Chapter 9
Epigenetics of Breast Cancer: DNA Methylome and Global Histone Modifications

Gulistan Mese and Ozden Yalcin-Ozuysal

9.1 Introduction

Breast cancer is the most frequently diagnosed cancer in women. In 2012, more than 1.6 million women were diagnosed with breast cancer worldwide. Despite the improvements in screening and therapeutic approaches, in 2012 more than half a million women died due to breast cancer, which is among the leading cause of cancer deaths in women [1].

Breast cancer is a heterogeneous disease comprised of tumors with different histological characteristics and clinical outcomes in terms of prognosis, drug response and metastatic potential. Heterogeneous nature of the breast cancer demands delicate approaches to diagnose and follow the most appropriate strategy for clinical management. Classical histological analysis including assessment of hormone receptor (HR) and receptor tyrosine-protein kinase erbB-2 (ERBB2) status, tumor size, histological grade and lymph node invasion was improved in the last decade with the gene expression profiling. Microarray analysis of mRNA expression revealed mainly four molecular subtypes of breast tumors: (a) luminal A, low grade estrogen receptor (ER)(+) tumors with good prognosis, (b) luminal B, high grade (ER)(+) tumors with poor prognosis, (c) basal-like, HR (−) and ERBB2(−), and (d) ERBB2(+), increased expression of several genes of ERBB2 amplicon [2–4]. Molecular subtypes not only provided additional significant information for better diagnosis, prognostic estimates and drug response predictions but also improved our understanding of breast tumor biology (reviewed in [5]). Effective therapeutic approaches could only be developed by unrevealing the mechanisms underlying tumorigenesis and metastasis. Despite the advancements due to molecular subtypes, there is still room for improvement for better diagnostic and therapeutic approaches.

G. Mese • O. Yalcin-Ozuysal (✉)
Department of Molecular Biology and Genetics, Izmir Institute of Technology, Izmir 35430, Turkey
e-mail: gulistanmese@iyte.edu.tr; ozdenyalcin@iyte.edu.tr

M.K. Mishra, K.S. Bishnupuri (eds.), *Epigenetic Advancements in Cancer*,
DOI 10.1007/978-3-319-24951-3_9

Cancer as a general was considered to be resulted from the accumulation of genetic mutations but numerous studies now implicate the contribution of abnormal epigenetic modifications in various cancer types including breast cancer. Contribution of epigenetic modifications in breast cancer development was further supported by the identification of several recurrent mutations in genome-wide sequencing analysis in the regulators of DNA methylation, post-translational histone modifications and chromatin remodeling in general [6–9]. Similar to genetic mutations, epigenetic alterations are also involved in abnormal regulation of oncogenes and tumor suppressors that play role in breast tumor development and progression. Further studies suggested global epigenetic alterations in breast tumors compared to normal samples that can affect the activities and/or regulation of hundreds to thousands genes involved in several cellular processes including DNA repair, signaling pathways, cell cycle, and transcription that are known to be important in tumorigenesis. In contrast to genetic mutations, epigenetic alterations are reversible that could provide an advantage for the correction of abnormalities in affected regions to revert to the normal genome state. Thus, epigenetic mechanisms are prominent candidates for the development of novel therapeutic approaches. Using both genetic and epigenetic factors would give a better definition of breast cancer subtypes that would increase the diagnostic and prognostic success for breast cancer patients with the development of better markers and treatment options.

In this chapter, we summarize the recent advances in global epigenetic changes, mainly DNA methylation and histone modification alterations, and their relevance to breast cancer classification and clinical observations.

9.1.1 DNA Methylome

Abnormal DNA methylation has long been observed in several cancer types including breast. Abnormalities include both hypermethylation and hypomethylation. Hypermethylation is described as gain of methyl residue at the 5′ position of cytosine nucleotide at regions that are normally not methylated, and mostly observed in promoter CpG islands. In general hypermethylation is associated with silencing of gene expression and observed in tumor suppressor genes (Reviewed in [10]). On the other hand, global hypomethylation is another characteristic property of cancer cells. Genome wide loss of methyl residues on cytosine nucleotides is observed usually in gene poor areas such as repetitive sequences and pericentromeric satellite DNA. Hypomethylation has long been known to be associated with genomic instability (Reviewed in [11]).

Analysis of individual genes or DNA sequences revealed many cancer-associated modifications in breast tumors. Pubmeth, a web based tool for automated text mining, lists 100 genes that are implicated as hypermethylated in breast cancer [12]. MeInfoText, another automated tool that extracts gene methylation and cancer relations from the literature, shows more than 150 genes that are supposedly hyper- or hypomethylated in breast cancer [13]. Both lists contain the most commonly known

genes, such as BRCA1, RASSF1, CCND2, APC, ESR1, TWIST1, and PTEN, which are reported to be frequently hypermethylated in breast tumors (Reviewed in [14, 15]).

Studies that focused on specific genes established the basis of our understanding on the role of DNA methylation in breast cancer. However, they were limited with the few number of genes or genomic regions due to technical inadequacies. Recent advances in high throughput technologies allowed the analysis of complete epigenome in numerous tumor samples. Similar to whole genome expression profiling, whole genome DNA methylation analysis brought insights into molecular basis of breast tumors, their classification, and prediction of prognosis (Table 9.1).

One of the first studies that analyzed high number of CpGs simultaneously in breast tumors made use of methylation sensitive restriction enzymes and hybridization on immobilized CpG sequences derived from human CpG island library [16]. Using this approach, 1104 CpG islands in 28 paired breast tumor and normal samples were analyzed. The results showed that 9 % of the CpG islands were hypermethylated in tumors compared to normal samples while there were little or no changes in the rest of the CpG islands. Furthermore, hypermethylation was found to be associated with poor differentiation. In this study, other clinical parameters did not show any correlation with the methylation status of the analyzed CpGs [16]. This initial screening comparing breast tumor and normal samples was expanded with extensive analysis of CpGs in several studies (Table 9.1). The most commonly used methods was applying DNA, which is bisulfite treated or immunoprecipitated with antibodies against methylated cytosines, on arrays that covers CpGs across the genome. Analysis of 27,578 CpGs in 19 inflammatory breast tumors, 43 non-inflammatory breast tumors and 10 normal samples revealed 1353 CpGs differentially methylated between normal and tumor samples. 77 % of these CpGs were hypermethylated, while the remaining 23 % were hypomethylated in tumors. Differentially methylated genes were related to the focal adhesion, ECM receptor interaction and cytokine-cytokine receptor interaction [17]. Using the same methodology, analysis of a larger set (119 tumor and 8 normal samples) identified larger number of CpGs (6309) that were differentially methylated between tumor and normal samples. The nature of the hyper- and hypo-methylated regions were consistent with the previous findings that hypermethylation was mostly observed within CpG islands, while hypomethylation was detected outside CpG islands [18]. In another study, 108 breast tumors and 11 adjacent normal breast tissue were analyzed and 100 gene associated CpGs were found to be differentially methylated between tumor and normal samples. The list contained both previously identified genes that are hypermethylated in breast tumors such as RUNX3 and PITX2, and novel genes such as GPR10, DRD5 and CDKN1C. Increased methylation of GPR10, DRD5 and CDKN1C in tumors were validated in an independent sample set indicating that those genes could be novel candidates as tumor markers in breast [19]. In addition to genes, CpGs within 5 kb of several microRNAs were also differentially methylated between normal and tumor samples. Among those microRNAs, miR335 has previously been shown to be lost in tumors, which was associated with increased risk of metastasis [19, 20]. In summary, high throughput analyses further confirmed

Table 9.1 Whole genome DNA methylation analyses

Main conclusions	Coverage	Samples	Reference
• Methylation profile reflects T-lymphocyte infiltration in breast tumors	27,578 CpGs 14,495 genes	DS: 119 IDC, 4 NB VS: 117 IDC, 8 NB	[18]
• Luminal progenitor signature is associated with methylation pattern of the cluster that is enriched in basal-like tumors			
• Methylation profile clusters tumors into five groups	27,578 CpGs-14,495 genes 482,421 CpGs-99 % of RefSeq genes	802 BT	[7]
• The cluster with hypermethylation is enriched in luminal B tumors			
• The cluster with lowest levels of DNA methylation is enriched in basal-like tumors			
• ERBB2(+) tumors are not associated with a specific cluster			
• Luminal B or basal-like specific methylation patterns, which are characterized by hyper- or hypomethylation, respectively, are generated	482,421 CpGs 99 % of RefSeq genes	DS: 40 IBT, 17 NB VS: 121 BT(TCGA), 310 BT	[29]
• Methylation based subtypes, Epi-LumB and Epi-Basal, are associated with unfavorable clinical parameters			
• 196 CpGs associated with molecular subtypes are more methylated in luminal B subtype than in basal-like subtype	1505 CpGs 807 cancer related genes	189 BT	[28]
• Distinct methylation patterns between normal and tumor tissues are observed	1505 CpGs 807 cancer related genes	215 IDC and aNB	[27]
• Differential methylation of 5 genes can discriminate between basal-like and ERBB2 overexpressing (ERBB2(+) and luminal B subtypes) tumors			
• Methylation profile could distinguish tumor tissue from normal tissue	151 cancer specific differentially methylated regions	27 BT, 9 NB	[21]
• Methylation profile predicts BRCA1, BRCA2 and non-BRCA1/2 mutation status more accurate than gene expression profile	25,500 promoters	DS: 33 FBT VS: 47 FBT	[33]
• EPCAM(−) SRAM genes can identify claudin-low tumors, which has mesenchymal characteristics	27,578 CpGs 14,495 genes	47 BT, 19 BCCL	[35]

Main conclusions	Coverage	Samples	Reference
• Breast CpG island methylator phenotype (B-CIMP), associated with HR(+) tumors, is described	27,578 CpGs	DS: 39 BT	[23]
• HR(+) B-CIMP(−) tumors have worse prognosis than HR(+) B-CIMP(+)tumors	14,495 genes	VS: 132 BT	
• Methylation profile clusters cell lines into two groups (i: ER(−), ii: ER(+))	21,570 CpG islands	7 BCCL	[24]
• 148 differentially methylated CpGs are identified in HR(+) and HR(−) tumors	27,578 CpGs 14,495 genes	12 (HR)(+) tumors, 12 (HR)(-) tumors	[25]
• Methylation profile is associated with tumor size, race, alcohol intake and total dietary folate	1505 CpGs 807 cancer related genes	162 IBT	[26]
• Luminal A tumors, which are classified into two different methylation groups, have different overall survival	1505 CpGs	80 BT	[37]
• p53 status, ER status and tumor grade influences methylation clusters	807 cancer related genes		
• 100 CpGs can distinguish tumor tissue from normal tissue	27,325 CpGs	108 BT, 11 aNB	[19]
• 921 CpGs can classify patients into poor or good prognosis groups			
• Methylation profile classifies tumors into two groups (i: high methylation phenotype, enriched in patients with distant metastasis, ii: low methylation phenotype)	27,578 CpGs 14,495 genes	19 InBT, 43 non-InBC, 10 NB	[17]
• Methylation profile clusters each metastasis with its own primary tumor	482,421 CpGs 99 % of RefSeq genes	44 paired BT and LNM	[39]
• Metastases, except luminal A subtype, have global hypermethylation compared to primary tumors			
• Luminal A metastases mainly confer hypomethylation			
• 90 differentially methylated CpGs can discriminate between subtypes of BBM	27,578 CpGs 14,495 genes	32 BBM, 12 NB, 15 NBr, 48 BT	[38]
• Basal-like metastases have the lowest methylation levels			

aNB Adjacent Normal Appearing Breast Tissue, *BBM* Breast Brain Metastasis, *BCCL* Breast Cancer Cell Lines, *BT* Breast Tumors, *DS* Discovery Set, *ER* Estrogen Receptor, *FBT* Familial breast tumors with BRCA1, BRCA2 or non-BRCA1/2 mutations, *HR* Hormone Receptor, *IBT* Invasive Breast Tumor, *IDC* Infiltrating/Invasive Ductal Carcinoma, *InBT* Inflammatory Breast Tumor, *LNM* Lymph Node Metastasis, *NB* Normal breast tissue, *NBr* Non-neoplastic Brain Tissue, *SRAM* Significantly Repressed in Association with Methylation, *TCGA* The Cancer Genome Atlas Network, *VS* Validation Set

the previous findings that there are significant differences in DNA methylation between tumor and normal breast tissues and that differences are not limited to a few gene associated CpGs or repetitive sequences but affects thousands of CpGs across the genome. Furthermore, a list of differentially methylated CpGs can distinguish breast tumors from normal breast tissues. Interestingly, the selected CpGs did not need to be derived from breast cancer, but 151 differentially methylated regions identified in colon cancer could successfully distinguish tumors from normal tissue in breast as well [21]. All these data indicate that methylation changes are involved in a global mechanism that differentiates tumor tissue from normal tissue.

Encouraged by success of mRNA profiling, an intriguing question whether DNA methylation profiling could improve our understanding of heterogeneity of breast cancer came up with the development of high throughput assays.

Analysis of genome wide DNA methylation of numerous samples in many studies derived clusters that are mainly enriched in a specific histological characteristic such as estrogen receptor (ER) expression, or a specific molecular subtype. Methylation profile of over 200 invasive ductal carcinomas successfully clustered the tumors into two: Cluster I: enriched in ER(−) and Cluster II: enriched in ER(+) tumors [18]. To understand whether this DNA methylation pattern is functionally important for tumor phenotype, the authors looked into the genes that were represented in differentially methylated CpGs. The genes that were shown to be positively correlated with ER expression were highly methylated in Cluster I, while the genes negatively correlated with ER expression were hypermethylated in Cluster II [18, 22]. These data indicate that expression of the gene sets that can distinguish ER(−) from ER(+) tumors are at least partially regulated by DNA methylation. In an independent study, methylation profile clustered more than 160 breast tumors into two as one of them being enriched in hormone receptor (HR) (+) tumors, and the other was a mixture of HR(+) and HR(−) tumors [23]. Cell line analysis provided similar results with the tumor studies. A panel of 7 breast cancer cell lines, 4 ER(+) and 3 ER(−), were analyzed for gene expression, gene copy number and DNA methylation. Gene expression analysis were able to cluster all the cell lines as ER(+) or ER(−), while methylation status of all 21,570 CpGs did the same except that one ER(−) and one ER(+) cell line failed to cluster with their respective groups. Among 21,570 CpGs, 444 that mapped within the 5 kb of the 5′ end of the gene were differentially methylated between ER(−) and ER(+) cell lines. Top 100 of this list illustrated a very robust methylation signature for ER(+) and ER(−) cell lines. The results obtained from cell lines were complemented with tumor data showing that 67 genes that are hypermethylated in ER(−) cell lines were overexpressed in ER(+) tumors and 17 genes that were hypermethylated in ER(+) cell lines were overexpressed in ER(−) tumors [24]. Another study identified only 148 CpGs out of 27,578 to be differentially methylated between 12 ER/PR(+) and 12 ER/PR(−) tumors [25]. Only one of the four genes that was found to be differentially methylated between this group of ER/PR(+) and ER/PR(−) tumors was found to be differentially methylated between ER(+) and ER(−) cell lines [24, 25]. Discrepancies between the studies could be due to low number of samples, different techniques that are applied or differences between the nature of cell line models and primary tumors. An analysis

of 162 invasive breast tumors, which were clustered into eight groups, initially showed no association between methylation clusters and ER status. However, when samples were restricted to post-menopausal patients, methylation clusters increased to 11, which were significantly associated with ER status showing the importance of sample diversity [26]. Overall, the data indicates that ER status not only affects the gene expression pattern of tumors but also represents a specific DNA methylation profile, which at least partially contributes to the regulation of gene expression.

Molecular subtypes are based on gene expression. Since DNA methylation is one of the mechanisms that regulate gene expression, many recent studies focused on whether it could represent distinct profiles in different molecular subtypes. Furthermore, DNA methylation profiling is considered as a promising strategy to improve classification of breast tumors and predictions on clinical outcomes such as survival and metastasis. Using a cancer specific panel screening 1505 CpGs from 807 cancer related genes in more than 200 tumors, 15 CpGs were found to be differentially methylated in different molecular subtypes. Basal-like tumors were shown to have the lowest methylation levels. Interestingly, methylation status of five genes (NPY, FGF2, HS3ST2, RASSF1, Let-7a) was able to discriminate basal-like and HER2-overexpressing tumors. The marker genes were hypermethylated mostly in luminal B and ERBB2+ tumors, were not methylated mostly in basal-like tumors and were variably methylated in a group of tumors composed of mainly luminal A with a few luminal B and ERBB2+ tumors [27]. Another study, using the same panel, clustered 189 tumors into three. Each cluster was associated with either luminal A and luminal B or basal-like subtypes, while ERBB2+ and normal-like tumors were not enriched in any of the clusters. 196 CpGs associated with molecular subtypes were identified. These CpGs were more frequently methylated in luminal B tumors and less methylated in basal-like tumors consistent with the finding that the cluster dominated by basal-like tumors had lower overall methylation levels compared to luminal B dominated ones [28]. Screening 27,578 CpGs in a different set of 119 tumors, revealed 6 distinct methylation clusters. Three of these clusters were enriched in ERBB2+, basal-like or luminal A tumors, pointing to a significant difference between the methylation profiles of these three subtypes [18]. 100 loci that were initially found to distinguish tumor from normal samples were not informative for molecular subtypes. However, 600 loci that were found in the same study clustered tumors into three classes, one of them being enriched in luminal A and other in basal-like [19]. The Cancer Genome Atlas (TCGA) Network had the largest data set with 802 breast tumors that were analyzed for whole genome DNA methylation [7]. In this study, 574 differentially methylated CpGs were found to cluster the breast tumors into five groups. Consistent with the previous findings, one of the clusters had hypermethylated phenotype and enriched in Luminal B subtype, while another one had the lowest methylation and was enriched in basal-like subtype. ERBB2+ group was not associated with any of the clusters [7]. A recent study initially analyzed 482,421 CpGs in 40 tumors and showed that DNA methylation pattern clustered the tumors into three, each of which were enriched in luminal B, basal-like or luminal A tumors [29]. Then, the authors compared their data with the

one released by TCGA Network and looked for the CpGs that were commonly changed in both groups. Interestingly, 254 CpGs were determined for luminal B, 202 for basal-like, while there were no common CpG changes detected for luminal A and ERBB2+. Luminal B specific CpG pattern, which was also observed in a number of luminal A and ERBB2+ tumors, predominantly consist of methylated CpGs in promoter region. On the other hand, basal-like specific CpG pattern was exclusively observed in basal-like tumors and dominated by hypomethylated CpGs in gene bodies, CpG shores and CpG poor promoter regions [29]. Differences between luminal B and basal-like subtypes were also observed in promoter specific analysis. Promoters of ten genes (APC1, BRCA1, BRCA2, CDH1, Cyclin D2, ESR1, HIN-1, RAR-β, RASSF1A and TWIST) that are involved in breast cancer were analyzed by quantitative multiplex methylation specific PCR in 114 primary breast tumor sample group, which is composed of basal-like, luminal and HER2+ tumors. Average methylation ratio of basal-like tumors were lower than the luminal and HER2+ tumors. Only BRCA1 methylation level were higher in basal-like subtype [30]. In another study, promoters of 15 genes (APC, DLEC1, GRIN2B, GSTP1, HOXA1, HOXA10, IGF2, MT1G, RARB, RASSF1A, RUNX3, SCGB3A1, SFRP1, SFRP4, and TMEFF20) were analyzed by PCR based assay in 179 primary breast tumors. 12 genes were differentially methylated between the subtypes and had the lowest values in basal-like group [31]. In conclusion, strong data accumulated by independent research groups showed that molecular subtypes have distinct methylation patterns. The most striking difference was repeatedly found between luminal B and basal-like subtypes, of which the former had hypermethylated phenotype while the latter had low methylation profile. High throughput data were further confirmed by promoter specific analysis of selected genes. ER status, which was shown to have a dominant effect on DNA methylation profile, is an important component of molecular subtypes. Thus, it is plausible to expect DNA methylation differences between luminal and basal-like subtypes, one of the main differences of which is being HR(+) or HR(−). However, ER status alone is unlikely to explain all the methylation changes.

Familial breast tumors could be assigned into one of the molecular subtypes based on their expression profile. A study analyzing 75 familial breast tumors showed that 74 % of BRCA1 tumors were classified as basal-like, 73 % of BRCA2 tumors were luminal A or luminal B, and 52 % of non-BRCA1/2 tumors were luminal A [32]. Analysis of 1505 CpGs from 807 cancer related genes in 71 familial breast tumors showed that BRCA2 tumors had higher methylation levels compared to BRCA1 or other familial tumors. In this tumor set, 50 % of the BRCA2 tumors were classified as luminal B, and 60 % of the BRCA1 tumors were basal-like. Since it was shown that luminal B subtype has higher methylation levels than basal-like, further analysis was required to eliminate the possibility that the methylation profile was an outcome of subtype differences in the particular sample set [7, 27–29]. The complementing data came from analysis of 25,500 transcription start sites in familial breast tumors including BRCA1, BRCA2 and non-BRCA1/2 groups [33]. In contrast to somatic tumors, molecular subtypes of familial tumors did not show any significant changes in methylation profile. However, 822 genes were found to be

differentially methylated between the mutation groups and methylation profile predicted BRCA1, BRCA2 or non-BRCA1/2 status more accurate than the gene expression profile. Furthermore, methylation profile clustered non-BRCA1/2 tumors into two groups, which were heterogeneous in molecular subtype composition [33]. Overall, the data showed that familial breast tumors represent unique DNA methylation profiles that are associated with the mutation type not the molecular subtype. Thus, DNA methylation profiling could expand our knowledge to understand the clinical phenotype of the familial breast tumors.

DNA methylation profile not only improved our understanding on the heterogeneity of breast tumors, but it also provided valuable information of their biology and clinical phenotypes. DNA methylation profile was shown to be able to reveal immune cell infiltration in breast tumors [18]. Gene ontology analysis showed that immunity related genes were differentially methylated between six different methylation clusters of breast tumors. These genes were hypomethylated in immune cells but hypermethylated in normal and cancer cell lines. Tumor clusters that have hypomethylation of the immunity related genes were found to be infiltrated by lymphocytes [18]. Furthermore the same group found out that DNA methylation pattern could reflect the cellular origin of the tumor. They showed that luminal progenitor signature was associated with one of the methylation clusters. This methylation cluster was enriched in basal-like tumors, for which candidate population for transformation was identified as luminal progenitor cells [18, 34]. Another study showed that, methylation profile reflects cell lineage origins of the breast tumors [35]. The authors analyzed 27,578 CpG islands in 19 breast cancer cell lines and found out that 120 genes that were significantly repressed in association with methylation (SRAM) clustered the cell lines into two; EPCAM(+) epithelial cells and EPCAM(−) mesenchymal cells. 71 SRAM genes that were methylated in EPCAM(+) breast cancer cell lines were also repressed in normal human luminal and luminal progenitor cells, which are EPCAM(+). Similarly SRAM genes that were methylated in EPCAM(−) mesenchymal cell lines were repressed in normal human basal/myoepithelial and stromal cells, which are EPCAM(−). Thus, methylation pattern of breast cancer cell lines matched expression patterns of specific lineages of normal breast cells. Furthermore, EPCAM(−) SRAM were able to identify claudin-low tumors, which carry mesenchymal characteristics, while EPCAM(+) SRAM were frequently methylated in primary tumors reflecting their epithelial origin [35]. Stem cell component of the breast tumors were also linked to methylation profile [31]. Stem cell rich tumors, determined by presence of CD44(+)/CD24(−) or ALDH1(+) cell content, had lower methylation levels. However, CD44(+)/CD24(−) tumors were enriched in basal-like subtype, which was shown to have lower methylation levels [7, 27–29]. To make sure that low methylation levels were a result of stem cell component not molecular subtype, the authors analyzed stem cell rich or poor basal-like tumors. They showed that basal-like tumors enriched in stem cells had lower methylation levels compared to basal-like tumors that do not have CD44(+)/CD24(−) cells indicating that stem cell phenotype contributes to DNA methylation profile independently of molecular subtype [31]. Breast cancer risk related exposures were also found to be associated with methylation profiles. Analysis of 1413

autosomal CpGs clustered 162 invasive breast tumors into 8 classes, which were significantly associated with alcohol intake and total dietary folate indicating the effect of environmental factors on DNA methylation pattern of breast tumors [26].

One of the practical benefits of DNA methylation profiling is its contribution to prediction of prognosis. Screening of 27,578 CpGs clustered breast tumors into two, one of which was defined as having breast CpG methylator phenotype (B-CIMP) [23]. B-CIMP was composed of 3297 differentially methylated CpGs of 2543 genes, more than two third of which was hypermethylated. B-CIMP(+) tumors had improved metastasis-free survival. However, B-CIMP(+) cluster was composed of HR(+) tumors, while the other cluster was a mix of HR(+) and HR(−) tumors. To validate that the association of B-CIMP(+) phenotype with better metastasis free survival was independent of HR status, the authors limited the analysis with HR(+) tumors. They showed that B-CIMP(−) HR(+) tumors had worse prognosis than B-CIMP(+) HR(+) tumors. B-CIMP status was a strong predictor of prognosis independent of tumor stage, age, nodal status and HR status. B-CIMP genes that has decreased mRNA expression, were enriched in genes listed in breast cancer prognostic expression signatures [23]. Thus, B-CIMP phenotype could partly explain the differential gene expression in poor vs. good prognosis tumors and could be used as a prognostic marker. Another study showed that DNA methylation profile was associated with the molecular signature that is an indicator of poor prognosis [17]. 500 differentially methylated CpGs clustered a tumor set that consists of 19 inflammatory and 43 non-inflammatory breast tumors into two. The cluster that has high methylation levels was enriched in tumors from patients with distant metastasis and poor prognosis as determined by 70 gene prognostic signature [17, 36]. There were no association between the methylation clusters and age, tumor stage, histological grade or HR and ERBB2 status of the tumors, but one of the clusters were enriched in inflammatory breast tumors. However, it is unlikely that the distribution of the inflammatory breast tumors would affect prognosis because only four CpGs corresponding to four genes were differentially methylated between inflammatory and non-inflammatory tumors [17]. It was also shown that methylation status of individual genes could cluster patients into poor or good prognosis groups [19]. Analysis of more than 100 tumors showed that 921 CpGs, including 490 genic and 431 non-genic loci, could derive two groups with different prognosis. 25 of the genes that were associated with these CpGs were able to cluster patients into poor and good prognosis groups individually. Interestingly, all of these genes were in a different methylation state in normal tissue compared to tumor tissue. Furthermore, these genes belong to gene ontology groups of vasculature development, cell death, proliferation, and cell cycle processes indicating that they could be functionally important in metastatic cascade [19]. DNA methylation profile could also reveal heterogeneity within the molecular subtypes. Analysis of 1505 CpGs clustered 80 breast tumors into three, none of which were strongly dominated by a specific group [37]. Interestingly, it was found that luminal A tumors, which were distributed between two different clusters, differ significantly in survival [37]. However, it should be noted that TP53 status and size of the tumors were also found to affect survival in different methylation clusters and status of luminal A tumors in these

parameters might be a factor in differential survival as well [37]. A recent study identified two new subtypes based on DNA methylation profile: Epi-LumB and Epi-Basal, which are luminal B and basal-like associated CpGs, respectively [29]. Both signatures were associated with reduced survival and observed in most but not all of the relevant molecular subtypes, reflecting the heterogeneity between them. Three genes for Epi-LumB and two genes for Epi-Basal phenotype were selected as proxy to analyze by locus specific assays. Analysis of 301 tumors showed that survival of patients with luminal B or basal-like tumors were not changed by having Epi-LumB or Epi-Basal phenotype. However, within ER(+) tumors, having luminal B or Epi-LumB phenotype significantly increased the death risk by approximately fivefold compared to having only luminal B phenotype. Similarly, within ER(−) tumors, having basal-like or Epi-Basal phenotype increased the death risk by tenfold compared to basal-like only phenotype although it did not reach to statistical significance [29]. In summary, DNA methylation profiling was shown to provide additional information to estimate prognosis and stands out as a prominent approach for clinical use.

Metastasis is another important mechanism that is tightly associated with the survival of the patient. Thus, its molecular characterization has long been a focus of interest. To understand the landscape of breast brain metastasis, DNA methylation profile was analyzed in 32 metastases, 12 non-neoplastic breast tissue and 15 non-neoplastic brain tissue [38]. 425 CpGs were found to be differentially methylated in metastases, majority being hypermethylated. Similar to primary tumors, subtypes of metastases were discriminated by 90 differentially methylated CpGs. Basal-like metastases were found to have the lowest methylation levels consistent with the primary tumor data [38]. The differences between metastases and matched primary tumors were not that striking. A recent study analyzed methylation pattern of more than 400,000 CpGs in 44 paired primary tumors and lymph node metastases [39]. Metastases did not form an independent group but each metastasis was grouped with the matching primary tumor instead, indicating that metastasis specific changes do not dominate the methylation profile. Metastases have global DNA hypermethylation compared to primary tumors mostly outside the core promoter regions, with the most significant increase being observed in basal-like subtype. Only luminal A metastases mainly confer hypomethylation leading to the question whether different subtypes acquire metastatic capacity via different mechanisms. Only 155 genes were differentially methylated between primary tumor and its matched metastasis in at least one of the molecular subtypes. Among these 155 genes, only 8 were differentially expressed indicating that majority of the methylation changes that are associated with gene expression must have occurred early in tumorigenesis [39].

Overall, high throughput DNA methylation analysis improved our understanding in breast tumor heterogeneity and clinical outcomes. However, how DNA methylation contributes to these processes is still an open question. Many of the studies combined DNA methylation profiling with mRNA profiling. Although there was an association, DNA methylation did not always correlate with gene expression. B-CIMP phenotype that is associated with better prognosis was composed of 1764

hypermethylated genes, only 102 of which were downregulated at the mRNA level [23]. 444 CpGs were found to be differentially methylated between ER(−) and ER(+) cell lines, but only 151 of them were inversely correlated with the expression of 149 genes [24]. Among 79 genes that were differentially methylated between three clusters of tumors, only 33 showed a significant anti-correlation between methylation and gene expression [19]. The tumor group with hypermethylation profile has 4283 genes that are differentially methylated, while only 1899 of them were differentially expressed [7]. All these data point to alternative mechanisms that regulate gene expression in coordination with or independent of DNA methylation to reach to a certain context in the cell. Loss of copy number could be one of the candidate mechanisms. In poorly metastatic MDA MB 468-GFP and highly metastatic MDA MB 468-LN cell lines, copy number alterations, gene expression and DNA methylation profiles were analyzed. It was shown that loss in copy number was correlated with hypermethylation, while copy number increase was associated with hypomethylation [40]. Gene expression and DNA methylation analysis in ERBB2(+) cell line HCC1954 showed that hypomethylated genes had greater tendency to be repressed [41]. In 50 ERBB2(+) breast tumors, the genes hypomethylated in HCC1954 were enriched in repressed genes compared to 23 normal breast samples. The repression mechanism of hypomethylated genes were revealed by the analysis of histone modifications. Regions showing hypomethylated DNA were enriched in H3K9me3 and H3K27me3. Allele specific analysis showed that in hypomethylated genes, one allele was methylated and the unmethylated allele was occupied by H3K9me3 and H3K27me3, resulting in repression of gene expression [41]. Thus, DNA hypomethylation was tightly linked to formation of closed chromatin structures and gene silencing.

In conclusion, DNA methylation profiling proved to be useful as a novel approach to understand the heterogeneity of breast cancer and classify breast tumors into better defined subtypes with predictable clinical characteristics. Since DNA could be well preserved in different kinds of clinical specimen and represent a more stable profile than mRNA, methylation pattern as a prognostic or diagnostic marker would benefit clinic.

9.1.2 Global Histone Modifications

Eukaryotes maintain their DNA content in a highly compact and ordered structure by the help of small, basic histone proteins. This condensed structure is formed by wrapping 147 bp of DNA around the octamer of four core histone proteins, composed of H2A, H2B, H3 and H4 dimers [6, 42–44]. As a result of this configuration, two types of chromatin regions are generated in the genome; transcriptionally inactive heterochromatin and transcriptionally active euchromatin regions. This chromatin structure is strictly regulated by post-translational covalent modifications of histone tails and at least 16 different types of histone modifications including the addition/removal of methyl, acetyl, phospho, ubiquitin, SUMO and

poly-ADP-ribose groups, have been identified that regulate DNA replication, transcription and repair mechanisms [45–50]. The specific function of each modification and their modifiers in DNA-based mechanisms have been extensively studied to determine their relevance to cellular functions and their contribution to the generation and progression of several human disorders including different types of cancer [6].

Identification of various types of histone modifications suggested the importance of combinations of histone marks in the genome that are referred to as the "histone code" for the regulation of gene function [50–52]. This also implicated the crosstalk among different histone modifier enzymes and their effector proteins. Examination of individual histone marks enabled the determination of the function of each mark and recent efforts now focus on the interplay between different modifications and their involvement in cellular functions [47, 50].

Among all modifications, acetylation and methylation (mono, double and triple) of basic lysine and arginine residues are the most widely investigated histone marks. The levels of acetylation and methylation are governed by the activities of histone acetyltransferases (HATs) and deacetylases (HDACs and Sirtuins); and by histone methyltransferases (HDMTs) and demethylases (HDMs), respectively [47, 53, 54]. Studies on the identification of functions of single histone modifications on gene activation and repression have shown that differential acetylation (ac), mono- (me), di- (me2) and tri- (me3) methylation of histones produce diverse functional outcomes in cells. For example, while H3K4me3, H3K36me3, H3K79me3, H3K9ac, H3K14ac, H3K20me and H2BK4me modifications were demonstrated to be involved in gene activation, H2BK5me3, H3K9me3, H3K27me3 and H4K20me3 were associated with gene repression [50, 52, 55–57]. Changes in the levels of histone modifications can affect their function in gene transcription and therefore can alter cell homeostasis.

Histone modifications are highly dynamic processes that can change in a cell-context dependent manner by altering the activities of modifier proteins that consist of modification inducers, erasers and readers that interact and bind to histone complexes. Changes in histone mark status will eventually alter histone-DNA, histone-histone, histone-non-histone protein interactions that control many DNA-template processes important for regulation of cellular events [49, 58, 59]. This type of crosstalk among DNA, histones, modifiers and effector proteins provide an enormous amount of complexity to the regulation of the gene function [6, 48]. That is why regulation and maintenance of histone modifications are crucial for normal cellular functions and abnormalities in histone modifications, alteration in modifier activities and alteration in the interaction with reader proteins might contribute to the generation of various human disorders including cancer [10, 15, 58, 60].

Cancer is a multistep process that involves both genetic and epigenetic changes of whom cooperation play essential part in the development and progression of cancer [6, 60, 61]. The contribution of epigenetic changes to cancer generation and progression was initially demonstrated by the observation of alterations in methylation status of CpG islands in promoters of cancer related genes [62–64]. This was followed by the identification of the involvement of histone modifiers that were

shown to be mutated to silence and/or activate cancer related-genes during cancer formation [6]. Since then, roles of modifiers and their specific histone modifications in various cancer types have been extensively studied. Although, initial studies examined functions of individual modifier and/or modification in cancer formation and progression, recent developments in global proteomic and genomic technologies enabled the study of global assessment of histone modifications and modifier enzymes in a genome-wide scale. Global changes in histone modifications in different cancer types including breast cancer have been focused on the potential usage of these modifications for correlation with prognostic and histopathological outcomes [65, 66].

Epigenetic factors are thought to be one of the contributors of the breast cancer heterogeneity. For this, effects of epigenetic modifications to the development of breast cancer and their association with breast cancer prognosis, recurrence risk and subtype classification have been extensively studied in recent years [66–68]. Similar to other cancers, breast cancer tissues and cells have been observed to have altered epigenetic modifications compared to normal tissues and cells. Although initial studies have focused on the analysis of modifications on a gene-by-gene basis, recent trends switched to investigation of global epigenetic changes and their relevance to breast cancer-specific subtypes, prognosis and patient outcomes [66]. In the following parts, the global changes in histone acetylation and histone methylations and their association with breast cancer will be mainly explained with a brief summary of involvement of other histone marks in breast cancer development and progression.

Analysis of normal tissues, cancer cell lines and primary tumors have shown that global loss of monoacetylation (H4K16ac) and trimethylation (H4K20me3) of histone 4 have been commonly observed in initial steps of different cancers and considered to be one of the hallmarks of cancer generation [69]. This suggested the involvement and usage of global histone modification changes for the prediction of risk assessment, prognosis, recurrence and overall survival in breast cancer [66, 70, 71]. The global changes in H3K4me2, H3K9ac, H3K18ac, H4R3me2, H4K12ac, H4K16ac and H4K20me3 histone modifications were compared using immunohistochemistry staining on tissue microarrays of 880 normal and primary invasive breast carcinoma cases [66]. Here, authors observed a correlation between histone modification status, breast cancer subtypes and clinical outcome in analyzed tissues. Although a heterogeneous staining of different histone marks were detected in tumor samples compared to normal tissue, majority of tumor cases (78.9 %) was shown to have a very low or undetectable acetylation of K16 on H4 (H4K16ac) levels. As a general observation, high levels of histone acetylation and methylation were associated with a better prognosis, ER(+) and PR(+) tumors and luminal-like breast tumors while low-to-moderate levels of all histone marks were generally implicated in poorer prognostic tumors such as basal-like and HER2-positive subtypes. Moreover, low levels of H3K18ac, H3K9ac, H3K4me2, H4R3me2, and H4K16ac were associated with unfavorable patient outcome while high levels were linked to more favorable breast cancer specific survival and metastatic specific survivals. Finally, high levels of H3K9ac, H3K18ac and H4R3me2 were associated

with longer disease free survival rate [66]. In summary, global hypermodified or hypomodified histones were differentially correlated with breast cancer subtypes, prognosis and patient outcomes.

Genome-wide changes in acetylation of K12 of H4 (H4K12ac) were also correlated with differential patient outcomes, where high levels of H4K12ac were implicated in better breast cancer specific and metastatic specific survivals and low levels were associated with unfavorable patient outcome [66]. In a similar way, immunostaining of global histone H4 acetylation (H4ac) and H4K12ac modification in 58 breast tissues containing concurrent non-cancerous breast epithelium, ductal carcinoma in situ (DCIS), and invasive ductal carcinoma (IDC) in the same block and 22 normal breast samples from reduction mammoplasties indicated that normal epithelium had a higher level of H4ac and H4K12ac expression compared to tumor samples. In tumors, on the other hand, a progressive hypoacetylation of H4ac and H4K12ac was observed as the tumor progresses from normal to DCIS and to IDC stages [72]. This suggested involvement of global histone H4 hypoacetylation, specifically H4K12ac, in the progression of cancer to more invasive stages.

Repressive histone modifications were among the widely studied histone marks that have been shown to be associated with different breast cancer subtypes [67, 73, 74]. The expression of trimethylation of histone 3 on K27 (H3K27me3) was observed to be higher in normal breast tissues (88 %) compared to breast tumor samples as shown by immunohistochemical staining of 142 primary breast tumor and 43 normal breast tissues [67]. Authors suggested a correlation between the reduction of H3K27me3 mark and ER negative tumors and also large tumor sizes which was supported by Holm et al. (2012) who observed a correlation between high expression of H3K27me3 and small tumor sizes [67, 75]. Moreover, low H3K27me3 levels were shown to have adverse prognostic values as low H3K27me3 levels were significantly correlated with poor prognosis in breast cancer patients with shorter overall survival time compared to patients with high H3K27me3. H3K27me3 was regulated by histone methyltransferase EZH2 (enhancer of zeste homolog 2) that was observed to be upregulated in breast cancers and this overexpression was shown to be mostly associated with aggressive breast cancer phenotypes [76–79]. Holm et al compared the global expression of H3K27me3 and EZH2 in more than 400 tumor samples and also in breast cancer cell lines to relate their expression pattern with different breast cancer subtypes [75]. There was an inverse correlation between H3K27me3 and EZH2 expression in different tumor types and tumor grades as high expression of EZH2 was associated with high grades and ER(−)/PR(−) tumors and these tumors were observed to have low H3K27me3 levels [75]. Consistent with this study, examination of H3K27me3 on tissue micro arrays in the Nurses' Health Study demonstrated a significant association of H3K27me3 mark with lower tumor grade and a positive association with ER(+) and PR(+) tumors [68]. Finally, H3K27me3 was also observed to be positively associated with luminal subtype A breast cancers [68, 75].

The levels of other repressor marks, H3K9me3 and H4K20me3, were also evaluated in a small group of primary tissues (15 breast cancer patients and 28 healthy individual) and both of them were observed to be increased in breast cancer patients

compared to the control group [46]. Assessment of global changes of H3K9me3 modification in a larger cohort suggested no correlation between H3K9me3 levels and clinical data as tumor samples had a diverse staining of H3K9me3 while non-cancerous cells of epithelium and myoepithelium had positive H3K9me3 staining [80]. Similarly, Healey et al. (2014) also observed no association between H3K9me3 and clinical outcomes even though there was a significant overlap between H3K9me3 and H3K27me3 positive tumors [68]. Studies in cell lines further supported this, where the global levels of H3K9Me3 were not observed to be significantly changed among non-cancerous (H16N2), atypical ductal hyperplasia (21PT), ductal in situ carcinoma (21NT) and metastatic carcinoma (21MT1) cell lines [81].

Loss of H4K20me3 is considered to be a common hallmark for different cancers [69, 82]. The studies in breast cancer tissues also suggested a global loss of H4K20me3 modification and its association with clinical outcomes [66, 80]. Low levels of H4K20me3 in breast tumor samples were correlated with poor prognostic features and higher tumor grades [66]. In the same way, Yokoyama et al. (2014) observed a high expression of H4K20me3 in noncancerous regions adjacent to tumor sites and benign cases and a low H4K20me3 levels in 63.9 % of cancer tissues in a study comprised of benign and tumor samples from 112 breast cancer patients. The heterogeneous H4K20me3 staining in cancer tissues was also correlated with different histopathological characteristics. For example, the loss of H4K20me3 mark was shown to be associated with poor prognosis and shorter disease-free survival while patients with high H4K20me3 expression had higher overall and disease-free survival rates [80]. Furthermore, a positive association between H4K20me3 levels and luminal subtypes was suggested in this study as the expression of H4K20me3 was correlated with ER and PR expression but not with HER2 expression. As a result, the loss of H4K20me3 might be a candidate for detection of poor prognostic cases in breast cancer patients. The loss of H4K20me3 in breast cancer was further verified by detection of a decrease in the expression of H4K20me3-specific methyltransferases, SUV420H1 and SUV420H2, as observed in Methylation and Expression of Normal and Tumor tissues, MENT, database as well as in invasive breast cancer cell lines, MDA MB 231 and BT474, compared to other cell lines [80]. Ectopic expression of SUV420H1 and SUV420H2 in these cell lines resulted in elevated H4K20me3 levels and limited invasive activities of cells while knockdown of SUV420H2 in MCF10A, immortalized non-tumorigenic mammary epithelial cell line, increased invasion potential of these cells, suggesting a possible role for SUV420H1 and SUV420H2 and their histone modifications, H4K20me3, in cancer invasion.

In addition to acetylation and methylation, histones are phosphorylated on various sites and histone phosphorylation has been shown to play role in mitosis, DNA repair and transcriptional regulation [83–86]. Phosphorylation of histone H3 (Phosphohistone H3, PPH3) has been used as a marker for proliferating cells and was suggested to have a prognostic value for rapidly proliferating tumor cases such as for early breast cancer cases [87, 88]. Skaland et al. (2007, 2009), observed a significant correlation between PPH3 levels and tumor size, estrogen receptor, histological grade, and mitotic activity index in lymph node-negative invasive breast

cancers in patients less than 55 years old and patients less than 71 years old, respectively [88, 89]. Further, PPH3 was suggested to have a strong prognostic value in all patients especially in ER-positive and histological grade 1 and 2 patients in lymph node-negative breast cancers [89, 90]. Another study investigated the global phosphorylation of linker histone H1 in breast cancer samples, where a correlation between the presence of pT146 of histone H1 staining and tumor grades and subtypes was observed by the immunohistochemical analysis of 242 primary breast tumors and 97 nonbreast cancer tissue [91]. An increase in pT146 staining of histone H1 was also observed in the nuclei of tumor cells as the tumor progressed from grade I to grade III. Further, an association between pT146 staining and tumor subtypes was suggested as higher pT146 intensity was correlated with triple negative breast tumors and lower staining was strongly linked to the luminal A subtype. A similar trend was also observed in cell lines where pT146 level was higher in MDA MB 231, a metastatic breast cancer cell line compared to the MCF-10A, immortalized non-tumorigenic mammary epithelial cell line and MCF7, non-invasive breast cancer cell line. This study suggested that pT146 staining of H1 might be used to differentiate tumor grades in breast cancer [91].

Aside from global changes in histone acetylation, methylation and phosphorylation levels, genome-wide alterations in other histone modifications such as ubiquitination and deimination/citrullination have been also investigated in breast cancer [92–94]. For example, a reversible histone ubiquitination process, governed by the activities of ubiquitin ligases, E1, E2, and E3 enzymes that adds ubiquitin and deubiquitinating enzymes (DUBs) that removes the ubiquitin moiety, regulates the activities of histones H2A and H2B [95, 96]. Ubiquitination of histones have been implicated in the transcriptional regulation and DNA repair mechanisms and their function in disease development including cancer have been recently started to be investigated [96, 97]. Monoubiquitination of histone H2B (H2Bub) levels were analyzed in 109 samples containing normal breast epithelial tissue samples, benign, malignant, and metastatic samples and it was shown that its level did not change in benign breast tumors compared to normal breast epithelium as observed in other cancers [93]. On the other hand, global loss of H2Bub expression was observed in malignant and metastatic samples, suggesting its role in breast cancer progression and metastasis [93, 98]. Further studies in a large group of breast cancer tissues can help to establish the association of H2Bub and other histone ubiquitination marks with different breast subtypes and clinical outcomes.

Citrullination is another histone modification that involves the conversion of positively charged arginine and methylated arginine residues to neutral citrulline by peptidylarginine deiminase (PAD) family of enzymes in a process called deamination or citrullination [94, 99]. Methylation of arginine at different sites of histones H3/H4 can have either repressive or activator effect on gene transcription and therefore, citrullination of these residues could alter chromatin structure and lead to transcriptional activation or repression, respectively [100–102]. The contribution of PADs and citrullinated histones (HCit) to the development of different tumors including breast cancer have been recently started to be investigated. For example, PAD4 was initially shown to be extensively expressed in various tumor types com-

pared to the normal or benign tissues that basically lacked the PAD4 expression [103, 104]. Guertin et al. (2014) further investigated the association of ER, PAD2 and H3R26Cit in tumor sections of 21 breast cancer patients [94]. They observed a correlation between the degree of ER staining, PAD2 and H3R26Cit staining. Additionally, a significant link between PAD2 expression and relapse free survival time of patients was detected in luminal A subtypes. These results might suggest the importance of PADs and citrullination in tumor development/progression but further studies will elicit their prognostic and histopathological values.

Genome also contains other rare modifications that include O-GlcNacetylation, sumoylation, ADP-ribosylation, proline isomerization, crotonylation, propionylation, butyrylation, formylation and biotinylation [6]. Involvement of some of these modifications (O-GlcNacetylation, sumoylation) in breast cancer development and progression have started to be documented recently [105, 106]. Development of new tools to detect these minor histone marks in the genome will enable the investigation of their global changes in different cancers including breast cancer and their relevance to prognosis and clinical outcomes.

In conclusion, breast cancer tissues contained altered global histone modification profiles compared to normal breast tissue, implicating the importance of chromatin regulation for tumorigenesis. Elucidation of genes affected by these changes would provide clearer picture about how histone modifications contribute to cellular functions as well as carcinogenesis.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide: IARC CancerBase No. 11. [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. http://globocan.iarc.fr.
2. Perou CM, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747–52.
3. Sorlie T, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;98(19):10869–74.
4. Sorlie T, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A. 2003;100(14):8418–23.
5. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat Rev Cancer. 2007;7(7):545–53.
6. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012;150(1):12–27.
7. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61–70.
8. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell. 2012;22(1):9–20.
9. Stephens PJ, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486(7403):400–4.
10. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128(4):683–92.

11. Veeck J, Esteller M. Breast cancer epigenetics: from DNA methylation to microRNAs. J Mammary Gland Biol Neoplasia. 2010;15(1):5–17.
12. Ongenaert M, et al. PubMeth: a cancer methylation database combining text-mining and expert annotation. Nucleic Acids Res. 2008;36(Database issue):D842–6.
13. Fang YC, et al. MeInfoText 2.0: gene methylation and cancer relation extraction from biomedical literature. BMC Bioinform. 2011;12:471.
14. Jones A, et al. Emerging promise of epigenetics and DNA methylation for the diagnosis and management of women's cancers. Epigenomics. 2010;2(1):9–38.
15. Barrow TM, Michels KB. Epigenetic epidemiology of cancer. Biochem Biophys Res Commun. 2014;455(1–2):70–83.
16. Yan PS, et al. CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. Clin Cancer Res. 2000;6(4):1432–8.
17. Van der Auwera I, et al. Array-based DNA methylation profiling for breast cancer subtype discrimination. PLoS One. 2010;5(9), e12616.
18. Dedeurwaerder S, et al. DNA methylation profiling reveals a predominant immune component in breast cancers. EMBO Mol Med. 2011;3(12):726–41.
19. Kamalakaran S, et al. DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. Mol Oncol. 2011;5(1):77–92.
20. Tavazoie SF, et al. Endogenous human microRNAs that suppress breast cancer metastasis. Nature. 2008;451(7175):147–52.
21. Hansen KD, et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet. 2011;43(8):768–75.
22. Desmedt C, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res. 2008;14(16):5158–65.
23. Fang F, et al. Breast cancer methylomes establish an epigenomic foundation for metastasis. Sci Transl Med. 2011;3(75):75ra25.
24. Sun Z, et al. Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing. PLoS One. 2011;6(2), e17490.
25. Li L, et al. Estrogen and progesterone receptor status affect genome-wide DNA methylation profile in breast cancer. Hum Mol Genet. 2010;19(21):4273–7.
26. Christensen BC, et al. Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. PLoS Genet. 2010;6(7), e1001043.
27. Bediaga NG, et al. DNA methylation epigenotypes in breast cancer molecular subtypes. Breast Cancer Res. 2010;12(5):R77.
28. Holm K, et al. Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. Breast Cancer Res. 2010;12(3):R36.
29. Stefansson OA, et al. A DNA methylation-based definition of biologically distinct breast cancer subtypes. Mol Oncol. 2015;9(3):555–68.
30. Lee JS, et al. Basal-like breast cancer displays distinct patterns of promoter methylation. Cancer Biol Ther. 2010;9(12):1017–24.
31. Park SY, et al. Distinct patterns of promoter CpG island methylation of breast cancer subtypes are associated with stem cell phenotypes. Mod Pathol. 2012;25(2):185–96.
32. Waddell N, et al. Subtypes of familial breast tumours revealed by expression and copy number profiling. Breast Cancer Res Treat. 2010;123(3):661–77.
33. Flanagan JM, et al. DNA methylome of familial breast cancer identifies distinct profiles defined by mutation status. Am J Hum Genet. 2010;86(3):420–33.
34. Lim E, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat Med. 2009;15(8):907–13.
35. Sproul D, et al. Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. Proc Natl Acad Sci U S A. 2011;108(11):4364–9.

36. Van't Veer LJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530–6.
37. Ronneberg JA, et al. Methylation profiling with a panel of cancer related genes: association with estrogen receptor, TP53 mutation status and expression subtypes in sporadic breast cancer. Mol Oncol. 2011;5(1):61–76.
38. Salhia B, et al. Integrated genomic and epigenomic analysis of breast cancer brain metastasis. PLoS One. 2014;9(1), e85448.
39. Reyngold M, et al. Remodeling of the methylation landscape in breast cancer metastasis. PLoS One. 2014;9(8), e103896.
40. Andrews J, et al. Multi-platform whole-genome microarray analyses refine the epigenetic signature of breast cancer metastasis with gene expression and copy number. PLoS One. 2010;5(1), e8665.
41. Hon GC, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. 2012;22(2):246–58.
42. Luger K, et al. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature. 1997;389(6648):251–60.
43. Richmond TJ, Davey CA. The structure of DNA in the nucleosome core. Nature. 2003; 423(6936):145–50.
44. Tsompana M, Buck MJ. Chromatin accessibility: a window into the genome. Epigenetics Chromatin. 2014;7(1):33.
45. Kingston RE, Bunker CA, Imbalzano AN. Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev. 1996;10(8):905–20.
46. Leszinski G, et al. Relevance of histone marks H3K9me3 and H4K20me3 in cancer. Anticancer Res. 2012;32(5):2199–205.
47. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693–705.
48. Tessarz P, Kouzarides T. Histone core modifications regulating nucleosome structure and dynamics. Nat Rev Mol Cell Biol. 2014;15(11):703–8.
49. Tan M, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 2011;146(6):1016–28.
50. Turner BM. Histone acetylation and an epigenetic code. Bioessays. 2000;22(9):836–45.
51. Munshi A, et al. Histone modifications dictate specific biological readouts. J Genet Genomics. 2009;36(2):75–88.
52. Strahl BD, Allis CD. The language of covalent histone modifications. Nature. 2000; 403(6765):41–5.
53. Barth TK, Imhof A. Fast signals and slow marks: the dynamics of histone modifications. Trends Biochem Sci. 2010;35(11):618–26.
54. Khorasanizadeh S. The nucleosome: from genomic organization to genomic regulation. Cell. 2004;116(2):259–72.
55. Rodriguez-Paredes M, Esteller M. A combined epigenetic therapy equals the efficacy of conventional chemotherapy in refractory advanced non-small cell lung cancer. Cancer Discov. 2011;1(7):557–9.
56. Schotta G, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 2004;18(11):1251–62.
57. Rice JC, et al. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell. 2003;12(6):1591–8.
58. Esteller M. Epigenetics in cancer. N Engl J Med. 2008;358(11):1148–59.
59. Basse C, Arock M. The increasing roles of epigenetics in breast cancer: Implications for pathogenicity, biomarkers, prevention and treatment. Int J Cancer. 2014. doi:10.1002/ijc.29347.
60. Berdasco M, Esteller M. Aberrant epigenetic landscape in cancer: how cellular identity goes awry. Dev Cell. 2010;19(5):698–711.
61. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5): 646–74.

62. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002;3(6):415–28.
63. Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet. 1999;21(2):163–7.
64. Feinberg AP, Tycko B. The history of cancer epigenetics. Nat Rev Cancer. 2004;4(2): 143–53.
65. Valk-Lingbeek ME, Bruggeman SW, van Lohuizen M. Stem cells and cancer; the polycomb connection. Cell. 2004;118(4):409–18.
66. Elsheikh SE, et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. Cancer Res. 2009;69(9):3802–9.
67. Wei Y, et al. Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. Mol Carcinog. 2008;47(9):701–6.
68. Healey MA, et al. Association of H3K9me3 and H3K27me3 repressive histone marks with breast cancer subtypes in the Nurses' Health Study. Breast Cancer Res Treat. 2014;147(3): 639–51.
69. Fraga MF, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet. 2005;37(4):391–400.
70. Seligson DB, et al. Global histone modification patterns predict risk of prostate cancer recurrence. Nature. 2005;435(7046):1262–6.
71. Park YS, et al. The global histone modification pattern correlates with cancer recurrence and overall survival in gastric adenocarcinoma. Ann Surg Oncol. 2008;15(7):1968–76.
72. Suzuki J, et al. Protein acetylation and histone deacetylase expression associated with malignant breast cancer progression. Clin Cancer Res. 2009;15(9):3163–71.
73. Cao R, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science. 2002;298(5595):1039–43.
74. Muller J, et al. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell. 2002;111(2):197–208.
75. Holm K, et al. Global H3K27 trimethylation and EZH2 abundance in breast tumor subtypes. Mol Oncol. 2012;6(5):494–506.
76. Leroy G, et al. A quantitative atlas of histone modification signatures from human cancer cells. Epigenetics Chromatin. 2013;6(1):20.
77. Bachmann IM, et al. EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. J Clin Oncol. 2006;24(2):268–73.
78. Collett K, et al. Expression of enhancer of zeste homologue 2 is significantly associated with increased tumor cell proliferation and is a marker of aggressive breast cancer. Clin Cancer Res. 2006;12(4):1168–74.
79. Pietersen AM, et al. EZH2 and BMI1 inversely correlate with prognosis and TP53 mutation in breast cancer. Breast Cancer Res. 2008;10(6):R109.
80. Yokoyama Y, et al. Loss of histone H4K20 trimethylation predicts poor prognosis in breast cancer and is associated with invasive activity. Breast Cancer Res. 2014;16(3):R66.
81. Santos Jr GC, et al. Epigenetic modifications, chromatin distribution and TP53 transcription in a model of breast cancer progression. J Cell Biochem. 2015;116(4):533–41.
82. Seligson DB, et al. Global levels of histone modifications predict prognosis in different cancers. Am J Pathol. 2009;174(5):1619–28.
83. Aihara H, et al. Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early Drosophila embryo. Genes Dev. 2004;18(8):877–88.
84. Harvey AC, Jackson SP, Downs JA. Saccharomyces cerevisiae histone H2A Ser122 facilitates DNA repair. Genetics. 2005;170(2):543–53.
85. Shroff R, et al. Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. Curr Biol. 2004;14(19):1703–11.
86. Maile T, et al. TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. Science. 2004;304(5673):1010–4.

87. Skaland I, et al. The prognostic value of the proliferation marker phosphohistone H3 (PPH3) in luminal, basal-like and triple negative phenotype invasive lymph node-negative breast cancer. Cell Oncol. 2009;31(4):261–71.
88. Skaland I, et al. Phosphohistone H3 expression has much stronger prognostic value than classical prognosticators in invasive lymph node-negative breast cancer patients less than 55 years of age. Mod Pathol. 2007;20(12):1307–15.
89. Skaland I, et al. Validating the prognostic value of proliferation measured by Phosphohistone H3 (PPH3) in invasive lymph node-negative breast cancer patients less than 71 years of age. Breast Cancer Res Treat. 2009;114(1):39–45.
90. Klintman M, et al. The prognostic value of mitotic activity index (MAI), phosphohistone H3 (PPH3), cyclin B1, cyclin A, and Ki67, alone and in combinations, in node-negative premenopausal breast cancer. PLoS One. 2013;8(12), e81902.
91. Harshman SW, et al. Histone H1 phosphorylation in breast cancer. J Proteome Res. 2014; 13(5):2453–67.
92. Jason LJ, et al. Histone ubiquitination: a tagging tail unfolds? Bioessays. 2002;24(2): 166–74.
93. Prenzel T, et al. Estrogen-dependent gene transcription in human breast cancer cells relies upon proteasome-dependent monoubiquitination of histone H2B. Cancer Res. 2011;71(17): 5739–53.
94. Guertin MJ, et al. Targeted H3R26 deimination specifically facilitates estrogen receptor binding by modifying nucleosome structure. PLoS Genet. 2014;10(9), e1004613.
95. Wilkinson KD. Ubiquitination and deubiquitination: targeting of proteins for degradation by the proteasome. Semin Cell Dev Biol. 2000;11(3):141–8.
96. Cao J, Yan Q. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. Front Oncol. 2012;2:26.
97. Zhang Y. Transcriptional regulation by histone ubiquitination and deubiquitination. Genes Dev. 2003;17(22):2733–40.
98. Cole AJ, Clifton-Bligh R, Marsh DJ. Histone H2B monoubiquitination: roles to play in human malignancy. Endocr Relat Cancer. 2015;22(1):T19–33.
99. Zhang Y. Molecular biology: no exception to reversibility. Nature. 2004;431(7009):637–9.
100. Strahl BD, et al. Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. Curr Biol. 2001;11(12):996–1000.
101. Bauer UM, et al. Methylation at arginine 17 of histone H3 is linked to gene activation. EMBO Rep. 2002;3(1):39–44.
102. Pal S, et al. Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. Mol Cell Biol. 2004;24(21):9630–45.
103. Chang X, Han J. Expression of peptidylarginine deiminase type 4 (PAD4) in various tumors. Mol Carcinog. 2006;45(3):183–96.
104. Chang X, et al. Increased PADI4 expression in blood and tissues of patients with malignant tumors. BMC Cancer. 2009;9:40.
105. Krzeslak A, et al. Gene expression of O-GlcNAc cycling enzymes in human breast cancers. Clin Exp Med. 2012;12(1):61–5.
106. Bogachek MV, et al. Sumoylation pathway is required to maintain the basal breast cancer subtype. Cancer Cell. 2014;25(6):748–61.