IVYSPRING INTERNATIONAL PUBLISHER

Journal of Cancer

2017; 8(16): 3131-3141. doi: 10.7150/jca.18457

Review

Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping

Xiaofeng Dai[1, 2]✉, Hongye Cheng[1, 2], Zhonghu Bai[1, 2], Jia Li[1, 2]✉

1. National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi, China;
2. The Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China.

✉ Corresponding authors: Xiaofeng Dai, National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi 214122, China, 86-0510-8532-9306, xiaofeng.dai@jiangnan.edu.cn; Jia Li, National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi 214122, China, 86-0510-8519-7681, lijia030111@sibs.ac.cn

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (https://creativecommons.org/licenses/by-nc/4.0/). See http://ivyspring.com/terms for full terms and conditions.

Received: 2016.11.22; Accepted: 2017.08.29; Published: 2017.09.12

Abstract

Breast cancer cell lines have been widely used for breast cancer modelling which encompasses a panel of diseases with distinct phenotypical associations. Though cell lines provide unlimited homogenous materials for tumor studies and are relatively easy to culture, they are known to accumulate mutations duringthe initial establishment and subsequent series of cultivations. Thus, whether breast cancer cell line heterogeneity reflects that of carcinoma remains an important issue to resolve before drawing any reliable conclusion at the tumor level using cell lines. Inconsistent nomenclatures used for breast cancer cell line subtyping and the different number of subtypes grouped for cell lines and tumors make their direct matching elusive. By analyzing the molecular features of 92 breast cancer cell lines as documented by different literatures, we categorize 84 cell lines into 5 groups to be consistent with breast tumor classification. After combing through these cell lines, we summarized the molecular features, genetically and epigenetically, of each subtype, and manually documented 10 cell lines lacking explicit information on subtyping. Nine cell lines, either found inconsistent on their primary molecular features from different studies or being contaminated at the origin, are not suggested as the first choice for experimental use. We conclude that breast tumor cell lines, though having a high mutational frequency with many uncertainties and could not fully capture breast cancer heterogeneity, are feasible but crude models for tumors of the same subtype. New cell lines with enriched interferon regulated genes need to be established to enlarge the coverage of cell lines on tumor heterogeneity.

Key words: breast cancer, cell line, subtyping, molecular feature

Introduction

Breast cancer is highly heterogeneous encompassing a group of genetically and epigenetically distinct diseases exhibiting diverse clinical features [1]. A large proportion of current knowledge on breast carcinomas are derived from *in vivo* and *in vitro* studies performed using breast cancer cell lines, given that they could provide an unlimited source of homogenous self-replicating materials using simple yet standard media and approaches [2]. Thus, whether these cell lines well capture the molecular features and reflect the heterogeneity of corresponding tumors remains an important issue to resolve before obtaining any clinically relevant results. Though it is concluded that breast cancer cell lines are representative of breast carcinoma to a large extent, with ER and HER2 being important stratifiers for their classification, continuous evidences have suggested dramatic genetic and epigenetic changes during the initial cell line establishment and subsequent serial passaging, suggesting that the resultant cell lines may have evolved significantly from the primary tumors [3]. Also, different studies

categorize breast cancer cell lines into different groups [1, 4-8], complicating our understandings towards cell line classification and their relevance with tumors. For example, Birnbaum et al. grouped 27 breast cancer cell lines into luminal, basal and mesenchymal subtypes [7]; Riaz et al. characterized 5 subtypes, i. e. , luminal, luminal-HER2+, ER-negative-HER2+, basal, normal-like, among 51 breast cancer cell lines using a panel of 496 genes identified by Perou [1]; Lehmann et al. subdivided triple negative cell lines into 7 categories, namely, basal-like 1 (BL1), basal-like 2 (BL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL), luminal androgen receptor (LAR) and unstable (UNS) [9]; and some studies discriminate breast cancer cell lines into luminal-like and basal-like, with the basal class further subdivided into basal-A and basal-B [5, 8]. Given the inconsistent nomenclatures, classification and even contradictory molecular characterization in different literatures [1, 4-8, 10], we are overwhelmed with cell lines lacking systematic feature documentation and consistent subtyping. On the other hand, the number of cell lines widely used for breast cancer studies is extremely small, with cell lines such as MCF7, T47D and MDAMB231 accounting for more than two-thirds of cell lines used in the associated studies [2]. This raises the issue on how representative these few cell lines are of the vast diverse spectrum of breast tumors with distinct clinical implications. We are thus motivated to sort out the molecular features and corresponding tumor subtype each cell line represents to facilitate breast cancer modeling using appropriate cell lines.

Nomenclature of breast cancer cell lines

Ever since the establishment of the first breast cancer cell line, BT-20, in 1958 [11], relatively few cell lines have been obtained due to technical difficulties in extracting viable tumor cells from the surrounding stroma [12, 13] and the bottleneck of long-term propagation during cultivation [12, 14]. Most cell lines were established in late 1970s.

Cell line naming does not, in general, reflect its phenotypical association, but rather how they are established regarding, e. g. whether they are derived from the same laboratory, the same patient, isolated by serial subculture from the same initial population, or cultured using the same approach. For instance, 'HCC series' cell lines, as represented by their names, were isolated at Hamon Cancer Centre[12]; 'MDA series' were established from M. D. Anderson Hospital and Tumor Institute[14]; '21 series' were established from the same patient diagnosed as having infiltrating and intra-ductal mammary carcinoma, despite their distinct phenotypes and genotypes [15]. 'HMT series' underwent a successive subcultivation under various conditions during their establishment, including P53 mutation, MYC amplification, EGF-independence accompanied by tumorigenicity in nude mice, EGFR and HER2 over expression. 'SUM series', though isolated from different tumor specimens, were established using the same selective media [16]. As cell lines are typically named by the scientist who derived them, there is no rule on how each cell line is named especially for those do not belong to any series. It is until recently that standardization on cell line nomenclature has been proposed [17], which helps regularize the naming of newly established cell lines and ultimately improvement in cell line annotation and scientific reproducibility [18].

Molecular classification of breast cancer cell lines

Gene expression profiling has been widely applied to catalogue breast cancer cell lines, with diverse number of clusters being matched into various schemes. We characterized 84 cell lines based on the status of three important receptors conventionally used for breast cancer subtyping, i. e. , estrogen receptor (ER), progesterone receptor (PR), and human epithelial receptor 2 (HER2), and classified them using the same nomenclature, i. e. , luminal A, luminal B, HER2 positive, and triple negative subtypes, with triple negative cells being further divided into A and B to capture its heterogeneity and provide an easy link to the widely used names, i. e. , basal A and B (Table 1).

Though, in most studies, luminal cell lines are not further differentiated into subtypes A and B according to their HER2 status as tissue samples, a few reports have classified ER+HER2+ cell lines, such as MDAMB330, as the luminal-ERBB2+ subtype [1, 6]. However, the majority of luminal-ERBB2+ cell lines including, e. g. , HCC202, SKBR3, UACC893, SUM190PT, SUM225CWN, are ER-HER2+ at the molecular level, corresponding to HER2 positive tumor patients. While some of these cell lines are referred to as luminal (HCC202, SKBR3, UACC893) [1, 2, 5-8, 12, 19], some are grouped as basal or basal A (SUM190PT, SUM225CWN) [1, 5-7]. HER2+ cell lines are given a diverse spectrum of names despite their considerable overlaps, e. g. , luminal (AU565, HCC2218, OCUB-F, SKBR5, HCC202, SKBR3, UACC893) [1, 5, 8], luminal-HRBB2+ (HCC202, SKBR3, UACC893, SUM190PT, SUM225CWN) [1], ER-negative-ERBB2+ (HCC1008, HCC1569, HCC1954) [1], and basal A (SUM190PT, SUM225CWN, HCC1008, HCC1569, HCC1954) [5, 8]. Cell lines with triple negative status of ER, PR and

HER2 are differentiated as basal A and basal B cell lines in most literatures, with basal A being more luminal-like and basal B more basal-like. For example, while MDAMB468 is classified as basal A in [5, 8], it is described as weakly luminal in [2]; Hs578T, MDAMB231 are grouped as basal B in [5, 8] but characterized as claudin-low in [20] or even mesenchymal-like in [2, 7]. We mapped the cell line subtyping nomenclatures from different literatures [1, 4, 5, 7, 21] to the most commonly used breast cancer classification structure (Figure 1), and summarized the molecular features and morphological characteristics of each subtype (Table 2) with details provided below.

Table 1. Categorization, molecular information and culture conditions of 84 breast cancer cell lines, and the clinical features of tumors where they derive.

Cell lines	ER	PR	HER2	BRCA1 Mutation	Subtype	Series	Basic Medium	Tumor	Source	References
BT483	+	+/-	-	WT	LA	BT	RPMI	IDC	L	[1, 5, 6, 8, 10]
CAMA1	+	+/-	-	WT	LA	NA	DMEM	AC	L	[1, 5, 6, 10]
EFM19	+	+	-	ND	LA	EFM	RPMI	IDC	L	[8, 10]
HCC1428	+	+	-	ND	LA	HCC	RPMI	AC	L	[5, 8]
HCC712	+	+/-	-	ND	LA	HCC	RPMI	DC	L	[8]
IBEP2	+	-	-	ND	LA	NA	DMEM	IDC	L	[10]
KPL1	+	-	-	ND	LA	NA	RPMI	IDC	L	[10, 61]
LY2	+	-	-	ND	LA	NA	DMEM	IDC	L	[5]
MCF7	+	+	-	WT	LA	MCF	RPMI, DMEM	IDC	L	[1, 5, 6, 8, 10, 58]
MDAMB134	+	-	-	ND	LA	MDA	RPMI	IDC	L	[8]
MDAMB134VI	+	-	-	WT	LA	MDA	DMEM	IDC	L	[1, 5, 6, 10]
MDAMB175	+	-	-	ND	LA	MDA	RPMI	IDC	L	[8]
MDAMB175VII	+	-	-	WT	LA	MDA	DMEM	IDC	L	[1, 5, 6, 10]
MDAMB415	+	+/-	-	WT	LA	MDA	DMEM	AC	L	[5, 6, 10]
T47D	+	+	-	WT	LA	NA	RPMI	IDC	L	[1, 5, 6, 8, 10]
ZR751	+	+/-	-	WT	LA	ZR75	RPMI	IDC	L	[1, 5, 6, 8, 10]
ZR75B	+	-	-	ND	LA	ZR75	RPMI	NA	L	[5]
BSMZ	+	+	+	ND	LB	NA	RPMI	IDC	L	[10, 19]
BT474	+	+	+	WT	LB	BT	RPMI	IDC	L	[1, 5, 6, 8, 10]
EFM192A	+	+	+	ND	LB	EFM	RPMI	AC	L	[8]
IBEP1	-	+	+	ND	LB	IBEP	DMEM	IDC	L	[10]
IBEP3	-	+	+	ND	LB	IBEP	DMEM	IDC	L	[10]
MDAMB330	+	-	+	WT	LB	MDA	RPMI	ILC	L	[1, 6, 10]
MDAMB361	+	+/-	+	WT	LB	MDA	RPMI, DMEM	AC	L	[1, 5, 6, 8, 10]
UACC812	+	+/-	+	WT	LB	UACC	RPMI, DMEM	IDC	L	[1, 5, 6, 8, 10]
ZR7527	+	-	+	WT	LB	ZR75	RPMI	IDC	L	[60]
ZR7530	+	-	+	WT	LB	ZR75	RPMI	IDC	L	[1, 5, 6, 8, 10]
21MT1	-	+/-	+	ND	H	21	α-MEM/DFC1	IDC	L	[15]
21MT2	-	+/-	+	ND	H	21	α-MEM/DFC1	IDC	L	[15]
21NT	-	+/-	+	ND	H	21	α-MEM/DFC1	IDC	L	[15]
21PT	-	+/-	+	ND	H	21	α-MEM/DFC1	IDC	L	[15]
AU565	-	-	+	WT	H	NA	RPMI	AC	L	[5, 10, 58]
HCC1008	-	-	+	ND	H	HCC	RPMI	IDC	L	[12]
HCC1569	-	-	+	WT	H	HCC	RPMI	MC	L	[1, 5, 8]
HCC1954	-	-	+	WT	H	HCC	RPMI	DC	L	[1, 5, 8]
HCC202	-	-	+	WT	H	HCC	RPMI	DC	L	[1, 5, 8]
HCC2218	-	-	+	ND	H	HCC	RPMI	DC	L	[8]
HH315	-	-	+	ND	H	HH	RPMI	C	L	[10, 59]
HH375	-	-	+	ND	H	HH	RPMI	C	L	[10, 59]
KPL-4	-	-	+	WT	H	KPL	DMEM	IDC	L	[10, 62]
MDAMB453	-	-	+	WT	H	MDA	RPMI, DMEM	AC	L	[1, 5, 6, 8, 10]
OCUB-F	-	-	+	WT	H	NA	RPMI	NA	L	[1, 6]
SKBR3	-	-	+	WT	H	SKBR	RPMI, McCoys	AC	L	[1, 5, 6, 8, 10]
SKBR5	-	-	+	WT	H	SKBR	RPMI	AC	L	[1, 6]
SUM190PT	-	-	+	WT	H	SUM	Ham's F12	Inf	L	[1, 5, 6, 8]
SUM225CWN	-	-	+	WT	H	SUM	Ham's F12	IDC	L	[1, 5, 6]
UACC893	-	-	+	WT	H	UACC	RPMI	IDC	L	[1, 6, 8, 10]
BT20	-	-	-	WT	TNA	BT	RPMI, DMEM	IDC	L	[1, 5, 6, 8, 10]
CAL148	-	-	-	WT	TNA	CAL	DMEM	AC	D	[71]
DU4475	-	-	-	WT	TNA	NA	RPMI	IDC	D	[1, 6, 10]
EMG3	-	-	-	ND	TNA	NA	DMEM	IDC	L	[25, 48, 57]
HCC1143	-	-	-	ND	TNA	HCC	RPMI	DC	L	[1, 5, 8]
HCC1187	-	-	-	ND	TNA	HCC	RPMI	DC	L	[5, 8]
HCC1599	-	-	-	ND	TNA	HCC	RPMI	DC	L	[8]

HCC1806	-	-	-	ND	TNA	HCC	RPMI	SqC	L	[1, 8]
HCC1937	-	-	-	MU	TNA	HCC	RPMI	DC	L	[1, 5, 6, 8]
HCC2157	-	-	-	ND	TNA	HCC	RPMI	DC	L	[5, 8]
HCC3153	-	-	-	MU	TNA	HCC	RPMI	DC	L	[5, 8]
HCC70	-	-	-	WT	TNA	HCC	RPMI	DC	L	[1, 5, 8]
HMT3522	-	-	-	WT	TNA	HMT	DMEM, F12	B	L	[49]
KPL-3C	-	-	-	ND	TNA	KPL	RPMI	IDC	D	[10, 36]
MA11	-	-	-	ND	TNA	NA	DMEM	ILC	D	[10, 37, 63]
MDAMB435	-	-	-	WT	TNA	MDA	DMEM	AC	L	[1, 5, 6, 10]
MDAMB436	-	-	-	MU	TNA	MDA	RPMI, L15	AC	L	[1, 5, 6, 8, 10]
MDAMB468	-	-	-	WT	TNA	MDA	RPMI, L15	AC	L	[1, 5, 6, 8, 10]
MFM223	-	-	-	WT	TNA	NA	MEM	C	D	[71]
SUM185PE	-	-	-	WT	TNA	SUM	Ham's F12	DC	L	[1, 5, 6]
SUM229PE	-	-	-	WT	TNA	SUM	RPMI	DC	L	[1, 6]
BT549	-	-	-	WT	TNB	BT	RPMI	IDC	L	[1, 5, 6, 8, 10]
CAL120	-	-	-	WT	TNB	CAL	DMEM	AC	D	[71]
CAL51	-	-	-	WT	TNB	CAL	DMEM	AC	D	[71]
CAL851	-	-	-	WT	TNB	CAL	DMEM	AC	D	[71]
HCC1395	-	-	-	ND	TNB	HCC	RPMI	DC	L	[1, 8]
HCC1739	-	-	-	ND	TNB	HCC	RPMI	DC	D	[12]
HCC38	-	-	-	ND	TNB	HCC	RPMI	DC	L	[1, 5, 8]
HDQ-P1	-	-	-	MU	TNB	NA	DMEM	IDC	D	[71]
Hs578T	-	-	-	WT	TNB	NA	RPMI, DMEM	IDC	L	[1, 5, 6, 8, 10]
MDAMB157	-	-	-	WT	TNB	MDA	RPMI, DMEM	MC	L	[1, 5, 6, 8, 10]
MDAMB231	-	-	-	WT	TNB	MDA	RPMI, DMEM	AC	L	[1, 5, 6, 8, 10]
SKBR7	-	-	-	WT	TNB	SKBR	RPMI	AC	L	[1, 6]
SUM102PT	-	-	-	WT	TNB	SUM	Ham's F12	IDC	L	[1, 6, 8]
SUM1315M02	-	-	-	MU	TNB	SUM	Ham's F12	IDC	L	[1, 5, 6]
SUM149PT	-	-	-	MU	TNB	SUM	Ham's F12	InfDC	L	[1, 5, 6, 8]
SUM159PT	-	-	-	WT	TNB	SUM	Ham's F12	AnC	L	[1, 5, 6]

The status on ER, PR and HER2 are the primary features used for breast cancer cell line subtyping [1, 5, 6, 8, 10, 12, 15, 19, 25, 36, 37, 48, 49, 57-63]. BRCA1 mutation status is annotated for the ease of sporadic core basal tumor modeling, where MU represents cell lines with 'BRCA1 mutation', WT means 'wild type', and ND is short for 'not decided' [4, 6, 64, 65]. We categorize these cell lines into luminal A (LA), luminal B (LB), HER2 positive (H), Triple negative A (TNA) and Triple negative B (TNB), according to literature available subtyping suggestions (L) or derived information from the original article where they are established (D). Such information are provided in the 'Subtype' and 'Source' columns, respectively, with references listed in column 'References'. The 'Series' column shows which cell line series each cell line comes from which are conventionally given by the establisher and represents, e. g. , the place where the cell line is created [12, 14, 15, 49, 60, 66-70]. 'Basic Medium' shows the common or suggested medium for the cultivation of each cell line, which may need the addition of other elements and does not eliminate the feasibility of other medium [1, 5, 8, 15, 66]. The 'Tumor' column shows the clinical property of the tumor where the cell lines are derived, with the following symbolic representations, i. e. , AC is 'adenocarcinoma', AnC is 'anaplastic carcinoma', B is 'benign tumor', C is 'carcinoma', CS is 'carcinosarcoma', DC is 'ductal carcinoma', IDC is 'invasive ductal carcinoma', ILC is 'invasive lobular carcinoma', InfC is 'inflammatory carcinoma', InfDC is 'inflammatory ductal carcinoma', MC is 'medullary carcinoma', SqC is 'squamous carcinoma', and NA means such information is not available [5, 8, 10, 15].

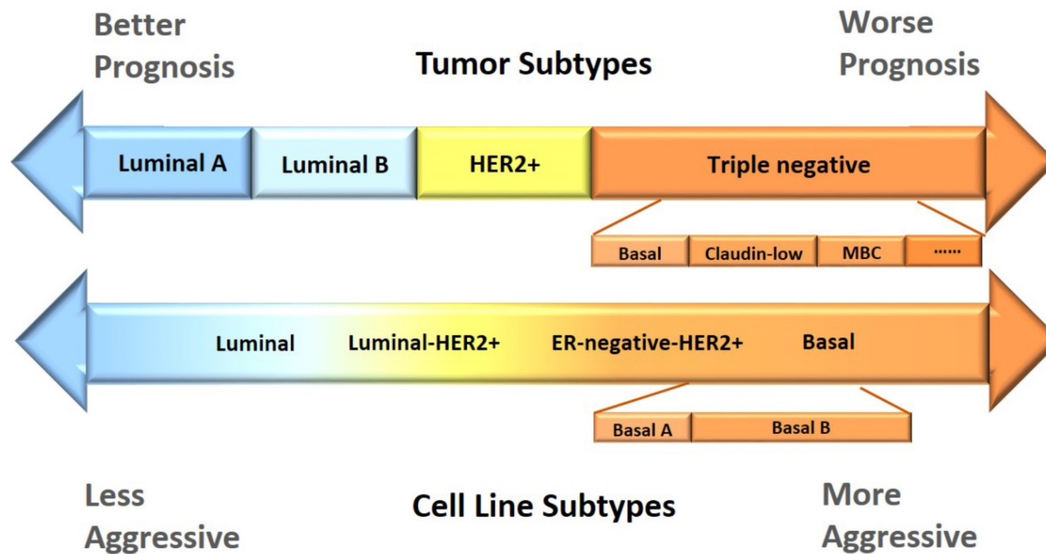

Figure 1. Comparison of the current subtyping schemes between breast cancer cell lines and tumors. According to the status of ER, PR, HER2, breast cancer is classified as luminal A, luminal B, HER2 positive, and triple negative, where triple negative tumors can be further differentiated into at least basal, claudin-low, MBC (metaplastic breast cancer) and interferon-rich. The current literatures, in general, do not differentiate luminal cell lines; HER2 positive cell lines are diffused into luminal B and triple negative cells, and named luminal-HER2+ and ER-negative-HER2+, respectively; and triple negative cells are called the 'Basal' subtype, with basal A and basal B being further differentiated to represent the basal tumors and claudin-low and/or MBC tumors, respectively. The morphological features of the subtypes in tumors and cell lines accord well, with luminal tumors having better prognosis and luminal cell lines less aggressive than that in triple negative tumors and cell lines.

Table 2. Molecular and morphological features of breast cancer cell line subtypes. The panels of mRNAs, miRNAs, and protein markers used for the identification of each breast cancer cell line subtype, as well as their morphological differences are summarized. In triple negative cell lines, the markers shared between sublines A and B are presented separately in the mRNA and protein columns. The literatures where such information is derived are provided right after.

Cell line subtype		mRNA		miRNA [1]	Protein		Morphology
Luminal		ER, GATA3, KRT19/KRT8/KRT18, XBP1, PBX1, ZNF278, SPDEF, CRABP2, MUC1, FOXA1, MYB, RET, EGR3, TFF1, HER3, TOB1, TFF3[1, 5, 7, 8]		hsa-miR-501-5p hsa-miR-202 hsa-miR-760 hsa-miR-626	ER, GATA3, KRT19 [7]		More differentiated; tight cell-cell junctions [5]
HER2 positive		HER2, GRB7, PERLD1, STARD3, C17ORF37[7]		hsa-let-7b hsa-miR-640 hsa-miR-200c hsa-miR-378 hsa-miR-141 hsa-miR-196a hsa-miR-29c hsa-miR-18a*	HER2		Breakdown of cell–cell junctions [10]
Triple negative	Triple negative A (Basal A)	EGFR CAV1/2 MSN ETS1 [5, 7, 8]	KRT4/5/6A/6B/13/14/15/16/17, ITGA6, ITGB4/6, LAMB3, LAMC2, TRIM29, S100A2, SLPI, LYN, ANXA8, COL17A1, BNC1, MET, CD133, GABRK, VTCN1, BST2, FABP7, CD10/14/58/59, [1, 7, 8]	hsa-miR-492 hsa-miR-26b hsa-miR-617 hsa-miR-155	EGFR CAV1/2 MSN ETS1 [7]	KRT5/6 CD10 MET [7]	Core basal-like
	Triple negative B (Basal B)		VIM, SPARC, FN1, FBN1, HAS2, PRG1, COL3A1, COL6A1/2/3, COL8A1, MMP2/14, TIMP1, CTSC, PLAU, PLAUR, AXL, PLAT, CD24(-), CD44, TGFBR2, SERPINE1/2, TGFB1[5, 7, 8]	hsa-miR-22 hsa-miR-532-3p hsa-miR-125b hsa-miR-501-5p hsa-miR-155*		CD44 [7]	Least differentiated and most stemness; more mesenchymal-like appearance and highly invasive [5]

Luminal breast cancer cell lines

Luminal breast cancer cell lines are characterized by positive ER and/or PR expression, despite the existence of a few exceptional cases, such as IBEP-1 and IBEP-3 [22], where PR positivity drives their luminal phenotype. This type of cell lines exhibits high expression of a panel of luminal feature associated genes/proteins including, e.g., ESR1 (ERα or ER), luminal keratins (KRT8/18/19), and transcription factors such as GATA3 and FOXA1. A systematic study on miRNA expression profiling has unveiled distinctive over-expression of has-miR-501-5p, has-miR-202, has-miR-760 and has-miR-626 in luminal cell lines [1]. Luminal cell lines are comparably more differentiated and have less propensity for migration due to tight cell-cell junctions, consistent with that at the tumor level.

Though the majority of the studies do not further stratify luminal cell lines into luminal A and B subtypes according to their HER2 status, we embrace such a differentiation not only for the sake of achieving consistent categorization with tumor subtyping to facilitate easy tumor modelling but also satisfying the requirement of drug response assays hinging on ER and HER2 status. For example, a study using BT474 (ER+HER2+) has revealed the synergistic advantage of tamoxifen and Herceptin in the treatment of breast tumors [23], and MCF7 (ER-HER2) has been conventionally used for testing tamoxifen-induced cell response [24]. Luminal B cell lines are, in principle, more invasive and consequently more aggressive than luminal A cells, as HER2 over-expression is shown to be associated with ER down-regulation [2]. Thus, luminal B cells participate in the phenotype attenuation of luminal A cells, and could be more appropriately described as weakly luminal [2]. Lots of information for luminal B tumors has been successfully deciphered using cell lines of this subtype with translational potential at clinics [25-28]. For example, the expression of quiescin-sulfhydryl oxidase 1 is reported to be associated with a poor prognosis in luminal B tumors using a ER+HER2+ cell line, BT474 [29], suggesting the usefulness and importance of separating luminal B cell lines from the luminal subtype.

HER2 positive breast cancer cell lines

HER2 positive cell lines, as featured by ER negativity and HER2 positivity, share the same over-represented genomic profile on chromosomal region 17q12 encompassing genes including HER2, GRB7, PERLD1, STARD3 and C17ORF37 [30]. Micro RNAs including has-let-7b, has-miR-640, has-miR-200c, has-miR-378, has-miR-141, has-miR-196a, has-miR-29c and has-miR-18a* are overtly expressed in these cell lines [1].

HER2 positive cell lines bridge the gap between luminal and basal cell lines, and are heterogeneous encompassing both luminal and basal features. They are grouped as luminal-ERBB2+ and ER-negative-ERBB2+ in [1] according to the

expression of luminal and basal markers besides ER and HER2 expression. Consistent with their molecular features, cells of this subtype are more aggressive with respect to cell migration as compared with luminal cells, since HER2 over-expression is associated with the breakdown of cell–cell junctions [2]; and more responsive to certain drugs and is an excellent model for the examination of Herceptin response, as a significant correlation is revealed between the molecular signature and biological response by Pearson's correlation test [5]. Proteins whose expression are strongly correlated with such a drug response include increased levels of ESR1, MAPK1/3, MEK, TYK2, FASN, and GRB7, which are mostly associated with cell proliferation, and up-regulated expression of SFN, CAV2, GRB2, RB1 and FLNA indicates a drug resistance [5]. Thus, by studying cell lines over-represented by HER2 expression, MAPK signaling is shown to predict Herceptin response, and mTOR pathway, Toll-like receptor pathway, N-glycan biosynthesis as well as inositol-phosphate signaling are associated with Herceptin resistance [5].

Triple negative breast cancer cell lines

Triple negative cell lines, as represented by the name, are featured by low or no expression of all three markers, i.e., ER-PR-HER2-. It is the most heterogeneous among all subtypes and referred to as basal A and basal B cell lines in many literatures. Triple negative A (basal A) lines are called basal-like as they are enriched with basal markers including cytokeratins (KRT4/5/6A/6B/13/14/15/16/17), integrins (ITGA6, ITGB4/6), LAMB3, LAMC2, TRIM29, S100A2, SLPI, ANXA8, COL17A1, BNC1, CD10/14/58/59, MET, LYN, CD133, GABRK, VTCN1, BST2, FABP7 [1, 7, 8], and resemble the core basal tumor subtype. Triple negative B (basal B) lines, designated the mesenchymal cluster or normal-like/claudin-low, over-express genes associated with tumor invasive and aggressive features such as VIM, MSN, PLAT, TGFB1, TGFBR2, AXL, COL3A1, COL6A1/2/3, COL8A1, MMP2/14, TIMP1, CTSC, PLAU, PLAUR, SERPINE1/2, SPARC, FN1, FBN1, HAS2, PRG1 [5, 7, 8]; and cancer stemness such as CD44(+) and CD24(-) [5]. Collagens (COL3A1, COL6A1/2/3, COL8A1), proteases (MMP2/14, TIMP1, CTSC, PLAU, PLAUR, SERPINE1/2, PLAT), and proteins stabilizing cytoskeletal interactions (VIM, MSN) are important players for extracellular matrix remodelling required for cell migration, and signalling factors (TGFB1, TGFBR2, AXL) are crucial in mediating such aggressive morphology. A selected panel of proteins (EGFR, CAV1/2, MSN, ETS1) are used for triple negative cell line characterization at the translational level as a whole, which characterize triple negative A cells when combined with basal keratins (KRT5/6), CD10, and MET, and identify triple negative B cells when jointly assessed with the stemness marker CD44 [7]. Distinct microRNAs are detected for triple negative A and B cells. In particular, hsa-miR-492, has-miR-26b, has-miR-617, has-miT-155 are overtly expressed in the triple negative A subtype, and hsa-miR-22, hsa-532-3p, hsa-miR-125b, hsa-miR-501-5p, hsa-miR-155* pop up in the triple negative B subtype [23]. Interestingly, hsa-miR-155 (triple negative A) and hsa-miR-155* (triple negative B) are derived from the same precursor, but show an opposing expression pattern in these two triple negative subtypes [1]. Phenotypically, triple negative A cells, as a comparatively more differentiated subtype within triple negative cell lines, may have either luminal-like or basal-like morphologies, whereas triple negative B cells has a more mesenchymal-like appearance and are more likely invasive. Thereby, triple negative A lines mostly resemble the core basal tumor subtype as featured by basal markers, and triple negative B cells could be used for claudin-low or metaplastic breast cancer modelling given their enrichment in epithelial mesenchymal transition (EMT) and stem-cell markers [31].

BRCA1 encodes a protein forming the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage [32]. Mutation of this gene predisposes hereditary breast cancer, which also resembles sporadic core basal tumors [33-35]. Triple negative A lines are characterized by BRCA1 signatures [8], with most of the currently available BRCA1 mutated commercial breast cancer cell lines (HCC1937, MDAMB436, SUM149PT, HCC3153) belonging to this subtype except for SUM1315MO2 (Table 1).

There are ten triple negative cell lines, DU4475, KPL-3C, MA11, HCC1739, CAL148, MFM223, CAL120, CAL51, CAL851, HDQ-P1, lacking direct information for further discrimination. We infer DU4475, KPL-3C, MA11, CAL148, MFM223 as triple negative A, and the rest as triple negative B based on the relevant literatures [5, 6, 12, 20, 36, 37, 39-41]. DU4475 cells carry a mutation in the gene MAP2K4 whose alteration is characteristic of luminal features [6], agreeing with the fact that triple negative A cells may have either luminal-like or basal-like morphologies [5]. Immunohistochemical studies reveal the expression of keratins but not vimentin in KPL-3C [36], fulfilling our classification criteria of triple negative A cells (Table 2). MA11 is characterized by MUC1 secretion [37], whose expression occurs at a high frequency in early stage

basal A triple negative breast cancer [20]. CAL148 and MFM223 are classified as the triple negative A cell line given their negative expression on vimentin [39]. CC1739 is a poorly differentiated cell line [12], satisfying the stemness feature of triple negative B cells. CAL120, CAL51 [40], CAL851 and HDQ-P1 [41] are positive on vimentin expression, falling into the triple negative B group. Our categorization of these cell lines is in agreement with other studies where different classification schemes were used. For example, Lehmann et al. categorized CAL148 and MFM223 to the LAR subtype which are heavily enriched with hormonally regulated pathways [9], suggesting their close relationship with hormone responsive cell lines and corroborating our classification of them as triple negative A than triple negative B cells.

Breast cancer cell lines are feasible models of breast cancer

The occurrence of DNA alterations was higher in breast cancer cell lines than in tumors. On average, alterations in cell lines are about two-fold more frequent than those in tumors [2, 21]. This might be because that breast cancer cell lines are mostly originated from invasive high-grade tumors, which are more easily to accumulate genomic aberrations during *in vitro* cultivation. Breast cancer cell lines may exhibit DNA alterations not detected in tumors, which may result from the sequential cultivation or reflect low frequency alteration in tumors that is hard to observe. Such newly gained mutations may cause phenotypical changes to cells. For example, MCF7 cell lines with variable sensitivities to tamoxifen are reported [42]. However, CGH analyses suggest that breast cancer cell lines are representative of tumors with regard to major DNA alterations [21]. A comprehensive study comparing the HCC series of cell lines with their original tumors reveals that almost all aberrations (except one) presented in a tumor also occur in the corresponding cell line [43]. This has been confirmed by another study comparing allelic losses at 18 chromosomal regions frequently deleted in breast tumors using 51 polymorphic micro-satellite markers [20]. Also, it seems that DNA alterations affecting specific genes under our interest remain highly consistent between cell lines and tissues. For example, a 75% concordance was seen on the mutation status in exons 5 to 10 of TP53 gene [44]. Point mutations identified in tumors were often found in at least one or some cell lines and vice versa [2]. At the chromosomal level, breast cancer cell lines, though, have a higher copy number changes, share a similar pattern regarding gains and losses with tumors [2, 21].

Epigenetic alterations are heritable modifications of gene expression that do not involve mutation, which are similar between breast cancer cell lines and tumors. A study has revealed that the CpG island methylation pattern is very close between cell lines and tumors. In their study, most of the genes harboring CpG island hypermethylation in the promoter region are present in both breast cancer cell lines and tumors such as ER, PR, HIC1, APC, ARHI, ASC, BRCA1, CCND2, CDH1, CDH13, CDKN2A, FABP3, FHIT, GIB2, GPC3, GSTP1, HOXA5, HSHIN1, KLK10, NME1, PRDM2, PRKCDBP, RARB, RASSF1, SFN, SYK, TFF1, TIMP3 and WT1 [45]; and quite a few cases show inconsistency between cell lines and tumors, e. g. , IL6, GSN, PLAU, PRSS8, SLC19A1, SNCG are hypermethylated in tumors but not in cell lines and the hypermethylation of PLAGL1, TGFB3 is observed in cell lines but not tumors [45].

At the molecular level, discriminative markers and associated phenotypic traits observed in breast cancer cell lines are also frequently discriminative features in tumors. Genes correlated to the ER+ or ER- phenotype in breast cancer cell lines are positively or negatively correlated to ER expression in tumors. The mRNA variants of ER containing precise truncations in various exons have been identified in both cell lines and tumors [46].

Breast cancer cell lines are crude models of breast cancer

Despite the considerable roles played by breast cancer cell lines in deciphering the mechanisms underlying tumor initiation and evolution, cell lines are still viewed as crude models of tumors and could not capture all tumor features and heterogeneity.

The clonal population comprising of any single cell line could not well capture the heterogeneity of breast tumors at the intra-tumoral level. Carcinogenesis is a sequential process through many clinical and pathological stages including a typical hyperproliferation, local invasion, invasive carcinomas and ultimately metastatic disease, and is believed to be accompanied with a sequential acquisition of various genetic and epigenetic alterations in a single cell followed by clonal selection and expansion. The establishment of cell lines, on the other hand, may eliminate some types of tumor cells initially present in the tissue samples under the environmental stress created by the cultivation condition. For instance, cells unable to grow properly on plastic or need specific factors secreted by tumor microenvironment may be eliminated during the cultivation process. Thereby, how well a permanent breast cancer cell line could model the tumor intra-heterogeneity becomes a questionable issue.

Success in long-term propagation has been a limiting factor during the establishment of breast cancer cell lines, resulting in less commercially available cell lines than needed to cover the inter-tumoral cancer heterogeneity. Take triple negative breast cancer as an example, it encompasses at least four subtypes, i.e., core basal, claudin-low, metaplastic breast cancer, and interferon-rich, each with distinct molecular features and clinical associations. However, we only have triple negative A and B cell lines. While type A may be well representative of the core basal tumor subtype, and type B is useful for the modelling of claudin-low and/or metaplastic subtype (further differentiation among type B lines is needed), we are left with no suitable cell line for interferon-rich tumor modelling according to their molecular features. Also and importantly, very few breast cancer cell lines (e. g. , MCF7, T-47D, MDAMB231), despite the total number of established ones, have been frequently used for research purpose for the sake of, e. g. , cultivation easiness, rendering the transportability of results obtained from such limited number of cell lines to the evolving tumors even more questionable.

Given the technical difficulties of extracting viable tumor cells from their surrounding stroma, most breast cancer cell lines are originated from invasive carcinoma (Table 1), raising the questions as to how they are representative of the primary tumors. Also, we could not exclude the possibility that certain breast cancer cell line, either from origin or due to cultivative selection, are not representative of breast cancer cell at all. For example, MDAMB435, previously identified as a breast cancer cell line, has been suggested as being originated from an occult melanoma [47].

Also important is how much tumor microenvironment affects cell signaling, and how much these changes alter cell molecular features and affect the experimental results. A study examining the influences of fibroblasts on cell line morphological changes using EMG3 reveals that fibroblasts could stimulate the expression of luminal keratins in basal cells and/or basal keratins in luminal cells [48]. Also reported is that cells developed in an environment with high/low EGFR activity tend to become ER negative/positive [49], suggestive of the selective power imposed by microenvironment on cells. This makes experiments convolving tumor microenvironment difficult to be modeled using cell lines.

Another concern regarding the usage of breast cancer cell lines in tumor modelling is that the same cell line, once cultured in different labs and/or under different conditions, may evolve into distinct populations. We do observe different categorizations of the same cell line into distinct groups according to their varied molecular or morphological descriptions annotated by different laboratories. Take the status of ER and HER2 for instance, HCC1007 is ER+HER2+ in [5], ER+HER2+ in [8], and ER-HER2+ in [12]; HCC1419 is ER+HER2+ in [1] but ER-HER2+ in [8, 12]; HCC1500 is ER+HER2- in [1, 8, 12] and ER-HER2- in [5]; HCC2185 is ER-HER2- in [5, 8] and ER-HER2+ in [12]; SUM52PE and SUM44PE are ER+HER2+ in [8] but ER+HER2- in [1, 5, 6]; EVSA-T is ER-HER2+ in [6] but ER-HER2- in [1]; and MPE600 is ER+HER2- in [5] but ER+HER2+ in [1, 6] (Table 3).

This renders not only studies hinging on cell lines difficult to be compared with but also our understandings towards these cell lines complicated and confusing at the first hand.

Table 3. Eight breast cancer cell lines with inconsistent annotations on the status of primary markers.

Cell lines	ER	PR	HER2	Subtype	Reference
HCC1007	+	-	-	LA	[5]
HCC1007	+	-	+	LB	[8]
HCC1007	-	-	+	H	[12]
HCC1419	+	-	+	LB	[1]
HCC1419	-	-	+	H	[8, 12]
HCC1500	+	+	-	LA	[1, 8, 12]
HCC1500	-	-	-	TNB	[5]
HCC2185	-	-	-	TNA	[5, 8]
HCC2185	-	-	+	H	[12]
SUM52PE	+	-	+	LB	[8]
SUM52PE	+	-	-	LA	[1, 5, 6]
SUM44PE	+	+	+	LB	[8]
SUM44PE	+	+/-	-	LA	[1, 5, 6]
EVSA-T	-	-	+	H	[6]
EVSA-T	-	+	-	LA	[1]
MPE600	+	-	-	LA	[5]
MPE600	+	-	+	LB	[1, 6]

We categorise these cell lines into luminal A (LA), luminal B (LB), HER2 positive (H), Triple negative A (TNA) and Triple negative B (TNB) according to the status of ER, PR and HER2. References where such information is derived are provided in the 'Reference' column.

Discussion

By summarizing and comparing the molecular features of breast cancer cell lines currently available to our knowledge with tumors, genetically and epigenetically, we categorize them into five subtypes, i. e. , luminal A, luminal B, HER2 positive, triple negative A and triple negative B. Unlike at the tissue level where luminal tumors dominate (63% [50]), cell lines of this subtype take the least percentage (31%) among all, i. e. , 17 are luminal A, 9 are luminal B, 20 HER2 positive, 21 triple negative A, and 17 triple negative B among the lines under study. This is due to the fact that most lines are derived from invasive carcinoma for the sake of easy cultivation, toppling the uneven distribution encountered at the tumor

level.

In our nomenclature, luminal A and B cell lines are differentiated from each other and HER2 positive lines are identified as a single subtype to facilitate studies on tissue subtyping and drug response experiments targeting ER and/or HER2 using cell lines. Triple negative lines are identified as two separate groups, which correspond to basal A and B as referred to in most literatures. While triple negative lines share a panel of feature markers such as EGFR which has been used as a conventional marker characterizing triple negative tumors in many studies [51-56], each has its own distinct molecular and phenotypical properties. Triple negative A is characterized by the expression of basal keratins (KRT4/5/6/13/14/15/16/17) which resembles the core basal tumors, and triple negative B is featured by the cancer stem cell pattern such as CD44+CD24- and migration markers such as VIM which could be used for modeling of claudin-low and/or metaplastic breast cancers. Among the collected 92 cell lines here, 74 are unanimously labeled with one subtype as suggested by existing literatures, 10 (DU4475, KPL-3C, MA11, HCC1739, MFM223, CAL148, CAL120, CAL51, CAL851, HDQ-P1) are categorized by their available information based on our subtyping criteria (Table 1), 8 (Table 3) have ambiguous and inconsistent categorization among different studies which may have evolved into different cell lines in different labs after serial cultivation, and one (MDAMB435) is found not of breast cancer origin.

Breast cancer cell lines, though having a higher frequency of DNA and/or copy number alterations, share a similar major genetic mutational spectrum and chromosome aberration profiles with tumors, allowing them still feasible for tumor modelling. However, cell lines are crude models for cancer research as they could not well capture the intra- and inter- tumor heterogeneities, and how well they could represent primary tumors and how much they evolve from their initial establishment are still questionable especially for a few cell lines whose characterization on the primary markers are inconsistent across studies. Thereby, it is encouraged to use breast cancer cell lines for initial tumor modelling but the results need to be further validated using tumor samples or mouse models. Also, cell line selection becomes critical, which needs to be representative and the relevant and/or determinant markers need to be pre-tested.

Cultivation medium is crucial for the success of cell line establishment, regarding how well cells maintain their molecular and physiological features in tumor tissue. Currently widely applied medium include RPMI, DMEM, /αMEM , Ham's F12, L15, McCoys etc. , with RPMI, DMEM, αMEM, Ham's F12 providing a full coverage of all the cell lines in Table 1. The amino acid concentration of these medium, ranging from the most enriched to the least are αMEM, DMEM, RPMI and Ham's F12. The observation that DMEM is rather frequently used for the cultivation of luminal cells, Ham's F12 is many times used in triple negative cells and occasionally in HER2 positive cells suggest that the more aggressive and invasive tumor cell lines are the fewer nutrition is needed in the medium.

Concluding Remarks

By summarizing and going through all the molecular features of currently available cell lines, we grouped them using a subtyping system compatible with that in tumors, and clarified some ambiguous information regarding cell molecular and morphological features. This helps us in the proper choice of and widens the selection spectrum of cell lines when conducting relevant studies, which is critical in the guarantee of final successof translating cell line based results to clinics.

Systematic analysis on the molecular differences between cell lines and tumors reveal that breast cell lines, on average, harbour more genetic mutations and approximately the same amount of epigenetic aberrations than the tumors they derive. The key molecular features of cell lines, however, remain the same with tissue tumors, except for a few exceptional cases whose primary markers have evolved into multiple versions. Also, despite the more frequent chromosomal alterations occurred in cell lines, their genomic profiles stay invariant with the tumors. These empower breast cancer cell lines feasible models for tumors of the same subtype. However, they are not accurate models to capture the full heterogeneity of tumors, both at a single line level or the whole cell population. Also, how much they are evolved from the corresponding primary tumors, either due to technical ease at the establishment or during serial cultivation, remain to be elucidated. In addition, without interactions with the microenvironment, whether or not cell lines could reproduce the signaling *in situ* worth serious considerations before cell lines are chosen for the experiments.

It is worth noting that, our claim that 'cell lines are feasible models for tumors' refers to tumors of the same subtype but not the tumor deriving the cell line. Some cell lines, after passage selection under, e. g. environmental stress, may alter their transcriptional profiles and exhibit the features of another subtype. For instance, EMG3, a triple negative A cell line showing KRT5/14 positivity in the

immunocytochemical analysis, is originated from tumors characteristic of luminal markers such as KRT18/19 [57].

Triple negative breast cancer, though, do not dominate the patient cases, are very heterogeneous, encompassing at least four subcategories. However, we only have two subgroups within cell lines of this category, leaving no cell line being appropriate for interferon-rich tumor modelling. As a subtype with approximately 10% prevalence among tumors [31], it is suggested to establish triple negative cell lines modeling such tumors which represent a 3rd triple negative subtype and enrich the heterogeneity of current cell line modality.

Acknowledgements

This study is supported by the National Natural Science Foundation of China (Grant No. 31471251), Natural Science Foundation of Jiangsu Province (Grant No. BK20161130), the Six Talent Peaks Project in Jiangsu Province (Grant No. SWYY-128), and the Fundamental Research Funds for the Central Universities (JUSRP11507).

Competing Interests

The authors have declared that no competing interest exists.

References

1. Riaz M, van Jaarsveld MT, Hollestelle A, Prager-van der Smissen WJ, Heine AA, Boersma AW, et al. miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. Breast cancer research : BCR. 2013; 15: R33.
2. Lacroix M, Haibe-Kains B, Hennuy B, Laes JF, Lallemand F, Gonze I, et al. Gene regulation by phorbol 12-myristate 13-acetate in MCF-7 and MDA-MB-231, two breast cancer cell lines exhibiting highly different phenotypes. Oncology reports. 2004; 12: 701-7.
3. Cope LM, Fackler MJ, Lopez-Bujanda Z, Wolff AC, Visvanathan K, Gray JW, et al. Do breast cancer cell lines provide a relevant model of the patient tumor methylome? PloS one. 2014; 9: e105545.
4. Elstrodt F, Hollestelle A, Nagel JH, Gorin M, Wasielewski M, van den Ouweland A, et al. BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. Cancer research. 2006; 66: 41-5.
5. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer cell. 2006; 10: 515-27.
6. Hollestelle A, Nagel JH, Smid M, Lam S, Elstrodt F, Wasielewski M, et al. Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines. Breast cancer research and treatment. 2010; 121: 53-64.
7. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adelaide J, Cervera N, et al. Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene. 2006; 25: 2273-84.
8. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PloS one. 2009; 4: e6146.
9. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of clinical investigation. 2011; 121: 2750-67.
10. Lacroix M, Leclercq G. Relevance of breast cancer cell lines as models for breast tumours: an update. Breast cancer research and treatment. 2004; 83: 249-89.
11. Lasfargues EY, Ozzello L. Cultivation of human breast carcinomas. Journal of the National Cancer Institute. 1958; 21: 1131-47.
12. Gazdar AF, Kurvari V, Virmani A, Gollahon L, Sakaguchi M, Westerfield M, et al. Characterization of paired tumor and non-tumor cell lines established from patients with breast cancer. International journal of cancer. 1998; 78: 766-74.
13. Amadori D, Bertoni L, Flamigni A, Savini S, De Giovanni C, Casanova S, et al. Establishment and characterization of a new cell line from primary human breast carcinoma. Breast cancer research and treatment. 1993; 28: 251-60.
14. Cailleau R, Olive M, Cruciger QV. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In vitro*. 1978; 14: 911-5.
15. Band V, Zajchowski D, Swisshelm K, Trask D, Kulesa V, Cohen C, et al. Tumor progression in four mammary epithelial cell lines derived from the same patient. Cancer research. 1990; 50: 7351-7.
16. Ethier SP, Mahacek ML, Gullick WJ, Frank TS, Weber BL. Differential isolation of normal luminal mammary epithelial cells and breast cancer cells from primary and metastatic sites using selective media. Cancer research. 1993; 53: 627-35.
17. [Internet] Committee ICLA. Naming a Cell Line http://iclac.org/resources/cell-line-names/ 2014.
18. Yu M, Selvaraj SK, Liang-Chu MM, Aghajani S, Busse M, Yuan J, et al. A resource for cell line authentication, annotation and quality control. Nature. 2015; 520: 307-11.
19. Watanabe M, Tanaka H, Kamada M, Okano JH, Takahashi H, Uchida K, et al. Establishment of the Human BSMZ Breast Cancer Cell Line, Which Overexpresses the erbB-2 and c-myc Genes. Cancer research. 1992; 52: 5178-82.
20. Alan S, Fadi W A, John M, Nancy F, Pingfu F, Hannah G, et al. MUC1 is expressed at high frequency in early-stage basal-like triple-negative breast cancer. Human pathology. 2013; 44: 2159-66.
21. Forozan F, Mahlamaki EH, Monni O, Chen YD, Veldman R, Jiang Y, et al. Comparative genomic hybridization analysis of 38 breast cancer cell lines: A basis for interpreting complementary DNA microarray data. Cancer research. 2000; 60: 4519-25.
22. Siwek B, Larsimont D, Lacroix M, Body JJ. Establishment and Characterization of Three New Breast-Cancer Cell Lines. International journal of cancer. 1998; 76:677-83.
23. Koay DC, Digiovanna MP. Synergistic interaction between all-trans retinoic acid, Herceptin, and tamoxifen in BT-474 human breast cancer cells. Cancer research. 2004; 64: 508.
24. Zheng AP, Kallio A, Harkonen P. Tamoxifen-induced rapid death of MCF-7 breast cancer cells is mediated via extracellulary signal-regulated kinase signaling and can be abrogated by estrogen. Endocrinology. 2007; 148: 2764-77.
25. Katchman BA, Ocal IT, Cunliffe HE, Chang YH, Hostetter G, Watanabe A, et al. Expression of quiescin sulfhydryl oxidase 1 is associated with a highly invasive phenotype and correlates with a poor prognosis in Luminal B breast cancer. Breast Cancer Research. 2013; 15: 7039-43.
26. Fang WB, Yao M, Jokar I, Alhakamy N, Berkland C, Chen J, et al. The CCL2 chemokine is a negative regulator of autophagy and necrosis in luminal B breast cancer cells. Breast cancer research and treatment. 2015; 150: 309-20.
27. De Iuliis F, Salerno G, Giuffrida A, Milana B, Taglieri L, Rubinacci G, et al. Breast cancer cells respond differently to docetaxel depending on their phenotype and on survivin upregulation. Tumour Biol. 2016; 37: 2603-11.
28. Das P, Siegers GM, Postovit LM. Illuminating luminal B: QSOX1 as a subtype-specific biomarker. Breast Cancer Research. 2013; 15: 104.
29. Katchman BA, Ocal IT, Cunliffe HE, Chang YH, Hostetter G, Watanabe A, et al. Expression of quiescin sulfhydryl oxidase 1 is associated with a highly invasive phenotype and correlates with a poor prognosis in Luminal B breast cancer. Breast cancer research : BCR. 2013; 15: R28.
30. Bertucci F, Borie N, Ginestier C, Groulet A, Charafe-Jauffret E, Adélaïde J, et al. Identification and validation of an ERBB2 gene expression signature in breast cancers. Oncogene. 2004; 23: 2564-75.
31. Dai XF, Xiang L, et al. Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes. Journal of Cancer. 2016; 7: 1281-94.
32. Wang B, Elledge SJ. Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. PNAS. 2007; 104: 20759-63.
33. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. P Natl Acad Sci USA. 2003; 100: 8418-23.
34. Cheang MCU, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, et al. Basal-Like Breast Cancer Defined by Five Biomarkers Has Superior Prognostic Value than Triple-Negative Phenotype. Clin Cancer Res. 2008; 14: 1368-76.
35. Turner N, Tutt A, Ashworth A. Hallmarks of 'BRCAness' in sporadic cancers. Nat Rev Cancer. 2004; 4: 814-9.
36. Kurebayashil J, Kurosumi M, Sonoo H. A new human breast cancer cell line, KPL-3C, secretes parathyroid hormone-related protein and produces tumours associated with microcalcifications in nude mice. British Journal of Cancer. 1996; 74: 200-7.
37. Rye PD, Norum L, Olsens D, Gakman-Vik S, Kaul S, Fodstad O. Brain metastasis model in athymic nude mice using a novel MUC1-secreting human breast-cancer cell line, MA11. International journal of cancer. 1996; 68: 682-7.
38. Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures.
39. Hackenberg R, Luttchens S, Hofmann J, Kunzmann R, Holzel F, Schulz KD. Androgen sensitivity of the new human breast cancer cell line MFM-223. Cancer research. 1991; 51: 5722-7.
40. Gioanni J, Le Francois D, Zanghellini E, Mazeau C, Ettore F, Lambert JC, et al. Establishment and characterisation of a new tumorigenic cell line with a normal karyotype derived from a human breast adenocarcinoma. Br J Cancer. 1990; 62: 8-13.

41. Wang CS, Goulet F, Lavoie J, Drouin R, Auger F, Champetier S, et al. Establishment and characterization of a new cell line derived from a human primary breast carcinoma. Cancer genetics and cytogenetics. 2000; 120: 58-72.
42. Hiscox S, Baruha B, Smith C, Bellerby R, Goddard L, Jordan N, et al. Overexpression of CD44 accompanies acquired tamoxifen resistance in MCF7 cells and augments their sensitivity to the stromal factors, heregulin and hyaluronan. BMC Cancer. 2012; 12.
43. Larramendy ML, Lushnikova T, Bjorkqvist AM, Wistuba II, Virmani AK, Shivapurkar N, et al. Comparative genomic hybridization reveals complex genetic changes in primary breast cancer tumors and their cell lines. Cancer Genet Cytogenet. 2000; 119: 132-8.
44. Wistuba II, Behrens C, Milchgrub S, Syed S, Ahmadian M, Virmani AK, et al. Comparison of features of human breast cancer cell lines and their corresponding tumors. Clin Cancer Res. 1998; 4: 2931–8.
45. Widschwendter M, Jones PA. DNA methylation and breast carcinogenesis. Oncogene. 2002; 21: 5462–82.
46. Castles CG, Klotz DM, Fuqua SA, Hill SM. Coexpression of wild-type and variant oestrogen receptor mRNAs in a panel of human breast cancer cell lines. Br J Cancer. 1995; 71: 974–80.
47. Ellison G, Klinowska T, Westwood RF, Docter E, French T, Fox JC. Further evidence to support the melanocytic origin of MDA-MB-435. Mol Pathol. 2002; 55: 294-9.
48. Dvorankova B, Szabo P, Lacina L, Kodet O, Matouskova E, Smetana K, Jr. Fibroblasts prepared from different types of malignant tumors stimulate expression of luminal marker keratin 8 in the EM-G3 breast cancer cell line. Histochem Cell Biol. 2012; 137: 679-85.
49. Briand P, Lykkesfeldt AE. An *in vitro* model of human breast carcinogenesis: epigenetic aspects. Breast cancer research and treatment. 2001; 65: 179–87.
50. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015; 5: 2929-43.
51. Kanapathy Pillai SK, Tay A, Nair S, Leong CO. Triple-negative breast cancer is associated with EGFR, CK5/6 and c-KIT expression in Malaysian women. BMC Clin Pathol. 2012; 12: 18.
52. Kim Y, Kim J, Lee HD, Jeong J, Lee W, Lee KA. Spectrum of EGFR gene copy number changes and KRAS gene mutation status in Korean triple negative breast cancer patients. PloS one. 2013; 8: e79014.
53. Li XR, Liu M, Zhang YJ, Wang JD, Zheng YQ, Li J, et al. CK5/6, EGFR, Ki-67, cyclin D1, and nm23-H1 protein expressions as predictors of pathological complete response to neoadjuvant chemotherapy in triple-negative breast cancer patients. Med Oncol. 2011; 28 Suppl 1: S129-34.
54. Park HS, Jang MH, Kim EJ, Kim HJ, Lee HJ, Kim YJ, et al. High EGFR gene copy number predicts poor outcome in triple-negative breast cancer. Mod Pathol. 2014; 27: 1212-22.
55. Ueno NT, Zhang D. Targeting EGFR in Triple Negative Breast Cancer. J Cancer. 2011; 2: 324-8.
56. Zhang M, Zhang X, Zhao S, Wang Y, Di W, Zhao G, et al. Prognostic value of survivin and EGFR protein expression in triple-negative breast cancer (TNBC) patients. Target Oncol. 2014; 9: 349-57.
57. Brozova M, Kleibl Z, Netikova I, Sevcik J, Scholzova E, Brezinova J, et al. Establishment, growth and *in vivo* differentiation of a new clonal human cell line, EM-G3, derived from breast cancer progenitors. Breast cancer research and treatment. 2006; 103: 247-57.
58. Bacus SS, Kiguchi K, Chin D, King CR, Huberman E. Differentiation of cultured human breast cancer cells (AU-565 and MCF-7) associated with loss of cell surface HER-2/neu antigen. Molecular Carcinogenesis. 1990; 3: 350-62.
59. Nayak SK, Kakati S, Harvey SR, Malone CC, Cornforth AN, Dillman RO. Characterization of cancer cell lines established from two human metastatic breast cancers. *In Vitro* Cellular & Developmental Biology - Animal. 2000; 36: 188-93.
60. Engel LW, Young NA. Human breast carcinoma cells in continuous culture: a review. Cancer research. 1978; 38: 4327-39.
61. Kurebayashi J, Kurosumi M, Sonoo H. A new human breast cancer cell line, KPL-1 secretes tumour-associated antigens and grows rapidly in female athymic nude mice. British Journal of Cancer. 1995; 71: 845-53.
62. Kurebayashi J, Otsuki T, Kurosumi M, Yamamoto S, Tanaka K, Mochizuki M, et al. Isolation and characterization of a new human breast cancer cell line, KPL-4, expressing the Erb B family receptors and interleukin-6. British Journal of Cancer. 1999; 79: 707-17.
63. Micci F, Teixeira MR, Heim S. Complete cytogenetic characterization of the human breast cancer cell line MA11 combining G-banding, comparative genomic hybridization, multicolor fluorescence *in situ* hybridization, RxFISH, and chromosome-specific painting. Cancer Genetics & Cytogenetics. 2001; 131: 25-30.
64. Arun B, Akar U, Gutierrez-Barrera AM, Hortobagyi GN, Ozpolat B. The PARP inhibitor AZD2281 (Olaparib) induces autophagy/mitophagy in BRCA1 and BRCA2 mutant breast cancer cells. Int J Oncol. 2015; 47: 262-8.
65. Gong C, Fujino K, Monteiro LJ, Gomes AR, Drost R, Davidson-Smith H, et al. FOXA1 repression is associated with loss of BRCA1 and increased promoter methylation and chromatin silencing in breast cancer. Oncogene. 2015; 34: 5012-24.
66. Rizki A, Weaver VM, Lee SY, Rozenberg GI, Chin K, Myers CA, et al. A Human Breast Cell Model of Preinvasive to Invasive Transition. Cancer research. 2008; 68: 1378-87.
67. Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT, et al. The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. Molecular Oncology. 2007; 1: 84-96.
68. Meltzer P, Leibovitz A, Dalton W, Villar H, Kute T, Davis J, et al. Establishment of two new cell lines derived from human breast carcinomas with HER-2/neu amplification. Br. J. Cancer 63: 727-735. British Journal of Cancer. 1991; 63: 727-35.
69. Forozan F, Veldman R, Ammerman CA, Parsa NZ, Kallioniemi A, Kallioniemi OP, et al. Molecular cytogenetic analysis of 11 new breast cancer cell lines. British Journal of Cancer. 1999; 81: 1328-34.
70. Ignatoski KMW, Ethier SP. Constitutive activation of pp125fak in newly isolated human breast cancer cell lines. Breast cancer research and treatment. 1999; 54: 173-82.
71. [Internet] GmbH DSvMuZ. Ddetails of Cell Lines http://www. dsmz. de. 2017.