

# Global Histone Modifications in Breast Cancer Correlate with Tumor Phenotypes, Prognostic Factors, and Patient Outcome

Somaia E. Elsheikh,<sup>1,7</sup> Andrew R. Green,<sup>1</sup> Emad A. Rakha,<sup>1</sup> Des G. Powe,<sup>1</sup> Rabab A. Ahmed,<sup>1,8</sup> Hilary M. Collins,<sup>2</sup> Daniele Soria,<sup>3</sup> Jonathan M. Garibaldi,<sup>3</sup> Claire E. Paish,<sup>1</sup> Amr A. Ammar,<sup>7</sup> Matthew J. Grainge,<sup>4</sup> Graham R. Ball,<sup>6</sup> Magdy K. Abdelghany,<sup>2,9</sup> Luisa Martinez-Pomares,<sup>5</sup> David M. Heery,<sup>2</sup> and Ian O. Ellis<sup>1</sup>

<sup>1</sup>Department of Histopathology, School of Molecular Medical Sciences, University of Nottingham and Nottingham Universities Hospital Trust, Schools of <sup>2</sup>Pharmacy and <sup>3</sup>Computer Science, <sup>4</sup>Division of Epidemiology and Public Health, and <sup>5</sup>School of Molecular Medical Sciences, Institute of Infection, Immunity and Inflammation, University of Nottingham; <sup>6</sup>Division of Life Sciences, Nottingham Trent University, Nottingham, United Kingdom; <sup>7</sup>Department of Pathology, Faculty of Medicine, Menoufya University, Menoufya, Egypt; <sup>8</sup>Department of Pathology, Faculty of Medicine, Asuit University, Asuit, Egypt; and <sup>9</sup>Department of Pathology, Faculty of Medicine, Suez Canal University, Ismailia, Egypt

## Abstract

Post-translational histone modifications are known to be altered in cancer cells, and loss of selected histone acetylation and methylation marks has recently been shown to predict patient outcome in human carcinoma. Immunohistochemistry was used to detect a series of histone lysine acetylation (H3K9ac, H3K18ac, H4K12ac, and H4K16ac), lysine methylation (H3K4me2 and H4K20me3), and arginine methylation (H4R3me2) marks in a well-characterized series of human breast carcinomas ( $n = 880$ ). Tissue staining intensities were assessed using blinded semiquantitative scoring. Validation studies were done using immunofluorescence staining and Western blotting. Our analyses revealed low or absent H4K16ac in the majority of breast cancer cases (78.9%), suggesting that this alteration may represent an early sign of breast cancer. There was a highly significant correlation between histone modifications status, tumor biomarker phenotype, and clinical outcome, where high relative levels of global histone acetylation and methylation were associated with a favorable prognosis and detected almost exclusively in luminal-like breast tumors (93%). Moderate to low levels of lysine acetylation (H3K9ac, H3K18ac, and H4K12ac), lysine (H3K4me2 and H4K20me3), and arginine methylation (H4R3me2) were observed in carcinomas of poorer prognostic subtypes, including basal carcinomas and HER-2-positive tumors. Clustering analysis identified three groups of histone displaying distinct pattern in breast cancer, which have distinct relationships to known prognostic factors and clinical outcome. This study identifies the presence of variations in global levels of histone marks in different grades, morphologic types, and phenotype classes of invasive breast cancer and shows that these differences have clinical significance. [Cancer Res 2009;69(9):3802–9]

## Introduction

Breast cancer is a heterogeneous disease ranging from premalignant hyperproliferation to invasive and metastatic carcinomas (1). Disease progression is poorly understood but is likely due to the accumulation of genetic mutations leading to widespread changes in gene expression and, in particular, affecting the expression of tumor suppressors and oncogenes (2). Consistent with this, recent studies have shown that different breast tumor subclasses display distinct gene expression profiles (3, 4). In addition to genetic mutations, there is increasing evidence for gross epigenetic alteration in tumor cells both at the levels of DNA methylation and histone marks (5, 6). Chromatin-modifying enzymes play a key role in gene regulation by catalyzing reversible post-translational modifications of histones, including lysine acetylation, lysine methylation, and arginine methylation (7). These and other modifications generate a combinatorial histone code that demarcates chromatin regions for transcription activation or repression (8). Although the “epigenetic” code is not fully understood, specific marks such as lysine acetylation (H3K9ac, H3K18ac, and H4K12ac), lysine trimethylation (H3K4me3), and arginine dimethylation (H4R3me2) are associated with transcriptionally active gene promoters (9–11), whereas other modifications such as lysine methylation (H3K9me2 or H3K9me3 and H4K20me3) are associated with repressed chromatin (7, 8). Global loss of acetylation (K16) and trimethylation (K20) of histone H4 have been shown to be hallmarks of human cancer (12). Such changes in global histone modification patterns can be predictive of clinical outcome as recently shown for prostate, lung, and gastric cancers (13–16); thus, there is a need to examine epigenetic alterations in breast cancers to define new prognostic markers and therapeutic targets. We therefore assessed the biological and clinical significance of global patterns of selected histone modifications in breast cancer using tissue microarray technology and immunohistochemistry. The relative levels of seven modified histones, including H3K18ac, H3K9ac, H4R3me2, H3K4me2, H4K12ac, H4K16ac, and H4K20me3, were determined in a large well-characterized series of 880 primary operable invasive breast carcinoma cases.

Digital image analysis was used for quantification of the epigenetic marker on a subset of tumors. Tumors were classified according to their relative level of the histone modifications using two different clustering methods, the K-means and the partitioning around medoids. The relationships between histone modifications, clinicopathologic data, and patient outcome were assessed for each marker individually and combinatorially.

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

**Requests for reprints:** Ian O. Ellis, Department of Histopathology, Nottingham Universities Hospitals City Campus, Hucknall Road, Nottingham NG5 1PB, United Kingdom. Phone: 44-115-9691169, ext. 56875; Fax: 440-115-9627768; E-mail: ian.ellis@nottingham.ac.uk

©2009 American Association for Cancer Research.  
doi:10.1158/0008-5472.CAN-08-3907

## Materials and Methods

**Patients.** The tissue microarrays comprised a cohort of 880 consecutive breast tumors diagnosed from 1986 to 1998 included in the Nottingham Tenovus Primary Breast Carcinoma Series. Histologic tumor types comprised 449 invasive ductal carcinomas of no special type, 182 tubular mixed carcinomas, 25 medullary carcinomas, 83 lobular carcinomas, 28 tubular carcinomas, 8 mucinous carcinomas, 6 cribriform carcinomas, 4 papillary carcinomas, 29 mixed no special type and lobular carcinomas, 23 mixed no special type and special type carcinomas, and 6 miscellaneous tumors. Full details of the characterization of the tissue microarray and the cohort of the patients are described elsewhere (17, 18). Patient management was based on tumor characteristics provided by the Nottingham prognostic index (NPI; ref. 19) and hormone receptor status. Patients with NPI score  $\leq 3.4$  received no adjuvant therapy and those with NPI score  $>3.4$  received tamoxifen if estrogen receptor-positive ( $\pm$ zoladex if premenopausal) or classic cyclophosphamide, methotrexate, and 5-fluorouracil if estrogen receptor-negative and fit enough to tolerate chemotherapy (20). Tumors were graded according to a modified Bloom-Richardson scoring system (21) and size was categorized according to the tumor-node-metastasis staging criteria (22). NPI was calculated as described previously (23). Survival data including disease-free survival (DFS), metastatic-specific survival, and breast cancer-specific survival (BCSS) were maintained on a prospective basis. DFS and metastatic-specific survival were defined as the interval (in months) from the date of the primary surgical treatment to the first locoregional or distant recurrence, respectively. BCSS survival was taken as the time (in months) from the date of the primary surgical treatment to the time of death from breast cancer. This study was approved by the Nottingham Research Ethics Committee 2 under the title of "Development of a Molecular Genetic Classification of Breast Cancer." None of the authors has any competing interests.

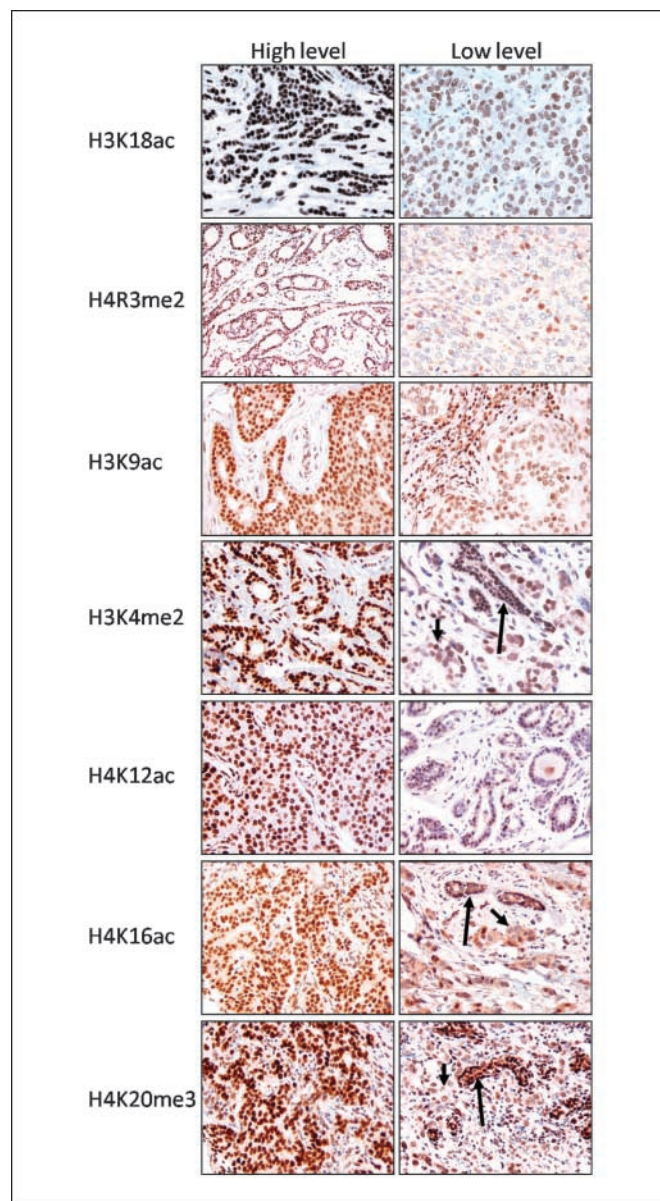
**Construction of the tissue microarray blocks.** Breast cancer tissue microarrays were prepared as described previously (24). Each case was sampled twice from both the center and the periphery of the tumor. Arrays of 150 cases per block were prepared.

**Immunohistochemistry on the tissue microarray.** Breast cancer tissue microarray slides were prepared and immunohistochemically stained to detect the seven histone marks and the other markers as described previously (24–27). The histone marks selected for study were those previously identified to be of significance in human cancer (12–14). The antibodies were initially optimized using 20 cases of whole-tissue sections to assess patterns of histone staining in both normal and malignant breast tissues. As a control, we also used an anti-H3 antibody, which detects both modified and unmodified H3 (1:1,000; Abcam). Primary rabbit anti-histone polyclonal antibodies were applied for 30 min at room temperature at the following dilutions: H3K18ac (1:2,000), H3K9ac (1:800), H3K4me2 (1:800), H4K12ac (1:200), H4K16ac (1:500), and H4K20me3 (1:500; all Abcam) and H4R3me2 (1:75; Upstate Biotechnology). See Supplementary Table S2 for full details of the antibodies. Prostate carcinoma tissue was used as a positive control and identical array sections stained in the absence of the primary antibody were used as negative controls.

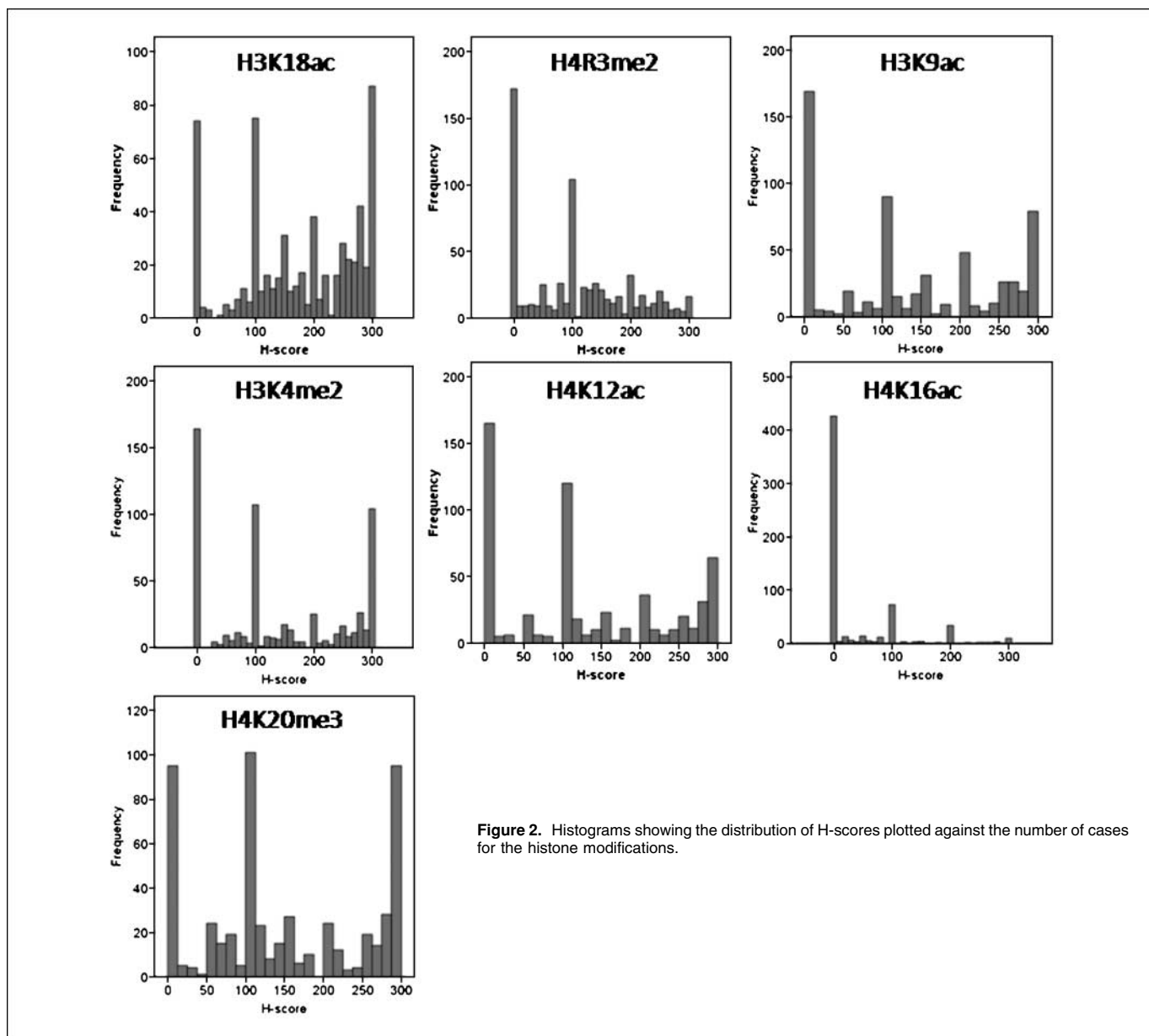
**Scoring of immunohistochemical reactivity.** After excluding any uninformative tissue microarray cores, which were lost, fragmented, or did not have invasive tumor, cases available for scoring were as follows: H3K18ac (613), H4R3me2 (683), H3K4me2 (610), H3K9ac (621), H4K12ac (597), H4K16ac (606), and H4K20me3 (557). Immunohistochemical scoring was done by the modified Histo-score (H-score; ref. 28), which involves semiquantitative assessment of both the intensity of staining (graded as 1–3) and the percentage of positive cells. The range of possible scores is 0 to 300, enabling us to explore rationalization of our cases into biologically relevant groups depending on different levels of detection, which could potentially be missed using simpler scoring methods (e.g., positive versus negative). Two observers (S.E.E. and A.R.G.) scored all cases and correlation between both scores was done. Tumor samples with H-score  $<100$  for individual chromatin marks were designated as low detection, where scores  $\geq 100$  were designated high detection. The distribution of

staining was assessed both in whole sections of normal and malignant breast carcinoma (20 cases) and in tissue microarray sections. As the distribution of staining was homogenous in the whole section, only one tumor core was stained from each tumor, as previous studies have validated the use of one core to study the expression of tumor markers even for those that have a heterogeneous distribution (29). Validation studies were done on a series of 9 cases using immunofluorescence staining and Western blotting for H3K18ac (see Supplementary Material and Methods). Because of the large number of negative cases observed with H4K16ac detection, we reassessed this result using a different antibody from a different supplier (see Supplementary Materials and Methods and Supplementary Table S2).

**Statistical analysis.** Statistical analysis was done using SPSS 13.0 statistical software. Median follow-up was defined as follow-up period for



**Figure 1.** Detection of histone modifications in invasive breast carcinoma as determined by immunohistochemistry and their histograms. Representative examples of breast tumor tissue cores presenting with high and low levels of the following histone modifications: H3K18ac, H4R3me2, H3K9ac, H3K4me2, H4K12ac, H4K16ac, and H4K20me3. Note the high level of H3K4me2, H4K16ac, and H4K20me3 in normal breast acini (long arrow) in comparison with the surrounding malignant cells (short arrow). Original magnification,  $\times 200$ .



**Figure 2.** Histograms showing the distribution of H-scores plotted against the number of cases for the histone modifications.

those patients still alive and disease-free at their latest hospital visit. Cutoff values for the different biomarkers included in this study were chosen before statistical analysis. Identical analysis was conducted for continuous noncategorical histone data, which showed similar results, suggesting that dichotomization had not weakened the evidence for an association in this instance. Standard cutoffs were used for established prognostic factors and were the same as for previously published patient series (25). All factors were used as dichotomous covariates in the statistical analysis, with the exception of histone clusters (high, moderate, and low), grade, NPI, and phenotypic groups proposed by Nielsen and colleagues (30) that were divided into three groups. Unweighted  $\kappa$  coefficient test was used to assess agreement between observers of the same variables. To test whether histone marks individually or in clusters differed according to clinicopathologic variables and biological markers, Pearson  $\chi^2$  and Spearman correlation was used to perform categorical and continuous statistical analysis, respectively. All  $P$  values were two-sided and  $P < 0.05$  was considered significant, with confidence interval of 95%.

Cumulative survival probabilities were calculated using the Kaplan-Meier method. Differences in DFS, metastatic-specific survival, and BCSS

based on modified histone expression were estimated using log-rank test. Multivariate analysis was done using the Cox multiple hazards model to test the statistical independence and significance of predictors on DFS, and BCSS, where the model was fitted to adjust for the effect of variables that are known to predict breast cancer prognosis based on information from the present cohort and elsewhere (31). Histone marks were then fitted individually or in clusters to the model to assess their association with BCSS and DFS independent of these key prognostic variables.

**Clustering methods.** Two different algorithms were used for cluster analysis: the K-means and the partitioning around medoids methods. They were both run for between 2 and 20 clusters. Six validity indices were calculated and recorded to determine the best number of clusters. The indices are Calinski and Harabasz, Hartigan, Scott and Symons, Marriot, TraceW, and TraceW-1B. For each index, the number of groups to be considered was chosen according to the rules reported in ref. 32.

**K-means.** The K-means technique aims to partition the data into  $k$ -groups such that the sum of squares from points to the assigned cluster centers is minimized. K-means clustering is dependent on the initial setting

**Table 1.** Detection of histone modifications in invasive breast carcinoma

Histone marks	Low detection (%)	High detection (%)
H3K18ac	114/613 (18.6)	499/613 (81.4)
H3K9ac	331/621 (50.1)	290/621 (49.9)
H4K12ac	328/597 (54.9)	269/597 (45.1)
H4R3me2	396/683 (58)	287/683 (42)
H3K4me2	316/610 (51.8)	294/610 (48.2)
H4K16ac	478/606 (78.9)	128/606 (21.1)
H4K20me3	168/557 (30.2)	389/557 (69.8)

of the cluster assignments, and for this study, we used a fixed initialization obtained with hierarchical clustering.

**Partitioning around medoids.** The partitioning around medoids algorithm is based on the search for *k*-representative objects (the so-called medoids) among the observations of the data set. These observations should represent the structure of the data. After finding a set of *k*-medoids, *k*-clusters are constructed by assigning each observation to the nearest medoid. The goal is to find *k*-representative objects, which minimize the sum of the dissimilarities of the observations to their closest representative object. Dissimilarities are nonnegative numbers that are close to zero when two points are near each other and that become large when they are very different (32). Biplots are generated by firstly transforming the original data space using principal component analysis and then plotting the points at their projected position on axes of the first and second principal components (33).

## Results

**Detection of histone marks in breast tumor tissue.** Tissue microarrays of 880 breast tumor cores and normal breast tissue controls were stained for the presence of seven distinct chromatin marks (H3K9ac, H3K18ac, H4K12ac, H4K16ac, H3K4me2, H4K20me3, and H4R3me2) using immunohistochemistry (Materials and Methods). Cases were scored for each mark using a modified H-score, and good agreement was found between two independent scorers (unweighted  $\kappa$  score = 0.57879; 0.4678-0.6896). In normal breast tissue, all seven histone marks were detected in the nuclei of normal mammary epithelial cells in the acini of terminal duct lobular units. Some positive nuclear staining was also observed in myoepithelial cells, stromal cells, and lymphocytes. There was observed heterogeneity of staining in these populations (Supplementary Fig. S1). Positive nuclear staining for all histone marks was also detected in breast tumor cells, and although the staining intensity was similar in the tumor cell

population within each case, it varied in intensity between individual cases (Fig. 1). Histograms showing the staining intensity and distribution of H-scores plotted against the number of cases are shown in Fig. 2 (*right column*). We identified that H4K16ac was different in its expression levels from the other chromatin marks assessed, in that it was found to be at low or undetectable levels in the majority of breast tumor cases. Similar results for H4K16ac were obtained using antibody from other supplier with a good correlation between the two antibodies ( $P < 0.001$ ; see Supplementary Data).

To explore such differences further, tumors with H-scores of  $<100$  for a particular chromatin mark were classified as a low detection group, whereas H-scores of  $\geq 100$  were scored as high detection. This revealed that most breast tumors (78.9%) scored low for H4K16ac, whereas H3K18ac and H4K20me3 were scored at relatively high levels in the majority of cases (81.4% and 69.8%, respectively; Table 1). There was a highly significant positive correlation between the levels all of the histone marks assessed within tumors ( $P < 0.001$ ; Supplementary Table S3). This suggests that global levels of histone marks vary among different breast tumors, which could be related to underlying biology or prognostic and clinicopathologic factors. The validation studies done on a series of 9 cases using immunofluorescence staining and Western blotting for H3K18ac showed comparable findings (see Supplementary Results and Supplementary Fig. S2).

**Histone marks correlate with clinicopathologic factors.** Complete clinical and follow-up data were available for 839 of 880 invasive breast cancer cases in this study (Supplementary Results and Supplementary Table S1). Significant correlations were found between low and high levels of individual histone marks with a range of clinicopathologic variables (Supplementary Table S4). In particular, high tumor grade was associated with low levels of all seven histone marks (Supplementary Fig. S3), and favorable NPI scores were also associated with high levels of detection of all the histone marks, except H3K18ac. Low detection levels of H4R3me2, H3K9ac, and H4K16ac were significantly associated with large tumor size and high detection levels of H4R3me2 and H3K9ac were associated with low lymph node stage. Positive vascular invasion was found to be associated with low levels of H4K16ac. These results may indicate that reduced detection of selected histone modifications correlates with poor prognostic characteristics.

**Relationship to histologic tumor type.** When histone modifications were examined with respect to their relationship with well-recognized morphologic tumor types (Supplementary Table S5), it was observed that there were widespread differences in their distribution between various tumor morphologic types. The most

**Table 2.** Histone modification clusters and phenotype groups of breast cancer as defined by Nielsen and colleagues (30)

Nielsen and colleagues' classes of breast cancer	Histone modification clusters*		
	Hypermodified cluster (%)	Intermediate-level cluster (%)	Hypomodified cluster (%)
Basal	2 (4.9)	8 (19.5)	31 (75.6)
Luminal	89 (45.4)	65 (33.2)	42 (21.4)
HER-2	4 (10.0)	13 (32.5)	23 (57.5)

\* $P < 0.001$ ,  $\chi^2$  test.

**Table 3.** Multivariate proportional hazards analysis including H3K18ac

Variables	BCSS		DFS	
	Hazard ratio (95% confidence interval)	<i>P</i>	Hazard ratio (95% confidence interval)	<i>P</i>
Grade*	1.01 (1.85-4.09)	<0.001	0.415 (1.22-1.88)	<0.001
Tumor size*	0.147 (0.94-1.41)	0.148	0.548 (0.90-1.24)	0.477
Lymph node stage (positive)	1.08 (1.89-4.59)	<0.001	0.059 (1.26-2.35)	0.001
H3K18ac status (positive)	-0.47 (1.01-2.57)	0.046	-0.448 (0.45-0.90)	0.011

\*Fitted as linear term, that is, increase in risk for change 1 unit in grade or size.

striking difference was the observed low-level detection of all histone marks in medullary carcinoma. In contrast, lobular, mucinous, tubular, and mixed tubular carcinomas were associated with high levels of global histone marks. This emphasizes the association between the high level of histone marks and the well-recognized good prognostic special types of invasive breast cancer.

**Relationship to biological markers and phenotypic groups of breast cancer.** Complementary evaluation of biological markers (Supplementary Table S6) showed that high-level detection of histone modifications was significantly correlated with steroid receptor (estrogen receptor, progesterone receptor, and androgen receptor)-positive tumors, high expression of luminal cytokeratins (CK7/8 and CK18), high E-cadherin and BRCA1 expression, and low expression of basal cytokeratins (CK5/6 and CK14), p53 and HER-2. These results were reflected in the relationships between histone marks either separately (Supplementary Table S7) or when combined in the cluster analysis (see below; Table 2). Similarly, in the key phenotype groups of breast cancer proposed by Nielsen and colleagues (30), the poor prognostic classes of basal and HER-2 were strongly represented in the low histone marks cluster at frequency levels of 75% and 57%, respectively ( $P < 0.001$ ).

These results may indicate the biological significance of histone modification and the association between low histone marks and poor prognostic classes of breast cancer.

**Histone marks and patient outcome.** Assessment of the relationship between histone marks and patient outcome showed that, in the whole cohort, low-level detection of histone marks was associated with adverse patient outcome and high detection was significantly associated with a more favorable BCSS (except H4K20me3), longer DFS with high level of H3K18ac, H4R3me2, and H3K9ac (Supplementary Fig. S4), and metastatic-specific survival (except H4K20me3; data not shown). Multivariate Cox regression analysis showed that only the prognostic effect of H3K18ac marks was independent of other key prognostic factors in breast cancer including histologic grade, lymph node stage, and tumor size with respect to BCSS and DFS (Table 3). We further observed that those patients who did not receive postoperative adjuvant therapy had better BCSS, longer DFS, and metastatic-specific survival when high levels of histone marks, particularly H3K18ac, H3K9ac, and H4K16ac, were present. However, in those patients receiving hormonal therapy, the relationships between histone marks and patient outcome were diminished. Furthermore, none of the histone marks was found to have a statistically significant relationship to the outcome of patients who received chemotherapy.

**Clustering of histone marks in breast tumors.** To explore the relationships between individual histone marks, unsupervised clustering analysis was done. A total of 408 cases with a complete data set for the seven epigenetic marks were available for clustering analysis using K-means and partitioning around medoids methods. Four of six validity indices indicated that acceptable classification could be obtained, which divided the data set into three groups (Supplementary Figs. S5 and S6). From the biplots of the clusters obtained for each method (Supplementary Figs. S7 and S8), it can be seen that a good separation among groups was achieved. The final three clusters were derived from the overlap of the two clustering methods used. It was found that 79.4% (324 of 408) of cases was assigned to the classes, whereas the remaining 20.6% were placed into a "not classified" group (Supplementary Fig. S9). Box plots of final clustering revealed the characterization of the three classes, where patients with higher H4K16ac H-scores than zero are classified in the hypermodified cluster, whereas those with H3K12ac and H3K4me2 H-scores equal to zero are more likely to be assigned to the hypomodified cluster. However, it was quite clear that, with the exception of H4K16ac, all six other histone modifications differ between the three groups, where hypermodified cluster showed high detection of the six marks, intermediate cluster showed moderate detection, and hypomodified cluster showed low detection of the six marks (Fig. 3A).

When associations with prognostic factors and clinical outcome were examined, cases in the hypermodified cluster were associated with low histologic grade, favorable NPI score (Supplementary Table S8), and longer BCSS and DFS (Fig. 3B). Moreover, these histone mark clusters were also associated with BCSS and DFS, which, in multivariate analysis, was shown to be independent of tumor size and lymph node stage (Supplementary Table S9). This analysis shows clearly that we have identified three groups of histone modifications that have distinct relationships to known prognostic factors and clinical outcome.

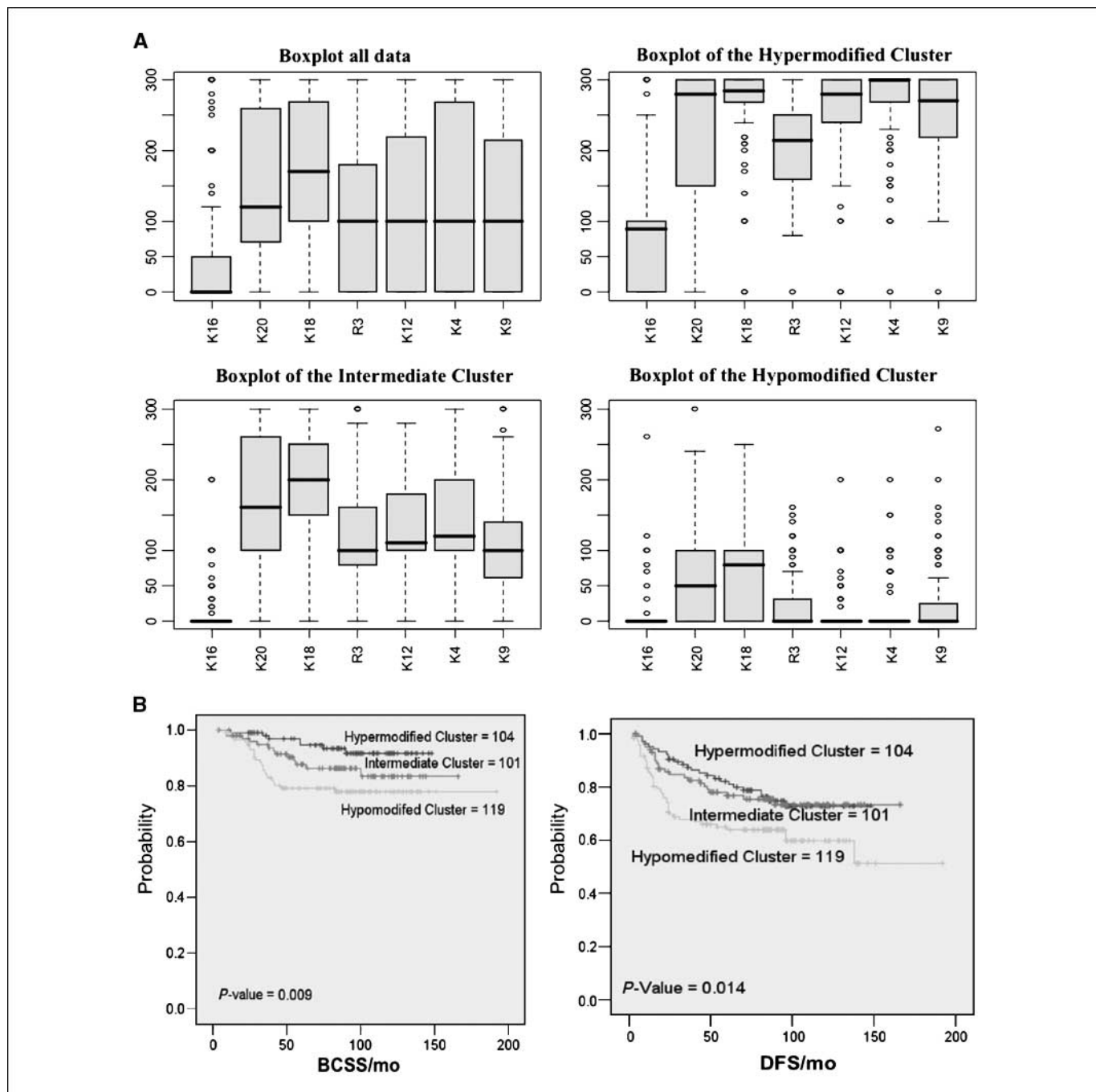
## Discussion

Our study has revealed differential levels of bulk histone acetylation and methylation modifications across a large cohort of breast tumors and that the occurrence of certain of these histone marks correlates with tumor morphology and biological subtype. In general, tumors with adverse traditional prognostic or phenotypic characteristics were found to have reduced levels of detectable H3 and H4 acetylation and methylation marks. In contrast, normal breast epithelium or tumors with good prognostic characteristics

were generally found to exhibit relatively higher levels of the chromatin modification subset assessed here. It is notable that H4K16ac, which was hypoacetylated in the majority of tumors, remained associated with better prognosis where high levels of acetylation of that mark were present. These observations suggest that global histone modifications play an important role in breast cancer progression and that loss of HK16ac may be an early event in the pathogenesis of invasive breast cancer.

Global reduction of histone H4K16ac and H4K20me3 has been reported to be hallmarks of human cancer, and loss of these

modifications has been detected in the promoters of tumor suppressor genes (12). Acetylation of H4K16 is implicated in both transcriptional activation and maintenance of euchromatin and is the only histone modification known to interfere with the formation of repressive heterochromatin and chromatin remodeling (34). H4K16 is acetylated by MYST acetyltransferases such as MOF and deacetylated by the NAD-dependent HDAC SIRT1 (35). Interestingly, MOF was recently reported to be frequently down-regulated in medulloblastomas and breast carcinomas (36). In addition, a gene designated “deleted in breast cancer” (DBC) was



**Figure 3.** A, box plots for the seven epigenetic markers grouped by clusters. B, unadjusted Kaplan-Meier curves for hypermodified cluster (high-level detection of the seven epigenetic marks), intermediate cluster (moderate-level detection), and hypomodified cluster (low-level detection) in invasive breast cancers with respect to BCSS and DFS.

recently shown to encode a negative regulator of SIRT1 (37, 38). It remains to be established whether aberrant function of these or other H4K16 regulatory factors is responsible for the hypoacetylation of H4K16 in breast cancers and whether loss of acH4K16 has a causal or priming role in tumorigenesis.

Methylation of H4K20 is complex and is catalyzed by several methyltransferases, including PRSET7, PRDM6, and NSD1. H4K20 trimethylation was reported to be reduced in malignant breast cancer cell lines compared with nontumorigenic breast epithelial cells (39). Loss of H4K20me3 has been observed in animal models of carcinogenesis (40), including estradiol-induced mammary carcinogenesis in the rat (41). Thus, our finding that H4K20me3 is detected at low levels in human breast and tumors with adverse prognostic (including higher tumor grade and NPI) and phenotypic characteristics shows that this epigenetic alteration appears to relate to or be a sign of biologically more aggressive forms of breast cancer.

Other histone modifications examined in the study are generally believed to be chromatin marks that are associated with gene transcription, which may facilitate the subsequent recruitment of chromatin modification complexes or have structural roles in chromatin remodeling. Although our study has detected differential levels of bulk histone modifications in tumor sections, the effect of this on gene-specific chromatin modifications on a genome-wide basis remains to be determined. However, the identification of hypermodified and hypomodified tumor clusters and their correlation with clinical and biological prognostic factors is of particular interest. The hypermodified cluster (with the exception of H4K16) includes tumors of the luminal phenotype class of breast cancer proposed by Nielsen and colleagues (30), those of low histologic grade, and those of favorable NPI. In contrast, the hypomodified cluster is associated with the basal and HER-2 phenotypic classes. These findings show a clear relationship between altered levels of modified histones and biological classes of breast cancer. In agreement with this, there is increasing evidence from recent gene profile analysis and protein expression studies in invasive breast cancer that breast cancer includes at least three main "biological" subclasses (luminal, HER-2, and basal classes), which can be defined according to their gene or protein expression characteristics (4, 25, 30). The traditional histologic types of breast cancer are associated with these gene and protein expression patterns. Tubular and lobular invasive cancer types cluster in the high estrogen receptor-expressing, luminal epithelial class of breast cancer. In contrast, medullary or medullary-like breast cancers cluster in the steroid receptor-negative, basal-like class of breast cancer. Our observations in this regard suggest that such histone marks could be of fundamental importance in defining the biological nature as well as clinical behavior in all subclasses of breast cancer.

Of clinical interest, and consistent with the above findings, is the apparent inverse relationship between the histone modification levels (assessed either individually or combinatorially) with patient outcome with respect to BCSS and DFS. The hypomodified and hypermodified clusters were able to predict the hazard of mortality independently of tumor size and lymph node stage in breast cancer (Supplementary Table S9). These findings are broadly similar to the recent study on prostate carcinoma (13) and lung cancer (14). Importantly, hypoacetylation of H3K18ac was found to be an independent prognostic marker, identifying patients with shorter BCSS and DFS when compared with other known prognostic factors. The inverse relation of H3K18ac with tumor grade was confirmed by digital image analysis and Western blots (Supplementary Fig. S2). Moreover, recent studies have shown that the adenoviral E1A protein, which can induce cell transformation, promotes global hypoacetylation of H3K18, reducing the detected levels of this modification by up to 3-fold probably by redirecting the recruitment of histone acetyltransferases p300/CBP and repressors such as Rb to specific gene promoters/enhancers (42, 43). The H3K18ac modification was detected only at genes that promote entry into cell cycle and inhibit cell differentiation and antiviral responses. Such studies highlight the importance of undertaking genome-wide studies to explore the epigenetic reprogramming that occurs in breast tumor cells and support the conclusion that global hypoacetylation of H3K18 is indicative of cell transformation and may therefore be an important prognostic indicator in breast cancer.

In conclusion, this study has identified variations in bulk histone modifications in different grades, morphologic types, and phenotype classes of invasive breast tumors. Furthermore, we have identified hypomodified and hypermodified tumor clusters, which correlate with known prognostic factors and clinical outcome. Further studies of histone modification patterns in breast tumors and the enzymes that regulate these processes will lead to a better understanding of epigenetic mechanisms in tumorigenesis and inform possible therapeutic strategies using small-molecule epigenetic modulators.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Acknowledgments

Received 10/15/08; revised 2/19/09; accepted 2/23/09; published OnlineFirst 4/14/09.

**Grant support:** Breast Cancer Campaign grant 2005/Nov08; Egyptian Government (S.E. Elsheikh, R.A. Ahmed, and M.K. Abdelghany).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Zahra Madjd and Ahmed El-Attar for the tissue microarray construction and Trevor Gray for help with the digital image analysis.

## References

- Campbell LL, Polyak K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle* 2007;6:2332–8.
- Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. *Science* 2007;318:1108–13.
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
- Santos-Rosa H, Caldas C. Chromatin modifier enzymes, the histone code and cancer. *Eur J Cancer* 2005;41:2381–402.
- Kurdistani SK. Histone modifications as markers of cancer prognosis: a cellular view. *Br J Cancer* 2007;97:1–5.
- Kouzarides T. Chromatin modifications and their function. *Cell* 2007;128:693–705.
- Jenuwein T, Allis CD. Translating the histone code. *Science* 2001;293:1074–80.
- Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* 2004;6:73–7.
- Bernstein BE, Kamal M, Lindblad-Toh K, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 2005;120:169–81.
- Pokholok DK, Harbison CT, Levine S, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 2005;122:517–27.
- Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys<sup>16</sup> and trimethylation at Lys<sup>20</sup> of

- histone H4 is a common hallmark of human cancer. *Nat Genet* 2005;37:391–400.
13. Seligson DB, Horvath S, Shi T, et al. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* 2005;435:1262–6.
  14. Barlesi F, Giaccone G, Gallegos-Ruiz MI, et al. Global histone modifications predict prognosis of resected non small-cell lung cancer. *J Clin Oncol* 2007;25:4358–64.
  15. Park YS, Jin MY, Kim YJ, Yook JH, Kim BS, Jang SJ. The global histone modification pattern correlates with cancer recurrence and overall survival in gastric adenocarcinoma. *Ann Surg Oncol* 2008;15:1968–76.
  16. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358:1148–59.
  17. Elbauomy Elsheikh S, Green AR, Lambros MB, et al. FGFR1 amplification in breast carcinomas: a chromogenic *in situ* hybridisation analysis. *Breast Cancer Res* 2007;9:R23.
  18. Elsheikh S, Green AR, Aleskandarany MA, et al. CCND1 amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome. *Breast Cancer Res Treat* 2007;9:325–35.
  19. Sevenet N, Lellouch-Tubiana A, Schofield D, et al. Spectrum of hSNF5/INI1 somatic mutations in human cancer and genotype-phenotype correlations. *Hum Mol Genet* 1999;8:2359–68.
  20. Madjd Z, Parsons T, Watson NF, Spendlove I, Ellis I, Durrant LG. High expression of Lewis y/b antigens is associated with decreased survival in lymph node negative breast carcinomas. *Breast Cancer Res* 2005;7:R780–7.
  21. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;19:403–10.
  22. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA Cancer J Clin* 2006;56:37–47.
  23. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham prognostic index in primary breast cancer. *Breast Cancer Res Treat* 1992;22:207–19.
  24. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression of luminal and basal cytokeratins in human breast carcinoma. *J Pathol* 2004;203:661–71.
  25. Abd El-Rehim DM, Ball G, Pinder SE, et al. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer* 2005;116:340–50.
  26. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br J Cancer* 2004;91:1532–42.
  27. Rakha EA, Putti TC, Abd El-Rehim DM, et al. Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *J Pathol* 2006;208:495–506.
  28. McCarty KS, Jr., Miller LS, Cox EB, Konrath J, McCarty KS Sr. Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. *Arch Pathol Lab Med* 1985;109:716–21.
  29. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000;80:1943–9.
  30. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;10:5367–74.
  31. Soerjomataram I, Louwman MW, Ribot JG, Roukema JA, Coebergh JW. An overview of prognostic factors for long-term survivors of breast cancer. *Breast Cancer Res Treat* 2008;107:309–30.
  32. Kaufman LaPj. Finding groups in data: an introduction to cluster analysis. Wiley series in probability and mathematical statistics. Applied probability and statistics. New York: Wiley; 1990.
  33. Jackson JE. A user's guide to principal components. Wiley series in probability and mathematical statistics. Applied probability and statistics. New York: Wiley; 1991.
  34. Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. Histone H4-16 acetylation controls chromatin structure and protein interactions. *Science* 2006;311:844–7.
  35. Vaquero A, Sternglanz R, Reinberg D. NAD<sup>+</sup>-dependent deacetylation of H4 lysine 16 by class III HDACs. *Oncogene* 2007;26:5505–20.
  36. Pfister S, Rea S, Taipale M, et al. The histone acetyltransferase hMOF is frequently downregulated in primary breast carcinoma and medulloblastoma and constitutes a biomarker for clinical outcome in medulloblastoma. *Int J Cancer* 2008;122:1207–13.
  37. Kim EJ, Kho JH, Kang MR, Um SJ. Active regulator of SIRT1 cooperates with SIRT1 and facilitates suppression of p53 activity. *Mol Cell* 2007;28:277–90.
  38. Zhao W, Kruse JP, Tang Y, Jung SY, Qin J, Gu W. Negative regulation of the deacetylase SIRT1 by DBC1. *Nature* 2008;451:587–90.
  39. Tryndyak VP, Kovalchuk O, Pogribny IP. Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins. *Cancer Biol Ther* 2006;5:65–70.
  40. Bagnyukova TV, Tryndyak VP, Montgomery B, et al. Genetic and epigenetic changes in rat preneoplastic liver tissue induced by 2-acetylaminofluorene. *Carcinogenesis* 2008;29:638–46.
  41. Kovalchuk O, Tryndyak VP, Montgomery B, et al. Estrogen-induced rat breast carcinogenesis is characterized by alterations in DNA methylation, histone modifications and aberrant microRNA expression. *Cell Cycle* 2007;6:2010–8.
  42. Horwitz GA, Zhang K, McBrien MA, Grunstein M, Kurdistani SK, Berk AJ. Adenovirus small e1a alters global patterns of histone modification. *Science* 2008;321:1084–5.
  43. Ferrari R, Pellegrini M, Horwitz GA, Xie W, Berk AJ, Kurdistani SK. Epigenetic reprogramming by adenovirus e1a. *Science* 2008;321:1086–8.