OPEN ACCESS Freely available online

PLOS ONE

Increased macroH2A1.1 Expression Correlates with Poor Survival of Triple-Negative Breast Cancer Patients

Anne-Claire Lavigne[1,2]*, Magali Castells[1,2], Jérôme Mermet[1,2¤a], Silvia Kocanova[1,2], Mathieu Dalvai[1,2¤b], Kerstin Bystricky[1,2]

1 Université de Toulouse; Laboratoire de Biologie Moléculaire Eucaryote (LBME); F-31062 Toulouse, France, **2** CNRS; LBME; F-31062 Toulouse, France

Abstract

Purpose: Epithelial-Mesenchymal Transition (EMT) features appear to be key events in development and progression of breast cancer. Epigenetic modifications contribute to the establishment and maintenance of cancer subclasses, as well as to the EMT process. Whether histone variants contribute to these transformations is not known. We investigated the relative expression levels of histone macroH2A1 splice variants and correlated it with breast cancer status/prognosis/types.

Methods: To detect differential expression of macroH2A1 variant mRNAs in breast cancer cells and tumor samples, we used the following databases: GEO, EMBL-EBI and publisher databases (may-august 2012). We extracted macroH2A1.1/macroH2A1 mRNA ratios and performed correlation studies on intrinsic molecular subclasses of breast cancer and on molecular characteristics of EMT. Associations between molecular and survival data were determined.

Results: We found increased macroH2A1.1/macroH2A1 mRNA ratios to be associated with the claudin-low intrinsic subtype in breast cancer cell lines. At the molecular level this association translates into a positive correlation between macroH2A1 ratios and molecular characteristics of the EMT process. Moreover, untreated Triple Negative Breast Cancers presenting a high macroH2A1.1 mRNA ratio exhibit a poor outcome.

Conclusion: These results provide first evidence that macroH2A1.1 could be exploited as an actor in the maintenance of a transient cellular state in EMT progress towards metastatic development of breast tumors.

Citation: Lavigne A-C, Castells M, Mermet J, Kocanova S, Dalvai M, et al. (2014) Increased macroH2A1.1 Expression Correlates with Poor Survival of Triple-Negative Breast Cancer Patients. PLoS ONE 9(6): e98930. doi:10.1371/journal.pone.0098930

Editor: Benjamin Haibe-Kains, Princess Margaret Cancer Centre, Canada

Received July 1, 2013; **Accepted** May 8, 2014; **Published** June 9, 2014

Funding: Funding provided by Institut National du Cancer - n°TRANSLA11-73 - URL: http://www.e-cancer.fr/and Fondation pour la Recherche Médicale - URL: http://www.frm.org/. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: lavigne@ibcg.biotoul.fr

¤a Current address: Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
¤b Current address: Laval University Cancer Research Center, Oncology Axis-CHU de Québec Research Center, Hotel-Dieu de Quebec, Quebec City, Quebec, Canada

Introduction

Triple-Negative Breast Cancer (TNBC) is clinically defined by the lack of expression of the estrogen (ER) and progesterone (PgR) receptor genes, and by the absence of amplification of human epidermal growth factor receptor-2 (HER2). Treatment of TNBC has been challenging due to its heterogeneity at the molecular level and the absence of well-defined molecular targets [1,2]. Despite a frequent complete response to neoadjuvant chemotherapy, TNBC patients also have a higher rate of long term recurrence and worse prognosis than ER-positive BC patients. Distinguishing chemoresistant TNBC patients at risk to relapse from those with a relatively favorable prognosis, would help to identify clinically relevant subgroups that could benefit from alternative treatments.

Advances in gene expression profiling have permitted characterization of different intrinsic molecular subtypes present in TNBC [3]. One of these, the claudin-low breast cancer subtype [4], is characterized by mesenchymal features, low expression of cell-cell junction proteins (i.e., E-cadherin), and intense immune infiltrates. Furthermore, claudin-low tumors have unique biological properties linked to mammary stem cells [5] and Epithelial-Mesenchymal Transition (EMT) features [6].

Gene expression during EMT is dependent on specific transcription factors that interact with enhancer or promoter elements, the accessibility of their binding sites which is regulated by epigenetic reprogramming [7,8]. Hence, chromatin reorganization could contribute to the regulation of epithelial plasticity [9–12]. To date however, the presence of histone variants has not been investigated with respect to the phenomenon of EMT. Gene expression accompanying EMT is also regulated at the post-transcriptional level via alternative splicing of RNA [13–15].

The histone variant macroH2A1 is a vertebrate-specific member of the H2A family and is unusual due to the presence of a C-terminal macro domain [16]. Two isoforms, macroH2A1.1 and macroH2A1.2 are produced by alternative splicing of the *H2AFY* gene. Both isoforms have been associated with silencing and transcriptional repression [17–19]. Regulation of macroH2A1 expression seems to be linked to self-renewal and commitment of ES cells, representing a barrier to reprogramming pluripotency

[20–22]. In melanoma, loss of macroH2A1 promoted progression of metastasis [23]. Moreover, high levels of macroH2A1.1 are associated with slowly proliferating cancers, whereas highly proliferating tumors have markedly decreased macroH2A1.1 levels. Conversely, macroH2A1.2 expression is independent of proliferation in all tumours [24–26]. Notably, expression of macroH2A1.1 has been identified as a novel biomarker in lung and colon cancer models [25,26].

In this study, we demonstrate that selective splicing of the *H2AFY* gene is correlated with EMT features linked to Claudin-low breast cancers. We propose that macroH2A1.1 expression levels could participate in the epigenetic program linked to poor clinical outcome of this molecular breast cancer subtype, and more generally in the EMT process.

Materials and Methods

Cell culture

MCF-7 and MDA-MB231 were obtained from ATCC. ZR-75, MDA-MB436 and Hs578T, were a gift from G. Freiss (Montpellier, France), originally purchased from ATCC [27]. MDA-MB231, MDA-MB436 and Hs578T cells were maintained in DMEM high glucose with glutamax. MCF-7 cells were maintained in DMEM/F12 with Glutamax. ZR-75 cells were maintained in RPMI-1640 supplemented with 10 mM Hepes. All these media were supplemented with 10% heat-inactivated fetal bovine serum and 1 mM sodium pyruvate.

Protein quantification

Antibodies against macroH2A1 (07-219; Upstate), macroH2A1.1 and macroH2A1.2 (gift by A. Ladurner), ERα (sc-543; Santa Cruz), GAPDH (MAB374; Millipore), H3 (ab1791; Abcam) were used for immunoblotting. To discriminate between the two splicing isoforms of macroH2A1, macroH2A1.1 and macroH2A1.2, total cell extracts were separated on low cross-linking (12% acrylamide,1:125 bisacrylamide) SDS-polyacrylamide gels and blotted with antibodies specific to one of the two isoforms (Fig.S1) specifically. Proteins were quantified using the Image Gauge software. Expression levels of each isoform and total macroH2A1 were normalized to GAPDH. ZR-75 expression was used as sample reference.

RNA extraction, reverse transcription and Quantitative PCR analysis

Total RNA was extracted using an RNeasy mini kit (Qiagen). First strand cDNA was generated using the ThermoScript RT-PCR system (Invitrogen) and used as the template for quantitative PCR (qPCR) using the platinium SYBR Green qPCR SuperMix (Invitrogen) according to the manufacturer's instructions. Gene- or splice variant-specific primers are shown in Fig.S2A. Relative levels of RNA were determined using the threshold cycle (ΔC_T) method [28]. Expression levels were normalized to the ribosomal RPLP0 gene. ZR-75 expression was used as sample calibrator.

Determination of the macroH2A1.1/macroH2A1 mRNA ratio

A summary of probe set IDs used is reported in Fig.S2. Among the probe set ID from HG-U133A, three detect the expression of *H2AFY* gene: 214501_s_at and 207168 _s_at probe set ID which are common to the two isoforms and 214500_at which recognized specifically the sequence of the exon 6a of macroH2A1.1 and 10 nucleotides in exons 5 and 7 common to macroH2A1.1 and macroH2A1.2 isoforms (Fig.S2). We extracted the corresponding log2 RMA values from the different GEO datasets studied and determined relative expression of macroH2A1 (mean value of 214501_s_at and 207168_s_at), macroH2A1.1 (214500_at) and calculated the macroH2A1.1/macroH2A1 mRNA ratio by the following formula:

$$\log 2(214500\_at) - [\log 2(214501\_s\_at) + \log 2(207168\_s\_at)]/2 = \log 2(macroH2A1.1/macroH2A1 ratio)$$

Among the probe set ID from Illumina Human-6 v1 expression bead chip, two of them detect the expression of *H2AFY* gene: 6620403 probe set ID which is common to the two isoforms and 6620403 which recognized specifically the sequence of the exon 6a of macroH2A1.1. Among the probe set ID from Illumina HumanHT-12 v4.0 expression bead chip, three detect the expression of *H2AFY* gene: ILMN_2373495 and ILMN_1746171 probe set ID which are common to the two isoforms and ILMN_1674034 which recognized specifically the sequence of the exon 6a of macroH2A1.1. We applied the same formula as above with corresponding log2 RMA values.

Analysis of correlation of macroH2A1.1/macroH2A1 mRNA ratio and breast cancer cell lines markers

For each GEO dataset analyzed, we conserved the intrinsic molecular subtype of breast cancer cell lines attributed in the original study or in absence of we attributed the molecular intrinsic subtype as defined in Table S1. Then we classified the different breast cancer cell lines into two groups, luminal/basal, or claudin-low/non claudin-low and compared the distribution of macroH2A1, macroH2A1.1 expression levels or macroH2A1.1/macroH2A1 mRNA ratio values (Table S2). The reported *p*-values are the results of a two-tailed Mann-Whitney test.

Survival analyses

A summary of affymetrix microarray datasets used in this study, including the number of patients included in each stage of the analysis, is given in Table 1.

Follow-up data was available for 383 of the 579 TNBC samples from the GSE31519 dataset. All survival intervals were measured from the time of surgery to the distinct survival endpoint used in the individual datasets. In the conduct of the presented analysis event free survival (EFS) was calculated as preferentially corresponding to the RFS endpoint, but measured with respect to the DMFS endpoint if RFS was not available. Rody A. et al., [29] have previously shown that the effect of using these different endpoints was rather small in the overall dataset. Follow up data for those women in whom the envisaged end point was not reached were censored as of the last follow-up date or at 120 months. Subjects with missing values were excluded from the analyses. For the analyses of untreated and adjuvant therapy treatment groups, we applied the Kolmogorov-Smirnov test to compare the cumulative distribution of the two data sets.

To determine the cutoff value of macroH2A1.1/macroH2A1 mRNA ratio, a receiver-operating characteristics (ROC) analysis was performed. We constructed Kaplan-Meier curves and used the log-rank test to determine the univariate significance of the variables. The predictive potential of macroH2A1.1/macroH2A1 mRNA ratio is assessed by its positive and negative predictive values (PPV and NPV) (Table S3). A Cox proportional-hazards model was used to examine the effects of multiple covariates on

Table 1. Summary of affymetrix microarray datasets used in this study.

		% of samples (complete datasets)						No of samples							
Dataset	**Data Source**	**Tumor size ≤ 2cm**	**LNN**	**G3**	**Systemic Treatment**			**Total**	**TNBC**	**DMFS**	**RFS**	**Used in**			***Ref***
					No	**Yes**									
						endocrine treatment	**chemotherapy**					**All samples**	**Untreated**	**Treated**	
San Francisco	E-TABM-158	35	52	45	17	13	70	118	23	x		23	7	16	[46]
Mainz	GSE11121	33	100	71	100	0	0	200	21	x		21	21		[47]
Veridex-Tam	GSE12093		100		0	100	0	136	1	x		1	1		[48]
Rotterdam-EMC204	GSE12276	22		71	63	25	13	204	56		x	56	42	6	[49]
Stockholm	GSE1456			74	45	55	0	159	25		x	25	22		[50]
Rotterdam-EMC344	GSE2034, GSE5327	35	100	72	100	0	0	344	82	x		82	82		[51] [52]
New York	GSE2603	11	49		0	4	96	99	35	x		29	1	26	[53]
Oxford-untreated	GSE2990		100		100	0	0	61	13		x	11	11		[54]
Hamburg-2	GSE31519	0	71	40	0	100	0	77	7	x		5	5		[55]
Hamburg-1	GSE31519	33	80	86	0	0	100	77	15		x	15		15	[56]
Franckfurt-2	GSE31519	0	39	56				67	19		x	18	14		[57]
Franckfurt-3	GSE31519	50	50	50	0	100	0	52	2		x	2	2		[58]
Frankfurt	GSE31519	42	83	83	0	0	100	119	24		x	24		24	[59]
Uppsala	GSE3494	41	68	59	86	14	0	251	27		x	27	21		[60]
London	GSE6532	0	0	100	0	100	0	87	2		x	2	2		[61]
TransBIG	GSE7390	18	100	75	100	0	0	198	40		x	40	40		[62]
London-2	GSE9195	100	100	100	0	100	0	77	2		x	2	2		[63]
TOTAL:												***383***	***273***	***87***	

Adapted from Rody A et al., [29].

doi:10.1371/journal.pone.0098930.t001

survival (Table S3). All *P*-values are two-sided and 0.05 was considered as a significant result.

All statistical analyses were performed using the XLSTAT version 2013.1.

Accession codes

Summary of public databases used: Gene Expression Omnibus (NCBI, Bethesda, MD, USA): GSE16795, GSE9691, GSE24202, GSE31519; ArrayExpress (EBI, Hinxton, UK): E-TABM-157, E-MTAB-183, E-MTAB-827, E-MTAB-884.

Results

Expression of MacroH2A1 splice variants in breast cancer cell lines

We quantified protein expression levels of macroH2A1, macroH2A1.1, macroH2A1.2 *per se* and each of the macroH2A1 isoforms relative to the total macroH2A1 protein pool in BC cell lines. MacroH2A1 isoforms or macroH2A1 expression levels did not differ significantly between cell lines representative of the ER positive luminal subtype, ZR-75 and MCF-7 cells, and cell lines of the ER negative basal subtype, MDA-MB436, Hs578T, and MDA-MB231 (Fig.1A). In contrast, expression levels of macroH2A1.1 protein relative to total macroH2A1 protein pool were greater in cells of the basal subtype compared to the luminal one (Fig.1B). Unlike macroH2A1.2, which showed no significant variation in expression levels, the increase in macroH2A1.1 was greatest in MDA-MB231 cells.

Using splice variant-specific primers (Fig.S2), we also determined that macroH2A1.1 transcription was greater in the basal compared to the luminal subgroup of BC cell lines, while mRNA expression levels of macroH2A1.2 or total macroH2A1 mRNA did not differ significantly (Fig.1C). As at the protein level, this differential expression was consolidated by analysis of the proportional expression of macroH2A1.1 relative to total macroH2A1 (Fig.1D).

To test whether this observation could be extended to a larger panel of BC cell lines, we analyzed macroH2A1 expression in 51 BC cell lines from data published by Neve *et al.*, [30] (Fig.1E). Total macroH2A1 expression levels were reduced in the basal BC cell lines ($p = 0.006$). Therefore, even if macroH2A1.1 expression levels *per se* did not vary ($p = 0.555$), the relative proportion of macroH2A1.1 to global macroH2A1 was significantly higher in the basal subtype ($p = 0.027$).

Because the three basal BC cell lines tested (Fig.1A–D) belonged to the claudin-low subtype, we subdivided the cell lines from Neve et al., study into two groups: claudin-low and non claudin-low BC cell lines. Cell line subtypes were attributed as in Prat et al., [31] (Table S1). Then we compared the relative expression levels of macroH2A1 mRNAs. Increased macroH2A1.1 expression levels appeared typical of claudin-low subtype BC cell lines (Fig.2A). This correlation was significant for the macroH2A1.1/macroH2A1 mRNA ratio (compare *p-values* Fig.2A center and right panels) and was further confirmed by our analysis of several independent studies that differed in the nature of the cell lines (Fig.S3) and the array platform used (Fig.2B).

Finally, the study of Lapuk et al. [32] allowed us to assess alternative splicing in 31 BC cell lines. We first analyzed relative expression levels of each exon of *H2AFY* gene except that of exon 8 (data unavailable). Globally each exon was less expressed in the claudin-low than in the non-claudin low BC cell lines (Fig.3A). One exception appeared to be the exon 6a of macroH2A1.1 which was expressed more strongly in Claudin-low BC cell lines, but the difference was not statistically significant (Fig.3B). We normalized the level of expression of each exon (log2 RMA values) relative to that of exon 9, the most expressed exon of the *H2AFY* gene. As shown in Fig.3C, expression levels of most of the exons of the *H2AFY* gene in claudin-low subtype cells decreased; and this was also true for the macroH2A1.2 specific exon 6b ($p = 0.001$) (Fig.3D). In contrast, expression levels of macroH2A1.1 specific exon 6a increased in the claudin-low compared to the non claudin-low subtype ($p = 0.007$). Moreover, determination of the splicing index showed that only exon 6a varied in all cell lines tested (Fig.3E).

We conclude that the expression of macroH2A1.1 relative to total macroH2A1 expression (macroH2A1.1/macroH2A1 mRNA ratio defined in materials and methods), not macroH2A1.1 expression *per se*, is correlated specifically with the claudin-low molecular subtype.

MacroH2A1.1 variant expression correlates with epithelial-mesenchymal transition

We classified the macroH2A1.1/macroH2A1 mRNA ratios of a set of 38 BC cell lines relative to the E-cadherin expression data determined by Hollestelle *et al.*, [33]. In E-cadherin^{negative} cell lines this ratio was generally greater and more diverse, but the difference between E-cadherin^{positive} and E-cadherin^{negative} cells was not statistically significant (Fig.4A). Different mechanisms for inactivating E-cadherin have been identified in human cancers: inherited and somatic mutations, increased promoter methylation, and induction of transcriptional repressors of the Twist, Snail and Zeb family members [6]. The latter induce EMT in parallel with induction of mesenchymal markers such as N-cadherin and/or vimentin. Interestingly, most of the cell lines exhibiting high macroH2A1.1/macroH2A1 mRNA ratios, expressed N-cadherin and vimentin (Fig.4B).

In order to determine whether enrichment of macroH2A1.1 could be related to the EMT process, we analyzed expression levels of this variant in different cellular models of EMT. Comparison of macroH2A1.1/macroH2A1 mRNA ratios in HMLE_shGFP and HMLE_shEcad, revealed that reduction of E-cadherin expression levels was accompanied by an increase in the macroH2A1.1 mRNA ratio in two independent data sets (Fig.4C) [34,35]. This increase was clearly associated with induction of EMT due to dysfunction in intracellular signaling caused by reduced E-cadherin levels. Indeed, expression of a truncated form of E-cadherin (ΔN-Ecad) lacking the extracellular domain of the wild-type protein normally responsible for E-cadherin cell-cell adhesion was not correlated with an increase in macroH2A1.1mRNA ratios (Fig.4C). Accordingly, overexpression of inducers of EMT, Twist1, Goosecoid or Snail in the HMLE cell line was accompanied by an increase in macroH2A1.1 mRNA (Fig.4D). In contrast, macroH2A1.1 mRNA ratios were not up-regulated by overexpression of TGF-β. This is in agreement with previous observations showing that induction of an EMT by Snail or Twist does not depend on TGF-β autocrine signaling [6]. Moreover, TGF-β signaling is not sufficient for an EMT conversion in primary normal, immortalized, and neoplastic HMECs [36], and is thus insufficient to induce an increase in macroH2A1.1/macroH2A1 mRNA ratio.

Extensive changes in alternative splicing play a role in shaping cellular behavior patterns that characterize EMT. Interestingly, the macroH2A1.2 specific exon was shown to be an Epithelial Splicing Regulatory Protein (ESRP)-regulated cassette [4]. Analysis of the genomic context of the exon specifically included in macroH2A1.1 identified potential binding sites for the EMT-associated splicing factors, ESRP1 and RBFOX2 (Fig.4E). ESRP binding sites located at the 5′ end and within the regulated exon

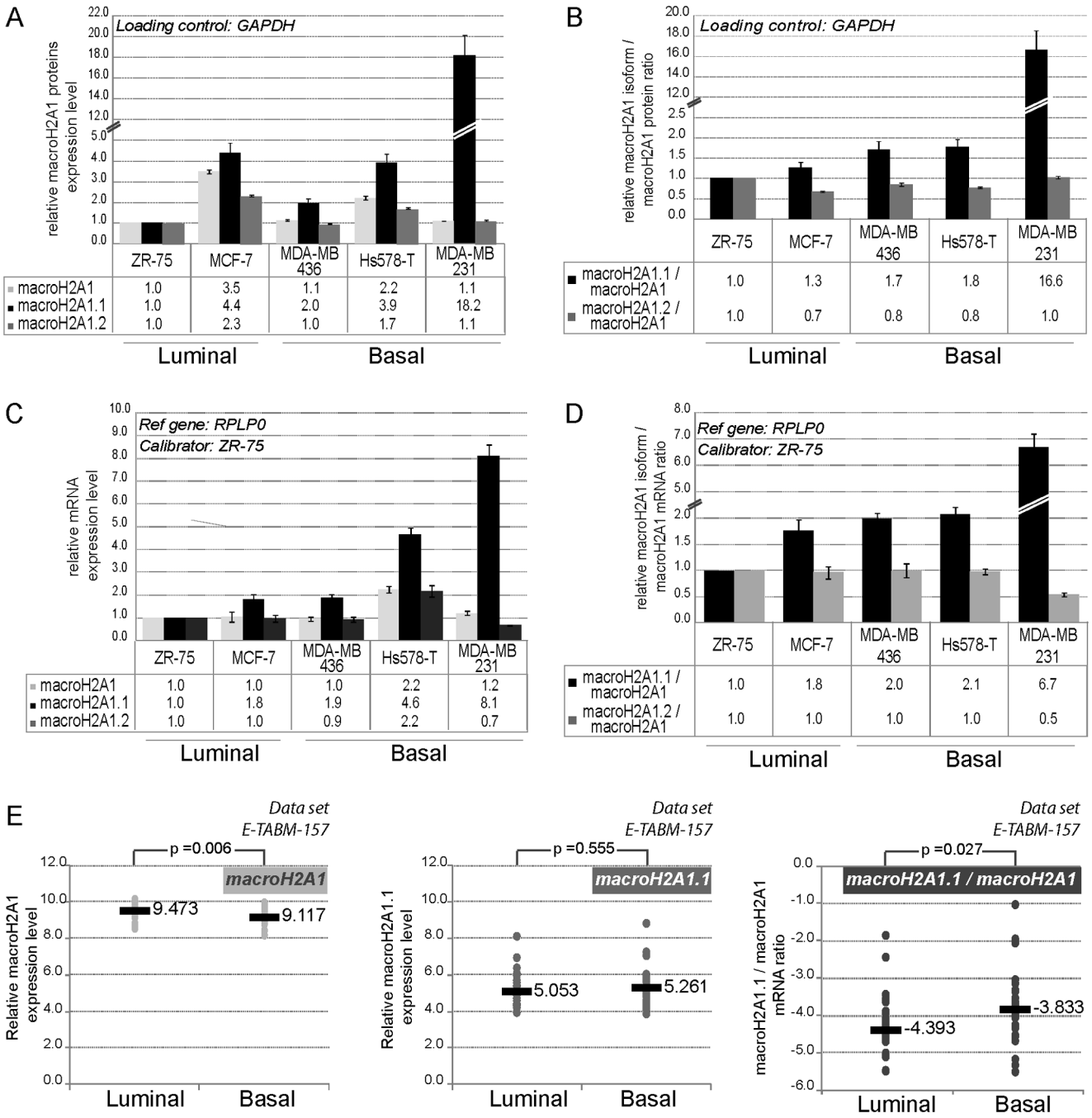


Figure 1. Expression levels of macroH2A1 splice variants in breast cancer cell lines. A-B- MacroH2A1.1, macroH2A1.2, and total macroH2A1 protein expression levels in five breast cancer cell lines. Quantification of each macroH2A1 splice variants, global macroH2A1 (A) and macroH2A1.1/macroH2A1, macroH2A1.2/macroH2A1 ratios (B) for each cell line normalized to GADPH is shown relative to the ZR-75 cell line. **C-D-** qPCR analysis of mRNA expression levels of macroH2A1, macroH2A1.1, macroH2A1.2 and macroH2A1 splice variants/macroH2A1 mRNA ratio. Each quantification was performed at minimum in biological triplicate. Expression levels were normalized to expression of RPLP0, and referred to the cell line ZR-75 as a sample calibrator. **E-** Analysis of expression data of macroH2A1 variants in 51 breast cancer cell lines on the basis of U133A array hybridization [30]. Log2 macroH2A1, MacroH2A1.1 and macroH2A1.1/macroH2A1 values are determined and classified according to luminal or basal molecular subtype as defined in [30]. Data of DU4475, HCC1008 and HCC1599 are included in the analysis with the molecular subtype assigned in the synthesis part of Table S1. The median of each subgroup is shown (grey bar). The reported *p*-values are the results of a two-tailed Mann-Whitney test.
doi:10.1371/journal.pone.0098930.g001

seem to be ESRP Binding Splicing Inhibitors (EBSI), and a RBFOX2 binding site found downstream of the alternatively spliced exon seems to be an RBFOX2 Binding Splicing Enhancer (RBSE). Exon 6a skipping could thus result from the interaction of ESRP1 with EBSI in an epithelial context. In a mesenchymal context, exon 6a would be preferentially included by enhanced binding of RBFOX2 (Fig.4E).

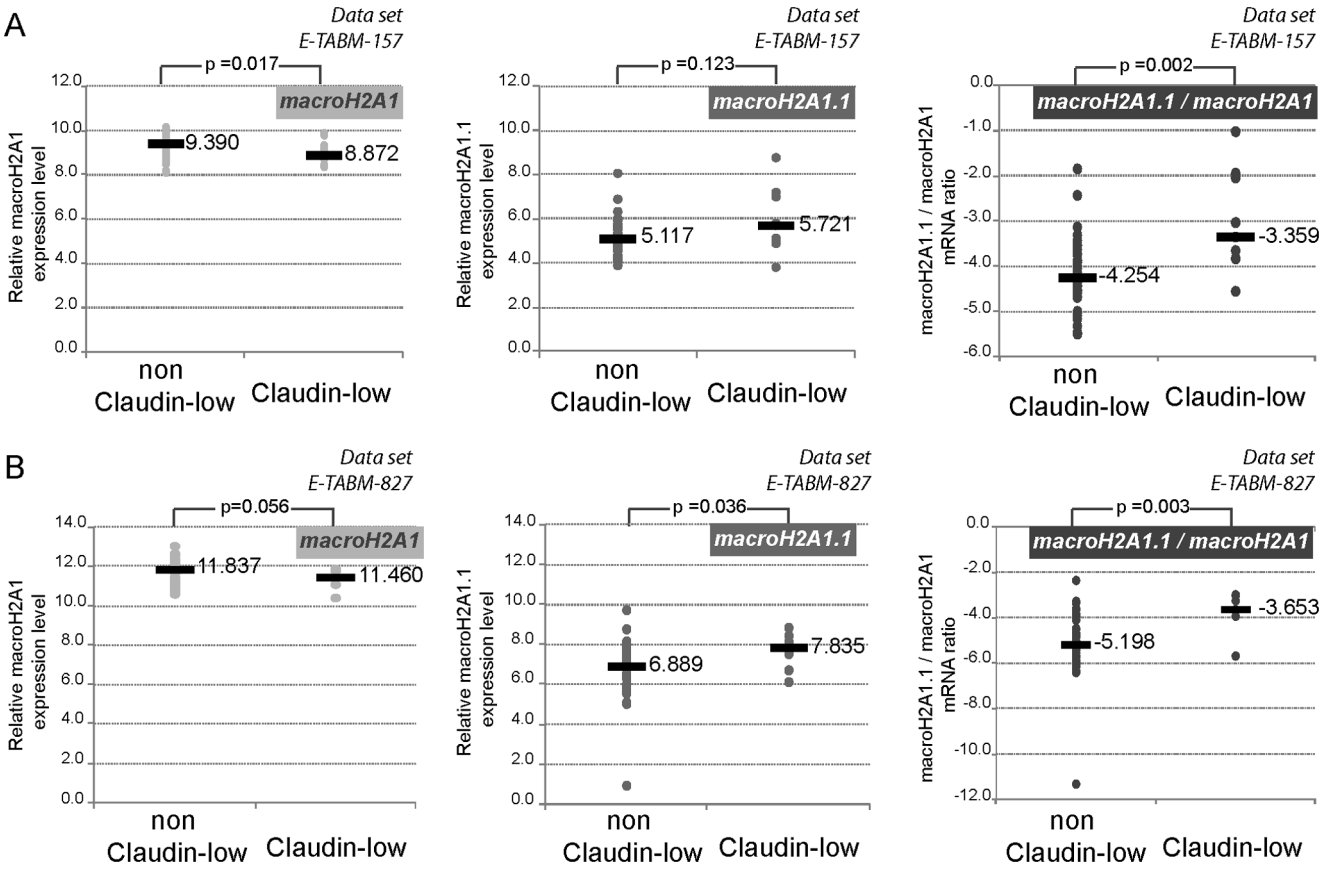


Figure 2. High macroH2A1.1 expression level in breast cancer cell lines characterizes Claudin-low molecular subtype. In two independent analyses, macroH2A1, macroH2A1.1 and macroH2A1.1/macroH2A1 mRNA ratios were determined for each cell line and classified according to claudin-low or non claudin-low molecular subtype assigned in the synthesis part of Table S1. The median of each subgroup are specified. In E-TABM-827 analysis, GI-101, HB4A, PMC42 and VP229 cell lines were omitted as the subgroup Basal A or B was not specified; HCC1509, MT3 and VP267 cell lines are omitted as their subtype were not assigned. The reported *p*-values are the result of a two-tailed Mann-Whitney test.
doi:10.1371/journal.pone.0098930.g002

Prognostic significance of macroH2A1.1 mRNA ratio in TNBC

To assess the potential prognostic value of macroH2A1.1 mRNA ratio in breast cancer, we analyzed the event-free survival of patients as a function of macroH2A1.1 mRNA ratio reported for the GSE31519 dataset, which provides access to a large cohort of TNBCs (Table 1).

We plotted Kaplan-Meier survival curves based on macroH2A1.1 mRNA ratios segmented into two groups, high and low macroH2A1.1 mRNA ratios. The Receiver Operating Characteristic (ROC) analysis was used to find the optimal cut-off level. We used the log-rank test to determine the univariate significance of the variables. Poor prognosis TNBCs had high macroH2A1.1 mRNA ratios ($p = 0.001$) (Fig.5A). The positive predictive value (PPV) and negative predictive value (NPV) were 49% and 69%, respectively. In multivariate Cox regression analysis, including age, histological grade, tumor size and lymph node status, the macroH2A1.1/macroH2A1 mRNA ratio showed a trend as an independent predictor (HR 3.457, 95%CI 1.087 to 10.990, $p = 0.036$) (Fig.5B).

Among the clinico-pathological characteristics of TNBC patients included in the GSE31519 dataset, the use of systemic treatment was specified. Hence, we sub-divided the cohort into one untreated sub-cohort, and a second treated sub-cohort which regroups all adjuvant treated patients. Analysis of the macroH2A1.1/macroH2A1 mRNA values in the untreated sub-cohort revealed a distribution comparable to the values of the total cohort ($p = 0.235$) (Fig.S4). MacroH2A1.1/macroH2A1 mRNA values of the group treated with an adjuvant therapy differed from those of the total cohort, with values shifted to higher values of the intervals ($p = 0.008$) (Fig.S4). Kaplan-Meier survival curves were plotted according to macroH2A1.1 mRNA ratios after segmentation into high and low groups as above for the two sub-cohorts. We observed that in contrast to treated tumors, high macroH2A1.1 mRNA ratios still correlated with reduced survival curves for untreated tumor ($p = 0.152$ vs. $p = 0.001$) (Fig.6).

Discussion

We provide evidence that overexpression of macroH2A1.1 correlates with major mesenchymal markers of the claudin-low breast cancer subtype. Notably, the increase in macroH2A1.1 seems to be a residual track of an EMT process, correlated with poor prognosis in TNBCs.

Claudin-low tumors are typically TNBCs with poor long-term prognosis, despite reduced expression of genes related to cell proliferation. Nevertheless, unlike prognostic signatures that rely heavily on proliferation-related genes, macroH2A1.1 preferentially associated with non-proliferative phenomena. It would belong to a new prognostic marker class independent of proliferative status, similar to factors related to the immune system response [37].

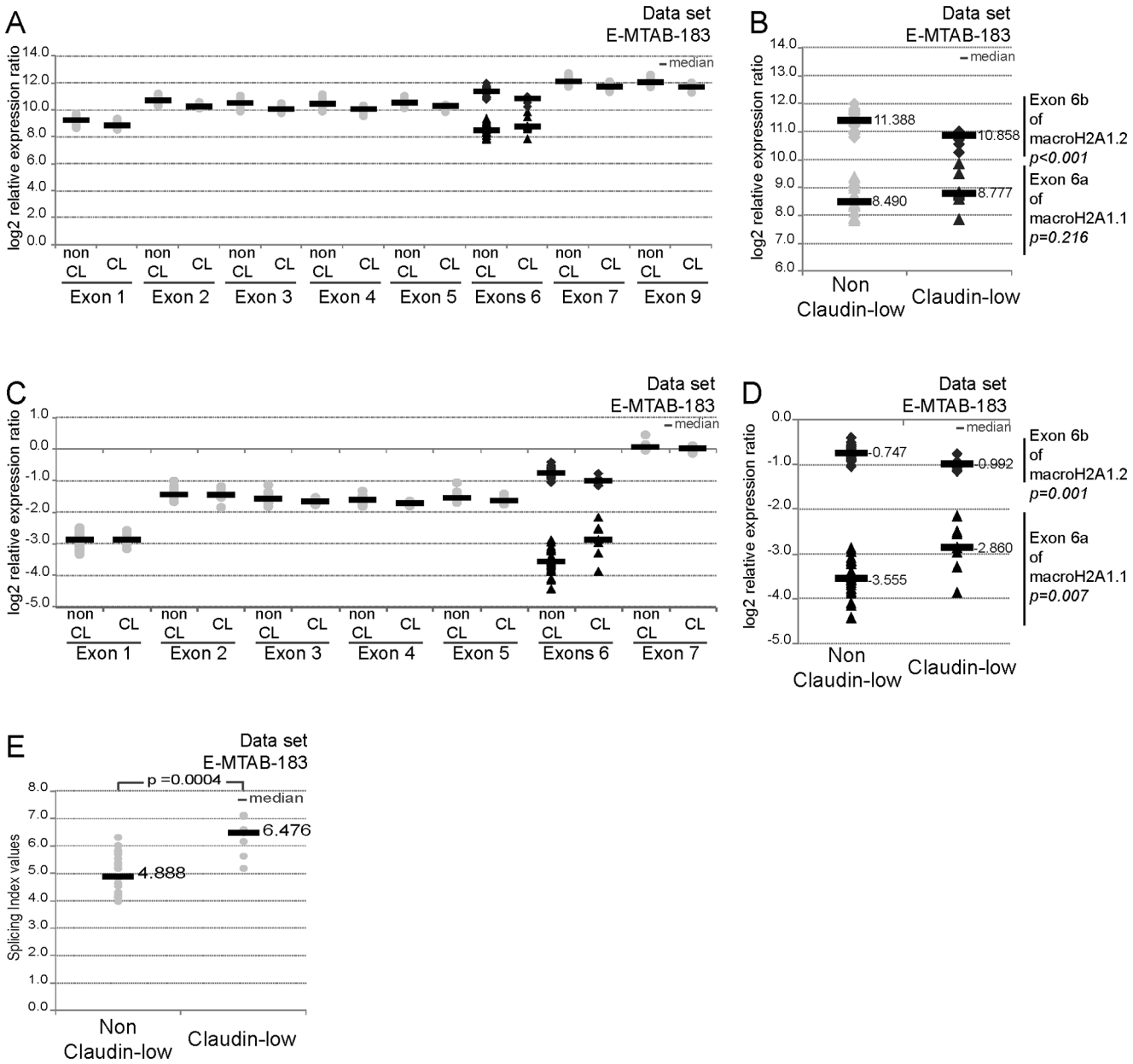
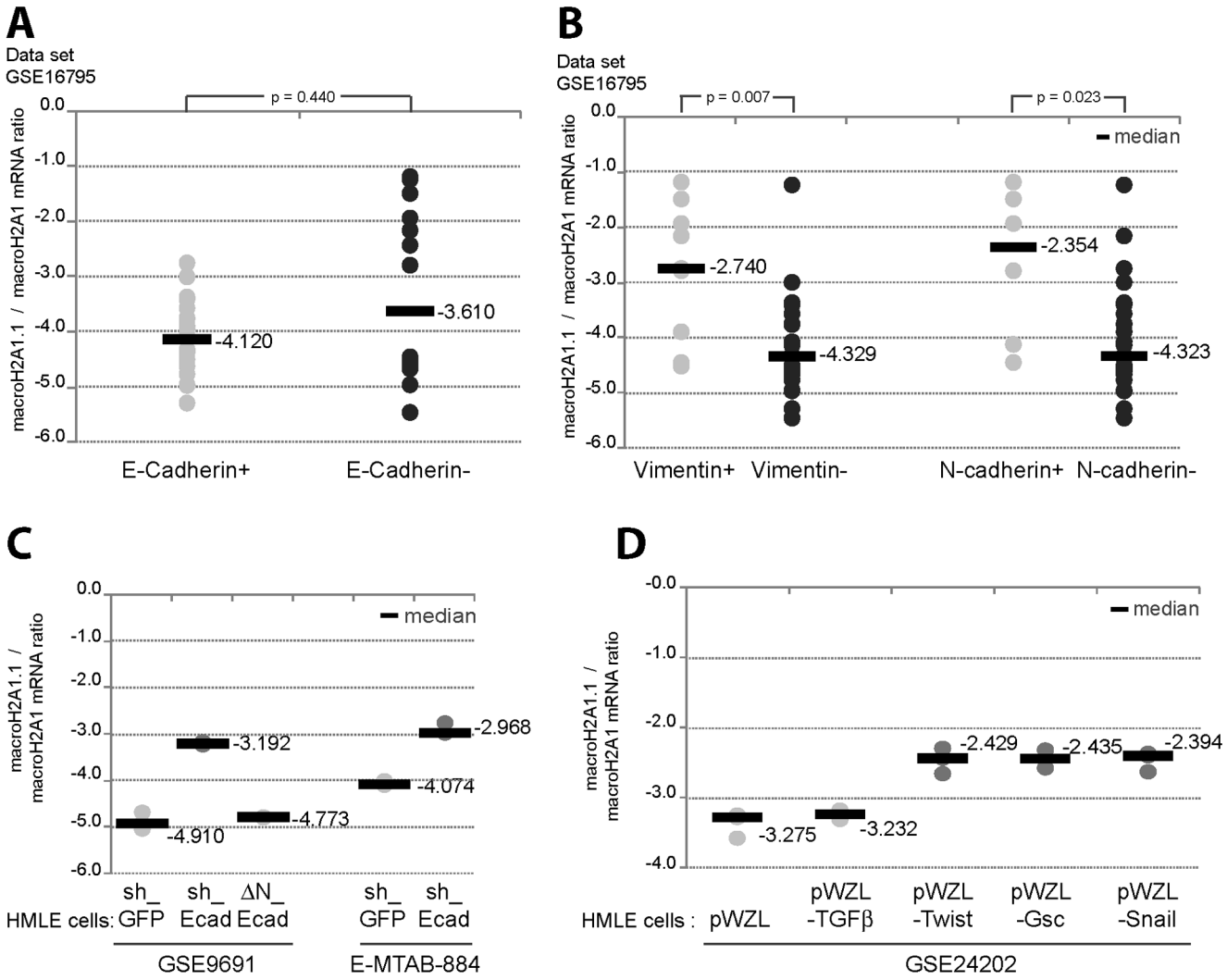


Figure 3. Analysis of expression data of macroH2A1 variants in breast cancer cell lines on the basis of Affymetrix Human Junction technology (E-MTAB-183[11]**).** A- Log2 expression values of exons of H2AFY gene are represented by molecular subtype non claudin-low (non CL) and claudin-low (CL). The analysis for exons 6a/b are highlight in B. C- Log2 expression values of exons of H2AFY gene normalized to the one of exon 9 are represented by molecular subtype non claudin-low (non CL) and claudin-low (CL). The analysis for exons 6a/b are highlight in D. E- Splicing index values for exon 6a included in macroH2A1.1 splice variant are represented by molecular subtype. The median of each subgroup is shown (grey bar). The reported *p*-values are the results of a two-tailed Mann-Whitney test.
doi:10.1371/journal.pone.0098930.g003

Interestingly, it was shown that upon entering EMT HMECs develop a stable, low proliferative mesenchymal phenotype. MacroH2A1 was identified as an epigenetic barrier which participates in the maintainance of cell identity and antagonizes induction of cell reprogramming to naive pluripotency [38,39]. Thus, macroH2A1.1 could be involved in the maintenance of a mesenchymal state, partial or complete, by establishing an epigenetic barrier against further de/differentiation.

The difficulty of identifying EMT-transitioning cells *in vivo* creates skepticism regarding the pathological relevance of EMT. One explanation for this is that cancer cells only undergo a transient EMT, reverting back to the epithelial state by a mesenchymal-epithelial transition (MET), making it difficult to isolate cells with true EMT markers. Studies in experimental mouse models have shown that a complete EMT-MET cascade is important for tumor metastasis [40,41]. If the EMT process is so transient and, in parallel, so important for the development of metastatic tumors, why do only claudin-low and, to a lesser extent, metaplastic intrinsic molecular subtypes of BC present molecular features of EMT? One explanation could be that in claudin-low tumors the EMT-MET turnover is trapped in an intermediate mesenchymal state, in which EMT markers are present. We



E

```
48791 attcctcctc ctcctgtcct cacccccatc cacatgtgcg cacactgcgt ggactgattg ccccttgggc tcatatgttg gaatcgacca ggtaggccag
                                   ESRP1 Binding Site
48891 ccctgccatt ggggcattag taaatgtgcc tgtgcgtggg tctcggtcca acacagttga tatacatttg tttacctgtt atagttgcaa gttgtacagg
                                                                                                    >>..............
                                                                   ESRP1 Binding Site
48991 ctgacattgc ctcgatcgac agtgatgctg tcgttcaccc gacaaacact gacttctaca tcggtggtga agtaggtaat gcctagccgg gtgctgccga
      ..................Exon 6a variant 1.....................>>
      RBFox2 Binding Site
49091 gtgtgtgtgt gcatggtcag tcggccgctg cagacagctt gatcctttga cagctatgca gggctgcatt catttcacac ttccagctgt cctgaccctg

49191 tctcctgtcc ctgggttgtg aacctgagta ataaagctag ctgtcagcca ggatgtcatg catcctggct gatgcagctc acaaatggac atttctggtt
```

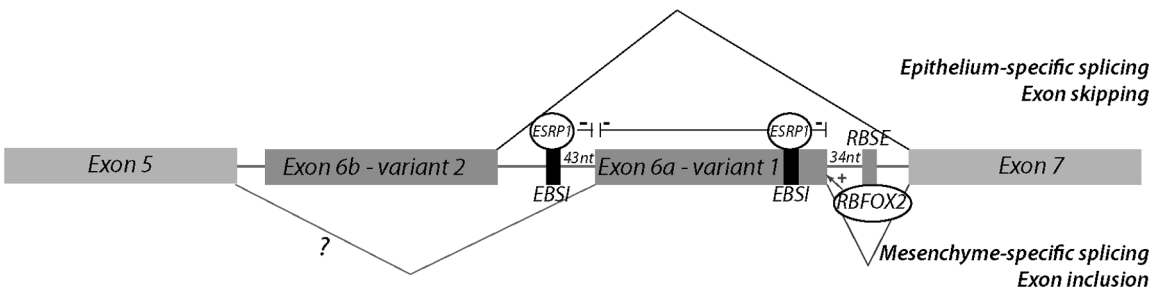
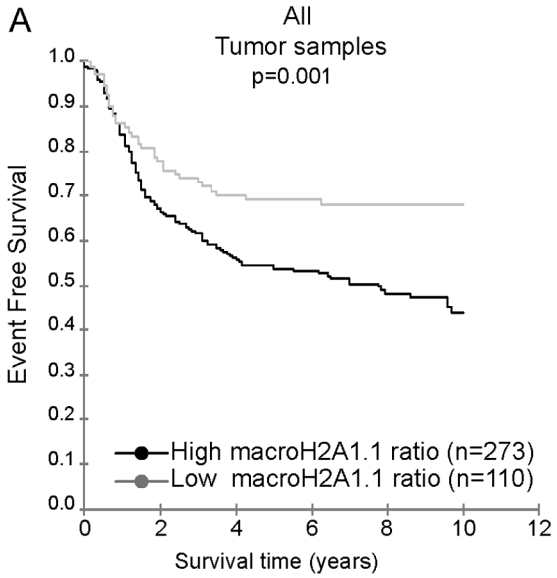


Figure 4. Overrepresentation of macroH2A1.1 correlated with mesenchymal features and induction of EMT features in HMLE cells is accompanied by an increase in macroH2A1.1/macroH2A1 mRNA ratio. A–B Analysis of GSE16795 data set [33]. MacroH2A1.1/macroH2A1 mRNA ratio are determined for each cell line and classified depending of the level of expression of E-cadherin (A); of the level of expression of Vimentin or N-cadherin (B). The median of macroH2A1.1/macroH2A1 values of each subgroup are specified. The reported *p*-values are the result of a two-tailed Mann-Whitney test. C- macroH2A1.1/macroH2A1 mRNA ratios were determined in immortalized human mammary epithelial cells with inhibiting E-cadherin function either shRNA-mediated (GSE9691 [34] and E-MTAB-884 [45]) or by expression of a truncated form of E-cadherin (ΔN-Ecad) (GSE9691 [34]) and compared. D- MacroH2A1.1/macroH2A1 mRNA ratios were determined for different breast cancer stem cell-like lines which overexpressed one EMT inducer, *i.e.* TGFβ, Twist, Gsc or Snail, and compared (GSE24202 [6]). The median of each subgroup is shown (grey bar). The reported *p*-values are the results of a two-tailed Mann-Whitney test. E- Upper panel- genomic sequence of H2AFY gene encompassing exon 6a. Potential ESRP1 and RBFox2 binding sites are represented with grey background and white letters. Bottom panel- Hypothetical schema for alternative splicing of exon 6a included in the macroH2A1.1 splice variant. Two cellular backgrounds are represented, epithelial with exon skipping of exon 6a related to the inhibitory binding of ESRP1 to EBSl, and mesenchymal with exon inclusion of exon 6a potentiated by binding of RBFox2 to RBSE.
doi:10.1371/journal.pone.0098930.g004

speculate that macroH2A1.1 stabilizes chromatin organizations characteristic of transcriptional programs linked to paused cell cycle progression. Hence, macroH2A1.1 expression could divert EMT-MET processes, stop progression and trap cells in such an intermediate state.

High macroH2A1.1 mRNA ratios in the slow cycling claudin-low molecular subtype [31] are correlated with earlier observations that macroH2A1.1 expression may be restricted to non-proliferative tissues [42], and that loss of its expression in lung and colon cancer was related to enhanced cell proliferation of cancer cells [24–26]. In the 67NR mouse model which formed primary carcinomas when implanted into mouse mammary fat pads, Dardenne et al., identified a high macroH2A1.1/macroH2A1 ratio. Inversely, in the 4T1 mouse model, reduced macroH2A1.1 expression was correlated with macroscopic metastatic capacity in the lung [43]. Our results point to high macroH2A1.1/macroH2A1 ratios as markers of engaged but paused intermediate cellular stages of the EMT. Because the metastasic power of a tumor clearly depends on a complete EMT-MET process, it is tempting to propose a model in which macroH2A1.1 is linked to the EMT process and macroH2A1.2 linked to the MET process.

A



Descriptive statistics (High) :		
Total observed	Events	Censored
273	135	138
Median survival time	Standart Deviation	95%CI
5.789	0.248	5.303-6.275

Descriptive statistics (Low) :		
Total observed	Events	Censored
110	34	76
Median survival time	Standart Deviation	95%CI
7.278	0.386	6.522-8.034

	Value	95%CI
Sensibility= TP/(TP+FN)	0.80	0.73-0.85
Specificity= TN/(TN+FP)	0.36	0.29-0.42
PPV= TP/(TP+FP)	0.49	0.44-0.55
NPV= TN/(TN+FN)	0.69	0.60-0.78
Odds ratio= TP*TN/(FP*FN)	2.19	1.37-3.49

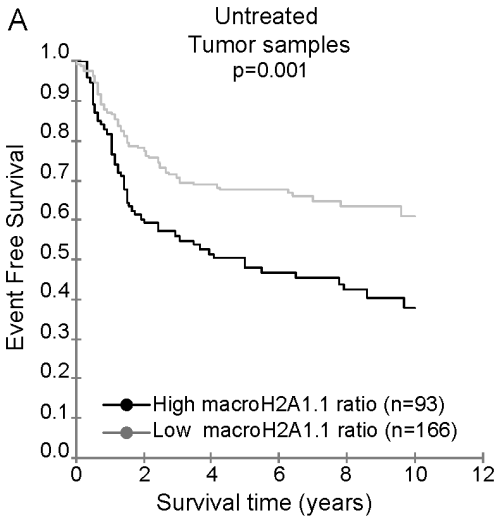
B

Multivariate analysis of Event Free survival according to standard parameters

Variable		No of patients*	Regression coefficient	Standard error	p-value	Hazard ratio	95%CI
Age	*mean 51.59*	*261*	*-0.012*	*0.009*	*0.194*	*0.988*	*0.970-1.006*
Tumor grade	*Grade 3 vs Grade 12*	*178 vs 83*	*0.205*	*0.241*	*0.395*	*1.227*	*0.765-1.969*
Tumor size	*>1cm vs ≤1cm*	*180 vs 81*	*0.126*	*0.243*	*0.604*	*1.134*	*0.704-1.827*
Lymph node status	*positive vs negative*	*39 vs 222*	*0.281*	*0.297*	*0.344*	*1.325*	*0.740-1.827*
macroH2A1.1 mRNA ratio	*cut-off determined by ROC*	*236 vs 25*	*1.240*	*0.425*	*0.036*	*3.457*	*1.087-10.990*

**Information of all parameters was available for 261 of the 383 TNBC samples with follow up data from GSE31519 data set*

Figure 5. Kaplan Meier analysis according to the macroH2A1.1 mRNA ratio. A- The 383 TNBC samples from the GSE31519 cohort were stratified according to the macroH2A1.1/macroH2A1 mRNA ratio. Kaplan Meier analysis of event free survival of 383 samples with follow up information is shown. Positive and negative predictive values (PPV and NPV) of macroH2A1.1/macroH2A1 mRNA ratio in the cohort are specified. **B-** Multivariate Cox proportional hazards models of disease-free survival.
doi:10.1371/journal.pone.0098930.g005



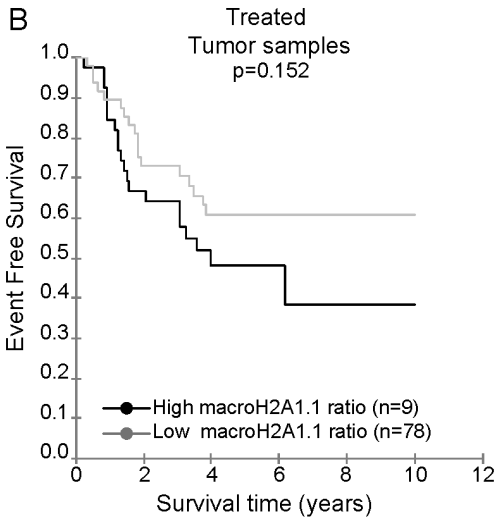
Descriptive statistics (High) :

Total observed	Events	Censored
93	54	39
Median survival time	**Standart Deviation**	**95%CI**
5.285	0.423	4.456-6.115

Descriptive statistics (Low) :

Total observed	Events	Censored
166	60	106
Median survival time	**Standart Deviation**	**95%CI**
7.017	0.308	6.414-7.621

	Value	95%CI
Sensibility= TP/(TP+FN)	0.47	0.38-0.56
Specificity= TN/(TN+FP)	0.73	0.65-0.80
PPV= TP/(TP+FP)	0.58	0.48-0.68
NPV= TN/(TN+FN)	0.64	0.57-0.71
*Odds ratio= TP*TN/(FP*FN)*	2.45	1.46-4.10



Descriptive statistics (High) :

Total observed	Events	Censored
39	20	19
Median survival time	**Standart Deviation**	**95%CI**
4.383	0.473	3.455-5.311

Descriptive statistics (Low) :

Total observed	Events	Censored
48	18	30
Median survival time	**Standart Deviation**	**95%CI**
6.791	0.600	5.615-7.968

	Value	95%CI
Sensibility= TP/(TP+FN)	0.53	0.37-0.67
Specificity= TN/(TN+FP)	0.61	0.47-0.74
PPV= TP/(TP+FP)	0.51	0.36-0.67
NPV= TN/(TN+FN)	0.63	0.49-0.76
*Odds ratio= TP*TN/(FP*FN)*	1.75	0.75-4.09

Figure 6. Kaplan Meier analysis according to the macroH2A1.1 mRNA ratio in untreated and treated sub-cohort. A- The 259 TNBC untreated samples from the GSE31519 cohort were stratified according to the macroH2A1.1/macroH2A1 mRNA ratio. Kaplan Meier analysis of event free survival of 259 samples with follow up information is shown. PPV and NPV of macroH2A1.1/macroH2A1 mRNA ratio in the untreated cohort are specified. **B-** The 87 TNBC adjuvant chemotherapy treated samples from the GSE31519 cohort were stratified according to the highest macroH2A1.1/macroH2A1 mRNA ratio. Kaplan Meier analysis of event free survival of 87 samples with follow up information is shown. PPV and NPV of macroH2A1.1/macroH2A1 mRNA ratio in the treated cohort are specified.
doi:10.1371/journal.pone.0098930.g006

TNBC is generally associated with a poor outcome, which is essentially not predicted by assessment of standard clinico-pathological variables, such as lymph node status or tumour size at initial presentation. The lack of identified molecular targets in the majority of TNBCs implies that chemotherapy remains the treatment of choice for patients with TNBCs. Here we show that, regardless of the reason that led to an absence of adjuvant therapy for patients involved in the GSE31519 study, those with a high macroH2A1.1/macroH2A1 mRNA ratio have a worse prognosis than those with a low one. Even if this observation clearly needs to be confirmed with a larger cohort, it is tempting to propose that assessing macroH2A1.1 expression levels will allow the identification of TNBC patients who, despite favorable clinic-pathological variables such as lymph node status or tumour size at initial presentation, will have a worse prognosis and may benefit from treatment. Interestingly, EMT and cell dissemination, although long associated with advanced stage of tumor progression, can be found at pre-neoplastic developmental stages of tumors [44]. Identifying early EMT process in primary tumors could then allow detection of tumors progressing towards metastasis. As expression of macroH2A1.1 seems to be correlated with EMT and unfavorable behavior in untreated TNBC patients, it is tempting to suggest macroH2A1.1 expression levels as an early biomarker of tumor genesis.

No difference in survival of patients who underwent adjuvant treatment was seen with respect to the macroH2A1.1/macroH2A1 mRNA ratio. But, as the ratios of this sub-cohort are already globally higher than in untreated tumors (Fig.S4), one could speculate that clinico-pathological parameters that initially led to treatment may already correlate with higher macroH2A1.1/

macroH2A1 mRNA ratio. It will be interesting to analyse this more in depth in a larger cohort.

In conclusion, it will be tempting to test if the correlation between macroH2A1.1 expression levels and EMT markers or poor prognosis in a TNBC cohort could be linked to a role for macroH2A1.1 in the maintenance of a transient cellular state in the early EMT process towards metastatic development of breast tumors.

Supporting Information

Figure S1 A- Characterization of α-macroH2A1 antibodies and cell lines. The specificity of α-macroH2A1 antibodies was verified using SDS-polyacrylamide gels low cross-linking (12.5% acrylamide, 1:125 bisacrylamide) to separate the two splice variants macroH2A1.1 and macroH2A1.2. Total extracts of breast cancer cell lines were first resolved in SDS-polyacrylamide gels (standard (left) or low cross-linking (right panel)) then immunoblotted with α-macroH2A1 and α-ERα antibodies. Left panel: ERα, macroH2A1 and H3 specific antibodies. Right: top panel; macroH2A1 specific; middle panel: macroH2A1.1 specific; bottom panel: macroH2A1.2 specific antibody. **B-** MacroH2A1.1, macroH2A1.2, and total macroH2A1 protein expression levels in five breast cancer cell lines. Total protein extracts were immunoblotted with α-macroH2A1 (bottom panel), α-macroH2A1.1 (top panel) or α-macroH2A1.2 antibodies (middle panel).
(TIF)

Figure S2 Primers, qPCR parameters and probe set IDs summary. A- Sequences, genomic location and size of amplicons generated by primers used in qPCR reactions are summarized. Parameters of the standard curve are reported for each pairs of primers in each cell lines used. For a given amplicon, efficiencies in the different cell lines are reported and compared each other (STDEV(Es)). **B-** Characterization of probe set ID of Affymetrix U133A, Illumina Human-6 v1 expression beadchip, IlluminaHuman HT-12 v3.0 expression beadchip arrays corresponding to macroH2A1 variants. For each probe, nucleotide reference sequences and macroH2A1 isoforms recognized are reported, as nucleotide and genomic localization of the sets of oligonucleotides presents at the probe ID. **C-** Clustal W multiple alignment of macroH2A1 variants sequences represented at the Probe Set ID 214500_at from U133A array. The sequences recognized by the probe are highlighted.
(TIF)

Figure S3 High macroH2A1.1 expression level in breast cancer cell lines characterizes Claudin-low molecular subtype. MacroH2A1.1/macroH2A1 mRNA ratios were determined for each cell line and classified according to molecular subtype assigned in the synthesis part of Table S1. In GSE16795 analysis [33], data from H3396 cell line are omitted as its subtype was not assigned. The median of macroH2A1.1/macroH2A1 values of each subgroup are specified. The reported *p*-values are the result of a two-tailed Mann-Whitney test.
(TIF)

Figure S4 Analysis of the distribution of macroH2A1.1/macroH2A1 mRNA values in the different groups of patients studied.
(TIF)

Table S1 Summary of molecular subtype of breast cancer cell lines used in this studies.
(PDF)

Table S2 Selected GEO dataset for cell lines analysis.
(XLSX)

Table S3 GSE31519 analysis.
(XLSX)

Acknowledgments

We would like to thank Andreas Ladurner (EMBL) for the macroH2A1.1 and 1.2 specific antibodies and Gilles Freiss (IRCM, Montpellier, France) for kindly providing ZR-75, MDA-MB436 and Hs578T cell lines. We would like to thank Pascal Martin (LBME) and Sébastien Déjean (IMT, Toulouse, France) for biostatistical advice.

Author Contributions

Conceived and designed the experiments: ACL KB. Performed the experiments: ACL MC JM SK MD. Analyzed the data: ACL. Wrote the paper: ACL KB.

References

1. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, et al. (2011) Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 121: 2750–2767.
2. Metzger-Filho O, Tutt A, de Azambuja E, Saini KS, Viale G, et al. (2012) Dissecting the heterogeneity of triple-negative breast cancer. J Clin Oncol 30: 1879–1887.
3. Perou CM (2011) Molecular stratification of triple-negative breast cancers. Oncologist 16 Suppl 1: 61–70.
4. Dittmar KA, Jiang P, Park JW, Amirikian K, Wan J, et al. (2012) Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. Mol Cell Biol 32: 1468–1482.
5. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, et al. (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med 15: 907–913.
6. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, et al. (2010) Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc Natl Acad Sci U S A 107: 15449–15454.
7. Stadler SC, Allis CD (2012) Linking epithelial-to-mesenchymal-transition and epigenetic modifications. Semin Cancer Biol.
8. Wang Y, Shang Y (2012) Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. Exp Cell Res.
9. Abell AN, Jordan NV, Huang W, Prat A, Midland AA, et al. (2011) MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. Cell Stem Cell 8: 525–537.
10. McDonald OG, Wu H, Timp W, Doi A, Feinberg AP (2011) Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat Struct Mol Biol 18: 867–874.
11. Onder TT, Kara N, Cherry A, Sinha AU, Zhu N, et al. (2012) Chromatin-modifying enzymes as modulators of reprogramming. Nature 483: 598–602.
12. Yang F, Sun L, Li Q, Han X, Lei L, et al. (2012) SET8 promotes epithelial-mesenchymal transition and confers TWIST dual transcriptional activities. EMBO J 31: 110–123.
13. De Craene B, Berx G (2013) Regulatory networks defining EMT during cancer initiation and progression. Nat Rev Cancer 13: 97–110.
14. Shapiro IM, Cheng AW, Flytzanis NC, Balsamo M, Condeelis JS, et al. (2011) An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. PLoS Genet 7: e1002218.
15. Warzecha CC, Carstens RP (2012) Complex changes in alternative pre-mRNA splicing play a central role in the epithelial-to-mesenchymal transition (EMT). Semin Cancer Biol.
16. Pehrson JR, Fried VA (1992) MacroH2A, a core histone containing a large nonhistone region. science 257: 1398–1400.
17. Costanzi C, Pehrson JR (1998) Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature 393: 599–601.
18. Rasmussen TP, Mastrangelo MA, Eden A, Pehrson JR, Jaenisch R (2000) Dynamic relocalization of histone MacroH2A1 from centrosomes to inactive X chromosomes during X inactivation. J Cell Biol 150: 1189–1198.
19. Zhang R, Poustovoitov MV, Ye X, Santos HA, Chen W, et al. (2005) Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev Cell 8: 19–30.

20. Barrero MJ, Sese B, Kuebler B, Bilic J, Boue S, et al. (2013) Macrohistone Variants Preserve Cell Identity by Preventing the Gain of H3K4me2 during Reprogramming to Pluripotency. Cell Rep.
21. Barrero MJ, Sese B, Marti M, Izpisua Belmonte JC (2013) Macro Histone Variants are Critical for the Differentiation of Human Pluripotent Cells. J Biol Chem.
22. Creppe C, Janich P, Cantarino N, Noguera M, Valero V, et al. (2012) MacroH2A1 regulates the balance between self-renewal and differentiation commitment in embryonic and adult stem cells. Mol Cell Biol 32: 1442–1452.
23. Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, et al. (2010) The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature 468: 1105–1109.
24. Novikov L, Park JW, Chen H, Klerman H, Jalloh AS, et al. (2011) QKI-mediated alternative splicing of the histone variant MacroH2A1 regulates cancer cell proliferation. Mol Cell Biol 31: 4244–4255.
25. Sporn JC, Jung B (2012) Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer. Am J Pathol 180: 2516–2526.
26. Sporn JC, Kustatscher G, Hothorn T, Collado M, Serrano M, et al. (2009) Histone macroH2A isoforms predict the risk of lung cancer recurrence. Oncogene 28: 3423–3428.
27. Belguise K, Milord S, Galtier F, Moquet-Torcy G, Piechaczyk M, et al. (2012) The PKCtheta pathway participates in the aberrant accumulation of Fra-1 protein in invasive ER-negative breast cancer cells. Oncogene 31: 4889–4897.
28. Schefe JH, Lehmann KE, Buschmann IR, Unger T, Funke-Kaiser H (2006) Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. J Mol Med (Berl) 84: 901–910.
29. Rody A, Karn T, Liedtke C, Pusztai L, Ruckhaeberle E, et al. (2011) A clinically relevant gene signature in triple negative and basal-like breast cancer. Breast Cancer Res 13: R97.
30. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, et al. (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10: 515–527.
31. Prat A, Parker JS, Karginova O, Fan C, Livasy C, et al. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 12: R68.
32. Lapuk A, Marr H, Jakkula L, Pedro H, Bhattacharya S, et al. (2010) Exon-level microarray analyses identify alternative splicing programs in breast cancer. Mol Cancer Res 8: 961–974.
33. Hollestelle A, Nagel JH, Smid M, Lam S, Elstrodt F, et al. (2010) Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines. Breast Cancer Res Treat 121: 53–64.
34. Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, et al. (2008) Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Cancer Res 68: 3645–3654.
35. Carmody LC, Germain AR, VerPlank L, Nag PP, Munoz B, et al. (2012) Phenotypic high-throughput screening elucidates target pathway in breast cancer stem cell-like cells. J Biomol Screen 17: 1204–1210.
36. Scheel C, Eaton EN, Li SH, Chaffer CL, Reinhardt F, et al. (2011) Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. Cell 145: 926–940.
37. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, et al. (2008) Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res 14: 5158–5165.
38. Gaspar-Maia A, Qadeer ZA, Hasson D, Ratnakumar K, Leu NA, et al. (2013) MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. Nat Commun 4: 1565.
39. Pasque V, Radzisheuskaya A, Gillich A, Halley-Stott RP, Panamarova M, et al. (2012) Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. J Cell Sci 125: 6094–6104.
40. Ocana OH, Corcoles R, Fabra A, Moreno-Bueno G, Acloque H, et al. (2012) Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer Cell 22: 709–724.
41. Tsai JH, Donaher JL, Murphy DA, Chau S, Yang J (2012) Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. Cancer Cell 22: 725–736.
42. Pehrson JR, Costanzi C, Dharia C (1997) Developmental and tissue expression patterns of histone macroH2A1 subtypes. J Cell Biochem 65: 107–113.
43. Dardenne E, Pierredon S, Driouch K, Gratadou L, Lacroix-Triki M, et al. (2012) Splicing switch of an epigenetic regulator by RNA helicases promotes tumor-cell invasiveness. Nat Struct Mol Biol 19: 1139–1146.
44. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, et al. (2012) EMT and dissemination precede pancreatic tumor formation. Cell 148: 349–361.
45. Germain AR, Carmody LC, Morgan B, Fernandez C, Forbeck E, et al. (2012) Identification of a selective small molecule inhibitor of breast cancer stem cells. Bioorg Med Chem Lett 22: 3571–3574.
46. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, et al. (2006) Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 10: 529–541.
47. Schmidt M, Bohm D, von Torne C, Steiner E, Puhl A, et al. (2008) The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res 68: 5405–5413.
48. Zhang Y, Sieuwerts AM, McGreevy M, Casey G, Cufer T, et al. (2009) The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. Breast Cancer Res Treat 116: 303–309.
49. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, et al. (2009) Genes that mediate breast cancer metastasis to the brain. Nature 459: 1005–1009.
50. Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, et al. (2005) Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res 7: R953–964.
51. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, et al. (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365: 671–679.
52. Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, et al. (2007) Lung metastasis genes couple breast tumor size and metastatic spread. Proc Natl Acad Sci U S A 104: 6740–6745.
53. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, et al. (2005) Genes that mediate breast cancer metastasis to lung. Nature 436: 518–524.
54. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, et al. (2006) Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 98: 262–272.
55. Ruckhaberle E, Karn T, Engels K, Turley H, Hanker L, et al. (2010) Prognostic impact of thymidine phosphorylase expression in breast cancer—comparison of microarray and immunohistochemical data. Eur J Cancer 46: 549–557.
56. Rody A, Karn T, Ruckhaberle E, Hanker L, Metzler D, et al. (2009) Loss of Plexin B1 is highly prognostic in low proliferating ER positive breast cancers—results of a large scale microarray analysis. Eur J Cancer 45: 405–413.
57. Rody A, Karn T, Solbach C, Gaetje R, Munnes M, et al. (2007) The erbB2+ cluster of the intrinsic gene set predicts tumor response of breast cancer patients receiving neoadjuvant chemotherapy with docetaxel, doxorubicin and cyclophosphamide within the GEPARTRIO trial. Breast 16: 235–240.
58. Ruckhaberle E, Rody A, Engels K, Gaetje R, von Minckwitz G, et al. (2008) Microarray analysis of altered sphingolipid metabolism reveals prognostic significance of sphingosine kinase 1 in breast cancer. Breast Cancer Res Treat 112: 41–52.
59. Rody A, Holtrich U, Gaetje R, Gehrmann M, Engels K, et al. (2007) Poor outcome in estrogen receptor-positive breast cancers predicted by loss of plexin B1. Clin Cancer Res 13: 1115–1122.
60. Miller LD, Smeds J, George J, Vega VB, Vergara L, et al. (2005) An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci U S A 102: 13550–13555.
61. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, et al. (2007) Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol 25: 1239–1246.
62. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, et al. (2007) Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res 13: 3207–3214.
63. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, et al. (2008) Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics 9: 239.