGs for each cell line. On average, ~3000 genes were higher expressed in 2D per cell line (log2 fold change >1) (Suppl. Fig. 4A), independent of cell line or subtype. Interestingly, in 3D cultures the number of DEGs was higher in basal compared to luminal cell lines (~2000 vs 4500; Fig. 2A and Suppl. Fig. 4B). This was confirmed by the number of significantly DEGs per subtype: for the luminal subtype ~200 genes were higher expressed in both 2D and 3D, while for the basal subtype also ~200 genes were higher expressed in 2D, but ~800 (basal A) and ~1100 (basal B) genes were upregulated in 3D (Fig. 2B). Furthermore, the majority of the DEGs for the basal subtypes were overlapping (Suppl. Fig. 4C). Interestingly, while in 2D cultures both basal A and basal B

subtypes displayed an active migratory behavior (Suppl. Fig. 5), in 3D cultures only basal B showed an invasive phenotype (Fig. 2C). In order to identify the underlying pathways that contribute to this differential invasive capacity, we performed an over-representation analysis for the subtype-specific DEGs in basal A and basal B cell lines (Fig. 2D/E, Suppl. Fig 4D/E). As shown earlier, both subtypes demonstrated higher levels of cell cycle related genes in 2D cultures (Fig. 2D and Suppl. Fig. 4E). Remarkably, pathways related to cell-matrix interactions, extracellular matrix organization, collagen synthesis and integrin cell surface interactions were higher expressed in 2D for the basal A subtype (Fig. 2D), while higher expressed in 3D for the basal B subtype (Fig. 2E). Only a small subset of the extracellular matrix organization pathway was upregulated in 3D settings in basal B cell lines, while equally expressed or down-regulated in basal A cells (Fig. 2F). Interestingly, this subset includes the genes PLOD1, P3H3 and P3H4, which together form a complex that catalyzes hydroxylation of lysine residues in collagen chains and is required for proper collagen biosynthesis and modification[394] (Fig. 2G). Other interesting members identified in this clusters are genes involved in cell adhesion, such as ITGB1 (Fig. 2G) and genes involved in breakdown of the extracellular matrix such as TIMP2 (Fig. 2G).

Metastatic breast cancer genes are downregulated in 3D cultures

Our analysis of the differential mRNA profiles between 2D and 3D cultures provided us with new insights on the transcriptional reprogramming related to basal cell line invasiveness in 3D cultures. Next, we wondered whether these newly identified DEGs would also play a role in breast cancer progression and metastasis formation in patients. Therefore, we examined the expression of DEGs in the different primary breast tumor subtypes using RNA sequencing data from 923 primary breast tumors (116 TNBC and 807 ER positive tumors) derived from The Cancer Genome Atlas (TCGA). Since we observed clear differences in culture-specific gene expression changes for the different subtypes, we performed these analyses for the DEGs obtained from the paired analysis for all cell lines, but also for basal A or basal B specific DEGs (Fig. 3A). Independent of the cell line subtype, DEGs higher expressed in 3D were higher expressed in the less aggressive ER positive subtype, while DEGs higher expressed in 2D were higher expressed in the more aggressive TNBC subtype (Fig. 3B-C). To investigate if these DEGs are also associated with metastatic progression in breast cancer patients, the hazard ratios (HRs) for distant metastasis-free survival (DMFS) were examined using microarray data of 867 untreated lymph-node negative breast cancer patients (Fig. 3A). DEGs with a significant HR were selected and the distribution of these HRs was plotted (Fig. 3D). DEGs higher expressed in 3D showed slightly more DEGs of which high expression is negatively related to metastatic progression in all subtypes (HR < 1). Interestingly, DEGs higher expressed in 2D were mainly positively related to metastatic progression (HR > 1) independent of the subtype. Examples of such DEGs down-regulated in 3D cultures and significantly related to metastatic progression were cyclin dependent kinases (CDKs) (Fig. 3E and F), aurora kinases (Fig. 3G and H) and proteasomal factors (Suppl. Fig. 6A-B). Since dysregulation of cell cycle is one of the hallmarks of cancer[21], CDKs and aurora kinases are attractive targets for new cancer therapy[395–398]. Moreover, cancer cells might be more dependent on the proteasome for the elimination of abnormal proteins due to high proliferation rate and genetic instability[399,400].

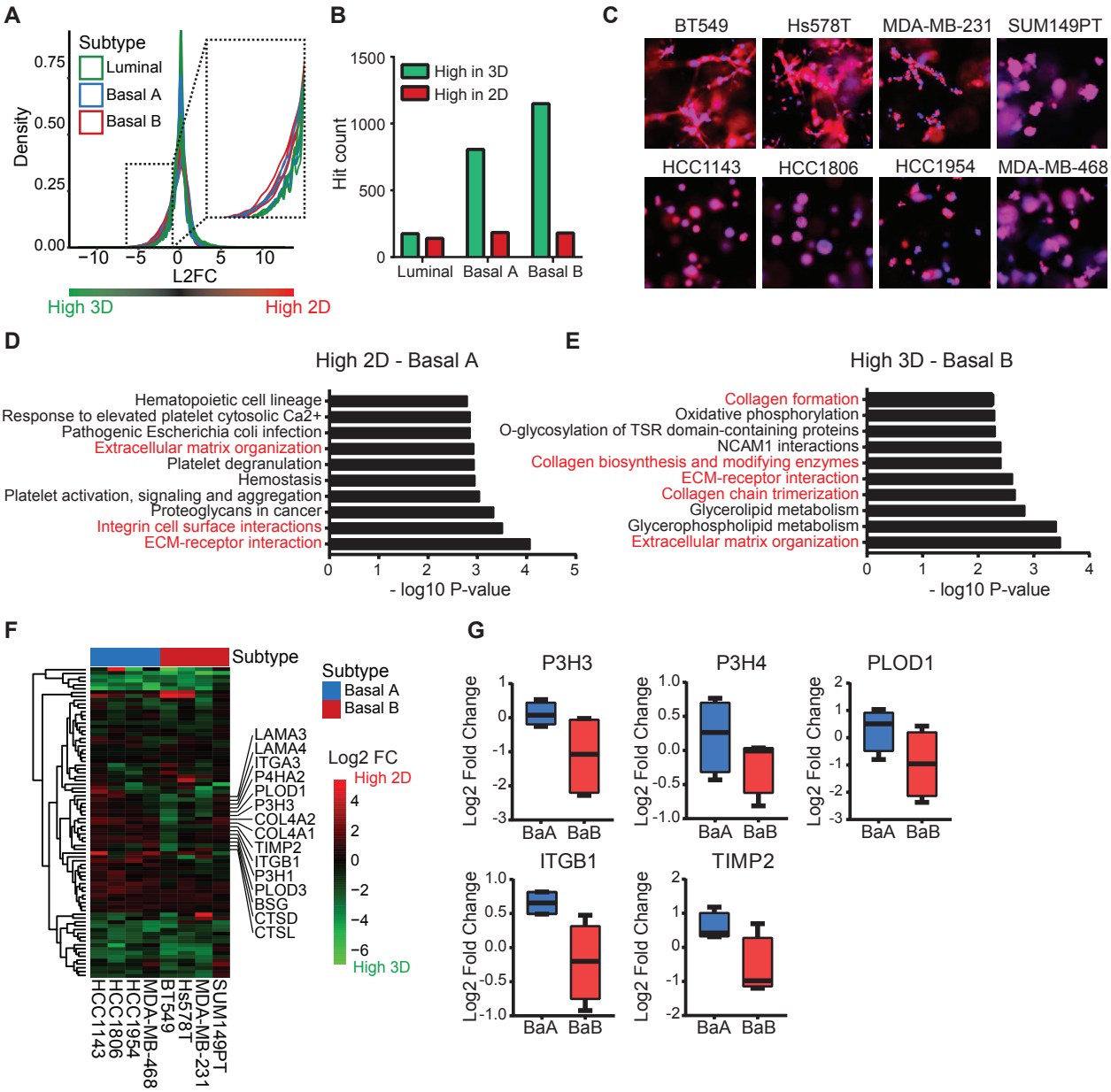
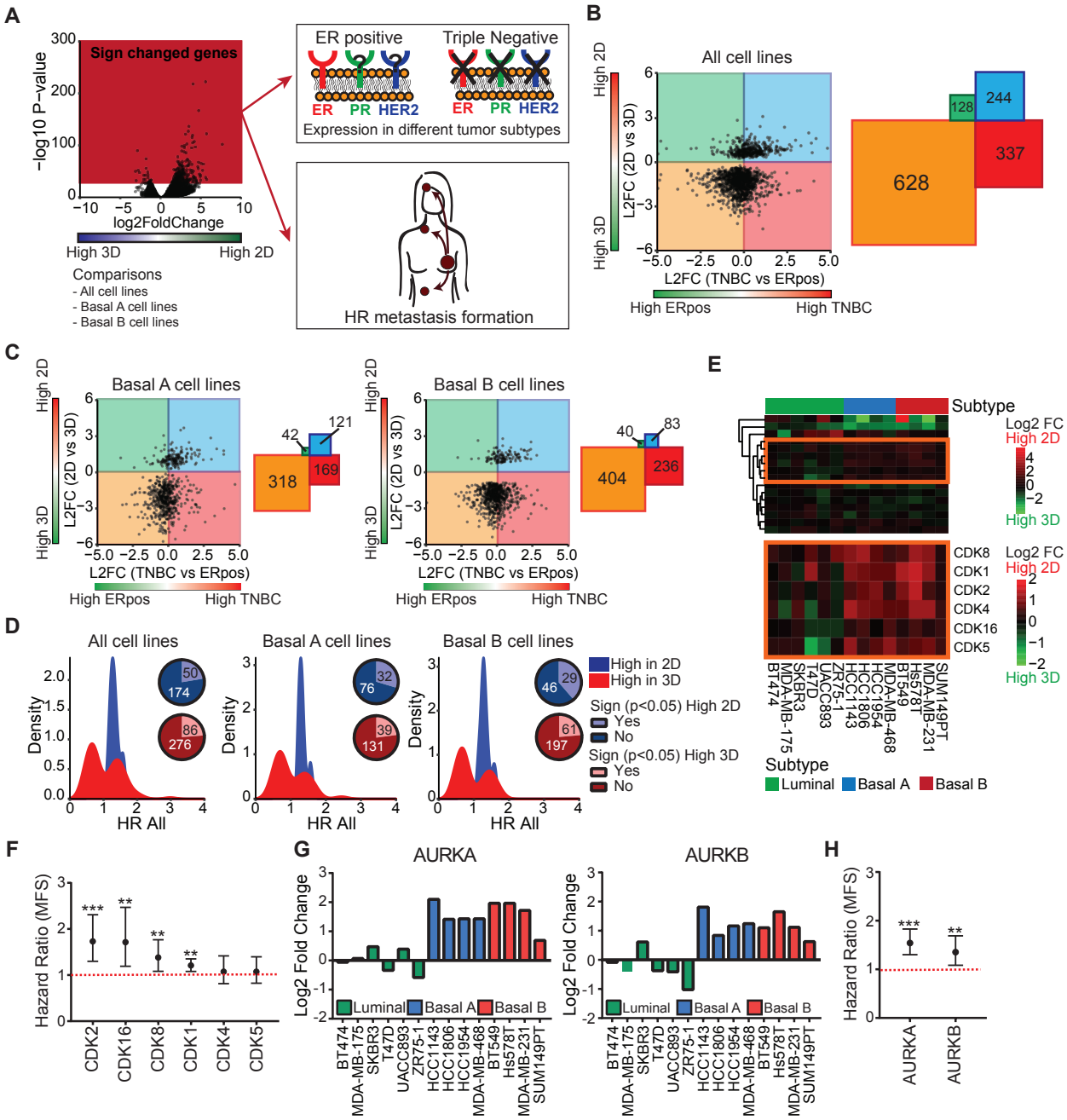


Figure 2. **Subtype specific differences in 2D and 3D breast cancer cultures.** (A) Density plot of log2 fold change in gene expression comparing 2D to 3D cultures of 14 cell lines (4 basal A, 4 basal B and 6 luminal). (B) Significantly DEG count for different subtypes based on paired analysis and $p < 0.01$. (C) Cellular morphology of basal A and basal B cell lines cultured in 3D. (D) Top 10 over-represented pathways in the high 2D DEGs in basal A cell lines. Migration and invasion related pathways are highlighted in red. DEGs were determined by paired analysis of basal A cell lines, $p < 0.01$. (E) Top 10 over-represented pathways in the high 3D DEGs in basal B cell lines. Migration and invasion related pathways are highlighted in red. DEGs were determined by paired analysis of basal B cell lines, $p < 0.01$. (F) Heatmap (complete linkage, Euclidean distance) of log2 fold change in expression of genes in the extracellular matrix organization pathway comparing 2D to 3D cultures. (G) Log2 fold change of expression levels of genes in the extracellular matrix pathway comparing 2D to 3D cultures in basal A and basal B cell lines. Selected genes were higher expressed in 2D in the basal A cell lines, but higher expressed in 3D in the basal B cell lines. Log2 fold change was determined using a paired analysis per subtype.



Altogether, DEGs higher expressed in 2D cultures were also higher expressed in the more aggressive TNBC subtype and positively associated with metastatic progression.

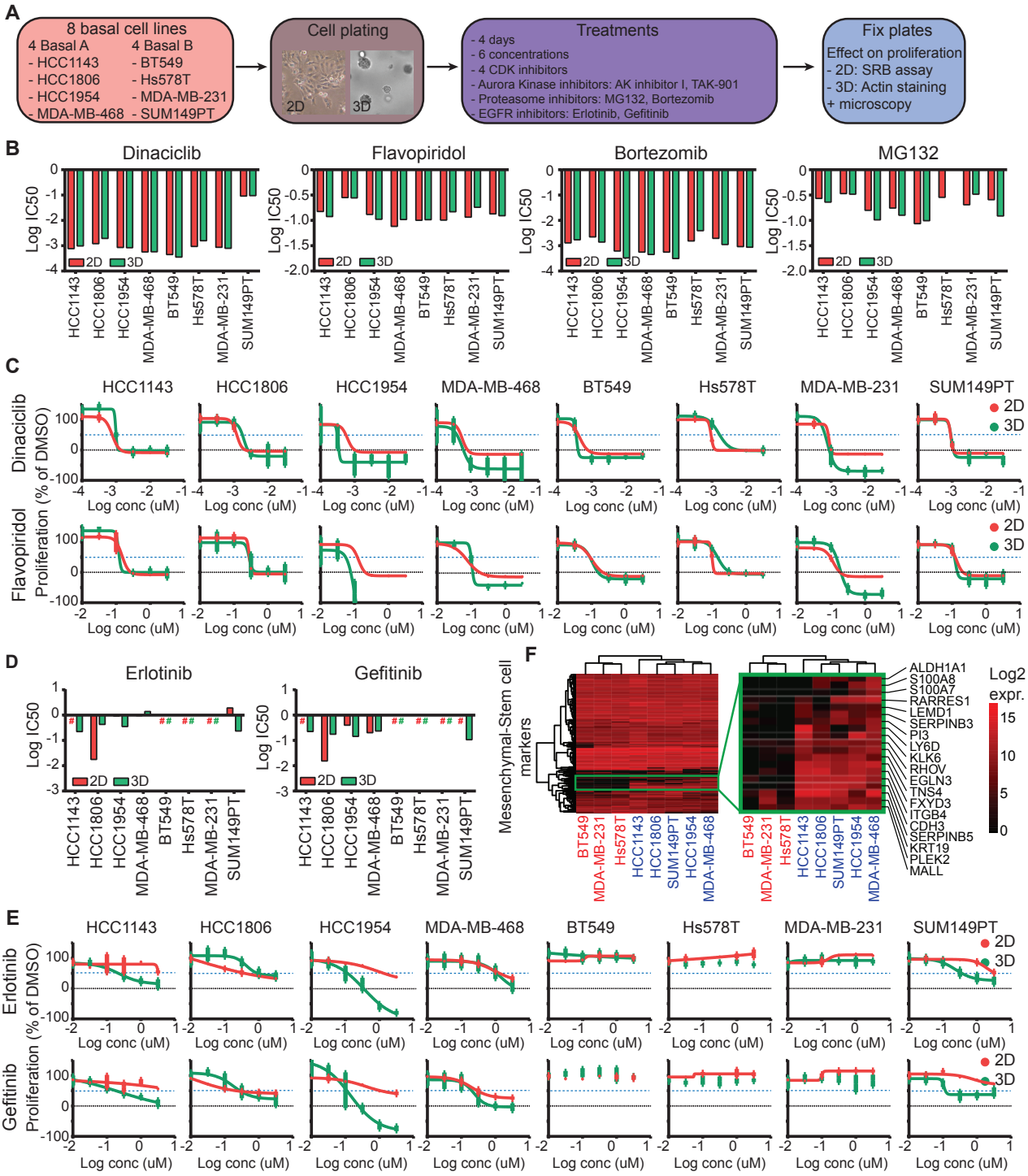
Differential sensitivity of 3D TNBC cultures to clinical relevant kinase inhibitors

CDKs, aurora kinases and proteasomal inhibitors are critical candidate drugs some of which are already in clinical trials or application; these inhibitors have mainly been studied in 2D cell culture systems[85,401–403]. The respective targets of these various inhibitors were significantly lower expressed in 3D cultures (Fig. 3E and 3G, Suppl. Fig. 6A), potentially leading to a change in sensitivity. Here, we tested whether BC cell line drug sensitivity is dependent on the differential transcriptome driven by the 2D and 3D culturing conditions. 4 basal A and 4 basal B breast cancer cell lines were cultured in both culturing systems, treated with several CDK

◄ ***Figure 3.*** **Role of DEGs in breast cancer aggressiveness and metastasis formation.** (A) Significantly DEGs comparing all cell lines, basal A cell lines or basal B cell lines were selected using a paired cell line analysis per subtype, $p < 0.01$. For these genes, the expression in different tumor subtypes as well as the relation to metastasis formation was determined. (B) Log2 fold change of DEGs (paired analysis of all cell lines, $p < 0.01$) comparing 2D to 3D cell cultures vs log2 fold change comparing TNBC to ER positive primary tumors (left). Number of genes in each quadrant (right). (C) Log2 fold change of DEGs (paired analysis of basal A and basal B cells, $p < 0.01$) comparing 2D to 3D cell cultures vs log2 fold change comparing TNBC to ER positive breast tumors (left). Number of genes in each quadrant (right). (D) Density plot of DEGs (paired analysis of all, basal A and basal B cell lines respectively, $p < 0.01$) with a significant hazard ratio (HR). Blue = density plot for DEGs with increased expression in 2D cultures. Red = density plot for DEGs with increased expression in 3D cultures. Venn diagrams show the fraction of DEGs showing a significant HR. (E) Heatmap (Euclidean distance, complete linkage) of log2 fold change of CDK expression levels comparing 2D to 3D cultures. Orange cluster is enlarged below. (F) Hazard ratios and 95% confidence interval (CI) for CDKs higher expressed in 2D. (G) Log2 fold change of aurora kinase A and B comparing 2D to 3D cultures for different cell lines. (H) Hazard ratios and 95% CI for aurora kinase A and B.

inhibitors with different CDK selectivity (dinaciclib, flavopiridol, milciclib or palbociclib, Suppl. Table 5), aurora kinase inhibitors (aurora Kinase Inhibitor I or TAK-901), proteasome inhibitors (MG132 or bortezomib) or EGFR inhibitors (erlotinib and gefitinib), after which the effect on proliferation was quantified using image-based analysis (Fig. 4A). Although the growth rates differed for the eight TNBC cell lines, there were neither subtype specific proliferation differences in either 2D or 3D cultures (Suppl. Fig. 7A-D) nor a correlation between growth rates in 2D and 3D cultures (Suppl. Fig. 7E). As expected by the decrease in expression of cell cycle related genes, all cell lines showed a significant decrease in growth rate in 3D cultures compared to 2D cultures (Suppl. Fig. 7F-G). Remarkably, this differential growth rate between 2D and 3D did not result in a significant change in sensitivity to anti-proliferative treatments such as CDK and aurora kinase inhibitors (Fig. 4B-C, Suppl. Fig. 8A-C). The same was true for the proteasome; although the expression of many proteasome factors was down-regulated in 3D, the sensitivity towards proteasome inhibitors MG132 and bortezomib did not change (Fig. 4B, Suppl. Fig. 8D).

EGFR is higher expressed in TNBC compared to ER positive BC and is another important drug target in BC treatment[404–406]. In contrast to CDKs, aurora kinases and proteasomal factors, EGFR was higher expressed in 3D for some basal cell lines (Suppl. Fig. 9A). Therefore we predicted that these cells would be more sensitive to EGFR inhibitors such as erlotinib and gefitinib. The majority of the basal cell lines used in this study were resistant to EGFR inhibitor treatment in 2D, except for HCC1806 and MDA-MB-468 (Fig. 4E). Interestingly, HCC1143, HCC1954 and SUM149PT cell lines became more sensitive to both erlotinib and gefitinib in 3D cultures, while BT549, Hs578T and MDA-MB-231 remained resistant. This increased sensitivity was likely not directly related to EGFR RNA expression or its signaling pathway (Suppl. Fig. 9A), since EGFR expression was also almost twice upregulated in the resistant MDA-MB-231 cell line. As shown before, basal cell lines in 3D cultures demonstrated differences in invasive behavior with a significantly higher invasion capacity for basal B (except for SUM149PT)



compared to basal A cell lines (Fig. 2C, Suppl. Fig. 9B-C). Interestingly, only the invasive cell lines remained resistant to EGFR inhibitors in 3D cultures. This was not related to differences in 3D growth rate (Suppl. Fig. 9D), EGFR expression (Suppl. Fig. 9E-F) or expression of genes involved in EGFR signaling (Suppl. Fig. 9G-H). A recent paper by *Savage et al* observed heterogeneous EGFR expression in breast cancer patient-derived xenografts (PDX) and showed that the sensitivity to EGFR inhibitors such as gefitinib is determined by the variation rather than the mean EGFR expression[407]. EGFR-high subpopulations were characterized by growth inhibition upon gefitinib treatment and enhanced expression of mesenchymal-stem cell

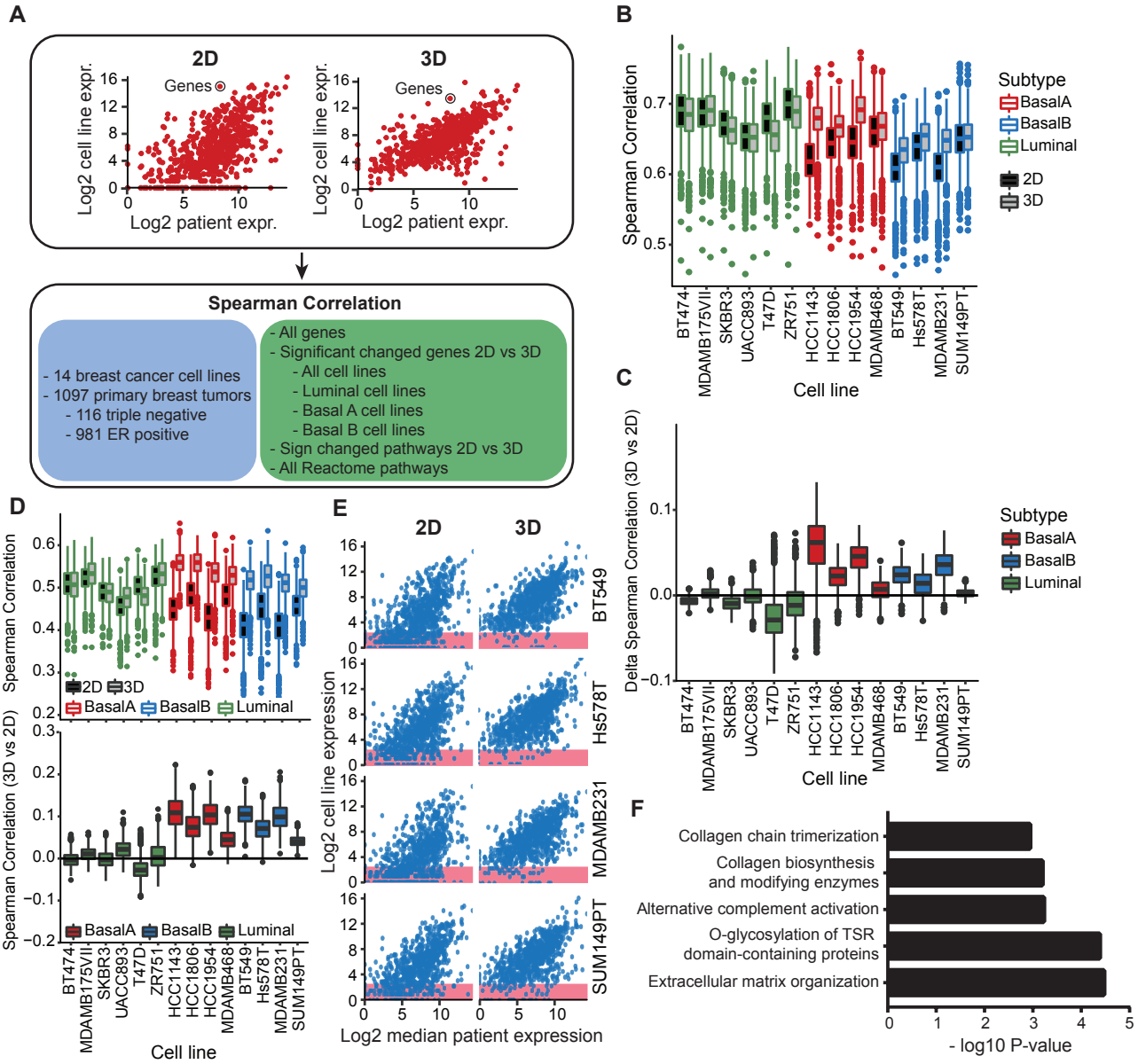
◄ ***Figure 4.*** **Effect of DEG inhibitors on cell line proliferation in 2D and 3D.** (A) Experimental setup. 4 basal A and 4 basal B cell lines were plated in 2D and 3D conditions, after which they were treated with 6 different concentrations of CDK, Aurora Kinase, Proteasome or EGFR inhibitors. 4 days after treatment, cells were fixed and proliferation was assessed. (B) Log IC50 (uM) for CDK inhibitors dinaciclib en flavopiridol and proteasome inhibitors bortezomib and MG132 in 2D and 3D cultures of basal A and basal B cell lines. (C) Dose response curves of dinaciclib en flavopiridol in 2D and 3D cultures of basal A and basal B cell lines. Mean + stdev for technical duplicates in 2D cultures. Mean + stdev for technical quadruplicates in 3D cultures. (D) Log IC50 of EGFR inhibitors erlotinib and gefitinib in basal A and basal B cell lines. (E) Dose-response curves of erlotinib and gefitinib in 2D and 3D cultures of basal A and basal B cell lines. Mean + stdev for technical duplicates in 2D cultures. Mean + stdev for technical quadruplicates in 3D cultures. (F) Heatmap (complete linkage, Euclidean distance) of log2 expression of mesenchymal-stem cell markers in basal A and basal B cell lines. Cell line color indicates 3D invasiveness; red = invasive, blue = non-invasive.

markers[407]. This relationship between EGFR and breast cancer stem cell properties has been demonstrated before[408,409]. Interestingly, the same mesenchymal-stem cell markers were also more highly expressed in non-invasive EGFR-inhibitor sensitive cells in both 2D (Figure 4F) and 3D cultures. We hypothesize that when losing oncogenic proliferative signaling pathways in 3D conditions, these cell lines became more dependent on EGFR related stem cell-mesenchymal pathways and therefore more vulnerable to EGFR inhibitor treatment.

Altogether, we demonstrated that differential gene expression does not always relate to changes in drug sensitivity, especially in case of cell cycle related genes. Non-invasive cells were sensitive to EGFR inhibitors in 3D systems irrespective of the EGFR expression levels, suggesting that other factors than single gene expression patterns might be involved in drug sensitivity.

Gene expression patterns of basal cell lines cultured in 3D are better correlated to human patient tumors than 2D cultured samples

For efficient drug development, it is vital to choose the optimal in vitro test system in which the regulation of pathways correlates as close as possible to the human primary tumor situation. To compare the translational value of both 2D and 3D culture systems for human patient BC tumors, we calculated the correlation between cell line and patient RNA expression levels derived from TCGA (Figure 5A). Correlations were calculated based on all genes and patients, but also focused on DEGs between 2D and 3D culturing systems, specific pathways or specific patient tumor subtypes (Figure 5A). Taking into consideration all genes and all patient tumors, luminal cell lines in 2D culture were better correlated to patient tumors than basal cell lines in 2D culture (Figure 5B). However, for the basal subtypes only, this correlation improved significantly for 3D cultures (Figure 5B-C). The same trend was observed when only analyzing ER positive tumors (Suppl. Fig. 10A). For TNBC, the correlation of basal and luminal cell lines was similar in 2D cultures, but much higher for basal (mainly basal A) cell lines in 3D cultures (Suppl. Fig. 10B). We next focused on the contribution of the DEGs in the correlation calculations (Suppl. Fig. 11 and Suppl. Fig. 12, Fig. 5D). Again, we only observed an improved



correlation in 3D conditions for the basal cell lines. Remarkably, the correlation of basal-specific DEGs showed a slightly larger increase in ER positive tumors compared to TNBC (Suppl. Fig. 11 and Suppl. Fig. 12) suggesting that the DEGs might have an important function in all patient tumors regardless of the subtype. Plotting the correlations of basal B subtype specific DEGs demonstrated that the correlation increase in 3D could be attributed to a small subset of DEGs; these genes were not or low expressed in 2D, but show increased levels in 3D (Fig. 5E, red box). Gene set enrichment analysis on this subset of genes revealed that the majority of these DEGs is involved in extracellular matrix organization and collagen formation (Fig. 5F), suggesting that studying these pathways in 3D cultures might be more relevant for human patient tumors. Furthermore, assessment of Reactome pathway expression correlations between all human patient tumors and 2D and 3D models emphasized that the optimal model is highly dependent on the pathway of interest (Suppl. Fig. 13A). As expected, the collagen formation pathway showed the best correlation in basal B cell lines, dominated by the Hs578T cell line (Suppl. Fig. 13B). However, the study of ERBB4 signaling might be preferably

◄ ***Figure 5.*** **Comparison the transcriptomes of 2D and 3D cell culture models with patient derived primary tumor transcriptomes.** (A) Analysis overview. Spearman correlations were calculated between the different transcriptomes using log2 normalized gene expression levels taking into consideration cell line subtypes, tumor subtypes and different gene sets. (B) Spearman correlation of log2 normalized RNA expression levels between cell lines and patient primary breast tumors (n = 1097). Boxes represent median and interquartile range of all tumor correlations. All genes and all tumors (irrespective of subtype) were included in the correlation calculations. (C) Difference in spearman correlation comparing 2D and 3D cell line cultures matched by patient tumor. All genes and all tumors were included in the correlation calculations. Boxes represent median and interquartile range of the difference in spearman correlation for all tumors. (D) Spearman correlation of log2 normalized RNA expression levels of basal B DEGs (paired analysis, $p < 0.01$) between cell lines and patient primary breast tumors (n = 1097). Boxes represent median and interquartile range of all tumor correlations. Only DEGs comparing 2D and 3D cultures for the basal B cell lines were included in the correlation calculations (top). Difference in spearman correlation comparing 2D and 3D cell line cultures matched by patient tumor. Only DEGs comparing 2D and 3D cultures for the basal B cell lines were included in the correlation calculations (bottom). Boxes represent median and interquartile range of the difference in spearman correlation for all tumors. (E) Log2 normalized RNA expression levels of basal B DEGs comparing 2D to 3D cultures (paired analysis, $p < 0.01$) in basal B cell lines and TCGA patient primary breast tumors. Red squares: genes with log2 RNA expression < 2. (F) Top 5 pathways of over-representation analysis of basal B DEGs comparing 2D to 3D cultures (paired analysis, $p < 0.01$) not or low expressed in 2D cultures and higher expressed in 3D cultures.

performed in luminal cell lines (Suppl. Fig. 13C), whereas for RNA polymerase II transcription 3D culture systems are preferred, irrespective of the cell line subtype. (Suppl. Fig. 13D).

Discussion

Although there are increasing efforts to use 3D tumor cell cultures to evaluate the efficacy of candidate target therapies, a systematic comparison between the cellular physiology of 2D and 3D cultures is still lacking. Here we compared the transcriptome of BC cell lines cultured in 2D and 3D conditions to 1) uncover the differential pathway expression linked to subtype-specific morphological changes and 2) identify the most suitable model for future drug screens. Consistent with previous comparisons of 2D and 3D cultures[410], we demonstrate that BC cells in a 3D matrigel environment have a decreased expression of genes important in cell cycle pathways, including various CDKs and aurora kinases, regardless of the subtype. Of relevance, despite the downregulation of these gene sets in 3D, we did not observe a differential potency of CDK and aurora kinase inhibitors comparing 2D and 3D cultures. Regardless of this lack of potency, the gene expression patterns of 3D cultures better resembled the BC patient transcriptome, in particular for TNBC subtypes.

Mitochondria-related genes were upregulated in 3D conditions, underlining the important effects of cell culture conditions on the cell metabolic phenotype. Deregulating cellular energetics is

one of the hallmarks of cancer[21] and possibilities for targeting oxidative phosphorylation as anti-cancer strategy are currently widely exploited[411]. Interestingly, mtDNA copy numbers are often decreased in BC[412] and low numbers are related to more metastasis formation[413,414]. Future studies should uncover if the elevated levels of mitochondrial genes in 3D cultures are part of a rewired metabolic program and more important, whether this change is representative for patient primary tumors.

We observed that cell lines of the basal A subtype showed a clear non-invasive spheroid morphology in 3D. In contrast, cell lines of the basal B subtype demonstrated a distinct invasive phenotype; an exception for the basal B phenotype was the SUM149PT, that shows ambiguous classification and has also sometimes been classified as basal A[42,390,391]. These contrasting phenotypes between basal A and basal B cells are likely related to differences in expression of a small subset of extracellular matrix interacting and cell-cell adhesion genes. Thus, the basal B cells upregulated genes involved in collagen biosynthesis and breakdown of the extracellular matrix in 3D cultures, while these genes were downregulated in basal A cells. Moreover, the higher basal levels of E-cadherin in the basal A subtype could contribute to the less invasive phenotype observed in 3D cultures.

CDKs, aurora kinases and proteasomal factors that were higher expressed in 2D, showed genuine higher expression levels in more aggressive breast tumors and cell lines, and were positively related to metastatic disease in breast cancer patients. Cell cycle-related CDKs are involved in almost all phases of cell cycle regulations, while aurora kinases are specifically involved in mitosis. Aurora kinases are associated with poor cancer outcome and often over-expressed in cancer, amongst which breast cancer[396]. Several aurora kinase inhibitors and Cell cycle-related CDK inhibitors have been developed of which some show promising results in clinical trials and are almost facing towards the clinic[395–398]. Next, cancer cells might be more dependent on the proteasome for the elimination of abnormal proteins due to high proliferation rate and genetic instability[399,400]. Moreover, the proteasome is also involved in the DNA damage response and DNA repair mechanisms[415,416]. Therefore, proteasomal inhibition might disrupt essential cellular pathways resulting in cancer cell death. To investigate whether transcriptomic differences change drug sensitivity and define the optimal in vitro model for future drug screening; we systematically treated cell lines in both 2D and 3D cultures with specific inhibitors against CDKs, aurora kinases and the proteasome. Whereas previous studies mainly identified reduced sensitivity towards established drugs (chemotherapeutics, endocrine agents and HER2-targeting agents) in 3D cultures in mostly HER2 positive cell lines[417–419], no change in sensitivity for CDKs, aurora kinases and proteasomal factors comparing 2D and 3D cultures was observed with our experimental setup. While the drugs that we tested were not evaluated in these studies, the discrepancies can also be explained by the differences in cell lines used and culture conditions. Regardless, we can conclude that in our culture systems drug sensitivity was not directly linked to specific drug-target gene expression levels. To determine the optimal future screening model to assess efficacy of new targeted therapies, a thorough understanding of the overall clinical translation of the test system for specific cancer types, including TNBC, will be essential.

For basal BC cell lines the transcriptome of 3D cultures demonstrated increased similarity towards human patient tumor transcriptomes, which was mainly caused by elevated levels of genes involved in extracellular matrix interactions. On the contrary, the transcriptome profile of luminal cell lines was already highly comparable to the patient tumor transcriptome in 2D conditions, but did not improve in 3D settings. Importantly, similarity of transcript levels was highly pathway and subtype dependent. For example, ERBB4 signaling in patient tumors is highly similar in luminal cell lines regardless of culture condition, while expression of the RNA polymerase II transcription pathway is better correlated to 3D cultures irrespective of cell line subtype. Next to the pathway of interest, the ideal in vitro model might change depending on the BC subtype of interest. In general, pathways in triple-negative patient tumors are better correlated to basal cell lines than the same pathways in ER-positive tumors.

Altogether, we can conclude that our systematic transcriptome analysis uncovered major differences between 2D and 3D models thereby increasing the global understanding of these in vitro test systems. Defining 'the best in vitro model' with efficient clinical translation is yet difficult and dependent on both the pathway and subtype of interest, although in general 3D cultures correlated better to human patient tumors. To support such a selection, we established an in silico web tool in to allow i) differential expression of genes between 2D and 3D BC cell lines, ii) collect gene expression levels in different breast cancer subtypes (cell lines as well as patient tumor data) and iii) calculate correlation between different cell lines/culture conditions with human patient tumors (https://gitlab.com/lacdr-dds/vandewater/koedoot_2d3d_rnaseq).

We performed a systematic comparison between 2D and 3D BC cell lines; we did not include other in vitro cell models that are of relevance for BC anticancer therapy development, including BC tumor organoids, short-term patient-derived xenograft cell cultures, stem cell reprogramming and tumors-on-a-chip[377,420–422]. Future systematic comparisons of the transcriptome of these test systems and their response to candidate drug targeted therapeutic approaches will be essential to provide clarity about the most optimal pre-clinical in vitro test systems.

Data availability

Supplemental Figures 1-13 and Supplemental Tables 1-5 are available upon request using the following link: https://doi.org/10.17026/dans-27n-r6bn. The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. Raw sequencing data will become publicly available upon publication and can currently be obtained by request.

Acknowledgements

This project was supported by the European Commission ERC Advanced grant Triple-BC (grant no. 322737) and ZonMW (grant no. 435002025).